Hardeo Sahai and Mario Miguel Ojeda

Analysis of Variance for Random Models

Volume II

Unbalanced Data

Theory, Methods, Applications, and Data Analysis

Birkhäuser

Hardeo Sahai
Mario Miguel Ojeda

Analysis of Variance for Random Models

Volume II: Unbalanced Data

Theory, Methods, Applications, and Data Analysis

Birkhäuser
Boston • Basel • Berlin

Hardeo Sahai
Center for Addiction Studies
School of Medicine
Universidad Central del Caribe
Bayamon, Puerto Rico 00960-6032
USA

Mario Miguel Ojeda
Director General del Área Académica
Económico Administrativa
Universidad Veracruzana
Xalapa, Veracruz C.P. 91090
México

Cover design by Alex Gerasev.

AMS Subject Classifications: 62H, 62J

Library of Congress Cataloging-in-Publication Data
Sahai, Hardeo.
Analysis of variance from random models : theory, methods, applications, and data analysis /Hardeo Sahai, Mario Miguel Ojeda.
p. cm.
Includes bibliographical references and index.
Contents: v.1. Balanced data.
ISBN 0-8176-3230-1 (v. 1: alk. paper)
1. Analysis of variance. I. Ojeda, Mario Miguel, 1959- II. Title.

QA279.S23 2003
519.5′38–dc22 20030630260

ISBN 0-8176-3229-8 Volume II Printed on acid-free paper.
ISBN 0-8176-3230-1 Volume I

Printed in the United States of America. (JLS/SB)

9 8 7 6 5 4 3 2 1

www.birkhauser.com

Contents

List of Figures **xiii**

List of Tables **xv**

Preface **xix**

Acknowledgments **xxiii**

9 Matrix Preliminaries and General Linear Model **1**

9.1 Generalized Inverse of a Matrix 3
9.2 Trace of a Matrix 4
9.3 Quadratic Forms 4
9.4 General Linear Model 6
9.4.1 Mathematical Model 6
9.4.2 Expectation Under Fixed Effects 7
9.4.3 Expectation Under Mixed Effects 8
9.4.4 Expectation Under Random Effects 9
Exercises 9
Bibliography 10

10 Some General Methods for Making Inferences about Variance Components **13**

10.1 Henderson's Method I 14
10.2 Henderson's Method II 16
10.3 Henderson's Method III 18
10.4 Analysis of Means Method 22
10.5 Symmetric Sums Method 23
10.6 Estimation of Population Mean in a Random Effects Model . 25
10.7 Maximum Likelihood Estimation 27
10.7.1 Hartley–Rao Estimation Procedure 27
10.7.2 Large Sample Variances 30
10.8 Restricted Maximum Likelihood Estimation 33
10.8.1 Numerical Algorithms, Transformations, and Computer Programs 38
10.9 Best Quadratic Unbiased Estimation 40
10.10 Minimum-Norm and Minimum-Variance Quadratic Unbiased Estimation 40
10.10.1 Formulation of MINQUE and MIVQUE 41

10.10.2 Development of the MINQUE 43
10.10.3 Development of the MIVQUE 47
10.10.4 Some Comments on MINQUE and MIVQUE . . . 48
10.11 Minimum Mean Squared Error Quadratic Estimation 50
10.12 Nonnegative Quadratic Unbiased Estimation 53
10.13 Other Models, Principles and Procedures 55
10.13.1 Covariance Components Model 55
10.13.2 Dispersion-Mean Model 56
10.13.3 Linear Models for Discrete and Categorical Data . . 56
10.13.4 Hierarchical or Multilevel Linear Models 56
10.13.5 Diallel Cross Experiments 57
10.13.6 Prediction of Random Effects 58
10.13.7 Bayesian Estimation 58
10.13.8 Gibbs Sampling 59
10.13.9 Generalized Linear Mixed Models 60
10.13.10 Nonlinear Mixed Models 61
10.13.11 Miscellany 61
10.14 Relative Merits and Demerits of General Methods of Estimation 62
10.15 Comparisons of Designs and Estimators 64
10.16 Methods of Hypothesis Testing 66
10.17 Methods for Constructing Confidence Intervals 69
Exercises 70
Bibliography 76

11 One-Way Classification 93
11.1 Mathematical Model 93
11.2 Analysis of Variance 94
11.3 Minimal Sufficient Statistics and Distribution Theory 96
11.4 Classical Estimation 97
11.4.1 Analysis of Variance Estimators 97
11.4.2 Fitting-Constants-Method Estimators 99
11.4.3 Symmetric Sums Estimators 100
11.4.4 Estimation of μ 102
11.4.5 Maximum Likelihood and Restricted Maximum Likelihood Estimators 103
11.4.6 Best Quadratic Unbiased Estimators 106
11.4.7 Naqvi's Goodness-of-Fit Estimators 108
11.4.8 Rao's MIVQUE and MINQUE 109
11.4.9 An Unbiased Estimator of $\sigma_\alpha^2/\sigma_e^2$ 113
11.4.10 Estimation of $\sigma_\alpha^2/(\sigma_e^2 + \sigma_\alpha^2)$ 113
11.4.11 A Numerical Example 114
11.5 Bayesian Estimation 117
11.5.1 Joint Posterior Distribution of $(\sigma_e^2, \sigma_\alpha^2)$ 119
11.5.2 Joint Posterior Distribution of $(\sigma_e^2, \sigma_\alpha^2/\sigma_e^2)$ 119

11.5.3 Conditional Posterior Distribution of σ_e^2 Given τ . . 120
11.5.4 Marginal Posterior Distributions of σ_e^2 and σ_α^2 . . . 120
11.5.5 Inferences About μ 121
11.6 Distribution and Sampling Variances of Estimators 122
11.6.1 Distribution of the Estimator of σ_e^2 122
11.6.2 Distribution of the Estimators of σ_α^2 123
11.6.3 Sampling Variances of Estimators 124
11.7 Comparisons of Designs and Estimators 131
11.8 Confidence Intervals 136
11.8.1 Confidence Interval for σ_e^2 136
11.8.2 Confidence Intervals for $\sigma_\alpha^2/\sigma_e^2$ and $\sigma_\alpha^2/(\sigma_e^2+\sigma_\alpha^2)$. 136
11.8.3 Confidence Intervals for σ_α^2. 143
11.8.4 A Numerical Example 145
11.9 Tests of Hypotheses 146
11.9.1 Tests for σ_e^2 and σ_α^2 146
11.9.2 Tests for τ 147
11.9.3 Tests for ρ 148
11.9.4 A Numerical Example 149
Exercises 149
Bibliography 155

12 Two-Way Crossed Classification without Interaction 165
12.1 Mathematical Model 165
12.2 Analysis of Variance 165
12.3 Expected Mean Squares 167
12.4 Estimation of Variance Components 169
12.4.1 Analysis of Variance Estimators 169
12.4.2 Fitting-Constants-Method Estimators 170
12.4.3 Analysis of Means Estimators 174
12.4.4 Symmetric Sums Estimators 178
12.4.5 Other Estimators 181
12.4.6 A Numerical Example 181
12.5 Variances of Estimators 186
12.5.1 Variances of Analysis of Variance Estimators 186
12.5.2 Variances of Fitting-Constants-Method Estimators . 189
12.6 Confidence Intervals 190
12.6.1 A Numerical Example 192
12.7 Tests of Hypotheses 193
12.7.1 Tests for $\sigma_\alpha^2 = 0$ and $\sigma_\beta^2 = 0$ 193
12.7.2 A Numerical Example 194
Exercises 195
Bibliography 198

13 Two-Way Crossed Classification with Interaction **201**
13.1 Mathematical Model 201
13.2 Analysis of Variances 201
13.3 Expected Mean Squares 204
13.4 Estimation of Variance Components 206
13.4.1 Analysis of Variance Estimators 207
13.4.2 Fitting-Constants-Method Estimators 208
13.4.3 Analysis of Means Estimators 213
13.4.4 Symmetric Sums Estimators 217
13.4.5 Other Estimators 221
13.4.6 A Numerical Example 221
13.5 Variances of Estimators 226
13.5.1 Variances of Analysis of Variance Estimators 227
13.5.2 Variances of Fitting-Constants-Method Estimators . 229
13.5.3 Variances of Analysis of Means Estimators 230
13.6 Comparisons of Designs and Estimators 230
13.7 Confidence Intervals 238
13.7.1 A Numerical Example 239
13.8 Tests of Hypotheses 240
13.8.1 Some Approximate Tests 240
13.8.2 Some Exact Tests 242
13.8.3 A Numerical Example 242
Exercises 243
Bibliography 250

14 Three-Way and Higher-Order Crossed Classifications **255**
14.1 Mathematical Model 255
14.2 Analysis of Variance 256
14.3 Expected Mean Squares 258
14.4 Estimation of Variance Components 261
14.4.1 Analysis of Variance Estimators 261
14.4.2 Symmetric Sums Estimators 262
14.4.3 Other Estimators 264
14.5 Variances of Estimators 264
14.6 General r-Way Crossed Classification 266
14.7 A Numerical Example 268
Exercises 272
Appendix: Coefficients A_{ij} of Products of Variance Components in Covariance Matrix of $\boldsymbol{T}$ 276
Bibliography 285

15 Two-Way Nested Classification **287**
15.1 Mathematical Model 287
15.2 Analysis of Variance 288
15.3 Expected Mean Squares 289
15.4 Distribution Theory 292
15.5 Unweighted Means Analysis 293
15.6 Estimation of Variance Components 295
15.6.1 Analysis of Variance Estimators 295
15.6.2 Unweighted Means Estimators 295
15.6.3 Symmetric Sums Estimators 296
15.6.4 Other Estimators 299
15.6.5 A Numerical Example 299
15.7 Variances of Estimators 303
15.7.1 Variances of Analysis of Variance Estimators 304
15.7.2 Large Sample Variances of Maximum Likelihood Estimators 306
15.8 Comparisons of Designs and Estimators 306
15.9 Confidence Intervals 310
15.9.1 Confidence Interval for σ_e^2 310
15.9.2 Confidence Intervals for σ_β^2 and σ_α^2 310
15.9.3 Confidence Intervals for $\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2$ 313
15.9.4 Confidence Intervals on $\sigma_\beta^2/\sigma_e^2$ and $\sigma_\alpha^2/\sigma_e^2$ 313
15.9.5 Confidence Intervals on $\sigma_\alpha^2/(\sigma_e^2+\sigma_\beta^2+\sigma_\alpha^2)$ and $\sigma_\beta^2/(\sigma_e^2+\sigma_\beta^2+\sigma_\alpha^2)$ 314
15.9.6 A Numerical Example 315
15.10 Tests of Hypotheses 317
15.10.1 Tests for $\sigma_\beta^2 = 0$ 317
15.10.2 Tests for $\sigma_\alpha^2 = 0$ 318
15.10.3 A Numerical Example 320
Exercises 321
Bibliography 325

16 Three-Way Nested Classification **329**
16.1 Mathematical Model 329
16.2 Analysis of Variance 330
16.3 Expected Mean Squares 331
16.4 Unweighted Means Analysis 334
16.5 Estimation of Variance Components 336
16.5.1 Analysis of Variance Estimators 336
16.5.2 Unweighted Means Estimators 337
16.5.3 Symmetric Sums Estimators 338
16.5.4 Other Estimators 342
16.5.5 A Numerical Example 342
16.6 Variances of Estimators 347

16.6.1 Variances of Analysis of Variance Estimators 347
16.6.2 Large Sample Variances of Maximum Likelihood Estimators 351
16.7 Comparisons of Designs and Estimators 353
16.8 Confidence Intervals and Tests of Hypotheses 359
16.8.1 Confidence Intervals 359
16.8.2 Tests of Hypotheses 359
16.8.3 A Numerical Example 360
Exercises 362
Bibliography 367

17 General r-Way Nested Classification 371
17.1 Mathematical Model 371
17.2 Analysis of Variance 372
17.3 Expected Mean Squares 375
17.4 Estimation of Variance Components 376
17.4.1 Analysis of Variance Estimators 376
17.4.2 Symmetric Sums Estimators 377
17.4.3 Other Estimators 381
17.5 Variances of Estimators 382
17.6 Confidence Intervals and Tests of Hypotheses 382
17.7 A Numerical Example 385
Exercises 387
Bibliography 388

Appendices 391
A Two Useful Lemmas in Distribution Theory 391
B Some Useful Lemmas for a Certain Matrix 393
Bibliography 393
C Incomplete Beta Function 394
Bibliography 395
D Incomplete Inverted Dirichlet Function 395
Bibliography 397
E Inverted Chi-Square Distribution 397
F The Satterthwaite Procedure 397
Bibliography 400
G Maximum Likelihood Estimation 400
H Some Useful Lemmas on the Invariance Property of the ML Estimators 402
I Complete Sufficient Statistics and the Rao–Blackwell and Lehmann–Sheffé Theorems 403
J Point Estimators and the MSE Criterion 403
Bibliography 404
K Likelihood Ratio Test 405
Bibliography 405

L Definition of Interaction 406
M Some Basic Results on Matrix Algebra 406
N Newton–Raphson, Fisher Scoring, and EM Algorithms 415
Bibliography 417
O Software for Variance Component Analysis 418
Bibliography 422

General Bibliography **425**

Author Index **459**

Subject Index **469**

List of Figures

11.1 Program instructions and output for the unbalanced one-way random effects analysis of variance: Data on the ratio of the electromagnetic to electrostatic units of electricity (Table 11.2) 116

13.1 An unconnected BD2 design 232
13.2 Values of n_{ij} for some examples of Bush–Anderson designs . 233

15.1 Program instructions and output for the two-way unbalanced nested random effects analysis of variance: Data on the blood pH readings of female mice (Table 15.3) 301
15.2 Three-stage Bainbridge design 307
15.3 Three-stage Anderson design 307
15.4 Three-stage Prairie–Anderson design 307
15.5 Basic structure of Goldsmith–Gaylor designs 308

16.1 Program instructions and output for the unbalanced three-way nested random effects analysis of variance: Data on insecticide residue on celery from plants sprayed with parathion solution (Table 16.3) 344
16.2 Anderson five-stage staggered nested design 354
16.3 Bainbridge four-stage inverted nested design with a single replicate 354
16.4 Bainbridge four-stage staggered nested design 354
16.5 Heckler–Rao four-stage extended staggered design 358
16.6 Heckler–Rao five-stage extended staggered design 358

List of Tables

11.1 Analysis of variance for the model in (11.1.1) 94
11.2 The ratio of the electromagnetic to electrostatic units of electricity 115
11.3 Analysis of variance for the ratio units of electricity data of Table 11.2 115
11.4 ML, REML, MINQUE(0), and MINQUE(1) estimates of the variance components using SAS®, SPSS®, and BMDP® software 118
11.5 Point estimates of some parametric functions of σ_α^2 and σ_e^2 . 118
11.6 Approximate 95% confidence intervals for τ, ρ, and σ_α^2 . . . 146

12.1 Analysis of variance for the model in (12.1.1) 166
12.2 Analysis of variance based on α adjusted for β 171
12.3 Analysis of variance based on β adjusted for α 173
12.4 Analysis of variance with unweighted sums of squares for the model in (12.1.1) 175
12.5 Analysis of variance with weighted sums of squares for the model in (12.1.1) 177
12.6 Proportions of symptomatic trees from five families and four test locations 182
12.7 Analysis of variance for the fusiform rust data of Table 12.6 . 182
12.8 Analysis of variance for the fusiform rust data of Table 12.6 (location adjusted for family) 184
12.9 Analysis of variance for the fusiform rust data of Table 12.6 (family adjusted for location) 184
12.10 Analysis of variance for the fusiform rust data of Table 12.6 (location adjusted for family and family adjusted for location) 184
12.11 Analysis of variance for the fusiform rust data of Table 12.6 (unweighted sums of squares) 185
12.12 Analysis of variance for the fusiform rust data of Table 12.6 (weighted sums of squares) 185
12.13 ML, REML, MINQUE(0), and MINQUE(1) estimates of the variance components using SAS®, SPSS®, and BMDP® software 186

13.1 Analysis of variance for the model in (13.1.1) 202
13.2 Analysis of variance based on α adjusted for β 209

13.3 Analysis of variance based on β adjusted for α 211
13.4 Analysis of variance with unweighted sums of squares for the model in (13.1.1) 214
13.5 Analysis of variance with weighted sums of squares for the model in (13.1.1) 216
13.6 Data on efficiency scores for assembly line workers 222
13.7 Analysis of variance for the worker efficiency-score data of Table 13.6 222
13.8 Analysis of variance for the efficiency-score data of Table 13.6 (worker adjusted for site) 224
13.9 Analysis of variance for the efficiency-score data of Table 13.6 (site adjusted for worker) 224
13.10 Analysis of variance for the efficiency-score data of Table 13.6 (worker adjusted for site and site adjusted for worker) 224
13.11 Analysis of variance for the efficiency-score data of Table 13.6 (unweighted sums of squares) 226
13.12 Analysis of variance for the efficiency-score data of Table 13.6 (weighted sums of squares) 226
13.13 ML, REML, MINQUE(0), and MINQUE(1) estimates of the variance components using SAS®, SPSS®, and BMDP® software 227
13.14 Efficiencies (E) of some two-way designs for estimating σ_α^2 and $\rho_\alpha (N = 30)$ 231
13.15 Incidence matrices for the Muse designs 235
13.16 Trace asymptotic variance results of Muse designs 235
13.17 Ratios of small sample MSE estimates (SS) and asymptotic variance (LS) for the BD2 and OD3 designs relative to the B design ($\sigma_e^2 = 1$) 236
13.18 Description of Thitakamol designs 237
13.19 Trace asymptotic variance results of Thitakamol 237

14.1 Analysis of variance for the model in (14.1.1) 256
14.2 Production output from an industrial experiment 269
14.3 Analysis of variance for the production output data of Table 14.2 271
14.4 ML, REML, MINQUE(0), and MINQUE(1) estimates of the variance components using SAS®, SPSS®, and BMDP® software 273

15.1 Analysis of variance for the model in (15.1.1) 288
15.2 Analysis of variance using unweighted means analysis for the model in (15.1.1) 294
15.3 Blood pH readings of female mice 300

15.4 Analysis of variance for the blood pH reading data of Table 15.3 . . . 301
15.5 Analysis of variance for the blood pH reading data Table 15.3 using unweighted sums of squares . . . 303
15.6 ML, REML, MINQUE(0), and MINQUE(1) estimates of the variance components using SAS®, SPSS®, and BMDP® software . . . 304
15.7 Test procedures for $\sigma_\alpha^2 = 0$ . . . 321

16.1 Analysis of variance for the model in (16.1.1) . . . 331
16.2 Analysis of variance with unweighted sums of squares for the model in (16.1.1) . . . 337
16.3 The insecticide residue on celery from plants sprayed with parathion solution . . . 343
16.4 Analysis of variance for the insecticide residue data of Table 16.3 . . . 344
16.5 Analysis of variance for the insecticide residue data of Table 16.3 (unweighted sums of squares) . . . 346
16.6 ML, REML, MINQUE(0), and MINQUE(1) estimates of the variance components using SAS®, SPSS®, and BMDP® software . . . 348
16.7 Sets of variance components included in the empirical study of the balanced, inverted, and staggered nested designs 356
16.8 Empirical percentages of negative estimates of the variance components . . . 357

17.1 Analysis of variance for the model in (17.1.1) . . . 373
17.2 Analysis of variance of the insecticide residue data of Table 16.3 . . . 385
17.3 Tests of hypotheses for $\sigma_i^2 = 0, i = 1, 2, 3$ . . . 387

Preface

Random effects models have found widespread applications in a variety of substantive fields requiring measurement of variance, including agriculture, biology, animal breeding, applied genetics, econometrics, quality control, medicine, engineering, education, and environmental and social sciences, among others.

The purpose of this monograph is to present a comprehensive coverage of different methods and techniques of point estimation, interval estimation, and tests of hypotheses for linear models involving random effects. Both Bayesian and repeated sampling procedures are considered. The book gives a survey of major theoretical and methodological developments in the area of estimation and testing of variance components of a random model and the related inference. It also includes numerical examples illustrating the use of these methods in analyzing data from research studies in agriculture, engineering, biology and other related fields. Many required computations can be readily performed with the assistance of a handheld scientific calculator. However, for large data sets and computationally complex procedures, the use of appropriate software is highly recommended. Most of the results being presented can be used by applied scientists and researchers with only a modest mathematical and statistical background. Thus, the work will appeal to graduate students and theoretical researchers as well as applied workers interested in using these methods in their respective fields of applications.

We consider a variety of experimental designs involving one factor, two factors, three factors, and other multifactor experiments. These include both crossed and nested designs with both balanced and unbalanced data sets. The analysis of variance models being presented include random models involving one-way, two-way, three-way, and other higher-order classifications. We illustrate the importance of these models and present a survey of their historical origins to a variety of substantive fields of research.

Many of the results being discussed are of relatively recent origin, and many of the books on linear models, analysis of variance, and experimental designs do not provide adequate coverage of these topics. Although there are a multitude of books and other publications giving a complete treatment of the fixed linear models, the number of such works devoted to random and mixed linear models is limited mainly to an abstract viewpoint and is not accessible for a wide readership. The present work is designed to rectify this situation, and we hope this monograph will fill a longstanding niche in this area and will serve the needs of both theoretical researchers and applied scientists. Applied readers can use the text with a judicious choice of topics and numerical examples of relevance to their work. Readers primarily interested in theoretical developments in the field will also find ample material and an abundance of references to guide them in their work.

Although the monograph includes some results and proofs requiring knowledge of advanced statistical theory, all of the theoretical developments have been kept to a minimal level. Most of the material can be read and understood by readers with basic knowledge of statistical inference and some background in analysis of variance and experimental design. The book can be used as a textbook for graduate-level courses in analysis of variance and experimental design. It will also serve as a handy reference for a broad spectrum of topics and results for applied scientists and practicing statisticians who need to use random models in their professional work.

The literature being surveyed in this volume is so vast, and the number of researchers and users so large that it is impossible to write a book which will satisfy the needs of all the workers in this field. Moreover, the number of papers both theoretical and methodological devoted to this topic is increasing so rapidly that it is not possible to provide a complete and up-to-date coverage. Nevertheless, we are confident that the present work provides a broad and comprehensive overview of all the basic developments in the field and will meet the professional needs of most of the researchers and practitioners interested in using the methodology presented here.

We have tried to elucidate in a unified way the basic results for the random effects analysis of variance. The work presents an introduction to many of the recently developed general results in the area of point and interval estimation and hypothesis testing on random effect models. Only the infinite population theory has been considered. The literature on the subject is vast and widely scattered over many books and periodicals. This monograph is an assemblage of the several publications on the subject and contains a considerable expansion and generalization of many ideas and results given in original works. Many of the results, we expect, will undergo considerable extension and revision in the future. Perhaps this presentation will help to stimulate the needed growth. For example, in the not too distant past, the estimation of variance components in many cases was limited to the so-called analysis of variance procedure. Today, a bewildering variety of new estimation procedures are available and many more are being developed. The entire work is devoted to the study of methods for balanced and unbalanced (i.e., unequal-subclass-numbers) data. Volume I deals with the analyses and results for balanced models, while Volume II is concerned with unbalanced models.

We have stated many theoretical results without proofs, in many cases, and referred readers to the literature for proofs. It is hoped that the sophisticated reader with a higher degree of scholarly interest will go through these sources to get a through grounding of the theory involved. At this time, it has not been possible to consider topics such as finite population models, multivariate generalizations, sequential methods, and nonparametric analogues to the random effects models, including experimental plans involving incomplete and mixed models. The omission of these topics is most sincerely regretted. It is hoped that many of these topics will be covered in a future volume, which is in preparation. The monograph also does not contain a complete bibliography. We have only given selected references for readers who desire to study some background material. Several bibliographies on the subject are currently available and the interested reader is referred to these publications for any additional work not included here.

The textbook contains an abundance of *footnotes* and *remarks*. They are intended for statistically sophisticated readers who wish to pursue the subject matter in greater depth, and it is not necessary that a novice studying the text for the first time read them. They often expand and elaborate on a particular topic, point the way to generalization and to other techniques, and make historical comments and remarks. In addition, they contain literature citations for further exploration of the topic and refer to finer points of theory and methods. We are confident that this two-tier approach will be pedagogically appealing and useful to readers with a higher degree of scholarly interest.

Hardeo Sahai and Mario Miguel Ojeda April 2004

Acknowledgments

The present work is an outgrowth of a number of courses and seminars that the authors have taught during the last twenty five years at the University of Puerto Rico, University of Veracruz (México), Federal University of Ceará (Brazil), National University of Colombia, National University of Trujillo (Perú), the University of Granada (Spain), and in various other forums and scientific meetings; and our sincere thanks go to students and others who have attended these courses and contributed many useful ideas to its development. Some of the results presented in the book have been adapted from the lecture notes which one of us transcribed, based on courses and seminars offered by Dr. Richard L. Anderson at the University of Kentucky, and we are deeply indebted to him; in many ways this work is his, too.

Although the material being presented here has been written by us and the book will bear our name, we do not make any claim to the authorship. The work is, indeed, a sprouting of the seeds and inspirations given to us by our parents, teachers, colleagues, and students, and the bounty of the crop sown by innumerable researchers, scientists, and professionals that we have lavishly harvested. In the words of Ralph W. Emerson, "Every book is a quotation; and every house is a quotation out of all forests and mines and stone quarries; and every man is a quotation from all his ancestors. . . ." Our sincere gratitude to the authors of papers, textbooks, monographs, lecture notes, technical reports, encyclopedias, and other publications that provided the basis for the development of this work, and who have thus contributed to its authorship. We have made every attempt to acknowledge results, formulas, data sources, or any other material utilized from the original sources and any subsequent works referring to these for the sake of wide accessibility. However, there is no guarantee for any accuracy or completeness, and any omission of due credit or priority is deeply regretted and would be rectified in any future revision of this work. Needless to say, any errors, omissions, or other shortcomings are our own demerits, for which we bear the sole responsibility.

We are especially thankful to the painstaking work of Janet Andrade, Margarita Caballero, Juliana Carmona, Guillermo Cruz, Diana González, Jaime Jiménez, Adalberto Lara, Idalia Lucero, Imelda Mendoza, Judith Montero, Edgar Morales, Hugo Salazar, Adrián Sánchez, Wendy Sánchez, and Lourdes Velazco of the Statistical Research and Consulting Laboratory, University of Veracruz, Xalapa, México, who with the assistance of other students carried out the arduous task of word processing the entire manuscript, in its numerous incarnations. Professor Lorena López and Dr. Anwer Khurshid assisted us in so many ways from the inception until the conclusion of the project, and we are immensely grateful for all their time, help, and cooperation, which they swiftly and cheerfully offered.

Parts of the manuscript were written and revised during the course of one author's secondment as the Patrimonial Professor of Statistics at the University of Veracruz (México), and he wishes to thank the Mexican National Council of Science and Technology (CONACYT) for extending the appointment and providing a stimulating environment for research and study. He would also like to acknowledge two sabbaticals (1978–1979 and 1993–1994) granted by the Administrative Board of the University of Puerto Rico, which provided the time to compile the material presented in this book.

Two anonymous reviewers provided several constructive comments and suggestions on the most recent draft, and undoubtedly the final text has greatly benefited from their input.

Dr. Rául Micchiavelli of the University of Puerto Rico and Mr. Guadalupe Hernádez Lira of the University of Veracruz (México) assisted us in running worked examples using statistical packages, and their helpful support is greatly appreciated.

The first author wishes to extend a warm appreciation to members and staff of the Puerto Rico Center for Addiction Research, especially Dr. Rafaela R. Robles, Dr. Héctor M. Colón, Ms. Carmen A. Marrero, M.P.H., Mr. Tomás L. Matos, M.S., and Dr. Juan C. Reyes, M.P.H., who as an innovative research group, for well over a decade, provided an intellectually stimulating environment and a lively research forum to discuss and debate the role of analysis of variance models in social and behavioral research.

Our grateful and special thanks go to our publisher, especially Ann Kostant, Executive Editor of Mathematics and Physics, and Tom Grasso, Statistics Editor, for their encouragement and support of the project. Equally, we would like to record our thanks to the editorial and production staff at Birkhäuser, especially Seth Barnes and Elizabeth Loew, for all their help and cooperation in bringing the project to its fruition. We particularly acknowledge the work of John Spiegelman, who worked long hours, above and beyond his normal call of duty, drawing on his considerable skills and experience in mathematical publishing to convert a highly complex manuscript to an elegant and camera-ready format using LaTeX with supreme care and accuracy. His discovery of techniques not commonly explained in manuals proved to be invaluable in typesetting many complex mathematical expressions and equations.

The authors and Birkhäuser would like to thank many authors, publishers, and other organizations for their kind permission to use the data and to reprint whole or parts of statistical tables from their previously published copyrighted materials, and the acknowledgments are made in the book where they appear.

Finally, we must make a special acknowledgment of gratitude to our families, who were patient during the many hours of daily work devoted to the book, in what seemed like an endless process of revisions for finalizing the manuscript, and we are greatly indebted for their continued help and support. Hardeo Sahai would like to thank his children Amogh, Mrisa, and Pankaj for their infinite patience and understanding throughout the time the work was in progress. Mario M. Ojeda owes an immense sense of appreciation to his dear wife Olivia for her patience and understanding during the countless hours spent on the project that truly belonged to her and the family.

The authors welcome any suggestions and criticisms of the book in regards to omissions, inaccuracies, corrections, additions, or ways of presentation that would be rectified in any further revision of this work.

Hardeo Sahai and Mario Miguel Ojeda April 2004

9 Matrix Preliminaries and General Linear Model

Volume I of the text was devoted to a study of various models with the common feature that the same numbers of observations were taken from each treatment group or in each submost subcell. When these numbers are the same, the data are referred to as balanced data; in contrast, when the numbers of observations in the cells are not all equal, the data are known as unbalanced data. In general, it is desirable to have equal numbers of observations in each subclass since the experiments with unbalanced data are much more complex and difficult to analyze and interpret than the ones with balanced data. However, in many practical situations, it is not always possible to have equal numbers of observations for the treatments or groups. Even if an experiment is well-thought-out and planned to be balanced, it may run into problems during execution due to circumstances beyond the control of the experimenter; for example, missing values or deletion of faulty observations may result in different sample sizes in different groups or cells. In many cases, the data may arise through a sample survey where the numbers of observations per group cannot be predetermined, or through an experiment designed to yield balanced data but which actually may result in unbalanced data because some plants or animals may die, patients may drop out or be taken out of the study. For example, in many clinical investigations involving a follow-up, patients may decide to discontinue their participation, they may withdraw due to side effects, they may die, or they are simply lost to follow-up. In many experiments, materials and other resources may be limited or accidentally destroyed, or observations misread or misrecorded that cannot be later used for any valid data analysis.

In the situations described above, of course, one might question the validity of the data, since the subjects or plants that are lost may be systematically different from those that survived. However, in many practical investigations, it is not always possible to meet all the assumptions precisely. Frequently, we have to rely on our good judgment and common sense to decide whether the departures from the assumptions are serious enough to make a difference. Although it is rather impossible to eliminate bias due to missing observations, there are ways to minimize its impact and to assess the likelihood and magnitude of such bias. For the purposes of our analyses, we will assume that the effects of the departures are negligible. Moreover, there are many situations in which, due to the nature of the experimental material, the treatment effects cannot be applied in a balanced way, and much unbalanced data then occur in a rather natural way. For example, in the production of milk involving a large number of cows classified according to the sire (male parent), the unbalanced data are the norm rather than the exception. Also there are situations, when a researcher may purposely design his experiment to have unequal numbers in the subclasses in order to estimate variance components with certain optimal properties. Such a situation is referred to as planned unbalancedness where no observations are obtained on certain, carefully planned combinations of levels of the factors involved in the experiment (see, e.g., Bainbridge, 1963; Bush and Anderson, 1963; Anderson and Crump, 1967; Muse and Anderson, 1978; Muse et al., 1982; Shen et al., 1996a, 1996b).

As mentioned earlier, inferences on variance components from unbalanced data are much more complicated than from balanced data. The reason is that the analysis of variance of balanced data is fairly straightforward since there exists a unique partitioning of the total sum of squares into component sums of squares, which under standard distributional assumptions follow a multiple of a chi-square distribution; this multiple being the product of the degrees of freedom and the expected mean square of one of the random effects. Thus the hypotheses about the treatment effects can be tested by dividing treatment mean squares by the appropriate error mean square to form a variance-ratio F-test. In contrast, analysis of unbalanced data lacks these properties since there does not exist a unique partitioning of the total sum of squares, and consequently there is no unique analysis of variance. In addition, in any given decomposition, the component sums of squares are not in general independent or distributed as chi-square type variables, and corresponding to any particular treatment mean square there does not exist an error mean square with equal expectation under the null hypothesis. Furthermore, the analysis of variance for unbalanced data involves relatively cumbersome and tedious algebra, and extensive numerical computations.

In this chapter, we briefly review some important results in matrix theory on topics such as generalized inverse, trace operation and quadratic forms and present an introduction to the general linear model. The results are extremely useful in the study of variance component models for unbalanced data. A more extensive review of basic results in matrix theory is given in Appendix M. Currently there are a number of textbooks on matrix algebra that are devoted entirely to the subject of modern matrix methods and their applications to statistics, particularly, the linear model. Among these are Pringle and Raynor (1971), Rao and Mitra (1971), Albert (1972), Ben-Israel and Greyville (1974), Seneta (1981), Searle (1982), Basilevsky (1983), Graybill (1983), Horn and Johnson (1985), Healy (2000), Berman and Plemmons (1994), Hadi (1996), Bapat and Raghavan (1997), Harville (1997), Schott (1997), Gentle (1998), Rao and Rao (1998), and Magnus and Neudecker (1999).

9.1 GENERALIZED INVERSE OF A MATRIX

The concept of *generalized inverse* of a matrix plays an important role in the study of linear models, though their application to such models is of relatively late origin. The use of such matrices as a mathematical tool greatly facilitates the understanding of certain aspects relevant to the analysis of linear models, especially the analysis of unbalanced data, which we will be concerned with in this volume of the text. In particular, they are very useful in the simplification of the development of the "less than full-rank" linear model. The topic of generalized inverse is discussed in many textbooks (see, e.g., Rao and Mitra, 1971; Pringle and Raynor, 1971; Ben-Israel and Greyville, 1974). Of its many definitions, we will make use of the following one.

Definition 9.1.1. Given an $m \times n$ matrix $\boldsymbol{A}$, its generalized inverse is a matrix denoted by $\boldsymbol{A}^-$, that satisfies the condition

$$\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A} = \boldsymbol{A}. \tag{9.1.1}$$

There are several other generalizations of the inverse matrix that have been proposed for a rectangular matrix of any rank. The definition given in (9.1.1) is useful for solving a system of linear equations and will suffice for our purposes. Several other alternative terms for "generalized inverse," such as "conditional inverse," "pseudoinverse," and "g-inverse," are sometimes employed in the literature. Further, note that the matrix $\boldsymbol{A}^-$ defined in (9.1.1) is not unique since there exists a whole class of matrices $\boldsymbol{A}^-$ that satisfy (9.1.1).

The study of linear models frequently leads to equations of the form

$$\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{Y}$$

that has to be solved for $\hat{\boldsymbol{\beta}}$. Therefore, the properties of a generalized inverse of the symmetric matrix $\boldsymbol{X}'\boldsymbol{X}$ are of particular interest. The following theorem gives some useful properties of a generalized inverse of $\boldsymbol{X}'\boldsymbol{X}$.

Theorem 9.1.1. *If* $(\boldsymbol{X}'\boldsymbol{X})^-$ *is a generalized inverse of* $\boldsymbol{X}'\boldsymbol{X}$, *then*

(i) $[(\boldsymbol{X}'\boldsymbol{X})^-]'$, *its transpose, is also a generalized inverse of* $\boldsymbol{X}'\boldsymbol{X}$;

(ii) $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}$, *i.e.,* $(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'$ *is a generalized inverse of* $\boldsymbol{X}$;

(iii) $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'$ *is invariant to* $(\boldsymbol{X}'\boldsymbol{X})^-$;

(iv) $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'$ *is symmetric, irrespective of whether* $(\boldsymbol{X}'\boldsymbol{X})^-$ *is symmetric or not;*

(v) $\text{rank}[(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X}] = \text{rank}(\boldsymbol{X})$.

Proof. See Searle (1971, p. 20). □

9.2 TRACE OF A MATRIX

The concept of *trace* of a matrix plays an important role in the analysis of linear models. The trace of a matrix is defined as follows.

Definition 9.2.1. The trace of a square matrix $\boldsymbol{A}$, denoted by $\mathrm{tr}(\boldsymbol{A})$, is the sum of its diagonal elements. More specifically, given a square matrix

$$\boldsymbol{A} = (a_{ij}), \quad i, j = 1, 2, \ldots, n;$$
$$\mathrm{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}.$$

The following theorem gives some useful properties associated with the trace operation of matrices.

Theorem 9.2.1. *Under the trace operation of matrices, the following results hold:*

(i) $\mathrm{tr}(\boldsymbol{A} + \boldsymbol{B} + \boldsymbol{C}) = \mathrm{tr}(\boldsymbol{A}) + \mathrm{tr}(\boldsymbol{B}) + \mathrm{tr}(\boldsymbol{C})$;

(ii) $\mathrm{tr}(\boldsymbol{ABC}) = \mathrm{tr}(\boldsymbol{BCA}) = \mathrm{tr}(\boldsymbol{CBA})$; *that is, under the trace operation, matrix products are cyclically commutative;*

(iii) $\mathrm{tr}(\boldsymbol{S}^{-1}\boldsymbol{AS}) = \mathrm{tr}(\boldsymbol{A})$;

(iv) $\mathrm{tr}(\boldsymbol{AA}^{-}) = \mathrm{rank}(\boldsymbol{A})$;

(v) $\mathrm{tr}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A})$, *if* $\boldsymbol{A}$ *is idempotent;*

(vi) $\mathrm{tr}(\boldsymbol{A}) = \sum_i \lambda_i$, *where* λ_i*s are latent roots of* $\boldsymbol{A}$;

(vii) $\boldsymbol{Y}'\boldsymbol{AY} = \mathrm{tr}(\boldsymbol{Y}'\boldsymbol{AY}) = \mathrm{tr}(\boldsymbol{AYY}')$;

(viii) $\mathrm{tr}(\boldsymbol{I}) = n$, *where* $\boldsymbol{I}$ *is an* $n \times n$ *identity matrix;*

(ix) $\mathrm{tr}(\boldsymbol{S}'\boldsymbol{AS}) = \mathrm{tr}(\boldsymbol{A})$, *if* $\boldsymbol{S}$ *is an orthogonal matrix.*

Proof. See Graybill (1983, Chapter 9). □

9.3 QUADRATIC FORMS

The methods for estimating variance components from unbalanced data employ, in one way or another, quadratic forms of the observations. The quadratic form associated with a column vector $\boldsymbol{Y}$ and a square matrix $\boldsymbol{A}$ is defined as follows.

Definition 9.3.1. An expression of the form $\boldsymbol{Y}'\boldsymbol{AY}$ is called a quadratic form and is a quadratic function of the elements of $\boldsymbol{Y}$.

The following theorem gives some useful results concerning the distribution of a quadratic form.

Theorem 9.3.1. *For a random vector* $\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *and positive definite matrices* $\boldsymbol{A}$ *and* $\boldsymbol{B}$, *we have the following results:*

(i) $E(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) = \text{tr}(\boldsymbol{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$;

(ii) $\text{Cov}(\boldsymbol{Y}, \boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) = 2\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\mu}$;

(iii) $\text{Var}(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) = 2\,\text{tr}(\boldsymbol{A}\boldsymbol{\Sigma})^2 + 4\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\mu}$;

(iv) $\text{Cov}(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}, \boldsymbol{Y}'\boldsymbol{B}\boldsymbol{Y}) = 2\,\text{tr}(\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B}\boldsymbol{\Sigma}) + 4\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B}\boldsymbol{\mu}$.

Proof.

(i) From Theorem 9.2.1, we have

$$\begin{aligned} E(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) &= E[\text{tr}(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y})] = E[\text{tr}(\boldsymbol{A}\boldsymbol{Y}\boldsymbol{Y}')] \\ &= \text{tr}[E(\boldsymbol{A}\boldsymbol{Y}\boldsymbol{Y}')] = \text{tr}[\boldsymbol{A}E(\boldsymbol{Y}\boldsymbol{Y}')]. \end{aligned} \tag{9.3.1}$$

Now, since $E(\boldsymbol{Y}) = \boldsymbol{\mu}$ and $\text{Var}(\boldsymbol{Y}) = \boldsymbol{\Sigma}$, we obtain

$$E(\boldsymbol{Y}\boldsymbol{Y}') = \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}'. \tag{9.3.2}$$

Substituting (9.3.2) into (9.3.1), we obtain

$$\begin{aligned} E(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) &= \text{tr}[\boldsymbol{A}(\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}')] = \text{tr}[(\boldsymbol{A}\boldsymbol{\Sigma}) + (\boldsymbol{A}\boldsymbol{\mu}\boldsymbol{\mu}')] \\ &= \text{tr}(\boldsymbol{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}. \end{aligned}$$

It is evident from the proof that this part of the theorem holds irrespective of whether $\boldsymbol{Y}$ is normal or not.

(ii) We have

$$\begin{aligned} \text{Cov}(\boldsymbol{Y}, \boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) &= E[(\boldsymbol{Y} - \boldsymbol{\mu})\{\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} - E(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y})\}] \\ &= E[(\boldsymbol{Y} - \boldsymbol{\mu})\{\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} - \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu} - \text{tr}(\boldsymbol{A}\boldsymbol{\Sigma})\}] \\ &= E[(\boldsymbol{Y} - \boldsymbol{\mu})\{(\boldsymbol{Y} - \boldsymbol{\mu})'\boldsymbol{A}(\boldsymbol{Y} - \boldsymbol{\mu}) \\ &\quad + 2(\boldsymbol{Y} - \boldsymbol{\mu})'\boldsymbol{A}\boldsymbol{\mu} - \text{tr}(\boldsymbol{A}\boldsymbol{\Sigma})\}] \\ &= 2\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\mu}, \end{aligned}$$

since the first and third moments of $\boldsymbol{Y} - \boldsymbol{\mu}$ are zero.

For the proofs of (iii) and (iv), see Searle (1971, pp. 65–66). Results similar to Theorem 9.3.1 with $\boldsymbol{\mu} = \boldsymbol{0}$ may be found several places in the literature (see, e.g., Lancaster, 1954; Anderson, 1961; Bush and Anderson, 1963). □

Theorem 9.3.2.

(i) *If the random vector* $\boldsymbol{Y} \sim N(\boldsymbol{0}, \boldsymbol{I}_n)$, *then a necessary and sufficient condition that the quadratic form* $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ *has a chi-square distribution with* v *degrees of freedom is that* $\boldsymbol{A}$ *be an idempotent matrix, of rank* v.

(ii) *If a random vector* $\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{I}_n)$*, then a necessary and sufficient condition that the quadratic form* $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ *has a noncentral chi-square distribution with* v *degrees of freedom and the noncentrality parameter* $\lambda = \frac{1}{2}\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$ *is that* $\boldsymbol{A}$ *be an idempotent matrix of rank* v.

Proof. See Graybill (1961, pp. 82–83). □

Theorem 9.3.3. *Suppose* $\boldsymbol{Y} \sim N(\boldsymbol{0}, \boldsymbol{I}_n)$ *and consider the quadratic forms* $Q_i = \boldsymbol{Y}'\boldsymbol{A}_i\boldsymbol{Y}$*, where* $\boldsymbol{Y}'\boldsymbol{Y} = \sum_{i=1}^{p} Q_i$ *and* $v_i = \text{rank}(\boldsymbol{A}_i)$, $i = 1, 2, \ldots, p$. *Then a necessary and sufficient condition that* Q_i *be independently distributed as* $\chi^2[v_i]$ *is that* $\sum_{i=1}^{p} \text{rank}(\boldsymbol{A}_i) = \text{rank}(\sum_{i=1}^{p} \boldsymbol{A}_i) = n$.

Proof. See Scheffé (1959, pp. 420–421), Graybill (1961, pp. 85–86). □

The theorem is popularly known as the Cochran–Fisher theorem and was first stated by Cochran (1934). A generalization of Theorem 9.3.3 for the distribution of quadratic forms in noncentral normal random variables was given by Madow (1940) and is stated in the theorem below.

Theorem 9.3.4. *Suppose* $\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{I}_n)$ *and consider the quadratic forms* $Q_i = \boldsymbol{Y}'\boldsymbol{A}_i\boldsymbol{Y}$*, where* $\boldsymbol{Y}'\boldsymbol{Y} = \sum_{i=1}^{p} Q_i$, $v_i = \text{rank}(A_i)$*, and* $\lambda_i = \frac{1}{2}\boldsymbol{\mu}'\boldsymbol{A}_i\boldsymbol{\mu}$, $i = 1, 2, \ldots, p$. *Then a necessary and sufficient condition that* Q_i *be independently distributed as* $\chi^2[v_i, \lambda_i]$ *is that* $\sum_{i=1}^{p} \text{rank}(\boldsymbol{A}_i) = \text{rank}(\sum_{i=1}^{p} \boldsymbol{A}_i) = n$.

Proof. See Graybill (1961, pp. 85–86). □

Theorem 9.3.5. *Suppose* $\boldsymbol{Y}$ *is an* N*-vector and* $\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{V})$*, then* $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ *and* $\boldsymbol{Y}'\boldsymbol{B}\boldsymbol{Y}$ *are independent if and only if* $\boldsymbol{A}\boldsymbol{V}\boldsymbol{B} = \boldsymbol{0}$.

Proof. See Graybill (1961, Theorem 4.21). □

Theorem 9.3.6. *Suppose* $\boldsymbol{Y}$ *is an* N*-vector and* $\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{V})$*, then* $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} \sim \chi^2[v, \lambda]$*, where* $\lambda = \frac{1}{2}\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$ *and* v *is the rank of* $\boldsymbol{A}$ *if and only if* $\boldsymbol{A}\boldsymbol{V}$ *is an idempotent matrix. In particular, if* $\boldsymbol{\mu} = \boldsymbol{0}$*, then* $\lambda = 0$*, that is,* $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} \sim \chi^2[v]$.

Proof. See Graybill (1961, Theorem 4.9). □

9.4 GENERAL LINEAR MODEL

The analysis of unbalanced data is more readily understood and appreciated in matrix terminology by considering what is known as the general linear model. In this chapter, we study certain salient features of such a model, which are useful in the problem of variance components estimation.

9.4.1 MATHEMATICAL MODEL

The equation of the general linear model can be written as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}, \tag{9.4.1}$$

where

$\boldsymbol{Y}$ is an N-vector of observations,

$\boldsymbol{X}$ is an $N \times p$ matrix of known fixed numbers, $p \leq N$,

$\boldsymbol{\beta}$ is a p-vector of fixed effects or random variables,

and

$\boldsymbol{e}$ is an N-vector of randomly distributed error terms with mean vector $\mathbf{0}$ and variance-covariance matrix $(\boldsymbol{\sigma}_e^2 \boldsymbol{I}_N)$. In general, the variance-covariance matrix would be $\boldsymbol{\sigma}_e^2 \boldsymbol{V}$, but here we consider only the case $\boldsymbol{\sigma}_e^2 \boldsymbol{I}_N$.

Under Model I, in the terminology of Eisenhart (1947), when the vector $\boldsymbol{\beta}$ represents all fixed effects, the normal equations for estimating $\boldsymbol{\beta}$ are (see, e.g., Graybill, 1961, p. 114, Searle, 1971, pp. 164–165)

$$\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{Y}. \tag{9.4.2}$$

A general solution of (9.4.2) is

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{Y}, \tag{9.4.3}$$

where $(\boldsymbol{X}'\boldsymbol{X})^{-}$ is a generalized inverse of $(\boldsymbol{X}'\boldsymbol{X})$. Now, it can be shown that in fitting the model in (9.4.1), the reduction in sum of squares is (see, e.g., Graybill 1961, pp. 138–139; Searle, 1971, pp. 246–247).

$$R(\boldsymbol{\beta}) = \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{Y} = \boldsymbol{Y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{Y}. \tag{9.4.4}$$

In estimating variance components, we are frequently interested in expected values of $R(\boldsymbol{\beta})$, which requires knowing the expected value of a quadratic form involving the response vector $\boldsymbol{Y}$. Thus we will consider the expected value of the quadratic form $\boldsymbol{Y}'\boldsymbol{Q}\boldsymbol{Y}$, when $\boldsymbol{\beta}$ represents (i) all fixed effects, (ii) all random effects, and (iii) a mixture of both. For some further discussion and details and a survey of mixed models for unbalanced data, see Searle (1971, pp. 421–424; 1988).

9.4.2 EXPECTATION UNDER FIXED EFFECTS

If $\boldsymbol{\beta}$ in (9.4.1) represents all fixed effects, we have

$$E(\boldsymbol{Y}) = \boldsymbol{X}\boldsymbol{\beta} \tag{9.4.5}$$

and

$$\text{Var}(\boldsymbol{Y}) = \sigma_e^2 \boldsymbol{I}_N.$$

Using Theorem 9.3.1 with $\boldsymbol{\mu} = \boldsymbol{X}\boldsymbol{\beta}$ and $\boldsymbol{\Sigma} = \sigma_e^2 \boldsymbol{I}_N$, we obtain

$$E(\boldsymbol{Y}'\boldsymbol{Q}\boldsymbol{Y}) = \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{Q}\boldsymbol{X}\boldsymbol{\beta} + \sigma_e^2 \operatorname{tr}(\boldsymbol{Q}). \tag{9.4.6}$$

Now, we consider two applications of (9.4.6).

(i) If $\boldsymbol{Q} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$, then $\boldsymbol{Y}'\boldsymbol{Q}\boldsymbol{Y}$ is the reduction in sum of squares $R(\boldsymbol{\beta})$ given by (9.4.4). Hence,

$$E[R(\boldsymbol{\beta})] = \boldsymbol{\beta}'\boldsymbol{X}'[\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}']\boldsymbol{X}\boldsymbol{\beta} + \sigma_e^2 \operatorname{tr}[\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}']. \quad (9.4.7)$$

Further, from Theorems 9.1.1 and 9.2.1, we have

$$\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X} \quad (9.4.8)$$

and

$$\begin{aligned}\operatorname{tr}[\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'] &= \operatorname{tr}[(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{X}] \\ &= \operatorname{rank}[(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{X}] \\ &= \operatorname{rank}(\boldsymbol{X}). \quad (9.4.9)\end{aligned}$$

Substituting (9.4.8) and (9.4.9) into (9.4.7) gives

$$E[R(\boldsymbol{\beta})] = \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} + \sigma_e^2 \operatorname{rank}(\boldsymbol{X}). \quad (9.4.10)$$

(ii) The expectation of the residual sum of squares is given by

$$E[\boldsymbol{Y}'\boldsymbol{Y} - R(\boldsymbol{\beta})] = E(\boldsymbol{Y}'\boldsymbol{Y}) - E[R(\boldsymbol{\beta})]. \quad (9.4.11)$$

When $\boldsymbol{Q} = \boldsymbol{I}_N$, the quadratic form $\boldsymbol{Y}'\boldsymbol{Q}\boldsymbol{Y}$ is $\boldsymbol{Y}'\boldsymbol{Y}$, and from (9.4.6), we have

$$E(\boldsymbol{Y}'\boldsymbol{Y}) = \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} + N\sigma_e^2. \quad (9.4.12)$$

Therefore, on substituting (9.4.10) and (9.4.12) into (9.4.11), we obtain

$$E[\boldsymbol{Y}'\boldsymbol{Y} - R(\boldsymbol{\beta})] = [N - \operatorname{rank}(\boldsymbol{X})]\sigma_e^2, \quad (9.4.13)$$

which is a familiar result (see, e.g., Searle, 1971, pp. 170–171).

9.4.3 EXPECTATION UNDER MIXED EFFECTS

If $\boldsymbol{\beta}$ in (9.4.1) represents mixed effects, it can be partitioned as

$$\boldsymbol{\beta}' = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \ldots, \boldsymbol{\beta}_k'),$$

where $\boldsymbol{\beta}_1$ represents all the fixed effects in the model (including the general mean) and $\boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \ldots, \boldsymbol{\beta}_k$ each represents a set of random effects having zero means and zero covariances with the effects of any other set. Then, on partitioning $\boldsymbol{X}$ in conformity with $\boldsymbol{\beta}$ as

$$\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_k),$$

the general linear model in (9.4.1) can be written as

$$\boldsymbol{Y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \cdots + \boldsymbol{X}_k\boldsymbol{\beta}_k + e, \tag{9.4.14}$$

where

$$E(\boldsymbol{Y}) = \boldsymbol{X}_1\boldsymbol{\beta}_1$$

and

$$\begin{aligned}\operatorname{Var}(\boldsymbol{Y}) = {} & \boldsymbol{X}_2 \operatorname{Var}(\boldsymbol{\beta}_2)\boldsymbol{X}_2' + \boldsymbol{X}_3 \operatorname{Var}(\boldsymbol{\beta}_3)\boldsymbol{X}_3' \\ & + \cdots + \boldsymbol{X}_k \operatorname{Var}(\boldsymbol{\beta}_k)\boldsymbol{X}_k' + \sigma_e^2 \boldsymbol{I}_N.\end{aligned} \tag{9.4.15}$$

If, in addition, $\operatorname{Var}(\boldsymbol{\beta}_i) = \sigma_i^2 \boldsymbol{I}_{N_i}$, $i = 2, 3, \ldots, k$, where N_i is the number of different effects of the ith factor, then (9.4.1) represents the usual variance components model. Now, from (9.4.14) and (9.4.15) and using Theorem 9.3.1 with $\boldsymbol{\mu} = E(\boldsymbol{Y}) = \boldsymbol{X}_1\boldsymbol{\beta}_1$ and $\boldsymbol{\Sigma} = \operatorname{Var}(\boldsymbol{Y}) = \sum_{i=2}^{k} \boldsymbol{X}_i\boldsymbol{X}_i'\sigma_i^2 + \sigma_e^2\boldsymbol{I}_N$, we obtain

$$E(\boldsymbol{Y}'\boldsymbol{Q}\boldsymbol{Y}) = (\boldsymbol{X}_1\boldsymbol{\beta}_1)'\boldsymbol{Q}(\boldsymbol{X}_1\boldsymbol{\beta}_1) + \sum_{i=2}^{k} \sigma_i^2 \operatorname{tr}(\boldsymbol{Q}\boldsymbol{X}_i\boldsymbol{X}_i') + \sigma_e^2 \operatorname{tr}(\boldsymbol{Q}). \tag{9.4.16}$$

9.4.4 EXPECTATION UNDER RANDOM EFFECTS

If $\boldsymbol{\beta}$ in (9.4.1) represents all random effects except μ, the result in (9.4.16) can be used to derive $E(\boldsymbol{Y}'\boldsymbol{Q}\boldsymbol{Y})$ under random effects, by simply letting $\boldsymbol{\beta}_1$ be the scalar μ and $\boldsymbol{X}_1$ a vector of 1s denoted by $\mathbf{1}$. Thus we have

$$E(\boldsymbol{Y}'\boldsymbol{Q}\boldsymbol{Y}) = \boldsymbol{\mu}^2\mathbf{1}'\boldsymbol{Q}\mathbf{1} + \sum_{i=2}^{k} \sigma_i^2 \operatorname{tr}(\boldsymbol{Q}\boldsymbol{X}_i\boldsymbol{X}_i') + \sigma_e^2 \operatorname{tr}(\boldsymbol{Q}). \tag{9.4.17}$$

EXERCISES

1. Prove results (i)–(v) of Theorem 9.1.1.
2. Prove results (i)–(ix) of Theorem 9.2.1.
3. Prove results (iii) and (iv) of Theorem 9.3.1.
4. Prove results (i) and (ii) of Theorem 9.3.2.
5. Prove Theorem 9.3.3.
6. Prove Theorem 9.3.4.
7. Prove Theorem 9.3.5.
8. Prove Theorem 9.3.6.

Bibliography

A. Albert (1972),*Regression and the Moore–Penrose Inverse*, Academic Press, New York.

R. L. Anderson (1961), Designs for estimating variance components, in *Proceedings of the Seventh Conference on Design of Experiments and Army Preservation and Development Testing*, 781–823; also published as Mim. Ser. 310, Institute of Statistics, North Carolina State University, Raleigh, NC.

R. L. Anderson and P. P. Crump (1967), Comparisons of designs and estimation procedures for estimating parameters in a two-stage nested process, *Technometrics*, **9**, 499–516.

T. R. Bainbridge (1963), Staggered nested designs for estimating variance components, in *American Society for Quality Control Annual Conference Transactions*, American Society for Quality Control, Milwaukee, 93–103.

R. B. Bapat and T. E. S. Raghavan (1997), *Nonnegative Matrices and Applications*, Cambridge University Press, Cambridge, UK.

A. Basilevsky (1983), *Applied Matrix Algebra in the Statistical Sciences*, North-Holland, Amsterdam.

A. Ben-Israel and T. Greyville (1974), *Generalized Inverses: Theory and Applications*, Wiley, New York.

A. Berman and R. J. Plemmons (1994), *Nonnegative Matrices in the Mathematical Sciences*, 2nd ed., Society for Industrial and Applied Mathematics, Philadelphia.

N. Bush and R. L. Anderson (1963), A comparison of three different procedures for estimating variance components, *Technometrics*, **5**, 421–440.

W. G. Cochran (1934), The distribution of quadratic forms in a normal system, *Proc. Cambridge Philos. Soc. Suppl.*, **4**, 102–118.

C. Eisenhart (1947), The assumptions underlying the analysis of variance, *Biometrics*, **3**, 1–21.

J. E. Gentle (1998), *Numerical Linear Algebra for Applications in Statistics*, Springer-Verlag, New York.

F. A. Graybill (1961), *An Introduction to Linear Statistical Models*, Vol. I, McGraw–Hill, New York.

F. A. Graybill (1983), *Introduction to Matrices with Applications in Statistics*, 2nd ed., Wadsworth, Belmont, CA; 1st ed., 1969.

A. S. Hadi (1996), *Matrix Algebra as a Tool*, Wadsworth, Belmont, CA.

D. A. Harville (1997), *Matrix Algebra from a Statistician's Perspective*, Springer-Verlag, New York.

M. J. R. Healy (2000), *Matrices for Statistics*, 2nd ed., Clarendon Press, Oxford, UK; 1st ed., 1986.

R. Horn and C. R. Johnson (1985), *Matrix Analysis*, Cambridge University Press, Cambridge, UK.

H. O. Lancaster (1954), Traces and cumulants of quadratic forms in normal variables, *J. Roy. Statist. Soc. Ser.* B, **16**, 247–254.

W. Madow (1940), The distribution of quadratic forms in noncentral normal random variables, *Ann. Math. Statist.*, **11**, 100–101.

J. R. Magnus and H. Neudecker (1999), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 2nd ed., Wiley, Chichester, UK.

H. D. Muse and R. L. Anderson (1978), Comparison of designs to estimate variance components in a two-way classification model, *Technometrics*, **20**, 159–166.

H. D. Muse, R. L. Anderson, and B. Thitakamol (1982), Additional comparisons of designs to estimate variance components in a two-way classification model, *Comm. Statist.* A *Theory Methods*, **11**, 1403–1425.

R. Pringle and A. Raynor (1971), *Generalized Inverse Matrices with Applications in Statistics*, Hafner, New York.

C. R. Rao and S. K. Mitra (1971), *Generalized Inverse of Matrices and Its Applications*, Wiley, New York.

C. R. Rao and M. B. Rao (1998), *Matrix Algebra and Its Applications to Statistics and Econometrics*, World Scientific, Singapore.

H. Scheffé (1959), *The Analysis of Variance*, Wiley, New York.

J. R. Schott (1997), *Matrix Analysis for Statistics*, Wiley, New York.

S. R. Searle (1971), *Linear Models*, Wiley, New York.

S. R. Searle (1982), *Matrix Algebra Useful for Statistics*, Wiley, New York.

S. R. Searle (1988), Mixed models and unbalanced data: Wherefrom, whereat, whereto?, *Comm. Statist.* A *Theory Methods*, **17**, 935–968.

E. Seneta (1981), *Non-Negative Matrices and Markov Chains*, 2nd ed., Springer-Verlag, New York.

P.-S. Shen, P. L. Cornelius, and R. L. Anderson (1996a), Planned unbalanced designs for estimation of quantitative genetic parameters I: Two-way matings, *Biometrics*, **52**, 56–70.

P.-S. Shen, P. L. Cornelius, and R. L. Anderson (1996b), Planned unbalanced designs for estimation of quantitative genetic parameters II, *J. Agricultural Biol. Environ. Sci.*, **1**, 490–505.

10 Some General Methods for Making Inferences about Variance Components

In the study of random and mixed effects models, our interest lies primarily in making inferences about the specific variance components. In this chapter, we consider some general methods for point estimation, confidence intervals, and hypothesis testing for linear models involving random effects. Most of the chapter is devoted to the study of various methods of point estimation of variance components. However, in the last two sections, we briefly address the problem of hypothesis testing and confidence intervals. There are now several methods available for estimation of variance components from unbalanced data. Henderson's (1953) paper can probably be characterized as the first attempt to systematically describe different adaptations of the ANOVA methodology for estimating variance components from unbalanced data. Henderson outlined three methods for obtaining estimators of variance components.

The first two methods are used for completely random models and the third method is most appropriate for a mixed model situation. The methods are basically moment estimation procedures where estimators are obtained by equating sample moments in the form of quadratic functions of observations to their respective expected values and the resulting equations are solved for the unknown variance components. Method I uses quadratic forms that are equivalent to analogous sums of squares obtained from the corresponding balanced analysis of variance; Method II is a variation of Method I that adjusts the data for the fixed effects in the model; and Method III uses reductions in sums of squares due to fitting different models and submodels.

The methods were critically reviewed and reformulated in elegant matrix notations by Searle (1968). Since then a bewildering variety of new procedures have been developed and the theory has been extended in a number of different directions. The principal developments include the adoption of an old and familiar method of maximum likelihood and its variant form, the so-called restricted maximum likelihood to the problem of variance components estimation. In addition, C. R. Rao (1970, 1971a, 1972) introduced the concept of minimum-norm quadratic unbiased estimation (MINQUE). Similarly, LaMotte (1973a) considered minimum-variance quadratic unbiased estimation and Pukelsheim (1981a, 1981b) has investigated the existence of nonnegative quadratic unbiased estimators using convex programming. Another interesting development is the least squares and the notion of quadratic subspace approach to estimate variance components used by Seely (1970a, 1970b, 1971) and the use of restricted generalized inverse operators such as given by Hartung (1981) who minimizes the bias subject to nonnegativity. We begin with a discussion of Henderson's procedures.

10.1 HENDERSON'S METHOD I

Of the three methods of Henderson, Method I is the easiest to compute and is probably the most frequently used method of estimation of variance components. The procedure involves evaluating sums of squares analogous to those used for the analysis of variance for balanced data. These are then equated to their respective expected values and solved for variance components. We illustrate the method in terms of the general linear model in (9.4.1) following closely the developments given in Searle (1971b, pp. 431–434). In subsequent chapters, we discuss the application of the method for special cases.

We write the general linear model in (9.4.1) as

$$\boldsymbol{Y} = \mu\mathbf{1} + \sum_{\theta=A}^{P} \boldsymbol{X}_\theta \beta_\theta + \boldsymbol{e}, \tag{10.1.1}$$

where

$$E(\boldsymbol{Y}) = \mu\mathbf{1}$$

and (10.1.2)

$$\operatorname{Var}(\boldsymbol{Y}) = \sum_{\theta=A}^{P} \boldsymbol{X}_\theta \operatorname{Var}(\beta_\theta)\boldsymbol{X}'_\theta + \sigma_e^2 \boldsymbol{I}_N.$$

Now, let $y_{.}(A_i)$ and $n(A_i)$ denote the total value and the number of observations in the ith level of the factor A. Then the raw sum of squares of the factor A is

$$T_A = \sum_{i=1}^{N_A} [y_{.}(A_i)]^2/n(A_i), \tag{10.1.3}$$

where N_A is the number of levels of the factor A. On ordering the elements in the observation vector $\boldsymbol{Y}$ appropriately, we can write

$$T_A = \boldsymbol{Y}'\boldsymbol{Q}_A\boldsymbol{Y}, \tag{10.1.4}$$

where

$$\boldsymbol{Q}_A = \sum_{i=1}^{N_A}{}^{+} \frac{1}{n(A_i)} \boldsymbol{J}_{n(A_i)}; \tag{10.1.5}$$

i.e., $\boldsymbol{Q}_A$ is the direct sum denoted by $\boldsymbol{\Sigma}^+$ (see, e.g., Appendix M) of N_A matrices $[1/n(A_i)]\boldsymbol{J}_{n(A_i)}$, $i = 1, 2, \ldots, N_A$.

On using the result in (9.4.17), we obtain

$$E(T_A) = N\mu^2 + \sum_{\theta=A}^{P} \left[\sum_{i=1}^{N_A} \frac{\sum_{j=1}^{N_\theta} [n(A_i, \theta_j)]^2}{n(A_i)} \sigma_\theta^2 \right] + N_A \sigma_e^2, \tag{10.1.6}$$

where $n(A_i, \theta_j)$ is the number of observations in the ith level of the factor A and the jth level of the factor θ. With appropriate definitions of $n(A_i, \theta_j)$, $n(A_i)$, and N_A, the result in (10.1.6) is generally applicable to any T in any random model. Thus, for T_0, the total sum of squares, it can be written as

$$E(T_0) = N\mu^2 + N \sum_{\theta=A}^{P} \sigma_\theta^2 + N\sigma_e^2; \tag{10.1.7}$$

and for T_μ, the correction factor for the mean, it is equal to

$$E(T_\mu) = N\mu^2 + \sum_{\theta=A}^{P} \left\{ \sum_{j=1}^{N_\theta} [n(\theta_j)]^2 \right\} \frac{\sigma_\theta^2}{N} + \sigma_e^2. \tag{10.1.8}$$

Thus the term $N\mu^2$ occurs in the expectation of every T. But since sums of squares (SSs) involve only differences between Ts, expectations of SSs do not contain $N\mu^2$, and their coefficients of σ_e^2 are equal to their corresponding degrees of freedom. Further, if the number of submost cells containing data in them is r, then the within-cell sum of squares SS_E has expectation given by

$$E(\text{SS}_E) = (N - r)\sigma_e^2. \tag{10.1.9}$$

Now, let

$\boldsymbol{S}$ = the vector of SSs, excluding SS_E,
$\boldsymbol{\sigma}^2$ = the vector of σ^2s, excluding σ_e^2,
$\boldsymbol{f}$ = the vector of degrees of freedom,

and

$\boldsymbol{R}$ = the matrix containing the elements of the coefficients of σ^2s excluding σ_e^2 in expectations of SSs.

Then the expected values of the SSs involved in any random effects model can be written as

$$E\begin{bmatrix} \boldsymbol{S} \\ \mathrm{SS}_E \end{bmatrix} = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{f} \\ \boldsymbol{0} & N-r \end{bmatrix}\begin{bmatrix} \boldsymbol{\sigma}^2 \\ \sigma_e^2 \end{bmatrix}. \tag{10.1.10}$$

Hence, the analysis of variance estimators (Henderson's Method I) of σ_e^2 and $\boldsymbol{\sigma}^2$ are

$$\hat{\sigma}_e^2 = \mathrm{SS}_E/(N-r)$$

and

$$\hat{\boldsymbol{\sigma}}^2 = \boldsymbol{R}^{-1}(\boldsymbol{S} - \hat{\sigma}_e^2 \boldsymbol{f}). \tag{10.1.11}$$

Note that the elements of $\boldsymbol{R}$ are of such a nature that there is no suitable form for expressing $\boldsymbol{R}^{-1}$ and so the estimators in (10.1.11) cannot be simplified any further. For any particular case, one first evaluates $\boldsymbol{R}$ using the relation

$$E(\boldsymbol{S}) = \boldsymbol{R}\boldsymbol{\sigma}^2 + \sigma_e^2 \boldsymbol{f}, \tag{10.1.12}$$

and then (10.1.11) are used to calculate the estimators.

In a random effects model, all variance components estimators obtained by Henderson's Method I are unbiased. For $\hat{\sigma}_e^2$ this result is quite obvious and for $\hat{\boldsymbol{\sigma}}^2$, we have

$$\begin{aligned} E(\hat{\boldsymbol{\sigma}}^2) &= \boldsymbol{R}^{-1}[E(\boldsymbol{S}) - \sigma_e^2 \boldsymbol{f}] \\ &= \boldsymbol{R}^{-1}[\boldsymbol{R}\boldsymbol{\sigma}^2 + \sigma_e^2 \boldsymbol{f} - \sigma_e^2 \boldsymbol{f}] \\ &= \boldsymbol{\sigma}^2. \end{aligned} \tag{10.1.13}$$

This property of unbiasedness generally holds for all estimators obtained from random effects models, but does not apply to estimators from mixed models (see e.g., Searle, 1971b, pp. 429–430). Further note that the method does not require the assumption of normality in order to obtain estimators. Recently, Westfall (1986) has shown that Henderson's Method I estimators of variance components in the nonnormal unbalanced hierarchical mixed model are asymptotically normal. In particular, Westfall (1986) provides conditions under which the ANOVA estimators from a nested mixed model have an asymptotic multivariate normal distribution.

10.2 HENDERSON'S METHOD II

Consider the general linear model in (9.4.1) written in the form

$$\boldsymbol{Y} = \mu \boldsymbol{1} + \boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{U}\boldsymbol{\beta} + \boldsymbol{e}, \tag{10.2.1}$$

where $\boldsymbol{\alpha}$ represents all the fixed effects except that the general constant μ and $\boldsymbol{\beta}$ represents all the random effects. Henderson's Method II consists of correcting the observation vector $\boldsymbol{Y}$ by an estimator $\hat{\boldsymbol{\alpha}} = \boldsymbol{L}\boldsymbol{Y}$ such that the corrected vector $\boldsymbol{Y}^* = \boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}}$ assumes the form (Searle, 1968)

$$\boldsymbol{Y}^* = \mu^*\mathbf{1} + \boldsymbol{U}\boldsymbol{\beta} + \boldsymbol{e}^*, \tag{10.2.2}$$

where μ^* is a new scalar and $\boldsymbol{e}^* = (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{L})\boldsymbol{e}$ is an error vector different from $\boldsymbol{e}$. Note that the structures of the random effects in both models (10.2.1) and (10.2.2) are the same. Now, the model equation in (10.2.2) represents a completely random model and Method I applied to $\boldsymbol{Y}^*$ will yield unbiased estimates of the variance components. It should be noticed that the crux of Method II lies in the choice of a matrix $\boldsymbol{L}$ such that the model equation in (10.2.1) is transformed to a completely random model in (10.2.2).

Given $\hat{\boldsymbol{\alpha}} = \boldsymbol{L}\boldsymbol{Y}$, from (10.2.1), the model equation for $\boldsymbol{Y}^*$ is

$$\boldsymbol{Y}^* = \mu(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{L})\mathbf{1} + (\boldsymbol{X} - \boldsymbol{X}\boldsymbol{L}\boldsymbol{X})\boldsymbol{\alpha} + (\boldsymbol{U} - \boldsymbol{X}\boldsymbol{L}\boldsymbol{U})\boldsymbol{\beta} + (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{L})\boldsymbol{e}. \tag{10.2.3}$$

It is immediately seen that (10.2.3) is free of the fixed effects if $\boldsymbol{X} = \boldsymbol{X}\boldsymbol{L}\boldsymbol{X}$, i.e., $\boldsymbol{L}$ is a generalized inverse of $\boldsymbol{X}$. Further, on comparing (10.2.2) and (10.2.3), it is evident that the model equation in (10.2.3) reduces to the form (10.2.2) if $\boldsymbol{X} = \boldsymbol{X}\boldsymbol{L}\boldsymbol{X}$, $\boldsymbol{X}\boldsymbol{L}\boldsymbol{U} = \mathbf{0}$, and $\mu(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{L})\mathbf{1} = \mu^*\mathbf{1}$. However, on a closer examination it is evident that the condition $\boldsymbol{X} = \boldsymbol{X}\boldsymbol{L}\boldsymbol{X}$ can be replaced by $(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{L})\boldsymbol{X}\boldsymbol{\alpha} = \lambda\mathbf{1}$ for some scalar λ. This way, λ and μ^* could be combined into a single general constant and (10.2.3) will be reduced to the form (10.2.2). Therefore, in order that the corrected vector $\boldsymbol{Y}^*$ be given by the model equation in (10.2.2), the matrix $\boldsymbol{L}$ should be chosen such that

(i) $\boldsymbol{X}\boldsymbol{L}\boldsymbol{U} = \mathbf{0}$; (10.2.4)

(ii) $\boldsymbol{X}\boldsymbol{L}\mathbf{1} = \lambda^*\mathbf{1}$ for some scalar λ^*, i.e., all row totals are the same; (10.2.5)

and

(iii) $\boldsymbol{X} - \boldsymbol{X}\boldsymbol{L}\boldsymbol{X} = \mathbf{1}\boldsymbol{\tau}'$ for some column vector $\boldsymbol{\tau}$, i.e., all the rows of $\boldsymbol{X} - \boldsymbol{X}\boldsymbol{L}\boldsymbol{X}$ are the same. (10.2.6)

Henderson's Method II chooses $\boldsymbol{L}$ in $\hat{\boldsymbol{\alpha}} = \boldsymbol{L}\boldsymbol{Y}$ such that the conditions (10.2.4), (10.2.5), and (10.2.6) are satisfied.

For a detailed discussion of the choice of $\boldsymbol{L}$ and its calculation, the reader is referred to Searle (1968), Henderson et al. (1974), and Searle et al. (1992, pp. 192–196). It should, however, be pointed out that Method II cannot be used on data from models that include interactions between the fixed and random effects. The reason being that the presence of such interactions is inconsistent with the conditions (10.2.4), (10.2.5), and (10.2.6). For a proof of this result, see Searle (1968) and Searle et al. (1992, pp. 199–201).

10.3 HENDERSON'S METHOD III

The procedure known as Henderson's Method III uses reductions in sums of squares due to fitting constants (due to fitting different models and submodels) in place of the analysis of variance sums of squares used in Methods I and II using a complete least squares analysis. Thus it is also commonly referred to as the method of fitting constants. We have seen that for fixed effects, having normal equations

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}'\boldsymbol{Y},$$

the reduction in sum of squares due to $\boldsymbol{\beta}$, denoted by $R(\boldsymbol{\beta})$, is

$$R(\boldsymbol{\beta}) = \boldsymbol{Y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{Y}. \tag{10.3.1}$$

In Method III, the reductions in sums of squares are calculated for a variety of submodels of the model under consideration, which may be either a random or a mixed model. Then the variance components are estimated by equating each computed reduction in sum of squares to its expected value under the full model, and solving the resultant equations for the variance components.

We illustrate the procedure in terms of the general linear model (9.1.1), following closely the developments given in Searle (1971a, Section 10.4). We first rewrite the model as

$$\boldsymbol{Y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{e}, \tag{10.3.2}$$

where $\boldsymbol{\beta}' = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')$, without any consideration as to whether they represent fixed or random effects. At the present, we are only interested in finding the expected values of the reductions in sum of squares due to fitting the model in (10.3.2) and the submodel (or the reduced model)

$$\boldsymbol{Y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{e}, \tag{10.3.3}$$

where both expectations are taken under the full model in (10.3.2).

Now, first we will find the value of $E(\boldsymbol{Y}'\boldsymbol{Q}\boldsymbol{Y})$, where the vector $\boldsymbol{Y}$ is given by (10.3.2). Using result (i) of Theorem 9.3.1, we have

$$E(\boldsymbol{Y}'\boldsymbol{Q}\boldsymbol{Y}) = E(\boldsymbol{Y}')\boldsymbol{Q}\boldsymbol{E}(\boldsymbol{Y}) + \text{tr}[\boldsymbol{Q}\,\text{Var}(\boldsymbol{Y})]. \tag{10.3.4}$$

For the model in (10.3.2), we obtain

$$\begin{aligned} E(\boldsymbol{Y}) &= \boldsymbol{X}_1 E(\boldsymbol{\beta}_1) + \boldsymbol{X}_2 E(\boldsymbol{\beta}_2) = (\boldsymbol{X}_1 \vdots \boldsymbol{X}_2)\begin{bmatrix} E(\boldsymbol{\beta}_1) \\ E(\boldsymbol{\beta}_2) \end{bmatrix} \\ &= \boldsymbol{X}E(\boldsymbol{\beta}), \end{aligned} \tag{10.3.5}$$

and

$$\text{Var}(\boldsymbol{Y}) = \boldsymbol{X}_1\,\textbf{Var}(\boldsymbol{\beta}_1)\boldsymbol{X}_1' + \boldsymbol{X}_2\,\text{Var}(\boldsymbol{\beta}_2)\boldsymbol{X}_2' + \sigma_e^2\boldsymbol{I}_N$$

$$= (\boldsymbol{X}_1 \vdots \boldsymbol{X}_2) \begin{bmatrix} \text{Var}(\boldsymbol{\beta}_1) & \\ & \text{Var}(\boldsymbol{\beta}_2) \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_1' \\ \boldsymbol{X}_2' \end{bmatrix} + \sigma_e^2 \boldsymbol{I}_N$$

$$= \boldsymbol{X}\,\text{Var}(\boldsymbol{\beta})\boldsymbol{X}' + \sigma_e^2 \boldsymbol{I}_N. \tag{10.3.6}$$

Substituting (10.3.5) and (10.3.6) into (10.3.4), we obtain

$$\begin{aligned} E(\boldsymbol{Y}'\boldsymbol{Q}\boldsymbol{Y}) &= E(\boldsymbol{\beta}')\boldsymbol{X}'\boldsymbol{Q}\boldsymbol{X}\boldsymbol{E}(\boldsymbol{\beta}) + \text{tr}[\boldsymbol{Q}\{\boldsymbol{X}\,\text{Var}(\boldsymbol{\beta})\boldsymbol{X}' + \sigma_e^2 \boldsymbol{I}_N\}] \\ &= \text{tr}[\boldsymbol{X}'\boldsymbol{Q}\boldsymbol{X}E(\boldsymbol{\beta}\boldsymbol{\beta}')] + \sigma_e^2\,\text{tr}(\boldsymbol{Q}). \end{aligned} \tag{10.3.7}$$

Result (10.3.7) is true, irrespective of whether $\boldsymbol{\beta}$ is fixed or random.

Now, let $R(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$ be the reduction in sum of squares due to fitting (10.3.2). Then from (10.3.1), we have

$$R(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = \boldsymbol{Y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{Y}. \tag{10.3.8}$$

Taking the expectation of (10.3.8) by using (10.3.7) with $\boldsymbol{Q} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'$ gives

$$\begin{aligned} E\{R(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)\} &= \text{tr}[\boldsymbol{X}'\boldsymbol{X}E(\boldsymbol{\beta}\boldsymbol{\beta}')] + \sigma_e^2\,\text{rank}(\boldsymbol{X}) \\ &= \text{tr}\left\{\begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{X}_1 & \vdots & \boldsymbol{X}_1'\boldsymbol{X}_2 \\ \cdots & & \cdots \\ \boldsymbol{X}_2'\boldsymbol{X}_1 & \vdots & \boldsymbol{X}_2'\boldsymbol{X}_2 \end{bmatrix} E(\boldsymbol{\beta}\boldsymbol{\beta}')\right\} + \sigma_e^2\,\text{rank}(\boldsymbol{X}). \end{aligned} \tag{10.3.9}$$

Similarly, let $R(\boldsymbol{\beta}_1)$ be the reduction in sum of squares due to fitting the reduced model in (10.3.3). Then

$$R(\boldsymbol{\beta}_1) = \boldsymbol{Y}'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^-\boldsymbol{X}_1'\boldsymbol{Y}. \tag{10.3.10}$$

Again, taking the expectation of (10.3.10) under the full model by using (10.3.7) with $\boldsymbol{Q} = \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^-\boldsymbol{X}_1'$ gives

$$\begin{aligned} E\{R(\boldsymbol{\beta}_1)\} &= \text{tr}\{\boldsymbol{X}'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^-\boldsymbol{X}_1'\boldsymbol{X}E(\boldsymbol{\beta}\boldsymbol{\beta}')\} + \sigma_e^2\,\text{rank}[\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^-\boldsymbol{X}_1'] \\ &= \text{tr}\left\{\begin{bmatrix} \boldsymbol{X}_1' & \vdots & \boldsymbol{X}_1 \\ \cdots & & \cdots \\ \boldsymbol{X}_2' & \vdots & \boldsymbol{X}_2 \end{bmatrix} (\boldsymbol{X}_1'\boldsymbol{X}_1)^-[\boldsymbol{X}_1'\boldsymbol{X}_1 \vdots \boldsymbol{X}_1'\boldsymbol{X}_2]E(\boldsymbol{\beta}\boldsymbol{\beta}')\right\} \\ &\quad + \sigma_e^2\,\text{rank}(\boldsymbol{X}_1) \\ &= \text{tr}\left\{\begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{X}_1 & \vdots & \boldsymbol{X}_1'\boldsymbol{X}_2 \\ \cdots & & \cdots \\ \boldsymbol{X}_2'\boldsymbol{X}_1 & \vdots & \boldsymbol{X}_2'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^-\boldsymbol{X}_1'\boldsymbol{X}_2 \end{bmatrix} E(\boldsymbol{\beta}\boldsymbol{\beta}')\right\} \\ &\quad + \boldsymbol{\sigma}_e^2\,\text{rank}(\boldsymbol{X}_1). \end{aligned} \tag{10.3.11}$$

Hence, the expected value of the difference between the reductions (10.3.8) and (10.3.10), known as the reduction due to $\boldsymbol{\beta}_2$ after adjusting for $\boldsymbol{\beta}_1$ and denoted by $R(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)$, is

$$
\begin{aligned}
E\{R(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)\} &= E\{R(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)\} - E\{R(\boldsymbol{\beta}_1)\} \\
&= \operatorname{tr}\left\{\begin{bmatrix} \mathbf{0} & \vdots & \mathbf{0} \\ & & \cdots \\ \mathbf{0} & \vdots & \boldsymbol{X}_2'[\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-}\boldsymbol{X}_1']\boldsymbol{X}_2 \end{bmatrix}\right. \\
&\qquad \left.\times \begin{bmatrix} E(\boldsymbol{\beta}_1\boldsymbol{\beta}_1') & \vdots & E(\boldsymbol{\beta}_1\boldsymbol{\beta}_2') \\ E(\boldsymbol{\beta}_2\boldsymbol{\beta}_1') & \vdots & E(\boldsymbol{\beta}_2\boldsymbol{\beta}_2') \end{bmatrix}\right\} \\
&\qquad + \sigma_e^2[\operatorname{rank}(\boldsymbol{X}) - \operatorname{rank}(\boldsymbol{X}_1)] \\
&= \operatorname{tr}\{\boldsymbol{X}_2'[\boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-}\boldsymbol{X}_1']\boldsymbol{X}_2 E(\boldsymbol{\beta}_2\boldsymbol{\beta}_2')\} \\
&\qquad + \sigma_e^2[\operatorname{rank}(\boldsymbol{X}) - \operatorname{rank}(\boldsymbol{X}_1)]. \qquad (10.3.12)
\end{aligned}
$$

It should be noted that (10.3.12) is a function only of $E(\boldsymbol{\beta}_2\boldsymbol{\beta}_2')$ and σ_e^2 and has been derived without any assumption on the form of $E(\boldsymbol{\beta}\boldsymbol{\beta}')$.

Result (10.3.12) states that if the vector $\boldsymbol{\beta}$ is partitioned as $(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$, where $\boldsymbol{\beta}_1$ represents all the fixed effects and $\boldsymbol{\beta}_2$ represents all the random effects, then $\boldsymbol{E}\{R(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)\}$ contains only σ_e^2 and the variance components associated with the random effects; it contains no terms due to the fixed effects. Thus, in a mixed effects model, Henderson's Method III yields unbiased estimates of the variance components unaffected by the fixed effects. Moreover, in a completely random model, where $\boldsymbol{\beta}_1$ also contains only random effects, $\boldsymbol{E}\{R(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)\}$ does not contain any variance components associated with $\boldsymbol{\beta}_1$; nor does it contain any covariance terms between the elements of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$. Thus, even for completely random models where $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are correlated, the method provides unbiased estimates unaffected by any correlative terms.

Note that in comparison to Methods I and II, Method III is more appropriate for the mixed model, in which case it yields unbiased estimates of the variance components free of any fixed effects. Its principal drawback is that it involves computing generalized inverses of matrices of very large dimensions in cases when the model contains a large number of effects. In addition, the method suffers from the lack of uniqueness since it can give rise to more quadratics than there are components to be estimated. For a more thorough and complete treatment of Henderson's Method III, see Searle et al. (1992, Section 5.5).

Rosenberg and Rhode (1971) have investigated the consequences of estimating variance components using the method of fitting constants when the hypothesized random model contains factors which do not belong in the true model. They have derived variance components estimators and their expectations and variances under both the true and the hypothesized model.

Remarks:

(i) An alternative formulation of Henderson's Method III can be given as follows (Verdooren, 1980). Consider the general linear model in (9.4.1) in the following form:

$$\boldsymbol{Y} = \boldsymbol{X\alpha} + \boldsymbol{U}_1\boldsymbol{\beta}_1 + \boldsymbol{U}_2\boldsymbol{\beta}_2 + \cdots + \boldsymbol{U}_p\boldsymbol{\beta}_p,$$

where

$\boldsymbol{X}$ is an $N \times q$ matrix of known fixed numbers, $q \leq N$,
$\boldsymbol{U}_i$ is an $N \times m_i$ matrix of known fixed numbers, $m_i \leq N$,
$\boldsymbol{\alpha}$ is a q-vector of fixed effects, and
$\boldsymbol{\beta}_i$ is an m_i-vector of random effects.

We further assume that $E(\boldsymbol{\beta}_i) = \boldsymbol{0}$, $\boldsymbol{\beta}_i$s are uncorrelated, and $E(\boldsymbol{\beta}_i\boldsymbol{\beta}_i') = \sigma_i^2\boldsymbol{I}_{\boldsymbol{m}_i}$. The assumptions imply that $\text{Var}(\boldsymbol{Y}) = \sum_{i=1}^p \sigma_i^2\boldsymbol{U}_i\boldsymbol{U}_i' = \sum_{i=1}^p \sigma_i^2\boldsymbol{V}_i$ where $\boldsymbol{V}_i = \boldsymbol{U}_i\boldsymbol{U}_i'$. Let $\boldsymbol{P}_i$ $(i = 1, 2, \ldots, p)$ be the orthogonal projection operator on the column space of $(\boldsymbol{X}, \boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_i)$. Note that $\boldsymbol{P}_p = \boldsymbol{I}_N$. Let $\boldsymbol{P}_0$ be the orthogonal projection operator on the column space of $\boldsymbol{X}$, i.e., $\boldsymbol{P}_0 = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$. Finally, let $\boldsymbol{Q}_j$ be the orthogonal projection on the orthogonal complement of the column space of $(\boldsymbol{X}, \boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_{j-1})$ (for $j = 1$, the column space of $\boldsymbol{X}$). Note that $\boldsymbol{Q}_j = \boldsymbol{P}_j - \boldsymbol{P}_{j-1} (j = 1, 2, \ldots, p)$ and $\boldsymbol{Q}_p = \boldsymbol{P}_p - \boldsymbol{P}_{p-1} = \boldsymbol{I}_N - \boldsymbol{P}_{p-1}$. Now, consider the following orthogonal decomposition of $\boldsymbol{Y}$:

$$\boldsymbol{Y} = \boldsymbol{P}_0\boldsymbol{Y} + \sum_{j=1}^p \boldsymbol{Q}_j\boldsymbol{Y},$$

which implies that

$$\boldsymbol{Y}'\boldsymbol{Y} = \boldsymbol{Y}'\boldsymbol{P}_0\boldsymbol{Y} + \sum_{j=1}^p \boldsymbol{Y}'\boldsymbol{Q}_j\boldsymbol{Y}.$$

Here, $\boldsymbol{P}_0\boldsymbol{Y}$ can be used as an estimator of $\boldsymbol{\alpha}$ and $\boldsymbol{Y}'\boldsymbol{Q}_j\boldsymbol{Y}$s can be used to yield unbiased estimators of σ_i^2s $(i = 1, 2, \ldots, p)$. Applying Theorem 9.3.1, and noting that $\boldsymbol{Q}_j\boldsymbol{X} = \boldsymbol{0}$, $\boldsymbol{Q}_j\boldsymbol{U}_i = \boldsymbol{0}$ for $i < j$, we have $E(\boldsymbol{Y}'\boldsymbol{Q}_j\boldsymbol{Y}) = \sum_{i=1}^p \sigma_i^2 \,\text{tr}(\boldsymbol{Q}_j\boldsymbol{V}_i) = \sum_{i=j}^p \sigma_i^2 \,\text{tr}(\boldsymbol{Q}_j\boldsymbol{V}_i)$. Now, Henderson's Method III consists of the hierarchical setup of the quadratic forms $(\boldsymbol{Y}'\boldsymbol{Q}_j\boldsymbol{Y})$ and by solving the following system of linear equations: $\boldsymbol{Y}'\boldsymbol{Q}_1\boldsymbol{Y} = \sum_{i=1}^p \sigma_i^2 \,\text{tr}(\boldsymbol{Q}_1\boldsymbol{V}_i), \boldsymbol{Y}'\boldsymbol{Q}_2\boldsymbol{Y} = \sum_{i=2}^p \sigma_i^2 \,\text{tr}(\boldsymbol{Q}_2\boldsymbol{V}_i), \ldots,$ $\boldsymbol{Y}'\boldsymbol{Q}_p\boldsymbol{Y} = \sigma_p^2 \,\text{tr}(\boldsymbol{Q}_p\boldsymbol{V}_p)$. Note that the procedure depends on the order of the $\boldsymbol{U}_j'$s in the definition of the projection operators $\boldsymbol{P}_i'$s.

(ii) For completely nested random models, Henderson's Methods I, II, and III reduce to the customary analysis of variance procedure.

(iii) A general procedure for the calculation of expected mean squares for the analysis of variance based on least squares fitting constants quadratics using the Abbreviated Doolittle and Square Root methods has been given by Gaylor et al. (1970). ♦

Finally, it should be noted that Henderson's methods may produce negative estimates. Khattree (1998, 1999) proposed some simple modifications of Henderson's procedures which ensure the nonnegativity of the estimates. The modifications entail seeking nonnegative estimates to Henderson's solution that are closest to the expected values of the quadratics being used for estimation. The resulting estimators are found to be superior in terms of various comparison criteria to Henderson's estimators except in the case of the error variance component.

10.4 ANALYSIS OF MEANS METHOD

In a fixed effects model, when data in every cell or subclass of the model contain at least one observation, an easily calculated analysis is to consider the means of these cells as individual observations and perform a balanced data analysis in terms of the means of the submost subclasses. The analysis can be based on the sums of squares of the means (unweighted), or can be performed by weighting the terms of the sums of squares in inverse proportion to the variance of the term concerned (weighted). The analysis was originally proposed by Yates (1934) and provides a simple and efficient method of analyzing data from experimental situations having unbalanced design structure with no empty cells.

The mean squares of these analyses (weighted or unweighted) can then be used for estimating variance components in random as well as mixed models. Estimators of the variance components are obtained in the usual manner of equating the mean squares to their expected values and solving the resulting equations for the variance components. The estimators, thus obtained, are unbiased. This is, of course, only an approximate procedure, with the degree of approximation depending on the extent to which the unbalanced data are not balanced. Several authors have investigated the adequacy of the unweighted mean squares empirically under various degrees of imbalance (see, e.g., Gosslee and Lucas, 1965; Hartwell and Gaylor, 1973; Knoke, 1985; Elliott, 1989). It has been found that their performance is fairly adequate except in cases of extreme imbalance and for certain values of the variance components for the models under consideration (see, e.g., Thomas and Hultquist, 1978; Burdick et al., 1986; Hernández et al., 1992). The use of the procedure is illustrated in subsequent chapters for certain specific experimental situations.

In recent years, unweighted cell means and unweighted means estimators have been used and studied by a number of authors, including Burdick and Graybill (1984), Tan et al. (1988), and Khuri (1990). Thomsen (1975) and Khuri and Littel (1987) have used unweighted cell means to test hypotheses that variance components are zero. Hocking et al. (1989) show that the unweighted means estimators reduce to a simple form that permits diagnostic analysis which can detect any problem with data and violations of model assumptions. Westfall and Bremer (1994) have made analytic investigation of some efficiency properties of the unweighted means estimators in general r-way unbalanced mixed models. They have shown that the efficiency approaches 1 when certain design parameters are increased, or when certain variance components become large.

10.5 SYMMETRIC SUMS METHOD

Koch (1967a) suggested a method of estimating variance components which utilizes symmetric sums of products (SSP) of the observations instead of sums of squares. In a variance component model, expected values of products of observations can be expressed as linear functions of the variance components. Hence, estimates of variance components can be obtained in terms of sums or means of these products. The resulting estimators are unbiased and consistent, and they are identical to the analysis of variance estimators for balanced data. However, for certain unbalanced experiments, the estimates obtained in this manner have an undesirable property that they may change in value if the same constant is added to all the observations, and their variances are functions of the general mean μ. This difficulty is overcome by Koch (1968), who suggested a modification of the above method to obtain estimators of the variance components, which are invariant under changes in location of the data. In the modified procedure, instead of using symmetric sums of products, symmetric sums of squares of differences are used.

Forthofer and Koch (1974) have extended the SSP method of the estimation of variance components to the general mixed model. Here, we illustrate the SSP method for the degenerate or one-stage design. In subsequent chapters, we consider the application of the method for specific experimental situations. To illustrate the SSP method for the degenerate or one-stage design, let the observations y_is $(i = 1, 2, \ldots, n)$ be given by the model

$$y_i = \mu + e_i, \tag{10.5.1}$$

where e_is are assumed to be independent and identically distributed random variables with mean zero and variance σ_e^2. Now, the expected values of the products of observations y_is from the model in (10.5.1) are

$$E(y_i y_j) = \begin{cases} \mu^2 + \sigma_e^2 & \text{if } i = j, \\ \mu^2 & \text{if } i \neq j. \end{cases} \tag{10.5.2}$$

The estimator of σ_e^2 is derived by taking means of the different products in (10.5.2). Thus we get

$$\hat{\mu}^2 = \frac{\sum\limits_{\substack{i=1 \\ i \neq j}}^{n} \sum\limits_{j=1}^{n} y_i y_j}{(n^2 - n)}$$

$$= \frac{\left(\sum_{i=1}^n y_i\right)^2 - \sum_{i=1}^n y_i^2}{n(n-1)}, \tag{10.5.3}$$

and

$$\hat{\mu}^2 + \hat{\sigma}_e^2 = \frac{\sum_{i=1}^n y_i^2}{n}. \tag{10.5.4}$$

Therefore, the estimator of σ_e^2 is given by

$$\hat{\sigma}_e^2 = \frac{\sum_{i=1}^n y_i^2}{n} - \hat{\mu}^2$$

$$= \frac{\sum_{i=1}^n y_i^2}{n} - \frac{\left(\sum_{i=1}^n y_i\right)^2 - \sum_{i=1}^n y_i^2}{n(n-1)}$$

$$= \frac{n\sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i\right)^2}{n(n-1)}$$

$$= \frac{\sum_{i=1}^n (y_i - \bar{y}_.)^2}{n-1}, \tag{10.5.5}$$

where

$$\bar{y}_. = \frac{\sum_{i=1}^n y_i}{n}.$$

Thus, in this case, the estimation procedure leads to the usual unbiased estimator of σ_e^2.

Using symmetric sums of squares of differences, we get

$$E(y_i - y_j)^2 = \begin{cases} 2\sigma_e^2 & \text{if } i \neq j, \\ 0 & \text{if } i = j. \end{cases} \tag{10.5.6}$$

By taking the means of the symmetric sums in (10.5.6), we obtain

$$2\hat{\sigma}_e^2 = \frac{\sum\limits_{\substack{i=1 \\ i \neq j}}^{n} \sum\limits_{j=1}^{n} y_i y_j}{(n^2 - n)}$$

$$= \frac{\sum_{i=1}^n \sum_{j=1}^n (y_i - y_j)^2}{n(n-1)}$$

$$= \frac{2}{(n-1)} \left(\sum_{i=1}^{n} y_i^2 - n\bar{y}_.^2 \right),$$

where

$$\bar{y}_. = \frac{\sum_{i=1}^{n} y_i}{n}.$$

Therefore, the estimator of σ_e^2 is given by

$$\hat{\sigma}_e^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y}_.)^2}{n-1}. \qquad (10.5.7)$$

Again, the procedure leads to the usual unbiased estimator of σ_e^2.

10.6 ESTIMATION OF POPULATION MEAN IN A RANDOM EFFECTS MODEL

In many random effects models, it is often of interest to estimate the population mean μ. For balanced data, as we have seen in Volume I, the "best" estimator of μ is the ordinary sample mean. However, for unbalanced data, the choice of a best estimator of μ is not that obvious. We have seen in Section 10.5 that the SSP method involved the construction of an unbiased estimate of the square of the population mean. As proposed by Koch (1967b), this procedure can be used to obtain an unbiased estimate of the mean itself by proceeding as follows.

Suppose that the unbiased estimator of μ^2 obtained by the SSP method is $Q(\boldsymbol{Y})$, a quadratic function of the observations, that is,

$$\hat{\mu}^2 = Q(\boldsymbol{Y}), \qquad (10.6.1)$$

where

$$E(\hat{\mu}^2) = E\{Q(\boldsymbol{Y})\} = \mu^2.$$

Now, consider the set of transformations on the data obtained by adding a constant θ to each observation. After making such a transformation, the SSP method is used to obtain the unbiased estimator of the square of the population mean of the transformed data, which will have the form $Q(\boldsymbol{Y} + \theta\mathbf{1})$. Then it follows that

$$\begin{aligned} \widehat{(\mu + \theta)}^2 &= Q(\boldsymbol{Y} + \theta\mathbf{1}) \\ &= Q(\boldsymbol{Y}) + 2G(\boldsymbol{Y})\theta + \theta^2, \end{aligned} \qquad (10.6.2)$$

where $G(\boldsymbol{Y})$ is a linear function of the observations.

Now, the function (10.6.2) is minimized as a function of θ when

$$\theta = -G(\boldsymbol{Y})$$

and the minimum value of (10.6.2) is

$$\begin{aligned}(\hat{\mu} - G(\boldsymbol{Y}))^2 &= Q(\boldsymbol{Y}) - \{G(\boldsymbol{Y})\}^2 \\ &= \hat{\mu}^2 - \{G(\boldsymbol{Y})\}^2.\end{aligned}$$

This suggests the estimator of the population mean as

$$\hat{\mu} = G(\boldsymbol{Y}). \tag{10.6.3}$$

It is easily shown that (10.6.3) is an unbiased estimator of μ. Thus, from (10.6.2), we have

$$\hat{\mu} = G(\boldsymbol{Y}) = [Q(\boldsymbol{Y} + \theta\mathbf{1}) - Q(\boldsymbol{Y}) - \theta^2]/2\theta,$$

so that

$$\begin{aligned}E(\hat{\mu}) &= [E\{Q(\boldsymbol{Y} + \theta\mathbf{1})\} - E\{Q(\boldsymbol{Y})\} - \theta^2]/2\theta \\ &= [(\mu + \theta)^2 - \mu^2 - \theta^2]/2\theta \\ &= \mu.\end{aligned}$$

We now illustrate the procedure for the degenerate or one-stage design. In subsequent chapters, we consider the application of the method for other experimental situations.

Let the observations y_is ($i = 1, 2, \ldots, n$) be given by the model

$$y_i = \mu + e_i, \tag{10.6.4}$$

where e_is are assumed to be independent and identically distributed random variables with mean zero and variance σ_e^2. From (10.5.3), we have

$$\hat{\mu}^2 = \frac{\left(y_.^2 - \sum_{i=1}^n y_i^2\right)}{n(n-1)}$$

where

$$y_. = \sum_{i=1}^n y_i.$$

Now, proceeding as described above, we obtain

$$\widehat{(\mu + \theta)}^2 = \frac{\left\{(y_. + n\theta)^2 - \sum_{i=1}^n (y_i + \theta)^2\right\}}{n(n-1)},$$

$$= \hat{\mu}^2 + 2\bar{y}_. \theta + \theta^2, \tag{10.6.5}$$

where

$$\bar{y}_. = \frac{y_.}{n}.$$

Hence, comparing (10.6.2) and (10.6.5), the desired estimate of μ is

$$\hat{\mu} = G(\boldsymbol{Y}) = \bar{y}_.$$

Thus, in this case, the estimate coincides with the usual sample mean.

10.7 MAXIMUM LIKELIHOOD ESTIMATION

Maximum likelihood (ML) equations for estimating variance components from unbalanced data cannot be solved explicitly. Thus, for unbalanced designs, explicit expressions for the ML estimators of variance components cannot be found in general and solutions have to obtained using some iterative procedures. The application of maximum likelihood estimation to the variance components problem in a general mixed model has been considered by Hartley and Rao (1967) and Miller (1977, 1979), among others. Hartley and Rao (1967) have developed a general set of equations, from which specific estimates can be obtained by iteration, involving extensive numerical computations. In this section, we consider the Hartley–Rao procedure of ML estimation and derive large sample variances of the ML estimators of variance components.

10.7.1 HARTLEY–RAO ESTIMATION PROCEDURE

We write the general linear model in (9.4.1) in the following form:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{U}_1\boldsymbol{\beta}_1 + \cdots + \boldsymbol{U}_p\boldsymbol{\beta}_p + \boldsymbol{e}, \tag{10.7.1}$$

where

$\boldsymbol{X}$ is an $N \times q$ matrix of known fixed numbers, $q \leq N$;
$\boldsymbol{U}_i$ is an $N \times m_i$ matrix of known fixed numbers, $m_i \leq N$;
$\boldsymbol{\alpha}$ is a q-vector of fixed effects;
$\boldsymbol{\beta}_i$ is an m_i-vector of random effects;

and

$\boldsymbol{e}$ is an N-vector of error terms.

We assume that the matrices $\boldsymbol{X}$ and $\boldsymbol{U}_i$, known as incidence or design matrices, are all of full rank; i.e., the rank of $\boldsymbol{X}$ is q and the rank of $\boldsymbol{U}_i$ is m_i. We further assume that $\boldsymbol{\beta}_i$ and $\boldsymbol{e}$ have multivariate normal distributions with mean vectors zero and variance-covariance matrices $\sigma_i^2 \boldsymbol{I}_{m_i}$ and $\sigma_e^2 \boldsymbol{I}_N$, respectively. Here, σ_i^2 $(i = 1, 2, \ldots, p)$ and σ_e^2 are the unknown variance components and the problem is to find their ML estimates.

From (10.7.1), it follows that

$$\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{V}),$$

where

$$\mu = E(\boldsymbol{Y}) = \boldsymbol{X\alpha}$$

and

$$\boldsymbol{V} = \operatorname{Var}(\boldsymbol{Y}) = \sum_{i=1}^{p} \sigma_i^2 \boldsymbol{U}_i \boldsymbol{U}_i' + \sigma_e^2 \boldsymbol{I}_N = \sigma_e^2 \boldsymbol{H}, \qquad (10.7.2)$$

with

$$\boldsymbol{H} = \sum_{i=1}^{p} \rho_i \boldsymbol{U}_i \boldsymbol{U}_i' + \boldsymbol{I}_N$$

and

$$\rho_i = \sigma_i^2 / \sigma_e^2.$$

Hence, the likelihood function is given by

$$L = \frac{\exp\left\{-\frac{1}{2\sigma_e^2}(\boldsymbol{Y} - \boldsymbol{X\alpha})' \boldsymbol{H}^{-1} (\boldsymbol{Y} - \boldsymbol{X\alpha})\right\}}{(2\pi\sigma_e^2)^{\frac{1}{2}N} |\boldsymbol{H}|^{\frac{1}{2}}}$$

and the natural logarithm of the likelihood is

$$\ell n L = -\frac{1}{2} N \ell n(2\pi) - \frac{1}{2} N \ell n \sigma_e^2 - \frac{1}{2} \ell n |\boldsymbol{H}| - \frac{1}{2\sigma_e^2} (\boldsymbol{Y} - \boldsymbol{X\alpha})' \boldsymbol{H}^{-1} (\boldsymbol{Y} - \boldsymbol{X\alpha}). \qquad (10.7.3)$$

Equating to zero the partial derivatives of (10.7.3) with respect to $\boldsymbol{\alpha}$, σ_e^2, and ρ_i yields

$$\frac{\partial \ell n L}{\partial \boldsymbol{\alpha}} = \frac{1}{\sigma_e^2} (\boldsymbol{X}' \boldsymbol{H}^{-1} \boldsymbol{Y} - \boldsymbol{X}' \boldsymbol{H}^{-1} \boldsymbol{X\alpha}) = \boldsymbol{0}, \qquad (10.7.4)$$

$$\frac{\partial \ell n L}{\partial \sigma_e^2} = -\frac{1}{2\sigma_e^2} N + \frac{1}{2\sigma_e^4} (\boldsymbol{Y} - \boldsymbol{X\alpha})' \boldsymbol{H}^{-1} (\boldsymbol{Y} - \boldsymbol{X\alpha}) = 0, \qquad (10.7.5)$$

and

$$\frac{\partial \ell n L}{\partial \rho_i} = -\frac{1}{2}\operatorname{tr}(\boldsymbol{H}^{-1}\boldsymbol{U}_i\boldsymbol{U}_i') + \frac{1}{2\sigma_e^2}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha})'\boldsymbol{H}^{-1}\boldsymbol{U}_i\boldsymbol{U}_i'\boldsymbol{H}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha})$$
$$= 0. \qquad (10.7.6)$$

Equations (10.7.4), (10.7.5), and (10.7.6) have to be solved for the elements of $\boldsymbol{\alpha}$, σ_e^2, and the ρ_is contained in $\boldsymbol{H}$ with the constraints that the σ_e^2 and ρ_is be nonnegative. Hartley and Rao (1967) indicate how this can be achieved, either by the method of steepest ascent or by obtaining an alternative form of (10.7.6), which are difficult equations to handle. The difficulty arises because the ML equations may yield multiple roots or the ML estimates may be on the boundary points. Equations (10.7.4) and (10.7.5) can be readily solved in terms of ρ_is. Thus we obtain

$$\hat{\boldsymbol{\alpha}} = (\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{Y} \qquad (10.7.7a)$$

and

$$\hat{\sigma}_e^2 = \frac{1}{N}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}})'\boldsymbol{H}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}})$$
$$= \frac{1}{N}[\boldsymbol{Y}'\boldsymbol{H}^{-1}\boldsymbol{Y} - (\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{Y})'(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{Y})]. \qquad (10.7.7b)$$

On substituting (10.7.7a) and (10.7.7b) into (10.7.6), we obtain the equation

$$\operatorname{tr}(\boldsymbol{H}^{-1}\boldsymbol{U}_i\boldsymbol{U}_i') = \frac{1}{\hat{\sigma}_e^2}\boldsymbol{Y}'\boldsymbol{R}'\boldsymbol{H}^{-1}\boldsymbol{U}_i\boldsymbol{U}_i'\boldsymbol{H}^{-1}\boldsymbol{R}\boldsymbol{Y}, \qquad (10.7.8)$$

where

$$\boldsymbol{R} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{H}^{-1}.$$

Therefore, an iterative procedure can be established using equations (10.7.7a), (10.7.7b), and (10.7.8).

For some alternative formulations of the likelihood functions and the ML equations, see Hocking (1985, pp. 239–244), Searle et al. (1992, pp. 234–237), and Rao (1997, pp. 93–96). Harville (1977) has presented a thorough review of ML estimation in terms of the general linear model in (10.7.1). Necessary and sufficient conditions for the existence of ML estimates of the variance components are considered by Demidenko and Massaam (1999). Miller (1977) discusses the asymptotic properties of the ML estimates. In particular, Miller proves a result of Cramér type consistency for the ML estimates of both fixed effects and the variance components. For a discussion of the ML estimation for various special models, see Thompson (1980). Hayman (1960) considered the problem of ML estimation of genetic components of variance and Thompson

(1977a, 1977b) discussed the application of the ML procedure for the estimation of heritability. Some iterative procedures and computational algorithms for solving the ML equations are presented in Section 10.8.1. As pointed out by Harville (1969a), however, there are several drawbacks of the Hartley and Rao procedure. Some of them are as follow:

(i) Though it produces a solution to the likelihood equations, over the constrained parameter space, there is no guarantee that the solution is an absolute maximum of the likelihood function over that space.

(ii) While it is true that the procedure yields a sequence estimator with the usual asymptotic properties of maximum likelihood estimators, it is hard to justify the choice of an estimator on the basis of its being a part of a "good" sequence.

(iii) The amount of computation required to apply the Hartley–Rao procedure may be undesirable or prohibitively large.

(iv) The sampling distribution of the estimates produced by the Hartley–Rao procedure can usually be investigated only by a Monte Carlo method. Such studies are awkward to carry out since the sampling distributions of the estimates vary with the true values of the underlying parameters. Moreover, since the likelihood equations may have multiple roots; the solution selected by the Hartley–Rao procedure is partially dependent on the estimate employed to start the iteration process. Thus the sampling distributions of such estimates will be different for each possible choice of the estimator employed to obtain these starting values. It appears likely that the "goodness" of their estimates is directly related to the goodness of their starting values.

10.7.2 LARGE SAMPLE VARIANCES

General expressions for large sample variances of the ML estimators of variance components can be derived, even though the estimators themselves cannot be obtained explicitly. Thus it is known that the large sample variance-covariance matrix of the ML estimators of any model is the inverse of the information matrix. This matrix is the negative of the expected value of second-order partial derivatives—the Hessian matrix—with respect to the parameters of the logarithm of the likelihood (see, e.g., Wald, 1943). The above results can be utilized in deriving large sample variances and covariances of the ML estimators. The presentation given here follows closely Searle (1970).

Consider the general linear model in the form (10.7.1) with the difference that the error vector $\boldsymbol{e}$ is now given by one of the $\boldsymbol{\beta}_i$s and σ_e^2 is one of the variance components $\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2$. The natural logarithm of the likelihood can now be written as

$$\ell n L = -\frac{1}{2} N \ell n(2\pi) - \frac{1}{2} \ell n|\boldsymbol{V}| - \frac{1}{2}(\boldsymbol{Y} - \boldsymbol{X\alpha})' \boldsymbol{V}^{-1} (\boldsymbol{Y} - \boldsymbol{X\alpha}), \quad (10.7.9)$$

where $\boldsymbol{V}$ is the variance-covariance matrix of the observation vector $\boldsymbol{Y}$. Now, let $\boldsymbol{\sigma}^2 = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2)'$ and define

$$\boldsymbol{L}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} = \left\{ \frac{\partial^2 \ell n(L)}{\partial \alpha_h \partial \alpha_k} \right\}, \quad h, k = 1, \ldots, q,$$
$$\boldsymbol{L}_{\boldsymbol{\alpha}\boldsymbol{\sigma}^2} = \left\{ \frac{\partial^2 \ell n(L)}{\partial \alpha_h \partial \sigma_j^2} \right\}, \quad h = 1, \ldots, q; \quad j = 1, \ldots, p,$$

and

$$\boldsymbol{L}_{\boldsymbol{\sigma}^2\boldsymbol{\sigma}^2} = \left\{ \frac{\partial^2 \ell n(L)}{\partial \sigma_i^2 \partial \sigma_j^2} \right\}, \quad i, j = 1, \ldots, p.$$

Then, upon taking the second-order partial derivatives of (10.7.9) with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\sigma}^2$, we obtain

$$\boldsymbol{L}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} = -\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}, \tag{10.7.10}$$
$$\boldsymbol{L}_{\boldsymbol{\alpha}\boldsymbol{\sigma}^2} = \left\{ \boldsymbol{X}' \frac{\partial \boldsymbol{V}^{-1}}{\partial \sigma_j^2} (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha}) \right\}, \quad j = 1, \ldots, p, \tag{10.7.11}$$

and

$$\boldsymbol{L}_{\boldsymbol{\sigma}^2\boldsymbol{\sigma}^2} = \left\{ -\frac{1}{2} \frac{\partial^2 \ell n|\boldsymbol{V}|}{\partial \sigma_i^2 \partial \sigma_j^2} - \frac{1}{2} (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha})' \frac{\partial^2 \boldsymbol{V}^{-1}}{\partial \sigma_i^2 \partial \sigma_j^2} (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha}) \right\}, \tag{10.7.12}$$
$$i, j = 1, \ldots, p.$$

Taking the expectations of (10.7.10), (10.7.11), and (10.7.12) yields

$$E(\boldsymbol{L}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}) = -\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}, \tag{10.7.13}$$
$$E(\boldsymbol{L}_{\boldsymbol{\alpha}\boldsymbol{\sigma}^2}) = \left\{ \boldsymbol{X}' \frac{\partial \boldsymbol{V}^{-1}}{\partial \sigma_j^2} E(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha}) \right\} = \boldsymbol{0}, \quad j = 1, \ldots, p, \tag{10.7.14}$$

and

$$\begin{aligned} E(\boldsymbol{L}_{\boldsymbol{\sigma}^2\boldsymbol{\sigma}^2}) &= \left\{ -\frac{1}{2} \frac{\partial^2 \ell n|\boldsymbol{V}|}{\partial \sigma_i^2 \partial \sigma_j^2} - \frac{1}{2} \operatorname{tr} \left[E(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha})(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha})' \frac{\partial^2 \boldsymbol{V}^{-1}}{\partial \sigma_i^2 \partial \sigma_j^2} \right] \right\} \\ &= \left\{ -\frac{1}{2} \frac{\partial^2 \ell n|\boldsymbol{V}|}{\partial \sigma_i^2 \partial \sigma_j^2} - \frac{1}{2} \operatorname{tr} \left[\frac{\boldsymbol{V} \partial^2 \boldsymbol{V}^{-1}}{\partial \sigma_i^2 \partial \sigma_j^2} \right] \right\}, \quad i, j = 1, \ldots, p. \end{aligned} \tag{10.7.15}$$

Now, using a result of Hartley and Rao (1967) which states that

$$\frac{\partial}{\partial \sigma_i^2} \{\ell n|\boldsymbol{V}|\} = \operatorname{tr}\left(\boldsymbol{V}^{-1} \frac{\partial}{\partial \sigma_i^2} \boldsymbol{V} \right), \tag{10.7.16}$$

we obtain on taking the partial derivative of (10.7.16) with respect to σ_j^2,

$$\frac{\partial^2}{\partial\sigma_i^2\partial\sigma_j^2}\{\ell n|\boldsymbol{V}|\} = \operatorname{tr}\left(\boldsymbol{V}^{-1}\frac{\partial^2\boldsymbol{V}}{\partial\sigma_i^2\partial\sigma_j^2} + \frac{\partial\boldsymbol{V}^{-1}}{\partial\sigma_j^2}\frac{\partial\boldsymbol{V}}{\partial\sigma_i^2}\right). \tag{10.7.17}$$

Again, since

$$\frac{\partial\boldsymbol{V}^{-1}}{\partial\sigma_j^2} = -\boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_j^2}\boldsymbol{V}^{-1}, \tag{10.7.18}$$

on substituting (10.7.18) into (10.7.17), we obtain

$$\frac{\partial^2\{\ell n|\boldsymbol{V}|\}}{\partial\sigma_i^2\partial\sigma_j^2} = \operatorname{tr}\left(\boldsymbol{V}^{-1}\frac{\partial^2\boldsymbol{V}}{\partial\sigma_i^2\partial\sigma_j^2} - \boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_j^2}\boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_i^2}\right). \tag{10.7.19}$$

Furthermore, taking the partial derivative of (10.7.18) with respect to σ_i^2, we obtain

$$\begin{aligned}\frac{\partial^2\boldsymbol{V}^{-1}}{\partial\sigma_i^2\partial\sigma_j^2} &= \boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_j^2}\boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_i^2}\boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\frac{\partial^2\boldsymbol{V}}{\partial\sigma_i^2\partial\sigma_j^2}\boldsymbol{V}^{-1}\\ &\quad + \boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_i^2}\boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_j^2}\boldsymbol{V}^{-1}.\end{aligned} \tag{10.7.20}$$

Multiplying (10.7.20) by $\boldsymbol{V}$ and taking the trace yields

$$\begin{aligned}\operatorname{tr}\left[\boldsymbol{V}\frac{\partial^2\boldsymbol{V}^{-1}}{\partial\sigma_i^2\partial\sigma_j^2}\right] &= \operatorname{tr}\left[\frac{\partial\boldsymbol{V}}{\partial\sigma_j^2}\boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_i^2}\boldsymbol{V}^{-1} - \frac{\partial^2\boldsymbol{V}}{\partial\sigma_i^2\partial\sigma_j^2}\boldsymbol{V}^{-1}\right.\\ &\quad\left. + \frac{\partial\boldsymbol{V}}{\partial\sigma_i^2}\boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_j^2}\boldsymbol{V}^{-1}\right]\\ &= \operatorname{tr}\left[2\boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_i^2}\boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_j^2} - \boldsymbol{V}^{-1}\frac{\partial^2\boldsymbol{V}}{\partial\sigma_i^2\partial\sigma_j^2}\right].\end{aligned} \tag{10.7.21}$$

Now, substituting (10.7.19) and (10.7.21) into (10.7.15), we obtain

$$\begin{aligned}E(\boldsymbol{L}_{\boldsymbol{\sigma}^2\boldsymbol{\sigma}^2}) &= \left\{-\frac{1}{2}\operatorname{tr}\left[\boldsymbol{V}^{-1}\frac{\partial^2\boldsymbol{V}}{\partial\sigma_i^2\partial\sigma_j^2} - \boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_j^2}\boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_i^2}\right]\right.\\ &\quad\left. - \frac{1}{2}\operatorname{tr}\left[2\boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_i^2}\boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_j^2} - \boldsymbol{V}^{-1}\frac{\partial^2\boldsymbol{V}}{\partial\sigma_i^2\sigma_j^2}\right]\right\}\\ &= \left\{-\frac{1}{2}\operatorname{tr}\left[\boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_i^2}\boldsymbol{V}^{-1}\frac{\partial\boldsymbol{V}}{\partial\sigma_j^2}\right]\right\}.\end{aligned}$$

Hence, letting $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\sigma}}^2$ denote the ML estimators of $\boldsymbol{\alpha}$ and $\boldsymbol{\sigma}^2$, their variance-covariance matrix is given by

$$\begin{bmatrix} \mathrm{Var}(\hat{\boldsymbol{\alpha}}) & \vdots & \mathrm{Cov}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\sigma}}^2) \\ \cdots & & \cdots \\ \mathrm{Cov}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\sigma}}^2) & \vdots & \mathrm{Var}(\hat{\boldsymbol{\sigma}}^2) \end{bmatrix} = \begin{bmatrix} -E(\boldsymbol{L}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}) & \vdots & -E(\boldsymbol{L}_{\boldsymbol{\alpha}\boldsymbol{\sigma}^2}) \\ \cdots & & \cdots \\ -E(\boldsymbol{L}_{\boldsymbol{\alpha}\boldsymbol{\sigma}^2}) & \vdots & -E(\boldsymbol{L}_{\boldsymbol{\sigma}^2\boldsymbol{\sigma}^2}) \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X} & \vdots & \boldsymbol{0} \\ \cdots & & \cdots \\ \boldsymbol{0} & \vdots & \frac{1}{2}\,\mathrm{tr}\left[\boldsymbol{V}^{-1}\frac{\partial \boldsymbol{V}}{\partial \sigma_i^2}\boldsymbol{V}^{-1}\frac{\partial \boldsymbol{V}}{\partial \sigma_j^2}\right] \end{bmatrix}^{-1}.$$

Thus we obtain the following results:

$$\begin{aligned} \mathrm{Var}(\hat{\boldsymbol{\alpha}}) &= (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}, \\ \mathrm{Cov}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\sigma}}^2) &= \boldsymbol{0}, \end{aligned} \tag{10.7.22}$$

and

$$\mathrm{Var}(\hat{\boldsymbol{\sigma}}^2) = 2\left\{\mathrm{tr}\left[\boldsymbol{V}^{-1}\frac{\partial \boldsymbol{V}}{\partial \sigma_i^2}\boldsymbol{V}^{-1}\frac{\partial \boldsymbol{V}}{\partial \sigma_j^2}\right],\ i, j = 1, \ldots, p\right\}^{-1}.$$

It should be remarked that the result in (10.7.22) represents the lower bound for the variance-covariance matrix of unbiased estimators. The above result can also be derived from the general procedure described by C. R. Rao (1973, p. 52).

10.8 RESTRICTED MAXIMUM LIKELIHOOD ESTIMATION

The ML procedure discussed in the preceding section yields simultaneous estimation of both the fixed effects and the variance components by maximizing the likelihood, or equivalently the log-likelihood (10.7.3) with respect to each element of the fixed effects and with respect to each of the variance components. Thus the ML estimators for the variance components do not take into account the loss in degrees of freedom resulting from estimating the fixed effects and may produce biased estimates. For example, in the particular case of the model in (10.7.1) with $p = 0$, $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{e}$, and $\boldsymbol{V} = \sigma_e^2\boldsymbol{I}_N$, the ML estimator for the single variance component σ_e^2 is

$$\hat{\sigma}_e^2 = \frac{1}{N}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}})'(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}}),$$

where

$$\hat{\boldsymbol{\alpha}} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}.$$

Clearly, $\hat{\sigma}_e^2$ is a biased estimator since $E(\hat{\sigma}_e^2) = \hat{\sigma}_e^2(N - q)/N$. In contrast, the restricted maximum likelihood (REML) procedure[1] is based on maximizing the restricted or marginal likelihood function that does not contain the fixed effects.[2] This is a generalization of the notion of the restricted maximum likelihood estimation of Thompson (1962) for balanced data, considered in Volume I. Patterson and Thompson (1971) extended this to the randomized block design with unequal block sizes. Following Patterson and Thompson (1975), the REML estimators of the variance components for the model in (10.7.1) are obtained by maximizing the likelihood not of the observation vector $\boldsymbol{Y}$ but the joint likelihood of all error contrasts which are linear combinations of the data having zero expectation.[3] It is to be noted that any linear combination $\boldsymbol{L}'\boldsymbol{Y}$ of the observation vector such that $E(\boldsymbol{L}'\boldsymbol{Y}) = \boldsymbol{0}$, i.e., $\boldsymbol{L}'\boldsymbol{X} = \boldsymbol{0}$, with $\boldsymbol{L}'$ independent of $\boldsymbol{\alpha}$ is an error contrast.[4] Thus the method consists of applying the ML estimation to $\boldsymbol{L}'\boldsymbol{Y}$ where $\boldsymbol{L}'$ is especially chosen so that it contains none of the fixed effects in the model in (10.7.1), i.e., $\boldsymbol{L}'\boldsymbol{X} = \boldsymbol{0}$.

The estimation procedure consists in partitioning the natural logarithm of the likelihood in (10.7.3) into two parts, one of which is free of $\boldsymbol{\alpha}$. This is achieved by adopting a transformation suggested by Patterson and Thompson (1971). In terms of the general linear model in (10.7.1), the transformation being used is

$$\boldsymbol{Y}^* = \boldsymbol{S}\boldsymbol{Y}, \tag{10.8.1}$$

where $\boldsymbol{Y}^*$ and $\boldsymbol{S}$ are partitioned as

$$\boldsymbol{Y}^* = \begin{bmatrix} \boldsymbol{Y}_1^* \\ \boldsymbol{Y}_2^* \end{bmatrix}, \qquad \boldsymbol{S} = \begin{bmatrix} \boldsymbol{S}_1 \\ \boldsymbol{S}_2 \end{bmatrix},$$

with

$$\boldsymbol{S}_1 = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}', \qquad \boldsymbol{S}_2 = \boldsymbol{X}'(\sigma_e^2\boldsymbol{H})^{-1}.$$

It then follows that

$$\boldsymbol{Y}^* \sim N(\boldsymbol{S}\boldsymbol{X}\boldsymbol{\alpha}, \sigma_e^2\boldsymbol{S}\boldsymbol{H}\boldsymbol{S}'), \tag{10.8.2}$$

where

$$\boldsymbol{S}\boldsymbol{X}\boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{X}'(\sigma_e^2\boldsymbol{H})^{-1}\boldsymbol{X}\boldsymbol{\alpha} \end{bmatrix}$$

[1] Some writers use the term residual maximum likelihood or marginal maximum likelihood to describe this procedure.

[2] Harville (1974) showed that the REML may be regarded as a Bayesian procedure where the posterior density is being integrated over fixed effects. In particular, in the case of a noninformative uniform prior, REML is the mode of variance parameters after integrating the fixed effects.

[3] Harville (1974) has shown that, from a Bayesian viewpoint, making inferences on variance components using only error contrasts is equivalent to ignoring any prior information on the unknown fixed parameters and using all the data to make those inferences.

[4] It can be readily seen that REML estimators are invariant to whatever set of error contrasts are chosen as $\boldsymbol{L}'\boldsymbol{Y}$ as long as $\boldsymbol{L}'$ is of full row rank, $N - \text{rank}(\boldsymbol{X})$, with $\boldsymbol{L}'\boldsymbol{X} = \boldsymbol{0}$.

and

$$\sigma_e^2 \boldsymbol{SHS}' = \begin{bmatrix} \boldsymbol{S}_1(\sigma_e^2 \boldsymbol{H})\boldsymbol{S}_1' & \mathbf{0} \\ \mathbf{0} & \boldsymbol{X}'(\sigma_e^2 \boldsymbol{H})^{-1}\boldsymbol{X} \end{bmatrix}.$$

Thus $\boldsymbol{Y}_1^*$ and $\boldsymbol{Y}_2^*$ are independent and the distribution of $\boldsymbol{Y}_1^*$ does not depend on $\boldsymbol{\alpha}$. Note that $\boldsymbol{S}_1$ is a symmetric, idempotent, and singular matrix of rank $N-q$ where q is the rank of $\boldsymbol{X}$.

Now, $\boldsymbol{Y}_1^* = \boldsymbol{S}_1\boldsymbol{Y}$ has a singular multivariate normal distribution with mean vector and variance-covariance matrix given by

$$E(\boldsymbol{S}_1\boldsymbol{Y}) = \boldsymbol{S}_1\boldsymbol{X}\boldsymbol{\alpha} = \mathbf{0}$$

and (10.8.3)

$$\mathrm{Var}(\boldsymbol{S}_1\boldsymbol{Y}) = \boldsymbol{S}_1(\sigma_e^2\boldsymbol{H})\boldsymbol{S}_1' = \sigma_e^2\boldsymbol{S}_1\boldsymbol{H}\boldsymbol{S}_1'.$$

Its likelihood function, therefore, forms the basis for derivation of the estimators of the variance components contained in $\sigma_e^2\boldsymbol{H}$. However, to avoid the singularity of $\boldsymbol{S}_1\boldsymbol{H}\boldsymbol{S}_1'$, arising from the singularity of $\boldsymbol{S}_1$, Corbeil and Searle (1976b) proposed an alternative form of $\boldsymbol{S}_1$. For this, they considered a special form of the incident matrix $\boldsymbol{X}$ given by

$$\boldsymbol{X} = \begin{bmatrix} \mathbf{1}_{n_1} & \mathbf{0}_{n_1} & \dots & \mathbf{0}_{n_1} \\ \mathbf{0}_{n_2} & \mathbf{1}_{n_2} & \dots & \mathbf{0}_{n_2} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{0}_{n_q} & \mathbf{0}_{n_q} & \dots & \mathbf{1}_{n_q} \end{bmatrix} = \sum_{i=1}^{q}{}^{+} \mathbf{1}_{n_i}, \tag{10.8.4}$$

where $\mathbf{1}_{n_i}$ is a vector of n_i ones and $\mathbf{0}_{n_i}$ is a vector of n_i zeros, with $n_i \neq 0$ being the number of observations corresponding to the ith level of the fixed effect; and where $\boldsymbol{\Sigma}^+$ represents a direct sum of matrices.

For many familiar designs, the incident matrix $\boldsymbol{X}$ has the form as given in (10.8.4). Then the matrix $\boldsymbol{S}_1$ defined in (10.8.1) is given by

$$\boldsymbol{S}_1 = \sum_{i=1}^{q}{}^{+} (\boldsymbol{I}_{n_i} - n_i^{-1}\boldsymbol{J}_{n_i}), \tag{10.8.5}$$

where $\boldsymbol{J}_{n_i}$ is an $n_i \times n_i$ matrix with every element unity. Now, the alternative form of $\boldsymbol{S}_1$, denoted by $\boldsymbol{T}$, is derived by deleting n_1th, (n_1+n_2)th, $\dots$, $(n_1 + n_2 + \cdots + n_q)$th rows of $\boldsymbol{S}_1$. Thus $\boldsymbol{T}$ has order $(N-q) \times N$ and is given by

$$\begin{aligned} \boldsymbol{T} &= \sum_{i=1}^{q}{}^{+} [\boldsymbol{I}_{n_i-1} \vdots \mathbf{0}_{n_i-1}] - n_i^{-1}\boldsymbol{J}_{(n_i-1)\times n_i} \\ &= \sum_{i=1}^{q}{}^{+} [(\boldsymbol{I}_{n_i-1} - n_i^{-1}\boldsymbol{J}_{n_i-1} \vdots -n_i^{-1}\mathbf{1}_{n_i-1})]. \end{aligned} \tag{10.8.6}$$

Now, instead of (10.8.1), the transformation being used is

$$Y^* = \begin{bmatrix} T \\ \cdots \\ X'H^{-1} \end{bmatrix} Y = \begin{bmatrix} TY \\ \cdots \\ X'H^{-1}Y \end{bmatrix}, \tag{10.8.7}$$

where Y^* has a multivariate normal distribution with mean vector and variance-covariance matrix given by

$$E(Y^*) = \begin{bmatrix} TE(Y) \\ X'H^{-1}E(Y) \end{bmatrix} = \begin{bmatrix} TX\alpha \\ X'H^{-1}X\alpha \end{bmatrix} \tag{10.8.8}$$

and

$$\text{Var}(Y^*) = \begin{bmatrix} T \\ \cdots \\ X'H^{-1} \end{bmatrix} (\sigma_e^2 H)[T' \vdots H^{-1}X]. \tag{10.8.9}$$

It can be verified that for X and T given by (10.8.4) and (10.8.6), respectively, $TX = \mathbf{0}$, so that (10.8.8) and (10.8.9) reduce to

$$E(Y^*) = \begin{bmatrix} \mathbf{0} \\ X'H^{-1}X\alpha \end{bmatrix} \tag{10.8.10}$$

and

$$\text{Var}(Y^*) = \sigma_e^2 \begin{bmatrix} THT' & \vdots & \mathbf{0} \\ \cdots & \vdots & \cdots \\ \mathbf{0} & \vdots & X'H^{-1}X \end{bmatrix}. \tag{10.8.11}$$

The transformation (10.8.7) is nonsingular, because each X' and T, given by (10.8.4) and (10.8.6), respectively, has full row rank; and from the relation $TX = \mathbf{0}$ it follows that the rows of T are linearly independent of those of X'. Now, from (10.8.7) and (10.8.11), it can be readily seen that the log-likelihood of Y^* is the sum of the log-likelihoods of TY and $X'H^{-1}Y$. Denoting these likelihoods by L_1 and L_2, we have

$$\begin{aligned} L_1 = &-\frac{1}{2}(N-q)\ell n(2\pi) - \frac{1}{2}(N-q)\ell n\sigma_e^2 \\ &- \frac{1}{2}\ell n|THT'| - \frac{1}{2\sigma_e^2}Y'T'(THT')^{-1}TY \end{aligned} \tag{10.8.12}$$

and

$$L_2 = -\frac{1}{2}q\ell n(2\pi) - \frac{1}{2}q\ell n\sigma_e^2 - \frac{1}{2}\ell n|X'H^{-1}X|$$

$$-\frac{1}{2\sigma_e^2}(\boldsymbol{Y}-\boldsymbol{X\alpha})'\boldsymbol{H}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{H}^{-1}(\boldsymbol{Y}-\boldsymbol{X\alpha}). \quad (10.8.13)$$

Now, L_1 does not involve $\boldsymbol{\alpha}$; so that the REML estimators of σ_e^2 and the variance ratios ρ_is contained in $\boldsymbol{H}$ are those values of σ_e^2 and ρ_is that maximize L_1 subject to the constraints that σ_e^2 and ρ_is are nonnegative. Equating to zero the partial derivatives of (10.8.12) with respect to σ_e^2 and ρ_is, the ML equations are

$$\frac{\partial L_1}{\partial \sigma_e^2} = -\frac{1}{2\sigma_e^2}(N-q) + \frac{1}{2\sigma_e^4}\boldsymbol{Y}'\boldsymbol{T}'(\boldsymbol{THT}')^{-1}\boldsymbol{TY} = 0, \quad (10.8.14)$$

and

$$\frac{\partial L_1}{\partial \rho_i} = -\frac{1}{2}\operatorname{tr}[\boldsymbol{U}_i'\boldsymbol{T}'(\boldsymbol{THT}')^{-1}\boldsymbol{TU}_i] + \frac{1}{2\sigma_e^2}\boldsymbol{Y}'\boldsymbol{T}'(\boldsymbol{THT}')^{-1}\boldsymbol{TU}_i\boldsymbol{U}_i'\boldsymbol{T}'(\boldsymbol{THT}')^{-1}\boldsymbol{TY} = 0. \quad (10.8.15)$$

Equations (10.8.14) and (10.8.15) clearly have no closed form analytic solutions and have to be solved numerically using some iterative procedures under the constraints that $\sigma_e^2 > 0$ and $\rho_i \geq 0$, for $i = 1, 2, \ldots, p$. An iterative procedure consists of assigning some initial values to ρ_is, and then (i) solve (10.8.14) for σ_e^2 giving

$$\hat{\sigma}_e^2 = \boldsymbol{Y}'\boldsymbol{T}'(\boldsymbol{THT}')^{-1}\boldsymbol{TY}/(N-q), \quad (10.8.16)$$

and (ii) use the ρ_i values and $\hat{\sigma}_e^2$ from (10.8.16) to compute new ρ_i values that make (10.8.15) closer to zero. Repetition of (i) and (ii) terminating at (i) is continued until a desired degree of accuracy is achieved. Corbeil and Searle (1976b) discuss some computing algorithms as well as the estimation of the fixed effects based on the restricted maximum likelihood estimators. They also consider the generalization of the method applicable for any $\boldsymbol{X}$ and $\boldsymbol{T}$ and derive the large sample variances of the estimators thus obtained. It should be remarked that Patterson and Thompson (1975) in their work do not take into consideration the constraints of nonnegativity for the variance components. Similarly, Corbeil and Searle (1976b) also do not incorporate these constraints in their development. Giesbrecht and Burrows (1978) have proposed an efficient method for computing REML estimates by an iterative application of MINQUE procedure where estimates obtained from each iteration are used as the prior information for the next iteration.

For some alternative formulations of the restricted likelihood functions and the REML equations, see Harville (1977), Hocking (1985, pp. 244–249), Lee and Kapadia (1991), Searle et al. (1992, pp. 249–253), and Rao (1997, pp. 99–102). Necessary and sufficient conditions for the existence of REML estimates of the variance components are considered by Demidenko and Massam (1999). Engel (1990) discussed the problem of statistical inference for the fixed effects and the REML estimation of the variance components in an unbalanced mixed model. Fellner (1986) and Richardson and Welsh (1995) have considered robust modifications of the REML estimation. For asymptotic behavior and other related properties of the REML estimation, see Das (1979), Cressie and Lahiri (1993), Richardson and Welsh (1994), and Jiang (1996). For some results on estimation of sampling variances and covariances of the REML estimators, see Ashida and Iwaisaki (1995).

10.8.1 NUMERICAL ALGORITHMS, TRANSFORMATIONS, AND COMPUTER PROGRAMS

As we have seen in Sections 10.7 and 10.8, the evaluation of the ML and REML estimators of variance components entails the use of numerical algorithms involving iterative procedures. There are many iterative algorithms that can be employed for computing ML and REML estimates. Some were developed specifically for this problem while others are adaptations of general procedures for the numerical solution of nonlinear optimization problems with constraints. There is no single algorithm that is best or even satisfactory for every application. An algorithm that may converge to an ML or REML estimate rather rapidly for one problem may converge slowly or even fail to converge in another. The solution of an algorithm for a particular application requires some judgement about the computational requirements and other properties as applied to a given problem.

Some of the most commonly used algorithms for this problem include the so-called, steepest ascent, Newton–Raphson, Fisher scoring, EM (expectation-maximization) algorithm, and various ad hoc algorithms derived by manipulating the likelihood equations and applying the method of successive approximations. Vandaele and Chowdhury (1971) proposed a revised method of scoring that will ensure convergence to a local maximum of the likelihood function, but there is no guarantee that the global maximum will be attained. Hemmerle and Hartley (1973) discussed the Newton–Raphson method for the mixed model estimation which is closely related to the method of scoring. Jennrich and Sampson (1976) presented a unified approach of the Newton–Raphson and scoring algorithms to the estimation and testing in the general mixed model analysis of variance and discussed their advantages and disadvantages. Harville (1977) and Hartley et al. (1978) discuss the iterative solution of the likelihood equations and Thompson (1980) describes the method of scoring using the expected values of second-order differentials.

Dempster et al. (1981), Laird (1982), Henderson (1984), and Raudenbush and Bryk (1986) discuss the use of an EM algorithm for computation of the ML and REML estimates of the variance and covariance components. In addition, Dempster et al. (1984) and Longford (1987) have described the Newton–Raphson and scoring algorithms for computing the ML estimates of variance components for a mixed model analysis. Thompson and Meyer (1986) proposed some efficient algorithms which for balanced data situations yield an exact solution in a single iteration. Graser et al. (1987) described a derivative-free algorithm for REML estimation of variance components in single-trait animal or reduced animal models that does not use matrix inversion. Laird et al. (1987) used Aitken's acceleration (Gerald, 1977) to improve the speed of convergence of the EM algorithm for ML and REML estimation and Lindstrom and Bates (1988) developed the implementation of the Newton–Raphsonindex and EM algorithms for ML and REML estimation of the parameters in mixed effects models for repeated measures data. More recently, Callanan (1985), Harville and Callanan (1990), and Callanan and Harville (1989, 1991) have proposed several new algorithms. Numerical results indicate that these algorithms improve on the method of successive approximation and the Newton–Raphson algorithm and are superior to other widely used algorithms like Fisher's scoring and the EM algorithm.

Robinson (1984, 1987) discussed a modification of an algorithm proposed by Thompson (1977a) which is similar to Fisher's scoring technique. Robinson (1984, 1987) noted that his algorithm compares favorably with the Newton–Raphson algorithm outlined by Dempster et al. (1984). Lin and McAllister (1984) and others have commented favorably on the algorithm, which generally converges faster than others with many jobs requiring three or fewer iterations. For some further discussion and details of computational algorithms for the ML and REML estimation of variance components, see Searle et al. (1992, Chapter 8). Utmost caution should be exercised in using these algorithms for problems that are fairly large and highly unbalanced. As Klotz and Putter (1970) have noted, the behavior of likelihood as a function of variance components is generally complex even for a relatively simple model. For example, the likelihood equation may have multiple roots or the ML estimate may lie at the boundary rather than a solution of any of these roots. In fact, J. N. K. Rao (1977) has commented that none of the existing algorithms guarantee a solution, which is indeed ML or REML.

In many practical problems, the use of a suitable transformation can ease much of the computational burden associated with determination of the ML and REML estimates. Various transformations have been suggested to improve the performance of the numerical algorithms in computing the ML and REML estimates. For example, Hemmerle and Hartley (1973) proposed a transformation known as W-transformation in order to reduce the problem of inversion of the variance-covariance matrix of order $N \times N$ to a smaller matrix of order $m \times m$, where $m = \sum_{i=1}^{p} m_i$. Thompson (1975) and Hemmerle and Lorens (1976) discussed some improved algorithms for the W-transformation. Corbeil and Searle (1976b) presented an adaptation of the W-transformation for computing the REML estimates of variance components in the general mixed model. Jennrich and Sampson (1976) used the W-transformation to develop a Newton–Raphson algorithm and a Fisher scoring algorithm, both distinct from the Newton–Raphson algorithm of Hemmerle and Hartley. Similarly, Harville (1977) suggested that the algorithms may be made more efficient by making the likelihood function more quadratic. Another class of transformations has been suggested by Thompson (1980) by consideration of orthogonal designs.

Hartley and Vaughn (1972) developed a computer program for computing the ML estimates using the Hartley–Rao procedure described in Section 10.7. Robinson (1984) developed a general purpose FORTRAN program, the REML program, which can be run without conversion on most modern computers. The user can specify the type of output required, which may range from estimates of variance components plus standard errors to a complete list of all parameters and standard errors of differences between all pairs including linear functions and ratios of linear functions of variance components such as heritability. Current releases of SAS®, SPSS®, BMDP®, and S-PLUS® compute the ML and REML estimates with great speed and accuracy simply by specifying the model in question (see, Appendix O).

10.9 BEST QUADRATIC UNBIASED ESTIMATION

The variance component analogue of the best linear unbiased estimator (BLUE) of a function of fixed effects is a best quadratic unbiased estimator (BQUE), that is, a quadratic function of the observations that is unbiased for the variance component and has minimum variance among all such estimators. As we have seen in Volume I of this text, for balanced data, the analysis of variance estimators are unbiased and have minimum variance. Derivation of BQUEs from unbalanced data, however, is much more difficult than from balanced data. Ideally, one would like estimators that are uniformly "best" for all values of the variance components. However, as Scheffé (1959), Harville (1969a), Townsend and Searle (1971), and LaMotte (1973b) have noted, uniformly best estimators (not functions of variance components) of variance components from unbalanced data do not exist even for the simple one-way random model. Townsend and Searle (1971) have obtained locally BQUEs for the variance components in a one-way classification with $\mu = 0$; and from these they have proposed approximate BQUEs for the $\mu \neq 0$ model. We will discuss their results in the next chapter. The BQUE procedure for the variance components in a general linear model is C. R. Rao's minimum-variance quadratic unbiased estimation (MIVQUE) to be discussed in the following section.

10.10 MINIMUM-NORM AND MINIMUM-VARIANCE QUADRATIC UNBIASED ESTIMATION

In a series of papers, C. R. Rao (1970, 1971a, 1971b, 1972) proposed some general procedures for deriving quadratic unbiased estimators, which have either the minimum-norm or minimum-variance property. Rao's (1970) paper is motivated by Hartley et al. (1969) paper, which considers the following problem on the estimation of heteroscedastic variances in a linear model. Let $y_1, y_2, \ldots, y_N$ be a random sample from the model

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}, \qquad (10.10.1)$$

where

$\boldsymbol{X}$ is a known $N \times m$ matrix,
$\boldsymbol{\beta}$ is an m-vector of unknown parameters,

and

$\boldsymbol{e}$ is an N-vector of random error terms.

It is further assumed that $\boldsymbol{e}$ has mean vector zero and variance-covariance matrix given by a diagonal matrix with diagonal terms given by $\sigma_1^2, \ldots, \sigma_N^2$. The problem is to estimate σ_i^2s when they may be all unequal. C. R. Rao (1970) derived the conditions on $\boldsymbol{X}$ which ensure unbiased estimability of the σ_i^2s. He further introduced an estimation principle, called the minimum-norm quadratic unbiased estimation (MINQUE), and showed that the estimators of Hartley et al. (1969) are in fact MINQUE. As noted by Rao (1972), the problem of estimation of heteroscedastic variances is, indeed, a special case of the estimation of variance components problem.

10.10.1 FORMULATION OF MINQUE[5] AND MIVQUE

Consider the general linear model in the form (10.7.1) with the difference that the error vector $\boldsymbol{e}$ is now given by one of the $\boldsymbol{\beta}_i$s and σ_e^2 is one of the variance components $\sigma_1^2, \ldots, \sigma_p^2$. The model in (10.7.1) can then be expressed in a more succinct form as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{U}\boldsymbol{\beta}, \qquad (10.10.2)$$

where

$$\boldsymbol{U} = [\boldsymbol{U}_1 \;\vdots\; \boldsymbol{U}_2 \;\vdots\; \cdots \;\vdots\; \boldsymbol{U}_p]$$

and

$$\boldsymbol{\beta}' = [\boldsymbol{\beta}_1' \;\vdots\; \boldsymbol{\beta}_2' \;\vdots\; \cdots \;\vdots\; \boldsymbol{\beta}_p'].$$

From (10.10.2), we have

$$E(\boldsymbol{Y}) = \boldsymbol{X}\boldsymbol{\alpha} \qquad (10.10.3)$$

[5]The acronym MINQUE (MIVQUE) is used both for minimum-norm (-variance) quadratic unbiased estimation and for minimum-norm (-variance) quadratic unbiased estimate/estimator.

and

$$\mathrm{Var}(\boldsymbol{Y}) = \sum_{i=1}^{p} \sigma_i^2 \boldsymbol{V}_i,$$

where

$$\boldsymbol{V}_i = \boldsymbol{U}_i \boldsymbol{U}_i', \quad i = 1, 2, \ldots, p.$$

For both MINQUE and MIVQUE, Rao (1972) proposed estimating

$$\sum_{i=1}^{p} \ell_i \sigma_i^2, \tag{10.10.4}$$

a linear combination of the variance components σ_i^2s, by a quadratic form $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$, where $\boldsymbol{A}$ is a symmetric matrix chosen subject to the conditions which guarantee the estimator's unbiasedness and invariance to changes in $\boldsymbol{\alpha}$. For unbiasedness, we must have

$$E(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) = \sum_{i=1}^{p} \ell_i \sigma_i^2. \tag{10.10.5}$$

Further, from result (i) of Theorem 9.3.1, we have

$$E(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) = E(\boldsymbol{Y}')\boldsymbol{A}E(\boldsymbol{Y}) + \mathrm{tr}[\boldsymbol{A}\,\mathrm{Var}(\boldsymbol{Y})],$$

which, after substitution from (10.10.3), becomes

$$E(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) = \boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\boldsymbol{\alpha} + \sum_{i=1}^{p} \sigma_i^2\, \mathrm{tr}[\boldsymbol{A}\boldsymbol{V}_i].$$

Therefore, the condition of unbiasedness in (10.10.5) is equivalent to

$$\boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\boldsymbol{\alpha} + \sum_{i=1}^{p} \sigma_i^2\, \mathrm{tr}[\boldsymbol{A}\boldsymbol{V}_i] = \sum_{i=1}^{p} \ell_i \sigma_i^2.$$

Thus the estimator $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ is unbiased if and only if $\boldsymbol{A}$ is chosen to satisfy

$$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X} = \boldsymbol{0} \quad \text{and} \quad \mathrm{tr}[\boldsymbol{A}\boldsymbol{V}_i] = \ell_i. \tag{10.10.6}$$

For invariance[6] to changes in $\boldsymbol{\alpha}$ (i.e., $\boldsymbol{\alpha}$ is transformed to $\boldsymbol{\alpha} + \boldsymbol{\alpha}_0$), we must have

$$(\boldsymbol{Y} + \boldsymbol{X}\boldsymbol{\alpha}_0)'\boldsymbol{A}(\boldsymbol{Y} + \boldsymbol{X}\boldsymbol{\alpha}_0) = \boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} \tag{10.10.7}$$

[6]For a discussion of various levels of invariance and invariant inference for variance components, see Harville (1988).

for all $\boldsymbol{\alpha}_0$. Now, (10.10.7) is true if and only if

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{0}. \tag{10.10.8}$$

Hence, from (10.10.6) and (10.10.8), the conditions for both unbiasedness and invariance to $\boldsymbol{\alpha}$ are

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{0} \quad \text{and} \quad \text{tr}[\boldsymbol{A}\boldsymbol{V}_i] = \ell_i. \tag{10.10.9}$$

10.10.2 DEVELOPMENT OF THE MINQUE

Suppose $\boldsymbol{\beta}_i$s in the model in (10.10.2) are observable random vectors. Then a *natural*[7] estimator of (10.10.4) is

$$\sum_{i=1}^{p} \ell_i \boldsymbol{\beta}_i' \boldsymbol{\beta}_i / n_i, \tag{10.10.10}$$

which can be written as

$$\boldsymbol{\beta}' \boldsymbol{\Delta} \boldsymbol{\beta}, \tag{10.10.11}$$

where $\boldsymbol{\Delta}$ is a suitably defined diagonal matrix. However, from (10.10.2), the proposed estimator of (10.10.4) is

$$\begin{aligned} \boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} &= (\boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{U}\boldsymbol{\beta})'\boldsymbol{A}(\boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{U}\boldsymbol{\beta}) \\ &= \boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}\boldsymbol{\alpha} + 2\boldsymbol{\alpha}'\boldsymbol{X}'\boldsymbol{A}\boldsymbol{U}\boldsymbol{\beta} + \boldsymbol{\beta}'\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U}\boldsymbol{\beta}. \end{aligned} \tag{10.10.12}$$

Under the conditions in (10.10.9), the estimator (10.10.12) reduces to

$$\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} = \boldsymbol{\beta}'\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U}\boldsymbol{\beta}. \tag{10.10.13}$$

Now, the difference between the proposed estimator (10.10.13) and the *natural* estimator (10.10.11) is

$$\boldsymbol{\beta}'(\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U} - \boldsymbol{\Delta})\boldsymbol{\beta}. \tag{10.10.14}$$

The MINQUE procedure seeks to minimize the difference (10.10.14) in some sense subject to the conditions in (10.10.9). One possibility is to minimize the Euclidean norm

$$\|\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U} - \boldsymbol{\Delta}\|, \tag{10.10.15}$$

where $\| \;\|$ denotes the norm of a matrix, and for any symmetric matrix $\boldsymbol{M}$,

$$\|\boldsymbol{M}\| = \{\text{tr}[\boldsymbol{M}^2]\}^{1/2}. \tag{10.10.16}$$

Equivalently, we can minimize the squared Euclidean norm given by

$$\|\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U} - \boldsymbol{\Delta}\|^2 = \text{tr}[(\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U} - \boldsymbol{\Delta})^2]$$

[7] The term *natural* was introduced by Rao himself.

$$= \text{tr}[(\boldsymbol{A}\boldsymbol{V})^2] - \text{tr}[\boldsymbol{\Delta}^2], \qquad (10.10.17)$$

where

$$\boldsymbol{V} = \boldsymbol{V}_1 + \cdots + \boldsymbol{V}_p$$

with $\boldsymbol{V}_i$ defined in (10.10.3). Inasmuch as $\text{tr}[\boldsymbol{\Delta}^2]$ does not involve $\boldsymbol{A}$, the problem of MINQUE reduces to minimizing $\text{tr}[(\boldsymbol{A}\boldsymbol{V})^2]$, subject to the conditions in (10.10.9).

Alternatively, Rao (1972) considers the standardization of $\boldsymbol{\beta}_i$s (since all may not have the same standard deviation) by

$$\boldsymbol{\eta}_i = \sigma_i^{-1}\boldsymbol{\beta}_i. \qquad (10.10.18)$$

Then the difference (10.10.14) is given by

$$\boldsymbol{\eta}'\boldsymbol{\Sigma}^{1/2}(\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U} - \boldsymbol{\Delta})\boldsymbol{\Sigma}^{1/2}\boldsymbol{\eta}, \qquad (10.10.19)$$

where

$$\boldsymbol{\eta}' = (\boldsymbol{\eta}_1' \vdots \boldsymbol{\eta}_2' \vdots \cdots \vdots \boldsymbol{\eta}_p')$$

and

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2\boldsymbol{I}_{m_1} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \sigma_p^2\boldsymbol{I}_{m_p} \end{bmatrix}. \qquad (10.10.20)$$

Now, the minimization of (10.10.19) using the Euclidean norm (10.10.16) is equivalent to minimizing $\text{tr}[(\boldsymbol{A}\boldsymbol{W})^2]$ subject to the conditions in (10.10.9), where

$$\boldsymbol{W} = \sigma_1^2\boldsymbol{V}_1 + \cdots + \sigma_p^2\boldsymbol{V}_p. \qquad (10.10.21)$$

In the definition of the matrix $\boldsymbol{W}$ in (10.10.21), the weights σ_i^2s are, of course, unknown. Rao (1972) suggested the following two amendments to this problem:

(i) If we have a priori knowledge of the approximate ratios $\sigma_1^2/\sigma_p^2, \ldots, \sigma_{p-1}^2/\sigma_p^2$, we can substitute them in (10.10.21) and use the $\boldsymbol{W}$ thus computed.

(ii) We can use a priori weights in (10.10.21) and obtain MINQUEs of σ_i^2s. These estimates then may be substituted in (10.10.21) and the MINQUE procedure repeated. The procedure is called iterative MINQUE or I-MINQUE (Rao and Kleffé, 1988, Section 9.1). In this iterative scheme, the property of unbiasedness is usually lost; but the estimates thus obtained may have some other interesting properties. Rao (1971a) also gives the conditions under which the MINQUE is independent of a priori weights σ_i^2s.

We now state a theorem due to Rao (1972), that can be employed to solve the minimization problem involved in the MINQUE procedure.

Theorem 10.10.1. *Define a matrix* $\boldsymbol{P}$ *as*

$$\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{H}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{H}^{-1}, \tag{10.10.22}$$

where $\boldsymbol{X}$ *is the matrix in the model in* (10.10.2) *and* $\boldsymbol{H}$ *is a positive definite matrix. Then the minimum of* $\text{tr}[(\boldsymbol{AH})^2]$, *subject to the conditions*

$$\boldsymbol{AX} = \boldsymbol{0} \quad \textit{and} \quad \text{tr}[\boldsymbol{AV}_i] = \ell_i, \quad i = 1, \ldots, p, \tag{10.10.23}$$

is attained at

$$\boldsymbol{A}^* = \sum_{i=1}^{p} \lambda_i \boldsymbol{RV}_i\boldsymbol{R}, \tag{10.10.24}$$

where

$$\boldsymbol{R} = \boldsymbol{H}^{-1}(\boldsymbol{I} - \boldsymbol{P})$$

and (10.10.25)

$$\boldsymbol{\lambda}' = (\lambda_1, \lambda_2, \ldots, \lambda_p)$$

is determined from the equations

$$\boldsymbol{S\lambda} = \boldsymbol{\ell} \tag{10.10.26}$$

with

$$\boldsymbol{S} = \{s_{ij}\} = \{\text{tr}\, \boldsymbol{RV}_i\boldsymbol{RV}_j\}, \quad i, j = 1, \ldots, p,$$

and (10.10.27)

$$\boldsymbol{\ell}' = (\ell_1, \ell_2, \ldots, \ell_p).$$

Proof. From (10.10.26), we note that

$$\boldsymbol{\lambda} = \boldsymbol{S}^{-}\boldsymbol{\ell},$$

so that $\boldsymbol{\lambda}$ exists if an unbiased estimator of $\sum_{i=1}^{p} \ell_i\sigma_i^2$ exists. Also $\boldsymbol{A}^*\boldsymbol{X} = \boldsymbol{0}$ and $\text{tr}[\boldsymbol{A}^*\boldsymbol{V}_i] = \ell_i$, in view of the choice of $\boldsymbol{\lambda}$ to satisfy (10.10.26). Now, let

$\boldsymbol{A} = \boldsymbol{A}^* + \boldsymbol{D}$ be an alternative matrix. Then $\text{tr}[\boldsymbol{D}\boldsymbol{V}_i] = 0, i = 1, \ldots, p$. Furthermore, $\boldsymbol{D}\boldsymbol{X} = \boldsymbol{0} \rightarrow \boldsymbol{R}\boldsymbol{H}\boldsymbol{D} = \boldsymbol{D}$. Then

$$\begin{aligned}\text{tr}[A^*\boldsymbol{H}\boldsymbol{D}\boldsymbol{H}] &= \sum_{i=1}^{p} \lambda_i \,\text{tr}[\boldsymbol{R}\boldsymbol{V}_i\boldsymbol{R}\boldsymbol{H}\boldsymbol{D}\boldsymbol{H}] \\ &= \sum_{i=1}^{p} \lambda_i \,\text{tr}[\boldsymbol{V}_i\boldsymbol{D}\boldsymbol{H}\boldsymbol{R}] \\ &= \sum_{i=1}^{p} \lambda_i \,\text{tr}[\boldsymbol{V}_i\boldsymbol{D}] \\ &= \boldsymbol{0}. \end{aligned} \tag{10.10.28}$$

Hence,

$$\text{tr}[(\boldsymbol{A}^* + \boldsymbol{D})\boldsymbol{H}(\boldsymbol{A}^* + \boldsymbol{D})\boldsymbol{H}] = \text{tr}[(\boldsymbol{A}^*\boldsymbol{H})^2] + \text{tr}[(\boldsymbol{D}\boldsymbol{H})^2], \tag{10.10.29}$$

which shows that the minimum is attained at $\boldsymbol{A}^*$.

Now, we can apply Theorem 10.10.1 for the problem of MINQUE, choosing $\boldsymbol{H} = \boldsymbol{V}_1 + \cdots + \boldsymbol{V}_p$ or $\boldsymbol{H} = \sigma_1^{*2}\boldsymbol{V}_1 + \cdots + \sigma_p^{*2}\boldsymbol{V}_p$, where $\sigma_1^{*2}, \ldots, \sigma_p^{*2}$ are a priori ratios of unknown variance components. Using formula (10.10.24), the MINQUE of $\sum_{i=1}^{p} \ell_i\sigma_i^2 = \boldsymbol{\ell}'\boldsymbol{\sigma}^2$, where $\boldsymbol{\sigma}'^2 = (\sigma_1^2, \ldots, \sigma_p^2)$, is given by

$$\boldsymbol{\ell}'\hat{\boldsymbol{\sigma}}^2 = \boldsymbol{Y}'\boldsymbol{A}^{\boldsymbol{*}}\boldsymbol{Y} = \sum_{i=1}^{p} \lambda_i \boldsymbol{Y}'\boldsymbol{R}\boldsymbol{V}_i\boldsymbol{R}\boldsymbol{Y} = \sum_{i=1}^{p} \lambda_i \gamma_i, \tag{10.10.30}$$

where

$$\gamma_i = \boldsymbol{Y}'\boldsymbol{R}\boldsymbol{V}_i\boldsymbol{R}\boldsymbol{Y}.$$

Letting

$$\boldsymbol{\gamma}' = (\gamma_1, \ldots, \gamma_p),$$

the estimator (10.10.30) can be written as

$$\boldsymbol{\ell}'\hat{\boldsymbol{\sigma}}^2 = \boldsymbol{\lambda}'\boldsymbol{\gamma}. \tag{10.10.31}$$

Further, on substituting $\boldsymbol{\lambda} = \boldsymbol{S}^-\boldsymbol{\ell}$ in (10.10.31), we have

$$\boldsymbol{\ell}'\hat{\boldsymbol{\sigma}}^2 = \boldsymbol{\ell}'\boldsymbol{S}^-\boldsymbol{\gamma}. \tag{10.10.32}$$

Therefore, the MINQUE vector of $\boldsymbol{\sigma}^2$ is given by

$$\hat{\boldsymbol{\sigma}}^2 = \boldsymbol{S}^-\boldsymbol{\gamma}. \tag{10.10.33}$$

The solution vector (10.10.33) is unique if and only if the individual components are unbiasedly estimable. However, if $\boldsymbol{\ell}'\boldsymbol{\sigma}^2$ is estimable, any solution to (10.10.33) would lead to a unique estimate. Furthermore, the solution vector (10.10.33) for MINQUE involves the computation of terms like tr$[\boldsymbol{R}\boldsymbol{V}_i\boldsymbol{R}\boldsymbol{V}_j]$ (the (i, j)th element of the matrix $\boldsymbol{S}$), $\boldsymbol{Y}'\boldsymbol{R}\boldsymbol{V}_i\boldsymbol{R}\boldsymbol{Y} = \text{tr}[\boldsymbol{R}\boldsymbol{V}_i\boldsymbol{R}\boldsymbol{Y}\boldsymbol{Y}']$ (the ith component of the vector $\boldsymbol{\gamma}$), the matrix $\boldsymbol{R} = \boldsymbol{H}^{-1}(\boldsymbol{I} - \boldsymbol{P})$, which in turn involves the computation of the matrix $\boldsymbol{P}$ defined by (10.10.22). □

Remark: One can also consider the problem of deriving MINQUE without the condition of invariance. Now the problem reduces to that of minimizing (10.10.14) subject to the conditions (10.10.6). Rao (1971a) gives an explicit solution for this problem and an alternative form is given by Pringle (1974) (see also Focke and Dewess, 1972). ♦

10.10.3 DEVELOPMENT OF THE MIVQUE

For MIVQUE, Rao (1971b) proposes to minimize the variance of $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ subject to the conditions in (10.10.9) for unbiasedness and invariance. In general, when the elements of $\boldsymbol{\beta}_i$ have a common variance σ_i^2 and common fourth moment μ_{4i}, the variance of $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ is given by

$$\text{Var}(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) = 2\,\text{tr}[(\boldsymbol{A}\boldsymbol{W})^2] + \sum_{i=1}^{p} \kappa_i \sigma_i^4\, \text{tr}(\boldsymbol{A}\boldsymbol{V}_i)^2, \tag{10.10.34}$$

where $\boldsymbol{W}$ is defined in (10.10.21) and κ_i is the common kurtosis of the variables in $\boldsymbol{\beta}_i$ i.e., $\kappa_i = \mu_{4i}/\sigma_i^4 - 3$. Under normality, i.e., when $\boldsymbol{\beta}_i$s are normally distributed, the kurtosis terms are zero; so that (10.10.34) simplifies to

$$\text{Var}(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) = 2\,\text{tr}[(\boldsymbol{A}\boldsymbol{W})^2]. \tag{10.10.35}$$

The MIVQUE procedure, under normality, therefore, consists of minimizing (10.10.35) subject to the conditions in (10.10.9). Thus MIVQUE under normality is identical to the alternative form of the MINQUE discussed earlier in this section (see also Kleffé, 1976). The problem of general MIVQUE, i.e., of minimizing (10.10.34) is considered by Rao (1971b). Furthermore, expression (10.10.34) can be written as

$$\text{Var}(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) = \sum_{i=1}^{p}\sum_{j=1}^{p} \lambda_{ij}\, \text{tr}[\boldsymbol{A}\boldsymbol{V}_i\boldsymbol{A}\boldsymbol{V}_j)],$$

where

$$\lambda_{ij} = \begin{cases} 2\sigma_i^2\sigma_j^2, & i \neq j, \\ (2+\kappa_i)\sigma_i^4, & i = j. \end{cases}$$

When λ_{ij}s are unknown, one may minimize

$$\sum_{i=1}^{p}\sum_{j=1}^{p} \text{tr}[(\boldsymbol{A}\boldsymbol{V}_i\boldsymbol{A}\boldsymbol{V}_j)]. \tag{10.10.36}$$

Note that expression (10.10.36) is precisely equivalent to $\mathrm{tr}[(\boldsymbol{A}\boldsymbol{V})^2]$. Thus, in this case MIVQUE is identical to MINQUE norm chosen in (10.10.17).

10.10.4 SOME COMMENTS ON MINQUE AND MIVQUE

It should be noted that the MIVQUEs are, in general, functions of the unknown variance components. Thus there are different MIVQUEs for different values of $(\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2)$; and they are sometimes called "locally" MIVQUE. As noted in Section 10.9, "uniformly" MIVQUEs (not functions of the variance components) from unbalanced data do not exist even for the simple one-way random model.

Mitra (1972) verified some of the MINQUE and MIVQUE results through derivations using least squares by considering variables whose expectations are linear functions of the variances. LaMotte (1973a) also arrived at many of these results, although he approaches the problem completely in terms of the minimum-variance criterion and without the use of minimum-norm principle. Some of the results of LaMotte are discussed in the next section. Brown (1977) derived the MINQUE using the weighted least squares approach. Verdooren (1980, 1988) also gives a derivation of the MINQUE using the generalized least squares estimation. Rao (1973, 1974, 1979) further elaborated some of the properties of the MINQUE such as its relationship to the ML and REML estimation. Hocking and Kutner (1975) and Patterson and Thompson (1975) have pointed out that the MINQUE estimates are equivalent to the REML estimates obtained using a single iteration. Note that the computation of iterative MINQUE under the assumption of normality until convergence is achieved (with appropriate constraints for nonnegative values) leads to REML. In practice, convergence tends to be very rapid and the estimates obtained from a single iteration can be interpreted as equivalent to REML estimates. Thus, for the balanced models, if the usual ANOVA estimates are nonnegative, they are equivalent to the MINQUE estimates (see also Anderson, 1979). Pukelsheim (1974, 1976) introduced the concept of dispersion mean model and showed that an application of generalized least squares to this model yields the MINQUE estimators. Chaubey (1977) considered various extensions, modifications, and applications of the MINQUE principle to estimate variance and covariance components in the univariate and multivariate linear models. Chaubey (1980b, 1982, 1985) used some modifications of the MINQUE procedure to estimate variances and covariances in intraclass covariance models and to derive some commonly used estimators of covariances in time series models. Henderson (1985) has discussed the relation between the REML and MINQUE in the context of a genetic application. For a general overview of the MINQUE theory and related topics, see P. S. R. S. Rao (1977, 2001), Kleffé (1977b, 1980), and Rao and Kleffé (1980); for a book-length treatment of the MINQUE and MIVQUE estimation, see Rao and Kleffé (1988).

It should be remarked that the MINQUE procedure is 'nonparametric', that is, it does not require any distributional assumptions of the underlying random effects. Liu and Senturia (1975) presented some results concerning the distribution of the MINQUE estimators. Brown (1976) has shown that in nonnormal models having a special balanced structure, the MINQUE and I-MINQUE estimators of variance components are asymptotically normal. Westfall (1987) has considered the MINQUE type estimators by taking identical values for the ratios of the a priori variance components to the error variance component and letting this common value tend to infinity. Westfall and Bremer (1994) have obtained cell means variance components estimates as special cases of the MINQUE estimates. A particularly simple form of the MINQUE estimator, as indicated by Rao (1972), arises when a priori weights σ_i^2 are chosen such that $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_p^2 = 0$ and $\sigma_e^2 = 1$. The estimator is commonly known as MINQUE(0). Hartley et al. (1978) have also obtained MINQUE estimates by treating all nonerror variance components to be zero; in which case the matrix $\boldsymbol{V}$ reduces to an identity matrix. These estimates are locally optimal when all nonerror variances are zero; otherwise, they are inefficient (see, e.g., Quass and Bolgiano, 1979).

MINQUEs and MIVQUEs like any other variance component estimators, may assume negative values. Rao (1972) proposed a modification of the MINQUE which would provide nonnegative estimates, but the resulting estimators would generally be neither quadratic nor unbiased. J. N. K. Rao and Subrahmaniam (1971), and J. N. K. Rao (1973) employed a modification of the MINQUE, resulting in truncated quadratic biased estimates of variance components. Brown (1978) discussed an iterative feedback procedure using residuals which ensures nonnegative estimation of variance components. P. S. R. S. Rao and Chaubey (1978) also considered a modification of the MINQUE by ignoring the condition for unbiasedness. They call the resulting procedure a minimum-norm quadratic estimation (MINQE), which also yields nonnegative estimates. Computational and other related issues of MINQE estimators have also been considered by Brockleban and Giesbrech (1984), Ponnuswamy and Subramani (1987) and Lee (1993). In as much as MINQE may entail large bias, Chaubey (1991) has considered nonnegative MINQE with minimum bias. Rich and Brown (1979) consider I-MINQUE estimators which are nonnegative. Nonnegative MINQUE estimates of variance components have also been considered by Massam and Muller (1985). Chaubey (1983) proposed a nonnegative estimator closest to MINQUE.

One difficulty with the MINQUE and MIVQUE procedures is that the expressions for the estimators are in a general matrix form and involve a number of matrix operations including the inversion of a matrix of order N (the number of observations). Since many variance component problems involve a large volume of data, this may be a serious matter. Schaeffer (1973) has shown that this problem may be eased somewhat by using Henserson's best linear unbiased predictor (BLUP) equations to obtain MINQUEs and MIVQUEs under normality. Liu and Senturia (1977) also discuss some computational procedures which reduce the number and order of matrix operations involved in the computation of MINQUE. They have developed a FORTRAN program with large capacity and high efficiency for the computation of the MINQUE vector. The program written for the UNIVAC 1110 computer requires 65K words of available memory and will handle linear models in which $1 + q + \sum_{i=1}^{p-1} m_i \leq 180$. Copies of the listing of the program are available from the authors.

Liu and Senturia (1977) reported that the MINQUE procedure is a rapidly convergent one; the estimates usually being obtained after two or three iterations. This is in contrast to the maximum likelihood method which provides only an implicit expression for the estimates, necessitating the use of approximations by iterative techniques. Wansbeck (1980) also reformulated the MINQUE estimates in such a manner that it requires an inversion of a matrix of order $m = \sum_{i=1}^{p} m_i$. In addition, Kaplan (1983) has shown the possibility of even further reduction in the order of the matrix to be inverted. Giesbrecht and Burrows (1978) have proposed an efficient method for computing MINQUE estimates of variance components for hierarchical classification models. Furthermore, Giesbrecht (1983), using modifications of the W-transformation, developed an efficient algorithm for computing MINQUE estimates of variance components and the generalized least squares (GLS) estimates of the fixed effects. Computational and other related issues of MINQUE and MIVQUE estimation have also been considered in the papers by P. S. R.S. Rao et al. (1981), Kleffé and Siefert (1980, 1986), and Lee and Kim (1989), among others. Finally, it should be remarked that although the theory of MINQUE estimation has generated a lot of theoretical interest and research activity in the field; the estimators have some intuitive appeal and under the assumption of normality reduce to well-known estimators, the use of *prior measure* is not well appreciated or understood by many statisticians.

10.11 MINIMUM MEAN SQUARED ERROR QUADRATIC ESTIMATION

For the general linear model in (10.10.2), LaMotte (1973a) has considered minimum mean squared error (MSE) quadratic estimators of linear combinations of variance components, i.e.,

$$\boldsymbol{\ell}'\boldsymbol{\sigma}^2 = \sum_{i=1}^{p} \ell_i \sigma_i^2, \tag{10.11.1}$$

for each of several classes of estimators of the form $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$. In the notation of Section 10.10, the classes of estimators being considered are

$$\begin{aligned}
C_0 &= \{\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} : \boldsymbol{A} \text{ unrestricted}\},\\
C_1 &= \{\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} : \boldsymbol{X}'\boldsymbol{A}\boldsymbol{X} = \boldsymbol{0}\},\\
C_2 &= \{\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} : \boldsymbol{A}\boldsymbol{X} = \boldsymbol{0}\},\\
C_3 &= \{\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} : \boldsymbol{X}'\boldsymbol{A}\boldsymbol{X} = \boldsymbol{0}, \text{tr}[\boldsymbol{A}\boldsymbol{V}_i] = \ell_i, i = 1, 2, \ldots, p\},
\end{aligned}$$

and

$$C_4 = \{\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} : \boldsymbol{A}\boldsymbol{X} = \boldsymbol{0}, \text{tr}[\boldsymbol{A}\boldsymbol{V}_i] = \ell_i, i = 1, 2, \ldots, p\}.$$

More specifically, the above classes of estimators are

(i) C_0 is the class of all quadratics;

(ii) C_1 is the class of all quadratics with expected value invariant to $\boldsymbol{\alpha}$;

(iii) C_2 is the class of all quadratics which are translation invariant;

(iv) C_3 is the class of all quadratics unbiased for $\boldsymbol{\ell}'\boldsymbol{\sigma}^2$;

(v) C_4 is the class of all quadratics, which are translation invariant and unbiased for $\boldsymbol{\ell}'\boldsymbol{\sigma}^2$.

A quadratic $Q_t(\boldsymbol{\alpha}, \boldsymbol{\sigma}^2)$ in the class $C_t (t = 0, 1, 2, 3, 4)$ is called "best" at $(\boldsymbol{\alpha}, \boldsymbol{\sigma}^2)$, provided that for any quadratic $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ in C_t,

$$\text{MSE}(Q_t(\boldsymbol{\alpha}, \boldsymbol{\sigma}^2)|\boldsymbol{\alpha}, \boldsymbol{\sigma}^2) \leq \text{MSE}(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}|\boldsymbol{\alpha}, \boldsymbol{\sigma}^2). \tag{10.11.2}$$

The best estimators in the class C_0, C_1, C_2, C_3, and C_4 as derived in LaMotte (1973a) are as follows.

(I) Best in C_0. The best estimator of $\boldsymbol{\ell}'\boldsymbol{\sigma}^2$ at $(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2)$ in C_0 is $Q_0(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2)$ defined by

$$\begin{aligned} Q_0(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2) &= \frac{\boldsymbol{\ell}'\boldsymbol{\sigma}_0^2}{\theta_0^2 + (N+2)(2\theta_0 + 1)} \boldsymbol{Y}'\{(2\theta_0 + 1)\boldsymbol{V}_0^{-1} - \boldsymbol{V}_0^{-1}\boldsymbol{X}\boldsymbol{\alpha}_0\boldsymbol{\alpha}_0'\boldsymbol{X}'\boldsymbol{V}_0^{-1}\}\boldsymbol{Y} \\ &= \frac{\boldsymbol{\ell}'\boldsymbol{\sigma}_0^2}{\theta_0^2 + (N+2)(2\theta_0 + 1)} \boldsymbol{Y}'\{\theta_0\boldsymbol{V}_0^{-1} + (\theta_0 + 1)(\boldsymbol{V}_0 + \boldsymbol{X}\boldsymbol{\alpha}_0\boldsymbol{\alpha}_0'\boldsymbol{X}')^{-1}\}\boldsymbol{Y}, \end{aligned}$$

where

$$\theta_0 = \boldsymbol{\alpha}_0'\boldsymbol{X}'\boldsymbol{V}_0^{-1}\boldsymbol{X}\boldsymbol{\alpha}_0, \qquad \boldsymbol{V}_0 = \boldsymbol{V}(\boldsymbol{\sigma}_0^2), \quad \text{and} \quad \boldsymbol{V}(\boldsymbol{\sigma}^2) = \text{Var}(\boldsymbol{Y}).$$

(II) Best in C_1. The best estimator of $\boldsymbol{\ell}'\boldsymbol{\sigma}^2$ at $(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2)$ in C_1 is $Q_1(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2)$ defined by

$$Q_1(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2) = \frac{\boldsymbol{\ell}'\boldsymbol{\sigma}_0^2}{\delta + 2} \boldsymbol{Y}'\boldsymbol{W}_0\boldsymbol{Y},$$

where

$$\delta = \text{tr}[\boldsymbol{W}_0\boldsymbol{V}_0] = \text{rank}(\boldsymbol{W}_0) = N - \text{rank}(\boldsymbol{X})$$

and

$$\boldsymbol{W}_0 = \boldsymbol{V}_0^{-1} - \boldsymbol{V}_0^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}_0^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}_0^{-1}$$

with

$$\boldsymbol{V}_0 = \boldsymbol{V}(\sigma_0^2).$$

(iii) Best in C_2. Q_1 is also in C_2 and is best at $(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2)$ in C_2.

(iv) Best in C_3. If $\boldsymbol{\ell}'\boldsymbol{\sigma}^2$ is estimable in C_3, then the best estimator of $\boldsymbol{\ell}'\boldsymbol{\sigma}^2$ at $(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2)$ is $Q_3(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2)$ defined by

$$Q_3(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2) = \boldsymbol{\ell}'\hat{\boldsymbol{\sigma}}^2,$$

where $\hat{\boldsymbol{\sigma}}^2$ is a solution of the consistent equation

$$\boldsymbol{G}_0\boldsymbol{\sigma}^2 = \boldsymbol{\psi}_0,$$

where

$\boldsymbol{G}_0$ is a $p \times p$ matrix with the (i, j)th element equal to $\mathrm{tr}[\boldsymbol{M}_i\boldsymbol{V}_j]$

and

$\boldsymbol{\psi}_0$ is a p-vector with the ith element equal to $\boldsymbol{Y}'\boldsymbol{M}_i\boldsymbol{Y}$,

with

$$\begin{aligned}\boldsymbol{M}_i &= \boldsymbol{W}_0\boldsymbol{V}_i\boldsymbol{W}_0 + \boldsymbol{W}_0\boldsymbol{V}_i\boldsymbol{H}_0^{-} + \boldsymbol{H}_0^{-}\boldsymbol{V}_i\boldsymbol{W}_0, \quad i = 1, 2, \ldots, p,\\ \boldsymbol{W}_0 &= \boldsymbol{V}_0^{-1} - \boldsymbol{V}_0^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}_0^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}_0^{-1},\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{H}_0^{-} = \boldsymbol{H}_0^{-}(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2) &= \boldsymbol{V}_0^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}_0^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}_0^{-1}\\ &\quad - (1 + \theta_0)^{-1}\boldsymbol{V}_0^{-1}\boldsymbol{X}\boldsymbol{\alpha}_0\boldsymbol{\alpha}_0'\boldsymbol{X}'\boldsymbol{V}_0^{-1},\end{aligned}$$

with

$$\theta_0 = \hat{\boldsymbol{\alpha}}_0'\boldsymbol{X}'\boldsymbol{V}_0^{-1}\boldsymbol{\alpha}_0.$$

(v) Best in C_4. If $\boldsymbol{\ell}'\boldsymbol{\sigma}^2$ is estimable in C_4, then the best estimator of $\boldsymbol{\ell}'\boldsymbol{\sigma}^2$ at $(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2)$ is $Q_4(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2)$ defined by

$$Q_4(\boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0^2) = \boldsymbol{\ell}'\hat{\boldsymbol{\sigma}}^2,$$

where $\hat{\boldsymbol{\sigma}}^2$ is a solution of the consistent equation

$$\boldsymbol{G}_0\boldsymbol{\sigma}^2 = \boldsymbol{\psi}_0,$$

where

$\boldsymbol{G}_0$ is a $p \times p$ matrix with the (i, j)th element equal to $\text{tr}[\boldsymbol{M}_i \boldsymbol{V}_j]$

and

$\boldsymbol{\psi}_0$ is a p-vector with the ith element equal to $\boldsymbol{Y}'\boldsymbol{M}_i\boldsymbol{Y}$,

with

$$\boldsymbol{M}_i = \boldsymbol{W}_0\boldsymbol{V}_i\boldsymbol{W}_0$$

and

$$\boldsymbol{W}_0 = \boldsymbol{V}_0^{-1} - \boldsymbol{V}_0^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}_0^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}_0^{-1}.$$

LaMotte (1973a) presents extensive derivations of the above results and also gives attainable lower bounds on MSEs of the estimator in each class. Since the property of ‘bestness’ is a local property, guidelines for amending and combining the best quadratics in order to achieve more uniform performance for the entire $(\boldsymbol{\alpha}, \boldsymbol{\sigma}^2)$ parameter space are presented. It is shown that whenever a uniformly best quadratic estimator exists, it is given by a “best” estimator. It should be noted that the best estimator in the class C_4 is C. R. Rao’s alternative form of MINQUE or MIVQUE under normality. Minimum mean square quadratic estimators (MIMSQE) are also considered by Rao (1971b). Chaubey (1980) considers minimum-norm quadratic estimators (MINQE) in the classes C_0, C_1 and C_2; and Volaufová and Witkovsky (1991) consider quadratic invariant estimators of the linear functions of variance components with locally minimum mean square error using least squares approach. MSE efficient estimators of the variance components have also been considered by Lee and Kapadia (1992).

10.12 NONNEGATIVE QUADRATIC UNBIASED ESTIMATION

LaMotte (1973b) has investigated the problem of nonnegative quadratic unbiased estimation of variance components. In particular, LaMotte (1973b) has characterized those linear functions of variance components in linear models for which there exist unbiased and nonnegative quadratic estimators. Pukelsheim (1981a) also presents some conditions for the existence of such estimators. In this section, we discuss some of these results briefly.

For the general linear model in (10.10.2), we know from Section 10.10 that the necessary and sufficient conditions that a linear function of the variance components, i.e.,

$$\boldsymbol{\ell}'\boldsymbol{\sigma}^2 = \sum_{i=1}^{p} \ell_i \sigma_i^2, \tag{10.12.1}$$

be estimated unbiasedly by a quadratic form $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ is that

$$\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X} = \boldsymbol{0} \quad \text{and} \quad \text{tr}[\boldsymbol{A}\boldsymbol{V}_i] = \ell_i. \tag{10.12.2}$$

Further, if the estimator is to be nonnegative, we require that

$$\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} \geq 0,$$

i.e., $\boldsymbol{A}$ be a nonnegative definite.

Now, we state a lemma due to LaMotte (1973a) that guarantees nonnegative unbiased estimability.

Lemma 10.12.1. *In order that there exist a nonnegative quadratic $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ unbiased for $\boldsymbol{\ell}'\boldsymbol{\sigma}^2 = \sum_{i=1}^{p} \ell_i \sigma_i^2$, it is necessary and sufficient that there exists a matrix $\boldsymbol{C}$ such that*

(i) $$\boldsymbol{A} = \boldsymbol{R}\boldsymbol{C}\boldsymbol{C}'\boldsymbol{R} \tag{10.12.3}$$

and

(ii) $$\text{tr}[\boldsymbol{C}'\boldsymbol{R}\boldsymbol{V}_i\boldsymbol{R}\boldsymbol{C}] = \ell_i, \quad i = 1, \ldots, p, \tag{10.12.4}$$

where

$$\boldsymbol{R} = \boldsymbol{V}^{-1}[\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}] \quad \textit{and} \quad \boldsymbol{V} = \text{Var}(\boldsymbol{Y}). \tag{10.12.5}$$

Note that the matrix $\boldsymbol{R}$ is the same as in Rao's MINQUE *procedure defined by* (10.10.25).

Proof. See LaMotte (1973b). □

An important consequence of Lemma 10.12.1 is the following corollary.

Corollary 10.12.1. *If for some $i\,(i = 1, \ldots, p)$, $\boldsymbol{V}_i$ is positive definite and $\ell_i = 0$, then the only vector $\boldsymbol{\ell}$ for which there is a nonnegative quadratic unbiased estimator of $\boldsymbol{\ell}'\boldsymbol{\sigma}^2$ is $\boldsymbol{\ell}' = \boldsymbol{0}$.*

As we have seen, for the analysis of variance model,

$$\boldsymbol{V}_i = \boldsymbol{U}_i\boldsymbol{U}_i' \quad \textit{for some } \boldsymbol{U}_i, \quad i = 1, \ldots, p-1,$$

$$V_p = I_N,$$

and

$$\sigma_i^2 \geq 0, \quad i = 1, \ldots, p.$$

Thus it follows from Corollary 10.12.1 that the only individual variance component which can be estimated unbiasedly by a nonnegative quadratic is σ_p^2 (the error variance component), and even σ_p^2 is so estimable only if all V_is ($i = 1, \ldots, p-1$) are singular. (Note that $V_p = I_N$ is nonsingular.)

For a survey of methods of estimation, without the restriction of nonnegativity of the quadratic estimator, see Kleffé (1977b). Although nonnegative quadratic unbiased estimators of variance components do not exist, Kleffé and J. N. K. Rao (1986) have investigated the existence of asymptotically unbiased nonnegative quadratic estimators. Similarly, Baksalary and Molinska (1984) have investigated nonnegative unbiased estimability of a linear combination of two variance components and Pukelsheim (1981a, 1981b) investigated the existence of nonnegative quadratic unbiased estimators using convex programming. In particular, Pukelsheim (1981a, 1981b) characterized nonnegative estimability of a linear combination of the variance components, $\sum \ell_i \sigma_i^2$, by means of the natural parameter set in the residual model. This leads to an alternative formulation that in the presence of a quadratic subspace condition either the usual unbiased estimators of the individual variance components, $\hat{\sigma}_i^2$, provide an unbiased nonnegative definite quadratic estimator, $\sum \ell_i \hat{\sigma}_i^2$, or no such estimator exists. The result was proven by Mathew (1984). In addition, for the same problem, Gnot et al. (1985) characterized nonnegative admissible invariant estimators. For some other related works on nonnegative estimation of variance components, see Mathew (1987), Mathew et al. (1992a, 1992b), Gao and Smith (1995), and Ghosh (1996).

10.13 OTHER MODELS, PRINCIPLES AND PROCEDURES

In addition to methods of estimation for the normal linear models described in earlier sections, there are a number of other models, principles and procedures and we will briefly outline some of them here.

10.13.1 COVARIANCE COMPONENTS MODEL

In the development of this text, we have been mainly concerned with the random effect models involving only variance components. Covariances between any two elements of a random effects or between every possible pair of random effects are assumed to be zero. The generalization of the variance components models to allow for covariances between any random effects leads to the so-called covariance components models. Covariance components models are useful in a variety of applications in biology, genetics, education, among others. Covariance components models are discussed in the works of Henderson (1953), C. R. Rao (1971a, 1972), Henderson (1986), and Searle et al. (1992, Section 11.1). In addition, there are several papers that describe the variance components and the related estimation procedures in terms of the covariances of the random effects (see, e.g., Smith and Murray, 1984; Green, 1988; Hocking et al., 1989); and in some cases a negative estimate can be interpreted as a negative covariance. Rocke (1983) suggested a robust analysis for a special class of problems.

10.13.2 DISPERSION-MEAN MODEL

In many situations, the general mixed model can be restructured in the form of a linear model in which the vector of mean is the vector of variance components parameters of the model to be estimated. It is called the dispersion-mean model and was first introduced by Pukelsheim (1974). The notion of the common structure for the mean and variance (mean-dispersion correspondence) has been elaborated by Pukelsheim (1977c). For a discussion of variance components estimation based on dispersion-mean model and other related works, see Pukelsheim (1976), Malley (1986), and Searle et al. (1992, Chapter 12).

10.13.3 LINEAR MODELS FOR DISCRETE AND CATEGORICAL DATA

The random effects models considered in this text are based on continuous data. In recent years, there has been some work on construction and estimation of models for binary, discrete, and categorical data. Cox (1955) gives some simple methods for estimating variance components in multiplicative models entailing Poisson variables. Landis and Koch (1977) discuss estimation of variance components for a one-way random effects model with categorical data. Similarly, binary, count, discrete, logit, probit, generalized linear and log-linear models have been discussed in the works of Hudson (1983), Harville and Mee (1984), Ochi and Prentice (1984), Stiratelli et al. (1984), Gilmour et al. (1985), Wong and Mason (1985), Gianola and Fernando (1986a), Zeger et al. (1988) , Conaway (1989), Zeger and Karim (1991), Hedeker and Gibbons (1994), McDonald (1994), Chan and Kuk (1997), Gibbons and Hedeker (1997), Lee (1997), Lin (1997), and Omori (1997), among others.

10.13.4 HIERARCHICAL OR MULTILEVEL LINEAR MODELS

A class of models closely related to variance components models considered in this text are linear models involving modeling in a hierarchy. Hierarchical or multilevel linear models constitute a general class of linear models which enable a more realistic modeling process in many common situations encountered in biology (growth curve fitting, analysis of genetic experiments), in educational research (achievement studies, school effectiveness), in social sciences (survey analysis, marketing research, contextual problem analysis), and in many other fields in which information is collected using observational or experimental studies that lead to complex databases. This formulation assumes a set of elementary level units nested or grouped within level two units, which may further be nested within level three units and so on. Hierarchical linear models are discussed in the works of Laird and Ware (1982), Goldstein (1995), Longford (1993), Hedeker and Gibbons (1994), Morris (1995), Kreft and deLeeuw (1998), Heck and Thomas (1999), and Raudenbush and Bryk (2002), among others.

10.13.5 DIALLEL CROSS EXPERIMENTS

The diallel cross, used to study the genetic properties of a set of inbred lines, is one of the most popular mating designs used in animal and plant breeding experiments. It is a very useful method for conducting animal and plant breeding experiments, especially for estimating combined ability effects of lines. A diallel crossing system consists of all possible crosses from a single set of parents. Diallel crosses in which all possible distinct crosses in pairs among the available lines are taken are called complete diallel crosses. Diallel crosses in which only a fraction of all possible crosses among the available lines are taken are called partial diallel crosses. Reciprocal crosses are utilized in an attempt to separate genetically determined variation. Yates (1947) first developed a method of analysis for diallel mating designs. Griffing (1956) introduced four choices of diallel mating system, known as Methods 1, 2, 3, and 4, and presented a detailed analysis for these designs laid out in a complete block design. In addition, Griffing himself developed the ANOVA method for the estimation of variance components for all the four methods. Diallel crosses are generally conducted using a completely randomized design or a randomized complete block design; however, incomplete block designs are also common. By diallel analysis, both additive and dominance variance components can be estimated. Some other approaches to diallel analysis are due to Hayman (1954), Topham (1966), and Cockerham and Weir (1977). Hayman (1954) developed and elaborated a method of analysis for studying the nature of gene action based on assumptions such as no genetic-environmental interaction. Topham (1966) considered maternal effects and maternal-paternal interaction effects in the same model. Cockerham and Weir (1977) introduced the biomodel of diallel crosses which is more attuned to the biological framework and provides a method for estimating maternal and paternal variance components. In diallel cross experiments, the estimation of general combining abilities and maternal effects has been commonly carried out on the basis of the fixed effects model. In most applications, however, the genetic and environment components are random leading to imprecise estimates. Recent research is being directed toward developing algorithms for obtaining the best linear unbiased predictors (BLUP) by using the methodology for the estimation of random effects in the mixed effects model.

Further developments on the estimation of variance components based on a biomodel of diallel crosses can be found in the works of Venkateswarlu (1996), Venkateswarlu and Ponnuswamy (1998), and Venkateswarlu et al. (1998).

10.13.6 PREDICTION OF RANDOM EFFECTS

In many applications of random effects models in biology, genetics, psychology, education, and other related fields, the interest often centers on predicting the (unobservable) realized value of a random effect. For example, in animal breeding, the researcher wants to predict the genetic merit of a dairy bull from the data on milk production of his daughters; in psychology, one may want to predict an individual's intelligence based on data from IQ scores. The term prediction is used for estimation of random effects to emphasize the distinction between a fixed and a random effect. Note that a fixed effect is considered to be a constant that we wish to estimate; but a random effect is just one of the infinite number of effects belonging to a population and we wish to predict it. Three methods of prediction of random effects which have received some attention in the published literature are, best prediction (BP), best linear prediction (BLP), and best linear unbiased prediction (BLUP). The BP method consists of deriving a best predictor in the sense that it minimizes the mean squared error of prediction. The BLP derives the best predictor by limiting the class of predictors that are linear in the observations. The BLUP attempts to derive the best linear predictor that is unbiased. BLUPs are linear combinations of the responses that are unbiased estimators of the random effects and minimize the mean squared error. In the prediction of random effects using BLUP, often, the variance components are unknown and need to be estimated. The traditional approach consists of first estimating the variance components and then using the estimated variances in the equation for the BLUP as if they were true values. This approach is often known as empirical BLUP. The procedures for BP, BLP, BLUP, and empirical BLUP are discussed in the works of Henderson et al. (1959), Golberger (1962), Henderson (1973, 1975, 1984), Harville (1990), Harville and Carriquiry (1992), and Searle et al. (1992, Chapter 7), among others. For an excellent review of BLUP methodology and related topics, see Kennedy (1991) and Robinson (1991).

10.13.7 BAYESIAN ESTIMATION

In the Bayesian approach, all parameters are regarded as "random" in the sense that all uncertainty about them should be expressed in terms of a probability distribution. The basic paradigm of Bayesian statistics involves a choice of a joint prior distribution of all parameters of interest that could be based on objective evidence or subjective judgment or a combination of both. Evidence from experimental data is summarized by a likelihood function, and the joint prior distribution multiplied by the likelihood function is the (unnormalized) joint posterior density. The (normalized) joint posterior distribution and its marginals form the basis of all Bayesian inference (see, e.g., Lee, 1998). The use of Bayesian methods in estimating variance components for some balanced random models was considered in Volume I. The seminal paper of Lindley and Smith (1972) provided a general formulation of a linear hierarchical Bayesian model that established a link between the Bayesian approach and the classical formulation of mixed models. Many recent developments in the Bayesian analysis of the mixed effects model took place in conjunction with animal breeding studies and appeared in genetics journals. Gianola and Fernando (1986b), Gianola and Foulley (1990), Wang et al. (1993), and Theobold et al. (1997) summarize posterior distributions resulting from several different Bayesian mixed models and discuss computational aspects of the problem. More recently, Gönen (2000) presents a Bayesian approach to the analysis of random effects in the mixed linear model in terms intraclass correlations as opposed to the traditional reparametrization in terms of variance components. Further developments on Bayesian methodology in estimating variance components can be found in the works of Box and Tiao (1973), Rudolph (1976), Gharaff (1979), Rajagopalan (1980), Rajagopalan and Broemeling (1983), Broemeling (1985), Cook et al. (1990), Schervish (1992), Searle et. al. (1992, Chapter 9), Harville and Zimmerman (1996), Sun et al. (1996), and Weiss et al. (1997), among others.

10.13.8 GIBBS SAMPLING

This is a popular procedure belonging to the family of Markov Chain Monte Carlo algorithms. The procedure is an iterative one and involves sampling of the parameters of a statistical model one by one from the joint density function which is conditional on the previous set of parameters already sampled. At each stage of iteration, the simulated posterior distribution is obtained and the sampling is continued until the distribution is considered to have converged to the true posterior. To illustrate the procedure, consider a one-way random effects model involving the parameters: overall mean (μ), between group variance (σ_α^2), and the error variance (σ_e^2). The procedure then proceeds as follow: (i) Specify an initial set of values of the parameters, say, $\mu_0, \sigma_{\alpha,0}^2$, and $\sigma_{e,0}^2$. The choice of initial values is rather arbitrary, but the convergence is much more rapid if they are closer to realistic values. (ii) Sample each parameter from its posterior distribution, conditional on the previous values sampled for other parameters. Thus μ_1 is sampled from $p_\mu(\mu|\sigma_\alpha^2 = \sigma_{\alpha,0}^2, \sigma_e^2 = \sigma_{e,0}^2, \boldsymbol{Y})$; $\sigma_{\alpha,1}^2$ is sampled from $p_{\sigma_\alpha^2}(\sigma_\alpha^2|\mu = \mu_1, \sigma_e^2 = \sigma_{e,0}^2, \boldsymbol{Y})$; $\sigma_{e,1}^2$ is sampled from $p_{\sigma_e^2}(\sigma_e^2|\mu = \mu_1, \sigma_\alpha^2 = \sigma_{\alpha,1}^2, \boldsymbol{Y})$. The parameter values, $\mu_1, \sigma_{\alpha,1}^2, \sigma_{e,1}^2$, constitute the first set of iteration. Sample a second set of parameter values from their respective posterior distributions conditional on the preceding set of parameters: Thus μ_2 is sampled from $p_\mu(\mu|\sigma_\alpha^2 = \sigma_{\alpha,1}^2, \sigma_e^2 = \sigma_{e,1}^2, \boldsymbol{Y})$; $\sigma_{\alpha,2}^2$ is sampled from $p_{\sigma_\alpha^2}(\sigma_\alpha^2|\mu = \mu_2, \sigma_e^2 = \sigma_{e,2}^2, \boldsymbol{Y})$; $\sigma_{e,2}^2$ is sampled from $p_{\sigma_e^2}(\sigma_e^2|\mu = \mu_2, \sigma_\alpha^2 = \sigma_{\alpha,2}^2, \boldsymbol{Y})$. The parameter values constitute the second set of iterations. Continue the iterative process until the convergence is achieved. After a suitable number of iterations, we obtain sample values from the distribution of any posterior component that can be used to derive the required set of estimates or any other characteristics of the distribution. Gibbs sampling is a complex and computationally demanding procedure and a very large number of iterations (hundreds if not thousands) may be required to ensure that convergence has been achieved. It is more useful for small and moderate size samples and when used in conjunction with a likelihood-based algorithm, such as EM. The procedure can be carried out using the package BUGS (see Appendix N). A comprehensive discussion with applications can be found in Gilks et al. (1993). Rates of convergence for variance component models are discussed by Rosenthal (1995). Applications to variance component estimation are considered by Baskin (1993), Kasim and Raudenbush (1998), Burton et al. (1999); and Bayesian analysis on variance components is illustrated in the works of Gelfand et al. (1990), Gelfand and Smith (1991), Wang et al. (1993), and Hobert and Casella (1996), among others.

10.13.9 GENERALIZED LINEAR MIXED MODELS

Generalized linear mixed models (GLMM) are generalizations of the fixed effects generalized linear models (GLM) to incorporate random coefficients and covariance patterns. GLMs and GLMMs allow the extension of classical normal models to certain types of nonnormal data with a distribution belonging to the exponential family; and provide an elegant unifying framework for a wide range of seemingly disparate problems of statistical modeling and inference, such as analysis of variance, analysis of covariance, normal, binomial and Poisson regressions, and so on. GLMs and GLMMs provide a flexible parametric approach for the estimation of covariate effects with clustered or longitudinal data. They are particularly useful for investigating multiple sources of variation, including components associated with measured factors, such as covariates, and variation attributed to measured factors or random effects, and provide the experimenter a rich and rewarding modeling environment. These models employ the concept of a link function as a way of mapping the response data from their original scale to the real scale $(-\infty, +\infty)$. For example, binary response data with parameter $p (0 < p < 1)$ employs the link function, $\log(\mu/(1 - \mu))$, to map this range to the real scale. The use of a link function allows the model parameters to be included in the model linearity in the same manner as the normal models. Both fixed and mixed effects models are fitted based on maximizing the likelihood for model parameters. Recent computational advances have made the routine fitting of the models possible and there are now numerous statistical packages available for fitting these models. GLIM and S-PLUS are especially designed for this purpose, while other packages such as SAS, SPSS, and BMDP have routines that facilitate fitting many types of generalized linear models. GLMs and GLMMs are relatively a new class of models and are still not widely used among researchers in substantive fields. The interested reader is referred to the works of McCullagh and Nelder (1989), Breslow and Clayton (1993), Littell et al. (1996), McCulloch and Searle (2001), and Dobson (2002) for further discussions and details. Estimation in GLMs with random effects is discussed by Schall (1991). A more applied treatment with application to medicine is given by Brown and Prescott (1999). For a brief overview of GLMMs, see Stroup and Kachman (1994).

10.13.10 NONLINEAR MIXED MODELS

Nonlinear mixed models are a newer family of models for analyzing experimental and research data. These are similar to mixed effects linear models where the mean response is assumed to consist of two parts: a mean function with fixed parameters and a set of random effects added to the mean function. The mean function is allowed to be nonlinear in the parameters. The covariance structure of the observation vector is defined by the random effects included in the model and our interest lies in estimating parameters of the model. This type of model is useful for observational studies as well as for designed experiments since the treatment levels need not be the same for different experimental units. Such models are often appropriate for analyzing data from nested or split-plot designs used in agricultural and environmental research. Nonlinear functions such as Weibull functions have been widely used to model the effect of ozone exposure on the yield of many crops. The model is related to nonlinear random coefficient models where coefficients are assumed to be random variables. Methods of estimation of variance components for nonlinear models have been described by Gumpertz and Pantula (1992), Gumpertz and Rawlings (1992), among others. It should be noted that GLMMs considered in Section 10.13.9 constitute a proper set of NLMMs. Detailed coverage of NLMMs for longitudinal data is given by Giltinan and Davidian (1995) and Vonesh and Chinchilli (1997). Solomon and Cox (1992) provide a discussion of nonlinear components of variance models.

10.13.11 MISCELLANY

Seely (1970a, 1970b, 1971) employed the quadratic least squares (QLS) theory and the notion of quadratic subspace to estimate variance components. Seely (1972, 1977) also used the notion of quadratic subspaces in the derivation of completeness of certain statistics for a family of multivariate normal distributions. Using the QLS approach of Seely (loc. cit.), Yuan (1977) developed a procedure to obtain the invariant quadratic unbiased estimator as a particular case of QLS principle and has shown that certain well-known procedures for estimating variance components, like symmetric sums, MINQUE, etc., are special cases of the QLS procedure by choosing appropriate weights. Following Yuan (1977) and Mitra (1972), Subramani (1991) has considered QLS, weighted QLS, and Mitra type estimators and compared them using different optimality criteria, namely, D-optimality, T-optimality, and M-optimality. It has been shown that Mitra type estimators have better optimal properties. Hartung (1981) developed generalized inverse operators to minimize the estimation bias subject to nonnegativity of the variance components, but the method is not order preserving for estimators of linear combinations of variance components. Verdooren (1980, 1988) introduced the concept of permissible estimation and underscored its importance as a necessary condition for an estimation procedure. Verdooren (1988) presented a unified procedure for the derivation of estimators of the variance components using the least squares theory and showed that they are unbiased but not always nonnegative. Under the condition of invariance, the least squares estimators are shown to be the MINQUE, which under the assumption of the multivariate normality for the observation vector are the same as the MIVQUE. More recently, Hoefer (1998) has reviewed a large body of literature on variance component estimation in animal breeding.

10.14 RELATIVE MERITS AND DEMERITS OF GENERAL METHODS OF ESTIMATION

The relative merits and demerits of different methods of estimation of variance components can be summarized as follows:

(i) The analysis of variance or Henderson's Method I commends itself because it is the obvious analogue of the ANOVA for balanced data and is relatively simple to use. It produces unbiased estimates of variance components which under the assumption of normality have known results for unbiased estimators of sampling variances. Its disadvantage lies in the fact that some of its terms are not sums of squares (and hence may be negative) and it produces biased estimates in mixed models.

(ii) Henderson's Method II corrects the deficiency of Method I and is uniquely defined, but it is difficult to use. In addition, the method cannot be used when there are interactions between fixed and random effects, and no analytic expressions are available for sampling variances of estimators.

(iii) The fitting-constants method or Henderson's Method III uses reductions in sums of squares, due to fitting different submodels, that have noncentral chi-square distributions in the fixed effects model. It produces unbiased estimates in mixed models, but it can give rise to more quadratics than there are components to be estimated and involves extensive numerical computations. No closed form expressions for sampling variances are generally available, though they can be calculated through a series of matrix operations using estimated values for the variance components. In addition, it has been shown that, for at least some unbalanced designs, there are estimators in the class of locally best translation invariant estimators that have uniformly smaller variance than Method III estimators.

(iv) The analysis of means method is straightforward to use and yields estimators that are unbiased. However, this is only an approximate method with the degree of approximation depending on the extent to which the unbalanced data are not balanced. Furthermore, the method is applicable only when every subclass of the model contains at least one observation.

(v) The symmetric sums of products (SSP) method has computational simplicity and utilizes all possible products of observations and their means. It yields unbiased estimators by construction. However, the procedure leads to estimates that do not have the μ-invariance property. The modified procedure, based on the symmetric sums of squares of differences rather than products, remedies this fault; but it has an even more serious defect, i.e., it yields estimators that are inquadmissible. Harville (1969a) showed that in the case of a one-way random effects model, ANOVA estimators of variance components have uniformly smaller variance than the modified SSP estimators. Moreover, there is not much difference between the ANOVA estimators and modified SSP estimators in terms of computational simplicity.

(vi) The maximum likelihood or restricted maximum likelihood methods of estimation have strong theoretical basis and yield estimates with known optimal properties. Furthermore, ML estimates of functions of variance components, such as heritability, are readily obtained, along with approximate standard errors. It has further been shown that for certain experimental designs, there exist variance components estimators, closely linked to the ML estimators, that have uniformly smaller variance than the ANOVA estimators (see Olsen et al. 1976). However, the estimators cannot be obtained explicitly and for large data sets may involve extensive and costly computations with iterative calculations often converging very slowly. In addition, ML estimates are biased downwards, sometimes quite markedly, with the bias being larger when the number of parameters in a model is a substantial fraction of the number of data items. The REML yields variance components estimates that are unaffected by the fixed effects by taking into account the degrees of freedom used for estimating fixed effects. It should also be noted that, although the difference between the ML and REML estimation is often quite small, each procedure has slightly different properties. Furthermore, for balanced designs, the REML gives the same results as the ANOVA procedure provided the estimates are nonnegative; but little is known about its properties for unbalanced data. The coincidence between the REML and ANOVA estimates for balanced data when the estimates are nonnegative and the possibility of limited replication in the higher strata of a design provide compelling reasons for preferring REML. It has also been found that REML estimators do not seem to be as sensitive to outliers in the data as are ML estimators (Verbyla, 1993). Huber et al. (1994) recommended the use of REML for mating design data structures typical in analysis problems in quantitative forest genetics, basing his conclusion on a simulation study, and noted that it has most desirable properties in terms of variance, MSE, and bias in comparison to MINQUE, MIVQUE, ML, and Henderson Method III. Finally, it should be noted that the optimal properties of the ML estimation are large sample properties, based on asymptotic arguments, and are generally not applicable in many experimental situations involving small samples.

(vii) The MINQUE and MIVQUE procedures are quite general and are applicable to all experimental situations. Furthermore, MIVQUE or BQUE has an intuitive appeal in the estimation of variance components similar to the BLUE for fixed effects. Unfortunately, MINQUEs and MIVQUEs are, in general, functions of the unknown variance components and require a priori knowledge of the variance components to be estimated. Since in application, the variance components are unknown, the MINQUEs and MIVQUEs are, in general, also unknown. This difficulty is alleviated using iterative or I-MINQUE, but the resultant estimators are neither unbiased nor minimum-variance. Another difficulty with the MINQUE and MIVQUE procedures is that the expressions for the estimators are in a general matrix form and involve the inversion of a matrix of order N (the number of observations). Since many variance component estimation problems involve large volumes of data, this may be a serious matter. However, there now exist many efficient methods of computing MINQUE and MIVQUE estimators which involve the inversion of a matrix of much lower order. Finally, it should be mentioned that for balanced data MINQUE under the Euclidean norm reduces to ANOVA estimation which truncated at zero is equivalent to REML under the assumption of normality of the random effects when the estimates are nonnegative.

Most of the procedures discussed in this chapter yield unbiased estimators and reduce to the ANOVA estimators for balanced data. However, they can all produce negative estimates. Rao (1972) proposed a modification of MINQUE, which would provide nonnegative estimates; but the resulting estimators would generally be neither quadratic nor unbiased. In the following section we consider the problem of the comparison of designs and estimators. The results on analytic and numerical comparisons of variances and mean square errors of different estimators for various experimental situations will be discussed in subsequent chapters.

10.15 COMPARISONS OF DESIGNS AND ESTIMATORS

The term 'design' has commonly been associated with the estimation of fixed effects in a given linear model. However, in a random or mixed effects model, the quality of estimation of variance components to a large extent depends on the design used to generate the response data. Moreover, for the most part, the choice of a design is related to some optimality criterion that depends on the particular method of estimation, the model used, and the values of the variance components themselves. In most experimental situations involving joint estimation of variance components, it is rather a common practice to use balanced designs for the reasons of simplicity of the analysis and interpretation of data under the standard normal theory assumption. However, under the constraint of limited experimental resources, the balanced plans may produce estimates of certain important parameters with comparatively low precision. For example, in a two-way classification with 10 rows and 10 columns and two observations per cell, there are only nine degrees of freedom for the row and column mean squares, in contrast to 100 degrees of freedom for the residual error. Thus the row and column components of variance, which are often large and of much greater interest, are estimated with comparatively low precision; while the error variance component, which is often small and of lesser interest, is estimated with comparatively higher precision. Similarly, in a balanced nested design, the degrees of freedom are too heavily concentrated in the last stage. For example, in the $5 \times 2 \times 2$ design, the variance of the first stage has only four degrees of freedom. In order to have 10 degrees of freedom in the first stage, it will require a total of 88 observations. In general, in order to increase the degrees of freedom associated with the first stage without increasing the size of the experiment, a design with unbalanced arrangement is required. For a further discussion of this problem, the reader is referred to Davies and Goldsmith (1972, Appendix 6D, pp. 168–173), who made approximate comparisons of the precision of five alternative designs each comprising 48 observations.

Thus, as mentioned earlier in Chapter 9, there are situations when the researcher may purposely choose an unbalanced plan in order to estimate all or certain specified functions of variance components with a desired level of precision. For a given experimental layout and cost of experimentation, there are usually many possible arrangements to choose from. On the other hand, variance components analysis from an unbalanced configuration is usually quite complicated. For example, the variances of the variance component estimators for the model in (10.1.1) are tractable only under the assumption of normality. Furthermore, as in the case of a balanced model, the variances themselves are functions of the true variance components. To study the behavior of such variances in terms of their being functions of the total number of observations, the number of levels of each factor, the number of observations in each cell, and of the variance components themselves appears to be an enormous task. The comparison of such functions with the equally complex functions that are variances of other estimators adds further to the complexity of the problem. Thus the analytic comparison of sampling variances of different estimators is beset with difficulties. However, Harville (1969b) has been able to obtain explicit expressions for the differences between the variances of ANOVA estimators and fitting-constants-method estimators for balanced incomplete block designs. These differences are functions of the variance components and thus can be compared for specified values of these components. Another result on analytic comparison seems to be that of Harville (1969a), where he notes that using Theorem 2 of Harville (1969c), it can be shown that the ANOVA estimators of σ_e^2 and σ_α^2 in the model in (11.1.1) have uniformly smaller variance than the estimators based on symmetric sums of squares of differences. Inasmuch as the analytic comparison of estimators appears fruitless, the other open recourse is that of numerical comparison. Unfortunately, such numerical studies are difficult to carry out and the amount of computation required to obtain numerical results may be prohibitively large. Although the literature on variance components is rather quite extensive, the number of publications devoted to design aspects is somewhat limited. In the succeeding chapters, we will discuss the results of some empirical studies on comparisons of designs and estimators for each one of the crossed and nested models separately.

10.16 METHODS OF HYPOTHESIS TESTING

In many experimental situations involving the mixed effects model, the experimenter wishes to determine if there is evidence to conclude that a fixed effect has a nonnull value or a particular variance component is greater than zero; i.e., she wishes to test the hypothesis $H_0 : \sigma_i^2 = 0$ vs. $H_1 : \sigma_i^2 > 0$. In this section, we briefly consider the problem of hypothesis testing for fixed and random factors involving unbalanced designs.

We have seen in Volume I that for most balanced models, the ratio of any two mean squares has sampling distribution proportional to the F-distribution and the usual F-tests for fixed effects and variance components are unbiased and optimum. In situations where there are no suitable mean squares to be used as the numerator and denominator of the F-ratio, approximate F-tests based on the Satterthwaite procedure provide a simple and effective alternative. For unbalanced models, however, the sums of squares in the analysis of variance table are no longer independent nor do they have a chi-square type distribution although for some special cases certain sets of sums of squares may be independent. An exception to this rule is the residual or error sum of squares which is always independent of the other sums of squares and has a scaled chi-square distribution. Testing contrasts of even a single fixed effect factor is a problem since the estimated error variances are not sums of squares with chi-square distributions. Giesbrecht and Burns (1985) proposed performing t-tests on selected orthogonal contrasts that are not statistically independent by assuming a chi-square to the distribution of variances of contrast estimates and estimating the degrees of freedom using Satterthwaite's (1946) procedure. The results of a Monte Carlo simulation study show that the resulting tests have rather an adequate performance. Similarly, for a single fixed-effect factor, McLean and Saunders (1988) used t-tests for contrasts involving levels of both fixed and random effects. On the other hand, the problem of simultaneous testing of fixed effects is even more complex. Berk (1987) proposed the Wald type statistic as a generalization of the Hotelling T^2, but the theoretical distribution of the test statistic is rather difficult to evaluate. For some further discussions and proposed solutions to the problem, the interested reader is referred to Brown and Kempton (1994), Welham and Thompson (1997), and Elston (1998).

To test for random-effect factors, any factor with expected mean square equal to $\sigma_e^2 + n_0\sigma_i^2$, where σ_i^2 is the corresponding variance component, the test statistic for the hypothesis $H_0 : \sigma_i^2 = 0$ vs. $H_1 : \sigma_i^2 > 0$ can be based on the ratio of the mean square to the error mean square and provides an exact F-test. Mean squares with expectations involving linear combinations of several variance components cannot be used to obtain test statistics having exact F-distributions. This is so since under the null hypothesis, as indicated above, we do not have two mean squares in the analysis of variance table that estimate the same quantity. Furthermore, as noted earlier, the mean squares other than the error mean square are not distributed as a multiple of a chi-square random variable and they are not statistically independent of other mean squares. In such situations, a common procedure is to ignore the assumption of independence and chi-squaredness and construct an approximate F-test using synthesis of mean squares based on the Satterthwaite procedure.

An alternative approach is to employ the likelihood-ratio test which is based on the ratio of the likelihood function under the full model to the likelihood under the null condition. For the general linear model in (10.7.1), the likelihood function is

$$L(\boldsymbol{\alpha}, \sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2) = \frac{\exp\left\{-\frac{1}{2}(\boldsymbol{Y} - \boldsymbol{X\alpha})'\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X\alpha})\right\}}{(2\pi)^{\frac{1}{2}N}|\boldsymbol{V}|^{\frac{1}{2}}},$$

where $\boldsymbol{V} = \prod_{i=1}^{p} \sigma_i^2 \boldsymbol{U}_i \boldsymbol{U}_i'$. Further, the likelihood function under the conditions of $H_0 : \sigma_1^2 = 0$ is

$$L_0(\boldsymbol{\alpha}, 0, \sigma_2^2, \ldots, \sigma_p^2) = \frac{\exp\left\{-\frac{1}{2}(\boldsymbol{Y} - \boldsymbol{X\alpha})'\boldsymbol{V}_{\mathbf{0}}^{-1}(\boldsymbol{Y} - \boldsymbol{X\alpha})\right\}}{(2\pi)^{\frac{1}{2}N}|\boldsymbol{V}_0|^{\frac{1}{2}}},$$

where $\boldsymbol{V}_0 = \prod_{i=2}^{p} \sigma_i^2 \boldsymbol{U}_i \boldsymbol{U}_i'$. Next, we obtain the ML estimators for the parameters of both likelihood functions and evaluate the likelihood functions at those estimators; and the likelihood-ratio statistic is

$$\lambda = \frac{L_0(\boldsymbol{\alpha}, 0, \hat{\sigma}_2^{2\prime}, \hat{\sigma}_3^{2\prime}, \ldots, \hat{\sigma}_p^{2\prime})}{L(\boldsymbol{\alpha}, \hat{\sigma}_1^{2\prime\prime}, \hat{\sigma}_2^{2\prime\prime}, \hat{\sigma}_3^{2\prime\prime}, \ldots, \hat{\sigma}_p^{2\prime\prime})},$$

where $\hat{\sigma_i}^{2\prime}$ and $\hat{\sigma}_i^{2\prime\prime}$ denote the ML estimates of $\hat{\sigma}_i^2$ under the conditions of H_0 and H_1, respectively. The exact distribution of the likelihood ratio statistic is generally intractable (Self and Liang, 1987). Under a number of regularity conditions, it can be proven that the statistic $-2\ell n\lambda$ is asymptotically distributed as a chi-square variable with one degree of freedom. Stram and Lee (1994) investigated the asymptotic behavior of the likelihood-ratio statistic for variance components in the linear mixed effects model and noted that it does not satisfy the usual regularity conditions of the likelihood-ratio test. They apply a result due to Self and Liang (1987) to determine the correct asymptotic distribution of $-2\ell n\lambda$. The use of higher-order asymptotics to the likelihood to construct confidence intervals and perform tests of single parameters are also discussed by Pierce and Peters (1992).

The determination of likelihood-ratio test is computationally complex and generally requires the use of a computer program. One can use SAS® PROC MIXED and BMDP 3V to apply the likelihood-ratio test. When the design is not too unbalanced and the sample size is small, the tests of hypotheses based on the Satterthwaite procedure are generally adequate. However, when the design is moderately unbalanced or the Satterthwaite procedure is expected to be very liberal, the likelihood ratio tests should be preferred. For extremely unbalanced designs, none of the two procedures seem to be appropriate. Recent research suggests that exact tests are possible (see Remark (ii) below), but there are no most powerful invariant tests when the model is unbalanced (Westfall, 1989). For a complete and authoritative treatment of methods of hypothesis testing for unbalanced data, the reader is referred to Khuri et al. (1998).

Remarks:

(i) In Section 10.8 we considered the REML estimators which arose by factoring the original likelihood function, and noted that these estimators have more appeal than the ML estimators. One can therefore develop a modified likelihood-ratio test in which the REML rather than the ML estimators are used. While there is no general result to support optimality of these tests, it appears that their general properties would be analogous to those of the likelihood-ratio test. Some recent research seems to support the said argument. The results of an extensive Monte Carlo study show that the REML has a reasonable agreement with the ML test (Morell, 1998). It is found that for the configuration of parameter values used in the study, the rejection rates in most cases are less than the nominal 5% for both test statistics; though, on the average, the rejection rates for the REML are closer to the nominal level than for the ML.

(ii) Öfversten (1993) presented a method for deriving exact tests for testing hypotheses concerning variance components of some unbalanced mixed linear models that are special cases of the model in (10.7.1). In particular, he developed methods for obtaining exact F-tests of variance components in three unbalanced mixed linear models, models with one random factor, with nested classifications and models with interaction between two random factors. The method is a generalization of a technique employed by Khuri (1987), Khuri and Littell (1987), and Khuri (1990) for testing variance components in random models. The procedure is based on an orthogonal transformation that reduces the model matrix to contain zero elements as the so-called row-echelon normal forms. The resulting tests are based on mutually independent sums of squares which under the null hypothesis are distributed as scalar multiples of chi-square variates. Although the actual value of the test statistic depends on the particular partitioning of the sums of squares, the distribution of the test statistic is invariant to this choice (see also Christiansen, 1996). Fayyad et al. (1996) have derived an inequality for setting a bound on the power of the procedure. For balanced data, these tests reduce to the traditional F-tests. ♦

10.17 METHODS FOR CONSTRUCTING CONFIDENCE INTERVALS

As mentioned in Section 10.16, the sums of squares in the analysis of variance table from an unbalanced model are generally not independent, neither do they have a chi-square type distribution. Thus the methods for constructing confidence intervals discussed in Volume 1 cannot be applied to unbalanced models without violating the assumptions of independence and chi-squaredness. An exception to this rule is the error sum of squares which has scalar multiple of a chi-square distribution. Thus an exact $100(1-\alpha)\%$ confidence interval for the error variance σ_e^2 is determined as

$$P\left\{\frac{\mathrm{SSE}}{\chi^2[\nu_e, 1-\alpha/2]} \le \sigma_e^2 \le \frac{\mathrm{SSE}}{\chi^2[\nu_e, \alpha/2]}\right\} \doteq 1-\alpha,$$

where SSE is the error sum of squares and ν_e is the corresponding degrees of freedom. For other variance components only approximate methods either based on the Satterthwaite procedure or large sample normal theory can be employed. In particular, for large sample sizes, the ML estimates and their asymptotic properties can be used to construct confidence intervals for the variance components. Thus, if $\hat{\sigma}_i^2$ is the ML estimate of σ_i^2 with asymptotic variance $\mathrm{Var}(\hat{\sigma}_i^2)$, then an approximate $100(1-\alpha)\%$ confidence interval for σ_i^2 is given by

$$P\left\{\hat{\sigma}_i^2 - Z_{1-\alpha/2}\sqrt{\mathrm{Var}(\hat{\sigma}_i^2)} \le \sigma_i^2 \le \hat{\sigma}_i^2 + Z_{1-\alpha/2}\sqrt{\mathrm{Var}(\hat{\sigma}_i^2)}\right\} \cong 1-\alpha. \tag{10.17.1}$$

Note that confidence intervals based on a likelihood method may contain negative values. For some further discussion and details of likelihood-based confidence intervals of variance components, see Jones (1989). The MINQUE procedure can also be used to provide an estimate of the asymptotic variance of the variance component and the normal theory confidence interval is constructed in the usual way. El-Bassiouni (1994) proposed four approximate methods to construct confidence intervals for the estimation of variance components in a general unbalanced mixed model with two variance components, one corresponding to residual effects and the other corresponding to a set of random main or interaction effects. More recently, Burch and Iyer (1997) have proposed a family of procedures to construct confidence intervals for a ratio of variance components and the heritability coefficient in a mixed linear model having two sources of variation. The best interval from the family of procedures can be obtained based on the criteria of bias and expected length. The results can be extended to mixed linear models having more than two variance components.

In cases where sample sizes are small, the large sample normal theory intervals presented above cannot always be recommended. In succeeding chapters, we discuss a number of ad hoc methods for deriving confidence intervals for a variety of statistical designs involving unbalanced random models. In contrast to the large sample intervals, these methods provide "good" confidence intervals for any sample size. A good confidence interval is one that has a coefficient equal to or close to specified confidence coefficient $1-\alpha$. Moreover, the confidence intervals presented above are one-at-a-time intervals. Khuri (1981) developed simultaneous confidence intervals for functions of variance components, and Fennech and Harville (1991) considered exact confidence sets for the variance components and the ratios of the variance components to the error variance in unbalanced mixed linear models.

EXERCISES

1. Consider the model (10.7.1) with $\boldsymbol{Y} \sim N(\boldsymbol{X\alpha}, \boldsymbol{V})$ and the error contrast $\boldsymbol{L}'\boldsymbol{Y}$, where $\boldsymbol{L}'$ is chosen such that $\boldsymbol{L}'\boldsymbol{X} = \boldsymbol{0}$ and $\boldsymbol{L}'$ has row rank equal to $N - \text{rank}(\boldsymbol{X})$.

   (a) Show that $\boldsymbol{L}'\boldsymbol{Y} \sim N(\boldsymbol{0}, \boldsymbol{L}'\boldsymbol{V}\boldsymbol{L})$ and the log-likelihood of $\boldsymbol{L}'\boldsymbol{Y}$ is

   $$\ell n L_1 = \text{constant} - \frac{1}{2}\ell n|\boldsymbol{L}'\boldsymbol{V}\boldsymbol{L}| - \frac{1}{2}\boldsymbol{Y}'\boldsymbol{L}(\boldsymbol{L}'\boldsymbol{V}\boldsymbol{L})^{-1}\boldsymbol{L}'\boldsymbol{Y}.$$

   (b) Show that the log-likelihood can also be written as (Kenward and Roger, 1997)

   $$\ell n \boldsymbol{L}_2 = \text{constant} - \frac{1}{2}\ell n|\boldsymbol{V}| - \frac{1}{2}\ell n|\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}| - \frac{1}{2}\boldsymbol{Y}'\boldsymbol{K}\boldsymbol{Y},$$

   where

   $$\boldsymbol{K} = \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}.$$

   (c) Use the results in parts (a) and (b) to derive the log-likelihood equations and indicate how they can be used to determine REML estimators of the variance components.

2. Consider the linear model $y_i = \mu + e_i$, where $e_i \sim N(0, \sigma^2)$, $i = 1, 2, \ldots, n$, and e_is are uncorrelated. Let $\boldsymbol{Y}' = (y_1, y_2, \ldots, y_n)$ and

$L' = [I_{n-1}\ 0 1_{n-1}] - \frac{1}{n} J_{n-1,n}$, where I_{n-1} is the identity matrix of order $n-1$, $\mathbf{1}_{n-1}$ is the $(n-1)$ component column vector of unity, and $J_{n-1,n}$ is the unity matrix of order $(n-1) \times n$. Prove the following results:

(a) $$L'\mathbf{1} = \mathbf{0}, \qquad L'L = I_{n-1} - \frac{1}{n} J_{n-1,n-1},$$

(b) $$f(L'Y) = \frac{\exp[-(Y'L(L'L)^{-1}L'Y)/(2\sigma^2)]}{(2\pi\sigma^2)^{(n-1)/2}|L'L|^{1/2}},$$

(c) $$\hat{\sigma}^2_{\text{REML}} = \frac{1}{n-1}[Y'L(L'L)^{-1}L'Y] = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2,$$

where

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

3. Consider the unbalanced one-way random model with unequal error variances, $y_{ij} = \mu + \alpha_i + e_{ij}$, $i = 1, 2, \ldots, a$; $j = 1, 2, \ldots, n_i$; $E(\alpha_i) = 0$, $E(e_{ij}) = 0$; $\text{Var}(\alpha_i) = \sigma^2_\alpha$, $\text{Var}(e_{ij}) = \sigma^2_i$; and α_is and e_{ij}s are assumed to be mutually and completely uncorrelated. Find the MINQUE and MIVQUE estimators for σ^2_α and σ^2_i. For $\sigma^2_i \equiv \sigma^2_e$ show that the estimators of σ^2_α and σ^2_e coincide with the estimators considered in Section 11.4.8.

4. For the model described in Exercise 3, show that an unbiased estimator of σ^2_α is given by

$$\frac{\sum_{i=1}^{a} w_i(\bar{y}_{i.} - \bar{y}_w)^2 - \sum_{i=1}^{a} w_i(w - w_i)S_i^2/n_i}{w - \sum_{i=1}^{a} w_i^2/w},$$

where $\bar{y}_{i.} = \sum_{j=1}^{n_i} y_{ij}/n_i$, $\bar{y}_w = \sum_{i=1}^{a} w_i\bar{y}_{i.}/w$, $S_i^2 = \sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2/(n_i - 1)$, $w = \sum_{i=1}^{a} w_i$, and w_is designate a set of arbitrary weights. For the corresponding balanced model with equal error variances, i.e., $n_i \equiv n$ and $\sigma^2_i \equiv \sigma^2_e$, show that the above estimator reduces to the ANOVA estimator of σ^2_α.

5. Spell out details of the derivation of the MINQUE and MIVQUE estimators of σ^2_α and σ^2_e considered in Section 11.4.8.

6. For the model described in Exercise 3, show that the MINQEs of σ^2_α and σ^2_i are given by (Rao and Chaubey, 1978)

$$\hat{\sigma}^2_{\alpha,\text{MINQE}} = (\gamma^4_\alpha/a)\sum_{i=1}^{a} w_i^2(\bar{y}_{i.} - \bar{y}_w)^2$$

and

$$\hat{\sigma}^2_{i,\text{MINQE}} = (n_i - 1)S_i^2/n_i + w_i^2\gamma_i^4(\bar{y}_{i.} - \bar{y}_w)^2/n_i^2,$$

where

$$\bar{y}_{i.} = \sum_{j=1}^{n_i} y_{ij}/n_i, \quad \bar{y}_w = \sum_{i=1}^{a} w_i\bar{y}_{i.}/\sum_{i=1}^{a} w_i,$$

$$S_i^2 = \sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2/(n_i - 1), \quad \text{and} \quad w_i = n_i/(n_i\gamma_\alpha^2 + \gamma_i^2),$$

and γ_α^2 and γ_i^2 denote a priori values of σ_α^2 and σ_i^2. If $\sigma_i^2 \equiv \sigma_e^2$, show that $\hat{\sigma}^2_{\alpha,\text{MINQE}}$ is obtained by replacing γ_i^2 with a common a priori value γ^2 and

$$\hat{\sigma}^2_{e,\text{MINQE}} = \frac{\sum_{i=1}^{a}(n_i - 1)S_i^2}{N} + \frac{\gamma^2}{N}\sum_{i=1}^{a}\frac{w_i^2}{n_i}(\bar{y}_{i.} - \bar{y}_w)^2$$

where $N = \sum_{i=1}^{a} n_i$.

7. Consider the model $y_{ij} = \mu + e_{ij}, i = 1, 2, \ldots, a;\ j = 1, 2, \ldots, n_i$; $E(e_{ij}) = 0$, $\text{Var}(e_{ij}) = \sigma_i^2$; and e_{ij}s are uncorrelated. Show that the MINQE of σ_i^2 is given by (Rao and Chaubey, 1978)

$$\hat{\sigma}^2_{i,\text{MINQE}} = \frac{1}{n_i}[(n_i - 1)S_i^2 + n_i(\bar{y}_{i.} - \bar{y}_w)^2],$$

where

$$\bar{y}_{i.} = \sum_{j=1}^{n_i} y_{ij}/n_i, \bar{y}_w = \sum_{i=1}^{a} w_i\bar{y}_{i.}/\sum_{i=1}^{a} w_i,$$

$$S_i^2 = \sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2/(n_i - 1), w_i = n_i/\gamma_i^2,$$

and γ_i^2 denote a priori values of σ_i^2s. If $\sigma_i^2 \equiv \sigma_e^2$, show that the MINQE of σ_e^2 is given by

$$\hat{\sigma}^2_{e,\text{MINQE}} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2}{N},$$

where

$$\bar{y}_{..} = \sum_{i=1}^{a} n_i\bar{y}_{i.}/N \quad \text{and} \quad N = \sum_{i=1}^{a} n_i.$$

8. In Exercise 6, when $n_i \equiv n$ and $\sigma_i^2 \equiv \sigma_e^2$, show that (Conerly and Webster, 1987)

$$\hat{\sigma}^2_{\alpha,\text{MINQE}} = (w^2/a)\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2,$$

where

$$w = n/(n + \gamma_e^2/\gamma_\alpha^2).$$

Furthermore,

$$E(\hat{\sigma}^2_{\alpha,\text{MINQE}}) = \frac{w^2(a-1)}{a}\left(\sigma_\alpha^2 + \frac{\sigma_e^2}{n}\right)$$

and

$$\text{Var}(\hat{\sigma}^2_{\alpha,\text{MINQE}}) = \frac{2w^2(a-1)}{a^2}\left(\sigma_\alpha^2 + \frac{\sigma_e^2}{n}\right)^2.$$

9. For the balanced one-way random model in (2.1.1) show that the MIMSQE for σ_α^2 considered in Section 10.11 is given by (Rao, 1997)

$$\sigma^2_{\alpha,\text{MIMSQE}} = \frac{1}{N+1}\left[\frac{N-a}{\gamma_e^2/\gamma_\alpha^2} + w\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2\right],$$

where $N = an$, $w = n(n+\gamma_e^2/\gamma_\alpha^2)$, and γ_α^2 and γ_e^2 denote a priori values of σ_α^2 and σ_e^2. Furthermore,

$$E(\hat{\sigma}^2_{\alpha,\text{MIMSQE}}) = \frac{1}{N+1}\left[\frac{N-a}{\gamma_e^2/\gamma_\alpha^2}\sigma_e^2 + w(a-1)\left(\sigma_\alpha^2 + \frac{\sigma_e^2}{n}\right)\right]$$

and

$$\text{Var}(\hat{\sigma}^2_{\alpha,\text{MIMSQE}}) = \frac{2}{(N+1)^2}\left[\frac{N-a}{\gamma_e^2/\gamma^2}\sigma_e^4 + w^2(a-1)\left(\sigma_\alpha^2 + \frac{\sigma_e^2}{n}\right)^2\right].$$

10. Consider the model in (10.10.2) where σ_p^2 represents the error variance. Note that $\boldsymbol{U}_p = \boldsymbol{I}$ and define $\boldsymbol{U}^* = [\boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_{p-1}]$. Show that the MINQUE of σ_p^2, $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$, is the usual error mean square and can be obtained by minimizing $\text{tr}(\boldsymbol{A}^2)$ subject to the conditions that $\text{tr}(\boldsymbol{A}) = 1$ and $\boldsymbol{A}(\boldsymbol{X} : \boldsymbol{U}^*) = \boldsymbol{0}$ (Rao, 1997).

11. Consider an application of Lemma 10.12.1 to show that there exists a nonnegative unbiased estimator for $\boldsymbol{\ell}'\boldsymbol{\sigma}^2$ if $\sum_{i=1}^{p} \ell_i \geq 0$. In particular, show that for the balanced one-way random model in (2.1.1), $\ell_1\sigma_\alpha^2 + \ell_2\sigma_e^2$ can have a nonnegative unbiased estimator if $\ell_1 \geq 0$ and $\ell_2 \geq \ell_1/n$ (Verdooren, 1988).

12. For the model described in Exercise 3, show that under the assumption of normality for the random effects, the log-likelihood function of $(\mu, \sigma_\alpha^2, \sigma_i^2)$ is give by

$$\ell n(L) = C - \frac{1}{2}\left[\sum_{i=1}^{a} \ell n(n_i\sigma_\alpha^2 + \sigma_i^2) + \sum_{i=1}^{a}(n_i - 1)\ell n(\sigma_i^2) + \sum_{i=1}^{a}[(\bar{y}_{i.} - \mu)^2/(\sigma_\alpha^2 + \sigma_i^2/n_i)] + \sum_{i=1}^{a}(n_i - 1)S_i^2/\sigma_i^2\right],$$

where $\bar{y}_{i.} = \sum_{j=1}^{n_i} y_{ij}/n_i$ and $S_i^2 = \sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2/(n_i - 1)$, and C is a constant.

13. In Exercise 12 above, (a) find the likelihood equations for estimating μ, σ_i^2, and σ_α^2; (b) find the likelihood equations for estimating μ and σ_i^2 when $\sigma_\alpha^2 = 0$ and $n_i \equiv n$; (c) find the likelihood equations for estimating μ and σ_α^2 when σ_i^2 are replaced by S_i^2.

14. Use equations (10.7.7a), (10.7.7b), and (10.7.8) to derive the ML solutions of variance components for (a) one-way classification model (2.1.1), (b) two-way nested classification model (6.1.1), and (d) three-way nested classification model (7.1.1).

15. Use equations (10.8.15) and (10.8.16) to derive the REML solutions of variance components for (a) one-way classification model (2.1.1), (b) two-way classification model (3.1.1), (c) two-way classification model (4.1.1), (d) two-way nested classification model (6.1.1), and (e) three-way nested classification model (7.1.1).

16. Use equations (10.7.4), (10.7.5), and (10.7.6) to derive the likelihood equations given by (11.4.15), (11.4.16), and (11.4.17).

17. Use equations (10.8.14) and (10.8.15) to derive the restricted likelihood equations given by (11.4.18) and (11.4.19).

18. Use equation (10.7.7a) to show that in any balanced random model the ML estimator of μ is the grand (overall) mean.

19. Consider the unbalanced one-way random model with a covariate, $y_{ij} = \mu + \alpha_i + \beta X_{ij} + e_{ij}$, $i = 1, 2, \ldots, a$; $j = 1, 2, \ldots, n_{ij}$; $E(\alpha_i) = 0$, $E(e_{ij}) = 0$; $\text{Var}(\alpha_i) = \sigma_\alpha^2$, $\text{Var}(e_{ij}) = \sigma_i^2$; and α_is and e_{ij}s are assumed to be mutually and completely uncorrelated. Derive the MINQUE and MIVQUE estimators of σ_α^2 and σ_i^2 (P. S. R. S. Rao and Miyawaki, 1989).

20. Consider the regression model, $y_{ij} = \alpha + \beta X_i + e_{ij}$, $i = 1, 2, \ldots, a$; $j = 1, 2, \ldots, n_i$; $e_{ij} \sim N(0, \sigma_i^2)$. Find the ML, REML, MINQUE, and MIVQUE estimators of σ_i^2 (Chaubey and Rao, 1976).

21. Consider the regression model with random intercept, $y_{ij} = \mu + \alpha_i + \beta X_i + e_{ij}$, $i = 1, 2, \ldots, a$; $j = 1, 2, \ldots, n_i$; $E(\alpha_i) = 0$, $\text{Var}(\alpha_i) = \sigma_\alpha^2$, $E(e_{ij}) = 0$, $\text{Var}(e_{ij}) = \sigma_i^2$; and α_is and e_{ij}s are mutually and completely uncorrelated. Derive the expressions for the MIVQUE estimators of σ_α^2 and σ_i^2 (P. S. R. S. Rao and Kuranchie, 1988). Show that for the balanced model with equal error variances, i.e., $n_i \equiv n$, and $\hat{\sigma}_i^2 \equiv \hat{\sigma}_e^2$, the MIVQUE estimators of $\sigma_i^2 \equiv \sigma_e^2$ and σ_α^2 reduce to

$$\hat{\sigma}^2_{e,\text{MIVQUE}} = \frac{1}{a}\sum_{i=1}^{a} S_i^2$$

and

$$\hat{\sigma}^2_{\alpha,\text{MIVQUE}} = \frac{\sum_{i=1}^{a}[(\bar{y}_{i.} - \bar{y}_{..})^2 - \hat{\beta}(x_i - \bar{x})]^2}{a-2} - \frac{\hat{\sigma}^2_{e,\text{MIVQUE}}}{n},$$

where

$$\bar{y}_{i.} = \sum_{j=1}^{n} y_{ij}/n, \qquad \bar{y}_{..} = \sum_{i=1}^{a} \bar{y}_{i.}/a, \qquad \bar{x} = \sum_{i=1}^{a} x_i/a,$$

$$S_i^2 = \sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2/(n-1),$$

and

$$\hat{\beta} = \sum_{i=1}^{a}(x_i - \bar{x})(\bar{y}_{i.} - \bar{y}_{..}) \Big/ \sum_{i=1}^{a}(x_i - \bar{x})^2.$$

22. For the log-likelihood function (10.7.3), verify the results on first-order partial derivatives given by (10.7.4) through (10.7.6).

23. For the log-likelihood function (10.7.3) verify the results on second-order partial derivatives given by (10.7.10) through (10.7.12).

24. Consider the linear model in (10.10.2) and let $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ be the MINQUE of σ_i^2, where $\boldsymbol{A}$ is a real symmetric matrix not necessarily nonnegative definite. Define a nonnegative estimator as $\boldsymbol{Y}'\boldsymbol{B}'\boldsymbol{B}\boldsymbol{Y}$ and assume it to be "close" to $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ if the Euclidean norm of the difference $|\boldsymbol{A} - \boldsymbol{B}'\boldsymbol{B}|$ is minimum. Show that the solution for $\boldsymbol{B}'\boldsymbol{B}$ is given by

$$\boldsymbol{B}'\boldsymbol{B} = \sum_{i=1}^{r} e_i \boldsymbol{u}_i \boldsymbol{u}_i',$$

where $e_1 \geq e_2 \geq \cdots \geq e_r$ are the nonnegative eigenvalues of $\boldsymbol{A}$, and $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r$ are the corresponding orthonormal eigenvectors (Chaubey, 1983).

Bibliography

R. D. Anderson (1979), Estimating variance components from balanced data: Optimum properties of REML solutions and MIVQUE estimation, in L. D. Van Vleck and S. R. Searle, eds., *Variance Components and Animal Breeding: Proceedings of a Conference in Honor of C. R. Henderson*, Animal Science Department, Cornell University, Ithaca, NY, 205–206.

I. Ashida and H. Iwaisaki (1995), Estimation of sampling co(variance) of the REML estimators for variance components in a mixed linear model, *Japan J. Biometrics*, **16**, 9–17.

J. K. Baksalary and A. Molinska (1984), Nonnegative unbiased estimability of linear combinations of two variance components, *J. Statist. Plann. Inference*, **10**, 1–8.

R. M. Baskin (1993), Estimation of variance components for the U. S. Consumer Price Index via Gibbs sampling, in *ASA Proceedings of Survey Research Section*, American Statistical Association, Alexandria, VA, 803–813.

K. Berk (1987), Computing for incomplete repeated measures, *Biometrics*, **43**, 385–398.

G. E. P. Box and G. C. Tiao (1973), *Bayesian Inference in Statistical Analysis*, Addison–Wesley, Reading, MA.

N. E. Breslow and D. G. Clayton (1993), Approximate inference in generalized linear mixed models, *J. Amer. Statist. Assoc.*, **88**, 9–25.

J. C. Brockleban and F. G. Giesbrech (1984), Estimating variance components using alternative MINQUEs in selected unbalanced design, in K. Hinkelman, ed., *Experimental Design, Statistical Methods and Genetic Statistics*, Marcel Dekker, New York, 172–211.

L. D. Broemeling (1985), *Bayesian Analysis of Linear Models*, Marcel Dekker, New York.

H. Brown and R. Prescott (1999), *Applied Mixed Models in Medicine*, Wiley, New York.

H. K. Brown and R. A. Kempton (1994), The application of REML in clinical trials, *Statist. Med.*, **13**, 1601–1617.

K. G. Brown (1976), Asymptotic behavior of MINQUE-type estimators of variance components, *Ann. Statist.*, **4**, 746–754.

K. G. Brown (1977), On estimation of diagonal covariance matrices by MINQUE, *Comm. Statist.* A *Theory Methods*, **6**, 471–484.

K. G. Brown (1978), Estimation of variance components using residuals, *J. Amer. Statist. Assoc.*, **73**, 141–146.

B. D. Burch and H. K. Iyer (1997), Exact confidence intervals for a variance ratio (or heritability) in a mixed linear model, *Biometrics*, **53**, 1318–1333.

R. K. Burdick, H. J. Birch, and F. A. Graybill (1986), Confidence intervals on measures of variability in an unbalanced two-fold nested design with equal subsampling, *J. Statist. Comput. Simul.*, **25**, 259–272.

R. K. Burdick and F. A. Graybill (1984), Confidence intervals on linear combinations of variance components in the unbalanced one-way classification, *Technometrics*, **26**, 131–136.

P. R. Burton, K. J. Tiller, L. C. Gurrin, W. A. C. M. Cookson, A. W. Musk, and L. J. Palmer (1999), Genetic variance components analysis for binary phenotypes using generalized linear mixed models (GLMMs) and Gibbs sampling, *Genetic Epidemiology*, **17**, 118–140.

T. P. Callanan (1985), *Restricted Maximum Likelihood Estimation of Variance Components: Computational Aspects*, Ph.D.. dissertation, Iowa State University, Ames, IA.

T. P. Callanan and D. A. Harville (1989), Some new algorithms for computing maximum likelihood estimates of variance components, in K. Berk and L. Malone, eds., *Computer Science and Statistics: Proceedings of the 21st Annual Symposium on the Interface*, American Statistical Association, Alexandria, VA, 435–444.

T. P. Callanan and D. A. Harville (1991), Some new algorithms for computing restricted maximum likelihood estimates of variance components, *J. Statist. Comput. Simul.*, **38**, 239–259.

J. S. K. Chan and A. Y. C. Kuk (1997), Maximum likelihood estimation for probit-linear mixed models with correlated random effects, *Biometrics*, **53**, 86–97.

Y. B. Chaubey (1977), *The Principle of MINQUE in Linear Models: Modifications, Extensions and Applications*, Ph.D. dissertation, University of Rochester, Rochester, NY.

Y. B. Chaubey (1980a), Application of the method of MINQUE for estimation in regression with intraclass covariance matrix, *Sankhyā*, **42**, 28–32.

Y. B. Chaubey (1980b), Minimum norm quadratic estimators of variance components, *Metrika*, **27**, 225–262.

Y. B. Chaubey (1982), Minimum norm invariant quadratic estimation of a covariance matrix in linear model, *Biometrical J.*, **24**, 457–461.

Y. B. Chaubey (1983), A non-negative estimator of variance component closest to MINQUE, *Sankhyā*, **45**, 201–211.

Y. B. Chaubey (1985), Estimators of covariances in time series model, *Statist. Probab. Lett.*, **3**, 51–53.

Y. B. Chaubey (1991), A note on nonnegative minimum bias MINQE in variance components model, *Statist. Probab. Lett.*, **11**, 395–397.

Y. B. Chaubey and P. S. R. S. Rao (1976), Efficiencies of five estimators for the parameters of two linear models with unequal variances, *Sankhyā Ser.* B, **38**, 364–370.

R. Christiansen (1996), Exact tests for variance components, *Biometrics*, **52**, 309–314.

C. C. Cockerham and B. S. Weir (1977), Quadratic analysis of reciprocal crosses, *Biometrics*, **33**, 187–203.

M. R. Conaway (1989), Analysis of repeated categorical measurements with conditional likelihood methods, *J. Amer. Statist. Assoc.*, **84**, 53–62.

M. D. Conerly and J. T. Webster (1987), MINQE for the one-way classification, *Technometrics*, **29**, 229–236.

R. R. Corbeil and S. R. Searle (1976a), A comparison of variance component estimators, *Biometrics*, **32**, 779–791.

R. R. Corbeil and S. R. Searle (1976b), Restricted maximum likelihood (REML) estimation of variance components in the mixed model, *Technometrics*, **18**, 31–38.

P. Cook, L. D. Broemeling, and M. K. Gharaff (1990), A Bayesian analysis of mixed linear model, *Comm. Statist.* A *Theory Methods*, **19**, 987–1002.

D. R. Cox (1955), Some statistical methods connected with series of events (with discussion), *J. Roy. Statist. Soc. Ser.* B, **17**, 129–164.

N. Cressie and S. N. Lahiri (1993), The asymptotic distribution of REML estimators, *J. Multivariate Anal.*, **45**, 217–233.

K. Das (1979), Asymptotic optimality of restricted maximum likelihood estimates for the mixed model, *Calcutta Statist. Assoc. Bull.*, **23**, 125–142.

O. L. Davies and P. L. Goldsmith, eds. (1972), *Statistical Methods in Research and Production*, 4th ed., Oliver and Boyd, Edinburgh.

E. Demidenko and H. Massam (1999), On the existence of the maximum likelihood estimates in variance components models, *Sankhyā Ser.* A, **61**, 431–443.

A. P. Dempster, D. B. Rubin, and R. K. Tsutakawa (1981), Estimation in covariance components models, *J. Amer. Statist. Assoc.*, **76**, 341–353.

A. P. Dempster, M. R. Selwyn, C. M. Patel, and A. J. Roth (1984), Statistical and computational aspects of mixed model analysis, *Appl. Statist.*, **33**, 203–214.

A. J. Dobson (2002), *An Introduction to Generalized Linear Models*, 2nd ed., Chapman and Hall–CRC Press, Boca Raton, FL.

M. Y. El-Bassiouni (1994), Short confidence intervals for variance components, *Comm. Statist.* A *Theory Methods*, **23**, 1915–1933.

S. D. Elliott (1989), The method of unweighted means in univariate and multivariate analysis of variance, *Educ. Psychology Meas.*, **49**, 399–405.

D. A. Elston (1998), Estimation of denominator degrees of freedom of *F*-distributions for assessing Wald statistics for fixed-effect factors in unbalanced mixed models, *Biometrics*, **54**, 1085–1096.

B. Engel (1990), The analysis of unbalanced linear models with variance components, *Statist. Neerland.*, **44**, 195–219.

R. Fayyad, F. A. Graybill, and R. Burdick (1996), A note on exact tests for variance components in unbalanced random and mixed linear models, *Biometrics*, **52**, 306–308.

W. H. Fellner (1986), Robust estimation of variance components, *Technometrics*, **28**, 51–60.

A. R. Fennech and D. A. Harville (1991), Exact confidence sets for variance components in unbalanced mixed linear models, *Ann. Statist.*, **4**, 1771–1785.

J. Focke and G. Dewess (1972), Uber die achätz methode Minque von C. R. Rao and ihre verallgemeinerung, *Math. Oper. Statist. Ser. Statist.*, **3**, 129–143.

R. N. Forthofer and G. G. Koch (1974), An extension of the symmetric sum approach to estimation of variance components, *Biomet. Z.*, **16**, 3–14.

S. Gao and T. M. F. Smith (1995), On the non-existence of a global nonnegative minimum bias invariant quadratic estimator of variance components, *Statist. Probab. Lett.*, **25**, 117–120.

D. W. Gaylor, H. L. Lucas, and R. L. Anderson (1970), Calculation of expected mean squares by the abbreviated Doolittle and square root methods, *Biometrics*, **20**, 641–656.

A. E. Gelfand and A. F. M. Smith (1991), Gibbs sampling for marginal posterior expectations, *Comm. Statist.* A *Theory Methods*, **20**, 1747–1766.

A. E. Gelfand, S. E. Hills, S. Racine-Poon, and A. F. M. Smith (1990), Illustration of Bayesian inference in normal data models using Gibbs sampling, *J. Amer. Statist. Assoc.*, **85**, 972–985.

C. F. Gerald (1977), *Applied Numerical Analysis*, Addison–Wesley, Reading, MA.

M. K. Gharaff (1979), *A General Solution to Making Inferences about the Parameters of Mixed Linear Models*, Ph.D. dissertation, Oklahoma State University, Stillwater, OK.

M. Ghosh (1996), On the nonexistence of nonnegative unbiased estmators of variance components, *Sankhyā Ser.* B, **58**, 360–362.

D. Gianola and R. L. Fernando (1986a), Random effects models for binary responses, *Biometrics*, **42**, 217–218.

D. Gianola and R. L. Fernando (1986b), Bayesian methods in animal breeding theory, *J. Animal Sci.*, **63**, 217–244.

D. Gianola and J. L. Foulley (1990), Variance estimation from integrated likelihoods (VEIL), *Genetic Selection Evolution*, **22**, 403–417.

R. D. Gibbons and D. Hedeker (1997), Random effects probit and logic regression models for three-level data, *Biometrics*, **53**, 1527–1537.

F. G. Giesbrecht (1983), An efficient procedure for computing MINQUE of variance components and generalized least squares estimates of fixed effects, *Comm. Statist.* A *Theory Methods*, **12**, 2169–2177.

F. G. Giesbrecht and J. C. Burns (1985), Two-stage analysis based on a mixed model: Large sample asymptotic theory and small-sample simulation results, *Biometrics*, **41**, 477–486.

F. G. Giesbrecht and P. M. Burrows (1978), Estimating variance components in hierarchical structures using MINQUE and restricted maximum likelihood, *Comm. Statist.* A *Theory Methods*, **7**, 891–904.

W. R. Gilks, D. G. Clayton, D. J. Spiegelhalter, N. G. Best, A. J. McNiel, L. D. Sharples, and A. J. Kirby (1993), Modeling complexity: Applications of Gibbs sampling in medicine (with discussion), *J. Roy. Statist. Soc. Ser.* B, **55**, 39–102.

A. R. Gilmour, R. D. Anderson, and A. L. Rae (1985), The analysis of binomial data by a generalized linear mixed model, *Biometrika*, **72**, 593–599.

D. Giltinan and M. Davidian (1995), *Nonlinear Models for Repeated Measure-*

ment Data, Chapman and Hall, London.

S. Gnot, J. Kleffé, and R. Zmyslony (1985), Nonnegativity of admissible invariant quadratic estimates in mixed linear models with two variance components, *J. Statist. Plann. Inference*, **12**, 249–258.

A. S. Golberger (1962), Best linear unbiased prediction in the generalized linear regression model, *J. Amer. Statist. Assoc.*, **57**, 369–375.

H. Goldstein (1995), *Multilevel Statistical Models*, 2nd ed., Arnold, London.

M. Gönen (2000), A Bayesian analysis of the intraclass correlations in the mixed linear model, *Comm. Statist.* A *Theory Methods*, **29**, 1451–1464.

J. H. Goodnight and W. J. Hemmerle (1979), A simplified algorithm for the W-transformation in variance component estimation, *Technometrics*, **21**, 265–268.

D. G. Gosslee and H. L. Lucas (1965), Analysis of variance of disproportionate data when interaction is present, *Biometrics*, **21**, 115–133.

H.-U. Graser, S. P. Smith, and B. Tier (1987), A derivative-free approach for estimating variance components in animal models by restricted maximum likelihood, *J. Animal Sci.*, **64**, 1362–1370.

J. W. Green (1988), *Diagnostic Methods for Repeated Measures Experiments with Missing Cell*, Technical Report, Department of Mathematical Sciences, University of Delware, Newark, DE.

B. Griffing (1956), Concept of general and specific combining ability in relation to diallel crossing system, *Australian J. Biol. Sci.*, **9**, 463–493.

M. L. Gumpertz and S. G. Pantula (1992), Nonlinear regression with variance components, *Amer. Statist. Assoc.*, **87**, 201–209.

M. L. Gumpertz and J. O. Rawlings (1992), Nonlinear regression with variance components: Modeling effects of ozone on crop yield, *Crop Sci.*, **32**, 219–224.

H. O. Hartley and J. N. K. Rao (1967), Maximum likelihood estimation for the mixed analysis of variance model, *Biometrika*, **54**, 93–108.

H. O. Hartley and W. K. Vaughn (1972), A computer program for the mixed analysis of variance model based on maximum likelihood, in T. A. Bancroft, ed., *Statistical Papers in Honor of George W. Snedecor*, Iowa State University Press, Ames, IA, 129–144.

H. O. Hartley, J. N. K. Rao, and G. Keifer (1969), Variance estimation with one unit per statum, *J. Amer. Statist. Assoc.*, **64**, 841–851.

H. O. Hartley, J. N. K. Rao, and L. R. LaMotte (1978), A simple synthesis-based method of estimating variance components, *Biometrics*, **34**, 233–243.

J. Hartung (1981), Non-negative minimum biased invariant estimation in variance components models, *Ann. Statist.*, **9**, 272–292.

T. D. Hartwell and D. W. Gaylor (1973), Estimating variance components for two-way disproportionate data with missing cells by the method of unweighted means, *J. Amer. Statist. Assoc.*, **68**, 379–383.

D. A. Harville (1969a), *Variance Component Estimation for the Unbalanced One-Way Random Classification*, Technical Report ARL-69-0180, Aerospace Research Laboratories, Wright-Patterson Air Force Base, Dayton, OH.

D. A. Harville (1969b), Variances of variance component estimators for the unbalanced two-way cross classification with application to balanced incomplete block designs, *Ann. Math. Statist.*, **40**, 408–416.

D. A. Harville (1969c), Quadratic unbiased estimation of variance components for the one-way classification, *Biometrika*, **56**, 313–326.

D. A. Harville (1974), Bayesian inference for variance components using only error contrast, *Biometrika*, **61**, 383–385.

D. A. Harville (1977), Maximum-likelihood approaches to variance component estimation and to related problems, *J. Amer. Statist. Assoc.*, **72**, 320–340.

D. A. Harville (1988), Invariant inference for variance components, in J. N. Srivastava, ed., *Probability and Statistics: Essays in Honor of F. A. Graybill*, North-Holland, Amsterdam, 117–133.

D. A. Harville (1990), BLUP (best linear unbiased estimation) and beyond, in D. Gianola and K. Hammond, eds., *Advances in Statistical Methods for Genetic Improvement of Livestock*, Springer-Verlag, New York, 136–176.

D. A. Harville and T. P. Callanan (1990), Computational aspects of likelihood-based inference for variance components, in D. Gianola and K. Hammond, eds., *Advances in Statistical Methods for Genetic Improvement of Livestock*, Springer-Verlag, New York, 136–176.

D. A. Harville and A. L. Carriquiry (1992), Classical and Bayesian prediction as applied to an unbalanced mixed linear model, *Biometrics*, **48**, 987–1003.

D. A. Harville and R. W. Mee (1984), A mixed model procedure for analyzing ordered categorical data, *Biometrics*, **40**, 393–408.

D. A. Harville and A. G. Zimmerman (1996), The posterior-distribution of the fixed and random effects in a mixed effects linear model, *J. Statist. Comput. Simul.*, **54**, 211–219.

B. I. Hayman (1954), The theory and analysis of diallel crosses, *Genetics*, **39**, 789–809.

B. I. Hayman (1960), Maximum likelihood estimation of genetic components of variation, *Biometrics*, **16**, 369–381.

R. H. Heck and S. L. Thomas (1999), *An Introduction to Multilevel Modeling Techniques*, Lawrence Erlbaum, Mahwah, NJ.

D. Hedecker and R. D. Gibbons (1994), A random effects ordinal regression model for multi-level analysis, *Biometrics*, **50**, 933–944.

W. J. Hemmerle and H. O. Hartley (1973), Computing maximum likelihood for the mixed A.O.V. estimation using the *W* transformation, *Technometrics*, **15**, 819–831.

W. J. Hemmerle and J. A. Lorens (1976), Improved algorithm for the *W*-transformation in variance component estimation, *Technometrics*, **18**, 207–211.

C. R. Henderson (1953), Estimation of variance and covariance components, *Biometrics*, **9**, 226–252.

C. R. Henderson (1973), Sire evaluation and genetic trends, in *Proceedings of the Animal Breeding and Genetics Symposium in Honor of Dr. Jay L.*

Lush, American Society of Animal Science and American Dairy Science Association, Champaign, IL, 10–41.

C. R. Henderson (1975), Best linear unbiased estimation and prediction under a selection model, *Biometrics*, **31**, 423–447.

C. R. Henderson (1984), *Applications of Linear Models in Animal Breeding*, University of Guelph, Guelph, ON, Canada.

C. R. Henderson (1985), MIVQUE and REML estimation of additive and non-additive genetic variances, *J. Animal Sci.*, **61**, 113–121.

C. R. Henderson (1986), Recent developments in variance and covariance estimation, *J. Animal Sci.*, **63**, 208–216.

C. R. Henderson, O. Kempthorne, S. R. Searle, and C. N. von Krosigk (1959), Estimation of environmental and genetic trends from records subject to culling, *Biometrics*, **15**, 192–218.

C. R. Henderson, S. R. Searle, and L. R. Schaffer (1974), The invariance and calculation of method 2 for estimating variance components, *Biometrics*, **30**, 583–588.

R. P. Hernández, R. K. Burdick, and N. J. Birch (1992), Confidence intervals and tests of hypotheses on variance components in an unbalanced two-fold nested design, *Biometrical J.*, **34**, 387–402.

J. P. Hobert and G. Casella (1996), The effect of improper priors on Gibbs sampling in hierarchical linear mixed models, *J. Amer. Statist. Assoc.*, **91**, 1461–1473.

R. R. Hocking (1985), *The Analysis of Linear Models*, Brooks–Cole, Monterey, CA.

R. R. Hocking, J. W. Green, and R. H. Bremer (1989), Variance component estimation with model-based diagnostics, *Technometrics*, **31**, 227–240.

R. R. Hocking and M. H. Kutner (1975), Some analytical and numerical comparisons of estimators for the mixed A. O. V. model, *Biometrics*, **31**, 19–28.

A. Hoefer (1998), Variance component estimation in animal breeding: A review, *J. Animal Breeding Genetics*, **115**, 247–265.

S. D. Horn and R. A. Horn (1975), Comparison of estimators of heteroscedastic variances in linear models, *J. Amer. Statist. Assoc.*, **70**, 872–879.

D. A. Huber, T. L. White, and G. R. Hodge (1994), Variance component estimation techniques compared for two mating designs with forest genetic architecture through computer simulation, *Theoret. Appl. Genetics*, **88**, 236–242.

H. M. Hudson (1983), Random effects in log-linear models, *J. Statist. Comput. Simul.*, **17**, 43–50.

R. L. Jennrich and P. F. Sampson (1976), Newton–Raphson and related algorithms for maximum likelihood variance component estimation, *Technometrics*, **18**, 11–17.

J. Jiang (1996), REML estimation: Asymptotic behavior and related topics, *Ann. Statist.*, **24**, 255–286.

J. Jiang (1997), Wald consistency and the method of sieves in REML estimation, *Ann. Statist.*, **25**, 1781–1803.

A. C. Jones (1989), Likelihood based confidence intervals for variance components, in *ASA Proceedings of Statistical Graphics Section*, American Statistical Association, Alexandria, VA, 36–39.

J. S. Kaplan (1983), A method for calculating MINQUE estimators of variance components, *J. Amer. Statist. Assoc.*, **78**, 476–477.

R. M. Kasim and S. W. Raudenbush (1998), Application of Gibbs sampling to nested variance components models with heterogeneous within-group variance, *J. Educ. Behavioral Statist.*, **23**, 93–116.

B. W. Kennedy (1991), C. R. Henderson: The unfinished legacy, *J. Dairy Sci.*, **74**, 4067–4081.

M. G. Kenward and J. H. Roger (1997), Small sample inference for fixed effects from restricted maximum likelihood, *Biometrics*, **53**, 983–997.

R. Khattree (1998), Some practical estimation procedures for variance components, *Comput. Statist. Data Anal.*, **28**, 1–32.

R. Khattree (1999), Nonnegative estimation of variance components: A modification to Henderson's ANOVA methodology, *Sankhyā Ser.* B, **61**, 261–265.

A. I. Khuri (1981), Simultaneous confidence intervals for functions of variance components in random models, *J. Amer. Statist. Assoc.*, **76**, 878–885.

A. I. Khuri (1987), An exact test for the nesting effect's variance component in an unbalanced random two-fold nested model, *Statist. Probab. Lett.*, **5**, 305–311.

A. I. Khuri (1990), Exact test for random models with unequal cell frequencies in the last stage, *J. Statist. Plann. Inference*, **24**, 177–193.

A. I. Khuri and R. C. Littell (1987), Exact tests for the main effects variance components in an unbalanced random two-way model, *Biometrics*, **43**, 545–560.

A. I. Khuri, T. Mathew, and B. K. Sinha (1998), *Statistical Tests for Mixed Linear Models*, Wiley, New York.

J. Kleffé (1976), A note on MINQUE for normal models, *Math. Oper. Statist. Ser. Statist.*, **7**, 707–714.

J. Kleffé (1977a), Invariant methods for estimating variance components in mixed linear models, *Math. Oper. Statist. Ser. Statist.*, **8**, 233–250.

J. Kleffé (1977b), Optimal estimation of variance components: A survey, *Sankhyā Ser.* B, **39**, 211–244.

J. Kleffé (1980), On recent progress of MINQUE theory: Nonnegative estimation, consistency, asymptotic normality and explicit formulae, *Math. Oper. Statist. Ser. Statist.*, **11**, 563–588.

J. Kleffé and J. N. K. Rao (1986), The existence of asymptotically unbiased nonnegative quadratic estimates of variance components in ANOVA models, *J. Amer. Statist. Assoc.*, **81**, 692–698.

J. Kleffé and B. Seifert (1980), Matrix free computation of C. R. Rao's MINQUE for unbalanced nested classification models, *Comput. Statist. Data Anal.*, **2**, 215–228.

J. Kleffé and B. Seifert (1986), Computation of variance components by the MINQUE method, *J. Multivariate Anal.*, **18**, 107–116.

J. Klotz and J. Putter (1970), *Remarks on Variance Components: Likelihood Summits and Flats*, Technichal Report 247, Department of Statistics, University of Wisconsin, Madison, WI.

J. D. Knoke (1985), The method of unweighted means in unbalanced analysis of variance, *Biometrical J.*, **29**, 383–396

G. G. Koch (1967a), A general approach to the estimation of variance components, *Techometrics*, **9**, 93–118.

G. G. Koch (1967b), A procedure to estimate the population mean in random effects models, *Technometrics*, **9**, 577–586.

G. G. Koch (1968), Some further remarks concerning "A general approach to estimation of variance components," *Technometrics*, **10**, 551–558.

I. Kreft and J. deLeeuw (1998), *Introducing Multilevel Modeling*, Sage, Thousand Oaks, CA.

N. M. Laird (1982), Computation of variance components using the EM algorithm, *J. Statist. Comput. Simul.*, **14**, 295–303.

N. M. Laird, N. Lange, and D. Stram (1987), Maximum likelihood computation with repeated measures: Application of the EM algorithm, *J. Amer. Statist. Assoc.*, **82**, 97–105.

N. M. Laird and J. H. Ware (1982), Random-effects models for longitudinal data, *Biometrics*, **38**, 963–974.

L. R. LaMotte (1973a), Quadratic estimation of variance components, *Biometrics*, **29**, 311–330.

L. R. LaMotte (1973b), On non-negative quadratic unbiased estimation of variance components, *J. Amer. Statist. Assoc.*, **68**, 728–730.

J. R. Landis and G. G. Koch (1977), A one-way component of variance model for categorical data, *Biometrics*, **33**, 671–679.

H. S. Lee (1997), Random effects model for count data, *Comm. Statist. A Theory Methods*, **26**, 1893–1904.

J. T. Lee (1993), Efficiency of MINQE for arbitrary distribution under one-way random effects model, *Korean J. Appl. Statist.*, **6**, 355–37.

J. T. Lee and B. C. Kim (1989), An efficient method for computing MINQUE estimators in the mixed models, *J. Korean Statist. Soc.*, **18**, 4–12.

K. R. Lee and C. H. Kapadia (1991), Restricted maximum likelihood estimation under Eisenhart model III, *Statist. Neerland.*, **45**, 271–282.

K. R. Lee and C. H. Kapadia (1992), Mean squared error efficient estimators of the variance components, *Metrika*, **39**, 21–26.

P. M. Lee (1998), *Bayesian Statistics*, 2nd ed., Arnold, London.

C. Y. Lin and A. J. McAllister (1984), Monte Carlo comparison of four methods for estimation of genetic parameters in the univariate case, *J. Dairy Sci.*, **67**, 2389–2398.

X. Lin (1997), Variance components in generalized linear models with random effects, *Biometrika*, **34**, 309–326.

D. V. Lindley and A. F. M. Smith (1972), Bayes estimates for the linear model (with discussion), *J. Roy. Statist. Soc. Ser.* B, **34**, 1–41.

M. J. Lindstrom and D. M. Bates (1988), Newton–Raphson and EM algorithms for linear mixed effects models for repeated measures data, *J. Amer. Statist. Assoc.*, **83**, 1014–1022.

R. C. Littell, G. A. Milliken, W. W. Stroup, and R. D. Wolfinger (1996), *SAS System for Mixed Models*, SAS Institute, Cary, NC.

L.-M. Liu and J. Senturia (1975), *The Distribution of MINQUE Estimators*, Technical Report 406, Department of Statistics, University of Wisconsin, Madison, WI.

L.-M. Liu and J. Senturia (1977), Computation of MINQUE variance components estimates, *J. Amer. Statist. Assoc.*, **72**, 867–868.

N. T. Longford (1987), A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects, *Biometrika*, **74**, 817–827.

N. T. Longford (1993), *Random Coefficient Models*, Oxford University Press, New York.

J. D. Malley (1986), *Optimal Unbiased Estimation of Variance Components*, Lecture Notes in Statistics 39, Springer-Verlag, New York.

H. Massam and J. Muller (1985), The nonnegative MINQUE estimate, *J. Multivariate Anal.*, **16**, 253–259.

T. Mathew (1984), On nonnegative quadratic unbiased estimability of variance components, *Ann. Statist.*, **12**, 1566–1569.

T. Mathew (1987), On the characterization of nonnegative estimable linear combinations of variance components, *Statistics*, **18**, 227–236.

T. Mathew, B. K. Sinha, and B. C. Sutradhar (1992a), Nonnegative estimation of variance components in unbalanced mixed models with two variance components, *J. Multivariate Anal.*, **42**, 77–101.

T. Mathew, B. K. Sinha, and B. C. Sutradhar (1992b), Nonnegative estimation of variance components in general balanced mixed models, in A. Md. E. Saleh, ed., *Nonparametric Statistics and Related Topics*, Elsevier, New York, 281–295.

B. W. McDonald (1994), Two random effects models for multivariate binary data, *Biometrics*, **50**, 164–172.

P. McCullagh and J. Nelder (1989), *Generalized Linear Models*, 2nd ed., Chapman and Hall, London.

C. E. McCulloch and S. R. Searle (2001), *Generalized, Linear, and Mixed Models*, Wiley, New York.

R. A. McLean and W. L. Saunders (1988), Approximating degrees of freedom for standard errors in mixed linear models, in *ASA Proceedings of Statistical Computing Section*, American Statistical Association, Arlington, VA, 50–59.

J. J. Miller (1977), Asymptotic properties of maximum likelihood estimates in the mixed model of the analysis of variance, *Ann. Statist.*, **5**, 746–762.

J. J. Miller (1979), Maximum likelihood estimation of variance components (a Monte Carlo study), *J. Statist. Comput. Simul.*, **8**, 175–190.

S. K. Mitra (1972), Another look at Rao's MINQUE of variance components, *Bull. Internat. Statist. Inst.*, **44**, 279–283.

C. H. Morrel (1998), Likelihood ratio testing of variance components in the linear mixed-effects model using restricted maximum likelihood, *Biometrics*, **54**, 1560–1568.

C. N. Morris (1995), Hierarchical models for educational data: An Overview, *J. Educ. Statist.*, **20**, 190–200.

Y. Ochi and R. L. Prentice (1984), Likelihood inference in a correlated probit regression model, *Biometrika*, **71**, 531–543.

A. Olsen, J. Seely, and D. Birkes (1976), Invariant quadratic unbiased estimation for two variance components, *Ann. Statist.*, **4**, 878–890.

Y. Omori (1997), Comparing two means in count models having random effects, *Statist. Probab. Lett.*, **31**, 225–235.

J. Öfversten (1993), Exact tests for variance components in unbalanced mixed linear models, *Biometrics*, **49**, 45–57.

H. D. Patterson and R. Thompson (1971), Recovery of inter-block information when block sizes are unequal, *Biometrika*, **58**, 545–554.

H. D. Patterson and R. Thompson (1975), Maximum likelihood estimation of components of variance, in L. C. A. Corsten and T. Postelnicu, eds., *Proceedings of the* 8*th International Biometrics Conference*, Editoria Academie, Constanta, Romania, 197–207.

D. A. Pierce and D. Peters (1992), Practical use of higher order asymptotics for multiparameter exponential families, *J. Roy. Statist. Soc. Ser.* B, **54**, 701–725.

K. N. Ponnuswamy and J. Subramani (1987), A simplified procedure for calculating MINQUE of variance components, in *Modeling and Simulation: Proceedings of the International AMSA Conference*, Vol. B, ASMA, New Delhi, 75–90.

R. M. Pringle (1974), Some results on the estimation of variance components by MINQUE, *J. Amer. Statist. Assoc.*, **69**, 987–989.

F. Pukelsheim (1974), *Schätzen von Mittelwert und Streuungsmatrix in Gauss–Markov Modellen*, Diplomarbeit, Albert-Ludwigs-Universität, Freiburg im Breisgau, Germany.

F. Pukelsheim (1976), Estimating variance components in linear models, *J. Multivariate Anal.*, **6**, 626–629.

F. Pukelsheim (1977a), *Linear Models and Convex Programs: Unbiased Nonnegative Estimation in Variance Components Models*, Technical Report 104, Department of Statistics, Stanford University, Stanford, CA.

F. Pukelsheim (1977b), *Examples for Unbiased Nonnegative Estimation in Variance Components Models*, Technical Report 113, Department of Statistics, Stanford University, Stanford, CA.

F. Pukelsheim (1977c), *Punktschätzungen in linearen statistischen Modellen*, Dissertation zur Erlangung des Doktorgrades der Mathematischen, Albert-Ludwigs-Universität, Freiburg im Breisgau, Germany.

F. Pukelsheim (1981a), On the existence of unbiased nonnegative estimates of variance-covariance components, *Ann. Statist.*, **9**, 293–299.

F. Pukelsheim (1981b), Linear models and convex geometry: Aspects of nonnegative variance estimation, *Math. Oper. Statist. Ser. Statist.*, **12**, 271–286.

R. L. Quaas and D. C. Bolgiano (1979), Sampling variances of the MINQUE and method 3 estimators of sire components of variance: A numerical comparison, in L. D. Van Vleck and S. R. Searle, eds., *Variance Components and Animal Breeding: Proceedings of a Conference in Honor of C. R. Henderson*, Cornell University, Ithaca, NY, 99–106.

M. Rajagopalan (1980), *Bayesian Inference for the Variance Components in Mixed Linear Models*, Ph.D. dissertation, Oklahoma State University, Stillwater, OK.

M. Rajagopalan and L. D. Broemeling (1983), Bayesian inference for the variance components of the general mixed model, *Comm. Statist.* A *Theory Methods*, **12**, 701–723.

C. R. Rao (1970), Estimation of heterocedastic variances in linear models, *J. Amer. Statist. Assoc.*, **65**, 161–172.

C. R. Rao (1971a), Estimation of variance and covariance components: MINQUE theory, *J. Multivariate Anal.*, **1**, 257–275.

C. R. Rao (1971b), Minimum variance quadratic unbiased estimation of variance components, *J. Multivariate Anal.*, **1**, 445–456.

C. R. Rao (1972), Estimation of variance and covariance components in linear models, *J. Amer. Statist. Assoc.*, **67**, 112–115.

C. R. Rao (1973), *Linear Statistical Inference and Its Applications*, 2nd ed., Wiley, New York.

C. R. Rao (1974), Projections, generalized inverse and the BLUE, *J. Roy. Statist. Soc. Ser.* B, **36**, 442–448.

C. R. Rao (1979), MINQUE theory and its relation to ML and RML estimation of variance components, *Sankhyā Ser.* B, **41**, 138–153.

C. R. Rao and J. Kleffé (1980), Estimation of variance components, in P. R. Krishnaiah, ed., *Handbook of Statistics*, Vol. 1, North-Holland, Amsterdam, 1–40.

C. R. Rao and J. Kleffé (1988), *Estimation of Variance Components and Applications*, North-Holland, Amsterdam.

J. N. K. Rao (1973), On the estimation of heterocedastic variances, *Biometrics*, **29**, 11–24.

J. N. K. Rao (1977), Contribution to the discussion of the paper by D. A. Harville, *J. Amer. Statist. Assoc.*, **72**, 338–339.

J. N. K. Rao and K. Subrahmaniam (1971), Combining independent estimators and estimation in linear regression with unequal variances, *Biometrics*, **27**, 971–990.

P. S. R. S. Rao (1977), Theory of the MINQUE: A review, *Sankhyā Ser.* B, **39**, 201–210.

P. S. R. S. Rao (1997), *Variance Components Estimation: Mixed Models, Methodologies, and Applications*, Chapman and Hall, London.

P. S. R. S. Rao (2001), The MINQUE: Investigations, modifications and extensions, *J. Statist. Res.*, **35**-1, 29–34.

P. S. R. S. Rao and Y. B. Chaubey (1978), Three modifications of the principle of MINQUE, *Comm. Statist.* A *Theory Methods*, **7**, 767–778.

P. S. R. S. Rao, J. Kaplan, and W. G. Cochran (1981), Estimators for the one-way random effects model with unequal error variances, *J. Amer. Statist. Assoc.*, **76**, 89–97.

P. S. R. S. Rao and P. Kuranchie (1988), Variance components of linear regression model with a random intercept, *Comm. Statist.* A *Theory Methods*, **17**, 1011–1026.

P. S. R. S. Rao and N. Miyawaki (1989), Estimation procedures for unbalanced one-way variance components model, *Bull. Internat. Statist. Inst.*, **2**, 243–244.

S. Raudenbush and A. S. Bryk (1986), A hierarchical model for studying school effects, *Sociol. Educ.*, **59**, 1–17.

S. Raudenbush and A. S. Bryk (2002), *Hierarchical Linear Models: Applications and Data Analysis Methods*, 2nd ed., Sage, Thousand Oaks, CA.

D. K. Rich and K. G. Brown (1979), Estimation of variance components using residuals: Some empirical evidence, in in L. D. Van Vleck and S. R. Searle, eds., *Variance Components and Animal Breeding: Proceedings of a Conference in Honor of C. R. Henderson*, Cornell University, Ithaca, NY, 139–154.

A. M. Richardson and A. H. Welsh (1994), Asymptotic properties of restricted maximum likelihood (REML) estimates for hierarchical mixed linear models, *Australian J. Statist.*, **36**, 31–43.

A. M. Richardson and A. H. Welsh (1995), Robust restricted maximum likelihood in mixed linear models, *Biometrics*, **51**, 1429–1439.

D. L. Robinson (1984), *The REML Program Manual*, unpublished report; available from the author.

D. L. Robinson (1987), Estimation and use of variance components, *Statistician*, **36**, 3–14.

G. K. Robinson (1991), The BLUP is a good thing: The estimation of random effects, *Statist. Sci.*, **6**, 15–51.

D. M. Rocke (1983), Robust statistical analysis of interlaboratory studies, *Biometrika*, **70**, 421–431.

S. H. Rosenberg and C. A. Rhode (1971), Robustness properties of variance component estimates when the underlying model is incorrectly specified, *Biometrics*, **27**, 1091–1097.

J. S. Rosenthal (1995), Rates of convergence for Gibbs sampling for variance components models, *Ann. Statist.*, **23**, 740–761.

P. Rudolph (1976), Bayesian estimation in the linear models under different assumptions about the covariance structure; variance components model, equicorrelation model, *Math. Oper. Statist. Ser. Statist.*, **7**, 649–665.

F. E. Satterthwaite (1946), An approximate distribution of estimates of variance components, *Biometrics Bull.*, **2**, 110–114.

L. R. Schaeffer (1973), *Factors Affecting the Accuracy of Variance Components Estimation*, Ph.D. dissertation, Cornell University, Ithaca, NY.

R. Schall (1991), Estimation in generalized linear models with random effects, *Biometrika*, **78**, 719–727.

H. Scheffé (1959), *The Analysis of Variance*, Wiley, New York.

M. L. Schervish (1992), Bayesian analysis of linear models, in J. M. Bernardo,

J. V. Berger, A. P. David, and A. F. M. Smith, eds., *Bayesian Statistics* IV, Oxford University Press, New York, 419–434.

S. R. Searle (1968), Another look at Henderson's methods of estimating variance components, *Biometrics*, **24**, 749–778.

S. R. Searle (1970), Large sample variances of maximum likelihood estimators of variance components using unbalanced data, *Biometrics*, **26**, 505–525.

S. R. Searle (1971a), Topics in variance components estimation, *Biometrics*, **27**, 1–76.

S. R. Searle (1971b), *Linear Models*, Wiley, New York.

S. R. Searle, G. Casella, and C. E. McCulloch (1992), *Variance Components*, Wiley, New York.

J. Seely (1970a), Linear spaces and unbiased estimation, *Ann. Math. Statist.*, **41**, 1725–1734.

J. Seely (1970b), Linear spaces and unbiased estimation: Applications to the mixed linear model, *Ann. Math. Statist.*, **41**, 1735–1748.

J. Seely (1971), Quadratic subspaces and complements, *Ann. Math. Statist.*, **42**, 710–721.

J. Seely (1972), Completeness for a family of multivariate normal distributions, *Ann. Math. Statist.*, **43**, 1644–1647.

J. Seely (1977), Minimal sufficient statistics and completeness for multivariate normal families, *Sankhyā Ser.* A, **39**, 170–185.

S. G. Self and K.-Y. Liang (1987), Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions, *J. Amer. Statist. Assoc.*, **82**, 605–610.

D. W. Smith and L. W. Murray (1984), An alternative to Eisenhart's Model II and mixed model in the case of negative variance estimates, *J. Amer. Statist. Assoc.*, **79**, 145–151.

P. J. Solomon and D. R. Cox (1992), Nonlinear component of variance models, *Biometrika*, **79**, 1–11.

R. Stiratelli, N. M. Laird, and J. H. Ware (1984), Random-effects models for serial observations with binary response, *Biometrics*, **40**, 961–971.

D. O. Stram and J. W. Lee (1994), Variance components testing in the longitudinal mixed effects model, *Biometrics*, **50**, 1171–1177.

W. W. Stroup and S. D. Kachman (1994), Generalized linear mixed models: An overview, in J. R. Schewenke, ed., *Proceedings of the* 1994 *Kansas State University Conference on Applied Statistics in Agriculture*, Kansas State University, Manhattan, KS, 82–98.

J. Subramani (1991), On invariant quadratic unbiased estimation of variance components, *Comm. Statist.* A *Theory Methods*, **20**, 1705–1730.

L. Sun, J. S. J. Hsu, I. Guttman, and T. Leonard (1996), Bayesian methods for variance component models, *J. Amer. Statist. Assoc.*, **91**, 743–752.

W. Y. Tan, M. A. Tabatabai, and N. Balakrishnan (1988), Harmonic mean aproach to unbalanced random effects model under heteroscedasticity, *Comm. Statist.* A *Theory Methods*, **17**, 1261–1286.

C. M. Theobold, M. Z. Firat, and R. Thompson (1997), Gibbs Sampling, adaptive rejection sampling and robustness to prior specification for a mixed linear model, *Genetic Selection Evolution*, **29**, 57–72.

J. D. Thomas and R. A. Hultquist (1978), Interval estimation for the unbalanced case of the one-way random effects model, *Ann. Statist.*, **6**, 582–587.

R. Thompson (1975), A note on the *W* transformation, *Technometrics*, **17**, 511–512.

R. Thompson (1977a), The estimation of heritability with unbalanced data I: Observations available on parents and offspring, *Biometrics*, **33**, 485–495.

R. Thompson (1977b), The estimation of heritability with unbalanced data II: Data availability on more than two generations, *Biometrics*, **33**, 497–504.

R. Thompson (1980), Maximum likelihood estimation of variance components, *Math. Oper. Statist. Ser. Statist.*, **11**, 545–561.

R. Thompson and Meyer, K. (1986), Estimation of variance components: What is missing in the EM algorithm?, *J. Statist. Comput. Simul.*, **24**, 215–230.

W. A. Thompson, Jr. (1962), The problem of negative estimates of variance components, *Ann. Math. Statist.*, **33**, 273–289.

W. A. Thompson, Jr. and J. R. Moore (1963), Non-negative estimates of variance components, *Technometrics*, **5**, 441–450.

I. B. Thomsen (1975), Testing hypotheses in unbalanced variance components models for two-way layouts, *Ann. Statist.*, **3**, 257–265.

P. B. Topham (1966), Diallel analysis involving maternal and paternal interaction effects, *Heredity*, **21**, 665–674.

E. C. Townsend and S. R. Searle (1971), Best quadratic unbiased estimation of variance components from unbalanced data in the one-way classification, *Biometrics*, **27**, 643–657.

W. H. Vandaele and S. R. Chowdhury (1971), A revised method of scoring, *Statist. Neerland.*, **25**, 101–112.

K. Venkateswarlu (1996), *Estimations of Genetic Variance Components in Diallel Cross Models with Unbalanced Data*, Ph.D. dissertation, University of Madras, Madras, India.

K. Venkateswarlu and K. N. Ponnuswamy (1998), Estimation of variance components based on biomodel of diallel crosses (balanced), *Comm. Statist.* A *Theory Methods*, **27**, 139–152.

K. Venkateswarlu, K. N. Ponnuswamy, and M. R. Srinivasan (1998), Estimation of variance components based on diallel model, *Math. Biol. Sci.*, **150**, 105–112.

A. P. Verbyla (1993), Modeling variance heterogeneity: Residual maximum likelihood and diagnostics, *J. Roy. Statist. Soc. Ser.* B, **55**, 82–86.

L. R. Verdooren (1980), On estimation of variance components, *Statist. Neerland.*, **34**, 83–106.

L. R. Verdooren (1988), Least squares estimators and non-negative estimators of variance components, *Comm. Statist.* A *Theory Methods*, **17**, 1027–1051.

J. Volaufová and V. Witkovsky (1991), Least squares and minimum MSE estimators of variance components in mixed linear models, *Biometrical J.*, **33**, 923–936.

E. F. Vonesh and V. M. Chinchilli (1997), *Linear and Nonlinear Models for the Analysis of Repeated Measurements*, Marcel Dekker, New York.

C. S. Wang, J. J. Rutledge, and D. Gianola (1993), Marginal inference about variance components in a mixed linear model using Gibbs sampling, *Genetic Selection Evolution*, **25**, 41–62.

T. Wansbeck (1980) A regression interpretation of the computation of MINQUE variance component estimates, *J. Amer. Statist. Assoc.*, **75**, 375–376.

R. E. Weiss, Y. Wang, and J. G. Ibrahim (1997), Predictive model selection for repeated measures random effects models using Bayes factors, *Biometrics*, **53**, 592–602.

S. J. Welham and R. Thompson (1997), Likelihood ratio tests for fixed model terms using residual maximum likelihood, *J. Roy. Statist. Soc. Ser.* B, **59**, 701–714.

P. H. Westfall (1986), Asymptotic normality of the ANOVA estimates of components of variance in the non-normal hierarchical mixed model, *Ann. Statist.*, **14**, 1572–1582.

P. H. Westfall (1987), A comparison of variance component estimates for arbitrary underlying distributions, *J. Amer. Statist. Assoc.*, **82**, 866–874.

P. H. Westfall (1989), Power comparisons for invariant variance ratio tests in mixed ANOVA models, *Ann. Statist.*, **17**, 318–326.

P. H. Westfall and R. H. Bremer (1994), Efficiency properties of cell means variance component estimates, *J. Statist. Plann. Inference*, **38**, 159–178.

G. Y. Wong and W. M. Mason (1985), The hierarchical logistic regression model for multilevel analysis, *J. Amer. Statist. Assoc.*, **80**, 513–524.

F. Yates (1934), The analysis of multiple classifications with unequal numbers in the different classes, *J. Amer. Statist. Assoc.*, **29**, 51–66.

F. Yates (1947), Analysis of data from all possible reciprocal crosses between sets of parental lines, *Heredity*, **1**, 287–302.

C. H. Yuan (1977), *A Least Squares Approach to Quadratic Unbiased Estimation of Variance Components in General Mixed Model*, Ph.D. dissertation, North Carolina State University, Raleigh, NC.

S. L. Zeger and M. R. Karim (1991), Generalized linear models with random effects: A Gibbs sampling approach, *J. Amer. Statist. Assoc.*, **86**, 79–86.

S. L. Zeger, K.-Y. Liang, and D. S. Albert (1988), Models for longitudinal data: A generalized estimating equation approach, *Biometrics*, **44**, 1049–1060.

11 One-Way Classification

In Chapter 2, we considered the so-called balanced one-way random effects model where n_is are all equal. Equal numbers of observations for each treatment group or factor level are desirable because of the simplicity of organizing the experimental data and subsequent analysis. However, as indicated in Chapter 9, for a variety of reasons, more data may be available for some levels than for others. In this chapter, we consider a one-way random effects model involving unequal numbers of observations for different groups. This model is widely used in a number of applications in science and engineering.

11.1 MATHEMATICAL MODEL

The random effects model for the unbalanced one-way classification is given by

$$y_{ij} = \mu + \alpha_i + e_{ij}, \quad i = 1, \ldots, a; \quad j = 1, \ldots, n_i, \tag{11.1.1}$$

where y_{ij} is the jth observation in the ith treatment group, μ is the overall mean, α_i is the effect due to the ith level of the treatment factor and e_{ij} is the customary error term. It is assumed that $-\infty < \mu < \infty$ is a constant, and α_is and e_{ij}s are mutually and completely uncorrelated random variables with zero means and variances σ_α^2 and σ_e^2, respectively. Here, σ_α^2 and σ_e^2 are known as the components of variance and in this context inferences are sought about them or certain of their parametric functions.

Remark: A variation of the model in (11.1.1) arises due to lack of homogeneity of error variances in different groups. This model, first considered by Cochran (1937, 1954) and Yates and Cochran (1938), is often used for combination of results of randomized experiments conducted at different times or different places, and for comparing randomly chosen groups with heterogeneous error variances (see also Rao, 1997). ♦

TABLE 11.1 Analysis of variance for the model in (11.1.1).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected square mean
Between	$a-1$	SS_B	MS_B	$\sigma_e^2 + n_0\sigma_\alpha^2$
Within	$N-a$	SS_W	MS_W	σ_e^2

11.2 ANALYSIS OF VARIANCE

The analysis of variance technique involves the partitioning of the total variation, defined by $\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})$, into two components by the following identity:

$$\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{a} n_i(\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2, \tag{11.2.1}$$

where

$$\bar{y}_{i.} = \sum_{j=1}^{n_i} y_{ij}/n_i \quad \text{and} \quad \bar{y}_{..} = \sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}/N$$

with

$$N = \sum_{i=1}^{a} n_i.$$

The quantity on the left side of the identity in (11.2.1) is known as the total sum of squares and the first and second terms on the right side of the identity are called the between group sum of squares, abbreviated as SS_B, and the within group sum of squares, abbreviated as SS_W, respectively. The corresponding mean squares, denoted by MS_B and MS_W, are obtained by dividing SS_B and SS_W by $a - 1$ and $N - a$, respectively. Now, the conventional analysis of variance for the model in (11.1.1) is summarized in Table 11.1.

The expected mean squares can be obtained as follows:

$$\begin{aligned} E(MS_W) &= \frac{1}{N-a} E\left[\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2\right] \\ &= \frac{1}{N-a} E\left[\sum_{i=1}^{a}\sum_{j=1}^{n_i}(\mu + \alpha_i + e_{ij} - \mu - \alpha_i - \bar{e}_{i.})^2\right] \\ &= \frac{1}{N-a} E\left[\sum_{i=1}^{a}\sum_{j=1}^{n_i}(e_{ij} - \bar{e}_{i.})^2\right]. \end{aligned} \tag{11.2.2}$$

Now using Lemma A.1 with $z_j = e_{ij}$ and $\bar{z} = \bar{e}_{i.}$, we have

$$E\left\{\sum_{j=1}^{n_i}(e_{ij} - \bar{e}_{i.})^2\right\} = (n_i - 1)\sigma_e^2. \tag{11.2.3}$$

Substituting (11.2.3) into (11.2.2), we get

$$\begin{aligned}E(\mathrm{MS}_W) &= \frac{1}{N-a}\left[\sum_{i=1}^{a}(n_i - 1)\sigma_e^2\right]\\ &= \sigma_e^2.\end{aligned} \tag{11.2.4}$$

Similarly,

$$\begin{aligned}
E(\mathrm{MS}_B) &= \frac{1}{a-1}E\left[\sum_{i=1}^{a} n_i(\bar{y}_{i.} - \bar{y}_{..})^2\right]\\
&= \frac{1}{a-1}E\left[\sum_{i=1}^{a} n_i\left(\mu + \alpha_i + \bar{e}_{i.} - \mu - \frac{1}{N}\sum_{r=1}^{a} n_r\alpha_r\right.\right.\\
&\qquad\qquad \left.\left. - \frac{1}{N}\sum_{r=1}^{a} n_r\bar{e}_{r.}\right)^2\right]\\
&= \frac{1}{a-1}E\left[\sum_{i=1}^{a} n_i\left(\alpha_i - \frac{1}{N}\sum_{r=1}^{a} n_r\alpha_r + \bar{e}_{i.} - \frac{1}{N}\sum_{r=1}^{a} n_r\bar{e}_{r.}\right)^2\right]\\
&= \frac{1}{a-1}\left[\sum_{i=1}^{a} n_i E\left\{\alpha_i^2 - \frac{2}{N}\alpha_i\sum_{r=1}^{a} n_r\alpha_r + \frac{1}{N^2}\left(\sum_{r=1}^{a} n_r\alpha_r\right)^2\right\}\right.\\
&\qquad \left. + \sum_{i=1}^{a} n_i E\left\{\bar{e}_{i.}^2 - \frac{2}{N}\bar{e}_{i.}\sum_{r=1}^{a} n_r\bar{e}_{r.} + \frac{1}{N^2}\left(\sum_{r=1}^{a} n_r\bar{e}_{r.}\right)^2\right\}\right]\\
&= \frac{1}{a-1}\left[\sum_{i=1}^{a} n_i\left\{\sigma_\alpha^2 - \frac{2}{N}n_i\sigma_\alpha^2 + \frac{1}{N^2}\sum_{r=1}^{a} n_r^2\sigma_\alpha^2\right\}\right.\\
&\qquad \left. + \sum_{i=1}^{a} n_i\left\{\frac{\sigma_e^2}{n_i} - \frac{2n_i\sigma_e^2}{Nn_i} + \frac{1}{N^2}\sum_{r=1}^{a} n_r^2\frac{\sigma_e^2}{n_r}\right\}\right]\\
&= \frac{1}{a-1}\left[\left\{\sum_{i=1}^{a} n_i - \frac{2}{N}\sum_{i=1}^{a} n_i^2 + \frac{1}{N^2}\left(\sum_{i=1}^{a} n_i\right)\left(\sum_{r=1}^{a} n_r^2\right)\right\}\sigma_\alpha^2\right.\\
&\qquad \left. + \left\{\sum_{i=1}^{a}\frac{n_i}{n_i} - \frac{2}{N}\sum_{i=1}^{a} n_i + \frac{1}{N^2}\left(\sum_{i=1}^{a} n_i\right)\left(\sum_{r=1}^{a} n_r\right)\right\}\sigma_e^2\right]
\end{aligned}$$

$$= \frac{1}{a-1}\left[\left(N - \frac{1}{N}\sum_{i=1}^{a} n_i^2\right)\sigma_\alpha^2 + (a-1)\sigma_e^2\right]$$

$$= \sigma_e^2 + n_0\sigma_\alpha^2, \tag{11.2.5}$$

where $n_0 = (N^2 - \sum_{i=1}^{a} n_i^2)/N(a-1)$.

Results (11.2.4) and (11.2.5) seem to have been first given by Cochran (1939) and are derived explicitly in several places, e.g., Winsor and Clarke (1940), Baines (1943), Hammersley (1949), and Graybill (1961, Section 16.5). Searle et al. (1992, pp. 70–71) present a simple derivation of these results using matrix formulation.

11.3 MINIMAL SUFFICIENT STATISTICS AND DISTRIBUTION THEORY

Let $\upsilon_1, \upsilon_2, \ldots, \upsilon_p$ denote those distinct integer values assumed by more than one of the n_is, i.e., $\upsilon_k = n_i = n_j$ for at least one (i, j) pair having $i \neq j, k = 1, \ldots, p$; and let $\upsilon_{p+1}, \ldots, \upsilon_q$ represent those assumed by just one of the n_is. Furthermore, define

$$S_i = \{j | n_j = \upsilon_i\}$$

and let η_i be the number of elements in S_i. Note that $\eta_i = 1$, for $i = p + 1, \ldots, q$. Also, define

$$\bar{y}_{i.} = \sum_{j=1}^{n_i} y_{ij}/n_i \quad \text{and} \quad \bar{y}_r^* = \frac{1}{\eta_r}\sum_{i \in S_r} \bar{y}_{i.}.$$

Now, under the assumption of normality and independence of the random effects, it follows from Hultquist and Graybill (1965) that the $q + p + 1$ dimensional vector

$$\left\{\bar{y}_1^*, \ldots, \bar{y}_q^*; \sum_{i \in S_1}(\bar{y}_{i.} - \bar{y}_1^*)^2, \ldots, \sum_{i \in S_p}(\bar{y}_{i.} - \bar{y}_p^*)^2, \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2\right\}, \tag{11.3.1}$$

is a minimal sufficient statistic for the parameter vector $(\mu, \sigma_\alpha^2, \sigma_e^2)$. We have seen that if the model is balanced, this statistic is complete, otherwise not. It can be shown, using the arguments given in Graybill (1961, pp. 339–346), that the components of the minimal sufficient statistic vector are stochastically independent. Furthermore, it follows that

$$\bar{y}_r^* \sim N(\mu, (\sigma_e^2 + \upsilon_r\sigma_\alpha^2)/\eta_r\upsilon_r), \tag{11.3.2}$$

$$\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2 \sim \sigma_e^2\chi^2[N-a], \tag{11.3.3}$$

and

$$\upsilon_r \sum_{i \in S_r} (\bar{y}_{i.} - \bar{y}_r^*)^2 \sim (\sigma_e^2 + \upsilon_r \sigma_\alpha^2)\chi^2[\eta_r - 1]. \tag{11.3.4}$$

Three important functions of the minimal sufficient statistics are, the sample mean and the between and within group sums of squares of the analysis of variance Table 11.1 given by[1]

$$\bar{y}_{..} = \sum_{i=1}^{a} n_i \bar{y}_{i.}/N,$$

$$\mathrm{SS}_B = \sum_{r=1}^{q} \upsilon_r \sum_{i \in S_r} (\bar{y}_{i.} - \bar{y}_{..})^2, \tag{11.3.5}$$

and

$$\mathrm{SS}_W = \sum_{i=1}^{a} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2.$$

11.4 CLASSICAL ESTIMATION

In this section, we consider various classical methods of estimation of variance components σ_e^2 and σ_α^2.

11.4.1 ANALYSIS OF VARIANCE ESTIMATORS

The analysis of variance (ANOVA) method of estimating variance components σ_e^2 and σ_α^2 consists of equating observed values of the mean squares MS_B and MS_W to their expected values, and solving the resulting equations for σ_e^2 and σ_α^2. The estimators thus obtained are[2]

$$\hat{\sigma}_{e,\mathrm{ANOV}}^2 = \mathrm{MS}_W$$

and (11.4.1)

$$\hat{\sigma}_{\alpha,\mathrm{ANOV}}^2 = \frac{\mathrm{MS}_B - \mathrm{MS}_W}{n_0}.$$

[1] The sum of squares between groups, SS_B, does not follow a constant times a chi-square distribution unless $\sigma_\alpha^2 = 0$. However, it can be shown that it is the weighted sum of $a-1$ independent chi-square variables each with one degree of freedom and SS_B and SS_W are stochastically independent.

[2] Cochran (1939) seems to have employed this procedure for the model in (11.1.1) while discussing sampling strategies for observations in fields taken from farms within regions.

By definition, the estimators in (11.4.1) are unbiased. However, they do not satisfy the usual optimal properties of the ANOVA estimators in the case of balanced data. It was first stated by Scheffé (1959, p. 224) and later proved by Read (1961) that there does not exist a quadratic unbiased estimator of σ_α^2, whose variance is uniformly smaller than that of every other quadratic unbiased estimator. Kleffé (1975) proves a similar result for the two-way classification model. Note that the estimators in (11.4.1) do of course, reduce to those for the balanced data and can produce a negative estimate for σ_α^2.

Remarks:

(i) An unbiased estimator of σ_α^2 based on an unweighted between group mean square defined as

$$\mathrm{MS}_B^* = \frac{\bar{n}_h \sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..}^*)^2}{a-1},$$

where

$$\bar{y}_{i.} = \sum_{j=1}^{n_i} y_{ij}/n_i, \qquad \bar{y}_{..}^* = \sum_{i=1}^{a} \bar{y}_{i.}/a, \quad \text{and} \quad \bar{n}_h = a/\sum_{i=1}^{a} n_i^{-1},$$

is given by

$$\hat{\sigma}_{\alpha,\mathrm{UNW}}^2 = (\mathrm{MS}_B^* - \mathrm{MS}_W)/\bar{n}_h.$$

(ii) In choosing a variance component estimator for unbalanced data, although one cannot find a single estimator in the class of quadratic unbiased estimators that is "better" than all others; one can exclude from considerations those estimators that are inquadmissible. A quadratic estimator is called inquadmissible if there exists a second quadratic estimator, having the same expectation, whose sampling variance is less than or equal to that of the first for all points in the parameter space with strict inequality for at least one such point. Otherwise, the estimator is said to be quadmissible. Harville (1969a) has considered the problem of determining whether an arbitrary quadratic form in the one-way normally distributed data is inquadmissible.

(iii) The problem of weighting in the estimation of variance components is discussed by Robertson (1962). It is found that the correct weighting is dependent on the F-value of the analysis of variance.

(iv) The estimator $\hat{\sigma}_{\alpha,\mathrm{ANOV}}^2$ in (11.4.1) can yield a negative estimate. Mathew et al. (1992) consider nonnegative estimators from unbalanced models with two variance components, of which the model in (11.1.1) is a special case. Chatterjee and Das (1983) develop best asymptotically normal (BAN) estimators for the variance components. Kelly and Mathew (1993) discuss an invariant quadratic estimator of σ_α^2 that has smaller MSE and smaller probability of yielding a negative estimate than $\hat{\sigma}_{\alpha,\mathrm{ANOV}}^2$.

(v) From (11.4.1) the probability of a negative $\hat{\sigma}^2_{\alpha,\text{ANOV}}$ is

$$P(\hat{\sigma}^2_{\alpha,\text{ANOV}} < 0) = P(\text{MS}_B < \text{MS}_W).$$

Further, from (11.3.4) and (11.3.5), it follows that MS_B can be expressed as a linear combination of independent central chi-square variables. Thus the distribution of MS_B can be approximated by a central chi-square variable using the Satterthwaite approximation; and the probability of a negative estimate can be evaluated in terms of the central F-distribution. Singh (1989a) developed an expression for determining an exact value of $P(\hat{\sigma}^2_{\alpha,\text{ANOV}} < 0)$ using an infinite weighted sum of incomplete beta functions. An exact value of the probability of a negative estimate can also be evaluated from Davies (1980) who gives an algorithm for computing the distribution of a linear combination of independent chi-square variables (possibly noncentral) with arbitrary degrees of freedom. Lee and Khuri (2001) investigated the behavior of $P(\hat{\sigma}^2_{\alpha,\text{ANOV}} < 0)$ by modeling its values for different values of n, intraclass correlation, $\rho = \sigma^2_\alpha/(\sigma^2_e + \sigma^2_\alpha)$, and an imbalance measure, $\phi = N^2/a\sum_{i=1}^a n_i^2$, using the generalized linear model technique. ♦

11.4.2 FITTING-CONSTANTS-METHOD ESTIMATORS

The reduction in sum of squares in fitting the fixed effects version of the model in (11.1.1) is

$$R(\mu, \alpha) = \sum_{i=1}^a n_i \bar{y}^2_{i.}.$$

The submodel

$$y_{ij} = \mu + e_{ij}$$

has the normal equation

$$\hat{\mu} = \bar{y}_{..}$$

and the corresponding reduction in sum of squares is

$$R(\mu) = N\bar{y}^2_{..}.$$

Now, the quadratics to be equated to their respective expected values in the fitting-constants-method of estimating variance components are

$$R(\alpha|\mu) = R(\mu, \alpha) - R(\mu) = \sum_{i=1}^a n_i \bar{y}^2_{i.} - N\bar{y}^2_{..}$$

and

$$\text{SS}_E = \sum_{i=1}^a \sum_{j=1}^{n_i} y^2_{ij} - R(\mu, \alpha) = \sum_{i=1}^a \sum_{j=1}^{n_i} y^2_{ij} - \sum_{i=1}^a n_i \bar{y}^2_{i.}.$$

The quadratics $R(\alpha|\mu)$ and SS_E are the same as the sum of squares terms SS_B and SS_W defined in (11.2.1). Thus, in this case, the method of fitting constants would give estimators of variance components identical to the analysis of variance procedure.

11.4.3 SYMMETRIC SUMS ESTIMATORS

In this section, we consider symmetric sums estimators based on products and squares of differences of observations (Koch, 1967a, 1968). The expected values of products of observations from the model in (11.1.1) are

$$E(y_{ij}y_{i'j'}) = \begin{cases} \mu^2, & i \neq i', \\ \mu^2 + \sigma_\alpha^2, & i = i', j \neq j', \\ \mu^2 + \sigma_\alpha^2 + \sigma_e^2, & i = i', j = j'. \end{cases} \tag{11.4.2}$$

We now estimate μ^2, $\mu^2 + \sigma_\alpha^2$, and $\mu^2 + \sigma_\alpha^2 + \sigma_e^2$ by taking the means of the symmetric sums of products of observations in (11.4.2). Thus we obtain

$$\begin{aligned} \hat{\mu}^2 = g_m &= \sum_{\substack{i,i' \\ i \neq i'}}^{a} \sum_{j,j'} y_{ij}y_{i'j'} \Big/ \sum_{\substack{i,i' \\ i \neq i'}} n_i n_{i'} \\ &= \left(y_{..}^2 - \sum_{i=1}^{a} y_{i.}^2 \right) \Big/ \left(N^2 - \sum_{i=1}^{a} n_i^2 \right), \\ \hat{\mu}^2 + \hat{\sigma}_\alpha^2 = g_A &= \sum_{i=1}^{a} \sum_{j \neq j'}^{n_i} y_{ij}y_{ij'} \Big/ \sum_{i=1}^{a} n_i(n_i - 1) \\ &= \left(\sum_{i=1}^{a} y_{i.}^2 - \sum_{i=1}^{a} \sum_{j=1}^{n_i} y_{ij}^2 \right) \Big/ \left(\sum_{i=1}^{a} n_i^2 - N \right), \end{aligned}$$

and

$$\begin{aligned} \hat{\mu}^2 + \hat{\sigma}_\alpha^2 + \hat{\sigma}_e^2 = g_E &= \sum_{i=1}^{a} \sum_{j=1}^{n_i} y_{ij}^2 \Big/ \sum_{i=1}^{a} n_i \\ &= \sum_{i=1}^{a} \sum_{j=1}^{n_i} y_{ij}^2 / N. \end{aligned}$$

The estimators of σ_α^2 and σ_e^2, therefore, are given by

$$\hat{\sigma}_{\alpha,\text{SSP}}^2 = g_A - g_m$$

and (11.4.3)

$$\hat{\sigma}_{e,\text{SSP}}^2 = g_E - g_A.$$

The estimators in (11.4.3), by construction, are unbiased; and they reduce to the analysis of variance estimators in the case of balanced data. However, they are not translation invariant, i.e., they may change in values if the same constant is added to all the observation and their variances are functions of μ. This drawback is overcome by using the symmetric sums of squares of differences rather than products.

From the model in (11.1.1), the expected values of squares of differences of observations are

$$E(y_{ij} - y_{i'j'})^2 = \begin{cases} 2\sigma_e^2, & i = i', \quad j \neq j', \\ 2(\sigma_e^2 + \sigma_\alpha^2), & i \neq i'. \end{cases} \tag{11.4.4}$$

Now, we estimate $2\sigma_e^2$ and $2(\sigma_e^2 + \sigma_\alpha^2)$ by taking the means of the symmetric sums of squares of differences in (11.4.4). Thus we obtain

$$\begin{aligned} 2\hat{\sigma}_e^2 = h_E &= \sum_{i=1}^{a} \sum_{\substack{j,j' \\ j \neq j'}}^{n_i} (y_{ij} - y_{ij'})^2 \Big/ \sum_{i=1}^{a} n_i(n_i - 1) \\ &= \sum_{i=1}^{a} n_i \left(\sum_{j=1}^{n_i} y_{ij}^2 - n_i \bar{y}_{i.}^2 \right) \Big/ \left(\sum_{i=1}^{a} n_i^2 - N \right) \end{aligned} \tag{11.4.5}$$

and

$$\begin{aligned} 2(\hat{\sigma}_e^2 + \hat{\sigma}_\alpha^2) = h_A &= \sum_{\substack{i,i' \\ i \neq i'}}^{a} \sum_{j=1}^{n_i} \sum_{j'=1}^{n_{i'}} (y_{ij} - y_{i'j'})^2 \Big/ \sum_{i=1}^{a} n_i(N - n_i) \\ &= \frac{2}{N^2 - \sum_{i=1}^{a} n_i^2} \sum_{i=1}^{a} (N - n_i) \sum_{j=1}^{n_i} y_{ij}^2 - 2g_m, \end{aligned} \tag{11.4.6}$$

where

$$g_m = \frac{1}{N^2 - \sum_{i=1}^{a} n_i^2} \left(y_{..}^2 - \sum_{i=1}^{a} y_{i.}^2 \right).$$

Note that the quantity g_m represents the unbiased estimator of μ^2 given earlier. The estimators of the variance components are obtained by solving equations (11.4.5) and (11.4.6) for $\hat{\sigma}_e^2$ and $\hat{\sigma}_\alpha^2$. The resulting estimators are

$$\hat{\sigma}_{e,\text{SSS}}^2 = h_E/2$$

and

$$\hat{\sigma}_{\alpha,\text{SSS}}^2 = (h_A - h_E)/2. \tag{11.4.7}$$

The estimators in (11.4.7) are unbiased and translation invariant, and their variances contain no terms in μ. Further, they reduce to the analysis of variance estimators for the case of balanced data.

11.4.4 ESTIMATION OF μ

In many investigations the researcher is often interested in estimating the general mean μ. The usual sample mean $\bar{y}_{..}$ is unbiased for μ with variance

$$\mathrm{Var}(\bar{y}_{..}) = \sum_{i=1}^{a} \frac{n_i(\sigma_e^2 + n_i\sigma_\alpha^2)}{N^2}.$$

The unweighted mean $\bar{y}_{\mathrm{UNW}}$ is also unbiased with variance

$$\mathrm{Var}(\bar{y}_{\mathrm{UNW}}) = \sum_{i=1}^{a} \frac{(\sigma_e^2 + n_i\sigma_\alpha^2)}{a^2 n_i}.$$

The weighted least squares estimator for μ is

$$\bar{y}_{\mathrm{WLS}} = \sum_{i=1}^{a} \frac{n_i \bar{y}_{i.}}{(\sigma_e^2 + n_i\sigma_\alpha^2)} \Big/ \sum_{i=1}^{a} \frac{n_i}{(\sigma_e^2 + n_i\sigma_\alpha^2)},$$

which is the minimum variance unbiased estimator with variance

$$\mathrm{Var}(\bar{y}_{\mathrm{WLS}}) = 1 \Big/ \sum_{i=1}^{a} \frac{n_i}{(\sigma_e^2 + n_i\sigma_\alpha^2)}.$$

Note that $\bar{y}_{\mathrm{WLS}}$ is a weighted estimator of $\bar{y}_{i.}$s with weights w_is determined as $w_i = 1/\mathrm{Var}(\bar{y}_{i.})$. Furthermore, w_is are functions of unknown variance components which in practice are unknown and must be estimated. The use of variance component estimates results in an estimator which is no longer unbiased or minimum variance. For a discussion of relative advantages of $\bar{y}_{\mathrm{UNW}}$ and $\bar{y}_{\mathrm{WLS}}$, see Cochran (1937, 1954), Cochran and Carroll (1953), and Rao (1997, Section 10.3). The maximum likelihood estimator of μ does not have an explicit closed form expression and has to be obtained using an iterative procedure (see Section 11.4.5.1).

To obtain an unbiased estimator of μ in the model in (11.1.1) by the method discussed in Section 10.6, we note from Section 11.4.3 that an unbiased estimator of μ^2 is

$$\hat{\mu}^2 = \left(y_{..}^2 - \sum_{i=1}^{a} y_{i.}^2\right) \Big/ \left(N^2 - \sum_{i=1}^{a} n_i^2\right).$$

Now, proceeding as in Section 10.6, we get

$$\begin{aligned}\widehat{(\mu+\theta)}^2 &= \left\{(y_{..} + N\theta)^2 - \sum_{i=1}^{a}(y_{i.} + n_i\theta)^2\right\} \Big/ \left(N^2 - \sum_{i=1}^{a} n_i^2\right) \\ &= \hat{\mu}^2 + \frac{2\left(Ny_{..} - \sum_{i=1}^{a} n_i y_{i.}\right)}{N^2 - \sum_{i=1}^{a} n_i^2}\theta + \theta^2. \qquad (11.4.8)\end{aligned}$$

Hence, comparing (11.4.8) with (10.6.2), the desired estimator of μ is

$$\hat{\mu} = \left(N y_{..} - \sum_{i=1}^{a} n_i y_{i.}\right) \Big/ \left(N^2 - \sum_{i=1}^{a} n_i^2\right)$$
$$= \frac{1}{(N^2 - \sum_{i=1}^{a} n_i^2)} \sum_{i=1}^{a} \sum_{j=1}^{n_i} (N - n_i) y_{ij}. \qquad (11.4.9)$$

The variance of the estimator (11.4.9) is given by

$$\text{Var}(\hat{\mu}) = \frac{1}{(N^2 - \sum_{i=1}^{a} n_i^2)^2} \left\{ \sum_{i=1}^{a} n_i^2 (N - n_i)^2 \sigma_\alpha^2 + \sum_{i=1}^{a} n_i (N - n_i)^2 \sigma_e^2 \right\}.$$

Koch (1967b) has made a numerical comparison of the variances of the estimators $\bar{y}_{..}$, $\bar{y}_{\text{UNW}}$, and $\hat{\mu}$.

11.4.5 MAXIMUM LIKELIHOOD AND RESTRICTED MAXIMUM LIKELIHOOD ESTIMATORS

Under the assumption of normality for the random effects α_is and e_{ij}s, one can proceed to obtain the maximum likelihood (ML) and the restricted maximum likelihood (REML) estimators of σ_e^2 and σ_α^2. However, as we have seen in Sections 10.7 and 10.8, the ML and REML estimators of variance components from unbalanced data cannot be obtained explicitly. In this section, we consider the problem of deriving the ML estimators of the parameters for the model in (11.1.1) and the REML estimators of the variance components. It should be remarked that Crump (1947, 1951) seems to have been the first to consider the ML estimators of the variance components for this problem.

11.4.5.1 The Maximum Likelihood Estimators

The likelihood function for the sample observations y_{ij}s from the model in (11.1.1) is

$$L = f(y_{11}, \ldots, y_{1n_1}; y_{21}, \ldots, y_{2n_2}; \ldots; y_{a1}, \ldots, y_{an_a})$$
$$= f(\boldsymbol{Y}_1) f(\boldsymbol{Y}_2) \ldots f(\boldsymbol{Y}_a), \qquad (11.4.10)$$

where $\boldsymbol{Y}_i' = (y_{i1}, y_{i2}, \ldots, y_{in_i})$ is an n_i-vector having a multivariate normal distribution, with mean vector and variance-covariance matrix given by

$$\boldsymbol{\mu}_i = \mu \mathbf{1}_{n_i}$$

and (11.4.11)

$$\boldsymbol{V}_i = \sigma_e^2 \boldsymbol{I}_{n_i} + \sigma_\alpha^2 \boldsymbol{J}_{n_i},$$

with $\mathbf{1}_{n_i}$ being an n_i-vector having every element unity, $\boldsymbol{I}_{n_i}$ being an identity matrix of order n_i, and $\boldsymbol{J}_{n_i}$ being a square matrix of order n_i having every element unity. Hence,

$$f(\boldsymbol{Y}_i) = \frac{1}{(2\pi)^{n_i/2}|\boldsymbol{V}_i|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)'\boldsymbol{V}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)\right\}$$

and the likelihood function (11.4.10) is given by

$$L = \frac{1}{(2\pi)^{N/2}\Pi_{i=1}^{a}|\boldsymbol{V}_i|^{1/2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{a}(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)'\boldsymbol{V}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{\mu}_i)\right\}. \tag{11.4.12}$$

Now, from Lemmas B.1 and B.2, we obtain

$$|\boldsymbol{V}_i| = (\sigma_e^2)^{n_i-1}(\sigma_e^2 + n_i\sigma_\alpha^2)$$

and

$$\boldsymbol{V}_i^{-1} = \frac{1}{\sigma_e^2}\boldsymbol{I}_{n_i} - \frac{\sigma_\alpha^2}{\sigma_e^2(\sigma_e^2 + n_i\sigma_\alpha^2)}\boldsymbol{J}_{n_i}.$$

On substituting for $\boldsymbol{\mu}_i$, $|\boldsymbol{V}_i|$, and $\boldsymbol{V}_i^{-1}$ in (11.4.12), and after some simplifications, the likelihood function reduces to

$$L = \frac{\exp\left[-\frac{1}{2}\left\{\sum_{i=1}^{a}\sum_{j=1}^{n_i}\frac{(y_{ij}-\bar{y}_{i.})^2}{\sigma_e^2} + \sum_{i=1}^{a}\frac{n_i(\bar{y}_{i.}-\mu)^2}{(\sigma_e^2+n_i\sigma_\alpha^2)}\right\}\right]}{(2\pi)^{\frac{1}{2}N}(\sigma_e^2)^{\frac{1}{2}(N-a)}\Pi_{i=1}^{a}(\sigma_e^2 + n_i\sigma_\alpha^2)^{1/2}}. \tag{11.4.13}$$

The likelihood function in (11.4.13) is given explicitly in Henderson et al. (1957) and Hill (1965). It can also be obtained as a special case of the general results given by Hartley and Rao (1967). The natural logarithm of the function (11.4.13) is

$$\begin{aligned}\ell n(L) = &-\frac{1}{2}N\ell n(2\pi) - \frac{1}{2}(N-a)\ell n(\sigma_e^2) - \frac{1}{2}\sum_{i=1}^{a}\ell n(\sigma_e^2 + n_i\sigma_\alpha^2)\\ &- \frac{1}{2\sigma_e^2}\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2 - \frac{1}{2}\sum_{i=1}^{a}\frac{n_i(\bar{y}_{i.} - \mu)^2}{(\sigma_e^2 + n_i\sigma_\alpha^2)}.\end{aligned} \tag{11.4.14}$$

Equating to zero the partial derivatives of (11.4.14) with respect to μ, σ_e^2, and σ_α^2, and denoting the solutions by $\hat{\mu}$, $\hat{\sigma}_e^2$, and $\hat{\sigma}_\alpha^2$, we obtain, after some simplifications, the following system of equations:

$$\sum_{i=1}^{a}\left(\frac{n_i}{\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2}\right)\hat{\mu} - \sum_{i=1}^{a}\frac{n_i\bar{y}_{i.}}{(\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2)} = 0, \tag{11.4.15}$$

$$\frac{(N-a)}{\hat{\sigma}_e^2} + \sum_{i=1}^{a} \frac{1}{(\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2)} - \frac{1}{\hat{\sigma}_e^4}\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2$$
$$- \sum_{i=1}^{a} \frac{n_i(\bar{y}_{i.} - \hat{\mu})^2}{(\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2)^2} = 0, \qquad (11.4.16)$$

and

$$\sum_{i=1}^{a} \frac{n_i}{(\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2)} - \sum_{i=1}^{a} \frac{n_i^2(\bar{y}_{i.} - \hat{\mu})^2}{(\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2)^2} = 0, \qquad (11.4.17)$$

where circumflex accents (hats) over σ_α^2 and σ_e^2 indicate "estimates of" the corresponding parameters. It is immediately seen that equations (11.4.15), (11.4.16), and (11.4.17) have no explicit solutions for $\hat{\mu}$, $\hat{\sigma}_e^2$, and $\hat{\sigma}_\alpha^2$ and need to be solved using some iterative procedure. They do, of course, reduce to the simpler equations in the case of balanced data, i.e., when $n_1 = \cdots = n_a$. Moreover, even if the solutions could be found using an iterative procedure, the problem of using them to obtain a nonnegative estimate of σ_α^2 in the restricted parameter space must also be considered. Chatterjee and Das (1983) discuss relationship of the ML estimators with that obtained using a weighted least squares approach. For some properties of the ML estimator, see Yu et al. (1994). More recently, Vangel and Rukhin (1999) have considered ML estimation of the parameters for the case involving heteroscedastic error variances.

11.4.5.2 Restricted Maximum Likelihood Estimator

Proceeding from the general case considered in Section 10.8 or otherwise, the restricted log-likelihood function for the sample observations y_{ij}s, from the model in (11.1.1), is obtained as[3]

$$\ell n(L') = -\frac{1}{2}\Bigg[(N-a)\ell n(\sigma_e^2) + \sum_{i=1}^{a} \ell n(\sigma_e^2 + n_i\sigma_\alpha^2) + \ell n\left(\sum_{i=1}^{a} \frac{n_i}{(\sigma_e^2 + n_i\sigma_\alpha^2)}\right)$$
$$+ \frac{(N-a)\text{MS}_W}{\sigma_e^2} + \sum_{i=1}^{a} \frac{n_i(\bar{y}_{i.} - \bar{y}_W)^2}{(\sigma_e^2 + n_i\sigma_\alpha^2)}\Bigg],$$

where

$$\bar{y}_W = \sum_{i=1}^{a} \frac{n_i\bar{y}_{i.}}{(\sigma_e^2 + n_i\sigma_\alpha^2)} \Bigg/ \sum_{i=1}^{a} \frac{n_i}{(\sigma_e^2 + n_i\sigma_\alpha^2)}.$$

[3]It can be readily observed that the likelihood function (11.4.13) does not permit a straightforward factoring to separate out a function of μ similar to the case of balanced data.

It is readily verified that for $n_i = n$, $\ell n(L')$ reduces to

$$-\frac{1}{2}\left[\ell n(an) + a(n-1)\ell n(\sigma_e^2) + (a-1)\ell n(\sigma_e^2 + n\sigma_\alpha^2) + \frac{a(n-1)\mathrm{MS}_W}{\sigma_e^2} + \frac{a(n-1)\mathrm{MS}_B}{(\sigma_e^2 + n\sigma_\alpha^2)}\right],$$

which is equivalent to the restricted log-likelihood function for the balanced model given in (2.4.14).

Equating to zero the partial derivatives of $\ell n(L')$ with respect to σ_e^2 and σ_α^2 and denoting the solutions by $\hat{\sigma}_e^2$ and $\hat{\sigma}_\alpha^2$, we obtain, after some simplifications, the following system of equations:

$$\frac{N-a}{\hat{\sigma}_e^2} + \sum_{i=1}^{a}\frac{1}{(\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2)} - \sum_{i=1}^{a}\frac{n_i}{(\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2)^2}\Big/\sum_{i=1}^{a}\frac{n_i}{(\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2)} = \frac{(N-a)\mathrm{MS}_W}{\hat{\sigma}_e^4} + \sum_{i=1}^{a}\frac{n_i(\bar{y}_{i.} - \bar{y}_W)^2}{(\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2)^2} \tag{11.4.18}$$

and

$$\sum_{i=1}^{a}\frac{n_i}{(\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2)} - \sum_{i=1}^{a}\frac{n_i^2}{(\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2)^2}\Big/\sum_{i=1}^{a}\frac{n_i}{(\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2)} = \sum_{i=1}^{a}\frac{n_i^2(\bar{y}_{i.} - \bar{y}_W)^2}{(\hat{\sigma}_e^2 + n_i\hat{\sigma}_\alpha^2)^2}. \tag{11.4.19}$$

It is immediately seen that equations (11.4.18) and (11.4.19) have no explicit solutions for $\hat{\sigma}_e^2$ and $\hat{\sigma}_\alpha^2$. They do, of course, reduce to the simpler equations in the case of balanced data. Moreover, to obtain the REML estimators, equations (11.4.18) and (11.4.19) need to be solved for $\hat{\sigma}_e^2$ and $\hat{\sigma}_\alpha^2$, subject to the constraints that $\hat{\sigma}_e^2 > 0$ and $\hat{\sigma}_\alpha^2 \geq 0$, using some iterative procedure (see Section 10.8.1).

11.4.6 BEST QUADRATIC UNBIASED ESTIMATORS

As we have seen in Section 10.9, the uniformly best quadratic unbiased estimators (BQUE) for the variance components in the case of unbalanced data do not exist. Townsend (1968) and Townsend and Searle (1971) have obtained locally BQUEs for σ_e^2 and σ_α^2 in the case of the model in (11.1.1) with $\mu = 0$. In this section, we outline their development briefly.

With $\mu = 0$, the model in (11.1.1) reduces to

$$y_{ij} = \alpha_i + e_{ij}, \quad i = 1, \ldots, a; \quad j = 1, \ldots, n_i,$$

which in the notation of the general linear model in (10.7.1) can be written as

$$\boldsymbol{Y} = \boldsymbol{U\beta} + \boldsymbol{e}$$

with

$$E(\boldsymbol{Y}) = \mathbf{0}$$

and

$$\text{Var}(\boldsymbol{Y}) = \boldsymbol{V} = \sigma_\alpha^2 \boldsymbol{U}\boldsymbol{U}' + \sigma_e^2 \boldsymbol{I}_N.$$

Now, let the estimators of σ_e^2 and σ_α^2 be given by

$$\hat{\sigma}_e^2 = \boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$$

and (11.4.20)

$$\hat{\sigma}_\alpha^2 = \boldsymbol{Y}'\boldsymbol{B}\boldsymbol{Y},$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are symmetric matrices chosen subject to the conditions that (11.4.20) are unbiased and have minimum variances. From the results on the mean and variance of a quadratic form, as given in Theorem 9.3.1, we have

$$E(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) = \text{tr}(\boldsymbol{A}\boldsymbol{V}), \qquad E(\boldsymbol{Y}'\boldsymbol{B}\boldsymbol{Y}) = \text{tr}(\boldsymbol{B}\boldsymbol{V}),$$
$$\text{Var}(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) = 2\,\text{tr}(\boldsymbol{A}\boldsymbol{V})^2, \quad \text{and} \quad \text{Var}(\boldsymbol{Y}'\boldsymbol{B}\boldsymbol{Y}) = 2\,\text{tr}(\boldsymbol{B}\boldsymbol{V})^2.$$

Then the conditions for the estimators in (11.4.20) to be BQUEs are

$$\begin{aligned} E(\hat{\sigma}_e^2) &= \text{tr}(\boldsymbol{A}\boldsymbol{V}) = \sigma_e^2, \\ E(\hat{\sigma}_\alpha^2) &= \text{tr}(\boldsymbol{B}\boldsymbol{V}) = \sigma_\alpha^2, \end{aligned} \tag{11.4.21}$$

and

$$\left.\begin{aligned} \text{Var}(\hat{\sigma}_e^2) &= 2\,\text{tr}(\boldsymbol{A}\boldsymbol{V})^2 \\ \text{Var}(\hat{\sigma}_\alpha^2) &= 2\,\text{tr}(\boldsymbol{B}\boldsymbol{V})^2 \end{aligned}\right\} \text{are minimum.} \tag{11.4.22}$$

Hence, the problem of determining BQUEs is to find the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ such that (11.4.22) is satisfied subject to the conditions in (11.4.21). After some lengthy algebraic manipulations, the BQUEs are obtained as (Townsend and Searle, 1971)

$$\sigma_{e,\text{BQUE}}^2 = \frac{1}{rs - t^2}\left[\sum_{i=1}^{a} \frac{s - tn_i}{(1 + n_i\tau)^2}\frac{y_{i.}^2}{n_i} + s\left\{\sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{a} n_i \bar{y}_{i.}^2\right\}\right]$$

and (11.4.23)

$$\hat{\sigma}_{\alpha,\text{BQUE}}^2 = \frac{1}{rs - t^2}\left[\sum_{i=1}^{a} \frac{rn_i - t}{(1 + n_i\tau)^2}\frac{y_{i.}^2}{n_i} - t\left\{\sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{a} n_i \bar{y}_{i.}^2\right\}\right],$$

where

$$\tau = \frac{\sigma_\alpha^2}{\sigma_e^2}, \qquad r = \sum_{i=1}^{a}(1 + n_i\tau)^{-2} + N - a,$$

$$s = \sum_{i=1}^{a} n_i^2/(1+n_i\tau)^2, \quad \text{and} \quad t = \sum_{i=1}^{a} n_i/(1+n_i\tau)^2,$$

with

$$N = \sum_{i=1}^{a} n_i.$$

It should be noted that the estimators in (11.4.23) are functions of $\tau = \sigma_\alpha^2/\sigma_e^2$ and not of the components individually. Further, it can be shown that as $\tau \to \infty$, i.e., when σ_α^2 is quite large compared to σ_e^2, we have (Townsend, 1968)

$$\begin{aligned} \lim \hat{\sigma}^2_{e,\text{BQUE}} &= \frac{1}{N-a}\left[\sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{a} n_i \bar{y}_{i.}^2\right] \\ &= \hat{\sigma}^2_{e,\text{ANOV}} \quad \text{(for the zero mean model)} \end{aligned}$$

and (11.4.24)

$$\lim \hat{\sigma}^2_{\alpha,\text{BQUE}} = \frac{1}{a}\left[\sum_{i=1}^{a} \bar{y}_{i.}^2 - \sum_{i=1}^{a} n_i^{-1}\hat{\sigma}^2_{e,\text{ANOV}}\right].$$

Thus the $\lim \hat{\sigma}^2_{\alpha,\text{BQUE}}$ is not $\hat{\sigma}^2_{\alpha,\text{ANOV}}$ as is the case with $\lim \hat{\sigma}^2_{e,\text{BQUE}}$. Also, on the other end of the scale, when $\tau \to 0$, we get (Townsend, 1968)

$$\begin{aligned} \lim \hat{\sigma}^2_{e,\text{BQUE}} &= \frac{1}{N\left(\sum_{i=1}^{a} n_i^2 - N\right)}\left[\sum_{i=1}^{a}\left(\sum_{i=1}^{a} n_i^2 - Nn_i\right)\frac{y_{i.}^2}{n_i}\right. \\ &\qquad \left. + (N-a)\left(\sum_{i=1}^{a} n_i^2\right)\hat{\sigma}^2_{e,\text{ANOV}}\right] \\ &= \hat{\sigma}^2_{e,\text{SSP}} \quad \text{(for the zero mean model)} \end{aligned}$$

and (11.4.25)

$$\begin{aligned} \lim \hat{\sigma}^2_{\alpha,\text{BQUE}} &= \frac{1}{\left(\sum_{i=1}^{a} n_i^2 - N\right)}\left[\sum_{i=1}^{a}(n_i^2 - n_i)\bar{y}_{i.}^2 - (N-a)\hat{\sigma}^2_{e,\text{ANOV}}\right] \\ &= \hat{\sigma}^2_{\alpha,\text{SSP}} \quad \text{(for the zero mean model).} \end{aligned}$$

11.4.7 NAQVI'S GOODNESS-OF-FIT ESTIMATORS

Naqvi's goodness-of-fit procedure for obtaining estimators of σ_e^2 and σ_α^2 in the balanced model discussed in Section 2.4.7 can in principle be extended to the unbalanced case. However, for the model in (11.1.1), the between sum squares SS_B is not distributed as a constant times a chi-square variate, and consequently explicit expressions for the variance component estimators cannot be obtained by this method. However, approximate estimators can be developed by approximating SS_B by a chi-square variate (see Section 11.6.2).

11.4.8 RAO'S MIVQUE AND MINQUE

The general theory of C.R. Rao's MIVQUE and MINQUE procedures was discussed in Section 10.10. In this section, we present the MIVQUE and MINQUE estimators of σ_e^2 and σ_α^2 for the model in (11.1.1).

Writing the vector of observations y_{ij}s in lexicon order as

$$\boldsymbol{Y}' = (y_{11}, \ldots, y_{1n_1}; \ldots; \ldots; y_{a1}, \ldots, y_{an_a}),$$

the model in (11.1.1) can be written as

$$\boldsymbol{Y} = \mu \boldsymbol{X} + \boldsymbol{U}_1 \beta + \boldsymbol{U}_2 \boldsymbol{e}, \tag{11.4.26}$$

where

$$\boldsymbol{X} = \boldsymbol{1}_N, \qquad \boldsymbol{U}_1 = \sum_{i=1}^{a}{}^{+} \boldsymbol{1}_{n_i}, \qquad \boldsymbol{U}_2 = \boldsymbol{I}_N,$$
$$\boldsymbol{\beta}' = (\beta_1, \ldots, \beta_a), \qquad \boldsymbol{e}' = (e_{11}, \ldots, e_{an_a}),$$

$\boldsymbol{1}_N$ is an N-vector containing all 1s, $\boldsymbol{I}_N$ is an identity matrix of order N, and Σ^+ denotes a direct sum of matrices. Furthermore, the mean vector and variance-covariance matrix of $\boldsymbol{Y}$ in (11.4.26) are given by

$$\left.\begin{aligned} E(\boldsymbol{Y}) &= \mu \boldsymbol{X} \\ \text{and} \qquad \operatorname{Cov}(\boldsymbol{Y}) &= \boldsymbol{V} = \sigma_\alpha^2 \boldsymbol{V}_1 + \sigma_e^2 \boldsymbol{V}_2 \end{aligned}\right\} \tag{11.4.27}$$

where

$$\boldsymbol{V}_1 = \boldsymbol{U}_1 \boldsymbol{U}_1' = \sum_{i=1}^{a}{}^{+} \boldsymbol{J}_{n_i},$$
$$\boldsymbol{V}_2 = \boldsymbol{U}_2 \boldsymbol{U}_2' = \boldsymbol{I}_N,$$

and $\boldsymbol{J}_{n_i}$ is a square matrix of order n_i containing all 1s.

11.4.8.1 The MIVQUE

In the general notation of Section 10.10, assuming normality, the MIVQUE vector of $\boldsymbol{\sigma}^2 = (\sigma_\alpha^2, \sigma_e^2)'$ is given by

$$\boldsymbol{\sigma}^2 = \boldsymbol{S}^{-1} \boldsymbol{\gamma}, \tag{11.4.28}$$

where

$$\boldsymbol{S} = \{s_{ij}\} = \operatorname{tr}(\boldsymbol{V}_i \boldsymbol{R} \boldsymbol{V}_j \boldsymbol{R}), \quad i, j = 1, 2, \tag{11.4.29}$$

$$\left.\begin{aligned} \boldsymbol{\gamma}' &= (\gamma_1, \gamma_2), \\ \gamma_i &= \boldsymbol{Y}' \boldsymbol{R} \boldsymbol{V}_i \boldsymbol{R} \boldsymbol{Y}, \quad i = 1, 2, \end{aligned}\right. \tag{11.4.30}$$

$$\boldsymbol{R} = \boldsymbol{V}^{-1}[\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}]. \tag{11.4.31}$$

From (11.4.28), the MIVQUEs of σ_e^2 and σ_α^2 are given by

$$\hat{\sigma}^2_{e,\text{MIVQ}} = \frac{1}{|\boldsymbol{S}|}(s_{11}\gamma_2 - s_{12}\gamma_1)$$

and (11.4.32)

$$\hat{\sigma}^2_{\alpha,\text{MIVQ}} = \frac{1}{|\boldsymbol{S}|}(s_{22}\gamma_1 - s_{12}\gamma_2),$$

where

$$|\boldsymbol{S}| = |s_{11}s_{22} - s_{12}^2|.$$

After evaluating for $\boldsymbol{R}$ from (11.4.31) and substituting it in (11.4.29) and (11.4.30), one obtains, after some simplifications (see Swallow, 1974; Swallow and Searle, 1978),

$$s_{11} = \sum_{i=1}^{a} k_i^2 - 2k\sum_{i=1}^{a} k_i^3 + k^2\left(\sum_{i=1}^{a} k_i^2\right)^2,$$

$$s_{12} = \sum_{i=1}^{a} \frac{k_i^2}{n_i} - 2k\sum_{i=1}^{a} \frac{k_i^3}{n_i} + k^2\sum_{i=1}^{a} k_i^2 \sum_{i=1}^{a} \frac{k_i^2}{n_i},$$

$$s_{22} = \frac{N-a}{\sigma_e^4} + \sum_{i=1}^{a} \frac{k_i^2}{n_i} - 2k\sum_{i=1}^{a} \frac{k_i^3}{n_i^2} + k^2\left(\sum_{i=1}^{a} \frac{k_i^2}{n_i}\right)^2,$$

$$\gamma_1 = \sum_{i=1}^{a} k_i^2\left(\bar{y}_{i.} - k\sum_{i=1}^{a} k_i\bar{y}_{i.}\right)^2,$$

and

$$\gamma_2 = \frac{1}{\sigma_e^4}\left[\sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{a} n_i\bar{y}_{i.}^2\right] + \sum_{i=1}^{a} \frac{k_i^2}{n_i}\left(\bar{y}_{i.} - k\sum_{i=1}^{a} k_i\bar{y}_{i.}\right)^2,$$

where

$$k_i = \frac{n_i}{\sigma_e^2 + n_i\sigma_\alpha^2}, \qquad k = 1/\sum_{i=1}^{a} k_i, \quad \text{and} \quad \bar{y}_{i.} = \sum_{j=1}^{n_i} y_{ij}/n_i.$$

The MIVQUEs for the case with $\mu = 0$ will also be given by (11.4.32) except that now $\boldsymbol{X} = \boldsymbol{0}$, so that considerable simplifications result in the expressions of s_{ij}s and γ_is. Thus, with $\boldsymbol{X} = \boldsymbol{0}$, $\boldsymbol{R} = \boldsymbol{V}^{-1}$, we obtain

$$s_{ij} = \text{tr}(\boldsymbol{V}_i\boldsymbol{V}^{-1}\boldsymbol{V}_j\boldsymbol{V}^{-1}), \quad i, j = 1, 2,$$

and

$$\gamma_i = \boldsymbol{Y}'\boldsymbol{V}^{-1}\boldsymbol{V}_i\boldsymbol{V}^{-1}\boldsymbol{Y}, \quad i = 1, 2,$$

which, after some simplifications, lead to

$$s_{11} = \left(\sum_{i=1}^{a} n_i^2 q_i^2\right) / \sigma_e^4,$$

$$s_{12} = \left(\sum_{i=1}^{a} n_i^2 q_i^2\right) / \sigma_e^4,$$

$$s_{22} = \left(\sum_{i=1}^{a} q_i^2 + N - a\right) / \sigma_e^4,$$

$$\gamma_1 = \left(\sum_{i=1}^{a} q_i^2 y_{i.}^2\right) / \sigma_e^4,$$

and

$$\gamma_2 = \left(\sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{a} \frac{y_{i.}^2}{n_i} + \sum_{i=1}^{a} q_i^2 \frac{y_{i.}^2}{n_i}\right) / \sigma_e^4,$$

where

$$q_i = \frac{\sigma_e^2}{\sigma_e^2 + n_i\sigma_\alpha^2} \quad \text{and} \quad y_{i.} = \sum_{j=1}^{n_i} y_{ij}.$$

It can be seen that the resulting estimators are identical to the BQUEs given by (11.4.23). P. S. R.S. Rao (1982) discussed the use of prior information for MIVQUE estimators and Rao (2001) proposed some nonnegative modifications of MIVQUE.

11.4.8.2 The MINQUE

As we know from Section 10.10, the MINQUEs of σ_e^2 and σ_α^2 are also given by (11.4.32) except that the matrix $\boldsymbol{V}$ is now replaced by $\boldsymbol{V}^*$, defined as

$$\boldsymbol{V}^* = \boldsymbol{V}_1 + \boldsymbol{V}_2 = \sum_{i=1}^{a}{}^{+} (\boldsymbol{I}_{n_i} + \boldsymbol{J}_{n_i}),$$

and the matrix $\boldsymbol{R}$ is given by

$$\boldsymbol{R} = \boldsymbol{V}^{*-1}[\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{*-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{*-1}]. \tag{11.4.33}$$

After evaluating for $\boldsymbol{R}$ from (11.4.33) and substituting it into (11.4.29) and (11.4.30), we obtain, after some simplifications (see Swallow, 1974; Swallow and Searle, 1978),

$$\begin{aligned}
s_{11} &= \sum_{i=1}^{a} \theta_i^2 - 2\theta \sum_{i=1}^{a} \theta_i^3 + \theta^2 \left(\sum_{i=1}^{a} \theta_i^2 \right)^2, \\
s_{12} &= -s_{11} - \theta \sum_{i=1}^{a} \theta_i^2 + \sum_{i=1}^{a} \theta_i, \\
s_{22} &= s_{11} + N - 1 - 2 \sum_{i=1}^{a} \theta_i + 2\theta \sum_{i=1}^{a} \theta_i^2, \qquad (11.4.34) \\
\gamma_1 &= \sum_{i=1}^{a} \theta_i^2 \left(\bar{y}_{i.} - \theta \sum_{i=1}^{a} \theta_i \bar{y}_{i.} \right)^2,
\end{aligned}$$

and

$$\gamma_2 = -\gamma_1 + \sum_{i=1}^{a} \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{a} n_i \bar{y}_{i.}^2 + \sum_{i=1}^{a} \theta_i \left(\bar{y}_{i.} - \theta \sum_{i=1}^{a} n_i \bar{y}_{i.} \right)^2,$$

where

$$\theta_i = n_i/(1 + n_i) \quad \text{and} \quad \theta = 1/\sum_{i=1}^{a} \theta_i.$$

The resulting estimators are obtained by substituting s_{ij}s and γ_is given in (11.4.34) into (11.4.32).

Explicit expressions for the MINQUE estimators are also developed by Ahrens (1978). Hess (1979) has investigated the sensitivity of the MINQUE estimators with respect to a priori weights. Rao et al. (1981) discuss the MINQUE estimators when the common value of the relative a priori weight is equal to unity. Chaubey (1984) and Rao (2001) have considered various modifications of the MINQUEs so that they yield nonnegative estimates. For some further results on the MINQUE estimators of the one-way model, see Rao and Sylvestre (1984).

The MINQUEs for the case with $\mu = 0$ will also be given by (11.4.32) except that now $\boldsymbol{X} = \boldsymbol{0}$, so that, as in the case of MIVQUEs, s_{ij}s and γ_is simplify to

$$\begin{aligned}
s_{11} &= \sum_{i=1}^{a} n_i^2 \phi_i^2, \\
s_{12} &= \sum_{i=1}^{a} n_i \phi_i^2,
\end{aligned}$$

$$s_{22} = \sum_{i=1}^{a} \phi_i^2 + N - a,$$

$$\gamma_1 = \sum_{i=1}^{a} \phi_i^2 y_{i.}^2,$$

and

$$\gamma_2 = \sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{a} \frac{y_{i.}^2}{n_i} + \sum_{i=1}^{a} \frac{\phi_i^2 y_{i.}^2}{n_i},$$

where

$$\phi_i = 1/(1 + n_i).$$

11.4.9 AN UNBIASED ESTIMATOR OF $\sigma_\alpha^2/\sigma_e^2$

The estimator of $\sigma_\alpha^2/\sigma_e^2$ given by

$$\frac{\hat{\sigma}^2_{\alpha,\text{ANOV}}}{\hat{\sigma}^2_{e,\text{ANOV}}} = \frac{\text{MS}_B - \text{MS}_W}{n_0 \text{MS}_W},$$

is biased. An unbiased estimator of $\sigma_\alpha^2/\sigma_e^2$, assuming normality, is obtained as

$$\frac{1}{n_0}\left[\frac{(N-a-2)}{(N-a)} \cdot \frac{\text{MS}_B}{\text{MS}_W} - 1\right]. \qquad (11.4.35)$$

Result (11.4.35) is given in Crump (1954) and Anderson and Crump (1967). When the model is balanced, we saw in Section 2.4.9 that the estimator (11.4.35) has uniformly minimum variance among all unbiased estimators. The sampling variance of the estimator in (11.4.35) is given in (11.6.16)

11.4.10 ESTIMATION OF $\sigma_\alpha^2/(\sigma_e^2 + \sigma_\alpha^2)$

An unbiased estimator for the intraclass correlation $\rho = \sigma_\alpha^2/(\sigma_e^2 + \sigma_\alpha^2)$ does not have a closed form expression (Olkin and Pratt, 1958). A simple biased estimator based on the ANOVA estimators of σ_α^2 and σ_e^2 is

$$\hat{\rho}_{\text{ANOV}} = \frac{\hat{\sigma}^2_{\alpha,\text{ANOV}}}{\hat{\sigma}^2_{e,\text{ANOV}} + \hat{\sigma}^2_{\alpha,\text{ANOV}}} = \frac{\text{MS}_B - \text{MS}_W}{\text{MS}_B + (n_0 - 1)\text{MS}_W}. \qquad (11.4.36)$$

The estimator (11.4.36) is known as the analysis of variance estimator. Although not unbiased, it is consistent for ρ and the degree of bias is very slight (Van der Kemp, 1972). A serious drawback of the estimator (11.4.36) is that it can assume a negative value whenever $\text{MS}_B < \text{MS}_W$. In practice, a negative value is often set equal to zero, resulting in a truncated estimator.[4] Another biased estimator of $\sigma_\alpha^2/(\sigma_e^2 + \sigma_\alpha^2)$ obtained by using the unbiased estimator of $\sigma_\alpha^2/\sigma_e^2$ from (11.4.35) is

$$\hat{\rho}' = \frac{\widehat{\left(\frac{\sigma_\alpha^2}{\sigma_e^2}\right)}_{\text{UNB}}}{1 + \widehat{\left(\frac{\sigma_\alpha^2}{\sigma_e^2}\right)}_{\text{UNB}}} = \frac{\frac{1}{n_0}\left[\frac{(N-a-2)}{(N-a)} \cdot \frac{\text{MS}_B}{\text{MS}_W} - 1\right]}{1 + \frac{1}{n_0}\left[\frac{(N-a-2)}{(N-a)} \cdot \frac{\text{MS}_B}{\text{MS}_W} - 1\right]}.$$

The oldest estimator of ρ was proposed by Karl Pearson as the product moment correlation computed over all possible pairs of observations that can be constructed within groups. Rao (1973, p. 268) also considered an estimator of ρ as the sample correlation of sibling pairs. Karlin et al. (1981) and Namboodiri et al. (1984) have considered modifications to the Pearson estimator where each pair is weighted according to some weighting scheme. As in the case of the variance components for the model in (11.1.1), the ML estimator of ρ cannot be obtained in explicit form. However, from the invariance property of the ML estimation, it follows that the ML estimator of ρ can be obtained as a direct function of the ML estimators of σ_e^2 and σ_α^2. Donner and Koval (1980a) provide an algorithm for computing the ML estimator of ρ under the common correlation model. Some results on efficiency calculation show that the ML estimator is more accurate than the ANOVA estimator (11.4.36) for very small and very large values of ρ ($\rho \leq 0.1$, $\rho \geq 0.8$) while the two estimators are about equally accurate for $0.1 < \rho < 0.8$. Kleffé (1993) derived computable expressions for MINQUE estimators of ρ and their limiting sample variances and covariances. Bansal and Bhandary (1994) have considered robust M-estimation. For some other estimation procedures for ρ, see Smith (1980a,1980b) and Bener and Huda (1993).

11.4.11 A NUMERICAL EXAMPLE

Brownlee (1965, p. 133) reported some of the results of Rosa and Dorsey (A new determination of the ratio of the electromagnetic to the electrostatic unit of electricity, *Bull. Nat. Bur. Standards*, **3** (1990), pp. 433–604) on the ratio of the electromagnetic to electrostatic units of electricity, a constant which equals the velocity of light. The five groups in the study correspond to successive dismantling and reassembly of the apparatus and can be considered a sample of a large number of such groups. The data are given in Table 11.2.

We will use the one-way random effects model in (11.1.1) to analyze the data in Table 11.2. In this example, $a = 5, n_1 = 11, n_2 = 8, n_3 = 6, n_4 = 24, n_5 = 15$; $i = 1, 2, 3, 4, 5$ refer to the groups; and $j = 1, 2, \ldots, n_i$ refer to replications within the groups. Further, σ_α^2 designates the variance component due to group and σ_e^2 denotes the error variance component which includes variability in

[4] Singh (1991) has investigated the probability of obtaining a negative estimate for the estimator (11.4.36).

TABLE 11.2 The ratio of the electromagnetic to electrostatic units of electricity.

Groups						
1	**2**	**3**	**4**		**5**	
62	65	65	62	65	66	64
64	64	64	66	63	65	65
62	63	67	64	63	65	64
62	62	62	64	63	66	
65	65	65	63	61	67	
64	63	62	62	56	66	
65	64		64	64	69	
62	63		64	64	70	
62			66	65	68	
63			64	64	69	
64			66	64	63	
			63	65	65	

All figures had 2.99 subtracted from them and then multiplied by 10,000.

Source: Brownlee (1965); used with permission.

TABLE 11.3 Analysis of variance for the ratio units of electricity data of Table 11.2.

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Group	4	80.4011	20.1003	$\sigma_e^2 + 12.008\sigma_\alpha^2$
Error	59	198.0364	3.3565	σ_e^2
Total	63	278.4375		

measurement as well as the sampling error. The calculations leading to the analysis of variance are readily performed and the results are summarized in Table 11.3. The selected outputs using SAS®GLM, SPSS®GLM, and BMDP® 3V are displayed in Figure 11.1.

We now illustrate the calculations of point estimates of the variance components $\sigma_e^2, \sigma_\alpha^2$, and certain of their parametric functions. The analysis of variance estimates in (11.4.1) are

$$\hat{\sigma}^2_{e,\text{ANOV}} = \frac{198.0364}{59} = 3.357$$

```
DATA SAHAIC11;
INPUT GROUP YIELD;
CARDS;
1 62
1 64
1 62
1 62
1 65
1 64
.  .
5 64
;
PROC GLM;
CLASS GROUP;
MODEL YIELD =GROUP;
RANDOM GROUP;
RUN;
CLASS  LEVELS   VALUES
GROUP    5      1 2 3 4 5
NUMBER OF OBSERVATIONS IN
DATA SET=64
```

```
                          The SAS System
                   General Linear Models Procedure
Dependent Variable: YIELD
                          Sum of        Mean
Source       DF      Squares        Square     F Value   Pr > F
Model         4      80.401136    20.100284      5.99    0.0004
Error        59     198.036363     3.356548
Corrected    63     278.437500
Total
           R-Square      C.V.      Root MSE     YIELD Mean
           0.288758    2.855667   1.83208857     64.15625

Source      DF    Type I SS    Mean Square   F Value    Pr > F
GROUP        4    80.401136     20.100284     5.99     0.0004
Source      DF    Type III SS  Mean Square   F Value    Pr > F
GROUP        4    80.401136     20.100284     5.99     0.0004

Source      Type III Expected Mean Square
GROUP       Var(Error) + 12.008 Var(GROUP)
```

SAS application: This application illustrates SAS GLM instructions and output for the unbalanced one-way random effects analysis of variance.[a,b]

```
DATA SAHAIC11
/GROUP 1 YIELD 3-6(1)
BEGIN DATA.
1 62
1 64
1 62
1 62
1 65
1 64
1 65
.  .
5 64
END DATA.
GLM YIELD BY
 GROUP
/DESIGN GROUP
/METHOD SSTYPE(1)
/RANDOM GROUP.
```

```
                    Tests of Between-Subjects Effects
Dependent Variable: YIELD

Source                  Type I SS    df   Mean Square    F       Sig.
GROUP     Hypothesis      80.401      4     20.100     5.988    0.000
          Error          198.036     59      3.357(a)
a MS(Error)

                      Expected Mean Squares(b,c)
                          Variance Component
Source                Var(GROUP)                 Var(ERROR)
GROUP                   12.008                     1.000
ERROR                     .000                     1.000

b For each source, the expected mean square equals the sum of the
  coefficients in the cells times the variance components, plus a
  quadratic term involving effects in the Quadratic Term cell.
c Expected Mean Squares are based on the Type I Sums of Squares.
```

SPSS application: This application illustrates SPSS GLM instructions and output for the unbalanced one-way random effects analysis of variance.[a,b]

```
/INPUT    FILE='C:\SAHAIC11.TXT'.
          FORMAT=FREE.
          VARIABLES=2.
/VARIABLE NAMES=GROUP,YIELD.
/GROUP    CODES(GROUP)=1,2,3,4,5.
          NAMES(GROUP)=G1,G2,G3,
                       G4,G5.
/DESIGN DEPENDENT=YIELD.
           RANDOM=GROUP.
           METHOD=REML.
/END
1 62
1 64
1 62
1 62
.  .
5 64
```

```
    BMDP3V - GENERAL MIXED MODEL ANALYSIS OF VARIANCE
           Release: 7.0        (BMDP/DYNAMIC)
        DEPENDENT VARIABLE YIELD

PARAMETER   ESTIMATE   STANDARD   EST/      TWO-TAILPROB.
                       ERROR      ST.DEV.   (ASYM.THEORY)
 ERR.VAR.     3.345     0.614
 CONSTANT    64.144     0.554     115.870       0.000
 RAND(1)      1.218     1.032

   TESTS OF FIXED EFFECTS BASED ON ASYMPTOTIC VARIANCE
                  -COVARIANCE MATRIX

 SOURCE    F-STATISTIC   DEGREES OF     PROBABILITY
                         FREEDOM
CONSTANT    13425.92      1    63         0.00000
```

BMDP application: This application illustrates BMDP 8V instructions and output for the unbalanced one-way random effects analysis of variance.[a,b]

[a] Several portions of the output were extensively edited and doctored to economize space and may not correspond to the original printout.

[b] Results on significance tests may vary from one package to the other.

FIGURE 11.1 Program instructions and output for the unbalanced one-way random effects analysis of variance: Data on the ratio of the electromagnetic to electrostatic units of electricity (Table 11.2).

and

$$\hat{\sigma}^2_{\alpha,\text{ANOV}} = \frac{1}{12.008}\left(\frac{80.4011}{4} - \frac{198.0364}{59}\right) = 1.394.$$

We used SAS® VARCOMP, SPSS® VARCOMP, and BMDP® 3V to estimate the variance components using the ML, REML, MINQUE(0), and MINQUE(1) procedures.[5] The desired estimates using these software are given in Table 11.4. Note that all three software produce nearly the same results except for some minor discrepancy in rounding decimal places.

Finally, we can obtain estimates of $\sigma_\alpha^2/\sigma_e^2$, $\sigma_\alpha^2/(\sigma_e^2 + \sigma_\alpha^2)$, and $\sigma_e^2 + \sigma_\alpha^2$ based on the ANOVA, ML, REML, MINQUE(0), and MINQUE(1) estimates of the variance components and the results are summarized in Table 11.5.

11.5 BAYESIAN ESTIMATION

In this section, we consider some results on the Bayesian analysis of the model in (11.1.1) given in Hill (1965). Hill obtained an expression for the joint and marginal posterior densities of σ_e^2 and σ_α^2 under the assumption that the prior opinion for μ is diffuse and effectively independent of that for σ_e^2 and σ_α^2, i.e., roughly

$$p(\mu, \sigma_e^2, \sigma_\alpha^2) = p(\sigma_e^2, \sigma_\alpha^2), \tag{11.5.1}$$

where $p(\sigma_e^2, \sigma_\alpha^2)$ is the subjective prior density of σ_e^2 and σ_α^2. Hill also obtained the joint posterior density of σ_e^2 and $\sigma_\alpha^2/\sigma_e^2$, the marginal posterior density of $\sigma_\alpha^2/\sigma_e^2$, and the conditional posterior density of σ_e^2 given $\sigma_\alpha^2/\sigma_e^2$ based on the prior

$$p(\sigma_e^2, \sigma_\alpha^2) = p_e(\sigma_e^2)p_\alpha(\sigma_\alpha^2), \tag{11.5.2}$$

where $p_e(\sigma_e^2) \propto 1/\sigma_e^2$ and $p_\alpha(\sigma_\alpha^2)$ is such that $1/\sigma_\alpha^2$ has a gamma distribution.

Hill (1965) gave special consideration to the problem of approximating the marginal posterior distributions of σ_e^2 and σ_α^2. He showed that if the posterior density of $\sigma_\alpha^2/\sigma_e^2$ points sharply to some positive value, then the posterior probability that σ_e^2 or σ_α^2 assumes a value in a given interval can be obtained from the chi-square distribution. More generally, he pointed out that these posterior probabilities can be evaluated by using tables of chi-square and F distributions and by performing a numerical integration. Hill considered both situations where the likelihood function is sharp or highly concentrated relative to the prior distributions, and also where although the likelihood function is expected to be relatively sharp (on the basis of, say, Fisherian information) before the experiment, it is in actual fact not, as for example when $\text{MS}_B \leq \text{MS}_W$.

[5]The computations for ML and REML estimates were also carried out using SAS® PROC MIXED and some other programs to assess their relative accuracy and convergence rate. There did not seem to be any appreciable differences between the results from different software.

TABLE 11.4 ML, REML, MINQUE(0), and MINQUE(1) estimates of the variance components using SAS®, SPSS®, and BMDP® software.

Variance component	SAS®		
	ML	REML	MINQUE(0)
σ_e^2	3.339368	3.344529	3.204551
σ_α^2	0.936709	1.217570	1.593770

Variance component	SPSS®			
	ML	REML	MINQUE(0)	MINQUE(1)
σ_e^2	3.333937	3.344529	3.204551	3.354307
σ_α^2	0.936706	1.217570	1.593770	1.122615

Variance component	BMDP®	
	ML	REML
σ_e^2	3.333937	3.344529
σ_α^2	0.936706	1.217570

SAS® VARCOMP does not compute MINQUE(1). BMDP3V does not compute MINQUE(0) and MINQUE(1).

TABLE 11.5 Point estimates of some parametric functions of σ_α^2 and σ_e^2.

Parametric function	Method of estimation	Point estimate
$\sigma_\alpha^2/\sigma_e^2$	ANOVA	0.415
	ML	0.281
	REML	0.364
	MINQUE(0)	0.497
	MINQUE(1)	0.335
$\sigma_\alpha^2/(\sigma_e^2+\sigma_\alpha^2)$	ANOVA	0.293
	ML	0.219
	REML	0.267
	MINQUE(0)	0.332
	MINQUE(1)	0.251
$\sigma_e^2+\sigma_\alpha^2$	ANOVA	4.751
	ML	4.276
	REML	4.563
	MINQUE(0)	4.799
	MINQUE(1)	4.477

11.5.1 JOINT POSTERIOR DISTRIBUTION OF $(\sigma_e^2, \sigma_\alpha^2)$

We know from (11.4.13) that the likelihood function is given by

$$L(\mu, \sigma_e^2, \sigma_\alpha^2 | \text{ data } y_{ij}) \propto \frac{\exp\left\{-\frac{1}{2}\frac{SS_W}{\sigma_e^2}\right\}\exp\left\{-\frac{1}{2}\sum_{i=1}^{a}\frac{n_i(\bar{y}_{i.}-\mu)^2}{\sigma_e^2+n_i\sigma_\alpha^2}\right\}}{(\sigma_e^2)^{\frac{1}{2}(N-a)}\prod_{i=1}^{a}(\sigma_e^2+n_i\sigma_\alpha^2)^{1/2}}. \tag{11.5.3}$$

Combining the likelihood function in (11.5.3) with the prior distribution in (11.5.1), the approximate marginal posterior density of $(\sigma_e^2, \sigma_\alpha^2)$ is

$$\begin{aligned}
p(\sigma_e^2, \sigma_\alpha^2 | \text{ data } y_{ij}) &= \int_{-\infty}^{\infty} p(\mu, \sigma_e^2, \sigma_\alpha^2 | \text{ data } y_{ij}) d\mu \\
&= \int_{-\infty}^{\infty} p(\mu, \sigma_e^2, \sigma_\alpha^2) L(\mu, \sigma_e^2, \sigma_\alpha^2 | \text{ data } y_{ij}) d\mu \\
&\propto \frac{p(\sigma_e^2,\sigma_\alpha^2)\exp\left\{-\frac{1}{2}\frac{SS_W}{\sigma_e^2}\right\}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2}\sum_{i=1}^{a}\frac{n_i(\bar{y}_{i.}-\mu)^2}{\sigma_e^2+n_i\sigma_\alpha^2}\right\}d\mu}{(\sigma_e^2)^{\frac{1}{2}(N-a)}\prod_{i=1}^{a}(\sigma_e^2+n_i\sigma_\alpha^2)^{1/2}} \\
&\propto \frac{p(\sigma_e^2,\sigma_\alpha^2)\exp\left\{-\frac{1}{2}\frac{SS_W}{\sigma_e^2}\right\}\exp\left\{-\frac{1}{2}\sum_{i=1}^{a}\frac{n_i(\bar{y}_{i.}-\hat{\mu})^2}{\sigma_e^2+n_i\sigma_\alpha^2}\right\}}{(\sigma_e^2)^{\frac{1}{2}(N-a)}\prod_{i=1}^{a}(\sigma_e^2+n_i\sigma_\alpha^2)^{1/2}\left\{\sum_{i=1}^{a}\frac{n_i}{\sigma_e^2+n_i\sigma_\alpha^2}\right\}^{1/2}},
\end{aligned} \tag{11.5.4}$$

where

$$\hat{\mu} = \sum_{i=1}^{a}\frac{n_i\bar{y}_{i.}}{\sigma_e^2+n_i\sigma_\alpha^2} \Big/ \sum_{i=1}^{a}\frac{n_i}{\sigma_e^2+n_i\sigma_\alpha^2}. \tag{11.5.5}$$

11.5.2 JOINT POSTERIOR DISTRIBUTION OF $(\sigma_e^2, \sigma_\alpha^2/\sigma_e^2)$

Using the joint posterior density in (11.5.4) and taking the prior distribution of $(\sigma_e^2, \sigma_\alpha^2)$ as

$$p(\sigma_e^2, \sigma_\alpha^2) \propto (\sigma_e^2)^{-1}(\sigma_\alpha^2)^{-\frac{1}{2}\lambda_\alpha - 1}\exp\{-c_\alpha/2\sigma_\alpha^2\}, \tag{11.5.6}$$

so that $1/\sigma_\alpha^2$ has a gamma distribution with parameters $\lambda_\alpha/2$ and $c_\alpha/2$, the posterior density of (σ_e^2, τ), where $\tau = \sigma_\alpha^2/\sigma_e^2$, is given by

$$\begin{aligned}
&p(\sigma_e^2, \tau | \text{ data } y_{ij}) \\
&\quad\propto \sigma_e^2 p(\sigma_e^2, \sigma_e^2\tau | \text{ data } y_{ij}) \\
&\quad\propto \frac{\sigma_e^2 p(\sigma_e^2, \sigma_e^2\tau)\exp\left\{-\frac{1}{2}\frac{SS_W}{\sigma_e^2}\right\}\exp\left\{-\frac{1}{2\sigma_e^2}\sum_{i=1}^{a}\frac{n_i(\bar{y}_{i.}-\hat{\mu})^2}{1+n_i\tau}\right\}}{(\sigma_e^2)^{\frac{1}{2}(N-1)}\prod_{i=1}^{a}(1+n_i\tau)^{1/2}\left\{\sum_{i=1}^{a}\frac{n_i}{1+n_i\tau}\right\}^{1/2}}
\end{aligned}$$

$$\propto \frac{\exp\left[-\frac{1}{2\sigma_e^2}\left\{\frac{\tau \mathrm{SS}_W + c_\alpha}{\tau} + \sum_{i=1}^{a} \frac{n_i(\bar{y}_{i.}-\hat{\mu})^2}{1+n_i\tau}\right\}\right]}{(\sigma_e^2)^{\frac{1}{2}(N-1+\lambda_\alpha)+1}(\tau)^{\frac{1}{2}}\lambda_\alpha + 1 \prod_{i=1}^{a}(1+n_i\tau)^{1/2}\left\{\sum_{i=1}^{a}\frac{n_i}{1+n_i\tau}\right\}^{1/2}}. \tag{11.5.7}$$

11.5.3 CONDITIONAL POSTERIOR DISTRIBUTION OF σ_e^2 GIVEN τ

From (11.5.7), the marginal posterior density of τ is

$$p(\tau \mid \text{data } y_{ij}) \propto \frac{\left\{\mathrm{SS}_W + c_\alpha/\tau + \sum_{i=1}^{a} \frac{n_i(\bar{y}_{i.}-\hat{\mu})^2}{1+n_i\tau}\right\}^{-\frac{1}{2}(N-1+\lambda_\alpha)}}{(\tau)^{\frac{1}{2}\lambda_\alpha+1}\prod_{i=1}^{a}(1+n_i\tau)^{1/2}\left\{\sum_{i=1}^{a}\frac{n_i}{1+n_i\tau}\right\}^{1/2}}. \tag{11.5.8}$$

Now, the conditional posterior density of σ_e^2 given τ is obtained as

$$\begin{aligned} p(\sigma_e^2|\tau) &= p(\sigma_e^2, \tau \mid \text{data } y_{ij})/p(\tau \mid \text{data } y_{ij}) \\ &\propto \frac{\exp\left[-\frac{1}{2\sigma_e^2}\left\{\mathrm{SS}_W + c_\alpha/\tau + \sum_{i=1}^{a}\frac{n_i(\bar{y}_{i.}-\hat{\mu})^2}{1+n_i\tau}\right\}\right]}{(\sigma_e^2)^{\frac{1}{2}(N-1+\lambda_\alpha)+1}}. \end{aligned} \tag{11.5.9}$$

Hence, given τ, it follows that the variable

$$\frac{1}{\sigma_e^2}\left\{\mathrm{SS}_W + c_\alpha/\tau + \sum_{i=1}^{a}\frac{n_i(\bar{y}_{i.}-\hat{\mu})^2}{1+n_i\tau}\right\}$$

has a chi-square distribution with $N + \lambda_\alpha - 1$ degrees of freedom.

11.5.4 MARGINAL POSTERIOR DISTRIBUTIONS OF σ_e^2 and σ_α^2

Hill (1965) devoted considerable efforts to the problem of approximating the marginal posterior densities of σ_e^2 and σ_α^2. Employing the marginal and conditional densities (11.5.8) and (11.5.9), it follows that

$$\begin{aligned} P\{t_1 \le \sigma_e^2 \le t_2 \mid \text{data } y_{ij}\} &= \int_0^\infty P\{t_1 \le \sigma_e^2 \le t_2|\tau\} p(\tau \mid \text{data } y_{ij}) d\tau \\ &= \int_0^\infty P\left\{\frac{1}{t_1}\left[\mathrm{SS}_W + \frac{c_\alpha}{\tau} + \sum_{i=1}^{a}\frac{n_i(\bar{y}_{i.}-\hat{\mu})^2}{1+n_i\tau}\right]\right. \\ &\le \chi^2[N+\lambda_\alpha-1] \\ &\left.\le \frac{1}{t_2}\left[\mathrm{SS}_W + \frac{c_\alpha}{\tau} + \sum_{i=1}^{a}\frac{n_i(\bar{y}_{i.}-\hat{\mu})^2}{1+n_i\tau}\right]\right\} \end{aligned}$$

$$\times\, p(\tau|\,\text{data } y_{ij})d\tau. \qquad (11.5.10)$$

Similarly,

$$\begin{aligned} P\{t_1 \le \sigma_\alpha^2 \le t_2|\,\text{data } y_{ij}\} &= \int_0^\infty P\{t_1 \le \sigma_\alpha^2 \le t_2|\tau\}p(\tau|\,\text{data } y_{ij})d\tau \\ &= \int_0^\infty P\{t_1/\tau \le \sigma_\alpha^2 \le t_2/\tau|\tau\}p(\tau|\,\text{data } y_{ij})d\tau. \end{aligned} \qquad (11.5.11)$$

Now, each of the posterior probabilities in (11.5.10) and (11.5.11) is the integral of the chi-square probability of an interval whose endpoints are functions of τ with respect to the posterior distribution of τ. Hill mentions various approximations to evaluate (11.5.10) and (11.5.11). In particular, when posterior density of τ points sharply to some positive value, Hill shows that these probabilities can be determined in terms of a chi-square distribution. In general, even if the posterior density of τ is not particularly sharp, Hill observes that the posterior probabilities (11.5.10) and (11.5.11) can be approximated using chi-square and F distributions and by performing a numerical integration.

11.5.5 INFERENCES ABOUT μ

The joint posterior density of (μ, σ_e^2, τ) can be written as

$$p(\mu, \sigma_e^2, \tau|\,\text{data } y_{ij}) \propto \frac{\exp\left[-\frac{1}{2\sigma_e^2}\left\{\text{SS}_W + c_\alpha/\tau + \sum_{i=1}^a \frac{n_i(\bar{y}_{i.}-\mu)^2}{1+n_i\tau}\right\}\right]}{(\sigma_e^2)^{\frac{1}{2}(N+\lambda_\alpha)+1}(\tau)^{\frac{1}{2}\lambda_\alpha+1}\prod_{i=1}^a(1+n_i\tau)^{1/2}}. \qquad (11.5.12)$$

From (11.5.12), the joint posterior density of (μ, τ) is given by

$$\begin{aligned} p(\mu, \tau|\,\text{data } y_{ij}) &\propto (\tau)^{-\frac{1}{2}\lambda_\alpha-1}\prod_{i=1}^a(1+n_i\tau)^{-\frac{1}{2}} \\ &\quad\times \int_0^\infty \frac{\exp\left[-\frac{1}{2\sigma_e^2}\left\{\text{SS}_W + \frac{c_\alpha}{\tau} + \sum_{i=1}^a \frac{n_i(\bar{y}_{i.}-\mu)^2}{1+n_i\tau}\right\}\right]}{(\sigma_e^2)^{\frac{1}{2}(N+\lambda_\alpha)+1}}d\sigma_e^2 \\ &\propto \frac{\left\{\text{SS}_W + \frac{c_\alpha}{\tau} + \sum_{i=1}^a \frac{n_i(\bar{y}_{i.}-\mu)^2}{1+n_i\tau}\right\}^{-\frac{1}{2}(N+\lambda_\alpha)}}{(\tau)^{\frac{1}{2}\lambda_\alpha+1}\prod_{i=1}^a(1+n_i\tau)^{1/2}}. \end{aligned} \qquad (11.5.13)$$

Now, from (11.5.8) and (11.5.13) and the fact that

$$\sum_{i=1}^a \frac{n_i(\bar{y}_{i.}-\mu)^2}{1+n_i\tau} = \sum_{i=1}^a \frac{n_i(\bar{y}_{i.}-\hat{\mu})^2}{1+n_i\tau} + (\mu-\hat{\mu})^2\sum_{i=1}^a \frac{n_i}{1+n_i\tau},$$

the conditional posterior density of μ given τ is

$$\begin{aligned} p(\mu|\tau|\,\text{data}\; y_{ij}) &= p(\mu, \tau)/p(\tau|\,\text{data}\; y_{ij}) \\ &\propto \{1 + H^2(\mu - \hat{\mu})^2\}^{-\frac{1}{2}(N+\lambda_\alpha)}, \end{aligned} \tag{11.5.14}$$

where

$$H = \left\{ \frac{\sum_{i=1}^{a} \frac{n_i}{1+n_i\tau}}{\text{SS}_W + \frac{c_\alpha}{\tau} + \sum_{i=1}^{a} \frac{n_i(\bar{y}_{i.}-\hat{\mu})^2}{1+n_i\tau}} \right\}^{1/2}.$$

From (11.5.14), it follows that the conditional posterior density of $\psi = (N + \lambda_\alpha - 1)^{1/2} H(\mu - \hat{\mu})$ given τ is

$$p(\psi|\tau) \propto \{1 + \psi^2/(N + \lambda_\alpha - 1)\}^{-\frac{1}{2}(N+\lambda_\alpha)}. \tag{11.5.15}$$

The density function in (11.5.15) is the same as that of Student's t-distribution with $N + \lambda_\alpha - 1$ degrees of freedom. From (11.5.15), as Hill (1965) points out, one can obtain the unconditional posterior distribution of μ using a numerical integration or other approximation methods.

To conclude this development we note, as Hill (1965) remarked, that the unbalanced model in (11.1.1) presents "only more complexity in the form of the posterior distributions and no fundamental difficulties."

Reference priors such as the ones considered by Hill represent minimal prior information and allow the user to specify which parameters are of interest and which ones are considered as nuisance parameters. Berger and Bernardo (1992) considered several configurations of interest-nuisance parameters for reference priors in a variance components problem. More recently, Belzile and Angers (1995) have considered several noninformative priors for the model in (11.1.1) and derived their posterior distributions.

11.6 DISTRIBUTION AND SAMPLING VARIANCES OF ESTIMATORS

In this section, we briefly describe some results on distribution and sampling variances of estimators of variance components σ_e^2 and σ_α^2.

11.6.1 DISTRIBUTION OF THE ESTIMATOR OF σ_e^2

From the distribution law in (11.3.3), the distribution of the ANOVA estimator of σ_e^2 is

$$\hat{\sigma}^2_{e,\text{ANOV}} = \text{MS}_W \sim \left(\frac{\sigma_e^2}{N - a}\right) \chi^2[N - a]. \tag{11.6.1}$$

Result (11.6.1) is rather an exception to the otherwise complicated distribution theory of the other variance component estimators.

11.6.2 DISTRIBUTION OF THE ESTIMATORS OF σ_α^2

The distribution of the quadratic estimators of σ_α^2 for the unbalanced model in (11.1.1) is much more complicated than in the balanced case. Press (1966) has shown that the probability density function for any linear combination of independent noncentral chi-square variables can be expressed as a mixture of those density functions obtained as a linear difference of two independent chi-squares. It is also known that any quadratic form in a random vector with nondegenrate multivariate normal distribution is distributed as a linear combination of independent noncentral chi-square variables. Thus, in a given quadratic form, if one can obtain the appropriate linear combination, one can apply Press' results, or similar results of Robinson (1965), to determine an expression for its probability density function.

Harville (1969b) has considered the problem of determining the linear combination having a distribution identical to that of a given quadratic form. He has shown that for μ-invariant quadratic estimators, this problem can be reduced to that of finding eigenvalues of a matrix whose elements are functions of σ_e^2 and σ_α^2, of p, of $\eta_1, \ldots, \eta_p, \upsilon_1, \ldots, \upsilon_q$, and of the coefficient matrix of the quadratic form. (See Section 11.3 for the definition of p, q, υ_is, and η_is.) In the case of a non-μ-invariant estimator, one must also find a set of eigenvectors for that matrix that satisfy certain orthogonality properties. Further, if one can determine the probability density function for a quadratic estimator, we can also obtain the density function for the corresponding truncated estimator.

Wang (1967) has proposed the following approximation to the probability density function of the ANOVA estimator of σ_α^2 given by (11.4.1), i.e.,

$$\hat{\sigma}_{\alpha,\text{ANOV}}^2 = \frac{1}{n_0}(\text{MS}_B - \text{MS}_W). \tag{11.6.2}$$

From the results in (11.3.4) and (11.3.5), it can be seen that MS_B is the weighted sum of independent chi-square variables. Wang proposed that MS_B be approximated by multiples of a chi-square variable employing the Satterthwaite procedure (see Appendix F), i.e., by equating the first two moments of SS_B to that of the proposed chi-square distribution to determine the multiplicity constant and the degrees of freedom. Since MS_W is distributed as a multiple of a chi-square distribution and MS_B and MS_W are independent, the approximate probability density of (11.6.2) can be determined using the results of Section 2.6.2.

Searle (1967) compared the probability density functions for the $\hat{\sigma}_{\alpha,\text{ANOV}}^2$ yielded by Wang's approximation with those obtained by the Monte Carlo simulation. For the values of n_is, σ_α^2, and σ_e^2 included in the study, he found a good agreement between simulation and approximate results. One of the difficulties in using Monte Carlo simulation results, as pointed out by Harville (1969a), is that it is rather impossible to be complete. There are literally infinite numbers of different unbalanced patterns of n_is that should be considered. In this context, Searle (1967) has emphasized the need for a measure of unbalancedness that was particularly suited to the problem of characterizing its effect on the estimation of σ_α^2. Unfortunately, as Harville (1969a) has noted, the existence of such a measure of unbalancedness appears somewhat doubtful because of the manner in which the distribution of an estimator of σ_α^2 depends on n_is which in turn is dependent, in a rather complex way, on the values of the variance components themselves. Furthermore, these dependent relations are not the same for different estimators, i.e., changes in the patterns of n_is affect different estimators of σ_α^2 in different ways. For some further discussion of measures of unbalancedness, see Khuri (1987, 1996) and Ahrens and Sánchez (1988, 1992).

11.6.3 SAMPLING VARIANCES OF ESTIMATORS

In this section, we present some results on sampling variances of the estimators of variance components σ_e^2 and σ_α^2.

11.6.3.1 Sampling Variances of the ANOVA Estimators

The result on sampling variance of the ANOVA estimator of σ_e^2 follows immediately from the distribution law in (11.6.1) and is given by

$$\mathrm{Var}(\hat{\sigma}_{e,\mathrm{ANOV}}^2) = \frac{2\sigma_e^4}{N-a}. \tag{11.6.3}$$

As noted in the preceding section, the sampling distribution of the ANOVA estimator of σ_α^2 is largely unknown except in cases when the estimators can be expressed as linear functions of independent chi-square variables. However, some advances have been made in deriving sampling variances of the estimator, although the results are much more complicated than with the balanced data.

For any arbitrary distribution of the observations from the model in (11.1.1), the only available results on sampling variances are those of Hammersley (1949) and Tukey (1957). Hammersley considered infinite populations and Tukey, using polykays, treated both infinite and finite populations cases. Tukey (1957) derived sampling variances and covariances for a large class of quadratic unbiased estimators, including ANOVA estimators of σ_α^2 and σ_e^2. In addition, he compared the sampling variances for several estimators of σ_α^2 for different types of nonnormality and sample sizes n_is. In a subsequent work, Arvesen (1976) verified some of the results of Tukey (1957) using the technique of U-statistics.

Under the assumption of normality, Crump (1951) and Searle (1956) derived the expressions for the variance of $\hat{\sigma}_{\alpha,\mathrm{ANOV}}^2$. Searle used a matrix method to arrive at his result. However, both results contain certain typographical errors. Searle (1971a) gives the corrected versions for both the Crump (1951) and Searle (1956) results. Crump's corrected result is

$$\mathrm{Var}(\hat{\sigma}_{\alpha,\mathrm{ANOV}}^2) = \frac{2\sigma_e^4}{n_0^2}\left[\frac{1}{(a-1)^2}\left\{\left(\frac{1}{N}\sum_{i=1}^{a}\frac{n_i^2}{w_i^2}\right)^2\right.\right.$$

$$+\sum_{i=1}^{a}\frac{n_i^2}{w_i^2}-\frac{2}{N}\sum_{i=1}^{a}\frac{n_i^3}{w_i^3}\Bigg\}+\frac{1}{N-a}\Bigg], \qquad (11.6.4)$$

where

$$w_i = n_i\sigma_e^2/(\sigma_e^2 + n_i\sigma_\alpha^2).$$

The original form of (11.6.4) omits $1/N$ from inside the first term within the braces and contains w_i in place of w_i^2 in the same term.

The corrected version of Searle's result is

$$\begin{aligned}\text{Var}(\hat{\sigma}^2_{\alpha,\text{ANOV}}) &= \frac{2\sigma_e^4}{n_0^2}\left[\frac{1}{N(a-1)^2}\left\{\frac{\tau^2}{N}\left(N^2\sum_{i=1}^{a}n_i^2+\left(\sum_{i=1}^{a}n_i^2\right)^2-2N\sum_{i=1}^{a}n_i^3\right)\right.\right.\\ &\quad\left.\left.+2\tau\left(N^2-\sum_{i=1}^{a}n_i^2\right)\right\}+\frac{N-1}{(a-1)(N-a)}\right], \qquad (11.6.5)\end{aligned}$$

where

$$\tau = \sigma_\alpha^2/\sigma_e^2.$$

The published version of (11.6.5) has 1 instead of 2 in the middle term within the braces. An alternative form of (11.6.5) given in Searle (1971a, 1971b, p. 474) is

$$\begin{aligned}\text{Var}(\hat{\sigma}^2_{\alpha,\text{ANOV}}) = 2\sigma_e^4&\left[\frac{\tau^2\left\{N^2\sum_{i=1}^{a}n_i^2+\left(\sum_{i=1}^{a}n_i^2\right)^2-2N\sum_{i=1}^{a}n_i^3\right\}}{(N^2-\sum_{i=1}^{a}n_i^2)^2}\right.\\ &\left.+\frac{2\tau N}{N^2-\sum_{i=1}^{a}n_i^2}+\frac{N^2(N-1)(a-1)}{\left(N^2-\sum_{i=1}^{a}n_i^2\right)^2(N-a)}\right]. \qquad (11.6.6)\end{aligned}$$

For an alternate version of this result, see Rao (1997, p. 20).

Remarks:

(i) Singh (1989b) develops formulas for higher-order moments and cumulants of the sampling distribution of $\hat{\sigma}^2_{\alpha,\text{ANOV}}$.

(ii) It is seen from (11.6.3) that for a given value of a and N, $\text{Var}(\hat{\sigma}^2_{e,\text{ANOV}})$ is unaffected by the degree of unbalancedness in the data. However, the behavior of $\text{Var}(\hat{\sigma}^2_{\alpha,\text{ANOV}})$ with respect to changes in the n_i-values is much more complicated. Singh (1992) carried out a numerical study for different configurations of the a priori values of σ_α^2, σ_e^2, and n_is and found that for a given value of a and N, imbalance results in an increase in the variance. Similar results were reached by Caro et al. (1985) who studied the effects of unbalancedness on estimation of heritability. Khuri et al. (1998, pp. 56–57) prove a theorem which states that for a given N, $\text{Var}(\hat{\sigma}^2_{\alpha,\text{ANOV}})$ attains its minimum for all values of σ_α^2 and σ_e^2 if and only if $n_i = n$.

(iii) The variance of the estimator of σ_α^2 based on the unweighted mean square as defined in Remark (i) of Section 11.4.1 is given by

$$\text{Var}(\hat{\sigma}^2_{\alpha,\text{UNW}}) = 2\left[\frac{\sum_{i=1}^{a}(\sigma_e^2 + n_i\sigma_\alpha^2)^2/n_i^2}{a(a-1)} + \frac{\sigma_e^4}{(N-a)\bar{n}_h^2}\right].$$

It should be noted that $\text{Var}(\hat{\sigma}^2_{\alpha,\text{ANOV}})$ and $\text{Var}(\hat{\sigma}^2_{\alpha,\text{UNW}})$ will be close to each other if n_is do not differ greatly from each other. However, the relative magnitude of these variances in general depends on $\tau = \sigma_\alpha^2/\sigma_e^2$.

(iv) Koch (1967a, 1968) developed formulas for variances and covariances of symmetric sums estimators of σ_α^2 and σ_e^2 given in (11.4.3) and (11.4.7). ♦

11.6.3.2 Sampling Variances of the ML Estimators

In Section 11.4.6, we have seen that the ML estimators of σ_e^2 and σ_α^2 cannot be obtained explicitly. The exact sampling variances of the ML estimators also cannot be obtained. However, the expressions for large sample variances have been derived by Crump (1946, 1947) and Searle (1956). Searle used a matrix method to derive these results. Crump's results, as given in Crump (1951), are

$$\text{Var}(\hat{\sigma}^2_{e,\text{ML}}) = \frac{2\sigma_e^4}{(N-a)\left[1 + \frac{a}{N-a}\left\{1 - \left(\sum_{i=1}^{a} w_i\right)^2\left(a\sum_{i=1}^{a} w_i^2\right)^{-1}\right\}\right]} \tag{11.6.7}$$

and

$$\text{Var}(\hat{\sigma}^2_{\alpha,\text{ML}}) = \frac{2\sigma_e^4\left[(N-a) + \sum_{i=1}^{a}\left(\frac{w_i}{n_i}\right)^2\right]}{N\sum_{i=1}^{a} w_i^2 - \left(\sum_{i=1}^{a} w_i\right)^2} \times \frac{\left\{\frac{1}{N}\sum_{i=1}^{a}\frac{n_i^2}{w_i^2}\right\}^2 + \sum_{i=1}^{a}\frac{n_i^2}{w_i^2} - \frac{2}{N}\sum_{i=1}^{a}\frac{n_i^3}{w_i^3}}{\left\{\frac{1}{N}\sum_{i=1}^{a}\frac{n_i^2}{w_i^2}\right\}^2 + \sum_{i=1}^{a}\frac{n_i^2}{w_i^2} - \frac{2}{N}\sum_{i=1}^{a}\frac{n_i^3}{w_i^3} + \frac{(a-1)^2}{N-a}}, \tag{11.6.8}$$

where

$$w_i = n_i\sigma_e^2/(\sigma_e^2 + n_i\sigma_\alpha^2).$$

The Searle results are

$$\operatorname{Var}(\hat{\sigma}^2_{e,\mathrm{ML}}) = \frac{2\sigma_e^4 \sum_{i=1}^a w_i^2}{N\sum_{i=1}^a w_i^2 - \left(\sum_{i=1}^a w_i\right)^2}, \tag{11.6.9}$$

$$\operatorname{Var}(\hat{\sigma}^2_{\alpha,\mathrm{ML}}) = \frac{2\sigma_e^4 \left[N - a + \sum_{i=1}^a w_i^2/n_i^2\right]}{N\sum_{i=1}^a w_i^2 - \left(\sum_{i=1}^a w_i\right)^2}, \tag{11.6.10}$$

and

$$\operatorname{Cov}(\hat{\sigma}^2_{\alpha,\mathrm{ML}}, \hat{\sigma}^2_{e,\mathrm{ML}}) = \frac{-2\sigma_e^4 \sum_{i=1}^a w_i^2/n_i}{N\sum_{i=1}^a w_i^2 - \left(\sum_{i=1}^a w_i\right)^2}. \tag{11.6.11}$$

The results in (11.6.9) through (11.6.11) are readily derived from the matrix obtained as inverse of the matrix whose elements are negatives of the expected value of the matrix of second-order partial derivatives of $\ell n(L)$ in (11.4.14) with respect to μ, σ_e^2, and σ_α^2. They can also be obtained as special cases of the general results in (10.7.22).

11.6.3.3 Sampling Variances of the BQUE

The sampling variances of BQUEs of σ_e^2 and σ_α^2 defined in (11.4.23) are (Towsend and Searle, 1971)

$$\operatorname{Var}(\hat{\sigma}^2_{e,\mathrm{BQUE}}) = \frac{2s\sigma_e^4}{rs - t^2}$$

and (11.6.12)

$$\operatorname{Var}(\hat{\sigma}^2_{\alpha,\mathrm{BQUE}}) = \frac{2r\sigma_e^4}{rs - t^2},$$

where

$$r = \sum_{i=1}^a \frac{1}{(1+n_i\tau)^2} + N - a, \quad s = \sum_{i=1}^a \frac{n_i^2}{(1+n_i\tau)^2}, \quad \text{and} \quad t = \sum_{i=1}^a \frac{n_i}{(1+n_i\tau)^2},$$

with

$$\tau = \sigma_\alpha^2/\sigma_e^2.$$

11.6.3.4 Sampling Variances of the MIVQUE and MINQUE

Sampling variances of MIVQUEs and MINQUEs of σ_e^2 and σ_α^2 defined in (11.4.32) are (Swallow, 1974, 1981)

$$\begin{aligned}\operatorname{Var}(\hat{\sigma}^2_{e,\mathrm{MIV(N)Q}}) &= \frac{1}{|\boldsymbol{S}|^2}[s_{12}^2 \operatorname{Var}(\gamma_1) + s_{11}^2 \operatorname{Var}(\gamma_2) \\ &\quad - 2s_{11}s_{12}\operatorname{Cov}(\gamma_1, \gamma_2)], \\ \operatorname{Var}(\hat{\sigma}^2_{\alpha,\mathrm{MIV(N)Q}}) &= \frac{1}{|\boldsymbol{S}|^2}[s_{22}^2 \operatorname{Var}(\gamma_1) + s_{12}^2 \operatorname{Var}(\gamma_2) \\ &\quad - 2s_{12}s_{22}\operatorname{Cov}(\gamma_1, \gamma_2)], \end{aligned} \tag{11.6.13}$$

and

$$\begin{aligned}\operatorname{Cov}(\hat{\sigma}^2_{e,\mathrm{MIV(N)Q}}, \hat{\sigma}^2_{\alpha,\mathrm{MIV(N)Q}}) &= \frac{1}{|\boldsymbol{S}|^2}[-s_{12}s_{22} \operatorname{Var}(\gamma_1) - s_{11}s_{12} \operatorname{Var}(\gamma_2) \\ &\quad + (s_{11}s_{22} + s_{12}^2)\operatorname{Cov}(\gamma_1, \gamma_2)],\end{aligned}$$

where $|\boldsymbol{S}| = |s_{11}s_{22} - s_{12}^2|$ and s_{ij}s and γ_is are defined in (11.4.29) and (11.4.30).

As we have seen, γ_is are quadratic forms in the observation vector. Their variances and covariances under normality can therefore be obtained by familiar results on variances and covariances of quadratic forms as stated in Theorem 9.3.1. The results have been obtained by Swallow (1974) and Swallow and Searle (1978). For MIVQUEs, we have

$$\begin{aligned}\operatorname{Var}(\gamma_1) &= 2\left[\sum_{i=1}^a k_i^2 - 2k\sum_{i=1}^a k_i^3 + k^2\left(\sum_{i=1}^a k_i^2\right)^2\right], \\ \operatorname{Var}(\gamma_2) &= 2\left[\frac{N-a}{\sigma_e^4} + \sum_{i=1}^a \frac{k_i^2}{n_i^2} - 2k\sum_{i=1}^a \frac{k_i^3}{n_i^2} + k^2\left(\sum_{i=1}^a \frac{k_i^2}{n_i}\right)^2\right],\end{aligned} \tag{11.6.14}$$

and

$$\operatorname{Cov}(\gamma_1, \gamma_2) = 2\left[\sum_{i=1}^a \frac{k_i^2}{n_i} - 2k\sum_{i=1}^a \frac{k_i^3}{n_i} + k^2\left(\sum_{i=1}^a k_i^2\right)\left(\sum_{i=1}^a \frac{k_i^2}{n_i}\right)\right],$$

where

$$k_i = n_i/(\sigma_e^2 + n_i\sigma_\alpha^2) \quad \text{and} \quad k = 1/\sum_{i=1}^a k_i.$$

For MINQUEs, we have

$$\begin{aligned}\text{Var}(\gamma_1) = 2\Bigg[&\sum_{i=1}^{a}\frac{\theta_i^4}{k_i^2} + 2\theta^2\sum_{i=1}^{a}\theta_i^2\sum_{i=1}^{a}\frac{\theta_i^4}{k_i^2} - 4\theta\sum_{i=1}^{a}\frac{\theta_i^5}{k_i^2}\\ &+ \theta^4\left(\sum_{i=1}^{a}\theta_i^2\right)^2\left(\sum_{i=1}^{a}\frac{\theta_i^2}{k_i}\right)^2 + 2\theta^2\sum_{i=1}^{a}\frac{\theta_i^2}{k_i}\sum_{i=1}^{a}\frac{\theta_i^4}{k_i}\\ &- 4\theta^3\sum_{i=1}^{a}\theta_i^2\sum_{i=1}^{a}\frac{\theta_i^2}{k_i}\sum_{i=1}^{a}\frac{\theta_i^3}{k_i} + 2\theta^2\left(\sum_{i=1}^{a}\frac{\theta_i^3}{k_i}\right)^2\Bigg],\\ \text{Var}(\gamma_2) = 2\Bigg[&-\frac{1}{2}\text{Var}(\gamma_1) - \text{Cov}(\gamma_1,\gamma_2) + N\sigma_e^4\\ &+ 2N\sigma_e^2\sigma_\alpha^2 + \sum_{i=1}^{a}n_i\sigma_\alpha^4 - \sum_{i=1}^{a}\frac{n_i^2}{k_i^2}\\ &+ \sum_{i=1}^{a}\frac{\theta_i^2}{k_i^2} - 2\theta\sum_{i=1}^{a}\frac{\theta_i^3}{k_i^2} + \theta^2\left(\sum_{i=1}^{a}\frac{\theta_i^2}{k_i}\right)^2\Bigg], \qquad (11.6.15)\end{aligned}$$

and

$$\begin{aligned}\text{Cov}(\gamma_1,\gamma_2) = 2\Bigg[&-\frac{1}{2}\text{Var}(\gamma_1) + \sum_{i=1}^{a}\frac{\theta_i^3}{k_i^2} - 3\theta\sum_{i=1}^{a}\frac{\theta_i^4}{k_i^2} + \theta^2\sum_{i=1}^{a}\theta_i^2\sum_{i=1}^{a}\frac{\theta_i^3}{k_i^2}\\ &- \theta^3\sum_{i=1}^{a}\theta_i^2\left(\sum_{i=1}^{a}\frac{\theta_i^2}{k_i}\right)^2 + 2\theta^2\sum_{i=1}^{a}\frac{\theta_i^2}{k_i}\sum_{i=1}^{a}\frac{\theta_i^3}{k_i}\Bigg],\end{aligned}$$

where

$$\theta_i = \frac{n_i}{1+n_i}, \qquad k_i = \frac{n_i}{\sigma_e^2 + n_i\sigma_\alpha^2}, \quad \text{and} \quad \theta = 1/\sum_{i=1}^{a}\theta_i.$$

Sampling variances for MINQUEs and MIVQUEs of σ_e^2 and σ_α^2 are also developed in the papers by Ahrens (1978), Swallow (1981), and Sánchez (1983).

11.6.3.5 Sampling Variance of an Unbiased Estimator of $\sigma_\alpha^2/\sigma_e^2$

An unbiased estimator of $\tau = \sigma_\alpha^2/\sigma_e^2$, under normality, was given in (11.4.35). The sampling variance of the estimator is given by

$$\begin{aligned}\text{Var}(\hat{\tau}_{\text{UNB}}) = \frac{2}{N-a-4}\Bigg[&\{(N-a-2)A+1\}\tau^2 + (N-3)B\tau\\ &+ \frac{(N-3)(N-a)}{N-1}C\Bigg], \qquad (11.6.16)\end{aligned}$$

where

$$A = \frac{N^2 S_2 - 2NS_3 + S_2^2}{N^2 - S_2^2}, \qquad B = \frac{2N}{N^2 - S_2},$$

and

$$C = \frac{N^2(N-1)(a-1)}{(N-a)(N^2 - S_2)^2}$$

with

$$N = \sum_{i=1}^{a} n_i, \qquad S_2 = \sum_{i=1}^{a} n_i^2, \quad \text{and} \quad S_3 = \sum_{i=1}^{a} n_i^3.$$

The result is given in Anderson and Crump (1967).

11.6.3.6 Sampling Variances of the ANOVA and ML Estimators of $\sigma_\alpha^2/(\sigma_e^2 + \sigma_\alpha^2)$

A large sample variance of the analysis of variance estimator of ρ in (11.4.36) was derived by Smith (1956), under the assumption of normality, and is given by

$$\begin{aligned}\text{Var}(\hat{\rho}_{\text{ANOV}}) = \frac{2(1-\rho)^2}{n_0^2}\Bigg\{&\frac{[1+\rho(n_0-1)]^2}{N-a} \\ &+ \frac{(a-1)(1-\rho)[1+\rho(2n_0-1)] + \rho^2[S_2 - 2N^{-1}S_3 + N^{-2}S_2^2]}{(a-1)^2}\Bigg\},\end{aligned} \tag{11.6.17}$$

where

$$N = \sum_{i=1}^{a} n_i, \qquad S_2 = \sum_{i=1}^{a} n_i^2, \qquad S_3 = \sum_{i=1}^{a} n_i^3,$$

and

$$n_0 = (N^2 - S_2)/N(a-1).$$

A simpler expression for the large sample variance, applicable when the variation in the group sizes is small, has been derived by Swiger et al. (1964) and is given by

$$\text{Var}(\hat{\rho}_{\text{ANOV}}) \doteq \frac{2(N-1)(1-\rho)^2[1+(n_0-1)\rho]^2}{n_0^2(N-a)(a-1)}. \tag{11.6.18}$$

Some simulation work by the authors show that the variance expression (11.6.18) is quite accurate if $\rho \leq 0.1$.

The large sample variance of the ML estimator of ρ has been derived by Donner and Koval (1980b) and is given by

$$\text{Var}(\hat{\rho}_{\text{ML}}) = \frac{2N(1-\rho)^2}{N\sum_{i=1}^{a} n_i(n_i-1)u_i^{-2}v_i - \rho^2\left[\sum_{i=1}^{a} n_i(n_i-1)u_i^{-1}\right]^2},$$

where

$$u_i = 1 + (n_i - 1)\rho \quad \text{and} \quad v_i = 1 + (n_i - 1)\rho^2.$$

Some calculations on the large sample relative efficiency of the ANOVA estimator of ρ compared to the ML show that it is very similar to the large sample relative efficiency of the ANOVA estimator of σ_α^2 for all values of ρ; although the former tends to be slightly higher for values of $\rho \geq 0.3$ (Donner and Koval, 1983).

11.7 COMPARISONS OF DESIGNS AND ESTIMATORS

It seems Hammersley (1949) was the first to consider the problem of optimal designs for the model in (11.1.1). Hammersley (1949) showed that for a fixed N, $\text{Var}(\hat{\sigma}^2_{\alpha,\text{ANOV}})$ is minimized by allocating an equal number, n, of observations to each class where $n = (N\tau + N + 1)/(N\tau + 2)$. Since this formula may not yield an integer value, it was suggested that the closest integer value be chosen for n. Subsequently, Crump (1954) and Anderson and Crump (1967) compared various designs and estimators using the usual ANOVA method for estimating μ and σ_α^2. They, however, concentrated on the problem of optimal allocation of a fixed total sample of N to different classes in order to estimate more efficiently σ_α^2 or $\sigma_\alpha^2/\sigma_e^2$. They proved that for fixed a and N, $\text{Var}(\hat{\sigma}^2_{\alpha,\text{ANOV}})$ would be minimized if $N = (q-1)a_{q-1} + qa_q$ where $a_{q-1} + a_q = a, n_i = q - 1$ for $i = 1, 2, \ldots, a_{q-1}$ and $n_i = q$ for $i = a_{q-1} + 1, \ldots, a_{q-1} + a_q$. Thus there should be $p + 1$ elements in each of r groups and p elements in each of $a - r$ groups, where $N = ap + r, 0 \leq r < a$. The optimal number of groups (a) was found to be the integer closest to

$$\begin{aligned} a' &= N(N\tau + 2)/(N\tau + N + 1) && \text{for } \sigma_\alpha^2, \\ a'' &= 1 + [(N-5)(N\tau + 1)/(2N\tau + N - 3)] && \text{for } \tau. \end{aligned}$$

For large N, a', and a'' are approximately given as follows. For estimating σ_α^2, $a' = N\tau/(1+\tau)$ with an average of $1 + 1/\tau$ observations per group; and for $\tau, a'' = N\tau/(1+2\tau)$ with an average of $2 + 1/\tau$ observations per group. Note that $a'/a'' = (1+2\tau)/(1+\tau)$, indicating higher number of groups to estimate σ_α^2 than to estimate τ. Thus the determination of the optimal design depends on τ, the ratio of variance components themselves. They also compared the ANOVA method of estimation of σ_α^2 with the unweighted means method and found that for extremely unbalanced data the unweighted means estimator appears to be poorer (has large variances) than the ANOVA estimator for small values of $\tau = \sigma_\alpha^2/\sigma_e^2$, but that it is superior (has smaller variance) for large τ.

Kussmaul and Anderson (1967) considered a special form of the one-way classification in which the compositing of samples in a two-way nested classification is envisaged. As a result, the jth observation in the ith class is thought to be an average of the n_{ij} observations that the sample would have provided separately had it not been composited prior to measurement. In this situation, the ANOVA method of estimating the between class variance component is compared numerically with two other unbiased estimation procedures for a variety of n_{ij} values and for various values of the ratio of the variance components. It may be advantageous to composite some of the samples and take measurements on the composited samples in cases where the measurement cost is high; for example, in sampling for many bulk materials, especially those requiring chemical assays. The problem was first discussed by Cameron (1951) who proposed a number of compositing plans for estimating the components of variance when sampling baled wool for percent clean content. Kussmaul and Anderson also considered the problem of estimating μ where the knowledge of σ_e^2 and σ_α^2 is needed for the sole purpose of determining the best design for estimating μ (see also Anderson, 1981).

Thompson and Anderson (1975) compared certain optimal designs from Anderson and Crump (1967) using the truncated ANOVA, the maximum likelihood (ML), and the modified ML for both balanced and unbalanced situations. Exact values of mean squared error (MSE) were calculated for balanced designs to compare the optimality of different designs. The MSEs for unbalanced designs were obtained from Monte Carlo simulations. For balanced designs the modified ML estimator was found to be superior and for this estimator the optimal design was less sensitive to the intraclass correlation ρ for $\rho < 0.5$, than those designs based on minimizing the variance of the usual ANOVA estimator. For $\rho > 0.5$, where an unbalanced design is preferable, asymptotic results were obtained to indicate optimal designs for the ML and ANOVA estimators. It was found that the ML estimators have smaller MSEs than the truncated ANOVA or the iterated least squares estimators. Further, they reported that for $\rho \geq 0.2$, the optimum allocation from Anderson and Crump (1967), which was developed by minimizing $\text{Var}(\hat{\sigma}^2_{\alpha,\text{ANOV}})$, is also quite good for the ML estimators. Large sample results for unbalanced designs were compared with small sample results obtained by simulation for a wide range of values of the intraclass correlation and for several selected designs.

Mukerjee and Huda (1988) established optimality of the balanced design under an unweighted means method for a multifactor random effects model of which the model (11.1.1) is a special case. A similar conclusion regarding optimality of the balanced design was arrived at by Mukerjee (1990) in the context of MINQUE estimation. Optimality of the balanced design under ML estimation was similarly assessed by Giovagnoli and Sebastiani (1989), who also investigated optimality of the design for a general mixed model involving estimation of the variance components as well as the model's fixed effects.

Townsend and Searle (1971) made numerical comparisons of the BQUEs with the ANOVA estimators by using a range of values of τ, both for actual BQUEs (assuming τ known) and for approximate BQUEs (using a prior estimate or guess τ_0 of τ). The conclusions are that in certain situations considerable reduction in the variance of estimates of σ_α^2 can be achieved if the approximate BQUE is used rather than the ANOVA estimator. Furthermore, this reduction occurs even when rather inaccurate prior estimates of τ are used. The reduction in variance appears to be greatest when the data are severely unbalanced and τ is either small or large; and it appears smallest for values of τ that are moderately small. In some cases when the ANOVA is a BQUE for some specific τ_0, there is actually no reduction in variance.

Swallow (1974) and Swallow and Searle (1978) compared numerically the variances of the ANOVA, MINQUE and MIVQUE estimators assuming normality for the random effects. The large sample variances of the ML estimators were not included because those estimators are biased and many of the design patterns included in the study involved rather small samples. They found that for each of the estimators under consideration, the variance depends on $(\sigma_\alpha^2, \sigma_e^2)$ only through σ_e^2 (a function τ), where $\tau = \sigma_\alpha^2/\sigma_e^2$. Therefore, the variance components values included in the study involved different values of τ with $\sigma_e^2 = 1$. The comparisons were made for $\tau = \sigma_\alpha^2 = 1/2, 1, 2, 3, 5, 10,$ and 20; and for various design patterns, which intuitively ranged from balanced to severely unbalanced. In evaluating the results of these comparisons, it should be noted that the MIVQUE variances are lower bounds for the variances of quadratic unbiased estimators. The MINQUEs are MIVQUEs for balanced data or when $\tau = 1$, otherwise not. The ANOVA estimators are MIVQUEs only for balanced data. The conclusions of the comparisons can be summarized as follows:

(i) The MINQUE and MIVQUE variances, in nearly all cases, are reasonably similar. The ratio of the MINQUE variance to the MIVQUE variance increases with τ and under severe unbalancedness; but for the cases with fixed a and N, the MINQUE and MIVQUE of σ_α^2 are much less dependent than the ANOVA estimator on the n_is.

(ii) The ANOVA estimator of σ_α^2 may have a very large variance when most of the n_is are equal to one, especially if $\tau \gg 1$. The MINQUE of σ_e^2 is again much worse than the MIVQUE or ANOVA estimator when $\tau \ll 1$.

Hess (1979) made numerical comparisons of the variance efficiency of the MINQUE and ANOVA estimators in order to investigate the sensitivity of the MINQUEs of σ_α^2 and σ_e^2 to their prior weights w_α and w_e. It was found that for w_α/w_e in a neighborhood of $\sigma_\alpha^2/\sigma_e^2$, the variance of $\hat{\sigma}^2_{\alpha,\text{MINQUE}}$ is quite stable.

Further, for designs with moderate imbalance, if the ratio $\sigma_\alpha^2/\sigma_e^2 < 1$, then $\hat{\sigma}^2_{\alpha,\text{ANOV}}$ is more efficient; for $\sigma_\alpha^2/\sigma_e^2 > 1$, $\hat{\sigma}^2_{\alpha,\text{MINQUE}}$ has superior performance. However, regardless of the magnitude of $\sigma_\alpha^2/\sigma_e^2$, $\hat{\sigma}^2_{\alpha,\text{ANOV}}$ is preferred since $\hat{\sigma}^2_{\alpha,\text{MINQUE}}$ is very sensitive for $w_\alpha/w_e < \sigma_\alpha^2/\sigma_e^2$ and offers little improvement for $w_\alpha/w_e \geq \sigma_\alpha^2/\sigma_e^2$. Similarly, Swallow (1981) made numerical comparisons of the ANOVA, MIVQUEs, and "MIVQUEs," where "MIVQUEs" designate MIVQUEs obtained by replacing σ_α^2 and σ_e^2 by $\sigma_{\alpha_0}^2$ and $\sigma_{e_0}^2$ as prior estimates. The results show that when $\sigma_\alpha^2/\sigma_e^2 > 1$ (and unless $\sigma_{\alpha_0}^2/\sigma_{e_0}^2 \ll \sigma_\alpha^2/\sigma_e^2$): (a) The "MIVQUEs" have variances near their lower bounds and (b) $\hat{\sigma}^2_{\alpha,\text{“MIVQUE”}}$ is more efficient than $\hat{\sigma}^2_{\alpha,\text{ANOV}}$. When $\sigma_\alpha^2/\sigma_e^2 \leq 1$, the "MIVQUEs" are more dependent on accurate choice of $\sigma_{\alpha_0}^2/\sigma_{e_0}^2$. Further, $\hat{\sigma}^2_{e,\text{“MIVQUE”}}$ and $\hat{\sigma}^2_{e,\text{ANOV}}$ have nearly equal variances unless $\sigma_{\alpha_0}^2/\sigma_{e_0}^2 \ll \sigma_\alpha^2/\sigma_e^2$ when $\hat{\sigma}^2_{e,\text{ANOV}}$ has smaller variance.

Swallow and Monaham (1984) compared five estimators of σ_α^2 and σ_e^2 in terms of biases and efficiencies for $\sigma_e^2 = 1$; $\sigma_\alpha^2 = 0.1, 0.2, 0.5, 1.0, 2.0, 5.0$; and for 13 different design patterns. The estimators being compared are the ANOVA, ML, REML, MIVQUE(0), and MIVQUE with ANOVA estimates as a prioris, called MIVQUE(A). The results indicate that for estimating σ_α^2, (i) the ANOVA estimates are reasonably efficient except for the cases when $\sigma_\alpha^2/\sigma_e^2 > 1$ and the design is severely unbalanced; (ii) the ML has superior performance (low MSE and small bias) when $\sigma_\alpha^2/\sigma_e^2 < 0.5$ but has large bias when $\sigma_\alpha^2/\sigma_e^2 \geq 1.5$; (iii) the MIVQUE(0) as expected perfoms well when $\sigma_\alpha^2 \approx 0$ but not as well as the ML estimator; when $\sigma_\alpha^2/\sigma_e^2 = 0.1$, it is no better than the ML estimator; (iv) when $\sigma_\alpha^2/\sigma_e^2 \geq 1.0$, the MIVQUE(0) has poor performance for σ_α^2 and is exceedingly bad for σ_e^2 even for mildly unbalanced designs; (v) the MIVQUE(A) and REML have similar performance and it is not improved by iterating to the REML estimates. For estimating σ_e^2, all the five estimators have negligible bias and except for the MIVQUE(0) all have comparable MSEs. Conerly and Webster (1987) compared the MSEs of the Rao–Chaubey's MINQE along with the estimators studied by Swallow and Monahan (1984) and found that the MINQE has smaller MSE than other estimators when $\sigma_\alpha^2 > \sigma_e^2$.

Westfall (1987) performed analytic as well as numerical comparisons between the ANOVA, ML, REML, MINQUE(0), MINQUE (1), MINQUE(∞), and MIVQUE and showed that the ML and REML are asymptotically equivalent to MIVQUE and have relatively good performance in nonnormal situations even for data with moderate sample sizes. Westfall (1994) also investigated the asymptotic behavior of the ANOVA and MINQUE estimators of variance components in the nonnormal random models. Khattree and Gill (1988) made a numerical comparison between the ANOVA, ML, REML, MINQUE, and MIVQUE(0) using the relative MSE and Pitman nearness criteria and they found the MIVQUE(0) to have the worst performance among all the methods being compared. Their conclusions are that the ANOVA is the preferred method for estimating σ_e^2 whereas the REML is favored for estimating σ_α^2. For simultaneous estimation of σ_α^2 and σ_e^2, they recommended the use of the ML procedure which, however, entails considerable amount of bias. Chaloner (1987) compared the ANOVA, ML, and a Bayesian estimator given by the mode of the joint posterior distribution of the variance components using a noninformative prior distribution. The simulation results indicate that the posterior modes have good sampling properties and are generally superior to other estimators in terms of the mean squared error. Rao et al. (1981) considered the model in (11.1.1) where α_is and e_is are normal with means of zero and variances σ_α^2 and σ_i^2. They compared various methods, including ANOVA, MINQUE, MIVQE, and USS, of estimating σ_α^2 and σ_i^2 in terms of biases and mean square errors for different configurations of the values σ_α^2, σ_i^2, a, and n_i. For the same model, Heine (1993) developed nonnegative minimum norm quadratic minimum biased estimators (MNQMBE) and compared them with MINQE estimators in terms of bias and MSE criteria.

Mathew et al. (1992) proposed and compared four nonnegative invariant quadratic estimators of σ_α^2 along with the ANOVA and MINQUE estimators. The results seem to indicate that the proposed estimators offer significant reduction in MSE over the ANOVA and MINQUE estimators, although they may entail a substantial amount of bias. Kelly and Mathew (1994) proposed and compared nine nonnegative estimators of σ_α^2, along with the truncated ANOVA, ML, and REML estimators, in terms of their biases and MSE efficiencies. The results of a Monte Carlo comparison seem to indicate that some of the proposed estimators provide substantial MSE improvement over the truncated ANOVA, ML, and REML estimators. Belzile and Angers (1995) have compared the posterior means of variance components based on different noninformative priors with the REML estimators using a Monte Carlo simulation study. It is found that under the squared error loss, the invariant noninformative priors lead to the optimal estimators of the variance components. More recently, Rao (2001) has compared ten estimators including six that yield nonnegative estimates and found that the MINQE adjusted for reducing bias (MINQE*) and the nonnegative minimum MSE estimator (MIMSQUE) in general have much smaller mean square error but entail a greater amount of bias.

Ahrens (1978) developed comprehensive formulas for the risk function of the MINQUE estimators and made extensive numerical studies to compare them with the ANOVA procedure. He also established the equivalence between the ANOVA and MINQUE estimators for the balanced design. In addition, Ahrens (1978) derived explicit expressions for the minimum norm quadratic (MINQ) estimators which may be biased. Sánchez (1983) developed formulas for sampling variances of several MINQ type estimators. Furthermore, Ahrens et al. (1981) made extensive MSE comparisons between the MINQUE, MINQ, ANOVA, and two alternative estimators of σ_α^2. Ahrens and Pincus (1981) proposed two measures of imbalance for the special case of the model in (11.1.1). Ahrens and Sánchez (1982, 1986, 1992) also studied measures of unbalancedness and investigated the relative efficiencies of the

ANOVA and MINQUE estimators as functions of the measures of unbalancedness (see also Ahrens and Sánchez, 1988; Singh, 1992; Lera, 1994). Lee and Khuri (1999) developed graphical techniques involving plots of the so-called quantile dispersion graphs based on ANOVA and ML estimation of the variance components. The quantiles are functions of the unknown variance components and are assessed by computing their maxima and minima over some specified parameter space. Their plots provide a comprehensive picture of the quality of estimation for a given design and a given estimator. The results are extended to the two-way random model without interaction by Lee and Khuri (2000).

11.8 CONFIDENCE INTERVALS

In this section, we consider some results on confidence intervals for the variance components σ_e^2 and σ_α^2 and certain of their parametric functions such as the ratio $\sigma_\alpha^2/\sigma_e^2$ and the intraclass correlation $\sigma_\alpha^2/(\sigma_e^2 + \sigma_\alpha^2)$.

11.8.1 CONFIDENCE INTERVAL FOR σ_e^2

As in Section 2.8.1, using the distribution law in (11.3.3), an exact $100(1-\alpha)\%$ confidence interval for σ_e^2 is given by

$$P\left\{\frac{\mathrm{SS}_E}{\chi^2[N-a, 1-\alpha/2]} \leq \sigma_e^2 \leq \frac{\mathrm{SS}_E}{\chi^2[N-a, \alpha/2]}\right\} = 1-\alpha, \qquad (11.8.1)$$

where $\chi^2[N-a, \alpha/2]$ and $\chi^2[N-a, 1-\alpha/2]$ denote lower- and upper-tail $\alpha/2$-level critical values of the $\chi^2[N-a]$ distribution. Note that the interval in (11.8.1) is the same as given in (2.8.4) for the case of the balanced data, where the degrees of freedom ν_e is replaced by $N-a$.

11.8.2 CONFIDENCE INTERVALS FOR $\sigma_\alpha^2/\sigma_e^2$ AND $\sigma_\alpha^2/(\sigma_e^2 + \sigma_\alpha^2)$

In this section, we consider the problem of constructing confidence intervals for the variance ratio $\tau = \sigma_\alpha^2/\sigma_e^2$ and the intraclass correlation $\rho = \sigma_\alpha^2/(\sigma_e^2+\sigma_\alpha^2)$. Wald (1940) developed an exact procedure for constructing confidence intervals for τ and ρ, but the method requires the numerical solution of two nonlinear equations and is computationally somewhat difficult to carry out. Thomas and Hultquist (1978) proposed a simplified procedure based on unweighted mean squares which yields approximate confidence intervals for τ and ρ. The procedure gives satisfactory results unless $\tau < 0.25$ in which case it produces liberal intervals. In addition, several other approximate procedures have been proposed in the literature, and we will discuss them briefly here.

11.8.2.1 Wald's Procedure

We will describe the procedure for determining the limits for τ, and the limits for ρ are, of course, obtained from the limits of τ by an appropriate transformation. First, we need the following lemma.

Lemma 11.8.1. *Define*

$$H = \frac{1}{\sigma_e^2} \sum_{i=1}^{a} w_i \left(\bar{y}_{i.} - \frac{\sum_{i=1}^{a} w_i \bar{y}_{i.}}{\sum_{i=1}^{a} w_i} \right)^2,$$

where

$$w_i = n_i/(1 + n_i \tau).$$

Then H has a chi-square distribution with $a - 1$ degrees of freedom.

Proof. Let $x_i = \sqrt{w_i}\,\bar{y}_{i.}$, $i = 1, 2, \ldots, a$, and consider the orthogonal transformation

$$\begin{aligned} x'_1 &= L_1(x_1, \ldots, x_a), \\ x'_2 &= L_2(x_1, \ldots, x_a), \\ &\vdots \\ x'_{a-1} &= L_{a-1}(x_1, \ldots, x_a), \end{aligned}$$

and

$$x'_a = \frac{\sum_{i=1}^{a} \sqrt{w_i}\, x_i}{\sqrt{\sum_{i=1}^{a} w_i}},$$

where $L_i(x_1, \ldots, x_a)$ $(i = 1, \ldots, a - 1)$ denote arbitrary homogeneous linear functions subject to the only condition that the transformation should be orthogonal.

Now,

$$E(x_i) = \sqrt{w_i}\,\mu$$

and

$$\mathrm{Var}(x_i) = w_i(\sigma_\alpha^2 + \sigma_e^2/n_i) = \sigma_e^2.$$

Furthermore,

$$E(x'_i) = 0, \qquad i = 1, 2, \ldots, a - 1,$$

and

$$\mathrm{Var}(x_i') = \sigma_e^2, \quad i = 1, 2, \ldots, a.$$

It therefore follows that

$$\frac{1}{\sigma_e^2}\sum_{i=1}^{a-1} x'^2_i \sim \chi^2[a-1].$$

Thus, to prove the lemma, it suffices to show that

$$H = \frac{1}{\sigma_e^2}\sum_{i=1}^{a-1} x'^2_i.$$

Again, by definition, we have

$$\begin{aligned} H &= \frac{1}{\sigma_e^2}\sum_{i=1}^{a} w_i \left(\bar{y}_{i.} - \frac{\sum_{i=1}^a w_i \bar{y}_{i.}}{\sum_{i=1}^a w_i}\right)^2 \\ &= \frac{1}{\sigma_e^2}\left[\sum_{i=1}^{a} w_i \bar{y}_{i.}^2 - 2\frac{\left(\sum_{i=1}^a w_i \bar{y}_{i.}\right)^2}{\sum_{i=1}^a w_i} + \frac{\left(\sum_{i=1}^a w_i \bar{y}_{i.}\right)^2}{\sum_{i=1}^a w_i}\right] \\ &= \frac{1}{\sigma_e^2}\left[\sum_{i=1}^{a} w_i \bar{y}_{i.}^2 - \frac{\left(\sum_{i=1}^a w_i \bar{y}_{i.}\right)^2}{\sum_{i=1}^a w_i}\right]. \end{aligned}$$

On substituting $\bar{y}_{i.} = x_i/\sqrt{w_i}$, we get

$$\begin{aligned} H &= \frac{1}{\sigma_e^2}\left[\sum_{i=1}^{a} x_i^2 - \frac{\left(\sum_{i=1}^a \sqrt{w_i}x_i\right)^2}{\sum_{i=1}^a w_i}\right] \\ &= \frac{1}{\sigma_e^2}\left[\sum_{i=1}^{a} x_i'^2 - x_a'^2\right] \\ &= \frac{1}{\sigma_e^2}\sum_{i=1}^{a-1} x_i'^2. \end{aligned}$$

This proves the lemma. □

Now, since

$$\frac{\mathrm{SS}_W}{\sigma_e^2} = \frac{1}{\sigma_e^2}\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2 \sim \chi^2[N-a],$$

it follows that

$$F^* = \frac{(N-a)\sum_{i=1}^a w_i\left(\bar{y}_{i.} - \frac{\sum_{i=1}^a w_i \bar{y}_{i.}}{\sum_{i=1}^a w_i}\right)^2}{(a-1)\sum_{i=1}^a \sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2}$$

$$= \frac{f(\tau)}{(a-1)\mathrm{MS}_W} \sim F[a-1, N-a],$$

where

$$f(\tau) = \sum_{i=1}^{a} w_i \left(\bar{y}_{i.} - \frac{\sum_{i=1}^{a} w_i \bar{y}_{i.}}{\sum_{i=1}^{a} w_i} \right)^2.$$

Further, let F_L^* and F_U^* denote the lower and upper confidence limits of F^*. Then we shall show that the set of values of τ for which F^* lies between its confidence limits F_L^* and F_U^*, is an interval. For this purpose, it suffices to show only that $f(\tau)$ is monotonically decreasing in τ. Now, we have

$$\frac{df(\tau)}{d\tau} = \sum_{i=1}^{a} \frac{dw_i}{d\tau} \left(\bar{y}_{i.} - \frac{\sum_{i=1}^{a} w_i \bar{y}_{i.}}{\sum_{i=1}^{a} w_i} \right)^2 - 2\frac{d}{d\tau}\left(\frac{\sum_{i=1}^{a} w_i \bar{y}_{i.}}{\sum_{i=1}^{a} w_i} \right)$$
$$\times \left[\sum_{i=1}^{a} w_i \left(\bar{y}_{i.} - \frac{\sum_{i=1}^{a} w_i \bar{y}_{i.}}{\sum_{i=1}^{a} w_i} \right) \right].$$

Since

$$\sum_{i=1}^{a} w_i \left(\bar{y}_{i.} - \frac{\sum_{i=1}^{a} w_i \bar{y}_{i.}}{\sum_{i=1}^{a} w_i} \right) = 0,$$

it follows that

$$\frac{df(\tau)}{d\tau} = \sum_{i=1}^{a} \frac{dw_i}{d\tau} \left(\bar{y}_{i.} - \frac{\sum_{i=1}^{a} w_i \bar{y}_{i.}}{\sum_{i=1}^{a} w_i} \right)^2$$
$$= -\sum_{i=1}^{a} w_i^2 \left(\bar{y}_{i.} - \frac{\sum_{i=1}^{a} w_i \bar{y}_{i.}}{\sum_{i=1}^{a} w_i} \right)^2 < 0,$$

which proves our statement.

Hence, the lower confidence limit τ_L^* of τ is given by the root of the equation in τ_L,

$$f(\tau_L) = (a-1)\mathrm{MS}_W F_U^*$$

and the upper confidence limit τ_U^* of τ is given by the root of the equation in τ_U,

$$f(\tau_U) = (a-1)\mathrm{MS}_W F_L^*.$$

Since $f(\tau)$ is monotonically decreasing, the above equations have at most one root in τ_L and τ_U. If one of the equations has no root, the corresponding confidence limit has to be put equal to zero. If neither of the two equations has a root, then at least one of the assumptions of the model in (11.1.1) is violated. Furthermore, with $f(0) = \mathrm{MS}_B/\mathrm{MS}_W$ and $f(\infty) = 0$, it may be verified that there may be no roots to either or both of these equations when $f(0)$ is less than F_L^* or F_U^*. Thus the confidence limits τ_L^* and τ_U^* are

$$\tau_L^* = \begin{cases} \tau_L & \text{when } f(0) > F_U^*, \\ 0 & \text{otherwise} \end{cases} \tag{11.8.2}$$

and

$$\tau_U^* = \begin{cases} \tau_U & \text{when } f(0) > F_L^*, \\ 0 & \text{otherwise.} \end{cases} \tag{11.8.3}$$

The corresponding confidence limits for ρ, say, ρ_L and ρ_U, are given by

$$\rho_L = \frac{\tau_L^*}{1+\tau_L^*} \quad \text{and} \quad \rho_U = \frac{\tau_U^*}{1+\tau_U^*}.$$

It should be borne in mind that the equations in question are complicated algebraic equations in τ. For the actual calculation of the roots of these equations, well-known approximation methods can be employed. In applying any such approximation method it is very useful to start with two limits of the root which do not lie far apart. One of the methods of finding such limits is discussed by Wald (1940). Seely and El-Basiouni (1983) obtained Wald's limits via reductions in sums of squares for the random effects adjusted for the fixed effects in a general mixed model. They also presented necessary and sufficient conditions for the applicability of Wald's interval in a mixed model. Further computational and other issues, including its generalization to higher-order random models, are discussed in the papers by Verdooren (1976, 1988), Harville and Fenech (1985), Burdick et al. (1986), El-Bassiouni and Seely (1988), LaMotte et al. (1988), Westfall (1988, 1989), Lin and Harville (1991), and Lee and Seely (1996). For a numerical example using SAS® codes, see Burdick and Graybill (1992, Appendix B).

11.8.2.2 The Thomas–Hultquist Procedure

Define a statistic H as

$$H = \frac{(a-1)\bar{n}_h S_{\bar{y}}^2}{\sigma_e^2 + \bar{n}_h \sigma_\alpha^2}, \tag{11.8.4}$$

where

$$\bar{n}_h = a / \sum_{i=1}^{a} n_i^{-1}$$

and

$$S_{\bar{y}}^2 = \frac{1}{a-1}\left[\sum_{i=1}^{a} \bar{y}_{i.}^2 - \frac{1}{a}\left(\sum_{i=1}^{a} \bar{y}_{i.}\right)^2\right]$$

is the sample variance for the group means. The statistic H forms the basis for the construction of confidence intervals for τ and ρ.

It is readily seen that for the balanced case $\bar{n}_h S_{\bar{y}}^2 = \text{MS}_B$ with $\bar{n}_h = n$ and thus H will have an exact $\chi^2[a-1]$ distribution. For the unbalanced case, Thomas and Hultquist (1978) discuss an exact distribution of H, which is somewhat intractable. They also show empirically that the statistic H can be approximated by a $\chi^2[a-1]$ distribution, and, therefore, approximately, the statistic

$$G = \frac{H/(a-1)}{\text{SS}_W/\{\sigma_e^2(N-a)\}} \sim F[a-1, N-a]. \tag{11.8.5}$$

For the balanced case, the statistic G reduces to the customary F-statistic and thus has an exact $F[a-1, N-a]$ distribution.

The confidence intervals for τ and ρ are thus obtained by substituting $\bar{n}_h$ for n and $F' = \bar{n}_h S_{\bar{y}}^2/\text{MS}_W$ for $F^* = \text{MS}_B/\text{MS}_W$ in the corresponding formulas for the balanced case. Thus, substituting these quantities in (2.8.9) and (2.8.15), the interval for τ is obtained as

$$\left[\frac{F' - F[\nu_\alpha, \nu_e; 1-\alpha/2]}{\bar{n}_h F[\nu_\alpha, \nu_e; 1-\alpha/2]}\right] \le \tau \le \left[\frac{F' - F[\nu_\alpha, \nu_e; \alpha/2]}{\bar{n}_h F[\nu_\alpha, \nu_e; \alpha/2]}\right], \tag{11.8.6}$$

and the interval for the intraclass correlation becomes

$$\frac{F' - F[\nu_\alpha, \nu_e; 1-\alpha/2]}{F' + (\bar{n}_h - 1)F[\nu_\alpha, \nu_e; 1-\alpha/2]} \le \rho \le \frac{F' - F[\nu_\alpha, \nu_e; \alpha/2]}{F' + (\bar{n}_h - 1)F[\nu_\alpha, \nu_e; \alpha/2]}, \tag{11.8.7}$$

where $\nu_\alpha = a - 1$ and $\nu_e = N - a$.

Thomas and Hultquist also carried out some Monte Carlo studies to evaluate the goodness of the proposed intervals in terms of percentage coverage and average width for certain selected values of τ and the design parameters (a, n_i). In the case of τ and ρ intervals given by (11.8.6) and (11.8.7) the results show that the proposed formulas do give $1-\alpha$ coverage. Thus, considering the ease with which the proposed interval estimates are calculated, they may be used in preference to Wald's procedure. However, for designs with extreme imbalance and $\tau < 0.25$, its coverage can fall below the prescribed level of confidence.

11.8.2.3 The Burdick–Maqsood–Graybill Procedure

Burdick, Maqsood, and Graybill (1986) have proposed a simple noniterative interval for τ based on the unweighted mean square. The desired $100(1-\alpha)\%$

confidence interval is given by

$$P\left\{\frac{S_{\bar{y}}^2}{\mathrm{MS}_W F[\nu_\alpha, \nu_e; 1-\alpha/2]} - \frac{1}{n_{\min}} \le \tau \le \frac{S_{\bar{y}}^2}{\mathrm{MS}_W F[\nu_\alpha, \nu_e; \alpha/2]} - \frac{1}{n_{\max}}\right\} \doteq 1-\alpha, \tag{11.8.8}$$

where

$$n_{\min} = \min(n_1, n_2, \ldots, n_a) \quad \text{and} \quad n_{\max} = \max(n_1, n_2, \ldots, n_a).$$

The interval in (11.8.8) is known to be conservative and can produce a much wider interval than the exact Wald interval when $\tau < 0.25$ and the design is extremely unbalanced. The interval on ρ is obtained from (11.8.8) by using the relationship $\rho = \tau/(1+\tau)$.

11.8.2.4 The Thomas–Hultquist–Donner Procedure

Thomas and Hultquist (1978) and Donner (1979) suggested that adequate confidence intervals for τ and ρ may be obtained by using the corresponding formulas for the balanced case with the term n_0 replacing n. Making the appropriate substitutions in (2.8.9) and (2.8.15), the desired $100(1-\alpha)\%$ confidence intervals for τ and ρ are given by

$$P\left\{\frac{F^* - F[v_\alpha, v_e; 1-\alpha/2]}{n_0 F[v_\alpha, v_e; 1-\alpha/2]} \le \tau \le \frac{F^* - F[v_\alpha, v_e; \alpha/2]}{n_0 F[v_\alpha, v_e; \alpha/2]}\right\} \doteq 1-\alpha \tag{11.8.9}$$

and

$$P\left\{\frac{F^* - F[v_\alpha, v_e; 1-\alpha/2]}{F^* + (n_0-1)F[v_\alpha, v_e; 1-\alpha/2]} \le \rho \le \frac{F^* - F[v_\alpha, v_e; \alpha/2]}{F^* + (n_0-1)F[v_\alpha, v_e; \alpha/2]}\right\} \doteq 1-\alpha. \tag{11.8.10}$$

The approximation in (11.8.9) and (11.8.10) arises because the variance ratio statistic F^* under the model in (11.1.1) is not distributed according to the (central) F-distribution unless $\sigma_\alpha^2 = 0$. Thus the adequacy of the approximation presumably declines as the values of τ and ρ depart from the null value.

11.8.2.5 The Donner–Wells Procedure

Donner and Wells (1986) have proposed that an accurate approximation of the $100(1-\alpha)\%$ confidence interval for ρ for a moderately large value of a is given by

$$P\left\{\hat{\rho}_{\mathrm{ANOV}} - z_{\alpha/2}\sqrt{\widehat{\mathrm{Var}}(\hat{\rho}_{\mathrm{ANOV}})} \le \rho \le \hat{\rho}_{\mathrm{ANOV}} + z_{\alpha/2}\sqrt{\widehat{\mathrm{Var}}(\hat{\rho}_{\mathrm{ANOV}})}\right\} \doteq 1-\alpha, \tag{11.8.11}$$

where $\hat{\rho}_{\text{ANOV}}$ is defined in (11.4.36), $\widehat{\text{Var}}(\hat{\rho}_{\text{ANOV}})$ is defined in (11.6.17) with $\hat{\rho}_{\text{ANOV}}$ replacing ρ, and $z_{\alpha/2}$ is the $100(1-\alpha/2)$th percentile of the standard normal distribution. The confidence interval on τ can be obtained by the transformation $\tau = \rho/(1-\rho)$.

In addition, several other approximate procedures have been proposed in the literature which involve only simple noniterative calculations. The interested reader is referred to the papers by Shoukri and Ward (1984), Donner and Wells (1986), Groggel et al. (1988), Donner et al. (1989), Mian et al. (1989), and Kala et al. (1990). For a review of various procedures and their properties, see Donner (1986) and Donner and Wells (1986).

11.8.3 CONFIDENCE INTERVALS FOR σ_α^2.

As we have seen for the case of balanced data, there does not exist an exact confidence interval for σ_α^2. However, there are several approximate procedures available for this problem. In this section, we briefly describe two such procedures.

11.8.3.1 The Thomas–Hultquist Procedure

As in the case of confidence intervals for τ and ρ considered in Section 11.8.2.2, an approximate confidence interval for σ_α^2 can be obtained by substituting $\bar{n}_h$ for n and $F' = \bar{n}_h S_{\bar{y}}^2/\text{MS}_W$ for $F^* = \text{MS}_B/\text{MS}_W$ in any one of the approximate intervals for σ_α^2 described in Section 2.8.2.2. In particular, a Tukey–Moriguti–Williams interval is obtained as

$$P\left\{\frac{\text{MS}_W(F' - F[\nu_\alpha, \nu_e; 1-\alpha/2])}{\bar{n}_h F[\nu_\alpha, \infty; 1-\alpha/2]} \le \sigma_\alpha^2 \le \frac{\text{MS}_W(F' - F[\nu_\alpha, \nu_e; \alpha/2])}{\bar{n}_h F[\nu_\alpha, \infty; \alpha/2]}\right\} \doteq 1-\alpha. \tag{11.8.12}$$

Based on some Monte Carlo simulation results reported by Thomas and Hultquist (1978), in using the Moriguti–Bulmer procedure, it is found that the procedure gives satisfactory results in comparison to the exact method. However, the chi-square approximation used in the Thomas–Hultquist procedure does not perform well when $\tau < 0.25$ and the design is highly unbalanced. In such situations, the procedure can produce liberal intervals for σ_α^2.

11.8.3.2 The Burdick–Eickman Procedure

Burdick and Eickman (1986) have developed a procedure which seems to perform well over the entire range of values for τ. The desired $100(1-\alpha)\%$ confidence interval for σ_α^2 is given by

$$P\left\{\frac{LS_{\bar{y}}^2}{(1+\bar{n}_h L)F[\nu_\alpha, \infty; 1-\alpha/2]} \le \sigma_\alpha^2 \le \frac{US_{\bar{y}}^2}{(1+\bar{n}_h U)F[\nu_\alpha, \infty; \alpha/2]}\right\}$$

$$\doteq 1 - \alpha, \tag{11.8.13}$$

where

$$L = \frac{S_{\bar{y}}^2}{\bar{n}_h \mathrm{MS}_W F[\nu_\alpha, \nu_e; 1-\alpha/2]} - \frac{1}{n_{\min}},$$

$$U = \frac{S_{\bar{y}}^2}{\bar{n}_h \mathrm{MS}_W F[\nu_\alpha, \nu_e; \alpha/2]} - \frac{1}{n_{\max}},$$

with

$$n_{\min} = \min(n_1, n_2, \ldots, n_a) \quad \text{and} \quad n_{\max} = \max(n_1, n_2, \ldots, n_a).$$

The interval in (11.8.13), however, may produce negative limits which are defined to be zero. For balanced designs, the interval reduces to the Tukey–Moriguti–Williams interval described in Section 2.8.3.2. Some simulation work by Burdick and Eickman (1986) indicates that the interval is more conservative than the one in (11.8.12); however, the average lengths of the two intervals do not differ appreciably.

11.8.3.3 The Hartung–Knapp Procedure

Instead of using an approximate interval of τ as in the Thomas–Hultquist and Burdick–Eickman approach, Hartung and Knapp (2000) proposed using the exact interval for τ considered in Section 11.8.2.1. Let τ_L^* and τ_U^* be the lower and upper confidence limits of τ defined in (11.8.2) and (11.8.3). Then the proposed interval for σ_α^2 is

$$P\left\{ \frac{\mathrm{SS}_W \tau_L^*}{\chi^2[N-a, 1-\alpha/2]} \le \sigma_\alpha^2 \le \frac{\mathrm{SS}_W \tau_U^*}{\chi^2[N-a, \alpha/2]} \right\} \ge 1 - 2\alpha. \tag{11.8.14}$$

Using Bonferroni's inequality, it follows that the interval in (11.8.14) has confidence coefficient at least $1 - 2\alpha$.

Since the interval in (11.8.14) may be very conservative, Hartung and Knapp also proposed an approximate interval for σ_α^2 given by

$$P\left\{ \frac{\mathrm{SS}_W \tau_L^*}{N-a} \le \sigma_\alpha^2 \le \frac{\mathrm{SS}_W \tau_U^*}{N-a} \right\} \doteq 1 - \alpha. \tag{11.8.15}$$

Note that the interval in (11.8.15) is based on the MVU estimator of σ_e^2 instead of its confidence bounds.

Some simulation work by the authors confirm the results of Burdick and Eickman (1986) that the Thomas–Hultquist interval may be very liberal for small values of σ_α^2 while the Burdick–Eickman interval may be very conservative. The interval in (11.8.14) always has a confidence coefficient of at least $1 - \alpha$, but for large values of σ_α^2 this interval can be very conservative. The interval in (11.8.15) has a confidence coefficient of at least $1 - \alpha$ for small values of σ_α^2 and is only moderately conservative for large values of σ_α^2; and thus may be a good compromise for the whole range of values of σ_α^2.

11.8.4 A NUMERICAL EXAMPLE

In this section, we illustrate computations of confidence intervals on the variance components σ_e^2, σ_α^2, and contain of their parametric functions, using methods described in Sections 11.8.1 through 11.8.3, for the ratio units of electricity data of the numerical example in Section 11.4.11. From the results of the analysis of variance given in Table 11.3, we have

$$a = 5, \quad n_1 = 11, \quad n_2 = 8, \quad n_3 = 6, \quad n_4 = 24, \quad n_5 = 15,$$
$$\nu_e = 59, \quad \nu_\alpha = 4, \quad \text{MS}_W = 3.3565, \quad \text{MS}_B = 20.003.$$

Further, for $\alpha = 0.05$, we obtain

$$\chi^2[\nu_e, \alpha/2] = 39.6619, \qquad \chi^2[\nu_e, 1-\alpha/2] = 82.1174.$$

Substituting the appropriate quantities in (11.8.1), the desired 95% confidence interval for σ_e^2 is given by

$$P\{2.412 \leq \sigma_e^2 \leq 4.993\} = 0.95.$$

Now, we proceed to determine approximate 95% confidence intervals for $\tau = \sigma_\alpha^2/\sigma_e^2$, $\rho = \sigma_\alpha^2/(\sigma_e^2 + \sigma_\alpha^2)$, and σ_α^2 using the Thomas–Hultquist, Thomas–Hultquist–Donner, Burdick–Maqsood–Graybill, Donner–Wells, Burdick–Eickman, and Hartung–Knapp procedures. First, we compute the following quantities:

$$F[\nu_\alpha, \nu_e; \alpha/2] = 0.120, \qquad F[\nu_\alpha, \nu_e; 1-\alpha/2] = 3.012,$$
$$n_0 = 12.008, \quad \bar{n}_h = 10.1852, \quad S_2 = 1022, \quad S_3 = 19258,$$
$$S_{\bar{y}}^2 = 1.3787, \quad F' = 4.1836, \quad \hat{\rho}_{\text{ANOV}} = 0.293, \quad \text{and} \quad F^* = 5.9885.$$

Substituting the appropriate quantities in formulas (11.8.6) through (11.8.15), the desired intervals for τ, ρ, and σ_α^2 are readily calculated and are summarized in Table 11.6. For the purpose of comparison Wald's exact interval computed using SAS® code given in Burdick and Graybill (1992, Appendix B) is also included. It is understood that a negative limit is defined to be zero. Note that all the procedures except Donner–Wells produce somewhat wider intervals. This is typically the case; since the chi-square approximation used in the Thomas–Hultquist procedure does not perform well when $\tau < 0.25$ and the design is highly unbalanced. The Burdick–Maqsood–Graybill procedure can produce much wider intervals when $\tau < 0.25$ and the design is extremely unbalanced; and the Burdick–Eickman interval tends to be more conservative. The Donner–Wells interval may be slightly liberal; however, due to the small value of a, the accuracy of the approximation in (11.8.11) is somewhat unreliable. Hartung–Knapp I is known to be conservative and gives rise to a wider interval, while Hartung–Knapp II seems to be slightly tighter than expected.

TABLE 11.6 Approximate 95% confidence intervals for τ, ρ, and σ_α^2.

Parameter	Method	Confidence interval*
τ	Wald	(0.092, 2.265)
	Thomas–Hultquist	(0.038, 3.325)
	Burdick–Maqsood–Graybill	(−0.030, 3.381)
	Thomas–Hultquist–Donner	(0.082, 4.073)
	Donner–Wells	(−0.091, 2.021)
ρ	Wald	(0.084, 0.694)
	Thomas–Hultquist	(0.037, 0.769)
	Burdick–Maqsood–Graybill	(−0.029, 0.772)
	Thomas–Hultquist–Donner	(0.076, 0.803)
	Donner–Wells	(−0.083, 0.669)
σ_α^2	Thomas–Hultquist	(0.128, 11.160)
	Burdick–Eickman	(−0.201, 11.165)
	Hartung–Knapp I	(0.222, 11.310)
	Hartung–Knapp II	(0.309, 7.603)

*The negative bounds are defined to be zero.

11.9 TESTS OF HYPOTHESES

In this section, we briefly review the problem of testing hypotheses on variance components and certain of their parametric functions for the model in (11.1.1).

11.9.1 TESTS FOR σ_e^2 AND σ_α^2

The test of $H_0 : \sigma_e^2 = \sigma_{e0}^2$ vs. $H_1 : \sigma_e^2 \neq \sigma_{e0}^2$ (or $H_1 : \sigma_e^2 \geq (\leq)\sigma_{e0}^2$) for any specified value of σ_{e0}^2 can be based on MS_W which has multiple of a chi-squares distribution with $N - a$ degrees of freedom. This test is exactly the same as described in Section 2.9.1 for the case of balanced data. The test of $H_0 : \sigma_\alpha^2 = 0$ vs. $H_1 : \sigma_\alpha^2 > 0$ is performed using the ratio $\mathrm{MS}_B/\mathrm{MS}_W$ which has an F-distribution with $a - 1$ and $N - a$ degrees of freedom. Again, this test is the same test as discussed in Section 2.9.2 for the balanced model. The nonnull distribution of the test statistic has been investigated by Singh (1987) and can be employed to evaluate the power of the test. Tan and Wong (1980) have investigated the null and nonnull distribution of the F-ratio under the assumption of nonnormality. Donner and Koval (1989) have studied the performance of the F-test vis-à-vis the likelihood ratio test. Their results seem to indicate that the F-test is more powerful for testing nonzero values of σ_α^2 even for highly unbalanced data. However, the likelihood ratio test can be appreciably more powerful than the F-test in testing a null value of σ_α^2 if the design is extremely unbalanced. Othman (1983) and Jeyaratnam and Othman

(1985) have proposed an approximate test for this hypothesis for the data with heteroscedastic error variances.

It should be noted that although the F-test is exact, it is not uniformly optimum as was the case for the balanced design. Uniformly optimum tests, such as uniformly most powerful (UMP), uniformly most powerful unbiased (UMPU), uniformly most powerful invariant (UMPI), or uniformly most powerful invariant unbiased (UMPIU), generally do not exist in the case of unbalanced models. In such situations, the usual practice is to derive the so-called locally optimum tests, such as locally best unbiased (LBU), locally best invariant (LBI), and locally best invariant unbiased (LBIU) tests. Spjøtvoll (1967) has discussed optimal invariant tests for this hypothesis studying, in particular, tests which give high power for alternatives distant from the null hypothesis. Das and Sinha (1987) derived an LBIU test for this problem (see also Khuri et al. 1998, pp. 96–100). The LBIU test is based on the statistic

$$L = \frac{\sum_{i=1}^{a} n_i^2(\bar{y}_{i.} - \bar{y}_{..})^2}{\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2 + \sum_{i=1}^{a} n_i(\bar{y}_{i.} - \bar{y}_{..})^2}.$$

The LBIU is expected to yield a higher power compared to the F-test; and it reduces to the usual F test for balanced data.

11.9.2 TESTS FOR τ

An exact test for the hypothesis $H_0 : \tau = 0$ vs. $H_1 : \tau > 0$ is the usual F-test for testing σ_α^2. Spjøtvoll (1967) showed that the test is near optimal for alternatives in τ which are distant from the null hypothesis in the class of invariant and similar tests. Spjøtvoll (1968) presented several examples of exact power function calculations against the alternative $\tau = 0.1$ for the special case with $a = 3$ and $\alpha = 0.01$. Mostafa (1967) obtained a locally most powerful test for this hypothesis and compared the power function of this test with that of the F-test. Power comparisons of the F-test with other exact tests have also been made by Westfall (1988, 1989) and LaMotte et al. (1988). Westfall (1988) has considered a locally optimal test for this hypothesis based on the statistic

$$\sum_{i=1}^{a} n_i^2(\bar{y}_{i.} - \bar{y}_{..})^2 \Big/ \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2$$

and has investigated its robustness and power properties for nonnormal data. An exact test for the hypothesis $H_0 : \tau \leq \tau_0$ vs. $H_1 : \tau > \tau_0$ can be obtained by the Wald interval discussed in Section 11.8.2.1. For an unbalanced design in (11.1.1) there does not exist a uniformly most powerful test in the class of invariant and unbiased tests.

Spjøtvoll (1967) derived the most powerful invariant test for the hypothesis $\tau = \tau_0$ vs. $\tau = \tau_1$ (simple null, against simple alternative). Since the resulting test depends on τ_1, he also derived a test by letting $\tau_1 \to \infty$ which is independent of τ_1. This test is equivalent to the conventional Wald test and achieves high power for distant alternatives. Mostafa (1967) also considered a locally most powerful test for the hypothesis: $\tau = \tau_0$ vs. $\tau = \tau_0 + \Delta$, where Δ is small. Westfall (1989) has made a power comparison of Wald's test and the locally most powerful test under Pitman alternatives. Some robust tests for this hypothesis using jackknife statistics have been developed by Arvesen and Schmitz (1970), Arvesen and Layard (1975), and Prasad and Rao (1988), which are asymptotically distribution free. Donner and Koval (1989) developed the likelihood ratio test of $H_0 : \tau = \tau_0$ vs. $H_1 : \tau > \tau_0$ and compared its performance with that of the F-test. Hypothesis tests for τ including generalizations to higher-order mixed models are also discussed by LaMotte et al. (1988) and Lin and Harville (1991). For some additional results and a bibliography, see Verdooren (1988).

11.9.3 TESTS FOR ρ

An exact test for the hypothesis $H_0 : \rho = 0$ vs. $H_1 : \rho > 0$ is the usual F-test for testing σ_α^2. A significant value of F implies that $\rho > 0$, i.e., the proportion of variability attributable to the grouping factor is statistically significant, or, in other words, the elements of the same group tend to be similar with respect to the given characteristic. An exact test for the hypothesis[6] $H_0 : \rho = \rho_0$ vs. $H_1 : \rho > \rho_0$, where ρ_0 is a nonzero constant, was derived independently by Bhargava (1946) and Spjøtvoll (1967). The test is based on the statistic

$$F_E = \frac{\sum_{i=1}^{a} n_i(n_i\theta_0 + 1)^{-1}(\bar{y}_{i.} - \bar{y}_0)^2}{\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2},$$

where

$$\bar{y}_0 = \frac{\sum_{i=1}^{a} n_i(n_i\theta_0 + 1)^{-1}\bar{y}_{i.}}{\sum_{i=1}^{a} n_i(n_i\theta_0 + 1)^{-1}},$$

and $\theta_0 = \rho_0/(1 - \rho_0)$. The statistic F_E has an F-distribution with $a - 1$ and $N - a$ degrees of freedom, and the test is performed by rejecting the null hypothesis for large values of F_E. Note that at $\rho_0 = 0$, F_E reduces to the usual F-statistic for the standard one-way analysis of variance. Donner et al. (1989) considered two approximate tests for the above hypothesis based on the statistics

$$F' = (\mathrm{MS}_B/\mathrm{MS}_W)/(1 + n_0\theta_0) \quad \text{and} \quad F'' = (\bar{n}_h S_{\bar{y}}^2/\mathrm{MS}_W)/(1 + \bar{n}_h\theta_0).$$

Note that F' is the analogue of the F-statistic based on balanced data where n_0 is substituted for n and F'' is the same as the G-statistic defined in (11.8.5). The

[6]This hypothesis arises frequently in family studies where ρ measures the degree of resemblance among siblings with respect to a certain attribute or trait.

results on empirical significance values associated with F' and F'' and the corresponding exact p-values based on F_E show that the approximate methods may give very unsatisfactory results, and exact methods are therefore recommended for general use. Donner and Koval (1989) developed the likelihood-ratio test of $H_0 : \rho = \rho_0$ vs. $H_1 : \rho > \rho_0$ and compared its performance with that of the F-test.

Remark: Young and Bhandary (1998) derived a likelihood-ratio test and two large sample z tests for testing the equality of two intraclass correlation coefficients based on two independent samples drawn from multivariate normal distributions. ♦

11.9.4 A NUMERICAL EXAMPLE

In this section, we outline the results for testing the hypothesis $H_0 : \sigma_\alpha^2 = 0$ vs. $\sigma_\alpha^2 > 0$, or equivalently $H_0 : \tau = 0$ vs. $\tau > 0$ for the ratio units of electricity data of the numerical example in Section 11.4.11. Here, σ_α^2 and σ_e^2 correspond to variations among groups and replications, respectively. The usual F-test based on the ratio $\mathrm{MS}_B/\mathrm{MS}_W$ yields an F-value of 5.99 ($p < 0.001$). The results are highly significant and we reject H_0 and conclude that $\sigma_\alpha^2 > 0$, or that the data from different groups differ significantly. The test is exact but it is not uniformly optimum as was the case for the balanced design. Further, note that the results on confidence intervals support this conclusion.

EXERCISES

1. Show that minimal sufficient statistics in (11.3.1) are not complete.
2. From the log-likelihood equation (11.4.14) show that $L \to -\infty$ as $\sigma_e^2 \to 0$ and as $\sigma_e^2 \to \infty$, so that L must have a maximum for $\sigma_e^2 > 0$.
3. For the one-way random model in (11.1.1) derive the expression for the restricted log-likelihood function considered in Section 11.4.5.2 under the assumption of normality for the random effects.
4. Find the second-order partial derivatives of the log-likelihood function in (11.4.14) and examine whether the solutions from (11.4.15) through (11.4.17) maximize the likelihood function.
5. From the second-order partial derivatives in Exercise 4 determine the information matrix and the Cramér–Rao lower bounds for the variances of the estimators of μ, σ_α^2, and σ_e^2.
6. Show that for the balanced one-way random model the MINQUE and MIVQUE estimators of σ_α^2 and σ_e^2 considered in Section 11.4.8 coincide with the ANOVA estimators in (11.4.1).
7. Apply the method of "synthesis" to derive the expected mean squares given in Table 11.1.

8. Describe how the statistic H in (11.8.4) can have a $\chi^2[a-1]$ distribution for large $\sigma_\alpha^2/\sigma_e^2$ or large n_is.
9. Show that the ANOVA estimators of σ_e^2 and σ_α^2 in (11.4.1) are unbiased.
10. Spell out details of the derivation of the likelihood function in (11.4.13) and show that for the balanced design it reduces to the likelihood function in (2.4.2).
11. Derive expressions for sampling variances of the ANOVA and ML estimators of the variance components given in Sections 11.6.3.1 and 11.6.3.2.
12. Consider the model $y_{ij} = \mu + e_{ij}$, $i = 1, 2, \ldots, a$; $j = 1, 2, \ldots, n_i$, and $e_{ij} \sim N(0, \sigma_i^2)$. Show that the ML and REML estimators of σ_i^2 are given by

$$\hat{\sigma}^2_{i,\text{ML}} = \frac{\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2}{n_i}$$

and

$$\hat{\sigma}^2_{i,\text{REML}} = \frac{\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2}{n_i - 1},$$

where

$$\bar{y}_{i.} = \sum_{j=1}^{n_i} y_{ij}/n_i.$$

For $\sigma_i^2 \equiv \sigma_e^2$, show that the ML and REML estimators of σ_e^2 are given by

$$\hat{\sigma}^2_{e,\text{ML}} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2}{N}$$

and

$$\hat{\sigma}^2_{e,\text{REML}} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2}{N - 1},$$

where

$$\bar{y}_{..} = \sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}/N \quad \text{and} \quad N = \sum_{i=1}^{a} n_i.$$

13. Show that the unbiased estimators of the variances and covariance of the ANOVA estimators given in Section 11.6.3.1 are

$$\widehat{\text{Var}}(\hat{\sigma}^2_{e,\text{ANOV}}) = \frac{2\hat{\sigma}^2_{e,\text{ANOV}}}{N-a+2},$$

$$\widehat{\text{Var}}(\hat{\sigma}^2_{\alpha,\text{ANOV}}) = \frac{1}{1+h_5}\left(\frac{h_3 - h_2h_4}{1+h_1}\hat{\sigma}^2_{e,\text{ANOV}} + h_4\hat{\sigma}^2_{e,\text{ANOV}}\hat{\sigma}^2_{\alpha,\text{ANOV}} + h_5\hat{\sigma}^2_{\alpha,\text{ANOV}}\right),$$

and

$$\widehat{\text{Cov}}(\hat{\sigma}^2_{e,\text{ANOV}}, \hat{\sigma}^2_{\alpha,\text{ANOV}}) = \left(\frac{h_2}{h_1}\right)\widehat{\text{Var}}(\hat{\sigma}^2_{e,\text{ANOV}}),$$

where

$$h_1 = \frac{2}{N-a}, \qquad h_2 = \frac{-2N(a-1)}{(N-a)\left(N^2 - \sum_{i=1}^a n_i^2\right)},$$

$$h_3 = \frac{2N^2(N-1)(a-1)}{(N-a)\left(N^2 - \sum_{i=1}^a n_i^2\right)^2}, \qquad h_4 = \frac{4N}{N^2 - \sum_{i=1}^a n_i^2},$$

and

$$h_5 = \frac{2\left(N^2\sum_{i=1}^a n_i^2 + \left(\sum_{i=1}^a n_i^2\right)^2 - 2N\sum_{i=1}^a n_i^3\right)}{\left(N^2 - \sum_{i=1}^a n_i^2\right)^2}.$$

14. Consider the unbalanced one-way random model with unequal error variances, $y_{ij} = \mu + \alpha_i + e_{ij}$, $i = 1, 2, \ldots, a$; $j = 1, 2, \ldots, n_i$, $\alpha_i \sim N(0, \sigma^2_\alpha)$, and $e_{ij} \sim N(0, \sigma^2_i)$. Define the following quantities:

$$\hat{\sigma}^2_i = \sum_{j=1}^{n_i}\frac{(y_{ij} - \bar{y}_{i.})^2}{n_i - 1},$$

$$\hat{\sigma}^2_\alpha = \sum_{i=1}^{a}\frac{(\bar{y}_{i.} - \bar{y}^*_{..})^2}{a-1} - \frac{1}{a}\sum_{i=1}^{a}\frac{\hat{\sigma}^2_i}{n_i},$$

$$\text{MS}_B = \sum_{i=1}^{a}\frac{(\bar{y}_{i.} - \bar{y}^*_{..})^2}{a-1},$$

and

$$\text{MS}_W = \sum_{i=1}^{a}\sum_{j=1}^{n_i}\frac{(y_{ij} - \bar{y}_{i.})^2}{an_i(n_i-1)},$$

where

$$\bar{y}_{i.} = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij} \quad \text{and} \quad \bar{y}^*_{..} = \frac{1}{a}\sum_{i=1}^{a}\bar{y}_{i.}.$$

Show that

(a) $E(\hat{\sigma}_i^2) = \sigma_i^2$.

(b) $E(\hat{\sigma}_\alpha^2) = \sigma_\alpha^2$.

(c) $E(\text{MS}_B) = \sigma_\alpha^2 + \frac{1}{a}\sum_{i=1}^a \frac{\sigma_i^2}{n_i}$.

(d) MS_B and MS_W are stochastically independent.

(e) $\frac{(n_i-1)\hat{\sigma}_i^2}{\sigma_i^2}$ has a chi-square distribution with $n_i - 1$ degrees of freedom.

(f) $(k-1)\text{MS}_B$ has the same distribution as $U = \sum \lambda_i U_i$, where U_is are independent chi-square variates each with one degree of freedom and λ_is are characteristic roots of the matrix $\boldsymbol{AS}$, where $\boldsymbol{A}$ is an $a \times a$ matrix with diagonal elements $1 - 1/a$ and off-diagonal elements $-1/a$, and $\boldsymbol{S}$ is an $a \times a$ diagonal matrix with diagonal elements $\sigma_\alpha^2 + \sigma_i^2/n_i$ $(i = 1, 2, \ldots, a)$.

(g) $\nu_1 U/[(a-1)\sigma_\alpha^2 + \sum_{i=1}^a \sigma_i^2/an_i]$ has an approximate chi-square distribution with ν_1 degrees of freedom, where

$$\nu_1 = \frac{(a-1)^2\left[\sigma_\alpha^2 + \sum_{i=1}^a \sigma_i^2/an_i\right]^2}{\sum_{i=1}^a \lambda_i^2}.$$

(Hint: Use Satterthwaite approximation.)

(h) $\nu_2 \text{MS}_W/[\sum_{i=1}^a \sigma_i^2/an_i]$ has an approximate chi-square distribution with ν_2 degrees of freedom, where

$$\nu_2 = \frac{(\sum_{i=1}^a \sigma_i^2/n_i)^2}{\sum_{i=1}^a \sigma_i^4/n_i^2(n_i-1)}.$$

15. Refer to Exercise 14 above and show that an approximate α-level test for testing $H_0 : \sigma_\alpha^2 = 0$ vs. $H_1 : \sigma_\alpha^2 > 0$ is obtained by rejecting H_0 if and only if $\text{MS}_B/\text{MS}_W > F[v_1, v_2; 1-\alpha]$, where v_1 and v_2 are estimated by (Jeyaratnam and Othman, 1985)

$$\hat{v}_1 = \frac{\left[(a-1)\sum_{i=1}^a \frac{\hat{\sigma}_i^2}{an_i}\right]^2}{\left(\sum_{i=1}^a \frac{\hat{\sigma}_i^2}{an_i}\right)^2 + (a-2)\sum_{i=1}^a \frac{\hat{\sigma}_i^4}{an_i^2}}$$

and

$$\hat{v}_2 = \frac{\left(\sum_{i=1}^a \frac{\hat{\sigma}_i^2}{n_i}\right)^2}{\sum_{i=1}^a \frac{\hat{\sigma}_i^4}{n_i^2(n_i-1)}}.$$

16. Show that the formulas for confidence intervals of variance components and their parametric functions given in Section 11.8 reduce to the corresponding formulas given in Section 2.8.

17. Show that the procedures for testing hypotheses on the between group variance component (σ_α^2) considered in Section 11.9.1 reduce to the usual F-test for the balanced model.

18. Sokal and Rohlf (1995, p. 210) reported data on morphological measurements of the width of the scutum (dorsal shield) of samples of tick larvae obtained from four different host individuals of the cottontail rabit. The hosts were obtained at random from certain localities and can be considered to be a representative sample of the host individuals from the given locality. The data are given below.

Host			
1	**2**	**3**	**4**
380	350	354	376
376	356	360	344
360	358	362	342
368	376	352	372
372	338	366	374
366	342	372	360
374	366	362	
382	350	344	
	344	342	
	364	358	
		351	
		348	
		348	

Source: Sokal and Rohlf (1995); used with permission.

(a) Describe the mathematical model and the assumptions involved.
(b) Analyze the data and report the analysis of variance table.
(c) Perform an appropriate F-test to determine whether the morphological measurements differ from host to host.
(d) Find point estimates of the variance components, the ratio of the variance components, the intraclass correlation, and the total variance using the ANOVA, ML, and REML procedures.
(e) Calculate 95% confidence intervals associated with the point estimates in part (d), using the methods described in the text.

19. An experiment was designed to test the variation in cycles at which failure occurred on beams from different batches of concrete. A sample of five batches was randomly selected and the data cycles rounded to 10 are given below.

Batch				
1	2	3	4	5
800	850	810	650	840
600	810	880	770	950
760	960	880	840	

(a) Describe the mathematical model and the assumptions involved.

(b) Analyze the data and report the analysis of variance table.

(c) Perform an appropriate F-test to determine whether the failure cycles differ from batch to batch.

(d) Find point estimates of the variance components, the ratio of the variance components, the intraclass correlation, and the total variance using the ANOVA, ML, and REML procedures.

(e) Calculate 95% confidence intervals associated with the point estimates in part (d), using the methods described in the text.

20. An experiment was conducted to test batch to batch variation in luminous flux of lamps. A sample of 5 batches was selected and the data on the results of testing lamps for luminous flux (lumens per watt) are given below.

Batch				
1	2	3	4	5
8.48	9.81	9.38	9.66	8.55
8.01	10.29	9.43	9.34	7.63
8.13	10.16	9.29	8.78	7.95
8.28	9.87	9.65	8.58	
8.29	10.31			
8.26				

(a) Describe the mathematical model and the assumptions involved.

(b) Analyze the data and report the analysis of variance table.

(c) Perform an appropriate F-test to determine whether the luminous flux differs from batch to batch.

(d) Find point estimates of the variance components, the ratio of the variance components, the intraclass correlation, and the total variance using the ANOVA, ML, and REML procedures.

(e) Calculate 95% confidence intervals associated with the point estimates in part (d), using the methods described in the text.

21. Snedecor and Cochran (1989, p. 246) described the results of an investigation on artificial insemination of cows to test for their ability to produce conception. Semen samples from bulls were taken and tested to determine percentages of conceptions to services for successive samples. The results on six bulls from a larger data set are given below.

Bull					
1	**2**	**3**	**4**	**5**	**6**
46	70	52	42	42	35
31	59	44	21	64	68
37		57	70	50	59
62		40	46	69	38
30		67	14	77	57
		64		81	76
		70		87	57
					29
					60

Source: Snedecor and Cochran (1989); used with permission.

(a) Describe the mathematical model and the assumptions involved.

(b) Analyze the data and report the analysis of variance table.

(c) Perform an appropriate F-test to determine whether the percentages of conceptions differ from bull to bull.

(d) Find point estimates of the variance components, the ratio of the variance components, the intraclass correlation, and the total variance using the ANOVA, ML, and REML precedures.

(e) Calculate 95% confidence intervals associated with the point estimates in part (d), using the methods described in the text.

Bibliography

H. J. Ahrens (1978), MINQUE and ANOVA estimator for one-way classification: A risk comparison, *Biometrical J.*, **20**, 535–556.

H. J. Ahrens and R. Pincus (1981), On two measures of unbalancedness in a one-way model and their relation to efficiency, *Biometrical J.*, **23**, 227–235.

H. J. Ahrens and J. E. Sánchez (1982), Unbalancedness and efficiency in estimating components of variance: MINQUE and ANOVA procedure, *Biometrical J.*, **24**, 649–661.

H. J. Ahrens and J. E. Sánchez (1986), *Measures of Unbalancedness and Their Relations to Variance Component Estimation*, Technical Report P-Math-34/8, Akademie der Wissenschaften der DDR, Karl-Weierstrass-Insitut für Matematik, Berlin.

H. J. Ahrens and J. E. Sánchez (1988), Unbalancedness of designs, Measures of, in S. Kotz and N. L. Johnson, *Encyclopedia of Statistical Sciences*, Vol. 9, Wiley, New York, 383–386.

H. J. Ahrens and J. E. Sánchez (1992), Imbalance and its influence on variance component estimation, *Biometrical J.*, **34**, 539–555.

H. J. Ahrens, J. Kleffé, and R. Tenzler (1981), Mean square error comparison of MINQUE, ANOVA and two alternative estimators under the unbalanced one-way random effects model, *Biometrical J.*, **23**, 323–342.

R. L. Anderson (1981), Recent developments in designs and estimators for variance components, in M. Csörgö, D. A. Dawson, J. N. K. Rao, and A. K. Md. E. Saleh, eds., *Statistics and Related Topics*, North-Holland, Amsterdam, 3–22.

R. L. Anderson and P. P. Crump (1967), Comparisons of designs and estimation procedures for estimating parameters in a two-stage nested process, *Technometrics*, **9**, 499–516.

J. N. Arvesen (1976), A note on the Tukey–Hooke variance component results, *Ann. Inst. Statist. Math.* (*Japan*), **28**, 111–121.

J. N. Arvesen and M. W. J. Layard (1975), Asymptotically robust tests in unbalanced variance components models, *Ann. Statist.*, **3**, 1122–1134.

J. N. Arvesen and T. H. Schmitz (1970), Robust procedures for variance component problems using the jacknife, *Biometrics*, **26**, 677–686.

A. H. J. Baines (1943), *On the Allocation of Observations in Certains Types of Factorial Experiments when the Number of Observations is Limited*, Technical Report Q.C./R/4, Ministry of Supply Advisory Service on Statistical Method and Quality Control, London.

N. K. Bansal and M. Bhandary (1994), Robust M-estimation of the intraclass correlation coefficient, *Australian J. Statist.*, **36**, 287–301.

E. Belzile and J.-F. Angers (1995), Choice of noninformative priors for the variance components of an unbalanced one-way random model, *Comm. Statist.* A *Theory Methods*, **24**, 1325–1341.

A. Bener and S. Huda (1993), Estimation of intraclass correlation when family sizes are unequal, *Comput. Statist. Data Anal.*, **15**, 273–284.

J. O. Berger and J. M. Bernardo (1992), Reference priors in a variance components problem, in P. K. Goel and N. S. Iyenger, eds., *Bayesian Analysis in Statistics and Econometrics*, Springer-Verlag, New York, 177–194.

R. P. Bhargava (1946), Tests of significance for intraclass correlation when family sizes are not equal, *Sankhyā*, **7**, 435–438.

K. A. Brownlee (1965), *Statistical Theory and Methodology in Science and Engineering*, 2nd ed., Wiley, New York.

R. K. Burdick and J. E. Eickman (1986), Confidence intervals on the among group variance component in the unbalanced one-fold nested design, *J. Statist. Comput. Simul.*, **26**, 205–219.

R. K. Burdick and F. A. Graybill (1992), *Confidence Intervals on Variance Components*, Marcel Dekker, New York.

R. K. Burdick, F. Maqsood, and F. A. Graybill (1986), Confidence intervals on the intraclass correlation in the unbalanced one-way classification, *Comm. Statist.* A *Theory Methods*, **15**, 3353–3378.

J. M. Cameron (1951), Use of components of variance in preparing schedules for the sampling of baled wool, *Biometrics*, **7**, 83–96.

R. F. Caro, M. Grossman, and R. L. Fernando (1985), Effects of data imbalance on estimation of heritability, *Theoret. Appl. Genetics*, **69**, 523–530.

K. Chaloner (1987), A Bayesian approach to the estimation of variance components for the unbalanced one-way model, *Technometrics*, **29**, 323–337.

S. K. Chatterjee and K. Das (1983), Estimation of variance components in an unbalanced one-way classification, *J. Statist. Plann. Inference*, **8**, 27–41.

Y. B. Chaubey (1984), On the comparison of some non-negative estimators of variance components for two models, *Comm. Statist.* B *Simul. Comput.*, **13**, 619–633.

W. G. Cochran (1937), Problems arising in the analysis of a series of similar experiments, *J. Roy. Statist. Soc. Suppl.*, **4**, 102–118.

W. G. Cochran (1939), The use of the analysis of variance in enumeration by sampling, *J. Amer. Statist. Assoc.*, **34**, 492–510.

W. G. Cochran (1954), The combination of estimates from different experiments, *Biometrics*, **10**, 109–129.

W. G. Cochran and S. P. Carroll (1953), A sampling investigation of the efficiency of weighting inversely as the estimated variance, *Biometrics*, **9**, 447–459.

M. D. Conerly and J. T. Webster (1987), MINQE for the one-way classification, *Technometrics*, **29**, 229–236.

P. P. Crump (1954), *Optimal Designs to Estimate the Parameters of a Variance Component Model*, Ph.D. dissertation, North Carolina State University, Raleigh, NC.

S. L. Crump (1946), The estimation of variance components in analysis of variance, *Biometrics Bull.*, **2**, 7–11.

S. L. Crump (1947), *The Estimation of Components of Variance in Multiple Classifications*, Ph.D. dissertation, Iowa State University, Ames, IA.

S. L. Crump (1951), The present status of variance components analysis, *Biometrics*, **7**, 1–16.

R. Das and B. K. Sinha (1987), Robust optimum invariant unbiased test for variance components, in T. Pukkila and S. Puntanen, eds., *Proceedings of the 2nd International Tampere Conference on Statistics*, University of Tampere, Tampere, Finland, 317–342.

R. B. Davies (1980), The distribution of a linear combination of χ^2 random variables, *Appl. Statist.*, **29**, 323–333.

A. Donner (1979), The use of correlation and regression in the analysis of family resemblance, *Amer. J. Epidemiology*, **110**, 335–342.

A. Donner (1986), A review of inference procedures for the intraclass correlation coefficient in the one-way random effects model, *Internat. Statist. Rev.*, **54**, 67–82.

A. Donner, M. Eliasziw, and M. Shoukri (1998), Review of inference procedures for the interclass coefficient with emphasis on applications to family studies, *Genetic Epidemiology*, **15**, 627–646.

A. Donner and J. J. Koval (1980a), The estimation of intraclass correlation in the analysis of family data, *Biometrics*, **36**, 19–25.

A. Donner and J. J. Koval (1980b), The large sample variance of an intraclass correlation, *Biometrika*, **67**, 719–722.

A. Donner and J. J. Koval (1983a), Variance component estimation from human sibship data, *Biometrics*, **39**, 599–605.

A. Donner and J. J. Koval (1983b), A note on the accuracy of Fisher's approximation to the large sample variance of an intraclass correlation, *Comm. Statist.* B *Simul. Comput.*, **12**, 443–449.

A. Donner and J. J. Koval (1989), The effect of imbalance on significance-testing in one-way model II analysis of variance, *Comm. Statist.* A *Theory Methods*, **18**, 1239–1250.

A. Donner and G. A. Wells (1986), A comparison of confidence interval methods for the intraclass correlation coefficient, *Biometrics*, **42**, 401–412.

A. Donner, G. A. Wells, and M. Eliasziw (1989), On two aproximations to the F-distribution: Application to testing for intraclass correlation in family studies, *Canadian J. Statist.*, **17**, 209–215.

M. Y. El-Bassiouni and J. F. Seely (1988), On the power of Wald's variance component test in the unbalanced random one-way model, in Y. Dodge, V. V. Fedorov, and H. P. Wynn, eds., *Optimal Design and Analysis of Experiments*, North-Holland, Amsterdam, 157–165.

A. Giovagnoli and P. Sebastiani (1989), Experimental designs for mean and variance estimation in variance components models, *Comput. Statist. Data Anal.*, **8**, 21–28.

F. A. Graybill (1961), *An Introduction to Linear Statistical Models*, Vol. I, McGraw–Hill, New York.

D. J. Groggel, D. D. Wackerly, and P. V. Rao (1988), Nonparametric estimation in one-way random effects models, *Comm. Statist.* B *Simul. Comput.*, **17**, 887–903.

J. M. Hammersley (1949), The unbiased estimate and standard error of the interclass variance, *Metron*, **15**, 189–205.

H. O. Hartley and J. N. K. Rao (1967), Maximum likelihood estimation for the mixed analysis of variance model, *Biometrika*, **54**, 93–108.

J. Hartung and G. Knapp (2000), Confidence intervals for the between group variance in the unbalanced one-way random effects model, *J. Statist. Comput. Simul.*, **65**, 311–323.

D. A. Harville (1969a), *Variance Component Estimation for the Unbalanced One-Way Random Classification*, Technical Report ARL-69-0180, Aerospace Research Laboratories, Wright-Patterson Air Force Base, Dayton, OH.

D. A. Harville (1969b), Expression of variance component estimators as linear combinations of independent non-central chi-square variates, *Ann. Math. Statist.*, **40**, 2189–2194.

D. A. Harville and A. P. Fenech (1985), Confidence intervals for a variance ratio, or for heritability, in an unbalanced mixed linear model, *Biometrics*, **41**, 137–152.

B. Heine (1993), Nonnegative estimation of variance components in an unbalanced one-way random effects model, *Comm. Statist.* A *Theory Methods*, **22**, 2351–2372.

C. R. Henderson, S. R. Searle, and D. S. Robson (1957), *Seminar on Unbalanced Data*, mimeographed class notes, Cornell University, Ithaca, NY.

J. L. Hess (1979), Sensitivity of the MINQUE with respect to prior weights, *Biometrics*, **5**, 645–649.

B. M. Hill (1965), Inference about variance component in the one-way model, *J. Amer. Statist. Assoc.*, **60**, 806–825.

R. A. Hultquist and F. A. Graybill (1965), Minimal sufficient statistics for the two-way classification mixed model design, *J. Amer. Statist. Assoc.*, **60**, 183–192.

S. Jeyaratnam and A. R. Othman (1985), Test of hypothesis in the one-way random effects model with unequal error variances, *J. Statist. Comput. Simul.*, **21**, 51–57.

R. Kala, A. Molińska, and K. Moliński (1990), Approximate confidence intervals for a ratio of two variance components, *Comm. Statist.* A *Theory Methods*, **19**, 1889–1898.

S. Karlin, P. E. Cameron, and P. Williams (1981), Sibling and parent-offspring correlation estimation with variable family age, *Proc. Nat. Acad. Sci.* (*U.S.A.*), **61**, 302–315.

R. J. Kelly and T. Mathew (1993), Improved estimators of variance components with smaller probability of negativity, *J. Roy. Statist. Soc. Ser.* B, **4**, 897–911.

R. J. Kelly and T. Mathew (1994), Improved non-negative estimation of variance components in some mixed models with unbalanced data, *Technometrics*, **36**, 171–181.

R. Khattree and D. S. Gill (1988), Comparison of some estimates of variance components using Pitman's nearness criterion, in *ASA Proceedings of Statistical Computing Section*, American Statistical Association, Alexandria, VA, 133–136.

A. I. Khuri (1987), Measures of imbalance for unbalanced data, *Biometrical J.*, **29**, 383–396.

A. I. Khuri (1996), A method for determining the effect of imbalance, *J. Statist. Plann. Inference*, **55**, 115–129.

A. I. Khuri, T. Mathew, and B. K. Sinha (1998), *Statistical Tests for Mixed Linear Models*, Wiley, New York.

J. Kleffé (1975), Best quadratic unbiased estimators for variance components in the balanced two-way classification model, *Math. Oper. Statist. Ser. Statist.*, **6**, 189–196.

J. Kleffé (1993), Parent-offspring and sibling correlation estimation based on MINQUE theory, *Biometrika*, **80**, 393–404.

G. G. Koch (1967a), A general approach to the estimation of variance components, *Techometrics*, **9**, 93–118.

G. G. Koch (1967b), A procedure to estimate the population mean in random effects models, *Technometrics*, **9**, 577–586.

G. G. Koch (1968), Some further remarks concerning "A general approach to estimation of variance components," *Technometrics*, **10**, 551–558.

K. Kussmaul and R. L. Anderson (1967), Estimation of variance components in two-stage nested designs with composite samples, *Technometrics*, **9**, 373–389.

L. R. LaMotte, A. McWhorter, and R. A. Prasad (1988), Confidence intervals and tests on the variance ratio in random models with two variance components, *Comm. Statist.* A *Theory Methods*, **17**, 1135–1164.

J. Lee and A. I. Khuri (1999), Graphical technique for comparing designs for random models, *J. Appl. Statist.*, **26**, 933–947.

J. Lee and A. I. Khuri (2000), Quantile dispersion graphs for the comparison of designs for a random two-way model, *J. Statist. Plann. Inference*, **91**, 123–137.

J. Lee and A. I. Khuri (2001), Modeling the probability of a negative ANOVA estimate of a variance component, *Calcutta Assoc. Bull.*, **51**, 201–202.

Y. Lee and J. Seely (1996), Computing the Wald interval for a variance ratio, *Biometrics*, **52**, 1486–1491.

M. L. Lera (1994), Measures of imbalance for higher-order designs, *Biometrical J.*, **36**, 481–490.

T.-H. Lin and D. A. Harville (1991), Some alternatives to Wald's confidence interval and test, *J. Amer. Statist. Assoc.*, **86**, 179–187.

T. Mathew, B. K. Sinha, and B. C. Sutradhar (1992a), Nonnegative estimation of variance components in unbalanced mixed models with two variance components, *J. Multivariate Anal.*, **42**, 77–101.

I. U. H. Mian, M. M. Shoukri, and D. S. Tracy (1989), A comparison of significance testing procedures for the intraclass correlation from family data, *Comm. Statist.* B *Simul. Comput.*, **18**, 613–631.

M. G. Mostafa (1967), Note on testing hypotheses in an unbalanced random effects model, *Biometrika*, **54**, 659–662.

R. Mukerjee (1990), Optimality of balanced designs for minimum norm quadratic unbiased estimation of variance components, *Metrika*, **37**, 181–188.

R. Mukerjee and S. Huda (1988), Optimal design for the estimation of variance components, *Biometrika*, **75**, 75–80.

K. K. Namboodiri, P. P. Green, E. B. Kaplan, J. A. Morrison, G. A. Chase, R. C. Elston, A. R. G. Owen, B. M. Rifkind, C. J. Glueck, and H. A. Tyroler (1984), The collaborative lipid research clinics program: Family study IV: Familial associations of plasma lipids and lipoproteins, *Amer. J. Epidemiology*, **119**, 975–996.

I. Olkin and J. W. Pratt (1958), Unbiased estimation of certain correlation coefficients, *Ann. Math. Statist.*, **29**, 201–211.

A. R. Othman (1983), *One-Way Random Effects Model with Unequal Error Variances*, M.S. thesis, Southern Illinois University, Carbondale, IL.

N. G. N. Prasad and J. N. K. Rao (1988), Robust test and confidence intervals for error variance in a regression model and for functions of variance components in an unbalanced random one-way model, *Comm. Statist.* A *Theory Methods*, **17**, 1111–1133.

S. J. Press (1966), Linear combinations of non-central chi-square variates, *Ann. Math. Statist.*, **37**, 480–487.

C. R. Rao (1973), *Linear Statistical Inference and Its Applications*, 2nd ed., Wiley, New York.

P. S. R. S. Rao (1982), Use of prior information for estimating the variance components, *Comm. Statist.* A *Theory Methods*, **11**, 1659–1669.

P. S. R. S. Rao (1997), *Variance Components Estimation: Mixed Models, Methodologies, and Applications*, Chapman and Hall, London.

P. S. R. S. Rao (2001), Nonnegative estimators for the one-way random effects model, *Comm. Statist.* A *Theory Methods*, **30**, 1605–1613.

P. S. R. S. Rao, J. Kaplan, and W. G. Cochran (1981), Estimators for the one-way random effects model with unequal error variances, *J. Amer. Statist. Assoc.*, **76**, 89–97.

P. S. R. S. Rao and E. A. Sylvestre (1984), ANOVA and MINQUE type of estimators for the one-way random effects model, *Comm. Statist.* A *Theory Methods*, **13**, 1667–1673.

R. R. Read (1961), On quadratic estimates of intraclass variance for unbalanced designs, *J. Roy. Statist. Soc. Ser.* B, **23**, 493–497.

A. Robertson (1962), Weighting in the estimation of variance components in the unbalanced single classification, *Biometrics*, **18**, 413–417.

J. Robinson (1965), The distribution of general quadratic form in normal variables, *Australian J. Statist.*, **7**, 110–114.

J. Sánchez (1983), Sample variances for several MINQ-estimators in the one-way classification, *Biometrical J.*, **25**, 259–267.

H. Scheffé (1959), *The Analysis of Variance*, Wiley, New York.

S. R. Searle (1956), Matrix methods in components of variance and covariance analysis, *Ann. Math. Statist.*, **27**, 737–748.

S. R. Searle (1967), *Computer Simulation of Variance Component Estimates*, Mim. Ser. Paper BU-233-M, Biometrics Unit, Cornell University, Ithaca, NY.

S. R. Searle (1971a), Topics in variance components estimation, *Biometrics*, **27**, 1–76.

S. R. Searle (1971b), *Linear Models*, Wiley, New York.

S. R. Searle (1978), A univariate formulation of the multivariate linear model, in H. A. David, ed., *Contribution to Survey Sampling and Applied Statistics: Papers in Honor of H. O. Hartley*, Academic Press, San Diego, 181–189.

S. R. Searle, G. Casella, and C. E. McCulloch (1992), *Variance Components*, Wiley, New York.

J. F. Seely and M. Y. El-Bassiouni (1983), Applying Wald's variance component test, *Ann. Statist.*, **11**, 197–201.

M. M. Shoukri and R. H. Ward (1984), On the estimation of the intraclass correlation, *Comm. Statist.* A *Theory Methods*, **13**, 1239–1258.

B. Singh (1987), On the non-null distribution of ANOVA F-ratio in one-way unbalanced random model, *Calcutta Statist. Assoc. Bull.*, **36**, 57–62.

B. Singh (1989a), A comparison of variance component estimators under unbalanced situations, *Sankhyā Ser.* B, **51**, 323–330.

B. Singh (1989b), Moments of group variance estimator in one-way unbalanced classfication, *J. Indian Soc. Agricultural Statist.*, **41**, 77–84.

B. Singh (1991), Probability of negative estimates of heritability in one-way unbalanced random model, *J. Indian Soc. Agricultural Statist.*, **43**, 74–80.

B. Singh (1992), On effect of unbalancedness and heteroscedasticity on the ANOVA estimator of group variance component in one-way random model, *Biometrical J.*, **34**, 91–96.

R. A. Singhal (1984), Effect of non-normality on the estimation of functions of variance components, *J. Indian Soc. Agricultural Statist.*, **35**, 89–98.

C. A. B. Smith (1956), On the estimation of intraclass correlation, *Ann. Human Genetics*, **21**, 363–373.

C. A. B. Smith (1980a), Estimating genetic correlations, *Ann. Human Genetics*, **43**, 265–284.

C. A. B. Smith (1980b), Further remarks on estimating genetic correlations, *Ann. Human Genetics*, **44**, 95–105.

G. W. Snedecor and W. G. Cochran (1989), *Statistical Methods*, 8th ed., Iowa State University Press, Ames, IA; 6th ed., 1967; 7th ed., 1980.

R. R. Sokal and F. J. Rohlf (1995), *Biometry*, 3rd ed., W. H. Freeman, New York; 1st ed., 1969; 3rd ed., 1981.

E. Spjøtvoll (1967), Optimum invariant tests in unbalanced variance components models, *Ann. Math. Statist.*, **38**, 422–428.

E. Spjøtvoll (1968), Most powerful tests for some non-exponential families, *Ann. Math. Statist.*, **39**, 772–784.

W. H. Swallow (1974), *Minimum Norm and Minimum Variance Quadratic Unbiased Estimation of Variance Components*, Ph.D. dissertation, Cornell University, Ithaca, NY.

W. H. Swallow (1981), Variances of locally minimum variance quadratic unbiased estimators (MINQUE) of variance components, *Technometrics*, **23**, 271–283.

W. H. Swallow and J. F. Monaham (1984), Monte Carlo comparison of ANOVA4, MINQUE, REML, and ML estimators of variance components, *Technometrics*, **26**, 47–57.

W. H. Swallow and S. R. Searle (1978), Minimum variance quadratic unbiased estimation (MINQUE) of variance components, *Technometrics*, **20**, 265–272.

L. A. Swiger, L. R. Harvey, D. O. Everson, and K. E. Gregory (1964), The variance of intraclass correlation involving groups with one observation, *Biometrics*, **20**, 818–826.

W. Y. Tan and S. P. Wong (1980), On approximating the null and non-null distributions of the F-ratio in unbalanced random effects models from non-normal universes, *J. Amer. Statist. Assoc.*, **75**, 655–662.

J. D. Thomas and R. A. Hultquist (1978), Interval estimation for the unbalanced case of the one-way random effects model, *Ann. Statist.*, **6**, 582–587.

W. O. Thompson and R. L. Anderson (1975), A comparison of designs and estimators for the two-stage nested random model, *Technometrics*, **17**, 37–44.

E. C. Townsend (1968), *Unbiased Estimators of Variance Components in Simple*

Unbalanced Designs, Ph.D. dissertation, Cornell University, Ithaca, NY.

E. C. Townsend and S. R. Searle (1971), Best quadratic unbiased estimation of variance components from unbalanced data in the one-way classification, *Biometrics*, **27**, 643–657.

J. W. Tukey (1957), Variances of variance components II: The unbalanced single classification, *Ann. Math. Statist.*, **28**, 43–56.

L. J. Van der Kemp (1972), Reliability coefficients via intraclass correlation: Biased and unbiased estimators, *Neederland. Tidschrift Psychol. Haar Grensgebieden*, **27**, 447–459.

M. G. Vangel and A. L. Rukhin (1999), Maximum likelihood analysis for heteroscedastic one-way random effects ANOVA in interlaboratory studies, *Biometrics*, **55**, 129–136.

L. R. Verdooren (1976), Exact confidence intervals for the variance ratio in nested designs, in J. Gordesch and P. Naeve, eds., *COMPSTAT: Proceedings of the* 2*nd Symposium on Computational Statistics*, Physica-Verlag, Wien, Germany, 221–228.

L. R. Verdooren (1988), Exact tests and confidence intervals for the ratio of variance components in unbalanced two- and three-stage nested designs, *Comm. Statist.* A *Theory Methods*, **17**, 1197–1230.

A. Wald (1940), A note on analysis of variance with unequal class frequencies, *Ann. Math. Statist.*, **11**, 96–100.

W. Y. Wang (1967), A comparison of several variance component estimators, *Biometrika*, **54**, 301–305.

P. H. Westfall (1987), A comparison of variance component estimates for arbitrary underlying distributions, *J. Amer. Statist. Assoc.*, **82**, 866–874.

P. H. Westfall (1988), Robustness and power of test for a null variance ratio, *Biometrika*, **75**, 207–214.

P. H. Westfall (1989), Power comparisons for invariant variance ratio tests in mixed ANOVA models, *Ann. Statist.*, **17**, 318–326.

P. H. Westfall (1994), Asymptotic efficiencies of MINQUE and ANOVA variance component estimates in the nonnormal random model, in T. Calinski and R. Kal, eds., *Proceedings of the International Conference on Linear Statistical Inference*, Kluwer, Amsterdam, 113–119.

C. P. Winsor and G. L. Clarke (1940), A statistical study of variation in the catch of plankton nets, *J. Marine Res.*, **3**, 1–34.

F. Yates and W. G. Cochran (1938), The analysis of groups of experiments, *J. Agricultural Sci.* (*Cambridge*), **28**, 556–580.

D. J. Young and M. Bhandary (1998), Test for equality of intraclass correlation coefficients under unequal family sizes, *Biometrics*, **54**, 1363–1373.

H. Yu, S. R. Searle, and C. E. McCulloch (1994), Properties of maximum likelihood estimators of variance components in the one-way classification, *Comm. Statist.* B *Simul. Comput.*, **23**, 897–914.

12 Two-Way Crossed Classification without Interaction

In this chapter, we consider the random effects model involving two factors in a factorial arrangement where the numbers of observations in each cell are different. We further assume that the model does not involve any interaction terms. Consider two factors A and B and let there be n_{ij} (≥ 0) observations corresponding to the (i, j)th cell. The model for this design is known as the two-way crossed classification without interaction.

12.1 MATHEMATICAL MODEL

The random effects model for the unbalanced two-way crossed classification without interaction is given by

$$y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}; \quad i = 1, \ldots, a; \quad j = 1, \ldots, b; \quad k = 0, \ldots, n_{ij}, \tag{12.1.1}$$

where y_{ijk} is the kth observation corresponding to the ith level of factor A and the jth level of factor B, μ is the overall mean, α_is and β_js are main effects, i.e., α_i is the effect of the ith level of factor A, β_j is the effect of the jth level of factor B, and e_{ijk} is the customary error term. It is assumed that $-\infty < \mu < \infty$ is a constant and α_is, β_js, and e_{ijk}s are mutually and completely uncorrelated random variables with means zero and variances σ_α^2, σ_β^2, and σ_e^2, respectively. The parameters σ_α^2, σ_β^2, and σ_e^2 are known as the variance components.

12.2 ANALYSIS OF VARIANCE

For the model in (12.1.1) there is no unique analysis of variance. The conventional analysis of variance obtained by an analogy with the corresponding balanced design is given in Table 12.1.

TABLE 12.1 Analysis of variance for the model in (12.1.1).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Factor A	$a-1$	SS_A	MS_A	$\sigma_e^2 + r_5\sigma_\beta^2 + r_6\sigma_\alpha^2$
Factor B	$b-1$	SS_B	MS_B	$\sigma_e^2 + r_3\sigma_\beta^2 + r_4\sigma_\alpha^2$
Error	$N-a-b+1$	SS_E	MS_E	$\sigma_e^2 + r_1\sigma_\beta^2 + r_2\sigma_\alpha^2$

The sums of squares in Table 12.1 are defined as follows:

$$SS_A = \sum_{i=1}^{a} n_{i.}(\bar{y}_{i..} - \bar{y}_{...})^2 = \sum_{i=1}^{a} \frac{y_{i..}^2}{n_{i.}} - \frac{y_{...}^2}{N},$$

$$SS_B = \sum_{j=1}^{b} n_{.j}(\bar{y}_{.j.} - \bar{y}_{...})^2 = \sum_{j=1}^{b} \frac{y_{.j.}^2}{n_{.j}} - \frac{y_{...}^2}{N},$$

and (12.2.1)

$$\begin{aligned} SS_E &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 \\ &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} y_{ijk}^2 - \sum_{i=1}^{a} \frac{y_{i..}^2}{n_{i.}} - \sum_{j=1}^{b} \frac{y_{.j.}^2}{n_{.j}} + \frac{y_{...}^2}{N}, \end{aligned}$$

where

$$y_{ij.} = \sum_{k=1}^{n_{ij}} y_{ijk}, \qquad \bar{y}_{ij.} = \frac{y_{ij.}}{n_{ij}},$$

$$y_{i..} = \sum_{j=1}^{b} y_{ij.}, \qquad \bar{y}_{i..} = \frac{y_{i..}}{n_{i.}},$$

$$y_{.j.} = \sum_{i=1}^{a} y_{ij.}, \qquad \bar{y}_{.j.} = \frac{y_{.j.}}{n_{.j}},$$

and

$$y_{...} = \sum_{i=1}^{a} y_{i..} = \sum_{j=1}^{b} y_{.j.}, \qquad \bar{y}_{...} = \frac{y_{...}}{N},$$

with

$$n_{i.} = \sum_{j=1}^{b} n_{ij}, \qquad n_{.j} = \sum_{i=1}^{a} n_{ij},$$

and

$$N = \sum_{i=1}^{a} n_{i.} = \sum_{j=1}^{b} n_{.j} = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij}.$$

The SS_A, SS_B, and SS_E terms in (12.2.1) have been defined by establishing an analogy with the corresponding terms for the balanced case.

Define the uncorrected sums of squares as

$$T_A = \sum_{i=1}^{a} \frac{y_{i..}^2}{n_{i.}}, \qquad T_B = \sum_{j=1}^{b} \frac{y_{.j.}^2}{n_{.j}},$$

$$T_{AB} = \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{y_{ij.}^2}{n_{ij}}, \qquad T_0 = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n_{ij}} y_{ijk}^2,$$

and

$$T_\mu = \frac{y_{...}^2}{N}.$$

Then the corrected sums of squares defined in (12.2.1) can be written as

$$SS_A = T_A - T_\mu,$$
$$SS_B = T_B - T_\mu,$$

and

$$SS_E = T_0 - T_A - T_B + T_\mu.$$

It should be pointed out that the expressions in (12.2.1) have been defined solely by an analogy with the analysis of variance for balanced data. In general, not all such analogous expressions are sums of squares. For example, SS_E of (12.2.1) can be negative and so it is not a sum of squares. We might therefore refer to the terms in (12.2.1) and their counterparts in other unbalanced models as analogous sums of squares. The mean squares as usual are obtained by dividing the sums of squares by the corresponding degrees of freedom. The results on expected mean squares are outlined in the following section.

12.3 EXPECTED MEAN SQUARES

The expected sums of squares or mean squares are readily obtained by first calculating the expected values of the quantities T_0, T_A, T_B, and T_μ. First, note that by the assumptions of the model in (12.1.1),

$$E(\alpha_i) = 0, \qquad E(\alpha_i^2) = \sigma_\alpha^2, \quad \text{and} \quad E(\alpha_i \alpha_{i'}) = 0, \quad i \neq i',$$

with similar results for the β_js and e_{ijk}s. Also, all covariances between pairs of nonidentical random variables are zero. Now, the following results are readily derived:

$$E(T_0) = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} E(\mu + \alpha_i + \beta_j + e_{ijk})^2 = N(\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_e^2),$$

$$\begin{aligned} E(T_A) &= \sum_{i=1}^{a} \frac{1}{n_{i.}} E\left(n_{i.}\mu + n_{i.}\alpha_i + \sum_{j=1}^{b} n_{ij}\beta_j + \sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} e_{ijk}\right)^2 \\ &= N\mu^2 + N\sigma_\alpha^2 + k_3\sigma_\beta^2 + a\sigma_e^2; \end{aligned}$$

$$\begin{aligned} E(T_B) &= \sum_{j=1}^{b} \frac{1}{n_{.j}} E\left(n_{.j}\mu + \sum_{i=1}^{a} n_{ij}\alpha_i + n_{.j}\beta_j + \sum_{i=1}^{a}\sum_{k=1}^{n_{ij}} e_{ijk}\right)^2 \\ &= N\mu^2 + N\sigma_\beta^2 + k_4\sigma_\alpha^2 + b\sigma_e^2; \end{aligned}$$

and

$$\begin{aligned} E(T_\mu) &= \frac{1}{N} E\left(N\mu + \sum_{i=1}^{a} n_{i.}\alpha_i + \sum_{j=1}^{b} n_{.j}\beta_j + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} e_{ijk}\right)^2 \\ &= N\mu^2 + k_1\sigma_\alpha^2 + k_2\sigma_\beta^2 + \sigma_e^2; \end{aligned}$$

where

$$k_1 = \frac{1}{N}\sum_{i=1}^{a} n_{i.}^2, \qquad k_2 = \frac{1}{N}\sum_{j=1}^{b} n_{.j}^2,$$

$$k_3 = \sum_{i=1}^{a} \frac{\sum_{j=1}^{b} n_{ij}^2}{n_{i.}}, \quad \text{and} \quad k_4 = \sum_{j=1}^{b} \frac{\sum_{i=1}^{a} n_{ij}^2}{n_{.j}}.$$

Hence, expected sums of squares and mean squares are given as follows:

$$\begin{aligned} E(\text{SS}_E) &= E(T_0 - T_A - T_B + T_\mu) \\ &= (N - a - b + 1)\sigma_e^2 + (k_2 - k_3)\sigma_\beta^2 + (k_1 - k_4)\sigma_\alpha^2, \\ E(\text{MS}_E) &= \frac{1}{N - a - b + 1} E(\text{SS}_E) = \sigma_e^2 + r_1\sigma_\beta^2 + r_2\sigma_\alpha^2; \\ E(\text{SS}_B) &= E(T_B - T_\mu) = (b - 1)\sigma_e^2 + (N - k_2)\sigma_\beta^2 + (k_4 - k_1)\sigma_\alpha^2, \\ E(\text{MS}_B) &= \frac{1}{b - 1} E(\text{SS}_B) = \sigma_e^2 + r_3\sigma_\beta^2 + r_4\sigma_\alpha^2; \end{aligned}$$

$$E(\mathrm{SS}_A) = E(T_A - T_\mu) = (a-1)\sigma_e^2 + (k_3 - k_2)\sigma_\beta^2 + (N - k_1)\sigma_\alpha^2,$$

and

$$E(\mathrm{MS}_A) = \frac{1}{a-1}E(\mathrm{SS}_A) = \sigma_e^2 + r_5\sigma_\beta^2 + r_6\sigma_\alpha^2,$$

where

$$\begin{aligned} r_1 &= \frac{k_2 - k_3}{N-a-b-1}, & r_2 &= \frac{k_1 - k_4}{N-a-b+1}, \\ r_3 &= \frac{N - k_2}{b-1}, & r_4 &= \frac{k_4 - k_1}{b-1}, \\ r_5 &= \frac{k_3 - k_2}{a-1}, \quad \text{and} & r_6 &= \frac{N - k_1}{a-1}. \end{aligned}$$

12.4 ESTIMATION OF VARIANCE COMPONENTS

In this section, we consider some results on estimation of variance components.

12.4.1 ANALYSIS OF VARIANCE ESTIMATORS

The analysis of variance or Henderson's Method I for estimating variance components is to equate the sums of squares or mean squares in Table 12.1 to their respective expected values. The resulting equations are

$$\begin{aligned} \mathrm{SS}_A &= (N-k_1)\sigma_\alpha^2 + (k_3 - k_2)\sigma_\beta^2 + (a-1)\sigma_e^2, \\ \mathrm{SS}_B &= (k_4 - k_1)\sigma_\alpha^2 + (N - k_2)\sigma_\beta^2 + (b-1)\sigma_e^2, \end{aligned} \tag{12.4.1}$$

and

$$\mathrm{SS}_E = (k_1 - k_4)\sigma_\alpha^2 + (k_2 - k_3)\sigma_\beta^2 + (N-a-b+1)\sigma_e^2.$$

The variance component estimators are obtained by solving the equations in (12.4.1) for σ_α^2, σ_β^2, and σ_e^2. The estimators thus obtained are given by

$$\begin{bmatrix} \hat{\sigma}^2_{\alpha,\mathrm{ANOV}} \\ \hat{\sigma}^2_{\beta,\mathrm{ANOV}} \\ \hat{\sigma}^2_{e,\mathrm{ANOV}} \end{bmatrix} = \begin{bmatrix} N-k_1 & k_3-k_2 & a-1 \\ k_4-k_1 & N-k_2 & b-1 \\ k_1-k_4 & k_2-k_3 & N-a-b+1 \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{SS}_A \\ \mathrm{SS}_B \\ \mathrm{SS}_E \end{bmatrix}. \tag{12.4.2}$$

Further simplification of (12.4.2) yields (see, e.g., Searle, 1958; 1971, p. 487)

$$\hat{\sigma}^2_{e,\mathrm{ANOV}} = \frac{\theta_1(\mathrm{SS}_E + \mathrm{SS}_A) + \theta_2(\mathrm{SS}_E + \mathrm{SS}_B) - (\mathrm{SS}_E + \mathrm{SS}_B + \mathrm{SS}_A)}{\theta_1(N-b) + \theta_2(N-a) - (N-1)},$$

$$\hat{\sigma}^2_{\beta,\text{ANOV}} = \frac{\text{SS}_E + \text{SS}_B - (N-a)\hat{\sigma}^2_{e,\text{ANOV}}}{N-k_3}, \tag{12.4.3}$$

and

$$\hat{\sigma}^2_{\alpha,\text{ANOV}} = \frac{\text{SS}_E + \text{SS}_A - (N-b)\hat{\sigma}^2_{e,\text{ANOV}}}{N-k_4},$$

where

$$\theta_1 = \frac{N-k_1}{N-k_4} \quad \text{and} \quad \theta_2 = \frac{N-k_2}{N-k_3}.$$

12.4.2 FITTING-CONSTANTS-METHOD ESTIMATORS

Let $R(\mu, \alpha, \beta)$ be the reduction in sum of squares due to fitting the fixed version of the model in (12.1.1) and let $R(\mu, \alpha)$, $R(\mu, \beta)$, and $R(\mu)$ be the reductions in sums of squares due to fitting the submodels

$$\begin{aligned} y_{ijk} &= \mu + \alpha_i + e_{ijk}, \\ y_{ijk} &= \mu + \beta_j + e_{ijk}, \end{aligned} \tag{12.4.4}$$

and

$$y_{ijk} = \mu + e_{ijk},$$

respectively. Then it can be shown that (see, Searle, 1971, pp. 292–293; 1987, pp. 124–125)

$$\begin{aligned} R(\mu, \alpha, \beta) &= T_A + \boldsymbol{r}'\boldsymbol{C}^{-1}\boldsymbol{r}, \\ R(\mu, \alpha) &= T_A, \\ R(\mu, \beta) &= T_B, \end{aligned} \tag{12.4.5}$$

and

$$R(\mu) = T_\mu,$$

where[1]

$$\boldsymbol{C} = \{c_{jj'}\} \quad \text{for } j, j' = 1, 2, \ldots, b-1,$$

with

$$c_{jj} = n_{.j} - \sum_{i=1}^{a} \frac{n_{ij}^2}{n_{i.}},$$

[1] For a numerical example illustrating the computation of the elements of matrix $\boldsymbol{C}$, see Searle and Henderson (1961).

TABLE 12.2 Analysis of variance based on α adjusted for β.

Source of variation	Degrees of freedom	Sum of squares
Mean μ	1	$R(\mu)$
β **adjusted for** μ	$b-1$	$R(\beta\|\mu)$
α **adjusted for** μ **and** β	$a-1$	$R(\alpha\|\mu,\beta)$
Error	$N-a-b+1$	SS_E

$$c_{jj'} = -\sum_{i=1}^{a} \frac{n_{ij}n_{ij'}}{n_{i.}}, \quad j \neq j' \quad \left(\sum_{j'=1}^{b} c_{jj'} = 0 \quad \text{for all } j\right),$$

and

$$\boldsymbol{r} = \{r_j\} = \left\{y_{.j.} - \sum_{i=1}^{a} n_{ij}\bar{y}_{i..}\right\} \quad \text{for } j = 1, 2, \ldots, b-1 \quad \left(\sum_{j=1}^{b} r_j = 0\right).$$

The analysis of variance based on α adjusted for β (fitting β before α) is given in Table 12.2. From Table 12.2, the terms (quadratics) needed in the fitting-constants-method of estimating variance components are

$$\begin{aligned} R(\mu) &= T_\mu, \\ R(\beta|\mu) &= R(\mu,\beta) - R(\mu) = T_B - T_\mu, \\ R(\alpha|\mu,\beta) &= R(\mu,\alpha,\beta) - R(\mu,\beta) = R(\mu,\alpha,\beta) - T_B, \end{aligned} \tag{12.4.6}$$

and

$$\mathrm{SS}_E = R(0) - R(\mu,\alpha,\beta) = T_0 - R(\mu,\alpha,\beta).$$

Remarks:

(i) The quadratics in (12.4.6) lead to the following partitioning of the total sum of squares (uncorrected for the mean):

$$\mathrm{SS}_T = R(\mu) + R(\beta|\mu) + R(\alpha|\mu,\beta) + \mathrm{SS}_E.$$

(ii) The quadratics in (12.4.6) are equivalent to SAS Type I sums of squares when ordering the factors as B, A. ♦

The expected values of the sums of squares in Table 12.2 are (see, e.g., Searle, 1958; Low, 1964, 1976; Searle et al. 1992, pp. 209–210)

$$E\{\mathrm{SS}_E\} = (N-a-b+1)\sigma_e^2,$$

$$E\{R(\alpha|\mu, \beta)\} = (N - k_4)\sigma_\alpha^2 + (a - 1)\sigma_e^2, \tag{12.4.7}$$

and

$$E\{R(\beta|\mu)\} = (N - k_2)\sigma_\beta^2 + (k_4 - k_1)\sigma_\alpha^2 + (b - 1)\sigma_e^2.$$

The variance components estimators are obtained by equating the sums of squares in Table 12.2 to their respective expected values given in (12.4.7). The resulting estimators are

$$\begin{aligned} \sigma_{e,\text{FTC1}}^2 &= \frac{\text{SS}_E}{N - a - b + 1}, \\ \hat{\sigma}_{\alpha,\text{FTC1}}^2 &= \frac{R(\alpha|\mu, \beta) - (a - 1)\hat{\sigma}_{e,\text{FTC1}}^2}{N - k_4}, \end{aligned} \tag{12.4.8}$$

and

$$\hat{\sigma}_{\beta,\text{FTC1}}^2 = \frac{R(\beta|\mu) - (k_4 - k_1)\hat{\sigma}_{\alpha,\text{FTC1}}^2 - (b - 1)\hat{\sigma}_{e,\text{FTC1}}^2}{N - k_2}.$$

The analysis of variance Table 12.2 carries with it a sequential concept of first fitting μ, then μ and β, and then μ, β, and α. Because of the symmetry of the crossed classification model in (12.1.1), an alternative approach for the analysis of variance would be to consider the following sums of squares:

$$\begin{aligned} R(\mu) &= T(\mu), \\ R(\alpha|\mu) &= R(\mu, \alpha) - R(\mu) = T_A - T_\mu, \\ R(\beta|\mu, \alpha) &= R(\mu, \alpha, \beta) - R(\mu, \alpha) = R(\mu, \alpha, \beta) - T_A, \end{aligned} \tag{12.4.9}$$

and

$$\text{SS}_E = R(0) - R(\mu, \alpha, \beta) = T_0 - R(\mu, \alpha, \beta).$$

The resulting analysis of variance is given in Table 12.3.

Remarks:

(i) The quadratics in (12.4.9) lead to the following partitioning of the total sum of squares (uncorrected for the mean):

$$\text{SS}_T = R(\mu) + R(\alpha|\mu) + R(\beta|\mu, \alpha) + \text{SS}_E.$$

(ii) The quadratics in (12.4.9) are equivalent to SAS Type I sums of squares when ordering the factors as A, B. ♦

TABLE 12.3 Analysis of variance based on β adjusted for α.

Source of variation	Degrees of freedom	Sum of squares
Mean μ	1	$R(\mu)$
α adjusted for μ	$b-1$	$R(\alpha\|\mu)$
β adjusted for μ and α	$a-1$	$R(\beta\|\mu,\alpha)$
Error	$N-a-b+1$	SS_E

From symmetry the results on expected sums of squares in Table 12.3 are easily obtained from the results in (12.4.7) and are given by

$$\begin{aligned} E\{\mathrm{SS}_E\} &= (N-a-b+1)\sigma_e^2, \\ E\{R(\beta|\mu,\alpha)\} &= (N-k_3)\sigma_\beta^2 + (b-1)\sigma_e^2, \end{aligned} \tag{12.4.10}$$

and

$$E[R(\alpha|\mu)] = (N-k_1)\sigma_\alpha^2 + (k_3-k_2)\sigma_\beta^2 + (a-1)\sigma_e^2.$$

The variance component estimators are obtained by equating the sums of squares in Table 12.3 to their respective expected values given in (12.4.10). The resulting estimators of the variance components are

$$\begin{aligned} \hat{\sigma}^2_{e,\mathrm{FTC2}} &= \frac{\mathrm{SS}_E}{N-a-b+1}, \\ \hat{\sigma}^2_{\beta,\mathrm{FTC2}} &= \frac{R(\beta|\mu,\alpha) - (b-1)\hat{\sigma}^2_{e,\mathrm{FTC2}}}{N-k_3}, \end{aligned} \tag{12.4.11}$$

and

$$\hat{\sigma}^2_{\alpha,\mathrm{FTC2}} = \frac{R(\alpha|\mu) - (k_3-k_2)\hat{\sigma}^2_{\beta,\mathrm{FTC2}} - (a-1)\hat{\sigma}^2_{e,\mathrm{FTC2}}}{N-k_1}.$$

Inasmuch as the variances of the estimators based on the "adjusted" quadratics contain only σ_e^2, design constants, and the parameter in question, they are often used to obtain estimators. Such quadratics and their expectations are

$$\begin{aligned} E\{\mathrm{SS}_E\} &= (N-a-b+1)\sigma_e^2, \\ E\{R(\alpha|\mu,\beta)\} &= (N-k_4)\sigma_\alpha^2 + (a-1)\sigma_e^2, \end{aligned} \tag{12.4.12}$$

and

$$E\{R(\beta|\mu,\alpha)\} = (N-k_3)\sigma_\beta^2 + (b-1)\sigma_e^2.$$

The resulting estimators are then given by

$$\hat{\sigma}^2_{e,\text{FTC3}} = \frac{\text{SS}_E}{N-a-b+1},$$
$$\hat{\sigma}^2_{\alpha,\text{FTC3}} = \frac{R(\alpha|\mu,\beta) - (a-1)\hat{\sigma}^2_{e,\text{FTC3}}}{N-k_4}, \tag{12.4.13}$$

and

$$\hat{\sigma}^2_{\beta,\text{FTC3}} = \frac{R(\beta|\mu,\alpha) - (b-1)\hat{\sigma}^2_{e,\text{FTC3}}}{N-k_3}.$$

Remarks:

(i) The quadratics in (12.4.12) do not lead to the following partitioning of the total sum of squares (corrected for the mean):

$$\text{SS}_T = R(\mu) + R(\beta|\mu,\alpha) + R(\alpha|\mu,\beta) + \text{SS}_E.$$

(ii) The quadratics in (12.4.12) are equivalent to SAS Type II sums of squares. ♦

12.4.3 ANALYSIS OF MEANS ESTIMATORS

As indicated in Section 10.4, the approach of the analysis of means method, when all $n_{ij} \geq 1$, is to treat the means of those cells as observations and then carry out a balanced data analysis. The calculations for the analysis are rather straightforward as illustrated below. We first discuss the unweighted analysis and then the weighted analysis.

12.4.3.1 Unweighted Means Analysis

For the observations y_{ijk}s from the model in (12.1.1), let x_{ij} be the cell mean defined by

$$x_{ij} = \bar{y}_{ij.} = \sum_{k=1}^{n_{ij}} \frac{y_{ijk}}{n_{ij}}. \tag{12.4.14}$$

Further define

$$\bar{x}_{i.} = \frac{\sum_{j=1}^{b} x_{ij}}{b}, \qquad \bar{x}_{.j} = \frac{\sum_{i=1}^{a} x_{ij}}{a},$$

and

$$\bar{x}_{..} = \frac{\sum_{i=1}^{a}\sum_{j=i}^{b} x_{ij}}{ab}.$$

TABLE 12.4 Analysis of variance with unweighted sums of squares for the model in (12.1.1).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Factor A	$a-1$	SS_{Au}	MS_{Au}	$\sigma_e^2 + b\bar{n}_h\sigma_\alpha^2$
Factor B	$b-1$	SS_{Bu}	MS_{Bu}	$\sigma_e^2 + a\bar{n}_h\sigma_\beta^2$
Error	$N-a-b+1$	SS_{Eu}	MS_{Eu}	σ_e^2

Then the analysis of variance for the unweighted means analysis is shown in Table 12.4.

The quantities in the sum of squares column are defined by

$$SS_{Au} = b\bar{n}_h \sum_{i=1}^{a} (\bar{x}_{i.} - \bar{x}_{..})^2,$$

$$SS_{Bu} = a\bar{n}_h \sum_{j=1}^{b} (\bar{x}_{.j} - \bar{x}_{..})^2, \tag{12.4.15}$$

and

$$SS_{Eu} = \bar{n}_h \sum_{i=1}^{a}\sum_{j=1}^{b} (x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij.})^2,$$

where

$$\bar{n}_h = \frac{1}{\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}^{-1}/ab}.$$

The mean squares are obtained in the usual way by dividing the sums of squares by the corresponding degrees of freedom. For a method of derivation of the results on expected mean squares, see Hirotsu (1966) and Mostafa (1967).

The following features of the above analysis are worth noting:

(i) The means of the x_{ij}s are calculated in the usual manner, i.e., $\bar{x}_{i.} = \sum_{j=1}^{b} x_{ij}/b$, $\bar{x}_{.j} = \sum_{i=1}^{a} x_{ij}/a$, and $\bar{x}_{..} = \sum_{i=1}^{a}\sum_{j=1}^{b} x_{ij}/ab$.

(ii) The error sum of squares, SS_{Eu}, is calculated by pooling the interaction and error sums of squares in the unweighted means analysis of the two-way random model with interaction (see Section 13.4.3.1).

(iii) The individual sums of squares do not add up to the total sum of squares.

(iv) The sums of squares SS_{Au}, SS_{Bu}, and SS_{Eu} do not have a scaled chi-square distribution, as in the case of the balanced analogue of the model in (12.1.1); nor are SS_{Au} and SS_{Bu} in general independent of SS_{Eu}.

The estimators of the variance components, as usual, are obtained by equating the means squares to their respective expected values and solving the resulting equations for the variance components. The resulting estimators are given as follows:

$$\hat{\sigma}^2_{e,\mathrm{UME}} = \mathrm{MS}_E,$$
$$\hat{\sigma}^2_{\beta,\mathrm{UME}} = \frac{\mathrm{MS}_{Bu} - \mathrm{MS}_E}{a\bar{n}_h}, \tag{12.4.16}$$

and

$$\hat{\sigma}^2_{\alpha,\mathrm{UME}} = \frac{\mathrm{MS}_{Au} - \mathrm{MS}_E}{b\bar{n}_h}.$$

12.4.3.2 Weighted Means Analysis

The weighted square of means analysis consists of weighting the terms in the sums of squares SS_{Au} and SS_{Bu}, defined in (12.4.15) in the unweighted means analysis, in inverse proportion to the variance of the term concerned. Thus, instead of SS_A and SS_B given by

$$\mathrm{SS}_{Au} = b\bar{n}_h \sum_{i=1}^{a} (\bar{x}_{i.} - \bar{x}_{..})^2, \qquad \mathrm{SS}_{Bu} = a\bar{n}_h \sum_{j=1}^{b} (\bar{x}_{.j} - \bar{x}_{..})^2,$$

we use

$$\mathrm{SS}_{Aw} = \sum_{i=1}^{a} w_i (\bar{x}_{i.} - \bar{x}^w_{..})^2, \qquad \mathrm{SS}_{Bw} = \sum_{j=1}^{b} \nu_j (\bar{x}_{.j} - \bar{x}^\nu_{..})^2,$$

where

$$w_i = \sigma^2/\operatorname{var}(\bar{x}_{i.}), \qquad \nu_j = \sigma^2/\operatorname{var}(\bar{x}_{.j})$$

and $\bar{x}^w_{..}$ and $\bar{x}^\nu_{..}$ are weighted means of $\bar{x}_{i.}$s and $\bar{x}_{.j}$s weighted by w_i and v_j, respectively; i.e.,

$$\bar{x}^w_{..} = \sum_{i=1}^{a} w_i \bar{x}_{i.} / \sum_{i=1}^{a} w_i, \qquad \bar{x}^\nu_{..} = \sum_{j=1}^{b} \nu_j \bar{x}_{.j} / \sum_{j=1}^{b} v_j.$$

There are a variety of weights that can be used for w_i and ν_j as discussed by Gosslee and Lucas (1965). A weighted analysis of variance based on weights

TABLE 12.5 Analysis of variance with weighted sums of squares for the model in (12.1.1).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Factor A	$a-1$	SS_{Aw}	MS_{Aw}	$\sigma_e^2 + b\theta_1\sigma_\alpha^2$
Factor B	$b-1$	SS_{Bw}	MS_{Bw}	$\sigma_e^2 + a\theta_2\sigma_\beta^2$
Error	$N-a-b+1$	SS_E	MS_E	σ_e^2

originally proposed by Yates (1934) (for fixed effects model) is shown in Table 12.5. (See also Searle et al. (1992, pp. 220–221).) It is calculated by the SAS® GLM or SPSS® GLM procedures using Type III sums of squares.

The quantities in the sum of squares column are given by

$$\text{SS}_{Aw} = \sum_{i=1}^{a} \phi_i(\bar{x}_{i.} - \bar{x}_{..}^{\phi})^2,$$

$$\text{SS}_{Bw} = \sum_{j=1}^{b} \psi_j(\bar{x}_{.j} - \bar{x}_{..}^{\psi})^2, \tag{12.4.17}$$

and

$$\text{SS}_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} y_{ijk}^2 - R(\mu, \alpha, \beta),$$

where

$$\bar{x}_{..}^{\phi} = \sum_{i=1}^{a} \phi_i \bar{x}_{i.} \Big/ \sum_{i=1}^{a} \phi_i,$$

$$\bar{x}_{..}^{\psi} = \sum_{j=1}^{b} \psi_j \bar{x}_{.j} \Big/ \sum_{j=1}^{b} \psi_j,$$

$$\phi_i = b^2 \Big/ \sum_{j=1}^{b} n_{ij}^{-1}, \qquad \psi_j = a^2 \Big/ \sum_{i=1}^{a} n_{ij}^{-1},$$

and $R(\mu, \alpha, \beta)$ is defined in (12.4.5). The quantities θ_1 and θ_2 in the expected mean square column are defined as

$$\theta_1 = \left\{ \sum_{i=1}^{a} \phi_i - \sum_{i=1}^{a} \phi_i^2 / \sum_{i=1}^{a} \phi_i \right\} / b(a-1)$$

and (12.4.18)

$$\theta_2 = \left\{ \sum_{j=1}^{b} \psi_j - \sum_{j=1}^{b} \psi_j^2 / \sum_{j=1}^{b} \psi_j \right\} / a(b-1).$$

The estimators of the variance components obtained using the weighted analysis are

$$\hat{\sigma}^2_{e,\text{WME}} = \text{MS}_E,$$
$$\hat{\sigma}^2_{\beta,\text{WME}} = \frac{\text{MS}_{Bw} - \text{MS}_E}{a\theta_2}, \tag{12.4.19}$$

and

$$\hat{\sigma}^2_{\alpha,\text{WME}} = \frac{\text{MS}_{Aw} - \text{MS}_E}{b\theta_1}.$$

12.4.4 SYMMETRIC SUMS ESTIMATORS

We consider symmetric sums estimators for the special case when $n_{ij} = 0$ or 1. For this case, the model equation (12.1.1) becomes

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}, \quad i = 1, 2, \ldots, a; \quad j = 1, 2, \ldots, b. \tag{12.4.20}$$

From (12.4.20) the expected values of the products of the observations are

$$E(y_{ij} y_{i'j'}) = \begin{cases} \mu^2, & i \neq i', \ \ j \neq j', \\ \mu^2 + \sigma_\beta^2, & i \neq i', \ \ j = j', \\ \mu^2 + \sigma_\alpha^2, & i = i', \ \ j \neq j', \\ \mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_e^2, & i = i', \ \ j = j', \end{cases} \tag{12.4.21}$$

where $i, i' = 1, 2, \ldots, a$; $j, j' = 1, 2, \ldots, b$, provided $y_{ij} y_{i'j'}$ is defined. Now, the normalized symmetric sums of the terms in (12.4.21) are

$$g_m = \frac{\sum_{\substack{i,i' \\ i \neq i'}} \sum_{\substack{j,j' \\ j \neq j'}} y_{ij} y_{i'j'}}{N^2 - \sum_{i=1}^a n_{i.}^2 - \sum_{j=1}^b n_{.j}^2 + N}$$
$$= \frac{y_{..}^2 - \sum_{i=1}^a y_{i.}^2 - \sum_{j=1}^b y_{.j}^2 + \sum_{i=1}^a \sum_{j=1}^b y_{ij}^2}{N^2 - k_1 - k_2 + N},$$

$$g_B = \frac{\sum_{\substack{i,i' \\ i \neq i'}} \sum_{j=1}^b y_{ij} y_{i'j}}{\sum_{j=1}^b n_{.j}^2 - N} = \frac{\sum_{j=1}^b y_{.j}^2 - \sum_{i=1}^a \sum_{j=1}^b y_{ij}^2}{k_2 - N},$$

$$g_A = \frac{\sum_{i=1}^{a} \sum_{\substack{j,j' \\ j \neq j'}} y_{ij} y_{ij'}}{\sum_{i=1}^{a} n_{i.}^2 - N} = \frac{\sum_{i=1}^{a} y_{i.}^2 - \sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij}^2}{k_1 - N},$$

and

$$g_E = \frac{\sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij} y_{ij}}{\sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij}} = \frac{\sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij}^2}{N},$$

where

$$n_{ij} = \begin{cases} 1 & \text{if an observation appears in the } (i, j)\text{th cell,} \\ 0 & \text{otherwise,} \end{cases}$$

$$n_{i.} = \sum_{j=1}^{b} n_{ij}, \qquad n_{.j} = \sum_{i=1}^{a} n_{ij}, \qquad N = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij},$$

$$k_1 = \sum_{i=1}^{a} n_{i.}^2, \qquad k_2 = \sum_{j=1}^{b} n_{.j}^2.$$

Equating g_m, g_B, g_A, and g_E to their respective expected values, we obtain

$$\begin{aligned} \mu^2 &= g_m, \\ \mu^2 + \sigma_\beta^2 &= g_B, \\ \mu^2 + \sigma_\alpha^2 &= g_A, \end{aligned} \tag{12.4.22}$$

and

$$\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 = g_E.$$

The variance component estimators obtained by solving the equations in (12.4.22) are (Koch, 1967)

$$\begin{aligned} \hat{\sigma}_{\alpha,\text{SSP}}^2 &= g_A - g_m, \\ \hat{\sigma}_{\beta,\text{SSP}}^2 &= g_B - g_m, \end{aligned} \tag{12.4.23}$$

and

$$\hat{\sigma}_{e,\text{SSP}}^2 = g_E - g_A - g_B + g_m.$$

The estimators in (12.4.23), by construction, are unbiased; and they reduce to the analysis of variance estimators in the case of balanced data. However, they are not translation invariant, i.e., they may change in values if the same constant is added to all the observations and their variances are functions of

μ. This drawback is overcome by using the symmetric sums of squares of differences rather than the products.

For symmetric sums based on the expected values of the squares of differences of the observations, we have

$$E\{(y_{ij} - y_{i'j'})^2\} = \begin{cases} 2(\sigma_e^2 + \sigma_\beta^2), & i = i', \quad j \neq j', \\ 2(\sigma_e^2 + \sigma_\alpha^2), & i \neq i', \quad j = j', \\ 2(\sigma_e^2 + \sigma_\alpha^2 + \sigma_\beta^2), & i \neq i', \quad j \neq j'. \end{cases} \tag{12.4.24}$$

The results in (12.4.24), of course, only hold for those cases where the observations in both cells (i, j) and (i', j') exist. The normalized symmetric sums of the terms in (12.4.24) are

$$\begin{aligned} h_A &= \frac{1}{\sum_{j=1}^b n_{.j}(n_{.j} - 1)} \sum_{\substack{i,i' \\ i \neq i'}}^{a} \sum_{j=1}^{b} (y_{ij} - y_{i'j})^2 \\ &= \frac{2}{k_2 - N} \sum_{j=1}^{b} \left(n_{.j} \sum_{i=1}^{a} y_{ij}^2 - y_{.j}^2 \right), \\ h_B &= \frac{1}{\sum_{i=1}^a n_{i.}(n_{i.} - 1)} \sum_{i=1}^{a} \sum_{\substack{j,j' \\ j \neq j'}}^{b} (y_{ij} - y_{ij'})^2 \\ &= \frac{2}{k_1 - N} \sum_{i=1}^{a} \left(n_{i.} \sum_{j=1}^{b} y_{ij}^2 - y_{i.}^2 \right), \end{aligned}$$

and

$$\begin{aligned} h_E &= \frac{1}{\sum_{i=1}^a \sum_{j=1}^b n_{ij}(N - n_{i.} - n_{.j} + n_{ij})} \sum_{\substack{i,i' \\ i \neq i'}}^{a} \sum_{\substack{j,j' \\ j \neq j'}}^{b} (y_{ij} - y_{i'j'})^2 \\ &= \frac{2}{N^2 - k_1 - k_2 + N} \sum_{i=1}^{a} \sum_{j=1}^{b} (n_{ij} - n_{i.} - n_{.j} + N) y_{ij}^2 - 2g_m, \end{aligned}$$

where n_{ij}, $n_{i.}$, $n_{.j}$, N, k_1, k_2 and g_m are defined as before.

Equating h_A, h_B, and h_E to their respective expected values, we obtain

$$\begin{aligned} 2(\sigma_e^2 + \sigma_\alpha^2) &= h_A, \\ 2(\sigma_e^2 + \sigma_\beta^2) &= h_B, \end{aligned} \tag{12.4.25}$$

and

$$2(\sigma_e^2 + \sigma_\alpha^2 + \sigma_\beta^2) = h_E.$$

The variance component estimators obtained by solving the equations in (12.4.25) are (Koch, 1968)

$$\hat{\sigma}^2_{\alpha,\text{SSS}} = \frac{h_E - h_B}{2},$$
$$\hat{\sigma}^2_{\beta,\text{SSS}} = \frac{h_E - h_A}{2}, \qquad (12.4.26)$$

and

$$\hat{\sigma}^2_{e,\text{SSS}} = \frac{h_A + h_B - h_E}{2}.$$

It can be readily seen that if the model in (12.1.1) is balanced, i.e., if $n_{ij} = 1$ for all (i, j), then the estimators (12.4.26) reduce to the usual analysis of variance estimators.

12.4.5 OTHER ESTIMATORS

The ML, REML, MINQUE, and MIVQUE estimators can be developed as special cases of the results for the general case considered in Chapter 10 and their special formulations for this model are not amenable to any simple algebraic expressions. With the advent of the high-speed digital computer, the general results on these estimators involving matrix operations can be handled with great speed and accuracy and their explicit algebraic evaluation for this model seems to be rather unnecessary. In addition, some commonly used statistical software packages, such as SAS®, SPSS®, and BMDP®, have special routines to compute these estimates rather conveniently simply by specifying the model in question.

12.4.6 A NUMERICAL EXAMPLE

Khuri and Littell (1987, p. 147) reported results of an experiment designed to study the variation in fusiform rust in Southern pine tree plantations, due to different families and test locations. The proportions of symptomatic trees from several plots for different families and test locations were recorded. The data given in Table 12.6 represent the results coming from a sample of five different families and four test locations. We will use the two-way unbalanced crossed model in (12.1.1) to analyze the data in Table 12.6. Here $a = 4$, $b = 5$; $i = 1, 2, \ldots, 4$ refer to the locations; and $j = 1, 2, \ldots, 5$ refer to the families. Further, σ^2_α and σ^2_β designate variance components due to location and family as factors, and σ^2_e denotes the error variance component. The calculations leading to the conventional analysis of variance based on Henderson's Method I were performed using the SAS®GLM procedure and the results are summarized in Table 12.7.[2]

[2]Since data are reported in terms of proportions, it would be more appropriate to analyze them using the arcsine transformation in order to stabilize the variance.

TABLE 12.6 Proportions of symptomatic trees from five families and four test locations.

Location	Family 1	2	3	4	5
1	0.804	0.734	0.967	0.917	0.850
	0.967	0.817	0.930		
	0.970	0.833	0.889		
		0.304			
2	0.867	0.407	0.896	0.952	0.486
	0.667	0.511	0.717		0.467
	0.793	0.274			
	0.458	0.428			
3	0.409	0.411	0.919	0.408	0.275
	0.569	0.646	0.669	0.435	0.256
	0.715	0.310	0.669	0.500	
	0.487		0.450		
4	0.587	0.304	0.928	0.367	0.525
	0.538	0.428	0.855		
	0.961		0.655		
	0.300		0.800		

Source: Khuri and Littell (1987); used with permission.

TABLE 12.7 Analysis of variance for the fusiform rust data of Table 12.6.

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Location	3	0.7404356	0.2468	$\sigma_e^2+0.238\sigma_\beta^2+13.182\sigma_\alpha^2$
Family	4	0.776807	0.1942	$\sigma_e^2+10.255\sigma_\beta^2+0.234\sigma_\alpha^2$
Error	45	1.256110	0.0279	$\sigma_e^2-0.016\sigma_\beta^2-0.028\sigma_\alpha^2$
Total	52	2.773273		

We now illustrate the calculations of point estimates of the variance components $\sigma_\alpha^2, \sigma_\beta^2, \sigma_e^2$, and certain of their parametric functions.

The analysis of variance (ANOVA) estimates in (12.4.3) based on Henderson's Method I are obtained as the solution to the following simultaneous equations:

$$\sigma_e^2 + 0.238\sigma_\beta^2 + 13.182\sigma_\alpha^2 = 0.2468,$$
$$\sigma_e^2 + 10.255\sigma_\beta^2 + 0.234\sigma_\alpha^2 = 0.1942,$$
$$\sigma_e^2 - 0.016\sigma_\beta^2 - 0.028\sigma_\alpha^2 = 0.0279.$$

Therefore, the desired ANOVA estimates of the variance components are given by

$$\begin{bmatrix} \hat{\sigma}^2_{e,\text{ANOV}} \\ \hat{\sigma}^2_{\beta,\text{ANOV}} \\ \hat{\sigma}^2_{\alpha,\text{ANOV}} \end{bmatrix} = \begin{bmatrix} 1 & 0.238 & 13.182 \\ 1 & 10.255 & 0.234 \\ 1 & -0.016 & -0.028 \end{bmatrix}^{-1} \begin{bmatrix} 0.2468 \\ 0.1942 \\ 0.0279 \end{bmatrix} = \begin{bmatrix} 0.0286 \\ 0.0158 \\ 0.0162 \end{bmatrix}.$$

These variance components account for 47.2%, 26.1%, and 26.7% of the total variation in the fusiform rust in this experiment.

To obtain variance component estimates based on fitting-constants-method estimators (12.4.8), (12.4.11), and (12.4.13), we calculated analysis of variance tables based on reductions in sums of squares due to fitting the submodels. The results are summarized in Tables 12.8, 12.9, and 12.10.

Now, the estimates in (12.4.8) based on Table 12.8 (location adjusted for family) are

$$\hat{\sigma}^2_{e,\text{FTC1}} = 0.025752,$$
$$\hat{\sigma}^2_{\alpha,\text{FTC1}} = \frac{0.279215 - 0.025752}{12.870} = 0.019694,$$

and

$$\hat{\sigma}^2_{\beta,\text{FTC1}} = \frac{0.194202 - 0.025752 - 0.234 \times 0.019694}{10.255} = 0.015977.$$

Similarly, the estimates in (12.4.11) based on Table 12.9 (family adjusted for location) are

$$\hat{\sigma}^2_{e,\text{FTC2}} = 0.025752,$$
$$\hat{\sigma}^2_{\beta,\text{FTC2}} = \frac{0.218524 - 0.025752}{10.076} = 0.019132,$$

and

$$\hat{\sigma}^2_{\alpha,\text{FTC2}} = \frac{0.246785 - 0.025752 - 0.238 \times 0.019132}{13.182} = 0.016422.$$

Finally, the estimates in (12.4.13) based on Table 12.10 (location adjusted for family and family adjusted for location) are

$$\hat{\sigma}^2_{e,\text{FTC3}} = 0.025752,$$
$$\hat{\sigma}^2_{\beta,\text{FTC3}} = \frac{0.218524 - 0.025752}{10.076} = 0.019132,$$

and

$$\hat{\sigma}^2_{\alpha,\text{FTC3}} = \frac{0.279215 - 0.025752}{12.870} = 0.019694.$$

TABLE 12.8 Analysis of variance for the fusiform rust data of Table 12.6 (location adjusted for family).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Location	4	0.776807	0.194202	$\sigma_e^2+10.255\sigma_\beta^2+0.234\sigma_\alpha^2$
Family	3	0.837645	0.279215	$\sigma_e^2+12.870\sigma_\alpha^2$
Error	45	1.158821	0.025752	σ_e^2
Total	52	2.773273		

TABLE 12.9 Analysis of variance for the fusiform rust data of Table 12.6 (family adjusted for location).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Location	3	0.740356	0.246785	$\sigma_e^2+0.238\sigma_\beta^2+13.182\sigma_\alpha^2$
Family	4	0.874096	0.218524	$\sigma_e^2+10.076\sigma_\beta^2$
Error	45	1.158821	0.025752	σ_e^2
Total	52	2.773273		

TABLE 12.10 Analysis of variance for the fusiform rust data of Table 12.6 (location adjusted for family and family adjusted for location).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Location	3	0.837645	0.279215	$\sigma_e^2+12.870\sigma_\alpha^2$
Family	4	0.874096	0.218524	$\sigma_e^2+10.076\sigma_\beta^2$
Error	45	1.158821	0.025752	σ_e^2
Total	52	2.773273		

For the analysis of means estimates in (12.4.16) and (12.4.19), we computed analysis of variance based on cell means using unweighted and weighted sums of squares and the results are summarized in Tables 12.11 and 12.12.

Now, the unweighted means estimates in (12.4.16) based on Table 12.11 are

$$\hat{\sigma}^2_{e,\text{UME}} = 0.02758,$$

$$\hat{\sigma}^2_{\beta,\text{UME}} = \frac{0.13884 - 0.02758}{7.932} = 0.01403,$$

TABLE 12.11 Analysis of variance for the fusiform rust data of Table 12.6 (unweighted sums of squares).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Location	3	0.84298	0.28099	$\sigma_e^2 + 9.915\sigma_\alpha^2$
Family	4	0.55537	0.13884	$\sigma_e^2 + 7.932\sigma_\beta^2$
Error	45	1.24097	0.02758	σ_e^2

TABLE 12.12 Analysis of variance for the fusiform rust data of Table 12.6 (weighted sums of squares).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Location	3	0.837645	0.279215	$\sigma_e^2 + 12.870\sigma_\alpha^2$
Family	4	0.874096	0.218524	$\sigma_e^2 + 10.076\sigma_\beta^2$
Error	45	1.158821	0.025752	σ_e^2

and

$$\hat{\sigma}^2_{\alpha,\text{UME}} = \frac{0.28099 - 0.02758}{9.915} = 0.02556.$$

Similarly, the weighted means estimates in (12.4.19) based on Table 12.12 are

$$\hat{\sigma}^2_{e,\text{WME}} = 0.025752,$$
$$\hat{\sigma}^2_{\beta,\text{WME}} = \frac{0.218524 - 0.025752}{10.076} = 0.019132,$$

and

$$\hat{\sigma}^2_{\alpha,\text{WME}} = \frac{0.279215 - 0.025752}{12.870} = 0.019694.$$

We used SAS® VARCOMP, SPSS® VARCOMP, and BMDP® 3V to estimate the variance components using the ML, REML, MINQUE(0), and MINQUE(1) procedures.[3] The desired estimates using these software are given in Table 12.13. Note that all three software produce nearly the same results except for some minor discrepancy in rounding decimal places.

[3]The computations for ML and REML estimates were also carried out using SAS® PROC MIXED and some other programs to assess their relative accuracy and convergence rate. There did not seem to be any appreciable differences between the results from different software.

TABLE 12.13 ML, REML, MINQUE(0), and MINQUE(1) estimates of the variance components using SAS®, SPSS®, and BMDP® software.

Variance component	SAS® ML	SAS® REML	SAS® MINQUE(0)
σ_e^2	0.025774	0.025736	0.027744
σ_β^2	0.015154	0.017258	0.016762
σ_α^2	0.015916	0.019505	0.016260

Variance component	SPSS® ML	SPSS® REML	SPSS® MINQUE(0)	SPSS® MINQUE(1)
σ_e^2	0.025719	0.025681	0.027700	0.025690
σ_β^2	0.015234	0.017343	0.016850	0.017164
σ_α^2	0.016020	0.019629	0.016361	0.019746

Variance component	BMDP® ML	BMDP® REML
σ_e^2	0.025774	0.025736
σ_β^2	0.015154	0.017258
σ_α^2	0.015916	0.019505

SAS® VARCOMP does not compute MINQUE(1). BMDP®3V does not compute MINQUE(0) and MINQUE(1).

12.5 VARIANCES OF ESTIMATORS

In this section, we present some results on sampling variances of estimators of variance components.

12.5.1 VARIANCES OF ANALYSIS OF VARIANCE ESTIMATORS

To find the variances and covariances of $\hat{\sigma}^2_{\alpha,\text{ANOV}}$, $\hat{\sigma}^2_{\beta,\text{ANOV}}$, and $\hat{\sigma}^2_{e,\text{ANOV}}$, we write the equations in (12.4.2) as

$$\hat{\boldsymbol{\sigma}}^2_{\text{ANOV}} = \boldsymbol{P}^{-1}\left[\boldsymbol{Ht} + \begin{bmatrix} 0 \\ 0 \\ T_0 - T_{AB} \end{bmatrix}\right], \qquad (12.5.1)$$

where

$$\hat{\boldsymbol{\sigma}}^2_{\text{ANOV}} = (\hat{\sigma}^2_{\alpha,\text{ANOV}}, \hat{\sigma}^2_{\beta,\text{ANOV}}, \hat{\sigma}^2_{e,\text{ANOV}}),$$
$$\boldsymbol{t}' = (T_A, T_B, T_{AB}, T_\mu),$$

$$\boldsymbol{P} = \begin{bmatrix} N-k_1 & k_3-k_2 & a-1 \\ k_4-k_1 & N-k_2 & b-1 \\ k_1-k_4 & k_2-k_3 & N-a-b+1 \end{bmatrix},$$

and

$$\boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix}.$$

When $n_{ij} = 0$ or 1, $T_{AB} = T_0$, so the equations in (12.5.1) are

$$\hat{\boldsymbol{\sigma}}^2_{\text{ANOV}} = \boldsymbol{P}^{-1}\boldsymbol{Ht}. \qquad (12.5.2)$$

From (12.5.2), the variance-covariance matrix of $\hat{\sigma}^2_{\text{ANOV}}$ is then given by

$$\text{Var}(\hat{\boldsymbol{\sigma}}^2_{\text{ANOV}}) = \boldsymbol{P}^{-1}\boldsymbol{H}\,\text{Var}(\boldsymbol{t})\boldsymbol{H}'\boldsymbol{P}^{-1\prime}. \qquad (12.5.3)$$

When $n_{ij} \geq 0$, T_{AB} exists although it is not used in the estimation of variance components. Further, it can be shown that $T_0 - T_{AB}$ is distributed independently of every element in $\boldsymbol{Ht}$ and has a scaled chi-square distribution with N-s degrees of freedom, where s is the number of nonempty cells. In this case,

$$\text{Var}(\hat{\sigma}^2_{\text{ANOV}}) = \boldsymbol{P}^{-1}\boldsymbol{H}\,\text{Var}(\boldsymbol{t})\boldsymbol{H}'\boldsymbol{P}^{-1\prime} + 2\boldsymbol{p}_3\boldsymbol{p}_3'\sigma_e^4(N-s), \qquad (12.5.4)$$

where $\boldsymbol{p}_3$ designates the third column of $\boldsymbol{P}^{-1}$ (Searle, 1958; 1971, p. 488).

Thus, to evaluate $\text{Var}(\hat{\sigma}^2_{\text{ANOV}})$ given by (12.5.3) or (12.5.4), we only need to find $\text{Var}(\boldsymbol{t})$, whose elements are variances and covariances of the uncorrected sums of squares T_A, T_B, T_{AB}, and T_μ. They have been obtained by Searle (1958) and are given as follows (see also Searle, 1971, pp. 487–488; Searle et al., 1992, pp. 439–440):

$$\begin{aligned}
\text{Var}(T_A) &= 2[k_1\sigma_\alpha^4 + (k_{21}+k_9)\sigma_\beta^4 + a\sigma_e^4 \\
&\quad + 2(k_{23}\sigma_\alpha^2\sigma_\beta^2 + N\sigma_\alpha^2\sigma_e^2 + k_3\sigma_\beta^2\sigma_e^2)], \\
\text{Var}(T_B) &= 2[(k_{22}+k_{10})\sigma_\alpha^4 + k_2\sigma_\beta^4 + b\sigma_e^4 \\
&\quad + 2(k_{23}\sigma_\alpha^2\sigma_\beta^2 + k_4\sigma_\alpha^2\sigma_e^2 + N\sigma_\beta^2\sigma_e^2)], \\
\text{Var}(T_{AB}) &= 2[k_1\sigma_\alpha^4 + k_2\sigma_\beta^4 + s\sigma_e^4 \\
&\quad + 2(k_{23}\sigma_\alpha^2\sigma_\beta^2 + N\sigma_\alpha^2\sigma_e^2 + N\sigma_\beta^2\sigma_e^2)], \\
\text{Var}(T_\mu) &= \frac{2}{N^2}[k_1^2\sigma_\alpha^4 + k_2^2\sigma_\beta^4 + N^2\sigma_e^4 \\
&\quad + 2(k_1k_2\sigma_\alpha^2\sigma_\beta^2 + Nk_1\sigma_\alpha^2\sigma_e^2 + Nk_2\sigma_\beta^2\sigma_e^2)], \\
\text{Cov}(T_A, T_B) &= 2[k_{18}\sigma_\alpha^4 + k_{17}\sigma_\beta^4 + k_{26}\sigma_e^4 \\
&\quad + 2(k_{23}\sigma_\alpha^2\sigma_\beta^2 + k_4\sigma_\alpha^2\sigma_e^2 + k_3\sigma_\beta^2\sigma_e^2)],
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(T_A, T_{AB}) &= 2[k_1\sigma_\alpha^4 + k_{17}\sigma_\beta^4 + a\sigma_e^4 \\
&\quad + 2(k_{23}\sigma_\alpha^2\sigma_\beta^2 + N\sigma_\alpha^2\sigma_e^2 + k_3\sigma_\beta^2\sigma_e^2)], \\
\mathrm{Cov}(T_A, T_\mu) &= \frac{2}{N}[k_5\sigma_\alpha^4 + k_{15}\sigma_\beta^4 + N\sigma_e^4 \\
&\quad + 2(k_{25}\sigma_\alpha^2\sigma_\beta^2 + k_1\sigma_\alpha^2\sigma_e^2 + k_2\sigma_\beta^2\sigma_e^2)], \\
\mathrm{Cov}(T_B, T_{AB}) &= 2[k_{18}\sigma_\alpha^4 + k_2\sigma_\beta^4 + b\sigma_e^4 \\
&\quad + 2(k_{23}\sigma_\alpha^2\sigma_\beta^2 + k_4\sigma_\alpha^2\sigma_e^2 + N\sigma_\beta^2\sigma_e^2)], \\
\mathrm{Cov}(T_B, T_\mu) &= \frac{2}{N}[k_{16}\sigma_\alpha^4 + k_6\sigma_\beta^4 + N\sigma_e^4 \\
&\quad + 2(k_{25}\sigma_\alpha^2\sigma_\beta^2 + k_1\sigma_\alpha^2\sigma_e^2 + k_2\sigma_\beta^2\sigma_e^2)],
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Cov}(T_{AB}, T_\mu) &= \frac{2}{N}[k_5\sigma_\alpha^4 + k_6\sigma_\beta^4 + N\sigma_e^4 \\
&\quad + 2(k_{25}\sigma_\alpha^2\sigma_\beta^2 + k_1\sigma_\alpha^2\sigma_e^2 + k_2\sigma_\beta^2\sigma_e^2)],
\end{aligned}$$

where[4]

$$k_1 = \sum_{i=1}^{a} n_{i.}^2, \qquad k_2 = \sum_{j=1}^{b} n_{.j}^2, \qquad k_3 = \sum_{i=1}^{a} \frac{\left(\sum_{j=1}^{b} n_{ij}^2\right)}{n_{i.}},$$

$$k_4 = \sum_{j=1}^{b} \frac{\left(\sum_{i=1}^{a} n_{ij}^2\right)}{n_{.j}}, \qquad k_5 = \sum_{i=1}^{a} n_{i.}^3, \qquad k_6 = \sum_{j=1}^{b} n_{.j}^3,$$

$$k_7 = \sum_{i=1}^{a} \frac{\left(\sum_{j=1}^{b} n_{ij}^2\right)^2}{n_{i.}}, \qquad k_8 = \sum_{j=1}^{b} \frac{\left(\sum_{i=1}^{a} n_{ij}^2\right)^2}{n_{.j}},$$

$$k_9 = \sum_{i=1}^{a} \frac{\left(\sum_{j=1}^{b} n_{ij}^2\right)^2}{n_{i.}^2}, \qquad k_{10} = \sum_{j=1}^{b} \frac{\left(\sum_{i=1}^{a} n_{ij}^2\right)^2}{n_{.j}^2},$$

$$k_{11} = \sum_{i=1}^{a} \frac{\left(\sum_{j=1}^{b} n_{ij}^3\right)}{n_{i.}}, \qquad k_{12} = \sum_{j=1}^{b} \frac{\left(\sum_{i=1}^{a} n_{ij}^3\right)}{n_{.j}},$$

$$k_{13} = \sum_{i=1}^{a} \frac{\left(\sum_{j=1}^{b} n_{ij}^2\right)\left(\sum_{j=1}^{b} n_{ij}n_{.j}\right)}{n_{i.}},$$

$$k_{14} = \sum_{j=1}^{b} \frac{\left(\sum_{i=1}^{a} n_{ij}^2\right)\left(\sum_{i=1}^{a} n_{ij}n_{i.}\right)}{n_{.j}}, \qquad k_{15} = \sum_{i=1}^{a} \frac{\left(\sum_{j=1}^{b} n_{ij}n_{.j}\right)^2}{n_{i.}},$$

[4] Some of the k-terms being defined, although not used here, are employed in Section 13.5.1.

$$k_{16} = \sum_{j=1}^{b} \frac{\left(\sum_{i=1}^{a} n_{ij} n_{i.}\right)^2}{n_{.j}}, \qquad k_{17} = \sum_{i=1}^{a} \frac{\left(\sum_{j=1}^{b} n_{ij}^2 n_{.j}\right)}{n_{i.}},$$

$$k_{18} = \sum_{j=1}^{b} \frac{\left(\sum_{i=1}^{a} n_{ij}^2 n_{i.}\right)}{n_{.j}}, \qquad k_{19} = \sum_{i=1}^{a} \left(\sum_{j=1}^{b} n_{ij}^2\right) n_{i.},$$

$$k_{20} = \sum_{j=1}^{b} \left(\sum_{i=1}^{a} n_{ij}^2\right) n_{.j}, \qquad k_{21} = \mathop{\sum_{i=1}^{a} \sum_{i'=1}^{a}}_{i \neq i'} \frac{\left(\sum_{j=1}^{b} n_{ij} n_{i'j}\right)^2}{n_{i.} n_{i'.}},$$

$$k_{22} = \sum_{j=1}^{b} \sum_{j'=1}^{b} \frac{\left(\sum_{i=1}^{a} n_{ij} n_{ij'}\right)^2}{n_{.j} n_{.j'}}, \qquad k_{23} = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij}^2, \qquad k_{24} = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij}^3,$$

$$k_{25} = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} n_{i.} n_{.j}, \qquad k_{26} = \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{n_{ij}^2}{n_{i.} n_{.j}},$$

$$k_{27} = \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{n_{ij}^3}{n_{i.} n_{.j}}, \quad \text{and} \quad k_{28} = \sum_{i=1}^{a} \sum_{j=1}^{b} \frac{n_{ij}^4}{n_{i.} n_{.j}}.$$

12.5.2 VARIANCES OF FITTING-CONSTANTS-METHOD ESTIMATORS

For estimators of the variance components using fitting-constants method as given by (12.4.13), Low (1964) has developed the expressions for variances and covariances. The desired results are (see also Searle, 1971, p. 489)

$$\begin{aligned}
\operatorname{Var}(\hat{\sigma}_e^2) &= \frac{2\sigma_e^4}{\nu_e}, \\
\operatorname{Var}(\hat{\sigma}_\beta^2) &= \frac{2}{h_1^2}\left[(N-a)(b-1)\frac{\sigma_e^4}{\nu_e} + 2h_1\sigma_e^2\sigma_\beta^2 + f_2\sigma_\beta^4\right], \\
\operatorname{Var}(\hat{\sigma}_\alpha^2) &= \frac{2}{h_2^2}\left[(N-b)(a-1)\frac{\sigma_e^4}{\nu_e} + 2h_2\sigma_e^2\sigma_\alpha^2 + f_1\sigma_\alpha^4\right], \\
\operatorname{Cov}(\hat{\sigma}_\beta^2, \hat{\sigma}_e^2) &= \frac{-2(b-1)\sigma_e^4}{\nu_e h_1}, \\
\operatorname{Cov}(\hat{\sigma}_\alpha^2, \hat{\sigma}_e^2) &= \frac{-2(a-1)\sigma_e^4}{\nu_e h_2},
\end{aligned}$$

and

$$\operatorname{Cov}(\hat{\sigma}_\alpha^2, \hat{\sigma}_\beta^2) = \frac{2\sigma_e^4}{h_1 h_2}\left[k_{26} - 1 + \frac{(a-1)(b-1)}{\nu_e}\right],$$

where

$$h_1 = N - k_3, \qquad h_2 = N - k_4,$$
$$f_1 = k_1 - 2k_{18} + \Sigma_i \Sigma_{i'} \left(\frac{\Sigma_j n_{ij} n_{i'j}}{n_{.j}} \right)^2,$$
$$f_2 = k_2 - 2k_{17} + \Sigma_j \Sigma_{j'} \left(\frac{\Sigma_i n_{ij} n_{ij'}}{n_{i.}} \right)^2,$$
$$\nu_e = N - a - b + 1,$$

and k_1, k_2, k_3, k_{17}, k_{18}, and k_{26} are defined in Section 12.5.1.

12.6 CONFIDENCE INTERVALS

Exact confidence intervals on $\sigma_\alpha^2/\sigma_e^2$ and $\sigma_\beta^2/\sigma_e^2$ can be constructed using Wald's procedure discussed in Section 11.8 (see also Spjøtvoll, 1968). Burdick and Graybill (1992, pp. 143–144) provide a numerical example illustrating Wald's procedure using SAS® code. However, there do not exist exact intervals on other functions of variance components. For the design with no empty cells, Burdick and Graybill (1992, pp. 142–143) recommend using intervals for the corresponding balanced case where the usual mean squares are replaced by the mean squares in the unweighted analysis presented in Section 12.4.3 and n is substituted by $\bar{n}_h$.

For example, approximate confidence intervals for σ_β^2 and σ_α^2 based on Ting et al. (1990) procedure are given by

$$P\left\{ \frac{1}{a\bar{n}_h}(\text{MS}_{Bu} - \text{MS}_{Eu}) - \sqrt{L_{\beta u}} \le \sigma_\beta^2 \le \frac{1}{a\bar{n}_h}(\text{MS}_{Bu} - \text{MS}_{Eu}) + \sqrt{U_{\beta u}} \right\} \doteq 1 - \alpha \tag{12.6.1}$$

and

$$P\left\{ \frac{1}{b\bar{n}_h}(\text{MS}_{Au} - \text{MS}_{Eu}) - \sqrt{L_{\alpha u}} \le \sigma_\alpha^2 \le \frac{1}{b\bar{n}_h}(\text{MS}_{Au} - \text{MS}_{E}) + \sqrt{U_{\alpha u}} \right\} \doteq 1 - \alpha, \tag{12.6.2}$$

where

$$L_{\beta u} = \frac{1}{a^2\bar{n}_h^2}[G_2^2\text{MS}_{Bu}^2 + H_3^2\text{MS}_{Eu}^2 + G_{23}\text{MS}_{Bu}\text{MS}_{Eu}],$$
$$U_{\beta u} = \frac{1}{a^2\bar{n}_h^2}[H_2^2\text{MS}_{Bu}^2 + G_3^2\text{MS}_{Eu}^2 + H_{23}\text{MS}_{Bu}\text{MS}_{Eu}],$$
$$L_{\alpha u} = \frac{1}{b^2\bar{n}_h^2}[G_1^2\text{MS}_{Au}^2 + H_3^2\text{MS}_{Eu}^2 + G_{13}\text{MS}_{Au}\text{MS}_{Eu}],$$

$$U_{\alpha u} = \frac{1}{b^2\bar{n}_h^2}[H_1^2\mathrm{MS}_{Au}^2 + G_3^2\mathrm{MS}_{Eu}^2 + H_{13}\mathrm{MS}_{Au}\mathrm{MS}_{Eu}],$$

with

$$G_1 = 1 - F^{-1}[\nu_\alpha, \infty; 1-\alpha/2], \qquad G_2 = 1 - F^{-1}[\nu_\beta, \infty; 1-\alpha/2],$$

$$G_3 = 1 - F[\nu_e, \infty; 1-\alpha/2], \qquad H_1 = F^{-1}[\nu_\alpha, \infty; \alpha/2] - 1,$$

$$H_2 = F^{-1}[\nu_\beta, \infty; \alpha/2] - 1, \qquad H_3 = F^{-1}[\nu_e, \infty; \alpha/2] - 1,$$

$$G_{13} = \frac{(F[\nu_\alpha, \nu_e; 1-\alpha/2] - 1)^2 - G_1^2F^2[\nu_\alpha, \nu_e; 1-\alpha/2] - H_3^2}{F[\nu_\alpha, \nu_e; 1-\alpha/2]},$$

$$G_{23} = \frac{(F[\nu_\beta, \nu_e; 1-\alpha/2] - 1)^2 - G_2^2F^2[\nu_\beta, \nu_e; 1-\alpha/2] - H_3^2}{F[\nu_\beta, \nu_e; 1-\alpha/2]},$$

$$H_{13} = \frac{(1 - F[\nu_\alpha, \nu_e; \alpha/2])^2 - H_1^2F^2[\nu_\alpha, \nu_e; \alpha/2] - G_3^2}{F[\nu_\alpha, \nu_e; \alpha/2]},$$

$$H_{23} = \frac{(1 - F[\nu_\beta, \nu_e; \alpha/2])^2 - H_2^2F^2[\nu_\beta, \nu_e; \alpha/2] - G_3^2}{F[\nu_\beta, \nu_e; \alpha/2]},$$

$$\nu_\alpha = a - 1, \qquad \nu_\beta = b - 1, \quad \text{and} \quad \nu_e = N - a - b + s.$$

Similarly, an approximate confidence interval for $\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2$ is given by

$$\begin{aligned} P\Bigg\{\hat{\gamma} &- \frac{1}{ab\bar{n}_h}\sqrt{[a^2G_1^2\mathrm{MS}_{Au}^2 + b^2G_2^2\mathrm{MS}_{Bu}^2 + (ab\bar{n}_h - a - b)G_3^2\mathrm{MS}_{Eu}^2]} \\ &\leq \sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2 \\ &\leq \hat{\gamma} + \frac{1}{ab\bar{n}_h}\sqrt{[a^2H_1^2\mathrm{MS}_{Au}^2 + b^2H_2^2\mathrm{MS}_{Bu}^2 + (ab\bar{n}_h - a - b)H_3^2\mathrm{MS}_{Eu}^2]}\Bigg\} \\ &\doteq 1 - \alpha, \end{aligned} \tag{12.6.3}$$

where

$$\hat{\gamma} = \frac{1}{ab\bar{n}_h}[a\mathrm{MS}_{Au} + b\mathrm{MS}_{Bu} + (ab\bar{n}_h - a - b)\mathrm{MS}_{Eu}]$$

and G_1, G_2, G_3, H_1, H_2, and H_3 are defined following (12.6.2). Other formulas can similarly be developed. On the basis of some simulation studies by Srinivasan (1986), Hernández (1991), and Srinivasan and Graybill(1991), the authors report that these intervals provide reasonably good coverage. For data sets with some empty cells, where the design is connected, Burdick and Graybill (1992, pp. 144–145) recommend the use of adjusted sums of squares considered in Section 12.4.2. These sums of squares are equivalent to Type II, Type III, or Type IV sums of squares produced in an analysis using PROC GLM in SAS®. This approach for constructing confidence intervals on σ_α^2 and σ_β^2 has been used by Kazempour and Graybill (1992); and Kazempour and Graybill (1989) have used it for constructing confidence intervals on $\rho_\alpha = \sigma_\alpha^2/(\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2)$, $\rho_\beta = \sigma_\beta^2/(\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2)$, and $\rho_e = \sigma_e^2/(\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2)$. The interested reader is referred to these works for further details and insight into the problem.

12.6.1 A NUMERICAL EXAMPLE

In this section, we illustrate computations of confidence intervals on the variance components σ_β^2 and σ_α^2 and the total variance $\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2$ using formulas (12.6.1), (12.6.2), and (12.6.3). Now, from the results of the analysis of variance given in Table 12.11, we have

$$\text{MS}_{Eu} = 0.02758, \qquad \text{MS}_{Bu} = 0.13884, \qquad \text{MS}_{Au} = 0.28099,$$
$$a = 4, \quad b = 5, \quad \bar{n}_h = 1.983, \quad \nu_e = 45, \quad \nu_\beta = 4, \quad \nu_\alpha = 3.$$

Further, for $\alpha = 0.05$, we obtain

$$\begin{aligned}
F[\nu_\alpha, \infty; \alpha/2] &= 0.072, & F[\nu_\alpha, \infty; 1-\alpha/2] &= 3.116,\\
F[\nu_\beta, \infty; \alpha/2] &= 0.121, & F[\nu_\beta, \infty; 1-\alpha/2] &= 2.786,\\
F[\nu_e, \infty; \alpha/2] &= 0.630, & F[\nu_e, \infty; 1-\alpha/2] &= 1.454,\\
F[\nu_\alpha, \nu_e; \alpha/2] &= 0.071, & F[\nu_\alpha, \nu_e; 1-\alpha/2] &= 3.422,\\
F[\nu_\beta, \nu_e; \alpha/2] &= 0.119, & F[\nu_\beta, \nu_e; 1-\alpha/2] &= 3.086.
\end{aligned}$$

In addition, to determine approximate confidence intervals for σ_β^2 and σ_α^2 using formulas (12.6.1) and (12.6.2), we evaluate the following quantities:

$$G_1 = 0.67907574, \qquad G_2 = 0.64106245, \qquad G_3 = 0.31224209,$$
$$H_1 = 12.88888889, \qquad H_2 = 7.26446281, \qquad H_3 = 0.58730159,$$
$$\begin{aligned}
G_{13} &= 0.03539671, & G_{23} &= 0.03004790,\\
H_{13} &= -1.01242914, & H_{23} &= -0.57684338,\\
L_{\beta u} &= 1.31910089 \times 10^{-4}, & U_{\beta u} &= 0.01613461,\\
L_{\alpha u} &= 3.75826605 \times 10^{-4}, & U_{\alpha u} &= 0.13334288.
\end{aligned}$$

Substituting the appropriate quantities in (12.6.1) and (12.6.2), the desired 95% confidence intervals for σ_β^2 and σ_α^2 are given by

$$P\{0.0025 \le \sigma_\beta^2 \le 0.1411\} \doteq 0.95$$

and

$$P\{0.0062 \le \sigma_\alpha^2 \le 0.3907\} \doteq 0.95.$$

To determine an approximate confidence interval for the total variance $\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2$ using formula (12.6.3), we obtain

$$\begin{aligned}
\hat{\gamma} = \frac{1}{4 \times 5 \times 1.983}&[4 \times 0.28099 + 5 \times 0.13884\\
&+ (4 \times 5 \times 1.983 - 4 - 5)0.02758] = 0.0672.
\end{aligned}$$

Substituting the appropriate quantities in (12.6.3), the desired 95% confidence interval for $\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2$ is given by

$$P\{0.0449 \leq \sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2 \leq 0.4539\} \doteq 0.95.$$

Confidence intervals for other parametric functions of the variance components can similarly be determined.

12.7 TESTS OF HYPOTHESES

In this section, we consider briefly some tests of the hypotheses

$$H_0^A : \sigma_\alpha^2 = 0 \quad \text{vs.} \quad H_1^A : \sigma_\alpha^2 > 0$$

and (12.7.1)

$$H_0^B : \sigma_\beta^2 = 0 \quad \text{vs.} \quad H_1^B : \sigma_\beta^2 > 0.$$

12.7.1 TESTS FOR $\sigma_\alpha^2 = 0$ AND $\sigma_\beta^2 = 0$

Exact tests for the hypotheses in (12.7.1) were first proposed by Wald (1941, 1947) in the context of construction of confidence intervals. Spjøtvoll (1968) and Thomsen (1975) using two different approaches also derived exact tests. For designs with no empty cells, Seely and El-Bassiouni (1983) have shown that the Spjøtvoll–Thomsen test is equivalent to Wald's test and is given by the usual ANOVA F-tests for main effects in the fixed effects model[5] (see also Khuri et al., 1998, pp. 101–103). Approximate F-tests can be constructed using synthesis of mean squares obtained from the conventional analysis of variance given in Section 12.2. For example, to test $H_0^A : \sigma_\alpha^2 = 0$ vs. $H_1^A : \sigma_\alpha^2 > 0$, the test procedure can be based on MS_A/MS_D, where MS_D is given by

$$\text{MS}_D = \left(\frac{r_5 - r_1}{r_3 - r_1}\right)\text{MS}_B + \left(\frac{r_3 - r_5}{r_3 - r_1}\right)\text{MS}_E. \qquad (12.7.2)$$

Similarly, to test $H_0^B : \sigma_\beta^2 = 0$ vs. $H_1^B : \sigma_\beta^2 > 0$, the test procedure can be based on MS_B/MS'_D, where MS'_D is given by

$$\text{MS}'_D = \left(\frac{r_4 - r_2}{r_6 - r_2}\right)\text{MS}_A + \left(\frac{r_6 - r_4}{r_6 - r_2}\right)\text{MS}_E. \qquad (12.7.3)$$

The test statistics MS_A/MS_D and MS_B/MS'_D are approximated by F-variables with $(a-1, \nu_D)$ and $(b-1, \nu'_D)$ degrees of freedom, respectively, where ν_D and

[5]The F-test for $\sigma_\alpha^2 = 0$ is based on the ANOVA decomposition when ordering the factors as B, A and can be obtained using SAS Type I sums of squares. A similar F-test is obtained for testing the significance of σ_β^2 by ordering the factors as A, B. Alternatively, one can perform both tests more directly using SAS Type II sums of squares.

ν'_D are estimated using the Satterthwaite formula. Similar psuedo F-tests can also be constructed using synthesized mean squares based on unweighted and weighted means analyses considered in Section 12.4.3. Hussein and Milliken (1978) discuss tests for hypotheses in (12.7.1) involving heterogeneous error variances. As noted in Section 11.9, uniformly optimum tests for testing the hypotheses in (12.7.1) do not exist. In the special case when $n_{i.}$s are all equal and $n_{.j}$s are all equal, Mathew and Sinha (1988) derived a locally best invariant unbiased (LBIU) test. However, the LBIU test requires obtaining information on certain conditional distributions and is difficult to use in practice. For a concise discussion of some of these tests, see Khuri et al. (1998, pp. 101–104).

12.7.2 A NUMERICAL EXAMPLE

In this example, we outline results for testing the hypotheses in (12.7.1) using the fusiform rust data of the numerical example in Section 12.4.6. First, we use the Wald test, which is the usual ANOVA F-test for main effects in the fixed effects model. The F-test for $H_0^A : \sigma_\alpha^2 = 0$ using Type I sums of squares when ordering the factors as (family, location) gives an F-value of 10.84 ($p < 0.001$). The results are highly significant and we reject H_0^A and conclude that $\sigma_\alpha^2 > 0$ or the fusiform rusts in trees from different locations differ significantly. Similarly, the F-test for $H_0^B : \sigma_\beta^2 = 0$ using Type I sums of squares when ordering the factors as (location, family) gives an F-value of 8.48 ($p < 0.001$). Again, the results are highly significant and we reject H_0^B and conclude that $\sigma_\beta^2 > 0$, or fusiform rusts in trees from different families differ significantly. Now, we illustrate the application of F-tests based on the Satterthwaite procedure using the conventional analysis of variance. From Table 12.7, we have

$$\begin{aligned} &r_1 = -0.016, \qquad r_2 = -0.028, \qquad r_3 = 10.255,\\ &r_4 = 0.234, \qquad r_5 = 0.238, \qquad r_6 = 13.182. \end{aligned}$$

Further, from (12.7.2) and (12.7.3), the synthesized mean squares MS_D and MS'_D and the corresponding degrees of freedom ν_D and ν'_D are given by

$$\begin{aligned} \text{MS}_D &= \left(\frac{0.238 + 0.016}{10.255 + 0.016}\right) 0.1942 + \left(\frac{10.255 - 0.238}{10.255 + 0.016}\right) 0.0279,\\ &= 0.0048 + 0.0272 = 0.0320,\\ \text{MS}'_D &= \left(\frac{0.234 + 0.028}{13.182 + 0.028}\right) 0.2468 + \left(\frac{13.182 - 0.234}{13.182 + 0.028}\right) 0.0279,\\ &= 0.0049 + 0.0273 = 0.0322,\\ \nu_D &= \frac{(0.0320)^2}{\frac{(0.0048)^2}{4} + \frac{(0.0272)^2}{45}} = 46.1, \end{aligned}$$

and

$$\nu'_D = \frac{(0.0322)^2}{\frac{(0.0049)^2}{3} + \frac{(0.0273)^2}{45}} = 42.2.$$

The test statistics $\mathrm{MS}_A/\mathrm{MS}_D$ and $\mathrm{MS}_B/\mathrm{MS}'_D$ yield F-values of 7.71 and 6.03 which are to be compared against the theoretical F-values with (3, 46.1) and (4, 42.2) degrees of freedom, respectively. The corresponding p-values are < 0.001 and < 0.001, respectively, and both the results are highly significant. Finally, these tests can also be based on analysis of variance on cell means using unweighted or weighted sums of squares given in Tables 12.11 and 12.12. Using unweighted analysis, the F-values for testing $\sigma_\alpha^2 = 0$ and $\sigma_\beta^2 = 0$ are 10.19 ($p < 0.001$) and 5.03 ($p < 0.001$), respectively. Using weighted analysis, the corresponding F-values are 10.84 ($p < 0.001$) and 8.49 ($p < 0.001$), respectively. Thus all the tests, exact as well as approximate, lead to the same conclusion.

EXERCISES

1. Apply the method of "synthesis" to derive the expected mean squares given in Section 12.3.
2. Derive the results on expected values of reductions in sums of squares given in (12.4.7) and (12.4.10).
3. Derive the results on expected values of the unweighted sums of squares given in (12.4.15).
4. Derive the results on expected values of the weighted sums of squares given in (12.4.17).
5. Derive the expressions for the variances and covariances of T_A, T_B, T_{AB}, and T_μ given in Section 12.5.1.
6. Show that the fitting-constants-method estimators (12.4.8), (12.4.11), and (12.4.13) reduce to the ANOVA estimators (3.4.1) for balanced data.
7. Show that the ANOVA estimators (12.4.3) reduce to the corresponding estimators (3.4.1) for balanced data.
8. Show that the unweighted means estimators (12.4.16) reduce to the ANOVA estimators (3.4.1) for balanced data.
9. Show that the weighted means estimators (12.4.19) reduce to the ANOVA estimators (3.4.1) for balanced data.
10. Show that the symmetric sums estimators (12.4.23) and (12.4.26) reduce to the ANOVA estimators (3.4.1) for balanced data.
11. An experiment was designed to study the variation in the intensity of radiation from an open earth furnace at different locations and the time

of the day. The locations and three time periods were randomly chosen and the data on the radiation intensity measured in milivolts are given below.

Location	**Time**		
	1	**2**	**3**
1	48, 49	53	43, 44
2	49, 51	55, 56	52
3	46	57	48

(a) Describe the mathematical model with additive effect and the assumptions involved.

(b) Analyze the data and report the analysis of variance table.

(c) Perform an appropriate F-test to determine whether the intensity of radiation differs from location to location.

(d) Perform an appropriate F-test to determine whether the intensity of radiation differs between different time periods of the day.

(e) Find point estimates of the variance components and the total variance using the methods described in the text.

(f) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

12. An experiment was designed to determine the length of development period (in days) for different strains of house flies at different densities of container. A sample of three strains was taken and each was bred at four densities. For each strain $\times$ density combination varying number of measurements were taken to measure the mean length of development period. The data are given below.

Strain	**Density**			
	1	**2**	**3**	**4**
1	20.6	20.0	10.0	18.2
	26.8	15.8	12.5	19.5
	20.6		17.6	
			18.7	
			18.7	
2	12.3	23.2	16.4	28.0
	12.0	20.5		28.9
	17.2	20.2		27.9
	13.9	17.7		
	13.0			
3	16.9	12.5	6.8	14.4
	16.1	5.9		13.0
	20.8	5.2		11.0
		9.4		
		8.7		

(a) Describe the mathematical model with additive effect and the assumptions involved.

(b) Analyze the data and report the analysis of variance table.

(c) Perform an appropriate F-test to determine whether the development period differs from strain to strain.

(d) Perform an appropriate F-test to determine whether the development period differs from density to density.

(e) Find point estimates of the variance components and the total variance using the methods described in the text.

(f) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

13. Samples of new growth in hybrid polars of four varieties grown in three soil conditions were taken and analyzed for oven-dry weights (in grams). The data are given below.

Soil	Variety			
	1	2	3	4
1	57.6	46.1	51.1	59.1
	58.8	48.5		
	57.2	47.3		
	54.2	43.5		
	55.4			
2	47.1	41.9	38.0	39.6
	44.2		41.4	36.0
	41.5		40.3	41.2
	49.2		41.5	37.4
			35.7	
3	37.1	47.8	45.9	51.4
		45.9	39.7	50.1
				52.1
				48.4

(a) Describe the mathematical model with additive effect and the assumptions involved.

(b) Analyze the data and report the analysis of variance table.

(c) Perform an appropriate F-test to determine whether the oven-dry weight differs from soil to soil.

(d) Perform an appropriate F-test to determine whether the oven-dry weight differs from variety to variety.

(e) Find point estimates of the variance components and the total variance using the methods described in the text.

(f) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

14. An experiment was designed with 12 rabbits who received an injection of insulin. Two factors were involved, the preparation of insulin at three levels and the dose at four levels. The levels of preparation and dose were randomly selected from a large number of such levels available for the experiment. Five blood samples were taken and analyzed to determine the percent of reduction in blood sugar. However, for certain combinations of levels of preparation and dose, a number of analyses could not be performed because of insufficient quantity of blood. The data are given below.

Preparation	**Dose**			
	1	**2**	**3**	**4**
1	32.6	46.5	37.9	33.6
		50.1	43.5	32.7
		51.9	38.9	37.7
		47.9	42.8	38.0
		45.9		32.5
2	22.7	21.7	28.4	32.1
		22.1	30.4	29.9
		22.4	33.8	31.9
		21.4	27.7	32.6
			31.2	28.6
3	32.3	30.9	46.3	27.7
	32.1	34.0	42.4	28.6
	32.0	32.2	42.0	27.2
		33.4	45.8	28.8
		31.5		30.5

(a) Describe the mathematical model with additive effect and the assumptions involved.

(b) Analyze the data and report the analysis of variance table.

(c) Perform an appropriate F-test to determine whether the percent reduction in blood sugar differs from preparation to preparation.

(d) Perform an appropriate F-test to determine whether the percent reduction in blood sugar differs from dose to dose.

(e) Find point estimates of the variance components and the total variance using the methods described in the text.

(f) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

Bibliography

R. K. Burdick and F. A. Graybill (1992), *Confidence Intervals on Variance Components*, Marcel Dekker, New York.

D. G. Gosslee and H. L. Lucas (1965), Analysis of variance of disproportionate data when interaction is present, *Biometrics*, **21**, 115–133.

R. P. Hernández (1991), *Confidence Intervals on Linear Combinations of Variance Components in Unbalanced Designs*, Ph.D. dissertation, Arizona State University, Tempe, AZ.

C. Hirotsu (1966), Estimating variance, components in a two-way layout with unequal numbers of observations, *Rep. Statist. Appl. Res.* (*JUSE*), **13**, 29–34.

M. Hussein and G. A. Milliken (1978), An unbalanced two-way model with random effects having unequal variances, *Biometrical J.*, **20**, 203–213.

M. K. Kazempour and F. A. Graybill (1989), Confidence bounds for proportion of σ_E^2 in unbalanced two-way crossed models, in Y. Dodge, ed., *Statistical Data Analysis*, North-Holland, Amsterdam, 389–396.

M. K. Kazempour and F. A. Graybill (1992), Confidence intervals on individual variances in two-way models, *Comput. Statist. Data Anal.*, **14**, 29–37.

A. I. Khuri and R. C. Littell (1987), Exact tests for the main effects variance components in an unbalanced random two-way model, *Biometrics*, **43**, 545–560.

A. I. Khuri, T. Mathew, and B. K. Sinha (1998), *Statistical Tests for Mixed Linear Models*, Wiley, New York.

G. G. Koch (1967), A general approach to the estimation of variance components, *Techometrics*, **9**, 93–118.

G. G. Koch (1968), Some further remarks concerning "A general approach to estimation of variance components," *Technometrics*, **10**, 551–558.

L. Y. Low (1964), Sampling variances of estimates of components of variance from a non-orthogonal two-way classification, *Biometrika*, **51**, 491–494.

L. Y. Low (1976), Some properties of variance component estimators, *Technometrics*, **18**, 39–46.

T. Mathew and B. K. Sinha (1988), Optimum tests in unbalanced two-way models without interaction, *Ann. Statist.*, **16**, 1727–1740.

M. G. Mostafa (1967), Note on testing hypotheses in an unbalanced random effects model, *Biometrika*, **54**, 659–662.

S. R. Searle (1958), Sampling variances of estimates of components of variance, *Ann. Math. Statist.*, **29**, 167–178.

S. R. Searle (1971), *Linear Models*, Wiley, New York.

S. R. Searle (1987), *Linear Models for Unbalanced Data*, Wiley, New York.

S. R. Searle, G. Casella, and C. E. McCulloch (1992), *Variance Components*, Wiley, New York.

S. R. Searle and C. R. Henderson (1961), Computing procedures for estimating components of variance in the two-way classification, mixed model, *Biometrics*, **17**, 607–616; corrigenda, 1967, **23**, 852.

J. F. Seely and M. Y. El-Bassiouni (1983), Applying Wald's variance component test, *Ann. Statist.*, **11**, 197–201.

E. Spjøtvoll (1968), Confidence intervals and tests for variance ratios in unbalanced variance components models, *Rev. Internat. Statist. Inst.*, **36**, 37–42.

S. Srinivasan (1986), *Confidence Intervals on Functions of Variance Compo-*

nents in Unbalanced Two-Way Design Models, Ph.D. dissertation, Colorado State University, Fort Collins, CO.

S. Srinivasan and F. A. Graybill (1991), Confidence intervals for proportions of total variation in unbalanced two-way components of variance models using unweighted means, *Comm. Statist.* A *Theory Methods*, **20**, 511–526.

I. B. Thomsen (1975), Testing hypotheses in unbalanced variance components models for two-way layouts, *Ann. Statist.*, **3**, 257–265.

N. Ting, R. K. Burdick, F. A. Graybill, S. Jeyaratnam, and T.-F. C. Lu (1990), Confidence intervals on linear combinations of variance components that are unrestricted in sign, *J. Statist. Comput. Simul.*, **35**, 135–143.

A. Wald (1941), On the analysis of variance in case of multiple classifications with unequal class frequencies, *Ann. Math. Statist.*, **12**, 346–350.

A. Wald (1947), A note on regression analysis, *Ann. Math. Statist.*, **18**, 586–589.

F. Yates (1934), The analysis of multiple classifications with unequal numbers in the different classes, *J. Amer. Statist. Assoc.*, **29**, 51–66.

13 Two-Way Crossed Classification with Interaction

Consider two factors A and B with a and b levels, respectively, involving a factorial arrangement. Assume that $n_{ij}\,(\geq 0)$ observations are taken corresponding to the (i, j)th cell. The model for this design is known as the unbalanced two-way crossed classification. This model is the same as the one considered in Chapter 4 except that now the number of observations per cell is not constant but varies from cell to cell. Models of this type frequently occur in many experiments and surveys since many studies cannot guarantee the same number of observations for each cell. This chapter is devoted to the study of a random effects model for unbalanced two-way crossed classification with interaction.

13.1 MATHEMATICAL MODEL

The random effects model for the unbalanced two-way crossed classification with interaction is given by

$$\begin{aligned} y_{ijk} &= \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}, \qquad (13.1.1)\\ i &= 1, \ldots, a; \quad j = 1, \ldots, b; \quad k = 0, \ldots, n_{ij}, \end{aligned}$$

where y_{ijk} is the kth observation at the ith level of factor A and the jth level of factor B, μ is the overall mean, α_is and β_js are main effects, i.e., α_i is the effect of the ith level of factor A and β_j is the effect of the jth level of factor B, $(\alpha\beta)_{ij}$s are the interaction terms, and e_{ijk}s are the customary error terms. It is assumed that $-\infty < \mu < \infty$ is a constant and α_is, β_js, $(\alpha\beta)_{ij}$s, and e_{ijk}s are mutually and completely uncorrelated random variables with means zero and variances σ_α^2, σ_β^2, $\sigma_{\alpha\beta}^2$, and σ_e^2, respectively. The parameters σ_α^2, σ_β^2, $\sigma_{\alpha\beta}^2$, and σ_e^2 are the variance components of the model in (13.1.1).

13.2 ANALYSIS OF VARIANCES

For the two-way model in (13.1.1) there is no unique analysis of variance. The conventional analysis of variance obtained by an analogy with the correspond-

TABLE 13.1 Analysis of variance for the model in (13.1.1).

Source of variation	Degrees of freedom*	Sum of squares	Mean square	Expected mean square
Factor A	$a-1$	SS_A	MS_A	$\sigma_e^2+r_7\sigma_{\alpha\beta}^2+r_8\sigma_\beta^2+r_9\sigma_\alpha^2$
Factor B	$b-1$	SS_B	MS_B	$\sigma_e^2+r_4\sigma_{\alpha\beta}^2+r_5\sigma_\beta^2+r_6\sigma_\alpha^2$
Interaction AB	$s-a-b+1$	SS_{AB}	MS_{AB}	$\sigma_e^2+r_1\sigma_{\alpha\beta}^2+r_2\sigma_\beta^2+r_3\sigma_\alpha^2$
Error	$N-s$	SS_E	MS_E	σ_e^2

*s = number of nonempty cells, i.e., $n_{ij} > 0$ for $s(i, j)$ cells.

ing balanced analysis is given in Table 13.1.

The sums of squares in Table 13.1 are defined as follows:[1]

$$\begin{aligned}
SS_A &= \sum_{i=1}^{a} n_{i.}(\bar{y}_{i..}-\bar{y}_{...})^2 = \sum_{i=1}^{a}\frac{y_{i..}^2}{n_{i.}} - \frac{y_{...}^2}{N},\\
SS_B &= \sum_{j=1}^{b} n_{.j}(\bar{y}_{.j.}-\bar{y}_{...})^2 = \sum_{j=1}^{b}\frac{y_{.j.}^2}{n_{.j}} - \frac{y_{...}^2}{N},\\
SS_{AB} &= \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}\bar{y}_{ij.}^2 - \sum_{i=1}^{a} n_{i.}\bar{y}_{i..}^2 - \sum_{j=1}^{b} n_{.j}\bar{y}_{.j.}^2 + N\bar{y}_{...}^2 \qquad (13.2.1)\\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{y_{ij.}^2}{n_{ij}} - \sum_{i=1}^{a}\frac{y_{i..}^2}{n_{i.}} - \sum_{j=1}^{b}\frac{y_{.j.}^2}{n_{.j}} + \frac{y_{...}^2}{N},
\end{aligned}$$

and

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}}(y_{ijk}-\bar{y}_{ij.})^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} y_{ijk}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{y_{ij.}^2}{n_{ij}},$$

where

$$y_{ij.} = \sum_{k=1}^{n_{ij}} y_{ijk}, \qquad \bar{y}_{ij.} = \frac{y_{ij.}}{n_{ij}},$$

$$y_{i..} = \sum_{j=1}^{b} y_{ij.}, \qquad \bar{y}_{i..} = \frac{y_{i..}}{n_{i.}},$$

[1]Note that SS_{AB} defined in (13.2.1) is not equal to $\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}(\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}+\bar{y}_{...})^2$ (see Exercise 13.3).

$$y_{.j.} = \sum_{i=1}^{a} y_{ij.}, \qquad \bar{y}_{.j.} = \frac{y_{.j.}}{n_{.j}},$$

and

$$y_{...} = \sum_{i=1}^{a} y_{i..} = \sum_{j=1}^{b} y_{.j.}, \qquad \bar{y}_{...} = \frac{y_{...}}{N},$$

with

$$n_{i.} = \sum_{j=1}^{b} n_{ij}, \qquad n_{.j} = \sum_{i=1}^{a} n_{ij}$$

and

$$N = \sum_{i=1}^{a} n_{i.} = \sum_{j=1}^{b} n_{.j} = \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}.$$

The SS_A, SS_B, SS_{AB}, and SS_E terms in (13.2.1) have been defined by establishing an analogy with the corresponding terms for the balanced case.

Define the uncorrected sums of squares as

$$T_A = \sum_{i=1}^{a} \frac{y_{i..}^2}{n_{i.}}, \qquad T_B = \sum_{j=1}^{b} \frac{y_{.j.}^2}{n_{.j}},$$

$$T_{AB} = \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{y_{ij.}^2}{n_{ij}}, \qquad T_0 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} y_{ijk}^2,$$

and

$$T_\mu = \frac{y_{...}^2}{N}.$$

Then the corrected sums of squares defined in (13.2.1) can be written as

$$\mathrm{SS}_A = T_A - T_\mu, \qquad \mathrm{SS}_B = T_B - T_\mu,$$
$$\mathrm{SS}_{AB} = T_{AB} - T_A - T_B + T_\mu,$$

and

$$\mathrm{SS}_E = T_0 - T_{AB}.$$

As remarked in Section 12.2, not all the expressions defined in (13.2.1) are in fact sums of squares, notably the SS_{AB} term which can be negative. The mean squares as usual are obtained by dividing the sums of squares values by the corresponding degrees of freedom. The results on expected mean squares are outlined in the following section.

13.3 EXPECTED MEAN SQUARES

The expected sums of squares or mean squares are readily obtained by first calculating the expected values of the quantities T_0, T_{AB}, T_A, T_B, and T_μ. First note that by the assumption of the model in (13.1.1),

$$E(\alpha_i) = 0, \qquad E(\alpha_i^2) = \sigma_\alpha^2, \quad \text{and} \quad E(\alpha_i \alpha_{i'}) = 0, \quad i \neq i',$$

with similar results for the β_js, $(\alpha\beta)_{ij}$s, and e_{ijk}s. Also, all covariances between pairs of nonidentical random variables are zero.

Now, we have

$$\begin{aligned}
E(T_0) &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}} E[\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}]^2 \\
&= N(\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_{\alpha\beta}^2 + \sigma_e^2), \\
E(T_{AB}) &= \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{1}{n_{ij}} E[\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ij.}]^2. \\
&= N(\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_{\alpha\beta}^2) + s\sigma_e^2, \\
E(T_A) &= \sum_{i=1}^{a} \frac{1}{n_{i.}} E\left[n_{i.}\mu + n_{i.}\alpha_i + \sum_{j=1}^{b} n_{ij}\beta_j + \sum_{j=1}^{b} n_{ij}(\alpha\beta)_{ij} + e_{i..} \right]^2 \\
&= N(\mu^2 + \sigma_\alpha^2) + k_3(\sigma_\beta^2 + \sigma_{\alpha\beta}^2) + a\sigma_e^2, \\
E(T_B) &= \sum_{j=1}^{b} \frac{1}{n_{.j}} E\left[n_{.j}\mu + n_{.j}\beta_j + \sum_{i=1}^{a} n_{ij}\alpha_i + \sum_{i=1}^{a} n_{ij}(\alpha\beta)_{ij} + e_{.j.} \right]^2 \\
&= N(\mu^2 + \sigma_\beta^2) + k_4(\sigma_\alpha^2 + \sigma_{\alpha\beta}^2) + b\sigma_e^2,
\end{aligned}$$

and

$$\begin{aligned}
E(T_\mu) &= \frac{1}{N} E\left[N\mu + \sum_{i=1}^{a} n_{i.}\alpha_i + \sum_{j=1}^{b} n_{.j}\beta_j + \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}(\alpha\beta)_{ij} + e_{...} \right]^2 \\
&= N\mu^2 + k_1\sigma_\alpha^2 + k_2\sigma_\beta^2 + k_{23}\sigma_{\alpha\beta}^2 + \sigma_e^2,
\end{aligned}$$

where

$$k_1 = \frac{1}{N}\sum_{i=1}^{a} n_{i.}^2, \qquad k_2 = \frac{1}{N}\sum_{j=1}^{b} n_{.j}^2,$$

$$k_3 = \sum_{i=1}^{a} \frac{\left(\sum_{j=1}^{b} n_{ij}^2\right)}{n_{i.}}, \qquad k_4 = \sum_{j=1}^{b} \frac{\left(\sum_{i=1}^{a} n_{ij}^2\right)}{n_{.j}},$$

$$k_{23} = \frac{1}{N}\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}^2,$$

and s is the number of nonempty cells, i.e., $n_{ij} > 0$ for $s(i, j)$ cells.

Hence, expected values of sums of squares and mean squares are given as follows:

$$\begin{aligned}
E(\mathrm{SS}_E) &= E[T_0 - T_{AB}] = (N - s)\sigma_e^2,\\
E(\mathrm{MS}_E) &= \frac{1}{N-s}E(\mathrm{SS}_E) = \sigma_e^2;\\
E(\mathrm{SS}_{AB}) &= E[T_{AB} - T_A - T_B + T_\mu]\\
&= (s - a - b + 1)\sigma_e^2 + (N - k_3 - k_4 + k_{23})\sigma_{\alpha\beta}^2\\
&\quad + (k_2 - k_3)\sigma_\beta^2 + (k_1 - k_4)\sigma_\alpha^2,\\
E(\mathrm{MS}_{AB}) &= \frac{1}{s-a-b+1}E(\mathrm{SS}_{AB}) = \sigma_e^2 + r_1\sigma_{\alpha\beta}^2 + r_2\sigma_\beta^2 + r_3\sigma_\alpha^2;\\
E(\mathrm{SS}_B) &= E[T_B - T_\mu]\\
&= (b-1)\sigma_e^2 + (k_4 - k_{23})\sigma_{\alpha\beta}^2 + (N - k_2)\sigma_\beta^2 + (k_4 - k_1)\sigma_\alpha^2,\\
E(\mathrm{MS}_B) &= \frac{1}{b-1}E(\mathrm{SS}_B) = \sigma_e^2 + r_4\sigma_{\alpha\beta}^2 + r_5\sigma_\beta^2 + r_6\sigma_\alpha^2,\\
E(\mathrm{SS}_A) &= E[T_A - T_\mu]\\
&= (a-1)\sigma_e^2 + (k_3 - k_{23})\sigma_{\alpha\beta}^2 + (k_3 - k_2)\sigma_\beta^2 + (N - k_1)\sigma_\alpha^2,
\end{aligned}$$

and

$$E(\mathrm{MS}_A) = \frac{1}{a-1}E(\mathrm{SS}_A) = \sigma_e^2 + r_7\sigma_{\alpha\beta}^2 + r_8\sigma_\beta^2 + r_9\sigma_\alpha^2,$$

where

$$\begin{aligned}
r_1 &= \frac{N - k_3 - k_4 + k_{23}}{s - a - b + 1}, & r_2 &= \frac{k_2 - k_3}{s - a - b + 1},\\
r_3 &= \frac{k_1 - k_4}{s - a - b + 1}, & r_4 &= \frac{k_4 - k_{23}}{b - 1},\\
r_5 &= \frac{N - k_2}{b - 1}, & r_6 &= \frac{k_4 - k_1}{b - 1},\\
r_7 &= \frac{k_3 - k_{23}}{a - 1}, & r_8 &= \frac{k_3 - k_2}{a - 1},
\end{aligned}$$

and

$$r_9 = \frac{N - k_1}{a - 1}.$$

A noticeable aspect of the result $E(\text{MS}_A)$ is that it has a nonzero coefficient for every variance component in the model, whereas with balanced data the comparable expected value contains no terms in σ_β^2. The term of σ_β^2 in $E(\text{MS}_A)$ does, of course, reduce to zero for balanced data. Thus, when $n_{ij} = n, n_{i.} = bn$, $n_{.j} = an$, and $N = abn$, the coefficient of σ_β^2 in $E(\text{MS}_A)$ is

$$\begin{aligned} r_8 = \frac{k_3 - k_2}{a-1} &= \frac{\sum_{i=1}^a \frac{\sum_{j=1}^b n_{ij}^2}{n_{i.}} - \frac{1}{N}\sum_{j=1}^b n_{.j}^2}{a-1} \\ &= \frac{a\left(\frac{bn^2}{bn}\right) - b\left(\frac{a^2n^2}{abn}\right)}{a-1} \\ &= 0. \end{aligned}$$

Similarly, the coefficient of σ_α^2 in $E(\text{MS}_A)$ becomes

$$\begin{aligned} r_9 = \frac{N - k_1}{a-1} &= \frac{N - \frac{1}{N}\sum_{i=1}^a n_{i.}^2}{a-1} \\ &= \frac{abn - \frac{ab^2n^2}{abn}}{a-1} \\ &= bn, \end{aligned}$$

and that of $\sigma_{\alpha\beta}^2$ reduces to

$$\begin{aligned} r_7 = \frac{k_3 - k_{23}}{a-1} &= \frac{\sum_{i=1}^a \frac{\sum_{j=1}^b n_{ij}^2}{n_{i.}} - \frac{1}{N}\sum_{i=1}^a\sum_{j=1}^b n_{ij}^2}{a-1} \\ &= \frac{a\left(\frac{bn^2}{bn}\right) - a\left(\frac{bn^2}{abn}\right)}{a-1} \\ &= n. \end{aligned}$$

Hence, for balanced data

$$E(\text{MS}_A) = \sigma_e^2 + n\sigma_{\alpha\beta}^2 + bn\sigma_\alpha^2,$$

which is the same result as given in Table 4.2. Similar remarks and simplifications apply for the MS_B and MS_{AB} terms. The results on expected mean squares seem to have been first derived by Crump (1947). Gaylor et al. (1970) discuss the procedures for calculating expected mean squares using the abbreviated Dolittle and square root methods.

13.4 ESTIMATION OF VARIANCE COMPONENTS

In this section, we consider some methods of estimation of variance components $\sigma_e^2, \sigma_{\alpha\beta}^2, \sigma_\beta^2$, and σ_α^2.

13.4.1 ANALYSIS OF VARIANCE ESTIMATORS

The analysis of variance or Henderson's Method I for estimating variance components is to equate the sums of squares or mean squares in Table 13.1 to their respective expected values. The resulting equations are

$$\begin{aligned}
\mathrm{SS}_A &= T_A - T_\mu = (N - k_1)\sigma_\alpha^2 + (k_3 - k_2)\sigma_\beta^2 + (k_3 - k_{23})\sigma_{\alpha\beta}^2 + (a-1)\sigma_e^2, \\
\mathrm{SS}_B &= T_B - T_\mu = (k_4 - k_1)\sigma_\alpha^2 + (N - k_2)\sigma_\beta^2 + (k_4 - k_{23})\sigma_{\alpha\beta}^2 + (b-1)\sigma_e^2, \\
\mathrm{SS}_{AB} &= T_{AB} - T_A - T_B + T_\mu \qquad (13.4.1) \\
&= (k_1 - k_4)\sigma_\alpha^2 + (k_2 - k_3)\sigma_\beta^2 + (N - k_3 - k_4 + k_{23})\sigma_{\alpha\beta}^2 \\
&\quad + (s - a - b + 1)\sigma_e^2,
\end{aligned}$$

and

$$\mathrm{SS}_E = T_0 - T_{AB} = (N - s)\sigma_e^2.$$

The variance component estimators are obtained by solving the equations in (13.4.1) for σ_α^2, σ_β^2, $\sigma_{\alpha\beta}^2$, and σ_e^2. The estimators thus obtained are given by

$$\hat{\sigma}_{e,\mathrm{ANOV}}^2 = \frac{\mathrm{SS}_E}{N - s}, \qquad (13.4.2)$$

and

$$\begin{bmatrix} \hat{\sigma}_{\alpha,\mathrm{ANOV}}^2 \\ \hat{\sigma}_{\beta,\mathrm{ANOV}}^2 \\ \hat{\sigma}_{\alpha\beta,\mathrm{ANOV}}^2 \end{bmatrix} = \boldsymbol{P}^{-1} \begin{bmatrix} \mathrm{SS}_A - (a-1)\hat{\sigma}_{e,\mathrm{ANOV}}^2 \\ \mathrm{SS}_B - (b-1)\hat{\sigma}_{e,\mathrm{ANOV}}^2 \\ \mathrm{SS}_{AB} - (s - a - b + 1)\hat{\sigma}_{e,\mathrm{ANOV}}^2 \end{bmatrix}, \qquad (13.4.3)$$

where

$$\boldsymbol{P} = \begin{bmatrix} N - k_1 & k_3 - k_2 & k_3 - k_{23} \\ k_4 - k_1 & N - k_2 & k_4 - k_{23} \\ k_1 - k_4 & k_2 - k_3 & N - k_3 - k_4 + k_{23} \end{bmatrix}.$$

Further simplification of (13.4.3) yields (Searle, 1958; 1971, p. 481)

$$\begin{aligned}
\hat{\sigma}_{\alpha\beta,\mathrm{ANOV}}^2 = \Bigg[& \frac{\mathrm{N} - k_1}{N - k_4}\{\mathrm{SS}_{AB} + \mathrm{SS}_A - (s - b)\hat{\sigma}_{e,\mathrm{ANOV}}^2\} \\
& + \frac{k_3 - k_2}{N - k_3}\{\mathrm{SS}_{AB} + \mathrm{SS}_B - (s - a)\hat{\sigma}_{e,\mathrm{ANOV}}^2\} \\
& - \{\mathrm{SS}_A - (a-1)\hat{\sigma}_{e,\mathrm{ANOV}}^2\}\Bigg] \Big/ (N - k_1 - k_2 + k_{23}), \qquad (13.4.4)
\end{aligned}$$

$$\hat{\sigma}_{\beta,\mathrm{ANOV}}^2 = \frac{1}{N - k_3}\{\mathrm{SS}_{AB} + \mathrm{SS}_B - (s - a)\hat{\sigma}_{e,\mathrm{ANOV}}^2\} - \hat{\sigma}_{\alpha\beta,\mathrm{ANOV}}^2,$$

and

$$\hat{\sigma}^2_{\alpha,\text{ANOV}} = \frac{1}{N - k_4}\{\text{SS}_{AB} + \text{SS}_A - (s - b)\hat{\sigma}^2_{e,\text{ANOV}}\} - \hat{\sigma}^2_{\alpha\beta,\text{ANOV}}.$$

13.4.2 FITTING-CONSTANTS-METHOD ESTIMATORS

The model in (13.1.1) involves the terms μ, α_i, β_j, and $(\alpha\beta)_{ij}$. The sum of squares for fitting it is therefore denoted by $R(\mu, \alpha, \beta, \alpha\beta)$. Similarly, let $R(\mu, \alpha, \beta)$, $R(\mu, \alpha)$, $R(\mu, \beta)$, and $R(\mu)$ be the reductions due to fitting the submodels

$$\begin{aligned} y_{ijk} &= \mu + \alpha_i + \beta_j + e_{ijk}, \\ y_{ijk} &= \mu + \alpha_i + e_{ijk}, \\ y_{ijk} &= \mu + \beta_j + e_{ijk}, \end{aligned} \tag{13.4.5}$$

and

$$y_{ijk} = \mu + e_{ijk},$$

respectively. Then it can be shown that (see, e.g., Searle, 1971, pp. 292–293; 1987, pp. 124–125)

$$\begin{aligned} R(\mu, \alpha, \beta, \alpha\beta) &= T_{AB}, \\ R(\mu, \alpha, \beta) &= T_A + \boldsymbol{r}'\boldsymbol{C}^{-1}\boldsymbol{r}, \\ R(\mu, \alpha) &= T_A, \\ R(\mu, \beta) &= T_B, \end{aligned} \tag{13.4.6}$$

and

$$R(\mu) = T_\mu,$$

where $\boldsymbol{r}$ and $\boldsymbol{C}$ are as defined following (12.4.5).

Now, the analysis of variance based on α adjusted for β (fitting β before α) is as given in Table 13.2. From Table 13.2, the terms (quadratics) needed in the fitting-constants-method of estimating variance components are

$$\begin{aligned} R(\mu) &= T_\mu, \\ R(\beta|\mu) &= R(\mu, \beta) - R(\mu) = T_B - T_\mu, \\ R(\alpha|\mu, \beta) &= R(\mu, \alpha, \beta) - R(\mu, \beta) = R(\mu, \alpha, \beta) - T_B, \\ R(\alpha\beta|\mu, \alpha, \beta) &= R(\mu, \alpha, \beta, \alpha\beta) - R(\mu, \alpha, \beta) = T_{AB} - R(\mu, \alpha, \beta), \end{aligned} \tag{13.4.7}$$

and

$$\text{SS}_E = R_0 - R(\mu, \alpha, \beta, \alpha\beta) = T_0 - T_{AB}.$$

TABLE 13.2 Analysis of variance based on α adjusted for β.

Source of variation	Degrees of freedom	Sum of squares
Mean μ	1	$R(\mu)$
β **adjusted for** μ	$b-1$	$R(\beta\|\mu)$
α **adjusted for** μ **and** β	$a-1$	$R(\alpha\|\mu,\beta)$
$(\alpha\beta)$ **adjusted for** $\mu, \alpha,$ **and** β	$s-a-b+1$	$R(\alpha\beta\|\mu,\alpha,\beta)$
Error	$N-s$	SS_E

Remarks:

(i) The quadratics in (13.4.7) lead to the following partitioning of the total sum of squares (uncorrected for the mean):

$$SS_T = R(\mu) + R(\beta|\mu) + R(\alpha|\mu,\beta) + R(\alpha\beta|\mu,\alpha,\beta) + SS_E,$$

(ii) The quadratics in (13.4.7) are equivalent to SAS Type I sums of squares when ordering the factors as B, A, and $A \times B$. ♦

The expected values of the sums of squares in Table 13.2 are (see, e.g., Searle, 1958; Searle et al., 1992, pp. 214–217)

$$\begin{aligned} E\{SS_E\} &= (N-s)\sigma_e^2, \\ E\{R(\alpha\beta|\mu,\alpha,\beta)\} &= (s-a-b+1)\sigma_e^2 + h\sigma_{\alpha\beta}^2, \\ E\{R(\alpha|\mu,\beta)\} &= (a-1)\sigma_e^2 + (N-k_4-h)\sigma_{\alpha\beta}^2 + (N-k_4)\sigma_\alpha^2, \end{aligned} \tag{13.4.8}$$

and

$$E\{R(\beta|\mu)\} = (b-1)\sigma_e^2 + (k_4-k_{23})\sigma_{\alpha\beta}^2 + (N-k_2)\sigma_\beta^2 + (k_4-k_1)\sigma_\alpha^2,$$

where[2]

$$h = N - \sum_{i=1}^{a}\lambda_i - \text{tr}\left\{C^{-1}\sum_{i=1}^{a} F_i\right\}$$

with the matrix $\boldsymbol{C}$ being defined following (12.4.5) and the matrix $\boldsymbol{F}_i$ is defined as

$$\boldsymbol{F}_i = \{f_{i,jj'}\}, \tag{13.4.9}$$

[2]For a numerical example illustrating the computation of h, see Searle and Henderson (1961).

with

$$f_{i,jj} = \frac{n_{ij}^2}{n_{i.}}(\lambda_i + n_{i.} - 2n_{ij}),$$
$$f_{i,jj'} = \frac{n_{ij}n_{ij'}}{n_{i.}}(\lambda_i - n_{ij} - n_{ij'}) \quad \text{for } j \neq j' \left(\sum_{j=1}^{b} f_{i,jj'} = 0 \right),$$

and

$$\lambda_i = \sum_{j=1}^{b} \frac{n_{ij}^2}{n_{i.}},$$

for $i = 1, \ldots, a$ and $j, j' = 1, \ldots, b-1$. The variance component estimators are obtained by equating the sums of squares in Table 13.2 to their respective expected values given in (13.4.8). The resulting estimators are

$$\begin{aligned}
\hat{\sigma}^2_{e,\text{FTC1}} &= \frac{\text{SS}_E}{N-s}, \\
\hat{\sigma}^2_{\alpha\beta,\text{FTC1}} &= \frac{R(\alpha\beta|\mu,\alpha,\beta) - (s-a-b+1)\hat{\sigma}^2_{e,\text{FTC1}}}{h}, \\
\hat{\sigma}^2_{\alpha,\text{FTC1}} &= \frac{R(\alpha|\mu,\beta) - (N-k_4-h)\hat{\sigma}^2_{\alpha\beta,\text{FTC1}} - (a-1)\hat{\sigma}^2_{e,\text{FTC1}}}{N-k_4},
\end{aligned} \tag{13.4.10}$$

and

$$\hat{\sigma}^2_{\beta,\text{FTC1}} = \frac{R(\beta|\mu) - (k_4-k_1)\hat{\sigma}^2_{\alpha} - (k_4-k_{23})\hat{\sigma}^2_{\alpha\beta,\text{FTC1}} - (b-1)\hat{\sigma}^2_{e,\text{FTC1}}}{N-k_2}.$$

The analysis of variance given in Table 13.2 carries with it a sequential connotation of first fitting μ, then μ and β, and then μ, β, and α. Because of the symmetry of the crossed-classification model in (13.1.1), an alternative approach for the analysis of variance would be to fit α before β. The resulting sum of squares terms for the analysis of variance are

$$\begin{aligned}
R(\mu) &= T(\mu), \\
R(\alpha|\mu) &= R(\mu,\alpha) - R(\mu) = T_A - T_\mu, \\
R(\beta|\mu,\alpha) &= R(\mu,\alpha,\beta) - R(\mu,\alpha) = R(\mu,\alpha,\beta) - T_A, \\
R(\alpha\beta|\mu,\alpha,\beta) &= R(\mu,\alpha,\beta,\alpha\beta) - R(\mu,\alpha,\beta) = T_{AB} - R(\mu,\alpha,\beta),
\end{aligned} \tag{13.4.11}$$

and

$$\text{SS}_E = R(0) - R(\mu,\alpha,\beta,\alpha\beta) = T_0 - T_{AB}.$$

The analysis of variance based on β adjusted for α can then be written as in Table 13.3.

TABLE 13.3 Analysis of variance based on β adjusted for α.

Source of variation	Degrees of freedom	Sum of squares
Mean μ	1	$R(\mu)$
α **adjusted for** μ	$a-1$	$R(\alpha\|\mu)$
β **adjusted for** μ **and** α	$b-1$	$R(\beta\|\mu,\alpha)$
$(\alpha\beta)$ **adjusted for** $\mu, \alpha,$ **and** β	$s-a-b+1$	$R(\alpha\beta\|\mu,\alpha,\beta)$
Error	$N-s$	SS_E

Remarks:

(i) The quadratics in (13.4.11) lead to the following partitioning of the total sum of squares (uncorrected for the mean):

$$SS_T = R(\mu) + R(\alpha|\mu) + R(\beta|\mu,\alpha) + R(\alpha\beta|\mu,\alpha,\beta) + SS_E.$$

(ii) The quadratics in (13.4.11) are equivalent to SAS Type I sums of squares when ordering the factors as A, B, and $A \times B$ ♦

In view of symmetry, the expected values of sums of squares in Table 13.3 follow readily from the results in (13.4.8) and are given by

$$\begin{aligned} E\{SS_E\} &= (N-s)\sigma_e^2, \\ E\{R(\alpha\beta|\mu,\alpha,\beta)\} &= (s-a-b+1)\sigma_e^2 + h\sigma_{\alpha\beta}^2, \\ E\{R(\beta|\mu,\alpha)\} &= (b-1)\sigma_e^2 + (N-k_3-h)\sigma_{\alpha\beta}^2 + (N-k_3)\sigma_\beta^2, \end{aligned} \tag{13.4.12}$$

and

$$\begin{aligned} E\{R(\alpha|\mu)\} = {} & (a-1)\sigma_e^2 + (k_3-k_{23})\sigma_{\alpha\beta}^2 \\ & + (k_3-k_2)\sigma_\beta^2 + (N-k_1)\sigma_\alpha^2. \end{aligned}$$

The variance component estimators are obtained by equating the sums of squares in Table 13.3 to their respective expected values given in (13.4.12). The resulting estimators of the variance components are

$$\begin{aligned} \hat\sigma_{e,\text{FTC2}}^2 &= \frac{SS_E}{N-s}, \\ \hat\sigma_{\alpha\beta,\text{FTC2}}^2 &= \frac{R(\alpha\beta|\mu,\alpha,\beta) - (s-a-b+1)\hat\sigma_{e,\text{FTC2}}^2}{h}, \\ \hat\sigma_{\beta,\text{FTC2}}^2 &= \frac{R(\beta|\mu,\alpha) - (N-k_3-h)\hat\sigma_{\alpha\beta,\text{FTC2}}^2 - (b-1)\hat\sigma_{e,\text{FTC2}}^2}{N-k_3}, \end{aligned} \tag{13.4.13}$$

and

$$\hat{\sigma}^2_{\alpha,\text{FTC2}} = \frac{R(\alpha|\mu) - (k_3 - k_2)\hat{\sigma}^2_{\beta,\text{FTC2}} - (k_3 - k_{23})\hat{\sigma}^2_{\alpha\beta,\text{FTC2}} - (a-1)\hat{\sigma}^2_e}{N - k_1}.$$

It should be noted that the estimators $\hat{\sigma}^2_{e,\text{FTC2}}$ and $\hat{\sigma}^2_{\alpha\beta,\text{FTC2}}$ given in (13.4.13) are the same as those given in (13.4.10), but $\hat{\sigma}^2_{\alpha,\text{FTC2}}$ and $\hat{\sigma}^2_{\beta,\text{FTC2}}$ are not. This is an obvious disadvantage of the fitting-constants-method; that it does not yield a unique set of estimators of variance components.

A third possible set of estimators of variance components would be to consider the estimators based on adjusted quadratics by adjusting each term by all other terms that do not contain the effect in question. Such quadratics and their expectations are

$$\begin{aligned} E\{\text{SS}_E\} &= (N-s)\sigma^2_e, \\ E\{R(\alpha\beta|\mu,\alpha,\beta)\} &= (s-a-b+1)\sigma^2_e + h\sigma^2_{\alpha\beta}, \\ E\{R(\alpha|\mu,\beta)\} &= (a-1)\sigma^2_e + (N-k_4-h)\sigma^2_{\alpha\beta} + (N-k_4)\sigma^2_\alpha, \end{aligned} \tag{13.4.14}$$

and

$$E\{R(\beta|\mu,\alpha)\} = (b-1)\sigma^2_e + (N-k_3-h)\sigma^2_{\alpha\beta} + (N-k_3)\sigma^2_\beta.$$

The resulting estimators are then given by

$$\begin{aligned} \hat{\sigma}^2_{e,\text{FTC3}} &= \frac{\text{SS}_E}{N-s}, \\ \hat{\sigma}^2_{\alpha\beta,\text{FTC3}} &= \frac{R(\alpha\beta|\mu,\alpha,\beta) - (s-a-b+1)\hat{\sigma}^2_{e,\text{FTC3}}}{h}, \\ \hat{\sigma}^2_{\alpha,\text{FTC3}} &= \frac{R(\alpha|\mu,\beta) - (N-k_4-h)\hat{\sigma}^2_{\alpha\beta,\text{FTC3}} - (a-1)\hat{\sigma}^2_{e,\text{FTC3}}}{N-k_4}, \end{aligned} \tag{13.4.15}$$

and

$$\hat{\sigma}^2_{\beta,\text{FTC3}} = \frac{R(\beta|\mu,\alpha) - (N-k_3-h)\hat{\sigma}^2_{\alpha\beta,\text{FTC3}} - (b-1)\hat{\sigma}^2_{e,\text{FTC3}}}{N-k_3}.$$

Remarks:

(i) The quadratics in (13.4.14) do not lead to the following partitioning of the total sum of squares (uncorrected for the mean):

$$\text{SS}_T = R(\mu) + R(\beta|\mu,\alpha) + R(\alpha|\mu,\beta) + R(\alpha\beta|\mu,\alpha,\beta) + \text{SS}_E.$$

(ii) The quadratics in (13.4.14) are equivalent to SAS Type II sums of squares. ♦

It should be mentioned that in addition to the three sets of sums of squares considered above, there are a number of other sets that could be used; e.g., $R(\alpha|\mu)$, $R(\beta|\mu)$, $R(\alpha\beta|\mu, \alpha, \beta)$, and SS_E; and so on. The number of such sets of sums of squares that can be used in a higher-order model increases rather rapidly. For example, in an unbalanced crossed classification model involving three, four, or five factors, even without interactions, there would be 6, 24, and 120 sets of sums of squares that can be used. Moreover, there is no theoretical basis whatsoever for deciding which set of sums of squares is to be preferred. Thus the procedure suffers from the lack of uniqueness which is a serious drawback limiting its usefulness.

13.4.3 ANALYSIS OF MEANS ESTIMATORS

As discussed in Section 10.4, the approach of the analysis of means method, when all $n_{ij} \geq 1$, is to treat the means of those cells as observations and then carry out a balanced data analysis.[3] The calculations for the analysis are rather straightforward as illustrated below. We first discuss the unweighted analysis and then the weighted analysis.

13.4.3.1 Unweighted Means Analysis

For the observations y_{ijk}s from the model in (13.1.1), let x_{ij} be the cell mean defined by

$$x_{ij} = \bar{y}_{ij.} = \sum_{k=1}^{n_{ij}} \frac{y_{ijk}}{n_{ij}}. \tag{13.4.16}$$

Further, define

$$\bar{x}_{i.} = \frac{\sum_{j=1}^{b} x_{ij}}{b}, \qquad \bar{x}_{.j} = \frac{\sum_{i=1}^{a} x_{ij}}{a},$$

and

$$\bar{x}_{..} = \frac{\sum_{i=1}^{a} \sum_{j=i}^{b} x_{ij}}{ab}.$$

Then the analysis of variance for the unweighted means analysis is shown in Table 13.4.

The quantities in the sum of squares column are defined by

$$\text{SS}_{Au} = b\bar{n}_h \sum_{i=1}^{a} (\bar{x}_{i.} - \bar{x}_{..})^2,$$

[3]The procedure can be used if there is one empty cell. For an example with one empty cell in a 3 × 3 design, see Bush and Anderson (1963).

TABLE 13.4 Analysis of variance with unweighted sums of squares for the model in (13.1.1).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected square mean
Factor A	$a-1$	SS_{Au}	MS_{Au}	$\sigma_e^2+\bar{n}_h\sigma_{\alpha\beta}^2+b\bar{n}_h\sigma_\alpha^2$
Factor B	$b-1$	SS_{Bu}	MS_{Bu}	$\sigma_e^2+\bar{n}_h\sigma_{\alpha\beta}^2+a\bar{n}_h\sigma_\beta^2$
Interaction AB	$(a-1)(b-1)$	SS_{ABu}	MS_{ABu}	$\sigma_e^2+\bar{n}_h\sigma_{\alpha\beta}^2$
Error	$N-ab$	SS_E	MS_E	σ_e^2

$$\begin{aligned} SS_{Bu} &= a\bar{n}_h\sum_{j=1}^{b}(\bar{x}_{.j}-\bar{x}_{..})^2, \\ SS_{ABu} &= \bar{n}_h\sum_{i=1}^{a}\sum_{j=1}^{b}(x_{ij}-\bar{x}_{i.}-\bar{x}_{.j}+\bar{x}_{..})^2, \end{aligned} \tag{13.4.17}$$

and

$$SS_E=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}}(y_{ijk}-\bar{y}_{ij.})^2,$$

where

$$\bar{n}_h=\frac{1}{\sum_{i=1}^{a}\sum_{j=1}^{b}n_{ij}^{-1}/ab}.$$

Note that in the fixed effects version of the model in (13.1.1), $\sigma_e^2/\bar{n}_h$ represents the average variance of the cell means. Thus $\bar{n}_h$ acts like n, the common cell frequency for the case corresponding to the balanced model. For some further discussion on the use of $\bar{n}_h$ in the definition of unweighted sums of squares, see Khuri (1998). The mean squares are obtained in the usual way by dividing the sums of squares by the corresponding degrees of freedom. For a method of derivation of the results on expected mean squares, see Hirotsu (1966) and Mostafa (1967).

The following features of the above analysis are worth noting:

(i) The means of the x_{ij}s are calculated in the usual manner, i.e.,

$$\bar{x}_{i.}=\sum_{j=1}^{b}\frac{x_{ij}}{b},\qquad \bar{x}_{.j}=\sum_{i=1}^{a}\frac{x_{ij}}{a},\quad\text{and}\quad \bar{x}_{..}=\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{x_{ij}}{ab}.$$

(ii) The error sum of squares, SS_E, is calculated exactly as in the conventional analysis of variance given in Section 13.2.

(iii) The individual sums of squares do not add up to the total sum of squares. The first three sums of squares, i.e., SS_{Au}, SS_{Bu}, and SS_{ABu} add up to $\bar{n}_h \sum_{i=1}^{a} \sum_{j=1}^{b} (x_{ij} - \bar{x}_{..})^2$, but all four do not add up to the total sum of squares.

(iv) The error sum of squares SS_E has σ_e^2 times a chi-square distribution with N-ab degrees of freedom.

(v) The sums of squares SS_{Au}, SS_{Bu}, and SS_{ABu} do not have a scaled chi-square distribution, as in the case of the balanced analogue of the model in (13.1.1); nor are they independent of SS_E. However, it can be shown that $\mathrm{SS}_{Au}/(\sigma_e^2 + \bar{n}_h\sigma_{\alpha\beta}^2 + b\bar{n}_h\sigma_\alpha^2)$, $\mathrm{SS}_{Bu}/(\sigma_e^2 + \bar{n}_h\sigma_{\alpha\beta}^2 + a\bar{n}_h\sigma_\beta^2)$, and $\mathrm{SS}_{ABu}/(\sigma_e^2 + \bar{n}_h\sigma_{\alpha\beta}^2)$ are approximately distributed as independent chi-square variates with $a-1$, $b-1$, and $(a-1)(b-1)$ degrees of freedom, respectively (see, e.g., Hirotsu, 1968; Khuri, 1998).

The estimators of the variance components, as usual, are obtained by equating the means squares to their respective expected values and solving the resulting equations for the variance components. The estimators are given as follows:

$$\begin{aligned}
\hat{\sigma}^2_{e,\mathrm{UME}} &= \mathrm{MS}_E, \\
\hat{\sigma}^2_{\alpha\beta,\mathrm{UME}} &= \frac{\mathrm{MS}_{ABu} - \mathrm{MS}_E}{\bar{n}_h}, \\
\hat{\sigma}^2_{\beta,\mathrm{UME}} &= \frac{\mathrm{MS}_{Bu} - \mathrm{MS}_{ABu}}{a\bar{n}_h},
\end{aligned} \tag{13.4.18}$$

and

$$\hat{\sigma}^2_{\alpha,\mathrm{UME}} = \frac{\mathrm{MS}_{Au} - \mathrm{MS}_{ABu}}{b\bar{n}_h}.$$

13.4.3.2 Weighted Means Analysis

The weighted square of means analysis consists of weighting the terms in the sums of squares SS_{Au} and SS_{Bu}, defined in (13.4.17) in the unweighted analysis, in inverse proportion to the variance of the term concerned. Thus, instead of SS_A and SS_B given by

$$\mathrm{SS}_{Au} = b\bar{n}_h \sum_{i=1}^{a} (\bar{x}_{i.} - \bar{x}_{..})^2, \qquad \mathrm{SS}_{Bu} = a\bar{n}_h \sum_{j=1}^{b} (\bar{x}_{.j} - \bar{x}_{..})^2,$$

we use

$$\mathrm{SS}_{Aw} = \sum_{i=1}^{a} w_i (\bar{x}_{i.} - \bar{x}^{w}_{..})^2, \qquad \mathrm{SS}_{Bw} = \sum_{j=1}^{b} \nu_j (\bar{x}_{.j} - \bar{x}^{\nu}_{..})^2,$$

TABLE 13.5 Analysis of variance with weighted sums of squares for the model in (13.1.1).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected square mean
Factor A	$a-1$	SS_{Aw}	MS_{Aw}	$\sigma_e^2+\theta_1(\sigma_{\alpha\beta}^2+b\sigma_\alpha^2)$
Factor B	$b-1$	SS_{Bw}	MS_{Bw}	$\sigma_e^2+\theta_2(\sigma_{\alpha\beta}^2+a\sigma_\beta^2)$
Interaction AB	$(a-1)(b-1)$	SS_{ABw}	MS_{ABw}	$\sigma_e^2+\theta_3\sigma_{\alpha\beta}^2$
Error	$N-ab$	SS_E	MS_E	σ_e^2

where

$$w_i = \sigma^2/\operatorname{var}(\bar{x}_{i.}), \qquad \nu_j = \sigma^2/\operatorname{Var}(\bar{x}_{.j}),$$

and $\bar{x}_{..}^w$ and $\bar{x}_{..}^\nu$ are weighted means of $\bar{x}_{i.}$s and $\bar{x}_{.j}$s weighted by w_i and ν_j, respectively, i.e.,

$$\bar{x}_{..}^w = \sum_{i=1}^a w_i\bar{x}_{i.}\Big/\sum_{i=1}^a w_i, \qquad \bar{x}_{..}^\nu = \sum_{j=1}^b \nu_j\bar{x}_{.j}\Big/\sum_{j=1}^b \nu_{j}.$$

There are a variety of weights that can be used for w_i and ν_j as discussed by Gosslee and Lucas (1965). A weighted analysis of variance based on weights originally proposed by Yates (1934) (for a fixed effects model) is shown in Table 13.5. (See also Searle et al. (1992, pp. 220–221).) It is calculated by the SAS® GLM or SPSS® GLM procedures using Type III sums of squares.

The quantities in the sum of squares column are given by

$$\begin{aligned}
\mathrm{SS}_{Aw} &= \sum_{i=1}^a \phi_i(\bar{x}_{i.} - \bar{x}_{..}^\phi)^2,\\
\mathrm{SS}_{Bw} &= \sum_{j=1}^b \psi_j(\bar{x}_{.j} - \bar{x}_{..}^\psi)^2,\\
\mathrm{SS}_{ABw} &= R(\alpha\beta|\mu,\alpha,\beta),
\end{aligned} \tag{13.4.19}$$

and

$$\mathrm{SS}_E = \sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^{n_{ij}}(y_{ijk}-\bar{y}_{ij.})^2,$$

where

$$\bar{x}_{..}^\phi = \sum_{i=1}^a \phi_i\bar{x}_{i.}\Big/\sum_{i=1}^a \phi_i,$$

$$\bar{x}_{..}^{\psi} = \sum_{j=1}^{b} \psi_j \bar{x}_{.j} \Big/ \sum_{j=1}^{b} \psi_j,$$

$$\phi_i = b^2 \Big/ \sum_{j=1}^{b} n_{ij}^{-1}, \qquad \psi_j = a^2 \Big/ \sum_{i=1}^{a} n_{ij}^{-1},$$

and $R(\alpha\beta|\mu, \alpha, \beta)$ is defined in (13.4.11). The quantities θ_1, θ_2, and θ_3 in the expected mean square column are defined as

$$\theta_1 = \left\{ \sum_{i=1}^{a} \phi_i - \sum_{i=1}^{a} \phi_i^2 \Big/ \sum_{i=1}^{a} \phi_i \right\} \Big/ b(a-1),$$
$$\theta_2 = \left\{ \sum_{j=1}^{b} \psi_j - \sum_{j=1}^{b} \psi_j^2 \Big/ \sum_{j=1}^{b} \psi_j \right\} \Big/ a(b-1), \tag{13.4.20}$$

and

$$\theta_3 = h/(a-1)(b-1),$$

where h is defined in (13.4.8). Note that the sum of squares for the error term is the same in both unweighted and weighted analyses.

The estimators of the variance components obtained using the weighted analysis are

$$\begin{aligned} \hat{\sigma}_{e,\text{WME}}^2 &= \text{MS}_E, \\ \hat{\sigma}_{\alpha\beta,\text{WME}}^2 &= \frac{\text{MS}_{ABw} - \text{MS}_E}{\theta_3}, \\ \hat{\sigma}_{\beta,\text{WME}}^2 &= \frac{\text{MS}_{Bw} - (\theta_2/\theta_3)\text{MS}_{ABw} - (1 - \theta_2/\theta_3)\text{MS}_E}{a\theta_2}, \end{aligned} \tag{13.4.21}$$

and

$$\hat{\sigma}_{\alpha,\text{WME}}^2 = \frac{\text{MS}_{Aw} - (\theta_1/\theta_3)\text{MS}_{ABw} - (1 - \theta_1/\theta_3)\text{MS}_E}{b\theta_1}.$$

It can be seen that the estimator of σ_e^2 is the same in both unweighted and weighted analyses, but those of $\theta_{\alpha\beta}^2$, σ_β^2, and σ_α^2 are different. Further, in this case, $\hat{\theta}_{\alpha\beta,\text{WME}}^2$ is the same as $\hat{\theta}_{\alpha\beta,\text{FTC1}}^2$ and $\hat{\theta}_{\alpha\beta,\text{FTC2}}^2$ in the fitting-constant method estimation.

13.4.4 SYMMETRIC SUMS ESTIMATORS

For symmetric sums estimators, we consider expected values for products and squares of differences of observations. From the model in (13.1.1), expected values of products of observations are

$$E(y_{ijk}\, y_{i'j'k'}) = \begin{cases} \mu^2, & i \neq i', \quad j \neq j', \\ \mu^2 + \sigma_\alpha^2, & i = i', \quad j \neq j', \\ \mu^2 + \sigma_\beta^2, & i \neq i', \quad j = j', \\ \mu^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2 + \sigma_\alpha^2, & i = i', \quad j = j', \quad k \neq k', \\ \mu^2 + \sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2 + \sigma_\alpha^2, & i = i', \quad j = j', \quad k = k', \end{cases} \tag{13.4.22}$$

where $i, i' = 1, 2, \ldots, a$; $j, j' = 1, 2 \ldots, b$; $k = 1, 2 \ldots, n_{ij}$; $k' = 1, 2, \ldots, n_{i'j'}$. Now, the normalized symmetric sums of the terms in (13.4.22) are

$$\begin{aligned} g_m &= \frac{\sum_{\substack{i,i' \\ i \neq i'}} \sum_{\substack{j,j' \\ j \neq j'}} y_{ij.} y_{i'j'.}}{N^2 - \sum_{i=1}^a n_{i.}^2 - \sum_{j=1}^b n_{.j}^2 + \sum_{i=1}^a \sum_{j=1}^b n_{ij}^2} \\ &= \frac{\sum_{i=1}^a \sum_{j=1}^b y_{ij.}^2 - \sum_{i=1}^a y_{i..}^2 - \sum_{j=1}^b y_{.j.}^2 + y_{...}^2}{N^2 - k_1 - k_2 + k_{12}}, \\ g_A &= \frac{\sum_{i=1}^a \sum_{\substack{j,j' \\ j \neq j'}} y_{ij.} y_{ij.'}}{\sum_{i=1}^a n_{i.}^2 - \sum_{i=1}^a \sum_{j=1}^b n_{ij}^2} = \frac{\sum_{i=1}^a y_{i..}^2 - \sum_{i=1}^a \sum_{j=1}^b y_{ij.}^2}{k_1 - k_{12}}, \\ g_B &= \frac{\sum_{\substack{i,i' \\ i \neq i'}} \sum_{j=1}^b y_{ij.} y_{i'j.}}{\sum_{j=1}^b n_{.j}^2 - \sum_{i=1}^a \sum_{j=1}^b n_{ij}^2} = \frac{\sum_{j=1}^b y_{.j.}^2 - \sum_{i=1}^a \sum_{j=1}^b y_{ij.}^2}{k_2 - k_{12}}, \\ g_{AB} &= \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{\substack{k,k' \\ k \neq k'}} y_{ijk} y_{ijk'}}{\sum_{i=1}^a \sum_{j=1}^b n_{ij}^2 - \sum_{i=1}^a \sum_{j=1}^b n_{ij}} \\ &= \frac{\sum_{i=1}^a \sum_{j=1}^b y_{ij.}^2 - \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{n_{ij}} y_{ijk}^2}{k_{12} - N}, \end{aligned}$$

and

$$g_E = \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{n_{ij}} y_{ijk} y_{ijk}}{\sum_{i=1}^a \sum_{j=1}^b n_{ij}} = \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{n_{ij}} y_{ijk}^2}{N},$$

where

$$n_{i.} = \sum_{j=1}^b n_{ij}, \qquad n_{.j} = \sum_{i=1}^a n_{ij}, \qquad N = \sum_{i=1}^a \sum_{j=1}^b n_{ij},$$

$$k_1 = \sum_{i=1}^a n_{i.}^2, \qquad k_2 = \sum_{j=1}^b n_{.j}^2, \qquad k_{12} = \sum_{i=1}^a \sum_{j=1}^b n_{ij}^2.$$

Equating g_m, g_A, g_B, g_{AB}, and g_E to their respective expected values, we obtain

$$\begin{aligned} \mu^2 &= g_m, \\ \mu^2 + \sigma_\alpha^2 &= g_A, \\ \mu^2 + \sigma_\beta^2 &= g_B, \\ \mu^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2 + \sigma_\alpha^2 &= g_{AB}, \end{aligned} \tag{13.4.23}$$

and

$$\mu^2 + \sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2 + \sigma_\alpha^2 = g_E.$$

The variance component estimators obtained by solving the equations in (13.4.23) are (Koch, 1967)

$$\begin{aligned} \hat{\sigma}_{\alpha,\mathrm{SSP}}^2 &= g_A - g_m, \\ \hat{\sigma}_{\beta,\mathrm{SSP}}^2 &= g_B - g_m, \\ \hat{\sigma}_{\alpha\beta,\mathrm{SSP}}^2 &= g_{AB} - g_A - g_B + g_m, \end{aligned} \tag{13.4.24}$$

and

$$\hat{\sigma}_{e,\mathrm{SSP}}^2 = g_E - g_{AB}.$$

The estimators in (13.4.24), by construction, are unbiased; and they reduce to the analysis of variance estimators in the case of balanced data. However, they are not translation invariant, i.e., they may change in values if the same constant is added to all the observations and their variances are functions of μ. This drawback is overcome by using the symmetric sums of squares of differences rather than the products.

For symmetric sums based on expected values of the squares of differences of the observations, we have

$$E\{(y_{ijk} - y_{i'j'k'})^2\} = \begin{cases} 2\sigma_e^2, & i = i', \quad j = j', \quad k \neq k', \\ 2(\sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2), & i = i', \quad j \neq j', \\ 2(\sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\alpha^2), & i \neq i', \quad j = j', \\ 2(\sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2 + \sigma_\alpha^2), & i \neq i', \quad j \neq j'. \end{cases} \tag{13.4.25}$$

The normalized (mean) symmetric sums of the terms in (13.4.25) are given by

$$h_E = \frac{1}{(k_{12} - N)} \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{\substack{kk' \\ k \neq k'}}^{n_{ij}} (y_{ijk} - y_{ijk'})^2$$

$$= \frac{2}{(k_{12} - N)} \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \left(\sum_{k=1}^{n_{ij}} y_{ijk}^2 - n_{ij} \bar{y}_{ij.}^2 \right),$$

$$h_B = \frac{1}{(k_1 - k_{12})} \sum_{i=1}^{a} \sum_{\substack{j,j' \\ j \neq j'}}^{b} \sum_{k,k'}^{n_{ij}} (y_{ijk} - y_{ij'k'})^2$$

$$= \frac{2}{(k_1 - k_{12})} \sum_{i=1}^{a} \sum_{j=1}^{b} (n_{i.} - n_{ij}) \sum_{k=1}^{n_{ij}} y_{ijk}^2 - 2g_A,$$

$$h_A = \frac{1}{(k_2 - k_{12})} \sum_{\substack{i,i' \\ i \neq i'}}^{a} \sum_{j=1}^{b} \sum_{k,k'} (y_{ijk} - y_{i'jk'})^2$$

$$= \frac{2}{(k_2 - k_{12})} \sum_{i=1}^{a} \sum_{j=1}^{b} (n_{.j} - n_{ij}) \sum_{k=1}^{n_{ij}} y_{ijk}^2 - 2g_B,$$

and

$$h_{AB} = \frac{1}{(N^2 - k_1 - k_2 + k_{12})} \sum_{\substack{i,i' \\ i \neq i'}}^{a} \sum_{\substack{j,j' \\ j \neq j'}}^{b} \sum_{k,k'} (y_{ijk} - y_{i'j'k'})^2$$

$$= \frac{2}{(N^2 - k_1 - k_2 + k_{12})} \sum_{i=1}^{a} \sum_{j=1}^{b} (N - n_{i.} - n_{.j} + n_{ij}) \sum_{k=1}^{n_{ij}} y_{ijk}^2 - 2g_m,$$

where $n_{i.}$, $n_{.j}$, N, k_1, k_2, k_{12}, g_m, g_A, and g_B are defined as before.

Equating h_A, h_B, h_{AB}, and h_E terms to their respective expected values, we obtain

$$\begin{aligned} 2\sigma_e^2 &= h_E, \\ 2(\sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2) &= h_B, \\ 2(\sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\alpha^2) &= h_A, \end{aligned} \tag{13.4.26}$$

and

$$2(\sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2 + \sigma_\alpha^2) = h_{AB}.$$

The estimators of variance components are obtained by solving the equations in (13.4.26), yielding (Koch, 1968)

$$\begin{aligned} \hat{\sigma}_{e,\text{SSS}}^2 &= h_E/2, \\ \hat{\sigma}_{\alpha,\text{SSS}}^2 &= (h_{AB} - h_B)/2, \\ \hat{\sigma}_{\beta,\text{SSS}}^2 &= (h_{AB} - h_A)/2, \end{aligned} \tag{13.4.27}$$

and

$$\hat{\sigma}^2_{\alpha\beta,\text{SSS}} = (h_A + h_B - h_{AB} - h_E)/2.$$

It can be verified that if $n_{ij} = n$ for all (i, j) then the estimators in (13.4.27) reduce to the analysis of variance estimators.

13.4.5 OTHER ESTIMATORS

The ML, REML, MINQUE, and MIVQUE estimators can be developed as special cases of the results for the general case considered in Chapter 10 and their special formulations for this model are not amenable to any simple algebraic expressions. Simple numerical techniques for computing MINQUE for several unbalanced two-way classification models have been discussed by Kleffé (1980). With the advent of the high-speed digital computer, the general results on these estimators involving matrix operations can be handled with great speed and accuracy and their explicit algebraic evaluation for this model seems to be rather unnecessary. In addition, some commonly used statistical software packages, such as SAS®, SPSS®, and BMDP®, have special routines to compute these estimates rather conveniently simply by specifying the model in question.

13.4.6 A NUMERICAL EXAMPLE

Milliken and Johnson (1992, p. 265) reported results of an experiment conducted to study the efficiency of workers in assembly lines. Three assembly plants were chosen for the experiment. Three assembly sites within each plant were then selected and a sample of four workers was taken from a large pool of available workers from each plant. Each worker was scheduled to work at each site five times, but because of logistics and other priorities, some tasks could not be completed. The data shown in Table 13.6 correspond to efficiency scores taken from only one of the three plants. We will use a two-way crossed model in (13.1.1) to analyze the data in Table 13.6. Here, $a = 3$, $b = 4$, $i = 1, 2, 3$ refer to the sites, and $j = 1, 2, 3, 4$ refer to the workers. Further, σ^2_α and σ^2_β designate variance components due to site and worker as factors, $\sigma^2_{\alpha\beta}$ is the interaction variance component, and σ^2_e denotes the error variance component. The calculations leading to the conventional analysis of variance based on Henderson's Method I were performed using SAS®GLM procedure and the results are summarized in Table 13.7.

We will now illustrate the calculations of point estimates of the variance components $\sigma^2_\alpha, \sigma^2_\beta, \sigma^2_{\alpha\beta}, \sigma^2_e$ using methods described in this section.

The analysis of variance (ANOVA) estimates in (13.4.4) based on Henderson's Method I are obtained as the solution to the following system of equations:

$$\sigma^2_e + 4.449\sigma^2_{\alpha\beta} + 0.130\sigma^2_\beta + 15.638\sigma^2_\alpha = 638.209,$$

TABLE 13.6 Data on efficiency scores for assembly line workers.

Site	Worker 1	Worker 2	Worker 3	Worker 4
	82.6	96.5	87.9	83.6
		100.1	93.5	82.7
1		101.9	88.9	87.7
		97.9	92.8	88.0
		95.9		82.5
	72.7	71.7	78.4	82.1
		72.1	80.4	79.9
2		72.4	83.4	81.9
		71.4	77.7	82.6
			81.2	78.6
	82.5	80.9	96.3	77.7
	82.1	84.0	92.4	78.6
3	82.0	82.2	92.0	77.2
		83.4	95.8	78.8
		81.5		80.5

Source: Milliken and Johnson (1992); used with permission.

TABLE 13.7 Analysis of variance for the worker efficiency-score data of Table 13.6.

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected square mean
Site	2	1,276.418	638.209	$\sigma_e^2 + 4.449\sigma_{\alpha\beta}^2 + 0.130\sigma_\beta^2 + 15.638\sigma_\alpha^2$
Worker	3	361.437	120.479	$\sigma_e^2 + 3.951\sigma_{\alpha\beta}^2 + 11.305\sigma_\beta^2 + 0.192\sigma_\alpha^2$
Interaction	6	1,002.108	167.018	$\sigma_e^2 + 3.634\sigma_{\alpha\beta}^2 - 0.043\sigma_\beta^2 - 0.096\sigma_\alpha^2$
Error	35	142.087	4.060	σ_e^2
Total	46	2,782.0497		

$$\sigma_e^2 + 3.951\sigma_{\alpha\beta}^2 + 11.305\sigma_\beta^2 + 0.192\sigma_\alpha^2 = 120.479,$$
$$\sigma_e^2 + 3.634\sigma_{\alpha\beta}^2 - 0.043\sigma_\beta^2 - 0.096\sigma_\alpha^2 = 167.018,$$
$$\sigma_e^2 = 4.060.$$

Therefore, the desired ANOVA estimates of the variance components are given by

$$\hat{\sigma}^2_{e,\text{ANOV}} = 4.060,$$

and

$$\begin{bmatrix} \hat{\sigma}^2_{\alpha\beta,\text{ANOV}} \\ \hat{\sigma}^2_{\beta,\text{ANOV}} \\ \hat{\sigma}^2_{\alpha,\text{ANOV}} \end{bmatrix} = \begin{bmatrix} 4.449 & 0.130 & 15.638 \\ 3.951 & 11.305 & 0.192 \\ 3.634 & -0.043 & -0.096 \end{bmatrix}^{-1} \begin{bmatrix} 634.149 \\ 116.419 \\ 162.958 \end{bmatrix}$$
$$= \begin{bmatrix} 45.501 \\ -6.074 \\ 27.657 \end{bmatrix}.$$

To obtain variance component estimates based on fitting-constants-method estimators (13.4.10), (13.4.13), and (13.4.15), we calculated analysis of variance based on reductions in sums of squares due to fitting the submodels. The results are summarized in Tables 13.8, 13.9, and 13.10.

Now, the estimates in (13.4.10) based on Table 13.8 (worker adjusted for site) are

$$\hat{\sigma}^2_{e,\text{FTC1}} = 4.060,$$
$$\hat{\sigma}^2_{\alpha\beta,\text{FTC1}} = \frac{160.211 - 4.060}{3.652} = 42.758,$$
$$\hat{\sigma}^2_{\beta,\text{FTC2}} = \frac{134.093 - 4.060 - 3.915 \times 42.758}{11.218} = -3.331,$$

and

$$\hat{\sigma}^2_{\alpha,\text{FTC2}} = \frac{638.209 - 4.060 - 4.449 \times 42.758 - 0.130 \times (-3.331)}{15.638}$$
$$= 28.415.$$

Similarly, the estimates in (13.4.13) based on Table 13.9 (site adjusted for worker) are

$$\hat{\sigma}^2_{e,\text{FTC2}} = 4.060,$$
$$\hat{\sigma}^2_{\alpha\beta,\text{FTC2}} = \frac{160.211 - 4.060}{3.652} = 42.758,$$
$$\hat{\sigma}^2_{\alpha,\text{FTC1}} = \frac{658.630 - 4.060 - 4.395 \times 42.758}{15.351} = 30.399,$$

and

$$\hat{\sigma}^2_{\beta,\text{FTC1}} = \frac{120.479 - 3.951 \times 42.758 - 0.192 \times 30.399}{11.305} = -4.803.$$

TABLE 13.8 Analysis of variance for the efficiency-score data of Table 13.6 (worker adjusted for site).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Site	2	1, 276.41771	638.209	$\sigma_e^2 + 4.449\sigma_{\alpha\beta}^2 + 0.130\sigma_\beta^2 + 15.638\sigma_\alpha^2$
Worker	3	402.28005	134.093	$\sigma_e^2 + 3.915\sigma_{\alpha\beta}^2 + 11.218\sigma_\beta^2$
Interaction	6	961.26503	160.211	$\sigma_e^2 + 3.652\sigma_{\alpha\beta}^2$
Error	35	142.08700	4.060	σ_e^2
Total	46	2, 782.0497		

TABLE 13.9 Analysis of variance for the efficiency-score data of Table 13.6 (site adjusted for worker).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Worker	3	361.43723	120.479	$\sigma_e^2 + 3.951\sigma_{\alpha\beta}^2 + 11.305\sigma_\beta^2 + 0.192\sigma_\alpha^2$
Site	2	1, 317.26052	658.630	$\sigma_e^2 + 4.395\sigma_{\alpha\beta}^2 + 15.351\sigma_\alpha^2$
Interaction	6	961.26503	160.211	$\sigma_e^2 + 3.652\sigma_{\alpha\beta}^2$
Error	35	142.08700	4.060	σ_e^2
Total	46	2, 782.0497		

TABLE 13.10 Analysis of variance for the efficiency-score data of Table 13.6 (worker adjusted for site and site adjusted for worker).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Site	2	1, 317.26052	658.630	$\sigma_e^2 + 4.395\sigma_{\alpha\beta}^2 + 15.351\sigma_\alpha^2$
Worker	3	402.28005	134.093	$\sigma_e^2 + 3.915\sigma_{\alpha\beta}^2 + 11.218\sigma_\beta^2$
Interaction	6	961.26503	160.211	$\sigma_e^2 + 3.652\sigma_{\alpha\beta}^2$
Error	35	142.08700	4.060	σ_e^2
Total	46	2, 782.0497		

Finally, the estimates in (13.4.15) based on Table 13.10 (worker adjusted for site and site adjusted for worker) are

$$\hat{\sigma}^2_{e,\text{FTC3}} = 4.060,$$
$$\hat{\sigma}^2_{\alpha\beta,\text{FTC3}} = \frac{160.211 - 4.060}{3.652} = 42.758,$$
$$\hat{\sigma}^2_{\beta,\text{FTC3}} = \frac{134.093 - 4.060 - 3.915 \times 42.758}{11.218} = -3.331,$$

and

$$\hat{\sigma}^2_{\alpha,\text{FTC3}} = \frac{658.630 - 4.060 - 4.395 \times 42.758}{15.351} = 30.399.$$

The negative estimates for σ^2_β is probably an indication that the variance component may be zero.

For the analysis of means estimates in (13.4.18) and (13.4.21), we computed analysis of variance using unweighted and weighted sums of squares and the results are summarized in Tables 13.11 and 13.12.

Now, the estimates in (13.4.18) based on Table 13.11 (unweighted sums of squares) are

$$\hat{\sigma}^2_{e,\text{UME}} = 4.060,$$
$$\hat{\sigma}^2_{\alpha\beta,\text{UME}} = \frac{101.068 - 4.060}{2.802} = 34.621,$$
$$\hat{\sigma}^2_{\beta,\text{UME}} = \frac{132.415 - 101.068}{8.406} = 3.729,$$

and

$$\hat{\sigma}^2_{\alpha,\text{UME}} = \frac{460.423 - 101.068}{11.208} = 32.062.$$

Similarly, the estimates in (13.4.21) based on Table 13.12 (weighted sums of squares) are

$$\hat{\sigma}^2_{e,\text{WME}} = 4.060,$$
$$\hat{\sigma}^2_{\alpha\beta,\text{WME}} = \frac{160.211 - 4.060}{3.652} = 42.758,$$
$$\hat{\sigma}^2_{\beta,\text{WME}} = \frac{144.613 - 4.060 - 3.607 \times 42.758}{10.820} = -1.264,$$

and

$$\hat{\sigma}^2_{\alpha,\text{WME}} = \frac{401.050 - 4.060 - 2.871 \times 42.758}{11.484} = 23.879.$$

TABLE 13.11 Analysis of variance for the efficiency-score data of Table 13.6 (unweighted sums of squares).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Site	2	920.847	460.423	$\sigma_e^2 + 2.802\sigma_{\alpha\beta}^2 + 11.208\sigma_\alpha^2$
Worker	3	397.244	132.415	$\sigma_e^2 + 2.802\sigma_{\alpha\beta}^2 + 8.406\sigma_\beta^2$
Interaction	6	606.406	101.068	$\sigma_e^2 + 2.802\sigma_{\alpha\beta}^2$
Error	35	142.087	4.060	σ_e^2
Total	46			

TABLE 13.12 Analysis of variance for the efficiency-score data of Table 13.6 (weighted sums of squares).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Site	2	802.101	401.050	$\sigma_e^2 + 2.871\sigma_{\alpha\beta}^2 + 11.484\sigma_\alpha^2$
Worker	3	433.838	144.613	$\sigma_e^2 + 3.607\sigma_{\alpha\beta}^2 + 10.820\sigma_\beta^2$
Interaction	6	961.265	160.211	$\sigma_e^2 + 3.652\sigma_{\alpha\beta}^2$
Error	35	142.087	4.060	σ_e^2
Total	46			

We used SAS® VARCOMP, SPSS® VARCOMP, and BMDP®3V to estimate the variance components using the ML, REML, MINQUE(0), and MINQUE(1) procedures.[4] The desired estimates using these software packages are given in Table 13.13. Note that all three software packages produce nearly the same results except for some minor discrepancy in rounding decimal places.

13.5 VARIANCES OF ESTIMATORS

In this section, we present some results on sampling variances of estimators considered in the preceding section.

[4]The computations for ML and REML estimates were also carried out using SAS® PROC MIXED and some other programs to assess their relative accuracy and convergence rate. There did not seem to be any appreciable differences between the results from different software.

TABLE 13.13 ML, REML, MINQUE(0), and MINQUE(1) estimates of the variance components using SAS®, SPSS®, and BMDP® software.

Variance component	SAS®		
	ML	REML	MINQUE(0)
σ_e^2	4.056567	4.054976	5.282301
$\sigma_{\alpha\beta}^2$	37.752135	35.497787	45.460059
σ_β^2	0.650768	3.101930	−6.046132
σ_α^2	17.790144	31.984173	25.886682

Variance component	SPSS®		
	ML	REML	MINQUE(0)
σ_e^2	4.056561	4.054968	5.282301
$\sigma_{\alpha\beta}^2$	37.752303	35.497991	45.460059
σ_β^2	0.650774	3.101952	−6.046132
σ_α^2	17.790225	31.984354	25.886682

Variance component	BMDP®		
	MINQUE(1)	ML	REML
σ_e^2	3.794245	4.056561	4.054968
$\sigma_{\alpha\beta}^2$	39.824284	37.752303	35.497991
σ_β^2	0.587333	0.650774	3.101952
σ_α^2	31.550930	17.790225	31.984354

SAS® VARCOMP does not compute MINQUE(1). BMDP®3V does not compute MINQUE(0) and MINQUE(1).

13.5.1 VARIANCES OF ANALYSIS OF VARIANCE ESTIMATORS

In the analysis of variance given in Section 13.2, SS_E has σ_e^2 times a chi-square distribution with $N - s$ degrees of freedom and is distributed independently of SS_A, SS_B, and SS_{AB}. Hence, the variance of $\hat{\sigma}_e^2$ is given by

$$\text{Var}(\hat{\sigma}_{e,\text{ANOV}}^2) = 2\sigma_e^4/(N - s), \tag{13.5.1}$$

and the covariances of $\hat{\sigma}_{e,\text{ANOV}}^2$ with $\hat{\sigma}_{\alpha,\text{ANOV}}^2$, $\hat{\sigma}_{\beta,\text{ANOV}}^2$, and $\hat{\sigma}_{\alpha\beta,\text{ANOV}}^2$ are zero. To find the variances and covariances of $\hat{\sigma}_{\alpha,\text{ANOV}}^2$, $\hat{\sigma}_{\beta,\text{ANOV}}^2$, $\hat{\sigma}_{\alpha\beta,\text{ANOV}}^2$, we rewrite the equations in (13.4.3) as

$$\hat{\boldsymbol{\sigma}}_{\text{ANOV}}^2 = \boldsymbol{P}^{-1}[\boldsymbol{H}\boldsymbol{t} - \hat{\sigma}_{e,\text{ANOV}}^2 \boldsymbol{f}], \tag{13.5.2}$$

where

$$\begin{aligned}\hat{\boldsymbol{\sigma}}'^{2}_{\text{ANOV}} &= (\hat{\sigma}^2_{\alpha,\text{ANOV}}, \hat{\sigma}^2_{\beta,\text{ANOV}}, \hat{\sigma}^2_{\alpha\beta,\text{ANOV}}),\\ \boldsymbol{t}' &= (T_A, T_B, T_{AB}, T_\mu),\\ \boldsymbol{f}' &= (a-1, b-1, s-a-b+1),\end{aligned}$$

and

$$\boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix}.$$

From (13.5.2), the variance-covariance matrix of $\hat{\sigma}^2_{\alpha,\text{ANOV}}$, $\hat{\sigma}^2_{\beta,\text{ANOV}}$, and $\hat{\sigma}^2_{\alpha\beta,\text{ANOV}}$ is given by (Searle, 1958).

$$\text{Var}(\hat{\boldsymbol{\sigma}}^2_{\text{ANOV}}) = \boldsymbol{P}^{-1}[\boldsymbol{H}\,\text{Var}(\boldsymbol{t})\boldsymbol{H}' + \text{Var}(\hat{\sigma}^2_{e,\text{ANOV}})\boldsymbol{f}\boldsymbol{f}']\boldsymbol{P}^{-1'}. \qquad (13.5.3)$$

Thus, to evaluate $\text{Var}(\hat{\boldsymbol{\sigma}}^2_{\text{ANOV}})$ given by (13.5.3), we only need to find the variance-covariance matrix, $\text{Var}(\boldsymbol{t})$, whose elements are variances and covariances of the uncorrected sums of squares T_A, T_B, T_{AB}, and T_μ. The expressions for variances and covariances of T_A, T_B, T_{AB}, and T_μ have been obtained by Searle (1958) and involve some extensive and tedious algebra. The results are given as follows (see also Searle, 1971, pp. 481–483; Searle et al., 1992, pp. 434–437):

$$\begin{aligned}
\text{Var}(T_A) &= 2[k_1\sigma^4_\alpha + (k_{21}+k_9)\sigma^4_\beta + k_9\sigma^2_{\alpha\beta} + a\sigma^4_e\\
&\quad + 2(k_{23}\sigma^2_\alpha\sigma^2_\beta + k_{23}\sigma^2_\alpha\sigma^2_{\alpha\beta} + N\sigma^2_\alpha\sigma^2_e + k_9\sigma^2_\beta\sigma^2_{\alpha\beta}\\
&\quad + k_3\sigma^2_\beta\sigma^2_e + k_3\sigma^2_{\alpha\beta}\sigma^2_e)],\\
\text{Var}(T_B) &= 2[(k_{22}+k_{10})\sigma^4_\alpha + k_2\sigma^4_\beta + k_{10}\sigma^4_{\alpha\beta} + b\sigma^4_e\\
&\quad + 2(k_{23}\sigma^2_\alpha\sigma^2_\beta + k_{10}\sigma^2_\alpha\sigma^2_{\alpha\beta} + k_4\sigma^2_\alpha\sigma^2_e + k_{23}\sigma^2_\beta\sigma^2_{\alpha\beta}\\
&\quad + N\sigma^2_\beta\sigma^2_e + k_4\sigma^2_{\alpha\beta}\sigma^2_e)],\\
\text{Var}(T_{AB}) &= 2[k_1\sigma^4_\alpha + k_2\sigma^4_\beta + k_{23}\sigma^4_{\alpha\beta} + s\sigma^4_e\\
&\quad + 2(k_{23}\sigma^2_\alpha\sigma^2_\beta + k_{23}\sigma^2_\alpha\sigma^2_{\alpha\beta} + N\sigma^2_\alpha\sigma^2_e + k_{23}\sigma^2_\beta\sigma^2_{\alpha\beta}\\
&\quad + N\sigma^2_\beta\sigma^2_e + N\sigma^2_{\alpha\beta}\sigma^2_e)],\\
\text{Var}(T_\mu) &= \frac{2}{N^2}[k_1^2\sigma^4_\alpha + k_2^2\sigma^4_\beta + k_{23}^2\sigma^4_{\alpha\beta} + N^2\sigma^4_e\\
&\quad + 2(k_1k_2\sigma^2_\alpha\sigma^2_\beta + k_1k_{23}\sigma^2_\alpha\sigma^2_{\alpha\beta} + Nk_1\sigma^2_\alpha\sigma^2_e + k_2k_{23}\sigma^2_\beta\sigma^2_{\alpha\beta}\\
&\quad + Nk_2\sigma^2_\beta\sigma^2_e + Nk_{23}\sigma^2_{\alpha\beta}\sigma^2_e)],\\
\text{Cov}(T_A, T_B) &= 2[k_{18}\sigma^4_\alpha + k_{17}\sigma^4_\beta + k_{28}\sigma^4_{\alpha\beta} + k_{26}\sigma^4_e
\end{aligned}$$

$$
\begin{aligned}
&+ 2(k_{23}\sigma_\alpha^2\sigma_\beta^2 + k_{12}\sigma_\alpha^2\sigma_{\alpha\beta}^2 + k_4\sigma_\alpha^2\sigma_e^2 + k_{11}\sigma_\beta^2\sigma_{\alpha\beta}^2 \\
&+ k_3\sigma_\beta^2\sigma_e^2 + k_{27}\sigma_{\alpha\beta}^2\sigma_e^2)],
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Cov}(T_A, T_{AB}) = {} & 2[k_1\sigma_\alpha^4 + k_{17}\sigma_\beta^4 + k_{11}\sigma_{\alpha\beta}^4 + a\sigma_e^4 \\
&+ 2(k_{23}\sigma_\alpha^2\sigma_\beta^2 + k_{23}\sigma_\alpha^2\sigma_{\alpha\beta}^2 + N\sigma_\alpha^2\sigma_e^2 + k_{11}\sigma_\beta^2\sigma_{\alpha\beta}^2 \\
&+ k_3\sigma_\beta^2\sigma_e^2 + k_3\sigma_{\alpha\beta}^2\sigma_e^2)],
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Cov}(T_A, T_\mu) = {} & \frac{2}{N}[k_5\sigma_\alpha^4 + k_{15}\sigma_\beta^4 + k_7\sigma_{\alpha\beta}^4 + N\sigma_e^4 \\
&+ 2(k_{25}\sigma_\alpha^2\sigma_\beta^2 + k_{19}\sigma_\alpha^2\sigma_{\alpha\beta}^2 + k_1\sigma_\alpha^2\sigma_e^2 + k_{13}\sigma_\beta^2\sigma_{\alpha\beta}^2 \\
&+ k_2\sigma_\beta^2\sigma_e^2 + k_{23}\sigma_{\alpha\beta}^2\sigma_e^2)],
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Cov}(T_B, T_{AB}) = {} & 2[k_{18}\sigma_\alpha^4 + k_2\sigma_\beta^4 + k_{12}\sigma_{\alpha\beta}^4 + b\sigma_e^4 \\
&+ 2(k_{23}\sigma_\alpha^2\sigma_\beta^2 + k_{12}\sigma_\alpha^2\sigma_{\alpha\beta}^2 + k_4\sigma_\alpha^2\sigma_e^2 + k_{23}\sigma_\beta^2\sigma_{\alpha\beta}^2 \\
&+ N\sigma_\beta^2\sigma_e^2 + k_4\sigma_{\alpha\beta}^2\sigma_e^2)],
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Cov}(T_B, T_\mu) = {} & \frac{2}{N}[k_{16}\sigma_\alpha^4 + k_6\sigma_\beta^4 + k_8\sigma_{\alpha\beta}^4 + N\sigma_e^4 \\
&+ 2(k_{25}\sigma_\alpha^2\sigma_\beta^2 + k_{14}\sigma_\alpha^2\sigma_{\alpha\beta}^2 + k_1\sigma_\alpha^2\sigma_e^2 + k_{20}\sigma_\beta^2\sigma_{\alpha\beta}^2 \\
&+ k_2\sigma_\beta^2\sigma_e^2 + k_{23}\sigma_{\alpha\beta}^2\sigma_e^2)],
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{Cov}(T_{AB}, T_\mu) = {} & \frac{2}{N}[k_5\sigma_\alpha^4 + k_6\sigma_\beta^4 + k_{24}\sigma_{\alpha\beta}^4 + N\sigma_e^4 \\
&+ 2(k_{25}\sigma_\alpha^2\sigma_\beta^2 + k_{19}\sigma_\alpha^2\sigma_{\alpha\beta}^2 + k_1\sigma_\alpha^2\sigma_e^2 + k_{20}\sigma_\beta^2\sigma_{\alpha\beta}^2 \\
&+ k_2\sigma_\beta^2\sigma_e^2 + k_{23}\sigma_{\alpha\beta}^2\sigma_e^2)],
\end{aligned}
$$

where k_is and k_{ij}s are defined in Section 12.5.1. It should be mentioned that Crump (1947) seems to have been the first to derive the sampling variances of this class of estimators for the two-way crossed classification random model.

13.5.2 VARIANCES OF FITTING-CONSTANTS-METHOD ESTIMATORS

Rhode and Tallis (1969) give formulas for expectations and covariances of sums of squares and products in a two-way crossed analysis of covariance model in a general computable form using matrix notations. The results can be simplified to yield variances and covariances of fitting-constants-method estimators. However, explicit algebraic evaluation of these expressions seems to be too involved. The interested reader is referred to the original paper for any further details.

13.5.3 VARIANCES OF ANALYSIS OF MEANS ESTIMATORS

For the unweighted analysis of means estimators (13.4.18), the expressions for variances have been developed by Hirotsu (1966). The results are given as follows:

$$\operatorname{Var}(\hat{\sigma}^2_{e,\text{UME}}) = \frac{2\sigma_e^4}{N-ab},$$

$$\begin{aligned}\operatorname{Var}(\hat{\sigma}^2_{\alpha\beta,\text{UME}}) &= \frac{2n_h^2\sigma_e^4}{(N-ab)} + \frac{2(\sigma^4_{\alpha\beta} + 2n_h\sigma^2_{\alpha\beta}\sigma^2_e)}{(a-1)(b-1)} \\ &\quad + \frac{2[(a-2)(b-2)N_2 + (a-2)N_3 + (b-2)N_4 + N_5]\sigma_e^4}{[ab(a-1)(b-1)]^2},\end{aligned}$$

$$\begin{aligned}\operatorname{Var}(\hat{\sigma}^2_{\beta,\text{UME}}) &= \frac{2[(1-a^{-1})\sigma^4_\beta + \{\sigma^2_\beta + a^{-1}(\sigma^2_{\alpha\beta} - \sigma^2_\beta)\}^2]}{(a-1)(b-1)} \\ &\quad + \frac{4n_h[\sigma^2_\beta + a^{-1}(\sigma^2_{\alpha\beta} - \sigma^2_\beta)]\sigma^2_e}{(a-1)(b-1)} \\ &\quad + \frac{2[N_5 - N_3a^{-1} + (b-2)(N_4 - N_2a^{-1})]\sigma_e^4}{[ab(a-1)(b-1)]^2},\end{aligned}$$

and

$$\begin{aligned}\operatorname{Var}(\hat{\sigma}^2_{\alpha,\text{UME}}) &= \frac{2[(1-b^{-1})\sigma^4_\alpha + \{\sigma^2_\alpha + b^{-1}(\sigma^2_{\alpha\beta} - \sigma^2_\alpha)\}^2]}{(a-1)(b-1)} \\ &\quad + \frac{4n_h[\sigma^2_\alpha + b^{-1}(\sigma^2_{\alpha\beta} - \sigma^2_\alpha)]\sigma^2_e}{(a-1)(b-1)} \\ &\quad + \frac{2[N_5 - N_4b^{-1} + (a-2)(N_3 - N_2b^{-1})]\sigma_e^4}{[ab(a-1)(b-1)]^2},\end{aligned}$$

where

$$N_1 = ab\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}^{-1}, \qquad N_2 = ab\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}^{-2},$$

$$N_3 = a\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{j'=1}^{b} n_{ij}^{-1}n_{ij'}^{-1}, \qquad N_4 = b\sum_{i=1}^{a}\sum_{i'=1}^{a}\sum_{j=1}^{b} n_{ij}^{-1}n_{i'j}^{-1},$$

$$N_5 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{i'=1}^{a}\sum_{j'=1}^{b} n_{ij}^{-1}n_{i'j'}^{-1}, \quad \text{and} \quad n_h = \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}^{-1}/ab.$$

13.6 COMPARISONS OF DESIGNS AND ESTIMATORS

The problem of constructing a two-way crossed unbalanced design in order to estimate variance components with greater precision seems to have been

TABLE 13.14 Efficiencies (E) of some two-way designs for estimating σ_α^2 and $\rho_\alpha (N = 30)$.

ρ_α	a	b	s	$E(\hat{\sigma}_\alpha^2)$	$E(\hat{\rho}_\alpha)$	ρ_α	a	b	s	$E(\hat{\sigma}_\alpha^2)$	$E(\hat{\rho}_\alpha)$
0.25	3	10	0	0.69	0.74	2.0	10	3	0	0.74	1.00
	5	6	0	0.95	0.98		12	3	6	0.82	0.97
	6	5	0	1.00	1.00		15	2	0	0.97	0.93
	7	5	2	0.98	0.95		19	2	11	1.00	0.63
	8	4	6	1.00	0.94						
1.0	10	3	0	0.88	1.00	4.0	10	3	0	0.59	1.00
	11	3	8	0.90	0.97		15	2	0	0.84	0.99
	14	3	2	0.96	0.83		20	2	10	0.94	0.53
	15	2	0	1.00	0.84		24	2	6	1.00	—

a = number of rows, b = number of columns; $N = ab$ if $s = 0$; $N = a(b-1) + s$ if $s > 0$.

Source: Anderson (1975); used with permission.

considered first by Gaylor (1960). Gaylor considered methods of sampling to minimize the variance of certain estimators of variance components. He also investigated optimal designs for estimating certain specified functions of the variance components. The two-way design with equal numbers of observations could produce very inefficient estimates of variance components corresponding to the main effects and may sometimes be considered extravagant for this purpose. Gaylor showed that if the design were restricted to a class of designs in which $n_{ij} = 0$ or n (an integer), then for an optimal estimate of σ_α^2 the value of n should be equal to one. Hence, each cell would either be empty or contain only one observation. In this case, only σ_α^2, σ_β^2, and $\sigma_e^2 + \sigma_{\alpha\beta}^2$ could be estimated. Based on the fitting-constants method, Gaylor (1960) recommended the following design: (i) If $\sigma_\alpha^2/\sigma_e^2 > \sqrt{2}$, one would use one column with $a = N - a'$ rows and a second column with a' of these rows where a' is the integer (≥ 2) which is closest to $1 + (N-2)/(\sqrt{2}\sigma_\alpha^2/\sigma_e^2)$; (ii) when $\sigma_\alpha^2/\sigma_e^2 \leq \sqrt{2}$, one would use a balanced design with number of columns b as the integer closest to $[(N-1/2)(\sigma_\alpha^2/\sigma_e^2) + N + 1/2]/[(N-1/2)(\sigma_\alpha^2/\sigma_e^2) + 2]$. In general, N/b will not be an integer, hence it would be advisable to use a few more or less observations to obtain a balanced plan.

Efficiency factors of various designs considered by Gaylor are shown in Table 13.14 These show that if $\sigma_\alpha^2/\sigma_e^2 > 1$ the design should be unbalanced with three columns to estimate $\rho_\alpha = \sigma_\alpha^2/\sigma_e^2$ and two columns to estimate σ_α^2. For example, if $\rho_\alpha = 4.0$ and $N = 30$, the optimal design to estimate σ_α^2 would have one column with 24 rows and a second column with only six of these rows. From the foregoing results it is evident that in order to obtain "good" estimates of both σ_α^2 and σ_β^2, one must modify the design since an

	C_1	C_2	C_3	C_4
R_1	1	1		
R_2	1	1		
R_3			1	1
R_4			1	1

Source: Gaylor (1960); used with permission.

FIGURE 13.1 An unconnected BD2 design.

optimal plan for σ_α^2 would rather be inefficient for σ_β^2. For this purpose, Gaylor proposed an L design that consists of an optimal design for σ_α^2 superimposed on an optimal design for σ_β^2. He also considered a series of unconnected designs, called balanced disjoint (BD) designs, such as the one shown in Figure 13.1. This design is known as BD2 design. If each block has r rows and r columns, then the design would be called BDr design. It is also possible to use a design having a series of rectangles with r rows and c columns known as BD $(r \times c)$ design. These designs do not provide separate estimates of $\sigma_{\alpha\beta}^2$ and σ_e^2 unless two observations are taken from some of the cells. In addition, estimation procedures such as the iterated least squares or ML must be used because there are more mean squares in the ANOVA table than parameters to be estimated and pooling is not possible for a disconnected design.

Bush (1962) and Bush and Anderson (1963) compared the variances of the estimators obtained from the analysis of variance method, the fitting-constants method, and the weighted means method. Comparisons were made between the estimation procedures and between the designs themselves, using a variety of values of the true components and for several sets of n_{ij}-values for a number of unbalanced designs with three and six rows and columns, representing what might be termed not wholly unbalanced but designed unbalancedness. In particular, they considered Gaylor L designs and modified BD designs, called S and C designs.Some examples of Bush–Anderson designs are shown in Figure 13.2. Eighteen sets of parameter values were used: σ_α^2 ranged from $1/2$ to 16, σ_β^2 ranged from 0 to 16, $\sigma_{\alpha\beta}^2$ from 0 to 16 and $\sigma_e^2 = 1$. It was found that when σ_e^2 was larger than σ_α^2 and σ_β^2, a balanced design was preferable; otherwise a nonbalanced S or C design was preferred to estimate σ_α^2 and σ_β^2. If the experimenter does not have any prior information concerning the values of the variance components, then the use of an S design is probably the best first choice. The results further indicate that at least for the designs included in the study, the ANOVA method yields estimates with the smallest variances only when $\sigma_{\alpha\beta}^2$ was larger than σ_α^2 and σ_β^2; however, in this situation, Bush and Anderson (1963) recomended the use of a balanced design. The method of fitting constants was found to have a slightly better performance when $\sigma_{\alpha\beta}^2$ was smaller than σ_α^2 and σ_β^2. The authors also provided a generalization of their results to higher-order classification models.

D_1			D_2			D_3			D_4			D_5		
2	2	2	1	1	1	1	2	3	1	1	4	1	1	1
2	2	2	2	2	2	2	3	1	1	4	1	1	2	3
2	2	2	3	3	3	3	1	2	4	1	1	1	3	5

equal 6 × 6						$S16$						$S22$						$C18$					
1	1	1	1	1	1	1	1	0	0	0	0	2	1	0	0	0	0	1	1	1	0	0	0
1	1	1	1	1	1	1	1	1	0	0	0	1	2	1	0	0	0	1	1	1	0	0	0
1	1	1	1	1	1	0	1	1	1	0	0	0	1	2	1	0	0	0	1	1	1	0	0
1	1	1	1	1	1	0	0	1	1	1	0	0	0	1	2	1	0	0	0	1	1	1	0
1	1	1	1	1	1	0	0	0	1	1	1	0	0	0	1	2	1	0	0	0	1	1	1
1	1	1	1	1	1	0	0	0	0	1	1	0	0	0	0	1	2	0	0	0	1	1	1

$C24$						$L20$						$L24$					
2	1	1	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0
2	1	1	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0
0	2	1	1	0	0	1	1	0	0	0	0	2	1	0	0	0	0
0	0	1	1	2	0	1	1	0	0	0	0	1	2	0	0	0	0
0	0	0	1	1	2	1	1	1	1	1	1	1	1	2	1	1	1
0	0	0	1	1	2	1	1	1	1	1	1	1	1	1	2	1	1

Source: Bush and Anderson (1963).

FIGURE 13.2 Values of n_{ij} for some examples of Bush–Anderson designs.

Hirotsu (1966) found that for the five designs having three levels in each classification and with no empty cells, many of the unweighted means estimators have still smaller variance. Mostafa (1967) also compared certain unbalanced square designs that have all the n_{ij}s equal to 1 or 2 with the corresponding balanced plans having the same number of observations. In particular, Mostafa considered the designs M_1 and M_2 described as follows. The M_1 design contains r_1 rows and r_1 columns where a single observation is taken from each of $r_1(r_1 - 1)$ cells and two observations are taken from each diagonal cell. The M_2 design contains r_2 rows and r_2 columns where two observations are taken from each of the two cells in each row and column and one observation is taken from each of the other cells. As noted by the author, an interesting feature of these designs is that the sums of squares for rows and columns are each distributed as multiple of a chi-square variate with respective degrees of freedom. Mostafa investigated the problem of joint estimation of the parameters $\sigma_\alpha^2, \sigma_\beta^2, \sigma_{\alpha\beta}^2, \sigma_e^2$, $\sigma_\alpha^2/\sigma_e^2, \sigma_\beta^2/\sigma_e^2$, and $\sigma_{\alpha\beta}^2/\sigma_e^2$. The variances of the estimators are compared with those based on a balanced design with the same number of observations for certain selected values of the ratios of the variance components. The estimation procedure employed was the unweighted means method. The results show that the estimates of $\sigma_\alpha^2, \sigma_\beta^2$, and $\sigma_{\alpha\beta}^2$ obtained from the unbalanced designs are much more efficient than those obtained from a balanced plan with the same number of observations, particularly in situations where the ratios $\sigma_\alpha^2/\sigma_e^2$, $\sigma_\beta^2/\sigma_e^2$, and $\sigma_{\alpha\beta}^2/\sigma_e^2$ are much greater than unity.

For the model in (13.1.1) with $\sigma_{\alpha\beta}^2 = 0$, Haile and Webster (1975) compared four designs for estimating the variance components $\sigma_e^2, \sigma_\beta^2$, and σ_α^2. The designs being compared are the disjoint rectangle, the generalized L-shaped, the generalized staggered, and the balanced incomplete. It is found that the optimum selection of type of design depends upon the ratio of the main effect variance to the error variance. Furthermore, in estimating σ_β^2 and σ_α^2, the choice between the disjoint rectangle, the backed-up staggered, or the balanced incomplete block is of minor importance as compared to the choice of the number of levels of the random effects, which is a function of unknown variance ratios $\sigma_\alpha^2/\sigma_e^2$ and $\sigma_\beta^2/\sigma_e^2$. For the same model, Muse (1974) and Muse and Anderson (1978) compared various designs with 0 or 1 observation per cell using mean squared error criterion and the method of maximum likelihood for the estimation of variance components. The mean squared errors were determined for each variance component and the sum of the mean squares for all components (the trace of the matrix of mean square errors).

They investigated both large and small sample properties for various connected and disconnected designs. The large sample results are based on asymptotic variances of the ML estimators and small sample results were obtained by 5000 simulated runs for each parameter set. The designs considered are 2×2 BD2 (nine squares each 2×2), 3×3BD3 (four squares each 3×3), BD2 $\times$ 3 (six rectangles each 2×3); a new design 3×3 OD3 (six squares each 3×3 with empty diagonals), 6×6 balanced, 10×10 $L36$, 13×13 S37 and 12×12 modified S36 (obtained by adding one observation to the upper right and lower left corners, yielding three observations in each row and column). The incidence matrices for these designs with 36 or 37 observations are shown in Table 13.15. Muse and Anderson (1978) compared the asymptotic variances for the designs in Table 13.15 for σ_α^2 and σ_β^2 values ranging from 0 to 8. The criteria used in the comparison were trace asymptotic variance (trace($\boldsymbol{AV}$)) of the vector of ML estimates, Var($\hat{\sigma}_t^2$), Var($\hat{\sigma}_\alpha^2$), and Var($\hat{\sigma}_\beta^2$), where $\sigma_t^2 = \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_e^2$. For Var($\hat{\sigma}_t^2$), it was found that OD3 design is generally superior. The results on trace($\boldsymbol{AV}$) are summarized in Table 13.16. Thus it is seen that the balanced design is best for σ_α^2 and $\sigma_\beta^2 < \sigma_e^2$; when $\sigma_\alpha^2 = \sigma_\beta^2 \geq \sigma_e^2$, the S37, MS36, and OD3 had superior performance. When $\sigma_\alpha^2 = \sigma_\beta^2 \gg \sigma_e^2$, the OD3 design becomes definitely superior. Further, the L design should not be used. When σ_e^2 is not the dominant variance component, all the five nonbalanced designs have similar performance, except that OD3 is superior when both σ_α^2 and σ_β^2 are quite large. When $\sigma_\alpha^2 < \sigma_e^2 \ll \sigma_\beta^2$, there is a slight advantage in using a design such as BD2 $\times$ 3.

A comparison of large and small sample results for BD2-36 and OD3-36 relative to the B36 design, using $\sigma_\alpha^2 = \sigma_\beta^2 = 0.5$, $\sigma_e^2 = 1$; $\sigma_\alpha^2 = \sigma_\beta^2 = 8.0$, $\sigma_e^2 = 1$; and $\sigma_\alpha^2 = 0.5$, $\sigma_\beta^2 = 8.0$, $\sigma_e^2 = 1$ is given in Table 13.17. It is seen that although the small sample and asymptotic comparisons of BD2 and OD3 designs with the balanced design do not agree as closely as desired, they point toward the same design preferences. Furthermore, these results support the viewpoint that the large sample results provide a reasonable indication of design preference provided the asymptotic ratio of interest for the two designs

TABLE 13.15 Incidence matrices for the Muse designs.

Design	Incidence matrix
S37	$\begin{bmatrix}\text{Let } i, j = 1, 2, \ldots, 13, \quad \text{where} \\ n_{11} = n_{12} = n_{13,12} = n_{13,13} = 1; \\ n_{ij} = 1 \quad \text{if } i-1 \le j \le i+1 \quad \text{for} \\ i = 2, \ldots, 12; \quad n_{ij} = 0 \quad \text{otherwise.}\end{bmatrix}$
MS36	$\begin{bmatrix}\text{Let } i, j = 1, 2, \ldots, 12, \quad \text{where} \\ n_{11} = n_{12} = n_{1,12} = 1; \\ n_{12,1} = n_{12,11} = n_{12,12} = 1, \\ n_{ij} = 1 \quad \text{if } i-1 \le j \le i+1, \\ \text{for } i = 2, \ldots, 11; \quad n_{ij} = 0 \quad \text{otherwise.}\end{bmatrix}$
OD3-36	$\left[\underset{6\times 6}{\boldsymbol{I}} \otimes \begin{bmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0\end{bmatrix}\right]$
L36	$\begin{bmatrix}\text{Let } i, j = 1, 2, \ldots, 10, \\ n_{ij} = 0 \quad \text{if } i \ge 3 \text{ and } j \ge 3; \\ n_{ij} = 1 \quad \text{otherwise.}\end{bmatrix}$
BD3-36	$\left[\underset{4\times 4}{\boldsymbol{I}} \otimes [\boldsymbol{J}_3 \boldsymbol{J}_3']\right]$
BD2-36	$\left[\underset{9\times 9}{\boldsymbol{I}} \otimes (\boldsymbol{J}_2 \boldsymbol{J}_2')\right]$
BD2 × 3-36	$\left[\underset{6\times 6}{\boldsymbol{I}} \otimes (\boldsymbol{J}_2 \boldsymbol{J}_3')\right]$
BD3 × 2-36	$\left[\underset{6\times 6}{\boldsymbol{I}} \otimes (\boldsymbol{J}_3 \boldsymbol{J}_2')\right]$
B36	$[(\boldsymbol{J}_6 \boldsymbol{J}_6')]$

$\boldsymbol{J}_n$ denotes an n-vector of 1s.

Source: Muse and Anderson (1978); used with permission.

TABLE 13.16 Trace asymptotic variance results of Muse designs.

Condition	Preferred design
σ_e^2 is dominant	Balanced
$\max(\sigma_\alpha^2/\sigma_e^2, \sigma_\beta^2/\sigma_e^2) \doteq 1$	BD3
$1 \le \sigma_\alpha^2/\sigma_e^2 \doteq \sigma_\beta^2/\sigma_e^2 < 2$	MS
$2 \le \sigma_\alpha^2/\sigma_e^2 \doteq \sigma_\beta^2/\sigma_e^2$	S or OD3
$\sigma_\alpha^2 \ne \sigma_\beta^2$ and one larger than σ_e^2	BD2
$\sigma_\alpha^2 \ne \sigma_\beta^2$ and both larger than σ_e^2	OD3

Source: Muse and Anderson (1978); used with permission.

is not too close to 1. It should be mentioned that the authors obtained closed form analytic solutions for ML equations for the B36 and BD2-36 designs. This work was further extended by Thitakamol (1977) and Muse et al. (1982) who compared designs with 0, 1, or 2 observations in order to estimate both σ_e^2 and $\sigma_{\alpha\beta}^2$. As before, the comparisons were based on trace asymptotic variance of the ML estimates.

A description of these designs is given in Table 13.18 and the trace asymptotic variance results are summarized in Table 13.19. Thus, as before, OD is preferred when both σ_α^2 and σ_β^2 are large; BDI is most desirable when either σ_α^2 or σ_β^2 is large; and a balanced design is the best when both σ_α^2 and σ_β^2 are small. For some similar results in the case of completely random balanced incom-

TABLE 13.17 Ratios of small sample MSE estimates (SS) and asymptotic variance (LS) for the BD2 and OD3 designs relative to the B design ($\sigma_e^2 = 1$).

σ_α^2	σ_β^2	Ratio	Type	R($\hat{\sigma}_\alpha^2$)	R($\hat{\sigma}_\beta^2$)	R($\hat{\sigma}_e^2$)	R(trace)	R($\hat{\sigma}_t^2$)
0.5	0.5	BD2:B	SS	1.21	1.21	2.17	1.42	0.706
			LS	1.18	1.18	2.62	1.46	0.701
		OD3:B	SS	1.25	1.25	2.58	1.55	1.13
			LS	1.21	1.21	3.04	1.56	0.703
8.0	8.0	BD2:B	SS	0.653	0.653	2.81	0.655	0.436
			LS	0.523	0.523	2.78	0.527	0.433
		OD3:B	SS	0.510	0.510	5.56	0.520	0.446
			LS	0.442	0.442	4.14	0.448	0.404
0.5	8.0	BD2:B	SS	1.68	0.385	2.41	0.403	0.378
			LS	1.53	0.377	2.77	0.395	0.375
		OD3:B	SS	1.62	0.395	3.02	0.414	0.402
			LS	1.95	0.374	3.98	0.399	0.373

Source: Muse and Anderson (1978); used with permission.

plete block designs, see Stroup et al. (1980). For a clear and concise review of some of these designs, see Anderson (1975, 1981). More recently, Shen et al. (1996) have compared a number of balanced and unbalanced two-way designs for estimation of genetic parameters using simulated and asymptotic variances of the ML estimates computed via an iterative least squares method. The results indicate that except when the error variance is quite large, certain unbalanced designs can yield more efficient estimates of the additive genetic variance, heritability and predicted gain for selection, but not for dominance variance (σ_d^2) or degree of dominance (d). Balanced designs are preferred for σ_d^2 and d. For some other results on estimation of heritability for unbalanced data, see Pederson (1972) and Thompson (1976, 1977).

Schaeffer (1973) compared numerically the sampling variances of estimators obtained from the ANOVA method, fitting-constants-method, Koch's symmetric sums method and the MINQUE procedure for both random and mixed model cases. For the random model he found that MINQUEs had the smallest variances for all components in a majority of combinations of n_{ij} patterns and parameter sets $\{\sigma_\alpha^2, \sigma_\beta^2, \sigma_{\alpha\beta}^2, \sigma_e^2\}$ considered in the study. For the mixed model case, MINQUE$'$s were best when the variance components including σ_e^2 were approximately of equal size, but not otherwise. This last result is not surprising since MINQUE with all σ_i^2s equal is MIVQUE under normality. For the model in (13.1.1) with $\sigma_{\alpha\beta}^2 = 0$, Low (1976) investigated some small sample properties of the ANOVA and fitting-constants-method estimators and noted that each of them yields estimates with smaller variance in respective subspaces of the parameter space.

Bremer (1989) made an extensive numerical comparison of small sam-

TABLE 13.18 Description of Thitakamol designs.

Design	Total number of observations	Description
BI	50	B25 with 2 observations per cell
BII	72	B36 with 2 observations per cell
BDI	60	BD2-12 with 2 observations per cell BD2-36 with 1 observation per cell
BDII	60	BD2-24 with 2 observations in the upper-left and lower-right cells and 1 observation in each of the remaining cells; BD2-24 with 1 observation per cell
BDIII	60	BD2-48 with 2 observations in the upper-left cell and 1 observation in each of the remaining cells
BDIV	64	BD4-32 with 2 observations per cell
OD	60	OD3-12 with 2 observations per off-diagonal cell; OD3-36 with 1 observation per off-diagonal cell

Source: Muse et al. (1982); used with permission.

TABLE 13.19 Trace asymptotic variance results of Thitakamol.

Condition			Preferred design		
σ_e^2	σ_α^2	σ_β^2	$\sigma_{\alpha\beta}^2$		
			0.5	**1.0**	**2.0**
1	1/2	1/2	BI	BII	BII
1	≤ 1	1	BDIV	BDIV	BI or BII
1	≤ 1	2	BDI	BDI	BDIV
1	≤ 1	8	BDI	BDI	BDI
1	2	2	BDI	BDI	BDI
1	2	8	OD	OD	BDI
1	8	8	OD	OD	OD

Source: Muse et al. (1982); used with permission.

ple variances of eight variance component estimators, which included several ANOVA- and MINQUE-type estimators, using Bhattacharya's lower bound. He reported that the only estimators that performed with relative uniform efficiency were the ANOVA (Henderson's Method I) and the MINQUE(1) estimators and recommended the use of ANOVA estimates for most situations. In contrast to asymptotic results, Bremer (1990) also investigated the small sample variance efficiency of different MINQUE-type estimators for varying number of levels and the n_{ij} patterns. He found that increasing the number of levels of a factor increases the efficiency of the estimators. Similarly, taking larger sample sizes in most of the cells resulted in greater efficiency. However, having too many cells with very few observations had adverse effect on efficiency.

Lin and McAllister (1984) made a simulation study to compare the ML and REML estimates of the variance components from the two-way mixed model using the MSE criterion. Factor A consisted of 480 herds with fixed effects and factor B consisted of 120 sires with random effects, having 5 and 100 daughters per sire. For each simulation run, typical parameters were chosen for sire variance and heritability. They found that MSEs of the ML and REML estimators of sire variance were quite similar (10.999 and 10.600, respectively); however, for the error variance, the REML had an MSE of 1.0 while that of the ML was 316.7. Thus they recommended the use of the REML estimator if a large or moderately large number of degrees of freedom are required for the fixed effect. For a completely random model, the ML and REML estimators give comparable results; the ML estimates being biased downward and in general smaller than the REML estimates.

13.7 CONFIDENCE INTERVALS

An exact interval for σ_e^2 can be based on the error mean square in a weighted or unweighted analysis. An exact interval for $\sigma_{\alpha\beta}^2/\sigma_e^2$ can be obtained using Wald's procedure considered in Section 11.8. The application of Wald's procedure for this model is shown by Spjøtvoll (1968) and Thomson (1975). The confidence interval, however, must be determined using an iterative procedure. For the balanced situation when $n_{ij} = n$, the Wald interval reduces to the interval given in (4.7.13). For a discussion and a numerical example of Wald's procedure using SAS® codes, see Burdick and Graybill (1992, pp. 140–141). However, there do not exist exact intervals for other functions of variance components. For a design with no missing cells, Srinivasan (1986), Burdick and Graybill (1992, pp. 137–139), and Hernández and Burdick (1993) recommended using intervals for the corresponding balanced case discussed in Section 4.7, where the usual mean squares are replaced by the mean squares obtained in the unweighted means analysis and $\bar{n}_h$ is substituted for n. On the basis of some simulation work by Hernández (1991) and Hernández and Burdick (1993), the authors report that these intervals maintain their coverage at the stated confidence level. Although this approach violates the assumptions of the chi-squaredness and independence of mean squares, they seem to have cancellation effects on the confidence coefficient. For designs with some empty cells, Kazempour and Graybill (1991) have considered using the intervals (4.7.14) and (4.7.15) for $\rho_\alpha = \sigma_\alpha^2/(\sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2 + \sigma_\alpha^2)$ and $\rho_\beta = \sigma_\beta^2/(\sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2 + \sigma_\alpha^2)$, for the corresponding balanced situation, by using alternate sums of squares which are equivalent to Type II sums of squares reported by the SAS®GLM.

13.7.1 A NUMERICAL EXAMPLE

In this section, we illustrate computations of confidence intervals on the variance components $\sigma_{\alpha\beta}^2$, σ_β^2, σ_α^2, and the total variance, $\sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2 + \sigma_\alpha^2$, using formulas (4.7.5) through (4.7.8) by replacing MS_A, MS_B, MS_{AB}, and n with MS_{Au}, MS_{Bu}, MS_{ABu}, and $\bar{n}_h$, respectively. Now, from the results of the analysis of variance given in Table 13.11, we have

$$\text{MS}_E = 4.060, \quad \text{MS}_{ABu} = 101.068, \quad \text{MS}_{Bu} = 132.415, \quad \text{MS}_{Au} = 460.423,$$
$$a = 3, \quad b = 4, \quad \bar{n}_h = 2.802, \quad \nu_e = 35, \quad \nu_{\alpha\beta} = 6, \quad \nu_\beta = 3, \quad \nu_\alpha = 2.$$

Further, for $\alpha = 0.05$, we obtain

$$\begin{aligned}
F[\nu_\alpha, \infty; \alpha/2] &= 0.025, & F[\nu_\alpha, \infty; 1-\alpha/2] &= 3.689,\\
F[\nu_\beta, \infty; \alpha/2] &= 0.072, & F[\nu_\beta, \infty; 1-\alpha/2] &= 3.116,\\
F[\nu_{\alpha\beta}, \infty; \alpha/2] &= 0.210, & F[\nu_{\alpha\beta}, \infty; 1-\alpha/2] &= 2.408,\\
F[\nu_e, \infty; \alpha/2] &= 0.587, & F[\nu_e, \infty; 1-\alpha/2] &= 1.520,\\
F[\nu_\alpha, \nu_{\alpha\beta}; \alpha/2] &= 0.025, & F[\nu_\alpha, \nu_{\alpha\beta}; 1-\alpha/2] &= 7.260,\\
F[\nu_\beta, \nu_{\alpha\beta}; \alpha/2] &= 0.068, & F[\nu_\beta, \nu_{\alpha\beta}; 1-\alpha/2] &= 6.599,\\
F[\nu_{\alpha\beta}, \nu_e; \alpha/2] &= 0.199, & F[\nu_{\alpha\beta}, \nu_e; 1-\alpha/2] &= 2.796.
\end{aligned}$$

In addition, to determine approximate confidence intervals for $\sigma_{\alpha\beta}^2$, σ_β^2, and σ_α^2, using formulas (4.7.5) through (4.7.7), we evaluate the following quantities:

$$\begin{aligned}
G_1 &= 0.72892383, & G_2 &= 0.67907574,\\
G_3 &= 0.58471761, & G_4 &= 0.34210526,\\
H_1 &= 39, & H_2 &= 12.88888889,\\
H_3 &= 3.76190476, & H_4 &= 0.70357751,
\end{aligned}$$

$$\begin{aligned}
G_{13} &= -0.40901566, & G_{23} &= -0.43710657, & G_{34} &= 0.02067001,\\
H_{13} &= -13.67578724, & H_{23} &= -3.55037567, & H_{34} &= -0.18022859,
\end{aligned}$$

$$\begin{aligned}
L_{\alpha\beta} &= 446.9386954, & U_{\alpha\beta} &= 18{,}403.06824,\\
L_\beta &= 2,077.446113, & U_\beta &= 40{,}598.75802,\\
L_\alpha &= 1,895.89874, & U_\alpha &= 2{,}561{,}731.992.
\end{aligned}$$

Substituting the appropriate quantities in (4.7.5) through (4.7.7), the desired 95% confidence intervals for $\sigma_{\alpha\beta}^2$, σ_β^2, and σ_α^2 are given by

$$P\{13.480 \le \sigma_{\alpha\beta}^2 \le 170.279\} \doteq 0.95,$$

$$P\{-41.850 \leq \sigma_\beta^2 \leq 205.220\} \doteq 0.95,$$

and

$$P\{-11.480 \leq \sigma_\alpha^2 \leq 1,632.603\} \doteq 0.95.$$

It is understood that the negative limits are defined to be zero.

To determine an approximate confidence interval for the total variance $\sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2 + \sigma_\alpha^2$ using formula (4.7.8), we obtain

$$\begin{aligned}\hat{\gamma} &= \frac{1}{3 \times 4 \times 2.802}[3 \times 460.423 + 4 \times 132.415 + 5 \times 101.068 + 35 \times 4.060] \\ &= 76.088.\end{aligned}$$

Substituting the appropriate quantities in (4.7.8), the desired 95% confidence interval for $\sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2 + \sigma_\alpha^2$ is given by

$$P\{43.067 \leq \sigma_e^2 + \sigma_{\alpha\beta}^2 + \sigma_\beta^2 + \sigma_\alpha^2 \leq 1,692.008\} \doteq 0.95.$$

Confidence intervals for other parametric functions of the variance components can similarly be computed.

13.8 TESTS OF HYPOTHESES

In this section, we consider briefly some tests of the hypotheses:

$$\begin{aligned} H_0^A &: \sigma_\alpha^2 = 0 \quad \text{vs.} \quad H_1^A : \sigma_\alpha^2 > 0, \\ H_0^B &: \sigma_\beta^2 = 0 \quad \text{vs.} \quad H_1^B : \sigma_\beta^2 > 0, \end{aligned} \tag{13.8.1}$$

and

$$H_0^{AB} : \sigma_{\alpha\beta}^2 = 0 \quad \text{vs.} \quad H_1^{AB} : \sigma_{\alpha\beta}^2 > 0.$$

13.8.1 SOME APPROXIMATE TESTS

Hirotsu (1968) proposed approximate F-tests for testing the hypotheses in (13.8.1) by using the test statistics analogous to those in the balanced case where now the mean squares are those obtained in the unweighted means analysis considered in Section 13.4.3.1. Denoting these mean squares as MS_{Au}, MS_{Bu}, MS_{ABu}, and MS_E; the test statistics used are

$$\begin{aligned} &\mathrm{MS}_{Au}/\mathrm{MS}_{ABu} \quad \text{for } H_0^A, \\ &\mathrm{MS}_{Bu}/\mathrm{MS}_{ABu} \quad \text{for } H_0^B, \end{aligned} \tag{13.8.2}$$

and

$$\mathrm{MS}_{ABu}/\mathrm{MS}_E \quad \text{for } H_0^{AB}.$$

The test statistics in (13.8.2) are to be compared with the $100(1-\alpha)$ percentage points of the F-distribution with degrees of freedom $[(a-1), (a-1)(b-1)]$, $[(b-1), (a-1)(b-1)]$, and $[(a-1)(b-1), (N-ab)]$, respectively. Hirotsu (1968) gives the expressions for the power functions of these tests with numerical examples, which, however, tend to be very complex. Hirotsu reported that α-levels of the approximate F-tests in (13.8.2) do not differ greatly from the nominal value when it is taken as 0.05 and the powers of the tests are close to that of the usual F-tests whose cell frequencies are all equal to the harmonic mean of the original cell frequencies, provided that the coefficient of variation of n_{ij}^{-1}s (inverses of cell frequencies) is small.

Approximate F-tests for $\sigma_\alpha^2 = 0$, $\sigma_\beta^2 = 0$, and $\sigma_{\alpha\beta}^2 = 0$ in (13.8.1) can also be constructed by determining linear combinations of means squares to be used as the F-ratios in the conventional analysis of variance given in Section 13.2. For example, from Table 13.1, to test $H_0^A : \sigma_\alpha^2 = 0$ vs. $H_1^A : \sigma_\alpha^2 > 0$, the test procedure can be based on $\mathrm{MS}_A/\mathrm{MS}_D$, where MS_D is given by

$$\mathrm{MS}_D = \ell_1 \mathrm{MS}_B + \ell_2 \mathrm{MS}_{AB} + (1 - \ell_1 - \ell_2)\mathrm{MS}_E,$$

with

$$\ell_1 = \frac{r_2 r_7 - r_1 r_8}{r_2 r_4 - r_1 r_5} \quad \text{and} \quad \ell_2 = \frac{r_4 r_8 - r_5 r_7}{r_2 r_4 - r_1 r_5}.$$

Similarly, to test $H_0^B : \sigma_\beta^2 = 0$ vs. $H_1^B : \sigma_\beta^2 > 0$, the test procedure can be based on $\mathrm{MS}_B/\mathrm{MS}'_D$, where MS'_D is given by

$$\mathrm{MS}'_D = \ell'_1 \mathrm{MS}_A + \ell'_2 \mathrm{MS}_{AB} + (1 - \ell'_1 - \ell'_2)\mathrm{MS}_E$$

with

$$\ell'_1 = \frac{r_3 r_4 - r_1 r_5}{r_3 r_7 - r_1 r_9} \quad \text{and} \quad \ell'_2 = \frac{r_5 r_7 - r_4 r_9}{r_3 r_7 - r_1 r_9}.$$

Finally, to test $H_0^{AB} : \sigma_{\alpha\beta}^2 = 0$ vs. $H_1^{AB} : \sigma_{\alpha\beta}^2 > 0$, the test procedure can be based on $\mathrm{MS}_{AB}/\mathrm{MS}''_D$, where MS''_D is given by

$$\mathrm{MS}''_D = \ell''_1 \mathrm{MS}_A + \ell''_2 \mathrm{MS}_B + (1 - \ell''_1 - \ell''_2)\mathrm{MS}_E,$$

with

$$\ell''_1 = \frac{r_2 r_6 - r_3 r_5}{r_6 r_8 - r_5 r_9} \quad \text{and} \quad \ell''_2 = \frac{r_3 r_8 - r_2 r_9}{r_6 r_8 - r_5 r_9}.$$

The test statistics $\mathrm{MS}_A/\mathrm{MS}_D$, $\mathrm{MS}_B/\mathrm{MS}'_D$, and $\mathrm{MS}_{AB}/\mathrm{MS}''_D$ are approximated as F-variables with $(a-1, \nu_D)$, $(b-1, \nu'_D)$, and $((a-1)(b-1), \nu''_D)$ degrees of freedom, respectively, where ν_D, ν'_D, and ν''_D are estimated using the Satterthwaite formula. Similar, pseudo F-tests can also be considered using synthesized mean squares based on weighted means analysis considered in Section 13.4.3.2.

13.8.2 SOME EXACT TESTS

Spjøtvoll (1968) and Thomsen (1975) have derived exact tests for $\sigma^2_{\alpha\beta} = 0$ in (13.8.1) which are equivalent. An exact test for $\sigma^2_{\alpha\beta} = 0$ in (13.8.1) can also be based on the Wald interval for $\sigma^2_{\alpha\beta}/\sigma^2_e$ mentioned in the preceding section. Burdick and Graybill (1992, pp. 140–141) illustrate this procedure with a numerical example using SAS® code. It has been shown by Seely and El-Bassiouni (1983) that the Wald test is equivalent to the Spjøtvoll–Thomsen test. Seely and El-Bassiouni (1983) have also shown that it is not possible to construct Wald-type exact tests for $\sigma^2_\alpha = 0$ or $\sigma^2_\beta = 0$ in (13.8.1) unless $\sigma^2_{\alpha\beta} = 0$. Spjøtvoll (1968) and Thomsen (1975) proposed exact tests for these hypotheses under the assumption that $\sigma^2_{\alpha\beta} = 0$. Khuri and Littell (1987) proposed exact tests for $\sigma^2_\alpha = 0$ and $\sigma^2_\beta = 0$ without assuming that $\sigma^2_{\alpha\beta} = 0$ by employing appropriate orthogonal transformations to the model for the cell means $\bar{y}_{ij.}$s. The procedure leads to a decomposition of independent sums of squares that are scalar multiples of chi-square random variables and can be used to obtain F-ratios similar to those in the balanced case. Through the results of a simulation study, the authors have noted that Satterthwaite-type tests can be highly unreliable and their exact tests have superior power properties. The procedure, however, requires a nonunique partitioning of the error sum of squares and is difficult to implement in practice. Burdick and Graybill (1992, p. 139) also consider an approximate test for $\sigma^2_\alpha = 0$ or $\sigma^2_\beta = 0$ in (13.8.1) based on lower bounds formed using the mean squares obtained in the unweighted means analysis. The test statistics, however, are the same as given in (13.8.2). Hernández (1991) has investigated the power function of this test vis-à-vis the Khuri–Littell test and has reported similar power properties for the two tests. Tan et al. (1988), using a harmonic mean approach, have reported tests for the hypotheses in (13.8.1) for the case involving heteroscedastic error variances. For a concise discussion and derivation of some of these tests, see Khuri et al. (1998, pp. 104–112).

13.8.3 A NUMERICAL EXAMPLE

In this example, we illustrate the application of psuedo F-tests for testing the hypotheses in (13.8.1) using the analysis of variance for the efficiency score data based on unweighted and weighted sums of squares given in Tables 13.11 and 13.12. From Table 13.11, the F-tests for testing $\sigma^2_{\alpha\beta} = 0, \sigma^2_\beta = 0$, and

$\sigma_\alpha^2 = 0$ yield F-values of 24.89, 1.31, and 4.56, which are to be compared against the theoretical F-values with (6,35), (3,6), and (2,6) degrees of freedom, respectively. The corresponding p-values are < 0.0001, 0.355, and 0.062, respectively. Thus there is very strong evidence of interaction effects between workers and sites; however, there are no significant differences between workers as well as between the sites. Note that the variance component estimate for sites is rather large. However, the F-test for $\sigma_\alpha^2 = 0$ has so few degrees of freedom that it may not be able to detect significant differences even if there are really important differences among them.

From Table 13.12, the F-value for testing $\sigma_{\alpha\beta}^2 = 0$ is 39.46, with a p-value of < 0.0001, which is again highly significant. Now, for testing $\sigma_\beta^2 = 0$ and $\sigma_\alpha^2 = 0$, the synthesized mean squares to be used as the denominators for site and worker mean squares, and the corresponding degrees of freedom are

$$\mathrm{MS}_D = 0.988 \times 160.211 + 0.012 \times 4.060 = 158.337,$$
$$\mathrm{MS}'_D = 0.786 \times 160.211 + 0.214 \times 4.060 = 126.795,$$
$$\nu_D = \frac{(158.337)^2}{\dfrac{(0.988 \times 160.211)^2}{6} + \dfrac{(0.0124 \times 4.060)^2}{35}} = 6.0,$$

and

$$\nu'_D = \frac{(126.795)^2}{\dfrac{(0.786 \times 160.211)^2}{6} + \dfrac{(0.214 \times 4.060)^2}{35}} = 6.1.$$

The F-tests for $\sigma_\beta^2 = 0$ and $\sigma_\alpha^2 = 0$ based on MS_D and MS'_D yield F-values of 0.91 and 3.16, with the corresponding p-values of 0.489 and 0.115, respectively. Thus the conclusions based on the unweighted as well as the weighted means analyses are the same.

EXERCISES

1. Apply the method of "synthesis" to derive the expected mean squares given in Table 13.1
2. For the model in (4.1.1) with proportional frequencies, i.e., $n_{ij} = (n_{i.}n_{.j})/N$, do the following:

   (a) Show that the expected mean squares are given by (Wilk and Kempthorne, 1955)

$$E(\mathrm{MS}_A) = \sigma_e^2 + \frac{N}{a-1}\left(1 - \sum_{i=1}^{a} \frac{n_{i.}^2}{N^2}\right)\left(\sum_{j=1}^{b} \frac{n_{.j}^2}{N^2}\sigma_{\alpha\beta}^2 + \sigma_\alpha^2\right),$$

$$E(\mathrm{MS}_B) = \sigma_e^2 + \frac{N}{b-1}\left(1 - \sum_{j=1}^{b} \frac{n_{.j}^2}{N^2}\right)\left(\sum_{i=1}^{a} \frac{n_{i.}^2}{N^2}\sigma_{\alpha\beta}^2 + \sigma_\beta^2\right),$$

$$E(\mathrm{MS}_{AB}) = \sigma_e^2 + \frac{N}{(a-1)(b-1)}\left\{\sum_{i=1}^{a} \frac{n_{i.}}{N}\left(1 - \frac{n_{i.}}{N}\right)\right\} \times \left\{\sum_{j=1}^{b} \frac{n_{.j}}{N}\left(1 - \frac{n_{.j}}{N}\right)\right\}\sigma_{\alpha\beta}^2,$$

and

$$E(\mathrm{MS}_E) = \sigma_e^2.$$

(b) Find estimators of the variance components using the results in part (a) and derive expressions for the variances of these estimators.

(c) Describe the procedures for testing the hypotheses regarding the variance components using the results in parts (a) and (b).

3. Proceeding from the definition of SS_{AB} given in (13.2.1), show that (Searle, 1987, p. 129)

$$\mathrm{SS}_{AB} = \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 - 2\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}(\bar{y}_{i..} - \bar{y}_{...})(\bar{y}_{.j.} - \bar{y}_{...}).$$

Hence, show that SS_{AB} is not a sum of squares and, in fact it can assume a negative value.

4. Consider the estimators of the variance components σ_α^2 and σ_β^2 given by

$$\hat{\sigma}_{\alpha,\mathrm{AVE}}^2 = \frac{1}{b(a-1)(b-1)}\sum_{i=1}^{a}\sum_{\substack{j,j'\\ j\neq j'}}^{b}(\bar{y}_{ij.} - \bar{y}_{.j.})(\bar{y}_{ij'.} - \bar{y}_{.j'.}),$$

$$\hat{\sigma}_{\beta,\mathrm{AVE}}^2 = \frac{1}{a(a-1)(b-1)}\sum_{j=1}^{b}\sum_{\substack{i,i'\\ i\neq i'}}^{a}(\bar{y}_{ij.} - \bar{y}_{i..})(\bar{y}_{i'j.} - \bar{y}_{i'..}),$$

where

$$\bar{y}_{ij.} = \frac{1}{\bar{n}_h}\sum_{k=1}^{n_{ij}} y_{ijk}, \qquad \bar{y}_{i..} = \frac{1}{b}\sum_{j=1}^{b}\bar{y}_{ij.}, \qquad \bar{y}_{.j.} = \frac{1}{a}\sum_{i=1}^{a}\bar{y}_{ij.},$$

with

$$\bar{n}_h = ab/\left(\sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij}^{-1}\right).$$

Show that (Hocking et al., 1989)

$$E(\hat{\sigma}^2_{\alpha,\text{AVE}}) = \sigma^2_\alpha \quad \text{and} \quad E(\hat{\sigma}^2_{\beta,\text{AVE}}) = \sigma^2_\beta.$$

5. Refer to Exercise 4.16 and suppose that the observations (block 1, variety 2, replication 3) and (block 2, variety 3, replication 1) are missing due to mishaps. For the resulting two-way factorial design, respond to the following questions:

   (a) Describe the mathematical model with interaction effect and the assumptions involved.
   (b) Analyze the data and report the analysis of variance table based on Henderson's Method I.
   (c) Perform an appropriate F-test to determine whether the dry matter content varies from variety to variety.
   (d) Perform an appropriate F-test to determine whether the dry matter content varies from block to block.
   (e) Perform an appropriate F-test for interaction effects between blocks and varieties.
   (f) Find point estimates of the variance components and the total variance using the methods described in the text.
   (g) Calculate 95% confidence intervals associated with the point estimates in part (f) using the methods described in the text.

6. Refer to Exercise 4.17 and suppose that the observations (block 3, variety 1, replication 2) and (block 1, variety 2, replication 1) are missing due to mishaps. For the resulting two-way factorial design, respond to the following questions:

   (a) Describe the mathematical model with interaction effect and the assumptions involved.
   (b) Analyze the data and report the analysis of variance table based on Henderson's Method I.
   (c) Perform an appropriate F-test to determine whether the plant height varies from block to block.
   (d) Perform an appropriate F-test to determine whether the plant height varies from variety to variety.

(e) Perform an appropriate F-test for interaction effects between blocks and varieties.

(f) Find point estimates of the variance components and the total variance using the methods described in the text.

(g) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

7. Refer to Exercise 4.18 and suppose that the observations (reagent 2, catalyst 3, replication 2) and (reagent 3, catalyst 1, replication 1) are missing due to mishaps. For the resulting two-way factorial design, respond to the following questions:

   (a) Describe the mathematical model with interaction effect and the assumptions involved.

   (b) Analyze the data and report the analysis of variance table based on Henderson's Method I.

   (c) Perform an appropriate F-test to determine whether the production rate varies from reagent to reagent.

   (d) Perform an appropriate F-test to determine whether the production rate varies from catalyst to catalyst.

   (e) Perform an appropriate F-test for interaction effects between reagents and catalysts.

   (f) Find point estimates of the variance components and the total variance using the methods described in the text.

   (g) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

8. Refer to Exercise 4.19 and suppose that the observations (therapist 1, patient 2, replication 1) and (therapist 2, patient 3, replication 2) are missing due to mishaps. For the resulting two-way factorial design, respond to the following questions:

   (a) Describe the mathematical model with interaction effect and the assumptions involved.

   (b) Analyze the data and report the analysis of variance table based on Henderson's Method I.

   (c) Perform an appropriate F-test to determine whether the anxiety reduction differs from therapist to therapist.

   (d) Perform an appropriate F-test to determine whether the anxiety reduction differs from patient to patient.

   (e) Perform an appropriate F-test for interaction effects between therapists and patients.

   (f) Find point estimates of the variance components and the total variance using the methods described in the text.

(g) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

9. Refer to Exercise 4.20 and suppose that the observations (machine 1, operator 2, replication 2) and (machine 2, temperature 3, replication 1) are missing due to mishaps. For the resulting two-way factorial design, respond to the following questions:

   (a) Describe the mathematical model with interaction effect and the assumptions involved.
   (b) Analyze the data and report the analysis of variance table based on Henderson's Method I.
   (c) Perform an appropriate F-test to determine whether the absolute diameter difference differs from machine to machine.
   (d) Perform an appropriate F-test to determine whether the absolute diameter difference differs from operator to operator.
   (e) Perform an appropriate F-test for interaction effects between machines and operators.
   (f) Find point estimates of the variance components and the total variance using the methods described in the text.
   (g) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

10. Refer to Exercise 4.21 and suppose that the observations (oven 2, temperature 3, replication 1) and (oven 3, temperature 4, replication 2) are missing due to mishaps. For the resulting two-way factorial design, respond to the following questions:

    (a) Describe the mathematical model with interaction effect and the assumptions involved.
    (b) Analyze the data and report the analysis of variance table based on Henderson's Method I.
    (c) Perform an appropriate F-test to determine whether the quality of texture differs from oven to oven.
    (d) Perform an appropriate F-test to determine whether the quality of texture differs from temperature to temperature.
    (e) Perform an appropriate F-test for interaction effects between ovens and temperatures.
    (f) Find point estimates of the variance components and the total variance using the methods described in the text.
    (g) Calculate percent confidence intervals of the variance components and the total variance using the methods described in the text.

11. Refer to Exercise 4.22 and suppose that the observations (projectile 3, propeller 4, replication 2) and (projectile 4, propeller 1, replication 1)

are missing due to mishaps. For the resulting two-way factorial design, respond to the following questions:

(a) Describe the mathematical model with interaction effect and the assumptions involved.
(b) Analyze the data and report the analysis of variance table based on Henderson's Method I.
(c) Perform an appropriate F-test to determine whether the muzzle velocity differs from projectile to projectile.
(d) Perform an appropriate F-test to determine whether the muzzle velocity differs from propeller to propeller.
(e) Perform an appropriate F-test for interaction effects between projectiles and propellers.
(f) Find point estimates of the variance components and the total variance using the methods described in the text.
(g) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

12. Refer to Exercise 4.23 and suppose that the observations (gauger 1, breaker 2, replication 1) and (gauger 2, breaker 3, replication 2) are missing due to mishaps. For the resulting two-way factorial design, respond to the following questions:

(a) Describe the mathematical model with interaction effect and the assumptions involved.
(b) Analyze the data and report the analysis of variance table based on Henderson's Method I.
(c) Perform an appropriate F-test to determine whether the testing strength differs from gauger to gauger.
(d) Perform an appropriate F-test to determine whether the testing strength differs from breaker to breaker.
(e) Perform an appropriate F-test for interaction effects between gaugers and breakers.
(f) Find point estimates of the variance components and the total variance using the methods described in the text.
(g) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

13. Refer to Exercise 12.11 and respond to the following questions:

(a) Describe the mathematical model with interaction effect and the assumptions involved.
(b) Analyze the data and report the analysis of variance table based on Henderson's Method I.

(c) Perform an appropriate F-test to determine whether the intensity of radiation differs from location to location.
(d) Perform an appropriate F-test to determine whether the intensity of radiation differs for different time periods of the day.
(e) Perform an appropriate F-test for interaction effects between locations and time periods of the day.
(f) Find point estimates of the variance components and the total variance using the methods described in the text.
(g) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

14. Refer to Exercise 12.12 and respond to the following questions:

(a) Describe the mathematical model with interaction effect and the assumptions involved.
(b) Analyze the data and report the analysis of variance table based on Henderson's Method I.
(c) Perform an appropriate F-test to determine whether the development period differs from strain to strain.
(d) Perform an appropriate F-test to determine whether the development period differs from density to density.
(e) Perform an appropriate F-test for interaction effects between strains and densities.
(f) Find point estimates of the variance components and the total variance using the methods described in the text.
(g) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

15. Refer to Exercise 12.13 and respond to the following questions:

(a) Describe the mathematical model with interaction effect and the assumptions involved.
(b) Analyze the data and report the analysis of variance table based on Henderson's Method I.
(c) Perform an appropriate F-test to determine whether the oven-dry weight differs from soil to soil.
(d) Perform an appropriate F-test to determine whether the oven-dry weight differs from variety to variety.
(e) Perform an appropriate F-test for interaction effects between soils and varieties.
(f) Find point estimates of the variance components and the total variance using the methods described in the text.
(g) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

16. Refer to Exercise 12.14 and respond to the following questions:

    (a) Describe the mathematical model with interaction effect and the assumptions involved.
    (b) Analyze the data and report the analysis of variance table based on Henderson's Method I.
    (c) Perform an appropriate F-test to determine whether the percent reduction in blood sugar differs from preparation to preparation.
    (d) Perform an appropriate F-test to determine whether the percent reduction in blood sugar differs from dose to dose.
    (e) Perform an appropriate F-test for interaction effects between levels of preparation and dose.
    (f) Find point estimates of the variance components and the total variance using the methods described in the text.
    (g) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

Bibliography

R. L. Anderson (1961), Designs for estimating variance components, in *Proceedings of the Seventh Conference on Design of Experiments and Army Preservation and Development Testing*, 781–823; also published as Mim. Ser. 310, Institute of Statistics, North Carolina State University, Raleigh, NC.

R. L. Anderson (1975), Designs and estimators for variance components, in J. N. SrivastavaJ. N., ed., *Statistical Design and Linear Model*, North-Holland, Amsterdam, 1–30.

R. L. Anderson (1981), Recent developments in designs and estimators for variance components, in M. Csörgö, D. A. Dawson, J. N. K. Rao, and A. K. Md. E. Saleh, eds., *Statistics and Related Topics*, North-Holland, Amsterdam, 3–22.

R. Bremer (1989), Numerical study of small sample variances of estimators of variance components in the two-way factorial model, *Comm. Statist.* B *Simul. Comput.*, **18**, 985–1009.

R. Bremer (1990), Numerical study of the small sample variance of the MINQUE estimators of variance components in the two-way factorial design, *Comm. Statist.* A *Theory Methods*, **19**, 1261–1279.

R. K. Burdick and F. A. Graybill (1992), *Confidence Intervals on Variance Components*, Marcel Dekker, New York.

N. Bush (1962), *Estimating Variance Components in a Multi-Way Classification*, Ph.D. dissertation, Mim. Ser. 324, North Carolina State University, Raleigh, NC.

N. Bush and R. L. Anderson (1963), A comparison of three different procedures for estimating variance components, *Technometrics*, **5**, 421–440.

S. L. Crump (1947), *The Estimation of Components of Variance in Multiple Classifications*, Ph.D. dissertation, Iowa State University, Ames, IA.

D. W. Gaylor (1960), *The Construction and Evaluation of Some Designs for the Estimation of Parameters in Random Models*, Ph.D. dissertation, Mim. Ser. 256, Institute of Statistics, North Carolina State University, Raleigh, NC.

D. W. Gaylor, H. L. Lucas, and R. L. Anderson (1970), Calculation of expected mean squares by the abbreviated Doolittle and square root methods, *Biometrics*, **20**, 641–656.

D. G. Gosslee and H. L. Lucas (1965), Analysis of variance of disproportionate data when interaction is present, *Biometrics*, **21**, 115–133.

A. Haile and J. T. Webster (1975), A comparison of four designs for estimating the variances of random effects in a two-way classification, *Technometrics*, **17**, 495–500.

R. P. Hernández (1991), *Confidence Intervals on Linear Combinations of Variance Components in Unbalanced Designs*, Ph.D. dissertation, Arizona State University, Tempe, AZ.

R. P. Hernández and R. K. Burdick (1993), Confidence intervals and tests of hypotheses on variance components in an unbalanced two-factor crossed design with interaction, *J. Statist. Comput. Simul.*, **47**, 67–77.

C. Hirotsu (1966), Estimating variance, components in a two-way layout with unequal numbers of observations, *Rep. Statist. Appl. Res.* (*JUSE*), **13**, 29–34.

C. Hirotsu (1968), An approximate test for the case of random effects model in a two-way layout with unequal cell frequencies, *Rep. Statist. Appl. Res.* (*JUSE*), **15**, 13–26.

R. R. Hocking, J. W. Green, and R. H. Bremer (1989), Variance component estimation with model-based diagnostics, *Technometrics*, **31**, 227–240.

M. K. Kazempour and F. A. Graybill (1991), Approximate confidence bounds for ratios of variance components in two-way unbalanced crossed models with interactions, *Comm. Statist.* B *Simul. Comput.*, **20**, 955–967.

A. I. Khuri (1998), On unweighted sums of squares in unbalanced analysis of variance, *J. Statist. Plann. Inference*, **74**, 135–147.

A. I. Khuri and R. C. Littell (1987), Exact tests for the main effects variance components in an unbalanced random two-way model, *Biometrics*, **43**, 545–560.

A. I. Khuri, T. Mathew, and B. K. Sinha (1998), *Statistical Tests for Mixed Linear Models*, Wiley, New York.

J. Kleffé (1980), C. R. Rao's MINQUE under four two-way ANOVA models, *Biometrical J.*, **22**, 93–104.

G. G. Koch (1967), A general approach to the estimation of variance components, *Techometrics*, **9**, 93–118.

G. G. Koch (1968), Some further remarks concerning "A general approach to estimation of variance components," *Technometrics*, **10**, 551–558.

C. Y. Lin and A. J. McAllister (1984), Monte Carlo comparison of four methods for estimation of genetic parameters in the univariate case, *J. Dairy Sci.*, **67**, 2389–2398.

L. Y. Low (1964), Sampling variances of estimates of components of variance from a non-orthogonal two-way classification, *Biometrika*, **51**, 491–494.

G. A. Milliken and D. E. Johnson (1992), *Analysis of Messy Data*, Vol. I, Chapman and Hall, London.

M. G. Mostafa (1967), Note on testing hypotheses in an unbalanced random effects model, *Biometrika*, **54**, 659–662.

H. D. Muse (1974), *Comparison of Designs to Estimate Variance Components in a Two-Way Classification Model*, Ph.D. dissertation, University of Kentucky, Lexington, KY.

H. D. Muse and R. L. Anderson (1978), Comparison of designs to estimate variance components in a two-way classification model, *Technometrics*, **20**, 159–166.

H. D. Muse, R. L. Anderson, and B. Thitakamol (1982), Additional comparisons of designs to estimate variance components in a two-way classification model, *Comm. Statist.* A *Theory Methods*, **11**, 1403–1425.

D. G. Pederson (1972), A comparison of four experimental designs for the estimation of heritability, *Theoret. Appl. Genetics*, **42**, 371–377.

C. A. Rhode and G. M. Tallis (1969), Exact first- and second-order moments of estimates of components of covariance, *Biometrika*, **56**, 517–525.

L. R. Schaeffer (1973), *Factors Affecting the Accuracy of Variance Components Estimation*, Ph.D. dissertation, Cornell University, Ithaca, NY.

S. R. Searle (1958), Sampling variances of estimates of components of variance, *Ann. Math. Statist.*, **29**, 167–178.

S. R. Searle (1971), *Linear Models*, Wiley, New York.

S. R. Searle (1987), *Linear Models for Unbalanced Data*, Wiley, New York.

S. R. Searle, G. Casella, and C. E. McCulloch (1992), *Variance Components*, Wiley, New York.

S. R. Searle and C. R. Henderson (1961), Computing procedures for estimating components of variance in the two-way classification, mixed model, *Biometrics*, **17**, 607–616; corrigenda, 1967, **23**, 852.

J. F. Seely and M. Y. El-Bassiouni (1983), Applying Wald's variance component test, *Ann. Statist.*, **11**, 197–201.

P.-S. Shen, P. L. Cornelius, and R. L. Anderson (1996), Planned unbalanced designs for estimation of quantitative genetic parameters I: Two-way matings, *Biometrics*, **52**, 56–70.

E. Spjøtvoll (1968), Confidence intervals and tests for variance ratios in unbalanced variance components models, *Rev. Internat. Statist. Inst.*, **36**, 37–42.

S. Srinivasan (1986), *Confidence Intervals on Functions of Variance Components in Unbalanced Two-Way Design Models*, Ph.D. dissertation, Colorado State University, Fort Collins, CO.

W. W. Stroup, J. W. Evans, and R. L. Anderson (1980), Maximum likelihood estimation of variance components in a completely random BIB design, *Comm.*

Statist. A *Theory Methods*, **9**, 725–756.

W. Y. Tan, M. A. Tabatabai, and N. Balakrishnan (1988), Harmonic mean aproach to unbalanced random effects model under heteroscedasticity, *Comm. Statist.* A *Theory Methods*, **17**, 1261–1286.

B. Thitakamol (1977), *Extension of Previous Results on Properties of Estimators of Variance Components*, Ph.D. dissertation, University of Kentucky, Lexingon, KY.

R. Thompson (1976), Design of experiments to estimate heritability when observations are available on parents and offspring, *Biometrics*, **32**, 283–304.

R. Thompson (1977), The estimation of heritability with unbalanced data I: Observations available on parents and offspring, *Biometrics*, **33**, 485–495.

I. B. Thomsen (1975), Testing hypotheses in unbalanced variance components models for two-way layouts, *Ann. Statist.*, **3**, 257–265.

M. B. Wilk and O. Kempthorne (1955), Fixed, mixed, and random models, *J. Amer. Statist. Assoc.*, **50**, 1144–1167; corrigenda, **51**, 652.

F. Yates (1934), The analysis of multiple classifications with unequal numbers in the different classes, *J. Amer. Statist. Assoc.*, **29**, 51–66.

14 Three-Way and Higher-Order Crossed Classifications

Crossed classifications involving several factors are common in experiments and surveys in many substantive fields of research. Consider three factors A, B, and C with a, b, and c levels, respectively, involving a factorial arrangement. Assume that n_{ijk} (≥ 0) observations are taken corresponding to the (i, j, k)th cell. The model for this design is known as the unbalanced three-way crossed-classification model. This model is the same as the one considered in Chapter 5 except that now the number of observations per cell is not constant but varies from cell to cell including some cells with no data. Models of this type frequently occur in many experiments and surveys since many investigations cannot guarantee the same number of observations for each cell. In this chapter, we briefly outline the analysis of random effects model for the unbalanced three-way crossed-classification with interaction and indicate its extension to higher-order classifications.

14.1 MATHEMATICAL MODEL

The random effects model for the unbalanced three-way crossed classification with interactions is given by

$$\begin{aligned} y_{ijk\ell} &= \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} \\ &\quad + (\alpha\beta\gamma)_{ijk} + e_{ijk\ell} \begin{cases} i = 1, 2, \ldots, a, \\ j = 1, 2, \ldots, b, \\ k = 1, 2, \ldots, c, \\ \ell = 1, 2, \ldots, n_{ijk}, \end{cases} \end{aligned} \tag{14.1.1}$$

where $y_{ijk\ell}$ is the ℓth observation at the ith level of factor A, the jth level of factor B, and the kth level of factor C, μ is the overall mean, α_is, β_js, and γ_ks are main effects; $(\alpha\beta)_{ij}$s, $(\alpha\gamma)_{ik}$s, and $(\beta\gamma)_{jk}$s are two-factor interaction terms; $(\alpha\beta\gamma)_{ijk}$s are three-factor interaction terms, and $e_{ijk\ell}$s are customary error terms. It is assumed that $-\infty < \mu < \infty$ is a constant and α_is, β_js, γ_ks,

TABLE 14.1 Analysis of variance for the model in (14.1.1).

Source of variation	Degrees of freedom*	Sum of squares	Mean square	Expected mean square
Factor *A*	$a-1$	SS_A	MS_A	δ_α^2
Factor *B*	$b-1$	SS_B	MS_B	δ_β^2
Factor *C*	$c-1$	SS_C	MS_C	δ_γ^2
Interaction *A* x *B*	$s_{ab}-a-b+1$	SS_{AB}	MS_{AB}	$\delta_{\alpha\beta}^2$
Interaction *A* x *C*	$s_{ac}-a-c+1$	SS_{AC}	MS_{AC}	$\delta_{\alpha\gamma}^2$
Interaction *B* x *C*	$s_{bc}-b-c+1$	SS_{BC}	MS_{BC}	$\delta_{\beta\gamma}^2$
Interaction *A* x *B* x *C*	$s-s_{ab}-s_{ac}-s_{bc}+a+b+c-1$	SS_{ABC}	MS_{ABC}	$\delta_{\alpha\beta\gamma}^2$
Error	$N-s$	SS_E	MS_E	δ_E^2

*s = number of nonempty ABC-subclasses, s_{ab} = number of nonempty AB-subclasses, s_{ac} = number of nonempty AC-subclasses, s_{bc} = number of nonempty BC-subclasses.

$(\alpha\beta)_{ij}$s, $(\alpha\gamma)_{ik}$s, $(\beta\gamma)_{jk}$s, $(\alpha\beta\gamma)_{ijk}$s, and e_{ijk}s are mutually and completely uncorrelated random variables with means zero and variances σ_α^2, σ_β^2, σ_γ^2, $\sigma_{\alpha\beta}^2$, $\sigma_{\alpha\gamma}^2$, $\sigma_{\beta\gamma}^2$, $\sigma_{\alpha\beta\gamma}^2$, and σ_e^2, respectively. The parameters σ_α^2, σ_β^2, σ_γ^2, $\sigma_{\alpha\beta}^2$, $\sigma_{\alpha\gamma}^2$, $\sigma_{\beta\gamma}^2$, $\sigma_{\alpha\beta\gamma}^2$, and σ_e^2 are the variance components of the model in (14.1.1).

14.2 ANALYSIS OF VARIANCE

For the model in (14.1.1) there is no unique analysis of variance. The conventional analysis of variance obtained by an analogy with corresponding balanced analysis can be given in the form of Table 14.1. The sums of squares in Table 14.1 are defined as follows:

$$\mathrm{SS}_A = \sum_{i=1}^{a} n_{i..}(\bar{y}_{i...} - \bar{y}_{....})^2 = \sum_{i=1}^{a} \frac{y_{i...}^2}{n_{i..}} - \frac{y_{....}^2}{N},$$

$$\mathrm{SS}_B = \sum_{j=1}^{b} n_{.j.}(\bar{y}_{.j..} - \bar{y}_{....})^2 = \sum_{j=1}^{b} \frac{y_{.j..}^2}{n_{.j.}} - \frac{y_{....}^2}{N},$$

$$\mathrm{SS}_C = \sum_{k=1}^{c} n_{..k}(\bar{y}_{..k.} - \bar{y}_{....})^2 = \sum_{k=1}^{c} \frac{y_{..k.}^2}{n_{..k}} - \frac{y_{....}^2}{N},$$

$$
\begin{aligned}
\mathrm{SS}_{AB} &= \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij.}\bar{y}_{ij..}^2 - \sum_{i=1}^{a} n_{i..}\bar{y}_{i...}^2 - \sum_{j=1}^{b} n_{.j.}\bar{y}_{.j..}^2 + N\bar{y}_{....}^2 \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{y_{ij..}^2}{n_{ij.}} - \sum_{i=1}^{a} \frac{y_{i...}^2}{n_{i..}} - \sum_{j=1}^{b} \frac{y_{.j..}^2}{n_{.j.}} + \frac{y_{....}^2}{N}, \\
\mathrm{SS}_{AC} &= \sum_{i=1}^{a}\sum_{k=1}^{c} n_{i.k}\bar{y}_{i.k.}^2 - \sum_{i=1}^{a} n_{i..}\bar{y}_{i...}^2 - \sum_{k=1}^{c} n_{..k}\bar{y}_{..k.}^2 + N\bar{y}_{....}^2 \\
&= \sum_{i=1}^{a}\sum_{k=1}^{c} \frac{y_{i.k.}^2}{n_{i.k}} - \sum_{i=1}^{a} \frac{y_{i...}^2}{n_{i..}} - \sum_{k=1}^{c} \frac{y_{..k.}^2}{n_{..k}} + \frac{y_{....}^2}{N}, \\
\mathrm{SS}_{BC} &= \sum_{j=1}^{b}\sum_{k=1}^{c} n_{.jk}\bar{y}_{.jk.}^2 - \sum_{j=1}^{b} n_{.j.}\bar{y}_{.j..}^2 - \sum_{k=1}^{c} n_{..k}\bar{y}_{..k.}^2 + N\bar{y}_{....}^2 \\
&= \sum_{j=1}^{b}\sum_{k=1}^{c} \frac{y_{.jk.}^2}{n_{.jk}} - \sum_{j=1}^{b} \frac{y_{.j..}^2}{n_{.j.}} - \sum_{k=1}^{c} \frac{y_{..k.}^2}{n_{..k}} + \frac{y_{....}^2}{N}, \\
\mathrm{SS}_{ABC} &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} n_{ijk}\bar{y}_{ijk.}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij.}\bar{y}_{ij..}^2 - \sum_{i=1}^{a}\sum_{k=1}^{c} n_{i.k}\bar{y}_{i.k.}^2 \\
&\quad - \sum_{j=1}^{b}\sum_{k=1}^{c} n_{.jk}\bar{y}_{.jk.}^2 + \sum_{i=1}^{a} n_{i..}\bar{y}_{i...}^2 + \sum_{j=1}^{b} n_{.j.}\bar{y}_{.j..}^2 \\
&\quad + \sum_{k=1}^{c} n_{..k}\bar{y}_{..k.}^2 - N\bar{y}_{....}^2 \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} \frac{y_{ijk.}^2}{n_{ijk}} - \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{y_{ij..}^2}{n_{ij.}} - \sum_{i=1}^{a}\sum_{k=1}^{c} \frac{y_{i.k.}^2}{n_{i.k}} - \sum_{j=1}^{b}\sum_{k=1}^{c} \frac{y_{.jk.}^2}{n_{.jk}} \\
&\quad + \sum_{i=1}^{a} \frac{y_{i...}^2}{n_{i..}} + \sum_{j=1}^{b} \frac{y_{.j..}^2}{n_{.j.}} + \sum_{k=1}^{c} \frac{y_{..k.}^2}{n_{..k}} - \frac{y_{....}^2}{N},
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{SS}_E &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{\ell=1}^{n_{ijk}} (y_{ijk\ell} - \bar{y}_{ijk.})^2 \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{\ell=1}^{n_{ijk}} y_{ijk\ell}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} \frac{y_{ijk.}^2}{n_{ijk}},
\end{aligned}
$$

where the customary notations for totals and means are employed.

Now, define the uncorrected sums of squares as

$$T_A = \sum_{i=1}^{a} \frac{y_{i...}^2}{n_{i..}}, \qquad T_B = \sum_{j=1}^{b} \frac{y_{.j..}^2}{n_{.j.}}, \qquad T_C = \sum_{k=1}^{c} \frac{y_{..k.}^2}{n_{..k}},$$

$$T_{AB} = \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{y_{ij..}^2}{n_{ij.}}, \quad T_{AC} = \sum_{i=1}^{a}\sum_{k=1}^{c} \frac{y_{i.k.}^2}{n_{i.k}}, \quad T_{BC} = \sum_{j=1}^{b}\sum_{k=1}^{c} \frac{y_{.jk.}^2}{n_{.jk}},$$

$$T_{ABC} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} \frac{y_{ijk.}^2}{n_{ijk}}, \quad T_\mu = \frac{y_{....}^2}{N}, \quad \text{and} \quad T_0 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{\ell=1}^{n_{ijk}} y_{ijk\ell}^2.$$

Then the corrected sums of squares defined earlier can be written as

$$\begin{aligned}
\text{SS}_A &= T_A - T_\mu, \qquad \text{SS}_B = T_B - T_\mu, \qquad \text{SS}_C = T_C - T_\mu,\\
\text{SS}_{AB} &= T_{AB} - T_A - T_B + T_\mu, \qquad \text{SS}_{AC} = T_{AC} - T_A - T_C + T_\mu,\\
\text{SS}_{BC} &= T_{BC} - T_B - T_C + T_\mu,\\
\text{SS}_{ABC} &= T_{ABC} - T_{AB} - T_{AC} - T_{BC} + T_A + T_B + T_C - T_\mu,
\end{aligned}$$

and

$$\text{SS}_E = T_0 - T_{ABC}.$$

The mean squares are obtained by dividing the sums of squares by the corresponding degrees of freedom. The results on expected mean squares are outlined in the following section.

14.3 EXPECTED MEAN SQUARES

Proceeding directly or using the algebraic results in (10.1.6), (10.1.7), and (10.1.8), or the "synthesis" method[1] of Hartley (1967), it can be shown that the results on expectations of uncorrected sums of squares are (see, e.g., Searle, 1971, Chapter 11):

$$\begin{aligned}
E(T_A) &= N\mu^2 + N\sigma_\alpha^2 + c_{i(j)}\sigma_\beta^2 + c_{i(k)}\sigma_\gamma^2 + c_{i(j)}\sigma_{\alpha\beta}^2 + c_{i(k)}\sigma_{\alpha\gamma}^2\\
&\quad + c_{i(jk)}\sigma_{\beta\gamma}^2 + c_{i(jk)}\sigma_{\alpha\beta\gamma}^2 + a\sigma_e^2,\\
E(T_B) &= N\mu^2 + c_{j(i)}\sigma_\alpha^2 + N\sigma_\beta^2 + c_{j(k)}\sigma_\gamma^2 + c_{j(i)}\sigma_{\alpha\beta}^2 + c_{j(ik)}\sigma_{\alpha\gamma}^2\\
&\quad + c_{j(k)}\sigma_{\beta\gamma}^2 + c_{j(ik)}\sigma_{\alpha\beta\gamma}^2 + b\sigma_e^2,\\
E(T_C) &= N\mu^2 + c_{k(i)}\sigma_\alpha^2 + c_{k(j)}\sigma_\beta^2 + N\sigma_\gamma^2 + c_{k(ij)}\sigma_{\alpha\beta}^2 + c_{k(i)}\sigma_{\alpha\gamma}^2
\end{aligned}$$

[1]The method of "synthesis" can be used to evaluate the expectations of sums of squares for any model involving random effects or multiple error terms. The method has been extended by Rao (1968) to general incidence matrices and to mixed models that also described how to evaluate the variances of the estimators (see also Searle, 1971, pp. 432–433; Hocking, 1985, pp. 336–339).

$$
\begin{aligned}
&+ c_{k(j)}\sigma^2_{\beta\gamma} + c_{k(ij)}\sigma^2_{\alpha\beta\gamma} + c\sigma^2_e,\\
E(T_{AB}) &= N\mu^2 + N\sigma^2_\alpha + N\sigma^2_\beta + c_{ij(k)}\sigma^2_\gamma + N\sigma^2_{\alpha\beta} + c_{ij(k)}\sigma^2_{\alpha\gamma}\\
&\quad + c_{ij(k)}\sigma^2_{\beta\gamma} + c_{ij(k)}\sigma^2_{\alpha\beta\gamma} + s_{ab}\sigma^2_e,\\
E(T_{AC}) &= N\mu^2 + N\sigma^2_\alpha + c_{ik(j)}\sigma^2_\beta + N\sigma^2_\gamma + c_{ik(j)}\sigma^2_{\alpha\beta} + N\sigma^2_{\alpha\gamma}\\
&\quad + c_{ik(j)}\sigma^2_{\beta\gamma} + c_{ik(j)}\sigma^2_{\alpha\beta\gamma} + s_{ac}\sigma^2_e,\\
E(T_{BC}) &= N\mu^2 + c_{jk(i)}\sigma^2_\alpha + N\sigma^2_\beta + N\sigma^2_\gamma + c_{jk(i)}\sigma^2_{\alpha\beta} + c_{jk(i)}\sigma^2_{\alpha\gamma}\\
&\quad + N\sigma^2_{\beta\gamma} + c_{jk(i)}\sigma^2_{\alpha\beta\gamma} + s_{bc}\sigma^2_e,\\
E(T_{ABC}) &= N\mu^2 + N\sigma^2_\alpha + N\sigma^2_\beta + N\sigma^2_\gamma + N\sigma^2_{\alpha\beta}\\
&\quad + N\sigma^2_{\alpha\gamma} + N\sigma^2_{\beta\gamma} + N\sigma^2_{\alpha\beta\gamma} + s\sigma^2_e,\\
E(T_\mu) &= N\mu^2 + d_i\sigma^2_\alpha + d_j\sigma^2_\beta + d_k\sigma^2_\gamma + d_{ij}\sigma^2_{\alpha\beta} + d_{ik}\sigma^2_{\alpha\gamma}\\
&\quad + d_{jk}\sigma^2_{\beta\gamma} + d_{ijk}\sigma^2_{\alpha\beta\gamma} + \sigma^2_e,
\end{aligned}
$$

and

$$
\begin{aligned}
E(T_0) &= N\mu^2 + N\sigma^2_\alpha + N\sigma^2_\beta + N\sigma^2_\gamma + N\sigma^2_{\alpha\beta}\\
&\quad + N\sigma^2_{\alpha\gamma} + N\sigma^2_{\beta\gamma} + N\sigma^2_{\alpha\beta\gamma} + N\sigma^2_e,
\end{aligned}
$$

where

$$
c_{i(j)} = \sum_{i=1}^{a} \frac{\sum_{j=1}^{b} n^2_{ij.}}{n_{i..}}, \qquad c_{k(ij)} = \sum_{k=1}^{c} \frac{\sum_{i=1}^{a}\sum_{j=1}^{b} n^2_{ij.}}{n_{..k}}, \qquad \text{etc.},
$$

$$
d_i = \frac{\sum_{i=1}^{a} n^2_{i..}}{N}, \qquad d_{jk} = \frac{\sum_{j=1}^{b}\sum_{k=1}^{c} n^2_{.jk}}{N}, \qquad \text{etc.},
$$

s is the number of ABC subclasses containing data and s_{ab}, s_{ac}, s_{bc} are, respectively, the number of AB-, AC-, and BC-subclasses containing nonempty cells.

Hence, expected sums of squares are given as follows:

$$
\begin{aligned}
E(\text{SS}_E) &= (N - s)\sigma^2_e,\\
E(\text{SS}_{ABC}) &= k^1_1\sigma^2_\alpha + k^1_2\sigma^2_\beta + k^1_3\sigma^2_\gamma + k^1_{12}\sigma^2_{\alpha\beta} + k^1_{13}\sigma^2_{\alpha\gamma} + k^1_{23}\sigma^2_{\beta\gamma}\\
&\quad + k^1_{123}\sigma^2_{\alpha\beta\gamma} + k^1_0\sigma^2_e,\\
E(\text{SS}_{BC}) &= k^2_1\sigma^2_\alpha + k^2_2\sigma^2_\beta + k^2_3\sigma^2_\gamma + k^2_{12}\sigma^2_{\alpha\beta} + k^2_{13}\sigma^2_{\alpha\gamma} + k^2_{23}\sigma^2_{\beta\gamma}\\
&\quad + k^2_{123}\sigma^2_{\alpha\beta\gamma} + k^2_0\sigma^2_e,\\
E(\text{SS}_{AC}) &= k^3_1\sigma^2_\alpha + k^3_2\sigma^2_\beta + k^3_3\sigma^2_\gamma + k^3_{12}\sigma^2_{\alpha\beta} + k^3_{13}\sigma^2_{\alpha\gamma} + k^3_{23}\sigma^2_{\beta\gamma}\\
&\quad + k^3_{123}\sigma^2_{\alpha\beta\gamma} + k^3_0\sigma^2_e,
\end{aligned}
$$

$$E(\mathrm{SS}_{AB}) = k_1^4\sigma_\alpha^2 + k_2^4\sigma_\beta^2 + k_3^4\sigma_\gamma^2 + k_{12}^4\sigma_{\alpha\beta}^2 + k_{13}^4\sigma_{\alpha\gamma}^2 + k_{23}^4\sigma_{\beta\gamma}^2 + k_{123}^4\sigma_{\alpha\beta\gamma}^2 + k_0^4\sigma_e^2,$$
$$E(\mathrm{SS}_{C}) = k_1^5\sigma_\alpha^2 + k_2^5\sigma_\beta^2 + k_3^5\sigma_\gamma^2 + k_{12}^5\sigma_{\alpha\beta}^2 + k_{13}^5\sigma_{\alpha\gamma}^2 + k_{23}^5\sigma_{\beta\gamma}^2 + k_{123}^5\sigma_{\alpha\beta\gamma}^2 + k_0^5\sigma_e^2,$$
$$E(\mathrm{SS}_{B}) = k_1^6\sigma_\alpha^2 + k_2^6\sigma_\beta^2 + k_3^6\sigma_\gamma^2 + k_{12}^6\sigma_{\alpha\beta}^2 + k_{13}^6\sigma_{\alpha\gamma}^2 + k_{23}^6\sigma_{\beta\gamma}^2 + k_{123}^6\sigma_{\alpha\beta\gamma}^2 + k_0^6\sigma_e^2,$$

and

$$E(\mathrm{SS}_{A}) = k_1^7\sigma_\alpha^2 + k_2^7\sigma_\beta^2 + k_3^7\sigma_\gamma^2 + k_{12}^7\sigma_{\alpha\beta}^2 + k_{13}^7\sigma_{\alpha\gamma}^2 + k_{23}^7\sigma_{\beta\gamma}^2 + k_{123}^7\sigma_{\alpha\beta\gamma}^2 + k_0^7\sigma_e^2,$$

where

$$k_1^1 = c_{j(i)} + c_{k(i)} - c_{jk(i)} - d_i, \quad k_2^1 = c_{i(j)} + c_{k(j)} - c_{ik(j)} - d_j,$$
$$k_3^1 = c_{i(k)} + c_{j(k)} - c_{ij(k)} - d_k,$$
$$k_{12}^1 = c_{i(j)} + c_{j(i)} + c_{k(ij)} - c_{ik(j)} - c_{jk(i)} - d_{ij},$$
$$k_{13}^1 = c_{i(k)} + c_{k(i)} + c_{j(ik)} - c_{ij(k)} - c_{jk(i)} - d_{ik},$$
$$k_{23}^1 = c_{j(k)} + c_{k(j)} + c_{i(jk)} - c_{ij(k)} - c_{ik(j)} - d_{jk},$$
$$k_{123}^1 = N + c_{i(jk)} + c_{j(ik)} + c_{k(ij)} - c_{ij(k)} - c_{ik(j)} - c_{jk(i)} - d_{ijk},$$
$$k_0^1 = s - s_{ab} - s_{ac} - s_{bc} + a + b + c - 1, \quad k_1^2 = c_{jk(i)} - c_{j(i)} - c_{k(i)} + d_i,$$
$$k_2^2 = d_j - c_{k(j)}, \quad k_3^2 = d_k - c_{j(k)}, \quad k_{12}^2 = c_{jk(i)} - c_{k(ij)} - c_{j(i)} + d_{ij},$$
$$k_{13}^2 = c_{jk(i)} - c_{j(ik)} - c_{k(i)} + d_{ik}, \quad k_{23}^2 = N - c_{j(k)} - c_{k(j)} + d_{jk},$$
$$k_{123}^2 = c_{jk(i)} - c_{j(ik)} - c_{k(ij)} + d_{ijk}, \quad k_0^2 = s_{bc} - b - c + 1,$$
$$k_1^3 = d_i - c_{k(i)}, \quad k_2^3 = c_{ik(j)} - c_{i(j)} - c_{k(j)} + d_j, \quad k_3^3 = d_k - c_{i(k)},$$
$$k_{12}^3 = c_{ik(j)} - c_{k(ij)} - c_{i(j)} + d_{ij}, \quad k_{13}^3 = N - c_{i(k)} - c_{k(i)} + d_{ik},$$
$$k_{23}^3 = c_{ik(j)} - c_{i(jk)} - c_{k(j)} + d_{jk}, \quad k_{123}^3 = c_{ik(j)} - c_{i(jk)} - c_{k(ij)} + d_{ijk},$$
$$k_0^3 = s_{ac} - a - c + 1, \quad k_1^4 = d_i - c_{j(i)}, \quad k_2^4 = d_j - c_{i(j)},$$
$$k_3^4 = c_{ij(k)} - c_{i(k)} - c_{j(k)} + d_k, \quad k_{12}^4 = N - c_{i(j)} - c_{j(i)} + d_{ij},$$
$$k_{13}^4 = c_{ij(k)} - c_{j(ik)} - c_{i(k)} + d_{ik}, \quad k_{23}^4 = c_{ij(k)} - c_{i(jk)} - c_{j(k)} + d_{jk},$$
$$k_{123}^4 = c_{ij(k)} - c_{i(jk)} - c_{j(ik)} + d_{ijk}, \quad k_0^4 = s_{ab} - a - b + 1,$$
$$k_1^5 = c_{k(i)} - d_i, \quad k_2^5 = c_{k(j)} - d_j, \quad k_3^5 = N - d_k, \quad k_{12}^5 = c_{k(ij)} - d_{ij},$$
$$k_{13}^5 = c_{k(i)} - d_{ik}, \quad k_{23}^5 = c_{k(j)} - d_{jk}, \quad k_{123}^5 = c_{k(ij)} - d_{ijk}, \quad k_0^5 = c - 1,$$
$$k_1^6 = c_{j(i)} - d_i, \quad k_2^6 = N - d_j, \quad k_3^6 = c_{j(k)} - d_k, \quad k_{12}^6 = c_{j(i)} - d_{ij},$$
$$k_{13}^6 = c_{j(ik)} - d_{ik}, \quad k_{23}^6 = c_{j(k)} - d_{jk}, \quad k_{123}^6 = c_{j(ik)} - d_{ijk}, \quad k_0^6 = b - 1,$$

$$k_1^7 = N - d_i, \quad k_2^7 = c_{i(j)} - d_j, \quad k_3^7 = c_{i(k)} - d_k, \quad k_{12}^7 = c_{i(j)} - d_{ij},$$
$$k_{13}^7 = c_{i(k)} - d_{ik}, \quad k_{23}^7 = c_{i(jk)} - d_{jk}, \quad k_{123}^7 = c_{i(jk)} - d_{ijk}, \quad k_0^7 = a - 1.$$

The expected mean squares are obtained by dividing the expected sums of squares by their corresponding degrees of freedom.

14.4 ESTIMATION OF VARIANCE COMPONENTS

In this section, we outline briefly some methods of estimation of variance components.

14.4.1 ANALYSIS OF VARIANCE ESTIMATORS

The analysis of variance estimators of variance components can be obtained by equating mean squares or equivalently the sums of squares in the analysis of variance Table 14.1 to their respective expected values expressed as linear combinations of the variance components. The resulting equations are then solved for the variance components to produce the required estimates.

Denote the estimators as $\hat{\sigma}_\alpha^2$, $\hat{\sigma}_\beta^2$, $\hat{\sigma}_\gamma^2$, $\hat{\sigma}_{\alpha\beta}^2$, $\hat{\sigma}_{\alpha\gamma}^2$, $\hat{\sigma}_{\beta\gamma}^2$, $\hat{\sigma}_{\alpha\beta\gamma}^2$, and $\hat{\sigma}_e^2$, and define

$$\begin{aligned}
\hat{\boldsymbol{\sigma}}'^2 &= (\hat{\sigma}_\alpha^2, \hat{\sigma}_\beta^2, \hat{\sigma}_\gamma^2, \hat{\sigma}_{\alpha\beta}^2, \hat{\sigma}_{\alpha\gamma}^2, \hat{\sigma}_{\beta\gamma}^2, \hat{\sigma}_{\alpha\beta\gamma}^2), \\
\boldsymbol{S}' &= (\mathrm{SS}_A, \mathrm{SS}_B, \mathrm{SS}_C, \mathrm{SS}_{AB}, \mathrm{SS}_{AC}, \mathrm{SS}_{BC}, \mathrm{SS}_{ABC}), \\
\boldsymbol{f} &= \begin{bmatrix} a-1 \\ b-1 \\ c-1 \\ s_{ab} - a - b + 1 \\ s_{ac} - a - c + 1 \\ s_{bc} - b - c + 1 \\ s - s_{ab} - s_{ac} - s_{bc} + a + b + c - 1 \end{bmatrix},
\end{aligned}$$

and $\boldsymbol{P}$ as the matrix of coefficients of variance components (other than σ_e^2) in the expected sums of squares given below:

$$\boldsymbol{P} = \begin{bmatrix}
k_1^7 & k_2^7 & k_3^7 & k_{12}^7 & k_{13}^7 & k_{23}^7 & k_{123}^7 \\
k_1^6 & k_2^6 & k_3^6 & k_{12}^6 & k_{13}^6 & k_{23}^6 & k_{123}^6 \\
k_1^5 & k_2^5 & k_3^5 & k_{12}^5 & k_{13}^5 & k_{23}^5 & k_{123}^5 \\
k_1^4 & k_2^4 & k_3^4 & k_{12}^4 & k_{13}^4 & k_{23}^4 & k_{123}^4 \\
k_1^3 & k_2^3 & k_3^3 & k_{12}^3 & k_{13}^3 & k_{23}^3 & k_{123}^3 \\
k_1^2 & k_2^2 & k_3^2 & k_{12}^2 & k_{13}^2 & k_{23}^2 & k_{123}^2 \\
k_1^1 & k_2^1 & k_3^1 & k_{12}^1 & k_{13}^1 & k_{23}^1 & k_{123}^1
\end{bmatrix}.$$

Then the equations giving the desired estimates can be written as

$$\begin{bmatrix} \hat{\boldsymbol{\sigma}}^2 \\ \hat{\sigma}_e^2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{P} & \boldsymbol{f} \\ 0 & N-s \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{S} \\ \mathrm{SS}_E \end{bmatrix},$$

which yields

$$\hat{\sigma}_e^2 = \frac{\mathrm{SS}_E}{N-s} \tag{14.4.1}$$

and

$$\hat{\boldsymbol{\sigma}}^2 = \boldsymbol{P}^{-1}[\boldsymbol{S} - \hat{\sigma}_e^2 \boldsymbol{f}]. \tag{14.4.2}$$

14.4.2 SYMMETRIC SUMS ESTIMATORS

For symmetric sums estimators based on products of the observations, we have

$$E(y_{ijk\ell}y_{i'j'k'\ell'}) = \begin{cases} \mu^2, \quad i \neq i', \ j \neq j', \ k \neq k', \\ \mu^2 + \sigma_\alpha^2, \quad i = i', \ j \neq j', \ k \neq k', \\ \mu^2 + \sigma_\beta^2, \quad i \neq i', \ j = j', \ k \neq k', \\ \mu^2 + \sigma_\gamma^2, \quad i \neq i', \ j \neq j', \ k = k', \\ \mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_{\alpha\beta}^2, \quad i = i', \ j = j', \ k \neq k', \\ \mu^2 + \sigma_\alpha^2 + \sigma_\gamma^2 + \sigma_{\alpha\gamma}^2, \quad i = i', \ j \neq j', \ k = k', \\ \mu^2 + \sigma_\beta^2 + \sigma_\gamma^2 + \sigma_{\beta\gamma}^2, \quad i \neq i', \ j = j', \ k = k', \\ \mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 + \sigma_{\alpha\beta}^2 + \sigma_{\alpha\gamma}^2 + \sigma_{\beta\gamma}^2 + \sigma_{\alpha\beta\gamma}^2, \\ \qquad i = i', \ j = j', \ k = k', \ \ell \neq \ell', \\ \mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 + \sigma_{\alpha\beta}^2 + \sigma_{\alpha\gamma}^2 + \sigma_{\beta\gamma}^2 + \sigma_{\alpha\beta\gamma}^2 + \sigma_e^2, \\ \qquad i = i', \ j = j', \ k = k', \ \ell = \ell', \end{cases} \tag{14.4.3}$$

where $i, i' = 1, 2, \ldots, a$; $j, j' = 1, 2, \ldots, b$; $k, k' = 1, 2, \ldots, c$; $\ell = 1, 2, \ldots, n_{ijk}$; $\ell' = 1, 2, \ldots, n_{i'j'k'}$. Now, the normalized symmetric sums of the terms in (14.4.3) are

$$g_m = \frac{y_{....}^2 - \sum_{i=1}^{a} y_{i...}^2 - \sum_{j=1}^{b} y_{.j..}^2 - \sum_{k=1}^{c} y_{..k.}^2}{N^2 - k_1 - k_2 - k_3 + k_{12} + k_{13} + k_{23} - k_{123}}$$
$$+ \frac{\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij..}^2 + \sum_{i=1}^{a}\sum_{k=1}^{c} y_{i.k.}^2 + \sum_{j=1}^{b}\sum_{k=1}^{c} y_{.jk.}^2 - \sum_{i=1}^{a}\sum_{j=i}^{b}\sum_{k=1}^{c} y_{ijk.}^2}{N^2 - k_1 - k_2 - k_3 + k_{12} + k_{13} + k_{23} - k_{123}},$$

$$g_A = \frac{\sum_{i=1}^{a} y_{i...}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij..}^2 - \sum_{i=1}^{a}\sum_{k=1}^{c} y_{i.k.}^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk.}^2}{k_1 - k_{12} - k_{13} + k_{123}},$$

$$g_B = \frac{\sum_{j=1}^{b} y_{.j..}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij..}^2 - \sum_{j=1}^{b}\sum_{k=1}^{c} y_{.jk.}^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk.}^2}{k_2 - k_{12} - k_{23} + k_{123}},$$

$$g_C = \frac{\sum_{k=1}^{c} y_{..k.}^2 - \sum_{i=1}^{a}\sum_{k=1}^{c} y_{i.k.}^2 - \sum_{j=1}^{b}\sum_{k=1}^{c} y_{.jk.}^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk.}^2}{k_3 - k_{13} - k_{23} + k_{123}},$$

$$g_{AB} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij..}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk.}^2}{k_{12} - k_{123}},$$

$$g_{AC} = \frac{\sum_{i=1}^{a}\sum_{k=1}^{c} y_{i.k.}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk.}^2}{k_{13} - k_{123}},$$

$$g_{BC} = \frac{\sum_{j=1}^{b}\sum_{k=1}^{c} y_{.jk.}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk.}^2}{k_{23} - k_{123}},$$

$$g_{ABC} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk.}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk\ell}^2}{k_{123} - N},$$

and

$$g_E = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk\ell}^2}{N},$$

where

$$k_1 = \sum_{i=1}^{a} n_{i..}^2, \qquad k_2 = \sum_{j=1}^{b} n_{.j.}^2, \qquad k_3 = \sum_{k=1}^{c} n_{..k}^2,$$

$$k_{12} = \sum_{i=1}^{a}\sum_{j=1}^{b} n_{ij.}^2, \qquad k_{13} = \sum_{i=1}^{a}\sum_{k=1}^{c} n_{i.k}^2, \qquad k_{23} = \sum_{j=1}^{b}\sum_{k=1}^{c} n_{.jk}^2,$$

$$k_{123} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} n_{ijk}^2, \qquad \text{and} \qquad N = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} n_{ijk}.$$

By equating g_m, g_A, g_B, g_C, g_{AB}, g_{AC}, g_{BC}, g_{ABC}, and g_E to their respective expected values and solving the resulting equations, we obtain the estimators of variance components as (Koch, 1967)

$$\begin{aligned}
\hat{\sigma}_\alpha^2 &= g_A - g_m,\\
\hat{\sigma}_\beta^2 &= g_B - g_m,\\
\hat{\sigma}_\gamma^2 &= g_C - g_m,\\
\hat{\sigma}_{\alpha\beta}^2 &= g_{AB} - g_A - g_B + g_m,\\
\hat{\sigma}_{\alpha\gamma}^2 &= g_{AC} - g_A - g_C + g_m,\\
\hat{\sigma}_{\beta\gamma}^2 &= g_{BC} - g_B - g_C + g_m,\\
\hat{\sigma}_{\alpha\beta\gamma}^2 &= g_{ABC} - g_{AB} - g_{AC} - g_{BC} + g_A + g_B + g_C - g_m,
\end{aligned} \tag{14.4.4}$$

and

$$\hat{\sigma}_e^2 = g_E - g_{ABC}.$$

The estimators in (14.4.4) are, by construction, unbiased, and they reduce to the analysis of variance estimators in the case of balanced data. However, they are not translation invariant; i.e., they may change in values if the same constant is added to all the observations and their variances are functions of μ. This drawback is overcome by using the symmetric sums of squares of differences rather than the products (Koch, 1968) (see Exercise 14.2).

14.4.3 OTHER ESTIMATORS

The ML, REML, MINQUE, and MIVQUE estimators can be developed as special cases of the results for the general case considered in Chapter 10 and their special formulation for this model are not amenable to any simple algebraic expressions. With the advent of the high-speed digital computer, the general results on these estimators involving matrix operations can be handled with great speed and accuracy and their explicit algebraic evaluation for this model seems to be rather unnecessary. In addition, some commonly used statistical software packages, such as SAS®, SPSS®, and BMDP®, have special routines to compute these estimates rather conveniently simply by specifying the model in question.

14.5 VARIANCES OF ESTIMATORS

Under the assumption of normality, it can be shown that

$$\frac{\mathrm{SS}_E}{\sigma_e^2} \sim \chi^2[N-s],$$

so that from (14.4.1), we have

$$\mathrm{Var}(\hat{\sigma}_e^2) = \frac{2\sigma_e^4}{N-s}. \tag{14.5.1}$$

Further, SS_E has zero covariance with every element of $\boldsymbol{S}$, i.e., with every other sum of squares term. Therefore, from (14.4.1) and (14.4.2), we have

$$\mathrm{Cov}(\hat{\boldsymbol{\sigma}}^2, \hat{\boldsymbol{\sigma}}_e^2) = -\boldsymbol{P}^{-1}\boldsymbol{f}\,\mathrm{Var}(\hat{\sigma}_e^2) \tag{14.5.2}$$

and

$$\mathrm{Var}(\hat{\boldsymbol{\sigma}}^2) = \boldsymbol{P}^{-1}[\mathrm{Var}(\boldsymbol{S}) + \mathrm{Var}(\hat{\sigma}_e^2)\boldsymbol{f}\boldsymbol{f}']\boldsymbol{P}'^{-1}. \tag{14.5.3}$$

If we define the vector $\boldsymbol{t}$ and the matrix $\boldsymbol{G}$ as

$$\boldsymbol{t}' = (T_A, T_B, T_C, T_{AB}, T_{AC}, T_{BC}, T_{ABC}, T_\mu),$$
$$\boldsymbol{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ -1 & -1 & 0 & 1 & 0 & 0 & 0 & 1 \\ -1 & 0 & -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 \end{bmatrix},$$

then $\boldsymbol{S}$ can be expressed as

$$\boldsymbol{S} = \boldsymbol{G}\boldsymbol{t}. \tag{14.5.4}$$

From (14.5.4) we note that

$$\mathrm{Var}(\boldsymbol{S}) = \boldsymbol{G}\,\mathrm{Var}(\boldsymbol{t})\boldsymbol{G}', \tag{14.5.5}$$

so that, on substituting (14.5.5) into (14.5.3), we obtain

$$\mathrm{Var}(\hat{\boldsymbol{\sigma}}^2) = \boldsymbol{P}^{-1}[\boldsymbol{G}\,\mathrm{Var}(\boldsymbol{t})\boldsymbol{G}' + \mathrm{Var}(\hat{\sigma}_e^2)\boldsymbol{f}\boldsymbol{f}']\boldsymbol{P}^{-1\prime}.$$

Thus, in order to compute the variance-covariance matrix of $\hat{\boldsymbol{\sigma}}^2$, it suffices to know only $\mathrm{Var}(\boldsymbol{t})$, the variance-covariance matrix of $\boldsymbol{t}$. Now, $\mathrm{Var}(\boldsymbol{t})$ is an 8×8 matrix with 36 distinct elements. Each element is a linear combination of the 36 possible squares and products of the eight variance components in $\boldsymbol{\sigma}^2$, viz., $\sigma_\alpha^4, \sigma_\beta^4, \ldots, \sigma_\alpha^2\sigma_\beta^2, \ldots, \ldots$. The 36×36 matrix of coefficients of these products has been prepared by Blischke (1968) as an unpublished appendix to the paper and is reproduced in Searle (1971, Table 11.6). It is also reprinted in this chapter as an appendix with the kind permission of Dr. Blischke.

The three factors A, B, C in the appendix table are denoted by the numbers 1, 2, 3. Thus T_1 and T_{13} stand for T_A and T_{AC}, respectively, and σ_1^2 and σ_{13}^2 stand for σ_α^2 and $\sigma_{\alpha\gamma}^2$, respectively. The entries of the table are given in terms of the n_{ijh} employing the customary dot notation and the additional notation

$$w_{ijhstu} = n_{ijh}n_{stu},$$

where an asterisk in the fourth, fifth, or sixth subscript indicates that the subscript is equated to the first, second, or third subscript, respectively, prior to summation. Thus, for example,

$$w_{ij.st^*} = \sum_h n_{ijh}n_{sth},$$

$$w_{i..i.^*} = \sum_{j,h,t} n_{ijh}n_{ith} = \sum_h n_{i.h}^2,$$

$$\sum_i \frac{w_{i..i.^*}^2}{w_{i..i..}} = \sum_i \frac{\left(\sum_h w_{i.hi.h}\right)^2}{w_{i..i..}} = \sum_i \frac{1}{n_{i..}^2}\left(\sum_h n_{i.h}^2\right)^2.$$

The entry in the ith row and jth column is denoted by A_{ij}. Unless otherwise indicated, the summation is understood to be extended to all the subscripts.

14.6 GENERAL r-WAY CROSSED CLASSIFICATION

In this section, we shall briefly indicate the analysis of variance model for an r-way crossed classification. Let $y_{i_1,i_2...i_rk}$ be the kth observation at the treatment combination comprising the i_1th level of factor 1, i_2th level of factor 2,..., and i_rth level of factor R, where $i_j = 1, 2, \ldots, a_j$ for $j = 1, 2, \ldots, r$, and $k = 1, 2, \ldots, n_{i_1i_2...i_r}$. Then the random effects model for the general r-way crossed classification can be written as

$$\begin{aligned} y_{i_1i_2...i_rk} &= \mu + \alpha(i_1, 1) + \alpha(i_2, 2) + \cdots + \alpha(i_r, r) + \alpha(i_1, i_2; 1, 2) \\ &\quad + \cdots + \alpha(i_{r-1}, i_r; r-1, r) \\ &\quad + \cdots + \cdots + \alpha(i_1, \ldots, i_r; 1, \ldots, r) + e_{i_1i_2...i_rk}, \end{aligned} \tag{14.6.1}$$

where μ = overall or general mean; $\alpha(i_v, v)$ = the main effect corresponding to the i_vth level of factor $v (v = 1, 2, \ldots, r)$; $\alpha(i_{v_1}, \ldots, i_{v_j}; v_1, \ldots, v_j)$ = the $(j-1)$th-order interaction effect corresponding to the combination of the i_{v_1}th level of factor $v_1, \ldots,$ and the i_{v_j}th level of factor v_j, and $e_{i_1i_2...i_rk}$ = the experimental error or residual effect. It is further assumed that $\alpha(i_v, v)$, $\alpha(i_{v_1}, \ldots, i_{v_j}; v_1, \ldots, v_j)$, and the $e_{i_1i_2...i_rk}$ are mutually and completely uncorrelated random variables with zero means and variances σ_v^2, $\sigma_{v_1...v_j}^2$, and σ_e^2, respectively. Note that there are 2^r variance components, viz, $\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2, \sigma_{12}^2$, $\ldots, \ldots,$ and $\sigma_{12...r}^2$. Because of the generality of the problem, the algebraic notation required for analytic results including mean squares, expectations, variances, and covariances of moment-type estimators become quite complex, and it becomes extremely tedious to work out the expected values of individual mean squares. The principles, however, remain the same involving straightforward algebra. For further discussion concerning the analysis of this model the reader is referred to Blischke (1968). In the following, we outline briefly the algebraic expressions for the symmetric sums estimators for the model in (14.6.1) (Koch, 1967).

For symmetric sums estimators based on products of observations, we have

$$E(y_{i_1 i_2 \ldots i_r k} y_{i'_1 i'_2 \ldots i'_r k'}) = \begin{cases} \mu^2, & i_1 \neq i'_1, i_2 \neq i'_2, \ldots, i_r \neq i'_r, \\ \mu^2 + \sigma_1^2, & i_1 = i'_1, i_2 \neq i'_2, \ldots, i_r \neq i'_r, \\ \mu^2 + \sigma_2^2, & i_1 \neq i'_1, i_2 = i'_2, \ldots, i_r \neq i'_r, \\ \mu^2 + \sigma_1^2 + \sigma_2^2 + \sigma_{12}^2, & i_1 = i'_1, i_2 = i'_2, i_3 \neq i'_3, \ldots, i_r \neq i'_r, \\ \vdots & \vdots \qquad \vdots \\ \mu^2 + \sigma_1^2 + \sigma_2^2 + \sigma_{12}^2 + \cdots + \sigma_{12\ldots r}^2, & \\ \qquad i_1 = i'_1, i_2 = i'_2, \ldots, i_r = i'_r, k \neq k', & \\ \mu^2 + \sigma_1^2 + \sigma_2^2 + \sigma_{12}^2 + \cdots + \sigma_{12\ldots r}^2 + \sigma_e^2, & \\ \qquad i_1 = i'_1, i_2 = i'_2, \ldots, i_r = i'_r, k = k', & \end{cases} \tag{14.6.2}$$

where $i_1, i'_1 = 1, 2, \ldots, a_1$; $i_2, i'_2 = 1, 2, \ldots, a_2$; $\ldots, i_r, i'_r = 1, 2, \ldots, a_r$; $k = 1, 2, \ldots, n_{i_1 i_2 \ldots i_r}$; $k' = 1, 2, \ldots, n_{i'_1 i'_2 \ldots i'_r}$. Now, the normalized symmetric sums of the terms in (14.6.2) are

$$g_m = \frac{y_{\ldots\ldots}^2 - \sum_{i_1=1}^{a_1} y_{i_1\ldots\ldots}^2 - \cdots = \sum_{i_r=1}^{a_r} y_{\ldots\ldots i_r .}^2 + \sum_{i_1=1}^{a_1}\sum_{i_2=1}^{a_2} y_{i_1 i_2 \ldots\ldots}^2 + \cdots + (-1)^r \sum_{i_1=1}^{a_1}\sum_{i_2=1}^{a_2}\cdots\sum_{i_r=1}^{a_r} y_{i_1 i_2 \ldots i_r .}^2}{N^2 - k_1 - \cdots - k_r + k_{12} + \cdots + (-1)^r k_{12\ldots r}},$$

$$g_1 = \frac{\sum_{i_1=1}^{a_1} y_{i_1\ldots\ldots}^2 - \sum_{i_1=1}^{a_1}\sum_{i_2=1}^{a_2} y_{i_1 i_2 \ldots\ldots} - \cdots - \sum_{i_1=1}^{a_1}\sum_{i_r=1}^{a_r} y_{i_1 \ldots\ldots i_r .} + \cdots + (-1)^{r-1} \sum_{i_1=1}^{a_1}\sum_{i_2=1}^{a_2}\cdots\sum_{i_r=1}^{a_r} y_{i_1 i_2 \ldots i_r .}^2}{k_1 - k_{12} - \cdots - k_{1r} + \cdots + (-1)^{r-1} k_{12\ldots r}},$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$g_{12\ldots r} = \frac{\sum_{i_1=1}^{a_1}\sum_{i_2=1}^{a_2}\cdots\sum_{i_r=1}^{a_r} y_{i_1 i_2 \ldots i_r .}^2 - \sum_{i_1=1}^{a_1}\sum_{i_2=1}^{a_2}\cdots\sum_{i_r=1}^{a_r}\sum_{k=1}^{n_{i_1 i_2 \ldots i_r}} y_{i_1 i_2 \ldots i_r k}^2}{k_{12\ldots r} - N},$$

and

$$g_E = \frac{\sum_{i_1=1}^{a_1}\sum_{i_2=1}^{a_2}\cdots\sum_{i_r=1}^{a_r}\sum_{k=1}^{n_{i_1 i_2 \ldots i_r}} y_{i_1 i_2 \ldots i_r k}^2}{N},$$

where

$$k_1 = \sum_{i_1=1}^{a_1} n^2_{i_1\ldots\ldots}, \qquad k_2 = \sum_{i_2=1}^{a_2} n^2_{.i_2\ldots\ldots}, \ldots, \qquad k_r = \sum_{i_r=1}^{a_r} n^2_{\ldots\ldots i_r},$$

$$k_{12} = \sum_{i_1=1}^{a_1}\sum_{i_2=1}^{a_2} n^2_{i_1 i_2\ldots\ldots}, \qquad k_{13} = \sum_{i_1=1}^{a_1}\sum_{i_3=1}^{a_3} n^2_{i_1 . i_3\ldots\ldots}, \ldots, \qquad k_{1r} = \sum_{i_1=1}^{a_1}\sum_{i_r=1}^{a_r} n^2_{i_1\ldots\ldots i_r},$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$k_{12\ldots r} = \sum_{i_1=1}^{a_1}\sum_{i_2=1}^{a_2}\cdots\sum_{i_r=1}^{a_r} n^2_{i_1 i_2\ldots i_r}, \quad \text{and} \quad N = \sum_{i_1=1}^{a_1}\sum_{i_2=1}^{a_2}\cdots\sum_{i_r=1}^{a_r} n_{i_1 i_2\ldots i_r}.$$

By equating $g_m, g_1, \ldots, g_{12\ldots r}$, and g_E to their respective expected values and solving the resulting equations, we obtain the estimators of the variance components as (Koch, 1967)

$$\begin{aligned}
\hat{\sigma}_1^2 &= g_1 - g_m, \\
\hat{\sigma}_2^2 &= g_2 - g_m, \\
\hat{\sigma}_{12}^2 &= g_{12} - g_1 - g_2 + g_m, \\
&\vdots \quad \vdots \quad \vdots \\
\hat{\sigma}_{12\ldots r}^2 &= g_{12\ldots r} - g_{12\ldots r-1} - \cdots + (-1)^{r-1} g_1 \\
&\quad + \cdots + (-1)^{r-1} g_r + (-1)^r g_m,
\end{aligned} \tag{14.6.3}$$

and

$$\hat{\sigma}_e^2 = g_E - g_{12\ldots r}.$$

The estimators in (14.6.3), by construction, are unbiased, and they reduce to the analysis of variance estimators in the case of balanced data. However, they are not translation invariant; i.e., they may change in values if the same constant is added to all the observations and their variances are functions of μ. This drawback is overcome by using the symmetric sums of squares of differences rather than the products.

Symmetric sums estimators based on the expected values of the squares of differences of the observations can be developed analogously (Koch, 1968); however, no algebraic expressions will be given for these because of the notational complexity.

14.7 A NUMERICAL EXAMPLE

Consider a factorial experiment involving machines and operators. Three machines were randomly selected from a large number of machines available for

TABLE 14.2 Production output from an industrial experiment.

Day	Machine	Operator			
		1	2	3	4
		201.7	211.1	201.1	199.3
	1	207.9	206.9	203.6	200.6
		201.7		198.7	
				199.8	
				199.8	
		193.4	204.3	197.5	209.1
		193.1	201.6		210.0
		198.3	201.3		209.0
1	2	195.0	198.8		
		194.1			
		198.0	193.6	187.9	195.5
		197.2	187.0		194.1
		101.9	186.3		192.1
	3		190.5		
			189.8		
		183.7	197.6	189.0	184.7
			201.2	194.6	183.8
	1		203.0	190.0	188.8
			199.0	193.9	189.1
			197.0		183.6
		173.8	172.8	179.5	183.2
2	2		173.2	181.5	181.0
			173.5	184.9	183.0
			172.5	178.8	183.7
				182.3	179.7
		183.6	182.0	197.4	178.8
		183.2	185.1	193.5	179.7
	3	183.1	183.3	193.1	178.3
			184.5	194.9	179.9
			182.6		181.6
		208.7	197.2	202.2	210.2
		209.9	199.6		
	1	208.3	198.4		
		205.3	194.6		
		206.5			
		198.2	193.0	189.1	190.7
3	2	199.3		192.5	187.1
		192.6		191.4	192.3
		200.3		192.6	188.5
				186.8	
		188.2	198.9	197.0	202.5
	3		197.0	190.8	201.2
					203.2
					199.5

the experiment, and four operators were chosen at random from a pool of operators. The purpose of the experiment is to investigate the variation in the output due to different machines and operators. The whole experiment was replicated by repeating it on three days. Each operator was to work five times with each machine and on each day. However, because of logistics and other scheduling problems, it was not possible to accomplish all the tasks. The relevant data on production output are given in Table 14.2. We will use the three-way unbalanced crossed model in (14.1.1) to analyze the data in Table 14.2. Here, $a = 3$, $b = 3$, $c = 4$; $i = 1, 2, 3$ refer to the days; $b = 1, 2, 3$ refer to the machines; and $k = 1, 2, 3, 4$ refer to the operators. Further, σ_α^2, σ_β^2, and σ_γ^2 designate the variance components due to day, machine, and operator as factors; $\sigma_{\alpha\beta}^2$, $\sigma_{\alpha\gamma}^2$, $\sigma_{\beta\gamma}^2$ are two-factor interaction variance components; $\sigma_{\alpha\beta\gamma}^2$ is the three-factor interaction variance component; and σ_e^2 denotes the error variance component. The calculations leading to the conventional analysis of variance based on Henderson's Method I were performed using the SAS®GLM procedure and the results are summarized in Table 14.3.

We now illustrate the calculations of point estimates of the variance components σ_e^2, $\sigma_{\alpha\beta\gamma}^2$, $\sigma_{\beta\gamma}^2$, $\sigma_{\alpha\gamma}^2$, $\sigma_{\alpha\beta}^2$, σ_γ^2, σ_β^2, and σ_α^2. The analysis of variance (ANOVA) estimates based on Henderson's Method I are obtained as the solution to the following system of equations:

$$
\begin{aligned}
&\sigma_e^2 + 3.928\sigma_{\alpha\beta\gamma}^2 + 0.851\sigma_{\beta\gamma}^2 + 10.307\sigma_{\alpha\gamma}^2 + 13.136\sigma_{\alpha\beta}^2 + 0.799\sigma_\gamma^2 \\
&\quad + 0.178\sigma_\beta^2 + 38.941\sigma_\alpha^2 = 1,847.782, \\
&\sigma_e^2 + 3.990\sigma_{\alpha\beta\gamma}^2 + 10.128\sigma_{\beta\gamma}^2 + 0.634\sigma_{\alpha\gamma}^2 + 13.480\sigma_{\alpha\beta}^2 + 0.340\sigma_\gamma^2 \\
&\quad + 39.288\sigma_\beta^2 + 0.175\sigma_\alpha^2 = 1,281.247, \\
&\sigma_e^2 + 3.984\sigma_{\alpha\beta\gamma}^2 + 10.050\sigma_{\beta\gamma}^2 + 10.607\sigma_{\alpha\gamma}^2 + 0.814\sigma_{\alpha\beta}^2 + 29.424\sigma_\gamma^2 \\
&\quad + 0.293\sigma_\beta^2 + 0.805\sigma_\alpha^2 = 7.997, \\
&\sigma_e^2 + 3.857\sigma_{\alpha\beta\gamma}^2 + 0.788\sigma_{\beta\gamma}^2 + 0.668\sigma_{\alpha\gamma}^2 + 12.814\sigma_{\alpha\beta}^2 + 0.815\sigma_\gamma^2 \\
&\quad - 0.089\sigma_\beta^2 - 0.088\sigma_\alpha^2 = 479.141, \\
&\sigma_e^2 + 3.620\sigma_{\alpha\beta\gamma}^2 + 0.586\sigma_{\beta\gamma}^2 + 9.142\sigma_{\alpha\gamma}^2 + 0.549\sigma_{\alpha\beta}^2 - 0.266\sigma_\gamma^2 \\
&\quad + 0.810\sigma_\beta^2 - 0.403\sigma_\alpha^2 = 236.522, \\
&\sigma_e^2 + 3.841\sigma_{\alpha\beta\gamma}^2 + 9.573\sigma_{\beta\gamma}^2 + 0.530\sigma_{\alpha\gamma}^2 + 0.678\sigma_{\alpha\beta}^2 - 0.113\sigma_\gamma^2 \\
&\quad - 0.146\sigma_\beta^2 - 0.683\sigma_\alpha^2 = 345.258, \\
&\sigma_e^2 + 2.162\sigma_{\alpha\beta\gamma}^2 - 0.699\sigma_{\beta\gamma}^2 - 0.594\sigma_{\alpha\gamma}^2 - 0.819\sigma_{\alpha\beta}^2 - 0.272\sigma_\gamma^2 \\
&\quad - 0.405\sigma_\beta^2 - 0.342\sigma_\alpha^2 = 31.801, \\
&\sigma_e^2 = 79.247.
\end{aligned}
$$

TABLE 14.3 Analysis of variance for the production output data of Table 14.2.

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Day	2	3,695.564	1,847.782	$\sigma_e^2 + 3.928\sigma_{\alpha\beta\gamma}^2 + 0.851\sigma_{\beta\gamma}^2 + 10.307\sigma_{\alpha\gamma}^2 + 13.136\sigma_{\alpha\beta}^2 + 0.799\sigma_\gamma^2 + 0.178\sigma_\beta^2 + 38.941\sigma_\alpha^2$
Machine	2	2,562.494	1,281.247	$\sigma_e^2 + 3.990\sigma_{\alpha\beta\gamma}^2 + 10.128\sigma_{\beta\gamma}^2 + 0.634\sigma_{\alpha\gamma}^2 + 13.480\sigma_{\alpha\beta}^2 + 0.340\sigma_\gamma^2 + 39.288\sigma_\beta^2 + 0.175\sigma_\alpha^2$
Operator	3	23.990	7.997	$\sigma_e^2 + 3.984\sigma_{\alpha\beta\gamma}^2 + 10.050\sigma_{\beta\gamma}^2 + 10.607\sigma_{\alpha\gamma}^2 + 0.814\sigma_{\alpha\beta}^2 + 29.424\sigma_\gamma^2 + 0.293\sigma_\beta^2 + 0.805\sigma_\alpha^2$
Day × Machine	4	1,916.565	479.141	$\sigma_e^2 + 3.857\sigma_{\alpha\beta\gamma}^2 + 0.788\sigma_{\beta\gamma}^2 + 0.668\sigma_{\alpha\gamma}^2 + 12.814\sigma_{\alpha\beta}^2 + 0.815\sigma_\gamma^2 - 0.089\sigma_\beta^2 - 0.088\sigma_\alpha^2$
Day × Operator	6	1,419.129	236.522	$\sigma_e^2 + 3.620\sigma_{\alpha\beta\gamma}^2 + 0.586\sigma_{\beta\gamma}^2 + 9.142\sigma_{\alpha\gamma}^2 + 0.549\sigma_{\alpha\beta}^2 - 0.266\sigma_\gamma^2 + 0.810\sigma_\beta^2 - 0.403\sigma_\alpha^2$
Machine × Operator	6	2,071.548	345.258	$\sigma_e^2 + 3.841\sigma_{\alpha\beta\gamma}^2 + 9.573\sigma_{\beta\gamma}^2 + 0.530\sigma_{\alpha\gamma}^2 + 0.678\sigma_{\alpha\beta}^2 - 0.113\sigma_\gamma^2 - 0.146\sigma_\beta^2 - 0.683\sigma_\alpha^2$
Day × Machine × Operator	12	381.610	31.801	$\sigma_e^2 + 2.162\sigma_{\alpha\beta\gamma}^2 - 0.699\sigma_{\beta\gamma}^2 - 0.594\sigma_{\alpha\gamma}^2 - 0.819\sigma_{\alpha\beta}^2 - 0.272\sigma_\gamma^2 - 0.405\sigma_\beta^2 - 0.342\sigma_\alpha^2$
Error	82	6,498.250	79.247	σ_e^2
Total	117			

Therefore, the desired ANOVA estimates of the variance components are given by

$$\begin{bmatrix} \hat{\sigma}^2_{e,\text{ANOV}} \\ \hat{\sigma}^2_{\alpha\beta\gamma,\text{ANOV}} \\ \hat{\sigma}^2_{\beta\gamma,\text{ANOV}} \\ \hat{\sigma}^2_{\alpha\gamma,\text{ANOV}} \\ \hat{\sigma}^2_{\alpha\beta,\text{ANOV}} \\ \hat{\sigma}^2_{\gamma,\text{ANOV}} \\ \hat{\sigma}^2_{\beta,\text{ANOV}} \\ \hat{\sigma}^2_{\alpha,\text{ANOV}} \end{bmatrix} = \begin{bmatrix} 79.247 \\ 5.571 \\ 25.153 \\ 11.351 \\ 28.828 \\ -17.663 \\ 13.481 \\ 31.876 \end{bmatrix}.$$

We used SAS® VARCOMP, SPSS® VARCOMP, and BMDP® 3V procedures to estimate variance components using the ML, REML, MINQUE(0), and MINQUE(1) methods.[2] The relevant estimates using these software are given in Table 14.4. Note that all three software produce nearly the same results except for some minor discrepancy in rounding decimal places.

EXERCISES

1. Apply the method of "synthesis" to derive the expected mean squares given in Section 14.3.

2. Derive symmetric sums estimators based on squares of differences of the observations in the model in (14.1.1).

3. Show that the ANOVA estimators (14.4.1) and (14.4.2) reduce to the corresponding estimators (5.1.2) for balanced data.

4. Show that the symmetric sums estimators (14.4.4) reduce to the ANOVA estimators (5.1.2) for balanced data.

5. Consider a three-way crossed classification model that includes a fixed main effect, two random main effects, and a random interaction effect. Derive algebraic expressions for the ANOVA estimators of the variance components and variances of the resulting estimators (Blischke, 1966).

6. Refer to Exercise 5.15 and suppose that the observations (day 1, operator 1, machine 1) and (day 3, operator 3, machine 3) are missing due to mishaps. For the resulting three-way factorial design, respond to the following questions:

---

[2]The computations for ML and REML estimates were also carried out using SAS® PROC MIXED and some other programs to assess their relative accuracy and convergence rate. There did not seem to be any appreciable differences between the results from different software.

TABLE 14.4 ML, REML, MINQUE(0), and MINQUE(1) estimates of the variance components using SAS®, SPSS®, and BMDP® software.

Variance component	SAS®		
	ML	REML	MINQUE(0)
σ_e^2	77.274006	77.234658	82.581345
$\sigma_{\alpha\beta\gamma}^2$	18.436095	18.513179	6.378687
$\sigma_{\beta\gamma}^2$	0	0	12.230421
$\sigma_{\alpha\gamma}^2$	4.242237	4.255799	11.880401
$\sigma_{\alpha\beta}^2$	35.624683	31.567263	24.665982
σ_{γ}^2	0	0	−13.591424
σ_{β}^2	9.993819	18.818808	17.155330
σ_{α}^2	17.401254	31.726776	36.731359

Variance component	SPSS®			
	ML	REML	MINQUE(0)	MINQUE(1)
σ_e^2	77.273876	77.234425	82.581345	78.853462
$\sigma_{\alpha\beta\gamma}^2$	18.436303	18.513562	6.378687	12.236978
$\sigma_{\beta\gamma}^2$	0	0	12.230421	0.810241
$\sigma_{\alpha\gamma}^2$	4.242258	4.255839	11.880401	3.690244
$\sigma_{\alpha\beta}^2$	35.624855	31.567547	24.665982	31.835595
σ_{γ}^2	0	0	−13.591424	−0.710133
σ_{β}^2	9.993866	18.818971	17.155330	20.14874
σ_{α}^2	17.401337	31.727057	36.731359	33.285089

Variance component	BMDP®	
	ML	REML
σ_e^2	77.273876	77.234425
$\sigma_{\alpha\beta\gamma}^2$	18.436303	18.513562
$\sigma_{\beta\gamma}^2$	0	0
$\sigma_{\alpha\gamma}^2$	4.242258	4.255839
$\sigma_{\alpha\beta}^2$	35.624855	31.567547
σ_{γ}^2	0	0
σ_{β}^2	9.993866	18.818971
σ_{α}^2	17.401337	31.727057

SAS® VARCOMP does not compute MINQUE(1). BMDP® 3V does not compute MINQUE(0) and MINQUE(1).

(a) Describe the mathematical model and the assumptions involved.

(b) Analyze the data and report the conventional analysis of variance table based on Henderson's Method I.

(c) Test whether there are differences in the dry film thickness among different days.

(d) Test whether there are differences in the dry film thickness among different operators.

(e) Test whether there are differences in the dry film thickness among different machines.

(f) Test the significance of two-factor interaction effects.

(g) Find point estimates of each of the variance components of the model and the total variance using the ANOVA, ML, REML, MINQUE(0), and MINQUE(1) procedures and appropriate computing software.

7. Refer to Exercise 5.16 and suppose that the observations (day 1, soil 1, variety 1) and (day 2, soil 2, variety 2) are missing due to mishaps. For the resulting three-way factorial design, respond to the following questions:

   (a) Describe the mathematical model and the assumptions involved.

   (b) Analyze the data and report the conventional analysis of variance table based on Henderson's Method I.

   (c) Test whether there are differences in the residue weight among different days.

   (d) Test whether there are differences in the residue weight among different soil types.

   (e) Test whether there are differences in the residue weight among different varieties.

   (f) Test the significance of two-factor interaction effects.

   (g) Find point estimates of each of the variance components of the model and the total variance using the ANOVA, ML, REML, MINQUE(0), and MINQUE(1) procedures and appropriate computing software.

8. Refer to Exercise 5.17 and suppose that the observations (day 1, analyst 2, preparation 3) and (day 3, analyst 2, preparation 3) are missing due to mishaps. For the resulting three-way factorial design, respond to the following questions:

(a) Describe the mathematical model and the assumptions involved.

(b) Analyze the data and report the conventional analysis of variance table based on Henderson's Method I.

(c) Test whether there are differences in the weight among different days.

(d) Test whether there are differences in the weight among different analysts.

(e) Test whether there are differences in the weight among different preparations.

(f) Test the significance of two-factor interaction effects.

(g) Find point estimates of each of the variance components of the model and the total variance using the ANOVA, ML, REML, MINQUE(0), and MINQUE(1) procedures and appropriate computing software.

9. Refer to Exercise 5.18 and suppose that the observations (year 2, block 3, variety 1) and (year 4, block 2, variety 3) are missing due to mishaps. For the resulting three-way factorial design, respond to the following questions:

(a) Describe the mathematical model and the assumptions involved.

(b) Analyze the data and report the conventional analysis of variance table based on Henderson's Method I.

(c) Test whether there are differences in the yield among different days.

(d) Test whether there are differences in the yield among different blocks.

(e) Test whether there are differences in the yield among different varieties.

(f) Test the significance of two-factor interaction effects.

(g) Find point estimates of each of the variance components of the model and the total variance using the ANOVA, ML, REML, MINQUE(0), and MINQUE(1) procedures and appropriate computing software.

APPENDIX: COEFFICIENTS A_{ij} OF PRODUCTS OF VARIANCE COMPONENTS IN COVARIANCE MATRIX OF T

Row	Column (1/2)Cov Product	1 T_1	2 T_2	3 T_3	4 T_{12}	5 T_{13}
1	σ_1^4	$w_{...*..}$	$\sum \frac{w_{ij.st.}^2}{w_{.j..t.}}$	$\sum \frac{w_{i.hs.u}^2}{w_{..h..u}}$	$A_{1,1}$	$A_{1,1}$
2	σ_2^4	$\sum \frac{w_{ij.st.}^2}{w_{i..s..}}$	$w_{....*.}$	$\sum \frac{w_{.jh.tu}^2}{w_{..h..u}}$	$A_{2,2}$	$\sum \frac{w_{i.hs*u}^2}{w_{i.hs.u}}$
3	σ_3^4	$\sum \frac{w_{i.hs.u}^2}{w_{i..s..}}$	$\sum \frac{w_{.jh.tu}^2}{w_{.j..t.}}$	$w_{.....*}$	$\sum \frac{w_{ij.st*}^2}{w_{ij.st.}}$	$A_{3,3}$
4	σ_{12}^4	$\sum \frac{w_{i..i*.}^2}{w_{i..i..}}$	$\sum \frac{w_{.j.*j.}^2}{w_{.j..j.}}$	$\sum \frac{w_{ijhstu}^2}{w_{..h..u}}$	$A_{9,1}$	$\sum \frac{w_{i.hi*u}^2}{w_{i.hi.u}}$
5	σ_{13}^4	$\sum \frac{w_{i..i.*}^2}{w_{i..i..}}$	$\sum \frac{w_{ijhstu}^2}{w_{.j..t.}}$	$\sum \frac{w_{i.hs.u}^2}{w_{..h..h}}$	$\sum \frac{w_{ij.it*}^2}{w_{ij.it.}}$	$A_{10,1}$
6	σ_{23}^4	$\left(\sum \frac{w_{i..i**}}{n_{i..}}\right)^2$	$\sum \frac{w_{.j..j*}^2}{w_{.j..j.}}$	$\sum \frac{w_{.jh.tu}^2}{w_{..h..h}}$	$\sum \frac{w_{ij.sj*}^2}{w_{ij.sj.}}$	$\sum \frac{w_{i.hs*h}^2}{w_{i.hs.h}}$
7	σ_{123}^4	$\sum \frac{w_{i..i**}^2}{w_{i..i..}}$	$\sum \frac{w_{.j.*j*}^2}{w_{.j..j.}}$	$\sum \frac{w_{..h**h}^2}{w_{..h..h}}$	$\sum \frac{w_{ij.ij*}^2}{w_{ij.ij.}}$	$\sum \frac{w_{i.hi*h}^2}{w_{i.hi.h}}$
8	σ_0^4	I	J	H	m_{12}	m_{13}
9	$2\sigma_1^2\sigma_2^2$	$w_{...**.}$	$A_{1,1}$	$\sum_{i,j}\left(\sum_h \frac{w_{i.h.jh}}{n_{..h}}\right)^2$	$A_{9,1}$	$A_{9,1}$
10	$2\sigma_1^2\sigma_3^2$	$w_{...*.*}$	$\sum_{i,h}\left(\sum_j \frac{w_{ij..jh}}{n_{.j.}}\right)^2$	$A_{10,1}$	$A_{10,1}$	$A_{10,1}$
11	$2\sigma_1^2\sigma_{12}^2$	$A_{9,1}$	$A_{9,1}$	$\sum_{i,s,j}\left(\sum_h \frac{w_{i.hsjh}}{n_{..h}}\right)^2$	$A_{9,1}$	$A_{9,1}$
12	$2\sigma_1^2\sigma_{13}^2$	$A_{10,1}$	$\sum_{i,s,h}\left(\sum_j \frac{w_{ij.sjh}}{n_{.j.}}\right)^2$	$A_{5,3}$	$A_{10,1}$	$A_{10,1}$

Row	Column (1/2)Cov Product	6 T_{23}	7 T_{123}	8 T_f	9 T_1, T_2	10 T_1, T_3	11 T_1, T_{12}
1	σ_1^4	$\sum \frac{w_{.jh*tu}^2}{w_{.jh.tu}}$	$A_{1,1}$	$\left(\frac{A_{1,1}}{n}\right)^2$	$\sum \frac{w_{ij.i..}^2}{w_{.j.i..}}$	$\sum \frac{w_{i.hi..}^2}{w_{..hi..}}$	$A_{1,1}$
2	σ_2^4	$A_{2,2}$	$A_{2,2}$	$\left(\frac{A_{2,2}}{n}\right)^2$	$\sum \frac{w_{ij..j.}^2}{w_{i...j.}}$	$\sum \frac{w_{i.h.*.}^2}{w_{i....h}}$	$A_{2,9}$
3	σ_3^4	$A_{3,3}$	$A_{3,3}$	$\left(\frac{A_{3,3}}{n}\right)^2$	$\sum \frac{w_{ij...*}^2}{w_{i...j.}}$	$\sum \frac{w_{i.h..h}^2}{w_{i....h}}$	$\sum \frac{w_{ij.s.*}^2}{w_{ij.s..}}$
4	σ_{12}^4	$\sum \frac{w_{.jh*ju}^2}{w_{.jh.ju}}$	$A_{9,1}$	$\left(\frac{A_{9,1}}{n}\right)^2$	$\sum \frac{w_{ij.ij.}^2}{w_{i...j.}}$	$\sum \frac{w_{i.hi*.}^2}{w_{i....h}}$	$A_{17,9}$
5	σ_{13}^4	$\sum \frac{w_{.jh*th}^2}{w_{.jh.th}}$	$A_{10,1}$	$\left(\frac{A_{10,1}}{n}\right)^2$	$\sum \frac{w_{ij.i.*}^2}{w_{i...j.}}$	$\sum \frac{w_{i.hi.h}^2}{w_{i....h}}$	$\sum \frac{w_{ij.i.*}^2}{w_{ij.i..}}$
6	σ_{23}^4	$A_{16,2}$	$A_{16,2}$	$\left(\frac{A_{16,2}}{n}\right)^2$	$\sum \frac{w_{ij..j*}^2}{w_{i...j.}}$	$\sum \frac{w_{i.h.*h}^2}{w_{i....h}}$	$\sum \frac{w_{ij.sj*}^2}{w_{ij.s..}}$
7	σ_{123}^4	$\sum \frac{w_{.jh*jh}^2}{w_{.jh.jh}}$	$A_{13,1}$	$\left(\frac{A_{13,1}}{n}\right)^2$	$\sum \frac{w_{ij.ij*}^2}{w_{i...j.}}$	$\sum \frac{w_{i.hi*h}^2}{w_{i....h}}$	$\sum \frac{w_{ij.ij*}^2}{w_{ij.i..}}$
8	σ_0^4	m_{23}	m_{123}	1	$\sum \frac{w_{ij.ij.}}{w_{i...j.}}$	$\sum \frac{w_{i.hi.h}}{w_{i....h}}$	I
9	$2\sigma_1^2\sigma_2^2$	$A_{9,1}$	$A_{9,1}$	$\frac{A_{1,1}A_{2,2}}{n^2}$	$A_{9,1}$	$\sum \frac{w_{ijhi.h}n_{.j.}}{n_{..h}}$	$A_{9,1}$
10	$2\sigma_1^2\sigma_3^2$	$A_{10,1}$	$A_{10,1}$	$\frac{A_{1,1}A_{3,3}}{n^2}$	$\sum \frac{w_{ijhij.}n_{..h}}{n_{.j.}}$	$A_{10,1}$	$A_{10,1}$
11	$2\sigma_1^2\sigma_{12}^2$	$A_{4,6}$	$A_{9,1}$	$\frac{A_{1,1}A_{9,1}}{n^2}$	$\sum \frac{w_{ij.ij.}^2}{w_{.j.ij.}}$	$\sum \frac{w_{i.hijh}n_{ij.}}{n_{i.h}}$	$A_{9,1}$
12	$2\sigma_1^2\sigma_{13}^2$	$A_{5,6}$	$A_{10,1}$	$\frac{A_{1,1}A_{10,1}}{n^2}$	$\sum \frac{w_{ij.ijh}n_{i.h}}{n_{ij.}}$	$\sum \frac{w_{i.hi.h}^2}{w_{i.h..h}}$	$A_{10,1}$

Row	Column (1/2)Cov Product	12 T_1, T_{13}	13 T_1, T_{23}	14 T_1, T_{123}	15 T_1, T_f	16 T_2, T_3
1	σ_1^4	$A_{1,1}$	$\sum \frac{w^2_{ijhi..}}{w_{.jhi..}}$	$A_{1,1}$	$\sum \frac{w^2_{i..i..}}{w_{i.....}}$	$\sum \frac{w^2_{.jh*..}}{w_{.j...h}}$
2	σ_2^4	$\sum \frac{w^2_{i.hs*.}}{w_{i.hs..}}$	$\sum \frac{w^2_{ijh.j.}}{w_{.jhi..}}$	$A_{2,9}$	$\sum \frac{w^2_{i...*.}}{w_{i.....}}$	$\sum \frac{w^2_{ijh.j.}}{w_{i.h.j.}}$
3	σ_3^4	$A_{3,10}$	$\sum \frac{w^2_{ijh..h}}{w_{.jhi..}}$	$A_{3,10}$	$\sum \frac{w^2_{i....*}}{w_{i.....}}$	$\sum \frac{w^2_{.jh..h}}{w_{.j...h}}$
4	σ_{12}^4	$\sum \frac{w^2_{i.hi*.}}{w_{i.hs..}}$	$\sum \frac{w^2_{ijhij.}}{w_{i...jh}}$	$A_{17,9}$	$\sum \frac{w^2_{i..i*.}}{w_{i.....}}$	$\sum \frac{w^2_{.jh*j.}}{w_{.j...h}}$
5	σ_{13}^4	$A_{23,10}$	$\sum \frac{w^2_{ijhi.h}}{w_{i...jh}}$	$A_{23,10}$	$\sum \frac{w^2_{i..i.*}}{w_{i.....}}$	$\sum \frac{w^2_{.jh*.h}}{w_{.j...h}}$
6	σ_{23}^4	$\sum \frac{w^2_{i.hs*h}}{w_{i.hs..}}$	$\sum \frac{w^2_{ijh.jh}}{w_{i...jh}}$	$A_{6,13}$	$\sum \frac{w^2_{i...**}}{w_{i.....}}$	$\sum \frac{w^2_{.jh.jh}}{w_{.j...h}}$
7	σ_{123}^4	$\sum \frac{w^2_{i.hi*h}}{w_{i..i.h}}$	$\sum \frac{w^2_{ijhijh}}{w_{i...jh}}$	$A_{34,13}$	$\sum \frac{w^2_{i..i**}}{w_{i.....}}$	$\sum \frac{w^2_{.jh*jh}}{w_{.j...h}}$
8	σ_0^4	I	$\sum \frac{w_{ijhijh}}{w_{i...jh}}$	I	1	$\sum \frac{w_{.jh.jh}}{w_{.j...h}}$
9	$2\sigma_1^2\sigma_2^2$	$A_{9,1}$	$\sum \frac{w^2_{ijh.j.}}{w_{.jh.j.}}$	$A_{9,1}$	$\sum \frac{w_{.j.*..}n_{.j.}}{n}$	$\sum \frac{w_{i.h.*h}n_{i..}}{n_{..h}}$
10	$2\sigma_1^2\sigma_3^2$	$A_{10,1}$	$\sum \frac{w^2_{ijh..h}}{w_{.jh..h}}$	$A_{10,1}$	$\sum \frac{w_{..h*..}n_{..h}}{n}$	$\sum \frac{w_{ij..j*}n_{i..}}{n_{.j.}}$
11	$2\sigma_1^2\sigma_{12}^2$	$A_{9,1}$	$\sum \frac{w^2_{ijhij.}}{w_{.jhij.}}$	$A_{9,1}$	$\sum \frac{w^2_{ij.i..}}{w_{i.....}}$	$\sum \frac{w_{.jh*..}w_{.jh*j.}}{w_{.j...h}}$
12	$2\sigma_1^2\sigma_{13}^2$	$A_{10,1}$	$\sum \frac{w^2_{ijhi.h}}{w_{.jhi.h}}$	$A_{10,1}$	$\sum \frac{w^2_{i.hi..}}{w_{i.....}}$	$\sum \frac{w_{.jh*..}w_{.jh*.h}}{w_{.j...h}}$

Row	Column (1/2)Cov Product	17 T_2, T_{12}	18 T_2, T_{13}	19 T_2, T_{23}	20 T_2, T_{123}
1	σ_1^4	$A_{1,9}$	$\sum \frac{w^2_{ijhi..}}{w_{i.h.j.}}$	$\sum \frac{w^2_{.jh*t.}}{w_{.jh..u}}$	$A_{1,9}$
2	σ_2^4	$A_{2,2}$	$\sum \frac{w^2_{ijh.j.}}{w_{i.h.j.}}$	$A_{2,2}$	$A_{2,2}$
3	σ_3^4	$\sum \frac{w^2_{ij..t*}}{w_{ij..t.}}$	$\sum \frac{w^2_{ijh..h}}{w_{i.h.j.}}$	$A_{3,16}$	$A_{3,16}$
4	σ_{12}^4	$A_{11,9}$	$\sum \frac{w^2_{ijhij.}}{w_{.j.i.h}}$	$\sum \frac{w^2_{.jh*j.}}{w_{.jh.j.}}$	$A_{11,9}$
5	σ_{13}^4	$\sum \frac{w^2_{ij.it*}}{w_{ij..t.}}$	$\sum \frac{w^2_{ijhi.h}}{w_{.j.i.h}}$	$\sum \frac{w^2_{.jh*th}}{w_{.jh.t.}}$	$A_{5,18}$
6	σ_{23}^4	$\sum \frac{w^2_{ij..j*}}{w_{ij..j.}}$	$\sum \frac{n^2_{ijh}n_{.jh}}{w_{.j.i.h}}$	$A_{24,16}$	$A_{24,16}$
7	σ_{123}^4	$\sum \frac{w^2_{ij.ij*}}{w_{ij..j.}}$	$\sum \frac{w^2_{ijhijh}}{w_{.j.i.h}}$	$\sum \frac{w^2_{.jh*jh}}{w_{.j..jh}}$	$A_{32,18}$
8	σ_0^4	J	$\sum \frac{w_{ijhijh}}{w_{.j.i.h}}$	J	J
9	$2\sigma_1^2\sigma_2^2$	$A_{9,1}$	$\sum \frac{w^2_{ijhi..}}{w_{i.hi..}}$	$A_{9,1}$	$A_{9,1}$
10	$2\sigma_1^2\sigma_3^2$	$\sum \frac{w_{.j.*.h}n_{.jh}}{n_{.j.}}$	$\sum \frac{w_{i....h}n^2_{ijh}}{w_{i.h.j.}}$	$A_{10,17}$	$A_{10,17}$
11	$2\sigma_1^2\sigma_{12}^2$	$A_{11,9}$	$\sum \frac{w_{i..ij.}n^2_{ijh}}{w_{i.h.j.}}$	$A_{4,19}$	$A_{11,9}$
12	$2\sigma_1^2\sigma_{13}^2$	$A_{12,9}$	$\sum \frac{w^2_{ijhi..}}{w_{.j.i..}}$	$\sum \frac{w_{.jh*t.}w_{.jh*th}}{w_{.jh.t.}}$	$A_{12,9}$

Row	Column (1/2)Cov Product	21 T_2, T_f	22 T_3, T_{12}	23 T_3, T_{13}	24 T_3, T_{23}
1	σ_1^4	$\sum \frac{w^2_{.j.*..}}{w_{.j....}}$	$\sum \frac{w^2_{ijhi..}}{w_{ij...h}}$	$A_{1,10}$	$\sum \frac{w^2_{.jh*.u}}{w_{.jh..u}}$
2	σ_2^4	$\sum \frac{w^2_{.j..j.}}{w_{.j....}}$	$\sum \frac{w^2_{ijh.j.}}{w_{ij...h}}$	$\sum \frac{w^2_{i.h.*u}}{w_{i.h..u}}$	$A_{2,16}$
3	σ_3^4	$\sum \frac{w^2_{.j...*}}{w_{.j....}}$	$\sum \frac{w^2_{ijh..h}}{w_{ij...h}}$	$A_{3,3}$	$A_{3,3}$
4	σ_{12}^4	$\sum \frac{w^2_{.j.*j.}}{w_{.j....}}$	$\sum \frac{w^2_{ijhij.}}{w_{ij...h}}$	$\sum \frac{w^2_{.jh*ju}}{w_{.jh..u}}$	$\sum \frac{w^2_{i.hi*u}}{w_{i.h..u}}$
5	σ_{13}^4	$\sum \frac{w^2_{.j.*.*}}{w_{.j....}}$	$\sum \frac{w^2_{ijhi.h}}{w_{ij...h}}$	$A_{12,10}$	$\sum \frac{w^2_{.jh*.h}}{w_{.jh..h}}$
6	σ_{23}^4	$\sum \frac{w^2_{.j..j*}}{w_{.j....}}$	$\sum \frac{w^2_{ijh.jh}}{w_{ij...h}}$	$\sum \frac{w^2_{i.h.*h}}{w_{i.h..h}}$	$A_{19,16}$
7	σ_{123}^4	$\sum \frac{w^2_{.j.*j*}}{w_{.j....}}$	$\sum \frac{w^2_{ijhijh}}{w_{ij...h}}$	$\sum \frac{w^2_{i.hi*h}}{w_{i.h..h}}$	$\sum \frac{w^2_{.jh*jh}}{w_{.jh..h}}$
8	σ_0^4	1	$\sum \frac{w_{ijhijh}}{w_{ij...h}}$	H	H
9	$2\sigma_1^2\sigma_2^2$	$A_{9,15}$	$\sum \frac{w_{i...j.}n^2_{ijh}}{w_{ij...h}}$	$\sum \frac{w_{ij.i.h}n_{.jh}}{n_{..h}}$	$A_{9,23}$
10	$2\sigma_1^2\sigma_3^2$	$\sum \frac{w_{.j.*..}w_{j...*}}{w_{.j....}}$	$\sum \frac{w^2_{ijhi..}}{w_{ij.i..}}$	$A_{10,1}$	$A_{10,1}$
11	$2\sigma_1^2\sigma_{12}^2$	$\sum \frac{w_{.j.*..}w_{.j.*j.}}{w_{.j....}}$	$\sum \frac{w^2_{ijhi..}}{w_{..hi..}}$	$A_{11,10}$	$\sum \frac{w_{.jh*.u}w_{.jh*ju}}{w_{.jh..h}}$
12	$2\sigma_1^2\sigma_{13}^2$	$\sum \frac{w_{.j.*..}w_{.j.*.*}}{w_{.j....}}$	$\sum \frac{w_{ijhi..}n^2_{ijh}}{w_{ij...h}}$	$A_{12,10}$	$A_{5,24}$

Row	Column (1/2)Cov Product	25 T_3, T_{123}	26 T_3, T_f	27 T_{12}, T_{13}	28 T_{12}, T_{23}	29 T_{12}, T_{123}	30 T_{12}, T_f
1	σ_1^4	$A_{1,10}$	$\sum \frac{w^2_{..h*..}}{w_{..h...}}$	$A_{1,1}$	$A_{3,22}$	$A_{1,1}$	$A_{1,15}$
2	σ_2^4	$A_{2,16}$	$\sum \frac{w^2_{..h.*.}}{w_{..h...}}$	$A_{2,18}$	$A_{2,2}$	$A_{2,2}$	$A_{2,21}$
3	σ_3^4	$A_{3,3}$	$\sum \frac{w^2_{..h..h}}{w_{..h...}}$	$A_{1,13}$	$A_{1,13}$	$A_{1,13}$	$\sum \frac{w^2_{ij...*}}{w_{ij....}}$
4	σ_{12}^4	$A_{4,22}$	$\sum \frac{w^2_{..h**.}}{w_{..h...}}$	$A_{17,18}$	$A_{11,13}$	$A_{9,1}$	$\sum \frac{w^2_{ij.ij.}}{w_{ij....}}$
5	σ_{13}^4	$A_{12,10}$	$\sum \frac{w^2_{..h*.h}}{w_{..h...}}$	$A_{23,22}$	$\sum \frac{w^2_{ijhith}}{w_{ij..th}}$	$A_{23,22}$	$\sum \frac{w^2_{ij.i.*}}{w_{ij....}}$
6	σ_{23}^4	$A_{19,16}$	$\sum \frac{w^2_{..h.*h}}{w_{..h...}}$	$\sum \frac{w^2_{ijhsjh}}{w_{ij.s.h}}$	$A_{24,22}$	$A_{24,22}$	$\sum \frac{w^2_{ij..j*}}{w_{ij....}}$
7	σ_{123}^4	$A_{29,22}$	$\sum \frac{w^2_{..h**h}}{w_{..h...}}$	$\sum \frac{w^2_{ijhijh}}{w_{ij.i.h}}$	$\sum \frac{w^2_{ijhijh}}{w_{ij..jh}}$	$A_{25,22}$	$\sum \frac{w^2_{ij.ij*}}{w_{ij....}}$
8	σ_0^4	H	1	$\sum \frac{w_{ijhijh}}{w_{ij.i.h}}$	$\sum \frac{w_{ijhijh}}{w_{ij..jh}}$	m_{12}	1
9	$2\sigma_1^2\sigma_1^2$	$A_{9,23}$	$\sum \frac{w_{..h*..}w_{..h.*.}}{w_{..h...}}$	$A_{9,1}$	$A_{9,1}$	$A_{9,1}$	$A_{9,15}$
10	$2\sigma_1^2\sigma_2^2$	$A_{10,1}$	$A_{10,15}$	$A_{10,1}$	$A_{10,1}$	$A_{10,1}$	$A_{10,15}$
11	$2\sigma_1^2\sigma_{12}^2$	$A_{11,10}$	$\sum \frac{w_{..h*..}w_{..h**.}}{w_{..h...}}$	$A_{9,1}$	$A_{11,13}$	$A_{9,1}$	$A_{11,15}$
12	$2\sigma_1^2\sigma_{13}^2$	$A_{12,10}$	$\sum \frac{w_{..h*..}w_{..h*.h}}{w_{..h...}}$	$A_{10,1}$	$A_{11,14}$	$A_{10,1}$	$A_{12,15}$

Row	Column (1/2)Cov Product	31 T_{13}, T_{23}	32 T_{13}, T_{123}	33 T_{13}, T_f	34 T_{23}, T_{123}	35 T_{23}, T_f	36 T_{123}, T_f
1	σ_1^4	$A_{3,22}$	$A_{1,1}$	$A_{1,15}$	$A_{3,22}$	$\sum \frac{w^2_{.jh*..}}{w_{.jh...}}$	$A_{1,15}$
2	σ_2^4	$A_{2,18}$	$A_{2,18}$	$\sum \frac{w^2_{i.h.*.}}{w_{i.h...}}$	$A_{2,2}$	$A_{2,21}$	$A_{2,21}$
3	σ_3^4	$A_{3,3}$	$A_{3,3}$	$A_{3,26}$	$A_{3,3}$	$A_{3,26}$	$A_{3,26}$
4	σ_{12}^4	$\sum \frac{w^2_{ijhiju}}{w_{i.h.ju}}$	$A_{17,18}$	$\sum \frac{w^2_{i.hi*.}}{w_{i.h...}}$	$A_{11,13}$	$\sum \frac{w^2_{.jh*j.}}{w_{.jh...}}$	$A_{4,30}$
5	σ_{13}^4	$A_{11,14}$	$A_{10,1}$	$\sum \frac{w^2_{i.hi.h}}{w_{i.h...}}$	$A_{11,14}$	$\sum \frac{w^2_{.jh*.h}}{w_{.jh...}}$	$A_{5,33}$
6	σ_{23}^4	$A_{19,18}$	$A_{19,18}$	$\sum \frac{w^2_{i.h.*h}}{w_{i.h...}}$	$A_{16,2}$	$\sum \frac{w^2_{.jh.jh}}{w_{.jh...}}$	$A_{6,35}$
7	σ_{123}^4	$\sum \frac{w^2_{ijhijh}}{w_{i.h.jh}}$	$A_{21,5}$	$\sum \frac{w^2_{i.hi*h}}{w_{i.h...}}$	$A_{15,6}$	$\sum \frac{w^2_{.jh*jh}}{w_{.jh...}}$	$\sum \frac{w^2_{ijhijh}}{w_{ijh...}}$
8	σ_0^4	$\sum \frac{w_{ijhijh}}{w_{i.h.jh}}$	m_{13}	1	m_{23}	1	1
9	$2\sigma_1^2\sigma_2^2$	$A_{9,1}$	$A_{9,1}$	$A_{9,15}$	$A_{9,1}$	$A_{9,15}$	$A_{9,15}$
10	$2\sigma_1^2\sigma_3^2$	$A_{10,1}$	$A_{10,1}$	$A_{10,15}$	$A_{10,1}$	$A_{10,15}$	$A_{10,15}$
11	$2\sigma_1^2\sigma_{12}^2$	$A_{11,13}$	$A_{9,1}$	$A_{11,15}$	$A_{11,13}$	$\sum \frac{w_{.jh*..}w_{.jh*j.}}{w_{.jh...}}$	$A_{11,15}$
12	$2\sigma_1^2\sigma_{13}^2$	$A_{11,14}$	$A_{10,1}$	$A_{12,15}$	$A_{11,14}$	$\sum \frac{w_{.jh*..}w_{.jh*j.}}{w_{.jh...}}$	$A_{12,15}$

Row	Column (1/2)Cov Product	1 T_1	2 T_2	3 T_3	4 T_{12}	5 T_{13}
13	$2\sigma_1^2\sigma_{23}^2$	$w_{...***}$	$\sum \frac{w^2_{ij..jh}}{w_{.j..j.}}$	$\sum \frac{w^2_{i.h.jh}}{w_{..h..h}}$	$A_{13,1}$	$A_{13,1}$
14	$2\sigma_1^2\sigma_{123}^2$	$A_{13,1}$	$\sum \frac{w^2_{ij.sjh}}{w_{.j..j.}}$	$\sum \frac{w^2_{i.hsjh}}{w_{..h..h}}$	$A_{13,1}$	$A_{13,1}$
15	$2\sigma_1^2\sigma_0^2$	n	$\sum \frac{w_{.j.*j.}}{n_{.j.}}$	$\sum \frac{w_{..h*.h}}{n_{..h}}$	n	n
16	$2\sigma_2^2\sigma_3^2$	$\sum_{j,h}\left(\sum_i \frac{w_{ij.i.h}}{n_{i..}}\right)^2$	$w_{....**}$	$A_{16,2}$	$A_{16,2}$	$A_{16,2}$
17	$2\sigma_2^2\sigma_{12}^2$	$A_{2,1}$	$A_{9,1}$	$\sum_{i,j,t}\left(\sum_h \frac{w_{.jhith}}{n_{..h}}\right)^2$	$A_{9,1}$	$A_{4,5}$
18	$2\sigma_2^2\sigma_{13}^2$	$\sum \frac{w^2_{ij.i.h}}{w_{i..i..}}$	$A_{13,1}$	$A_{18,1}$	$A_{13,1}$	$A_{13,1}$
19	$2\sigma_2^2\sigma_{23}^2$	$\sum_{j,t,h}\left(\sum_i \frac{w_{ij.ith}}{n_{i..}}\right)^2$	$A_{16,2}$	$A_{16,2}$	$A_{16,2}$	$A_{6,5}$
20	$2\sigma_2^2\sigma_{123}^2$	$\sum \frac{w^2_{ij.ith}}{w_{i..i..}}$	$A_{13,1}$	$A_{17,3}$	$A_{13,1}$	$A_{7,5}$
21	$2\sigma_2^2\sigma_0^2$	$\sum \frac{w_{i..i*.}}{n_{i..}}$	n	$\sum \frac{w_{..h.*h}}{n_{..h}}$	n	$\sum \frac{w_{i.hi*h}}{n_{i.h}}$
22	$2\sigma_3^2\sigma_{12}^2$	$A_{18,1}$	$A_{13,2}$	$A_{13,1}$	$A_{13,1}$	$A_{13,1}$
23	$2\sigma_3^2\sigma_{13}^2$	$A_{5,1}$	$\sum_{i,h,u}\left(\sum_j \frac{w_{.jhiju}}{n_{.j.}}\right)^2$	$A_{10,1}$	$A_{5,4}$	$A_{10,1}$
24	$2\sigma_3^2\sigma_{23}^2$	$\sum_{j,h,u}\left(\sum_i \frac{w_{i.hiju}}{n_{i..}}\right)^2$	$\sum \frac{w^2_{.j..j*}}{w_{.j..j.}}$	$A_{16,2}$	$A_{6,4}$	$A_{16,2}$

Row	Column (1/2)Cov Product	6 T_{23}	7 T_{123}	8 T_f	9 T_1, T_2	10 T_1, T_{13}
13	$2\sigma_1^2\sigma_{23}^2$	$A_{13,1}$	$A_{13,1}$	$\frac{A_{1,1}A_{16,2}}{n^2}$	$\sum \frac{w_{ij..j*}n_{ij.}}{n_{.j.}}$	$\sum \frac{w_{..h*jh}n_{.jh}}{n_{..h}}$
14	$2\sigma_1^2\sigma_{123}^2$	$A_{7,6}$	$A_{13,1}$	$\frac{A_{1,1}A_{13,1}}{n^2}$	$\sum \frac{w_{ij.ijh}^2}{w_{ij..j.}}$	$\sum \frac{w_{ijhi.h}^2}{w_{i.h..h}}$
15	$2\sigma_1^2\sigma_0^2$	$\sum \frac{w_{.jh*jh}}{n_{.jh}}$	n	$\frac{A_{1,1}}{n}$	$A_{15,2}$	$A_{15,3}$
16	$2\sigma_2^2\sigma_3^2$	$A_{16,2}$	$A_{16,2}$	$\frac{A_{2,2}A_{3,3}}{n^2}$	$\sum \frac{w_{ij...*}n_{ij.}}{n_{i..}}$	$\sum \frac{w_{i..ij*}n_{.j.}}{n_{i..}}$
17	$2\sigma_2^2\sigma_{12}^2$	$A_{9,1}$	$A_{9,1}$	$\frac{A_{2,2}A_{9,1}}{n^2}$	$\sum \frac{w_{ij.ij.}^2}{w_{i..ij.}}$	$\sum \frac{w_{i.h.*.}w_{i.hi*.}}{w_{i....h}}$
18	$2\sigma_2^2\sigma_{13}^2$	$A_{13,1}$	$A_{13,1}$	$\frac{A_{2,2}A_{10,1}}{n^2}$	$\sum \frac{w_{ij.i.*}n_{.ij.}}{n_{i..}}$	$\sum \frac{w_{i.hi.h}w_{i.h.*.}}{w_{i....h}}$
19	$2\sigma_2^2\sigma_{23}^2$	$A_{16,2}$	$A_{16,2}$	$\frac{A_{2,2}A_{16,2}}{n^2}$	$\sum \frac{w_{ij..j*}n_{.ij.}}{n_{i..}}$	$\sum \frac{w_{i.h.*.}w_{i.h.*h}}{w_{i....h}}$
20	$2\sigma_2^2\sigma_{123}^2$	$A_{13,1}$	$A_{13,1}$	$\frac{A_{2,2}A_{13,1}}{n^2}$	$\sum \frac{w_{ijhij.}^2}{w_{ij.i..}}$	$\sum \frac{w_{i.hi*h}w_{i.h.*.}}{w_{i....h}}$
21	$2\sigma_2^2\sigma_0^2$	n	n	$\frac{A_{2,2}}{n}$	$A_{21,1}$	$\sum \frac{w_{i.h.*.}n_{i.h}}{w_{i....h}}$
22	$2\sigma_3^2\sigma_{12}^2$	$A_{13,1}$	$A_{13,1}$	$\frac{A_{3,3}A_{4,1}}{n^2}$	$\sum \frac{w_{ij.ij.}w_{ij...*}}{w_{i...j.}}$	$A_{18,9}$
23	$2\sigma_3^2\sigma_{13}^2$	$A_{10,1}$	$A_{10,1}$	$\frac{A_{3,3}A_{10,1}}{n^2}$	$\sum \frac{w_{ij...*}w_{ij.i.*}}{w_{i...j.}}$	$\sum \frac{w_{i.hi.h}^2}{w_{i..i.h}}$
24	$2\sigma_3^2\sigma_{23}^2$	$A_{16,2}$	$A_{16,2}$	$\frac{A_{3,3}A_{16,2}}{n^2}$	$\sum \frac{w_{ij..j*}w_{ij...*}}{w_{i...j.}}$	$\sum \frac{w_{i.h.*h}n_{i.h}}{n_{i..}}$

Row	Column (1/2)Cov Product	11 T_1, T_{12}	12 T_1, T_{13}	13 T_1, T_{23}	14 T_1, T_{123}	15 T_1, T_f
13	$2\sigma_1^2\sigma_{23}^2$	$A_{13,1}$	$A_{13,1}$	$A_{13,1}$	$A_{13,1}$	$\sum \frac{w_{i...**}n_{i..}}{n}$
14	$2\sigma_1^2\sigma_{123}^2$	$A_{13,1}$	$A_{13,1}$	$\sum \frac{w_{ijhijh}^2}{w_{.jhijh}}$	$A_{13,1}$	$\sum \frac{w_{ijhi..}}{w_{...i..}}$
15	$2\sigma_1^2\sigma_0^2$	n	n	$A_{15,6}$	n	$\frac{A_{1,1}}{n}$
16	$2\sigma_2^2\sigma_3^2$	$\sum \frac{w_{i...j*}n_{.ij.}}{n_{i..}}$	$A_{16,11}$	$\sum \frac{w_{.j...h}n_{ijh}^2}{w_{i...jh}}$	$A_{16,11}$	$\sum \frac{w_{i...*.}w_{i....*}}{w_{...i..}}$
17	$2\sigma_2^2\sigma_{12}^2$	$A_{17,9}$	$A_{4,12}$	$\sum \frac{w_{ij..j.}n_{ijh}^2}{w_{i...jh}}$	$A_{17,9}$	$\sum \frac{w_{i...*.}w_{i..i*.}}{w_{...i..}}$
18	$2\sigma_2^2\sigma_{13}^2$	$A_{18,9}$	$A_{18,9}$	$\sum \frac{w_{i.h.j.}n_{ijh}^2}{w_{i...jh}}$	$A_{18,9}$	$\sum \frac{w_{i...*.}w_{i..i.*}}{w_{...i..}}$
19	$2\sigma_2^2\sigma_{23}^2$	$A_{19,9}$	$\sum \frac{w_{i.hs*.}w_{i.hs*h}}{w_{i.hs..}}$	$\sum \frac{w_{ijh.j.}^2}{w_{i...j.}}$	$A_{19,9}$	$\sum \frac{w_{i...*.}w_{i...**}}{w_{...i..}}$
20	$2\sigma_2^2\sigma_{123}^2$	$A_{20,9}$	$\sum \frac{w_{i.hi*h}w_{i.hi*.}}{w_{i.hi..}}$	$\sum \frac{n_{ijh}^3 n_{.j.}}{w_{i...jh}}$	$A_{20,9}$	$\sum \frac{w_{i...*.}w_{i..i**}}{w_{...i..}}$
21	$2\sigma_2^2\sigma_0^2$	$A_{21,1}$	$A_{21,1}$	$\sum \frac{n_{ijh}^2 n_{.j.}}{w_{i...jh}}$	$A_{21,1}$	$\frac{A_{2,2}}{n}$
22	$2\sigma_3^2\sigma_{12}^2$	$A_{18,9}$	$A_{18,9}$	$\sum \frac{w_{ij...h}n_{ijh}^2}{w_{i...jh}}$	$A_{18,9}$	$\sum \frac{w_{i....*}w_{i..i*.}}{w_{...i..}}$
23	$2\sigma_3^2\sigma_{13}^2$	$A_{5,11}$	$A_{23,10}$	$\sum \frac{w_{i.h..h}n_{ijh}^2}{w_{i...jh}}$	$A_{23,10}$	$\sum \frac{w_{i....*}w_{i..i.*}}{w_{...i..}}$
24	$2\sigma_3^2\sigma_{23}^2$	$\sum \frac{w_{ij.s.*}w_{ij.sj*}}{w_{ij.s..}}$	$A_{24,10}$	$\sum \frac{w_{ijh..h}^2}{w_{i....h}}$	$A_{24,10}$	$\sum \frac{w_{i....*}w_{i..i**}}{w_{...i..}}$

Row	Column (1/2)Cov Product	16 T_2, T_3	17 T_2, T_{12}	18 T_2, T_{13}	19 T_2, T_{23}	20 T_2, T_{123}
13	$2n_1^2 n_{23}^2$	$\sum \frac{w_{.jh.jh} w_{.jh*..}}{w_{.j...h}}$	$A_{13,9}$	$\sum \frac{w_{i...jh} n_{ijh}^2}{w_{.j.i.h}}$	$A_{13,9}$	$A_{13,9}$
14	$2n_1^2 n_{123}^2$	$\sum \frac{w_{.jh*..} w_{.jh*jh}}{w_{.j...h}}$	$A_{14,9}$	$\sum \frac{n_{ijh}^3 n_{i..}}{w_{.j.i.h}}$	$\sum \frac{w_{.jh*j.} w_{.jh*jh}}{w_{.jh.j.}}$	$A_{14,9}$
15	$2n_1^2 n_0^2$	$\sum \frac{w_{.jh*..} n_{.jh}}{w_{.j...h}}$	$A_{15,2}$	$\sum \frac{n_{ijh}^2 n_{i..}}{w_{.j.i.h}}$	$A_{15,2}$	$A_{15,2}$
16	$2n_2^2 n_3^2$	$A_{16,2}$	$A_{16,2}$	$\sum \frac{w_{ijh..h}^2}{w_{i.h..h}}$	$A_{16,2}$	$A_{16,2}$
17	$2n_2^2 n_{12}^2$	$\sum \frac{w_{.jh*j.} n_{.jh}}{n_{..h}}$	$A_{9,1}$	$\sum \frac{w_{ijhij.}^2}{w_{i.hij.}}$	$A_{9,1}$	$A_{9,1}$
18	$2n_2^2 n_{13}^2$	$A_{13,9}$	$A_{23,1}$	$A_{13,1}$	$A_{13,1}$	$A_{13,1}$
19	$2n_2^2 n_{23}^2$	$\sum \frac{w_{.jh.jh}^2}{w_{..h.jh}}$	$A_{16,2}$	$\sum \frac{w_{ijh.jh}^2}{w_{i.h.jh}}$	$A_{16,2}$	$A_{16,2}$
20	$2n_2^2 n_{123}^2$	$\sum \frac{w_{ijh.jh}^2}{w_{..h.jh}}$	$A_{13,1}$	$\sum \frac{w_{ijhijh}^2}{w_{i.hijh}}$	$A_{13,1}$	$A_{13,1}$
21	$2n_2^2 n_0^2$	$A_{21,3}$	n	$A_{21,5}$	n	n
22	$2n_3^2 n_{12}^2$	$A_{13,9}$	$A_{13,9}$	$\sum \frac{w_{ij...h} n_{ijh}^2}{w_{.j.i.h}}$	$A_{17,16}$	$A_{13,9}$
23	$2n_3^2 n_{13}^2$	$\sum \frac{w_{.jhijh} n_{i.h}}{n_{.j.}}$	$\sum \frac{w_{ij..t*} w_{ij.it*}}{w_{ij..t.}}$	$\sum \frac{w_{ijh..h}^2}{w_{.j...h}}$	$A_{23,16}$	$A_{23,16}$
24	$2n_3^2 n_{23}^2$	$\sum \frac{w_{.jh.jh}^2}{w_{.j..jh}}$	$A_{6,17}$	$\sum \frac{w_{ijh}^2 w_{.jh..h}}{w_{.j.i.h}}$	$A_{24,16}$	$A_{24,16}$

Row	Column (1/2)Cov Product	21 T_2, T_3	22 T_3, T_{12}	23 T_3, T_{13}	24 T_3, T_{23}	25 T_3, T_{123}
13	$2\sigma_1^2 \sigma_{23}^2$	$\sum \frac{w_{.j.*..} w_{.j..j*}}{w_{....j.}}$	$\sum \frac{w_{i...jh} n_{ijh}^2}{w_{ij...h}}$	$A_{13,10}$	$A_{13,10}$	$A_{13,10}$
14	$2\sigma_1^2 \sigma_{123}^2$	$\sum \frac{w_{.j.*..} w_{.j.*j*}}{w_{....j.}}$	$\sum \frac{n_{ijh}^3 n_{i..}}{w_{ij...h}}$	$A_{14,10}$	$\sum \frac{w_{.jh*.h} w_{.jh*jh}}{w_{.jh..h}}$	$A_{14,10}$
15	$2\sigma_1^2 \sigma_0^2$	$\frac{A_{1,1}}{n}$	$\sum \frac{n_{ijh}^2 n_{i..}}{w_{ij...h}}$	$A_{15,3}$	$A_{15,3}$	$A_{15,3}$
16	$2\sigma_2^2 \sigma_3^2$	$\sum \frac{w_{..hi*.} n_{..h}}{n}$	$\sum \frac{w_{ijh.j.}^2}{w_{ij..j.}}$	$A_{16,2}$	$A_{16,2}$	$A_{16,2}$
17	$2\sigma_2^2 \sigma_{12}^2$	$\sum \frac{w_{ij..j.}^2}{w_{....j.}}$	$\sum \frac{w_{ijh.j.}^2}{w_{.j...h}}$	$\sum \frac{w_{i.h.*u} w_{i.hi*u}}{w_{i.h..u}}$	$A_{17,16}$	$A_{17,16}$
18	$2\sigma_2^2 \sigma_{13}^2$	$\sum \frac{w_{.j.*.*} n_{.j.}}{n}$	$\sum \frac{w_{i.h.j.} n_{ijh}^2}{w_{ij...h}}$	$A_{13,10}$	$A_{13,10}$	$A_{13,10}$
19	$2\sigma_2^2 \sigma_{23}^2$	$\sum \frac{w_{.jh.j.}^2}{w_{....j.}}$	$\sum \frac{w_{.j..jh} n_{ijh}^2}{w_{ij...h}}$	$A_{6,23}$	$A_{19,16}$	$A_{19,16}$
20	$2\sigma_2^2 \sigma_{123}^2$	$\sum \frac{w_{ijh.j.}^2}{w_{....j.}}$	$\sum \frac{n_{ijh}^3 n_{.j.}}{w_{ij...h}}$	$\sum \frac{w_{i.h.*h} w_{i.hi*h}}{w_{i.h..h}}$	$A_{20,16}$	$A_{20,16}$
21	$2\sigma_2^2 \sigma_0^2$	$\frac{A_{2,2}}{n}$	$\sum \frac{n_{ijh}^2 n_{.j.}}{w_{ij...h}}$	$A_{21,3}$	$A_{21,3}$	$A_{21,3}$
22	$2\sigma_3^2 \sigma_{12}^2$	$\sum \frac{w_{.j...*} w_{.j.*j.}}{w_{....j.}}$	$A_{13,1}$	$A_{13,1}$	$A_{13,1}$	$A_{13,1}$
23	$2\sigma_3^2 \sigma_{13}^2$	$\sum \frac{w_{.j...*} w_{.j.*.*}}{w_{....j.}}$	$\sum \frac{w_{ijhi.h}^2}{w_{i.hij.}}$	$A_{10,2}$	$A_{10,2}$	$A_{10,1}$
24	$2\sigma_3^2 \sigma_{23}^2$	$\sum \frac{w_{.j...*} w_{.j..j*}}{w_{....j.}}$	$\sum \frac{w_{ijh.jh}^2}{w_{.jhij.}}$	$A_{16,2}$	$A_{16,2}$	$A_{16,2}$

Row	Column (1/2)Cov Product	26 T_3, T_f	27 T_{12}, T_{13}	28 T_{12}, T_{23}	29 T_{12}, T_{123}	30 T_{12}, T_f	31 T_{13}, T_{23}
13	$2\sigma_1^2\sigma_{23}^2$	$\sum \frac{w_{..h*..}w_{..h.*h}}{w_{..h...}}$	$A_{13,1}$	$A_{13,1}$	$A_{13,1}$	$A_{13,15}$	$A_{13,1}$
14	$2\sigma_1^2\sigma_{123}^2$	$\sum \frac{w_{..h*..}w_{..h**h}}{w_{..h...}}$	$A_{13,1}$	$A_{14,13}$	$A_{13,1}$	$A_{14,15}$	$A_{14,13}$
15	$2\sigma_1^2\sigma_0^2$	$\frac{A_{1,1}}{n}$	n	$A_{15,6}$	n	$\frac{A_{1,1}}{n}$	$A_{15,6}$
16	$2\sigma_2^2\sigma_3^2$	$A_{16,21}$	$A_{16,2}$	$A_{16,2}$	$A_{16,2}$	$A_{16,21}$	$A_{16,2}$
17	$2\sigma_2^2\sigma_{12}^2$	$\sum \frac{w_{..h.*.}w_{..h**.}}{w_{..h...}}$	$A_{17,18}$	$A_{9,1}$	$A_{9,1}$	$A_{17,21}$	$A_{17,18}$
18	$2\sigma_2^2\sigma_{13}^2$	$\sum \frac{w_{..h.*.}w_{..h*.h}}{w_{..h...}}$	$A_{13,1}$	$A_{13,1}$	$A_{13,1}$	$A_{18,21}$	$A_{13,1}$
19	$2\sigma_2^2\sigma_{23}^2$	$\sum \frac{w_{..h.*.}w_{..h.*h}}{w_{..h...}}$	$A_{19,18}$	$A_{16,2}$	$A_{16,2}$	$A_{19,21}$	$A_{19,18}$
20	$2\sigma_2^2\sigma_{123}^2$	$\sum \frac{w_{..h.*.}w_{..h**h}}{w_{..h...}}$	$A_{20,18}$	$A_{13,1}$	$A_{13,1}$	$A_{20,21}$	$A_{20,18}$
21	$2\sigma_2^2\sigma_0^2$	$\frac{A_{2,2}}{n}$	$A_{21,5}$	n	n	$\frac{A_{2,2}}{n}$	$A_{21,5}$
22	$2\sigma_3^2\sigma_{12}^2$	$\sum \frac{w_{..h**.}n_{..h}}{n}$	$A_{13,1}$	$A_{13,1}$	$A_{13,1}$	$A_{22,26}$	$A_{13,1}$
23	$2\sigma_3^2\sigma_{13}^2$	$\sum \frac{w_{..h*.h}n_{..h}}{n}$	$A_{23,22}$	$A_{23,22}$	$A_{23,22}$	$\sum \frac{w_{ij...*}w_{ij.i.*}}{w_{...ij.}}$	$A_{10,1}$
24	$2\sigma_3^2\sigma_{23}^2$	$\sum \frac{w_{..h.*h}n_{..h}}{n}$	$A_{24,22}$	$A_{24,22}$	$A_{24,22}$	$\sum \frac{w_{ij...*}w_{ij..j*}}{w_{...ij.}}$	$A_{16,2}$

Row	Column (1/2)Cov Product	32 T_{13}, T_{123}	33 T_{13}, T_f	34 T_{23}, T_{123}	35 T_{23}, T_f	36 T_{123}, T_f
13	$2\sigma_1^2\sigma_{23}^2$	$A_{13,1}$	$A_{13,15}$	$A_{13,1}$	$A_{13,15}$	$A_{13,15}$
14	$2\sigma_1^2\sigma_{123}^2$	$A_{13,1}$	$A_{14,15}$	$A_{14,13}$	$\sum \frac{w_{.jh*..}w_{.jh*jh}}{w_{....jh}}$	$A_{14,15}$
15	$2\sigma_1^2\sigma_0^2$	n	$\frac{A_{1,1}}{n}$	$A_{15,6}$	$\frac{A_{1,1}}{n}$	$\frac{A_{1,1}}{n}$
16	$2\sigma_2^2\sigma_3^2$	$A_{16,2}$	$A_{16,21}$	$A_{16,2}$	$A_{16,21}$	$A_{16,21}$
17	$2\sigma_2^2\sigma_{12}^2$	$A_{17,18}$	$\sum \frac{w_{i.h.*.}w_{i.hi*.}}{w_{...i.h}}$	$A_{9,1}$	$A_{17,21}$	$A_{17,21}$
18	$2\sigma_2^2\sigma_{13}^2$	$A_{13,1}$	$A_{18,21}$	$A_{13,1}$	$A_{18,21}$	$A_{15,21}$
19	$2\sigma_2^2\sigma_{23}^2$	$A_{19,18}$	$\sum \frac{w_{i.h.*.}w_{i.h.*h}}{w_{...i.h}}$	$A_{16,2}$	$A_{19,21}$	$A_{19,21}$
20	$2\sigma_2^2\sigma_{123}^2$	$A_{20,18}$	$\sum \frac{w_{i.h.*.}w_{i.hi*h}}{w_{...i.h}}$	$A_{13,1}$	$A_{20,21}$	$A_{20,21}$
21	$2\sigma_2^2\sigma_0^2$	$A_{21,5}$	$\frac{A_{2,2}}{n}$	n	$\frac{A_{2,2}}{n}$	$\frac{A_{2,2}}{n}$
22	$2\sigma_3^2\sigma_{12}^2$	$A_{13,1}$	$A_{22,26}$	$A_{13,1}$	$A_{22,26}$	$A_{22,26}$
23	$2\sigma_3^2\sigma_{13}^2$	$A_{10,1}$	$A_{23,26}$	$A_{10,1}$	$A_{23,26}$	$A_{23,26}$
24	$2\sigma_3^2\sigma_{23}^2$	$A_{16,2}$	$A_{24,26}$	$A_{16,2}$	$A_{24,26}$	$A_{24,26}$

Row	Column (1/2)Cov Product	1 T_1	2 T_2	3 T_3	4 T_{12}	5 T_{13}	6 T_{23}	7 T_{123}
25	$2\sigma_3^2\sigma_{123}^2$	$\sum \frac{w_{i.hiju}^2}{w_{i..i..}}$	$\sum \frac{w_{.jhiju}^2}{w_{.j..j.}}$	$A_{13,1}$	$A_{7,4}$	$A_{13,1}$	$A_{13,1}$	$A_{13,1}$
26	$2\sigma_3^2\sigma_0^2$	$\sum \frac{w_{i..i.*}}{n_{i..}}$	$\sum \frac{w_{.j..j*}}{n_{.j.}}$	n	$\sum \frac{w_{ij.ij*}}{n_{ij.}}$	n	n	n
27	$2\sigma_{12}^2\sigma_{13}^2$	$A_{18,1}$	$A_{14,2}$	$A_{14,3}$	$A_{13,1}$	$A_{13,1}$	$A_{7,6}$	$A_{13,1}$
28	$2\sigma_{12}^2\sigma_{23}^2$	$A_{20,1}$	$A_{13,2}$	$\sum \frac{w_{.jhith}^2}{w_{..h..h}}$	$A_{13,1}$	$A_{7,5}$	$A_{13,1}$	$A_{13,1}$
29	$2\sigma_{12}^2\sigma_{123}^2$	$A_{20,1}$	$A_{14,2}$	$A_{4,3}$	$A_{13,1}$	$A_{7,5}$	$A_{7,6}$	$A_{13,1}$
30	$2\sigma_{12}^2\sigma_0^2$	$A_{21,1}$	$A_{15,2}$	$[A_{4,3}]^{1/2}$	n	$A_{21,5}$	$A_{15,6}$	n
31	$2\sigma_{13}^2\sigma_{23}^2$	$A_{25,1}$	$A_{25,2}$	$A_{13,3}$	$A_{7,4}$	$A_{13,1}$	$A_{13,1}$	$A_{13,1}$
32	$2\sigma_{13}^2\sigma_{123}^2$	$A_{25,1}$	$A_{5,2}$	$\sum \frac{w_{i.hsjh}^2}{w_{..h..h}}$	$A_{7,4}$	$A_{13,1}$	$A_{7,6}$	$A_{13,1}$
33	$2\sigma_{13}^2\sigma_0^2$	$A_{26,1}$	$[A_{5,2}]^{1/2}$	$A_{15,3}$	$A_{26,4}$	n	n	n
34	$2\sigma_{23}^2\sigma_{123}^2$	$A_{6,1}$	$A_{5,2}$	$A_{28,3}$	$A_{7,4}$	$A_{7,5}$	$A_{13,1}$	$A_{13,1}$
35	$2\sigma_{23}^2\sigma_0^2$	$[A_{6,1}]^{1/2}$	$A_{26,2}$	$A_{15,3}$	$A_{26,4}$	$A_{21,5}$	$A_{15,6}$	n
36	$2\sigma_{123}^2\sigma_0^2$	$[A_{6,1}]^{1/2}$	$[A_{5,2}]^{1/2}$	$[A_{4,3}]^{1/2}$	$A_{26,4}$	$A_{21,5}$	$A_{15,6}$	n

Row	Column (1/2)Cov Product	8 T_f	9 T_1, T_2	10 T_1, T_3	11 T_1, T_{12}	12 T_1, T_{13}
25	$2\sigma_3^2\sigma_{123}^2$	$\frac{A_{3,3}A_{13,1}}{n^2}$	$\sum \frac{n_{iju}^2 w_{ij...*}}{w_{i...j.}}$	$\sum \frac{w_{i.hijh}^2}{w_{i..i.h}}$	$\sum \frac{n_{iju}^2 w_{ij.i.*}}{w_{ij.i..}}$	$A_{25,10}$
26	$2\sigma_3^2\sigma_0^2$	$\frac{A_{3,3}}{n}$	$\sum \frac{w_{ij...*} n_{ij\cdot}}{w_{i...j.}}$	$A_{26,1}$	$A_{26,1}$	$A_{26,1}$
27	$2\sigma_{12}^2\sigma_{13}^2$	$\frac{A_{9,1}A_{10,1}}{n^2}$	$\sum \frac{n_{ij.}^2 w_{ij.i.*}}{w_{i...j.}}$	$\sum \frac{n_{i.h}^2 w_{i.hi*.}}{w_{i....h}}$	$A_{18,9}$	$A_{18,9}$
28	$2\sigma_{12}^2\sigma_{23}^2$	$\frac{A_{9,1}A_{16,2}}{n^2}$	$\sum \frac{n_{ij.}^2 w_{ij..j*}}{w_{i...j.}}$	$\sum \frac{w_{i.hi*.} w_{i.h.*h}}{w_{i....h}}$	$A_{20,9}$	$A_{20,13}$
29	$2\sigma_{12}^2\sigma_{123}^2$	$\frac{A_{9,1}A_{13,1}}{n^2}$	$\sum \frac{w_{ij.ijh}^2}{w_{i...j.}}$	$\sum \frac{n_{ijh}^2 w_{i.hi*.}}{w_{i....h}}$	$A_{20,9}$	$A_{20,13}$
30	$2\sigma_{12}^2\sigma_0^2$	$\frac{A_{9,1}}{n}$	$\sum \frac{n_{ij.}^3}{w_{i...j.}}$	$\sum \frac{w_{i.hi*.} n_{i.h}}{w_{i....h}}$	$A_{21,1}$	$A_{21,1}$
31	$2\sigma_{13}^2\sigma_{23}^2$	$\frac{A_{10,1}A_{16,2}}{n^2}$	$\sum \frac{w_{ij..j*} w_{ij.i.*}}{w_{i...j.}}$	$\sum \frac{n_{i.h}^2 w_{i.h.*h}}{w_{i....h}}$	$A_{25,11}$	$A_{25,10}$
32	$2\sigma_{13}^2\sigma_{123}^2$	$\frac{A_{10,1}A_{13,1}}{n^2}$	$\sum \frac{n_{ijh}^2 w_{ij.i.*}}{w_{i...j.}}$	$\sum \frac{w_{ijhi.h}^2}{w_{i....h}}$	$A_{25,11}$	$A_{25,10}$
33	$2\sigma_{13}^2\sigma_0^2$	$\frac{A_{10,1}}{n}$	$\sum \frac{w_{ij.i.*} n_{ij.}}{w_{i...j.}}$	$\sum \frac{n_{i.h}^3}{w_{i....h}}$	$A_{26,1}$	$A_{26,1}$
34	$2\sigma_{23}^2\sigma_{123}^2$	$\frac{A_{16,2}A_{13,1}}{n^2}$	$\sum \frac{n_{ijh}^2 w_{ij..j*}}{w_{i...j.}}$	$\sum \frac{n_{ijh}^2 w_{i.h.*h}}{w_{i....h}}$	$A_{7,11}$	$A_{7,12}$
35	$2\sigma_{23}^2\sigma_0^2$	$\frac{A_{16,2}}{n}$	$\sum \frac{w_{ij..j*} n_{ij.}}{w_{i...j.}}$	$\sum \frac{w_{i.h.*h} n_{i.h}}{w_{i....h}}$	$A_{35,1}$	$A_{35,1}$
36	$2\sigma_{123}^2\sigma_0^2$	$\frac{A_{13,1}}{n}$	$\sum \frac{n_{ijh}^2 n_{ij.}}{w_{i...j.}}$	$\sum \frac{n_{ijh}^2 n_{i.h}}{w_{i....h}}$	$A_{35,1}$	$A_{35,1}$

Row	Column (1/2)Cov Product	13 T_1, T_{23}	14 T_1, T_{123}	15 T_1, T_f	16 T_2, T_3	17 T_2, T_{12}
25	$2\sigma_3^2\sigma_{123}^2$	$\sum \frac{n_{ijh}^3 n_{..h}}{w_{i...jh}}$	$A_{25,10}$	$\sum \frac{w_{i....*} w_{i..i**}}{w_{...i..}}$	$\sum \frac{w_{.jhijh}^2}{w_{.jh.j.}}$	$\sum \frac{w_{ij..j*} w_{ij.ij*}}{w_{ij..j.}}$
26	$2\sigma_3^2\sigma_0^2$	$\sum \frac{n_{ijh}^2 n_{..h}}{w_{i...jh}}$	$A_{26,1}$	$\frac{A_{3,3}}{n}$	$A_{26,2}$	$A_{26,2}$
27	$2\sigma_{12}^2\sigma_{13}^2$	$\sum \frac{n_{ijh}^2 w_{ij.i.h}}{w_{i...jh}}$	$A_{10,9}$	$\sum \frac{w_{ij.i.h}^2}{w_{...i..}}$	$\sum \frac{w_{.jh*j.} w_{.jh*.h}}{w_{.j...h}}$	$A_{14,9}$
28	$2\sigma_{12}^2\sigma_{23}^2$	$A_{20,9}$	$A_{20,9}$	$\sum \frac{w_{i..i*.} w_{i...**}}{w_{...i..}}$	$\sum \frac{w_{.jh.jh} w_{.jh*j.}}{w_{.j...h}}$	$A_{13,9}$
29	$2\sigma_{12}^2\sigma_{123}^2$	$\sum \frac{n_{ijh}^3 n_{ij.}}{w_{i...jh}}$	$A_{20,9}$	$\sum \frac{w_{ij.ith}^2}{w_{...i..}}$	$\sum \frac{w_{.jh*j.} w_{.jh*jh}}{w_{.j...h}}$	$A_{14,9}$
30	$2\sigma_{12}^2\sigma_0^2$	$\sum \frac{n_{ijh}^2 n_{ij.}}{w_{i...jh}}$	$A_{21,1}$	$\frac{A_{9,1}}{n}$	$\sum \frac{w_{.jh*j.} n_{.jh}}{w_{.j...h}}$	$A_{15,2}$
31	$2\sigma_{13}^2\sigma_{23}^2$	$A_{25,10}$	$A_{25,10}$	$\sum \frac{w_{i..i.*} w_{i...**}}{w_{...i..}}$	$\sum \frac{n_{.jh}^2 w_{.jh*.h}}{w_{.j...h}}$	$A_{25,17}$
32	$2\sigma_{13}^2\sigma_{123}^2$	$\sum \frac{n_{ijh}^3 n_{i.h}}{w_{i...jh}}$	$A_{25,10}$	$\sum \frac{w_{i.hiju}^2}{w_{...i..}}$	$\sum \frac{w_{.jh*jh} w_{.jh*.h}}{w_{.j...h}}$	$A_{17,17}$
33	$2\sigma_{13}^2\sigma_0^2$	$\sum \frac{n_{ijh}^2 n_{i.h}}{w_{i...jh}}$	$A_{26,1}$	$\frac{A_{10,1}}{n}$	$\sum \frac{w_{.jh*.h} n_{.jh}}{w_{.j...h}}$	$A_{33,2}$
34	$2\sigma_{23}^2\sigma_{123}^2$	$\sum \frac{w_{ijhijh}^2}{w_{ijhi..}}$	$A_{34,13}$	$\sum \frac{w_{i...**} w_{i..i**}}{w_{...i..}}$	$\sum \frac{w_{ijh.jh}^2}{w_{.j...h}}$	$A_{25,17}$
35	$2\sigma_{23}^2\sigma_0^2$	$A_{35,1}$	$A_{35,1}$	$\frac{A_{16,2}}{n}$	$\sum \frac{n_{.jh}^3}{w_{.j...h}}$	$A_{26,2}$
36	$2\sigma_{123}^2\sigma_0^2$	$\sum \frac{n_{ijh}^3}{w_{i...jh}}$	$A_{35,1}$	$\frac{A_{13,1}}{n}$	$\sum \frac{w_{.jh*jh} n_{.jh}}{w_{.j...h}}$	$A_{33,2}$

Row	Column (1/2)Cov Product	18 T_2, T_{13}	19 T_2, T_{23}	20 T_2, T_{123}	21 T_2, T_f	22 T_3, T_{12}	23 T_3, T_{13}
25	$2\sigma_3^2\sigma_{123}^2$	$\sum \frac{n_{ijh}^3 n_{..h}}{w_{.j.i.h}}$	$A_{25,16}$	$A_{25,16}$	$\sum \frac{w_{.j...*} w_{.j.*j*}}{w_{....j.}}$	$\sum \frac{w_{ijhijh}^2}{w_{ijhij.}}$	$A_{13,1}$
26	$2\sigma_3^2\sigma_0^2$	$\sum \frac{n_{ijh}^2 n_{..h}}{w_{.j.i.h}}$	$A_{26,2}$	$A_{26,2}$	$\frac{A_{3,3}}{n}$	$A_{26,4}$	n
27	$2\sigma_{12}^2\sigma_{13}^2$	$A_{14,9}$	$A_{14,19}$	$A_{14,9}$	$\sum \frac{w_{.j.*j.} w_{.j.*.*}}{w_{....j.}}$	$A_{14,10}$	$A_{14,10}$
28	$2\sigma_{12}^2\sigma_{23}^2$	$\sum \frac{n_{ijh}^2 w_{ij..jh}}{w_{.j.i.h}}$	$A_{13,9}$	$A_{13,9}$	$\sum \frac{w_{ij..jh}^2}{w_{....j.}}$	$A_{20,16}$	$A_{20,23}$
29	$2\sigma_{12}^2\sigma_{123}^2$	$\sum \frac{n_{ijh}^3 n_{ij.}}{w_{.j.i.h}}$	$A_{14,19}$	$A_{14,9}$	$\sum \frac{w_{ij..sjh}^2}{w_{....j.}}$	$\sum \frac{w_{ijhijh}^2}{w_{..hijh}}$	$A_{17,23}$
30	$2\sigma_{12}^2\sigma_0^2$	$\sum \frac{n_{ijh}^2 n_{ij.}}{w_{.j.i.h}}$	$A_{15,2}$	$A_{15,2}$	$\frac{A_{9,1}}{n}$	$A_{30,3}$	$A_{30,3}$
31	$2\sigma_{13}^2\sigma_{23}^2$	$A_{25,16}$	$A_{25,16}$	$A_{25,16}$	$\sum \frac{w_{.j..j*} w_{.j.*.*.}}{w_{....j.}}$	$\sum \frac{n_{ijh}^2 w_{i.h.jh}}{w_{ij...h}}$	$A_{13,10}$
32	$2\sigma_{13}^2\sigma_{123}^2$	$\sum \frac{w_{ijhijh}^2}{w_{ijh.j.}}$	$A_{7,19}$	$A_{32,18}$	$\sum \frac{w_{.j.*.*} w_{.j.*j*}}{w_{....j.}}$	$\sum \frac{n_{ijh}^3 n_{i.h}}{w_{ij...h}}$	$A_{14,10}$
33	$2\sigma_{13}^2\sigma_0^2$	$A_{33,2}$	$A_{33,2}$	$A_{33,2}$	$\frac{A_{10,1}}{n}$	$\sum \frac{n_{ijh}^2 n_{i.h}}{w_{ij...h}}$	$A_{15,3}$
34	$2\sigma_{23}^2\sigma_{123}^2$	$\sum \frac{n_{ijh}^3 n_{.jh}}{w_{.j.i.h}}$	$A_{25,16}$	$A_{25,16}$	$\sum \frac{w_{.jhijh}^2}{w_{....j.}}$	$\sum \frac{n_{ijh}^3 n_{.jh}}{w_{ij...h}}$	$A_{20,23}$
35	$2\sigma_{23}^2\sigma_0^2$	$\sum \frac{n_{ijh}^2 n_{.jh}}{w_{.j.i.h}}$	$A_{26,2}$	$A_{26,2}$	$\frac{A_{16,2}}{n}$	$\sum \frac{n_{ijh}^2 n_{.jh}}{w_{ij...h}}$	$A_{21,3}$
36	$2\sigma_{123}^2\sigma_0^2$	$\sum \frac{n_{ijh}^3}{w_{.j.i.h}}$	$A_{33,2}$	$A_{33,2}$	$\frac{A_{13,1}}{n}$	$\sum \frac{w_{ijhijh}^2}{w_{ij...h}}$	$A_{30,3}$

Row	Column (1/2)Cov Product	24 T_3, T_{23}	25 T_3, T_{123}	26 T_3, T_f	27 T_{12}, T_{13}	28 T_{12}, T_{23}	29 T_{12}, T_{123}
25	$2\sigma_3^2\sigma_{123}^2$	$A_{13,1}$	$A_{13,1}$	$\sum \frac{w_{ijh..h}^2}{w_{.....h}}$	$A_{25,22}$	$A_{25,22}$	$A_{25,22}$
26	$2\sigma_3^2\sigma_0^2$	n	n	$\frac{A_{3,3}}{n}$	$A_{26,4}$	$A_{26,4}$	$A_{26,4}$
27	$2\sigma_{12}^2\sigma_{13}^2$	$A_{14,24}$	$A_{14,10}$	$\sum \frac{w_{..h*.h} w_{..h**.}}{w_{.....h}}$	$A_{13,1}$	$A_{14,13}$	$A_{13,1}$
28	$2\sigma_{12}^2\sigma_{23}^2$	$A_{20,16}$	$A_{20,16}$	$\sum \frac{w_{..h.*h} w_{..h**.}}{w_{.....h}}$	$A_{20,18}$	$A_{13,1}$	$A_{13,1}$
29	$2\sigma_{12}^2\sigma_{123}^2$	$A_{7,24}$	$A_{29,22}$	$\sum \frac{w_{..h**.} w_{..h**h}}{w_{.....h}}$	$A_{20,18}$	$A_{14,13}$	$A_{13,1}$
30	$2\sigma_{12}^2\sigma_0^2$	$A_{30,3}$	$A_{30,3}$	$\frac{A_{9,1}}{n}$	$A_{21,5}$	$A_{15,6}$	n
31	$2\sigma_{13}^2\sigma_{23}^2$	$A_{13,10}$	$A_{13,10}$	$\sum \frac{w_{i.h.jh}^2}{w_{.....h}}$	$A_{25,22}$	$A_{25,22}$	$A_{25,22}$
32	$2\sigma_{13}^2\sigma_{123}^2$	$A_{14,24}$	$A_{14,10}$	$\sum \frac{w_{i.hsjh}^2}{w_{.....h}}$	$A_{25,22}$	$A_{7,28}$	$A_{25,22}$
33	$2\sigma_{13}^2\sigma_0^2$	$A_{15,3}$	$A_{15,3}$	$\frac{A_{10,1}}{n}$	$A_{26,4}$	$\sum \frac{n_{ijh}^3}{w_{ij..jh}}$	$A_{26,4}$
34	$2\sigma_{23}^2\sigma_{123}^2$	$A_{20,16}$	$A_{20,16}$	$\sum \frac{w_{.jhith}^2}{w_{.....h}}$	$A_{7,27}$	$A_{25,22}$	$A_{25,22}$
35	$2\sigma_{23}^2\sigma_0^2$	$A_{21,3}$	$A_{21,3}$	$\frac{A_{16,2}}{n}$	$\sum \frac{n_{ijh}^3}{w_{ij.i.h}}$	$A_{26,4}$	$A_{26,4}$
36	$2\sigma_{123}^2\sigma_0^2$	$A_{30,3}$	$A_{30,3}$	$\frac{A_{13,1}}{n}$	$A_{35,27}$	$A_{33,28}$	$A_{26,4}$

Row	Column (1/2)Cov Product	30 T_{12}, T_f	31 T_{13}, T_{23}	32 T_{13}, T_{123}	33 T_{13}, T_f	34 T_{23}, T_{123}
25	$2\sigma_3^2\sigma_{123}^2$	$\sum \frac{w_{ij...*}w_{ij.ij*}}{w_{...ij.}}$	$A_{13,1}$	$A_{13,1}$	$A_{25,26}$	$A_{13,1}$
26	$2\sigma_3^2\sigma_0^2$	$\frac{A_{3,3}}{n}$	n	n	$\frac{A_{3,3}}{n}$	n
27	$2\sigma_{12}^2\sigma_{13}^2$	$\sum \frac{w_{ij.ijh}n_{i.h}}{n}$	$A_{14,13}$	$A_{13,1}$	$A_{27,3}$	$A_{14,3}$
28	$2\sigma_{12}^2\sigma_{23}^2$	$\sum \frac{w_{ij.ijh}n_{.jh}}{n}$	$A_{20,18}$	$A_{20,18}$	$\sum \frac{w_{i.hi*.}w_{i.h.*h}}{w_{...i.h}}$	$A_{13,1}$
29	$2\sigma_{12}^2\sigma_{123}^2$	$\sum \frac{w_{ijhij.}^2}{w_{...ij.}}$	$A_{7,31}$	$A_{20,18}$	$\sum \frac{w_{i.hi*.}w_{i.hi*h}}{w_{...i.h}}$	$A_{14,13}$
30	$2\sigma_{12}^2\sigma_0^2$	$\frac{A_{9,1}}{n}$	$\sum \frac{n_{ijh}^3}{w_{i.h.jh}}$	$A_{21,5}$	$\frac{A_{9,1}}{n}$	$A_{15,6}$
31	$2\sigma_{13}^2\sigma_{23}^2$	$\sum \frac{w_{ij.i.*}w_{ij..j*}}{w_{...ij.}}$	$A_{13,1}$	$A_{13,1}$	$\sum \frac{w_{i.hijh}n_{.jh}}{n}$	$A_{13,1}$
32	$2\sigma_{13}^2\sigma_{123}^2$	$\sum \frac{w_{ij.i.*}w_{ij.ij*}}{w_{...ij.}}$	$A_{14,13}$	$A_{13,1}$	$\sum \frac{w_{ijhi.h}^2}{w_{...i.h}}$	$A_{14,13}$
33	$2\sigma_{13}^2\sigma_0^2$	$\frac{A_{10,1}}{n}$	$A_{15,6}$	n	$\frac{A_{10,1}}{n}$	$A_{15,6}$
34	$2\sigma_{23}^2\sigma_{123}^2$	$\sum \frac{w_{ij..j*}w_{ij.ij*}}{w_{...ij.}}$	$A_{20,18}$	$A_{20,18}$	$\sum \frac{w_{i.h.*h}w_{i.hi*h}}{w_{...i.h}}$	$A_{13,1}$
35	$2\sigma_{23}^2\sigma_0^2$	$\frac{A_{16,2}}{n}$	$A_{21,5}$	$A_{21,5}$	$\frac{A_{16,2}}{n}$	n
36	$2\sigma_{123}^2\sigma_0^2$	$\frac{A_{13,1}}{n}$	$A_{30,31}$	$A_{21,5}$	$\frac{A_{13,1}}{n}$	$A_{15,6}$

Row	Column (1/2)Cov Product	35 T_{23}, T_f	36 T_{123}, T_f
25	$2\sigma_3^2\sigma_{123}^2$	$A_{25,26}$	$A_{25,26}$
26	$2\sigma_3^2\sigma_0^2$	$\frac{A_{3,3}}{n}$	$\frac{A_{3,3}}{n}$
27	$2\sigma_{12}^2\sigma_{13}^2$	$\sum \frac{w_{.jh*j.}w_{.jh*.h}}{w_{....jh}}$	$A_{27,30}$
28	$2\sigma_{12}^2\sigma_{23}^2$	$A_{28,30}$	$A_{28,3}$
29	$2\sigma_{12}^2\sigma_{123}^2$	$\sum \frac{w_{.jh*j.}w_{.jh*jh}}{w_{....jh}}$	$A_{29,30}$
30	$2\sigma_{12}^2\sigma_0^2$	$\frac{A_{9,1}}{n}$	$\frac{A_{9,1}}{n}$
31	$2\sigma_{13}^2\sigma_{23}^2$	$A_{31,33}$	$A_{31,33}$
32	$2\sigma_{13}^2\sigma_{123}^2$	$\sum \frac{w_{.jh*.h}w_{.jh*jh}}{w_{....jh}}$	$A_{32,33}$
33	$2\sigma_{13}^2\sigma_0^2$	$\frac{A_{10,1}}{n}$	$\frac{A_{10,1}}{n}$
34	$2\sigma_{23}^2\sigma_{123}^2$	$\sum \frac{w_{ijh.jh}^2}{w_{....jh}}$	$A_{34,33}$
35	$2\sigma_{23}^2\sigma_0^2$	$\frac{A_{16,2}}{n}$	$\frac{A_{16,2}}{n}$
36	$2\sigma_{123}^2\sigma_0^2$	$\frac{A_{13,1}}{n}$	$\frac{A_{13,1}}{n}$

Source: Unpublished appendix to Blischke (1968) made available to one of the authors thanks to the courtesy of Dr. Blischke.

Bibliography

W. R. Blischke (1966), Variances of estimates of variance components in a three-way classification, *Biometrics*, **22**, 553–565.

W. R. Blischke (1968), Variances of moment estimators of variance components in the unbalanced r-way classification, *Biometrics*, **24**, 527–540.

H. O. Hartley (1967), Expectations, variances and covariances of ANOVA mean squares by "synthesis," *Biometrics*, **23**, 105–114; corrigenda, **23**, 853.

R. R. Hocking (1985), *The Analysis of Linear Models*, Brooks–Cole, Monterey, CA.

G. G. Koch (1967), A general approach to the estimation of variance components, *Techometrics*, **9**, 93–118.

G. G. Koch (1968), Some further remarks concerning "A general approach to estimation of variance components," *Technometrics*, **10**, 551–558.

J. N. K. Rao (1968), On expectations, variances, and covariances of ANOVA mean squares by "synthesis," *Biometrics*, **24**, 963–978.

S. R. Searle (1971), *Linear Models*, Wiley, New York.

15 Two-Way Nested Classification

Consider an experiment with two factors A and B where the levels of B are nested within the levels of A. Assume that there are a levels of A and within the ith level of A there are b_i levels of B and n_{ij} observations are taken at the jth level of B. The model for this design is known as the unbalanced two-way nested classification. This model is the same as the one considered in Chapter 6 except that now b_is and n_{ij}s rather than being constants vary from one level to the other. Models of this type are frequently used in many experiments and surveys since the sampling plans cannot be balanced because of the availability of limited resources. In addition, unless the number of levels of factor A is very large, the estimate of its variance component may be very imprecise for a balanced design. In this chapter, we consider the random effects model for the unbalanced two-way nested classification.

15.1 MATHEMATICAL MODEL

The random effects model for the unbalanced two-way nested classification is given by

$$\begin{aligned} y_{ijk} &= \mu + \alpha_i + \beta_{j(i)} + e_{k(ij)}, \\ & i = 1, \ldots, a; \quad j = 1, \ldots, b_i; \quad k = 0, \ldots, n_{ij}, \end{aligned} \tag{15.1.1}$$

where y_{ijk} is the kth observation at the jth level of factor B within the ith level of factor A, μ is the overall mean, α_i is the effect due to the ith level of factor A, $\beta_{j(i)}$ is the effect due to the jth level of factor B nested within the ith level of factor A, and $e_{k(ij)}$ is the residual error. It is assumed that α_is, $\beta_{j(i)}$s, and $e_{k(ij)}$s are mutually and completely uncorrelated random variables with means zero and variances σ_α^2, σ_β^2, and σ_e^2, respectively. The parameters σ_α^2, σ_β^2, and σ_e^2 are known as the variance components. Note that the model in (15.1.1) implies that the number of levels of factor A (main classification) is a and there are b_i levels of factor B (subclasses) within each level of A. Let $b_.$ denote the total

TABLE 15.1 Analysis of variance for the model in (15.1.1).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Factor A	$a-1$	SS_A	MS_A	$\sigma_e^2 + r_2\sigma_\beta^2 + r_3\sigma_\alpha^2$
Factor B within A	$b_. - a$	SS_B	MS_B	$\sigma_e^2 + r_1\sigma_\beta^2$
Error	$N - b_.$	SS_E	MS_E	σ_e^2

number of such subclasses, giving $b_. = \sum_{i=1}^{a} b_i$. The number of observations in the jth subclass of the ith class is n_{ij}.

15.2 ANALYSIS OF VARIANCE

For the model in (15.1.1), there is no unique analysis of variance and the conventional analysis of variance is shown in Table 15.1. The sum of squares terms in Table 15.1, known as Type I sums of squares, are defined by establishing an analogy with the corresponding terms for balanced data and are given as follows:

$$SS_A = \sum_{i=1}^{a} n_{i.}(\bar{y}_{i..} - \bar{y}_{...})^2 = \sum_{i=1}^{a} \frac{y_{i..}^2}{n_{i.}} - \frac{y_{...}^2}{N},$$

$$SS_B = \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij}(\bar{y}_{ij.} - \bar{y}_{i..})^2 = \sum_{i=1}^{a}\sum_{j=1}^{b_i} \frac{y_{ij.}^2}{n_{ij}} - \sum_{i=1}^{a} \frac{y_{i..}^2}{n_{i.}}, \tag{15.2.1}$$

and

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij.})^2 = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{n_{ij}} y_{ijk}^2 - \sum_{i=1}^{a}\sum_{j=1}^{b_i} \frac{y_{ij.}^2}{n_{ij}},$$

where

$$y_{ij.} = \sum_{k=1}^{n_{ij}} y_{ijk}, \qquad \bar{y}_{ij.} = \frac{y_{ij.}}{n_{ij}},$$

$$y_{i..} = \sum_{j=1}^{b_i} y_{ij.}, \qquad \bar{y}_{i..} = \frac{y_{i..}}{n_{i.}},$$

and

$$y_{...} = \sum_{i=1}^{a} y_{i..}, \qquad \bar{y}_{...} = \frac{y_{...}}{N},$$

with

$$n_{i.} = \sum_{j=1}^{b_i} n_{ij}, \quad \text{and} \quad N = \sum_{i=1}^{a} n_{i.}.$$

Define the uncorrected sums of squares as

$$T_A = \sum_{i=1}^{a} \frac{y_{i..}^2}{n_{i.}}, \qquad T_B = \sum_{i=1}^{a}\sum_{j=1}^{b_i} \frac{y_{ij.}^2}{n_{ij}},$$

$$T_0 = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{n_{ij}} y_{ijk}^2, \quad \text{and} \quad T_\mu = \frac{y_{...}^2}{N}.$$

Then the corrected sums of squares defined in (15.2.1) can be written as

$$\mathrm{SS}_A = T_A - T_\mu, \qquad \mathrm{SS}_B = T_B - T_A,$$

and

$$\mathrm{SS}_E = T_0 - T_B.$$

The mean squares are obtained by dividing the sums of squares by the corresponding degrees of freedom. The expected mean squares are readily obtained and the derivations are presented in the following section.

15.3 EXPECTED MEAN SQUARES

The expected values of the sums of squares or equivalently the mean squares can be readily obtained by first calculating the expected values of the quantities T_0, T_μ, T_A, and T_B. First, note that by the assumptions of the model in (15.1.1),

$$E(\alpha_i) = E(\beta_{j(i)}) = E(e_{k(ij)}) = 0,$$
$$E(\alpha_i^2) = \sigma_\alpha^2, \qquad E(\beta_{j(i)}^2) = \sigma_\beta^2, \quad \text{and} \quad E(e_{k(ij)}^2) = \sigma_e^2.$$

Further, all covariances between the elements of the same random variable and any pair of nonidentical random variables are equal to zero.

Now, we have

$$\begin{aligned} E(T_0) &= \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{n_{ij}} E(y_{ijk}^2) \\ &= \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{n_{ij}} E[\mu + \alpha_i + \beta_{j(i)} + e_{k(ij)}]^2 \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{n_{ij}}[\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_e^2] \\
&= N(\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_e^2), \\
E(T_\mu) &= E\left(\frac{y_{...}^2}{N}\right) \\
&= N^{-1}E\left[N\mu + \sum_{i=1}^{a} n_{i.}\alpha_i + \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij}\beta_{j(i)} + \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{n_{ij}} e_{k(ij)}\right]^2 \\
&= N^{-1}\left[N^2\mu^2 + \sum_{i=1}^{a} n_{i.}^2\sigma_\alpha^2 + \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij}^2\sigma_\beta^2 + N\sigma_e^2\right] \\
&= N\mu^2 + k_1\sigma_\alpha^2 + k_3\sigma_\beta^2 + \sigma_e^2, \\
E(T_A) &= \sum_{i=1}^{a} E\left(\frac{y_{i..}^2}{n_{i.}}\right) \\
&= \sum_{i=1}^{a} n_{i.}^{-1}E\left[n_{i.}(\mu + \alpha_i) + \sum_{j=1}^{b_i} n_{ij}\beta_{j(i)} + \sum_{j=1}^{b_i}\sum_{k=1}^{n_{ij}} e_{k(ij)}\right]^2 \\
&= \sum_{i=1}^{a} n_{i.}^{-1}\left[n_{i.}^2(\mu^2 + \sigma_\alpha^2) + \sum_{j=1}^{b_i} n_{ij}^2\sigma_\beta^2 + n_{i.}\sigma_e^2\right] \\
&= \sum_{i=1}^{a}\left[n_{i.}(\mu^2 + \sigma_\alpha^2) + \sum_{j=1}^{b_i}\frac{n_{ij}^2}{n_{i.}}\sigma_\beta^2 + \sigma_e^2\right] \\
&= N(\mu^2 + \sigma_\alpha^2) + k_{12}\sigma_\beta^2 + a\sigma_e^2,
\end{aligned}
$$

and

$$
\begin{aligned}
E(T_B) &= \sum_{i=1}^{a}\sum_{j=1}^{b_i} E\left(\frac{y_{ij.}^2}{n_{ij}}\right) \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij}^{-1}E\left[n_{ij}(\mu + \alpha_i + \beta_{j(i)}) + \sum_{k=1}^{n_{ij}} e_{k(ij)}\right]^2 \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij}^{-1}\left[n_{ij}^2(\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2) + n_{ij}\sigma_e^2\right] \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b_i}[n_{ij}(\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2) + \sigma_e^2]
\end{aligned}
$$

$$= N(\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2) + b_. \sigma_e^2,$$

where

$$k_1 = \frac{1}{N} \sum_{i=1}^{a} n_{i.}^2, \qquad k_3 = \frac{1}{N} \sum_{i=1}^{a} \sum_{j=1}^{b_i} n_{ij}^2,$$

and

$$k_{12} = \sum_{i=1}^{a} \frac{\sum_{j=1}^{b_i} n_{ij}^2}{n_{i.}}.$$

Hence, expected values of sums of squares and mean squares are given as follows:

$$\begin{aligned}
E(\mathrm{SS}_E) &= E[T_0 - T_B] \\
&= (N - b_.)\sigma_e^2, \\
E(\mathrm{MS}_E) &= \frac{1}{N - b_.} E[\mathrm{SS}_E] \\
&= \sigma_e^2, \\
E(\mathrm{SS}_B) &= E[T_B - T_A] \\
&= (b_. - a)\sigma_e^2 + (N - k_{12})\sigma_\beta^2, \\
E(\mathrm{MS}_B) &= \frac{1}{b_. - a} E[\mathrm{SS}_B] \\
&= \sigma_e^2 + r_1 \sigma_\beta^2, \\
E(\mathrm{SS}_A) &= E[T_A - T_\mu] \\
&= (a - 1)\sigma_e^2 + (k_{12} - k_3)\sigma_\beta^2 + (N - k_1)\sigma_\alpha^2,
\end{aligned}$$

and

$$\begin{aligned}
E(\mathrm{MS}_A) &= \frac{1}{a - 1} E[\mathrm{SS}_A] \\
&= \sigma_e^2 + r_2 \sigma_\beta^2 + r_3 \sigma_\alpha^2,
\end{aligned}$$

where

$$r_1 = \frac{N - k_{12}}{b_. - a}, \qquad r_2 = \frac{k_{12} - k_3}{a - 1},$$

and

$$r_3 = \frac{N - k_1}{a - 1}.$$

The results on expected mean squares were first derived by Ganguli (1941) and can also be found in Graybill (1961, pp. 354–357) and Searle (1961). A simple derivation of these results using matrix formulation is given by Verdooren (1988). Note that in the case of a balanced design, $b_i = b$, $n_{ij} = n$ for all i and j, $r_1 = r_2 = n$, and $r_3 = bn$.

15.4 DISTRIBUTION THEORY

Under the assumption of normality, MS_E is statistically independent of MS_B and MS_A and

$$\mathrm{MS}_E \sim \frac{\sigma_e^2}{v_e}\chi^2[v_e], \tag{15.4.1}$$

where $v_e = N - b_.$. However, in general, MS_A and MS_B do not have a chi-square type distribution and neither are they statistically independent. In the special case when $n_{ij} = n_i$ $(i = 1, 2, \ldots, a)$, it has been shown by Cummings (1972) that MS_A and MS_B are jointly independent but they do not have a chi-square type distribution due to different numbers of observations in the subclasses. Cummings (1972) has also shown that designs with $b_i = 2$, $n_{i1} = n_1$, $n_{i2} = n_2$ $(i = 1, 2, \ldots, a)$ have mean squares MS_A and MS_B with a chi-square type distribution but are dependent. Further, if $n_{ij} = n$ for all i and j, then MS_A and MS_B are jointly independent[1] and

$$\mathrm{MS}_B \sim \frac{(\sigma_e^2 + n\sigma_\beta^2)}{v_\beta}\chi^2[v_\beta], \tag{15.4.2}$$

where $v_\beta = b_. - a$; but MS_A in general does not have a scaled chi-square distribution (see, e.g., Scheffé, 1959, p. 252). It has a chi-square type distribution, if and only if $\sigma_\alpha^2 = 0$. Finally, if $b_i = b$ so that the imbalance occurs only at the last stage, a method proposed by Khuri (1990) can be used to construct a set of jointly independent sums of squares each having an exact chi-square distribution. Khuri and Ghosh (1990) have considered minimal sufficient statistics for the model in (15.1.1).

Remark: An outline of the proof that SS_A and SS_B are dependent and do not have a chi-square type distribution can be traced as follows. First note that the covariance structure of the observations in the model in (15.1.1) is

$$\mathrm{Cov}(y_{ijk}, y_{i'j'k'}) = \begin{cases} 0 & \text{if } i \neq i', \\ \sigma_\alpha^2 & \text{if } i = i',\quad j \neq j',\quad k \neq k', \\ \sigma_\alpha^2 + \sigma_\beta^2 & \text{if } i = i',\quad j = j',\quad k \neq k', \\ \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 & \text{if } i = i',\quad j = j',\quad k = k'. \end{cases}$$

[1] In this restricted case, MS_E is independent of $\bar{y}_{...}$ but MS_A and MS_B are not independent of $\bar{y}_{...}$.

Thus, if we let $\boldsymbol{Y}$ denote the vector of observations, i.e.,

$$\boldsymbol{Y}' = (y_{111}, y_{112}, \ldots, y_{11n_{11}}; \ldots; y_{ab_a1}, y_{ab_a2}, \ldots, y_{ab_an_{ab_a}}),$$

then $\boldsymbol{Y} \sim N(\mu\mathbf{1}, \boldsymbol{V})$ where $\mathbf{1}$ is the N-vector with each element equal to unity and $\boldsymbol{V}$ is a nondiagonal matrix. From Theorem 9.3.5, $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ and $\boldsymbol{Y}'\boldsymbol{B}\boldsymbol{Y}$ are independent if and only if $\boldsymbol{A}\boldsymbol{V}\boldsymbol{B} = \mathbf{0}$. Similarly, from Theorem 9.3.6, $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} \sim \chi^2[v]$ if and only if $\boldsymbol{A}\boldsymbol{V}$ is an idempotent matrix of rank v. A computation of $\boldsymbol{A}\boldsymbol{V}\boldsymbol{B}$ and $\boldsymbol{A}\boldsymbol{V}$ for some simple cases reveals that they do not satisfy the conditions of Theorems 9.3.5 and 9.3.6. In particular, it is readily verified that MS_E is independent of MS_A and MS_B and has a chi-square type distribution. For a rigorous proof, see Scheffé (1959, pp. 255–258). ♦

15.5 UNWEIGHTED MEANS ANALYSIS

In the unweighted means analysis, mean squares are obtained using the unweighted means of the observations. In particular, let

$$\bar{y}^*_{ij.} = \sum_{k=1}^{n_{ij}} y_{ijk}/n_{ij}, \qquad \bar{y}^*_{i..} = \sum_{j=1}^{b_i} \bar{y}^*_{ij.}/b_i,$$

and

$$\bar{y}^*_{...} = \sum_{i=1}^{a} \bar{y}^*_{i..}/a.$$

Then the unweighted sums of squares are defined as follows:

$$\text{SS}_{Au} = r^*_3 \sum_{i=1}^{a} (\bar{y}^*_{i..} - \bar{y}^*_{...})^2,$$

$$\text{SS}_{Bu} = r^*_1 \sum_{i=1}^{a}\sum_{j=1}^{b_i} (\bar{y}^*_{ij.} - \bar{y}^*_{i..})^2, \tag{15.5.1}$$

and

$$\text{SS}_E = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}^*_{ij.})^2,$$

where

$$r^*_1 = \left[\frac{1}{b_. - a}\sum_{i=1}^{a} \bar{n}_i^{-1}(b_i - 1)\right]^{-1}, \qquad r^*_3 = \left[\frac{1}{a}\sum_{i=1}^{a} \bar{n}_i^{-1} b_i^{-1}\right]^{-1},$$

TABLE 15.2 Analysis of variance using unweighted means analysis for the model in (15.1.1).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Factor *A*	$a-1$	SS_{Au}	MS_{Au}	$\sigma_e^2+r_2^*\sigma_\beta^2+r_3^*\sigma_\alpha^2$
Factor *B* within *A*	$b_.-1$	SS_{Bu}	MS_{Bu}	$\sigma_e^2+r_1^*\sigma_\beta^2$
Error	$N-b_.$	SS_E	MS_E	σ_e^2

with

$$\bar{n}_i = \left[\frac{1}{b_i}\sum_{j=1}^{b_i} n_{ij}^{-1}\right]^{-1}.$$

Note that $\bar{n}_i$ represents the harmonic mean of the n_{ij} values at the ith level of factor A. In addition, note that the definition of SS_E is the same as in the Type I sums of square.

The mean squares are obtained by dividing the sums of squares by the corresponding degrees of freedom. The results on expected values of unweighted mean squares are obtained as follows:

$$E(MS_{Au}) = \sigma_e^2 + r_2^*\sigma_\beta^2 + r_3^*\sigma_\alpha^2,$$
$$E(MS_{Bu}) = \sigma_e^2 + r_1^*\sigma_\beta^2, \tag{15.5.2}$$

and

$$E(MS_E) = \sigma_e^2,$$

where r_1^* and r_3^* are defined in (15.5.1) and

$$r_2^* = \left[\frac{1}{\sum_{i=1}^a b_i^{-1}}\sum_{i=1}^a \bar{n}_i^{-1}b_i^{-1}\right]^{-1}.$$

Notice that, with $n_{ij} = n$, $b_i = b$; SS_{Au}, SS_{Bu}, and SS_E reduce to the sums of squares in the corresponding balanced case; $r_1^* = r_2^* = n$, and $r_3^* = bn$. For a detailed derivation of the results on expected mean squares, the reader is referred to Sen (1988). The analysis of variance table for the unweighted means analysis is shown in Table 15.3.

15.6 ESTIMATION OF VARIANCE COMPONENTS

In this section, we consider some methods of estimation of variance components σ_e^2, σ_β^2, and σ_α^2.

15.6.1 ANALYSIS OF VARIANCE ESTIMATORS

The analysis of variance estimators of variance components are obtained by equating each sum of squares or equivalently the mean square in the analysis of variance Table 15.1 to its expected value. Denoting the estimators in question as $\hat{\sigma}^2_{\alpha,\text{ANOV}}$, $\hat{\sigma}^2_{\beta,\text{ANOV}}$, and $\hat{\sigma}^2_{e,\text{ANOV}}$, the resulting equations are

$$\text{MS}_A = \hat{\sigma}^2_{e,\text{ANOV}} + r_2\hat{\sigma}^2_{\beta,\text{ANOV}} + r_3\hat{\sigma}^2_{\alpha,\text{ANOV}},$$
$$\text{MS}_B = \hat{\sigma}^2_{e,\text{ANOV}} + r_1\hat{\sigma}^2_{\beta,\text{ANOV}}, \tag{15.6.1}$$

and

$$\text{MS}_E = \hat{\sigma}^2_{e,\text{ANOV}}.$$

Solving the equations in (15.6.1) we obtain the following estimators:

$$\hat{\sigma}^2_{e,\text{ANOV}} = \text{MS}_E,$$
$$\hat{\sigma}^2_{\beta,\text{ANOV}} = \frac{1}{r_1}(\text{MS}_B - \text{MS}_E), \tag{15.6.2}$$

and

$$\hat{\sigma}^2_{\alpha,\text{ANOV}} = \frac{1}{r_2}(\text{MS}_A - r_2\hat{\sigma}^2_{\beta,\text{ANOV}} - \hat{\sigma}^2_{e,\text{ANOV}}).$$

The estimator $\hat{\sigma}^2_{e,\text{ANOV}}$ is the minimum variance unbiased estimator under the assumption of normality, but other estimators lack any optimal property other than unbiasedness.

15.6.2 UNWEIGHTED MEANS ESTIMATORS

The unweighted means estimators are obtained by equating the unweighted mean squares in Table 15.2 to their corresponding expected values. Denoting the estimators as $\hat{\sigma}^2_{e,\text{UME}}$, $\hat{\sigma}^2_{\beta,\text{UME}}$, and $\hat{\sigma}^2_{\alpha,\text{UME}}$, the resulting equations are

$$\text{MS}_{Au} = \hat{\sigma}^2_{e,\text{UME}} + r_2^*\hat{\sigma}^2_{\beta,\text{UME}} + r_3^*\hat{\sigma}^2_{\alpha,\text{UME}},$$
$$\text{MS}_{Bu} = \hat{\sigma}^2_{e,\text{UME}} + r_1^*\hat{\sigma}^2_{\beta,\text{UME}}, \tag{15.6.3}$$

and

$$\text{MS}_E = \hat{\sigma}^2_{e,\text{UME}}.$$

Solving the equations in (15.6.3), we obtain the following estimators:

$$\begin{aligned}\hat{\sigma}^2_{e,\text{UME}} &= \text{MS}_E,\\ \hat{\sigma}^2_{\beta,\text{UME}} &= \frac{1}{r_1^*}(\text{MS}_{Bu} - \text{MS}_E),\end{aligned} \tag{15.6.4}$$

and

$$\hat{\sigma}^2_{\alpha,\text{UME}} = \frac{1}{r_2^*}(\text{MS}_A - r_2^*\hat{\sigma}^2_{\beta,\text{UME}} - \text{MS}_E).$$

Note that the ANOVA and the unweighted means estimators for the error variance component are the same.

15.6.3 SYMMETRIC SUMS ESTIMATORS

For symmetric sums estimators we consider expected values for products and squares of differences of observations. From the model in (15.1.1), the expected values of products of the observations are

$$E(y_{ijk}y_{i'j'k'}) = \begin{cases} \mu^2, & i \neq i', \\ \mu^2 + \sigma^2_\alpha, & i = i', \quad j \neq j', \\ \mu^2 + \sigma^2_\alpha + \sigma^2_\beta, & i = i', \quad j = j', \quad k \neq k', \\ \mu^2 + \sigma^2_\alpha + \sigma^2_\beta + \sigma^2_e, & i = i', \quad j = j', \quad k = k', \end{cases} \tag{15.6.5}$$

where $i', i' = 1, 2, \ldots, a$; $j = 1, 2, \ldots, b_i$; $j' = 1, 2, \ldots, b_{i'}$; $k = 1, 2, \ldots, n_{ij}$; $k' = 1, 2, \ldots, n_{i'j'}$. Now, the normalized symmetric sums of the terms in (15.6.5) are

$$g_m = \frac{\sum_{\substack{i,i'\\ i\neq i'}} y_{i..}y_{i'..}}{\sum_{i=1}^a n_{i.}(N - n_{i.})} = \frac{\left(y^2_{...} - \sum_{i=1}^a y^2_{i..}\right)}{N^2 - k_2},$$

$$g_A = \frac{\sum_{i=1}^a \sum_{\substack{j,j'\\ j\neq j'}} y_{ij.}y_{ij'.}}{\sum_{i=1}^a \sum_{j=1}^{b_i} n_{ij}(n_{i.} - n_{ij})} = \frac{\sum_{i=1}^a \left(y^2_{i..} - \sum_{j=1}^{b_i} y^2_{ij.}\right)}{k_2 - k_1},$$

$$g_B = \frac{\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{\substack{k,k'\\ k\neq k'}} y_{ijk}y_{ijk'}}{\sum_{i=1}^a \sum_{j=1}^{b_i} n_{ij}(n_{ij} - 1)} = \frac{\sum_{i=1}^a \sum_{j=1}^{b_i} \left(y^2_{ij.} - \sum_{k=1}^{n_{ij}} y^2_{ijk}\right)}{k_1 - N},$$

and

$$g_E = \frac{\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{n_{ij}} y_{ijk}y_{ijk}}{\sum_{i=1}^a \sum_{j=1}^{b_i} n_{ij}} = \frac{\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{n_{ij}} y^2_{ijk}}{N},$$

where

$$n_{i.} = \sum_{j=1}^{b_i} n_{ij}, \quad N = \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij}, \quad k_1 = \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij}^2, \quad k_2 = \sum_{i=1}^{a} n_{i.}^2.$$

Equating g_m, g_A, g_B, and g_E to their respective expected values, we obtain

$$\begin{aligned} \mu^2 &= g_m, \\ \mu^2 + \sigma_\alpha^2 &= g_A, \\ \mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 &= g_B, \end{aligned} \tag{15.6.6}$$

and

$$\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_e^2 = g_E.$$

The variance component estimator obtained by solving the equations in (15.6.6) are (Koch, 1967)

$$\begin{aligned} \hat{\sigma}^2_{\alpha,\text{SSP}} &= g_A - g_m, \\ \hat{\sigma}^2_{\beta,\text{SSP}} &= g_B - g_A, \end{aligned} \tag{15.6.7}$$

and

$$\hat{\sigma}^2_{e,\text{SSP}} = g_E - g_B.$$

The estimators in (15.6.7) are, by construction, unbiased, and they reduce to the analysis of variance estimators in the case of balanced data. However, they are not translation invariant, i.e., they may change in values if the same constant is added to all the observations and their variances are functions of μ. This drawback is overcome by using the symmetric sums of squares of differences rather than products.

From the model in (15.1.1), the expected values of the squares of differences of the observations are

$$E[(y_{ijk} - y_{i'j'k'})^2] = \begin{cases} 2\sigma_e^2, & i = i', \quad j = j', \quad k \neq k', \\ 2(\sigma_e^2 + \sigma_\beta^2), & i = i', \quad j \neq j', \\ 2(\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2), & i \neq i', \end{cases} \tag{15.6.8}$$

where $i, i' = 1, \ldots, a$; $j = 1, \ldots, b_i$; $j' = 1, \ldots, b_{i'}$; $k = 1, \ldots, n_{ij}$; $k' = 1, \ldots, n_{i'j'}$. Now, we estimate $2\sigma_e^2$, $2(\sigma_e^2 + \sigma_\beta^2)$, and $2(\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2)$ by taking the normalized symmetric sums of their respective unbiased estimators in (15.6.8), i.e.,

$$h_E = \frac{\sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{\substack{k,k' \\ k \neq k'}} (y_{ijk} - y_{ijk'})^2}{\sum_{i=1}^{a} \sum_{j=1}^{b_i} n_{ij}(n_{ij} - 1)}$$

$$= \frac{2}{(k_1 - N)} \sum_{i=1}^{a} \sum_{j=1}^{b_i} n_{ij} \left\{ \sum_{k=1}^{n_{ij}} y_{ijk}^2 - n_{ij} \bar{y}_{ij.}^2 \right\},$$

$$h_B = \frac{\sum_{i=1}^{a} \sum_{\substack{j,j' \\ j \neq j'}} \sum_{k,k'} (y_{ijk} - y_{ij'k'})^2}{\sum_{i=1}^{a} \sum_{j=1}^{b_i} n_{ij}(n_{i.} - n_{ij})}$$

$$= \frac{2}{(k_2 - k_1)} \sum_{i=1}^{a} \sum_{j=1}^{b_i} (n_{i.} - n_{ij}) \sum_{k=1}^{n_{ij}} y_{ijk}^2 - 2g_A,$$

and

$$h_A = \frac{\sum_{\substack{i,i' \\ i \neq i'}}^{a} \sum_{j,j'} \sum_{k,k'} (y_{ijk} - y_{i'j'k'})^2}{\sum_{i=1}^{a} n_{i.}(N - n_{i.})}$$

$$= \frac{2}{(N^2 - k_2)} \sum_{i=1}^{a} (N - n_{i.}) \sum_{j=1}^{b_i} \sum_{k=1}^{n_{ij}} y_{ijk}^2 - 2g_m,$$

where $n_{i.}$, N, k_1, k_2, g_m, and g_A are defined as before.

Now, the estimators of the variance components are obtained by setting h_E, h_B, and h_A to their respective expected values. Denoting $\hat{\sigma}_{\alpha,\text{SSE}}^2$, $\hat{\sigma}_{\beta,\text{SSE}}^2$, and $\hat{\sigma}_{e,\text{SSE}}^2$ as the estimators in question, the equations are

$$\begin{aligned} 2\hat{\sigma}_{e,\text{SSE}}^2 &= h_E, \\ 2(\hat{\sigma}_{e,\text{SSE}}^2 + \hat{\sigma}_{\beta,\text{SSE}}^2) &= h_B, \end{aligned} \tag{15.6.9}$$

and

$$2(\hat{\sigma}_{e,\text{SSE}}^2 + \hat{\sigma}_{\beta,\text{SSE}}^2 + \hat{\sigma}_{\alpha,\text{SSE}}^2) = h_A.$$

The estimators obtained as solutions to (15.6.9) are (Koch, 1967)

$$\begin{aligned} \hat{\sigma}_{e,\text{SSE}}^2 &= \frac{1}{2} h_E, \\ \hat{\sigma}_{\beta,\text{SSE}}^2 &= \frac{1}{2}(h_B - h_E), \end{aligned} \tag{15.6.10}$$

and

$$\hat{\sigma}_{\alpha,\text{SSE}}^2 = \frac{1}{2}(h_A - h_B).$$

It can be readily verified that for balanced data, the estimators in (15.6.10) reduce to the usual analysis of variance estimators.

15.6.4 OTHER ESTIMATORS

The ML, REML, MINQUE, and MIVQUE estimators can be developed as special cases of the results for the general case considered in Chapter 10 and their special formulations for this model are not amenable to any simple algebraic expressions. With the advent of the high-speed digital computer, the general results on these estimators involving matrix operations can be handled with great speed and accuracy, and their explicit algebraic evaluation for this model seems to be rather unnecessary. In addition, some commonly used statistical software packages, such as SAS®, SPSS®, and BMDP®, have special routines to compute these estimates rather conveniently simply by specifying the model in question. Rao and Heckler (1997) discuss computational algorithms for the ML, REML, and MIVQUE estimators of the variance components. In addition, they consider a new noniterative procedure called the weighted analysis of means (WAM) estimator which utilizes prior information on the variance components. Sen (1988) and Sen et al. (1992) consider estimators for $\sigma_\alpha^2/(\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2)$ and $\sigma_\beta^2/(\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2)$.

15.6.5 A NUMERICAL EXAMPLE

Sokal and Rohlf (1995, pp. 294–295) reported data from an experiment designed to investigate variation in the blood pH of female mice. The experiment was carried out on 15 dams which were mated over a period of time with either two or three sires. Each sire was mated to different dams and measurements were made on the blood pH reading of a female offspring. The data are given in Table 15.3. We will use the two-way nested model in (15.1.1) to analyze the data in Table 15.3. Here, $i = 1, 2, \ldots, 5$ refer to the dams, $j = 1, 2, \ldots, b_i$ refer to the sires within dams, and $k = 1, 2, \ldots, n_{ij}$ refer to the blood pH readings of a female. Further, σ_α^2 and σ_β^2 designate variance components due to dam and sire within dam as factors, and σ_e^2 denotes the error variance component. The calculations leading to the conventional analysis of variance using Type I sums of squares are readily performed and the results are summarized in Table 15.4. The selected outputs using SAS® GLM, SPSS® GLM, and BMDP® 3V are displayed in Figure 15.1.

We now illustrate the calculations of point estimates of the variance components $\sigma_\alpha^2, \sigma_\beta^2, \sigma_e^2$ using methods described in this section.

The analysis of variance (ANOVA) estimates based on Henderson's Method I are obtained as the solution to the following simultaneous equations:

$$\begin{aligned} \sigma_e^2 &= 0.002474, \\ \sigma_e^2 + 4.2868\sigma_\beta^2 &= 0.003637, \\ \sigma_e^2 + 4.3760\sigma_\beta^2 + 10.6250\sigma_\alpha^2 &= 0.012716. \end{aligned}$$

Therefore, the desired ANOVA estimates of the variance components are

TABLE 15.3 Blood pH readings of female mice.

Dam	**1**		**2**		**3**			**4**		**5**		
Sire	**1**	2	**1**	**2**	**1**	**2**	**3**	**1**	2	**1**	**2**	**3**
pH Reading	7.48	7.48	7.45	7.50	7.40	7.45	7.40	7.38	7.37	7.44	7.49	7.48
	7.48	7.53	7.43	7.45	7.45	7.33	7.47	7.48	7.31	7.51		7.59
	7.52	7.43	7.49	7.43	7.42	7.40	7.40	7.46	7.45	7.49	7.49	7.59
	7.54	7.39	7.40	7.36	7.48	7.46	7.47		7.41	7.51	7.49	
			7.40				7.47			7.52	7.50	

Dam	**6**			**7**			**8**		**9**		**10**	
Sire	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**1**	**2**	**1**	**2**
pH Reading	7.54	7.44	7.43	7.41	7.47	7.53	7.52	7.40	7.40	7.42	7.39	7.50
	7.36	7.47	7.52	7.42	7.36	7.40	7.53	7.48	7.34	7.37	7.31	7.44
	7.36	7.48	7.50	7.36	7.43	7.44	7.48	7.50	7.37	7.46	7.30	7.40
	7.40	7.48	7.46	7.47	7.38	7.40		7.40	7.45	7.40	7.41	7.45
			7.39		7.41	7.45		7.51			7.48	

Dam	**11**		**12**		**13**			**14**			**15**		
Sire	**1**	**2**	**1**	**2**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
pH Reading	7.52	7.56	7.50	7.52	7.39	7.43	7.46	7.50	7.44	7.42	7.47	7.45	7.51
	7.54	7.39	7.45	7.43	7.37	7.38	7.44	7.53	7.45	7.48	7.49	7.42	7.51
	7.52	7.52	7.43	7.38	7.33	7.44	7.37	7.51	7.39	7.45	7.45	7.52	7.53
	7.56	7.49	7.44	7.33	7.43		7.54	7.43	7.52	7.51	7.43	7.51	7.45
	7.53	7.48	7.49		7.42					7.48	7.42	7.32	7.51

Source: Sokal and Rohlf (1995); used with permission.

TABLE 15.4 Analysis of variance for the blood pH reading data of Table 15.3.

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Dams	14	0.178017	0.012716	$\sigma_e^2 + 4.3760\sigma_\beta^2 + 10.6250\sigma_\alpha^2$
Sires within dams	22	0.080024	0.003637	$\sigma_e^2 + 4.2868\sigma_\beta^2$
Error	123	0.304253	0.002474	σ_e^2
Total	159	0.562294		

```
DATA SAHAIC15;
INPUT DAM SIRE PH;
CARDS;
 1 1 7.48
 1 1 7.48
 1 1 7.52
 1 1 7.54
 1 2 7.48
 1 2 7.53
 1 2 7.43
 1 2 7.39
 2 1 7.45
 2 1 7.43
 2 1 7.49
 2 1 7.40
 2 1 7.40
 2 2 7.50
 2 2 7.45
 2 2 7.43
 . . .
15 3 7.51
;
PROC GLM;
CLASS DAM SIRE;
MODEL PH = DAM SIRE(DAM);
RANDOM DAM SIRE(DAM)/TEST;
RUN;
CLASS LEVELS   VALUES
DAM    15   1 2 3  4 5 6 7
            8 9 10 11 12 13
            14 15
SIRE    3   1 2 3
NUMBER OF OBSERVATIONS IN
DATA SET=160
```

```
                          The SAS System
                 General Linear Models Procedure
Dependent Variable: PH
                              Sum of           Mean
Source         DF            Squares          Square        F Value     Pr > F
Model          36         0.25804104      0.00716781          2.90      0.0001
Error         123         0.30425333      0.00247360
Corrected     159         0.56229437
Total
          R-Square        C.V.        Root MSE          PH Mean
          0.458907     0.667605     0.04973534       7.44981250

Source        DF     Type I SS       Mean Square     F Value    Pr > F
DAM           14     0.17801736      0.01271553       5.14      0.0001
SIRE(DAM)     22     0.08002368      0.00363744       1.47      0.0966

Source        DF     Type III SS     Mean Square     F Value    Pr > F
DAM           14     0.17940454      0.01281461       5.18      0.0001
SIRE(DAM)     22     0.08002368      0.00363744       1.47      0.0966

Source                 Type III Expected Mean Square
DAM          Var(Error)+4.2391 Var(SIRE(DAM))+10.453 Var(DAM)
SIRE(DAM)    Var(Error)+4.2868 Var(SIRE(DAM))

     Tests of Hypotheses for Random Model Analysis of Variance
Source: DAM            Error: 0.9889*MS(SIRE(DAM)) + 0.0111*MS(Error)
                       Denominator Denominator
  DF    Type III MS       DF           MS        F Value    Pr > F
  14    0.0128146097     22.34     0.00362448    3.5356     0.0039

Source: SIRE(DAM)      Error: MS(Error)
                       Denominator Denominator
  DF    Type III MS       DF           MS        F Value    Pr > F
  22    0.0036374401      123      0.00247360    1.4705     0.0966
```

SAS application: This application illustrates SAS GLM instructions and output for the two-way unbalanced nested random effects analysis of variance.[a,b]

[a] Several portions of the output were extensively edited and doctored to economize space and may not correspond to the original printout.

[b] Results on significance tests may vary from one package to the other.

FIGURE 15.1 Program instructions and output for the two-way unbalanced nested random effects analysis of variance: Data on the blood pH readings of female mice (Table 15.3).

given by

$$\hat{\sigma}^2_{e,\text{ANOV}} = 0.002474,$$

$$\hat{\sigma}^2_{\beta,\text{ANOV}} = \frac{0.003637 - 0.002474}{4.2868} = 0.000271,$$

```
DATA SAHAIC15
/DAM 1 SIRE 3
 PH 5-6.
BEGIN DATA.
 1 1 7.48
 1 1 7.48
 1 1 7.52
 1 1 7.54
 1 2 7.48
 1 2 7.53
 1 2 7.43
 1 2 7.39
 2 1 7.45
  . . .
15 3 7.51
END DATA.
GLM PH BY DAM SIRE
/DESIGN DAM
 SIRE(DAM)
/METHOD SSTYPE(1)
/RANDOM DAM SIRE.
```

```
                        Tests of Between-Subjects Effects
Dependent Variable: PH

Source                    Type I SS      df       Mean Square     F       Sig.
DAM         Hypothesis     0.178         14        1.3E-02      3.473    0.005
            Error          7.8E-02       21.394    3.7E-03(a)
SIRE(DAM)   Hypothesis     8.0E-02       22        3.6E-03      1.471    0.097
            Error          0.304        123        2.5E-03(b)
a 1.021 MS(S(D))-2.081E-02 MS(E) b MS(Error)

                        Expected Mean Squares(c,d)
                           Variance Component
Source            Var(DAM)           Var(SIRE(DAM))          Var(Error)
DAM                10.625                4.376                  1.000
SIRE (DAM)          0.000                4.287                  1.000
Error               0.000                0.000                  1.000

c  For each source, the expected mean square equals the sum of the
   coefficients in the cells times the variance components, plus a quadratic
   term involving effects in the Quadratic Term cell.
d  Expected Mean Squares are based on the Type I Sums of Squares.
```

SPSS application: This application illustrates SPSS GLM instructions and output for the two-way unbalanced nested random effects analysis of variance.[a,b]

```
/INPUT    FILE='C:\SAHAIC15.TXT'.
          FORMAT=FREE.
          VARIABLES=3.
/VARIABLE NAMES= DAM,SIRE,PH.
/GROUP    CODES(SIRE)=1,2,3.
          NAMES(SIRE)=S1,S2,S3.
          CODES(DAM)=1,2,3,…,15.
          NAMES(DAM)=D1,D2,…,D15.
/DESIGN   DEPENDENT=PH.
          RANDOM=DAM.
          RANDOM=DAM, SIRE.
          RNAMES=D,'S(D)'.
          METHOD=REML.
/END
 1 1 7.48
 .  .  .
15 3 7.51
```

```
     BMDP3V - GENERAL MIXED MODEL ANALYSIS OF VARIANCE
              Release: 7.0     (BMDP/DYNAMIC)
        DEPENDENT VARIABLE PH

PARAMETER   ESTIMATE    STANDARD     EST/.    TWO-TAIL PROB.
                          ERROR      ST.DEV   (ASYM. THEORY)
ERR.VAR.     .002481    .000317
CONSTANT    7.449179    .009105     818.152       0.000
D            .000889    .000487
S(D)         .000265    .000267

    TESTS OF FIXED EFFECTS BASED ON ASYMPTOTIC VARIANCE
                   -COVARIANCE MATRIX
SOURCE       F-STATISTIC     DEGREES OF      PROBABILITY
                               FREEDOM
CONSTANT      669372.13        1   159         0.00000
```

BMDP application: This application illustrates BMDP 3V instructions and output for the two-way unbalanced nested random effects analysis of variance.[a,b]

[a] Several portions of the output were extensively edited and doctored to economize space and may not correspond to the original printout.

[b] Results on significance tests may vary from one package to the other.

FIGURE 15.1 (*continued*)

and

$$\hat{\sigma}^2_{\alpha,\text{ANOV}} = \frac{0.012716 - 0.002474 - 4.3760 \times 0.000271}{10.6250} = 0.000852.$$

These variance components account for 6.8%, 7.5%, and 23.7% of the total variation in the blood pH readings in this experiment.

To obtain variance components estimates based on unweighted means squares, we performed analysis of variance on the cell means and the results are summarized in Table 15.5. The analysis of means estimators are obtained as the solution to the following system of equations:

$$\sigma_e^2 = 0.002474,$$

TABLE 15.5 Analysis of variance for the blood pH reading data Table 15.3 using unweighted sums of squares.

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Dams	14	0.180040	0.012860	$\sigma_e^2 + 4.186\sigma_\beta^2 + 9.914\sigma_\alpha^2$
Sires within dams	22	0.080608	0.003664	$\sigma_e^2 + 4.222\sigma_\beta^2$
Error	123	0.304302	0.002474	σ_e^2
Total	159	0.564950		

$$\sigma_e^2 + 4.222\sigma_\beta^2 = 0.003664,$$

and

$$\sigma_e^2 + 4.186\sigma_\beta^2 + 9.914\sigma_\alpha^2 = 0.012860.$$

Therefore, the desired estimates are given by

$$\hat{\sigma}^2_{e,\text{UNME}} = 0.002474,$$

$$\hat{\sigma}^2_{\beta,\text{UNME}} = \frac{0.003664 - 0.002474}{4.222} = 0.000282,$$

and

$$\hat{\sigma}^2_{\alpha,\text{UNME}} = \frac{0.012860 - 0.002474 - 4.186 \times 0.000282}{9.914} = 0.000929.$$

We used SAS®VARCOMP, SPSS®VARCOMP, and BMDP®3V to estimate the variance components using the ML, REML, MINQUE(0), and MIVQUE(1) procedures.[2] The desired estimates using these software are given in Table 15.6. Note that all three software produce nearly same results except for some minor discrepancy in rounding decimal places.

15.7 VARIANCES OF ESTIMATORS

In this section, we present some results on sampling variances of the variance component estimators considered in the preceding section.

[2]The computations for ML and REML estimates were also carried out using SAS® PROC MIXED and some other programs to assess their relative accuracy and convergence rate. There did not seem to be any appreciable differences between the results from different software.

TABLE 15.6 ML, REML, MINQUE(0), and MINQUE(1) estimates of the variance components using SAS®, SPSS®, and BMDP® software.

Variance component	SAS® ML	SAS® REML	SAS® MINQUE(0)
σ_e^2	0.002481	0.002481	0.002514
σ_β^2	0.000265	0.000265	0.000308
σ_α^2	0.000805	0.000890	0.000772

Variance component	SPSS® ML	SPSS® REML	SPSS® MINQUE(0)	SPSS® MINQUE(1)
σ_e^2	0.002481	0.002481	0.002514	0.002475
σ_β^2	0.000265	0.000265	0.000308	0.000280
σ_α^2	0.000805	0.000890	0.000772	0.000897

Variance component	BMDP® ML	BMDP® REML
σ_e^2	0.002481	0.002481
σ_β^2	0.000265	0.000265
σ_α^2	0.000805	0.000890

SAS®VARCOMP does not compute MINQUE(1). MBDP®3V does not compute MINQUE(0) and MINQUE(1).

15.7.1 VARIANCES OF ANALYSIS OF VARIANCE ESTIMATORS

In the analysis of variance presented in Section 15.2, SS_E/σ_e^2 has a chi-square distribution with $N - b_.$ degrees of freedom. Hence, the variance of $\hat{\sigma}_e^2$ is

$$\text{Var}(\hat{\sigma}^2_{e,\text{ANOV}}) = \frac{2\sigma_e^4}{N - b_.}.$$

Furthermore, SS_E is distributed independently of SS_A and SS_B, so that the covariances of $\hat{\sigma}_\alpha^2$ and $\hat{\sigma}_\beta^2$ with $\hat{\sigma}_e^2$ can be obtained directly from (15.6.2). The expressions for variances and covariances have been developed by Searle (1961). The results are given as follows (see also Searle, 1971, p. 476; Searle et al., 1992, pp. 429–430):

$$\text{Var}(\hat{\sigma}^2_{\alpha,\text{ANOV}}) = \frac{2(\lambda_1\sigma_\alpha^4 + \lambda_2\sigma_\beta^4 + \lambda_3\sigma_e^4 + 2\lambda_4\sigma_\alpha^2\sigma_\beta^2)}{(N - k_1)^2(N - k_{12})^2}$$

$$+ \frac{4(\lambda_5\sigma_\alpha^2\sigma_e^2 + \lambda_6\sigma_\beta^2\sigma_e^2)}{(N-k_1)^2(N-k_{12})^2},$$

$$\mathrm{Var}(\hat{\sigma}^2_{\beta,\mathrm{ANOV}}) = \frac{2(k_7 + Nk_3 - 2k_5)\sigma_\beta^4 + 4(N-k_{12})\sigma_\beta^2\sigma_e^2}{(N-k_{12})^2} + \frac{2(N-b_.)^{-1}(b_. - a)(N-a)\sigma_e^4}{(N-k_{12})^2},$$

$$\mathrm{Cov}(\hat{\sigma}^2_{\alpha,\mathrm{ANOV}}, \hat{\sigma}^2_{e,\mathrm{ANOV}}) = \left[\frac{(k_{12}-k_3)(b_. - a)}{(N-k_{12})} - (a-1)\right]\left[\frac{\mathrm{Var}(\hat{\sigma}^2_{e,\mathrm{ANOV}})}{(N-k_1)}\right],$$

$$\mathrm{Cov}(\hat{\sigma}^2_{\beta,\mathrm{ANOV}}, \hat{\sigma}^2_{e,\mathrm{ANOV}}) = \frac{-(b_. - a)\,\mathrm{Var}(\hat{\sigma}^2_{e,\mathrm{ANOV}})}{(N-k_{12})},$$

$$\mathrm{Cov}(\hat{\sigma}^2_{\alpha,\mathrm{ANOV}}, \hat{\sigma}^2_{\beta,\mathrm{ANOV}}) = \frac{2\left[k_5 - k_7 + \frac{(k_6-k_4)}{N}\right]\sigma_\beta^4 + 2\left[\frac{(a-1)(b_.-a)\sigma_e^4}{(N-b_.)}\right]}{(N-k_1)(N-k_{12})} - \frac{(N-k_{12})(k_{12}-k_3)\,\mathrm{Var}(\hat{\sigma}^2_{\beta,\mathrm{ANOV}})}{(N-k_1)(N-k_{12})},$$

where

$$k_1 = \frac{1}{N}\sum_{i=1}^a n_{i.}^2, \qquad k_3 = \frac{1}{N}\sum_{i=1}^a\sum_{j=1}^{b_i} n_{ij}^2, \qquad k_{12} = \sum_{i=1}^a \frac{\left(\sum_{j=1}^{b_i} n_{ij}^2\right)}{n_{i.}},$$

$$k_4 = \sum_{i=1}^a\sum_{j=1}^{b_i} n_{ij}^3, \qquad k_5 = \sum_{i=1}^a\left(\frac{\sum_j^{b_i} n_{ij}^3}{n_{i.}}\right), \qquad k_6 = \sum_{i=1}^a \frac{\left(\sum_{j=1}^{b_i} n_{ij}^2\right)^2}{n_{i.}},$$

$$k_7 = \sum_{i=1}^a \frac{\left(\sum_{j=1}^{b_i} n_{ij}^2\right)^2}{n_{i.}^2}, \quad k_8 = \sum_{i=1}^a n_{i.}\left(\sum_{j=1}^{b_i} n_{ij}^2\right), \quad k_9 = \sum_{i=1}^a n_{i.}^3,$$

$$\lambda_1 = (N-k_{12})^2\left[k_1(N+k_1) - \frac{2k_9}{N}\right],$$

$$\lambda_2 = k_3[N(k_{12}-k_3)^2 + k_3(N-k_{12})^2] + (N-k_3)^2k_7 - 2(N-k_3)\left[(k_{12}-k_3)k_5 + \frac{(N-k_{12})k_6}{N}\right] + \frac{2(N-k_{12})(k_{12}-k_3)k_4}{N},$$

$$\lambda_3 = \frac{(N-k_{12})^2(N-1)(a-1) - (N-k_3)^2(a-1)(b_. - a)}{N - b_.} + \frac{(k_{12}-k_3)^2(N-1)(b_. - a)}{N - b_.},$$

$$\lambda_4 = (N-k_{12})^2\left[k_3(N+k_1) - \frac{2k_8}{N}\right],$$

$$\lambda_5 = (N-k_{12})^2(N-k_1),$$

and

$$\lambda_6 = (N - k_{12})(N - k_3)(k_{12} - k_3).$$

15.7.2 LARGE SAMPLE VARIANCES OF MAXIMUM LIKELIHOOD ESTIMATORS

The explicit expressions for large sample variances of maximum likelihood estimators of the variance components σ_α^2, σ_β^2, and σ_e^2 have been developed by Searle (1970) using the general results of Section 10.7.2. The result on variance-covariance matrix of the vector of the maximum likelihood estimators of $(\sigma_\alpha^2, \sigma_\beta^2, \sigma_e^2)$ is given by (see also Searle, 1971, p. 477; Searle et al., 1992, pp. 430–431):

$$\text{Var}\begin{bmatrix}\hat{\sigma}^2_{\alpha,\text{ML}}\\ \hat{\sigma}^2_{\beta,\text{ML}}\\ \hat{\sigma}^2_{e,\text{ML}}\end{bmatrix} = 2\begin{bmatrix} t_{\alpha\alpha} & t_{\alpha\beta} & t_{\alpha e}\\ t_{\alpha\beta} & t_{\beta\beta} & t_{\beta e}\\ t_{\alpha e} & t_{\beta e} & t_{ee}\end{bmatrix}^{-1},$$

where

$$t_{\alpha\alpha} = \sum_{i=1}^{a}\frac{A_{i11}^2}{q_i^2}, \qquad t_{\alpha\beta} = \sum_{i=1}^{a}\frac{A_{i22}}{q_i^2}, \qquad t_{\alpha e} = \sum_{i=1}^{a}\frac{A_{i12}}{q_i^2},$$

$$t_{\beta\beta} = \sum_{i=1}^{a}\left(A_{i22} - \frac{2\sigma_\alpha^2 A_{i33}}{q_i} + \frac{\sigma_\alpha^4 A_{i22}^2}{q_i^2}\right),$$

$$t_{\beta e} = \sum_{i=1}^{a}\left(A_{i12} - \frac{2\sigma_\alpha^2 A_{i23}}{q_i} + \frac{\sigma_\alpha^4 A_{i12}A_{i22}}{q_i^2}\right),$$

$$t_{ee} = \sum_{i=1}^{a}\left(A_{i02} - \frac{2\sigma_\alpha^2 A_{i13}}{q_i} + \frac{\sigma_\alpha^4 A_{i12}^2}{q_i^2}\right) + \frac{(N - b.)}{\sigma_e^4},$$

$$A_{ipq} = \sum_{j=1}^{b_i}\left(\frac{n_{ij}^p}{m_{ij}^q}\right) \quad \text{for integers } p \text{ and } q,$$

$$q_i = 1 + \sigma_\alpha^2 A_{i11}, \quad \text{and} \quad m_{ij} = \sigma_e^2 + n_{ij}\sigma_\beta^2.$$

15.8 COMPARISONS OF DESIGNS AND ESTIMATORS

There have been a number of studies to investigate the performance of variance component estimators using planned unbalanced in two-way nested designs. Some of the earlier authors who have considered the problem include Anderson and Bancroft (1952), Prairie (1962), Prairie and Anderson (1962), Bainbridge (1965), Goldsmith (1969), and Goldsmith and Gaylor (1970). The Bainbridge design consists of a replications of the basic design shown in Figure 15.2. For

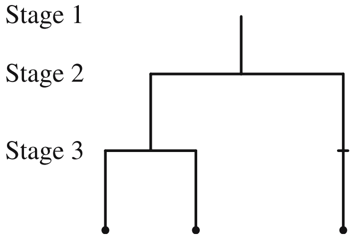


FIGURE 15.2 Three-stage Bainbridge design.

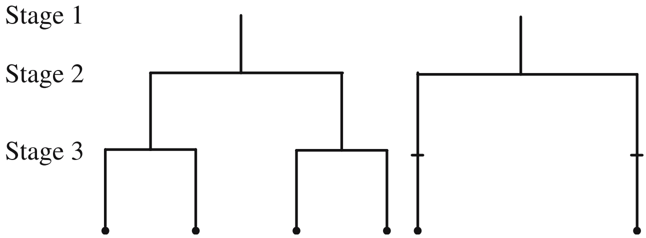


FIGURE 15.3 Three-stage Anderson design.

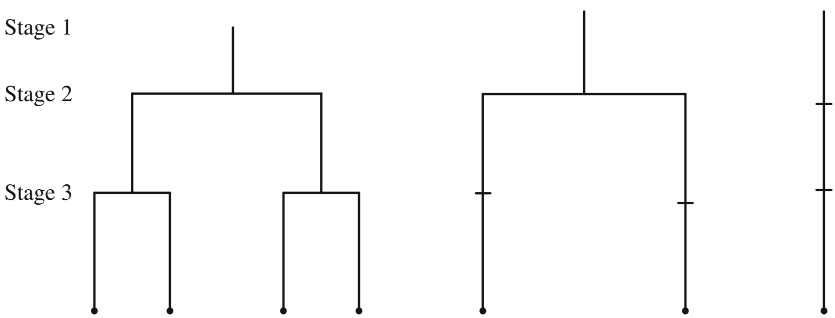
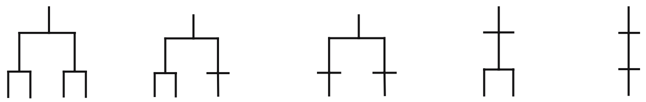


FIGURE 15.4 Three-stage Prairie–Anderson design.

each replication, this design has two second-stage units; two samples are drawn from the first one and one sample is drawn from the other. Hence, there are a first stage units, $2a$ second stage units, involving a total of $N = 3a$ samples. The Anderson design consists of the basic design shown in Figure 15.3. To make the size of the design comparable to that of Bainbridge's, one would require $a/2$ replications of the basic design. The Bainbridge design has one distinct advantage, that it is amenable to simple ANOVA estimation procedure. The Anderson design requires an unrealistic pooling of ANOVA sums of squares. The Prairie–Anderson design consists of a replications of the basic design shown in Figure 15.4. For each replication, there are a_1 first-stage units of the first type (in the left), a_2 of the second type (in the middle), and a_3 of the third

Structure number

Stage	I	II	III	IV	V
1					
2					
3					



	I	II	III	IV	V
n_{ij}	2	2,1	1	2	1
$n_{i.}$	4	3	2	2	1
b_i	2	2	2	1	1

Source: Goldsmith and Gaylor (1970); used with permission.

FIGURE 15.5 Basic structure of Goldsmith–Gaylor designs.

type (on the right) ($a_1 + a_2 + a_3 = a$). Hence, there are a first-stage units, $2a_1 + 2a_2 + a_3$ second-stage units, involving a total of $N = 4a_1 + 2a_2 + a_3$ samples. For a clear and concise review of these designs, see Anderson (1975, 1981).

Goldsmith (1969) and Goldsmith and Gaylor (1970) studied 61 connected designs. Each design comprised no more than three of the five structures of design shown in Figure 15.5. In the figure shown, b_i is the number of second-stage classes and $n_{i.} = \sum_{j=1}^{b_i} n_{ij}$. Each of the designs contains $12r$ observations, $r = 1(1)10$, where $\sum_{i=1}^{a} n_{i.} = 12$. For example, a basic design may have three replications of structure I (balanced design); four replications of structure II (a Bainbridge staggered design); and two replications of structure III (an Anderson design). Prairie (1962) and Prairie and Anderson (1962) used the following combination to generate 48 observations ($r = 4$), where the number in parentheses refers to the number of replications of the given structure: II(16), III(16), IV(8), IV*(8). The design IV* had $n_{i.} = n_{ij} = 3$.

Goldsmith and Gaylor (1970) carried out an extensive investigation of designs for estimation of the variance components. In order to find an optimal design, one can either state the conclusions in terms of each variance component or work with functions of the variance component estimators. Goldsmith and Gaylor chose the latter approach and considered several functions of the variance-covariance matrix (Σ) of the variance component estimators. The criteria used for optimality included tr[Σ], $|\Sigma|$, and adjusted tr[Σ], where the sampling variances of the variance component estimates are scaled according to the magnitude of variance components. For each criterion, optimal unbalanced designs were identified and tabulated. No single design was found to be the "best" for all the situations and the conclusions varied depending on the relative magnitudes of the population variance components and the optimality criterion used. In general, however, the Bainbridge staggered design fared well, although it was not optimal. The authors remarked that the trace criterion was probably the best since it tended to concentrate the sampling at a stage with a large variance component relative to the others. The determinant criterion was found to be the worst since it was relatively insensitive to changes in the sample size and values of the variance components. Furthermore, when the variance components for the first two stages were small compared to the error variance, the balanced design was considered to be optimum since it tended to concentrate the sampling at the last stage. Finally, if any one stage has a large variance component, then designs yielding the highest degrees of freedom are preferred.

In a later work, Schwartz (1980) used large sample asymptotic results as well as Monte Carlo simulation for some unbalanced designs from Bainbridge (1965) and Anderson (1975). Prairie, Bainbridge, and Goldsmith considered only ANOVA estimators, whereas Schwartz used iterated least squares (ITLS) and ML estimators. The ITLS estimator is based on the sums of squares for each component of the model. It makes use of weighted regression where a sum of squares is the dependent variable and the coefficients of the variance components in the expected sum of squares are taken as independent variables. The weights are determined via the variance-covariance matrix of the sums of squares and are functions of the variance components. Inasmuch as the variance components are usually unknown, one proceeds with initial approximate values of the variance components and makes repeated use of the variance components from the previous iteration. The process is continued until it converges yielding the ITLS estimates. If the overall mean is estimated from the same data, then ITLS and ML estimates are equivalent. Schwartz conducted 5000 simulated experiments for each parameter set to compare the performance of the Bainbridge, Anderson, and balanced designs in estimating the variance components. He found that there is very little difference between the Bainbridge and Anderson designs; the balanced design is best when σ_α^2 is small and least efficient when σ_α^2 is large. In addition, it was found that for large σ_α^2 asymptotic and simulation results are quite similar for all the designs included in his study. However, when σ_α^2 is smaller than σ_e^2, asymptotic variances are much larger than the simulated MSEs, especially for balanced designs. For such situations, he recommended a balanced design. When σ_α^2 is much larger than σ_β^2, he recommended using as many first-stage units as possible.

More recently, Rao and Heckler (1997) have compared the ANOVA, ML, REML, MIVQUE, and WAM estimators of the variance components, for some selected unbalanced designs and $(\sigma_\alpha^2, \sigma_\beta^2, \sigma_e^2) = (10, 1, 1), (1, 10, 1), (10, 10, 1)$, and $(1, 1, 1)$ using exact and empirical results. It was found that the biases and MSEs of the MIVQUE, REML, and WAM estimators of σ_α^2 and σ_β^2 are comparable. The ML estimators of σ_α^2 and σ_β^2 in general have smaller MSEs than the remaining four estimators, but they entail considerably greater absolute biases; however, the biases and MSEs of all the five estimators of σ_e^2 are comparable. They also evaluated the probability of negativity of these estimates for the design with $a = 4$; $b_i = 2, 3, 4, 5$, and $c_{ij} = (6, 6), (4, 4, 4), (3, 3, 3, 3), (2, 2, 2, 2, 2)$; and $(\sigma_\alpha^2, \sigma_\beta^2, \sigma_e^2) = (10, 1, 1), (1, 10, 1)$. The results indicate that for the ANOVA, MIVQUE, and WAM estimates, the chances of an estimate of σ_β^2 assuming a negative value are small if $\sigma_\beta^2/\sigma_e^2$ and c_{ij} are large;

similarly, the chances of an estimate of σ_α^2 assuming a negative value are small if $(\sigma_\alpha^2/\sigma_\beta^2, \sigma_\alpha^2/\sigma_e^2)$ as well as (b_i, c_{ij}) are large.

15.9 CONFIDENCE INTERVALS

In this section, we briefly discuss some results on confidence intervals for variance components $\sigma_\alpha^2, \sigma_\beta^2, \sigma_e^2$, and certain of their parametric functions.

15.9.1 CONFIDENCE INTERVAL FOR σ_e^2

From the result in (15.4.1) on distribution theory of MS_E, an exact $100(1-\alpha)\%$ normal theory confidence interval for σ_e^2 is given by

$$P\left\{\frac{v_e \text{MS}_E}{\chi^2[v_e, 1-\alpha/2]} \leq \sigma_e^2 \leq \frac{v_e \text{MS}_E}{\chi^2[v_e, \alpha/2]}\right\} = 1-\alpha, \qquad (15.9.1)$$

where $\chi^2[v_e, \alpha/2]$ and $\chi^2[v_e, 1-\alpha/2]$ denote the lower- and upper-tail $\alpha/2$-level critical values of the $\chi^2[v_e]$ distribution.

15.9.2 CONFIDENCE INTERVALS FOR σ_β^2 AND σ_α^2

Hernández et al. (1992) have proposed constructing confidence intervals on σ_β^2 and σ_α^2 based on unweighted means squares by using an approach similar to that of balanced designs considered in Section 6.8.2. They report that although unweighted means squares are not jointly independent and do not have scaled chi-square distributions, the procedure seems to perform well for most of the unbalanced designs encountered in practice, The resulting $100(1-\alpha)\%$ confidence interval for σ_β^2 is given by

$$P\left\{\frac{1}{r_1^*}\left\{\text{MS}_{Bu} - \text{MS}_E - \sqrt{L_\beta}\right\} \leq \sigma_\beta^2 \leq \frac{1}{r_1^*}\left\{\text{MS}_{Bu} - \text{MS}_E + \sqrt{U_\beta}\right\}\right\}$$
$$\doteq 1-\alpha, \qquad (15.9.2)$$

where

$$L_\beta = G_2^2 \text{MS}_{Bu}^2 + H_3^2 \text{MS}_E^2 + G_{23} \text{MS}_{Bu} \text{MS}_E,$$
$$U_\beta = H_2^2 \text{MS}_{Bu}^2 + G_3^2 \text{MS}_E^2 + H_{23} \text{MS}_{Bu} \text{MS}_E,$$

and G_2, G_3, H_2, H_3, G_{23}, and H_{23} are defined as in Section 6.8.2. It can be seen that when $b_i = b$ and $n_{ij} = n$, the interval in (15.9.2) reduces to the corresponding balanced formula given in Section 6.8.2. Hernández et al. (1992) note that although the interval (15.9.2) provides satisfactory coverage under most of the conditions, its coverage may drop below the stated level when

$\sigma_\beta^2/\sigma_e^2$ is small and the design is highly unbalanced. For these situations, they recommended an alternate interval given by

$$P\left\{\frac{L'_\beta \mathrm{MS}_{Bu}}{(1+r_1^* L'_\beta)F[v_\beta,\infty;1-\alpha/2]} \leq \sigma_\beta^2 \leq \frac{U'_\beta \mathrm{MS}_{Bu}}{(1+r_1^* U'_\beta)F[v_\beta,\infty;\alpha/2]}\right\} \doteq 1-\alpha, \tag{15.9.3}$$

where

$$L'_\beta = \frac{\mathrm{MS}_{Bu}}{r_1^* \mathrm{MS}_E F[v_\beta, v_e; 1-\alpha/2]} - \frac{1}{n_{\min}},$$
$$U'_\beta = \frac{\mathrm{MS}_{Bu}}{r_1^* \mathrm{MS}_E F[v_\beta, v_e; \alpha/2]} - \frac{1}{n_{\max}},$$
$$n_{\min} = \min(n_{ij}) \quad \text{and} \quad n_{\max} = \max(n_{ij}).$$

Note that both the lower and upper limits in (15.9.2) and (15.9.3) may assume negative values which are defined to be zero.

For σ_α^2, the resulting $100(1-\alpha)\%$ approximate confidence interval is given by

$$P\left\{\frac{1}{r_3^*}(\mathrm{MS}_{Au} - \ell_2 \mathrm{MS}_{Bu} + \ell_3 \mathrm{MS}_E - \sqrt{L_\alpha}) \leq \sigma_\alpha^2 \leq \frac{1}{r_3^*}(\mathrm{MS}_{Au} - \ell_2 \mathrm{MS}_{Bu} + \ell_3 \mathrm{MS}_E + \sqrt{U_\alpha})\right\} \doteq 1-\alpha, \tag{15.9.4}$$

where

$$L_\alpha = \begin{cases} G_1^2 \mathrm{MS}_{Au}^2 + \ell_3^2 G_3^2 \mathrm{MS}_E^2 + \ell_2^2 H_2^2 \mathrm{MS}_{Bu}^2 + \ell_2 G_{12} \mathrm{MS}_{Au} \mathrm{MS}_{Biu} \\ \quad + \ell_2 \ell_3 G_{32} \mathrm{MS}_{Bu} \mathrm{MS}_E + \ell_3 G_{13}^* \mathrm{MS}_{Au} \mathrm{MS}_E \quad \text{if } \ell_3 \geq 0, \\ G_1^2 \mathrm{MS}_{Au}^2 + \ell_2^2 H_2^2 \mathrm{MS}_{Bu}^2 + \ell_3^2 H_3^2 \mathrm{MS}_E^2 + \ell_2 G_{12} \mathrm{MS}_{Au} \mathrm{MS}_{Bu} \\ \quad + |\ell_3| G_{13} \mathrm{MS}_{Au} \mathrm{MS}_E \quad \text{if } \ell_3 < 0, \end{cases}$$

$$U_\alpha = \begin{cases} H_1^2 \mathrm{MS}_{Au}^2 + \ell_3^2 H_3^2 \mathrm{MS}_E^2 + \ell_2^2 G_2^2 \mathrm{MS}_{Bu}^2 + \ell_2 H_{12} \mathrm{MS}_{Au} \mathrm{MS}_{Bu} \\ \quad + \ell_2 \ell_3 H_{32} \mathrm{MS}_{Bu} \mathrm{MS}_E \quad \text{if } \ell_3 \geq 0, \\ H_1^2 \mathrm{MS}_{Au}^2 + \ell_2^2 G_2^2 \mathrm{MS}_{Bu}^2 + \ell_3^2 G_3^2 \mathrm{MS}_E^2 + \ell_2 H_{12} \mathrm{MS}_{Au} \mathrm{MS}_{Bu} \\ \quad + |\ell_3| H_{13} \mathrm{MS}_{Au} \mathrm{MS}_E + \ell_2 |\ell_3| H_{23}^* \mathrm{MS}_{Bu} \mathrm{MS}_E \quad \text{if } \ell_3 < 0, \end{cases}$$

$$\ell_2 = r_2^*/r_1^*, \qquad \ell_3 = \ell_2 - 1,$$
$$G_1 = 1 - F^{-1}[v_\alpha, \infty; 1-\alpha/2], \qquad G_2 = 1 - F^{-1}[v_\beta, \infty; 1-\alpha/2],$$
$$G_3 = 1 - F^{-1}[v_e, \infty; 1-\alpha/2],$$
$$H_1 = F^{-1}[v_\alpha, \infty; \alpha/2] - 1, \qquad H_2 = F^{-1}[v_\beta, \infty; \alpha/2] - 1,$$
$$H_3 = F^{-1}[v_e, \infty; \alpha/2] - 1,$$

$$
\begin{aligned}
G_{12} &= F^{-1}[v_\alpha, v_\beta; 1-\alpha/2]\{(F[v_\alpha, v_\beta; 1-\alpha/2]-1)^2 \\
&\quad - G_1^2 F^2[v_\alpha, v_\beta; 1-\alpha/2] - H_2^2\}, \\
G_{13} &= F^{-1}[v_\alpha, v_e; 1-\alpha/2]\{(F[v_\alpha, v_e; 1-\alpha/2]-1)^2 \\
&\quad - G_1^2 F^2[v_\alpha, v_e; 1-\alpha/2] - H_3^2\}, \\
G_{32} &= F^{-1}[v_e, v_\beta; 1-\alpha/2]\{(F[v_e, v_\beta; 1-\alpha/2]-1)^2 \\
&\quad - G_3^2 F^2[v_e, v_\beta; 1-\alpha/2] - H_2^2\}, \\
G_{13}^* &= (1 - F^{-1}[v_\alpha + v_e, \infty; 1-\alpha/2])^2 \frac{(v_\alpha + v_e)^2}{v_\alpha v_e} - \frac{v_\alpha G_1^2}{v_e} - \frac{v_e G_3^2}{v_\alpha}, \\
H_{12} &= F^{-1}[v_\alpha, v_\beta; \alpha/2]\{(1 - F[v_\alpha, v_\beta; \alpha/2])^2 - H_1^2 F^2[v_\alpha, v_\beta; \alpha/2] - G_2^2\}, \\
H_{13} &= F^{-1}[v_\alpha, v_e; \alpha/2]\{(1 - F[v_\alpha, v_e; \alpha/2])^2 - H_1^2 F^2[v_\alpha, v_e; \alpha/2] - G_3^2\}, \\
H_{32} &= F^{-1}[v_e, v_\beta; \alpha/2]\{(1 - F[v_e, v_\beta; \alpha/2])^2 - H_3^2 F^2[v_e, v_\beta; \alpha/2] - G_2^2\},
\end{aligned}
$$

and

$$
H_{23}^* = (1 - F^{-1}[v_\beta + v_e, \infty; \alpha/2])^2 \frac{(v_\beta + v_e)^2}{v_\beta v_e} - \frac{v_\beta G_2^2}{v_e} - \frac{v_e G_3^2}{v_\beta}.
$$

Hernández et al. (1992) noted that the interval in (15.9.4) provides satisfactory coverage for a wide variety of unbalanced designs. In addition, for the special case of the design with $n_{ij} = n$, Burdick et al. (1986) performed some simulation work that seems to indicate that the interval performs well. However, when $\rho_\alpha = \sigma_\alpha^2/(\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2)$ is small and b_is differ greatly with $b_i = 1$ or $b_i = 2$ for some $i = 1, 2, \ldots, a$, the resultant intervals can be liberal. Furthermore, note that when $b_i = b$ and $n_{ij} = n$, the interval (15.9.4) reduces to the corresponding balanced interval (6.8.3). The intervals in (15.9.2) and (15.9.4) can also be based on Type I sums of squares considered in Section 15.2. Hernández et al. (1992) have investigated the performance of these intervals and report them to be slightly inferior in comparison to the intervals in (15.9.2) and (15.9.4). For a special case of the design when $n_{ij} = n$ for all i and j, confidence intervals on σ_β^2, $\sigma_\beta^2/\sigma_e^2$, and on parameters that involve only σ_e^2 and $\sigma_e^2 + n\sigma_\beta^2$ can be constructed using balanced design formulas of Section 6.8. For other parameters, an unweighted sum of squares estimator recommended by Burdick and Graybill (1985) can be substituted for MS_A in the balanced design intervals. Further, when $b_i = b$, thus only the last stage of the design is unbalanced, confidence intervals on σ_β^2 and σ_α^2 can be constructed using a method due to Khuri (1990). For a discussion of the pros and cons of this procedure, see Burdick and Graybill(1992, p. 103). Jain et al. (1991) have reported some additional results for confidence intervals on σ_β^2 and σ_α^2.

15.9.3 CONFIDENCE INTERVALS FOR $\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2$

Hernández and Burdick (1993) have proposed constructing a confidence interval for $\gamma = \sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2$ based on unweighted means squares by using an approach similar to that of balanced designs considered in Section 6.8.3. The resulting $100(1-\alpha)\%$ confidence interval is given by

$$\begin{aligned} P\Bigg\{ & \frac{1}{r_3^*}(\mathrm{MS}_{Au} + \ell_2\mathrm{MS}_{Bu} + \ell_3\mathrm{MS}_E - \sqrt{L_\gamma}) \\ & \leq \gamma \leq \frac{1}{r_3^*}(\mathrm{MS}_{Au} + \ell_2\mathrm{MS}_{Bu} + \ell_3\mathrm{MS}_E + \sqrt{U_\gamma})\Bigg\} \\ & \doteq 1-\alpha, \end{aligned} \tag{15.9.5}$$

where

$$\begin{aligned} \ell_2 &= r_1^{*-1}(r_3^* - r_2^*) \geq 0, \qquad \ell_3 = r_3^* - 1 - \ell_2 \geq 0, \\ L_\gamma &= G_1^2\mathrm{MS}_{Au}^2 + \ell_2^2G_2^2\mathrm{MS}_{Bu}^2 + \ell_3^2G_3^2\mathrm{MS}_E^2, \\ U_\gamma &= H_1^2\mathrm{MS}_{Au}^2 + \ell_2^2H_2^2\mathrm{MS}_{Bu}^2 + \ell_3^2H_3^2\mathrm{MS}_E^2, \end{aligned}$$

and G_i, H_i $(i = 1, 2, 3)$ are defined in (15.9.4).

Based on some simulation work, Hernández and Burdick (1993) report that although the unweighted means squares violate the assumptions of independence and chi-squaredness, the interval in (15.9.5) maintains its coverage close to the stated confidence level. Burdick and Graybill (1985) have also considered the problem of setting a confidence interval on γ for the special case of the design with equal subsampling or the last stage uniformity. They consider an approximation for the distribution of a sum of squares and use it to obtain an approximate confidence interval for γ. Confidence intervals on γ can also be based on *Type* I sums of squares.

15.9.4 CONFIDENCE INTERVALS ON $\sigma_\beta^2/\sigma_e^2$ AND $\sigma_\alpha^2/\sigma_e^2$

An exact confidence interval on $\sigma_\beta^2/\sigma_e^2$ can be constructed by using Wald's procedure as described in Section 11.8.2 and is illustrated in a paper by Verdooren (1976). Similarly, approximate procedures of Thomas and Hultquist (1978) and Burdick and Eickman (1986) based on unweighted means squares can also be used for this problem. However, as indicated earlier, Wald's procedure cannot be used to construct an exact interval on $\sigma_\alpha^2/\sigma_e^2$. An approximate interval on $\sigma_\alpha^2/\sigma_e^2$ can be based on unweighted means squares or Type I sums of squares similar to that of balanced designs considered in Section 6.7. Verdooren (1988) has proposed an exact interval on $\sigma_\alpha^2/\sigma_e^2$ for a known value of $\sigma_\beta^2/\sigma_e^2$.

15.9.5 CONFIDENCE INTERVALS ON $\sigma_\alpha^2/(\sigma_e^2+\sigma_\beta^2+\sigma_\alpha^2)$ AND $\sigma_\beta^2/(\sigma_e^2+\sigma_\beta^2+\sigma_\alpha^2)$

Sen (1988) and Sen et al. (1992) have proposed constructing confidence intervals on $\rho_\alpha = \sigma_\alpha^2/(\sigma_e^2+\sigma_\beta^2+\sigma_\alpha^2)$ and $\rho_\beta = \sigma_\beta^2/(\sigma_e^2+\sigma_\beta^2+\sigma_\alpha^2)$ based on unweighted means squares using an approach similar to that of balanced designs considered in Section 6.8.5. The resulting $100(1-\alpha)\%$ confidence interval for ρ_α is given by

$$P\left\{\frac{p_1\mathrm{MS}_{Au}-p_2\mathrm{MS}_{Bu}-p_3\mathrm{MS}_E}{p_4\mathrm{MS}_{Au}-p_5\mathrm{MS}_{Bu}-p_6\mathrm{MS}_E}\le\rho_\alpha\le\frac{p_1'\mathrm{MS}_{Au}-p_2'\mathrm{MS}_{Bu}-p_3'\mathrm{MS}_E}{p_4'\mathrm{MS}_{Au}-p_5'\mathrm{MS}_{Bu}-p_6'\mathrm{MS}_E}\right\}$$
$$\doteq 1-\alpha, \tag{15.9.6}$$

where

$$\begin{aligned}
&p_1 = r_1^*, \qquad p_2 = r_2^* F[v_\alpha, v_\beta; 1-\alpha/2],\\
&p_3 = (r_1^* - r_2^*)F[v_\alpha, v_e; 1-\alpha/2],\\
&p_4 = r_1^*, \qquad p_5 = (r_2^* - r_3^*)F[v_\alpha, v_\beta; 1-\alpha/2],\\
&p_6 = (r_1^* - r_2^* + r_3^* - r_1^* r_3^*)F[v_\alpha, v_e; 1-\alpha/2],\\
&p_1' = r_1^*, \qquad p_2' = r_2^* F[v_\alpha, v_\beta; \alpha/2], \qquad p_3' = (r_1^* - r_2^*)F[v_\alpha, v_e; \alpha/2],\\
&p_4' = r_1^*, \qquad p_5' = (r_2^* - r_3^*)F[v_\alpha, v_\beta; \alpha/2],
\end{aligned}$$

and

$$p_6' = (r_1^* - r_2^* + r_3^* - r_1^* r_3^*)F[v_\alpha, v_e; \alpha/2].$$

Similarly, an approximate $100(1-\alpha)\%$ confidence interval for ρ_β is given by

$$P\left\{\frac{r_3^* L_\beta}{r_1^* + (r_3^* - r_2^*)L_\beta} \le \rho_\beta \le \frac{r_3^* U_\beta}{r_1^* + (r_3^* - r_2^*)U_\beta}\right\} \doteq 1-\alpha, \tag{15.9.7}$$

where

$$L_\beta = \frac{\mathrm{MS}_{Bu}^2 - q_1\mathrm{MS}_{Bu}\mathrm{MS}_E - q_2\mathrm{MS}_E^2}{q_3\mathrm{MS}_{Au}\mathrm{MS}_{Bu} + q_4\mathrm{MS}_{Bu}\mathrm{MS}_E},$$
$$U_\beta = \frac{\mathrm{MS}_{Bu}^2 - q_1'\mathrm{MS}_{Bu}\mathrm{MS}_E - q_2'\mathrm{MS}_E^2}{q_3'\mathrm{MS}_{Au}\mathrm{MS}_{Bu} + q_4'\mathrm{MS}_{Bu}\mathrm{MS}_E},$$

with

$$\begin{aligned}
&q_1 = F[v_\beta, \infty; 1-\alpha/2],\\
&q_2 = (F[v_\beta, v_e; 1-\alpha/2] - F[v_\beta, \infty; 1-\alpha/2])F[v_\beta, v_e; 1-\alpha/2],\\
&q_3 = F[v_\beta, v_\alpha; 1-\alpha/2], \qquad q_4 = (r_3^* - 1)F[v_\beta, \infty; 1-\alpha/2],
\end{aligned}$$

$$q_1' = F[\upsilon_\beta, \infty; \alpha/2],$$
$$q_2' = (F[\upsilon_\beta, \upsilon_e; \alpha/2] - F[\upsilon_\beta, \infty; \alpha/2])F[\upsilon_\beta, \upsilon_e; \alpha/2],$$
$$q_3' = F[\upsilon_\beta, \upsilon_\alpha; \alpha/2], \quad \text{and} \quad q_4' = (r_3^* - 1)F[\upsilon_\beta, \infty; \alpha/2].$$

If the limits in (15.9.6) and (15.9.7) are negative or greater than one, they are replaced by 0 and 1, respectively. For balanced designs, the intervals in (15.9.6) and (15.9.7) reduce to (6.8.20) and (6.8.21), respectively. Sen et al. (1992) present formulas where υ_α and υ_β are estimated using Satterthwaite's approximation. They report that such modifications of the degrees of freedom are not needed unless the design is extremely unbalanced. Burdick et al. (1986) have reported some additional results for designs with equal subsampling or the last-stage uniformity.

15.9.6 A NUMERICAL EXAMPLE

In this example, we illustrate computations of confidence intervals on the variance components and certain of their parametric functions using the procedures described in Sections 15.9.1 through 15.9.5. for the pH reading data of the numerical example in Section 15.6.5. Now, from the results of the analysis of variance given in Table 15.5, we have

$$\text{MS}_E = 0.002474, \qquad \text{MS}_{Bu} = 0.003664, \qquad \text{MS}_{Au} = 0.012860,$$
$$a = 15, \qquad b = 37, \qquad \nu_e = 123, \qquad \nu_\beta = 22, \qquad \nu_\alpha = 14,$$
$$r_1^* = 4.222, \qquad r_2^* = 4.186, \qquad r_3^* = 9.914.$$

Further, for $\alpha = 0.05$, we obtain the following quantities:

$$\chi^2[\nu_e, \alpha/2] = 94.1950, \qquad \chi^2[\nu_e, 1-\alpha/2] = 155.5892,$$
$$F[\nu_e, \infty; \alpha/2] = 0.766, \qquad F[\nu_e, \infty; 1-\alpha/2] = 1.265,$$
$$F[\nu_\beta, \infty; \alpha/2] = 0.499, \qquad F[\nu_\beta, \infty; 1-\alpha/2] = 1.672,$$
$$F[\nu_\alpha, \infty; \alpha/2] = 0.402, \qquad F[\nu_\alpha, \infty; 1-\alpha/2] = 1.866,$$
$$F[\nu_\beta, \nu_e; \alpha/2] = 0.482, \qquad F[\nu_\beta, \nu_e; 1-\alpha/2] = 1.787,$$
$$F[\nu_\alpha, \nu_\beta; \alpha/2] = 0.355, \qquad F[\nu_\alpha, \nu_\beta; 1-\alpha/2] = 2.529,$$
$$F[\nu_\alpha, \nu_e; \alpha/2] = 0.392, \qquad F[\nu_\alpha, \nu_e; 1-\alpha/2] = 1.975,$$
$$F[\nu_e, \nu_\beta; \alpha/2] = 0.560, \qquad F[\nu_e, \nu_\beta; 1-\alpha/2] = 2.074,$$
$$F[\nu_\beta, \nu_\alpha; \alpha/2] = 0.396, \qquad F[\nu_\beta, \nu_\alpha; 1-\alpha/2] = 2.814,$$
$$F[\nu_\alpha + \nu_e, \infty; \alpha/2] = 0.777, \qquad F[\nu_\alpha + \nu_e, \infty; 1-\alpha/2] = 1.250,$$
$$F[\nu_\beta + \nu_e, \infty; \alpha/2] = 0.783, \qquad F[\nu_\beta + \nu_e, \infty; 1-\alpha/2] = 1.243.$$

Substituting the appropriate quantities in (15.9.1), the desired 95% confidence interval for σ_e^2 is given by

$$P\{0.0020 \le \sigma_e^2 \le 0.0032\} = 0.95.$$

To determine approximate confidence intervals for σ_β^2 and σ_α^2 using formulas (15.9.2) and (15.9.3), we further evaluate the following quantities:

$$
\begin{aligned}
G_1 &= 0.46409432, & G_2 &= 0.40191388, & G_3 &= 0.20948617,\\
H_1 &= 1.48756219, & H_2 &= 1.00400802, & H_3 &= 0.30548303,\\
G_{12} &= -0.01888096, & G_{13} &= 0.00869606, & G_{23} &= 0.00571293,\\
G_{32} &= -0.02088911, & G_{13}^* &= 0.02590995, & &\\
H_{12} &= -0.06868474, & H_{13} &= -0.03636351, & H_{23} &= -0.02022926,\\
H_{32} &= 0.00570778, & H_{23}^* &= 0.02270031, & &
\end{aligned}
$$

$$
\begin{aligned}
L_\beta &= 2.791554394 \times 10^{-6}, & U_\beta &= 1.361795569 \times 10^{-5},\\
L_\alpha &= 4.804330612 \times 10^{-5}, & U_\alpha &= 3.64505229 \times 10^{-4}.
\end{aligned}
$$

Substituting the appropriate quantities in (15.9.2) and (15.9.4), the desired 95% confidence intervals for σ_β^2 and σ_α^2 are given by

$$P\{-0.00011 \leq \sigma_\beta^2 \leq 0.00116\} \doteq 0.95$$

and

$$P\{0.00023 \leq \sigma_\alpha^2 \leq 0.00285\} \doteq 0.95.$$

It is understood that a negative limit is defined to be zero.

To determine an approximate confidence interval for $\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2$ using formula (15.9.5), we further evaluate the following quantities:

$$
\begin{aligned}
\ell_2 &= 1.356702984, & \ell_3 &= 7.557297016,\\
L_\gamma &= 5.4952264 \times 10^{-5}, & U_\gamma &= 4.234894177 \times 10^{-4}.
\end{aligned}
$$

Substituting the appropriate quantities in (15.9.5), the desired 95% confidence interval for $\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2$ is given by

$$P\{0.00294 \leq \sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2 \leq 0.00577\} \doteq 0.95.$$

Finally, in order to determine approximate confidence intervals for $\rho_\alpha = \sigma_\alpha^2/(\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2)$ and $\rho_\beta = \sigma_\beta^2/(\sigma_e^2 + \sigma_\beta^2 + \sigma_\alpha^2)$ using formulas (15.9.6) and (15.9.7), we evaluated the following quantities:

$$
\begin{aligned}
p_1 &= 4.222, & p_2 &= 10.586394, & p_3 &= 0.0711, & &\\
p_4 &= 4.222, & p_5 &= -14.486112, & p_6 &= -63.0161433, & &\\
p_1' &= 4.222, & p_2' &= 1.48603, & p_3' &= 0.014112, & &\\
p_4' &= 4.222, & p_5' &= -2.03344, & p_6' &= -12.507508, & &\\
q_1 &= 1.672, & q_2 &= 0.205505, & q_3 &= 2.814, & q_4 &= 14.904208,\\
q_1' &= 0.499, & q_2' &= -0.008194, & q_3' &= 0.396, & q_4' &= 4.448086,
\end{aligned}
$$

$$L_\beta = -0.0111663065, \qquad U_\beta = 0.1517762975.$$

Substituting the appropriate quantities in (15.9.6) and (15.9.7), the desired 95% intervals for ρ_α and ρ_β are given by

$$P\{0.05823 \leq \rho_\alpha \leq 0.52666\} \doteq 0.95$$

and

$$P\{-0.02662 \leq \rho_\beta \leq 0.29554\} \doteq 0.95.$$

Since $0 \leq \rho_\alpha, \rho_\beta \leq 1$, any bound less than zero is defined to be zero and greater than 1 is defined to be 1.

15.10 TESTS OF HYPOTHESES

In this section, we consider the problem of testing the hypotheses

$$\begin{aligned} &H_0^B : \sigma_\beta^2 = 0 \quad \text{vs.} \quad H_1^B : \sigma_\beta^2 > 0 \\ \text{and} \\ &H_0^A : \sigma_\alpha^2 = 0 \quad \text{vs.} \quad H_1^A : \sigma_\alpha^2 > 0, \end{aligned} \tag{15.10.1}$$

using the results on analysis of variance based on Type I sums of squares.

15.10.1 TESTS FOR $\sigma_\beta^2 = 0$

To form a test for $\sigma_\beta^2 = 0$ in (15.10.1) note that MS_E and MS_B are independent, MS_E has a scaled chi-square distribution; and, in addition, under the null hypothesis, they have the same expectation and MS_B also has a scaled chi-square distribution. Therefore, a test statistic is constructed by the variance ratio

$$\text{MS}_B/\text{MS}_E, \tag{15.10.2}$$

which has an F distribution with v_β and v_e degrees of freedom. The test based on the statistic in (15.10.2) is exact and is equivalent to the corresponding test for balanced design. It has been shown that there does not exist a uniformly most powerful invariant or uniformly most powerful invariant unbiased test for this problem. Khattree and Naik (1990) consider some locally best invariant unbiased tests for the problem. Hussein and Milliken (1978) discuss an exact test for $\sigma_\beta^2 = 0$ in (15.10.1) when $\beta_{j(i)}$s have heterogeneous variance structure. A more general hypothesis of the type $H_0 : \rho_\beta \leq \rho_{\beta_0}$ vs. $H_1 : \rho_\beta > \rho_{\beta_0}$, where $\rho_\beta = \sigma_\beta^2/\sigma_e^2$, can be tested using Wald's procedure (Verdooren, 1976).

15.10.2 TESTS FOR $\sigma_\alpha^2 = 0$

In the unbalanced model in (15.1.1), there does not exist an exact test for $\sigma_\alpha^2 = 0$ in (15.10.1). Since MS_A and MS_B are not independent and do not have a scaled chi-square distribution, the usual test based on the statistic $\mathrm{MS}_A/\mathrm{MS}_B$ is no longer applicable.[3] A common procedure is to ignore the assumption of independence and chi-squaredness and construct a pseudo F-test using synthesis of mean squares based on Satterthwaite's procedure (see, e.g., Anderson, 1960; Eisen, 1966; Cummings and Gaylor, 1974). As we have seen, in constructing a pseudo F-test one can either obtain a numerator or a denominator component of the test statistic, or both. To construct a denominator component of the test statistic for $\sigma_\alpha^2 = 0$, we obtain a linear combination of MS_B and MS_E such that it has expected value equal to $\sigma_e^2 + r_2\sigma_\beta^2$. It is readily seen that the desired statistic is given by

$$\mathrm{MS}_D = (r_2/r_1)\mathrm{MS}_B + (1 - r_2/r_1)\mathrm{MS}_E. \tag{15.10.3}$$

We now assume (incorrectly) that MS_B has a scaled chi-square distribution and is independent of MS_A. Since MS_E has a scaled chi-square distribution and is independent of MS_B and MS_A, the linear combination (15.10.3) is approximated by a scaled chi-square distribution by fitting the first two moments (see Appendix F). Let v_D denote the degrees of freedom of the approximate chi-square statistic given by (15.10.3). Then the test procedure for testing $\sigma_\alpha^2 = 0$ in (15.10.1) is based on the statistic

$$F = \mathrm{MS}_A/\mathrm{MS}_D, \tag{15.10.4}$$

which is assumed to follow an approximate F-distribution with v_α and v_D degrees of freedom. Note that when $r_2 \geq r_1$, the coefficient $1 - r_2/r_1$, may assume a negative value which may affect the accuracy of the F-test. For some further discussion on the adequacy of approximation involving a negative coefficient, see Appendix F.

As mentioned earlier, an alternate test for $\sigma_\alpha^2 = 0$ can be obtained by constructing a numerator component of the test statistic such that under the null hypothesis it has expected value equal to $\sigma_e^2 + r_1\sigma_\beta^2$. It is readily seen that the desired component is given by

$$\mathrm{MS}_N = (r_1/r_2)\mathrm{MS}_A + (1 - r_1/r_2)\mathrm{MS}_E. \tag{15.10.5}$$

Thus, proceeding as above, the alternate test procedure is based on the statistic

$$\mathrm{MS}_N/\mathrm{MS}_B, \tag{15.10.6}$$

which is assumed to follow an approximate F-distribution with v_N and v_β degrees of freedom, where v_N is the degrees of freedom associated with the linear combination in (15.10.5).

[3]Some authors have ignored the unbalanced structure of the design and have used the conventional F-test based on the statistic $\mathrm{MS}_A/\mathrm{MS}_B$ with $a - 1$ and $b_. - a$ degrees of freedom (see, e.g., Bliss, 1967, p. 353).

Similarly, a test statistic for $\sigma_\alpha^2 = 0$ can be obtained by constructing both a numerator and a denominator component such that under the null hypothesis they have the same expected value. Again, it is seen that the desired numerator and denominator components are given by

$$\mathrm{MS}'_N = r_1\mathrm{MS}_A + r_2\mathrm{MS}_E \tag{15.10.7}$$

and

$$\mathrm{MS}'_D = r_2\mathrm{MS}_B + r_1\mathrm{MS}_E. \tag{15.10.8}$$

From which, the test procedure is based on the statistic

$$\mathrm{MS}'_N/\mathrm{MS}'_D, \tag{15.10.9}$$

which is assumed to follow an approximate F-distribution, with v'_N and v'_D degrees of freedom, where v'_N and v'_D are the degrees of freedom associated with the linear combinations in (15.10.7) and (15.10.8), respectively.

The Satterthwaite-like test procedures (15.10.4), (15.10.6), and (15.10.9) have been proposed by Cummings and Gaylor (1974) and Tan and Cheng (1984). In the special case when $n_{ij} = n$, we have seen that MS_A and MS_B are jointly independent and MS_B is distributed as the multiple of a chi-square variable. In addition, it readily follows that for this design[4] $r_1 = r_2 = n$, so that under the null hypothesis H_0^A in (15.10.1), MS_A and MS_B have identical expectations. Furthermore, under H_0^A, MS_A also has constant times a chi-square distribution (see, e.g., Johnson and Leone, 1964, Vol. 2, p. 32). Thus the usual variance ratio

$$\mathrm{MS}_A/\mathrm{MS}_B \tag{15.10.10}$$

provides an exact test of H_0^A in (15.10.1).

Tietjen (1974) investigated the test size and power of the test statistics in (15.10.4) and (15.10.10) for a variety of unbalanced designs taken from Goldsmith and Gaylor (1970) using Monte Carlo simulation. He found that the test size of the statistic in (15.10.10) was always in the interval (0.044–0.058) for all the 61 designs studied by him and in general its performance was far better than that of the statistic in (15.10.4). Cummings and Gaylor (1974) also investigated the effect of violation of assumptions of independence and chi-squaredness on test size in using the procedures based on the test statistics in (15.10.4) and (15.10.6) and reported that dependence and non-chi-squaredness seem to have cancellation effect and both procedures appear to be satisfactory. Their results appear to indicate that the test sizes of these statistics are only slightly affected for a wide range of variance component ratios and unbalanced designs. Tan and Cheng (1984) studied the performance of the test procedures in (15.10.4),

[4] This design has been called "last-stage uniformity" by Tietjen (1974) who attributes the term to Kruskal (1968).

(15.10.6), (15.10.9), and (15.10.10), using a better approximation for the distribution of the test statistic based on Laguerre polynomial expansion, and found that all of them had satisfactory performance, but the procedure in (15.10.10) is inferior for extremely unbalanced designs and cannot be recommended for general use. The exact probability level of these test procedures can be calculated using the method reported by Verdooren (1974).

Khuri (1987) proposed an exact test for this problem and compared it with the tests mentioned above. For the derivation of his test, Khuri considers the model for the cell means $\bar{y}_{ij.}$s and applies a series of orthogonal transformations to construct two independent sums of squares which under the null hypothesis have constant times a chi-square distribution. These sums of squares are then used to define an F-statistic for testing H_0^A in (15.10.1). He reports that his exact test has superior power properties over the others, but the test requires a nonunique partitioning of the error sum of squares (see also Khuri et al., 1998, pp. 113–117). Hernández et al. (1992) have proposed testing H_0^A in (15.10.1) by using the lower bound of σ_α^2 in (15.9.4). They report that this test has comparable power to other approximate tests mentioned earlier including Khuri's exact test and its test size is only slightly affected.

It has been shown that there does not exist a uniformly most powerful invariant or uniformly most powerful invariant unbiased test for this problem. Some locally best invariant unbiased tests are derived by Khattree and Naik (1990) and Naik and Khattree (1992) using partially balanced data. The latter paper considers a design with two-way mixed model when blocks are nested and random. Hussein and Milliken (1978) discuss an exact test for H_0^A in (15.10.1) when α_is have heterogeneous error structure given by $\text{Var}(\alpha_i) = d_i\sigma_\alpha^2$, $\text{Var}(\beta_{j(i)}) = \sigma_\beta^2$, and the design contains last stage uniformity. Verdooren (1988) outlined a procedure for testing a more general hypothesis of the type $H_0 : \rho_\alpha \leq \rho_{\alpha_0}$ vs. $H_1 : \rho_\alpha > \rho_{\alpha_0}$, where $\rho_\alpha = \sigma_\alpha^2/\sigma_e^2$. For some further results on tests of hypotheses in a two-way unbalanced nested random model, see Jain and Singh (1989).

15.10.3 A NUMERICAL EXAMPLE

In this section, we outline computations for testing the hypotheses in (15.10.1) for the blood pH reading data of the numerical example in Section 15.6.5. Note that in this example, the variance components σ_α^2 and σ_β^2 correspond to the variations among dams and among sires within dams, respectively. The hypothesis $H_0^B : \sigma_\beta^2 = 0$ is tested using the conventional F-test. The corresponding test statistic (15.10.2) gives an F-value of 1.47 ($p = 0.097$). The results are barely significant at a level of significance of 10% and there does not seem to be a strong evidence of variability among sires. Note that this F-test is exact. For testing the hypothesis $H_0^A : \sigma_\alpha^2 = 0$, however, there does not exist a simple exact test. We will therefore employ Satterthwaite-type tests given by (15.10.4), (15.10.6), (15.10.9); and the conventional F-test (15.10.10). The corresponding test procedures are readily evaluated and the resulting quantities

TABLE 15.7 Test procedures for $\sigma_\alpha^2 = 0$.

Test procedure	*F*-statistic		Degrees of freedom		*F*-value	*p*-value
	Numerator	Denominator	Numerator	Denominator		
(15.10.4)	0.012716	0.003620	14	21.9	3.513	0.004
(15.10.6)	0.012682	0.003637	13.9	22	3.508	0.004
(15.10.9)	0.065337	0.026521	19.6	56.6	2.464	0.004
(15.10.10)	0.012716	0.003736	14	22	3.496	0.004

including the numerator and denominator components, the corresponding degrees of freedom, the values of F-statistics, and the p-values are summarized in Table 15.7. Note that all the procedures lead to nearly the same result. Further, it is evident that the results are highly significant and we reject H_0^A and conclude that $\sigma_\alpha^2 > 0$, or pH readings among dams differ significantly.

EXERCISES

1. Express the coefficients of the variance components in the expected mean squares derived in Section 15.3 in terms of the formulation given in Section 17.3
2. Apply the method of "synthesis" to derive the expected mean squares given in Section 15.3.
3. Derive the results on expected values of unweighted mean squares given in (15.5.2).
4. Show that the ANOVA estimators (15.6.2) reduce to the corresponding estimators (6.4.1) for balanced data.
5. Show that the unweighted means estimators (15.6.4) reduce to the ANOVA estimators (6.4.1) for balanced data.
6. Show that the symmetric sums estimators (15.6.7) and (15.6.10) reduce to the ANOVA estimators (6.4.1) for balanced data.
7. Derive the expressions for variances and covariances of the analysis of variance estimators of the variance components as given in Section 15.7.1 (Searle, 1961).
8. Derive the expressions for large sample variances and covariances of the maximum likelihood estimators of the variance components as given in Section 15.7.2 (Searle, 1970).
9. For the model in (15.1.1) determine the minimal sufficient statistics (Khuri and Ghosh, 1990).
10. For the model in (15.1.1) show that SS_A and SS_B defined in (15.2.1) are not independent and do not have a chi-square–type distribution.

11. Show that for the three-stage Bainbridge design shown in Figure 15.2, the ANOVA estimators of the variance components are given by

$$\hat{\sigma}^2_{e,\text{ANOV}} = \text{MS}_E,\ \hat{\sigma}^2_{\beta,\text{ANOV}} = \frac{3}{4}(\text{MS}_B - \text{MS}_E),$$

and

$$\hat{\sigma}^2_{\alpha,\text{ANOV}} = \frac{1}{12}(4\text{MS}_A - 5\text{MS}_B + \text{MS}_E).$$

12. An experiment was conducted to investigate the variation in the blood pH of mice. Four female mice (dams) were successively mated over a period of time with either two, three, four, or five males (sires). Each sire was mated to different dams and the measurements were made on the blood pH reading of female offspring. The data are given below.

Dam	**1**		**2**			**3**				**4**				
Sire	**1**	**2**	**1**	**2**	**3**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**5**
pH Reading	7.76	7.97	7.97	8.08	8.18	8.13	8.23	8.33	8.42	8.37	8.43	8.48	8.53	8.57
	7.86	8.01	8.05	8.11	8.19	8.19	8.26	8.35	8.44	8.45	8.49	8.51	8.55	8.59
		8.05		8.14	8.22		8.29	8.37	8.46			8.53	8.57	8.61
					8.24			8.39	8.48				8.59	8.63
									8.50					8.65

(a) Describe the mathematical model and the assumptions involved.

(b) Analyze the data and report the conventional analysis of variance table based on Type I sums of squares.

(c) Perform an appropriate F-test to determine whether the blood pH readings differ from dam to dam.

(d) Perform an appropriate F-test to determine whether the blood pH readings differ from sire to sire.

(e) Find point estimates of the variance components, the ratios of the variance components to the error variance, the proportions of the variance components, and the total variance using the methods described in the text.

(f) Calculate 95% confidence intervals associated with the point estimates in part (e) using the methods described in the text.

13. Consider an experiment involving strain measurements from a large number of sealing machines. Three machines were randomly selected for the study. The first two machines each have two heads, while the third machine has three heads. The results in the form of coded raw data are given below.

Machine	**1**		**2**		**3**		
Head	**1**	**2**	**1**	**2**	**1**	**2**	**3**
Strain	8	7	6	4	5	4	6
	5	9	5	7	5	5	7
	7	8		5	7	6	
		9		6		8	
				3			

(a) Describe the mathematical model and the assumptions involved
(b) Analyze the data and report the conventional analysis of variance table based on Type I sums of squares.
(c) Perform an appropriate F-test to determine whether the strain measurements differ from machine to machine.
(d) Perform an appropriate F-test to determine whether the strain measurements differ from head to head.
(e) Find point estimates of the variance components, the ratios of the variance components to the error variance, the proportions of the variance components, and the total variance using the methods described in the text.
(f) Calculate 95% confidence intervals associated with the point estimates in part (e) using the methods described in the text.

14. Heckler and Rao (1985) reported the results of an experiment designed to measure the variation in enzyme measurements. Three laboratories preparing the enzyme were randomly selected and four weeks were randomly assigned for each of the laboratories. Two or three days were sampled from the selected weeks and two measurements were obtained on each day. The data containing the averages for the days are given below.

Laboratory	1				2				3			
Week	1	2	3	4	1	2	3	4	1	2	3	4
Enzyme	43.4	37.0	23.6	51.0	7.0	32.4	13.4	23.9	22.4	25.4	22.9	18.8
	46.2	16.6	33.6	52.4	7.8	16.8	9.6	19.3	15.5	23.1	0.6	3.7
	46.5				15.7				29.7			

Source: Heckler and Rao (1985); used with permission.

(a) Describe the mathematical model and the assumptions involved.
(b) Analyze the data and report the conventional analysis of variance table based on Type I sums of squares.
(c) Perform an appropriate F-test to determine whether the enzyme measurements differ from laboratory to laboratory.
(d) Perform an appropriate F-test to determine whether the enzyme measurements differ from week to week.
(e) Find point estimates of the variance components, the ratios of the variance components to the error variance, the proportions of the variance components, and the total variance using the methods described in the text.
(f) Calculate 95% confidence intervals associated with the point estimates in part (e) using the methods described in the text.

15. Snedecor and Cochran (1989, pp. 291–294) described an experiment designed to study variation in the wheat yield of the commercial wheat

fields in Great Britain. Six districts were chosen for the experiment and within each district a number of farms were selected. Finally, within each farm one to three fields were drawn and observed wheat yield of the field was recorded. The data are given below.

District	**1**				**2**			**3**			**4**										
Farm	**1**		**2**		**1**		**2**	**1**			**1**		**2**		**3**	**4**	**5**	**6**	**7**	**8**	**9**
Field	**1**	**2**	**1**	**2**	**1**	**2**	**1**	**1**	**2**	**3**	**1**	**2**	**1**	**2**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
Yield	23	19	31	37	33	29	29	36	29	33	11	21	23	18	33	23	26	39	20	24	36

District	**5**		**6**												
Farm	**1**		**1**		**2**		**3**		**4**	**5**	**6**	**7**	**8**	**9**	**10**
Field	**1**	**2**	**1**	**2**	**1**	**2**	**1**	**2**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
Yield	25	33	28	31	25	42	32	36	41	35	16	30	40	32	44

Source: Snedecor and Cochran (1989); used with permission.

(a) Describe the mathematical model and the assumptions involved.

(b) Analyze the data and report the conventional analysis of variance table based on Type I sums of squares.

(c) Perform an appropriate F-test to determine whether the wheat yields vary from district to district.

(d) Perform an appropriate F-test to determine whether the wheat yields vary from farm to farm.

(e) Find point estimates of the variance components, the ratios of the variance components to the error variance, the proportions of the variance components, and the total variance using the methods described in the text.

(f) Calculate 95% confidence intervals associated with the point estimates in part (e) using the methods described in the text.

16. Rosner (1982) described the analysis of a two-stage nested model used to analyze the data from certain measurements made in a routine ocular examination of an outpatient population of 218 persons aged 20–39 with *retinitis pigmentosa* (RP). The patients were classified into four genetic types: autosomal dominant (DOM), autosomal recessive (AR), sex-linked (SL), and isolate (ISO). The sample used for the analysis contained 212 persons, out of which 28 persons were in the DOM group, 20 persons in the AR group, 18 persons in the SL groups, and 146 persons in the ISO group. Of these persons, 210 had measurements taken for both eyes while two had information for only one eye. The following analysis of variance table gives results on sums of squares obtained from the data on "spherical refractive error." The results are based on 212 persons giving a total of 422 measurements; however, for the purpose of this exercise, they should be treated as coming from 210 persons (26 persons in the DOM group) who had information on both eyes. This assumption leads to the last-stage uniformity (i.e., two observations per person) and simplifies calculations for expected mean squares and tests of hypotheses.

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Groups		133.59		
Persons within groups		2,518.45		
Error		80.49		

Source: Rosner (1982); used with permission.

(a) Describe the mathematical model and the assumptions for the experiment. In the original analysis, the groups were considered to be fixed and the remaining two factors random. For the purpose of this exercise, you can assume a completely random model.

(b) Complete the remaining columns of the preceding analysis of variance table.

(c) Test the hypothesis that there are significant differences between the refractive errors of different genetic groups.

(d) Test the hypothesis that there are significant differences between the refractive errors of persons within groups.

(e) Find point estimates of the variance components, the ratios of the variance components to the error variance, the proportions of the variance components, and the total variance using the methods described in the text.

(f) Calculate 95% confidence intervals associated with the point estimates in part (a) using the methods described in the text.

Bibliography

R. L. Anderson (1960), Use of variance component analysis in the interpretation of biological experiments, Part 3, *Bull. Internat. Statist. Inst.*, **37**, 71–90.

R. L. Anderson (1975), Designs and estimators for variance components, in J. N. Srivastava, ed., *Statistical Design and Linear Model*, North-Holland, Amsterdam, 1–30.

R. L. Anderson (1981), Recent developments in designs and estimators for variance components, in M. Csörgö, D. A. Dawson, J. N. K. Rao, and A. K. Md. E. Saleh, eds., *Statistics and Related Topics*, North-Holland, Amsterdam, 3–22.

R. L. Anderson and T. A. Bancroft (1952), *Statistical Theory in Research*, McGraw–Hill, New York.

T. R. Bainbridge (1965), Staggered nested designs for estimating variance components, *Indust. Quality Control*, **22**-1, 12–20.

C. I. Bliss (1967), *Statistics in Biology*, McGraw–Hill, New York.

R. K. Burdick, H. J. Birch, and F. A. Graybill (1986), Confidence intervals on measures of variability in an unbalanced two-fold nested design with equal subsampling, *J. Statist. Comput. Simul.*, **25**, 259–272.

R. K. Burdick and J. E. Eickman (1986), Confidence intervals on the among group variance component in the unbalanced one-fold nested design, *J. Statist. Comput. Simul.*, **26**, 205–219.

R. K. Burdick and F. A. Graybill (1985), Confidence intervals on the total variance in an unbalanced two-fold nested classification with equal subsampling, *Comm. Statist.* A *Theory Methods*, **14**, 761–774.

R. K. Burdick and F. A. Graybill (1992), *Confidence Intervals on Variance Components*, Marcel Dekker, New York.

W. B. Cummings (1972), *Variance Component Testing in Unbalanced Nested Designs*, Mim. Ser. 243, Institute of Statistics, North Carolina State University, Raleigh, NC.

W. B. Cummings and D. W. Gaylor (1974), Variance component testing in unbalanced nested designs, *J. Amer. Statist. Assoc.*, **69**, 765–771.

E. J. Eisen (1966), The quasi-F test for an unnested fixed factor in an unbalanced hierarchal design with a mixed model, *Biometrics*, **22**, 937–942.

M. Ganguli (1941), A note on nested sampling, *Sankhyā*, **5**, 449–452.

C. H. Goldsmith (1969), *Three-Stage Nested Designs for Estimating Variance Components*, Ph.D. dissertation, Institute of Statistics, North Carolina State University, Raleigh, NC.

C. H. Goldsmith and D. W. Gaylor (1970), Three-stage nested designs for estimating variance components, *Technometrics*, **12**, 487–498.

F. A. Graybill (1961), *An Introduction to Linear Statistical Models*, Vol. I, McGraw–Hill, New York.

C. E. Heckler and P. S. R. S. Rao (1985), Efficient estimation of variance components using staggered nested designs, in *Fall Technical Conference*, American Society for Quality Control, Corning, NY.

R. P. Hernández and R. K. Burdick (1993), Confidence intervals on the total variance in an unbalanced two-fold nested design, *Biometrical J.*, **35**, 515–522.

R. P. Hernández, R. K. Burdick, and N. J. Birch (1992), Confidence intervals and tests of hypotheses on variance components in an unbalanced two-fold nested design, *Biometrical J.*, **34**, 387–402.

M. Hussein and G. A. Milliken (1978), An unbalanced nested model with random effects having unequal variances, *Biometrical J.*, **20**, 329–338.

R. C. Jain and J. Singh (1989), Tests of hypotheses in two-way unbalanced random model, *Gujrat Statist. Rev.*, **16**, 43–46.

R. C. Jain, J. Singh, and R. Agrawal (1991), Interval estimation in the two-way nested unbalanced random model, *Trab. Estadíst.*, **6**-2, 33–40.

N. L. Johnson and F. C. Leone (1964), *Statistics and Experimental Design in Engineering and the Physical Sciences*, 1st ed., Wiley, New York; 2nd ed., 1977,

R. Khattree and D. N. Naik (1990), Optimum tests for random effects in unbalanced nested designs, *Statistics*, **21**, 163–168.

A. I. Khuri (1987), An exact test for the nesting effect's variance component in an unbalanced random two-fold nested model, *Statist. Probab. Lett.*, **5**, 305–311.

A. I. Khuri (1990), Exact test for random models with unequal cell frequencies in the last stage, *J. Statist. Plann. Inference*, **24**, 177–193.

A. I. Khuri and M. Ghosh (1990), Minimal sufficient statistics for the unbalanced two-fold nested model, *Statist. Probab. Lett.*, **10**, 351–353.

A. I. Khuri, T. Mathew, and B. K. Sinha (1998), *Statistical Tests for Mixed Linear Models*, Wiley, New York.

G. G. Koch (1967), A general approach to the estimation of variance components, *Techometrics*, **9**, 93–118.

G. G. Koch (1968), Some further remarks concerning "A general approach to estimation of variance components," *Technometrics*, **10**, 551–558.

W. H. Kruskal (1968), Note to the Tietjen and Moore (1968) article, *Biometrics*, **24**, 1025–1026.

D. N. Naik and R. Khattree (1992), Optimum tests for treatments when blocks are nested and random, *Statistics*, **23**, 101–108.

R. R. Prairie (1962), *Optimal Designs to Estimate Variance Components and Reduce Product Variability for Nested Classifications*, Ph.D. dissertation, Institute of Statistics, North Carolina State University, Raleigh, NC.

R. R. Prairie and R. L. Anderson (1962), *Optimal Designs to Estimate Variance Components and to Reduce Product Variability for Nested Classifications*, Mim. Ser. 313, Institute of Statistics, North Carolina State University, Raleigh, NC.

P. S. R. S. Rao and C. E. Heckler (1997), Estimators for the three-fold nested random effects model, *J. Statist. Plann. Inference*, **64**, 341–352.

B. Rosner (1982), Statistical methods in ophthalmology: An adjustment for the intraclass correlation between eyes, *Biometrics*, **38**, 105–114.

H. Scheffé (1959), *The Analysis of Variance*, Wiley, New York.

J. H. Schwartz (1980), *Optimal Allocation of Resources to Reduce Product or Process Variability: A Comparison of Some Designs to Estimate the Allocation Parameters and the Variance Components*, Ph.D. dissertation, University of Kentucky, Lexington, KY.

S. R. Searle (1961), Variance components in the unbalanced two-way nested classification, *Ann. Math. Statist.*, **32**, 1161–1166.

S. R. Searle (1970), Large sample variances of maximum likelihood estimators of variance components using unbalanced data, *Biometrics*, **26**, 505–525.

S. R. Searle (1971), *Linear Models*, Wiley, New York.

S. R. Searle, G. Casella, and C. E. McCulloch (1992), *Variance Components*, Wiley, New York.

B. Sen (1988), *Confidence Intervals on Ratios of Variance Components in the Three-Stage Nested Unbalanced Model*, Ph.D. dissertation, Colorado State University, Fort Collins, CO.

B. Sen, F. A. Graybill, and N. Ting (1992), Confidence intervals on ratios of variance components for the unbalanced two-factor nested model, *Biometrical J.*, **3**, 259–274.

G. W. Snedecor and W. G. Cochran (1989), *Statistical Methods*, 8th ed., Iowa

State University Press, Ames, IA; 6th ed., 1967; 7th ed., 1980.

R. R. Sokal and F. J. Rohlf (1995), *Biometry*, 3rd ed., W. H. Freeman, New York; 1st ed., 1969; 3rd ed., 1981.

W. Y. Tan and S. S. Cheng (1984), On testing variance components in three-stage unbalanced nested random effects models, *Sankhyā Ser.* B, **46**, 188–200.

J. D. Thomas and R. A. Hultquist (1978), Interval estimation for the unbalanced case of the one-way random effects model, *Ann. Statist.*, **6**, 582–587.

G. L. Tietjen (1974), Exact and approximate tests for unbalanced random effects designs, *Biometrics*, **30**, 573–581.

L. R. Verdooren (1974), Exact tests about variance ratios in unbalanced two and three stages nested designs, in L. C. A. Corsten and T. Postelnicu, eds., *Proceedings of the* 8*th International Biometrics Conference*, Editoria Academie, Constanta, Romania, 221–228.

L. R. Verdooren (1976), Exact confidence intervals for the variance ratio in nested designs, in J. Gordesch and P. Naeve, eds., *COMPSTAT: Proceedings of the* 2*nd Symposium on Computational Statistics*, Physica-Verlag, Wien, Germany, 221–228.

L. R. Verdooren (1988), Exact tests and confidence intervals for the ratio of variance components in unbalanced two- and three-stage nested designs, *Comm. Statist.* A *Theory Methods*, **17**, 1197–1230.

16 Three-Way Nested Classification

Consider three factors A, B, and C, where B is nested within A and C is nested within B. Suppose that each A level has b_i B levels, each B level has c_{ij} C levels, and n_{ijk} observations are taken from each C level. This is an example of a three-way unbalanced nested design and is frequently encountered in many areas of scientific applications. For example, suppose a clinical study involves monthly blood analysis of patients participating in the study. Two blood tests are made on each patient and three analyses are made from each test. Here, tests are nested within patients and analyses are made within tests. It may happen that on certain occasions some patients fail to appear for their blood tests and this makes the design unbalanced. In this chapter, we will study the random effects model for the three-way nested classification involving an unbalanced design.

16.1 MATHEMATICAL MODEL

The random effects model for the unbalanced three-way nested classification is given by

$$y_{ijk\ell} = \mu + \alpha_i + \beta_{j(i)} + \gamma_{k(ij)} + e_{\ell(ijk)} \begin{cases} i = 1, 2, \ldots, a, \\ j = 1, 2, \ldots, b_i, \\ k = 1, 2, \ldots, c_{ij}, \\ \ell = 1, 2, \ldots, n_{ijk}, \end{cases} \tag{16.1.1}$$

where $y_{ijk\ell}$ is the ℓth observation within the kth level of factor C within the jth level of factor B within the ith level of factor A, μ is the overall mean, α_i is the effect due to the ith level of factor A, $\beta_{j(i)}$ is the effect due to the jth level of factor B nested within the ith level of factor A, $\gamma_{k(ij)}$ is the effect due to the kth level of factor C nested within the jth level of factor B within the ith level of factor A, and $e_{ijk\ell}$ is the residual error of the observation $y_{ijk\ell}$. It is assumed that $-\infty < \mu < \infty$ is a constant and α_is, $\beta_{j(i)}$s, $\gamma_{k(ij)}$s, and $e_{\ell(ijk)}$s are mutually and completely uncorrelated random variables with means zero and variances σ_α^2, σ_β^2, σ_γ^2, and σ_e^2, respectively. The parameters σ_α^2, σ_β^2, σ_γ^2, and σ_e^2 are known as the variance components.

16.2 ANALYSIS OF VARIANCE

For the model in (16.1.1) there in no unique analysis of variance. The conventional analysis of variance is given in Table 16.1, where $b_. = \sum_{i=1}^a b_i$, $c_{..} = \sum_{i=1}^a \sum_{j=1}^{b_i} c_{ij}$, and $N = \sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}$. The sums of squares in Table 16.1, commonly referred to as Type I sums of squares, are defined as follows:

$$\begin{aligned}
SS_A &= \sum_{i=1}^a n_{i..}(\bar{y}_{i...} - \bar{y}_{....})^2 = \sum_{i=1}^a \frac{y_{i...}^2}{n_{i..}} - \frac{y_{....}^2}{N}, \\
SS_B &= \sum_{i=1}^a \sum_{j=1}^{b_i} n_{ij.}(\bar{y}_{ij..} - \bar{y}_{i...})^2 = \sum_{i=1}^a \sum_{j=1}^{b_i} \frac{y_{ij..}^2}{n_{ij.}} - \sum_{i=1}^a \frac{y_{i...}^2}{n_{i..}}, \\
SS_C &= \sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}(\bar{y}_{ijk.} - \bar{y}_{ij..})^2 \\
&= \sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \frac{y_{ijk.}^2}{n_{ijk}} - \sum_{i=1}^a \sum_{j=1}^{b_i} \frac{y_{ij..}^2}{n_{ij.}},
\end{aligned} \tag{16.2.1}$$

and

$$\begin{aligned}
SS_E &= \sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \sum_{\ell=1}^{n_{ijk}} (y_{ijk\ell} - \bar{y}_{ijk.})^2 \\
&= \sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \sum_{\ell=1}^{n_{ijk}} y_{ijk\ell}^2 - \sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \frac{y_{ijk.}^2}{n_{ijk}},
\end{aligned}$$

with the customary notation for totals and means.

Define the uncorrected sums of squares as

$$T_A = \sum_{i=1}^a \frac{y_{i...}^2}{n_{i..}}, \qquad T_B = \sum_{i=1}^a \sum_{j=1}^{b_i} \frac{y_{ij..}^2}{n_{ij.}},$$

$$T_C = \sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ijk}} \frac{y_{ijk.}^2}{n_{ijk}}, \qquad T_0 = \sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \sum_{\ell=1}^{n_{ijk}} y_{ijk\ell}^2,$$

and

$$T_\mu = y_{....}^2/N.$$

TABLE 16.1 Analysis of variance for the model in (16.1.1).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Factor A	$a-1$	SS_A	MS_A	$\sigma_e^2+r_4\sigma_\gamma^2+r_5\sigma_\beta^2+r_6\sigma_\alpha^2$
Factor B within A	$b_.-a$	SS_B	MS_B	$\sigma_e^2+r_2\sigma_\gamma^2+r_3\sigma_\beta^2$
Factor C within B	$c_{..}-b_.$	SS_C	MS_C	$\sigma_e^2+r_1\sigma_\gamma^2$
Error	$N-c_{..}$	SS_E	MS_E	σ_e^2

Then the corrected sums of squares defined in (16.2.1) can be written as

$$SS_A = T_A - T_\mu, \qquad SS_B = T_B - T_A,$$
$$SS_C = T_C - T_B, \quad \text{and} \quad SS_E = T_0 - T_C.$$

The mean squares as usual are obtained by dividing the sums of squares by the respective degrees of freedom. The expected mean squares are readily obtained and the derivations are presented in the following section.

16.3 EXPECTED MEAN SQUARES

The expected values of the sums of squares, or, equivalently, the mean squares, are readily obtained by first calculating the expected values of the quantities T_0, T_μ, T_A, T_B, and T_C. First, note that by the assumptions of the model in (16.1.1),

$$E(\alpha_i) = E(\beta_{j(i)}) = E(\gamma_{k(ij)}) = E(e_{\ell(ijk)}) = 0,$$
$$E(\alpha_i^2) = \sigma_\alpha^2, \quad E(\beta_{j(i)}^2) = \sigma_\beta^2, \quad E(\gamma_{k(ij)}^2) = \sigma_\gamma^2, \quad \text{and} \quad E(e_{\ell(ijk)}^2) = \sigma_e^2.$$

Further, all covariances between the elements of the same random variable and any pair of nonidentical random variables are equal to zero.

Now, we have

$$\begin{aligned}
E(T_0) &= \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}}\sum_{\ell=1}^{n_{ijk}} E(y_{ijk\ell}^2) \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}}\sum_{\ell=1}^{n_{ijk}} E(\mu + \alpha_i + \beta_{j(i)} + \gamma_{k(ij)} + e_{\ell(ijk)})^2 \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}}\sum_{\ell=1}^{n_{ijk}} (\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 + \sigma_e^2)
\end{aligned}$$

$$
\begin{aligned}
&= N(\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 + \sigma_e^2),\\
E(T_\mu) &= E(y_{....}^2/N)\\
&= N^{-1}E\Bigg[N\mu + \sum_{i=1}^{a} n_{i..}\alpha_i + \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij.}\beta_{j(i)} + \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n_{ijk}\gamma_{k(ij)}\\
&\qquad + \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}}\sum_{\ell=1}^{n_{ijk}} e_{\ell(ijk)}\Bigg]^2\\
&= N^{-1}\Bigg[N^2\mu^2 + \sum_{i=1}^{a} n_{i..}^2\sigma_\alpha^2 + \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij.}^2\sigma_\beta^2\\
&\qquad + \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n_{ijk}^2\sigma_\gamma^2 + N\sigma_e^2\Bigg]\\
&= N\mu^2 + k_1\sigma_\alpha^2 + k_2\sigma_\beta^2 + k_3\sigma_\gamma^2 + \sigma_e^2,\\
E(T_A) &= \sum_{i=1}^{a} E(y_{i...}^2/n_{i..})\\
&= \sum_{i=1}^{a} n_{i..}^{-1}E\Bigg[n_{i..}(\mu + \alpha_i) + \sum_{j=1}^{b_i} n_{ij.}\beta_{j(i)} + \sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n_{ijk}\gamma_{k(ij)}\\
&\qquad + \sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}}\sum_{\ell=1}^{n_{ijk}} e_{\ell(ijk)}\Bigg]^2\\
&= \sum_{i=1}^{a} n_{i..}^{-1}\Bigg[n_{i..}^2(\mu^2 + \sigma_\alpha^2) + \sum_{j=1}^{b_i} n_{ij.}^2\sigma_\beta^2 + \sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n_{ijk}^2\sigma_\gamma^2 + n_{i..}\sigma_e^2\Bigg]\\
&= \sum_{i=1}^{a}\Bigg[n_{i..}(\mu^2 + \sigma_\alpha^2) + \sum_{j=1}^{b_i}\frac{n_{ij.}^2}{n_{i..}}\sigma_\beta^2 + \sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}}\frac{n_{ijk}^2}{n_{i..}}\sigma_\gamma^2 + \sigma_e^2\Bigg]\\
&= N(\mu^2 + \sigma_\alpha^2) + k_4\sigma_\beta^2 + k_5\sigma_\gamma^2 + a\sigma_e^2,\\
E(T_B) &= \sum_{i=1}^{a}\sum_{j=1}^{b_i} E(y_{ij..}^2/n_{ij.})\\
&= \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij.}^{-1}E\Bigg[n_{ij.}(\mu + \alpha_i + \beta_{j(i)}) + \sum_{k=1}^{c_{ij}} n_{ijk}\gamma_{k(ij)}\\
&\qquad + \sum_{k=1}^{c_{ij}}\sum_{\ell=1}^{n_{ijk}} e_{\ell(ijk)}\Bigg]^2
\end{aligned}
$$

$$= \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij.}^{-1}\left[n_{ij.}^2(\mu^2+\sigma_\alpha^2+\sigma_\beta^2)+\sum_{k=1}^{c_{ij}} n_{ijk}^2\sigma_\gamma^2 + n_{ij.}\sigma_e^2\right]$$
$$= \sum_{i=1}^{a}\sum_{j=1}^{b_i}\left[n_{ij.}(\mu^2+\sigma_\alpha^2+\sigma_\beta^2)+\sum_{k=1}^{c_{ij}}\frac{n_{ijk}^2}{n_{ij.}}\sigma_\gamma^2+\sigma_e^2\right]$$
$$= N(\mu^2+\sigma_\alpha^2+\sigma_\beta^2)+k_6\sigma_\gamma^2+b_.\sigma_e^2,$$

and

$$E(T_C) = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} E(y_{ijk.}^2/n_{ijk})$$
$$= \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n_{ijk}^{-1}E\left[n_{ijk}(\mu+\alpha_i+\beta_{j(i)}+\gamma_{k(ij)})+\sum_{\ell=1}^{n_{ijk}} e_{\ell(ijk)}\right]^2$$
$$= \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n_{ijk}^{-1}\left[n_{ijk}^2(\mu^2+\sigma_\alpha^2+\sigma_\beta^2+\sigma_\gamma^2)+n_{ijk}\sigma_e^2\right]$$
$$= \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}}\left[n_{ijk}(\mu^2+\sigma_\alpha^2+\sigma_\beta^2+\sigma_\gamma^2)+\sigma_e^2\right]$$
$$= N(\mu^2+\sigma_\alpha^2+\sigma_\beta^2+\sigma_\gamma^2)+c_{..}\sigma_e^2,$$

where

$$k_1 = \sum_{i=1}^{a} n_{i..}^2/N, \qquad k_2 = \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij.}^2/N,$$
$$k_3 = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n_{ijk}^2/N, \qquad k_4 = \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij.}^2/n_{i..},$$
$$k_5 = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n_{ijk}^2/n_{i..}, \quad \text{and} \quad k_6 = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n_{ijk}^2/n_{ij.}.$$

Hence, expected values of sums of squares and mean squares are given as follows:

$$E(\mathrm{SS}_E) = E[T_0 - T_C] = (N - c_{..})\sigma_e^2,$$
$$E(\mathrm{MS}_E) = \frac{1}{N-c_{..}}E(\mathrm{SS}_E)$$
$$= \sigma_e^2,$$
$$E(\mathrm{SS}_C) = E[T_C - T_B]$$

$$
\begin{aligned}
&= (c_{..} - b_{.})\sigma_e^2 + (N - k_6)\sigma_\gamma^2, \\
E(\mathrm{MS}_C) &= \frac{1}{c_{..} - b_{.}} E(\mathrm{SS}_C) \\
&= \sigma_e^2 + r_1\sigma_\gamma^2, \\
E(\mathrm{SS}_B) &= E[T_B - T_A] \\
&= (b_{.} - a)\sigma_e^2 + (k_6 - k_5)\sigma_\gamma^2 + (N - k_4)\sigma_\beta^2, \\
E(\mathrm{MS}_B) &= \frac{1}{b_{.} - a} E(\mathrm{SS}_B) \\
&= \sigma_e^2 + r_2\sigma_\gamma^2 + r_3\sigma_\beta^2, \\
E(\mathrm{SS}_A) &= E[T_A - T_\mu] \\
&= (a - 1)\sigma_e^2 + (k_5 - k_3)\sigma_\gamma^2 + (k_4 - k_2)\sigma_\beta^2 + (N - k_1)\sigma_\alpha^2,
\end{aligned}
$$

and

$$
\begin{aligned}
E(\mathrm{MS}_A) &= \frac{1}{a - 1} E(\mathrm{SS}_A) \\
&= \sigma_e^2 + r_4\sigma_\gamma^2 + r_5\sigma_\beta^2 + r_6\sigma_\alpha^2,
\end{aligned}
$$

where

$$
\begin{aligned}
r_1 &= \frac{N - k_6}{c_{..} - b_{.}}, \qquad r_2 = \frac{k_6 - k_5}{b_{.} - a}, \qquad r_3 = \frac{N - k_4}{b_{.} - a}, \\
r_4 &= \frac{k_5 - k_3}{a - 1}, \qquad r_5 = \frac{k_4 - k_2}{a - 1}, \quad \text{and} \quad r_6 = \frac{N - k_1}{a - 1}.
\end{aligned}
$$

The expected mean squares were first given by Ganguli (1941) and detailed derivations are also given in King and Henderson (1954), Mahamunulu (1963), and Leone et al. (1968). Gaylor and Hartwell (1969) give a general algorithm for expected mean square which is applicable to both finite and infinite populations.

16.4 UNWEIGHTED MEANS ANALYSIS

In the unweighted means analysis, the mean squares are obtained using the unweighted means of the observations. In particular, let

$$
\bar{y}^*_{ijk.} = \sum_{\ell=1}^{n_{ijk}} y_{ijk\ell}/n_{ijk}, \qquad \bar{y}^*_{ij..} = \sum_{k=1}^{c_{ij}} \bar{y}^*_{ijk.}/c_{ij},
$$

$$
\bar{y}^*_{i...} = \sum_{j=1}^{b_i} \bar{y}^*_{ij..}/b_i \quad \text{and} \quad \bar{y}^*_{....} = \sum_{i=1}^{a} \bar{y}^*_{i...}/a.
$$

Then the unweighted sums of squares are defined as follows:

$$\begin{aligned}
\mathrm{SS}_{Au} &= r_6^* \sum_{i=1}^{a} (\bar{y}_{i...}^* - \bar{y}_{....}^*)^2 \\
\mathrm{SS}_{Bu} &= r_3^* \sum_{i=1}^{a} \sum_{j=1}^{b_i} (\bar{y}_{ij..}^* - \bar{y}_{i...}^*)^2, \\
\mathrm{SS}_{Cu} &= r_1^* \sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} (\bar{y}_{ijk.}^* - \bar{y}_{ij..}^*)^2,
\end{aligned} \tag{16.4.1}$$

and

$$\mathrm{SS}_E = \sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \sum_{\ell=1}^{n_{ijk}} (y_{ijk\ell} - \bar{y}_{ijk.}^*)^2,$$

where

$$\begin{aligned}
r_1^* &= 1 \Big/ \left[\frac{1}{c_{..} - b_.} \sum_{i=1}^{a} \sum_{j=1}^{b_i} \frac{1}{\bar{n}_{ij}} (c_{ij} - 1) \right], \\
r_3^* &= 1 \Big/ \left[\frac{1}{b_. - a} \sum_{i=1}^{a} \frac{b_i - 1}{b_i} \left(\sum_{j=1}^{b_i} \frac{1}{\bar{n}_{ij} c_{ij}} \right) \right],
\end{aligned} \tag{16.4.2}$$

and

$$r_6^* = 1 \Big/ \left[\frac{1}{a} \sum_{i=1}^{a} \frac{1}{b_i^2} \left(\sum_{j=1}^{b_i} \frac{1}{\bar{n}_{ij} c_{ij}} \right) \right],$$

with

$$\bar{n}_{ij} = 1 \Big/ \left[\frac{1}{c_{ij}} \sum_{k=1}^{c_{ij}} \frac{1}{n_{ijk}} \right].$$

Note that $\bar{n}_{ij}$ represents the harmonic mean of the n_{ijk} values at the jth level of factor B within the ith level of factor A. In addition, note that the definition of SS_E is the same as in the Type I sums of squares.

The mean squares are obtained by dividing the sums of squares by the corresponding degrees of freedom. The results on expectations of the unweighted means squares are obtained as follows:

$$\begin{aligned}
E(\mathrm{MS}_{Au}) &= \sigma_e^2 + r_4^* \sigma_\gamma^2 + r_5^* \sigma_\beta^2 + r_6^* \sigma_\alpha^2, \\
E(\mathrm{MS}_{Bu}) &= \sigma_e^2 + r_2^* \sigma_\gamma^2 + r_3^* \sigma_\beta^2, \\
E(\mathrm{MS}_{Cu}) &= \sigma_e^2 + r_1^* \sigma_\gamma^2,
\end{aligned} \tag{16.4.3}$$

and

$$E(\mathrm{MS}_E) = \sigma_e^2,$$

where r_1^*, r_3^*, r_6^* are defined as in (16.4.2) and

$$r_2^* = \frac{1/\sum_{i=1}^{a} \frac{b_i-1}{b_i}\left(\sum_{j=1}^{b_i} \frac{1}{\bar{n}_{ij}c_{ij}}\right)}{1/\sum_{i=1}^{a} \frac{b_i-1}{\bar{c}_i}},$$

$$r_4^* = \frac{1/\left[\frac{1}{a}\sum_{i=1}^{a} \frac{1}{b_i^2}\left(\sum_{j=1}^{b_i} \frac{1}{\bar{n}_{ij}c_{ij}}\right)\right]}{1/\left[\frac{1}{a}\sum_{i=1}^{a} \frac{1}{\bar{c}_i b_i}\right]},$$

and

$$r_5^* = \frac{\left[\sum_{i=1}^{a} \frac{1}{b_i} / \sum_{i=1}^{a} \frac{1}{\bar{c}_i b_i}\right]\left[\frac{1}{a}\sum_{i=1}^{a} \frac{1}{\bar{c}_i b_i}\right]}{\left[\frac{1}{a}\sum_{i=1}^{a} \frac{1}{b_i^2}\left(\sum_{j=1}^{b_i} \frac{1}{\bar{n}_{ij}c_{ij}}\right)\right]},$$

with

$$\bar{c}_i = 1/\left[\frac{1}{b_i}\sum_{j=1}^{b_i} \frac{1}{c_{ij}}\right].$$

Note that $\bar{c}_i$ represents the harmonic mean of c_{ij} values at the ith level of factor A. Further, with $n_{ijk} = n$, $c_{ij} = c$, $b_{ij} = b$, $r_1^* = r_2^* = r_4^* = n$, $r_3^* = r_5^* = cn$, $r_6^* = bcn$, and SS_{Au}, SS_{Bu}, SS_{Cu}, and SS_E reduce to the sums of squares for the corresponding balanced case defined in Section 7.2. Finally, the analysis of variance table for the unweighted means analysis is shown in Table 16.2.

16.5 ESTIMATION OF VARIANCE COMPONENTS

In this section, we briefly consider some methods of estimation of the variance components σ_e^2, σ_γ^2, σ_β^2, and σ_α^2.

16.5.1 ANALYSIS OF VARIANCE ESTIMATORS

The analysis of variance estimates are obtained by equating each sum of squares or equivalently the mean square in the analysis of variance Table 16.1 to its expected value and solving the resultant equations for the variance components. Denoting the estimators as $\hat{\sigma}_{\alpha,\mathrm{ANOV}}^2$, $\hat{\sigma}_{\beta,\mathrm{ANOV}}^2$, $\hat{\sigma}_{\gamma,\mathrm{ANOV}}^2$, and $\hat{\sigma}_{e,\mathrm{ANOV}}^2$, the equations to be solved are

TABLE 16.2 Analysis of variance with unweighted sums of squares for the model in (16.1.1).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Factor *A*	$a-1$	SS_{Au}	MS_{Au}	$\sigma_e^2 + r_4^*\sigma_\gamma^2 + r_5^*\sigma_\beta^2 + r_6^*\sigma_\alpha^2$
Factor *B* within *A*	$b_. - a$	SS_{Bu}	MS_{Bu}	$\sigma_e^2 + r_2^*\sigma_\gamma^2 + r_3^*\sigma_\beta^2$
Factor *C* within *B*	$c_{..} - b_.$	SS_{Cu}	MS_{Cu}	$\sigma_e^2 + r_1^*\sigma_\gamma^2$
Error	$N - c_{..}$	SS_E	MS_E	σ_e^2

$$
\begin{aligned}
MS_A &= \hat{\sigma}^2_{e,\text{ANOV}} + r_4\hat{\sigma}^2_{\gamma,\text{ANOV}} + r_5\hat{\sigma}^2_{\beta,\text{ANOV}} + r_6\hat{\sigma}^2_{\alpha,\text{ANOV}},\\
MS_B &= \hat{\sigma}^2_{e,\text{ANOV}} + r_2\hat{\sigma}^2_{\gamma,\text{ANOV}} + r_3\hat{\sigma}^2_{\beta,\text{ANOV}},\\
MS_C &= \hat{\sigma}^2_{e,\text{ANOV}} + r_1\hat{\sigma}^2_{\gamma,\text{ANOV}},
\end{aligned}
\tag{16.5.1}
$$

and

$$
MS_E = \hat{\sigma}^2_{e,\text{ANOV}}.
$$

The solution to (16.5.1) yields the following estimators:

$$
\begin{aligned}
\hat{\sigma}^2_{e,\text{ANOV}} &= MS_E,\\
\hat{\sigma}^2_{\gamma,\text{ANOV}} &= \frac{1}{r_1}(MS_C - MS_E),\\
\hat{\sigma}^2_{\beta,\text{ANOV}} &= \frac{1}{r_3}(MS_B - r_2\hat{\sigma}^2_{\gamma,\text{ANOV}} - \hat{\sigma}^2_{e,\text{ANOV}}),
\end{aligned}
\tag{16.5.2}
$$

and

$$
\hat{\sigma}^2_{\alpha,\text{ANOV}} = \frac{1}{r_6}(MS_A - r_5\hat{\sigma}^2_{\beta,\text{ANOV}} - r_4\hat{\sigma}^2_{\gamma,\text{ANOV}} - \hat{\sigma}^2_{e,\text{ANOV}}).
$$

The estimator $\hat{\sigma}^2_{e,\text{ANOV}}$ is the minimum variance unbiased estimator under the assumption of normality, but other estimators lack any optimal property other than unbiasedness.

16.5.2 UNWEIGHTED MEANS ESTIMATORS

The unweighted means estimators are obtained by equating the unweighted mean squares in Table 16.2 to the corresponding expected values. Denoting the estimators as $\hat{\sigma}^2_{e,\text{UNME}}$, $\hat{\sigma}^2_{\gamma,\text{UNME}}$, $\hat{\sigma}^2_{\beta,\text{UNME}}$, and $\hat{\sigma}^2_{\alpha,\text{UNME}}$, the resulting equations are

$$\begin{aligned}
\text{MS}_{Au} &= \hat{\sigma}^2_{e,\text{UNME}} + r_4^*\hat{\sigma}^2_{\gamma,\text{UNME}} + r_5^*\hat{\sigma}^2_{\beta,\text{UNME}} + r_6^*\hat{\sigma}^2_{\alpha,\text{UNME}},\\
\text{MS}_{Bu} &= \hat{\sigma}^2_{e,\text{UNME}} + r_2^*\hat{\sigma}^2_{\gamma,\text{UNME}} + r_3^*\hat{\sigma}^2_{\beta,\text{UNME}},\\
\text{MS}_{Cu} &= \hat{\sigma}^2_{e,\text{UNME}} + r_1^*\hat{\sigma}^2_{\gamma,\text{UNME}},
\end{aligned} \tag{16.5.3}$$

and

$$\text{MS}_E = \hat{\sigma}^2_{e,\text{UNME}}.$$

Solving the equations in (16.5.3), we obtain the following estimators:

$$\begin{aligned}
\hat{\sigma}^2_{e,\text{UNME}} &= \text{MS}_E,\\
\hat{\sigma}^2_{\gamma,\text{UNME}} &= \frac{1}{r_1^*}(\text{MS}_{Cu} - \text{MS}_E),\\
\hat{\sigma}^2_{\beta,\text{UNME}} &= \frac{1}{r_3^*}(\text{MS}_{Bu} - r_2^*\hat{\sigma}^2_{\gamma,\text{UNME}} - \hat{\sigma}^2_{e,\text{UNME}}),
\end{aligned} \tag{16.5.4}$$

and

$$\hat{\sigma}^2_{\alpha,\text{UNME}} = \frac{1}{r_6^*}(\text{MS}_{Au} - r_5^*\hat{\sigma}^2_{\beta,\text{UNME}} - r_4^*\hat{\sigma}^2_{\gamma,\text{UNME}} - \hat{\sigma}^2_{e,\text{UNME}}).$$

Note that the ANOVA and the unweighted means estimators for σ_e^2 are the same.

16.5.3 SYMMETRIC SUMS ESTIMATORS

For symmetric sums estimators we consider expected values for products and squares of differences of observations. From the model in (16.1.1), the expected values of products of the observations are

$$E(y_{ijk\ell}y_{i'j'k'\ell'}) = \begin{cases}
\mu^2, & i \neq i',\\
\mu^2 + \sigma_\alpha^2, & i = i',\ \ j \neq j',\\
\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2, & i = i',\ \ j = j',\ \ k \neq k',\\
\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2, & i = i',\ \ j = j',\ \ k = k',\ \ \ell \neq \ell',\\
\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 + \sigma_e^2, & i = i',\ \ j = j',\ \ k = k',\ \ \ell = \ell',
\end{cases} \tag{16.5.5}$$

where $i, i' = 1, 2, \ldots, a$; $j = 1, 2, \ldots, b_i$; $j' = 1, 2, \ldots, b_{i'}$; $k = 1, 2, \ldots, c_{ij}$; $k' = 1, 2, \ldots, c_{i'j'}$; $\ell = 1, 2, \ldots, n_{ijk}$; $\ell' = 1, 2, \ldots, n_{i'j'k'}$. Now, the normalized symmetric sums of the terms in (16.5.5) are

$$g_m = \frac{\sum_{\substack{i,i' \\ i \neq i'}} y_{i\ldots} y_{i'\ldots}}{\sum_{i=1}^{a} n_{i..}(N - n_{i..})} = \frac{\left(y_{....}^2 - \sum_{i=1}^{a} y_{i\ldots}^2\right)}{N^2 - k_3},$$

$$g_A = \frac{\sum_{i=1}^{a} \sum_{\substack{j,j' \\ j \neq j'}} y_{ij..} y_{ij'..}}{\sum_{i=1}^{a} \sum_{j=1}^{b_i} n_{ij.}(n_{i..} - n_{ij.})} = \frac{\sum_{i=1}^{a} \left(y_{i\ldots}^2 - \sum_{j=1}^{b_i} y_{ij..}^2\right)}{k_3 - k_2},$$

$$g_B = \frac{\sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{\substack{k,k' \\ k \neq k'}} y_{ijk.} y_{ijk'.}}{\sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}(n_{ij.} - n_{ijk})}$$

$$= \frac{\sum_{i=1}^{a} \sum_{j=1}^{b_i} \left(y_{ij..}^2 - \sum_{k=1}^{c_{ij}} y_{ijk.}^2\right)}{k_2 - k_1},$$

$$g_C = \frac{\sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \sum_{\substack{\ell,\ell' \\ \ell \neq \ell'}} y_{ijk\ell} y_{ijk\ell'}}{\sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}(n_{ijk} - 1)}$$

$$= \frac{\sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \left(y_{ijk.}^2 - \sum_{\ell=1}^{n_{ijk}} y_{ijk\ell}^2\right)}{k_1 - N},$$

and

$$g_E = \frac{\sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \sum_{\ell=1}^{n_{ijk}} y_{ijk\ell} y_{ijk\ell}}{\sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}}$$

$$= \frac{\sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \sum_{l=1}^{n_{ijk}} y_{ijk\ell}^2}{N},$$

where

$$n_{ij.} = \sum_{k=1}^{c_{ij}} n_{ijk}, \qquad n_{i..} = \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}, \qquad N = \sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk},$$

$$k_1 = \sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}^2, \qquad k_2 = \sum_{i=1}^{a} \sum_{j=1}^{b_i} n_{ij.}^2, \qquad k_3 = \sum_{i=1}^{a} n_{i..}^2.$$

Equating g_m, g_A, g_B, g_C, and g_E to their respective expected values, we obtain

$$\begin{aligned} \mu^2 &= g_m, \\ \mu^2 + \sigma_\alpha^2 &= g_A, \\ \mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 &= g_B, \\ \mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 &= g_C, \end{aligned} \tag{16.5.6}$$

and

$$\mu^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 + \sigma_e^2 = g_E.$$

The variance component estimators obtained by solving the equations in (16.5.6) are (Koch, 1967)

$$\begin{aligned}\hat{\sigma}^2_{\alpha,\text{SSP}} &= g_A - g_m,\\ \hat{\sigma}^2_{\beta,\text{SSP}} &= g_B - g_A,\\ \hat{\sigma}^2_{\gamma,\text{SSP}} &= g_C - g_B,\end{aligned} \tag{16.5.7}$$

and

$$\hat{\sigma}^2_{e,\text{SSP}} = g_E - g_C.$$

The estimators in (16.5.7) are, by construction, unbiased, and they reduce to the analysis of variance estimators in the case of balanced data. However, they are not translation invariant, i.e., they may change in values if the same constant is added to all the observations and their variances are functions of μ. This drawback is overcome by using the symmetric sums of squares of differences rather than the products.

From the model in (16.1.1), the expected values of squares of differences of the observations are

$$E[(y_{ijk\ell} - y_{i'j'k'\ell'})^2] = \begin{cases} 2\sigma_e^2, & i = i',\ \ j = j',\ \ k = k',\ \ \ell \neq \ell',\\ 2(\sigma_e^2 + \sigma_\gamma^2), & i = i',\ \ j = j',\ \ k \neq k',\\ 2(\sigma_e^2 + \sigma_\gamma^2 + \sigma_\beta^2), & i = i',\ \ j \neq j',\\ 2(\sigma_e^2 + \sigma_\gamma^2 + \sigma_\beta^2 + \sigma_\alpha^2), & i \neq i'. \end{cases} \tag{16.5.8}$$

The normalized (mean) symmetric sums of the terms in (16.5.8) are given by

$$\begin{aligned} h_E &= \frac{\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \sum_{\substack{\ell,\ell'\\ \ell\neq\ell'}} (y_{ijk\ell} - y_{ijk\ell'})^2}{\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}(n_{ijk} - 1)}\\ &= \frac{2\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}\left(\sum_{\ell=1}^{n_{ijk}} y_{ijk\ell}^2 - n_{ijk}\bar{y}_{ijk.}^2\right)}{k_1 - N},\\ h_C &= \frac{\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{\substack{k,k'\\ k\neq k'}} \sum_{\ell,\ell'} (y_{ijk\ell} - y_{ijk'\ell'})^2}{\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}(n_{ij.} - n_{ijk})}\\ &= \frac{2\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} (n_{ij.} - n_{ijk}) \sum_{\ell=1}^{n_{ijk}} y_{ijk\ell}^2}{k_2 - k_1} - 2g_B, \end{aligned}$$

$$h_B = \frac{\sum_{i=1}^{a} \sum_{\substack{j,j' \\ j \neq j'}} \sum_{k,k'} \sum_{\ell,\ell'} (y_{ijk\ell} - y_{ij'k'\ell'})^2}{\sum_{i=1}^{a} \sum_{j=1}^{b_i} n_{ij.}(n_{i..} - n_{ij.})}$$
$$= \frac{2\sum_{i=1}^{a} \sum_{j=1}^{b_i} (n_{i..} - n_{ij.}) \sum_{k=1}^{c_{ij}} \sum_{\ell=1}^{n_{ijk}} y_{ijk\ell}^2}{k_3 - k_2} - 2g_A,$$

and

$$h_A = \frac{\sum_{\substack{i,i' \\ i \neq i'}} \sum_{j,j'} \sum_{k,k'} \sum_{\ell,\ell'} (y_{ijk\ell} - y_{i'j'k'\ell'})^2}{\sum_{i=1}^{a} n_{i..}(N - n_{i..})}$$
$$= \frac{2\sum_{i=1}^{a} (N - n_{i..}) \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \sum_{\ell=1}^{n_{ijk}} y_{ijk\ell}^2}{N^2 - k_3} - 2g_m,$$

where $n_{ij.}$, $n_{i..}$, N, k_1, k_2, k_3, g_m, and g_A are defined as before.

Equating h_A, h_B, h_C, and h_E to their respective expected values, we obtain

$$\begin{aligned} 2\sigma_e^2 &= h_E, \\ 2(\sigma_e^2 + \sigma_\gamma^2) &= h_C, \\ 2(\sigma_e^2 + \sigma_\gamma^2 + \sigma_\beta^2) &= h_B, \end{aligned} \tag{16.5.9}$$

and

$$2(\sigma_e^2 + \sigma_\gamma^2 + \sigma_\beta^2 + \sigma_\alpha^2) = h_A.$$

The variance component estimators obtained by solving the equations in (16.5.9) are (Koch, 1968)

$$\begin{aligned} \hat{\sigma}_{e,\text{SSD}}^2 &= \frac{1}{2} h_E, \\ \hat{\sigma}_{\gamma,\text{SSD}}^2 &= \frac{1}{2}(h_C - h_E), \\ \hat{\sigma}_{\beta,\text{SSD}}^2 &= \frac{1}{2}(h_B - h_C), \end{aligned} \tag{16.5.10}$$

and

$$\hat{\sigma}_{\alpha,\text{SSD}}^2 = \frac{1}{2}(h_A - h_B).$$

It can be readily seen that if the model in (16.1.1) is balanced, i.e., $b_i = b$, $c_{ij} = c$, $n_{ijk} = n$, then the estimators (16.5.10) reduce to the usual analysis of variance estimators.

16.5.4 OTHER ESTIMATORS

The ML, REML, MINQUE, and MIVQUE estimators can be developed as special cases of the results for the general case considered in Chapter 10 and their special formulations for this model are not amenable to any simple algebraic expressions. With the advent of the high-speed digital computer, the general results on these estimators involving matrix operations can be handled with great speed and accuracy and their explicit algebraic evaluation for this model seems to be rather unnecessary. In addition, some commonly used statistical software packages, such as SAS®, SPSS®, and BMDP®, have special routines to compute these estimates rather conveniently simply by specifying the model in question.

16.5.5 A NUMERICAL EXAMPLE

Consider a three-way nested analysis of variance described by Bliss (1967, pp. 352–357). The data come from an experiment reported by Sharpe and van Middelem (1955) on measurements of insecticide residue on celery. A parathion solution was uniformly sprayed on 11 field plots of celery selected in a three-stage nested design and, at maturity, three 10-plant samples (I,II,III) were collected from each plot. After each sample had been selected, chopped and mixed, two subsamples were taken from Sample I and the parathion content in parts per million (ppm) was determined from two aliquots from each subsample. Two subsamples were also taken from Sample II, but only one determination made from each subsample. Sample III was analyzed by a single subsample and determination on residue made. Consequently, the seven residue determinations from each plot accounted for three 10-plant samples. The data are shown in Table 16.3 where the observations ys are in units of $y = \text{ppm} - 0.70$.

We will use the three-way nested model in (16.1.1) to analyze the data in Table 16.3. Here, $i = 1, 2, \ldots, 11$ refer to the plots, $j = 1, 2, \ldots, b_i$ refer to the samples within plots, $k = 1, 2, \ldots, c_{ij}$ refer to the subsamples within samples, and $\ell = 1, 2, \ldots, n_{ijk}$ refer to measurements of residue on celery. Further, $\sigma_\alpha^2, \sigma_\beta^2, \sigma_\gamma^2$ designate variance components due to plot, sample, and subsample as factors, and σ_e^2 denotes the error variance component. The calculations leading to the conventional analysis of variance using Type I sums of squares are readily performed and the results are summarized in Table 16.4. The selected outputs using SAS®GLM, SPSS®GLM, and BMDP®3V are displayed in Figure 16.1.

We now illustrate the calculations of point estimates of the variance components $\sigma_\alpha^2, \sigma_\beta^2, \sigma_\gamma^2, \sigma_e^2$, and certain of their parametric functions.

The analysis of variance (ANOVA) estimates based on Henderson's Method I are obtained as the solution to the following system of equations:

TABLE 16.3 The insecticide residue on celery from plants sprayed with parathion solution.

Plot	1					2					3				
Sample	1		2		3	1		2		3	1		2		3
Subsample	1	2	1	2	1	1	2	1	2	1	1	2	1	2	1
Residue	.52	.40	.26	.54	.52	.50	.47	.04	.43	1.08	.34	.32	.25	.38	.29
	.43	.52				.59	.50				.26	.45			

Plot	4					5					6				
Sample	1		2		3	1		2		3	1		2		3
Subsample	1	2	1	2	1	1	2	1	2	1	1	2	1	2	1
Residue	.18	.31	.13	.25	.10	1.05	.60	.95	.84	.92	.52	.55	.33	.26	.41
	.24	.29				.66	.51				.66	.40			

Plot	7					8					9				
Sample	1		2		3	1		2		3	1		2		3
Subsample	1	2	1	2	1	1	2	1	2	1	1	2	1	2	1
Residue	.77	.51	.44	.50	.44	.89	.75	.64	.54	.36	.50	.60	.60	.71	.92
	.56	.60				.92	.58				.67	.53			

Plot	10					11				
Sample	1		2		3	1		2		3
Subsample	1	2	1	2	1	1	2	1	2	1
Residue	.58	.56	.46	.52	.52	.24	.48	.53	.50	.39
	.52	.44				.36	.30			

Source: Bliss (1967); used with permission.

TABLE 16.4 Analysis of variance for the insecticide residue data of Table 16.3.

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Plots	10	1.84041	0.18404	$\sigma_e^2 + 1.571\sigma_\gamma^2 + 3.000\sigma_\beta^2 + 7.000\sigma_\alpha^2$
Samples within plots	22	0.99175	0.04508	$\sigma_e^2 + 1.214\sigma_\gamma^2 + 2.000\sigma_\beta^2$
Subsamples within samples	22	0.35758	0.01625	$\sigma_e^2 + 1.500\sigma_\gamma^2$
Error	22	0.22085	0.01004	σ_e^2
Total	76	3.41058		

```
DATA SAHAIC16;
INPUT PLOT SAMPLE
 SUBSAMPL RESIDUE;
CARDS;
 1 1 1 0.52
 1 1 1 0.43
 1 1 2 0.40
 1 1 2 0.52
 1 2 1 0.26
 1 2 2 0.54
 1 3 1 0.52
 2 1 1 0.50
 2 1 1 0.59
 2 1 2 0.47
 2 1 2 0.50
 2 2 1 0.04
 2 2 2 0.43
 2 3 1 1.08
 3 1 1 0.34
 3 1 1 0.26
 3 1 2 0.32
 3 1 2 0.45
 3 2 1 0.25
 3 2 2 0.38
 3 3 1 0.29
 . . .  .
11 3 1 0.39
;
PROC GLM;
CLASS PLOT SAMPLE
 SUBSAMPL;
MODEL RESIDUE = PLOT
 SAMPLE(PLOT)
 SUBSAMPL(SAMPLE)
PLOT); RANDOM PLOT
 SAMPLE(PLOT)
 SUBSAMPL(SAMPLE
 PLOT)/TEST;
RUN;
CLASS  LEVELS  VALUES
PLOT      11   1 2 … 11
SAMPLE     3    1 2 3
SUBSAMPL   2    1 2
NUMBER OF OBSERVATIONS
IN DATA SET=77
```

```
                              The SAS System
                      General Linear Models Procedure
Dependent Variable: RESIDUE
                                  Sum of        Mean
Source                   DF      Squares       Square     F Value    Pr > F
Model                    54   3.18972922    0.05906906     5.88      0.0001
Error                    22   0.22085000    0.01003864
Corrected                76   3.41057922
Total
               R-Square      C.V.        Root MSE      RESIDUE Mean
               0.935246    19.94535     0.10019300      0.50233766
Source                    DF      Type I SS     Mean Square   F Value   Pr > F
PLOT                      10      1.84040779    0.18404078     18.33    0.0001
SAMPLE(PLOT)              22      0.99174643    0.04507938      4.49    0.0004
SUBSAMPL(PLOT*SAMPLE)     22      0.35757500    0.01625341      1.62    0.1331

Source                    DF      Type III SS   Mean Square   F Value   Pr > F
PLOT                      10      1.78685974    0.17868597     17.80    0.0001
SAMPLE(PLOT)              22      0.99174643    0.04507938      4.49    0.0004
SUBSAMPL(PLOT*SAMPLE)     22      0.35757500    0.01625341      1.62    0.1331

Source                   Type III Expected Mean Square
PLOT                Var(Error) + 1.1905 Var(SUBSAMPL(PLOT*SAMPLE))
                    + 1.9524 Var(SAMPLE(PLOT)) + 5.7619 Var(PLOT)
SAMPLE(PLOT)        Var(Error) + 1.2143 Var(SUBSAMPL(PLOT*SAMPLE))
                    + 2 Var(SAMPLE(PLOT))
SUBSAMPL(PLOT*SAMPLE) Var(Error) + 1.5 Var(SUBSAMPL(PLOT*SAMPLE))

          Tests of Hypotheses for Random Model Analysis of Variance
Source: PLOT          Error: 0.9762*MS(SAMPLE(PLOT)) +
                      0.0034*MS(SUBSAMPL(PLOT*SAMPLE)) + 0.0204*MS(Error)
                         Denominator   Denominator
   DF     Type III MS        DF             MS          F Value     Pr > F
   10     0.178685974      22.26      0.0442662183      4.0366      0.0030

Source: SAMPLE(PLOT)  Error:0.8095*MS(SUBSAMPL(PLOT*SAMPLE)) +
                      0.1905*MS(Error)
                         Denominator   Denominator
   DF     Type III MS        DF             MS          F Value     Pr > F
   22     0.0450793831     28.26      0.0150696429      2.9914      0.0034

Source: SUBSAMPL(PLOT*SAMPLE) Error: MS(Error)
                         Denominator   Denominator
   DF     Type III MS        DF             MS          F Value     Pr > F
   22     0.0162534091       22       0.0100386364      1.6191      0.1331
```

SAS application: This application illustrates SAS GLM instructions and output for the unbalanced three-way nested random effects analysis of variance.[a,b]

[a] Several portions of the output were extensively edited and doctored to economize space and may not correspond to the original printout.

[b] Results on significance tests may vary from one package to the other.

FIGURE 16.1 Program instructions and output for the unbalanced three-way nested random effects analysis of variance: Data on insecticide residue on celery from plants sprayed with parathion solution (Table 16.3).

```
DATA SAHAIC16
/PLOT 1
 SAMPLE 3 SUBSAMPL 5
 RESIDUE 7-9.
BEGIN DATA.
 1 1 1 0.52
 1 1 1 0.43
 1 1 2 0.40
 1 1 2 0.52
 1 2 1 0.26
 2 2 2 0.54
 . . .  .
11 3 1 0.39
END DATA.
GLM RESIDUE BY
 PLOT SAMPLE
 SUBSAMPL
/DESIGN PLOT
 SAMPLE(PLOT)
 SUBSAMPL(SAMPLE
 (PLOT))
/METHOD SSTYPE(1)
/RANDOM PLOT
 SAMPLE SUBSAMPL.
```

```
                      Tests of Between-Subjects Effects
Dependent Variable: RESIDUE

Source                     Type I SS   df        Mean Square     F       Sig
PLOT         Hypothesis      1.840     10        0.184         2.989    0.021
             Error           1.118     18.163    6.156E-02(a)
SAMPLE(PLOT) Hypothesis      0.992     22        4.508E-02     2.991    0.003
             Error           0.426     28.262    1.507E-02(b)
SUBSAMPL     Hypothesis      0.358     22        1.625E-02     1.619    0.133
(SAMPLE(PLOT)) Error         0.221     22        1.004E-02(c)
a 1.500MS(SAMPLE(PLOT))-0.167MS(SUBSAMPL(SAMPLE(PLOT)))-0.333MS(Error)
b 0.810MS(SUBSAMPL(SAMPLE(PLOT)))+ 0.190 MS(Error)
c MS(Error)
                                     Expected Mean Squares(d,e)
                                         Variance Component
Source                        Var(P)       Var(S(P))   Var(SB(S(P)))   Var(Error)
PLOT                          7.000          3.000        1.571          1.000
SAMPLE(PLOT)                  0.000          2.000        1.214          1.000
SUBSAMPL(SAMPLE(PLOT))        0.000          0.000        1.500          1.000
Error                         0.000          0.000        0.000          1.000
d  For  each  source,  the  expected  mean  square  equals  the  sum  of  the
   coefficients in the cells times the variance components, plus a quadratic
   term involving effects in the Quadratic Term  cell.
e  Expected Mean Squares are based on the Type I Sums of Squares.
```

SPSS application: This application illustrates SPSS GLM instructions and output for the unbalanced three-way nested random effects analysis of variance.[a,b]

```
/INPUT     FILE='C:\SAHAIC16.TXT'.
           FORMAT=FREE.
           VARIABLES=4.
/VARIABLE  NAMES=PLOT,SAMPLE,CUBSAMPLE,
                 RESIDUE.
/GROUP     CODES(PLOT)=1,2,…,11.
           NAMES(PLOT)=P1,P2,…,P11.
           CODES(SAMPLE)=1,2,3.
           NAMES(SAMPLE)=S1,S2,S3.
           CODES(CUBSAMPLE)=1,2.
           NAMES(CUBSAMPLE)=C1,C2.
/DESIGN    DEPENDENT=RESIDUE.
           RANDOM=PLOT.
           RANDOM=SAMPLE, PLOT.
           RANDOM=CUBSAMPLE, SAMPLE, PLOT.
           RNAMES=P,'S(P)','C(S)'.
           METHOD=REML.
/END
 1 1 1 0.52
 . . .  .
11 3 1 0.39
```

```
  BMDP3V - GENERAL MIXED MODEL ANALYSIS OF VARIANCE
         Release: 7.0        (BMDP/DYNAMIC)

        DEPENDENT VARIABLE RESIDUE

PARAMETER  ESTIMATE   STANDARD   ST/     TWO-TAIL PROB.
                       ERROR    ST.DEV. (ASYM. THEORY)

ERR.VAR.   0.010444   0.003216
CONSTANT   0.502113   0.052526   9.559      0.000
P          0.022343   0.013709
S(P)       0.015550   0.008427
C(S)       0.004598   0.004278

 TESTS OF FIXED EFFECTS BASED ON ASYMPTOTIC VARIANCE
                -COVARIANCE MATRIX

SOURCE         F-STATISTIC   DEGREES OF    PROBABILITY
                               FREEDOM

CONSTANT          91.38       1    76        0.00000
```

BMDP application: This application illustrates BMDP 3V instructions and output for the unbalanced three-way nested random effects analysis of variance.[a,b]

[a] Several portions of the output were extensively edited and doctored to economize space and may not correspond to the original printout.

[b] Results on significance tests may vary from one package to the other.

FIGURE 16.1 (*continued*)

$$\sigma_e^2 = 0.01004,$$
$$\sigma_e^2 + 1.500\sigma_\gamma^2 = 0.01625,$$
$$\sigma_e^2 + 1.214\sigma_\gamma^2 + 2.000\sigma_\beta^2 = 0.04508,$$

and

$$\sigma_e^2 + 1.571\sigma_\gamma^2 + 3.000\sigma_\beta^2 + 7.000\sigma_\alpha^2 = 0.18404.$$

TABLE 16.5 Analysis of variance for the insecticide residue data of Table 16.3 (unweighted sums of squares).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Plots	10	1.7073	0.1707	$\sigma_e^2 + 1.143\sigma_\gamma^2 + 1.714\sigma_\beta^2 + 5.143\sigma_\alpha^2$
Samples within plots	22	1.0894	0.0495	$\sigma_e^2 + 1.143\sigma_\gamma^2 + 1.714\sigma_\beta^2$
Subsamples within samples	22	0.3414	0.0155	$\sigma_e^2 + 1.333\sigma_\gamma^2$
Error	22	0.2209	0.0100	σ_e^2

Therefore, the desired ANOVA estimates of the variance components are given by

$$\hat{\sigma}^2_{e,\text{ANOV}} = 0.01004,$$
$$\hat{\sigma}^2_{\gamma,\text{ANOV}} = \frac{0.01625 - 0.01004}{1.500} = 0.00414,$$
$$\hat{\sigma}^2_{\beta,\text{ANOV}} = \frac{0.04508 - 0.01004 - 1.214 \times 0.00414}{2.000} = 0.01501,$$

and

$$\hat{\sigma}^2_{\alpha,\text{ANOV}} = \frac{0.18404 - 0.01004 - 1.571 \times 0.00414 - 3.000 \times 0.01501}{7.000} = 0.01750.$$

These variance components account for 21.5%, 8.9%, 32.1%, and 37.5% of the total variation in the residue data in this experiment.

To obtain variance component estimates based on unweighted means squares, we performed analysis of variance on the cell means and the results are summarized in Table 16.5. The analysis of means estimates are obtained as the solution to the following system of equations:

$$\sigma_e^2 = 0.0100,$$
$$\sigma_e^2 + 1.333\sigma_\gamma^2 = 0.0155,$$
$$\sigma_e^2 + 1.143\sigma_\gamma^2 + 1.714\sigma_\beta^2 = 0.0495,$$

and

$$\sigma_e^2 + 1.143\sigma_\gamma^2 + 1.714\sigma_\beta^2 + 5.143\sigma_\alpha^2 = 0.1707.$$

Therefore the desired estimates are given by

$$\hat{\sigma}^2_{e,\text{UNME}} = 0.0100,$$
$$\hat{\sigma}^2_{\gamma,\text{UNME}} = \frac{0.0155 - 0.0100}{1.333} = 0.0041,$$
$$\hat{\sigma}^2_{\beta,\text{UNME}} = \frac{0.0495 - 0.0100 - 1.143 \times 0.0041}{1.714} = 0.0225,$$

and

$$\hat{\sigma}^2_{\alpha,\text{UNME}} = \frac{0.1707 - 0.0100 - 1.143 \times 0.0041 - 1.714 \times 0.0225}{5.143} = 0.0228.$$

We used SAS®VARCOMP, SPSS®VARCOMP, and BMDP®3V to estimate the variance components using the ML, REML, MINQUE(0), and MINQUE(1) procedures.[1] The desired estimates are given in Table 16.6. Note that all three software produce nearly the same results except for some minor discrepancy in rounding decimal places.

16.6 VARIANCES OF ESTIMATORS

In this section, we present some results on sampling variances of the variance components estimators.

16.6.1 VARIANCES OF ANALYSIS OF VARIANCE ESTIMATORS

In the analysis of variance given in Table 16.1, SS_E/σ_e^2 has a chi-square distribution with $N - c_{..}$ degrees of freedom. Hence, the variance of $\hat{\sigma}_e^2$ is

$$\text{Var}(\hat{\sigma}^2_{e,\text{ANOV}}) = \frac{2\sigma_e^4}{N - c_{..}}.$$

Furthermore, SS_E is distributed independently of SS_A, SS_B, and SS_C. This property of independence can be used to derive the variances of $\hat{\sigma}_\alpha^2$, $\hat{\sigma}_\beta^2$, and $\hat{\sigma}_\gamma^2$, and covariances between them and $\hat{\sigma}_e^2$. These expressions for sampling variances and covariances have been derived by Mahamunulu (1963), and the results are given as follows (see also Searle, 1971, pp. 477–479; Searle et al., 1992, pp. 431–433):

$$\text{Var}(\hat{\sigma}^2_{\gamma,\text{ANOV}}) = 2[(Nk_3 + k_{19} - 2k_{11})\sigma_\gamma^4 + 2(N - b_.)v_9\sigma_e^4/v_{10} + 2(N - k_6)\sigma_\gamma^2\sigma_e^2]/v_8^2,$$

[1]The computations for ML and REML estimates were also carried out using SAS® PROC MIXED and some other programs to assess their relative accuracy and convergence rate. There did not seem to be any appreciable differences between the results from different software.

TABLE 16.6 ML, REML, MINQUE(0), and MINQUE(1) estimates of the variance components using SAS®, SPSS®, and BMDP® software.

Variance component	SAS®		
	ML	REML	MINQUE(0)
σ_e^2	0.010464	0.010444	0.021727
σ_γ^2	0.004615	0.004598	0.001223
σ_β^2	0.015425	0.015550	0.001436
σ_α^2	0.019606	0.022342	0.002230

Variance component	SPSS®			
	ML	REML	MINQUE(0)	MINQUE(1)
σ_e^2	0.010464	0.010444	0.021727	0.010563
σ_γ^2	0.004615	0.004598	0.001223	0.005765
σ_β^2	0.015425	0.015550	0.001436	0.012859
σ_α^2	0.019606	0.022343	0.002230	0.022256

Variance component	BMDP®	
	ML	REML
σ_e^2	0.010464	0.010444
σ_γ^2	0.004615	0.004598
σ_β^2	0.015425	0.015550
σ_α^2	0.019606	0.022343

SAS®VARCOMP does not compute MINQUE(1). BMDP®3V does not compute MINQUE(0) and MINQUE(1).

$$
\begin{aligned}
\mathrm{Var}(\hat{\sigma}^2_{\beta,\mathrm{ANOV}}) &= 2(d_1\sigma_\beta^4 + d_2\sigma_\gamma^4 + d_3\sigma_e^4 + 2d_4\sigma_\beta^2\sigma_\gamma^2 + 2d_5\sigma_\beta^2\sigma_e^2 \\
&\quad + 2d_6\sigma_\gamma^2\sigma_e^2)/v_5^2v_8^2, \\
\mathrm{Var}(\hat{\sigma}^2_{\alpha,\mathrm{ANOV}}) &= 2(g_1\sigma_\alpha^4 + g_2\sigma_\beta^4 + g_3\sigma_\gamma^4 + g_4\sigma_e^4 + 2g_5\sigma_\alpha^2\sigma_\beta^2 \\
&\quad + 2g_6\sigma_\alpha^2\sigma_\gamma^2 + 2g_7\sigma_\alpha^2\sigma_e^2 + 2g_8\sigma_\beta^2\sigma_\gamma^2 + 2g_9\sigma_\beta^2\sigma_e^2 \\
&\quad + 2g_{10}\sigma_\gamma^2\sigma_e^2)/v_1^2v_5^2v_8^2, \\
\mathrm{Cov}(\hat{\sigma}^2_{\gamma,\mathrm{ANOV}}, \hat{\sigma}^2_{e,\mathrm{ANOV}}) &= -(v_9/v_8)\,\mathrm{Var}(\hat{\sigma}^2_{e,\mathrm{ANOV}}) \\
\mathrm{Cov}(\hat{\sigma}^2_{\beta,\mathrm{ANOV}}, \hat{\sigma}^2_{e,\mathrm{ANOV}}) &= -(v_7v_8 - v_6v_9)\,\mathrm{Var}(\hat{\sigma}^2_{e,\mathrm{ANOV}})/v_5v_8, \\
\mathrm{Cov}(\hat{\sigma}^2_{\alpha,\mathrm{ANOV}}, \hat{\sigma}^2_{e,\mathrm{ANOV}}) &= [v_3v_5v_9 + v_2(v_7v_8 - v_6v_9) - v_4v_5v_8] \\
&\quad \times \mathrm{Var}(\hat{\sigma}^2_{e,\mathrm{ANOV}})/v_1v_5v_8,
\end{aligned}
$$

$$\begin{aligned}
\text{Cov}(\hat{\sigma}^2_{\beta,\text{ANOV}}, \hat{\sigma}^2_{\gamma,\text{ANOV}}) &= [2(k_{11} - k_{19} + k_{18} - k_{10})\sigma^4_\gamma + 2v_7 v_9 \sigma^4_e / v_{10} \\
&\quad - v_6 v_8 \,\text{Var}(\hat{\sigma}^2_{\gamma,\text{ANOV}})] / v_5 v_8, \\
\text{Cov}(\hat{\sigma}^2_{\alpha,\text{ANOV}}, \hat{\sigma}^2_{\gamma,\text{ANOV}}) &= [2[v_5\{(k_{10} - k_{18}) - (k_9 - k_{15})/N\} \\
&\quad - v_2\{(k_{11} - k_{19}) - (k_{10} - k_{18})\}]\sigma^4_\gamma \\
&\quad + 2v_9(v_4 v_5 - v_2 v_7)\sigma^4_e / v_{10} \\
&\quad - v_8(v_3 v_5 - v_2 v_6)\,\text{Var}(\hat{\sigma}^2_{\gamma,\text{ANOV}})] / v_1 v_5 v_8,
\end{aligned}$$

and

$$\begin{aligned}
\text{Cov}(\hat{\sigma}^2_{\alpha,\text{ANOV}}, \hat{\sigma}^2_{\beta,\text{ANOV}}) &= [2\{(k_{12} - k_{22}) - (k_8 - k_{13})/N\}\sigma^4_\beta \\
&\quad + 2[(k_{18} - k_{21}) - (k_{15} - k_{14})/N \\
&\quad - v_6\{(k_{10} - k_{18}) - (k_9 - k_{15})/N\} \\
&\quad - v_3(k_{11} - k_{19} + k_{18} - k_{10})]\sigma^4_\gamma \\
&\quad + 2\{(k_{16} - k_{20}) - (k_{25} - k_{17})/N\}\sigma^2_\beta \sigma^2_\gamma \\
&\quad + 2\{v_4 v_7 v_8 - v_9(v_4 v_6 + v_3 v_7)\}\sigma^4_e / v_{10} \\
&\quad - v_2 v_5 v_8 \,\text{Var}(\hat{\sigma}^2_{\beta,\text{ANOV}}) \\
&\quad + v_3 v_6 v_8 \,\text{Var}(\hat{\sigma}^2_{\gamma,\text{ANOV}})] / v_1 v_5 v_8,
\end{aligned}$$

where $k_1, \ldots, k_6$ are defined in Section 16.3 and

$$k_7 = \sum_{i=1}^{a} n^3_{i..}, \qquad k_8 = \sum_{i=1}^{a}\sum_{j=1}^{b_i} n^3_{ij.},$$

$$k_9 = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n^3_{ijk}, \qquad k_{10} = \sum_{i=1}^{a}\left(\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n^3_{ijk}\right)/n_{i..},$$

$$k_{11} = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\left(\sum_{k=1}^{c_{ij}} n^3_{ijk}\right)/n_{ij.}, \qquad k_{12} = \sum_{i=1}^{a}\left(\sum_{j=1}^{b_i} n^3_{ij.}\right)/n_{i..},$$

$$k_{13} = \sum_{i=1}^{a}\left(\sum_{j=1}^{b_i} n^2_{ij.}\right)^2/n_{i..}, \qquad k_{14} = \sum_{i=1}^{a}\left(\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n^2_{ijk}\right)^2/n_{i..},$$

$$k_{15} = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\left(\sum_{k=1}^{c_{ij}} n^2_{ijk}\right)^2/n_{ij.}, \qquad k_{16} = \sum_{i=1}^{a}\left\{\sum_{j=1}^{b_i} n_{ij.}\left(\sum_{k=1}^{c_{ij}} n^2_{ijk}\right)\right\}/n_{i..},$$

$$k_{17} = \sum_{i=1}^{a}\left(\sum_{j=1}^{b_i} n^2_{ij.}\right)\left(\sum_{j=1}^{b_i}\sum_{k=1}^{c_{ij}} n^2_{ijk}\right)/n_{i..},$$

$$k_{18} = \sum_{i=1}^{a} \left\{ \sum_{j=1}^{b_i} \left(\sum_{k=1}^{c_{ij}} n_{ijk}^2 \right)^2 / n_{ij} \right\} / n_{i..},$$

$$k_{19} = \sum_{i=1}^{a} \sum_{j=1}^{b_i} \left(\sum_{k=1}^{c_{ij}} n_{ijk}^2 \right)^2 / n_{ij.}^2, \qquad k_{20} = \sum_{i=1}^{a} \left(\sum_{j=1}^{b_i} n_{ij.}^2 \right) \left(\sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}^2 \right) / n_{i..}^2,$$

$$k_{21} = \sum_{i=1}^{a} \left(\sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}^2 \right)^2 / n_{i..}^2, \qquad k_{22} = \sum_{i=1}^{a} \left(\sum_{j=1}^{b_i} n_{ij.}^2 \right)^2 / n_{i..}^2,$$

$$k_{23} = \sum_{i=1}^{a} n_{i..} \left(\sum_{j=1}^{b_i} n_{ij.}^2 \right), \qquad k_{24} = \sum_{i=1}^{a} n_{i..} \left(\sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}^2 \right),$$

$$k_{25} = \sum_{i=1}^{a} \sum_{j=1}^{b_i} n_{ij.} \left(\sum_{k=1}^{c_{ij}} n_{ijk}^2 \right),$$

$$v_1 = N - k_1, \qquad v_2 = k_4 - k_2, \qquad v_3 = k_5 - k_3, \qquad v_4 = a - 1,$$

$$v_5 = N - k_4, \qquad v_6 = k_6 - k_5, \qquad v_7 = b_. - a, \qquad v_8 = N - k_6,$$

$$v_9 = c_{..} - b_., \qquad v_{10} = N - c_{..},$$

$$g_1 = v_5^2 v_8^2 [k_1(N + k_1) - 2k_7/N],$$

$$\begin{aligned} g_2 = {} & v_5^2 v_8^2 (k_{22} + k_2^2 - 2k_{13}/N) + v_2^2 v_8^2 (Nk_2 + k_{22} - 2k_{12}) \\ & - 2v_2 v_5 v_8^2 \{(k_{12} - k_{22}) - (k_8 - k_{13})/N\}, \end{aligned}$$

$$\begin{aligned} g_3 = {} & v_5^2 v_8^2 (k_{21} + k_3^2 - 2k_{14}/N) + v_2^2 v_8^2 (k_{19} + k_{21} - 2k_{18}) \\ & + (v_2 v_6 - v_3 v_5)^2 (Nk_3 + k_{19} - 2k_{11}) \\ & - 2v_2 v_5 v_8^2 [(k_{18} - k_{21}) - (k_{15} - k_{14})/N] \\ & + 2v_5 v_8 (v_2 v_6 - v_3 v_5)[(k_{10} - k_{18}) - (k_9 - k_{15})/N] \\ & - 2v_2 v_8 [(k_{11} - k_{19}) - (k_{10} - k_{18})], \end{aligned}$$

$$\begin{aligned} g_4 = {} & v_5^2 v_8^2 (a + 1 - 2N) + v_2^2 v_8^2 (b_. - a) + (v_2 v_6 - v_3 v_5)^2 (c_{..} - b_.) \\ & + [v_5 v_8 (a - 1) + v_2 v_8 (a - b_.) + (v_2 v_6 - v_3 v_5)(c_{..} - b_.)]^2 / v_{10}, \end{aligned}$$

$$g_5 = v_5^2 v_8^2 [k_2(N + k_1) - 2k_{23}/N],$$

$$g_6 = v_5^2 v_8^2 [k_3(N + k_1) - 2k_2/N],$$

$$g_7 = v_5^2 v_8^2 (N - k_1),$$

$$\begin{aligned} g_8 = {} & v_5^2 v_8^2 (k_{20} + k_2 k_3 - 2k_{17}/N) + v_2^2 v_8^2 (Nk_3 - k_{16}) \\ & - 2v_2 v_5 v_8^2 [(k_{16} - k_{20}) - (k_{25} - k_{17})/N], \end{aligned}$$

$$g_9 = v_5^2 v_8^2 (k_4 - k_2) + v_2^2 v_8^2 (N - k_4),$$

$$g_{10} = v_5^2 v_8^2 (k_5 - k_3) + v_2^2 v_8^2 (k_6 - k_5) + (v_2 v_6 - v_3 v_5)^2 (N - k_6),$$

$$d_1 = v_8^2 (Nk_2 + k_{22} - 2k_{12}),$$

$$d_2 = v_8^2(k_{19} + k_{21} - 2k_{18}) + v_6^2(Nk_3 + k_{19} - 2k_{11}) + 2v_6v_8(k_{10} - k_{18} + k_{11} - k_{19}),$$
$$d_3 = v_8^2(b_. - a) + v_6^2(c_{..} - b_.) + [v_8(a - b_.) + v_6(c_{..} - b_.)]^2/v_{10},$$
$$d_4 = v_8^2(Nk_3 + k_{20} - 2k_{16}), \qquad d_5 = (N - k_6)^2(N - k_4),$$

and

$$d_6 = (N - k_6)(N - k_5)(k_6 - k_5).$$

It should be observed that the expressions for variances and covariances of the variance components estimates involve products of the variance components σ_α^2, σ_β^2, σ_γ^2, and σ_e^2. Since in general the variance components are unknown, one needs to substitute the estimates of $\sigma_\alpha^2, \sigma_\beta^2, \sigma_\gamma^2$, and σ_e^2 from (16.5.2) for the parameters σ_α^2, σ_β^2, σ_γ^2, and σ_e^2, respectively, in the expressions for variances and covariances. The estimates thus obtained will in general be biased. In order to obtain unbiased estimates, one may proceed as follows.

In the formulas for variances and covariances of $\sigma_\alpha^2, \sigma_\beta^2, \sigma_\gamma^2$, and σ_e^2, every product of the type $\sigma_\theta^2\sigma_\phi^2$ is to be replaced by $\hat{\sigma}_\theta^2\hat{\sigma}_\phi^2 - \text{Cov}(\hat{\sigma}_\theta^2, \hat{\sigma}_\phi^2)$, whenever θ and ϕ are different. The terms of the type σ_θ^4 are to be replaced by $(\hat{\sigma}_\theta^2)^2 - \text{Var}(\hat{\sigma}_\theta^2)$. Then one can rewrite these formulas as 10 simultaneous equations for estimates of variances and covariances of variance components estimates. The solution of these equations would yield unbiased estimates.

It is interesting to note that the expressions for variances and covariances reduce to the simpler form for balanced data. For example, if $b_i = b$, $c_{ij} = c$, and $n_{ij} = n$, we obtain

$$\text{Var}(\hat{\sigma}_{\gamma,\text{ANOV}}^2) = 2\left[(abcn^2 + abn^2 - 2abn^2)\sigma_\gamma^4 + 2abn(c-1)\sigma_\gamma^2\sigma_e^2 + ab(c-1)(cn-1)\frac{\sigma_e^4}{c(n-1)}\right]\Big/a^2b^2n^2(c-1)^2,$$

which reduces to

$$\text{Var}(\hat{\sigma}_{\gamma,\text{ANOV}}^2) = \frac{2}{n^2}\left[\frac{(\sigma_e^2 + n\sigma_\gamma^2)^2}{ab(c-1)} + \frac{\sigma_e^4}{abc(n-1)}\right].$$

16.6.2 LARGE SAMPLE VARIANCES OF MAXIMUM LIKELIHOOD ESTIMATORS

The explicit expressions for the large sample variances of the maximum likelihood estimators of the variance components σ_α^2, σ_β^2, σ_γ^2, and σ_e^2 have been derived by Rudan and Searle (1971) using the general result on the information matrix of the variance components in a general linear model as given in Section 10.7.2. The results on variance-covariance matrix of the vector of maximum likelihood estimators of σ_α^2, σ_β^2, σ_γ^2, and σ_e^2 are given by

$$\text{Var}\begin{bmatrix}\hat{\sigma}^2_{\alpha,\text{ML}}\\ \hat{\sigma}^2_{\beta,\text{ML}}\\ \hat{\sigma}^2_{\gamma,\text{ML}}\\ \hat{\sigma}^2_{e,\text{ML}}\end{bmatrix} = 2\begin{bmatrix} t_{\alpha\alpha} & t_{\alpha\beta} & t_{\alpha\gamma} & t_{\alpha e}\\ t_{\alpha\beta} & t_{\beta\beta} & t_{\beta\gamma} & t_{\beta e}\\ t_{\alpha\gamma} & t_{\beta\gamma} & t_{\gamma\gamma} & t_{\gamma e}\\ t_{\alpha e} & t_{\beta e} & t_{\gamma e} & t_{ee}\end{bmatrix}^{-1},$$

where

$$t_{\alpha\alpha} = \sum_{i=1}^{a}\left[\sum_{j=1}^{b_i}(A_{ij11}/p_{ij})\right]^2/q_i^2, \qquad t_{\alpha\beta} = \sum_{i=1}^{a}\left[\sum_{j=1}^{b_i}(A_{ij11}/p_{ij})^2\right]/q_i^2,$$

$$t_{\alpha\gamma} = \sum_{i=1}^{a}\left[\sum_{j=1}^{b_i}A_{ij22}/p_{ij}^2\right]/q_i^2, \qquad t_{\alpha e} = \sum_{i=1}^{a}\left[\sum_{j=1}^{b_i}A_{ij12}/p_{ij}^2\right]/q_i^2,$$

$$t_{\beta\beta} = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\left\{(A_{ij11}/p_{ij})^2 - 2\sigma_\alpha^2 A_{ij11}^3/q_i p_{ij}^3 + \sigma_\alpha^4(A_{ij11}/p_{ij})^2\left[\sum_{j=1}^{b_i}(A_{ij11}/p_{ij})^2\right]/q_i^2\right\},$$

$$t_{\beta\gamma} = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\left\{A_{ij22}/p_{ij}^2 - 2\sigma_\alpha^2 A_{ij11}A_{ij22}/q_i p_{ij}^3 + \sigma_\alpha^4(A_{ij22}/p_{ij}^2)\left[\sum_{j=1}^{b_i}(A_{ij11}/p_{ij})^2\right]/q_i^2\right\},$$

$$t_{\beta e} = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\left\{A_{ij12}/p_{ij}^2 - 2\sigma_\alpha^2 A_{ij11}A_{ij12}/q_i p_{ij}^3 + \sigma_\alpha^4(A_{ij12}/p_{ij}^2)\left[\sum_{j=1}^{b_i}(A_{ij11}/p_{ij})^2\right]/q_i^2\right\},$$

$$t_{\gamma\gamma} = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\left\{A_{ij22} - 2\sigma_\alpha^2 A_{ij33}/q_i p_{ij}^2 - 2\sigma_\beta^2 A_{ij33}^3/p_{ij} + 2\sigma_\alpha^2\sigma_\beta^2 A_{ij22}^2/q_i p_{ij}^3 + \sigma_\beta^4(A_{ij22}/p_{ij})^2\right.$$

$$+ \sigma_\alpha^4 (A_{ij22}/p_{ij}^2) \left[\sum_{j=1}^{b_i} A_{ij22}/p_{ij}^2 \right] /q_i^2 \Bigg\},$$

$$t_{\gamma e} = \sum_{i=1}^{a} \sum_{j=1}^{b_i} \Bigg\{ A_{ij12} - 2\sigma_\alpha^2 A_{ij23}/q_i p_{ij}^2 - 2\sigma_\beta^2 A_{ij23}/p_{ij}$$

$$+ 2\sigma_\alpha^2 \sigma_\beta^2 A_{ij12} A_{ij22}/q_i p_{ij}^3 + \sigma_\beta^4 A_{ij12} A_{ij22}/p_{ij}^2$$

$$+ \sigma_\alpha^4 (A_{ij12}/p_{ij}^2) \left[\sum_{j=1}^{b_i} A_{ij22}/p_{ij}^2 \right] /q_i^2 \Bigg\},$$

and

$$t_{ee} = \sum_{i=1}^{a} \sum_{j=1}^{b_i} \Bigg\{ A_{ij02} - 2\sigma_\alpha^2 A_{ij13}/q_i p_{ij}^2 - 2\sigma_\beta^2 A_{ij13}/p_{ij}$$

$$+ 2\sigma_\alpha^2 \sigma_\beta^2 A_{ij12}^2/q_i p_{ij}^3 + \sigma_\beta^4 (A_{ij12}/p_{ij})^2$$

$$+ \sigma_\alpha^4 (A_{ij12}/p_{ij}^2) \left[\sum_{j=1}^{b_i} A_{ij12}/p_{ij}^2 \right] /q_i^2 \Bigg\} + (N - c_{..})/\sigma_e^4,$$

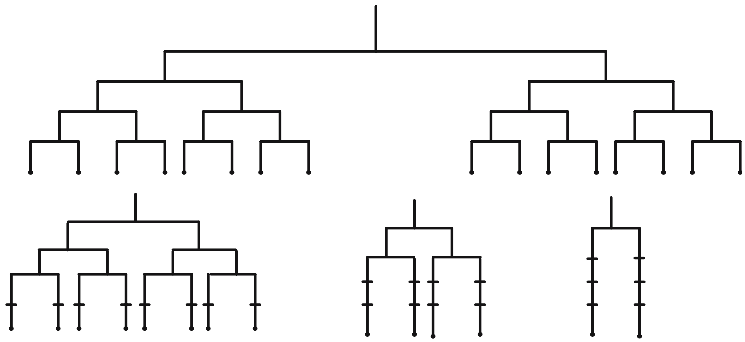
with

$$A_{ijpq} = \sum_{k=1}^{c_{ij}} (n_{ijk})^p/(m_{ijk})^q, \qquad m_{ijk} = n_{ijk}\sigma_\gamma^2 + \sigma_e^2,$$

$$p_{ij} = 1 + \sigma_\beta^2 A_{ij11}, \qquad \text{and} \qquad q_i = 1 + \sigma_\alpha^2 \sum_{j=1}^{b_i} A_{ij11}/p_{ij}.$$

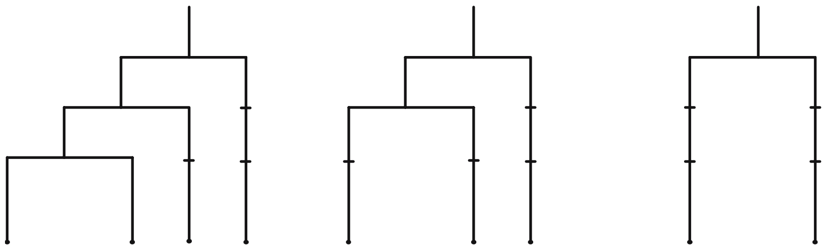
16.7 COMPARISONS OF DESIGNS AND ESTIMATORS

In a research project conducted at Purdue university in 1950, Dr. R. L. Anderson proposed a five-stage staggered nested design (Anderson and Bancroft, 1952, pp. 334–335) shown in Figure 16.2. This design was further elaborated by Prairie (1962), who proposed the following procedure for constructing a multistage nested design. If n_i is the number of samples in the ith first stage, n_{ij} in the (i, j)th second stage, etc., then one should try to get as near balance as possible by trying to achieve $|n_i - n_{i'}| = 0$ or 1, $|n_{ij} - n_{i'j'}| = 0$ or $1, \ldots,$ etc., $i \neq i'$. Subsequently, Calvin and Miller (1961) developed a four-stage unbalanced design, and Bainbridge (1965) proposed both four-, five-, and six-stage unbalanced designs which he called inverted and staggered nested designs. An example of a Bainbridge four-stage inverted design is shown in Figure 16.3.



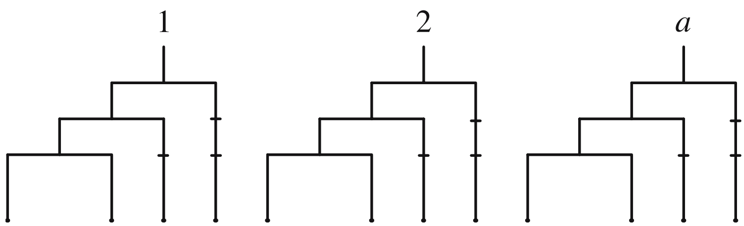
Source: Anderson and Bancroft (1952); used with permission.

FIGURE 16.2 Anderson five-stage staggered nested design.



Source: Bainbridge (1965); used with permission.

FIGURE 16.3 Bainbridge four-stage inverted nested design with a single replicate.



Source: Bainbridge (1965); used with permission.

FIGURE 16.4 Bainbridge four-stage staggered nested design.

The Bainbridge staggered nested design (BSN) consists of two levels at every stage except the first stage, which should have the maximum number of levels possible. Thus two levels at the second stage are nested within each level of the first stage. However, rather than selecting two levels at the third stage to be nested within each level of the second stage, two levels at the third stage occur with only one of the levels of the second stage. The other level of the second stage has only one level at the third stage. The process continues at each of the remaining stages of the design. An example of a four-stage BSN design is shown in Figure 16.4.

Staggered nested designs have certain definite advantages over other nested designs that make them popular in many scientific and industrial experiments. It assigns equal degrees of freedom (i.e., a) to all stages except to the first which receives $a - 1$ degrees of freedom. Thus it provides a much more even allocation of resources to estimate variance components. In addition, a p-stage design requires only p^a observations rather than $a2^{p-1}$ observations required by a balanced nested design with a levels at the first stage and two levels for each subsequent stage. Goss and Garret (1978) described an application of the use of the staggered nested designs in geology and Lamar and Zirk (1991) illustrated the usefulness of these designs in the chemical industry. Snee (1983) recommended their use in industrial process control to obtain robust estimates of variance components that affect a production process. More recently, Pettitt and McBratney (1993) explored the potential of these designs and recommended their use as the sampling design to estimate spatial variance components. Smith and Beverly (1981) extended the concept of evenly distributing the degrees of freedom among stages to designs where some factors have a factorial arrangement and others are nested within the factorial combinations or in levels of other factors, but where nesting is staggered. The estimation and testing problems associated with these designs have also been considered by Nelson (1983, 1995a, 1995b), Khattree and Naik (1995), Uhlig (1996), and Khattree et al. (1997).

Leone et al. (1968) compared the three designs, the balanced, Bainbridge inverted, and Bainbridge staggered, in terms of frequency of negative estimates and range of values as assumed by the traditional ANOVA estimates. For each type of design a sample of 40 was employed, since this is the smallest size which permits a convenient comparison among the unbalanced designs in Figures 16.3 and 16.4 and the balanced design considered earlier in Figure 7.1. Furthermore, it is the desired sample size which can be carried out within the constraints of industrial experimentation. Moreover, the unbalanced designs being proposed here provide useful alternatives to classical nested designs when the constraints of experimental resources and the relative precision of variance estimation are matters of utmost importance. The parameter values included in the study employed eight sets of variance components as shown in Table 16.7. Thus the models ranged from equal components to some components being nine times as large as the error component. Three underlying distributions, normal, rectangular, and exponential, were considered for each of the eight sets and three designs and a comparison was made. The exponential distribution was not used with inverted nested design. It was found that the type of design had very little effect on the shape of the resulting sampling distributions of the variance components. The descriptive statistics for normal and rectangular distributions were quite similar. However, the variance estimates for the long-tail exponential distribution were found to be quite imprecise. The sampling distributions of

TABLE 16.7 Sets of variance components included in the empirical study of the balanced, inverted, and staggered nested designs.

Variance component	Sets of parameter values for variance components							
	I	II	III	IV	V	VI	VII	VIII
σ_e^2	1	1	1	1	1	1	1	1
σ_γ^2	1	1	1	1	1	4	9	9
σ_β^2	1	1	1	4	9	9	9	9
σ_α^2	1	4	9	9	9	9	9	1

Source: Leone et al. (1968); used with permission.

variance component estimators were well approximated by Pearson Type III curves. The probability of obtaining negative estimates, which is an empirical percentage of negative estimates for $\hat{\sigma}^2_{\alpha,\text{ANOV}}$, $\hat{\sigma}^2_{\beta,\text{ANOV}}$, and $\hat{\sigma}^2_{\gamma,\text{ANOV}}$, is shown in Table 16.8. It was found that no single design performs the best for all the configuration of variance components; however, the choice of the Bainbridge staggered design appeared to be a good compromise.

Heckler and Rao (1985) have extended the concept of staggered nested designs to allow for more than two levels for any factor. An example of a four-stage extended staggered design (ESN) is shown in Figure 16.5. This ESN design has a levels at the first stage, four levels at the second stage, three levels at the third stage, and two levels at the fourth stage. The corresponding degrees of freedom are $a - 1$, $3a$, $2a$, and a, respectively. In general, a four-stage ESN design, with "staggering" commencing after the first stage, has b levels at the second stage, c levels at the third stage, and d levels at the fourth stage, with corresponding degrees of freedom $a - 1, a(b - 1), a(c - 1)$, and $a(d - 1)$, respectively. In order to obtain "good" estimates for variance components, the ESN design also allows for "balanced" levels for any number of upper stages. For example, Figure 16.6 shows a five-stage ESN design where the first two stages have balanced levels. The number of levels and degrees of freedom for different stages are given by $[a, 2, 3, 2, 3]$ and $[a - 1, a(1), 2a(2), 2a(1), 2a(2)]$, respectively. In general, a five-stage ESN design where the first two stages have balanced levels consists of a levels at the first stage, b levels at the second stage, c levels at the third stage, d levels at the fourth stage, and e levels at the fifth stage. The corresponding degrees of freedom are $a - 1, a(b - 1), ab(c - 1), ab(d - 1)$, and $ab(e - 1)$, respectively. Heckler and Rao performed an empirical study to assess the information loss due to smaller experimental size of the BSN design compared to the ESN design in a four-stage nested classification. The "best" and "worst" ESN designs were identified under a variety of combinations of population variance components and design

TABLE 16.8 Empirical percentages of negative estimates of the variance components.

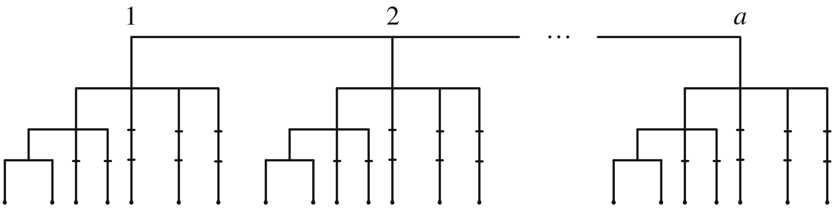
Stage	Design	Set I				Set II				Set III				Set IV			
		V	N	R	E	V	N	R	E	V	N	R	E	V	N	R	E
a	Balanced	1	24	23	25	4	6	5	12	9	2	1	5	9	8	6	14
	Staggered	1	20	20	23	4	2	2	7	9	0.2	0	2	9	2	2	7
	Inverted	1	19			4	1			9	0			9	1		
b	Balanced	1	18	19	20	1	18	17	21	1	18	18	20	4	3	2	5
	Staggered	1	21	19	24	1	21	20	24	1	21	21	23	4	2	2	4
	Inverted	1	20			1	20			1	20			4	3		
c	Balanced	1	4	3	6	1	4	4	6	1	4	4	7	1	4	2	6
	Staggered	1	10	9	16	1	10	9	16	1	10	9	13	1	10	9	12
	Inverted	1	16			1	16			1	16			1	16		

Stage	Design	Set I				Set II				Set III				Set IV			
		V	N	R	E	V	N	R	E	V	N	R	E	V	N	R	E
a	Balanced	9	17	15	22	9	19	17	22	9	22	20	25	1	46	43	44
	Staggered	9	10	8	11	9	12	11	14	9	16	14	18	1	44	42	44
	Inverted	9	6			9	9			9	14			1	43		
b	Balanced	9	1	0	2	9	4	5	7	9	13	11	14	9	13	11	15
	Staggered	9	0.2	0	2	9	4	3	7	9	12	11	17	9	12	12	16
	Inverted	9	0			9	3			9	11			9	11		
c	Balanced	1	4	3	6	4	0	0	0	9	0	0	0	9	0	0	0
	Staggered	1	10	9	14	4	0.4	0	2	9	0	0	0	9	0	0	0
	Inverted	1	16			4	2			9	0.2			9	0.2		

Column Heading Symbols: V = variance, N = normal*, R = rectangular, E = exponential.

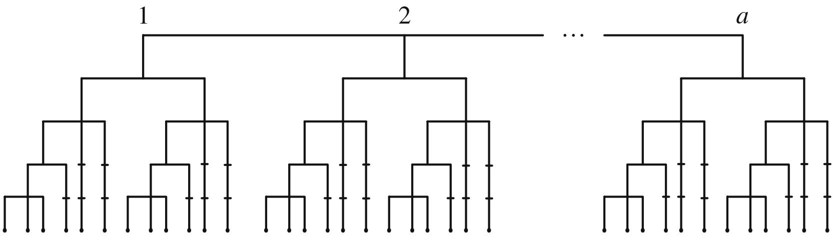
*All normal values are theoretical, except for stage *a* of the inverted design.

Source: Leone et al. (1968); used with permission.



Source: Heckler and Rao (1985); used with permission.

FIGURE 16.5 Heckler–Rao four-stage extended staggered design.



Source: Heckler and Rao (1985); used with permission.

FIGURE 16.6 Heckler–Rao five-stage extended staggered design.

parameters using the ANOVA estimates. Heckler and Rao also illustrate the cost-effectiveness of ESN designs via an example from an experiment designed to evaluate the performance of an assay for Lipase, a blood enzyme, on the EKTACHEM 400 clinical chemistry analyzer.

Khattree et al. (1997) performed a Monte Carlo study to compare the relative performance of the ANOVA, truncated ANOVA (TANOVA), and a new procedure known as principal components (PC) method for estimating variance components in staggered nested designs. Random samples were generated from a normal distribution for values of design parameters that included three-, six-, and ten-stage nested designs with $a = 10, 25$ to correspond to the staggered designs of interest. The mean was assumed to be zero and various values of the variance components were included in the simulation. The estimators were compared using the compound mean squared error (CMSE) and compound squared bias (CSB) criteria.[2] The PC method generally fared well in comparison to the ANOVA and TANOVA methods for six- and ten-stage designs with respect to the CSME criterion. For three-stage designs, when the sum of all the variance components is small, the PC method outperforms the TANOVA; however, for larger sums of variance components, TANOVA seems to have superior performance over the PC. The TANOVA yielded low CSB; in contrast, PC estimates were consistently biased. In addition to CMSE and CSB criteria,

[2]For a vector valued-estimator $\hat{\boldsymbol{\theta}}$ of a parameter vector $\boldsymbol{\theta}$, CMSE, and CSB are defined as $\text{CMSE}(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = E[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})'(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})]$ and $\text{CSB}(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = E[(\hat{\boldsymbol{\theta}}) - \boldsymbol{\theta}]'[E(\hat{\boldsymbol{\theta}}) - \boldsymbol{\theta})]$.

two alternative criteria for comparison based on Pitman's measure of closeness and probability of concentration were also used. Similar conclusions based on these criteria were observed.

16.8 CONFIDENCE INTERVALS AND TESTS OF HYPOTHESES

In this section, we briefly review the problem of constructing confidence intervals and testing hypotheses on variance components for the model in (16.1.1).

16.8.1 CONFIDENCE INTERVALS

In the three-way nested model in (16.1.1), MS_E has constant times a chi-square distribution, but SS_A, SS_B, and SS_C are neither independent nor distributed as chi-square type variables. An exact normal theory interval for σ_e^2 is constructed in the usual way. Similarly, an exact interval on $\sigma_\gamma^2/\sigma_e^2$ can be constructed using Wald's procedure described in Section 11.8.2. The approach of Hernández et al. (1992) based on unweighted and Type I sums of squares can be extended to construct approximate intervals for σ_γ^2, σ_β^2, and σ_α^2 and their parametric functions. However, it is not clear how the lack of independence and chi-squaredness of mean squares will affect the properties of intervals thus obtained.

16.8.2 TESTS OF HYPOTHESES

An exact test of $H_0^C : \sigma_\gamma^2 = 0$ vs. $H_1 : \sigma_\gamma^2 > 0$ can be performed using the conventional F-test. However, for testing the hypotheses on σ_β^2 and σ_α^2, approximate tests are generally needed. Satterthwaite-type test procedures can be constructed by the synthesis of mean squares using either a numerator component, a denominator component, or both. For example, to test $H_0^B : \sigma_\beta^2 = 0$ vs. $H_1^B : \sigma_\beta^2 > 0$, the synthesized mean squares for the numerator and the denominator components are given by

$$\mathrm{MS}_D = \left(\frac{r_2}{r_1}\right)\mathrm{MS}_C + \left(1 - \frac{r_2}{r_1}\right)\mathrm{MS}_E \tag{16.8.1}$$

and

$$\mathrm{MS}_N = \left(\frac{r_1}{r_2}\right)\mathrm{MS}_B + \left(1 - \frac{r_1}{r_2}\right)\mathrm{MS}_E. \tag{16.8.2}$$

Now, the test procedures based on linear combinations (16.8.1) and (16.8.2) are determined as

$$\mathrm{MS}_B/\mathrm{MS}_D \quad \text{and} \quad \mathrm{MS}_N/\mathrm{MS}_C,$$

which are approximated by F-statistics with $(b_. - a, \nu_D)$ and $(\nu_N, c_{..} - b_.)$ degrees of freedom, respectively, where ν_D and ν_N are estimated using Satterthwaite's procedure.

Similarly, to test $H_0^A : \sigma_\alpha^2 = 0$ vs. $H_1^A : \sigma_\alpha^2 > 0$, the synthesized mean squares for the denominator and the numerator components are given by

$$\text{MS}'_D = \left(\frac{r_5}{r_3}\right)\text{MS}_B + \left(\frac{r_4}{r_1} - \frac{r_2 r_5}{r_1 r_3}\right)\text{MS}_C + \left(1 - \frac{r_5}{r_3} - \frac{r_4}{r_1} + \frac{r_2 r_5}{r_1 r_3}\right)\text{MS}_E \tag{16.8.3}$$

and

$$\text{MS}'_N = \left(\frac{r_3}{r_5}\right)\text{MS}_A + \left(\frac{r_2}{r_1} - \frac{r_3 r_4}{r_1 r_5}\right)\text{MS}_C + \left(1 - \frac{r_3}{r_5} - \frac{r_2}{r_1} + \frac{r_3 r_4}{r_1 r_5}\right)\text{MS}_E. \tag{16.8.4}$$

The corresponding test procedures based on linear combinations (16.8.3) and (16.8.4) are determined as

$$\text{MS}_A/\text{MS}'_D \quad \text{and} \quad \text{MS}'_N/\text{MS}_B,$$

which are approximated by F-statistics with $(a - 1, \upsilon'_D)$ and $(\upsilon'_N, b_. - a)$ degrees of freedom, respectively, where, again, υ'_D and υ'_N are estimated using Satterthwaite's procedure.

16.8.3 A NUMERICAL EXAMPLE

In this section, we outline computations for constructing confidence intervals on σ_γ^2, σ_β^2, and σ_α^2 and for testing the hypotheses $H_0^C : \sigma_\gamma^2 = 0$, $H_0^B : \sigma_\beta^2 = 0$, and $H_0^A : \sigma_\alpha^2 = 0$ using the insecticide residue on celery data of the numerical example in Section 16.5.5. Here, σ_α^2, σ_β^2, σ_γ^2, and σ_e^2 correspond to the variation among plots, samples, subsamples, and error, respectively. Approximate confidence intervals can be calculated using the approach of Hernández et al. (1992) and employing the unweighted or Type I sums of squares (see Exercise 16.7). For example, it can be verified that an approximate two sided 95% confidence interval for σ_β^2 based on unweighted means squares is given by

$$P\{0.0065 \le \sigma_\beta^2 \le 0.0494\} \doteq 0.95.$$

The hypothesis $H_0^C : \sigma_\gamma^2 = 0$ is tested using the conventional F-statistic, $F_C = \text{MS}_C/\text{MS}_E$, giving an F-value of 1.62 ($p = 0.133$). The results are not significant at a level of 13.3% or lower, and we do not reject H_0^C and conclude that $\sigma_\gamma^2 \approx 0$, or the measurements on residue do not differ appreciably. Note that this F-test is exact.

For testing the hypothesis, $H_0^B : \sigma_\beta^2 = 0$, however, there is no simple exact test. An approximate F-test can be obtained by using Satterthwaite's procedure via test statistics $\mathrm{MS}_B/\mathrm{MS}_D$ or $\mathrm{MS}_N/\mathrm{MS}_C$, where MS_D and MS_N are evaluated using (16.8.1) and (16.8.2), respectively. Substituting the appropriate quantities, we have

$$\begin{aligned}\mathrm{MS}_D &= (1.214/1.500)(0.01625) + (1 - 1.214/1.500)(0.01004)\\ &= 0.01315 + 0.00191 = 0.01506\end{aligned}$$

and

$$\begin{aligned}\mathrm{MS}_N &= (1.500/1.214)(0.04508) + (1 - 1.500/1.214)(0.01004)\\ &= 0.05570 - 0.00237 = 0.05333.\end{aligned}$$

The degree of freedom ν_D and ν_N associated with MS_D and MS_N are estimated as

$$\nu_D = \frac{(0.01506)^2}{\dfrac{(0.01315)^2}{22} + \dfrac{(0.00191)^2}{22}} = 28.3$$

and

$$\nu_N = \frac{(0.05333)^2}{\dfrac{(0.05570)^2}{22} + \dfrac{(-0.00237)^2}{22}} = 20.1.$$

The test statistics $\mathrm{MS}_B/\mathrm{MS}_D$ and $\mathrm{MS}_N/\mathrm{MS}_C$ yield F-values of 2.99 and 3.28, which are to be compared against the theoretical F-values with (22, 28.3) and (20.1, 22) degrees of freedom, respectively. The corresponding p-values are 0.003 and 0.004, respectively, and the results are highly significant. Thus we reject H_0^B and conclude that $\sigma_\beta^2 > 0$, or different samples differ significantly.

Finally, for testing $H_0^A : \sigma_\alpha^2 > 0$, again, there is no simple exact test. An approximate F-test can be obtained by using Satterthwaite's procedure via test statistics $\mathrm{MS}_A/\mathrm{MS}'_D$ or $\mathrm{MS}'_N/\mathrm{MS}_B$ where MS'_D and MS'_N are evaluated using (16.8.3) and (16.8.4), respectively. Substituting the appropriate quantities, we have

$$\begin{aligned}\mathrm{MS}'_D &= \left(\frac{3.000}{2.000}\right)(0.04508) + \left(\frac{1.571}{1.500} - \frac{1.214 \times 3.000}{1.500 \times 2.000}\right)(0.01625)\\ &\quad + \left(1 - \frac{3.000}{2.000} - \frac{1.571}{1.500} + \frac{1.214 \times 3.000}{1.500 \times 2.000}\right)0.01004\\ &= 0.06762 - 0.00271 - 0.00335 = 0.06156\end{aligned}$$

and

$$\mathrm{MS}'_N = \left(\frac{2.000}{3.000}\right)(0.18404) + \left(\frac{1.214}{1.500} - \frac{2.000 \times 1.571}{1.500 \times 3.000}\right)(0.01625)$$

$$+\left(1-\frac{2.000}{3.000}-\frac{1.214}{1.500}+\frac{2.000\times 1.571}{1.500\times 3.000}\right)0.01004$$
$$= 0.12269 + 0.00181 + 0.00223 = 0.12673.$$

The degrees of freedom ν'_D and ν'_N associated with MS'_D and MS'_N are estimated as

$$\nu'_D = \frac{(0.06156)^2}{\frac{(0.06762)^2}{22}+\frac{(-0.00271)^2}{22}+\frac{(-0.00335)^2}{22}} = 18.2$$

and

$$\nu'_N = \frac{(0.12673)^2}{\frac{(0.12269)^2}{10}+\frac{(0.00181)^2}{22}+\frac{(0.00223)^2}{22}} = 10.7.$$

The test statistics MS_A/MS'_D and MS'_N/MS_B yield F-values of 2.99 and 2.81 which are to be compared against the theoretical F-values with (10, 18.2) and (10.7, 22) degrees of freedom, respectively. The corresponding p-values are 0.021 and 0.019, and the results are statistically significant at 5% level or higher. Thus we also reject H_0^A and conclude that $\sigma_\alpha^2 > 0$ or different plots differ significantly.

EXERCISES

1. Express the coefficients of the variance components in the expected mean squares derived in Section 16.3 in terms of the formulation given in Section 17.3.
2. Apply the method of "synthesis" to derive the expected mean squares given in Section 16.3.
3. Derive the results on expected values of unweighted mean squares given in (16.4.3).
4. Show that the ANOVA estimators (16.5.2) reduce to the corresponding estimators (7.3.1) for balanced data.
5. Show that the unweighted means estimators (16.5.4) reduce to the ANOVA estimators (7.3.1) for balanced data.
6. Show that the symmetric sums estimators (16.5.7) and (16.5.10) reduce to the ANOVA estimators (7.3.1) for balanced data.
7. For the numerical example in Section 16.8.3, calculate 95% confidence intervals on the variance components $\sigma_\alpha^2, \sigma_\beta^2, \sigma_\gamma^2$, and σ_e^2 using the method described in the text.

8. Derive the expressions for variances and covariances of the analysis of variance estimators of the variance components as given in Section 16.6.1 (Mahamunulu, 1963).

9. Derive the expressions for large sample variances and covariances of the maximum likelihood estimators of the variance components as given in Section 16.6.2 (Rudan and Searle, 1971).

10. Occasionally, in a nested design, the first-stage factor which is not nested may be fixed because all the levels of interest have been included in the experiment. This combination of fixed and random effects gives rise to a mixed model. In such a situation, the interest may lie in estimating the mean of a treatment level or a contrast between two treatment level means. Show that (Eisen, 1966):

    (i) $\bar{y}_{i...} = \mu + \alpha_i + \sum_{j=1}^{b_i} \frac{n_{ij.}}{n_{i..}} \beta_{j(i)} + \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \frac{n_{ijk}}{n_{i..}} \gamma_{k(ij)} + \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \sum_{\ell=1}^{n_{ijk}} \frac{e_{\ell(ijk)}}{n_{i..}}$,

    (ii) $E(\bar{y}_{i...}) = \mu + \alpha_i$,

    (iii) $E(\bar{y}_{i...} - \bar{y}_{i'...}) = \alpha_i - \alpha_{i'}$,

    (iv) $\text{Var}(\bar{y}_{i...}) = \frac{1}{n_{i..}^2}[\sum_{j=1}^{b_i} n_{ij.}^2 \sigma_\beta^2 + \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk}^2 \sigma_\gamma^2 + n_{i..}\sigma_e^2]$,

    (v) $\text{Var}(\bar{y}_{i...} - \bar{y}_{i'...}) = [\sum_{j=1}^{b_i} \frac{n_{ij.}^2}{n_{i..}^2} + \sum_{j'=1}^{b_{i'}} \frac{n_{i'j'.}^2}{n_{i'..}}]\sigma_\beta^2 + [\sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \frac{n_{ijk}^2}{n_{i..}^2}$
    $+ \sum_{j'=1}^{b_{i'}} \sum_{k'=1}^{c_{i'j'}} \frac{n_{i'j'k'}^2}{n_{i'..}^2}]\sigma_\gamma^2 + [\frac{1}{n_{i..}} + \frac{1}{n_{i'..}}]\sigma_e^2$.

    Thus the variance of each treatment level mean and each contrast may be different, depending on the imbalance structure of the data.

11. Bainbridge (1965) described a nested experiment designed to detect sources of variation occurring in industrial production through a chemical test on a specific textile material. The purpose of the experiment was to study variations in the chemical analysis due to changes in the raw material over the days, differences in the machines, long term testing at different shifts, and short term testing through the duplicate analyses. The experiment was conducted over a period of 42 days. From a large number of machines two were selected on each of the 42 days. Two samples were taken from one of the machines and one sample was taken from the other machine. The two samples from the first machine were tested by two analysts, one of them performing duplicate measurements, while only one measurement was made on the other sample. The sample from the second machine was tested only once by an analyst yielding a single measurement. Hence, there are 42 days, two machines per day, two samples from one machine and one sample from the second machine, two measurements from the first sample of the first machine, and only one measurement from the other two samples, giving a total of 88 machines and 168 observations. The relevant data are given below.

Day	1			2			3			4			5			6		
Machine	1		2	1		2	1		2	1		2	1		2	1		2
Sample	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1
Analysis	6.1	6.6	8.8	8.5	8.2	8.1	8.6	8.0	7.4	9.3	6.5	8.0	8.1	2.3	9.5	8.5	4.0	9.2
	6.6			9.6			6.7			7.2			7.1			9.0		

Day	7			8			9			10			11			12		
Machine	1		2	1		2	1		2	1		2	1		2	1		2
Sample	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1
Analysis	8.5	4.0	9.2	9.8	4.0	9.2	9.0	6.8	9.2	11.0	10.5	11.3	9.7	10.3	9.3	10.5	10.0	4.0
	9.0			9.8			8.0			10.9			10.6			8.4		

Day	13			14			15			16			17			18		
Machine	1		2	1		2	1		2	1		2	1		2	1		2
Sample	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1
Analysis	8.3	8.8	9.7	8.4	6.7	4.6	9.3	9.9	9.7	7.1	8.2	10.0	5.8	7.5	10.2	8.9	6.6	9.2
	10.6			7.2			8.7			8.7			6.8			6.6		

Day	19			20			21			22			23			24		
Machine	1		2	1		2	1		2	1		2	1		2	1		2
Sample	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1
Analysis	11.5	3.1	10.8	10.3	7.2	9.4	9.1	10.7	10.3	5.7	8.4	10.3	8.5	7.6	8.3	9.6	12.6	11.6
	7.1			10.0			9.5			7.7			8.8			12.2		

Day	25			26			27			28			29			30		
Machine	1		2	1		2	1		2	1		2	1		2	1		2
Sample	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1
Analysis	9.5	9.6	9.4	10.3	12.6	11.3	7.0	10.8	11.4	11.5	5.1	9.6	6.0	6.6	2.2	8.0	8.6	6.6
	10.4			10.6			10.6			7.3			7.0			7.0		

Day	31			32			33			34			35			36		
Machine	1		2	1		2	1		2	1		2	1		2	1		2
Sample	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1
Analysis	13.1	12.5	11.5	12.1	10.4	9.1	14.2	10.6	4.6	10.0	7.2	7.9	6.5	7.8	9.0	6.5	4.4	8.1
	9.2			11.7			10.6			10.4			8.4			6.8		

Day	37			38			39			40			41			42		
Machine	1		2	1		2	1		2	1		2	1		2	1		2
Sample	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1	1	2	1
Analysis	9.2	8.7	9.4	11.0	11.2	10.9	8.6	10.3	9.0	8.9	7.0	7.8	6.6	7.7	9.3	8.4	7.6	6.8
	10.1			11.0			10.0			8.0			7.2			8.8		

Source: Bainbridge (1965); used with permission.

(a) Describe the mathematical model and the assumptions of the experiment.

(b) Analyze the data and report the conventional analysis of variance table based on Type I sums of squares.

(c) Perform an appropriate F-test to determine whether the results of the chemical analysis vary from day to day.

(d) Perform an appropriate F-test to determine whether the results of the chemical analysis vary from machine to machine.

(e) Perform an appropriate F-test to determine whether the results of the chemical analysis vary from sample to sample.

(f) Find point estimates of the variance components and the total variance using the methods described in the text.

(g) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

12. Mason et al. (1989, pp. 366–367) described the use of a four-stage staggered nested design for a polymerization process that produces polyethelyne pellets. Thirty lots were chosen for the experiment and two boxes were selected from each lot. Two preparations were made from the first box whereas only one preparation was made from the second box. Finally, one strength test was made on preparation 1 from each box, but two tests were made on preparation 2 from the first box. The data are given below.

Lot	**1**			**2**			**3**			**4**			**5**		
Box	**1**		**2**	**1**		**2**	**1**		**2**	**1**		**2**	**1**		**2**
Prep.	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**
Test	11.91	9.76	9.02	10.0	10.65	13.69	8.02	6.50	7.95	9.15	8.08	7.46	7.43	7.84	6.11
		9.24			7.77			6.26			5.28			5.91	

Lot	**6**			**7**			**8**			**9**			**10**		
Box	**1**		**2**	**1**		**2**	**1**		**2**	**1**		**2**	**1**		**2**
Prep.	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**
Test	7.01	9.00	8.58	11.13	12.81	10.00	14.07	10.62	14.56	4.08	4.88	4.76	6.73	9.38	6.99
		8.38			13.58			11.71			4.96			8.02	

Lot	**11**			**12**			**13**			**14**			**15**		
Box	**1**		**2**	**1**		**2**	**1**		**2**	**1**		**2**	**1**		**2**
Prep.	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**
Test	6.59	5.91	6.55	5.77	7.19	8.33	8.12	7.93	7.43	3.95	3.70	5.92	5.96	4.64	5.88
		5.79			7.22			6.48			2.86			5.70	

Lot	**16**			**17**			**18**			**19**			**20**		
Box	**1**		**2**	**1**		**2**	**1**		**2**	**1**		**2**	**1**		**2**
Prep.	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**
Test	4.18	5.94	5.24	11.25	9.50	11.14	9.51	10.93	12.71	16.79	11.95	13.08	7.51	4.34	5.21
		6.28			8.00			12.16			10.58			5.45	

Lot	**21**			**22**			**23**			**24**			**25**		
Box	**1**		**2**	**1**		**2**	**1**		**2**	**1**		**2**	**1**		**2**
Prep.	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**
Test	6.51	7.60	6.35	6.31	5.12	8.74	4.53	5.28	5.07	4.35	5.44	7.04	2.57	3.50	3.76
		6.72			5.85			5.73			5.38			3.88	

Lot	**26**			**27**			**28**			**29**			**30**		
Box	**1**		**2**	**1**		**2**	**1**		**2**	**1**		**2**	**1**		**2**
Prep.	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**	**1**	**2**	**1**
Test	3.48	4.80	3.18	4.38	5.35	5.50	3.79	3.09	2.59	4.39	5.30	6.13	5.96	7.09	7.14
		4.46			6.39			3.19			4.72			7.82	

Source: Mason et al. (1989); used with permission.

(a) Describe the mathematical model and the assumptions of the experiment.

(b) Analyze the data and report the conventional analysis of variance table based on Type I sums of squares.

(c) Perform an appropriate F-test to determine whether the results of the strength test vary from lot to lot.

(d) Perform an appropriate F-test to determine whether the results of the strength test vary from box to box.

(e) Perform an appropriate F-test to determine whether the results of the strength test vary from preparation to preparation.

(f) Find point estimates of the variance components and the total variance using the methods described in the text.

(g) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

13. Eisen (1966) described an experiment conducted in the Mouse Genetics Laboratory at the North Carolina State University to compare the growth rates of different breeding structure lines of mice. Progeny of several dams were mated to a single sire and each line contained several sire families involving an unbalanced four-stage nested design. The analysis variance for the data on 21-day weaning of male progeny is given below.

Analysis of variance for 21-day weaning weight of male mice.

Source of variation	Degrees of freedom	Mean square	Expected mean squares
Lines	2	8.088	$\sigma_e^2 + 3.21\sigma_\gamma^2 + 6.26\sigma_\beta^2 + 126.50\sigma_\alpha^2$
Sires within lines	73	4.552	$\sigma_e^2 + 2.99\sigma_\gamma^2 + 5.03\sigma_\beta^2$
Dams within sires	58	4.388	$\sigma_e^2 + 2.73\sigma_\gamma^2$
Progeny within dams	252	0.758	σ_e^2

Source: Eisen (1966); used with permission.

(a) Describe the mathematical model and the assumption for the experiment. In the original experiment the lines of breeding were considered to be fixed. For the purpose of this exercise, you can assume a completely random model.

(b) Test the hypothesis that there are significant differences between different lines of breeding structure.

(c) Test the hypothesis that there are significant differences between different sires within lines.

(d) Test the hypothesis that there are significant differences between different dams within sires.

(e) Find point estimates of the variance components and the total variance using the methods described in the text.

(f) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

14. Consider the experiment on the growth rates of different breeding structure lines of mice described in Excercise 13 above. Eisen (1966) also analyzed the results for 56 day body weight of female mice and the analysis of variance is given below.

Analysis of variance for 56-day weaning weight of male mice.

Source of variation	Degrees of freedom	Mean square	Expected mean square
Lines	2	34.754	$\sigma_e^2 + 2.99\sigma_\gamma^2 + 5.84\sigma_\beta^2 + 113.33\sigma_\alpha^2$
Sires within lines	71	9.888	$\sigma_e^2 + 2.78\sigma_\gamma^2 + 4.65\sigma_\beta^2$
Dams within sires	56	5.955	$\sigma_e^2 + 2.40\sigma_\gamma^2$
Progeny within dams	211	2.474	σ_e^2

Source: Eisen (1966); used with permission.

(a) Describe the mathematical model and the assumption for the experiment. In the original experiment the lines of breeding were considered to be fixed. For the purpose of this exercise, you can assume a completely random model.

(b) Test the hypothesis that there are significant differences between different lines of breeding structure.

(c) Test the hypothesis that there are significant differences between different sires within lines.

(d) Test the hypothesis that there are significant differences between different dams within sires.

(e) Find point estimates of the variance components and the total variance using the methods described in the text.

(f) Calculate 95% confidence intervals of the variance components and the total variance using the methods described in the text.

Bibliography

R. L. Anderson and T. A. Bancroft (1952), *Statistical Theory in Research*, McGraw–Hill, New York.

T. R. Bainbridge (1965), Staggered nested designs for estimating variance components, *Indust. Quality Control*, **22**-1, 12–20.

C. I. Bliss (1967), *Statistics in Biology*, McGraw–Hill, New York.

L. D. Calvin and J. D. Miller (1961), A sampling design with incomplete dichotomy, *Agronomy J.*, **53**, 325–328.

E. J. Eisen (1966), The quasi-F test for an unnested fixed factor in an unbalanced hierarchal design with a mixed model, *Biometrics*, **22**, 937–942.

M. Ganguli (1941), A note on nested sampling, *Sankhyā*, **5**, 449–452.

D. W. Gaylor and T. D. Hartwell (1969), Expected mean squares for nested classifications, *Biometrics*, **25**, 427–430.

T. I. Goss and R. G. Garrett (1978), A new unbalanced nested ANOVA model in geology: A down-to-earth design, in *ASA Proceedings of Statistical Computing Section*, American Statistical Association, Alexandria, VA, 360–365.

C. E. Heckler and P. S. R. S. Rao (1985), Efficient estimation of variance components using staggered nested designs, in *Fall Technical Conference*, American Society for Quality Control, Corning, NY.

R. P. Hernández, R. K. Burdick, and N. J. Birch (1992), Confidence intervals and tests of hypotheses on variance components in an unbalanced two-fold nested design, *Biometrical J.*, **34**, 387–402.

R. Khattree and D. N. Naik (1995), Statistical tests for random effects in staggered nested designs, *J. Appl. Statist.*, **22**, 495–505.

R. Khattree, D. N. Naik, and R. L. Mason (1997), Estimation of variance components in staggered nested designs, *J. Appl. Statist.*, **24**, 395–408.

S. C. King and C. R. Henderson (1954), Variance component analysis in heritability studies, *Poultry Sci.*, **331**, 147–154.

G. G. Koch (1967), A general approach to the estimation of variance components, *Techometrics*, **9**, 93–118.

G. G. Koch (1968), Some further remarks concerning "A general approach to estimation of variance components," *Technometrics*, **10**, 551–558.

J. L. Lamar and W. E. Zirk (1991), Nested designs in the chemical industry, in *ASQC Quality Congress Transaction*, American Society for Quality Control, Milwaukee, 615–622.

F. C. Leone, L. S. Nelson, N. L. Johnson, and S. Eisenstat (1968), Sampling distributions of variance components II: Empirical studies of unbalanced nested designs, *Technometrics*, **10**, 719–738.

D. M. Mahamunulu (1963), Sampling variances of the estimates of variance components in the unbalanced three-way nested classification, *Ann. Math. Statist.*, **34**, 521–527.

R. L. Mason, R. F. Gunst, and J. L. Hess (1989), *Statistical Design and Analysis of Experiments with Applications to Engineering and Science*, Wiley, New York.

L. S. Nelson (1983), Variance estimation using staggered nested designs, *J. Quality Tech.*, **15**, 195–198.

L. S. Nelson (1995a), Using nested designs I: Estimation of standard deviations, *J. Quality Tech.*, **27**, 169–171.

L. S. Nelson (1995b), Using nested designs II: Confidence limits for standard deviations, *J. Quality Tech.*, **27**, 265–267.

A. N. Pettitt and A. B. McBratney (1993), Sampling designs for estimating spatial variance components, *Appl. Statist.*, **42**, 185–209.

R. R. Prairie (1962), *Optimal Designs to Estimate Variance Components and Reduce Product Variability for Nested Classifications*, Ph.D. dissertation, Institute of Statistics, North Carolina State University, Raleigh, NC.

J. W. Rudan and S. R. Searle (1971), Large sample variances of maximum likelihood estimators of variance components in the three-way nested classification, random model, with unbalanced data, *Biometrics*, **27**, 1087–1091.

S. R. Searle (1971), *Linear Models*, Wiley, New York.

S. R. Searle, G. Casella, and C. E. McCulloch (1992), *Variance Components*, Wiley, New York.

R. H. Sharpe and C. H. van Middelem (1955), Application of variance components to horticultural problems with special reference to a parathion residue study, *Proc. Amer. Soc. Horticultural Sci.*, **66**, 415–420.

J. R. Smith and J. M. Beverly (1981), The use and analysis of staggered nested factorial designs, *J. Quality Tech.*, **13**, 166–173.

R. D. Snee (1983), Graphical analysis of process variation studies, *J. Quality Tech.*, **15**, 76–88.

S. Uhlig (1996), Optimum two-way nested designs for estimation of variance components, *Tatra Mountains Math. Pub.*, **7**, 105–112.

17 General r-Way Nested Classification

The completely nested or hierarchical classification involving several stages arises in many areas of scientific research and applications. For example, in a large scale sample survey, experiments may be laid down on very many blocks, and the blocks are then naturally classified by cities, the cities by states in which they occur; and the states by the regions, and so forth. In a genetic investigation of dairy production, the units could be cattle classified by sires, sires classified by their dams, and so on. Frequently, the designs employed in these investigations are unbalanced, sometimes inadvertently. In this chapter, we shall briefly outline the analysis of variance for an unbalanced r-way nested classification.

17.1 MATHEMATICAL MODEL

The random effects model for the unbalanced r-way nested classification can be represented by

$$\begin{aligned} y_{i_1 i_2 \ldots i_{r-1} i_r i_{r+1}} &= \mu + \alpha_{i_1} + \beta_{i_2(i_1)} + \gamma_{i_3(i_2)} + \cdots + \eta_{i_{r-1}(i_{r-2})} \\ &\quad + \xi_{i_r(i_{r-1})} + e_{i_{r+1}(i_r)}, \end{aligned} \tag{17.1.1}$$

where

$$\begin{aligned} \mu &= \text{overall or general mean,} \\ \alpha_{i_1} &= \text{effect of } i_1\text{th factor,} \\ \beta_{i_2(i_1)} &= \text{effect of } i_2\text{th factor within } i_1\text{th factor,} \\ \gamma_{i_3(i_2)} &= \text{effect of } i_3\text{th factor within } i_2\text{th factor,} \\ \vdots \quad & \quad \vdots \\ \eta_{i_{r-1}(i_{r-2})} &= \text{effect of } i_{r-1}\text{th factor within } i_{r-2}\text{th factor,} \\ \xi_{i_r(i_{r-1})} &= \text{effect of } i_r\text{th factor within } i_{r-1}\text{th factor,} \end{aligned}$$

and

$$e_{i_{r+1}(i_r)} = \text{error or residual effect (among observations within } i_r\text{th factor).}$$

Here, the notation $\delta_{i_k(i_{k-1})}$ means that the i_kth factor is nested within the i_{k-1}th factor, and, consequently, in all the preceding factors. It is assumed that α_{i_1}s, $\beta_{i_2(i_1)}$s, $\gamma_{i_3(i_2)}$s, $\ldots$, $\eta_{i_{r-1}(i_{r-2})}$s, $\xi_{i_r(i_{r-1})}$s, and $e_{i_{r+1}(i_r)}$s are mutually and completely uncorrelated random variables with means zero and variances $\sigma_1^2, \sigma_2^2, \sigma_3^2, \ldots, \sigma_{r-1}^2, \sigma_r^2$, and σ_{r+1}^2, respectively. The parameters σ_i^2 $(i = 1, \ldots, r+1)$ are the variance components of the model in (17.1.1).

Now, let

$$\begin{aligned}
a &= \text{number of levels of the } i_1\text{th factor } (i_1 = 1, 2, \ldots, a), \\
b_{i_1} &= \text{number of levels of the } i_2\text{th factor } (i_2 = 1, 2, \ldots, b_{i_1}), \\
c_{i_1 i_2} &= \text{number of levels of the } i_3\text{th factor } (i_3 = 1, 2, \ldots, c_{i_1 i_2}), \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
\ell_{i_1 i_2 \ldots i_{r-2}} &= \text{number of levels of the } i_{r-1}\text{th factor } (i_{r-1} = 1, 2, \ldots, \ell_{i_1 i_2 \ldots i_{r-2}}), \\
m_{i_1 i_2 \ldots i_{r-1}} &= \text{number of levels of the } i_r\text{th factor } (i_r = 1, 2, \ldots, m_{i_1 i_2 \ldots i_{r-1}}),
\end{aligned}$$

and

$$n_{i_1 i_2 \ldots i_r} = \text{number of observations within the } i_{r+1}\text{th factor } i_{r+1} = 1, 2, \ldots, n_{i_1 i_2 \ldots i_r}).$$

Further, introduce the following notation for the sums of the number of levels for different factors:

$$b = \sum_{i_1} b_{i_1}, \qquad c = \sum_{i_1} \sum_{i_2} c_{i_1 i_2}, \ldots,$$

$$\ell = \sum_{i_1} \sum_{i_2} \cdots \sum_{i_{r-2}} \ell_{i_1 i_2 \ldots i_{r-2}}, \qquad m = \sum_{i_1} \sum_{i_2} \cdots \sum_{i_{r-1}} m_{i_1 i_2 \ldots i_{r-1}},$$

and

$$N = \sum_{i_1} \sum_{i_2} \cdots \sum_{i_r} n_{i_1 i_2 \ldots i_r}.$$

Hence, N represents the total number of observations in the sample.

17.2 ANALYSIS OF VARIANCE

The conventional analysis of variance based on Type I sums of squares can be represented as in Table 17.1. The sums of squares in Table 17.1 are defined as

TABLE 17.1 Analysis of variance for the model in (17.1.1).

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Factor A_1	$a-1$	SS_1	MS_1	$\sigma^2_{r+1} + c_{1,r}\sigma^2_r + c_{1,r-1}\sigma^2_{r-1} + \cdots + c_{1,3}\sigma^2_3 + c_{1,2}\sigma^2_2 + c_{1,1}\sigma^2_1$
Factor A_2 within A_1	$b-a$	SS_2	MS_2	$\sigma^2_{r+1} + c_{2,r}\sigma^2_r + c_{2,r-1}\sigma^2_{r-1} + \cdots + c_{2,3}\sigma^2_3 + c_{2,2}\sigma^2_2$
Factor A_3 within A_2	$c-b$	SS_3	MS_3	$\sigma^2_{r+1} + c_{3,r}\sigma^2_r + c_{3,r-1}\sigma^2_{r-1} + \cdots + c_{3,3}\sigma^2_3$
⋮	⋮	⋮	⋮	⋮
Factor A_R within A_{R-1}	$m-\ell$	SS_R	MS_R	$\sigma^2_{r+1} + c_{r,r}\sigma^2_r$
Error (factor A_{R+1})	N-m	SS_{R+1}	MS_{R+1}	σ^2_{r+1}

follows:

$$\mathrm{SS}_1 = \sum_{i_1} \frac{y_{i_1}^2}{n_{i_1}} - \frac{G^2}{N},$$

$$\mathrm{SS}_2 = \sum_{i_1}\sum_{i_2} \frac{y_{i_1 i_2}^2}{n_{i_1 i_2}} - \sum_{i_1} \frac{y_{i_1}^2}{n_{i_1}},$$

$$\mathrm{SS}_3 = \sum_{i_1}\sum_{i_2}\sum_{i_3} \frac{y_{i_1 i_2 i_3}^2}{n_{i_1 i_2 i_3}} - \sum_{i_1}\sum_{i_2} \frac{y_{i_1 i_2}^2}{n_{i_1 i_2}},$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$\mathrm{SS}_R = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_{r-1}}\sum_{i_r} \frac{y_{i_1 i_2 \ldots i_{r-1} i_r}^2}{n_{i_1 i_2 \ldots i_{r-1} i_r}} - \sum_{i_1}\sum_{i_2}\cdots\sum_{i_{r-2}}\sum_{i_{r-1}} \frac{y_{i_1 i_2 \ldots i_{r-2} i_{r-1}}^2}{n_{i_1 i_2 \ldots i_{r-2} i_{r-1}}},$$

and

$$\mathrm{SS}_{R+1} = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_r}\sum_{i_{r+1}} y_{i_1 i_2 \ldots i_r i_{r+1}}^2 - \sum_{i_1}\sum_{i_2}\cdots\sum_{i_{r-1}}\sum_{i_r} \frac{y_{i_1 i_2 \ldots i_{r-1} i_r}^2}{n_{i_1 i_2 \ldots i_{r-1} i_r}},$$

$$G = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_r}\sum_{i_{r+1}} y_{i_1 i_2 \ldots i_r i_{r+1}},$$

where

$$y_{i_1} = \sum_{i_2}\sum_{i_3}\cdots\sum_{i_r}\sum_{i_{r+1}} y_{i_1 i_2 \ldots i_r i_{r+1}},$$

$$y_{i_1 i_2} = \sum_{i_3}\sum_{i_4}\cdots\sum_{i_r}\sum_{i_{r+1}} y_{i_1 i_2 \ldots i_r i_{r+1}},$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$y_{i_1 i_2 \ldots i_{r-2} i_{r-1}} = \sum_{i_r}\sum_{i_{r+1}} y_{i_1 i_2 \ldots i_r i_{r+1}},$$

$$y_{i_1 i_2 \ldots i_{r-1} i_r} = \sum_{i_{r+1}} y_{i_1 i_2 \ldots i_r i_{r+1}};$$

$$n_{i_1} = \sum_{i_2}\sum_{i_3}\cdots\sum_{i_{r-1}}\sum_{i_r} n_{i_1 i_2 \ldots i_{r-1} i_r},$$

$$n_{i_1 i_2} = \sum_{i_3}\sum_{i_4}\cdots\sum_{i_{r-1}}\sum_{i_r} n_{i_1 i_2 \ldots i_{r-1} i_r},$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$n_{i_1 i_2 \ldots i_{r-3} i_{r-2}} = \sum_{i_{r-1}}\sum_{i_r} n_{i_1 i_2 \ldots i_{r-1} i_r},$$

and

$$n_{i_1 i_2 \ldots i_{r-2} i_{r-1}} = \sum_{i_r} n_{i_1 i_2 \ldots i_{r-1} i_r}.$$

The mean squares as usual are obtained by dividing the sums of squares by the corresponding degrees of freedom. The results on expected mean squares are outlined in the following section.

17.3 EXPECTED MEAN SQUARES

It can be shown that the coefficients of the variance components in the expected mean square column are given by (see, e.g., Gates and Shiue, 1962)

$$c_{k,t} = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_t} n^2_{i_1 i_2 \ldots i_t}\left[\frac{1}{n_{i_1 i_2 \ldots i_k}} - \frac{1}{n_{i_1 i_2 \ldots i_{k-1}}}\right]\frac{1}{v_k} \quad \text{for } k \le t \le r,$$
$$0 \quad \text{for } k > t, \tag{17.3.1}$$

where v_k is the degrees of freedom for the kth source of variation. From (17.3.1), it follows that

$$c_{k,k} = \left[N - \sum_{i_1}\sum_{i_2}\cdots\sum_{i_k}\frac{n^2_{i_1 i_2 \ldots i_k}}{n_{i_1 i_2 \ldots i_{k-1}}}\right]\frac{1}{v_k}, \tag{17.3.2}$$

where, if $k = 1$, $n_{i_1 i_2 \ldots i_{k-1}}$ should be replaced by N. Furthermore, the relation (17.3.1) can be written as

$$v_k c_{k,t} = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_t}\frac{n^2_{i_1 i_2 \ldots i_t}}{n_{i_1 i_2 \ldots i_k}} - \sum_{i_1}\sum_{i_2}\cdots\sum_{i_t}\frac{n^2_{i_1 i_2 \ldots i_t}}{n_{i_1 i_2 \ldots i_{k-1}}}. \tag{17.3.3}$$

The expression in (17.3.3) has the computational advantage in that for a given variance component, the first term is duplicated in the lower order and the second term in the next higher-order mean square.

Some simplifications in the formula in (17.3.1) occur if the number of subclasses at some particular stage of sampling are assumed to be equal. Thus, for example, if the number of observations in the last-stage of sampling is the same, i.e., $n_{i_1 i_2 \ldots i_r} = n$, then $c_{k,r} = n$, is the coefficient of σ_r^2 in all mean square expectations and

$$c_{k,t} = n \sum_{i_1} \sum_{i_2} \cdots \sum_{i_t} m^2_{i_1 i_2 \ldots i_t} \left[\frac{1}{m_{i_1 i_2 \ldots i_k}} - \frac{1}{m_{i_1 i_2 \ldots i_{k-1}}} \right] \frac{1}{v_k} \quad \text{for } k \leq t \leq r-1,$$
$$0 \quad \text{for } k > t$$

are the coefficients of the remaining variance components. Furthermore, if the number of levels of the last factor and the next to the last factor are the same, i.e., $n_{i_1 i_2 \ldots i_r} = n$ and $m_{i_1 i_2 \ldots i_{r-1}} = m$, then $c_{k,r} = n$, $c_{k,r-1} = mn$, and

$$c_{k,t} = mn \sum_{i_1} \sum_{i_2} \cdots \sum_{i_t} \ell^2_{i_1 i_2 \ldots i_t} \left[\frac{1}{\ell_{i_1 i_2 \ldots i_k}} - \frac{1}{\ell_{i_1 i_2 \ldots i_{k-1}}} \right] \quad \text{for } k \leq t \leq r-2,$$
$$0 \quad \text{for } k > t.$$

The results in (17.3.1) were first obtained by Ganguli (1941) and King and Henderson (1954) have given a detailed algebraic derivation. The expected values of the mean squares for hierarchical classifications have also been given by Finker et al. (1943) and Hetzer et al. (1944). A general procedure for determining the coefficients $c_{k,t}$ has also been given rather independently by Gates and Shiue (1962) and Gower (1962). Hartley (1967) has presented a very general method for the calculation of expected mean squares in an analysis of variance, based on manipulating vectors containing all zeros or unity as if they were observed values. Khattree et al. (1997) present a general formula for the expected mean square for an r-way staggered nested design considered in Section 16.7. General formulas for expectations, variances and covariances of the mean squares for staggered nested designs are also given by Ojima (1998). Pulley (1957) has developed a computer program to calculate the sums of squares, mean squares, and all coefficients in the expected mean squares for the unbalanced nested analysis of variance with as many as four factors. Another computer program that will analyze unbalanced nested designs up to nine factors and 99 observations and can be modified to accommodate more factors and observations has been given by Postma and White (1975).

17.4 ESTIMATION OF VARIANCE COMPONENTS

In this section, we briefly outline some general methods of estimating variance components.

17.4.1 ANALYSIS OF VARIANCE ESTIMATORS

The analysis of variance (ANOVA) estimators of variance components can be obtained by first equating observed mean squares in the analysis of variance Table 17.1 to their respective expected values expressed as linear combinations of the unknown variance components. The resulting equations are then solved for the variance components to yield the desired estimators.

Denoting the estimators as $\hat{\sigma}_1^2, \hat{\sigma}_2^2, \ldots, \hat{\sigma}_r^2$, and $\hat{\sigma}_{r+1}^2$, and defining

$$\hat{\boldsymbol{\sigma}}^2 = (\hat{\sigma}_1^2, \hat{\sigma}_2^2, \ldots, \hat{\sigma}_r^2, \hat{\sigma}_{r+1}^2)', \tag{17.4.1}$$

$$\boldsymbol{M} = (\mathrm{MS}_1, \mathrm{MS}_2, \ldots, \mathrm{MS}_R, \mathrm{MS}_{R+1})', \tag{17.4.2}$$

and an upper triangular matrix $\boldsymbol{C}$ as

$$\boldsymbol{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1k} & \cdots & c_{1r} & 1 \\ 0 & c_{22} & \cdots & c_{2k} & \cdots & c_{2r} & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & 1 \\ 0 & 0 & \cdots & c_{kk} & \cdots & c_{kr} & 1 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & 1 \\ 0 & 0 & \cdots & 0 & \cdots & c_{rr} & 1 \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_1 \\ \boldsymbol{C}_2 \\ \cdot \\ \vdots \\ \boldsymbol{C}_k \\ \vdots \\ \boldsymbol{C}_r \\ \boldsymbol{C}_{r+1} \end{bmatrix}, \tag{17.4.3}$$

the equations giving the desired estimates can be written as

$$\boldsymbol{M} = \boldsymbol{C}\boldsymbol{\sigma}^2. \tag{17.4.4}$$

The solution to (17.4.4) yields

$$\hat{\boldsymbol{\sigma}}^2 = \boldsymbol{C}^{-1}\boldsymbol{M}. \tag{17.4.5}$$

Postma and White (1975) published a computer program to calculate the ANOVA estimates of variance components in the general r-way unbalanced nested design. Nelson (1983) presented a BASIC computer program that calculates the ANOVA estimates of variance components up to $r = 5$. More recently, Naik and Khattree (1998) have given a computer program to estimate variance components in an r-way staggered nested design.

17.4.2 SYMMETRIC SUMS ESTIMATORS

For symmetric sums estimators, we consider expected values for products and squares of differences of the observations. From the model in (17.1.1), the expected values of products of the observations are

$$E(y_{i_1 i_2 \ldots i_r i_{r+1}} y_{i'_1 i'_2 \ldots i'_r i'_{r+1}})$$

$$
= \begin{cases}
\mu^2, & i_1 \neq i_1', \\
\mu^2 + \sigma_1^2, & i_1 = i_1',\ i_2 \neq i_2', \\
\mu^2 + \sigma_1^2 + \sigma_2^2, & i_1 = i_1',\ i_2 = i_2',\ i_3 \neq i_3', \\
\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots & \\
\mu^2 + \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_j^2, & i_1 = i_1',\ i_2 = i_2',\ \ldots,\ i_j = i_j',\ i_{j+1} \neq i_{j+1}', \\
\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots & \\
\mu^2 + \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2 + \sigma_{r+1}^2, & i_1 = i_1',\ i_2 = i_2',\ \ldots,\ i_r = i_r',\ i_{r+1} = i_{r+1}',
\end{cases}
\tag{17.4.6}
$$

where $i_1, i_1' = 1, 2, \ldots, a$; $i_2 = 1, 2, \ldots, b_{i_1}$; $i_2' = 1, 2, \ldots, b_{i_1'}$; $i_3 = 1, 2, \ldots, c_{i_1 i_2}$; $i_3' = 1, 2, \ldots, c_{i_1' i_2'}$; $\ldots$, $i_{r-1} = 1, 2, \ldots, \ell_{i_1 i_2 \ldots i_{r-2}}$; $i_{r-1}' = 1, 2, \ldots, \ell_{i_1' i_2' \ldots i_{r-2}'}$; $i_r = 1, 2, \ldots, m_{i_1 i_2 \ldots i_{r-1}}$; $i_r' = 1, 2, \ldots, m_{i_1' i_2' \ldots i_{r-1}'}$; $i_{r+1} = 1, 2, \ldots, n_{i_1 i_2 \ldots i_r}$; $i_{r+1}' = 1, 2, \ldots, n_{i_1' i_2' \ldots i_r'}$. From (17.4.6), the normalized symmetric sums are

$$
\begin{aligned}
g_m &= \frac{1}{\sum_{i_1} n_{i_1 \ldots\ldots}(N - n_{i_1 \ldots\ldots})} \sum_{\substack{i_1, i_1' \\ i_1 \neq i_1'}} y_{i_1 \ldots\ldots} y_{i_1' \ldots\ldots} \\
&= \frac{1}{N^2 - k_r} \left(y_{\ldots\ldots}^2 - \sum_{i_1} y_{i_1 \ldots\ldots}^2 \right), \\
g_1 &= \frac{\sum_{i_1} \sum_{\substack{i_2, i_2' \\ i_2 \neq i_2'}} y_{i_1 i_2 \ldots\ldots} y_{i_1 i_2' \ldots\ldots}}{\sum_{i_1} \sum_{i_2} n_{i_1 i_2 \ldots\ldots}(n_{i_1 \ldots\ldots} - n_{i_1 i_2 \ldots\ldots})} \\
&= \frac{\sum_{i_1} \left(y_{i_1 \ldots\ldots}^2 - \sum_{i_2} y_{i_1 i_2 \ldots\ldots}^2 \right)}{k_r - k_{r-1}}, \\
g_2 &= \frac{\sum_{i_1} \sum_{i_2} \sum_{\substack{i_3, i_3' \\ i_3 \neq i_3'}} y_{i_1 i_2 i_3 \ldots\ldots} y_{i_1 i_2 i_3' \ldots\ldots}}{\sum_{i_1} \sum_{i_2} \sum_{i_3} n_{i_1 i_2 i_3 \ldots\ldots}(n_{i_1 i_2 \ldots\ldots} - n_{i_1 i_2 i_3 \ldots\ldots})} \\
&= \frac{\sum_{i_1} \sum_{i_2} \left(y_{i_1 i_2 \ldots\ldots}^2 - \sum_{i_3} y_{i_1 i_2 i_3 \ldots\ldots}^2 \right)}{k_{r-1} - k_{r-2}}, \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
g_{r-1} &= \frac{\sum_{i_1} \sum_{i_2} \cdots \sum_{i_{r-1}} \sum_{\substack{i_r, i_r' \\ i_r \neq i_r'}} y_{i_1 i_2 \ldots i_{r-1} i_r .} y_{i_1 i_2 \ldots i_{r-1} i_r' .}}{\sum_{i_1} \sum_{i_2} \cdots \sum_{i_r} n_{i_1 i_2 \ldots i_r}(n_{i_1 i_2 \ldots i_{r-1} .} - n_{i_1 i_2 \ldots i_r})} \\
&= \frac{\sum_{i_1} \sum_{i_2} \cdots \sum_{i_{r-1}} \left(y_{i_1 i_2 \ldots i_{r-1} ..}^2 - \sum_{i_r} y_{i_1 i_2 \ldots i_{r-1} i_r .}^2 \right)}{k_2 - k_1},
\end{aligned}
$$

$$g_r = \frac{\sum_{i_1}\sum_{i_2}\cdots\sum_{i_r}\sum_{\substack{i_{r+1},i'_{r+1}\\ i_{r+1}\neq i'_{r+1}}} y_{i_1i_2\ldots i_ri_{r+1}}y_{i_1i_2\ldots i_ri'_{r+1}}}{\sum_{i_1}\sum_{i_2}\cdots\sum_{i_r} n_{i_1i_2\ldots i_r}(n_{i_1i_2\ldots i_r}-1)}$$
$$= \frac{\sum_{i_1}\sum_{i_2}\cdots\sum_{i_r}\left(y^2_{i_1i_2\ldots i_r.} - \sum_{i_{r+1}} y^2_{i_1i_2\ldots i_ri_{r+1}}\right)}{k_1 - k_0},$$

and

$$g_{r+1} = \frac{\sum_{i_1}\sum_{i_2}\cdots\sum_{i_r}\sum_{i_{r+1}} y_{i_1i_2\ldots i_ri_{r+1}}y_{i_1i_2\ldots i_ri_{r+1}}}{\sum_{i_1}\sum_{i_2}\cdots\sum_{i_r} n_{i_1i_2\ldots i_r}}$$
$$= \frac{\sum_{i_1}\sum_{i_2}\cdots\sum_{i_r}\sum_{i_{r+1}} y^2_{i_1i_2\ldots i_ri_{r+1}}}{N},$$

where

$$k_0 = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_r} n_{i_1i_2\ldots i_r} = N,$$
$$k_1 = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_r} n^2_{i_1i_2\ldots i_r},$$
$$k_2 = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_{r-1}} n^2_{i_1i_2\ldots i_{r-1}.},$$
$$\vdots \quad \vdots \quad \vdots$$
$$k_{r-2} = \sum_{i_1}\sum_{i_2}\sum_{i_3} n^2_{i_1i_2i_3\ldots.},$$
$$k_{r-1} = \sum_{i_1}\sum_{i_2} n^2_{i_1i_2\ldots.},$$

and

$$k_r = \sum_{i_1} n^2_{i_1\ldots.}.$$

By equating $g_m, g_1, g_2, \ldots, g_r, g_{r+1}$ to their respective expected values and solving the resulting equations, we obtain the estimators of variance components as (Koch, 1967)

$$\begin{aligned}
\hat{\sigma}^2_{1,\mathrm{SSP}} &= g_1 - g_m,\\
\hat{\sigma}^2_{2,\mathrm{SSP}} &= g_2 - g_1,\\
&\vdots\\
\hat{\sigma}^2_{r,\mathrm{SSP}} &= g_r - g_{r-1},
\end{aligned} \tag{17.4.7}$$

and

$$\hat{\sigma}^2_{r+1,\text{SSP}} = g_{r+1} - g_r.$$

For symmetric sums based on the expected values of the squares of differences of observations, we have

$$E(y_{i_1 i_2 \ldots i_r i_{r+1}} - y_{i'_1 i'_2 \ldots i'_r i'_{r+1}})^2$$
$$= \begin{cases} 2\sigma^2_{r+1}, & i_1 = i'_1,\ i_2 = i'_2,\ \ldots,\ i_r = i'_r,\ i_{r+1} \neq i'_{r+1}, \\ 2(\sigma^2_{r+1} + \sigma^2_r), & i_1 = i'_1,\ i_2 = i'_2,\ \ldots,\ i_{r-1} = i'_{r-1},\ i_r \neq i'_r, \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots & \\ 2(\sigma^2_{r+1} + \sigma^2_r + \cdots + \sigma^2_2), & i_1 = i'_1,\ i_2 \neq i'_2, \\ 2(\sigma^2_{r+1} + \sigma^2_r + \cdots + \sigma^2_2 + \sigma^2_1), & i_1 \neq i'_1, \end{cases} \tag{17.4.8}$$

where $i_1, i'_1 = 1, 2, \ldots, a$; $i_2 = 1, 2, \ldots, b_{i_1}$; $i'_2 = 1, 2, \ldots, b_{i'_1}$; $i_3 = 1, 2, \ldots, c_{i_1 i_2}$; $i'_3 = 1, 2, \ldots, c_{i'_1 i'_2}$; $\ldots$; $i_{r-1} = 1, 2, \ldots, \ell_{i_1 i_2 \ldots i_{r-2}}$; $i'_{r-1} = 1, 2, \ldots, \ell_{i'_1 i'_2 \ldots i'_{r-2}}$; $i_r = 1, 2, \ldots, m_{i_1 i_2 \ldots i_{r-1}}$; $i'_r = 1, 2, \ldots, m_{i'_1 i'_2 \ldots i'_{r-1}}$; $i_{r+1} = 1, 2, \ldots, n_{i_1 i_2 \ldots i_r}$; $i'_{r+1} = 1, 2, \ldots, n_{i'_1 i'_2 \ldots i'_r}$. From (17.4.8), the normalized symmetric sums are

$$h_{r+1} = \frac{\sum_{i_1} \sum_{i_2} \cdots \sum_{i_r} \sum_{\substack{i_{r+1}, i'_{r+1} \\ i_{r+1} \neq i'_{r+1}}} (y_{i_1 i_2 \ldots i_r i_{r+1}} - y_{i_1 i_2 \ldots i_r i'_{r+1}})^2}{\sum_{i_1} \sum_{i_2} \cdots \sum_{i_r} n_{i_1 i_2 \ldots i_r} (n_{i_1 i_2 \ldots i_r} - 1)}$$
$$= \frac{2 \sum_{i_1} \sum_{i_2} \cdots \sum_{i_r} n_{i_1 i_2 \ldots i_r} \left(\sum_{i_{r+1}} y^2_{i_1 i_2 \ldots i_r i_{r+1}} - n_{i_1 i_2 \ldots i_r} \bar{y}^2_{i_1 i_2 \ldots i_r .} \right)}{k_1 - k_0},$$

$$h_r = \frac{\sum_{i_1} \sum_{i_2} \cdots \sum_{i_{r-1}} \sum_{\substack{i_r, i'_r \\ i_r \neq i'_r}} \sum_{i_{r+1}, i'_{r+1}} (y_{i_1 i_2 \ldots i_r i_{r+1}} - y_{i_1 i_2 \ldots i'_r i'_{r+1}})^2}{\sum_{i_1} \sum_{i_2} \cdots \sum_{i_r} n_{i_1 i_2 \ldots i_r} (n_{i_1 i_2 \ldots i_{r-1} .} - n_{i_1 i_2 \ldots i_r})}$$
$$= \frac{2 \sum_{i_1} \sum_{i_2} \cdots \sum_{i_{r-1}} \sum_{i_r} (n_{i_1 i_2 \ldots i_{r-1} .} - n_{i_1 i_2 \ldots i_r}) \sum_{i_{r+1}} y^2_{i_1 i_2 \ldots i_r i_{r+1}}}{k_2 - k_1}$$
$$- 2g_{r-1},$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$h_2 = \frac{\sum_{i_1} \sum_{\substack{i_2, i'_2 \\ i_2 \neq i'_2}} \sum_{i_3, i'_3} \cdots \sum_{i_r, i'_r} \sum_{i_{r+1}, i'_{r+1}} (y_{i_1 i_2 \ldots i_r i_{r+1}} - y_{i_1 i'_2 \ldots i'_r i'_{r+1}})^2}{\sum_{i_1} \sum_{i_2} n_{i_1 i_2 \ldots .} (n_{i_1 \ldots .} - n_{i_1 i_2 \ldots .})}$$
$$= \frac{2 \sum_{i_1} \sum_{i_2} (n_{i_1 \ldots .} - n_{i_1 i_2 \ldots .}) \sum_{i_3} \cdots \sum_{i_r} \sum_{i_{r+1}} y^2_{i_1 i_2 \ldots i_r i_{r+1}}}{k_r - k_{r-1}} - 2g_1,$$

and

$$h_1 = \frac{\sum_{\substack{i_1,i'_1 \\ i_1 \neq i'_1}} \sum_{i_2,i'_2} \cdots \sum_{i_r,i'_r} \sum_{i_{r+1},i'_{r+1}} (y_{i_1 i_2 \ldots i_r i_{r+1}} - y_{i'_1 i'_2 \ldots i'_r i'_{r+1}})^2}{\sum_{i_1} n_{i_1 \ldots.}(N - n_{i_1 \ldots.})}$$
$$= \frac{2\sum_{i_1}(N - n_{i_1 \ldots.})\sum_{i_2} \cdots \sum_{i_r} \sum_{i_{r+1}} y^2_{i_1 i_2 \ldots i_r i_{r+1}}}{N^2 - k_r} - 2g_m,$$

where $N, k_0, k_1, \ldots, k_{r-1}, k_r, g_m, g_1, g_2, \ldots, g_{r-1}, g_r$ are defined as before.

Equating $h_1, h_2, \ldots, h_r$, and h_{r+1} to their respective expected values, we obtain

$$\begin{aligned} 2\sigma^2_{r+1} &= h_{r+1}, \\ 2(\sigma^2_{r+1} + \sigma^2_r) &= h_r, \\ \vdots \quad & \quad \vdots \\ 2(\sigma^2_{r+1} + \sigma^2_r + \cdots + \sigma^2_2) &= h_2, \\ 2(\sigma^2_{r+1} + \sigma^2_r + \cdots + \sigma^2_2 + \sigma^2_1) &= h_1. \end{aligned} \tag{17.4.9}$$

The variance component estimators obtained by solving the equations in (17.4.9) are (Koch, 1968)

$$\begin{aligned} \hat{\sigma}^2_{r+1,\mathrm{SSD}} &= \frac{1}{2}h_{r+1}, \\ \hat{\sigma}^2_{r,\mathrm{SSD}} &= \frac{1}{2}(h_r - h_{r+1}), \\ \vdots \quad & \quad \vdots \quad \vdots \\ \hat{\sigma}^2_{2,\mathrm{SSD}} &= \frac{1}{2}(h_2 - h_3), \end{aligned} \tag{17.4.10}$$

and

$$\hat{\sigma}^2_{1,\mathrm{SSD}} = \frac{1}{2}(h_1 - h_2).$$

17.4.3 OTHER ESTIMATORS

The ML, REML, MINQUE, and MIVQUE estimators can be developed as special cases of the results for the general case considered in Chapter 10 and are not treated separately for the nested models. With the advent of the high-speed digital computer, the general results on these estimators involving matrix operations can be handled with great speed and accuracy and their explicit algebraic evaluation for this model seems to be rather unnecessary. In addition, some commonly used statistical software packages, such as SAS®, SPSS®, and

BMDP®, have special routines to compute these estimates rather conveniently simply by specifying the model in question. The use of canonical forms for estimating variance components in unbalanced nested designs has been considered by Ojima (1984). More recently, Khattree et al. (1997) have proposed a new approach, known as the principal components method, that yields nonnegative estimates of variance components in an r-way random effects staggered nested design.

17.5 VARIANCES OF ESTIMATORS

From (17.4.5), the variance-covariance matrix of the ANOVA estimators of the variance components is obtained as

$$\mathrm{Var}(\hat{\boldsymbol{\sigma}}^2) = \boldsymbol{C}^{-1}\,\mathrm{Var}(\boldsymbol{M})\boldsymbol{C}^{-1'}.$$

Thus the sampling variances of the ANOVA estimators can be obtained from the knowledge of the variance-covariance matrix of the mean squares. In principle, the same methods as employed in Searle (1961) and Mahamunulu (1963) can be utilized to derive $\mathrm{Var}(\boldsymbol{M})$. However, for the higher-order nested classifications, the algebra tends to be extremely tedious and the notations become very complex to manage. More recently, Khattree et al. (1997) have given expressions for the approximate variances of the principal components estimators.

17.6 CONFIDENCE INTERVALS AND TESTS OF HYPOTHESES

Exact confidence intervals for the error variance component and the ratio of the penultimate component to the error component are constructed in the usual way. Conditions of partial balancedness will allow the use of balanced design formulas for some of the parameters. However, it is necessary to construct approximate intervals for other variance components and their parametric functions. As earlier, one can use unweighted or Type I sums of squares to construct approximate intervals for other components and their parametric functions. For four-way and higher-order nested designs an exact test for the penultimate variance component can be performed using the conventional F-test. However, approximate tests are needed for testing higher-order variance components. Satterthwaite-type tests can be constructed by synthesizing either a numerator component, a denominator component or both. Synthesis of only the denominator or numerator component generally involves a linear combination of correlated mean squares with possibly negative coefficients. However, these mean squares are neither independent nor distributed as multiples of a chi-square variable and it is not clear how the violations of these assumptions will affect the stated test size. One may expect the cancellation effect of the dependence and non-chi-squareness although there is a distinct possibility that rather than counterbalancing each other they may provide a reinforcement for test size disturbances. Generally, one may expect small disturbances for designs with small imbalances and large disturbances for designs with extreme imbalances.

One can also test the significance of variance components by first constructing confidence intervals on them as indicated above. In addition, one may use the general method of likelihood-ratio test described earlier in Section 10.16. Khuri (1990) proposed exact tests concerning the model's variance components when the imbalance occurs in the last stage of the associated design with no missing cells (see also Khuri et al., 1998, Chapter 5). The method is based on a particular transformation that reduces the analysis of the unbalanced model to that of a balanced one. A SAS® matrix software macro for testing the variance components using this procedure has been developed by Gallo et al. (1989). For the design with unbalanced cell frequencies in the last stage, Zhou and Mathew (1994) discuss some tests for variance components using the concept of a generalized p-value. More recently, Khattree and Naik (1995) have developed some new hypothesis testing procedures for variance components in a staggered nested design. Observations obtained from a staggered nested design are represented as a sample vector having a multivariate normal distribution with a certain covariance structure. One can then apply certain multivariate procedures to test the significance of variance components.

In the following, we describe Satterthwaite's (1946) procedure for testing a linear combination of mean squares that seems most convenient and easiest to implement among the available tests. Eisen (1966) used this method to test the significance of a fixed unnested main effect in an unbalanced analysis of variance where all other nested factors are random. However, the calculation of the degrees of freedom for the denominator of the test based on Satterthwaite's procedure tends to be very complicated in the general case. Tietjen and Moore (1968) have described a general and relatively easy method of calculating all the necessary quantities required in the test based on Satterthwaite's procedure. A similar method is also presented by Snee (1974). We outline their approach briefly here.

In order to test the hypothesis $H_0^k : \sigma_k^2 = 0$ vs $H_1^k : \sigma_k^2 > 0$, we take the numerator component of the pseudo F-test as the mean square MS_k with the expected value equal to (see Section 17.4.1)

$$\sigma_{r+1}^2 + c_{k,r}\sigma_r^2 + c_{k,r-1}\sigma_{r-1}^2 + \cdots + c_{k,k+1}\sigma_{k+1}^2 + c_{k,k}\sigma_k^2. \tag{17.6.1}$$

The appropriate denominator component is taken as a linear combination of the mean squares, i.e.,

$$D_k = \sum_i \ell_i \mathrm{MS}_i, \tag{17.6.2}$$

with the expected value equal to

$$\sigma_{r+1}^2 + c_{k,r}\sigma_r^2 + c_{k,r-1}\sigma_{r-1}^2 + \cdots + c_{k,k+1}\sigma_{k+1}^2. \tag{17.6.3}$$

Inasmuch as

$$E(\mathrm{MS}_k) = \boldsymbol{C}_k \boldsymbol{\sigma}^2,$$

it readily follows that

$$\begin{aligned} D_k &= \boldsymbol{C}_k \hat{\boldsymbol{\sigma}}^2 - c_{k,k}\hat{\sigma}_k^2 \\ &= \mathrm{MS}_k - c_{k,k}\hat{\sigma}_k^2, \end{aligned}$$

is the desired denominator component. The degrees of freedom for the numerator are v_k while those for the denominator are calculated as

$$v_k' = \frac{\left(\sum_i \ell_i \mathrm{MS}_i\right)^2}{\sum (\ell_i \mathrm{MS}_i)^2 / v_i}. \tag{17.6.4}$$

The expression in (17.6.4) can be further simplified by writing

$$\boldsymbol{C}^{-1} = (c^{i,j})$$

and noting from (17.6.2) and (17.6.3) that

$$\begin{aligned} D_k &= \hat{\sigma}_{r+1}^2 + c_{k,r}\hat{\sigma}_r^2 + c_{k,r-1}\hat{\sigma}_{r-1}^2 + \cdots + c_{k,k+1}\hat{\sigma}_{k+1}^2 \\ &= \sum_{j=1}^{r+1} c^{r+1,j}\mathrm{MS}_j + c_{k,r}\sum_{j=1}^{r+1} c^{r,j}\mathrm{MS}_j + c_{k,r-1}\sum_{j=1}^{r+1} c^{r-1,j}\mathrm{MS}_j \\ &\quad + \cdots + c_{k,k+1}\sum_{j=1}^{r+1} c^{k+1,j}\mathrm{MS}_j. \end{aligned} \tag{17.6.5}$$

By reorganizing the terms in (17.6.5), we note that $\sum_{i=k+1}^{r+1} c_{k,i}c^{i,j}$ is the coefficient of MS_j. It should be noted that except for a missing nonzero term $c_{k,k}c^{k,j}$ (and other $k-1$ terms each equal to zero), the expression $\sum_{i=k+1}^{r+1} c_{k,i}c^{i,j}$ is the element in the kth row and the jth column of $\boldsymbol{C}\boldsymbol{C}^{-1} = \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix of order $r+1$. By adding and subtracting the term $c_{k,k}c^{k,j}$, we obtain

$$D_k = -\sum_{j=k+1}^{r+1} c_{k,k}c^{k,j}\mathrm{MS}_j. \tag{17.6.6}$$

Inasmuch as the diagonal elements of $\boldsymbol{C}^{-1}$ are the reciprocals of the diagonal elements of $\boldsymbol{C}$ the coefficient of MS_k is zero. The degrees of freedom for D_k are then given by

$$v_k' = D_k^2 \bigg/ \left[\sum_{j=k+1}^{r+1} (c_{k,k}c^{k,j}\mathrm{MS}_j)^2 / v_j\right]. \tag{17.6.7}$$

By knowing the matrix $\boldsymbol{C}$ and $\boldsymbol{C}^{-1}$, the expression in (17.6.7) can be computed with relative ease.

TABLE 17.2 Analysis of variance of the insecticide residue data of Table 16.3.

Source of variation	Degrees of freedom	Sum of squares	Mean square	Expected mean square
Plots	10	1.84041	0.18404	$\sigma_4^2+1.571\sigma_3^2+3\sigma_2^2+7\sigma_1^2$
Samples	22	0.99175	0.04508	$\sigma_4^2+1.214\sigma_3^2+2\sigma_2^2$
Subsamples	22	0.35758	0.01625	$\sigma_4^2+1.500\sigma_3^2$
Error	22	0.22085	0.01004	σ_4^2
Total	76	3.41058		

17.7 A NUMERICAL EXAMPLE

Consider the insecticide residue on celery data reported by Bliss (1967, pp. 352–357) as given in the numerical example of Section 16.4. The analysis of variance given in Table 16.4 is reproduced in Table 17.2 in the notation of σ_i^2 $(i = 1, 2, 3, 4)$. Now, from Table 17.2, the matrix $\boldsymbol{C}$ and the vector $\boldsymbol{M}$ are given by

$$\boldsymbol{C} = \begin{bmatrix} 7 & 3 & 1.571 & 1 \\ 0 & 2 & 1.214 & 1 \\ 0 & 0 & 1.500 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \boldsymbol{M} = \begin{bmatrix} 0.18404 \\ 0.04508 \\ 0.01625 \\ 0.01004 \end{bmatrix}.$$

The variance component estimates are given as

$$\begin{bmatrix} \hat{\sigma}_1^2 \\ \hat{\sigma}_2^2 \\ \hat{\sigma}_3^2 \\ \hat{\sigma}_4^2 \end{bmatrix} = \begin{bmatrix} 7 & 3 & 1.571 & 1 \\ 0 & 2 & 1.214 & 1 \\ 0 & 0 & 1.500 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0.18404 \\ 0.04508 \\ 0.01625 \\ 0.01004 \end{bmatrix}$$

$$= \begin{bmatrix} 0.1429 & -0.2143 & 0.0238 & 0.0476 \\ 0 & 0.5000 & -0.4048 & -0.0952 \\ 0 & 0 & 0.6664 & -0.6661 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.18404 \\ 0.04508 \\ 0.01625 \\ 0.01004 \end{bmatrix}$$

$$= \begin{bmatrix} 0.01750 \\ 0.01501 \\ 0.00414 \\ 0.01004 \end{bmatrix}.$$

Now, we will illustrate the tests of hypotheses using the Satterthwaite procedure by first considering the hypothesis

$$H_0 : \sigma_1^2 = 0 \quad \text{vs.} \quad H_1 : \sigma_1^2 > 0. \tag{17.7.1}$$

The appropriate denominator component for testing the hypothesis in (17.7.1) is

$$\begin{aligned} D_1 &= \hat{\sigma}_4^2 + 1.571\hat{\sigma}_3^2 + 3\hat{\sigma}_2^2 \\ &= 0.01004 + 1.571(0.00414) + 3(0.01501) \\ &= 0.06156. \end{aligned}$$

To calculate the degrees of freedom associated with D_1 by Satterthwaite's procedure, we need to express it as a linear combination of the mean squares. This in turn requires expressing each of the $\hat{\sigma}_i^2$s in terms of the mean squares. Equating mean squares to their respective expected values and solving the resulting equations, we obtain

$$\hat{\sigma}_4^2 = \text{MS}_4, \qquad \hat{\sigma}_3^2 = 0.6664\text{MS}_3 - 0.6661\text{MS}_4,$$
$$\hat{\sigma}_2^2 = 0.5000\text{MS}_2 - 0.4048\text{MS}_3 - 0.0952\text{MS}_4.$$

Substituting these values into D_1, we obtain

$$D_1 = -0.3351\text{MS}_4 - 0.1670\text{MS}_3 + 1.5000\text{MS}_2.$$

The degrees of freedom for D_1 is then given by

$$\begin{aligned} v_1' &= (0.06156)^2 / \left[\frac{(-0.3351\text{MS}_4)^2}{22} + \frac{(-0.1670\text{MS}_3)^2}{22} + \frac{(1.5000\text{MS}_2)^2}{22} \right] \\ &= 18.2. \end{aligned}$$

From (17.6.7), the degrees of freedom corresponding to the denominator component is calculated as

$$\begin{aligned} v_1' &= D_1^2 / \left[\sum_{j=2}^{4} (c_{1,1}c^{1,j}\text{MS}_j)^2 / v_j \right] \\ &= (0.0616)^2 / \left\{ \frac{[7(-0.2143)(0.04508)]^2}{22} - \frac{[7(0.0238)(0.01625)]^2}{22} \right. \\ &\qquad \left. + \frac{[7(0.0476)(0.01004)]^2}{22} \right\} \\ &= 18.2. \end{aligned}$$

We can similarly construct the denominator components and the corresponding degrees of freedom for testing the hypotheses on σ_2^2 and σ_3^2. The resulting quantities including the values of test statistics and the p-values are outlined in Table 17.3. Note that both plots and samples within plots exceeded their errors significantly, but the mean square for the subsamples was not significant although appreciably larger than that for determinations.

TABLE 17.3 Tests of hypotheses for $\sigma_i^2 = 0, i = 1, 2, 3$.

Hypothesis	Tests statistic		Degrees of freedom		F-ratio*	p-value
	Numerator	Denominator	Numerator	Denominator		
$\sigma_1^2 = 0$	0.18404	0.06156	10	18.2	2.99	0.021
$\sigma_2^2 = 0$	0.04508	0.01507	22	28.3	2.99	0.003
$\sigma_3^2 = 0$	0.01625	0.01004	22	22.0	1.62	0.133

*Bliss (1967, p. 355) ignored the unbalanced structure of the design and computed F-ratios, based on conventional F-tests, by dividing each mean square by that in the next line in Table 17.2.

EXERCISES

1. Spell out proof of the result in (17.3.1).
2. Apply the method of "synthesis" to derive the expected mean squares given in Section 17.3.
3. Show that the ANOVA estimators (17.4.5) reduce to the corresponding estimators (7.7.3) for balanced data.
4. Show that the symmetric sums estimators (17.4.7) and (17.4.10) reduce to the ANOVA estimators (7.7.3) for balanced data.
5. Show that for a q-stage staggered nested design, the expected values of the mean squares are given by (Khattree et al., 1997)

$$E(\mathrm{MS}_i) = \sum_{j=1}^{q} d_{ij}\sigma_j^2,$$

where

$$d_{ij} = \begin{cases} 1 + \frac{j(j-1)}{q}, & i = 1;\ j = 1, 2, \ldots, q, \\ 1 + \frac{j(j-1)}{(q+1-i)(q+2-i)}, & i = 2, \ldots, q;\ j = 1, \ldots, q+1-i, \\ 0, & i = 2, \ldots, q;\ j = q+2-i, \ldots, q. \end{cases}$$

6. Refer to Exercise 5 above and show that the ANOVA estimators of the variance components in a q-stage staggered nested design are given by (Khattree et al., 1997).

$$\hat{\boldsymbol{\sigma}}^2 = \boldsymbol{D}^{-1}\boldsymbol{M},$$

where $\hat{\boldsymbol{\sigma}}^{2'} = (\hat{\sigma}_q^2, \ldots, \hat{\sigma}_1^2)$, $\boldsymbol{M}' = (\mathrm{MS}_1, \ldots, \mathrm{MS}_q)$, and $\boldsymbol{D} = (d_{ij})$.

Bibliography

C. I. Bliss (1967), *Statistics in Biology*, McGraw–Hill, New York.

E. J. Eisen (1966), The quasi-F test for an unnested fixed factor in an unbalanced hierarchal design with a mixed model, *Biometrics*, **22**, 937–942.

A. L. Finker, J. J. Morgan, and R. J. Monroe (1943), *Methods of Estimating Farm Employment from Sample Data in North Carolina*, Technical Bulletin 75, North Carolina State Agriculture Experimental Station, Raleigh, NC.

J. Gallo, R. MacElderry, and A. I. Khuri (1989), A SAS matrix software macro for estimating the random effects with unequal cell frequencies in the last stage, in *Procedings of the* 14*th SAS Users Group International Conference*, SAS Institute, Cary, NC, 1277–1282.

M. Ganguli (1941), A note on nested sampling, *Sankhyā*, **5**, 449–452.

C. E. Gates and C. Shiue (1962), The analysis of variance of the s-stage hierarchal classification, *Biometrics*, **18**, 529–536.

J. C. Gower (1962), Variance component estimation for unbalanced hierarchal classifications, *Biometrics*, **18**, 537–542.

H. O. Hartley (1967), Expectations, variances and covariances of ANOVA mean squares by "synthesis," *Biometrics*, **23**, 105–114; corrigenda, **23**, 853.

H. O. Hetzer, G. E. Dickerson, and J. H. Zeller (1944), Heritability of type in Poland China swine as evaluated by scoring, *J. Animal Sci.*, **3**, 390–398.

R. Khattree and D. N. Naik (1995), Statistical tests for random effects in staggered nested designs, *J. Appl. Statist.*, **22**, 495–505.

R. Khattree, D. N. Naik, and R. L. Mason (1997), Estimation of variance components in staggered nested designs, *J. Appl. Statist.*, **24**, 395–408.

A. I. Khuri (1990), Exact test for random models with unequal cell frequencies in the last stage, *J. Statist. Plann. Inference*, **24**, 177–193.

A. I. Khuri, T. Mathew, and B. K. Sinha (1998), *Statistical Tests for Mixed Linear Models*, Wiley, New York.

S. C. King and C. R. Henderson (1954), Variance component analysis in heritability studies, *Poultry Sci.*, **331**, 147–154.

G. G. Koch (1967), A general approach to the estimation of variance components, *Techometrics*, **9**, 93–118.

G. G. Koch (1968), Some further remarks concerning "A general approach to estimation of variance components," *Technometrics*, **10**, 551–558.

D. M. Mahamunulu (1963), Sampling variances of the estimates of variance components in the unbalanced three-way nested classification, *Ann. Math. Statist.*, **34**, 521–527.

D. N. Naik and R. Khattree (1998), A computer program to estimate variance components in staggered nested designs, *J. Quality Tech.*, **30**, 292–297.

L. S. Nelson (1983), Variance estimation using staggered nested designs, *J. Quality Tech.*, **15**, 195–198.

Y. Ojima (1984), The use of canonical forms for estimating variance components in unbalanced nested designs, *Rep. Statist. Appl. Res.* (*Japan*), **31**, 1–18.

Y. Ojima (1998), General formula for expectations, variances and covariances of the mean squares for staggered nested designs, *J. Appl. Statist.*, **25**, 785–799.

B. J. Postma and J. S. White (1975), Variance components for unbalanced N-level hierarchic design, *J. Quality Tech.*, **7**, 144–149; corrigenda, **8**, 208.

P. Pulley, Jr. (1957), *A Program for the Analysis of Variance of Hierarchal Classification Design*, M.S. thesis, Oklahoma State University, Stillwater, OK.

F. E. Satterthwaite (1946), An approximate distribution of estimates of variance components, *Biometrics Bull.*, **2**, 110–114.

S. R. Searle (1961), Variance components in the unbalanced two-way nested classification, *Ann. Math. Statist.*, **32**, 1161–1166.

R. D. Snee (1974), Computation and use of expected mean squares in analysis of variance, *J. Quality Tech.*, **6**, 128–137.

G. L. Tietjen and R. H. Moore (1968), On testing significance of variance in the unbalanced nested analysis of variance, *Biometrics*, **24**, 423–429.

L. Zhou and T. Mathew (1994), Some tests for variance components using generalized p-values, *Techometrics*, **36**, 394–402.

Appendices

A TWO USEFUL LEMMAS IN DISTRIBUTION THEORY

In this appendix, we give two lemmas frequently employed in the derivation of expected mean squares in an analysis of variance.

Lemma A.1. *Let $z_1, z_2, \ldots, z_n$ be uncorrelated random variables with mean μ and variance σ^2. Then $E[\sum_{i=1}^{n}(z_i - \bar{z})^2] = (n-1)\sigma^2$, where $\bar{z} = \sum_{i=1}^{n} z_i/n$.*

Proof. By definition, we have

$$E\left[\sum_{i=1}^{n}(z_i - \bar{z})^2\right] = \sum_{i=1}^{n} E(z_i^2) - nE(\bar{z})^2.$$

Now, noting that

$$E(z_i^2) = \mu^2 + \sigma^2$$

and

$$E(\bar{z})^2 = (n\mu^2 + \sigma^2)/n,$$

we obtain

$$E\left[\sum_{i=1}^{n}(z_i - \bar{z})^2\right] = n(\mu^2 + \sigma^2) - (n\mu^2 + \sigma^2)$$
$$= (n-1)\sigma^2. \qquad \square$$

Lemma A.2. *Let $z_1, z_2, \ldots, z_n$ be independent normal random variables with mean μ and variance σ^2. Then $\sum_{i=1}^{n}(z_i - \bar{z})^2 \sim \sigma^2\chi^2[n-1]$.*

Proof. First, we show that $(z_i - \bar{z})$ and $(\bar{z} - \mu)$ are statistically independent. Since

$$z_i \sim N(\mu, \sigma^2), \quad i = 1, 2, \ldots, n,$$

and

$$\bar{z} \sim N(\mu, \sigma^2/n),$$

it readily follows that

$$z_i - \bar{z} \sim N\left(0, \frac{n-1}{n}\sigma^2\right), \quad i = 1, 2, \ldots, n,$$

and

$$\bar{z} - \mu \sim N(0, \sigma^2/n).$$

Furthermore,

$$\begin{aligned}\text{Cov}\{z_i - \bar{z}, \bar{z} - \mu\} &= E\{(z_i - \bar{z})(\bar{z} - \mu)\} \\ &= E(z_i\bar{z}) - \mu E(z_i) - E(\bar{z}^2) + \mu E(\bar{z}) \\ &= \left(\mu^2 + \frac{\sigma^2}{n}\right) - \mu^2 - \left(\mu^2 + \frac{\sigma^2}{n}\right) + \mu^2 \\ &= 0, \quad i = 1, 2, \ldots, n.\end{aligned}$$

Thus $(z_i - \bar{z})$ and $(\bar{z} - \mu)$ are uncorrelated normal random variables, which implies that they are statistically independent.

Now,

$$\begin{aligned}\sum_{i=1}^{n}(z_i - \bar{z})^2 &= \sum_{i=1}^{n}\{(z_i - \mu) - (\bar{z} - \mu)\}^2 \\ &= \sum_{i=1}^{n}(z_i - \mu)^2 - 2(\bar{z} - \mu)\sum_{i=1}^{n}(z_i - \mu) + n(\bar{z} - \mu)^2 \\ &= \sum_{i=1}^{n}(z_i - \mu)^2 - n(\bar{z} - \mu)^2,\end{aligned}$$

so that

$$\sum_{i=1}^{n}(z_i - \mu)^2 = \sum_{i=1}^{n}(z_i - \bar{z})^2 + \frac{\sigma^2(\bar{z} - \mu)^2}{(\sigma/\sqrt{n})^2}.$$

The left-hand term in the above equation has a $\sigma^2\chi^2[n]$ distribution and the second term on the right side has a $\sigma^2\chi^2[1]$ distribution. Since both terms on the right side are independent, it immediately follows from the reproductive property of the chi-square distribution that

$$\sum_{i=1}^{n}(z_i - \bar{z})^2 \sim \sigma^2\chi^2[n-1]. \qquad \square$$

B SOME USEFUL LEMMAS FOR A CERTAIN MATRIX

In this appendix, we present two useful lemmas on the determinant of a matrix $\boldsymbol{A}$ and its inverse $\boldsymbol{A}^{-1}$, which frequently arise in many linear model problems. Let the matrix $\boldsymbol{A}$ be defined by

$$\boldsymbol{A} = \begin{bmatrix} a+b & a & a & \dots & a \\ a & a+b & a & \dots & a \\ \vdots & \vdots & \vdots & \vdots & \\ a & a & a & \dots & a+b \end{bmatrix}, \tag{B.1}$$

where a and b are either scalars or square matrices of the same order. If a and b are scalars, the matrix $\boldsymbol{A}$ can be written as

$$\boldsymbol{A} = b\boldsymbol{I}_n + a\boldsymbol{J}_n, \tag{B.2}$$

where $\boldsymbol{I}_n$ is an $n \times n$ identity matrix and $\boldsymbol{J}_n$ is an $n \times n$ matrix with each element equal to 1.

Lemma B.1. *For the matrix* $\boldsymbol{A}$ *defined by* (B.1), *we have*

$$|\boldsymbol{A}| = (|b + na|)(|b|^{n-1}),$$

and

$$|\boldsymbol{A}^{-1}| = (|b + na|^{-1})(|b|^{1-n}),$$

where $|\boldsymbol{D}|$ *designates the determinant of a matrix* $\boldsymbol{D}$.

Proof. For a and b scalars, the proof is given in Wilks (1962, p. 109). The proof when a and b are matrices follows readily from the case in which a and b are scalars. □

Lemma B.2. *For the matrix* A *given by* (B.2), *we have*

$$A^{-1} = \theta_1 \boldsymbol{I}_n + \theta_2 \boldsymbol{J}_n,$$

where

$$\theta_1 = 1/b$$

and

$$\theta_2 = -(a/b)(b + na)^{-1} \quad \textit{for } b \neq 0,\ b \neq -na.$$

Proof. See Graybill (1961, p. 340). □

Bibliography

F. A. Graybill (1961), *An Introduction to Linear Statistical Models*, Vol. I, McGraw–Hill, New York.

S. S. Wilks (1962), *Mathematical Statistics*, Wiley, New York.

C INCOMPLETE BETA FUNCTION

The function defined by

$$B_x(\ell, m) = \int_0^x t^{\ell-1}(1-t)^{m-1}dt,$$

where $0 < x < 1$ and ℓ and m are known constants, is called the incomplete beta function. Notice that

$$B_1(\ell, m) = \int_0^1 t^{\ell-1}(1-t)^{m-1}dt = B(\ell, m),$$

which is known as the complete beta function. In the normalized form

$$I_x(\ell, m) = B_x(\ell, m)/B(\ell, m),$$

the function represents the cumulative probability function of the beta distribution. Thus

$$I_1(\ell, m) = 1$$

and

$$I_x(\ell, m) = 1 - I_{1-x}(m, \ell).$$

The following recurrent relations hold for the normalized incomplete beta function:

$$\begin{aligned}
I_x(\ell, m+1) &= I_x(\ell, m) + [mB(\ell, m)]^{-1}x^{\ell}(1-x)^m, \\
I_x(\ell+1, m) &= I_x(\ell, m) - [\ell B(\ell, m)]^{-1}x^{\ell}(1-x)^m, \\
I_x(\ell, m) &= xI_x(\ell-1, m) + (1-x)I_x(\ell, m-1), \\
I_x(\ell, m) &= \frac{1}{x}\{I_x(\ell+1, m) - (1-x)I_x(\ell+1, m-1)\}, \\
I_x(\ell, m) &= \frac{\ell}{a(1-x)+m}\{I_x(\ell, m+1) \\
&\qquad + (1-x)I_x(\ell+1, m-1)\}, \\
I_x(\ell, m) &= \frac{1}{\ell+m}\{\ell I_x(\ell+1, m) + mI_x(\ell, m+1)\},
\end{aligned}$$

$$I_x(\ell, m) = [\ell B(\ell, m)]^{-1} x^{\ell}(1-x)^m + I_x(\ell+1, m),$$
$$(\ell + m) I_x(\ell, m) = \ell I_x(\ell+1, m) + m I_x(\ell, m+1),$$

and

$$(\ell + m - \ell m) I_x(\ell, m) = \ell(1-x) I_x(\ell+1, m-1) + m I_x(\ell, m+1).$$

The function $I_x(\ell, m)$ has the following binomial expansion:

$$I_x(\ell, m+1-\ell) = \sum_{r=\ell}^{m} \binom{m}{r} x^r (1-x)^{m-r},$$

where ℓ is a positive integer. The tables of incomplete beta function have been given by Pearson (1934). There are various algorithms currently available to evaluate the incomplete beta function (see, e.g., Aroian, 1941; Abramowitz and Stegun, 1965, p. 944). In addition, a number of computing software provide built-in routines to evaluate the incomplete beta function (see, e.g., SAS Institute, 1990; Wolfram, 1996; Visual Numerics, 1997).

Bibliography

M. Abramowitz and I. A. Stegun (1965), *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Applied Mathematics Series 55, National Bureau of Standards, Washington, DC.

L. A. Aroian (1941), Continued fraction for the incomplete beta function, *Ann. Math. Statist.*, **12**, 218–223.

K. Pearson (1934), *Tables of the Incomplete Beta Function*, Cambridge University Press, Cambridge, UK.

SAS Institute (1990), *SAS Language: Reference, Version* 6, 1st ed., SAS Institute, Cary, NC.

Visual Numerics (1997), *IMSL: International Mathematical and Statistical Libraries* (*FORTRAN Subroutines for Evaluating Special Functions*): *Version* 3, Visual Numerics, Houston.

S. Wolfram (1996), *The Mathematica Book*, 3rd ed., Cambridge University Press, Cambridge, UK.

D INCOMPLETE INVERTED DIRICHLET FUNCTION

The incomplete inverted Dirichlet function or cumulative distribution function of a k-variate inverted Dirichlet random variable is defined by

$$D_{x_1,\ldots,x_k}(v_1, \ldots, v_k; v_{k+1}) = \frac{\Gamma(\sum_{i=1}^{k+1} v_i)}{\prod_{i=1}^{k+1} \Gamma(v_i)} \int_0^{x_1} \cdots \int_0^{x_k} u_1^{v_1 - 1} \cdots u_k^{v_k - 1}$$

$$\times \left(1+\sum_{i=1}^{k} u_i\right)^{-\sum_{i=1}^{k+1} v_i} du_k \cdots du_1.$$

The exact evaluation of this integral is in general difficult, particularly if k is large. Tiao and Guttman (1965) considered some useful approximations to the integral of the above type. Yassaee (1976, 1981) developed an algorithm and a computer program to evaluate such an integral for any (finite) dimension and parameter values, whether integer are real. The above integral can also be readily evaluated by the use of Mathematica (Wolfram, 1996) and Scientific Workplace (Hardy and Walker, 1995) software for any finite k and integer or real values of v_is. In particular, if $k = 2$ and v_1 is an integer, the integral can be expressed in terms of incomplete beta integrals as follows:

$$\begin{aligned} D_{x_1,x_2}(v_1, v_2; v_3) = {} & I_{x_2/(1+x_2)}(v_2, v_3) - \left(\frac{1}{1+x_1}\right)^{v_3} \\ & \times \sum_{i=0}^{v_1-1} \frac{\Gamma(v_3+i)}{\Gamma(v_3)\Gamma(i+1)} \left(\frac{x_1}{1+x_1}\right)^{i} \\ & \times I_{x_2/(1+x_1+x_2)}(v_2, v_3+i), \end{aligned}$$

where $I_x(.,.)$ denotes the incomplete beta function.

The following recurrence relations on inverted Dirichlet functions are quite useful in the numerical evaluation of these functions:

$$\begin{aligned} D_{x,y}(\ell, m; n) = {} & \frac{n}{(\ell+m+n)} D_{x,y}(\ell, m; n+1) \\ & - \frac{\ell}{(n-1)} D_{x,y}(\ell+1, m; n-1) \\ & - \frac{m}{(n-1)} D_{x,y}(\ell, m+1; n-1), \\ D_{x,y}(\ell+1, m; n) = {} & D_{x,y}(\ell, m; n) - [\ell B(\ell, n)]^{-1} \\ & \times \frac{x^{\ell}}{(1+x)^{\ell+n}} I_{y/(1+x+y)}(m, \ell+n), \end{aligned}$$

and

$$\begin{aligned} D_{x,y}(\ell, m+1; n) = {} & D_{x,y}(\ell, m; n) - [mB(m, n)]^{-1} \\ & \times \frac{y^{m}}{(1+y)^{m+n}} I_{x/(1+x+y)}(\ell, m+n), \end{aligned}$$

where

$$B(\ell, m) = \int_0^1 u^{\ell-1}(1-u)^{m-1} du.$$

Bibliography

D. W. Hardy and C. L. Walker (1995), *Doing Mathematics with Scientific Workplace*, Brooks–Cole, Pacific Grove, CA.

G. C. Tiao and I. Guttman (1965), The inverted Dirichlet distribution with applications, *J. Amer. Statist. Assoc.*, **60**, 793–805.

S. Wolfram (1996), *The Mathematica Book*, 3rd ed., Cambridge University Press, Cambridge, UK.

H. Yassaee (1976), Probability integral of inverted Dirichlet distribution and its applications, in J. Gordesch and P. Naeve, eds., *COMPSTAT: Proceedings of the* 2nd *Symposium on Computational Statistics*, Physica-Verlag, Wien, Germany, 64–71.

H. Yassaee (1981), On integrals of Dirichlet distributions and their applications, *Comm. Statist.* A *Theory Methods*, **10**, 897–906.

E INVERTED CHI-SQUARE DISTRIBUTION

Let $\chi^2[v]$ represent a chi-square random variable having v degrees of freedom with probability density function

$$f(\chi^2[v]) = \frac{1}{2^{v/2}\Gamma(v/2)} \exp\left\{-\frac{1}{2}\chi^2\right\} (\chi^2)^{v/2-1}, \quad \chi^2 > 0. \tag{E.1}$$

The inverted chi-square distribution having v degrees of freedom is obtained from (E.1) by making the transformation

$$\chi^{-2} = 1/\chi^2. \tag{E.2}$$

The probability density function of (E.2) is given by

$$f(\chi^{-2}[v]) = \frac{1}{2^{v/2}\Gamma(v/2)} \exp\left\{-\frac{1}{2\chi^{-2}}\right\} (\chi^{-2})^{-(v/2+1)}.$$

F THE SATTERTHWAITE PROCEDURE

Let MS_i and v_i $(i = 1, \ldots, p)$ be the mean squares and the corresponding degrees of freedom in an analysis of variance model such that

$$v_i \mathrm{MS}_i \sim \sigma_i^2 \chi^2[v_i], \tag{F.1}$$

where $\chi^2[v_i]$ represents a (central) chi-square variable with v_i degrees of freedom. Consider a linear combination of mean squares given by

$$\eta = \sum_{i=1}^{p} \ell_i \mathrm{MS}_i. \tag{F.2}$$

The so-called Satterthwaite procedure consists of approximating the distribution of the quantity

$$g = v\eta / \sum_{i=1}^{p} \ell_i \sigma_i^2 \tag{F.3}$$

by that of a chi-square distribution with v degrees of freedom, where v is obtained by equating the first two moments of the left- and right-hand expressions in (F.3). Since they already have the same means, only the variances have to be equated. Now, from (F.1), we have

$$\text{Var}(\text{MS}_i) = 2\sigma_i^4/v_i$$

and

$$\text{Var}(g) = \frac{2v^2 \sum_{i=1}^{p} (\ell_i^2 \sigma_i^4 / v_i)}{\left(\sum_{i=1}^{p} \ell_i \sigma_i^2\right)^2}. \tag{F.4}$$

Equating $\text{Var}(g)$ in (F.4) to $2v$ yields

$$v = \left(\sum_{i=1}^{p} \ell_i \sigma_i^2\right)^2 / \sum_{i=1}^{p} (\ell_i^2 \sigma_i^4 / v_i). \tag{F.5}$$

The above approximation for the distribution of a linear combination of mean squares was first studied by Smith (1936) and later by Satterthwaite (1941, 1946). Generally, the σ_i^2s are not known, so they are replaced by their unbiased estimates MS_is giving an estimate of v as

$$\hat{v} = \left(\sum_{i=1}^{p} \ell_i \text{MS}_i\right)^2 / \sum_{i=1}^{p} (\ell_i^2 \text{MS}_i^2 / v_i). \tag{F.6}$$

Since MS_i^2 is not an unbiased estimator of σ_i^4, (F.6) is a biased estimator of v. Noting that an unbiased estimator of σ_i^4 is $v_i \text{MS}_i^2/(v_i + 2)$, a corrected estimator of v is given by

$$\hat{v} = \left(\sum_{i=1}^{p} \ell_i \text{MS}_i\right)^2 / \sum_{i=1}^{p} (\ell_i^2 \text{MS}_i^2 / (v_i + 2)). \tag{F.7}$$

The Satterthwaite procedure is frequently employed for constructing confidence intervals for the mean and the variance components in a random and mixed effects analysis of variance. For example, if a variance component σ^2 is estimated by $\text{MS} = \sum_{i=1}^{p} \ell_i \text{MS}_i$, then an approximate $100(1-\alpha)\%$ confidence interval for σ^2 is given by

$$\frac{\text{MS}}{\chi^2[v, \alpha/2]} < \sigma^2 < \frac{\text{MS}}{\chi^2[v, 1-\alpha/2]},$$

where $\chi^2[v, \alpha/2]$ and $\chi^2[v, 1-\alpha/2]$ are the $100(\alpha/2)$th lower and upper percentiles of the chi-square distribution with v degrees of freedom, where v is determined by formulas (F.6) or (F.7).

Another application of the Satterthwaite procedure involves the construction of a pseudo F-test when an exact F-test cannot be found from the ratio of two mean squares. In such cases, one can form linear combinations of mean squares for the numerator, for the denominator, or for both the numerator and the denominator such that their expected values are equal under the null hypothesis. For example, let

$$\text{MS}' = \ell_r \text{MS}_r + \cdots + \ell_s \text{MS}_s$$

and

$$\text{MS}'' = \ell_u \text{MS}_u + \cdots + \ell_v \text{MS}_v,$$

where the mean squares are chosen such that $E(\text{MS}') = E(\text{MS}'')$ under the null hypothesis that a particular variance component is zero. Now, an approximate F-test of the null hypothesis can be obtained by the statistic

$$F = \frac{\text{MS}'}{\text{MS}''},$$

which has an approximate F-distribution with v' and v'' degrees of freedom determined by

$$v' = \frac{(\ell_r \text{MS}_r + \cdots + \ell_s \text{MS}_s)^2}{\ell_r^2 \text{MS}_r^2/v_r + \cdots + \ell_s^2 \text{MS}_s^2/v_s}$$

and

$$v'' = \frac{(\ell_u \text{MS}_u + \cdots + \ell_v \text{MS}_v)^2}{\ell_u^2 \text{MS}_u^2/v_u + \cdots + \ell_v^2 \text{MS}_v^2/v_v}.$$

In many situations, it may not be necessary to approximate both the numerator and the denominator mean squares for obtaining an F-test. However, when both the numerator and the denominator mean squares are constructed, it is always possible to find additive combinations of mean squares and thereby avoid subtracting mean squares, which may result in a poor approximation. For some further discussion of approximate F-tests, see Anderson (1960) and Eisen (1966).

In many applications of the Satterthwaite procedure, some of the mean squares may involve negative coefficients. Satterthwaite remarked that care should be exercised in applying the approximation when some of the coefficients may be negative. When negative coefficients are involved, one can rewrite the linear combination as $\text{MS} = \text{MS}_A - \text{MS}_B$, where MS_A contains all the mean squares with positive coefficients and MS_B with negative coefficients. Now, the degrees of freedom associated with the approximate chi-square distribution of MS are determined by

$$f = (\mathrm{MS}_A + \mathrm{MS}_B)^2/(\mathrm{MS}_A^2/f_A + \mathrm{MS}_B^2/f_B),$$

where f_A and f_B are the degree of freedom associated with the approximate chi-square distributions of MS_A and MS_B, respectively. Gaylor and Hopper (1969) showed that the Satterthwaite approximation for MS with f degrees of freedom is an adequate one when

$$\mathrm{MS}_A/\mathrm{MS}_B > F[f_B, f_A; 0.975] \times F[f_A, f_B; 0.5],$$

if $f_A \leq 100$ and $f_B \geq f_A/2$. The approximation is usually adequate for the differences of mean squares when the mean squares being subtracted are relatively small. Khuri (1995) has developed a measure to evaluate the adequacy of the Satterthwaite approximation in balanced mixed models.

Bibliography

R. L. Anderson (1960), Use of variance component analysis in the interpretation of biological experiments, Part 3, *Bull. Internat. Statist. Inst.*, **37**, 71–90.

E. J. Eisen (1966), The quasi-F test for an unnested fixed factor in an unbalanced hierarchal design with a mixed model, *Biometrics*, **22**, 937–942.

D. W. Gaylor and F. N. Hopper (1969), Estimating the degrees of freedom for linear combinations of mean squares by Satterthwaite's formula, *Technometrics*, **11**, 691–706.

A. I. Khuri (1995), A test to detect inadequacy of Satterthwaite's approximation in balanced mixed models, *Statistics*, **27**, 45–54.

F. E. Satterthwaite (1941), Synthesis of variance, *Psychometrika*, **6**, 309–316.

F. E. Satterthwaite (1946), An approximate distribution of estimates of variance components, *Biometrics Bull.*, **2**, 110–114.

H. F. Smith (1936), The problem of comparing the results of two experiments with unequal errors, *J. Council Sci. Indust. Res.* (*U.K.*), **9**, 211–212.

G MAXIMUM LIKELIHOOD ESTIMATION

Consider a random sample $X_1, X_2, \ldots, X_n$ from a population with probability distribution $f_X(x|\theta_1, \theta_2, \ldots, \theta_k)$, where $f_X(x|\theta_1, \theta_2, \ldots, \theta_k)$ may stand for the density or for the probability function. If $x_1, x_2, \ldots, x_n$ denote actual realizations of the random sample, then the joint density function of $X_1, X_2, \ldots, X_n$, say, $f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n|\theta_1, \theta_2, \ldots, \theta_k)$, is

$$f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n|\theta_1, \theta_2, \ldots, \theta_k)$$

$$= f_{X_1}(x_1|\theta_1, \theta_2, \ldots, \theta_k) \ldots f_{X_n}(x_n|\theta_1, \theta_2, \ldots, \theta_k)$$
$$= \prod_{i=1}^{n} f_{X_i}(x_i|\theta_1, \theta_2, \ldots, \theta_k). \tag{G.1}$$

The product (G.1) regarded as a function of the parameters $\theta_1, \theta_2, \ldots, \theta_k$ is called the likelihood function of the sample, or simply the likelihood function. We shall use the symbol $L(\theta_1, \theta_2, \ldots, \theta_k|x_1, x_2, \ldots, x_n)$ for a likelihood function; i.e.,

$$L(\theta_1, \theta_2, \ldots, \theta_k|x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i|\theta_1, \theta_2, \ldots, \theta_k). \tag{G.2}$$

Letting $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)'$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)'$, and $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)'$, we may simply write

$$L(\boldsymbol{\theta}|\boldsymbol{x}) = \prod_{i=1}^{n} f_{X_i}(x_i|\boldsymbol{\theta}) = f_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}). \tag{G.3}$$

Note that $f_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta})$ and $L(\boldsymbol{\theta}|\boldsymbol{x})$ are mathematically equivalent. However, in the former $\boldsymbol{X}$ is a vector of random variables and $\boldsymbol{\theta}$ is assumed to be known, while in the latter $\boldsymbol{x}$ represents a known vector of data and $\boldsymbol{\theta}$ is taken to be unknown.

The maximum likelihood estimator (MLE) of the parameter vector $\boldsymbol{\theta}$ is the value $\hat{\boldsymbol{\theta}}(\boldsymbol{x})$ that maximizes the likelihood function $L(\boldsymbol{\theta}|\boldsymbol{x})$. Many likelihood functions satisfy certain regularity conditions, so that the ML estimator is the solution of the equation

$$\frac{dL(\boldsymbol{\theta}|\boldsymbol{x})}{d\boldsymbol{\theta}} = \mathbf{0}. \tag{G.4}$$

Equation (G.4) stands for the k equations

$$\frac{\partial L(\theta_1, \theta_2, \ldots, \theta_k|\boldsymbol{x})}{\partial \theta_i} = 0, \quad i = 1, 2, \ldots, k.$$

It should be observed that the functions $L(\boldsymbol{\theta}|\boldsymbol{x})$ and $\ell n L(\boldsymbol{\theta}|\boldsymbol{x})$ are maximized at the same value of $\boldsymbol{\theta}$, so that it may probably be easier to work with the natural logarithm of the likelihood function.

In some problems, the values of $\boldsymbol{\theta}$ are constrained to be in a restricted parameter space. For example, in a random or mixed effects linear model, variance components are nonnegative. In such cases, the MLE is the solution of the equation

$$\frac{dL(\boldsymbol{\theta}|\boldsymbol{x})}{d\boldsymbol{\theta}} = \mathbf{0},$$

subject to $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, where $\boldsymbol{\Theta}$ is a subset of the Euclidean space. In general the problem of maximization of a likelihood function with a constrained parameter space is quite difficult and entails the use of numerical algorithms involving iterative procedures.

A useful property of the MLE is that for a large sample it is asymptotically distributed as multivariate normal with mean vector $\boldsymbol{\theta}$ and the variance-covariance matrix $\boldsymbol{I}^{-1}(\boldsymbol{\theta})$, where $\boldsymbol{I}(\boldsymbol{\theta})$ is the information matrix; i.e.,

$$\boldsymbol{I}(\boldsymbol{\theta}) = E\left(\frac{\partial \ell n L(\boldsymbol{\theta}|\boldsymbol{x})}{\partial \boldsymbol{\theta}} \cdot \frac{\partial \ell n L(\boldsymbol{\theta}|\boldsymbol{x})}{\partial \boldsymbol{\theta}'}\right) = E\left\{\frac{\partial \ell n L(\boldsymbol{\theta}|\boldsymbol{x})}{\partial \theta_i} \cdot \frac{\partial \ell n L(\boldsymbol{\theta}|\boldsymbol{x})}{\partial \theta_j}\right\}.$$

Alternatively, the information matrix may be evaluated as

$$\boldsymbol{I}(\boldsymbol{\theta}) = -E\left(\frac{\partial^2 \ell n L(\boldsymbol{\theta}|\boldsymbol{x})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right) = -E\left\{\frac{\partial^2 \ell n L(\boldsymbol{\theta}|\boldsymbol{x})}{\partial \theta_i \partial \theta_j}\right\}.$$

The Cramér–Rao lower bound on the variance of an unbiased estimator of θ_i is given by the reciprocal of $-E[\partial^2 \ell n L(\boldsymbol{\theta}|\boldsymbol{x})/\partial \theta_i^2]$.

H SOME USEFUL LEMMAS ON THE INVARIANCE PROPERTY OF THE ML ESTIMATORS

In this appendix, we present two lemmas on the invariance property of the maximum likelihood (ML) estimators, which are useful in the derivation of the ML estimators of variance components. Consider the likelihood function $L(\theta|x_1, x_2, \ldots, x_n)$, where $(\theta \in \Theta)$, and consider first a mapping of the parameter space Θ onto some space Θ'. Suppose that we are interested in estimating $g(\theta)$ in the new parameter space Θ'.

Lemma H.1. *Let $\hat{\theta}$ be the* ML *estimator of the parameter θ. Let $g(\theta)$ be a single-valued function of θ. Then the* ML *estimate of $g(\theta)$ is obtained simply by replacing θ in $g(\theta)$ by the* ML *estimate of θ. That is $\hat{g}(\theta) = g(\hat{\theta})$.*

Proof. If $g(\theta)$ is a single-valued function of θ $(\theta \in \Theta')$, there exists the inverse mapping h of Θ' onto Θ such that $\theta' = g(\theta)$ whenever $\theta = h(\theta')$. If the likelihood $L(\theta|x_1, x_2, \ldots, x_n)$ is maximized at the point $\hat{\theta} = \hat{\theta}(x_1, x_2, \ldots, x_n)$, then the function $L(h(\theta'); x_1, \ldots, x_n)$ is maximized when $h(\theta') = \hat{\theta}(x_1, x_2, \ldots, x_n)$; hence when $\theta' = g[\hat{\theta}(x_1, x_2, \ldots, x_n)]$. □

The invariance property of the ML estimators is also valid in the multidimensional case.

Lemma H.2. *Let $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k$ be the* ML *estimates of $\theta_1, \theta_2, \ldots, \theta_k$. Let $g_1 = g_1(\theta_1, \theta_2, \ldots, \theta_k), \ldots, g_k = g_k(\theta_1, \theta_2, \ldots, \theta_k)$ be a set of transformations that are one-to-one. Then the* ML *estimates of $g_1, \ldots, g_k$ are $\hat{g}_1 = g_1(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k), \ldots, \hat{g}_k = g_k(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k)$.*

Proof. Let $(\theta_1, \theta_2, \ldots, \theta_k) \in \boldsymbol{\Theta}$, then it follows that the vector $\boldsymbol{G}(\theta_1, \theta_2, \ldots, \theta_k) = (g_1(\theta_1, \theta_2, \ldots, \theta_k), g_2(\theta_1, \theta_2, \ldots, \theta_k), \ldots, g_k(\theta_1, \theta_2, \ldots, \theta_k))$, defines a one-to-one mapping of k-dimensional parameter space $\boldsymbol{\Theta}$ into a subset $\boldsymbol{\Theta}'$ of k-dimensional space. By the argument of Lemma H.1, if $(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k)$ is the ML estimator of $(\theta_1, \theta_2, \ldots, \theta_k)$, then $(g_1(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k), g_2(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k), \ldots, g_k(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}))$ is the ML estimate of $(g_1(\theta_1, \theta_2, \ldots, \theta_k), g_2(\theta_1, \theta_2, \ldots, \theta_k), \ldots, g_k(\theta_1, \theta_2, \ldots, \theta_k))$. It follows therefore that $g_i(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k)$ is the ML estimator of $g_i(\theta_1, \theta_2, \ldots, \theta_k)(i = 1, 2, \ldots, k)$. □

I COMPLETE SUFFICIENT STATISTICS AND THE RAO–BLACKWELL AND LEHMANN–SHEFFÉ THEOREMS

Let $T_i(\boldsymbol{X})$, $i = 1, 2, \ldots, p$, be a set of statistics such that the conditional distribution of the random vector $\boldsymbol{X}$ given $T_i(\boldsymbol{X}) = t_i$ does not depend on the parameter vector $\boldsymbol{\theta}$. Then $T_i(\boldsymbol{X})$, $i = 1, 2 \ldots, p$, are said to be jointly sufficient for $\boldsymbol{\theta}$.

A set of sufficient statistics $T_i(\boldsymbol{X})$, $i = 1, 2, \ldots, p$, is called minimal if $T_i(\boldsymbol{X})$ are functions of any other sufficient statistics. The minimal set of sufficient statistics $\boldsymbol{T} = (T_1, T_2, \ldots, T_p)$ is said to be complete if there does not exist a function of $\boldsymbol{T}$ with expected value equal to zero.

If the ML estimator of $\boldsymbol{\theta}$ exists, then it is a function of the minimal sufficient set of statistics. If the ML estimator is not unique, then there exists an ML estimator that is a function of the minimal sufficient set of statistics.

Suppose $g(\boldsymbol{X})$ is an unbiased estimator of a scalar parametric function $h(\boldsymbol{\theta})$ and $T_i(\boldsymbol{X})$, $i = 1, 2, \ldots, p$, are jointly sufficient for $\boldsymbol{\theta}$; then there exists an estimator $u(\boldsymbol{T})$, depending on the data only through the sufficient statistics, such that $E[u(\boldsymbol{T})] = h(\boldsymbol{\theta})$ and $\mathrm{Var}[u(\boldsymbol{T})] \leq \mathrm{Var}[g(\boldsymbol{X})]$. The result is commonly known as the Rao–Blackwell theorem.

If $T_i(\boldsymbol{X})$, $i = 1, 2, \ldots, p$, are complete sufficient statistics for $\boldsymbol{\theta}$, then it follows from the Rao–Blackwell theorem that $u(\boldsymbol{T})$ is unique and therefore the minimum variance unbiased estimator of $h(\boldsymbol{\theta})$. The result is called the Lehmann–Sheffé theorem.

J POINT ESTIMATORS AND THE MSE CRITERION

Over a period of time, statisticians have attempted to obtain in some manner a single quantity determined as a function of sample data, which in some sense may be a representative value of a parameter of interest. In classical statistical literature, this is referred to as the problem of finding the most "probable" value, and sometimes that of finding an "average" or "mean" value. In modern sampling theory, a single quantity calculated from a sample data: $(x_1, x_2, \ldots, x_n)$ is known as a point estimator. The arguments proceed in the following manner.

Suppose an investigator is interested in a particular parameter θ. Then any function of the sample data, say $t(x_1, \ldots, x_n)$, which may provide some information on the value of θ may be considered an estimator of θ. In any given problem, generally speaking, a very large number of such estimators can be determined. For example, the variance of a population might be estimated by the sample variance, the sample range, the sample mean absolute deviation, and so forth. Therefore, to judge the merit of various estimators, a criterion of goodness is needed. Using such a criterion various estimators can be compared and the "best" one can be selected. It is argued that the "goodness" of an estimator should be measured by the average closeness of its values over all possible samples to the true value of θ. The criterion of the average closeness often suggested is the mean squared error (MSE). Thus, given a class of possible estimators, say, $T_1(x_1, \ldots, x_n), \ldots, T_i(x_1, \ldots, x_n), \ldots, T_k(x_1, \ldots, x_n)$, the goodness of a particular estimator is measured by the magnitude of the quantity

$$\text{MSE}(T_i) = E\{T_i(x_1, \ldots, x_n) - \theta\}^2,$$

where the expectation is taken over the sampling distribution of T_i. The estimator T_i would be considered "best" in the class of possible estimators, if its MSE is minimum for all values of θ compared to any other estimator. Such an estimator is called the minimum MSE estimator. An estimator with a small MSE is likely to have a high probability of concentration around the true value of θ.

The argument for the MSE criterion seems appealing, but the criterion itself is an arbitrary one and is easily shown to be rather unreliable. For example, consider the problem of estimating the reciprocal of the mean of the normal distribution with mean θ and unit variance. From a random sample $(x_1, \ldots, x_n)$, the maximum likelihood estimate of $1/\theta$ is $1/\bar{x}$, which is sufficient for $1/\theta$. However, it can be readily shown that the MSE of $1/\bar{x}$ is infinite. Moreover, it has been pointed out that MSE has a deficiency in that it cannot distinguish between cases of overestimation and underestimation (see, e.g., Pukelsheim, 1979). Thus it is recommended that both the biases of the estimators together with MSEs be considered. MSE efficient estimators of the variance components are discussed in the work of Lee and Kapadia(1992).

Bibliography

K. R. Lee and C. H. Kapadia (1992), Mean squared error efficient estimators of the variance components, *Metrika*, **39**, 21–26.

F. Pukelsheim (1979), Classes of linear models, in L. D. Van Vleck and S. R. Searle, eds., *Variance Components and Animal Breeding: Proceedings of a Conference in Honor of C. R. Henderson*, Cornell University, Ithaca, NY, 69–83.

K LIKELIHOOD RATIO TEST

Let $L(\boldsymbol{\theta}|\boldsymbol{x})$ denote the likelihood function defined in (G.3) of the parameter vector $\boldsymbol{\theta}$, where $\boldsymbol{\theta}$ is restricted by $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. Suppose we wish to test the hypothesis

$$H_0 : \boldsymbol{\theta} \in \boldsymbol{\Theta}_0 \quad \text{vs.} \quad H_1 : \boldsymbol{\theta} \notin \boldsymbol{\Theta}_0,$$

where $\boldsymbol{\Theta}_0$ is a subset of $\boldsymbol{\Theta}$. Consider the statistic defined by the ratio

$$\lambda(\boldsymbol{x}) = \frac{L(\hat{\boldsymbol{\Theta}}_0|\boldsymbol{x})}{L(\hat{\boldsymbol{\Theta}}|\boldsymbol{x})},$$

where $L(\hat{\boldsymbol{\Theta}}_0|x) = \max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0} L(\boldsymbol{\Theta}|\boldsymbol{x})$ and $L(\hat{\boldsymbol{\Theta}}|\boldsymbol{x}) = \max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} L(\boldsymbol{\Theta}|\boldsymbol{x})$. It follows that $0 \le \lambda(\boldsymbol{x}) \le 1$ since $L(\boldsymbol{\Theta}_0|\boldsymbol{x})$ and $L(\boldsymbol{\Theta}|\boldsymbol{x})$ are density functions and $\boldsymbol{\Theta}_0$ is a subset of $\boldsymbol{\Theta}$.

The function $\lambda(\boldsymbol{x})$ defines a random variable and a test based on the likelihood-ratio criterion such that it rejects H_0 for small values of $\lambda(\boldsymbol{x})$ is called the likelihood-ratio test. If we reject H_0 in favor of H_1 when $\lambda(\boldsymbol{x}) \le \lambda_0$, then the significance level and the power of the test are given by

$$\alpha = P\{\lambda(\boldsymbol{x}) \le \lambda_0 | H_0\} \quad \text{and} \quad 1 - \beta = P\{\lambda(\boldsymbol{x}) \le \lambda_0 | H_1\}.$$

The likelihood-ratio test principle is an intuitive one and does not always lead to the same test as given by the Neyman–Pearson Theorem. Thus the likelihood-ratio test is not necessarily a uniformly most powerful test, but it has been shown in the literature that such a test often has some desirable properties.

In general, it is difficult to determine the exact distribution of $\lambda(\boldsymbol{x})$ or an equivalent test statistic. Using the results in advanced statistical theory, however, it can be shown that under a number of regularity conditions, the asymptotic null distribution of -2 times the likelihood ratio criterion $(-2\ell n\lambda(\boldsymbol{x}))$ has an approximate chi-square distribution with ν degrees of freedom, where $\nu =$ the dimension of $\boldsymbol{\Theta}-$ dimension of $\boldsymbol{\Theta}_0$; i.e., ν is the number of independent constraints in the hypothesis (see, e.g., Cox and Hinkly, 1974; Lehmann 1986, p. 486).

Bibliography

D. R. Cox and D. V. Hinkly (1974), *Theoretical Statistics*, Chapman and Hall, London; softbound ed., 1986.

E. L. Lehmann (1986), *Testing Statistical Hypothesis*, 2nd ed., Wiley, New York; reprint, 1997, Springer-Verlag, New York.

L DEFINITION OF INTERACTION

In a factorial experiment, the interaction between two factors, say A and B, measures the failure of the effects of different levels of factor A to be the same for each level of the factor B, or equivalently the failure of the effects of different levels of factor B to be the same for each level of factor A. For example, in an agricultural experiment involving two factors, variety and fertilizer, some fertilizers may increase the yield of some varieties, but may decrease it for others.

To define the concept of interaction mathematically, let $f(x, y)$ be a function of two variables x and y. Then $f(x, y)$ is defined to be a function with no interaction if and only if there exist two functions, say, $g(x)$ and $h(y)$, such that

$$f(x, y) = g(x) + h(y).$$

For example, the functions $x^2 + 2xy^2$, $x^2 + \log \sqrt{y} + xy^3$, e^{xy} and e^{x+y} have interactions; but the functions $x + y$, $\log xy$, and $x^2 + 2x + y^2 + 2y$ have no interactions. To illustrate the above definition for an analysis of variance model, consider the two-way crossed classification model given by

$$y_{ij} = \mu_{ij} + e_{ij},$$

where μ_{ij} is the "true" total effect of the combination of the ith level of factor A and the jth level of factor B. If this "true" total effect is simply the sum of the effects of the ith level of A, which is α_i, and the jth of B, which is β_j, plus some constant μ, we say that there is no interaction between A and B.

M SOME BASIC RESULTS ON MATRIX ALGEBRA

In this appendix, we shall review some basic definitions and results on matrix algebra. For further details and proofs the reader is referred to any one of several books on matrix algebra given in Chapter 9.

SOME DEFINITIONS

An *$m \times n$ matrix* $\boldsymbol{A}$ is a rectangular array of order $m \times n$ with elements a_{ij}, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$, written as

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}.$$

The dimensions of a matrix are important and a matrix with m rows and n columns is referred to as $m \times n$ matrix.

In contrast to a matrix, a real number is called a *scalar*, which, of course, can be considered a 1×1 matrix.

A *vector* is a matrix with a single row or column. A p-component *column vector* with elements $a_1, a_2, \ldots, a_p$ is written as

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}$$

in lowercase boldface type. A p-component *row vector* consists of a single row with elements $a_1, a_2, \ldots, a_p$ and is written as

$$\boldsymbol{a}' = [a_1, a_2, \ldots, a_p].$$

The *transpose* of an $m \times n$ matrix is defined to be the $n \times m$ matrix $\boldsymbol{A}'$ which has in the jth row and the ith column the element that is in the ith row and jth column of $\boldsymbol{A}$. The matrix $\boldsymbol{A}'$ is formed by interchanging the roles of rows and columns of $\boldsymbol{A}$ and is written as

$$\boldsymbol{A}' = \begin{bmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & \ldots & a_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{bmatrix}.$$

Note that $(\boldsymbol{A}')' = \boldsymbol{A}$.

A matrix is called a *square matrix* if the number of rows and columns are the same.

A square matrix $\boldsymbol{A}$ is called *symmetric* if $\boldsymbol{A}' = \boldsymbol{A}$, that is $a_{ij} = a_{ji}$ for all pairs i and j.

A $p \times p$ square matrix is said to be *orthogonal* if and only if $\boldsymbol{A}\boldsymbol{A}' = \boldsymbol{I}$.

A $p \times p$ square matrix is said to be *idempotent* if $\boldsymbol{A}\boldsymbol{A} = \boldsymbol{A}$.

A *diagonal matrix* with elements $d_1, d_2, \ldots, d_p$ is a $p \times p$ square matrix with d_is in its main diagonal positions and zeros in other locations and is written as

$$\boldsymbol{D} = \begin{bmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & d_p \end{bmatrix}.$$

The *identity matrix* is a $p \times p$ square matrix with 1 in each diagonal position and 0 elsewhere and is written as

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}.$$

The identity matrix satisfies the relation:

$$\boldsymbol{I}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{I} = \boldsymbol{A}.$$

A *triangular matrix* is a $p \times p$ square matrix that has zeros or nonzero elements in its upper diagonal locations and zeros in its lower diagonal locations and is written as

$$\boldsymbol{T} = \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1p} \\ 0 & t_{22} & \dots & t_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & t_{pp} \end{bmatrix}.$$

A *null matrix* is an $m \times n$ rectangular matrix that has zeros in all the locations and is written as

$$\boldsymbol{0} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

A *unity matrix* is an $m \times n$ rectangular matrix that has 1s in all the positions and is written as

$$\boldsymbol{U} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix}.$$

A *unity vector* is a p-component column vector that has 1 in every position and is written as

$$\boldsymbol{J} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

SUMS AND PRODUCTS OF MATRICES

Two $m \times n$ matrices $\boldsymbol{A} = [a_{ij}]$ and $\boldsymbol{B} = [b_{ij}]$ are said to be equal if and only if $a_{ij} = b_{ij}$ for all pairs of i and j, $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$.

The sum of two $m \times n$ matrices $\boldsymbol{A} = [a_{ij}]$ and $\boldsymbol{B} = [b_{ij}]$ is the $m \times n$ matrix $\boldsymbol{A} + \boldsymbol{B} = [a_{ij} + b_{ij}]$, obtained by adding corresponding elements and is written as

$$\boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}.$$

The difference of two matrices of the same dimensions is defined similarly by forming the matrix of differences of the individual elements. It can be verified that these operations have the algebraic properties. Thus the addition or subtraction is commutative and associative:

$$\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A},$$
$$\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C},$$
$$\boldsymbol{A} - (\boldsymbol{B} - \boldsymbol{C}) = \boldsymbol{A} - \boldsymbol{B} + \boldsymbol{C},$$

and the transpose of a sum is the sum of the transpose:

$$(\boldsymbol{A} + \boldsymbol{B})' = \boldsymbol{A}' + \boldsymbol{B}'.$$

The product of an $m \times n$ matrix $\boldsymbol{A} = [a_{ij}]$ by a scalar (real number) c is the $m \times n$ matrix $c\boldsymbol{A} = [ca_{ij}]$, obtained by multiplying each element by c and is written as

$$c\boldsymbol{A} = \begin{bmatrix} ca_{11} & ca_{12} & \dots & ca_{1n} \\ ca_{21} & ca_{22} & \dots & ca_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ ca_{m1} & ca_{m2} & \dots & ca_{mn} \end{bmatrix}.$$

This multiplication is called scalar multiplication and has the same properties as scalar multiplication of a vector. Note that $(c\boldsymbol{A})' = c\boldsymbol{A}'$.

The *matrix product* of $\boldsymbol{A} = [a_{ij}]$ of dimension $m \times n$ and $\boldsymbol{B} = [b_{jk}]$ of dimension $n \times r$ is defined to be a matrix $\boldsymbol{C} = [c_{ik}]$ of dimension $m \times r$, where $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$, and is written as $\boldsymbol{AB} = \boldsymbol{C}$.

For the product $\boldsymbol{AB}$ to be defined it is necessary that the number of columns of $\boldsymbol{A}$ is equal to the number of rows of $\boldsymbol{B}$. The associative and distributive laws hold for matrix multiplication:

$$\boldsymbol{A}(\boldsymbol{BC}) = \boldsymbol{AB}(\boldsymbol{C})$$
$$\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}.$$

However, the commutative law does not hold, and in general it is not true that $\boldsymbol{AB} = \boldsymbol{BA}$. Further, the transposition of a matrix product has the following property:

$$(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}'.$$

More generally, if $\boldsymbol{A}_1, \boldsymbol{A}_2, \dots, \boldsymbol{A}_k$ are matrices with conformable dimensions, then

$$(\boldsymbol{A}_1 \boldsymbol{A}_2 \dots \boldsymbol{A}_k)' = \boldsymbol{A}_k' \dots \boldsymbol{A}_2' \boldsymbol{A}_1'.$$

The *direct sum* of matrices $\boldsymbol{A}_1, \boldsymbol{A}_2, \dots, \boldsymbol{A}_k$ is defined as the matrix

$$\sum_{i=1}^{k}{}^{+} \boldsymbol{A}_i = \begin{bmatrix} \boldsymbol{A}_1 & 0 & \dots & 0 \\ 0 & \boldsymbol{A}_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \boldsymbol{A}_k \end{bmatrix},$$

where these matrices can be of any order.

The *direct or Kronecker product* of an $m \times n$ matrix $\boldsymbol{A} = [a_{ij}]$ and a matrix $\boldsymbol{B}$ is the matrix

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \dots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \dots & a_{2n}\boldsymbol{B} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & \dots & a_{mn}\boldsymbol{B} \end{bmatrix}.$$

More generally, if $\boldsymbol{A}_1, \boldsymbol{A}_2, \dots, \boldsymbol{A}_k$ are matrices of any order, then

$$\boldsymbol{A}_1 \otimes \boldsymbol{A}_2 \otimes \boldsymbol{A}_3 = \boldsymbol{A}_1 \otimes (\boldsymbol{A}_2 \otimes \boldsymbol{A}_3)$$

and

$$\prod_{i=1}^{k} \otimes \boldsymbol{A}_i = \boldsymbol{A}_1 \otimes \boldsymbol{A}_2 \cdots \otimes \boldsymbol{A}_k.$$

Direct products have many properties. For example, assuming conformability for matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$, we have

$$\begin{aligned} (\boldsymbol{A} \otimes \boldsymbol{B})' &= \boldsymbol{A}' \otimes \boldsymbol{B}', \\ (\boldsymbol{A} \otimes \boldsymbol{B})^{-1} &= \boldsymbol{A}^{-1} \otimes \boldsymbol{B}^{-1}, \\ (\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) &= \boldsymbol{AC} \otimes \boldsymbol{BD}, \\ \operatorname{rank}(\boldsymbol{A} \otimes \boldsymbol{B}) &= \operatorname{rank}(\boldsymbol{A}) \operatorname{rank}(\boldsymbol{B}), \\ \operatorname{tr}(\boldsymbol{A} \otimes \boldsymbol{B}) &= \operatorname{tr}(\boldsymbol{A}) \operatorname{tr}(\boldsymbol{B}), \\ |\boldsymbol{A}_{a \times a} \otimes \boldsymbol{B}_{b \times b}| &= |\boldsymbol{A}|^b |\boldsymbol{B}|^a. \end{aligned}$$

THE RANK OF A MATRIX

The rank of a matrix $\boldsymbol{A}$ is the number of linearly independent rows and columns of $\boldsymbol{A}$. The rank has the following properties:

(i) $\operatorname{rank}(\boldsymbol{A}') = \operatorname{rank}(\boldsymbol{A})$,

(ii) $\operatorname{rank}(\boldsymbol{AA}') = \operatorname{rank}(\boldsymbol{A}'\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A})$,

(iii) $\operatorname{rank}(\boldsymbol{BAC}) = \operatorname{rank}(\boldsymbol{A})$, where $\boldsymbol{B}$ and $\boldsymbol{C}$ are nonsingular matrices with conformable dimensions,

(iv) $\operatorname{rank}(c\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A})$, where c is a nonzero scalar,

(v) $\operatorname{rank}(\boldsymbol{AB}) = \min(\operatorname{rank}(\boldsymbol{A}), \operatorname{rank}(\boldsymbol{B}))$,

(vi) $\operatorname{rank}(\boldsymbol{A})$ is unchanged by interchanging any two rows or columns of $\boldsymbol{A}$.

THE DETERMINANT OF A MATRIX

The determinant of a $p \times p$ matrix written as $|\boldsymbol{A}|$ is defined as

$$|\boldsymbol{A}| = \sum_{s} (-1)^{u} a_{1j_1} a_{2j_2} \ldots a_{pj_p},$$

where the summation is taken over the set s of all $p!$ permutations $j_1, j_2, \ldots, j_p$ of the set of integers $(1, 2, \ldots, p)$ and u is the number of inversions required to change $(1, 2, \ldots, p)$ into $j_1, j_2, \ldots, j_p$. It should be noted that the entries under the sum consist of all products of one element from each row and column and multiplied by -1 if u is odd. The number of inversions in a particular permutation is the total number of times in which an element is followed by numbers which would ordinarily precede it in natural order $1, 2, \ldots, p$.

The *minor* of an element a_{ij} of a matrix $\boldsymbol{A}$ is the determinant of the matrix obtained by deleting the ith row and the jth column of $\boldsymbol{A}$.

The *cofactor* of a_{ij} is the minor multiplied by $(-1)^{i+j}$ and is written as $\boldsymbol{A}_{ij}$. The determinant of the matrix $\boldsymbol{A}$ can be calculated more easily in terms of the cofactor by the following result:

$$\begin{aligned} |\boldsymbol{A}| &= \sum_{j=1}^{p} a_{ij} \boldsymbol{A}_{ij}, i = 1, 2, \ldots, p \\ &= \sum_{i=1}^{p} a_{ij} \boldsymbol{A}_{ij}, j = 1, 2, \ldots, p. \end{aligned}$$

The determinant of a diagonal matrix $\boldsymbol{D}$ with elements $d_1, d_2, \ldots, d_p$ is calculated by the formula

$$|\boldsymbol{D}| = d_1 d_2 \ldots d_p.$$

For any two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ with conformable dimensions,

$$|\boldsymbol{I} + \boldsymbol{AB}| = |\boldsymbol{I} + \boldsymbol{BA}|.$$

A matrix $p \times p$ is called *singular* if $|\boldsymbol{A}| = 0$; it is called *nonsingular* if $|\boldsymbol{A}| \neq 0$.

THE INVERSE OF A MATRIX

The *inverse* of a $p \times p$ nonsingular matrix $\boldsymbol{A}$ is the unique matrix $\boldsymbol{A}^{-1}$ such that

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}.$$

Note that $\boldsymbol{A}^{-1}$ does not exist if $\boldsymbol{A}$ is a singular matrix. The inverse of a matrix $\boldsymbol{A}$ can be expressed in terms of its cofactors as follows:

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \frac{\boldsymbol{A}_{11}}{|\boldsymbol{A}|}, & \frac{\boldsymbol{A}_{21}}{|\boldsymbol{A}|}, & \ldots, & \frac{\boldsymbol{A}_{p1}}{|\boldsymbol{A}|} \\ \frac{\boldsymbol{A}_{12}}{|\boldsymbol{A}|}, & \frac{\boldsymbol{A}_{22}}{|\boldsymbol{A}|}, & \ldots, & \frac{\boldsymbol{A}_{p2}}{|\boldsymbol{A}|} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\boldsymbol{A}_{1p}}{|\boldsymbol{A}|}, & \frac{\boldsymbol{A}_{2p}}{|\boldsymbol{A}|}, & \ldots, & \frac{\boldsymbol{A}_{pp}}{|\boldsymbol{A}|} \end{bmatrix}.$$

The inverse of a matrix has the following properties:

(i) If $\boldsymbol{A} = \boldsymbol{A}'$, $(\boldsymbol{A}^{-1})' = \boldsymbol{A}^{-1}$.

(ii) $(\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})'$.

(iii) If c is a nonzero scalar, $(c\boldsymbol{A})^{-1} = (1/c)\boldsymbol{A}^{-1}$.

(iv) If $\boldsymbol{D}$ is a diagonal matrix with elements $d_1, d_2, \ldots, d_p$, then

$$\boldsymbol{D}^{-1} = \begin{bmatrix} d_1^{-1} & 0 & \ldots & 0 \\ 0 & d_2^{-1} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & d_p^{-1} \end{bmatrix}.$$

(v) If $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_p$ are matrices with conformable dimensions, then

$$(\boldsymbol{A}_1\boldsymbol{A}_2 \ldots \boldsymbol{A}_p)^{-1} = \boldsymbol{A}_p^{-1} \ldots \boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1}.$$

(vi) For any two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ with conformable dimensions

$$(\boldsymbol{I} + \boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{I} + \boldsymbol{B}\boldsymbol{A})^{-1}\boldsymbol{B}.$$

PARTITIONED MATRICES

The partitioned matrices of a matrix $\boldsymbol{A}$ are submatrices written as an array:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \ldots & \boldsymbol{A}_{1c} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \ldots & \boldsymbol{A}_{2c} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{A}_{r1} & \boldsymbol{A}_{r2} & \ldots & \boldsymbol{A}_{rc} \end{bmatrix},$$

where $\boldsymbol{A}_{ij}$ contains m_i rows and n_j columns such that all submatrices in a given row must have the same number of rows and each column contains matrices with the same number of columns.

The sum of two partitioned matrices $\boldsymbol{A} = [\boldsymbol{A}_{ij}]$ and $\boldsymbol{B} = [\boldsymbol{B}_{ij}]$ with similar dimensions is the matrix

$$\boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} \boldsymbol{A}_{11} + \boldsymbol{B}_{11} & \boldsymbol{A}_{12} + \boldsymbol{B}_{12} & \dots & \boldsymbol{A}_{1c} + \boldsymbol{B}_{1c} \\ \boldsymbol{A}_{21} + \boldsymbol{B}_{21} & \boldsymbol{A}_{22} + \boldsymbol{B}_{22} & \dots & \boldsymbol{A}_{2c} + \boldsymbol{B}_{2c} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{A}_{r1} + \boldsymbol{B}_{r1} & \boldsymbol{A}_{r2} + \boldsymbol{B}_{r2} & \dots & \boldsymbol{A}_{rc} + \boldsymbol{B}_{rc} \end{bmatrix}.$$

The product of the partitioned matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ with conformable dimensions is the matrix

$$\boldsymbol{AB} = \begin{bmatrix} \sum_{j=1}^{c} \boldsymbol{A}_{1j}\boldsymbol{B}_{j1} & \sum_{j=1}^{c} \boldsymbol{A}_{1j}\boldsymbol{B}_{j2} & \dots & \sum_{j=1}^{c} \boldsymbol{A}_{1j}\boldsymbol{B}_{jp} \\ \sum_{j=1}^{c} \boldsymbol{A}_{2j}\boldsymbol{B}_{j1} & \sum_{j=1}^{c} \boldsymbol{A}_{2j}\boldsymbol{B}_{j2} & \dots & \sum_{j=1}^{c} \boldsymbol{A}_{2j}\boldsymbol{B}_{jp} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{j=1}^{c} \boldsymbol{A}_{rj}\boldsymbol{B}_{j1} & \sum_{j=1}^{c} \boldsymbol{A}_{rj}\boldsymbol{B}_{j2} & \dots & \sum_{j=1}^{c} \boldsymbol{A}_{rj}\boldsymbol{B}_{jp} \end{bmatrix}.$$

Note that if the submatrices of $\boldsymbol{A}$ have respective column numbers $n_1, n_2, \dots, n_c$, then $\boldsymbol{B}$ must have the respective row dimensions as $n_1, n_2, \dots, n_c$.

If $\boldsymbol{A}$ is a partitioned matrix

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix},$$

then

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \boldsymbol{A}^{11} & \boldsymbol{A}^{12} \\ \boldsymbol{A}^{21} & \boldsymbol{A}^{22} \end{bmatrix},$$

where

$$\begin{aligned} \boldsymbol{A}^{11} &= (\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21})^{-1}, \\ \boldsymbol{A}^{12} &= -(\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21})^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}, \\ \boldsymbol{A}^{21} &= -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}(\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21})^{-1}, \\ \boldsymbol{A}^{22} &= \boldsymbol{A}_{22}^{-1} + \boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}(\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21})^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}. \end{aligned}$$

DIFFERENTIATION OF MATRICES AND VECTORS

Let $f(\boldsymbol{x})$ be a continuous function of the elements of the vector $\boldsymbol{x}' = [x_1, x_2, \dots, x_p]$. Then $\partial f(\boldsymbol{x})/\partial \boldsymbol{x}$ is defined as

$$\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{bmatrix} \dfrac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \dfrac{\partial f(\boldsymbol{x})}{\partial x_2} \\ \vdots \\ \dfrac{\partial f(\boldsymbol{x})}{\partial x_p} \end{bmatrix}.$$

Some special functions and their derivatives are

(i) If $f(\boldsymbol{x}) = c$, where c is constant,

$$\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

(ii) If $f(\boldsymbol{x}) = \boldsymbol{a}'\boldsymbol{x}$, where $\boldsymbol{a}' = [a_1, a_2, \ldots, a_p]$,

$$\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}.$$

(iii) If $f(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$,

$$\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{A}'\boldsymbol{x}.$$

The matrix of second-order partial derivatives of $f(\boldsymbol{x})$, called the Hessian, is the matrix

$$\frac{\partial^2 f(\boldsymbol{x})}{\partial \boldsymbol{x} \partial \boldsymbol{x}'} = \begin{bmatrix} \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_1^2} & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_1 \partial x_p} \\ \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_2 \partial x_p} \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_p \partial x_1} & \dfrac{\partial^2 f(x)}{\partial x_p \partial x_2} & \cdots & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_p^2} \end{bmatrix}.$$

The derivative of the determinant of a matrix $\boldsymbol{A} = [a_{ij}]$ with respect to the element a_{ij} is

$$\frac{\partial |\boldsymbol{A}|}{\partial a_{ij}} = \boldsymbol{A}_{ij},$$

where A_{ij} is the cofactor of the element a_{ij}. If A is symmetric,

$$\frac{\partial|\boldsymbol{A}|}{\partial a_{ij}} = \begin{cases} \boldsymbol{A}_{ij}, & i = j, \\ 2\boldsymbol{A}_{ij}, & i \neq j. \end{cases}$$

Let $\boldsymbol{A}$ be an $m \times n$ matrix with elements a_{ij} as a function of $\boldsymbol{x}$. Then the derivative of $\boldsymbol{A}$ with respect to $\boldsymbol{x}$ is defined as the matrix of derivatives of its elements and is the matrix

$$\frac{\partial \boldsymbol{A}}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial(a_{11})}{\partial \boldsymbol{x}} & \frac{\partial(a_{12})}{\partial \boldsymbol{x}} & \cdots & \frac{\partial(a_{1n})}{\partial \boldsymbol{x}} \\ \frac{\partial(a_{21})}{\partial \boldsymbol{x}} & \frac{\partial(a_{22})}{\partial \boldsymbol{x}} & \cdots & \frac{\partial(a_{2n})}{\partial \boldsymbol{x}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial(a_{m1})}{\partial \boldsymbol{x}} & \frac{\partial(a_{m2})}{\partial \boldsymbol{x}} & \cdots & \frac{\partial(a_{mn})}{\partial \boldsymbol{x}} \end{bmatrix}.$$

If $\boldsymbol{A}$ is a nonsingular and square matrix,

$$\frac{\partial \boldsymbol{A}^{-1}}{\partial \boldsymbol{x}} = -\boldsymbol{A}^{-1}\frac{\partial \boldsymbol{A}}{\partial \boldsymbol{x}}\boldsymbol{A}^{-1}.$$

N NEWTON–RAPHSON, FISHER SCORING, AND EM ALGORITHMS

In this appendix, we briefly describe three commonly used iterative methods, Newton–Raphson, Fisher scoring, and EM algorithms, for calculating ML and REML estimates.

NEWTON–RAPHSON

The Newton–Raphson is an old and well-known method of maximizing or minimizing a function. It is an iterative method for finding a root of an equation. (More accurately, it is a method for finding stationary points of a function.) Given a function $f(\boldsymbol{\theta})$, the procedure attempts to derive a root of $f'(\boldsymbol{\theta}) = \partial f(\boldsymbol{\theta})/\partial \boldsymbol{\theta} = \mathbf{0}$ that may lead to a maximum. Using only the first-order (or linear) Taylor series approximation to the function $f'(\boldsymbol{\theta})$ about $\boldsymbol{\theta}_0$, we have

$$f'(\boldsymbol{\theta}) = f'(\boldsymbol{\theta}_0) + \frac{\partial^2 f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}(\boldsymbol{\theta} - \boldsymbol{\theta}_0).$$

Equating $f'(\boldsymbol{\theta})$ to $\mathbf{0}$ and solving for the root as $\boldsymbol{\theta}_1$, we obtain

$$\boldsymbol{\theta}_1 = \boldsymbol{\theta}_0 - \left[\frac{\partial^2 f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\right]^{-1} f'(\boldsymbol{\theta}_0).$$

Now, $\boldsymbol{\theta}_1$ can be substituted for $\boldsymbol{\theta}_0$ to set up an iterative scheme leading to $(m+1)$th iteration:

$$\boldsymbol{\theta}_{m+1} = \boldsymbol{\theta}_m - \boldsymbol{H}^{-1} f'(\boldsymbol{\theta}_m),$$

where

$$\boldsymbol{H} = \left[\frac{\partial^2 f(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]^{-1}_{\boldsymbol{\theta}=\boldsymbol{\theta}_m}.$$

The method has several drawbacks. It can fail to converge even to a local maximum when the linear approximation is a poor one, and the solution can be outside the parameter space. Finally, it should be noted that in applying the Newton–Raphson method to a log-likelihood function, different parametrizations can be expected to produce nonequivalent sequences of iterates.

FISHER SCORING

The method of scoring is an iterative method for maximizing a likelihood function. It is identical to the Newton–Raphson method except that the second-order partial derivatives of the log-likelihood function are replaced by their expected values. This way the computational effort required in evaluating the second derivative matrix in the Newton–Raphson is greatly reduced. As applied to the maximization of the log-likelihood function, $-\ell n L(\boldsymbol{\theta})$, the $(m+1)$th iterate of the method of scoring is determined using the form:

$$\boldsymbol{\theta}_{m+1} = \boldsymbol{\theta}_m + [\boldsymbol{I}(\boldsymbol{\theta}_m)]^{-1} f'(\boldsymbol{\theta}_m),$$

where $\boldsymbol{I}(\boldsymbol{\theta}_m)$ is the information matrix calculated using $\boldsymbol{\theta} = \boldsymbol{\theta}_m$. Jennrich and Sampson (1976) commented that the method of scoring is more robust to poor starting values than the Newton–Raphson procedure. They recommended a procedure in which scoring is used during the first few steps and then switches to Newton–Raphson. For some useful technical details and an overview of the connections between Newton–Raphson and Fisher Scoring, see Longford (1995).

THE EM ALGORITHM

The EM algorithm introduced by Dempster et al. (1977) is an elegant and popular technique for finding ML and REML estimates and posterior modes in missing data situations. It is an iterative procedure for calculating ML and REML estimates. The procedure alternates between calculating expected values and maximizing simplified likelihoods. The procedure is especially designed for situations where missing data are anticipated. It treats the observed data as incomplete and then attempts to fill in the missing data by calculating conditional expected values of the sufficient statistics given the observed data. The conditional expected values are then used in place of the sufficient statistics to improve estimates of the parameter. An iterative scheme is set up and convergence is guaranteed under relatively unrestricted conditions. In mixed effects models random effects are typically treated as “missing data” and are subsequently considered as fixed once they are filled in. The EM algorithm proceeds by evaluating the log-likelihood of the complete data, calculating its expectations with respect to the conditional distribution of the random effects given the observation vector $\boldsymbol{Y}$ and then maximizing with respect to the parameters. Now an iterative scheme can be set up since we can recalculate the log-likelihood of the complete data given the new parameter estimates, and so on.

The algorithm has several appealing properties relative to other iterative procedures such as Newton–Raphson. It is easily implemented since it relies on complete data computations and the M-step of each iteration involves taking expectations over complete-data ML estimation, which is often in closed form. It is numerically stable and convergence is nearly always to a local maximum for practically all important problems. However, if the M-step of this algorithm is not in closed form, EM loses some of its attractions. A number of modifications and extensions to the EM algorithm have been introduced to address this problem and to speed EM's convergence rate without losing its simplicity and monotone convergence properties. For a thorough and book-length coverage of the EM algorithm the reader is referred to McLachlan and Krishnan (1996). For a brief overview using minimum technical details, including a review of currently available software, see Longford (1995).

In recent years a rich variety of new algorithms such as quasi-Newton, Monte Carlo Newton–Raphson, Markov Chain Monte Carlo, Metropolis, among others, have been developed. The interested reader is referred to the works of Searle et al

indexCasella, G.. (1992, Chapter 8) and Kennedy and Gentle (1980) for a detailed treatment of these procedures. Callanan and Harville (1991) describe several new algorithms for computing REML estimates of variance components.

Bibliography

T. P. Callanan and D. A. Harville (1991), Some new algorithms for computing restricted maximum likelihood estimates of variance components, *J. Statist. Comput. Simul.*, **38**, 239–259.

A. P. Dempster, N. M. Laird, and D. B. Rubin (1977), Maximum likelihood from incomplete data via the EM algorithm, *J. Roy. Statist. Soc. Ser.* B, **39**, 1–38.

R. L. Jennrich and P. F. Sampson (1976), Newton–Raphson and related algorithms for maximum likelihood variance component estimation, *Technometrics*, **18**, 11–17.

W. J. Kennedy, Jr. and J. E. Gentle (1980), *Statistical Computing*, Marcel Dekker, New York.

N. T. Longford (1995), Random coefficient models, in G. Arminger, C. C. Clogg, and M. E. Sobel, eds., *Handbook of Statistical Modeling for the Social and Behavioral Sciences*, Plenum Press, London, 519–577.

G. J. McLachlan and T. Krishnan (1996), *The EM Algorithm and Extensions*, Wiley, New York.

S. R. Searle, G. Casella, and C. E. McCulloch (1992), *Variance Components*, Wiley, New York.

O SOFTWARE FOR VARIANCE COMPONENT ANALYSIS

Nowadays there is a host of computer software that can be used to perform many of the analyses described in the text. In this appendix, we briefly describe some major statistical packages (with their addresses) that can be used to analyze variance component models. Some of the packages described here, SAS, SPSS, BMDP, MINITAB, and GENMOD, were originally developed for mainframe computers, but are now available for personal computers. Microcomputers are now most commonly used for routine data analysis, and mainframe computers are needed only for very large data sets. It should, however, be borne in mind that the software industry is highly dynamic and all the packages are subject to ongoing development and frequent upgradings. Thus any attempt to describe them runs the immediate risk of being out of date by the time the information reaches the reader. Detailed reviews of statistical software, including history, categorization scheme and assessment criteria of software are given in Goldstein (1997, 1998). The Stata website (www.stata.com) contains links to these and other software.

SAS. There are several SAS procedures useful for analyzing random and mixed models, including PROC GLM, PROC NESTD, PROC VARCOMP, and PROC MIXED. PROC GLM is very general and can accommodate a variety of models, fixed, random, mixed. PROC NESTD is specially configured for anova designs where all factors are hierarchically nested and involve only random effects. PROC VARCOMP is especially designed for estimating variance components and currently implements four methods of variance component estimation. PROC MIXED, in addition to analyzing traditional variance component models, can also fit a variety of mixed models containing other covariance structures as well. To our knowledge, it is the most versatile software available for fitting all types of mixed models (random effects, random coefficients, and covariance pattern models, among others) to normal data. The procedure offers great flexibility and there are many options available for defining mixed models and for requesting output. It is also capable of performing a Bayesian analysis for random effects and random coefficient models. A complete description of all the SAS procedures and their features are available in the SAS/STAT manual published by SAS Institute, Inc. (2001). The package is available from the following address:

SAS Institute, Inc.
SAS Campus Drive
Cary, NC 27513
USA
www.sas.com/stat

SPSS. There are several SPSS procedures available for performing random and mixed effect analysis of variance, including MANOVA, GLM, and VARCOMP. In the MANOVA, special F-tests involving a random or mixed model analysis are performed by the use of a key word VS within the design statement. GLM procedure is probably the most versatile and complex of all the SPSS procedures and can accommodate both balanced and unbalanced designs, including nested or nonfactorial designs, multivariate data, and analyses involving random and mixed effects models. VARCOMP procedure is especially designed to estimate variance components and currently incorporates five methods of variance component estimation. A complete description of all the procedures and their features are available in the Advanced Statistics manual published by SPSS, Inc. (2001). The package is available from the following address:

SPSS, Inc.
233 S. Wacker Drive, 11th Floor
Chicago, IL 60608
USA
www.spss.com

BMDP. There are several BMDP procedures for analyzing normal mixed models, including 3V, 5V, and 8V. For designs involving balanced data, 8V is recommended since it is simpler to use and interpret. For designs with unbalanced data, 3V must be used. For random and mixed effect models, in addition to performing standard analysis of variance, 3V also provides variance component estimates using ML and REML procedures. Finally, 5V analyzes repeated measures data for a wide variety of models, and contains many modeling features such as a good number of options for the forms of the variance-covariance matrices, including unequal variances and covariances with specified patterns. The procedure processes unbalanced repeated measures models with structured covariance matrices, achieving an ANOVA model by way of ML estimation. It also permits the choice of several nonstandard designs such as unbalanced or partially missing data, and time-varying covariates. A complete description of all the procedures and their features are available in the BMDP manual by Dixon (1992). The package is no longer available from its former vendor BMDP Software, Inc., but its revivals, BMDP/PC and BMDP/Dynamic are available from the following address:

Statistical Solutions
Stone Hill Corporate Center
Suite 104

999 Broadway
Saugus, MA 01906
USA
www.statsolusa.com

S-PLUS. It is a general-purpose, command-driven, and highly interactive software package capable of analyzing mixed models. It includes hundreds of functions that operate on scalars, vectors, matrices, and more complex objects. The package is dramatically increasing in popularity because of its fantastic graphing capabilities. The procedures S-Plus LME and VARCOMP compute ML and REML estimates of the elements of the variance-covariance matrix for the random effects in a mixed model. It is available from the following address:

Mathsoft, Inc.
1700 West Lake Avenue
North Seattle, WA 98109
USA
www.splus.mathsoft.com

GENSTAT. It is a general-purpose software package capable of fitting normal mixed models. It provides a wide variety of data transformations and other manipulations to be carried out within the software with great ease and rapidity. The package incorporates generalized linear modeling and allows the application of linear regression, logistic and probit regression, log-linear models, and regression with skewed distributions, all in a unified and consistent manner. It has two programs, REML and VCOMPONENTS directives, which incorporate procedures for ML and REML estimation for normal response models. It is available from the following address:

Genstat Numerical Algorithms, Ltd.
Mayfield House
256 Banbury Road
Oxford OX2 7DE
UK
www.nag.com

BUGS. It is a special-purpose package designed to perform Bayesian analysis using Gibbs sampling. BUGS is an acronym for *Bayesian inference using Gibbs sampling* and was developed by the Biostatistics unit of the Medical Research Council in Cambridge, England. Gibbs sampling is a popular procedure belonging to the family of Markov chain Monte Carlo (MCMC) algorithms, which exploits the properties of Markov Chains where the probability of an event is conditionally dependent on a previous state. BUGS allows MCMC estimation for a wide range of models and can be used to fit random and mixed effect models to all types of data, including hierarchical linear and nonlinear models. The program determines the complete conditional distribution necessary for implementing Gibbs algorithm and uses S-like syntax for specifying hierarchical models. A description of the program has been given by Gilks et al. (1992) and a comprehensive easy-to-read user guide and a booklet of worked BUGS examples are also available (Spiegelhalter et al., 1995a, 1995b). It is available from the following address:

MRC Biostatistics Unit
Institute of Public Health
Robinson Way
Cambridge CB2 2SR
UK
www.mrc-bsu.cam.ac.uk/bugs

Other software. The GENMOD, HLM, ML3, and VARCL are special-purpose packages for fitting multilevel models and contain programs for performing mixed model analysis. A detailed review of these four packages has been given by Kreft et al. (1994). The review includes a comparison with respect to ease of use, documentation, error handling, execution speed, and accuracy and readability of the output. In their original forms, they were designed to fit normally distributed data and produced ML or REML estimates. One of these, GENMOD (Mason et al., 1988), though popular among demographers for whom it was originally developed, is no longer generally available. The other three, HLM (Bryk et al., 1988), ML3 (Prosser et al., 1991), and VARCL (Longford, 1988), are all capable to fit three-level models and ML3 and VARCL incorporate procedures for fitting binomial and Poisson response models. The two successors to ML3, Mln, and MlwiN (the Windows version) (Rasbash et al., 1995) are capable of fitting a very large number of levels, together with case weights, measurement errors, and robust estimates of standard errors. They also have a high level MACRO language that allows a wide range of special purpose computations that can be readily carried out. MlwiN also allows a wide variety of data manipulations that can be carried out within the software whereas others tend to require a somewhat rigid data structure. HLM is widely used by social scientists and educational researchers in the USA (where it was developed) while ML3 and VARCL are more popular in the UK (where they were developed) and are also used by social and educational researchers. HLM is available from the following address:

Scientific Software, Inc.
1525 East 53rd St., Suite 906
Chicago, IL 60615
USA

ML3, Mln, and MlwiN are available from the following address:

Hilary Williams
Institute of Education

University of London
20 Bedford Way
London WC1H 0AL
UK

VARCL (and also HLM, ML3, Mln, and MlwiN) are available from the following address:

ProGamma
P. O. B. Groningen
The Netherlands

Bibliography

A. S. Bryk, S. W. Raudenbush, M. Seltzer, and R. Congdon (1988), *An Introduction to HLM: Computer Program and User's Guide*, 2nd ed., Department of Education, University of Chicago, Chicago.

W. J. Dixon, ed. (1992), *BMDP Statistical Software Manual*, Vols. I, II, and III, University of California Press, Los Angeles.

W. R. Gilks, A. Thomas, and D. J. Spiegelhalter (1992), Software for the Gibbs sampler, *Comput. Sci. Statist.*, **24**, 439–448.

R. Goldstein (1997), Computer packages, in S. Kotz and N. L. Johnson, eds., *Encyclopedia of Statistical Sciences*, Update Vol. 1., Wiley, New York, 504–510.

R. Goldstein (1998), Software, biostatistical, in P. Armitage and T. Colton, eds., *Encyclopedia of Biostatistics*, Vol. 5, Wiley, Chichester, UK, 4180–4187.

I. G. Kreft, J. de Leeuw, and R. van der Leeden (1994), Comparing five different statistical packages for hierarchical linear regression: BMDP-5V, GENMOD, HLM, ML3, and VARCL, *Amer. Statist.*, **48**, 324–335.

N. T. Longford (1988), *VARCL: Software for Variance Component Analysis of Data with Hierarchically Nested Random Effects* (*Maximum Likelihood*), Educational Testing Service, Princeton, NJ.

W. M. Mason, A. F. Anderson, and N. Hayat (1988), *Manual for GENMOD*, Population Studies Center, University of Michigan, Ann Arbor, MI.

R. Prosser, J. Rasbash, and H. Goldstein (1991), *ML3 Software for Three-Level Analysis: User's Guide for Version* 2, Institute of Education, University of London, London.

J. Rasbash, M. Yang, G. Woodhouse, and H. Goldstein (1995), *Mln Command Reference*, Institute of Education, University of London, London.

SAS Institute (2001), *SAS/STAT User's Guide: Version* 8.2.0, Vols. I and II, 4th ed., SAS Institute, Cary, NC.

D. J. Spiegelhalter, A. Thomas, N. Best, and W. R. Gilks (1995a), *BUGS: Bayesian Inference Using Gibbs Sampling, Version* 0.50, Technical Report, Medical Research Council, Biostatistics Unit, Institute of Public Health, Cambridge University, Cambridge, UK.

D. J. Spiegelhalter, A. Thomas, N. Best, and W. R. Gilks (1995b), *BUGS Examples, Version* 0.50, Technical Report, Medical Research Council, Biostatistics Unit, Institute of Public Health, Cambridge University, Cambridge, UK.
SPSS, Inc. (2001), *SPSS Advanced Statistics, Release* 11.0., SPSS, Chicago.

General Bibliography

M. Abramowitz and I. A. Stegun (1965), *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Applied Mathematics Series 55, National Bureau of Standards, Washington, DC.

H. J. Ahrens (1978), MINQUE and ANOVA estimator for one-way classification: A risk comparison, *Biometrical J.*, **20**, 535–556.

H. J. Ahrens and R. Pincus (1981), On two measures of unbalancedness in a one-way model and their relation to efficiency, *Biometrical J.*, **23**, 227–235.

H. J. Ahrens and J. E. Sánchez (1982), Unbalancedness and efficiency in estimating components of variance: MINQUE and ANOVA procedure, *Biometrical J.*, **24**, 649–661.

H. J. Ahrens and J. E. Sánchez (1986), *Measures of Unbalancedness and Their Relations to Variance Component Estimation*, Technical Report P-Math-34/8, Akademie der Wissenschaften der DDR, Karl-Weierstrass-Insitut für Matematik, Berlin.

H. J. Ahrens and J. E. Sánchez (1988), Unbalancedness of designs, Measures of, in S. Kotz and N. L. Johnson, *Encyclopedia of Statistical Sciences*, Vol. 9, Wiley, New York, 383–386.

H. J. Ahrens and J. E. Sánchez (1992), Imbalance and its influence on variance component estimation, *Biometrical J.*, **34**, 539–555.

H. J. Ahrens, J. Kleffé, and R. Tenzler (1981), Mean square error comparison of MINQUE, ANOVA and two alternative estimators under the unbalanced one-way random effects model, *Biometrical J.*, **23**, 323–342.

A. Albert (1972), *Regression and the Moore–Penrose Inverse*, Academic Press, New York.

R. D. Anderson (1979), Estimating variance components from balanced data: Optimum properties of REML solutions and MIVQUE estimation, in L. D. Van Vleck and S. R. Searle, eds., *Variance Components and Animal Breeding: Proceedings of a Conference in Honor of C. R. Henderson*, Animal Science Department, Cornell University, Ithaca, NY, 205–206.

R. L. Anderson (1960), Use of variance component analysis in the interpretation of biological experiments, Part 3, *Bull. Internat. Statist. Inst.*, **37**, 71–90.

R. L. Anderson (1961), Designs for estimating variance components, in *Proceedings of the Seventh Conference on Design of Experiments and Army Preservation and Development Testing*, 781–823; also published as Mim. Ser. 310, Institute of Statistics, North Carolina State University, Raleigh, NC.

R. L. Anderson (1975), Designs and estimators for variance components, in J. N. Srivastava, ed., *Statistical Design and Linear Model*, North-Holland, Amsterdam, 1–30.

R. L. Anderson (1981), Recent developments in designs and estimators for variance components, in M. Csörgö, D. A. Dawson, J. N. K. Rao, and A. K. Md. E. Saleh, eds., *Statistics and Related Topics*, North-Holland, Amsterdam, 3–22.

R. L. Anderson and T. A. Bancroft (1952), *Statistical Theory in Research*, McGraw–Hill, New York.

R. L. Anderson and P. P. Crump (1967), Comparisons of designs and estimation procedures for estimating parameters in a two-stage nested process, *Technometrics*, **9**, 499–516.

L. A. Aroian (1941), Continued fraction for the incomplete beta function, *Ann. Math. Statist.*, **12**, 218–223.

J. N. Arvesen (1976), A note on the Tukey–Hooke variance component results, *Ann. Inst. Statist. Math.* (*Japan*), **28**, 111–121.

J. N. Arvesen and M. W. J. LayardArvesen, J. N. (1975), Asymptotically robust tests in unbalanced variance components models, *Ann. Statist.*, **3**, 1122–1134.

J. N. Arvesen and T. H. SchmitzArvesen, J. N. (1970), Robust procedures for variance component problems using the jacknife, *Biometrics*, **26**, 677–686.

I. Ashida and H. Iwaisaki (1995), Estimation of sampling co(variance) of the REML estimators for variance components in a mixed linear model, *Japan J. Biometrics*, **16**, 9–17.

T. R. Bainbridge (1963), Staggered nested designs for estimating variance components, in *American Society for Quality Control Annual Conference Transactions*, American Society for Quality Control, Milwaukee, 93–103.

T. R. Bainbridge (1965), Staggered nested designs for estimating variance components, *Indust. Quality Control*, **22**-1, 12–20.

A. H. J. Baines (1943), *On the Allocation of Observations in Certains Types of Factorial Experiments when the Number of Observations is Limited*, Technical Report Q.C./R/4, Ministry of Supply Advisory Service on Statistical Method and Quality Control, London.

J. K. Baksalary and A. Molinska (1984), Nonnegative unbiased estimability of linear combinations of two variance components, *J. Statist. Plann. Inference*, **10**, 1–8.

N. K. Bansal and M. Bhandary (1994), Robust M-estimation of the intraclass correlation coefficient, *Australian J. Statist.*, **36**, 287–301.

R. B. Bapat and T. E. S. Raghavan (1997), *Nonnegative Matrices and Applications*, Cambridge University Press, Cambridge, UK.

A. Basilevsky (1983), *Applied Matrix Algebra in the Statistical Sciences*, North-Holland, Amsterdam.

R. M. Baskin (1993), Estimation of variance components for the U. S. Consumer Price Index via Gibbs sampling, in *ASA Proceedings of Survey Research Section*, American Statistical Association, Alexandria, VA, 803–813.

E. Belzile and J.-F. Angers (1995), Choice of noninformative priors for the variance components of an unbalanced one-way random model, *Comm. Statist.* A *Theory Methods*, **24**, 1325–1341.

A. Ben-Israel and T. Greyville (1974), *Generalized Inverses: Theory and Applications*, Wiley, New York.

A. Bener and S. Huda (1993), Estimation of intraclass correlation when family sizes are unequal, *Comput. Statist. Data Anal.*, **15**, 273–284.

J. O. Berger and J. M. Bernardo (1992), Reference priors in a variance components problem, in P. K. Goel and N. S. Iyenger, eds., *Bayesian Analysis in Statistics and Econometrics*, Springer-Verlag, New York, 177–194.

K. Berk (1987), Computing for incomplete repeated measures, *Biometrics*, **43**, 385–398.

A. Berman and R. J. Plemmons (1994), *Nonnegative Matrices in the Mathematical Sciences*, 2nd ed., Society for Industrial and Applied Mathematics, Philadelphia.

R. P. Bhargava (1946), Tests of significance for intraclass correlation when family sizes are not equal, *Sankhyā*, **7**, 435–438.

W. R. Blischke (1966), Variances of estimates of variance components in a three-way classification, *Biometrics*, **22**, 553–565.

W. R. Blischke (1968), Variances of moment estimators of variance components in the unbalanced r-way classification, *Biometrics*, **24**, 527–540.

C. I. Bliss (1967), *Statistics in Biology*, McGraw–Hill, New York.

G. E. P. Box and G. C. Tiao (1973), *Bayesian Inference in Statistical Analysis*, Addison–Wesley, Reading, MA.

R. Bremer (1989), Numerical study of small sample variances of estimators of variance components in the two-way factorial model, *Comm. Statist.* B *Simul. Comput.*, **18**, 985–1009.

R. Bremer (1990), Numerical study of the small sample variance of the MINQUE estimators of variance components in the two-way factorial design, *Comm. Statist.* A *Theory Methods*, **19**, 1261–1279.

N. E. Breslow and D. G. Clayton (1993), Approximate inference in generalized linear mixed models, *J. Amer. Statist. Assoc.*, **88**, 9–25.

J. C. Brockleban and F. G. Giesbrech (1984), Estimating variance components using alternative MINQUEs in selected unbalanced design, in K. Hinkelman, ed., *Experimental Design, Statistical Methods and Genetic Statistics*, Marcel Dekker, New York, 172–211.

L. D. Broemeling (1985), *Bayesian Analysis of Linear Models*, Marcel Dekker, New York.

H. Brown and R. Prescott (1999), *Applied Mixed Models in Medicine*, Wiley, New York.

H. K. Brown and R. A. Kempton (1994), The application of REML in clinical trials, *Statist. Med.*, **13**, 1601–1617.

K. G. Brown (1976), Asymptotic behavior of MINQUE-type estimators of variance components, *Ann. Statist.*, **4**, 746–754.

K. G. Brown (1977), On estimation of diagonal covariance matrices by MINQUE, *Comm. Statist.* A *Theory Methods*, **6**, 471–484.

K. G. Brown (1978), Estimation of variance components using residuals, *J. Amer. Statist. Assoc.*, **73**, 141–146.

K. A. Brownlee (1965), *Statistical Theory and Methodology in Science and Engineering*, 2nd ed., Wiley, New York.

A. S. Bryk, S. W. Raudenbush, M. Seltzer, and R. Congdon (1988), *An Introduction to HLM: Computer Program and User's Guide*, 2nd ed., Department of Education, University of Chicago, Chicago.

B. D. Burch and H. K. Iyer (1997), Exact confidence intervals for a variance ratio (or heritability) in a mixed linear model, *Biometrics*, **53**, 1318–1333.

R. K. Burdick, H. J. Birch, and F. A. Graybill (1986), Confidence intervals on measures of variability in an unbalanced two-fold nested design with equal subsampling, *J. Statist. Comput. Simul.*, **25**, 259–272.

R. K. Burdick and J. E. Eickman (1986), Confidence intervals on the among group variance component in the unbalanced one-fold nested design, *J. Statist. Comput. Simul.*, **26**, 205–219.

R. K. Burdick and F. A. Graybill (1984), Confidence intervals on linear combinations of variance components in the unbalanced one-way classification, *Technometrics*, **26**, 131–136.

R. K. Burdick and F. A. Graybill (1985), Confidence intervals on the total variance in an unbalanced two-fold nested classification with equal subsampling, *Comm. Statist.* A *Theory Methods*, **14**, 761–774.

R. K. Burdick and F. A. Graybill (1992), *Confidence Intervals on Variance Components*, Marcel Dekker, New York.

R. K. Burdick, F. Maqsood, and F. A. Graybill (1986), Confidence intervals on the intraclass correlation in the unbalanced one-way classification, *Comm. Statist.* A *Theory Methods*, **15**, 3353–3378.

P. R. Burton, K. J. Tiller, L. C. Gurrin, W. A. C. M. Cookson, A. W. Musk, and L. J. Palmer (1999), Genetic variance components analysis for binary phenotypes using generalized linear mixed models (GLMMs) and Gibbs sampling, *Genetic Epidemiology*, **17**, 118–140.

N. Bush (1962), *Estimating Variance Components in a Multi-Way Classification*, Ph.D. dissertation, Mim. Ser. 324, North Carolina State University, Raleigh, NC.

N. Bush and R. L. Anderson (1963), A comparison of three different procedures for estimating variance components, *Technometrics*, **5**, 421–440.

T. P. Callanan (1985), *Restricted Maximum Likelihood Estimation of Variance Components: Computational Aspects*, Ph.D.. dissertation, Iowa State University, Ames, IA.

T. P. Callanan and D. A. Harville (1989), Some new algorithms for computing maximum likelihood estimates of variance components, in K. Berk and L. Malone, eds., *Computer Science and Statistics: Proceedings of the 21st Annual Symposium on the Interface*, American Statistical Association, Alexandria, VA, 435–444.

T. P. Callanan and D. A. Harville (1991), Some new algorithms for computing restricted maximum likelihood estimates of variance components, *J. Statist. Comput. Simul.*, **38**, 239–259.

L. D. Calvin and J. D. Miller (1961), A sampling design with incomplete dichotomy, *Agronomy J.*, **53**, 325–328.

J. M. Cameron (1951), Use of components of variance in preparing schedules for the sampling of baled wool, *Biometrics*, **7**, 83–96.

R. F. Caro, M. Grossman, and R. L. Fernando (1985), Effects of data imbalance on estimation of heritability, *Theoret. Appl. Genetics*, **69**, 523–530.

K. Chaloner (1987), A Bayesian approach to the estimation of variance components for the unbalanced one-way model, *Technometrics*, **29**, 323–337.

J. S. K. Chan and A. Y. C. Kuk (1997), Maximum likelihood estimation for probit-linear mixed models with correlated random effects, *Biometrics*, **53**, 86–97.

S. K. Chatterjee and K. Das (1983), Estimation of variance components in an unbalanced one-way classification, *J. Statist. Plann. Inference*, **8**, 27–41.

Y. B. Chaubey (1977), *The Principle of MINQUE in Linear Models: Modifications, Extensions and Applications*, Ph.D. dissertation, University of Rochester, Rochester, NY.

Y. B. Chaubey (1980a), Application of the method of MINQUE for estimation in regression with intraclass covariance matrix, *Sankhyā*, **42**, 28–32.

Y. B. Chaubey (1980b), Minimum norm quadratic estimators of variance components, *Metrika*, **27**, 225–262.

Y. B. Chaubey (1982), Minimum norm invariant quadratic estimation of a covariance matrix in linear model, *Biometrical J.*, **24**, 457–461.

Y. B. Chaubey (1983), A non-negative estimator of variance component closest to MINQUE, *Sankhyā*, **45**, 201–211.

Y. B. Chaubey (1984), On the comparison of some non-negative estimators of variance components for two models, *Comm. Statist.* B *Simul. Comput.*, **13**, 619–633.

Y. B. Chaubey (1985), Estimators of covariances in time series model, *Statist. Probab. Lett.*, **3**, 51–53.

Y. B. Chaubey (1991), A note on nonnegative minimum bias MINQE in variance components model, *Statist. Probab. Lett.*, **11**, 395–397.

Y. B. Chaubey and P. S. R. S. Rao (1976), Efficiencies of five estimators for the parameters of two linear models with unequal variances, *Sankhyā Ser.* B, **38**, 364–370.

R. Christiansen (1996), Exact tests for variance components, *Biometrics*, **52**, 309–314.

W. G. Cochran (1934), The distribution of quadratic forms in a normal system, *Proc. Cambridge Philos. Soc. Suppl.*, **4**, 102–118.

W. G. Cochran (1937), Problems arising in the analysis of a series of similar experiments, *J. Roy. Statist. Soc. Suppl.*, **4**, 102–118.

W. G. Cochran (1939), The use of the analysis of variance in enumeration by sampling, *J. Amer. Statist. Assoc.*, **34**, 492–510.

W. G. Cochran (1954), The combination of estimates from different experiments, *Biometrics*, **10**, 109–129.

W. G. Cochran and S. P. Carroll (1953), A sampling investigation of the efficiency of weighting inversely as the estimated variance, *Biometrics*, **9**, 447–459.

C. C. Cockerham and B. S. Weir (1977), Quadratic analysis of reciprocal crosses, *Biometrics*, **33**, 187–203.

M. R. Conaway (1989), Analysis of repeated categorical measurements with conditional likelihood methods, *J. Amer. Statist. Assoc.*, **84**, 53–62.

M. D. Conerly and J. T. Webster (1987), MINQE for the one-way classification, *Technometrics*, **29**, 229–236.

R. R. Corbeil and S. R. Searle (1976a), A comparison of variance component estimators, *Biometrics*, **32**, 779–791.

R. R. Corbeil and S. R. Searle (1976b), Restricted maximum likelihood (REML) estimation of variance components in the mixed model, *Technometrics*, **18**, 31–38.

P. Cook, L. D. Broemeling, and M. K. Gharaff (1990), A Bayesian analysis of mixed linear model, *Comm. Statist.* A *Theory Methods*, **19**, 987–1002.

D. R. Cox (1955), Some statistical methods connected with series of events (with discussion), *J. Roy. Statist. Soc. Ser.* B, **17**, 129–164.

D. R. Cox and D. V. Hinkly (1974), *Theoretical Statistics*, Chapman and Hall, London; softbound ed., 1986.

N. Cressie and S. N. Lahiri (1993), The asymptotic distribution of REML estimators, *J. Multivariate Anal.*, **45**, 217–233.

P. P. Crump (1954), *Optimal Designs to Estimate the Parameters of a Variance Component Model*, Ph.D. dissertation, North Carolina State University, Raleigh, NC.

S. L. Crump (1946), The estimation of variance components in analysis of variance, *Biometrics Bull.*, **2**, 7–11.

S. L. Crump (1947), *The Estimation of Components of Variance in Multiple Classifications*, Ph.D. dissertation, Iowa State University, Ames, IA.

S. L. Crump (1951), The present status of variance components analysis, *Biometrics*, **7**, 1–16.

W. B. Cummings (1972), *Variance Component Testing in Unbalanced Nested Designs*, Mim. Ser. 243, Institute of Statistics, North Carolina State University, Raleigh, NC.

W. B. Cummings and D. W. Gaylor (1974), Variance component testing in unbalanced nested designs, *J. Amer. Statist. Assoc.*, **69**, 765–771.

K. Das (1979), Asymptotic optimality of restricted maximum likelihood estimates for the mixed model, *Calcutta Statist. Assoc. Bull.*, **23**, 125–142.

R. Das and B. K. Sinha (1987), Robust optimum invariant unbiased test for variance components, in T. Pukkila and S. Puntanen, eds., *Proceedings of the 2nd International Tampere Conference on Statistics*, University of Tampere, Tampere, Finland, 317–342.

O. L. Davies and P. L. Goldsmith, eds. (1972), *Statistical Methods in Research and Production*, 4th ed., Oliver and Boyd, Edinburgh.

R. B. Davies (1980), The distribution of a linear combination of χ^2 random variables, *Appl. Statist.*, **29**, 323–333.

E. Demidenko and H. Massam (1999), On the existence of the maximum likelihood estimates in variance components models, *Sankhyā Ser.* A, **61**, 431–443.

A. P. Dempster, N. M. Laird, and D. B. Rubin (1977), Maximum likelihood from incomplete data via the EM algorithm, *J. Roy. Statist. Soc. Ser.* B, **39**, 1–38.

A. P. Dempster, D. B. Rubin, and R. K. Tsutakawa (1981), Estimation in covariance components models, *J. Amer. Statist. Assoc.*, **76**, 341–353.

A. P. Dempster, M. R. Selwyn, C. M. Patel, and A. J. Roth (1984), Statistical and computational aspects of mixed model analysis, *Appl. Statist.*, **33**, 203–214.

W. J. Dixon, ed. (1992), *BMDP Statistical Software Manual*, Vols. I, II, and III, University of California Press, Los Angeles.

A. J. Dobson (2002), *An Introduction to Generalized Linear Models*, 2nd ed., Chapman and Hall–CRC Press, Boca Raton, FL.

A. Donner (1979), The use of correlation and regression in the analysis of family resemblance, *Amer. J. Epidemiology*, **110**, 335–342.

A. Donner (1986), A review of inference procedures for the intraclass correlation coefficient in the one-way random effects model, *Internat. Statist. Rev.*, **54**, 67–82.

A. Donner, M. Eliasziw, and M. Shoukri (1998), Review of inference procedures for the interclass coefficient with emphasis on applications to family studies, *Genetic Epidemiology*, **15**, 627–646.

A. Donner and J. J. Koval (1980a), The estimation of intraclass correlation in the analysis of family data, *Biometrics*, **36**, 19–25.

A. Donner and J. J. Koval (1980b), The large sample variance of an intraclass correlation, *Biometrika*, **67**, 719–722.

A. Donner and J. J. Koval (1983a), Variance component estimation from human sibship data, *Biometrics*, **39**, 599–605.

A. Donner and J. J. Koval (1983b), A note on the accuracy of Fisher's approximation to the large sample variance of an intraclass correlation, *Comm. Statist.* B *Simul. Comput.*, **12**, 443–449.

A. Donner and J. J. Koval (1989), The effect of imbalance on significance-testing in one-way model II analysis of variance, *Comm. Statist.* A *Theory Methods*, **18**, 1239–1250.

A. Donner and G. A. Wells (1986), A comparison of confidence interval methods for the intraclass correlation coefficient, *Biometrics*, **42**, 401–412.

A. Donner, G. A. Wells, and M. Eliasziw (1989), On two aproximations to the F-distribution: Application to testing for intraclass correlation in family studies, *Canadian J. Statist.*, **17**, 209–215.

E. J. Eisen (1966), The quasi-F test for an unnested fixed factor in an unbalanced hierarchal design with a mixed model, *Biometrics*, **22**, 937–942.

C. Eisenhart (1947), The assumptions underlying the analysis of variance, *Biometrics*, **3**, 1–21.

M. Y. El-Bassiouni (1994), Short confidence intervals for variance components, *Comm. Statist.* A *Theory Methods*, **23**, 1915–1933.

M. Y. El-Bassiouni and J. F. Seely (1988), On the power of Wald's variance component test in the unbalanced random one-way model, in Y. Dodge,

V. V. Fedorov, and H. P. Wynn, eds., *Optimal Design and Analysis of Experiments*, North-Holland, Amsterdam, 157–165.

S. D. Elliott (1989), The method of unweighted means in univariate and multivariate analysis of variance, *Educ. Psychology Meas.*, **49**, 399–405.

D. A. Elston (1998), Estimation of denominator degrees of freedom of F-distributions for assessing Wald statistics for fixed-effect factors in unbalanced mixed models, *Biometrics*, **54**, 1085–1096.

B. Engel (1990), The analysis of unbalanced linear models with variance components, *Statist. Neerland.*, **44**, 195–219.

R. Fayyad, F. A. Graybill, and R. Burdick (1996), A note on exact tests for variance components in unbalanced random and mixed linear models, *Biometrics*, **52**, 306–308.

W. H. Fellner (1986), Robust estimation of variance components, *Technometrics*, **28**, 51–60.

A. R. Fennech and D. A. Harville (1991), Exact confidence sets for variance components in unbalanced mixed linear models, *Ann. Statist.*, **4**, 1771–1785.

A. L. Finker, J. J. Morgan, and R. J. Monroe (1943), *Methods of Estimating Farm Employment from Sample Data in North Carolina*, Technical Bulletin 75, North Carolina State Agriculture Experimental Station, Raleigh, NC.

J. Focke and G. Dewess (1972), Uber die achätz methode Minque von C. R. Rao and ihre verallgemeinerung, *Math. Oper. Statist. Ser. Statist.*, **3**, 129–143.

R. N. Forthofer and G. G. Koch (1974), An extension of the symmetric sum approach to estimation of variance components, *Biomet. Z.*, **16**, 3–14.

J. Gallo, R. MacElderry, and A. I. Khuri (1989), A SAS matrix software macro for estimating the random effects with unequal cell frequencies in the last stage, in *Procedings of the* 14*th SAS Users Group International Conference*, SAS Institute, Cary, NC, 1277–1282.

M. Ganguli (1941), A note on nested sampling, *Sankhyā*, **5**, 449–452.

S. Gao and T. M. F. Smith (1995), On the non-existence of a global nonnegative minimum bias invariant quadratic estimator of variance components, *Statist. Probab. Lett.*, **25**, 117–120.

C. E. Gates and C. Shiue (1962), The analysis of variance of the s-stage hierarchal classification, *Biometrics*, **18**, 529–536.

D. W. Gaylor (1960), *The Construction and Evaluation of Some Designs for the Estimation of Parameters in Random Models*, Ph.D. dissertation, Mim. Ser. 256, Institute of Statistics, North Carolina State University, Raleigh, NC.

D. W. Gaylor and T. D. Hartwell (1969), Expected mean squares for nested classifications, *Biometrics*, **25**, 427–430.

D. W. Gaylor and F. N. Hopper (1969), Estimating the degrees of freedom for linear combinations of mean squares by Satterthwaite's formula, *Technometrics*, **11**, 691–706.

D. W. Gaylor, H. L. Lucas, and R. L. Anderson (1970), Calculation of expected mean squares by the abbreviated Doolittle and square root methods, *Biometrics*, **20**, 641–656.

A. E. Gelfand and A. F. M. Smith (1991), Gibbs sampling for marginal posterior expectations, *Comm. Statist.* A *Theory Methods*, **20**, 1747–1766.

A. E. Gelfand, S. E. Hills, S. Racine-Poon, and A. F. M. Smith (1990), Illustration of Bayesian inference in normal data models using Gibbs sampling, *J. Amer. Statist. Assoc.*, **85**, 972–985.

J. E. Gentle (1998), *Numerical Linear Algebra for Applications in Statistics*, Springer-Verlag, New York.

C. F. Gerald (1977), *Applied Numerical Analysis*, Addison–Wesley, Reading, MA.

M. K. Gharaff (1979), *A General Solution to Making Inferences about the Parameters of Mixed Linear Models*, Ph.D. dissertation, Oklahoma State University, Stillwater, OK.

M. Ghosh (1996), On the nonexistence of nonnegative unbiased estmators of variance components, *Sankhyā Ser.* B, **58**, 360–362.

D. Gianola and R. L. Fernando (1986a), Random effects models for binary responses, *Biometrics*, **42**, 217–218.

D. Gianola and R. L. Fernando (1986b), Bayesian methods in animal breeding theory, *J. Animal Sci.*, **63**, 217–244.

D. Gianola and J. L. Foulley (1990), Variance estimation from integrated likelihoods (VEIL), *Genetic Selection Evolution*, **22**, 403–417.

R. D. Gibbons and D. Hedeker (1997), Random effects probit and logic regression models for three-level data, *Biometrics*, **53**, 1527–1537.

F. G. Giesbrecht (1983), An efficient procedure for computing MINQUE of variance components and generalized least squares estimates of fixed effects, *Comm. Statist.* A *Theory Methods*, **12**, 2169–2177.

F. G. Giesbrecht and J. C. Burns (1985), Two-stage analysis based on a mixed model: Large sample asymptotic theory and small-sample simulation results, *Biometrics*, **41**, 477–486.

F. G. Giesbrecht and P. M. Burrows (1978), Estimating variance components in hierarchical structures using MINQUE and restricted maximum likelihood, *Comm. Statist.* A *Theory Methods*, **7**, 891–904.

W. R. Gilks, D. G. Clayton, D. J. Spiegelhalter, N. G. Best, A. J. McNiel, L. D. Sharples, and A. J. Kirby (1993), Modeling complexity: Applications of Gibbs sampling in medicine (with discussion), *J. Roy. Statist. Soc. Ser.* B, **55**, 39–102.

W. R. Gilks, A. Thomas, and D. J. Spiegelhalter (1992), Software for the Gibbs sampler, *Comput. Sci. Statist.*, **24**, 439–448.

A. R. Gilmour, R. D. Anderson, and A. L. Rae (1985), The analysis of binomial data by a generalized linear mixed model, *Biometrika*, **72**, 593–599.

D. Giltinan and M. Davidian (1995), *Nonlinear Models for Repeated Measurement Data*, Chapman and Hall, London.

A. Giovagnoli and P. Sebastiani (1989), Experimental designs for mean and variance estimation in variance components models, *Comput. Statist. Data Anal.*, **8**, 21–28.

S. Gnot, J. Kleffé, and R. Zmyslony (1985), Nonnegativity of admissible invariant quadratic estimates in mixed linear models with two variance components, *J. Statist. Plann. Inference*, **12**, 249–258.

A. S. Golberger (1962), Best linear unbiased prediction in the generalized linear regression model, *J. Amer. Statist. Assoc.*, **57**, 369–375.

C. H. Goldsmith (1969), *Three-Stage Nested Designs for Estimating Variance Components*, Ph.D. dissertation, Institute of Statistics, North Carolina State University, Raleigh, NC.

C. H. Goldsmith and D. W. Gaylor (1970), Three-stage nested designs for estimating variance components, *Technometrics*, **12**, 487–498.

H. Goldstein (1995), *Multilevel Statistical Models*, 2nd ed., Arnold, London.

R. Goldstein (1997), Computer packages, in S. Kotz and N. L. Johnson, eds., *Encyclopedia of Statistical Sciences*, Update Vol. 1., Wiley, New York, 504–510.

R. Goldstein (1998), Software, biostatistical, in P. Armitage and T. Colton, eds., *Encyclopedia of Biostatistics*, Vol. 5, Wiley, Chichester, UK, 4180–4187.

M. Gönen (2000), A Bayesian analysis of the intraclass correlations in the mixed linear model, *Comm. Statist.* A *Theory Methods*, **29**, 1451–1464.

J. H. Goodnight and W. J. Hemmerle (1979), A simplified algorithm for the *W*-transformation in variance component estimation, *Technometrics*, **21**, 265–268.

T. I. Goss and R. G. Garrett (1978), A new unbalanced nested ANOVA model in geology: A down-to-earth design, in *ASA Proceedings of Statistical Computing Section*, American Statistical Association, Alexandria, VA, 360–365.

D. G. Gosslee and H. L. Lucas (1965), Analysis of variance of disproportionate data when interaction is present, *Biometrics*, **21**, 115–133.

J. C. Gower (1962), Variance component estimation for unbalanced hierarchal classifications, *Biometrics*, **18**, 537–542.

H.-U. Graser, S. P. Smith, and B. Tier (1987), A derivative-free approach for estimating variance components in animal models by restricted maximum likelihood, *J. Animal Sci.*, **64**, 1362–1370.

F. A. Graybill (1961), *An Introduction to Linear Statistical Models*, Vol. I, McGraw–Hill, New York.

F. A. Graybill (1983), *Introduction to Matrices with Applications in Statistics*, 2nd ed., Wadsworth, Belmont, CA; 1st ed., 1969.

J. W. Green (1988), *Diagnostic Methods for Repeated Measures Experiments with Missing Cell*, Technical Report, Department of Mathematical Sciences, University of Delware, Newark, DE.

B. Griffing (1956), Concept of general and specific combining ability in relation to diallel crossing system, *Australian J. Biol. Sci.*, **9**, 463–493.

D. J. Groggel, D. D. Wackerly, and P. V. Rao (1988), Nonparametric estimation in one-way random effects models, *Comm. Statist.* B *Simul. Comput.*, **17**, 887–903.

M. L. Gumpertz and S. G. Pantula (1992), Nonlinear regression with variance components, *Amer. Statist. Assoc.*, **87**, 201–209.

M. L. Gumpertz and J. O. Rawlings (1992), Nonlinear regression with variance components: Modeling effects of ozone on crop yield, *Crop Sci.*, **32**, 219–224.

A. S. Hadi (1996), *Matrix Algebra as a Tool*, Wadsworth, Belmont, CA.

A. Haile and J. T. Webster (1975), A comparison of four designs for estimating the variances of random effects in a two-way classification, *Technometrics*, **17**, 495–500.

J. M. Hammersley (1949), The unbiased estimate and standard error of the interclass variance, *Metron*, **15**, 189–205.

D. W. Hardy and C. L. Walker (1995), *Doing Mathematics with Scientific Workplace*, Brooks–Cole, Pacific Grove, CA.

H. O. Hartley (1967), Expectations, variances and covariances of ANOVA mean squares by "synthesis," *Biometrics*, **23**, 105–114; corrigenda, **23**, 853.

H. O. Hartley and J. N. K. Rao (1967), Maximum likelihood estimation for the mixed analysis of variance model, *Biometrika*, **54**, 93–108.

H. O. Hartley and W. K. Vaughn (1972), A computer program for the mixed analysis of variance model based on maximum likelihood, in T. A. Bancroft, ed., *Statistical Papers in Honor of George W. Snedecor*, Iowa State University Press, Ames, IA, 129–144.

H. O. Hartley, J. N. K. Rao, and G. Keifer (1969), Variance estimation with one unit per statum, *J. Amer. Statist. Assoc.*, **64**, 841–851.

H. O. Hartley, J. N. K. Rao, and L. R. LaMotte (1978), A simple synthesis-based method of estimating variance components, *Biometrics*, **34**, 233–243.

J. Hartung (1981), Non-negative minimum biased invariant estimation in variance components models, *Ann. Statist.*, **9**, 272–292.

J. Hartung and G. Knapp (2000), Confidence intervals for the between group variance in the unbalanced one-way random effects model, *J. Statist. Comput. Simul.*, **65**, 311–323.

T. D. Hartwell and D. W. Gaylor (1973), Estimating variance components for two-way disproportionate data with missing cells by the method of unweighted means, *J. Amer. Statist. Assoc.*, **68**, 379–383.

D. A. Harville (1969a), *Variance Component Estimation for the Unbalanced One-Way Random Classification*, Technical Report ARL-69-0180, Aerospace Research Laboratories, Wright-Patterson Air Force Base, Dayton, OH.

D. A. Harville (1969b), Variances of variance component estimators for the unbalanced two-way cross classification with application to balanced incomplete block designs, *Ann. Math. Statist.*, **40**, 408–416.

D. A. Harville (1969c), Expression of variance component estimators as linear combinations of independent non-central chi-square variates, *Ann. Math. Statist.*, **40**, 2189–2194.

D. A. Harville (1969d), Quadratic unbiased estimation of variance components for the one-way classification, *Biometrika*, **56**, 313–326.

D. A. Harville (1974), Bayesian inference for variance components using only error contrast, *Biometrika*, **61**, 383–385.

D. A. Harville (1977), Maximum-likelihood approaches to variance component estimation and to related problems, *J. Amer. Statist. Assoc.*, **72**, 320–340.

D. A. Harville (1988), Invariant inference for variance components, in J. N. Srivastava, ed., *Probability and Statistics: Essays in Honor of F. A. Graybill*, North-Holland, Amsterdam, 117–133.

D. A. Harville (1990), BLUP (best linear unbiased estimation) and beyond, in D. Gianola and K. Hammond, eds., *Advances in Statistical Methods for Genetic Improvement of Livestock*, Springer-Verlag, New York, 136–176.

D. A. Harville (1997), *Matrix Algebra from a Statistician's Perspective*, Springer-Verlag, New York.

D. A. Harville and T. P. Callanan (1990), Computational aspects of likelihood-based inference for variance components, in D. Gianola and K. Hammond, eds., *Advances in Statistical Methods for Genetic Improvement of Livestock*, Springer-Verlag, New York, 136–176.

D. A. Harville and A. L. Carriquiry (1992), Classical and Bayesian prediction as applied to an unbalanced mixed linear model, *Biometrics*, **48**, 987–1003.

D. A. Harville and A. P. Fenech (1985), Confidence intervals for a variance ratio, or for heritability, in an unbalanced mixed linear model, *Biometrics*, **41**, 137–152.

D. A. Harville and R. W. Mee (1984), A mixed model procedure for analyzing ordered categorical data, *Biometrics*, **40**, 393–408.

D. A. Harville and A. G. Zimmerman (1996), The posterior-distribution of the fixed and random effects in a mixed effects linear model, *J. Statist. Comput. Simul.*, **54**, 211–219.

B. I. Hayman (1954), The theory and analysis of diallel crosses, *Genetics*, **39**, 789–809.

B. I. Hayman (1960), Maximum likelihood estimation of genetic components of variation, *Biometrics*, **16**, 369–381.

M. J. R. Healy (2000), *Matrices for Statistics*, 2nd ed., Clarendon Press, Oxford, UK; 1st ed., 1986.

R. H. Heck and S. L. Thomas (1999), *An Introduction to Multilevel Modeling Techniques*, Lawrence Erlbaum, Mahwah, NJ.

C. E. Heckler and P. S. R. S. Rao (1985), Efficient estimation of variance components using staggered nested designs, in *Fall Technical Conference*, American Society for Quality Control, Corning, NY.

D. Hedecker and R. D. Gibbons (1994), A random effects ordinal regression model for multi-level analysis, *Biometrics*, **50**, 933–944.

B. Heine (1993), Nonnegative estimation of variance components in an unbalanced one-way random effects model, *Comm. Statist.* A *Theory Methods*, **22**, 2351–2372.

W. J. Hemmerle and H. O. Hartley (1973), Computing maximum likelihood for the mixed A.O.V. estimation using the *W* transformation, *Technometrics*, **15**, 819–831.

W. J. Hemmerle and J. A. Lorens (1976), Improved algorithm for the *W*-transformation in variance component estimation, *Technometrics*, **18**, 207–211.

C. R. Henderson (1953), Estimation of variance and covariance components, *Biometrics*, **9**, 226–252.

C. R. Henderson (1973), Sire evaluation and genetic trends, in *Proceedings of the Animal Breeding and Genetics Symposium in Honor of Dr. Jay L. Lush*, American Society of Animal Science and American Dairy Science Association, Champaign, IL, 10–41.

C. R. Henderson (1975), Best linear unbiased estimation and prediction under a selection model, *Biometrics*, **31**, 423–447.

C. R. Henderson (1984), *Applications of Linear Models in Animal Breeding*, University of Guelph, Guelph, ON, Canada.

C. R. Henderson (1985), MIVQUE and REML estimation of additive and non-additive genetic variances, *J. Animal Sci.*, **61**, 113–121.

C. R. Henderson (1986), Recent developments in variance and covariance estimation, *J. Animal Sci.*, **63**, 208–216.

C. R. Henderson, O. Kempthorne, S. R. Searle, and C. N. von Krosigk (1959), Estimation of environmental and genetic trends from records subject to culling, *Biometrics*, **15**, 192–218.

C. R. Henderson, S. R. Searle, and D. S. Robson (1957), *Seminar on Unbalanced Data*, mimeographed class notes, Cornell University, Ithaca, NY.

C. R. Henderson, S. R. Searle, and L. R. Schaffer (1974), The invariance and calculation of method 2 for estimating variance components, *Biometrics*, **30**, 583–588.

R. P. Hernández (1991), *Confidence Intervals on Linear Combinations of Variance Components in Unbalanced Designs*, Ph.D. dissertation, Arizona State University, Tempe, AZ.

R. P. Hernández and R. K. Burdick (1993a), Confidence intervals and tests of hypotheses on variance components in an unbalanced two-factor crossed design with interaction, *J. Statist. Comput. Simul.*, **47**, 67–77.

R. P. Hernández and R. K. Burdick (1993b), Confidence intervals on the total variance in an unbalanced two-fold nested design, *Biometrical J.*, **35**, 515–522.

R. P. Hernández, R. K. Burdick, and N. J. Birch (1992), Confidence intervals and tests of hypotheses on variance components in an unbalanced two-fold nested design, *Biometrical J.*, **34**, 387–402.

J. L. Hess (1979), Sensitivity of the MINQUE with respect to prior weights, *Biometrics*, **5**, 645–649.

H. O. Hetzer, G. E. Dickerson, and J. H. Zeller (1944), Heritability of type in Poland China swine as evaluated by scoring, *J. Animal Sci.*, **3**, 390–398.

B. M. Hill (1965), Inference about variance component in the one-way model, *J. Amer. Statist. Assoc.*, **60**, 806–825.

C. Hirotsu (1966), Estimating variance, components in a two-way layout with unequal numbers of observations, *Rep. Statist. Appl. Res.* (*JUSE*), **13**, 29–34.

C. Hirotsu (1968), An approximate test for the case of random effects model in a two-way layout with unequal cell frequencies, *Rep. Statist. Appl. Res.* (*JUSE*), **15**, 13–26.

J. P. Hobert and G. Casella (1996), The effect of improper priors on Gibbs sampling in hierarchical linear mixed models, *J. Amer. Statist. Assoc.*, **91**, 1461–1473.

R. R. Hocking (1985), *The Analysis of Linear Models*, Brooks–Cole, Monterey, CA.

R. R. Hocking, J. W. Green, and R. H. Bremer (1989), Variance component estimation with model-based diagnostics, *Technometrics*, **31**, 227–240.

R. R. Hocking and M. H. Kutner (1975), Some analytical and numerical comparisons of estimators for the mixed A. O. V. model, *Biometrics*, **31**, 19–28.

A. Hoefer (1998), Variance component estimation in animal breeding: A review, *J. Animal Breeding Genetics*, **115**, 247–265.

R. Horn and C. R. Johnson (1985), *Matrix Analysis*, Cambridge University Press, Cambridge, UK.

S. D. Horn and R. A. Horn (1975), Comparison of estimators of heteroscedastic variances in linear models, *J. Amer. Statist. Assoc.*, **70**, 872–879.

D. A. Huber, T. L. White, and G. R. Hodge (1994), Variance component estimation techniques compared for two mating designs with forest genetic architecture through computer simulation, *Theoret. Appl. Genetics*, **88**, 236–242.

H. M. Hudson (1983), Random effects in log-linear models, *J. Statist. Comput. Simul.*, **17**, 43–50.

R. A. Hultquist and F. A. Graybill (1965), Minimal sufficient statistics for the two-way classification mixed model design, *J. Amer. Statist. Assoc.*, **60**, 183–192.

M. Hussein and G. A. Milliken (1978a), An unbalanced two-way model with random effects having unequal variances, *Biometrical J.*, **20**, 203–213.

M. Hussein and G. A. Milliken (1978b), An unbalanced nested model with random effects having unequal variances, *Biometrical J.*, **20**, 329–338.

R. C. Jain and J. Singh (1989), Tests of hypotheses in two-way unbalanced random model, *Gujrat Statist. Rev.*, **16**, 43–46.

R. C. Jain, J. Singh, and R. Agrawal (1991), Interval estimation in the two-way nested unbalanced random model, *Trab. Estadíst.*, **6**-2, 33–40.

R. L. Jennrich and P. F. Sampson (1976), Newton–Raphson and related algorithms for maximum likelihood variance component estimation, *Technometrics*, **18**, 11–17.

S. Jeyaratnam and A. R. Othman (1985), Test of hypothesis in the one-way random effects model with unequal error variances, *J. Statist. Comput. Simul.*, **21**, 51–57.

J. Jiang (1996), REML estimation: Asymptotic behavior and related topics, *Ann. Statist.*, **24**, 255–286.

J. Jiang (1997), Wald consistency and the method of sieves in REML estimation, *Ann. Statist.*, **25**, 1781–1803.

N. L. Johnson and F. C. Leone (1964), *Statistics and Experimental Design in Engineering and the Physical Sciences*, 1st ed., Wiley, New York; 2nd ed., 1977,

A. C. Jones (1989), Likelihood based confidence intervals for variance components, in *ASA Proceedings of Statistical Graphics Section*, American Statistical Association, Alexandria, VA, 36–39.

R. Kala, A. Molińska, and K. Moliński (1990), Approximate confidence intervals for a ratio of two variance components, *Comm. Statist.* A *Theory Methods*, **19**, 1889–1898.

J. S. Kaplan (1983), A method for calculating MINQUE estimators of variance components, *J. Amer. Statist. Assoc.*, **78**, 476–477.

S. Karlin, P. E. Cameron, and P. Williams (1981), Sibling and parent-offspring correlation estimation with variable family age, *Proc. Nat. Acad. Sci.* (*U.S.A.*), **61**, 302–315.

R. M. Kasim and S. W. Raudenbush (1998), Application of Gibbs sampling to nested variance components models with heterogeneous within-group variance, *J. Educ. Behavioral Statist.*, **23**, 93–116.

M. K. Kazempour and F. A. Graybill (1989), Confidence bounds for proportion of σ_E^2 in unbalanced two-way crossed models, in Y. Dodge, ed., *Statistical Data Analysis*, North-Holland, Amsterdam, 389–396.

M. K. Kazempour and F. A. Graybill (1991), Approximate confidence bounds for ratios of variance components in two-way unbalanced crossed models with interactions, *Comm. Statist.* B *Simul. Comput.*, **20**, 955–967.

M. K. Kazempour and F. A. Graybill (1992), Confidence intervals on individual variances in two-way models, *Comput. Statist. Data Anal.*, **14**, 29–37.

R. J. Kelly and T. Mathew (1993), Improved estimators of variance components with smaller probability of negativity, *J. Roy. Statist. Soc. Ser.* B, **4**, 897–911.

R. J. Kelly and T. Mathew (1994), Improved non-negative estimation of variance components in some mixed models with unbalanced data, *Technometrics*, **36**, 171–181.

B. W. Kennedy (1991), C. R. Henderson: The unfinished legacy, *J. Dairy Sci.*, **74**, 4067–4081.

W. J. Kennedy, Jr. and J. E. Gentle (1980), *Statistical Computing*, Marcel Dekker, New York.

M. G. Kenward and J. H. Roger (1997), Small sample inference for fixed effects from restricted maximum likelihood, *Biometrics*, **53**, 983–997.

R. Khattree (1998), Some practical estimation procedures for variance components, *Comput. Statist. Data Anal.*, **28**, 1–32.

R. Khattree (1999), Nonnegative estimation of variance components: A modification to Henderson's ANOVA methodology, *Sankhyā Ser.* B, **61**, 261–265.

R. Khattree and D. S. Gill (1988), Comparison of some estimates of variance components using Pitman's nearness criterion, in *ASA Proceedings of Statistical Computing Section*, American Statistical Association, Alexandria, VA, 133–136.

R. Khattree and D. N. Naik (1990), Optimum tests for random effects in unbalanced nested designs, *Statistics*, **21**, 163–168.

R. Khattree and D. N. Naik (1995), Statistical tests for random effects in staggered nested designs, *J. Appl. Statist.*, **22**, 495–505.

R. Khattree, D. N. Naik, and R. L. Mason (1997), Estimation of variance components in staggered nested designs, *J. Appl. Statist.*, **24**, 395–408.

A. I. Khuri (1981), Simultaneous confidence intervals for functions of variance components in random models, *J. Amer. Statist. Assoc.*, **76**, 878–885.

A. I. Khuri (1987a), An exact test for the nesting effect's variance component in an unbalanced random two-fold nested model, *Statist. Probab. Lett.*, **5**, 305–311.

A. I. Khuri (1987b), Measures of imbalance for unbalanced data, *Biometrical J.*, **29**, 383–396.

A. I. Khuri (1990), Exact test for random models with unequal cell frequencies in the last stage, *J. Statist. Plann. Inference*, **24**, 177–193.

A. I. Khuri (1995), A test to detect inadequacy of Satterthwaite's approximation in balanced mixed models, *Statistics*, **27**, 45–54.

A. I. Khuri (1996), A method for determining the effect of imbalance, *J. Statist. Plann. Inference*, **55**, 115–129.

A. I. Khuri (1998), On unweighted sums of squares in unbalanced analysis of variance, *J. Statist. Plann. Inference*, **74**, 135–147.

A. I. Khuri and M. Ghosh (1990), Minimal sufficient statistics for the unbalanced two-fold nested model, *Statist. Probab. Lett.*, **10**, 351–353.

A. I. Khuri and R. C. Littell (1987), Exact tests for the main effects variance components in an unbalanced random two-way model, *Biometrics*, **43**, 545–560.

A. I. Khuri, T. Mathew, and B. K. Sinha (1998), *Statistical Tests for Mixed Linear Models*, Wiley, New York.

S. C. King and C. R. Henderson (1954), Variance component analysis in heritability studies, *Poultry Sci.*, **331**, 147–154.

J. Kleffé (1975), Best quadratic unbiased estimators for variance components in the balanced two-way classification model, *Math. Oper. Statist. Ser. Statist.*, **6**, 189–196.

J. Kleffé (1976), A note on MINQUE for normal models, *Math. Oper. Statist. Ser. Statist.*, **7**, 707–714.

J. Kleffé (1977a), Invariant methods for estimating variance components in mixed linear models, *Math. Oper. Statist. Ser. Statist.*, **8**, 233–250.

J. Kleffé (1977b), Optimal estimation of variance components: A survey, *Sankhyā Ser.* B, **39**, 211–244.

J. Kleffé (1980a), C. R. Rao's MINQUE under four two-way ANOVA models, *Biometrical J.*, **22**, 93–104.

J. Kleffé (1980b), On recent progress of MINQUE theory: Nonnegative estimation, consistency, asymptotic normality and explicit formulae, *Math. Oper. Statist. Ser. Statist.*, **11**, 563–588.

J. Kleffé (1993), Parent-offspring and sibling correlation estimation based on MINQUE theory, *Biometrika*, **80**, 393–404.

J. Kleffé and J. N. K. Rao (1986), The existence of asymptotically unbiased nonnegative quadratic estimates of variance components in ANOVA models, *J. Amer. Statist. Assoc.*, **81**, 692–698.

J. Kleffé and B. Seifert (1980), Matrix free computation of C. R. Rao's MINQUE for unbalanced nested classification models, *Comput. Statist. Data Anal.*, **2**, 215–228.

J. Kleffé and B. Seifert (1986), Computation of variance components by the MINQUE method, *J. Multivariate Anal.*, **18**, 107–116.

J. Klotz and J. Putter (1970), *Remarks on Variance Components: Likelihood Summits and Flats*, Technichal Report 247, Department of Statistics, University of Wisconsin, Madison, WI.

J. D. Knoke (1985), The method of unweighted means in unbalanced analysis of variance, *Biometrical J.*, **29**, 383–396

G. G. Koch (1967a), A general approach to the estimation of variance components, *Techometrics*, **9**, 93–118.

G. G. Koch (1967b), A procedure to estimate the population mean in random effects models, *Technometrics*, **9**, 577–586.

G. G. Koch (1968), Some further remarks concerning "A general approach to estimation of variance components," *Technometrics*, **10**, 551–558.

I. Kreft and J. deLeeuw (1998), *Introducing Multilevel Modeling*, Sage, Thousand Oaks, CA.

I. G. Kreft, J. deLeeuw, and R. van der Leeden (1994), Comparing five different statistical packages for hierarchical linear regression: BMDP-5V, GENMOD, HLM, ML3, and VARCL, *Amer. Statist.*, **48**, 324–335.

W. H. Kruskal (1968), Note to the Tietjen and Moore (1968) article, *Biometrics*, **24**, 1025–1026.

K. Kussmaul and R. L. Anderson (1967), Estimation of variance components in two-stage nested designs with composite samples, *Technometrics*, **9**, 373–389.

N. M. Laird (1982), Computation of variance components using the EM algorithm, *J. Statist. Comput. Simul.*, **14**, 295–303.

N. M. Laird, N. Lange, and D. Stram (1987), Maximum likelihood computation with repeated measures: Application of the EM algorithm, *J. Amer. Statist. Assoc.*, **82**, 97–105.

N. M. Laird and J. H. Ware (1982), Random-effects models for longitudinal data, *Biometrics*, **38**, 963–974.

J. L. Lamar and W. E. Zirk (1991), Nested designs in the chemical industry, in *ASQC Quality Congress Transaction*, American Society for Quality Control, Milwaukee, 615–622.

L. R. LaMotte (1973a), Quadratic estimation of variance components, *Biometrics*, **29**, 311–330.

L. R. LaMotte (1973b), On non-negative quadratic unbiased estimation of variance components, *J. Amer. Statist. Assoc.*, **68**, 728–730.

L. R. LaMotte, A. McWhorter, and R. A. Prasad (1988), Confidence intervals and tests on the variance ratio in random models with two variance components, *Comm. Statist.* A *Theory Methods*, **17**, 1135–1164.

H. O. Lancaster (1954), Traces and cumulants of quadratic forms in normal variables, *J. Roy. Statist. Soc. Ser.* B, **16**, 247–254.

J. R. Landis and G. G. Koch (1977), A one-way component of variance model for categorical data, *Biometrics*, **33**, 671–679.

H. S. Lee (1997), Random effects model for count data, *Comm. Statist.* A *Theory Methods*, **26**, 1893–1904.

J. Lee and A. I. Khuri (1999), Graphical technique for comparing designs for random models, *J. Appl. Statist.*, **26**, 933–947.

J. Lee and A. I. Khuri (2000), Quantile dispersion graphs for the comparison of designs for a random two-way model, *J. Statist. Plann. Inference*, **91**, 123–137.

J. Lee and A. I. Khuri (2001), Modeling the probability of a negative ANOVA estimate of a variance component, *Calcutta Assoc. Bull.*, **51**, 201–202.

J. T. Lee (1993), Efficiency of MINQE for arbitrary distribution under one-way random effects model, *Korean J. Appl. Statist.*, **6**, 355–37.

J. T. Lee and B. C. Kim (1989), An efficient method for computing MINQUE estimators in the mixed models, *J. Korean Statist. Soc.*, **18**, 4–12.

K. R. Lee and C. H. Kapadia (1991), Restricted maximum likelihood estimation under Eisenhart model III, *Statist. Neerland.*, **45**, 271–282.

K. R. Lee and C. H. Kapadia (1992), Mean squared error efficient estimators of the variance components, *Metrika*, **39**, 21–26.

P. M. Lee (1998), *Bayesian Statistics*, 2nd ed., Arnold, London.

Y. Lee and J. Seely (1996), Computing the Wald interval for a variance ratio, *Biometrics*, **52**, 1486–1491.

E. L. Lehmann (1986), *Testing Statistical Hypothesis*, 2nd ed., Wiley, New York; reprint, 1997, Springer-Verlag, New York.

F. C. Leone, L. S. Nelson, N. L. Johnson, and S. Eisenstat (1968), Sampling distributions of variance components II: Empirical studies of unbalanced nested designs, *Technometrics*, **10**, 719–738.

M. L. Lera (1994), Measures of imbalance for higher-order designs, *Biometrical J.*, **36**, 481–490.

C. Y. Lin and A. J. McAllister (1984), Monte Carlo comparison of four methods for estimation of genetic parameters in the univariate case, *J. Dairy Sci.*, **67**, 2389–2398.

T.-H. Lin and D. A. Harville (1991), Some alternatives to Wald's confidence interval and test, *J. Amer. Statist. Assoc.*, **86**, 179–187.

X. Lin (1997), Variance components in generalized linear models with random effects, *Biometrika*, **34**, 309–326.

D. V. Lindley and A. F. M. Smith (1972), Bayes estimates for the linear model (with discussion), *J. Roy. Statist. Soc. Ser.* B, **34**, 1–41.

M. J. Lindstrom and D. M. Bates (1988), Newton–Raphson and EM algorithms for linear mixed effects models for repeated measures data, *J. Amer. Statist. Assoc.*, **83**, 1014–1022.

R. C. Littell, G. A. Milliken, W. W. Stroup, and R. D. Wolfinger (1996), *SAS System for Mixed Models*, SAS Institute, Cary, NC.

L.-M. Liu and J. Senturia (1975), *The Distribution of MINQUE Estimators*, Technical Report 406, Department of Statistics, University of Wisconsin, Madison, WI.

L.-M. Liu and J. Senturia (1977), Computation of MINQUE variance components estimates, *J. Amer. Statist. Assoc.*, **72**, 867–868.

N. T. Longford (1987), A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects, *Biometrika*, **74**, 817–827.

N. T. Longford (1988), *VARCL: Software for Variance Component Analysis of Data with Hierarchically Nested Random Effects* (*Maximum Likelihood*), Educational Testing Service, Princeton, NJ.

N. T. Longford (1993), *Random Coefficient Models*, Oxford University Press, New York.

N. T. Longford (1995), Random coefficient models, in G. Arminger, C. C. Clogg, and M. E. Sobel, eds., *Handbook of Statistical Modeling for the Social and Behavioral Sciences*, Plenum Press, London, 519–577.

L. Y. Low (1964), Sampling variances of estimates of components of variance from a non-orthogonal two-way classification, *Biometrika*, **51**, 491–494.

L. Y. Low (1976), Some properties of variance component estimators, *Technometrics*, **18**, 39–46.

W. Madow (1940), The distribution of quadratic forms in noncentral normal random variables, *Ann. Math. Statist.*, **11**, 100–101.

J. R. Magnus and H. Neudecker (1999), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 2nd ed., Wiley, Chichester, UK.

D. M. Mahamunulu (1963), Sampling variances of the estimates of variance components in the unbalanced three-way nested classification, *Ann. Math. Statist.*, **34**, 521–527.

J. D. Malley (1986), *Optimal Unbiased Estimation of Variance Components*, Lecture Notes in Statistics 39, Springer-Verlag, New York.

R. L. Mason, R. F. Gunst, and J. L. Hess (1989), *Statistical Design and Analysis of Experiments with Applications to Engineering and Science*, Wiley, New York.

W. M. Mason, A. F. Anderson, and N. Hayat (1988), *Manual for GENMOD*, Population Studies Center, University of Michigan, Ann Arbor, MI.

H. Massam and J. Muller (1985), The nonnegative MINQUE estimate, *J. Multivariate Anal.*, **16**, 253–259.

T. Mathew (1984), On nonnegative quadratic unbiased estimability of variance components, *Ann. Statist.*, **12**, 1566–1569.

T. Mathew (1987), On the characterization of nonnegative estimable linear combinations of variance components, *Statistics*, **18**, 227–236.

T. Mathew and B. K. Sinha (1988), Optimum tests in unbalanced two-way models without interaction, *Ann. Statist.*, **16**, 1727–1740.

T. Mathew, B. K. Sinha, and B. C. Sutradhar (1992a), Nonnegative estimation of variance components in unbalanced mixed models with two variance components, *J. Multivariate Anal.*, **42**, 77–101.

T. Mathew, B. K. Sinha, and B. C. Sutradhar (1992b), Nonnegative estimation of variance components in general balanced mixed models, in A. Md. E. Saleh, ed., *Nonparametric Statistics and Related Topics*, Elsevier, New York, 281–295.

P. McCullagh and J. Nelder (1989), *Generalized Linear Models*, 2nd ed., Chapman and Hall, London.

C. E. McCulloch and S. R. Searle (2001), *Generalized, Linear, and Mixed Models*, Wiley, New York.

B. W. McDonald (1994), Two random effects models for multivariate binary data, *Biometrics*, **50**, 164–172.

G. J. McLachlan and T. Krishnan (1996), *The EM Algorithm and Extensions*, Wiley, New York.

R. A. McLean and W. L. Saunders (1988), Approximating degrees of freedom for standard errors in mixed linear models, in *ASA Proceedings of Statistical Computing Section*, American Statistical Association, Arlington, VA, 50–59.

I. U. H. Mian, M. M. Shoukri, and D. S. Tracy (1989), A comparison of significance testing procedures for the intraclass correlation from family data, *Comm. Statist.* B *Simul. Comput.*, **18**, 613–631.

J. J. Miller (1977), Asymptotic properties of maximum likelihood estimates in the mixed model of the analysis of variance, *Ann. Statist.*, **5**, 746–762.

J. J. Miller (1979), Maximum likelihood estimation of variance components (a Monte Carlo study), *J. Statist. Comput. Simul.*, **8**, 175–190.

G. A. Milliken and D. E. Johnson (1992), *Analysis of Messy Data*, Vol. I, Chapman and Hall, London.

S. K. Mitra (1972), Another look at Rao's MINQUE of variance components, *Bull. Internat. Statist. Inst.*, **44**, 279–283.

C. H. Morrel (1998), Likelihood ratio testing of variance components in the linear mixed-effects model using restricted maximum likelihood, *Biometrics*, **54**, 1560–1568.

C. N. Morris (1995), Hierarchical models for educational data: An Overview, *J. Educ. Statist.*, **20**, 190–200.

M. G. Mostafa (1967), Note on testing hypotheses in an unbalanced random effects model, *Biometrika*, **54**, 659–662.

R. Mukerjee (1990), Optimality of balanced designs for minimum norm quadratic unbiased estimation of variance components, *Metrika*, **37**, 181–188.

R. Mukerjee and S. Huda (1988), Optimal design for the estimation of variance components, *Biometrika*, **75**, 75–80.

H. D. Muse (1974), *Comparison of Designs to Estimate Variance Components in a Two-Way Classification Model*, Ph.D. dissertation, University of Kentucky, Lexington, KY.

H. D. Muse and R. L. Anderson (1978), Comparison of designs to estimate variance components in a two-way classification model, *Technometrics*, **20**, 159–166.

H. D. Muse, R. L. Anderson, and B. Thitakamol (1982), Additional comparisons of designs to estimate variance components in a two-way classification model, *Comm. Statist.* A *Theory Methods*, **11**, 1403–1425.

D. N. Naik and R. Khattree (1992), Optimum tests for treatments when blocks are nested and random, *Statistics*, **23**, 101–108.

D. N. Naik and R. Khattree (1998), A computer program to estimate variance components in staggered nested designs, *J. Quality Tech.*, **30**, 292–297.

K. K. Namboodiri, P. P. Green, E. B. Kaplan, J. A. Morrison, G. A. Chase, R. C. Elston, A. R. G. Owen, B. M. Rifkind, C. J. Glueck, and H. A. Tyroler (1984), The collaborative lipid research clinics program: Family study IV: Familial associations of plasma lipids and lipoproteins, *Amer. J. Epidemiology*, **119**, 975–996.

L. S. Nelson (1983), Variance estimation using staggered nested designs, *J. Quality Tech.*, **15**, 195–198.

L. S. Nelson (1995a), Using nested designs I: Estimation of standard deviations, *J. Quality Tech.*, **27**, 169–171.

L. S. Nelson (1995b), Using nested designs II: Confidence limits for standard deviations, *J. Quality Tech.*, **27**, 265–267.

Y. Ochi and R. L. Prentice (1984), Likelihood inference in a correlated probit regression model, *Biometrika*, **71**, 531–543.

Y. Ojima (1984), The use of canonical forms for estimating variance components in unbalanced nested designs, *Rep. Statist. Appl. Res.* (*Japan*), **31**, 1–18.

Y. Ojima (1998), General formula for expectations, variances and covariances of the mean squares for staggered nested designs, *J. Appl. Statist.*, **25**, 785–799.

I. Olkin and J. W. Pratt (1958), Unbiased estimation of certain correlation coefficients, *Ann. Math. Statist.*, **29**, 201–211.

A. Olsen, J. Seely, and D. Birkes (1976), Invariant quadratic unbiased estimation for two variance components, *Ann. Statist.*, **4**, 878–890.

Y. Omori (1997), Comparing two means in count models having random effects, *Statist. Probab. Lett.*, **31**, 225–235.

J. Öfversten (1993), Exact tests for variance components in unbalanced mixed linear models, *Biometrics*, **49**, 45–57.

A. R. Othman (1983), *One-Way Random Effects Model with Unequal Error Variances*, M.S. thesis, Southern Illinois University, Carbondale, IL.

H. D. Patterson and R. Thompson (1971), Recovery of inter-block information when block sizes are unequal, *Biometrika*, **58**, 545–554.

H. D. Patterson and R. Thompson (1975), Maximum likelihood estimation of components of variance, in L. C. A. Corsten and T. Postelnicu, eds., *Proceedings of the* 8*th International Biometrics Conference*, Editoria Academie, Constanta, Romania, 197–207.

K. Pearson (1934), *Tables of the Incomplete Beta Function*, Cambridge University Press, Cambridge, UK.

D. G. Pederson (1972), A comparison of four experimental designs for the estimation of heritability, *Theoret. Appl. Genetics*, **42**, 371–377.

A. N. Pettitt and A. B. McBratney (1993), Sampling designs for estimating spatial variance components, *Appl. Statist.*, **42**, 185–209.

D. A. Pierce and D. Peters (1992), Practical use of higher order asymptotics for multiparameter exponential families, *J. Roy. Statist. Soc. Ser.* B, **54**, 701–725.

K. N. Ponnuswamy and J. Subramani (1987), A simplified procedure for calculating MINQUE of variance components, in *Modeling and Simulation: Proceedings of the International AMSA Conference*, Vol. B, ASMA, New Delhi, 75–90.

B. J. Postma and J. S. White (1975), Variance components for unbalanced N-level hierarchic design, *J. Quality Tech.*, **7**, 144–149; corrigenda, **8**, 208.

R. R. Prairie (1962), *Optimal Designs to Estimate Variance Components and Reduce Product Variability for Nested Classifications*, Ph.D. dissertation, Institute of Statistics, North Carolina State University, Raleigh, NC.

R. R. Prairie and R. L. Anderson (1962), *Optimal Designs to Estimate Variance Components and to Reduce Product Variability for Nested Classifications*, Mim. Ser. 313, Institute of Statistics, North Carolina State University, Raleigh, NC.

N. G. N. Prasad and J. N. K. Rao (1988), Robust test and confidence intervals for error variance in a regression model and for functions of variance components in an unbalanced random one-way model, *Comm. Statist.* A *Theory Methods*, **17**, 1111–1133.

S. J. Press (1966), Linear combinations of non-central chi-square variates, *Ann. Math. Statist.*, **37**, 480–487.

R. Pringle and A. Raynor (1971), *Generalized Inverse Matrices with Applications in Statistics*, Hafner, New York.

R. M. Pringle (1974), Some results on the estimation of variance components by MINQUE, *J. Amer. Statist. Assoc.*, **69**, 987–989.

R. Prosser, J. Rasbash, and H. Goldstein (1991), *ML*3 *Software for Three-Level Analysis: User's Guide for Version* 2, Institute of Education, University of London, London.

F. Pukelsheim (1974), *Schätzen von Mittelwert und Streuungsmatrix in Gauss–Markov Modellen*, Diplomarbeit, Albert-Ludwigs-Universität, Freiburg im Breisgau, Germany.

F. Pukelsheim (1976), Estimating variance components in linear models, *J. Multivariate Anal.*, **6**, 626–629.

F. Pukelsheim (1977a), *Linear Models and Convex Programs: Unbiased Nonnegative Estimation in Variance Components Models*, Technical Report 104, Department of Statistics, Stanford University, Stanford, CA.

F. Pukelsheim (1977b), *Examples for Unbiased Nonnegative Estimation in Variance Components Models*, Technical Report 113, Department of Statistics, Stanford University, Stanford, CA.

F. Pukelsheim (1977c), *Punktschätzungen in linearen statistischen Modellen*, Dissertation zur Erlangung des Doktorgrades der Mathematischen, Albert-Ludwigs-Universität, Freiburg im Breisgau, Germany.

F. Pukelsheim (1979), Classes of linear models, in L. D. Van Vleck and S. R. Searle, eds., *Variance Components and Animal Breeding: Proceedings of a Conference in Honor of C. R. Henderson*, Cornell University, Ithaca, NY, 69–83.

F. Pukelsheim (1981a), On the existence of unbiased nonnegative estimates of variance-covariance components, *Ann. Statist.*, **9**, 293–299.

F. Pukelsheim (1981b), Linear models and convex geometry: Aspects of nonnegative variance estimation, *Math. Oper. Statist. Ser. Statist.*, **12**, 271–286.

P. Pulley, Jr. (1957), *A Program for the Analysis of Variance of Hierarchal Classification Design*, M.S. thesis, Oklahoma State University, Stillwater, OK.

R. L. Quaas and D. C. Bolgiano (1979), Sampling variances of the MINQUE and method 3 estimators of sire components of variance: A numerical comparison, in L. D. Van Vleck and S. R. Searle, eds., *Variance Components and Animal Breeding: Proceedings of a Conference in Honor of C. R. Henderson*, Cornell University, Ithaca, NY, 99–106.

M. Rajagopalan (1980), *Bayesian Inference for the Variance Components in Mixed Linear Models*, Ph.D. dissertation, Oklahoma State University, Stillwater, OK.

M. Rajagopalan and L. D. Broemeling (1983), Bayesian inference for the variance components of the general mixed model, *Comm. Statist.* A *Theory Methods*, **12**, 701–723.

C. R. Rao (1970), Estimation of heterocedastic variances in linear models, *J. Amer. Statist. Assoc.*, **65**, 161–172.

C. R. Rao (1971a), Estimation of variance and covariance components: MINQUE theory, *J. Multivariate Anal.*, **1**, 257–275.

C. R. Rao (1971b), Minimum variance quadratic unbiased estimation of variance components, *J. Multivariate Anal.*, **1**, 445–456.

C. R. Rao (1972), Estimation of variance and covariance components in linear models, *J. Amer. Statist. Assoc.*, **67**, 112–115.

C. R. Rao (1973), *Linear Statistical Inference and Its Applications*, 2nd ed., Wiley, New York.

C. R. Rao (1974), Projections, generalized inverse and the BLUE, *J. Roy. Statist. Soc. Ser.* B, **36**, 442–448.

C. R. Rao (1979), MINQUE theory and its relation to ML and RML estimation of variance components, *Sankhyā Ser.* B, **41**, 138–153.

C. R. Rao and J. Kleffé (1980), Estimation of variance components, in P. R. Krishnaiah, ed., *Handbook of Statistics*, Vol. 1, North-Holland, Amsterdam, 1–40.

C. R. Rao and J. Kleffé (1988), *Estimation of Variance Components and Applications*, North-Holland, Amsterdam.

C. R. Rao and S. K. Mitra (1971), *Generalized Inverse of Matrices and Its Applications*, Wiley, New York.

C. R. Rao and M. B. Rao (1998), *Matrix Algebra and Its Applications to Statistics and Econometrics*, World Scientific, Singapore.

J. N. K. Rao (1968), On expectations, variances, and covariances of ANOVA mean squares by "synthesis," *Biometrics*, **24**, 963–978.

J. N. K. Rao (1973), On the estimation of heterocedastic variances, *Biometrics*,

29, 11–24.

J. N. K. Rao (1977), Contribution to the discussion of the paper by D. A. Harville, *J. Amer. Statist. Assoc.*, **72**, 338–339.

J. N. K. Rao and K. Subrahmaniam (1971), Combining independent estimators and estimation in linear regression with unequal variances, *Biometrics*, **27**, 971–990.

P. S. R. S. Rao (1977), Theory of the MINQUE: A review, *Sankhyā Ser.* B, **39**, 201–210.

P. S. R. S. Rao (1982), Use of prior information for estimating the variance components, *Comm. Statist.* A *Theory Methods*, **11**, 1659–1669.

P. S. R. S. Rao (1997), *Variance Components Estimation: Mixed Models, Methodologies, and Applications*, Chapman and Hall, London.

P. S. R. S. Rao (2001a), The MINQUE: Investigations, modifications and extensions, *J. Statist. Res.*, **35**-1, 29–34.

P. S. R. S. Rao (2001b), Nonnegative estimators for the one-way random effects model, *Comm. Statist.* A *Theory Methods*, **30**, 1605–1613.

P. S. R. S. Rao and Y. B. Chaubey (1978), Three modifications of the principle of MINQUE, *Comm. Statist.* A *Theory Methods*, **7**, 767–778.

P. S. R. S. Rao and C. E. Heckler (1997), Estimators for the three-fold nested random effects model, *J. Statist. Plann. Inference*, **64**, 341–352.

P. S. R. S. Rao, J. Kaplan, and W. G. Cochran (1981), Estimators for the one-way random effects model with unequal error variances, *J. Amer. Statist. Assoc.*, **76**, 89–97.

P. S. R. S. Rao and P. Kuranchie (1988), Variance components of linear regression model with a random intercept, *Comm. Statist.* A *Theory Methods*, **17**, 1011–1026.

P. S. R. S. Rao and N. Miyawaki (1989), Estimation procedures for unbalanced one-way variance components model, *Bull. Internat. Statist. Inst.*, **2**, 243–244.

P. S. R. S. Rao and E. A. Sylvestre (1984), ANOVA and MINQUE type of estimators for the one-way random effects model, *Comm. Statist.* A *Theory Methods*, **13**, 1667–1673.

J. Rasbash, M. Yang, G. Woodhouse, and H. Goldstein (1995), *Mln Command Reference*, Institute of Education, University of London, London.

S. Raudenbush and A. S. Bryk (1986), A hierarchical model for studying school effects, *Sociol. Educ.*, **59**, 1–17.

S. Raudenbush and A. S. Bryk (2002), *Hierarchical Linear Models: Applications and Data Analysis Methods*, 2nd ed., Sage, Thousand Oaks, CA.

R. R. Read (1961), On quadratic estimates of intraclass variance for unbalanced designs, *J. Roy. Statist. Soc. Ser.* B, **23**, 493–497.

C. A. Rhode and G. M. Tallis (1969), Exact first- and second-order moments of estimates of components of covariance, *Biometrika*, **56**, 517–525.

D. K. Rich and K. G. Brown (1979), Estimation of variance components using residuals: Some empirical evidence, in in L. D. Van Vleck and S. R. Searle, eds., *Variance Components and Animal Breeding: Proceedings of a Confer-*

ence in Honor of C. R. Henderson, Cornell University, Ithaca, NY, 139–154.

A. M. Richardson and A. H. Welsh (1994), Asymptotic properties of restricted maximum likelihood (REML) estimates for hierarchical mixed linear models, *Australian J. Statist.*, **36**, 31–43.

A. M. Richardson and A. H. Welsh (1995), Robust restricted maximum likelihood in mixed linear models, *Biometrics*, **51**, 1429–1439.

A. Robertson (1962), Weighting in the estimation of variance components in the unbalanced single classification, *Biometrics*, **18**, 413–417.

D. L. Robinson (1984), *The REML Program Manual*, unpublished report; available from the author.

D. L. Robinson (1987), Estimation and use of variance components, *Statistician*, **36**, 3–14.

G. K. Robinson (1991), The BLUP is a good thing: The estimation of random effects, *Statist. Sci.*, **6**, 15–51.

J. Robinson (1965), The distribution of general quadratic form in normal variables, *Australian J. Statist.*, **7**, 110–114.

D. M. Rocke (1983), Robust statistical analysis of interlaboratory studies, *Biometrika*, **70**, 421–431.

S. H. Rosenberg and C. A. Rhode (1971), Robustness properties of variance component estimates when the underlying model is incorrectly specified, *Biometrics*, **27**, 1091–1097.

J. S. Rosenthal (1995), Rates of convergence for Gibbs sampling for variance components models, *Ann. Statist.*, **23**, 740–761.

B. Rosner (1982), Statistical methods in ophthalmology: An adjustment for the intraclass correlation between eyes, *Biometrics*, **38**, 105–114.

J. W. Rudan and S. R. Searle (1971), Large sample variances of maximum likelihood estimators of variance components in the three-way nested classification, random model, with unbalanced data, *Biometrics*, **27**, 1087–1091.

P. Rudolph (1976), Bayesian estimation in the linear models under different assumptions about the covariance structure; variance components model, equicorrelation model, *Math. Oper. Statist. Ser. Statist.*, **7**, 649–665.

J. Sánchez (1983), Sample variances for several MINQ-estimators in the one-way classification, *Biometrical J.*, **25**, 259–267.

SAS Institute (1990), *SAS Language: Reference, Version* 6, 1st ed., SAS Institute, Cary, NC.

SAS Institute (2001), *SAS/STAT User's Guide: Version* 8.2.0, Vols. I and II, 4th ed., SAS Institute, Cary, NC.

F. E. Satterthwaite (1941), Synthesis of variance, *Psychometrika*, **6**, 309–316.

F. E. Satterthwaite (1946), An approximate distribution of estimates of variance components, *Biometrics Bull.*, **2**, 110–114.

L. R. Schaeffer (1973), *Factors Affecting the Accuracy of Variance Components Estimation*, Ph.D. dissertation, Cornell University, Ithaca, NY.

R. Schall (1991), Estimation in generalized linear models with random effects, *Biometrika*, **78**, 719–727.

H. Scheffé (1959), *The Analysis of Variance*, Wiley, New York.

M. L. Schervish (1992), Bayesian analysis of linear models, in J. M. Bernardo, J. V. Berger, A. P. David, and A. F. M. Smith, eds., *Bayesian Statistics* IV, Oxford University Press, New York, 419–434.

J. R. Schott (1997), *Matrix Analysis for Statistics*, Wiley, New York.

J. H. Schwartz (1980), *Optimal Allocation of Resources to Reduce Product or Process Variability: A Comparison of Some Designs to Estimate the Allocation Parameters and the Variance Components*, Ph.D. dissertation, University of Kentucky, Lexington, KY.

S. R. Searle (1956), Matrix methods in components of variance and covariance analysis, *Ann. Math. Statist.*, **27**, 737–748.

S. R. Searle (1958), Sampling variances of estimates of components of variance, *Ann. Math. Statist.*, **29**, 167–178.

S. R. Searle (1961), Variance components in the unbalanced two-way nested classification, *Ann. Math. Statist.*, **32**, 1161–1166.

S. R. Searle (1967), *Computer Simulation of Variance Component Estimates*, Mim. Ser. Paper BU-233-M, Biometrics Unit, Cornell University, Ithaca, NY.

S. R. Searle (1968), Another look at Henderson's methods of estimating variance components, *Biometrics*, **24**, 749–778.

S. R. Searle (1970), Large sample variances of maximum likelihood estimators of variance components using unbalanced data, *Biometrics*, **26**, 505–525.

S. R. Searle (1971a), Topics in variance components estimation, *Biometrics*, **27**, 1–76.

S. R. Searle (1971b), *Linear Models*, Wiley, New York.

S. R. Searle (1978), A univariate formulation of the multivariate linear model, in H. A. David, ed., *Contribution to Survey Sampling and Applied Statistics: Papers in Honor of H. O. Hartley*, Academic Press, San Diego, 181–189.

S. R. Searle (1982), *Matrix Algebra Useful for Statistics*, Wiley, New York.

S. R. Searle (1987), *Linear Models for Unbalanced Data*, Wiley, New York.

S. R. Searle (1988), Mixed models and unbalanced data: Wherefrom, whereat, whereto?, *Comm. Statist.* A *Theory Methods*, **17**, 935–968.

S. R. Searle, G. Casella, and C. E. McCulloch (1992), *Variance Components*, Wiley, New York.

S. R. Searle and C. R. Henderson (1961), Computing procedures for estimating components of variance in the two-way classification, mixed model, *Biometrics*, **17**, 607–616; corrigenda, 1967, **23**, 852.

J. Seely (1970a), Linear spaces and unbiased estimation, *Ann. Math. Statist.*, **41**, 1725–1734.

J. Seely (1970b), Linear spaces and unbiased estimation: Applications to the mixed linear model, *Ann. Math. Statist.*, **41**, 1735–1748.

J. Seely (1971), Quadratic subspaces and complements, *Ann. Math. Statist.*, **42**, 710–721.

J. Seely (1972), Completeness for a family of multivariate normal distributions, *Ann. Math. Statist.*, **43**, 1644–1647.

J. Seely (1977), Minimal sufficient statistics and completeness for multivariate normal families, *Sankhyā Ser.* A, **39**, 170–185.

J. F. Seely and M. Y. El-Bassiouni (1983), Applying Wald's variance component test, *Ann. Statist.*, **11**, 197–201.

S. G. Self and K.-Y. Liang (1987), Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions, *J. Amer. Statist. Assoc.*, **82**, 605–610.

B. Sen (1988), *Confidence Intervals on Ratios of Variance Components in the Three-Stage Nested Unbalanced Model*, Ph.D. dissertation, Colorado State University, Fort Collins, CO.

B. Sen, F. A. Graybill, and N. Ting (1992), Confidence intervals on ratios of variance components for the unbalanced two-factor nested model, *Biometrical J.*, **3**, 259–274.

E. Seneta (1981), *Non-Negative Matrices and Markov Chains*, 2nd ed., Springer-Verlag, New York.

R. H. Sharpe and C. H. van Middelem (1955), Application of variance components to horticultural problems with special reference to a parathion residue study, *Proc. Amer. Soc. Horticultural Sci.*, **66**, 415–420.

P.-S. Shen, P. L. Cornelius, and R. L. Anderson (1996a), Planned unbalanced designs for estimation of quantitative genetic parameters I: Two-way matings, *Biometrics*, **52**, 56–70.

P.-S. Shen, P. L. Cornelius, and R. L. Anderson (1996b), Planned unbalanced designs for estimation of quantitative genetic parameters II, *J. Agricultural Biol. Environ. Sci.*, **1**, 490–505.

M. M. Shoukri and R. H. Ward (1984), On the estimation of the intraclass correlation, *Comm. Statist.* A *Theory Methods*, **13**, 1239–1258.

B. Singh (1987), On the non-null distribution of ANOVA F-ratio in one-way unbalanced random model, *Calcutta Statist. Assoc. Bull.*, **36**, 57–62.

B. Singh (1989a), A comparison of variance component estimators under unbalanced situations, *Sankhyā Ser.* B, **51**, 323–330.

B. Singh (1989b), Moments of group variance estimator in one-way unbalanced classfication, *J. Indian Soc. Agricultural Statist.*, **41**, 77–84.

B. Singh (1991), Probability of negative estimates of heritability in one-way unbalanced random model, *J. Indian Soc. Agricultural Statist.*, **43**, 74–80.

B. Singh (1992), On effect of unbalancedness and heteroscedasticity on the ANOVA estimator of group variance component in one-way random model, *Biometrical J.*, **34**, 91–96.

R. A. Singhal (1984), Effect of non-normality on the estimation of functions of variance components, *J. Indian Soc. Agricultural Statist.*, **35**, 89–98.

C. A. B. Smith (1956), On the estimation of intraclass correlation, *Ann. Human Genetics*, **21**, 363–373.

C. A. B. Smith (1980a), Estimating genetic correlations, *Ann. Human Genetics*, **43**, 265–284.

C. A. B. Smith (1980b), Further remarks on estimating genetic correlations, *Ann. Human Genetics*, **44**, 95–105.

D. W. Smith and L. W. Murray (1984), An alternative to Eisenhart's Model II and mixed model in the case of negative variance estimates, *J. Amer. Statist. Assoc.*, **79**, 145–151.

H. F. Smith (1936), The problem of comparing the results of two experiments with unequal errors, *J. Council Sci. Indust. Res.* (*U.K.*), **9**, 211–212.

J. R. Smith and J. M. Beverly (1981), The use and analysis of staggered nested factorial designs, *J. Quality Tech.*, **13**, 166–173.

G. W. Snedecor and W. G. Cochran (1989), *Statistical Methods*, 8th ed., Iowa State University Press, Ames, IA; 6th ed., 1967; 7th ed., 1980.

R. D. Snee (1974), Computation and use of expected mean squares in analysis of variance, *J. Quality Tech.*, **6**, 128–137.

R. D. Snee (1983), Graphical analysis of process variation studies, *J. Quality Tech.*, **15**, 76–88.

R. R. Sokal and F. J. Rohlf (1995), *Biometry*, 3rd ed., W. H. Freeman, New York; 1st ed., 1969; 3rd ed., 1981.

P. J. Solomon and D. R. Cox (1992), Nonlinear component of variance models, *Biometrika*, **79**, 1–11.

D. J. Spiegelhalter, A. Thomas, N. Best, and W. R. Gilks (1995a), *BUGS: Bayesian Inference Using Gibbs Sampling, Version* 0.50, Technical Report, Medical Research Council, Biostatistics Unit, Institute of Public Health, Cambridge University, Cambridge, UK.

D. J. Spiegelhalter, A. Thomas, N. Best, and W. R. Gilks (1995b), *BUGS Examples, Version* 0.50, Technical Report, Medical Research Council, Biostatistics Unit, Institute of Public Health, Cambridge University, Cambridge, UK.

E. Spjøtvoll (1967), Optimum invariant tests in unbalanced variance components models, *Ann. Math. Statist.*, **38**, 422–428.

E. Spjøtvoll (1968a), Most powerful tests for some non-exponential families, *Ann. Math. Statist.*, **39**, 772–784.

E. Spjøtvoll (1968b), Confidence intervals and tests for variance ratios in unbalanced variance components models, *Rev. Internat. Statist. Inst.*, **36**, 37–42.

SPSS, Inc. (2001), *SPSS Advanced Statistics, Release* 11.0., SPSS, Chicago.

S. Srinivasan (1986), *Confidence Intervals on Functions of Variance Components in Unbalanced Two-Way Design Models*, Ph.D. dissertation, Colorado State University, Fort Collins, CO.

S. Srinivasan and F. A. Graybill (1991), Confidence intervals for proportions of total variation in unbalanced two-way components of variance models using unweighted means, *Comm. Statist.* A *Theory Methods*, **20**, 511–526.

R. Stiratelli, N. M. Laird, and J. H. Ware (1984), Random-effects models for serial observations with binary response, *Biometrics*, **40**, 961–971.

D. O. Stram and J. W. Lee (1994), Variance components testing in the longitudinal mixed effects model, *Biometrics*, **50**, 1171–1177.

W. W. Stroup, J. W. Evans, and R. L. Anderson (1980), Maximum likelihood estimation of variance components in a completely random BIB design, *Comm. Statist.* A *Theory Methods*, **9**, 725–756.

W. W. Stroup and S. D. Kachman (1994), Generalized linear mixed models: An overview, in J. R. Schewenke, ed., *Proceedings of the* 1994 *Kansas State*

University Conference on Applied Statistics in Agriculture, Kansas State University, Manhattan, KS, 82–98.

J. Subramani (1991), On invariant quadratic unbiased estimation of variance components, *Comm. Statist.* A *Theory Methods*, **20**, 1705–1730.

L. Sun, J. S. J. Hsu, I. Guttman, and T. Leonard (1996), Bayesian methods for variance component models, *J. Amer. Statist. Assoc.*, **91**, 743–752.

W. H. Swallow (1974), *Minimum Norm and Minimum Variance Quadratic Unbiased Estimation of Variance Components*, Ph.D. dissertation, Cornell University, Ithaca, NY.

W. H. Swallow (1981), Variances of locally minimum variance quadratic unbiased estimators (MINQUE) of variance components, *Technometrics*, **23**, 271–283.

W. H. Swallow and J. F. Monaham (1984), Monte Carlo comparison of ANOVA4, MINQUE, REML, and ML estimators of variance components, *Technometrics*, **26**, 47–57.

W. H. Swallow and S. R. Searle (1978), Minimum variance quadratic unbiased estimation (MINQUE) of variance components, *Technometrics*, **20**, 265–272.

L. A. Swiger, L. R. Harvey, D. O. Everson, and K. E. Gregory (1964), The variance of intraclass correlation involving groups with one observation, *Biometrics*, **20**, 818–826.

W. Y. Tan and S. S. Cheng (1984), On testing variance components in three-stage unbalanced nested random effects models, *Sankhyā Ser.* B, **46**, 188–200.

W. Y. Tan, M. A. Tabatabai, and N. Balakrishnan (1988), Harmonic mean aproach to unbalanced random effects model under heteroscedasticity, *Comm. Statist.* A *Theory Methods*, **17**, 1261–1286.

W. Y. Tan and S. P. Wong (1980), On approximating the null and non-null distributions of the F-ratio in unbalanced random effects models from non-normal universes, *J. Amer. Statist. Assoc.*, **75**, 655–662.

C. M. Theobold, M. Z. Firat, and R. Thompson (1997), Gibbs Sampling, adaptive rejection sampling and robustness to prior specification for a mixed linear model, *Genetic Selection Evolution*, **29**, 57–72.

B. Thitakamol (1977), *Extension of Previous Results on Properties of Estimators of Variance Components*, Ph.D. dissertation, University of Kentucky, Lexingon, KY.

J. D. Thomas and R. A. Hultquist (1978), Interval estimation for the unbalanced case of the one-way random effects model, *Ann. Statist.*, **6**, 582–587.

R. Thompson (1975), A note on the W transformation, *Technometrics*, **17**, 511–512.

R. Thompson (1976), Design of experiments to estimate heritability when observations are available on parents and offspring, *Biometrics*, **32**, 283–304.

R. Thompson (1977a), The estimation of heritability with unbalanced data I: Observations available on parents and offspring, *Biometrics*, **33**, 485–495.

R. Thompson (1977b), The estimation of heritability with unbalanced data II: Data availability on more than two generations, *Biometrics*, **33**, 497–504.

R. Thompson (1980), Maximum likelihood estimation of variance components, *Math. Oper. Statist. Ser. Statist.*, **11**, 545–561.

R. Thompson and Meyer, K. (1986), Estimation of variance components: What is missing in the EM algorithm?, *J. Statist. Comput. Simul.*, **24**, 215–230.

W. A. Thompson, Jr. (1962), The problem of negative estimates of variance components, *Ann. Math. Statist.*, **33**, 273–289.

W. A. Thompson, Jr. and J. R. Moore (1963), Non-negative estimates of variance components, *Technometrics*, **5**, 441–450.

W. O. Thompson and R. L. Anderson (1975), A comparison of designs and estimators for the two-stage nested random model, *Technometrics*, **17**, 37–44.

I. B. Thomsen (1975), Testing hypotheses in unbalanced variance components models for two-way layouts, *Ann. Statist.*, **3**, 257–265.

G. C. Tiao and I. Guttman (1965), The inverted Dirichlet distribution with applications, *J. Amer. Statist. Assoc.*, **60**, 793–805.

G. L. Tietjen (1974), Exact and approximate tests for unbalanced random effects designs, *Biometrics*, **30**, 573–581.

G. L. Tietjen and R. H. Moore (1968), On testing significance of variance in the unbalanced nested analysis of variance, *Biometrics*, **24**, 423–429.

N. Ting, R. K. Burdick, F. A. Graybill, S. Jeyaratnam, and T.-F. C. Lu (1990), Confidence intervals on linear combinations of variance components that are unrestricted in sign, *J. Statist. Comput. Simul.*, **35**, 135–143.

P. B. Topham (1966), Diallel analysis involving maternal and paternal interaction effects, *Heredity*, **21**, 665–674.

E. C. Townsend (1968), *Unbiased Estimators of Variance Components in Simple Unbalanced Designs*, Ph.D. dissertation, Cornell University, Ithaca, NY.

E. C. Townsend and S. R. Searle (1971), Best quadratic unbiased estimation of variance components from unbalanced data in the one-way classification, *Biometrics*, **27**, 643–657.

J. W. Tukey (1957), Variances of variance components II: The unbalanced single classification, *Ann. Math. Statist.*, **28**, 43–56.

S. Uhlig (1996), Optimum two-way nested designs for estimation of variance components, *Tatra Mountains Math. Pub.*, **7**, 105–112.

L. J. Van der Kemp (1972), Reliability coefficients via intraclass correlation: Biased and unbiased estimators, *Neederland. Tidschrift Psychol. Haar Grensgebieden*, **27**, 447–459.

W. H. Vandaele and S. R. Chowdhury (1971), A revised method of scoring, *Statist. Neerland.*, **25**, 101–112.

M. G. Vangel and A. L. Rukhin (1999), Maximum likelihood analysis for heteroscedastic one-way random effects ANOVA in interlaboratory studies, *Biometrics*, **55**, 129–136.

K. Venkateswarlu (1996), *Estimations of Genetic Variance Components in Diallel Cross Models with Unbalanced Data*, Ph.D. dissertation, University of Madras, Madras, India.

K. Venkateswarlu and K. N. Ponnuswamy (1998), Estimation of variance components based on biomodel of diallel crosses (balanced), *Comm. Statist.* A *Theory Methods*, **27**, 139–152.

K. Venkateswarlu, K. N. Ponnuswamy, and M. R. Srinivasan (1998), Estimation of variance components based on diallel model, *Math. Biol. Sci.*, **150**, 105–112.

A. P. Verbyla (1993), Modeling variance heterogeneity: Residual maximum likelihood and diagnostics, *J. Roy. Statist. Soc. Ser.* B, **55**, 82–86.

L. R. Verdooren (1974), Exact tests about variance ratios in unbalanced two and three stages nested designs, in L. C. A. Corsten and T. Postelnicu, eds., *Proceedings of the* 8*th International Biometrics Conference*, Editoria Academie, Constanta, Romania, 221–228.

L. R. Verdooren (1976), Exact confidence intervals for the variance ratio in nested designs, in J. Gordesch and P. Naeve, eds., *COMPSTAT: Proceedings of the* 2*nd Symposium on Computational Statistics*, Physica-Verlag, Wien, Germany, 221–228.

L. R. Verdooren (1980), On estimation of variance components, *Statist. Neerland.*, **34**, 83–106.

L. R. Verdooren (1988a), Least squares estimators and non-negative estimators of variance components, *Comm. Statist.* A *Theory Methods*, **17**, 1027–1051.

L. R. Verdooren (1988b), Exact tests and confidence intervals for the ratio of variance components in unbalanced two- and three-stage nested designs, *Comm. Statist.* A *Theory Methods*, **17**, 1197–1230.

Visual Numerics (1997), *IMSL: International Mathematical and Statistical Libraries* (*FORTRAN Subroutines for Evaluating Special Functions*): *Version* 3, Visual Numerics, Houston.

J. Volaufová and V. Witkovsky (1991), Least squares and minimum MSE estimators of variance components in mixed linear models, *Biometrical J.*, **33**, 923–936.

E. F. Vonesh and V. M. Chinchilli (1997), *Linear and Nonlinear Models for the Analysis of Repeated Measurements*, Marcel Dekker, New York.

A. Wald (1940), A note on analysis of variance with unequal class frequencies, *Ann. Math. Statist.*, **11**, 96–100.

A. Wald (1941), On the analysis of variance in case of multiple classifications with unequal class frequencies, *Ann. Math. Statist.*, **12**, 346–350.

A. Wald (1947), A note on regression analysis, *Ann. Math. Statist.*, **18**, 586–589.

C. S. Wang, J. J. Rutledge, and D. Gianola (1993), Marginal inference about variance components in a mixed linear model using Gibbs sampling, *Genetic Selection Evolution*, **25**, 41–62.

W. Y. Wang (1967), A comparison of several variance component estimators, *Biometrika*, **54**, 301–305.

T. Wansbeck (1980) A regression interpretation of the computation of MINQUE variance component estimates, *J. Amer. Statist. Assoc.*, **75**, 375–376.

R. E. Weiss, Y. Wang, and J. G. Ibrahim (1997), Predictive model selection for repeated measures random effects models using Bayes factors, *Biometrics*, **53**, 592–602.

S. J. Welham and R. Thompson (1997), Likelihood ratio tests for fixed model terms using residual maximum likelihood, *J. Roy. Statist. Soc. Ser.* B, **59**, 701–714.

P. H. Westfall (1986), Asymptotic normality of the ANOVA estimates of components of variance in the non-normal hierarchical mixed model, *Ann. Statist.*, **14**, 1572–1582.

P. H. Westfall (1987), A comparison of variance component estimates for arbitrary underlying distributions, *J. Amer. Statist. Assoc.*, **82**, 866–874.

P. H. Westfall (1988), Robustness and power of test for a null variance ratio, *Biometrika*, **75**, 207–214.

P. H. Westfall (1989), Power comparisons for invariant variance ratio tests in mixed ANOVA models, *Ann. Statist.*, **17**, 318–326.

P. H. Westfall (1994), Asymptotic efficiencies of MINQUE and ANOVA variance component estimates in the nonnormal random model, in T. Calinski and R. Kal, eds., *Proceedings of the International Conference on Linear Statistical Inference*, Kluwer, Amsterdam, 113–119.

P. H. Westfall and R. H. Bremer (1994), Efficiency properties of cell means variance component estimates, *J. Statist. Plann. Inference*, **38**, 159–178.

M. B. Wilk and O. Kempthorne (1955), Fixed, mixed, and random models, *J. Amer. Statist. Assoc.*, **50**, 1144–1167; corrigenda, **51**, 652.

S. S. Wilks (1962), *Mathematical Statistics*, Wiley, New York.

C. P. Winsor and G. L. Clarke (1940), A statistical study of variation in the catch of plankton nets, *J. Marine Res.*, **3**, 1–34.

S. Wolfram (1996), *The Mathematica Book*, 3rd ed., Cambridge University Press, Cambridge, UK.

G. Y. Wong and W. M. Mason (1985), The hierarchical logistic regression model for multilevel analysis, *J. Amer. Statist. Assoc.*, **80**, 513–524.

H. Yassaee (1976), Probability integral of inverted Dirichlet distribution and its applications, in J. Gordesch and P. Naeve, eds., *COMPSTAT: Proceedings of the* 2*nd Symposium on Computational Statistics*, Physica-Verlag, Wien, Germany, 64–71.

H. Yassaee (1981), On integrals of Dirichlet distributions and their applications, *Comm. Statist.* A *Theory Methods*, **10**, 897–906.

F. Yates (1934), The analysis of multiple classifications with unequal numbers in the different classes, *J. Amer. Statist. Assoc.*, **29**, 51–66.

F. Yates (1947), Analysis of data from all possible reciprocal crosses between sets of parental lines, *Heredity*, **1**, 287–302.

F. Yates and W. G. Cochran (1938), The analysis of groups of experiments, *J. Agricultural Sci.* (*Cambridge*), **28**, 556–580.

D. J. Young and M. Bhandary (1998), Test for equality of intraclass correlation coefficients under unequal family sizes, *Biometrics*, **54**, 1363–1373.

H. Yu, S. R. Searle, and C. E. McCulloch (1994), Properties of maximum likelihood estimators of variance components in the one-way classification, *Comm. Statist.* B *Simul. Comput.*, **23**, 897–914.

C. H. Yuan (1977), *A Least Squares Approach to Quadratic Unbiased Estima-*

tion of Variance Components in General Mixed Model, Ph.D. dissertation, North Carolina State University, Raleigh, NC.

S. L. Zeger and M. R. Karim (1991), Generalized linear models with random effects: A Gibbs sampling approach, *J. Amer. Statist. Assoc.*, **86**, 79–86.

S. L. Zeger, K.-Y. Liang, and D. S. Albert (1988), Models for longitudinal data: A generalized estimating equation approach, *Biometrics*, **44**, 1049–1060.

L. Zhou and T. Mathew (1994), Some tests for variance components using generalized p-values, *Techometrics*, **36**, 394–402.

Author Index

Abramowitz, M., 395, 425
Agrawal, R., 312, 326, 438
Ahrens, H. J., 112, 124, 129, 135, 136, 155, 425
Albert, A., 2, 10, 425
Albert, D. S., 56, 91, 457
Anderson, A. F., 421, 422, 443
Anderson, R. D., 48, 56, 76, 79, 425, 433
Anderson, R. L., 2, 5, 10, 11, 22, 79, 113, 130–132, 156, 159, 162, 206, 213, 231–237, 250–252, 306–309, 318, 325, 327, 353, 354, 367, 399, 400, 425, 426, 428, 432, 441, 444, 446, 451, 452, 454
Angers, J. F., 122, 135, 156, 426
Arminger, G., 418, 443
Armitage, P., 422, 434
Aroian, L. A., 395, 426
Arvesen, J. N., 124, 148, 156, 426
Ashida, I., 38, 76, 426

Bainbridge, T. R., 2, 10, 306, 309, 325, 353, 354, 363, 364, 367, 426
Baines, A.H. J., 96, 156, 426
Baksalary, J. K., 55, 76, 426
Balakrishnan, N., 22, 89, 242, 253, 453
Bankroft, T. A., 80, 306, 325, 353, 354, 367, 426, 435
Bansal, N. K., 114, 156, 426
Bapat, R. B., 2, 10, 426
Basilevsky, A., 2, 10, 426
Baskin, R. M., 60, 76, 426
Bates, D. M., 39, 84, 442
Belzile, E., 122, 135, 156, 426
Ben-Israel, A., 2, 3, 10, 427
Bener, A., 114, 156, 427
Berger, J. O., 122, 156, 427
Berger, J. V., 89, 450
Berk, K., 66, 76, 77, 427, 428
Berman, A., 2, 10, 427
Bernardo, J. M., 88, 122, 156, 427, 450
Best, N. G., 60, 79, 421–423, 433, 452
Beverly, J. M., 355, 369, 452
Bhandary, M., 114, 149, 156, 163, 426, 456
Bhargava, R. P., 148, 156, 427
Birch, H. J., 22, 77, 312, 325, 428
Birch, N. J., 22, 82, 310, 312, 315, 320, 326, 359, 360, 368, 437
Birkes, D., 63, 86, 445
Blischke, W. R., 265, 267, 272, 285, 427
Bliss, C. I., 318, 325, 342, 343, 367, 385, 387, 388, 427
Bolgiano, D. C., 49, 86, 447
Box, G. E. P., 59, 76, 427
Bremer, R. H., 22, 23, 49, 56, 82, 91, 236, 238, 245, 250, 251, 427, 438, 456
Breslow, N. E., 61, 76, 427
Broemeling, L. D., 59, 76, 78, 87, 427, 430, 447
Brokleban, J. C., 49, 76, 427
Brown, H. K., 61, 67, 76, 427
Brown, K. G., 48, 49, 76, 88, 427, 448
Brownlee, K. A., 114, 115, 156, 428
Bryk, A. S., 38, 57, 88, 421, 422, 428, 448

Bulmer, M. G., 143
Burch, B. D., 69, 76, 428
Burdick, R. K., 22, 69, 77, 78, 82, 140, 141, 143–146, 156, 190, 191, 198, 200, 238, 242, 250, 251, 310, 312, 313, 315, 320, 325, 326, 359, 360, 368, 428, 432, 437, 454
Burns, J. C., 66, 79, 433
Burrows, P. M., 37, 50, 79, 433
Burton, P. R., 60, 77, 428
Bush, N., 2, 5, 10, 213, 232, 233, 250, 251, 428

Calinski, T., 163, 456
Callanan, T. P., 39, 77, 81, 417, 428, 436
Calvin, L. D., 353, 367, 428
Cameron, J. M., 132, 156, 429
Cameron, P. E., 114, 159, 439
Caro, R. F., 126, 156, 429
Carriquiry, A. L., 58, 81, 436
Carroll, S. P., 102, 157, 429
Casella, G., 17, 20, 29, 37, 39, 56, 58, 60, 82, 89, 96, 161, 171, 177, 187, 199, 209, 216, 228, 252, 304, 306, 327, 347, 369, 418, 438, 450
Chaloner, K., 135, 156, 429
Chan, J. S. K., 56, 77, 429
Chase, G. A., 114, 160, 445
Chatterjee, S. K., 98, 105, 157, 429
Chaubey, Y. P., 48, 49, 53, 71, 72, 75–77, 87, 112, 134, 157, 429, 448
Cheng, S. S., 319, 328, 453
Chinchilli, V. M., 61, 91, 455
Chowdhury, S. R., 38, 90, 454
Christiansen, R., 69, 77, 429
Clarke, G. L., 96, 163, 456
Clayton, D. G., 60, 61, 76, 79, 427, 433
Clogg, C. C., 418, 443
Cochran, W. G., 6, 10, 50, 87, 93, 96, 97, 102, 112, 135, 154, 155, 157, 161–163, 323, 324, 327, 429, 448, 452, 456
Cockerham, C. C., 57, 77, 430
Colton, T., 422, 434
Conaway, M. R., 56, 78, 430
Conerly, M. D., 73, 78, 134, 157, 430
Congdon, R., 421, 422, 428
Cook, P., 59, 78, 430
Cookson, W. A. C. M., 60, 77, 428
Corbeil, R. R., 35, 37, 39, 78, 430
Cornelius, P. L., 2, 11, 236, 252, 451
Corsten, L.C. A., 86, 328, 445, 455
Cox, D. R., 56, 61, 78, 89, 405, 430, 452
Cramér, H., 149, 402
Cressie, N., 38, 78, 430
Crump, P. P., 2, 10, 113, 130–132, 156, 157, 426, 430
Crump, S. L., 103, 124, 126, 157, 206, 229, 251, 430
Csörgö, M., 156, 250, 325, 426
Cummings, W. B., 292, 318, 319, 326, 430

Das, K., 38, 78, 98, 105, 157, 429, 430
Das, R., 147, 157, 430
David, A. P., 89, 450
David, H. A., 161, 450
Davidian, M., 61, 79, 433
Davies, O. L., 65, 78, 430
Davies, R. B., 99, 157, 430
Dawson, D. A., 156, 250, 325, 426
deLeeuw, J., 57, 84, 421, 422, 441
Demidenko, E., 29, 37, 78, 430
Dempster, A. P., 38, 78, 416, 417, 431
Dewess, G., 47, 78, 432
Dickerson, G. E., 376, 388, 437
Dixon, W. J., 419, 422, 431

Dobson, A. J., 61, 78, 431
Dodge, Y., 158, 199, 431, 439
Donner, A., 114, 131, 142, 143, 145, 146, 148, 149, 157, 158, 431

Eickman, J. E., 143–146, 156, 313, 326, 428
Eisen, E. J., 318, 326, 363, 366, 367, 383, 388, 399, 400, 431
Eisenhart, C., 7, 10, 89, 431, 442
Eisenstat, S., 334, 355–357, 368, 442
El-Bassiouni, M. Y., 69, 78, 140, 158, 161, 193, 199, 242, 252, 431, 451
Eliasziw, M., 143, 157, 158, 431
Elliott, S. D., 22, 78, 432
Elston, D. A., 67, 78, 432
Elston, R. C., 114, 160, 445
Engel, B., 37, 78, 432
Evans, J. W., 236, 252, 452
Everson, D. O., 130, 162, 453

Fayyad, R., 69, 78, 432
Fedorov, V. V., 158, 432
Fellner, W. H., 38, 78, 432
Fenech, A. P., 140, 158, 436
Fennech, A. R., 70, 78, 432
Fernando, R. L., 56, 59, 79, 156, 429, 433
Finker, A. L., 376, 388, 432
Firat, M. Z., 59, 89, 453
Fisher, R. A., 6, 38, 39, 158, 415, 416, 431
Focke, J., 47, 78, 432
Forthofer, R. N., 23, 79, 432
Foulley, J. L., 59, 79, 433

Gallo, J., 383, 388, 432
Ganguli, M., 292, 326, 334, 367, 376, 388, 432
Gao, S., 55, 79, 432
Garrett, R. G., 368, 434
Gates, C. E., 375, 376, 388, 432
Gaylor, D. W., 22, 79, 80, 206, 231, 232, 251, 306, 308, 318, 319, 326, 334, 367, 400, 430, 432, 434, 435
Gelfand, A. E., 60, 79, 433
Gentle, J. E., 2, 10, 417, 433, 439
Gerald, C. F., 39, 79, 433
Gharaff, M. K., 59, 78, 79, 430, 433
Ghosh, M., 55, 79, 292, 321, 327, 433, 440
Gianola, D., 56, 59, 60, 79, 81, 91, 433, 436, 455
Gibbons, R. D., 56, 57, 79, 81, 433, 436
Giesbrecht, F. G., 37, 50, 66, 79, 433
Gilks, W. R., 60, 79, 421–423, 433, 452
Gill, D. S., 134, 159, 439
Gilmour, A. R., 56, 79, 433
Giltinan, D., 61, 79, 433
Giovagnoli, A., 133, 158, 433
Glueck, C. J., 114, 160, 445
Gnot, S., 55, 80, 433
Goel, P. K., 156, 427
Golberger, A. S., 58, 80, 434
Goldsmith, C. H., 306, 308, 309, 319, 326, 434
Goldsmith, P. L., 65, 78, 430
Goldstein, H., 57, 80, 421, 422, 434, 446, 448
Goldstein, R., 418, 422, 434
Gönen, M., 59, 80, 434
Goodnight, J. H., 80, 434
Gordesch, J., 163, 328, 397, 455, 456
Goss, T. I., 355, 368, 434
Gosslee, D. G., 22, 80, 176, 199, 216, 251, 434
Gower, J. C., 376, 388, 434
Graser, H. U., 39, 80, 434
Graybill, F. A., 2, 4, 6, 7, 10, 22, 69, 77, 78, 81, 96, 140, 141, 145, 146, 156, 158,

159, 190, 191, 198–200, 238, 242, 250, 251, 292, 299, 312–315, 325–327, 393, 394, 428, 432, 434, 436, 438, 439, 451, 452, 454
Green, J. W., 56, 80, 82, 245, 251, 434, 438
Green, P. P., 114, 160, 445
Green, W. J., 22
Gregory, K. E., 130, 162, 453
Greyville, T., 2, 3, 10, 427
Griffing, B., 57, 80, 434
Groggel, D. J., 143, 158, 434
Grossman, M., 156, 429
Gumpertz, M. L., 61, 80, 434
Gunst, R. F., 365, 368, 443
Gurrin, L. C., 60, 77, 428
Guttman, I., 59, 89, 396, 397, 453, 454

Hadi, A. S., 2, 10, 435
Haile, A., 233, 251, 435
Hammersley, J. M., 96, 124, 131, 158, 435
Hammond, K., 81, 436
Hardy, D. W., 396, 397, 435
Hartley, H. O., 27, 29–31, 38–41, 49, 80, 81, 104, 158, 161, 258, 285, 376, 388, 435, 436, 450
Hartung, J., 14, 62, 80, 144–146, 158, 435
Hartwell, T. D., 22, 80, 334, 367, 432, 435
Harvey, L. R., 130, 162, 453
Harville, D. A., 2, 10, 29, 30, 34, 37–40, 42, 56, 58, 59, 63, 65, 66, 70, 77, 78, 80, 81, 87, 98, 123, 124, 140, 148, 158, 160, 417, 428, 432, 435, 436, 442, 448
Hayat, N., 421, 422, 443
Hayman, B. I., 29, 57, 81, 436
Healy, M. J. R., 2, 10, 436
Heck, R. H., 57, 81, 436
Heckler, C. E., 299, 309, 323, 326, 327, 356, 358, 368, 436, 448
Hedecker, D., 81, 436
Heine, B., 135, 158, 436
Hemmerle, W. J., 38, 39, 80, 81, 434, 436
Henderson, C. R., 13, 14, 16–18, 20–22, 38, 48, 56, 58, 62, 64, 76, 81–83, 87–89, 104, 158, 169, 170, 181, 182, 199, 207, 209, 221, 237, 245–250, 252, 270, 274, 275, 299, 334, 342, 368, 376, 388, 404, 425, 437, 439, 440, 446, 447, 449, 450
Hernández, R. P., 22, 82, 191, 199, 238, 242, 251, 310, 312, 313, 320, 326, 359, 360, 368, 437
Hess, J. L., 112, 133, 159, 365, 368, 437, 443
Hetzer, H. O., 376, 388, 437
Hill, B. M., 104, 117, 120–122, 159, 437
Hills, S. E., 60, 79, 433
Hinkelman, K., 76, 427
Hinkly, D. V., 405, 430
Hirotsu, C., 175, 199, 214, 215, 230, 233, 240, 241, 251, 437
Hobert, J. P., 60, 82, 438
Hocking, R. R., 22, 29, 37, 48, 56, 82, 245, 251, 258, 286, 438
Hodge, G. R., 63, 82, 438
Hoefer, A., 62, 82, 438
Hooke, R., 156, 426
Hopper, F. N., 400, 432
Horn, R., 2, 10, 438
Horn, R. A., 82, 438
Horn, S. D., 82, 438
Hsu, J. S. J., 59, 89, 453

Huber, D. A., 63, 82, 438
Huda, S., 114, 132, 156, 160, 427, 444
Hudson, H. M., 56, 82, 438
Hultquist, R. A., 22, 90, 96, 136, 140–146, 159, 162, 313, 328, 438, 453
Hussein, M., 194, 199, 317, 320, 326, 438

Ibrahim, J. G., 59, 91, 455
Iwaisaki, H., 38, 76, 426
Iyenger, N. S., 156, 427
Iyer, H. K., 69, 76, 428

Jain, R. C., 312, 320, 326, 438
Jayaratnam, S., 190
Jennrich, R. L., 38, 39, 82, 416, 417, 438
Jeyaratnam, S., 146, 152, 159, 200, 438, 454
Jiang, J., 38, 82, 438
Johnson, C. R., 2, 10, 438
Johnson, D. E., 221, 222, 252, 444
Johnson, N. L., 155, 319, 326, 334, 355–357, 368, 422, 425, 434, 438, 442
Jones, A. C., 69, 83, 439

Kachman, S. D., 61, 89, 452
Kal, R., 163, 456
Kala, R., 143, 159, 439
Kapadia, C. H., 37, 53, 84, 404, 442
Kaplan, E. B., 114, 160, 445
Kaplan, J., 50, 87, 112, 135, 161, 448
Kaplan, J. S., 50, 83, 439
Karim, M. R., 56, 91, 457
Karlin, S., 114, 159, 439
Kasim, R. M., 60, 83, 439
Kazempour, M. K., 191, 199, 238, 251, 439
Keifer, G., 40, 41, 80, 435
Kelly, R. J., 98, 135, 159, 439
Kempthorne, O., 58, 82, 253, 437, 456
Kempton, R. A., 67, 76, 427
Kennedy, B. W., 58, 83, 439
Kennedy, W. J., 417, 439
Kenward, M. G., 70, 83, 439
Khattree, R., 22, 83, 134, 159, 317, 320, 326, 327, 355, 358, 368, 376, 377, 382, 383, 387, 388, 439, 440, 445
Khuri, A. I., 22, 68, 70, 83, 99, 124, 126, 136, 147, 159, 160, 181, 182, 193, 194, 199, 214, 215, 242, 251, 292, 312, 320, 321, 326, 327, 383, 388, 400, 432, 440, 442
Kim, B. C., 50, 84, 442
King, S. C., 334, 368, 376, 388, 440
Kirby, A. J., 60, 79, 433
Kleffé, J., 44, 47, 48, 50, 55, 80, 83, 87, 98, 114, 135, 155, 159, 221, 251, 425, 433, 440, 441, 447
Klotz, J., 39, 84, 441
Knapp, G., 144–146, 158, 435
Knoke, J. D., 22, 84, 441
Koch, G. G., 23, 25, 56, 79, 84, 100, 103, 126, 159, 179, 181, 199, 219, 220, 236, 251, 264, 267, 268, 286, 297, 298, 327, 340, 341, 368, 379, 381, 388, 432, 441, 442
Kotz, S., 422, 434
Koval, J. J., 114, 131, 146, 148, 149, 157, 158, 431
Kreft, I. G., 57, 84, 421, 422, 441
Krishnaiah, P. R., 87, 447
Krishnan, T., 417, 418, 444
Kruskal, W. H., 319, 327, 441
Kuk, A. Y., 56, 77, 429
Kuranchie, P., 75, 88, 448
Kussmaul, K., 132, 159, 441
Kutner, M. H., 48, 82, 438

Lahiri, S. N., 38, 78, 430
Laird, N. M., 38, 39, 57, 84, 89, 416, 417, 431, 441, 452
Laird, R., 56
Lamar, J. L., 355, 368, 441
LaMotte, L. R., 13, 38, 40, 48–51, 53, 54, 80, 84, 140, 147, 148, 160, 435, 441
Lancaster, H. O., 5, 10, 441
Landis, J. R., 56, 84, 442
Lange, N., 84, 441
Layard, M. W. J., 148, 156, 426
Lee, H. S., 56, 84, 442
Lee, J., 99, 136, 160, 442
Lee, J. T., 49, 50, 84, 442
Lee, J. W., 67, 89, 452
Lee, K. R., 37, 53, 84, 404, 442
Lee, P. M., 59, 84, 442
Lee, Y., 140, 160, 442
Lehmann, E. L., 403, 405, 442
Leonard, T., 59, 89, 453
Leone, F. C., 319, 326, 334, 355–357, 368, 438, 442
Lera, M. L., 136, 160, 442
Liang, K. Y., 56, 67, 68, 89, 91, 451, 457
Lin, C. Y., 39, 84, 238, 252, 442
Lin, T. H., 140, 148, 160, 442
Lin, X., 56, 84, 442
Lindley, D. V., 59, 84, 442
Lindstrom, M. J., 39, 84, 442
Littell, R. C., 61, 68, 83, 85, 181, 182, 199, 242, 251, 440, 442
Liu, L. M., 49, 50, 85, 442, 443
Longford, N. T., 38, 57, 85, 416–418, 421, 422, 443
Lorens, J. A., 39, 81, 436
Low, L. Y., 171, 189, 199, 236, 252, 443
Lu, T.-F. C., 190, 200, 454
Lucas, H. L., 22, 79, 80, 176, 199, 206, 216, 251, 432, 434

MacElderry, R., 388, 432
Madow, W., 6, 11, 443
Magnus, J. R., 2, 11, 443
Mahamunulu, D. M., 334, 347, 363, 368, 382, 388, 443
Malley, J. D., 56, 85, 443
Malone, L., 77, 428
Maqsood, F., 140, 141, 145, 146, 156, 428
Mason, R. L., 355, 358, 365, 368, 376, 382, 387, 388, 422, 440, 443
Mason, W. M., 56, 91, 421, 443, 456
Massam, H., 37, 49, 78, 85, 430, 443
Mathew, T., 55, 68, 83, 85, 98, 126, 135, 147, 159, 160, 193, 194, 199, 242, 251, 320, 327, 383, 388, 389, 439, 440, 443, 457
McAllister, A. J., 39, 84, 238, 252, 442
McBratney, A. B., 355, 368, 445
McCullagh, P., 61, 85, 444
McCulloch, C. E., 17, 20, 29, 37, 39, 56, 58, 61, 85, 89, 96, 105, 161, 163, 171, 177, 187, 199, 209, 216, 228, 252, 304, 306, 327, 347, 369, 417, 418, 444, 450, 456
McDonald, B. W., 56, 85, 444
McLachlan, G. J., 417, 418, 444
McLean, R. A., 66, 85, 444
McNiel, A. J., 60, 79, 433
McWhorter, A., 147, 148, 160, 441
Mee, R. W., 56, 81, 436
Meyer, K., 38, 90, 454
Mian, I.U. H., 143, 160, 444
Miller, J. D., 353, 367, 428
Miller, J. J., 27, 29, 85, 444
Milliken, G. A., 61, 85, 194, 199, 221, 222, 252, 317, 320, 326, 438, 442, 444

Mitra, S. K., 2, 3, 11, 48, 61, 62, 85, 444, 447
Miyawaki, N., 74, 88, 448
Molińska, A., 55, 76, 159, 426, 439
Moliński, K., 159, 439
Monaham, J. F., 134, 162, 453
Monroe, R. J., 388, 432
Moore, J. R., 90, 454
Moore, R. H., 327, 383, 389, 441, 454
Morgan, J. J., 388, 432
Moriguti, S., 143, 144
Morrel, C. H., 85, 444
Morris, C. N., 57, 86, 444
Morrison, E. B., 114, 160, 445
Mostafa, M. G., 147, 148, 160, 175, 199, 214, 233, 252, 444
Mukerjee, R., 132, 160, 444
Muller, J., 49, 85, 443
Murray, L. W., 56, 89, 451
Muse, H. D., 2, 11, 234–237, 252, 444
Musk, A. W., 60, 77, 428

Naeve, P., 163, 328, 397, 455, 456
Naik, D. N., 317, 320, 326, 327, 355, 358, 368, 376, 377, 382, 383, 387, 388, 439, 440, 445
Namboodiri, K. K., 114, 160, 445
Naqvi, S.T. M., 108
Nelder, J., 61, 85, 444
Nelson, L. S., 334, 355–357, 368, 377, 388, 442, 445
Neudecker, H., 2, 11, 443
Neyman, J., 405

Ochi, Y., 56, 86, 445
Öfversten, 68, 86, 445
Ojima, Y., 376, 382, 388, 389, 445
Olkin, I., 113, 160, 445
Olsen, A., 63, 86, 445
Omori, Y., 56, 86, 445
Othman, A. R., 146, 152, 159, 160, 438, 445
Owen, A. R. G., 114, 160, 445

Palmer, L. J., 60, 77, 428
Pantula, S. G., 61, 80, 434
Patel, C. M., 38, 39, 78, 431
Patterson, H. D., 34, 37, 48, 86, 445
Pearson, K., 114, 395, 405, 445
Pederson, D. G., 236, 252, 445
Peters, D., 68, 86, 446
Pettitt, A. N., 355, 368, 445
Pierce, D. A., 68, 86, 446
Pincus, R., 135, 155, 425
Plemmons, R. J., 2, 10, 427
Ponnuswamy, K. N., 49, 58, 86, 90, 446, 454, 455
Postelnicu, T., 86, 328, 445, 455
Postma, B. J., 376, 377, 389, 446
Prairie, R. R., 306–309, 327, 353, 368, 446
Prasad, N. G. N., 148, 160, 446
Prasad, R. A., 147, 148, 160, 441
Pratt, J. W., 113, 160, 445
Prentice, R. L, 56, 86, 445
Prescott, R., 61, 76, 427
Press, S. J., 123, 160, 446
Pringle, R. M., 2, 3, 11, 47, 86, 446
Prosser, R., 421, 422, 446
Pukelsheim, F., 14, 48, 54–56, 86, 404, 446, 447
Pukkila, T., 157, 430
Pulley, P. Jr., 376, 389, 447
Puntanen, S., 157, 430
Putter, J., 39, 84, 441

Quaas, R. L., 86, 447

Racine-Poon, A., 60, 79, 433
Rae, A. L., 56, 79, 433
Raghavan, T.E. S., 2, 10, 426
Rajagopalan, M., 59, 87, 447
Rao, C. R., 2, 3, 11, 13, 33, 40–45, 47–49, 53, 56, 64, 78, 83, 85, 87, 109, 114, 149, 160, 402, 440, 447
Rao, J. N. K., 27, 29–31, 38–41, 49, 55, 80, 83, 87, 104,

148, 156, 158, 160, 250, 258, 286, 325, 426, 435, 440, 446–448
Rao, M. B., 2, 11, 447
Rao, P. S. R. S., 29, 37, 48–50, 71–75, 77, 87, 88, 93, 102, 111, 112, 125, 134, 135, 161, 299, 309, 323, 326, 327, 356, 358, 368, 429, 436, 448
Rao, P. V., 143, 158, 434
Rasbash, J., 421, 422, 446, 448
Raudenbush, S. W., 38, 57, 60, 83, 88, 421, 422, 428, 439, 448
Rawlings, J. O., 61, 80, 434
Raynor, A., 2, 3, 11, 446
Read, R. R., 98, 161, 448
Rhode, C. A., 20, 88, 229, 252, 448, 449
Rich, D. K., 49, 88, 448
Richardson, A. M., 38, 88, 449
Rifkind, B. M., 114, 160, 445
Robertson, A., 98, 161, 449
Robinson, D. L., 39, 40, 88, 449
Robinson, G. K., 58, 88, 449
Robinson, J., 123, 161, 449
Robson, D. S., 104, 158, 437
Rocke, D. M., 56, 88, 449
Roger, J. H., 70, 83, 439
Rohlf, F. J., 153, 162, 299, 300, 328, 452
Rosenberg, S. H., 20, 88, 449
Rosenthal, J. S., 60, 88, 449
Rosner, B., 324, 325, 327, 449
Roth, A. J., 38, 39, 78, 431
Rubin, D. B., 38, 78, 416, 417, 431
Rudan, J. W., 352, 363, 368, 449
Rudolph, P., 59, 88, 449
Rukhin, A. L., 105, 163, 454
Rutledge, J. J., 59, 60, 91, 455

Sánchez, J. E., 155, 161, 425, 449
Saleh, A. K. Md. E., 85, 156, 250, 325, 426, 443
Sampson, P. F., 38, 39, 82, 416, 417, 438
SAS Institute, 395, 418, 422, 449
Satterthwaite, F. E., 66–69, 88, 123, 194, 318, 319, 359–361, 383, 385, 386, 389, 397–400, 440, 449
Saunders, W. L., 66, 85, 444
Schaeffer, L. R., 49, 88, 236, 252, 449
Schaffer, L. R., 17, 82, 437
Schall, R., 61, 88, 449
Scheffé, H., 6, 11, 40, 88, 98, 161, 292, 293, 327, 449
Schervish, M. L., 59, 88, 450
Schewenke, J. R., 452
Schmitz, T. H., 148, 156, 426
Schott, J. R., 2, 11, 450
Schwartz, J. H., 309, 327, 450
Searle, S. R., 2, 3, 5, 7, 8, 11, 13, 14, 16–18, 20, 29, 30, 35, 37, 39, 40, 56, 58, 59, 61, 76, 78, 82, 85, 87–90, 96, 104–107, 110, 112, 123–128, 133, 158, 161–163, 169–171, 177, 187, 189, 199, 207–209, 216, 228, 244, 252, 258, 265, 286, 292, 304, 306, 321, 327, 347, 352, 363, 368, 369, 382, 389, 404, 417, 418, 425, 430, 437, 444, 446–450, 453, 454, 456
Sebastiani, P., 133, 158, 433
Seely, J. F., 14, 61, 63, 86, 89, 140, 158, 160, 161, 193, 199, 242, 252, 431, 442, 445, 450, 451
Seifert, B., 83, 441
Self, S. G., 67, 68, 89, 451
Seltzer, M., 421, 422, 428
Selwyn, M. R., 38, 39, 78, 431
Sen, B., 294, 299, 314, 315, 327, 451
Seneta, E., 2, 11, 451

Senturia, J., 49, 50, 85, 442, 443
Sharpe, R. H., 342, 369, 451
Sharples, L. D., 60, 79, 433
Sheffé, H., 403
Shen, P. S., 2, 11, 236, 252, 451
Shiue, C., 375, 376, 388, 432
Shoukri, M. M., 143, 157, 160, 161, 431, 444, 451
Singh, B., 99, 114, 125, 136, 146, 161, 162, 451
Singh, J., 312, 320, 326, 438
Singhal, R. A., 162, 451
Sinha, B. K., 55, 68, 83, 85, 98, 126, 135, 147, 157, 159, 160, 193, 194, 199, 242, 251, 320, 327, 383, 388, 430, 440, 443
Smith, A. F. M., 59, 60, 79, 84, 89, 433, 442, 450
Smith, C. A. B., 114, 130, 162, 451
Smith, D. W., 56, 89, 451
Smith, H. F., 398, 400, 452
Smith, J. R., 355, 369, 452
Smith, S. P., 80, 434
Smith, T. M. F., 55, 79, 432
Snedecor, G. W., 80, 154, 155, 162, 323, 324, 327, 435, 452
Snee, R. D., 355, 369, 383, 389, 452
Sobel, M. E., 418, 443
Sokal, R. R., 153, 162, 299, 300, 328, 452
Solomon, P. J., 61, 89, 452
Spiegelhalter, D. J., 60, 79, 421–423, 433, 452
Spjøtvoll, E., 147, 148, 162, 190, 193, 199, 238, 242, 252, 452
SPSS, Inc., 419, 423, 452
Srinivasan, M. R., 58, 90, 455
Srinivasan, S., 191, 199, 200, 238, 252, 452
Srivastava, J. N., 81, 325, 425, 436
Stegun, I. A., 395, 425
Stiratelli, R., 56, 89, 452
Stram, D. O., 67, 84, 89, 441, 452
Stroup, W. W., 61, 85, 89, 236, 252, 442, 452
Subrahmaniam, K., 49, 87, 448
Subramani, J., 49, 61, 86, 89, 446, 453
Sun, L., 59, 89, 453
Sutradhar, B. C., 55, 85, 98, 135, 160, 443
Swallow, W. H., 110, 112, 128, 129, 133, 134, 162, 453
Swiger, L. A., 130, 162, 453
Sylvestre, E. A., 112, 161, 448

Tabatabai, M. A., 22, 89, 242, 253, 453
Tallis, G. M., 229, 252, 448
Tan, W. Y., 22, 89, 146, 162, 242, 253, 319, 328, 453
Tenzler, R., 135, 155, 425
Theobold, C. M., 59, 89, 453
Thitakamol, B., 2, 11, 235, 237, 252, 253, 444, 453
Thomas, A., 421–423, 433, 452
Thomas, J. D., 22, 90, 136, 140–146, 162, 313, 328, 453
Thomas, S. L., 57, 81, 436
Thompson, R., 29, 34, 37–40, 48, 59, 67, 86, 89–91, 236, 253, 445, 453, 454, 456
Thompson, W. A. Jr., 34, 90, 454
Thompson, W. O., 132, 162, 454
Thomsen, I. B., 22, 90, 193, 200, 242, 253, 454
Tiao, G. C., 59, 76, 396, 397, 427, 454
Tier, B., 80, 434
Tietjen, G. L., 319, 327, 328, 383, 389, 441, 454
Tiller, K. J., 60, 77, 428
Ting, N., 190, 200, 299, 314, 327, 451, 454
Topham, P. B., 57, 90, 454
Townsend, D. S., 160, 444

Townsend, E. C., 40, 90, 106–108, 133, 162, 163, 454
Tracy, D. S., 143
Tsutakawa, R. K., 38, 78, 431
Tukey, J. W., 124, 143, 144, 156, 163, 426, 454
Tyroler, H. A., 114, 160, 445

Uhlig, S., 355, 369, 454

Van Der Kemp, L. J., 113, 163, 454
van der Leeden, R., 421, 422, 441
van Middelem, C. H., 342, 369, 451
Van Vleck, L. D., 76, 87, 88, 404, 425, 446–448
Vandaele, W. H., 38, 90, 454
Vangel, M. G., 105, 163, 454
Vaughn, W. K., 40, 80, 435
Venkateswarlu, K., 58, 90, 454, 455
Verbyla, A. P., 63, 90, 455
Verdooren, L. R., 21, 48, 62, 74, 90, 140, 148, 163, 292, 313, 317, 320, 328, 455
Visual Numerics, 395, 455
Volaufová, J., 53, 90, 455
von Krosigk, C. N., 58, 82, 437
Vonesh, E. F., 61, 91, 455

Wackerly, D. D., 143, 158, 434
Wald, A., 30, 66, 78, 82, 136, 137, 140–142, 145–148, 158, 160, 161, 163, 190, 193, 194, 199, 200, 238, 242, 252, 313, 317, 359, 431, 432, 438, 442, 451, 455
Walker, C. L., 396, 397, 435
Wang, C. S., 59, 60, 91, 455
Wang, W. Y., 123, 163, 455
Wang, Y., 59, 91, 455
Wansbeck, T., 50, 91, 455
Ward, R. H., 143, 161, 451
Ware, J. H., 56, 57, 84, 89, 441, 452
Webster, J. T., 73, 78, 134, 157, 233, 251, 430, 435
Weir, B. S., 57, 77, 430
Weiss, R. E., 59, 91, 455
Welham, S. J., 67, 91, 456
Wells, G. A., 142, 143, 145, 146, 158, 431
Welsh, A. H., 38, 88, 449
Westfall, P. H., 16, 23, 49, 68, 91, 134, 140, 147, 148, 163, 456
White, J. S., 376, 377, 389, 446
White, T. L., 63, 82, 438
Whitkovsky, V., 53, 90, 455
Wilk, M. B., 243, 253, 456
Wilks, S. S., 393, 394, 456
Williams, P., 114, 143, 144, 159, 439
Winsor, C. P., 96, 163, 456
Wolfinger, R. D., 61, 85, 442
Wolfram, S., 395–397, 456
Wong, G. Y., 56, 91, 456
Wong, S. P., 146, 162, 453
Woodhouse, G., 421, 422, 448
Wynn, H. P., 158, 432

Yang, M., 421, 422, 448
Yassaee, H., 396, 397, 456
Yates, F., 22, 57, 91, 93, 163, 177, 200, 216, 253, 456
Young, D. J., 149, 163, 456
Yu, H., 105, 163, 456
Yuan, C. H., 61, 91, 456

Zeger, S. L., 56, 91, 457
Zeller, J. H., 376, 388, 437
Zhou, L., 383, 389, 457
Zimmerman, A. G., 59, 81, 436
Zirk, W. E., 355, 368, 441
Zmyslony, R., 55, 80, 433

Subject Index

Analysis of means estimators
for two-way crossed classification with interaction, 213–217
for two-way crossed classification without interaction, 174–178
variances of, for two-way crossed classification with interaction, 230
Analysis of means method, 22–23, 62–63
Analysis of variance
for general r-way nested classification, 372–375
for one-way classification, 94–96
for three-way and higher-order crossed classifications, 256–258
for three-way nested classification, 330–331
for two-way crossed classification with interaction, 201–203
for two-way crossed classification without interaction, 165–167
for two-way nested classification, 288–289
Analysis of variance estimators, 113
for general r-way nested classification, 376–377
for one-way classification, 97–99
for three-way and higher-order crossed classifications, 261–262
for three-way nested classification, 336–337
for two-way crossed classification with interaction, 207–208
for two-way crossed classification without interaction, 169–170
for two-way nested classification, 295
variances of
for three-way nested classification, 347–351
for two-way crossed classification with interaction, 227–229
for two-way crossed classification without interaction, 186–189
Anderson design, 307
Anderson five-stage staggered nested design, 353, 354

Bainbridge design, 306–307
Bainbridge four-stage inverted nested design, 353, 354
Bainbridge four-stage staggered nested design, 354–355
Balanced data, 1–2
Balanced disjoint (BD) designs, 232
Bayesian estimation, 58–59
in one-way classification, 117–122
Best linear prediction (BLP), 58
Best linear unbiased prediction (BLUP), 57–58
empirical, 58

Best prediction (BP), 58
Best quadratic unbiased estimation (BQUE), 40
Best quadratic unbiased estimators, for one-way classification, 106–108
Beta function
  complete, 394
  incomplete, 394–395
Bibliography, general, 425-457
BMDP application, 419–420
  for one-way classification, 116
  for three-way and higher-order crossed classifications, 273
  for three-way nested classification, 345, 348
  for two-way crossed classification with interaction, 227
  for two-way crossed classification without interaction, 186
  for two-way nested classification, 302, 304
BUGS, 420–421
Burdick–Eickman procedure, 143–144
Burdick–Maqsood–Graybill procedure, 141–142
Bush–Anderson designs, 232, 233

C designs, 232
Chi-square distribution, inverted, 397
Classical estimation, for one-way classification, 97–117
Cochran–Fisher theorem, 6
Cofactor of element of a matrix, 411
Column vector, 407
Comparisons of designs and estimators
  for one-way classification, 131–136
  for three-way nested classification, 353–359
  for two-way crossed classification with interaction, 230–238
  for two-way nested classification, 306–310
Complete beta function, 394
Complete sufficient statistics, 403
Compound mean squared error (CMSE) criterion, 358
Compound squared bias (CSB) criterion, 358
Confidence intervals, 69
  for general r-way nested classification, 382–383
  methods for constructing, 69–70
  for one-way classification, 136–146
  for three-way nested classification, 359
  for two-way crossed classification with interaction, 238–240
  for two-way crossed classification without interaction, 190–193
  for two-way nested classification, 310–317
Covariance components model, 55–56
Covariance matrix of T, variance components in, coefficients of products of, 276–285
Crump's results for sampling variance, 126–127

Design matrices, 27
Designed unbalancedness, 232
Designs, 64–65

comparisons of, 64–66; *see also* Comparisons of designs and estimators
various, efficiency factors of, 231–232
Determinant of a matrix, 411
Diagonal matrix, 407
Diallel cross experiments, 57–58
Differentiation of matrices and vectors, 413–415
Direct product of matrices, 410
Direct sum of matrices, 409
Dirichlet function, incomplete inverted, 395–396
Dispersion-mean model, 56
Distribution and sampling variances of estimators, for one-way classification, 122–131
Distribution theory
for one-way classification, 96–97
two lemmas in, 391–392
for two-way nested classification, 292–293
Donner–Wells procedure, 142–143

Efficiency factors of various designs, 231–232
EM (expectation-maximization) algorithm, 38–39, 416–417
Empirical best linear unbiased prediction, 58
Estimation of variance components
for general r-way nested classification, 376–382
for three-way and higher-order crossed classifications, 261–264
for three-way nested classification, 336–347
for two-way crossed classification with interaction, 204–226
for two-way crossed classification without interaction, 169–186
for two-way nested classification, 295–303
Estimators, comparisons of, 64–66; *see also* Comparisons of designs and estimators
Expectation
under fixed effects, 7–8
under mixed effects, 8–9
under random effects, 9
Expectation-maximization (EM) algorithm, 38–39, 416–417
Expected mean squares
for general r-way nested classification, 375–376
for three-way and higher-order crossed classifications, 258–261, 289–292
for three-way nested classification, 331–334
for two-way crossed classification with interaction, 204–206
for two-way crossed classification without interaction, 167–169

F-tests, 66
pseudo, 399
Figures, list of, xiii
Fisher's scoring algorithm, 38–39, 416
Fitting-constants-method estimators
for one-way classification, 99–100

for two-way crossed classification with interaction, 208–213
for two-way crossed classification without interaction, 170–174
variances of
for two-way crossed classification with interaction, 229
for two-way crossed classification without interaction, 189–190
Fixed effects, 58
expectation under, 7–8

General linear model, 6–9
expectation under fixed effects, 7–8
expectation under mixed effects, 8–9
expectation under random effects, 9
mathematical model for, 6–7
General mean, for one-way classification, 102–103
General methods of estimation, relative merits and demerits of, 62–64
General r-way crossed classification, 266–268
General r-way nested classification, 371–387
analysis of variance estimators for, 376–377
analysis of variance for, 372–375
confidence intervals for, 382–383
estimation of variance components for, 376–382
expected mean squares for, 375–376
mathematical model for, 371–372
numerical example, 385–387
symmetric sums estimators for, 377–381
tests of hypotheses for, 382–384
variances of estimators for, 382
Generalized inverse of a matrix, 3
Generalized linear mixed models (GLMM), 60–61
Generalized linear models (GLM), 60
GENMOD, 421
Genstat, 420
Gibbs sampling, 59–60, 420
GLM procedure, 419
Goldsmith–Gaylor designs, 308

Hartley–Rao estimation procedure, 27–30
Hartung–Knapp procedure, 144
Heckler–Rao five-stage extended staggered design, 356, 358
Heckler–Rao four-stage extended staggered design, 356, 358
Henderson's Method I, 14–16, 62
Henderson's Method II, 16–17, 62
Henderson's Method III, 18–22, 62
alternative formulation of, 21–22
Hessian matrix, 414
Hierarchical linear models, 56–57
HLM package, 421–422
Hypothesis testing, methods of, 66–69

Idempotent square matrix, 407
Identity matrix, 407–408
Incidence matrices, 27

Incomplete beta function,
394–395
Incomplete inverted Dirichlet
function, 395–396
Inquadmissible estimators, 98
Interaction, definition of, 406
Intraclass correlation, for one-way
classification, 113–114
Invariance property of maximum
likelihood estimators,
lemmas on, 402–403
Inverse of a matrix, 411–412
generalized, 3
Inverted chi-square distribution,
397
Inverted Dirichlet function,
incomplete, 395–396
Iterated least squares (ITLS)
estimator, 309
Iterative MINQUE, 64

Khuri–Littell test, 242
Kronecker product of matrices,
410

L designs, 232
Large sample variances of
maximum likelihood
estimators, 30–33
for three-way nested
classification, 351–353
for two-way nested
classification, 306
"Last-stage uniformity," 319*n*
Lehmann–Sheffé theorem, 403
Likelihood function, 28–29, 401
Likelihood-ratio test, 67–68, 405
Locally optimum tests, 147

M_1 and M_2 designs, 233
MANOVA, 419
Marginal maximum likelihood,
34*n*
Markov chain Monte Carlo
(MCMC) algorithms,
420

Mathematical model, 6–7
for general linear model, 6–7
for general r-way nested
classification, 371–372
for one-way classification, 93
for three-way and
higher-order crossed
classifications, 255–256
for three-way nested
classification, 329–330
for two-way crossed
classification with
interaction, 201
for two-way crossed
classification without
interaction, 165
for two-way nested
classification, 287–288
Matrix algebra, 406–415
Matrix/matrices
definition of, 406
determinant of, 411
diagonal, 407
differentiation of, 414–415
direct product of, 410
direct sum of, 409
generalized inverse of, 3
idempotent square, 407
identity, 407–408
inverse of, 411–412
Kronecker product of, 410
lemmas for, 393
nonsingular, 411
null, 408
orthogonal square, 407
partitioned, 412–413
products of, 409–410
rank of, 410
singular, 411
square, 407
sums of, 408–410
symmetric square, 407
trace of, 4
transpose of, 407
triangular, 408

unity, 408
Matrix product, 409
Matrix theory, 2
Maximum likelihood (ML) estimation, 27–33, 63–64, 400–402
restricted, 33–40, 63–64
Maximum likelihood estimators
invariance property of, lemmas on, 402–403
large sample variances of
for three-way nested classification, 351–353
for two-way nested classification, 306
for one-way classification, 103–105
Mean squared error (MSE) criterion, 404
Mean squares
expected, *see* Expected mean squares
unweighted, 22–23
Means estimators, variances of analysis of, for two-way crossed classification with interaction, 230
Method of fitting constants, 18
Minimal sufficient statistics, for one-way classification, 96–97
Minimum mean square quadratic estimators (MIMSQE), 53
Minimum mean squared error (MSE) quadratic estimation, 50–53
Minimum-norm quadratic estimation (MINQE), 49
Minimum-norm quadratic unbiased estimation, *see* MINQUE
Minimum-variance quadratic unbiased estimation, *see* MIVQUE
Minor of element of a matrix, 411
MINQUE (minimum-norm quadratic unbiased estimation), 13, 40–50, 64, 69
comments on, 48–50
development of, 43–47
formulation of, 41–43
iterative, 64
for one-way classification, 109, 111–113
MIVQUE (minimum-variance quadratic unbiased estimation), 40–50, 64
comments on, 48–50
development of, 47–48
formulation of, 41–43
for one-way classification, 109–111
Mixed effects, expectation under, 8–9
ML3, 421–422
Mln and MlwiN, 421–422
Moriguti–Bulmere procedure, 143
Multilevel linear models, 56–57
Muse designs, 234, 235

Naqvi's goodness-of-fit estimators, for one-way classification, 108
Natural estimator, 43
Newton–Raphson method, 38–39, 415–416
Neyman–Pearson theorem, 405
Nonlinear mixed models (NLMM), 61
Nonnegative minimum MSE estimator (MIMSQUE), 135
Nonnegative quadratic unbiased estimation, 53–55
Nonsingular matrix, 411
Null matrix, 408
Numerical example

general r-way nested classification, 385–387
one-way classification, 114–117, 145–146
three-way and higher-order crossed classifications, 268–272
three-way nested classification, 342–347, 360–362
two-way crossed classification with interaction, 221–226, 227, 239–240, 242–243
two-way crossed classification without interaction, 181–186, 192–193
two-way nested classification, 299–303, 315–317, 320–321

OD3 design, 234
One-way classification
analysis of variance estimators for, 97–99
analysis of variance for, 94–96
Bayesian estimation in, 117–122
best quadratic unbiased estimators for, 106–108
BMDP application for, 116
classical estimation for, 97–117
comparisons of designs and estimators for, 131–136
confidence intervals for, 136–146
distribution and sampling variances of estimators for, 122–131
distribution theory for, 96–97
fitting-constants-method estimators for, 99–100
general mean for, 102–103
intraclass correlation for, 113–114
mathematical model for, 93
maximum likelihood estimators for, 103–105
minimal sufficient statistics for, 96–97
MINQUE for, 109, 111–113
MIVQUE for, 109–111
Naqvi's goodness-of-fit estimators for, 108
numerical example, 114–117, 145–146
restricted maximum likelihood estimators for, 105–106
sampling variances of estimators for, 124–131
SAS application for, 116
SPSS application for, 116
symmetric sums estimators for, 100–101
tests of hypotheses for, 146–149
unbiased estimator for, 113
Orthogonal square matrix, 407

Partitioned matrices, 412–413
Pearson estimator for the intraclass correlation, 114
Point estimators, 403–404
Population mean, 25
estimation of, in random effects models, 25–27
Prairie–Anderson design, 307–308
Principal components (PC) method, 358
Prior measure, 50
Probability density function, 123
PROC GLM, PROC NESTD, PROC VARCOMP, and PROC MIXED, 418–419

Products of matrices, 409–410
Pseudo F-test, 399

Quadmissible estimators, 98
Quadratic forms, 4–6
Quadratic least squares (QLS) theory, 61–62

Random effects, 58
    expectation under, 9
Random effects models, xix–xxi; *see also* Mathematical model
    for discrete and categorical data, 56
    estimation of population mean in, 25–27
Rank of a matrix, 410
Rao–Blackwell theorem, 403
Residual maximum likelihood, 34*n*
Restricted maximum likelihood (ML) estimation, 33–40, 63–64
Restricted maximum likelihood estimators, for one-way classification, 105–106
RMEL program, 40
Row-echelon normal forms, 68
Row vector, 407

S designs, 232
S-Plus, 420
Sampling variances of estimators, for one-way classification, 124–131
SAS application, 418–419
    for one-way classification, 116
    for three-way and higher-order crossed classifications, 273
    for three-way nested classification, 344, 348
    for two-way crossed classification with interaction, 227
    for two-way crossed classification without interaction, 186
    for two-way nested classification, 301, 304
Satterthwaite-like test procedures, 319
Satterthwaite procedure, 397–400
Satterthwaite-type tests, 242
Scalar, 407
Scalar multiplication of a matrix, 409
Searle's result for sampling variance, 125, 127
Singular matrix, 411
Software for variance component analysis, 418–422
Spjøtvoll–Thomsen test, 193, 242
SPSS application, 419
    for one-way classification, 116
    for three-way and higher-order crossed classifications, 273
    for three-way nested classification, 345, 348
    for two-way crossed classification with interaction, 227
    for two-way crossed classification without interaction, 186
    for two-way nested classification, 302, 304
Square matrix, 407
Sufficient statistics, complete, 403
Sums of matrices, 408–410
Sums of squares (SSs), 15–16
Symmetric square matrix, 407
Symmetric sums estimators
    for general r-way nested classification, 377–381

for one-way classification, 100–101
for three-way and higher-order crossed classifications, 262–263
for three-way nested classification, 338-341
for two-way crossed classification with interaction, 217–221
for two-way crossed classification without interaction, 178–181
for two-way nested classification, 296–298
Symmetric sums method, 23–25, 63
“Synthesis” method of Hartley, 258

Tables, list of, xv–xvii
Tests of hypotheses
for general r-way nested classification, 382–384
for one-way classification, 146–149
for three-way nested classification, 359–360
for two-way crossed classification with interaction, 240–243
for two-way crossed classification without interaction, 193–195
for two-way nested classification, 317–321
Thitakamol designs, 237
Thomas–Hultquist procedure, 140–141, 143
Thomas–Hultquist–Donner procedure, 142
Three-way and higher-order crossed classifications, 255–272
analysis of variance estimators for, 261–262
analysis of variance for, 256–258
BMDP application for, 273
estimation of variance components for, 261–264
expected mean squares for, 258–261, 289–292
mathematical model for, 255–256
numerical example, 268–272
SAS application for, 273
SPSS application for, 273
symmetric sums estimators for, 262–263
variances of estimators for, 264–266
Three-way nested classification, 329–362
analysis of variance estimators for, 336–337
analysis of variance for, 330–331
BMDP application for, 345, 348
comparisons of designs and estimators for, 353–359
confidence intervals for, 359
estimation of variance components for, 336–347
expected mean squares for, 331–334
large sample variances of maximum likelihood estimators for, 351–353
mathematical model for, 329–330
numerical example, 342–347, 360–362
SAS application for, 344, 348
SPSS application for, 345, 348

symmetric sums estimators for, 338-341
tests of hypotheses for, 359–360
unweighted means analysis for, 334–336
unweighted means estimators for, 337–338
variances of analysis of variance estimators for, 347–351
variances of estimators for, 347–353

Trace of a matrix, 4
Translation invariance, 101, 297
Transpose of a matrix, 407
Triangular matrix, 408
Truncated ANOVA (TANOVA), 358
Two-way crossed classification with interaction, 201–243
analysis of means estimators for, 213–217
analysis of variance estimators for, 207–208
analysis of variance for, 201–203
BMDP application for, 227
comparisons of designs and estimators for, 230–238
confidence intervals for, 238–240
estimation of variance components for, 204–226
expected mean squares for, 204–206
fitting-constants-method estimators for, 208–213
mathematical model for, 201
numerical example, 221–226, 227, 239–240, 242–243
SAS application for, 227
SPSS application for, 227
symmetric sums estimators for, 217–221
tests of hypotheses for, 240–243
unweighted means analysis for, 213–215
variances of analysis of means estimators for, 230
variances of analysis of variance estimators for, 227–229
variances of estimators for, 226–230
variances of fitting-constants-method estimators for, 229
weighted means analysis for, 215–217

Two-way crossed classification without interaction, 165–195
analysis of means estimators for, 174–178
analysis of variance estimators for, 169–170
analysis of variance for, 165–167
BMDP application for, 186
confidence intervals for, 190–193
estimation of variance components for, 169–186
expected mean squares for, 167–169
fitting-constants-method estimators for, 170–174
mathematical model for, 165
numerical example, 181–186, 192–193
SAS application for, 186
SPSS application for, 186
symmetric sums estimators for, 178–181

tests of hypotheses for, 193–195
unweighted means analysis for, 174–176
variances of analysis of variance estimators for, 186–189
variances of estimators for, 186–190
variances of fitting-constants-method estimators for, 189–190
weighted means analysis for, 176–178
Two-way nested classification, 287–321
analysis of variance estimators for, 295
analysis of variance for, 288–289
BMDP application for, 302, 304
comparisons of designs and estimators for, 306–310
confidence intervals for, 310–317
distribution theory for, 292–293
estimation of variance components for, 295–303
large sample variances of maximum likelihood estimators for, 306
mathematical model for, 287–288
numerical example, 299–303, 315–317, 320–321
SAS application for, 301, 304
SPSS application for, 302, 304
symmetric sums estimators for, 296–298
tests of hypotheses for, 317–321
unweighted means analysis for, 293–294
unweighted means estimators for, 295–296
variances of analysis of variance estimators for, 304–306
variances of estimators for, 303–306

Unbalanced data, 1–2
Unbalanced r-way nested classification, see General r-way nested classification
Unbalanced three-way crossed-classification with interaction, *see* Three-way and higher-order crossed classifications
Unbalanced two-way crossed classification with interaction, *see* Two-way crossed classification with interaction
Unbalanced two-way nested classification, *see* Two-way nested classification
Unbiased estimator of the ratio of variance components, for one-way classification, 113
Unbiasedness for the ANOVA estimators, property of, 16
Uniformly optimum tests, 147
Unity matrix, 408
Unity vector, 408
Unweighted mean squares, 22–23
Unweighted means analysis
for three-way nested classification, 334–336

for two-way crossed classification with interaction, 213–215
for two-way crossed classification without interaction, 174–176
for two-way nested classification, 293–294
Unweighted means estimators
for three-way nested classification, 337–338
for two-way nested classification, 295–296

VARCL, 421–422
VARCOMP procedure, 419
Variance component analysis, software for, 418–422
Variance components
in covariance matrix of T, coefficients of products of, 276–285
estimation of, *see* Estimation of variance components
inferences about, 13–70
Variances of analysis of means estimators, for two-way crossed classification with interaction, 230
Variances of analysis of variance estimators
for three-way nested classification, 347–351
for two-way crossed classification with interaction, 227–229
for two-way crossed classification without interaction, 186–189
for two-way nested classification, 304–306
Variances of estimators
for general r-way nested classification, 382
for three-way and higher-order crossed classifications, 264–266
for three-way nested classification, 347–353
for two-way crossed classification with interaction, 226–230
for two-way crossed classification without interaction, 186–190
for two-way nested classification, 303–306
Variances of fitting-constants-method estimators
for two-way crossed classification with interaction, 229
for two-way crossed classification without interaction, 189–190
Vectors, definition of, 407
differentiation of, 413–415
unity, 408

W-transformation, 39
Wald's procedure for the confidence interval, 137–140
Weighted analysis of means (WAM) estimator, 299
Weighted means analysis
for two-way crossed classification with interaction, 215–217
for two-way crossed classification without interaction, 176–178