

Inhaltsverzeichnis

1	Wahrscheinlichkeitsmodell und statistisches Modell	2
2	Lebensdauerverteilungen	7
2.1	Der Begriff der Lebensdauerverteilung	7
2.2	Momente von Lebensdauerverteilungen	10
2.3	Einige spezielle Verteilungen	11
2.3.1	Die Binomialverteilung	12
2.3.2	Die geometrische Verteilung	12
2.3.3	Die Poissonverteilung	13
2.3.4	Die Normalverteilung	13
2.3.5	Die logarithmische Normalverteilung	15
2.3.6	Die Exponentialverteilung	15
2.3.7	Die zweiparametrische Exponentialverteilung	16
2.3.8	Verteilungen vom Pareto-Typ	16
2.3.9	Weibull-Verteilungen	17
2.3.10	Verteilungen mit Lage- und Maßstabsparameter	18
2.3.11	Gomperz-Makeham-Verteilungen	19
2.3.12	Cox'sche Verteilungen	19
3	Nichtparametrische Lebensdauerverteilungen	21
3.1	IFR-Verteilungen	21
3.2	Weitere Familien von Verteilungen	31
4	Ausfallmodelle	39
4.1	Abnutzungsmodelle	39
4.1.1	Die Birnbaum-Saunders-Verteilung	39
4.1.2	Die Inverse Gaußverteilung als Lebensdauerverteilung bei speziellen Abnutzungsprozessen	40
4.2	Erneuerungstheorie	44
4.3	Poissonsche Punktprozesse	53
4.4	Perkolationsmodelle	66

5	Punktschätzungen	71
5.1	Datenstrukturen	71
5.2	Parameterschätzungen	79
5.3	Nichtparametrische Schätzungen	86
6	Tests und Konfidenzschätzungen	96
6.1	Graphische Methoden	96
6.2	Signifikanztests	102
6.3	Konfidenzschätzungen	109
7	Nichtparametrische Schätzungen für zensierte Daten	125
7.1	Produkt-Limit-Schätzer	125
7.2	Actuarial Schätzung	127
7.3	Schätzung der integrierten Hazard-Funktion	128
8	Proportional Hazard und Accelerated Life Testing	129
8.1	Accelerated life testing	129
8.2	Proportional Hazard	131
8.3	Mischverteilungen und Frailty-Modelle	135
A	Wahrscheinlichkeitsmodell Exponentialverteilung	137
A.1	Einparametrische Exponentialverteilung	137
A.1.1	Beendigung beim r^* -ten Ausfall ohne Ersetzen der ausgefallenen Elemente	137
A.1.2	Beendigung nach t^* Stunden ohne Ersetzen der ausgefallenen Elemente	142
A.1.3	Beendigung nach r^* Ausfällen mit Ersetzen der ausgefallenen Elemente	144
A.1.4	Beendigung nach t^* Stunden mit Ersetzen der ausgefallenen Elemente	146
A.2	Zweiparametrische Exponentialverteilung	149
A.2.1	Beendigung beim r^* -ten Ausfall ohne Ersetzen der ausgefallenen Elemente	149
A.2.2	Beendigung zum Zeitpunkt t^* ohne Ersetzen der ausgefallenen Elemente	152
A.2.3	Beendigung beim r^* -ten Ausfall mit Ersetzen der ausgefallenen Elemente	152
A.2.4	Beendigung zum Zeitpunkt t^* mit Ersetzen der ausgefallenen Elemente	153

Survival Analysis

Literatur:

1. Andersen, Borgan, Gill, Keiding: *Statistical Models Based on Counting Processes*. Springer-Verlag, Berlin, 1993.
2. Aven, Jensen: *Stochastic Models in Reliability*. Springer-Verlag, New-York, 1999.
3. Bain: *Statistical Analysis of Reliability and Life-Testing Models*. Marcel Dekker, New-York, 1978
4. Barlow, Proshan: *Statistische Theorie der Zuverlässigkeit*. Akademie-Verlag, Berlin, 1981.
5. Belyaev, Kahle: *Methoden der Wahrscheinlichkeitsrechnung und Statistik bei der Analyse von Zuverlässigkeitsdaten*. Teubner, Stuttgart, 2000.
6. Cox, Oakes: *Analysis of Survival Data*. Chapman & Hall, London, 1984.
7. Cox, Smith: *Renewal Theory*. John Wiley, New York, 1962.
8. Fleming, Harrington: *Counting Process and Survival Analysis*. Wiley, New-York, 1991.
9. Gertsbakh: *Reliability Theory*. Springer-Verlag, Berlin Heidelberg, 2000.
10. Härtler: *Statistische Methoden für die Zuverlässigkeitsanalyse*. VEB Verlag Technik, Berlin, 1983.
11. Hougaard: *Analysis of Multivariate Survival Data*. Springer-Verlag, New-York, 2000.
12. Kalbfleisch, Prentice: *The Statistical Analysis of Failure Time Data*. John Wiley, New York, 1980.
13. Klein, Moeschberger: *Survival Analysis: technics for censored and truncated data*. Springer-Verlag, New-York, 2003.
14. Lawless: *Statistical Models and Methods for Lifetime Data*. Wiley, New-York, 1991.
15. Mann, Schafer, Singpurwalla: *Methods for Statistical Analysis of Reliability and Life Data*. John Wiley, New York, 1974.

1 Wahrscheinlichkeitsmodell und statistisches Modell

Sei $(\Omega = (\omega), \mathfrak{B}(\Omega))$ ein meßbarer Raum. Ω ist die Menge von Elementarereignissen ω und $\mathfrak{B}(\Omega)$ ist eine σ -Algebra von Teilmengen (zufälligen Ereignissen) aus Ω .

Sei ferner $P_\Omega = P_\Omega(\cdot)$ ein Maß auf $\mathfrak{B}(\Omega)$ mit den Eigenschaften:

- (i) $P_\Omega(B) \geq 0 \quad \forall B \in \mathfrak{B}(\Omega)$,
- (ii) $P_\Omega(\Omega) = 1$,
- (iii) $P_\Omega(\cup_{k=1}^{\infty} B_k) = \sum_{k=1}^{\infty} P_\Omega(B_k) \quad \text{für } B_i \cap B_j = \emptyset, i \neq j$.

Definition 1.1 *Das Tripel $(\Omega, \mathfrak{B}(\Omega), P_\Omega)$ heißt allgemeiner Wahrscheinlichkeitsraum.*

Wir betrachten nun eine meßbare Abbildung $X : \Omega \rightarrow \mathfrak{X}$ und $x = X(\omega)$. Diese Abbildung ist in Bild 1.1 dargestellt.

Sei $\mathfrak{B}(\mathfrak{X})$ eine σ -Algebra von Teilmengen aus \mathfrak{X} . Dann bedeutet die Meßbarkeit der Abbildung X , daß

$$X^{-1}(A) = (\omega : X(\omega) \in A) \in \mathfrak{B}(\Omega) \quad \forall A \in \mathfrak{B}(\mathfrak{X}).$$

Die Funktion $X = X(\cdot)$ heißt Zufallsgröße. Die Zufallsgröße $X : \Omega \rightarrow \mathfrak{X}$ besitzt auf dem

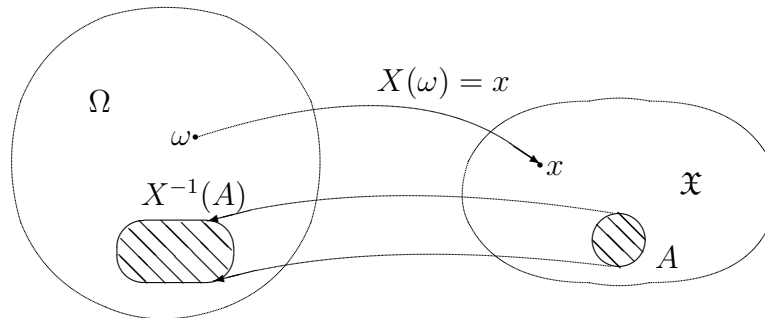


Abbildung 1.1: Wahrscheinlichkeitsraum und Zufallsgröße

meßbaren Raum $(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}))$ das induzierte Wahrscheinlichkeitsmaß $P(A) = P_\Omega(X^{-1}(A))$.

Definition 1.2 *Das Tripel $(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}), P)$ heißt Wahrscheinlichkeitsmodell.*

In der *Wahrscheinlichkeitsrechnung* ist das Maß $P(A)$ für jedes $A \in \mathfrak{B}(\mathfrak{X})$ wohlbekannt und ihr Hauptinhalt besteht darin, Wahrscheinlichkeiten für interessierende Ereignisse und darauf aufbauende Kenngrößen zu berechnen.

Eine ganz andere Situation liegt in der *mathematischen Statistik* vor. Hier nimmt man an, daß auf $(\Omega, \mathfrak{B}(\Omega))$ ein „wahres“ Wahrscheinlichkeitsmaß P^* existiert. Dieses Maß ist unbekannt. Die elementaren Ereignisse sind ebenfalls unbekannt und können nicht beobachtet werden. Beobachtet werden können nur die Ergebnisse eines zufälligen Experimentes. Sei $\mathfrak{X} = (x)$ der Stichprobenraum, der alle möglichen beobachtbaren Resultate enthält. Da jedem zufälligen Experiment eine meßbare Abbildung $X : \Omega \rightarrow \mathfrak{X}$ und $x = X(\omega)$ entspricht, betrachtet man wiederum den Raum $(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}), P)$, jedoch das Maß P ist nicht oder nur teilweise bekannt. Daher muß man eine Familie von Verteilungen $\mathfrak{P}_X = (P_\theta, \theta \in \Theta)$ betrachten, wobei Θ ein Parameterraum ist. Durch die Parameter werden die Verteilungen voneinander unterschieden. Diese Familie muß bestimmte Eigenschaften besitzen. So muß jedes θ eineindeutig eine Verteilung aus dieser Familie definieren, d.h. $\forall \theta_1 \neq \theta_2 \exists A : P_{\theta_1}(A) \neq P_{\theta_2}(A)$. Wir setzen stets voraus, daß die Familie $\mathfrak{P}_X = (P_\theta, \theta \in \Theta)$ die „wahre“ Verteilung P_{θ^*} enthält. Mittels eines statistischen Experimentes soll aus der Familie \mathfrak{P}_X eine Verteilung ausgewählt werden, die der „wahren“ Verteilung möglichst gut entspricht.

Auf der Grundlage des allgemeinen Wahrscheinlichkeitsraumes $(\Omega, \mathfrak{B}(\Omega), P)$ wird ein zufälliges Experiment durch die Abbildung $X : \Omega \rightarrow \mathfrak{X}$ durchgeführt; als Resultat erhält man die statistischen Daten $X(\omega) = x$. Hierbei sei x fixiert und bekannt, das wahre Maß P_{θ^*} ist unbekannt. Mit Hilfe von x werden Aussagen über P_{θ^*} , oder, was das gleiche ist, über θ^* getroffen. Diese Aussagen nennt man statistische Schlußweisen (statistical inference) und das Tripel $(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}), \mathfrak{P}_X = (P_{X,\theta}, \theta \in \Theta))$ nennt man *statistisches Modell*.

Definition 1.3 *Unter einem statistischen Modell versteht man das Tripel $(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}), \mathfrak{P})$ mit*

- $\mathfrak{X} = (x)$ - Stichprobenraum
- $\mathfrak{B}(\mathfrak{X})$ - σ -Algebra über \mathfrak{X}
- $\mathfrak{P} = (P_F, F \in \mathfrak{F})$ - eine Familie von Verteilungen auf $(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}))$.

Es ist günstig, jede Verteilung $F(\cdot) \in \mathfrak{F}$ eindeutig durch einen Parameter $\theta \in \Theta$ zu charakterisieren. Θ ist der Parameterraum und Eindeutigkeit der Zuordnung bedeutet, daß verschiedenen Werten von θ verschiedene Verteilungen aus der Familie \mathfrak{F} entsprechen. Dieser Zugang ist auch bei nichtparametrischen Aufgabenstellungen möglich. Darauf wird später in Zusammenhang mit nichtparametrischen Schätzungen eingegangen. Dann bezeichnen wir mit $F_\theta(\cdot)$ oder $F(\cdot, \theta)$ die Verteilungsfunktion $F(\cdot) \in \mathfrak{F}$, die durch den Parameter θ bestimmt wird. Im folgenden bezeichnen wir mit $P_\theta = P_{F_\theta}$ die entsprechende Wahrscheinlichkeitsverteilung. Damit besteht das statistische Modell aus dem Tripel

$$(\mathfrak{X} = (x), \mathfrak{B}(\mathfrak{X}), \mathfrak{P} = (P_\theta, \theta \in \Theta)) .$$

Häufig nimmt man an, daß die wahre Verteilung $P^* = P_{\theta^*} \in \mathfrak{P}$ ist. Es ist jedoch auch möglich, anhand statistischer Daten zu überprüfen, ob $P^* \in \mathfrak{P}$ gilt.

Bei statistischen Untersuchungen arbeitet man mit den erhaltenen statistischen Daten. Bei Berechnungen werden diese statistischen Daten $x \in \mathfrak{X}$ in neue Größen y umgeformt. Diesen Übergang kann man als Abbildung von \mathfrak{X} in \mathfrak{Y} , $y \in \mathfrak{Y}$ auffassen. Diese Abbildung muß ohne Kenntnis des unbekannten Verteilungsgesetzes erfolgen können. Damit kommen wir zu dem wichtigen Begriff der Statistik.

Definition 1.4 Die Abbildung $Y : \mathfrak{X} \rightarrow \mathfrak{Y}$ heißt Statistik, wenn

(i) sie nicht vom unbekannten Parameter abhängt,

(ii) sie $(\mathfrak{B}(\mathfrak{X}), \mathfrak{B}(\mathfrak{Y}))$ - meßbar ist, d.h. wenn

$$B = Y^{-1}(A) = \{x : Y(x) \in A\} \in \mathfrak{B}(\mathfrak{X}) \quad \forall A \in \mathfrak{B}(\mathfrak{Y}) ,$$

(iii) jede Einpunktmenge $\{y\}$ zu $\mathfrak{B}(\mathfrak{Y})$ gehört, d.h.

$$A_Y(y) = \{x : Y(x) = y\} \in \mathfrak{B}(\mathfrak{X}), \forall y \in \mathfrak{B}(\mathfrak{Y}).$$

Die Menge $A_Y(y)$ heißt Atom der Statistik Y zum Niveau y .

Die Genauigkeit statistischer Schlüsse hängt im wesentlichen vom Grad der Information ab, die aus einem statistischen Experiment gewonnen wird. In der klassischen Statistik ist im allgemeinen der Stichprobenumfang die wesentliche Einflußgröße. In der Zuverlässigkeitsstatistik spielen Stichproben anderer Struktur, z. B. zensierte Daten, eine wesentliche Rolle. Wir betrachten dazu ein einführendes Beispiel.

Beispiel 1.1 Gegeben seien n gleichartige Systeme, die jeweils aus m verschiedenen Elementen bestehen. Damit sind insgesamt $n \cdot m$ Elemente vorhanden, jeweils n von der Sorte i , $i = 1, \dots, m$. Die Elemente werden fortlaufend numeriert: $(i, 1), \dots, (i, n)$, $i = 1, \dots, m$, ist die Numerierung der i -ten Sorte von Elementen und das j -te System besteht aus den Elementen $(1, j), \dots, (m, j)$, $j = 1, \dots, n$. Jedes Element mit der Nummer (i, j) hat zwei Zustände:

$$Z_{ij} = \begin{cases} 0, & \text{wenn das Element ausgefallen ist,} \\ 1, & \text{wenn das Element intakt ist,} \end{cases}$$

und jedes System sei nur intakt, wenn alle seine Elemente intakt sind. Dieses ist der typische Fall einer Reihenschaltung. Dann gilt für das j -te System

$$\Psi_j = Z_{1j} \cdot Z_{2j} \cdot \dots \cdot Z_{mj} = \begin{cases} 0, & \text{wenn das } j\text{-te System ausgefallen ist,} \\ 1, & \text{wenn das } j\text{-te System intakt ist.} \end{cases}$$

Die Zustände der Elemente Z_{ij} seien unabhängige und für jede Sorte i identisch verteilte Zufallsgrößen mit $P(Z_{ij} = 1) = p_i$.

Für den Erwartungswert von Z_{ij} erhalten wir

$$EZ_{ij} = 1 \cdot P(Z_{ij} = 1) + 0 \cdot P(Z_{ij} = 0) = p_i .$$

Sei $p_s = P(\Psi_j = 1)$ die Wahrscheinlichkeit dafür, daß das j -te System intakt ist. Dann erhält man aus der Unabhängigkeit der Zustände

$$p_s = E\Psi_j = EZ_{1j} Z_{2j} \dots Z_{mj} = EZ_{1j} \cdot EZ_{2j} \dots EZ_{mj} = p_1 \dots p_m .$$

In der Wahrscheinlichkeitstheorie sind p_1, \dots, p_m bekannte Größen, mittels derer Wahrscheinlichkeiten für interessierende Ereignisse berechnet werden können. So ist z.B.

$$\sum_{l \geq k} \binom{n}{l} p_s^l (\bar{p}_s)^{n-l}, \quad \bar{p}_s = 1 - p_s$$

die Wahrscheinlichkeit, daß mindestens k Systeme intakt sind.

In der mathematischen Statistik dagegen sind die wahren Intaktwahrscheinlichkeiten p_1^*, \dots, p_m^* der Elemente oder p_s^* des Systems unbekannt und sollen mittels einer Beobachtung näherungsweise bestimmt werden. Wir nehmen an, daß uns die Intaktwahrscheinlichkeit p_s^* des Systems interessiert und betrachten zwei mögliche Stichprobenräume. Im ersten Fall, der im folgenden mit \mathfrak{X}_v bezeichnet wird, kann man alle Elemente der Matrix (Z_{ij}) , $i = 1, \dots, m$, $j = 1, \dots, n$, beobachten. Dabei steht v für vollständige Information. Die Anzahl der Elemente im Stichprobenraum \mathfrak{X}_v ist in diesem Fall $\text{card } \mathfrak{X}_v = 2^{n \cdot m}$, da $n \cdot m$ Elemente vorhanden sind und jedes Element 2 Zustände annehmen kann.

Im zweiten Fall beobachtet man nur die Zustände der n Systeme, d.h. den Vektor (Ψ_1, \dots, Ψ_n) . Der Stichprobenraum wird hier mit \mathfrak{X}_u bezeichnet und u steht für unvollständige Information. Die Anzahl der Elemente in diesem Stichprobenraum ist $\text{card } \mathfrak{X}_u = 2^n$.

Im Falle vollständiger Information kann die Intaktwahrscheinlichkeit des Systems aus den Intaktwahrscheinlichkeiten der Elemente ermittelt werden. Für die Elemente ist die relative Häufigkeit

$$\hat{p}_i = \frac{1}{n} \sum_{j=1}^n Z_{ij}, \quad i = 1, \dots, m, \quad \hat{p}_i : \mathfrak{X}_v \rightarrow [0, 1]$$

eine mögliche Punktschätzung der unbekannten Wahrscheinlichkeit.

Die Intaktwahrscheinlichkeit des Systems p_s^* ist eine Funktion der Intaktwahrscheinlichkeiten der Elemente: $p_s^* = p_1^* \cdot p_2^* \dots p_m^*$ und kann daher im Falle vollständiger Information durch

$$\hat{p}_{vs} = \prod_{i=1}^m \hat{p}_i = \prod_{i=1}^m \left(\frac{1}{n} \sum_{j=1}^n Z_{ij} \right)$$

geschätzt werden. Da alle Faktoren Z_{ij} unabhängig sind, gilt für den Erwartungswert dieser Schätzung

$$E\hat{p}_{vs} = \prod_{i=1}^m \left(\frac{1}{n} \sum_{j=1}^n EZ_{ij} \right) = \prod_{i=1}^m \left(\frac{1}{n} n p_i \right) = p_s$$

und damit ist die Punktschätzung \hat{p}_{vs} erwartungstreu.

Im Falle unvollständiger Information sind nur die Größen Ψ_j beobachtbar, daher muß man eine andere Punktschätzung betrachten:

$$\hat{p}_{us} = \frac{1}{n} \sum_{j=1}^n \Psi_j .$$

Diese Punktschätzung ist ebenfalls erwartungstreu. Da $\Psi_j = \prod_{i=1}^m Z_{ij}$ ist, gilt

$$\mathbb{E} \hat{p}_{us} = \mathbb{E} \left(\frac{1}{n} \sum_{j=1}^n \prod_{i=1}^m Z_{ij} \right) = \frac{1}{n} \sum_{j=1}^n \mathbb{E} \left(\prod_{i=1}^m Z_{ij} \right) = \frac{1}{n} \sum_{i=1}^n \prod_{i=1}^m p_i = p_s .$$

Damit haben wir gezeigt, daß beide Schätzungen erwartungstreu sind. Es ist jedoch klar, daß mehr Information zu einer besseren Schätzung führen muß. Das wird deutlich, wenn man die Varianzen der Schätzungen betrachtet. Man kann sich leicht davon überzeugen, daß $\text{Var}(\hat{p}_{vs}) \leq \text{Var}(\hat{p}_{us})$ ist und damit vollständige Information zu genaueren Schätzungen führt. \diamond

2 Lebensdauerverteilungen

2.1 Der Begriff der Lebensdauerverteilung

Definition 2.1 Die Funktion $F(s) = P(S \leq s)$, $s \geq 0$ heißt (Lebensdauer-) Verteilungsfunktion. $\bar{F}(s) = 1 - F(s) = P(S > s)$ ist die Wahrscheinlichkeit, daß bis zur Zeit einschließlich s kein Ausfall stattgefunden hat und heißt Überlebens- oder Zuverlässigkeitsfunktion (survival function).

Diese beiden Funktionen sind in Abbildung 2.1 dargestellt. $F(s)$ und $\bar{F}(s)$ sind rechtsseitig stetig. Bezeichnen wir mit $\bar{F}(s-) = \lim_{h \downarrow 0} \bar{F}(s-h)$, so erhalten wir $\bar{F}(s-) = P(S \geq s)$, $F(s-) = P(S < s)$. Die Funktion $\bar{F}(s-)$ ist linksseitig stetig. Die Verteilungsfunk-

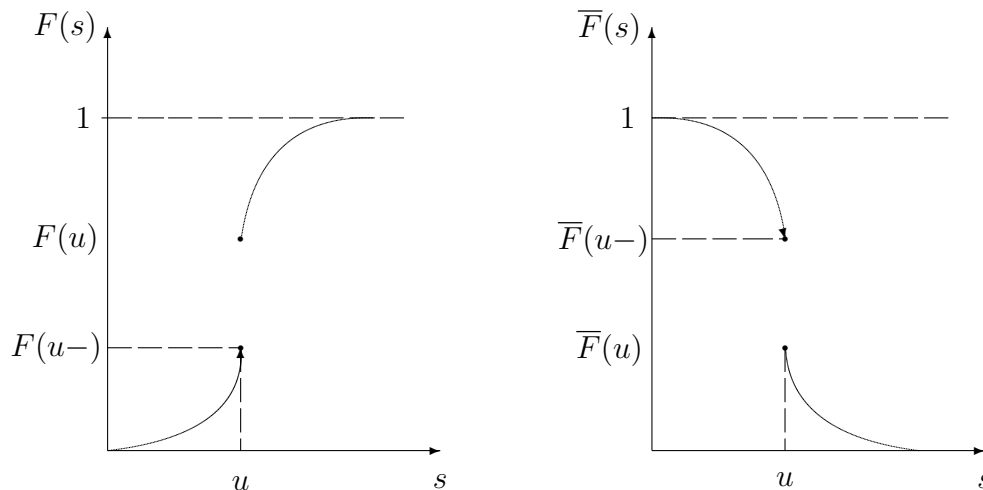


Abbildung 2.1: Lebensdauerverteilungsfunktion und Überlebensfunktion

tion $F(\cdot)$ besitzt einen Sprung im Punkt u , wenn $\Delta F(u) = F(u) - F(u-) > 0$. Es sei angemerkt, daß $P(S = u) = \Delta F(u)$ gilt (siehe auch Abbildung 2.1).

Alle Sprünge der Lebensdauerverteilungsfunktion können summiert werden und stellen den reinen Sprunganteil dieser Funktion dar:

$$F^{(d)}(s) = \sum_{u \leq s} \Delta F(u) .$$

Der stetige Anteil der Lebensdauerverteilungsfunktion kann nun folgendermaßen definiert werden:

$$F^{(c)}(s) = F(s) - \sum_{u \leq s} \Delta F(u) .$$

Damit besteht jede Verteilungsfunktion aus einem stetigen Anteil und einem reinen Sprunganteil: $F(\cdot) = F^{(d)}(\cdot) + F^{(c)}(\cdot)$.

Bemerkung: Man kann beweisen, daß der stetige Anteil der Verteilungsfunktion wiederum aus einem absolut stetigen Anteil $F^{(ac)}(\cdot)$ und einem singulären Anteil $F^{(sc)}(\cdot)$ besteht: $F^{(c)}(\cdot) = F^{(ac)}(\cdot) + F^{(sc)}(\cdot)$. Für jeden absolut stetigen Anteil existiert eine Dichte $f(\cdot)$ so, daß

$$F^{(ac)}(s) = \int_0^s f(u) du, \quad s \geq 0$$

gilt und jeder singuläre Anteil ist stetig, wobei alle Wachstumspunkte in einer Menge vom Lebesgue-Maß 0 liegen.

Damit läßt sich jede Lebensdauerverteilungsfunktion in folgender Form darstellen:

$$F(s) = F^{(d)}(s) + F^{(sc)}(s) + \int_0^s f(u) du, \quad s \geq 0.$$

In der Zuverlässigkeitstheorie arbeitet man häufig mit absolut stetigen Modellen. In diesem Fall gilt

$$F^{(d)}(\cdot) \equiv 0, \quad F^{(sc)}(\cdot) \equiv 0, \quad F(s) = \int_0^s f(u) du, \quad s \geq 0.$$

In der Statistik sind jedoch auch reine diskrete Modelle sehr wesentlich, bei denen

$$F(s) = F^d(s) = \sum_{u \leq s} \Delta F(u)$$

ist.

Im folgenden soll der wichtige Begriff der Ausfallrate (hazard intensity function) eingeführt werden. Dazu benötigen wir einen allgemeineren Begriff des Differentials, der bei Lebesgue–Stieltjes–Integralen Verwendung findet. Wir definieren das Differential als $dF(u) := P(u \leq S < u + du) = F(u + du-) - F(u-)$, wobei du infinitesimal klein ist. Wenn in einer Umgebung des Punktes u die Funktion F stückweise konstant mit einem Sprung in u ist, dann gilt $dF(u) = \Delta F(u)$. Bei absolut stetigen Funktionen F gilt $dF(u) = f(u)du$. Wir fixieren nun einen Zeitpunkt u und nehmen an, daß das Element bis zu einer beliebigen Zeit $u' < u$ intakt war. Die bedingte Wahrscheinlichkeit, daß das Element zur Zeit u ausfällt, unter der Bedingung, daß es vor u intakt war, ist dann

$$\begin{aligned} P(u \leq S < u + du \mid S \geq u) &= \frac{P(\{u \leq S < u + du\} \cap \{S \geq u\})}{P(S \geq u)} = \\ &= \frac{P(u \leq S < u + du)}{P(S \geq u)} = \frac{dF(u)}{\bar{F}(u-)}. \end{aligned}$$

Integriert man diese Beziehung, so erhält man eine neue Funktion $H(\cdot) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$:

$$H(s) := \int_0^s \frac{dF(u)}{\bar{F}(u-)}, \quad s \geq 0. \quad (2.1)$$

Als Spezialfälle ergeben sich im im Fall absolut stetiger Verteilungen das Riemannsche Integral:

$$H(s) = \int_0^s \frac{f(u)}{\bar{F}(u)} du$$

und im Fall diskreter Verteilungen die Summe:

$$H(s) = \sum_{u \leq s} \frac{\Delta F(u)}{\bar{F}(u-)} .$$

Es gilt also

$$\Delta H(u) = \frac{\Delta F(u)}{\bar{F}(u-)} = \frac{\bar{F}(u-) - \bar{F}(u)}{\bar{F}(u-)} \leq 1 . \quad (2.2)$$

Die Funktion $H(\cdot)$ nennt man *Hazardfunktion* oder *integrierte Ausfallrate*. $H(\cdot)$ ist nicht-fallend, rechtsseitig stetig und kann Sprünge der Höhe ≤ 1 besitzen. Wie die Verteilungsfunktion besitzt auch die Hazardfunktion folgende Darstellung:

$$H(s) = H^{(d)}(s) + H^{(c)}(s) = H^{(d)}(s) + H^{(sc)}(s) + H^{(ac)}(s) ,$$

wobei der Sprunganteil und der absolut stetige Anteil durch

$$H^{(d)}(s) = \sum_{u \leq s} \Delta H(u) \quad \text{bzw.} \quad H^{(ac)}(s) = \int_0^s h(u) du$$

gegeben sind. Wenn man mit absolut stetigen Modellen arbeitet, gilt wiederum

$$H^{(d)}(\cdot) \equiv H^{(sc)}(\cdot) \equiv 0$$

und man erhält

$$H(s) = \int_0^s h(u) du . \quad (2.3)$$

In diesem Falle nennt man $h(u)$ *Ausfallrate* (hazard intensity function). Wenn $F(\cdot)$ stetig ist, so folgt aus $\bar{F}(u-) = \bar{F}(u)$

$$H(s) = - \int_0^s \frac{d\bar{F}(u)}{\bar{F}(u)} = - \ln \bar{F}(s) = \ln \frac{1}{\bar{F}(s)} \quad (2.4)$$

und umgekehrt:

$$\bar{F}(s) = \exp(-H(s)) . \quad (2.5)$$

Die Gleichungen (2.4) und (2.5) bestimmen im Falle stetiger Modelle eine eindeutige Beziehung zwischen $H(\cdot)$ und $\bar{F}(\cdot)$.

Satz 2.1 *Im allgemeinen Fall gilt*

$$\bar{F}(s) = \exp(-H^{(c)}(s)) \prod_{u \leq s} (1 - \Delta H(u)) . \quad (2.6)$$

Bemerkung: Die Integraldarstellung (2.1) besitzt bezüglich $\overline{F}(\cdot)$ eine eindeutige Lösung. Diese Lösung ist die in Gleichung (2.6) dargestellte Funktion $\overline{F}(\cdot)$. Eine ausführliche Darstellung des Beweises und andere wesentliche Fakten findet man in GILL, JOHANNSEN.

Die Formeln (2.1) und (2.6) bestimmen im allgemeinen Fall eine eindeutige Beziehung zwischen $H(\cdot)$ und $\overline{F}(\cdot)$.

2.2 Momente von Lebensdauerverteilungen

In diesem Abschnitt sollen Momente einer zufälligen Lebensdauer sowie der bedingte Erwartungswert definiert werden.

Wir bezeichnen den Erwartungswert von S mit

$$m_F = m_{F,1} = \int_0^\infty t dF(t) = \int_0^\infty t f(t) dt ,$$

wobei der letzte Term für absolut stetige Verteilungen gilt. Das zweite Moment wird mit

$$m_{F,2} = \int_0^\infty t^2 dF(t) ,$$

die Varianz mit

$$\sigma_F^2 = m_{F,2} - (m_{F,1})^2$$

und das k -te Moment mit

$$m_{F,k} = \int_0^\infty t^k dF(t) .$$

bezeichnet. Für Lebesgue–Stieltjes–Integrale besitzt die Formel der partiellen Integration folgende Form:

$$\begin{aligned} G_1(s)G_2(s) - G_1(0)G_2(0) &= \int_0^s G_1(t) dG_2(t) + \int_0^s G_2(t) dG_1(t) = \\ &= \int_0^s G_1(t-) dG_2(t) + \int_0^s G_2(t-) dG_1(t) + \sum_{0 \leq t \leq s} \Delta G_1(t) \Delta G_2(t) . \end{aligned}$$

Wenn das k -te Moment existiert ($m_{F,k} < \infty$), so kann man durch partielle Integration mit $s = \infty$, $G_1(t) = F(t)$, $G_2(t) = t^k$, $t \geq 0$ eine nützliche Beziehung erhalten:

$$\begin{aligned} m_{F,k} &= \int_0^\infty t^k dF(t) = - \int_0^\infty t^k d\overline{F}(t) = -t^k \overline{F}(t) \Big|_0^\infty + \int_0^\infty k \cdot t^{k-1} \overline{F}(t-) dt = \\ &= k \int_0^\infty t^{k-1} \overline{F}(t-) dt = k \int_0^\infty t^{k-1} \overline{F}(t) dt , \end{aligned} \tag{2.7}$$

da $\lim_{t \rightarrow \infty} t^k \overline{F}(t) = 0$, wenn $m_{F,k} < \infty$ ist. Speziell erhält man also für den Erwartungswert

$$m_F = \int_0^\infty \overline{F}(t-) dt = \int_0^\infty \overline{F}(t) dt .$$

Wir führen nun die Begriffe *bedingte Verteilung* und *bedingter Erwartungswert* ein. Wir betrachten zwei zufällige Vektoren $X \in R^p$, $Y \in R^q$ und ihre gemeinsame Verteilung $P_{X,Y}(A, B)$, wobei A und B beliebige Borelmengen mit $A \in R^p$ und $B \in R^q$ sind. Die Randverteilung des zufälligen Vektors X ist durch $P_X(A) = P_{X,Y}(A, R^q)$ gegeben. Als Randverteilung für Y erhält man entsprechend $P_Y(B) = P_{X,Y}(R^p, B)$.

Definition 2.2 Wenn für beliebige Borelmengen A und B die Beziehung

$$P_{X,Y}(A, B) = \int_A P_{Y|X}(B, x) P_X(dx) \quad (2.8)$$

gilt, so nennt man den Integrationskern $P_{Y|X}(\cdot, x)$ bedingte Verteilung der Zufallsgröße Y bezüglich eines (beliebigen) gegebenen Wertes x der Zufallsgröße X .

Die Existenz der bedingten Verteilung $P_{Y|X}(\cdot, x)$ folgt dabei aus dem Satz von RADON–NIKODYM. Wenn eine Dichtefunktion existiert, so nimmt die Gleichung (2.8) folgende Form an:

$$p_{X,Y}(x, y) = p_{Y|X}(y, x) p_X(x), \quad \text{wobei} \quad P_{Y|X}(B, x) = \int_B p_{Y|X}(y, x) dy. \quad (2.9)$$

Da für beliebige feste x die bedingte Verteilung $P_{Y|X}(\cdot, x)$ eine Wahrscheinlichkeitsverteilung darstellt, kann man Erwartungswerte betrachten, z.B.

$$\int_{R^q} g(y) P_{Y|X}(dy, x). \quad (2.10)$$

Definition 2.3 Die für beliebige Werte x von X definierte Zufallsgröße $E(g(Y)|X)$ mit

$$E(g(Y)|x) := \int_{R^q} g(y) P_{Y|X}(dy, x) \quad (2.11)$$

heißt bedingte Erwartung der Zufallsgröße $g(Y)$ bezüglich X .

Als Beispiel einer bedingten Verteilung wird im nächsten Abschnitt eine bedingte Normalverteilung betrachtet.

2.3 Einige spezielle Verteilungen

In diesem Abschnitt betrachten wir einige spezielle Verteilungen, die nicht nur in der Zuverlässigkeitstheorie, sondern auch in vielen anderen Anwendungsgebieten eine wesentliche Rolle spielen.

2.3.1 Die Binomialverteilung

Wir betrachten eine Folge voneinander unabhängiger Versuche $\mathcal{E}_1, \mathcal{E}_2, \dots$. In jedem dieser Versuche interessieren uns nur zwei Versuchsausgänge (das Eintreten eines zufälligen Ereignisses A bzw. des komplementären Ereignisses \bar{A}). Wir setzen voraus, daß die Wahrscheinlichkeit von A in jedem Versuch die gleiche ist: $P(A) = p$ ($0 < p < 1$). Eine solche Folge von Versuchen heißt *Bernoullisches Versuchsschema*.

Ausgehend von diesem Versuchsschema untersuchen wir die Zufallsgröße X – zufällige Anzahl der Versuche (von insgesamt n Versuchen), in denen A eintritt, d.h. die absolute Häufigkeit des Ereignisses A in n unabhängigen Wiederholungen eines zufälligen Versuchs. X kann die Werte $0, 1, \dots, n$ mit den zugehörigen Einzelwahrscheinlichkeiten

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

annehmen.

Definition 2.4 Eine diskrete Zufallsgröße X unterliegt einer Binomialverteilung mit den Parametern n und p , falls sie die Einzelwahrscheinlichkeiten

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad (k = 0, 1, \dots, n)$$

besitzt.

Für Erwartungswert und Varianz gelten

$$E(X) = np, \quad \text{Var}(X) = np(1 - p).$$

2.3.2 Die geometrische Verteilung

Wir nehmen nun an, daß die Folge von Versuchen beim ersten Mißerfolg abgebrochen wird und untersuchen die Verteilung der Zufallsgröße Y : Anzahl der Versuche bis zum ersten Mißerfolg im Bernoulli-Versuchsschema. Offensichtlich kann Y die Realisierungen $1, 2, \dots$ mit den Einzelwahrscheinlichkeiten $P(Y = k) = (1 - p)p^{k-1}$ besitzen.

Definition 2.5 Eine Zufallsgröße Y unterliegt einer geometrischen Verteilung mit dem Parameter $0 < p < 1$, wenn sie die Einzelwahrscheinlichkeiten

$$P(Y = k) = (1 - p)p^{k-1} \quad (k = 1, 2, \dots)$$

besitzt.

Erwartungswert und Varianz einer geometrisch verteilten Zufallsgröße sind durch

$$E(Y) = \frac{1}{1 - p}, \quad \text{Var}(Y) = \frac{p}{(1 - p)^2}$$

gegeben.

2.3.3 Die Poissonverteilung

Wenn im Bernoullischen Versuchsschemas die Anzahl n der unabhängigen Versuche ist sehr groß, die Wahrscheinlichkeit $p_n = P(A)$ des interessierenden Ereignisses A in jedem einzelnen Versuch (bei einer Serie von n Versuchen) jedoch sehr klein werden, so lassen sich unter den Voraussetzungen

$$n \rightarrow \infty, \quad p_n \rightarrow 0, \quad np_n \rightarrow \lambda > 0$$

für X die Einzelwahrscheinlichkeiten

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (k = 0, 1, 2, \dots)$$

als Grenzwerte der Einzelwahrscheinlichkeiten der Binomialverteilung herleiten.

Definition 2.6 Eine diskrete Zufallsgröße X unterliegt einer Poissonverteilung mit dem Parameter $\lambda > 0$, wenn sie die Einzelwahrscheinlichkeiten

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (k = 0, 1, 2, \dots)$$

besitzt.

Für Erwartungswert und Varianz erhält man:

$$E(X) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda; \quad \text{Var}(X) = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} - \lambda^2 = \lambda,$$

d.h. Erwartungswert und Varianz dieser Verteilung sind gleich.

Bemerkung 1: Die Poissonverteilung hat besondere Bedeutung in der Theorie Poisson-scher Punktprozesse.

Bemerkung 2: Zwei unabhängige Zufallsgrößen sind genau dann poissonverteilt, wenn ihre Summe poissonverteilt ist.

Wir betrachten nun einige wesentliche stetige Verteilungen.

2.3.4 Die Normalverteilung

Die Normalverteilung ist eine der wesentlichsten Verteilungen in der Wahrscheinlichkeitstheorie.

Definition 2.7 Eine stetige Zufallsgröße W mit der Dichtefunktion

$$f_W(x) = \varphi_{\mu, \sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right), \quad x \in \mathbf{R}$$

heißt normalverteilt mit den Parametern μ und σ^2 .

Die beiden Parameter μ und σ^2 stellen Erwartungswert und Varianz der Normalverteilung dar.

Wir betrachten nun einen zufälligen Vektor $\underline{W}_m = (W_1, \dots, W_m)$. Wenn $E(W_k^2) < \infty$, so kann man den Erwartungswertvektor $\underline{\mu}_m = (\mu_1, \dots, \mu_m)$, $E(W_k) = \mu_k$ und die Kovarianzmatrix $\Sigma = (\sigma_{kl})$ mit

$$\sigma_{kl} = E(W_k - \mu_k)(W_l - \mu_l)$$

betrachten. Im Fall $m = 2$ besitzen die Elemente der Kovarianzmatrix folgendes Aussehen: $\sigma_{11} = \sigma_1^2$ und $\sigma_{22} = \sigma_2^2$ stellen die Varianzen der ersten bzw. zweiten Komponente dar und $\sigma_{12} = \rho\sigma_1\sigma_2$, wobei ρ der Korrelationskoeffizient mit $|\rho| \leq 1$ ist. Wenn die Kovarianzmatrix den Rang m besitzt, d.h. wenn $\det \Sigma > 0$, so existiert die inverse Matrix $\Sigma^{-1} = (\sigma^{kl})$ und man kann eine mehrdimensionale Dichte definieren.

Definition 2.8 Eine m -dimensionalen stetige Zufallsgröße \underline{W}_m unterliegt einer m -dimensionalen Normalverteilung mit dem Erwartungswertvektor $\underline{\mu} = (\mu_1, \dots, \mu_m)^T$ und der Kovarianzmatrix $\Sigma = (\sigma_{kl})$, $k, l = 1, \dots, m$, wenn sie die Dichtefunktion

$$\begin{aligned} \varphi_{\underline{\mu}, \Sigma}(s_1, \dots, s_m) &= \frac{1}{(2\pi)^{n/2}(\det \Sigma)^{1/2}} \times \\ &\times \exp \left(-\frac{1}{2} \sum_{k,l=1}^m (s_k - \mu_k) \sigma^{kl} (s_l - \mu_l) \right) \end{aligned}$$

und die Verteilungsfunktion

$$\Phi_{\underline{\mu}, \Sigma}(t_1, \dots, t_m) = \int_{-\infty}^{t_1} \dots \int_{-\infty}^{t_m} \varphi_{\underline{\mu}, \Sigma}(s_1, \dots, s_m) ds_1 \dots ds_m$$

besitzt.

Beispiel 2.1 Im Falle $m = 2$ besitzen die Kovarianzmatrix und ihre Inverse das folgende Aussehen:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad \Sigma^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}$$

und wir erhalten folgende Dichte der zweidimensionalen Normalverteilung:

$$\begin{aligned} \varphi_{\underline{\mu}, \Sigma}(s_1, s_2) &= \frac{1}{(2\pi)\sigma_1\sigma_2(1-\rho^2)^{1/2}} \times \\ &\times \exp \left(-\frac{\sigma_2^2(s_1 - \mu_1)^2 - 2\rho\sigma_1\sigma_2(s_1 - \mu_1)(s_2 - \mu_2) + \sigma_1^2(s_2 - \mu_2)^2}{2\sigma_1^2\sigma_2^2(1-\rho^2)} \right). \end{aligned}$$

Diese zweidimensionale Dichte läßt sich bezüglich der beiden Veränderlichen s_1 und s_2 folgendermaßen zerlegen:

$$\varphi_{\underline{\mu}, \Sigma}(s_1, s_2) = \varphi_{\mu_2 + \rho \frac{\sigma_2^2}{\sigma_1^2}(s_1 - \mu_1), \sigma_2^2(1-\rho^2)}(s_2) \varphi_{\mu_1, \sigma_1^2}(s_1).$$

Damit ist die gemeinsame Dichte darstellbar als Produkt der Dichte der ersten Komponente und einer bedingten Dichte der zweiten Komponente unter der Bedingung $W_1 = s_1$ (vergleiche Formel (2.9)). Die bedingte Dichte der zweiten Komponente ist wiederum normalverteilt mit dem Erwartungswert $\mu_2 + \rho_{\sigma_1}^2(s_1 - \mu_1)$ und der Varianz $\sigma_2^2(1 - \rho^2)$. \diamond

2.3.5 Die logarithmische Normalverteilung

Definition 2.9 Die zufällige Größe $S > 0$ heißt logarithmisch normalverteilt, wenn $\ln S$ einer Normalverteilung unterliegt.

Damit erhält man für die logarithmische Normalverteilung

$$F(s) = P(S \leq s) = P(\ln S \leq \ln s) = \Phi\left(\frac{\ln s - \mu}{\sigma}\right), s > 0.$$

2.3.6 Die Exponentialverteilung

Definition 2.10 Eine stetige Zufallsgröße X mit der Verteilungsfunktion:

$$F_\lambda(s) = \begin{cases} 1 - e^{-\lambda s} & \text{für } s \geq 0, \lambda > 0 \\ 0 & \text{sonst} \end{cases}$$

heißt exponentialverteilt mit dem Parameter λ .

Die Dichtefunktion dieser Verteilung ist:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{für } x \geq 0, \lambda > 0, \\ 0 & \text{sonst} \end{cases}.$$

Die Momente k -ter Ordnung ergeben sich aus Formel 2.7:

$$m_{F_{\lambda,k}} = k \int_0^\infty t^{k-1} \overline{F}(t) dt = k \int_0^\infty t^{k-1} e^{-\lambda t} dt = \frac{k!}{\lambda^k}$$

Speziell erhält man für Erwartungswert und Varianz

$$E(X) = \frac{1}{\lambda} \quad \text{und} \quad \text{Var}(X) = \frac{1}{\lambda^2}.$$

Die Exponentialverteilung verfügt über eine Reihe bemerkenswerter Eigenschaften. Die wichtigste dieser Eigenschaften ist die Gedächtnislosigkeit.

Betrachten wir die Restlebensdauer eines Elementes

$$S_t = \begin{cases} S - t, & t < S \\ 0, & t \geq S \end{cases} \quad (2.12)$$

Dabei gilt $\{S_t > 0\}$ genau dann, wenn $\{S > t\}$. Die bedingte Wahrscheinlichkeit $\bar{F}_t(s)$ des Ereignisses, daß in der Zeit von t bis $t+s$ kein Ausfall stattfindet unter der Bedingung, daß vor der Zeit t kein Ausfall war, kann dann folgendermaßen berechnet werden:

$$\bar{F}_t(s) = \frac{P(S_t > s, S > t)}{P(S > t)} = \frac{P(S > t+s)}{P(S > t)} = \frac{\bar{F}_\lambda(t+s)}{\bar{F}_\lambda(t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s}.$$

Damit hängt die Wahrscheinlichkeit davon, daß das Element nach t Zeiteinheiten weitere s Zeiteinheiten ausfallfrei arbeitet, nicht vom Zeitpunkt t ab. Die Restlebensdauer besitzt die gleiche Verteilung wie die Lebensdauer eines neuen Elementes.

2.3.7 Die zweiparametrische Exponentialverteilung

Definition 2.11 Eine Zufallsgröße S heißt zweiparametrisch exponentialverteilt mit den Parametern $\lambda > 0$ und $t_0 > 0$, wenn sie die Überlebensfunktion

$$\bar{F}(t) = P(S > t) = \exp(-\lambda(t - t_0))$$

für $t \geq t_0$ und $\bar{F}(t) = 1$ für $t < t_0$ besitzt.

Der Erwartungswert dieser Verteilung ist durch

$$E(S) = t_0 + \frac{1}{\lambda}$$

gegeben. Diese Familie von Verteilungen ist sehr gut geeignet, um das Ausfallverhalten technischer Erzeugnisse zu beschreiben, bei denen ein Ausfall erst nach einer „Inkubationszeit“ eintreten kann und nach dieser Inkubationszeit keine Alterungserscheinungen auftreten. Die Ausfallrate dieser Verteilung hat die Form

$$h(s) = \lambda I(s \geq t_0),$$

d.h. sie ist konstant für alle $s \geq t_0$.

2.3.8 Verteilungen vom Pareto-Typ

Definition 2.12 Eine Zufallsgröße S gehört zur Familie der Verteilungen vom Pareto-Typ $\mathfrak{F}_{P,\beta}$ mit den Parametern $\alpha > 0$, $\beta > 0$ und $\mu > 0$, wenn sie die Überlebensfunktion

$$\bar{F}(s) = \left(\frac{\alpha}{\alpha + (s - \mu)^+} \right)^\beta = \left(\frac{1}{1 + \left(\frac{s - \mu}{\alpha} \right)^+} \right)^\beta, \quad (s - \mu)^+ = (s - \mu) \vee 0$$

besitzt.

Hierbei sind α ein Maßstabsparameter, β ein Formparameter und μ ein Anfangsparameter. Für $\alpha = \mu$ erhält man die bekannte Paretoverteilung.

Verteilungen vom Pareto-Typ kann man als Mischung von Exponentialverteilungen ansehen. Betrachten wir eine Zufallsgröße S mit der Überlebensfunktion $\bar{F}(s) = \exp\{-\Lambda s\}$, wobei Λ wiederum eine zufällige Größe ist, die die Dichtefunktion $\exp(-\lambda)$ besitzt. Dann erhält man aus der Formel über die totale Wahrscheinlichkeit für die Verteilung von S

$$\begin{aligned} F(s) &= P(S \leq s) = \int_0^\infty P(S \leq s | \Lambda = \lambda) \exp(-\lambda) d\lambda = \\ &= \int_0^\infty (1 - \exp(-\lambda s)) \exp(-\lambda) d\lambda = \\ &= \int_0^\infty \exp(-\lambda) d\lambda - \int_0^\infty \exp(-\lambda(s+1)) d\lambda = 1 - \frac{1}{(s+1)^2}, \end{aligned}$$

d. h. eine Verteilung vom Pareto-Typ mit den Parametern $\alpha = 1$, $\beta = 1$ und $\mu = 0$.

2.3.9 Weibull-Verteilungen

Definition 2.13 Eine Zufallsgröße S gehört zur Familie der zweiparametrischen Weibull-Verteilungen $\mathfrak{F}_{w,2}$ mit den Parametern $\alpha > 0$ und $\beta > 0$, wenn sie die Überlebensfunktion

$$\bar{F}(s) = \exp\left(-\left(\frac{s}{\alpha}\right)^\beta\right)$$

für $s \geq 0$ besitzt und zur Familie der dreiparametrischen Weibull-Verteilungen $\mathfrak{F}_{w,3}$ mit den Parametern $\alpha > 0$, $\beta > 0$ und μ , wenn sie die Überlebensfunktion

$$\bar{F}(s) = \exp\left(-\left(\frac{s-\mu}{\alpha}\right)^\beta\right)$$

für $s \geq \mu$ besitzt.

Die Weibull-Verteilung gehört zur Klasse der Extremwertverteilungen¹. Diesen Verteilungen liegt folgende Modellvorstellung zugrunde: Ein System besteht aus vielen einzelnen Teilen mit den zufälligen Lebensdauern $S_i, i = 1, \dots, n$, von denen jedes für das Funktionieren des Systems wesentlich ist, d.h. der Ausfall des Systems erfolgt zur Zeit $S_{(1)} = \min(S_1, \dots, S_n)$. Wenn die Anzahl der Elemente des Systems und gleichzeitig die Lebensdauer jedes Elementes groß werden, so läßt sich die Weibullverteilung als Grenzlebensdauer des Systems herleiten (vgl. GUMBELL). Dabei muß man voraussetzen, daß alle $S_i, i = 1, 2, \dots$ identisch verteilt sind oder zumindest ihre Überlebensfunktionen für große Argumente asymptotisch identisch sind.

¹Da die vollständigsten Resultate in der Theorie der Extremwertverteilungen, zu denen die Weibullverteilung gehört, von GNEDENKO stammen, heißt diese Verteilung in der russischsprachigen Literatur Weibull-Gnedenko-Verteilung.

Die Zugehörigkeit der Weibullverteilung zur Klasse der Extremwertverteilungen wird deutlich, wenn die Überlebenswahrscheinlichkeit in folgender Form geschrieben wird:

$$\begin{aligned}\bar{F}(t) &= \exp\left(-\left(\frac{s}{\alpha}\right)^\beta\right) = \exp(-\exp(\beta(\ln s - \ln \alpha))) = \\ &= \exp\left(-\exp\left(\frac{s' - \alpha'}{\beta'}\right)\right),\end{aligned}$$

wobei $s' = \ln s$, $\alpha' = \ln \alpha$ und $\beta' = \beta^{-1}$ sind.

Außerdem ist offensichtlich, daß die Familie der Weibull-Verteilungen abgeschlossen bezüglich der Minumbildung ist. Wenn die zufälligen Lebensdauern S_i , $i = 1, \dots, n$ zur Familie $\mathfrak{F}_{W,2}$, gehören, so erhält man für das Minimum:

$$\begin{aligned}\bar{F}_{S_{(1)}}(t) &= P\left(\bigcup_{i=1}^n (S_i > s)\right) = (\bar{F}(s))^n = \\ &= \exp\left(-n\left(\frac{s}{\alpha}\right)^\beta\right) = \exp\left(-\left(\frac{s}{\alpha''}\right)^{\beta}\right) \quad \text{mit} \quad \alpha'' = \frac{\alpha}{n^{1/\beta}}.\end{aligned}$$

Damit gehört die Verteilung des Minimums ebenfalls zur Familie der zweiparametrischen Weibullverteilungen.

2.3.10 Verteilungen mit Lage- und Maßstabsparameter

Familien zweiparametrischer Verteilungen können u.a. folgendermaßen konstruiert werden. Es wird eine Verteilungsfunktion $F_0(s)$ (im allgemeinen mit bekannten Parametern oder in einer standardisierten Version) vorgegeben. Alle Verteilungen aus der entsprechenden Familie ergeben sich dann aus $F_0(s)$ durch Einführung zweier Parameter μ und α :

$$F(s) = F_0\left(\frac{s - \mu}{\alpha}\right)$$

oder, allgemeiner,

$$F(s) = F_0\left(\frac{g(s) - \mu}{\alpha}\right) \quad \text{mit} \quad g(s) \uparrow (s \uparrow).$$

In dieser Familie sind μ ein Lageparameter und α ein Maßstabsparameter. Diese Verteilungen werden verwendet, wenn man gewisse Eigenschaften des Verteilungsgesetzes voraussetzt (z.B. eine bestimmte Form der Ausfallrate) und durch die Parameter eine Anpassung der Daten an das Modell vornimmt. Einige der in diesem Abschnitt vorgestellten Verteilungen lassen sich so darstellen, daß sie zur Familie der Verteilungen mit Maßstabs- und Lageparameter gehören.

Zum Schluß sollen nun noch zwei Verteilungen vorgestellt werden, die durch ihre Ausfallrate definiert sind.

2.3.11 Gomperz–Makeham–Verteilungen

Definition 2.14 Eine Zufallsgröße S gehört zur Familie der Gomperz–Makeham–Verteilungen $\mathfrak{F}_{GM,s}$ mit den Parametern $h_0 > 0$, $h_1 > 0$ und $h_2 \neq 0$, wenn sie die Ausfallrate

$$h(s) = h_0 + h_1 \exp(h_2 s)$$

für $s \geq 0$ besitzt.

Verteilungen dieser Art erhält man z.B., wenn ein Ausfall auf zwei verschiedene Ursachen beruhen kann. Sei $S = \min(S_1, S_2)$, wobei die beiden Ausfallursachen den Verteilungsgesetzen

$$S_1 \sim \bar{F}_1(s) = \exp(-h_0 s), \quad S_2 \sim \bar{F}_2(s) = \exp\left(-\int_0^s h_1 \exp(h_2 u) du\right)$$

unterliegen und nicht bekannt ist, aus welcher Ursache der Ausfall erfolgte. Man kann sich leicht davon überzeugen, daß sich die Ausfallraten der Verteilung als Summe der Ausfallraten der beiden Ursachen ergibt, da

$$\begin{aligned} \bar{F}(s) &= \bar{F}_1(s) \bar{F}_2(s), \\ f(s) &= f_1(s) \bar{F}_2(s) + f_2(s) \bar{F}_1(s), \\ h(s) &= \frac{f(s)}{\bar{F}_1(s) \bar{F}_2(s)} = h_1(s) + h_2(s). \end{aligned}$$

2.3.12 Cox'sche Verteilungen

Definition 2.15 Eine Zufallsgröße S gehört zur Familie der Cox'schen Verteilungen \mathfrak{F}_C , wenn sie die Ausfallrate

$$h(s) = h_0(s) \exp(gx + hy)$$

für $s \geq 0$ besitzt.

Cox'sche Verteilungen sind als parametrische oder halbparametrische Verteilungen bekannt. Auch in diesem Fall wird die Verteilungsfamilie durch ihre Ausfallrate definiert. Als Beispiel für parametrische Cox'sche Verteilungen stellen wir uns vor, daß ein Element mit weibullverteilter Lebensdauer unter bekannten äußeren Bedingungen arbeitet, welche Einfluß auf den Zustand des Elementes haben. Seien z.B. x eine bestimmte Temperaturstufe und y eine bestimmte Luftfeuchtigkeit. Wir nehmen weiterhin an, daß die zufällige Lebensdauer S_i des Elementes unter diesen Bedingungen die Ausfallrate

$$h(s) = \left(\frac{s}{\alpha}\right)^\beta \exp(gx + hy)$$

besitzt. Der erste Faktor in dieser Funktion ist die Ausfallrate der Weibull-Verteilung und der zweite Faktor spiegelt den Einfluß der äußeren Bedingungen wieder. Wenn $g >$

0 und $h > 0$ sind, so spielen Temperatur und Feuchtigkeit eine negative Rolle und verschlechtern den Zustand des Elementes. Somit erhält man eine Verteilung mit den vier Parametern (α, β, g, h) .

Häufig werden Cox'schen Verteilungen als halbparametrische Verteilungsfamilie betrachtet, indem dem ersten Faktor h_0 keine vorgegebene Form zugeschrieben wird. In diesem Fall definiert man eine Cox'sche Verteilung über die Ausfallrate

$$h(s) = h_0(s) \exp(gx + hy) ,$$

wobei $h_0(s)$ unbekannt ist. Dieser Zugang wird häufig verwendet, um den Einfluß von Umweltbedingungen auf die Lebensdauer eines Erzeugnisses zu beschreiben.

3 Nichtparametrische Lebensdauerverteilungen

In diesem Kapitel sollen die wichtigsten in der Zuverlässigkeitstheorie verwendeten Lebensdauerverteilungen vorgestellt werden. In den ersten beiden Abschnitten werden nichtparametrische Familien von Verteilungen, die auf Alterungsbegriffen beruhen, vorgestellt. Der dritte Abschnitt ist einer kurzen Darstellung der wichtigsten parametrischen Verteilungsfamilien gewidmet.

3.1 Die Familie der IFR–Verteilungen

Wir beginnen die Betrachtungen mit der oft verwendeten Familie von Verteilungen mit wachsender Ausfallrate.

Sei S die zufällige Lebensdauer eines Elementes mit der Überlebensfunktion $\bar{F}(s)$. Wir nehmen an, daß das Element bereits eine Zeit t überlebt hat ($\bar{F}(t) > 0$) und bezeichnen mit $\bar{F}_t(s)$ die bedingte Wahrscheinlichkeit, daß in einem folgenden Intervall $(t, t+s]$ der Länge s kein Ausfall stattfindet:

$$\bar{F}_t(s) = \frac{\bar{F}(t+s)}{\bar{F}(t)}.$$

Definition 3.1 Die Verteilungsfunktion $F(\cdot)$ besitzt eine wachsende Ausfallrate, wenn für jedes $s \geq 0$ und $t_1 \leq t_2$

$$\frac{\bar{F}(t_1+s)}{\bar{F}(t_1)} \geq \frac{\bar{F}(t_2+s)}{\bar{F}(t_2)} \quad (3.1)$$

gilt, d.h. wenn die bedingte Ausfallwahrscheinlichkeit monoton wachsend ist. Die Familie solcher Verteilungsfunktionen heißt Familie von Verteilungen mit wachsender Ausfallrate (*increasing failure rate*) und wird mit $\mathfrak{F}_{\text{IFR}}$ bezeichnet.

Bemerkung: Einige Autoren definieren die Familie $\mathfrak{F}_{\text{IFR}}$ über die Monotonie der Ausfallrate $h(u)$ aus dem Abschnitt 2.1. Das setzt jedoch die Existenz einer Dichte voraus, so daß wir die obige allgemeinere Definition bevorzugen. Besitzt $F(t)$ eine Dichte, so ist (3.1) äquivalent dazu, daß $h(s)$ monoton wächst. Diese Aussage folgt sofort mittels der Beziehungen (2.3) und (2.5), da

$$\frac{\bar{F}(t+s)}{\bar{F}(t)} = \frac{\exp(-\int_0^{t+s} h(u)du)}{\exp(-\int_0^t h(u)du)} = \exp(-\int_t^{t+s} h(u)du).$$

Einen Spezialfall in dieser Familie stellen die altersunabhängigen Verteilungen dar. Von diesen Verteilungen fordern wir, daß

$$\overline{F}_t(s) = \frac{\overline{F}(t+s)}{\overline{F}(t)} = \overline{F}(s)$$

gilt.

Beispiel 3.1 Wir betrachten die Familie \mathfrak{F}_E von Exponentialverteilungen mit der Überlebensfunktion

$$\overline{F}(s) = e^{-\lambda s} = e^{-s/\theta} \in \mathfrak{F}_E .$$

Es handelt sich hier um eine einparametrische Familie von Verteilungen mit dem Parameter λ , $\Theta = (\lambda : \lambda > 0)$ oder mit dem Parameter $\theta = \frac{1}{\lambda}$, $\Theta = (\theta : \theta > 0)$. Es wurde schon gezeigt, daß diese Familie von Verteilungen über die Eigenschaft der Altersunabhängigkeit verfügt. Sie ist die einzige Familie unter den stetigen Lebensdauerverteilungen mit dieser Eigenschaft. Setzt man nämlich voraus, daß eine stetige Verteilungsfunktion $F(t)$ eine konstante Ausfallrate $h(u) = \lambda$, $u \geq 0$, besitzt, so erhält man aus (2.3) und (2.5)

$$\overline{F}(t) = e^{-\int_0^t \lambda du} = e^{-\lambda t}, \quad t \geq 0,$$

und damit die Exponentialverteilung. ◇

Beispiel 3.2 Nun betrachten wir eine weitere Familie von Verteilungen, die in der Zuverlässigkeitstheorie häufig verwendete Weibullverteilung $\mathfrak{F}_{W,2}$. Es handelt sich hier um eine zweiparametrische Familie mit der Überlebensfunktion

$$\overline{F}(s, \theta) = \overline{F}(s, \alpha, \beta) = e^{-(s/\alpha)^\beta}$$

und dem Parameter

$$\underline{\theta} = (\alpha, \beta), \quad \Theta = ((\alpha, \beta) : \alpha > 0, \beta > 0) .$$

Für diese Familie erhält man

$$\frac{\overline{F}(t+s)}{\overline{F}(t)} = \exp \left(- \left(\frac{t+s}{\alpha} \right)^\beta + \left(\frac{t}{\alpha} \right)^\beta \right) .$$

Um zu untersuchen, wann die Weibullverteilung zur IFR-Familie gehört, betrachten wir die Funktion $g(t) = (t+s)^\beta - t^\beta$.

Wenn $g(t)$ monoton wachsend für wachsende t ist, so liegt F in der Familie $\mathfrak{F}_{\text{IFR}}$. Bildet man die Ableitung dieser Funktion:

$$\frac{dg(t)}{dt} = \beta(t+s)^{\beta-1} - \beta t^{\beta-1} = \beta t^{\beta-1} \left(\left(1 + \frac{s}{t} \right)^{\beta-1} - 1 \right) ,$$

so gilt $\frac{dg(t)}{dt} \geq 0$ für $\beta \geq 1$, d.h. die Weibullverteilung gehört zur IFR-Familie, wenn $\beta \geq 1$ ist. Für $\beta = 1$ erhält man den Spezialfall der Exponentialverteilung, also eine konstante Ausfallrate oder Altersunabhängigkeit. Wenn $\beta < 1$ ist, gehört die Weibullverteilung nicht zur Familie der IFR-Verteilungen. ◇

Im folgenden sollen einige Eigenschaften der IFR-Familie aufgeführt werden.

Satz 3.1 *Eine Verteilungsfunktion $F \in \mathfrak{F}_{\text{IFR}}$ besitzt höchstens einen Sprung. Wenn sie einen Sprung besitzt, so handelt es sich um einen Sprung auf 1.*

Beweis: Wir nehmen an, daß die Verteilungsfunktion F im Punkte t einen Sprung $\Delta F(t) > 0$ besitzt, jedoch nicht auf 1 springt. Dann existiert ein Punkt $t' > t$ so, daß $F(t') < 1$ und $\bar{F}(t') > 0$ sind. Die Menge aller Sprünge ist abzählbar, jedoch die Menge der Punkte auf dem Intervall (t, t') ist überabzählbar. Daher kann man ein $s : t < t + s < t'$ finden, so daß $t + s$ ein Stetigkeitspunkt von $F(\cdot)$ ist, d.h.

$$\bar{F}((t + s)-) = \bar{F}(t + s) .$$

Da $F(t) < 1$, $\bar{F}(t) > 0$ und $\Delta F(t) > 0$ gilt, erhält man

$$\frac{\bar{F}(t + s)}{\bar{F}(t)} = \frac{\bar{F}((t + s)-)}{\bar{F}(t-) - \Delta F(t)} > \frac{\bar{F}((t + s)-)}{\bar{F}(t-)} = \lim_{h_n \downarrow 0} \frac{\bar{F}(t - h_n + s)}{\bar{F}(t - h_n)} .$$

Daher gilt für genügend große n

$$\frac{\bar{F}(t - h_n + s)}{\bar{F}(t - h_n)} < \frac{\bar{F}(t + s)}{\bar{F}(t)}$$

und man hat einen Widerspruch zur Annahme, daß $F \in \mathfrak{F}_{\text{IFR}}$.

Damit ist bewiesen, daß eine Verteilungsfunktion aus der Familie der IFR-Verteilungen keine Sprünge besitzen kann, außer einem letzten Sprung auf 1 (Abbildung 3.1). \square

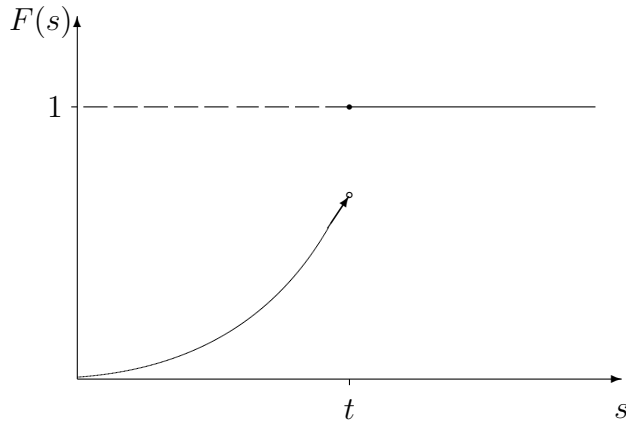


Abbildung 3.1: Mögliches Sprungverhalten einer Verteilungsfunktion aus der IFR-Familie

Folgerung 3.1 *Für Verteilungsfunktionen $F \in \mathfrak{F}_{\text{IFR}}$ und $F(s) < 1$ gilt*

$$\bar{F}(s) = e^{-H(s)} , \tag{3.2}$$

wobei $H(s)$ die im Abschnitt 2.1 definierte Hazardfunktion ist.

Beweis: Diese Eigenschaft folgt sofort aus der Stetigkeit der Funktion $F(s)$ auf dem Intervall $[0, s]$, wenn $F(s) < 1$. Selbstverständlich ist $H(s)$ ebenfalls eine stetige Funktion auf $[0, s]$. \square

Satz 3.2 Eine Verteilungsfunktion F gehört genau dann zur Familie $\mathfrak{F}_{\text{IFR}}$, wenn ihre Hazardfunktion $H(\cdot)$ für alle $s : F(s) < 1$ eine stetige konvexe Funktion ist.

Beweis: Wir zeigen zuerst die Notwendigkeit. Sei $F \in \mathfrak{F}_{\text{IFR}}$. Dann folgt aus Satz 3.1, daß $F(\cdot)$ eine stetige Funktion für alle $s : F(s) < 1$ ist. Aus Folgerung 3.1 erhält man, daß für diese s auch $H(\cdot)$ eine stetige Funktion ist. Weiterhin gilt

$$\frac{\overline{F}(t+s)}{\overline{F}(t)} = \exp(-(H(t+s) - H(t))) \quad \text{wobei} \quad H(t+s) - H(t) \uparrow (t \uparrow).$$

Dann erhält man

$$H\left(t + \frac{s}{2}\right) - H(t) \leq H(t+s) - H\left(t + \frac{s}{2}\right)$$

oder

$$2H\left(t + \frac{s}{2}\right) \leq H(t+s) + H(t).$$

Hieraus folgt

$$H\left(t + \frac{s}{2}\right) \leq \frac{H(t+s) + H(t)}{2}$$

und damit die Konvexität.

Sei nun $H(\cdot)$ eine stetige konvexe Funktion. Es soll gezeigt werden, daß diese Bedingung für die Zugehörigkeit zur IFR-Familie hinreichend ist. Zuerst betrachten wir den Fall zweier disjunkter Intervalle $(t_1, t_1 + s]$ und $(t_2, t_2 + s]$ mit $t_1 < t_1 + s \leq t_2 < t_2 + s$.

In Abbildung 3.2 ist eine konvexe Funktion dargestellt. Die Punkte A_1 und A_2 sowie

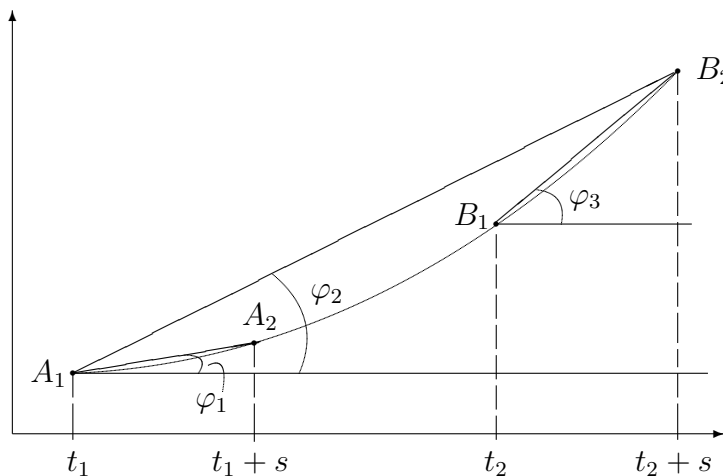


Abbildung 3.2: Das Verhalten einer konvexen Funktion

die Punkte B_1 und B_2 werden durch Strecken verbunden, ebenfalls die Punkte A_1 und B_2 . Für die dabei entstehenden Winkel gilt offensichtlich $\angle \varphi_1 \leq \angle \varphi_2 \leq \angle \varphi_3$ und $\tan \varphi_1 \leq \tan \varphi_3$. Hieraus folgt

$$\tan \varphi_1 = \frac{H(t_1 + s) - H(t_1)}{s} \leq \tan \varphi_3 = \frac{H(t_2 + s) - H(t_2)}{s},$$

$$H(t_1 + s) - H(t_1) \leq H(t_2 + s) - H(t_2), \quad t_1 < t_2$$

und aus Ungleichung (3.1) unter Berücksichtigung von (3.2) die Zugehörigkeit zur IFR-Familie. Es bleibt der Fall sich überschneidender Intervalle zu betrachten. Sei $t_1 < t_2 \leq t_1 + s < t_2 + s$. Wir haben schon bewiesen, daß $H(t_2) - H(t_1) \leq H(t_2 + s) - H(t_1 + s)$, da es sich hierbei um nichtüberschneidende Intervalle handelt.

Dann folgt

$$\begin{aligned} H(t_1 + s) - H(t_1) &= H(t_2) - H(t_1) + H(t_1 + s) - H(t_2) \leq \\ &\leq H(t_2 + s) - H(t_1 + s) + H(t_1 + s) - H(t_2) = H(t_2 + s) - H(t_2). \end{aligned}$$

Somit erhalten wir, daß die Funktionen $H(t + s) - H(t)$ monoton wachsend und $\bar{F}_t(s) = \exp(-(H(t + s) - H(t)))$ monoton fallend in t sind. Damit gilt $F \in \mathfrak{F}_{\text{IFR}}$ und der Satz ist bewiesen. \square

Folgerung 3.2 Eine Verteilungsfunktion $F(\cdot) \in \mathfrak{F}_{\text{IFR}}$ und die dazugehörige Hazardfunktion $H(\cdot)$ sind absolut stetige Funktionen für $s : F(s) < 1$.

Beweis: Es wurde schon bewiesen, daß H eine stetige konvexe Funktion ist. Jede stetige konvexe Funktion ist absolut stetig und hat eine monoton wachsende Ableitung (siehe z.B. WALTER). Sei $h(s) = \frac{dH(s)}{ds}$. Diese Ableitung existiert fast überall und ist wachsend, also kann man ohne Beschränkung der Allgemeinheit annehmen, daß $h(\cdot)$ eine rechtsseitig stetige Funktion ist und $h(s) < \infty$ wenn $F(s) < 1$. Aus Folgerung 3.1 erhält man

$$\frac{dF(s)}{ds} = h(s)\bar{F}(s),$$

also besitzt $F(\cdot)$ eine Ableitung und ist absolut stetig für alle s , für die $F(s) < 1$ ist. \square

Bemerkung: Wir haben IFR-Verteilungen allgemein definiert und im Gegensatz zur herkömmlichen Definition keine Existenz einer Dichte gefordert. Es zeigt sich jedoch, daß IFR-Verteilungen absolut stetig für alle $s : F(s) < 1$ sein müssen.

Folgerung 3.3 Für Verteilungen $F \in \mathfrak{F}_{\text{IFR}}$ existieren alle Momente $m_{F,k}$ von F .

Beweis: Seien $F_1(\cdot)$ und $F_2(\cdot)$ zwei Verteilungsfunktionen und $H_1(\cdot)$, $H_2(\cdot)$ die zugehörigen Hazardfunktionen. Wenn $H_1(s) \geq H_2(s)$ ist und die Momente der Ordnung k von F_2 existieren, dann gilt $m_{F_1,k} \leq m_{F_2,k}$, denn nach (2.7) ist

$$\begin{aligned} m_{F_1,k} &= k \int_0^\infty t^{k-1} \bar{F}_1(t) dt = k \int_0^\infty t^{k-1} e^{-H_1(t)} dt \leq \\ &\leq k \int_0^\infty t^{k-1} e^{-H_2(t)} dt = m_{F_2,k}. \end{aligned}$$

Seien nun $F_1 \in \mathfrak{F}_{\text{IFR}}$ und F_2 eine Verteilung mit der Hazardfunktion $(H_2(t) = H'_1(s)(t - s) + H_1(s)) \vee 0$, wobei $a \vee b = \max(a, b)$. Diese Hazardfunktion erhält man, indem man im Punkt $A = (s, H(s))$ eine Tangente an die Funktion $H_1(\cdot)$ legt (siehe Abbildung 3.3). Die Funktion $H_2(\cdot)$ ist eine stetige konvexe Funktion, die die Bedingung $H_2(t) \leq H_1(t)$

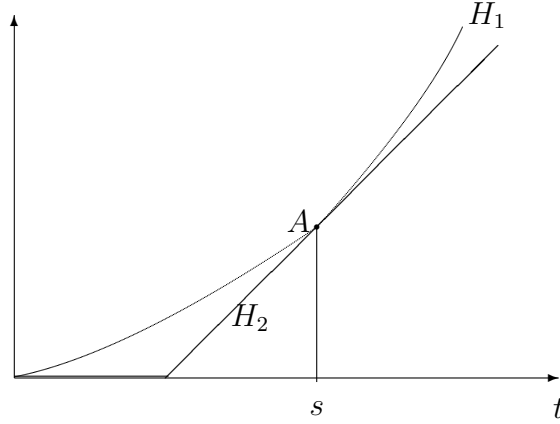


Abbildung 3.3: Zwei Hazardfunktionen

erfüllt. Da

$$m_{F_2, k} = k \int_0^\infty t^{k-1} \exp(-(H'_1(s)(t - s) + H_1(s)) \vee 0) dt < \infty$$

für alle $k = 1, 2, \dots$, und $m_{F_2, k} \geq m_{F_1, k}$, ist die Aussage bewiesen. \square

Wir wollen nun klären, ob es sich bei IFR-Verteilungen um parametrische Verteilungsfamilien handelt. Parametrische Verteilungsfamilien sind dadurch charakterisiert, daß ein endlichdimensionaler Parameterraum $\Theta \subseteq \mathbb{R}^m$, $m < \infty$ so existiert, daß jede Verteilungsfunktion aus dieser Familie eindeutig durch einen Punkt des Parameterraumes beschrieben werden kann (wie z. B. bei der Familie der Weibullverteilungen). Ist es nicht möglich, einen endlichdimensionalen Raum Θ zu finden, so liegt eine nichtparametrische Familie von Verteilungen vor.

Satz 3.3 *Die Familie $\mathfrak{F}_{\text{IFR}}$ ist eine nichtparametrische Familie von Verteilungen.*

Beweis: Wir betrachten eine parametrische Familie von Verteilungen \mathfrak{F}_r , wobei eine Verteilung aus dieser Familie durch

$$\overline{F}_\theta = \overline{F}_\theta(s) = \exp\left(-\int_0^s h_r(u) du\right)$$

mit $h_r(u) = h_k$ für $s_{k-1} \leq u < s_k$, $k = 1, \dots, r$ und der Parameterraum Θ_r durch

$$\Theta_r = ((h_1, \dots, h_r, s_1, \dots, s_r) : h_{k+1} > h_k, s_{k+1} \geq s_k, k = 0, \dots, r).$$

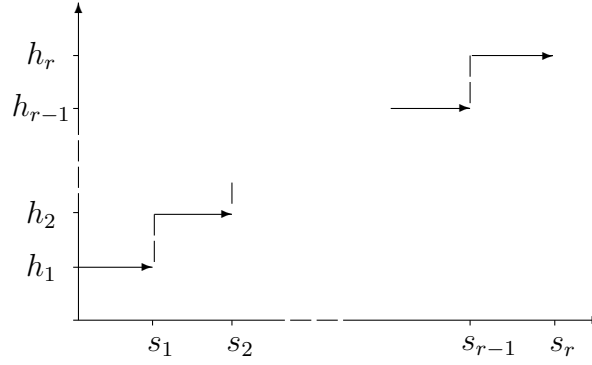


Abbildung 3.4: Der Parameter einer Verteilung aus \mathfrak{F}_r

gegeben sind (siehe Abbildung 3.4). Die Dimension des Parameters ist $2r$. Es ist offensichtlich, daß $\mathfrak{F}_r \subseteq \mathfrak{F}_{\text{IFR}}$. Da r beliebig groß gewählt werden kann, gehört die IFR-Familie zu den nichtparametrischen Familien von Verteilungen. \square

Satz 3.4 (Beichelt, Franken) Für $F \in \mathfrak{F}_{\text{IFR}}$ gilt

$$\frac{m_{F,k+1} \cdot m_{F,k-1}}{(k+1)!(k-1)!} \leq \left(\frac{m_{F,k}}{k!} \right)^2. \quad (3.3)$$

Speziell gilt für $k = 1$

$$m_{F,2} \leq 2m_{F,1}^2. \quad (3.4)$$

Bemerkung: Es ist gut bekannt, daß für eine konvexe Funktion φ , und eine nichtnegative Zufallsgröße S die Ungleichung $\varphi(\mathbb{E}(S)) \leq \mathbb{E}(\varphi(S))$ (Jensensche Ungleichung) gilt. Wählen wir $\varphi(t) = t^{(k+1)/k}$, so erhalten wir

$$\begin{aligned} (\mathbb{E}(S^k))^{(k+1)/k} &\leq \mathbb{E}((S^k)^{(k+1)/k}) \leq \mathbb{E}(S^{k+1}) \\ \text{und daraus} \quad (\mathbb{E}(S^k))^{1/k} &\leq (\mathbb{E}(S^{k+1}))^{1/(k+1)} \quad k = 1, 2, \dots \end{aligned}$$

Mit dieser Ungleichung ist es möglich, jeweils die niedrigeren Momente durch die höheren Momente abzuschätzen. Die im Satz 3.4 angegebene Ungleichung gestattet es jedoch, höhere Momente durch Momente niedrigerer Ordnung abzuschätzen, wenn die betrachtete Verteilung zur IFR-Familie gehört.

Satz 3.5 Sei $F \in \mathfrak{F}_{\text{IFR}}$. Dann gilt

$$\begin{aligned} (i) \quad \bar{F}(s) &\geq \exp\left(-\frac{s}{m_F}\right) \quad \text{für} \quad s \leq m_F = m_{F,1}, \\ (ii) \quad \bar{F}(s) &\leq \exp(-w(s)s) \quad \text{für} \quad s > m_F. \end{aligned}$$

Wobei $w(s)$ die Lösung der Gleichung $1 - wm_F = (1 - sw)e^{-ws}$ ist.

Beweis: Die Aussage (i) des Satzes ist ein Spezialfall des folgenden Satzes 3.6 aus BARLOW, PROSHAN. Wir bringen hier allerdings einen etwas anderen Beweis. Betrachten wir zwei Abbildungen $s \rightarrow z = H(s)$, $s > 0$ und $z \rightarrow s = R(z)$, $z > 0$. Hierbei ist $H(\cdot)$ die zu F gehörige Hazardfunktion und $R(\cdot)$ ist die Umkehrfunktion: $R(\cdot) = H^{-1}(\cdot)$. Somit gilt $H(R(z)) = z$. Diese beiden Funktionen sind in Abbildung 3.5 dargestellt. Wenn

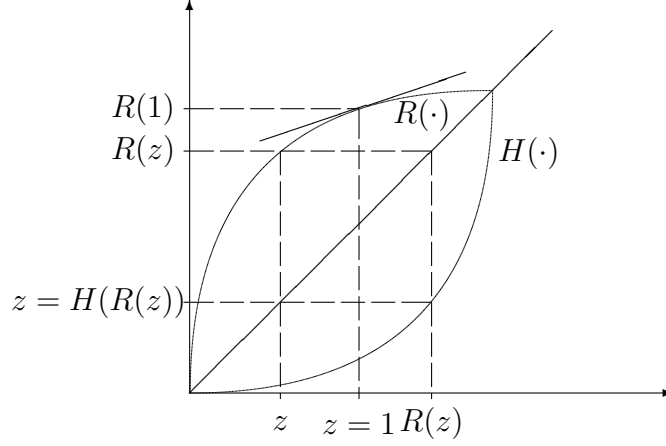


Abbildung 3.5: Die Funktionen $H(\cdot)$ und $R(\cdot)$

$F \in \mathfrak{F}_{\text{IFR}}$, so sind $H(\cdot)$ und $R(\cdot)$ rechtsseitig differenzierbar. $H(\cdot)$ ist konvex und $R(\cdot)$ ist konkav.

Nun betrachten wir den Punkt $(z = 1, s = R(1))$ und legen in diesem Punkt eine Tangente an die Kurve $R(\cdot)$. Da $R(\cdot)$ konkav ist, gilt die Ungleichung

$$R(z) \leq R(1) + R'(1)(z - 1) \quad \text{mit} \quad R'(1) = \lim_{h \downarrow 0} \frac{R(1+h) - R(1)}{h}.$$

Da

$$dF(s) = -d(e^{-H(s)}) = e^{-H(s)} dH(s)$$

erhalten wir für das erste Moment

$$m_F = \int_0^\infty s dF(s) = \int_0^\infty s e^{-H(s)} dH(s).$$

Führen wir jetzt im Integral die Substitution $H(s) = z$, $s = R(z)$ durch, so erhalten wir

$$\begin{aligned} m_F &= \int_0^\infty R(z) e^{-z} dz \leq \int_0^\infty [R(1) + R'(1)(z - 1)] e^{-z} dz = \\ &= R(1) \int_0^\infty e^{-z} dz + R'(1) \int_0^\infty (z - 1) e^{-z} dz = R(1). \end{aligned}$$

Somit gilt $m_F \leq R(1)$ oder $H(m_F) \leq H(R(1)) = 1$ und damit $\frac{H(m_F)}{m_F} \leq \frac{1}{m_F}$.
Aus der Konvexität von $H(\cdot)$ erhalten wir $\frac{H(s)}{s} \uparrow (\uparrow s)$, daraus folgt

$$\frac{H(s)}{s} \leq \frac{H(m_F)}{m_F} \leq \frac{1}{m_F} \quad \text{für } s \leq m_F.$$

Somit gilt

$$H(s) \leq \frac{s}{m_F}$$

und damit

$$\overline{F}(s) = e^{-H(s)} \geq e^{-s/m_F} \quad \text{für } s \leq m_F$$

und (i) ist bewiesen.

Zum Beweis der zweiten Ungleichung betrachten wir eine weitere Überlebensfunktion

$$\overline{G}_s(t) = \begin{cases} e^{-wt}, & t < s, \\ 0, & t \geq s. \end{cases}$$

Die zu dieser Überlebensfunktion gehörige Hazard-Funktion ist

$$H_s(t) = \begin{cases} wt, & t < s, \\ \infty, & t \geq s. \end{cases}$$

w wird nun so gewählt, daß m_F auch der Erwartungswert der Überlebensfunktion $\overline{G}_s(t)$ ist:

$$m_F = \int_0^\infty \overline{G}_s(t) dt = \int_0^s e^{-wt} dt + se^{-ws} = \frac{1 - e^{-ws}}{w} + se^{-ws}.$$

Daraus folgt, daß man w als eine Lösung der Gleichung $1 - wm_F = (1 - sw)e^{-ws}$ suchen muß. Die Gleichung $1 - ax = (1 - bx)e^{-bx}$ besitzt für $a < b$ eine eindeutige von Null verschiedene Lösung. In unserem Fall (siehe Abbildung 3.6) gilt das für $s > m_F$.

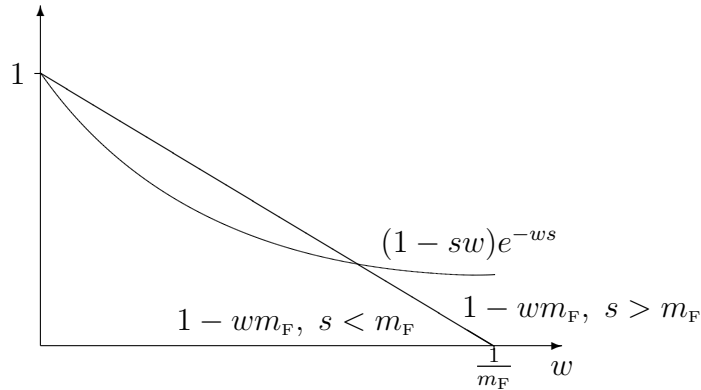


Abbildung 3.6: Die Lösung der Gleichung bei $s > m_F$

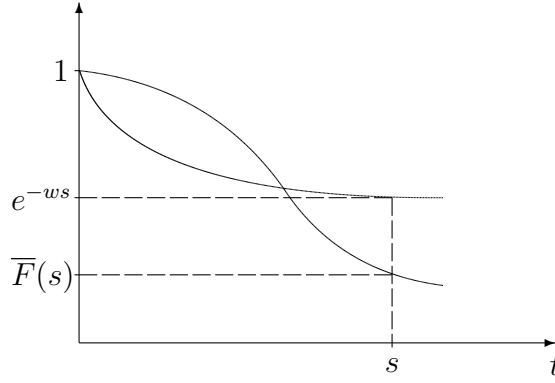


Abbildung 3.7: Zwei Überlebensfunktionen

Wenn wir nun annehmen, daß für alle $t < s$ $\bar{F}(t) > \bar{G}_s(t)$, dann gilt

$$m_F = \int_0^\infty \bar{F}(t) dt \geq \int_0^s \bar{F}(t) dt > \int_0^s \bar{G}_s(t) dt = m_F$$

und wir erhalten einen Widerspruch. Also schneiden sich $\bar{G}_s(\cdot)$ und $\bar{F}(\cdot)$ vor der Zeit s . Dann gilt $\bar{F}(s) \leq \bar{G}_s(s-) = e^{-w(s)s}$, $s > m_F$ und damit ist die zweite Ungleichung bewiesen. Abbildung 3.7 zeigt, wie die beiden Überlebensfunktionen $\bar{F}(s)$ und e^{-ws} aussehen. \square

Der nächste Satz zeigt, daß die Schranken (durch Anwendung der Jensenschen Ungleichung) verbessert werden können.

Satz 3.6 Wenn $F \in \mathfrak{F}_{IFR}$, dann existieren alle Momente und es gilt $m_{F,k}^{1/k} \leq m_{F,k+1}^{1/(k+1)} \leq \dots$ sowie

$$\bar{F}(t) \geq \begin{cases} \exp(-t(\frac{k!}{m_{F,k}})^{1/k}), & t \leq m_{F,k}^{1/k}, \\ 0, & t > m_{F,k}^{1/k}. \end{cases}$$

Den Beweis dieses Satzes findet man in BEICHELT, FRANKEN oder BARLOW, PROSHAN.

Der nächste Satz enthält eine wichtige Eigenschaft der IFR-Verteilungen.

Satz 3.7 (Barlow, Proshan) Wenn $F_1, F_2 \in \mathfrak{F}_{IFR}$, dann gehört die Faltung ebenfalls zu dieser Familie: $F_1 * F_2 \in \mathfrak{F}_{IFR}$.

Beispiel 3.3 Wir betrachten ein System aus zwei Elementen. Das erste Element besitzt die zufällige Lebensdauer S_1 . Wenn es ausgefallen ist, tritt an seine Stelle das zweite Element mit der zufälligen Lebensdauer S_2 . Diese Arbeitsweise heißt in der Zuverlässigkeitstheorie *kalte Reserve*. Das System besitzt dann die Lebensdauer $S = S_1 + S_2$ und die Wahrscheinlichkeit eines Ausfalls bis zur Zeit s ist

$$F(s) = \int_0^s F_1(s-u) dF_2(u) = \int_0^s F_2(s-u) dF_1(u).$$

Aus der Behauptung 3.7 folgt, daß die Lebensdauer eines Systems mit kalter Reserve zur IFR-Familie gehört, wenn die Verteilungsfunktionen F_1 und F_2 zu dieser Familie gehören. \diamond

3.2 Weitere nichtparametrische Familien von Lebensdauerverteilungen

In diesem Abschnitt sollen andere nichtparametrische Verteilungsfamilien betrachtet und ihr Zusammenhang mit der IFR-Familie bzw. untereinander dargestellt werden.

Definition 3.2 Eine Lebensdauerverteilung heißt IFRA-Verteilung (*increasing failure rate in average*), wenn $(\bar{F}(s))^{1/s}$ monoton fallend für wachsende s ist.

Alle Verteilungsfunktionen mit diesen Eigenschaften bilden die Familie $\mathfrak{F}_{\text{IFRA}}$. Interpretiert man diese Eigenschaft geometrisch, so bedeutet das, daß für Verteilungen aus $\mathfrak{F}_{\text{IFRA}}$ die Funktion $-\ln \bar{F}(\cdot)$ ein von Null ausgehendes Geradenbüschel von unten nach oben durchquert (siehe Abbildung 3.8). Diese Interpretation ergibt sich daraus, daß $-\frac{1}{s} \ln \bar{F}(\cdot)$ monoton wachsend ist und daher $-\ln \bar{F}(\cdot)$ mindestens so schnell wie s wächst.

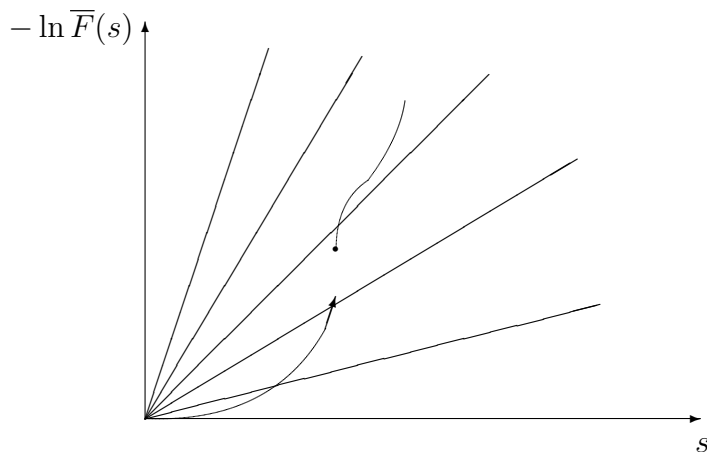


Abbildung 3.8: Das Verhalten der Überlebensfunktion einer IFRA-Verteilung

Satz 3.8 Die Familie $\mathfrak{F}_{\text{IFRA}}$ enthält die Familie $\mathfrak{F}_{\text{IFR}}$ ($\mathfrak{F}_{\text{IFR}} \subset \mathfrak{F}_{\text{IFRA}}$).

Beweis: Sei $F \in \mathfrak{F}_{\text{IFR}}$. Dann gilt $\bar{F}(s) = e^{-H(s)}$, wobei $H(s)$ konvex ist. Für konvexe Funktionen gilt

$$\frac{H(s)}{s} \uparrow (s \uparrow) \quad \text{und daher} \quad (\bar{F}(s))^{1/s} = \exp(-H(s)/s) \downarrow (s \uparrow).$$

□

Bemerkung: Eine Zufallsgröße mit der Ausfallrate

$$h(s) = \begin{cases} 0, & s < 0.5 \\ 2, & 0.5 \leq s < 1 \\ 1, & s \geq 1 \end{cases}$$

gehört zur IFRA-Familie. Da diese Ausfallrate nicht monoton wachsend ist, erhält man als Folgerung, daß $\mathfrak{F}_{\text{IFRA}} \setminus \mathfrak{F}_{\text{IFR}} \neq \emptyset$ ist.

Definition 3.3 Eine Lebensdauerverteilung $\bar{F}(t)$ heißt NBU-Verteilung (*new better than used*), wenn

$$\bar{F}(t+s) \leq \bar{F}(t) \cdot \bar{F}(s) \quad \forall t, s > 0 \quad (3.5)$$

Alle Verteilungen mit dieser Eigenschaft bilden die Familie von NBU-Verteilungen $\mathfrak{F}_{\text{NBU}}$.

Für alle $t : \bar{F}(t) > 0$ ist zu dieser Definition äquivalent, daß

$$\bar{F}_t(s) = \frac{\bar{F}(t+s)}{\bar{F}(t)} \leq \bar{F}(s) \quad \forall t, s > 0,$$

d.h., daß die bedingte Überlebenswahrscheinlichkeit eines Elementes, das bereits das Alter t erreicht hat, kleiner ist als die entsprechende Überlebenswahrscheinlichkeit eines neuen Elementes.

Satz 3.9 Die Familie $\mathfrak{F}_{\text{NBU}}$ enthält die Familie $\mathfrak{F}_{\text{IFRA}}$ ($\mathfrak{F}_{\text{IFRA}} \subset \mathfrak{F}_{\text{NBU}}$).

Beweis: Sei $F \in \mathfrak{F}_{\text{IFRA}}$. Dann gilt

$$(\bar{F}(t+s))^{1/(t+s)} \leq (\bar{F}(t))^{1/t}, \quad (\bar{F}(t+s))^{1/(t+s)} \leq (\bar{F}(s))^{1/s}.$$

Daraus folgt

$$(\bar{F}(t+s))^{t/(t+s)} \leq \bar{F}(t), \quad (\bar{F}(t+s))^{s/(t+s)} \leq \bar{F}(s)$$

und durch Multiplikation der beiden Ungleichungen

$$(\bar{F}(t+s))^{\frac{t}{t+s} + \frac{s}{t+s}} = \bar{F}(t+s) \leq \bar{F}(t) \cdot \bar{F}(s).$$

Das ist jedoch die charakteristische Eigenschaft für die Familie der NBU-Verteilungen. \square

Bemerkung: Eine Zufallsgröße mit der in Abbildung 3.9 dargestellten Ausfallrate

$$h(s) = \begin{cases} 0, & s < 1 \\ 1, & 1 \leq s < 1.5 \\ 0, & 1.5 \leq s < 2 \\ 2, & s \geq 2 \end{cases}$$

besitzt eine Verteilungsfunktion, für die $F \notin \mathfrak{F}_{\text{IFRA}}$, jedoch $F \in \mathfrak{F}_{\text{NBU}}$ gilt. Als Folgerung aus diesem Beispiel erhält man, daß $\mathfrak{F}_{\text{NBU}} \setminus \mathfrak{F}_{\text{IFRA}} \neq \emptyset$.

Wir betrachten nun bedingte Erwartungswerte für die Restlebensdauer $S_t = S - t$ unter der Bedingung, daß $S > t$ ist. Sei $m_{\text{F}}(t) = E(S_t | S > t)$ die erwartete Restlebensdauer.

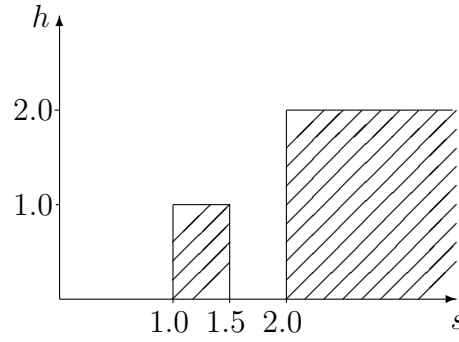


Abbildung 3.9: Eine Ausfallrate aus der NBU-Familie

Definition 3.4 Eine Lebensdauerverteilung F heißt NBUE-Verteilung (*new better than used in expectation*), wenn sie einen endlichen Erwartungswert m_F besitzt und für beliebige $t > 0$ die Ungleichung $m_F(t) \leq m_F(0) = m_F$ gilt.

Satz 3.10 Die Familie \mathfrak{F}_{NBUE} enthält die Familie \mathfrak{F}_{NBU} ($\mathfrak{F}_{NBU} \subset \mathfrak{F}_{NBUE}$).

Beweis: Sei $F \in \mathfrak{F}_{NBU}$. Dann folgt aus Gleichung (3.5)

$$\begin{aligned} m_F(t) &= E(S_t | S > t) = \int_0^\infty \bar{F}_t(s) ds = \int_0^\infty \frac{\bar{F}(t+s)}{\bar{F}(t)} ds \leq \\ &\leq \int_0^\infty \bar{F}(s) ds = m_F(0) = m_F. \end{aligned} \quad \square$$

Wir wollen nun einige interessante Eigenschaften der Familie \mathfrak{F}_{NBUE} betrachten, die durch die vorher aufgeführten Relationen der Verteilungsfamilien gleichzeitig auch Eigenschaften der anderen betrachteten Verteilungsfamilien sind.

Zu jeder Lebensdauerverteilung mit der Verteilungsfunktion F kann eine Größe $A_F(t)$ folgendermaßen eingeführt werden:

$$A_F(t) := \frac{1}{m_F} \int_0^t \bar{F}(u) du, \quad \bar{A}_F(t) = \frac{1}{m_F} \int_t^\infty \bar{F}(u) du$$

mit

$$m_F = \int_0^\infty u dF(u) = \int_0^\infty \bar{F}(u-) du = \int_0^\infty \bar{F}(u) du.$$

Es ist offensichtlich, daß $A_F(t)$ die Eigenschaften einer Verteilungsfunktion besitzt. Sie läßt sich als asymptotische Verteilung der Vorwärtsrekurrenzzeit eines stationären Erneuerungsprozesses interpretieren (siehe Abschnitt Erneuerungsprozesse).

Lemma 3.1 Für eine Verteilungsfunktion F gilt $F \in \mathfrak{F}_{NBUE}$ genau dann, wenn

$$\bar{A}_F(t) \leq \bar{F}(t) \quad \forall t \geq 0. \quad (3.6)$$

Beweis: Sei $F \in \mathfrak{F}_{NBUE}$. Dann gilt

$$m_F \geq m_F(t) = \int_0^\infty \bar{F}_t(s) ds = \int_0^\infty \frac{\bar{F}(t+s)}{\bar{F}(t)} ds = \int_t^\infty \frac{\bar{F}(u)}{\bar{F}(t)} du .$$

Hieraus folgt $\bar{F}(t) \geq \int_t^\infty \frac{\bar{F}(u)}{m_F} du = \bar{A}_F(t)$ und die Notwendigkeit der Ungleichung (3.6) ist bewiesen.

Nun sei $\bar{F}(t) \geq \int_t^\infty \frac{\bar{F}(u)}{m_F} du$. Dann erhält man

$$m_F \geq \int_t^\infty \frac{\bar{F}(u)}{\bar{F}(t)} du = \int_0^\infty \frac{\bar{F}(t+s)}{\bar{F}(t)} ds = \int_0^\infty \bar{F}_t(s) ds = m_F(t)$$

und damit ist die Ungleichung (3.6) auch hinreichend. \square

Folgerung 3.4 $\Delta(t) = \bar{F}(t) - \bar{A}_F(t) \geq 0$ für $F \in \mathfrak{F}_{NBUE}$.

Lemma 3.2 Aus $F \in \mathfrak{F}_{NBUE}$ folgt

(i) für die Ausfallrate h_{A_F} von A_F gilt $h_{A_F}(t) = \frac{1}{m_F(t)} \geq \frac{1}{m_F}$, falls h_{A_F} existiert,

(ii) $\bar{A}_F(t) \leq e^{-t/m_F}$, $\forall t \geq 0$.

Beweis:

(i) Die Dichte der Verteilung $A_F(t)$ ist $a_F(t) = \frac{dA_F(t)}{dt} = \frac{\bar{F}(t)}{m_F}$.

Damit erhält man für die Ausfallrate

$$h_{A_F}(t) = \frac{a_F(t)}{\bar{A}_F(t)} = \frac{\bar{F}(t)}{m_F \cdot \int_t^\infty \frac{\bar{F}(u)}{m_F} du} = \frac{1}{\int_t^\infty \frac{\bar{F}(u)}{\bar{F}(t)} du} = \frac{1}{m_F(t)} \geq \frac{1}{m_F} .$$

(ii) Unter Verwendung der in (i) gezeigten Ungleichung erhalten wir

$$\bar{A}_F(t) = \exp(-H_{A_F}(t)) = \exp\left(-\int_0^t h_{A_F}(u) du\right) \leq \exp\left(-\frac{t}{m_F}\right) .$$

\square

Folgerung 3.5 Alle Momente von $A_F(t)$ sind endlich:

$$m_{A_F,k} = \int_0^\infty t^k dA_F(t) < \infty \quad k = 1, 2, \dots$$

Beweis:

$$m_{A_F,k} = \int_0^\infty t^k dA_F(t) = k \int_0^\infty t^{k-1} \bar{A}_F(t) dt \leq k \int_0^\infty t^{k-1} \exp\left(-\frac{t}{m_F}\right) dt < \infty .$$

\square

Folgerung 3.6 Wenn $F \in \mathfrak{F}_{NBUE}$, dann sind alle Momente von F endlich: $m_{F,k} < \infty$, $k = 1, 2, \dots$.

Beweis:

$$m_{F,k} = \int_0^\infty t^k dF(t) = k \int_0^\infty t^{k-1} m_F \frac{\bar{F}(t)}{m_F} dt = k \cdot m_F \int_0^\infty t^{k-1} dA_F(t) < \infty.$$

□

Bemerkung: Für $k = 2$ erhält man

$$m_{F,2} = \int_0^\infty t^2 dF(t) = 2 m_F \int_0^\infty t dA_F(t) = 2 m_F \int_0^\infty \bar{A}_F(t) dt$$

und damit

$$m_{F,2} = 2 m_{F,1} \cdot m_{A_F,1}, \quad m_{A_F,1} = \frac{m_{F,2}}{2 m_{F,1}}. \quad (3.7)$$

Satz 3.11 Für $F \in \mathfrak{F}_{NBUE}$ gilt $m_{F,2} \leq 2 m_{F,1}^2$.

Beweis: Aus der vorhergehenden Bemerkung und Lemma 3.2 erhalten wir

$$m_{F,2} = 2 m_{F,1} \int_0^\infty \bar{A}_F(t) dt \leq 2 m_{F,1} \int_0^\infty \exp\left(-\frac{t}{m_{F,1}}\right) dt = 2 m_{F,1}^2. \quad (3.8)$$

□

Die folgenden Sätze beinhalten Abschätzungen des Abstandes einer Verteilung $F \in \mathfrak{F}_{NBUE}$ von der Exponentialverteilung $1 - e^{-t/m_F}$. Zunächst bezeichnen wir mit $\delta(F) = 1 - \frac{m_{F,2}}{2 m_{F,1}^2} \geq 0$ und beweisen folgendes Lemma:

Lemma 3.3 Für jede Verteilung $F \in \mathfrak{F}_{NBUE}$ gilt $\int_0^\infty \Delta(t) dt = \delta(F) \cdot m_{F,1}$, wobei $\Delta(t) = \bar{F}(t) - \bar{A}_F(t) \geq 0$ ist.

Beweis:

$$\int_0^\infty \Delta(t) dt = \int_0^\infty (\bar{F}(t) - \bar{A}_F(t)) dt = m_F - \int_0^\infty \bar{A}_F(t) dt = m_{F,1} - m_{A_F,1}$$

und aus (3.7) folgt weiter

$$\int_0^\infty \Delta(t) dt = m_{F,1} - \frac{m_{F,2}}{2 m_{F,1}} = m_{F,1} \left(1 - \frac{m_{F,2}}{2 m_{F,1}^2}\right) = m_{F,1} \delta(F).$$

□

Die folgenden Sätze dieses Abschnittes stammen von SOLOV'EV.

Satz 3.12 Für $F \in \mathfrak{F}_{NBUE}$ gilt

$$\left| \bar{F}(t) - \exp\left(-\frac{t}{m_F}\right) \right| \leq \Delta_0 = \sup_{t \geq 0} \Delta(t).$$

Beweis: Zunächst leiten wir eine geeignete Darstellung für $\bar{A}_F(t)$ her. Aus $\frac{d\bar{A}_F(t)}{dt} = -\frac{\bar{F}(t)}{m_F}$ folgt

$$\Delta(t) = \bar{F}(t) - \bar{A}_F(t) = m_F \frac{\bar{F}(t)}{m_F} - \bar{A}_F(t) = -m_F \frac{d\bar{A}_F(t)}{dt} - \bar{A}_F(t) .$$

Somit genügt $\bar{A}_F(t)$ der Differentialgleichung

$$m_F \frac{d\bar{A}_F(t)}{dt} + \bar{A}_F(t) = -\Delta(t)$$

mit der Anfangsbedingung $\bar{A}_F(0) = 1$.

Die Lösung dieser Differentialgleichung ist

$$\bar{A}_F(t) = \exp\left(-\frac{t}{m_F}\right) - \frac{1}{m_F} \int_0^t \exp\left(-\frac{t-s}{m_F}\right) \Delta(s) ds ,$$

wobei $\exp(-t/m_F)$ die Lösung der zugehörigen homogenen Differentialgleichung ist und die allgemeine Lösung mit Hilfe der Greenschen Funktion $\frac{1}{m_F} \exp(-(t-s)/m_F)$ ermittelt werden kann. Nun erhält man für $\Delta(t)$:

$$\Delta(t) = \bar{F}(t) - \bar{A}_F(t) = \bar{F}(t) - \exp\left(-\frac{t}{m_F}\right) + \frac{1}{m_F} \int_0^t \exp\left(-\frac{t-s}{m_F}\right) \Delta(s) ds$$

und durch Umstellen der Gleichung

$$\bar{F}(t) - \exp\left(-\frac{t}{m_F}\right) = \Delta(t) - \frac{1}{m_F} \int_0^t \exp\left(-\frac{t-s}{m_F}\right) \Delta(s) ds .$$

Hieraus folgt einerseits wegen $\frac{1}{m_F} \int_0^t e^{-\frac{t-s}{m_F}}(s) ds \geq 0$:

$$\bar{F}(t) - \exp\left(-\frac{t}{m_F}\right) \leq \Delta(t) \leq \Delta_0$$

und andererseits wegen $\Delta(t) \geq 0$:

$$\begin{aligned} \bar{F}(t) - \exp\left(-\frac{t}{m_F}\right) &\geq -\frac{1}{m_F} \int_0^t \exp\left(-\frac{t-s}{m_F}\right) \Delta(s) ds \geq \\ &\geq -\frac{1}{m_F} \int_0^t \exp\left(-\frac{t-s}{m_F}\right) \Delta_0 ds = -(1 - \exp\left(-\frac{t}{m_F}\right)) \Delta_0 \geq -\Delta_0 . \end{aligned}$$

Damit ist die Behauptung des Satzes bewiesen. □

Satz 3.13 Für jede Verteilungsfunktion $F \in \mathfrak{F}_{NBUE}$ gilt die Ungleichung

$$\left| \bar{F}(t) - \exp\left(-\frac{t}{m_F}\right) \right| \leq \sqrt{2\delta(F)} = \sqrt{2 \left(1 - \frac{m_{F,2}}{2m_{F,1}^2} \right)} .$$

Beweis: Wir verwenden die Bezeichnung $\Delta_0 = \sup_{t \geq 0} \Delta(t) = \Delta(t_0)$, wobei t_0 der Punkt ist, in dem $\Delta(t) = \bar{F}(t) - \bar{A}_F(t)$ sein Maximum besitzt. Für $t \leq t_0$ gilt

$$\begin{aligned} \Delta(t) &= \Delta(t_0) + \Delta(t) - \Delta(t_0) = \\ &= \Delta_0 + \bar{F}(t) - \int_t^\infty \frac{\bar{F}(s)}{m_F} ds - \bar{F}(t_0) + \int_{t_0}^\infty \frac{\bar{F}(s)}{m_F} ds = \\ &= \Delta_0 + (\bar{F}(t) - \bar{F}(t_0)) - \int_t^{t_0} \frac{\bar{F}(u)}{m_F} du \geq \Delta_0 - \int_t^{t_0} \frac{du}{m_F} = \Delta_0 - \frac{t_0 - t}{m_F}. \end{aligned}$$

Für $t = 0$ gilt $\Delta(0) = 0$ und $0 \geq \Delta_0 - \frac{t_0}{m_F}$, folglich ist $t_0 - \Delta_0 m_F \geq 0$. Aus Lemma 3.3 folgt

$$\begin{aligned} \delta(F) m_F &= \int_0^\infty \Delta(t) dt \geq \int_{t_0 - \Delta_0 m_F}^{t_0} \Delta(t) dt \geq \int_{t_0 - \Delta_0 m_F}^{t_0} \left(\Delta_0 - \frac{t_0 - t}{m_F} \right) dt = \\ &= \Delta_0 (\Delta_0 m_F) - \frac{1}{m_F} \int_{t_0 - \Delta_0 m_F}^{t_0} (t_0 - t) dt = \Delta_0^2 m_F - \frac{1}{m_F} \int_0^{\Delta_0 m_F} s ds = \\ &= \Delta_0^2 m_F - \frac{1}{m_F} \left(\frac{\Delta_0 m_F}{2} \right)^2 = \frac{\Delta_0^2 m_F}{2}. \end{aligned}$$

Damit erhalten wir $\delta(F) m_F \geq \frac{\Delta_0^2 m_F}{2}$ oder $\Delta_0^2 \leq 2\delta(F)$. Somit ist $\Delta_0 \leq \sqrt{2\delta(F)}$. Da auf Grund von Lemma 3.12 $\left| \bar{F}(t) - \exp(-\frac{t}{m_F}) \right| \leq \Delta_0$ ist, folgt hieraus die Aussage des Satzes. \square

Schränkt man sich auf die engere Familie $\mathfrak{F}_{\text{IFR}}$ ein, so läßt sich die Ungleichung in Satz 3.13 verschärfen:

Satz 3.14 Für $F \in \mathfrak{F}_{\text{IFR}}$ gilt die Ungleichung

$$\left| \bar{F}(t) - \exp(-\frac{t}{m_F}) \right| \leq 1 - \sqrt{1 - 2\delta(F)} \quad \forall t > 0.$$

Wenn $\delta(F) \rightarrow 0$, so erhält man für den Abstand

$$\sup_{t \geq 0} \left| \bar{F}(t) - \exp(-\frac{t}{m_F}) \right| \leq \delta(F) + o(\delta(F)).$$

Bemerkung: Gehört eine Verteilung $F(\cdot)$ zur Familie $\mathfrak{F}_{\text{NBUE}}$, so ist die Gleichung $\delta(F) = 0$ eine charakteristische Eigenschaft dafür, daß F zur Familie der Exponentialverteilungen gehört.

Zum Schluß dieses Abschnittes sei noch vermerkt, daß es möglich ist, nichtparametrische Familien von Verteilungen mit den umgekehrten Eigenschaften zu definieren, z.B. Verteilungen mit fallender Ausfallrate (DFR–Verteilungen oder decreasing failure rate Verteilungen), NWU–Verteilungen (new worse than used) usw.

Für diese Familien von Verteilungen können verschiedene andere Ungleichungen bewiesen werden, die jedoch über den Rahmen der Vorlesung hinausgehen.

4 Ausfallmodelle

4.1 Abnutzungsmodelle

4.1.1 Die Birnbaum–Saunders–Verteilung

Die Birnbaum–Saunders–Verteilung entsteht als Lebensdauerverteilung eines einfachen Abnutzungsmodelles. Nehmen wir an, daß eine zufällige Größe $X(t)$ den Zustand eines Bauteiles zur Zeit t charakterisiert. Dieser Zustand ändert sich nach der Beziehung

$$X(t) = \mu t + \sigma Z \sqrt{t}, \quad \mu > 0, \sigma > 0, \quad (4.1)$$

wobei Z eine standardnormalverteilte Zufallsgröße ist, d.h. $P(Z \leq z) = \Phi(z)$.

So ein Modell heißt auch Modell mit einmaliger Wirkung des Zufalls: Für jedes spezielle Bauteil wird die Realisierung einer normalverteilten Zufallsgröße bestimmt und dann verläuft der Abnutzungsprozeß deterministisch.

Betrachten wir nun

$$\frac{dX(t)}{dt} = \mu + \frac{\sigma Z}{2\sqrt{t}}.$$

Man kann zeigen, daß für große t diese Ableitung immer positiv ist. Damit werden alle Trajektorien des Abnutzungsprozesses nach einer gewissen Zeit monoton wachsend und überschreiten eine Grenze $h > 0$ zur zufälligen Zeit $S(h)$ (siehe Abbildung 4.1). Wir wollen annehmen, daß die Überschreitung von h zu einem Ausfall des Bauteiles

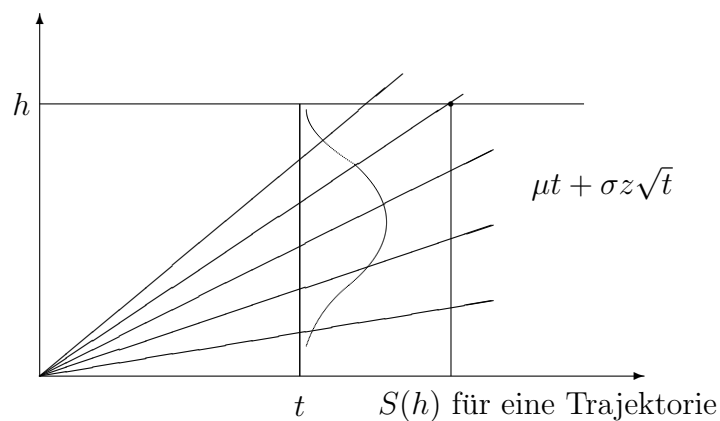


Abbildung 4.1: Trajektorien des Abnutzungsprozesses bis zur Erstüberschreitung

führt. Für die Verteilung der Erstüberschreitungszeit benutzt man die Beziehung zu

dem entsprechenden Abnutzungsprozeß: Die erste Überschreitung von h erfolgt vor dem Zeitpunkt s , wenn sich zur Zeit s der Prozeß oberhalb des Niveaus h befindet:

$$\begin{aligned} F_{S(h)}(s) &= P(S(h) \leq s) = P(X(s) \geq h) = P(\mu s + \sigma Z \sqrt{s} \geq h) \\ &= P(Z > \frac{h - \mu s}{\sigma \sqrt{s}}) = \bar{\Phi}(\frac{h - \mu s}{\sigma \sqrt{s}}). \end{aligned} \quad (4.2)$$

Definition 4.1 Eine zufällige Größe S mit der Verteilungsfunktion (4.2) heißt Birnbaum–Saunders–Verteilung.

Damit enthält die Familie der Birnbaum–Saunders–Verteilungen die drei Parameter $\underline{\theta} = (\mu, \sigma, h)$. Erwartungswert und Varianz dieser Verteilung sind durch

$$\begin{aligned} m_{F,1} &= \frac{h}{\mu} + \frac{\sigma^2}{2\mu^2}, \\ \sigma_F^2 &= m_{F,2} - (m_{F,1})^2 = h \frac{\sigma^2}{\mu^3} + \frac{5\sigma^4}{4\mu^4} \end{aligned}$$

gegeben.

4.1.2 Die Inverse Gaußverteilung als Lebensdauerverteilung bei speziellen Abnutzungsprozessen

Das Ausfallverhalten von technischen Erzeugnissen wird häufig durch Abnutzungsprozesse (Alterung, Verschleiß, Ermüdung usw.) bestimmt. Typische Verläufe dieser Abnutzungsprozesse sind in Abbildung 4.2 dargestellt.

Das Zeitintervall I beschreibt Einlaufphasen, bei denen die Abnutzung einen degres-

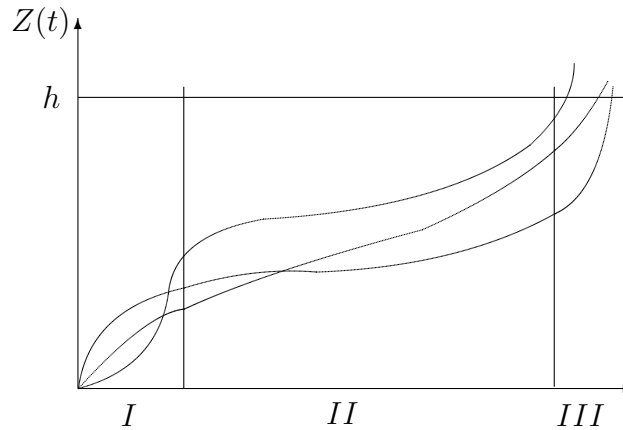


Abbildung 4.2: Verläufe von Abnutzungsprozessen

siven Verlauf aufweist, es folgt eine lineare Phase II und nach einer gewissen Zeit eine progressive Phase III. Der Ausfall des Erzeugnisses erfolgt dann, wenn die Abnutzungsentwicklung das Niveau h erreicht hat.

Mit $Z(t)$ bezeichnen wir einen stochastischen Prozeß, der die Abnutzungsentwicklung beschreibt. Für viele Anwendungsfälle ist der Wienerprozeß mit Drift ein geeignetes Modell, da dieser Prozeß normalverteilte Zuwächse besitzt und die Normalverteilung als Grenzverteilung bei der Überlagerung vieler kleiner zufälliger Einflüsse auftritt. Hier sollen kurz Modelle vorgestellt werden, die die Einlaufphase und die lineare Phase des Abnutzungsverlaufes beschreiben.

Dazu betrachten wir einen stochastischen Prozeß

$$Z(t) = \begin{cases} x_0 + X(t - t_0), & t > t_0 \\ x_0, & t \leq t_0 \end{cases}, \quad (4.3)$$

wobei die beiden Parameter x_0 und t_0 den Anfangszustand des Prozesses und den Abnutzungsbeginn darstellen. Durch diese Parameter ist es möglich, die nichtlineare Einlaufphase des Abnutzungsprozesses zu beschreiben. Darauf kommen wir etwas später noch einmal zurück.

$X(t)$ ist ein homogener Wienerprozeß mit Drift, das heißt

$$X(t) = \sigma W(t) + \mu t, \quad (4.4)$$

wobei $W(t)$ der Standard-Wienerprozeß ist. Damit sind die Zuwächse $X(t) - X(s)$ unabhängig und normalverteilt mit den Parametern

$$E(X(t) - X(s)) = \mu(t - s), \quad \text{Var}(X(t) - X(s)) = \sigma^2(t - s).$$

Der Mittelwert und die Varianzfunktion des Prozesses sind dann durch

$$\begin{aligned} E(Z(t)) &= x_0 + \mu(t - t_0), \\ \text{Var}(Z(t)) &= \sigma^2(t - t_0) \end{aligned}$$

gegeben, wobei μ die Abnutzungsintensität und σ^2 der Varianzparameter des Prozesses sind. Damit enthält der Abnutzungsprozeß vier Parameter, die eine gute Anpassung des Modells an die praktischen Gegebenheiten ermöglichen. Die Parameter μ und σ^2 kennzeichnen den Prozeßverlauf. Die beiden Anfangsparameter x_0 und t_0 ermöglichen es, nichtlineare Einlaufphasen zu berücksichtigen. Sie enthalten als Spezialfälle sowohl einen normalverteilten Anfangszustand des Abnutzungsteiles (z. B. herstellungsbedingt oder um damit die degressive Einlaufphase abzufangen) als auch eine „Inkubationszeit“, d.h. eine Zeitspanne bis zum Einsetzen des Abnutzungsverlaufes:

Ist $t_0 > 0$, so hat man eine „Inkubationsphase“ bis zur Zeit t_0 vorzuliegen; ist $t_0 < 0$, so erhält man zur Zeit $t = 0$ einen normalverteilten Anfangszustand mit den Parametern

$$\begin{aligned} E(Z(0)) &= x_0 - \mu t_0 = x_0 + \mu |t_0|, \\ \text{Var}(Z(0)) &= -\sigma^2 t_0 = \sigma^2 |t_0|. \end{aligned}$$

Die Unabhängigkeit der Zuwächse des Wienerprozesses ermöglicht es auch, relativ einfach Verfahren der klassischen Statistik auf die Beobachtung des Prozesses zu diskreten Zeitpunkten zu verallgemeinern.

Vergleicht man die Darstellung (4.4) mit der Formel (4.1), so sieht man, daß in beiden Fällen eine lineare Mittelwertfunktion μt mit einer „zufälligen“ Wirkung $\sigma W(t)$ bzw. $\sigma Z\sqrt{t}$ überlagert wurde. Die ersten beiden Momente des zufälligen Anteils sind gleich:

$$E\sigma W(t) = E\sigma Z\sqrt{t} = 0 ,$$

$$E(\sigma W(t))^2 = E(\sigma Z\sqrt{t})^2 = \sigma^2 t ,$$

jedoch unterscheiden sich die beiden Modelle wesentlich.

Während im Modell (4.1) für jede Realisierung der Zufall nur einmal wirkt und dann die Trajektorie deterministisch verläuft, wird im Modell (4.4) die Mittelwertfunktion durch einen Wienerprozeß überlagert.

Wir berechnen nun die Dichte der resultierenden Lebensdauerverteilung als Dichte der Erstüberschreitungszeit eines Niveaus h . Sei wieder $S(h)$ die zufällige Zeit des ersten Überschreitens des Niveaus h . Die einfache Herangehensweise aus dem vorigen Abschnitt ist hier ungeeignet. Wegen des dortigen deterministischen Verlaufes und der Monotonie der Trajektorien überschreitet jede Trajektorie genau einmal das Niveau h . Im hier betrachteten Fall ist es jedoch möglich, daß der Wienerprozeß das Niveau h überschreitet und danach wieder unterschreitet. Wir suchen also die Verteilung von

$$S(h) = \inf(t : Z(t) \geq h) .$$

Aus der Formel über die totale Wahrscheinlichkeit und der Markoveigenschaft für den Wienerprozeß erhalten wir folgende Integralgleichung:

$$\begin{aligned} P(Z(t) > x) &= \int_{t_0}^t P(Z(t) > x | S(h) = z) dF_{S(h)}(z) = \\ &= \int_{t_0}^t P(Z(t) > x | Z(z) = h) dF_{S(h)}(z) . \end{aligned} \quad (4.5)$$

Die Beziehung (4.5) sagt aus, daß der Wienerprozeß zum Zeitpunkt t die Höhe $x > h$ erreicht hat, wenn er zur Zeit z das Niveau h erstmalig überschreitet und in der Zeit $(t - z)$ den Zuwachs $(x - h)$ besitzt (siehe Abbildung 4.3).

Wir betrachten hier der Einfachheit halber die Lösung der Integralgleichung (4.5) für den Fall $\mu = 0, x_0 = 0, t_0 = 0$. In diesem Fall ist $Z(t) = \sigma W(t)$ und

$$P(Z(t) > x) = P(\sigma W(t) > x) = P\left(\frac{W(t)}{\sqrt{t}} > \frac{x}{\sigma\sqrt{t}}\right) = \bar{\Phi}\left(\frac{x}{\sigma\sqrt{t}}\right) ,$$

da $\frac{W(t)}{\sqrt{t}}$ einer Standardnormalverteilung unterliegt. Weiterhin gilt:

$$\begin{aligned} P(Z(t) > h | Z(z) = h) &= P(Z(t) - h > 0 | Z(z) = h) = \\ &= P(Z(t) - Z(z) > 0 | Z(z) = h) . \end{aligned}$$

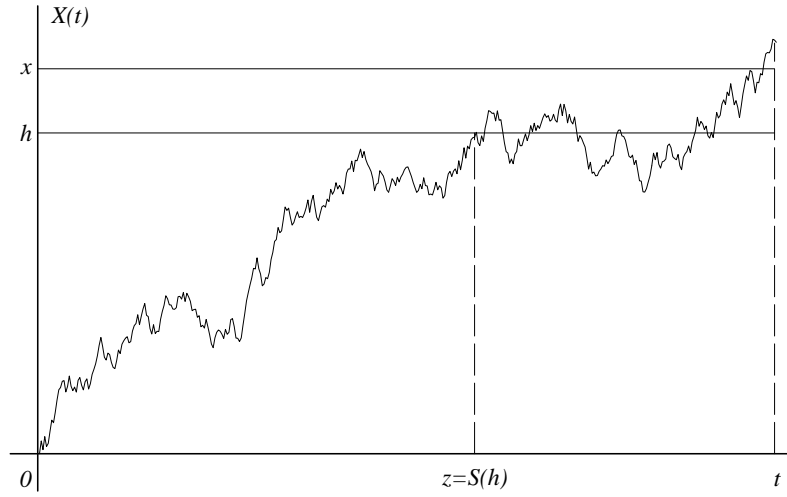


Abbildung 4.3: Eine Trajektorie des Wienerprozesses mit Drift bis zur Erstüberschreitung

Da der Wienerprozeß unabhängige Zuwächse besitzt, ist $Z(t) - Z(z)$ unabhängig von $Z(z)$. Der Zuwachs $Z(t) - Z(z) = \sigma(W(t) - W(z))$ ist normalverteilt mit dem Erwartungswert 0. Daher gilt

$$P(Z(t) - Z(z) > 0 | Z(z) = h) = P(Z(t) - Z(z) > 0) = \frac{1}{2}.$$

Aus Gleichung (4.5) erhält man nun

$$P(Z(t) > h) = \int_0^t \frac{1}{2} dF_{S(h)}(z) = \frac{1}{2} F_{S(h)}(t).$$

Dann folgt für die Verteilung der zufälligen Erstüberschreitungszeit

$$F_{S(h)}(t) = 2P(Z(t) > h) = 2P\left(\frac{W(t)}{\sigma\sqrt{t}} > \frac{h}{\sigma\sqrt{t}}\right) = 2\bar{\Phi}\left(\frac{h}{\sigma\sqrt{t}}\right)$$

und für die Dichte

$$\begin{aligned} f_{S(h)}(t) &= \frac{dF_{S(h)}(t)}{dt} = 2 \frac{d\bar{\Phi}\left(\frac{h}{\sigma\sqrt{t}}\right)}{dt} = \\ &= \varphi\left(\frac{h}{\sigma\sqrt{t}}\right) \frac{h}{\sigma t^{3/2}} = \frac{h}{\sqrt{2\pi\sigma^2 t^3}} \exp\left(-\frac{h^2}{2\sigma^2 t}\right), \quad t \geq 0. \end{aligned} \quad (4.6)$$

Definition 4.2 Eine Verteilung mit der Dichtefunktion (4.6) heißt Inverse Gaußverteilung.

Die Herleitung der Dichte für den allgemeinen Fall $\mu \neq 0$ ist wesentlich komplizierter. Man kann sich jedoch leicht davon überzeugen, daß

$$f_{S(h)}(t) = \frac{h - x_0}{\sqrt{2\pi\sigma^2(t - t_0)^3}} \exp\left(-\frac{(h - x_0 - \mu(t - t_0))^2}{2\sigma^2(t - t_0)}\right), \quad t \geq t_0 \quad (4.7)$$

Lösung der folgenden Differentialgleichung für die Dichte $f_{S(h)}(\cdot)$ ist:

$$f_{Z(t)}(x) = \int_{t_0}^t f_{S(h)}(z) f_{Z(t-z)}(x - h) dz. \quad (4.8)$$

Die Integralgleichung (4.8) ist das Analogon zu (4.5). Wenn eine Dichte existiert, so läßt sich Gleichung (4.5) (mit dem Argument x) ableiten und man erhält (4.8).

Integriert man die Dichte (4.7), so erhält man die Verteilungsfunktion

$$\begin{aligned} F_{S(h)}(t) &= P(S(h) \leq t) \\ &= \Phi\left(\frac{\mu(t - t_0) - h + x_0}{\sigma\sqrt{(t - t_0)}}\right) + \exp\left(\frac{2\mu(h - x_0)}{\sigma^2}\right) \Phi\left(-\frac{\mu(t - t_0) + h - x_0}{\sigma\sqrt{(t - t_0)}}\right), \quad t \geq t_0. \end{aligned}$$

Erwartungswert und Varianz dieser Verteilung sind im Fall $t_0 = x_0 = 0$

$$m_{F_{S(h)},1} = \frac{h}{\mu}, \quad \sigma_{F_{S(h)}}^2 = \frac{h\sigma^2}{\mu^3}.$$

Man beachte, daß im Fall $\mu = 0$ der Erwartungswert der inversen Gaußverteilung nicht existiert.

Die hier vorgestellten Modelle lassen sich verallgemeinern. Man kann statt einer linearen Mittelwertfunktion andere Funktionen oder Transformationen des Wienerprozesses betrachten. Statt des Wienerprozesses können auch andere Prozesse, z.B. spezielle Klassen von Punktprozessen zur Beschreibung des Abnutzungsverhaltens verwendet werden.

4.2 Einführung in die Erneuerungstheorie mit einigen Anwendungen in der Zuverlässigkeitstheorie

In diesem Abschnitt wollen wir das Ausfallverhalten und die Zuverlässigkeit reparierbarer Erzeugnisse betrachten. Dazu nehmen wir an, daß nach jedem Ausfall das betrachtete Element sofort (d.h. die Austauschzeit ist vernachlässigbar klein) durch ein neues Element desselben Typs ersetzt wird (bzw. so repariert wird, daß es wieder als neues Element angesehen werden kann).

Wir beschränken uns hier auf einige Aussagen der Erneuerungstheorie, soweit sie für die künftigen Abschnitte von Interesse sind. Eine ausführliche Darstellung der Erneuerungstheorie findet sich in COX, SMITH oder ASMUSSEN.

Sei S_1 die zufällige Lebensdauerzeit des ersten Elements, S_2 die des zweiten usw. Wir betrachten die Folge von Lebensdauerzeiten $(S_i)_{i \geq 1}$ und die Folge $(T_i)_{i \geq 1}$, $T_i = S_1 + \dots + S_i$, $T_0 = 0$ der zugehörigen Ausfallzeiten. Die beiden Folgen sind in Abbildung 4.4 dargestellt.

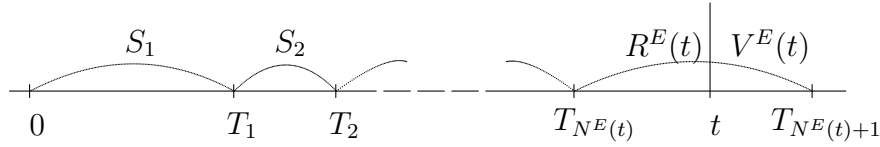


Abbildung 4.4: Ein Erneuerungsprozeß

Definition 4.3 Die Folge $(T_i)_{i \geq 1}, T_i < T_{i+1}$ heißt gewöhnlicher Erneuerungsprozeß, wenn die zugehörige Folge $(S_i)_{i \geq 1}$ mit $S_i = T_i - T_{i-1}, T_0 = 0$ eine Folge unabhängiger und identisch verteilter Zufallsgrößen mit der Verteilungsfunktion $P(S_i \leq s) = F(s), F(0) = 0$ ist.

Sind nur die Zeiten zwischen den Ausfällen S_2, S_3, \dots identisch mit $F(s)$ verteilt und besitzt die zufällige Zeit S_1 bis zum ersten Ausfall die Verteilungsfunktion $F_0(\cdot)$, so heißt $(T_i)_{i \geq 1}$ allgemeiner oder verzögerter Erneuerungsprozeß.

Im Zusammenhang mit einem Erneuerungsprozeß sind einige andere Prozesse von Interesse.

Definition 4.4 Die Anzahl der Erneuerungen (oder die Anzahl der ausgefallenen Elemente) im Intervall $[0, t]$

$$N^E(t) := \sum_{k=1}^{\infty} I(T_k \leq t)$$

heißt Zählprozeß. Die zufällige Zeit seit der letzten Erneuerung

$$R^E(t) := t - T_{N^E(t)}$$

heißt Rückwärtsrekurrenzzzeit des Prozesses. Die zufällige Zeit bis zur nächsten Erneuerung

$$V^E(t) := T_{N^E(t)+1} - t$$

heißt Vorwärtsrekurrenzzzeit des Prozesses.

Zunächst betrachten wir gewöhnliche Erneuerungsprozesse, in denen alle S_i die Verteilungsfunktion $F(s)$ besitzen. Dann erhalten wir für die Verteilung des k -ten Erneuerungszeitpunktes T_k

$$P(T_k \leq t) = P(S_1 + \dots + S_k \leq t) = F^{*k}(t),$$

wobei $F^{*k}(t)$ die k -fache Faltung der Verteilungsfunktion $F(t)$ ist, die aus einer Rekursionsbeziehung

$$F^{*(k)}(t) = F^{*(k-1)} * F(t) = \int_0^t F^{*(k-1)}(t-u) dF(u), \quad F^{*(0)}(u) \equiv 1$$

bestimmt werden kann.

Für den Erwartungswert der Anzahl der Erneuerungen bis zur Zeit t erhält man dann

$$H^E(t) := \mathbb{E}N^E(t) = \mathbb{E} \sum_{k=1}^{\infty} \mathbb{I}(T_k \leq t) = \sum_{k=1}^{\infty} \mathbb{P}(T_k \leq t) = \sum_{k=1}^{\infty} F^{*(k)}(t) . \quad (4.9)$$

Man kann beweisen, daß die Reihe (4.9) mit der Geschwindigkeit einer geometrischen Reihe konvergiert.

Definition 4.5 Die Funktion $H^E(t) = \mathbb{E}N^E(t)$ (erwartete Anzahl von Ausfällen bis zur Zeit t) heißt Erneuerungsfunktion des gewöhnlichen Erneuerungsprozesses.

Satz 4.1 Die Erneuerungsfunktion ist die eindeutige Lösung der Integralgleichung

$$H^E(t) = F(t) + \int_0^t H^E(t-u) dF(u) . \quad (4.10)$$

Beweis: Gemäß Formel (4.9) gilt

$$\begin{aligned} H^E(t) &= F(t) + \sum_{k=2}^{\infty} F^{*(k-1)} * F(t) = F(t) + \sum_{k=1}^{\infty} F^{*(k)} * F(t) = \\ &= F(t) + \sum_{k=1}^{\infty} \int_0^t F^{*(k)}(t-u) dF(u) = F(t) + \int_0^t \left(\sum_{k=1}^{\infty} F^{*(k)}(t-u) \right) dF(u) = \\ &= F(t) + \int_0^t H^E(t-u) dF(u) . \end{aligned}$$

Die Eindeutigkeit der Lösung beweisen wir nur für den Fall $F(t) < 1 \quad \forall t > 0$.

Seien $H_1^E(t)$ und $H_2^E(t)$ zwei Lösungen der Erneuerungsgleichung (4.10). Dann erhält man für die Differenz

$$D(t) = H_2^E(t) - H_1^E(t) = \int_0^t (H_2^E(t-u) - H_1^E(t-u)) dF(u)$$

und damit für alle $t' \leq t$

$$|D(t')| \leq \int_0^{t'} |D(t'-u)| dF(u) \leq \sup_{t' \leq t} |D(t')| F(t) \quad \forall t' \leq t .$$

Daraus folgt

$$\sup_{t' \leq t} |D(t')| \leq \sup_{t' \leq t} |D(t')| F(t) .$$

Das ist nur möglich, wenn $D(t') \equiv 0 \quad \forall t' \leq t$. □

Bemerkung: Analog zu diesem Satz kann bewiesen werden, daß für das zweite Moment

$$H_2^E(t) = E(N^E(t))^2 = \sum_{k=1}^{\infty} (2k-1) F^{*(k)}(t)$$

gilt (siehe COX). Man erhält dann die Gleichung

$$H_2^E(t) = H^E(t) + 2 \int_0^t H^E(t-u) dH^E(u),$$

und hieraus eine Ungleichung für das zweite Moment

$$H_2^E(t) \leq H^E(t) + 2(H^E(t))^2. \quad (4.11)$$

Satz 4.2 Seien $(T_i)_{i \geq 1}$ ein allgemeiner Erneuerungsprozeß und $H_0^E(t) := EN_0^E(t)$ die Erneuerungsfunktion dieses Prozesses. Dann gilt

$$H_0^E(t) = F_0(t) + \int_0^t H^E(t-u) dF(u),$$

wobei $H^E(t)$ die Erneuerungsfunktion des zugehörigen gewöhnlichen Erneuerungsprozesses ist.

Der Beweis dieses Satzes ist völlig analog zum Beweis der Integralgleichung (4.10), jedoch für die Verteilungsfunktion der Zeit bis zur ersten Erneuerung wird jeweils $F_0(t)$ eingesetzt.

Bemerkung: Für absolut stetige Funktionen $F_0(s)$ und $F(s)$ existieren die entsprechenden Dichten $f_0(s)$ und $f(s)$. In diesem Fall kann neben der Erneuerungsfunktion die Erneuerungsichte

$$h^E(s) = \frac{dH^E(s)}{ds}$$

betrachtet werden.

Für die Erneuerungsichten des gewöhnlichen und allgemeinen Erneuerungsprozesses erhält man dann die Integralgleichungen

$$h^E(t) = f(t) + \int_0^t h^E(t-u) f(u) du$$

bzw.

$$h_0^E(t) = f_0(t) + \int_0^t h^E(t-u) f_0(u) du.$$

Die Lösung der Integralgleichungen für Erneuerungsfunktionen und Erneuerungsichten ist u. a. mit Hilfe von Laplace-Transformationen möglich.

Ist $g(\cdot)$ eine reellwertige Funktion von s mit endlicher Variation, so wird ihre Laplace-Transformierte $\tilde{g}(z)$ durch

$$\tilde{g}(z) = \int_0^{\infty} e^{-zs} dg(s)$$

definiert. Die Laplace-Transformierte $\tilde{g}(z)$ existiert unter einigen schwachen Voraussetzungen, die für F und H^E erfüllt sind. Seien $\tilde{F}(z)$ und $\tilde{F}_0(z)$ die Laplace-Transformierten von den Verteilungen $F(s)$ und $F_0(s)$ und $\tilde{H}^E(z)$ bzw. $\tilde{H}_0^E(z)$ die Laplace-Transformierten der Erneuerungsfunktionen $H^E(s)$, $H_0^E(s)$ für den gewöhnlichen bzw. allgemeinen Erneuerungsprozeß.

$\tilde{F}(z)$ und $\tilde{F}_0(z)$ können als Erwartungswerte der zufälligen Größen e^{-zS_i} interpretiert werden :

$$\tilde{F}(z) = \text{E}e^{-zS_i}, \quad i = 2, 3, \dots, \quad \tilde{F}_0(z) = \text{E}e^{-zS_1}.$$

Diese Erwartungswerte existieren für beliebige $z \geq 0$, da $e^{-zS_i} \leq 1$ ist. Aus der Unabhängigkeit der S_i folgt für die Laplace-Transformierte der k -fachen Faltung

$$\tilde{F}^{*(k)}(z) = \int_0^{\infty} e^{-zs} dF^{*(k)}(s) = \text{E}e^{-z(S_1+\dots+S_k)} = \text{E}e^{-zS_1} \dots \text{E}e^{-zS_k} = (\tilde{F}(z))^k.$$

Daraus erhält man

$$\begin{aligned} \tilde{H}^E(z) &= \int_0^{\infty} e^{-zs} d\left(\sum_{k=1}^{\infty} F^{*(k)}(s)\right) = \sum_{k=1}^{\infty} \int_0^{\infty} e^{-zs} dF^{*(k)}(s) = \\ &= \sum_{k=1}^{\infty} (\tilde{F}(z))^k = \frac{\tilde{F}(z)}{1 - \tilde{F}(z)}. \end{aligned} \quad (4.12)$$

Damit existiert $\tilde{H}^E(z)$ und ist durch (4.12) definiert. Analog erhält man für den allgemeinen Erneuerungsprozeß

$$\tilde{H}_0^E(z) = \frac{\tilde{F}_0(z)}{1 - \tilde{F}(z)}. \quad (4.13)$$

Für einige Verteilungsfunktionen ist es möglich, die Laplace-Transformierten zu berechnen, die Erneuerungsfunktion im Bildraum zu bestimmen und die Rücktransformation durchzuführen. Die größte Schwierigkeit bereitet hierbei die Rücktransformation. Ist diese analytisch nicht möglich, kann man numerische Verfahren verwenden. Häufig beschränkt man sich auch auf die Laplace-Transformierte der Erneuerungsfunktion, da hieraus z.B. Momente berechnet werden können.

Beispiel 4.1 Wir betrachten einen gewöhnlichen Erneuerungsprozeß, in dem die Zeiten zwischen zwei Erneuerungen S_i , $i = 1, 2, \dots$ einer Exponentialverteilung unterliegen (dieser Prozeß heißt homogener Poissonprozeß und wird im nächsten Abschnitt aus

anderer Sicht näher untersucht). In diesem Fall erhält man für die Erneuerungsfunktion $H^E(t) = \lambda t$ (siehe Übungsaufgaben). Da in diesem Fall $m_{F,1} = \frac{1}{\lambda}$ ist, gilt also $H^E(t) = \frac{t}{m_{F,1}}$. Die Erneuerungsfunktion ist somit eine lineare Funktion in t . Diese Eigenschaft heißt Stationarität des Erneuerungsprozesses. \diamond

Wir betrachten nun allgemeine Erneuerungsprozesse hinsichtlich ihrer Stationarität.

Satz 4.3 (Satz von Blackwell) *Sei $F(\cdot)$ eine stetige Lebensdauerverteilung mit $m_{F,1} < \infty$. Dann gilt für beliebige s*

$$\lim_{t \rightarrow \infty} (H_0^E(t+s) - H_0^E(t)) = \frac{s}{m_{F,1}}. \quad (4.14)$$

Sei weiterhin $Q(t)$ eine beschränkte nichtwachsende Funktionen auf dem Intervall $[0, \infty)$ mit $\int_0^\infty Q(u) du < \infty$. Dann gilt

$$\lim_{t \rightarrow \infty} \int_0^t Q(t-u) dH_0(u) = \frac{1}{m_{F,1}} \int_0^\infty Q(u) du.$$

Den Beweis dieses Satzes findet man in FELLER, Bd. 2. Aus dem Satz von Blackwell lassen sich verschiedene asymptotische Aussagen über die Verteilungen der beiden Rekurrenzenzeiten, der Varianz des Zählprozesses u.a. ableiten. Dies wollen wir nun am Beispiel der Verteilung der Vorwärtsrekurrenzenzeiten demonstrieren.

Dazu betrachten wir die Größe $A_F(s) = \frac{1}{m_{F,1}} \int_0^s \bar{F}(u) du$.

Satz 4.4 *Unter den Voraussetzungen des Satzes von Blackwell gilt für die Verteilung der Vorwärtsrekurrenzenzeit eines Erneuerungsprozesses*

$$\lim_{t \rightarrow \infty} P(V^E(t) > s) = \bar{A}_F(s) := \frac{1}{m_{F,1}} \int_s^\infty \bar{F}(u) du.$$

Beweis: Sei $G_t(\cdot)$ die Verteilungsfunktion der Vorwärtsrekurrenzenzeit $G_t(s) = P(V^E(t) \leq s)$, d.h. die Wahrscheinlichkeit, daß der nächste Ausfall vor Ablauf der Zeit s erfolgt, wenn wir uns im Zeitpunkt t eines Erneuerungsprozesses befinden.

Dann folgt aus der Formel über die totale Wahrscheinlichkeit

$$P(V^E(t) > s) = P(S_1 > t+s) + \sum_{k=1}^{\infty} \int_0^t P(S_{k+1} > t-u+s) dF^{*(k)}(u),$$

wobei $F^{*(1)}(\cdot) = F_0(\cdot)$. Zur Interpretation dieser Formel dient Abbildung 4.5: $V^E(t)$ ist größer als s , wenn entweder die erste Erneuerung später als $t+s$ erfolgt oder wenn zum Zeitpunkt u die k -te Erneuerung stattgefunden hat ($0 < t \leq u, k = 1, 2, \dots$) und die $(k+1)$ -te Erneuerung erst nach Ablauf der Zeit $t-u+s$ erfolgt. Hieraus erhält man

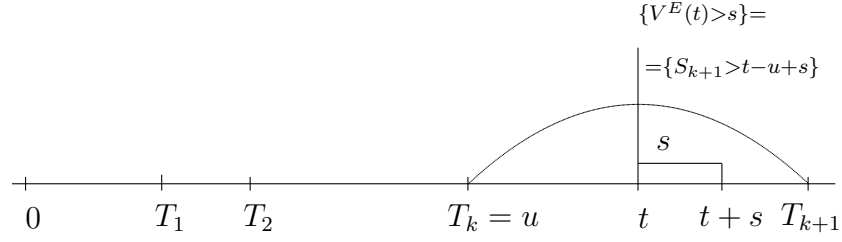


Abbildung 4.5: Die Vorwärtsrekurrenzenzeiten

$$\begin{aligned}
& P(V^E(t) > s) \\
&= P(S_1 > t + s) + \int_0^t P(S_{N^E(t)+1} > t - u + s | T_{N^E(t)} = u) dH_0^E(u) = \\
&= \bar{F}_0(t + s) + \int_0^t \bar{F}(t - u + s) dH_0^E(u) .
\end{aligned}$$

Wir wenden nun den Satz von Blackwell mit $Q(u) = \bar{F}(u + s)$ an. Für $t \rightarrow \infty$ konvergiert der erste Summand gegen 0 und der zweite gegen

$$\frac{1}{m_{F,1}} \int_0^\infty \bar{F}(u + s) du .$$

Daraus erhalten wir die geforderte Aussage. \square

Bemerkung 1: Der Satz von Blackwell läßt sich für den Beweis vieler anderer Aussagen verwenden. Zum Beispiel sind häufig genauere Aussagen über das Verhalten der Erneuerungsfunktion als in Formel (4.14) oder Aussagen über das Verhalten der Varianz des Zählprozesses von Interesse. Durch geeignete Wahl der Funktion $Q(t)$ im Satz von Blackwell erhält man z.B. für gewöhnliche Erneuerungsprozesse

$$H^E(t) = \frac{t}{m_{F,1}} + \left(\frac{m_{F,2}}{2m_{F,1}^2} - 1 \right) + g_1(t)$$

mit $g_1(t) \rightarrow 0$ für $t \rightarrow \infty$ und

$$\text{Var}(N^E(t)) = \frac{\sigma_F^2}{m_{F,1}^3} \cdot t + \frac{1}{2} \frac{m_{F,2}^2}{m_{F,1}^4} - \frac{m_{F,3}}{3m_{F,1}^3} + g_2(t)$$

mit $g_2(t) \rightarrow 0$ für $t \rightarrow \infty$.

Bemerkung 2: Weiterhin läßt sich zeigen, daß der Zählprozeß asymptotisch normalverteilt ist: Für beliebige x gilt

$$P\left(\frac{N^E(t) - H^E(t)}{\sqrt{\text{Var } N^E(t)}} \leq x\right) \rightarrow \Phi(x), \quad t \rightarrow \infty,$$

wobei $\Phi(\cdot)$ die Verteilungsfunktion der Standardnormalverteilung ist. Die asymptotische Normalverteilung kann zur näherungsweisen Bestimmung von oberen und unteren Grenzen für die Anzahl von Ausfällen in einem gegebenen Zeitintervall genutzt werden.

Wir betrachten nun eine sinnvolle Verallgemeinerung von Erneuerungsprozessen. Die Reparaturzeit soll nicht mehr vernachlässigbar klein sein, sondern eine zufällige Zeit U_i in Anspruch nehmen. Ein Prozeß dieser Art ist in Abbildung 4.6 dargestellt.

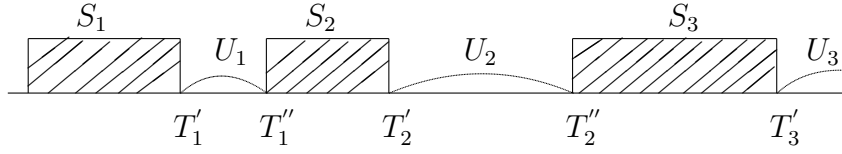


Abbildung 4.6: Ein alternierender Prozeß

Nach einer Arbeitszeit S_i folgt jeweils eine Reparaturzeit U_i . T'_i seien die Ausfallzeitpunkte und T''_i die Zeitpunkte, zu denen eine Reparatur beendet wird und die $(i+1)$ -te Arbeitsperiode beginnt. Wir nehmen an, daß die Reparatur eine vollständige ist, d.h. nach der Reparatur wird das Element als neu angesehen.

Die Arbeitszeiten besitzen wieder die Verteilungsfunktion $F(s) = P(S_i \leq s)$ und die Reparaturzeiten besitzen die Verteilungsfunktion $G(s) = P(U_i \leq s)$, $i = 1, 2, \dots$. Die U_i seien unabhängig von den S_i , $i = 1, 2, \dots$.

Definition 4.6 Die Folge $(T'_i, T''_i)_{i \geq 1}$ heißt gewöhnlicher alternierender Prozeß. Unterliegt S_1 einer anderen Verteilung $F_0(s) = P(S_1 \leq t)$, so nennt man $(T'_i, T''_i)_{i \geq 1}$ allgemeinen alternierenden Prozeß.

Analog zu Erneuerungsprozessen lassen sich auch hier verschiedene Aussagen beweisen. Betrachten wir einen gewöhnlichen alternierenden Prozeß mit stetigen Verteilungen $F(t)$ und $G(t)$. Die Erwartungswerte dieser beiden Verteilungen seien endlich: $m_{F,1} < \infty$, $m_{G,1} < \infty$.

Wenn wir nun nur die Folge $(T'_i)_{i \geq 1}$ betrachten, so erhalten wir einen allgemeinen Erneuerungsprozeß mit $F_0(s) = F(s)$ und der Verteilung der Wartezeiten zwischen zwei Erneuerungen $\tilde{F}(s) = F * G(s)$, da $\tilde{F}(s) = P(S_i + U_{i-1} \leq s)$, $i = 1, 2, \dots$. Die Erneuerungsfunktion dieses Prozesses bezeichnen wir mit $H_0^A(t)$ und aus dem Satz von Blackwell folgt

$$H_0^A(t+s) - H_0^A(t) \rightarrow \frac{s}{m_{F,1} + m_{G,1}}, \quad t \rightarrow \infty.$$

Ebenso einfach läßt sich die asymptotische Verfügbarkeit eines Elementes bestimmen. Nehmen wir an, daß ein Element die beiden Zustände $\{X = 0\}$ für „ausgefallen“ und $\{X = 1\}$ für „intakt“ annehmen kann. Die Verfügbarkeit des Elements ist die Wahrscheinlichkeit, daß sich das Element im arbeitsfähigem Zustand befindet. Aus der Formel über die totale Wahrscheinlichkeit erhalten wir

$$p(t) = P(X(t) = 1) = \bar{F}(t) + \int_0^t \bar{F}(t-u) dH_0^A(u) \quad (4.15)$$

Wendet man hierauf den Satz von Blackwell mit $Q(t-u) = \bar{F}(t-u)$ an, so erhält man, daß der zweite Term (und damit auch $p(t)$) für $t \rightarrow \infty$ gegen

$$\frac{1}{m_{\bar{F},1}} \int_0^\infty Q(u) du = \frac{m_{F,1}}{m_{\bar{F},1}} = \frac{m_{F,1}}{m_{F,1} + m_{G,1}}$$

konvergiert.

Bemerkung: Wenn in diesen Modellen $m_{F,1} \gg m_{G,1}$, so gilt $\frac{m_{F,1}}{m_{F,1} + m_{G,1}} \sim 1$ und damit erhält man wieder die Beziehungen der einfachen Erneuerungstheorie.

Beispiel 4.2 Betrachten wir einen alternierenden Prozeß mit $F(s) = 1 - e^{-\lambda s}$ und $G(s) = 1 - e^{-\mu s}$. Dann läßt sich die Gleichung (4.15) mittels Laplace-Transformation explizit lösen und man erhält

$$p(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\mu} e^{-\lambda t} \left(1 - \frac{\lambda}{\lambda + \mu} e^{-\mu t}\right).$$

In diesem Beispiel wird deutlich, inwieweit sich asymptotische und exakte Verfügbarkeit voneinander unterscheiden: Die Differenz beider Größen ist ein exponentiell schnell abklingender Term. Somit kann man schon nach relativ kurzer Einlaufphase die asymptotischen Ergebnisse zur Beschreibung der Arbeit des Bauteiles nutzen. \diamond

Bei Erneuerungsprozessen treten häufig Summen auf, wobei die Anzahl der Summanden zufällig ist. Für diese Summen spielt die folgende Waldsche Identität eine zentrale Rolle.

Satz 4.5 (Identität von Wald) Sei $(S_i)_{i \geq 1}$ eine Folge unabhängiger und identisch verteilter Zufallsgrößen mit der Verteilungsfunktion $F(\cdot)$ und $m_{F,1} < \infty$. Weiterhin sei C eine zufällige Größe mit Realisierungen aus den natürlichen Zahlen, $p_i = P(C = i)$ und $EC < \infty$. Das Ereignis $\{C \geq i\}$ sei unabhängig von S_i , $i = 1, 2, \dots$. Dann gilt

$$E\left(\sum_{i=1}^C S_i\right) = m_{F,1} \cdot EC.$$

Beweis:

$$\begin{aligned} E\left(\sum_{i=1}^C S_i\right) &= E \sum_{i=1}^{\infty} S_i I(C \geq i) = \sum_{i=1}^{\infty} E S_i I(C \geq i) = \\ &= \sum_{i=1}^{\infty} E S_i P(C \geq i) = \sum_{i=1}^{\infty} m_{F,1} \cdot \sum_{k=i}^{\infty} p_k = m_{F,1} \sum_{k=1}^{\infty} k \cdot p_k = m_{F,1} EC. \end{aligned}$$

□

Folgerung 4.1 Für die Vorwärtsrekurrenzzzeit eines gewöhnlichen Erneuerungsprozesses gilt $EV^E(t) = m_{F,1}(H^E(t) + 1) - t$.

Beweis: Das Ereignis $\{N^E(t) \geq i - 1\}$ hängt nur von S_1, \dots, S_{i-1} ab. Daher ist das Ereignis $\{N^E(t) + 1 \geq i\}$ unabhängig von S_i . Die Vorwärtsrekurrenzzeit ist durch

$$V^E(t) = T_{N^E(t)+1} - t \quad \text{mit} \quad T_{N^E(t)+1} = \sum_{i=1}^{N^E(t)+1} S_i$$

definiert. Dann erhält man aus der Waldschen Identität

$$EV^E(t) = E\left(\sum_{i=1}^{N^E(t)+1} S_i\right) - t = m_{F,1}(H^E(t) + 1) - t.$$

□

Erneuerungsprozesse stehen in engem Zusammenhang mit Poissonschen Punktprozessen, die im nächsten Abschnitt betrachtet werden.

4.3 Poissonsche Punktprozesse

Wir betrachten zunächst eine spezielle Verteilung der Wartezeiten in einem Erneuerungsprozeß. Sei S_1, S_2, \dots eine Folge unabhängiger und identisch exponentialverteilter Zufallsgrößen mit der Überlebensfunktion

$$\bar{F}(s) = P(S > s) = e^{-\lambda s}.$$

Mit $(T_n)_{n \geq 1}$ bezeichnen wir wieder die Folge der Ausfallzeitpunkte $T_n = \sum_{i=1}^n S_i$. Desweiteren bezeichnen wir für einen fixierten Zeitpunkt t mit $\Psi(0, t]$ die zufällige Anzahl von Ausfallszeitpunkten T_i im Intervall $(0, t]$, d.h. $\Psi(0, t] = \max(d : T_d \leq t)$. $T_{\Psi(0, t]}$ ist der letzte Erneuerungszeitpunkt vor der Zeit t und $T_{\Psi(0, t] + 1}$ ist der erste Erneuerungszeitpunkt nach der Zeit t (Abbildung 4.7). Dann ist $S_{\Psi(0, t]} = T_{\Psi(0, t]} - T_{\Psi(0, t] - 1}$ nicht exponentialverteilt.

Dieser zum Erneuerungsprozeß gehörende Zählprozeß $\Psi(0, t]$ entspricht der Größe $N(t)$

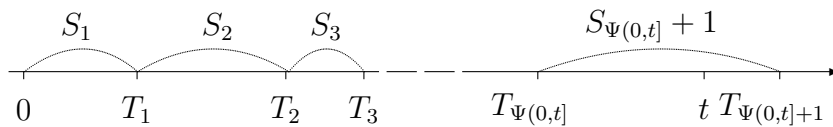


Abbildung 4.7: Ein Erneuerungsprozeß

aus dem letzten Abschnitt. Wir verwenden jedoch von jetzt an die für Punktprozesse übliche Bezeichnung. Durch diesen Zählprozeß kann der gesamte Erneuerungsprozeß beschrieben werden. Man kann sich leicht überlegen, daß sich aus alleiniger Kenntnis von $\Psi(0, t]$ (auch ohne Annahme exponentialverteilter Wartezeiten) alle Erneuerungszeitpunkte bestimmen lassen. Die Anzahl der Erneuerungen bis zur Zeit t und die Erneuerungszeitpunkte sind durch folgende Beziehung miteinander verbunden:

$$\{\Psi(0, t] \geq d\} = \{T_d \leq t\} = \{S_1 + S_2 + \dots + S_d \leq t\}.$$

Setzen wir nun eine Exponentialverteilung für die Zeiten S_j voraus, so läßt sich die Verteilung des Zählprozesses explizit bestimmen:

$$P(\Psi(0, t] \geq d) = P(T_d \leq t) = F^{*d}(t).$$

Es ist also die d -fache Faltung der Exponentialverteilung zu berechnen. Das kann mit Hilfe von charakteristischen Funktionen oder auch auf direktem Wege mit vollständiger Induktion erfolgen. Das Resultat ist wohlbekannt: Die Summe d unabhängiger exponentialverteilter Zufallsgrößen unterliegt einer Erlangverteilung der Stufe d :

$$P(\Psi(0, t] \geq d) = F^{*d}(t) = 1 - \sum_{k=0}^{d-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

Hieraus lassen sich die Einzelwahrscheinlichkeiten der Anzahl der Erneuerungen berechnen:

$$P(\Psi(0, t] = d) = P(\Psi(0, t] \geq d) - P(\Psi(0, t] \geq d+1) = \frac{(\lambda t)^d}{d!} e^{-\lambda t}.$$

Somit ist die Anzahl der Erneuerungen bis zur Zeit t poissonverteilt mit dem Parameter λt .

Es ist leicht zu sehen, daß die Anzahl der Erneuerungen $\Psi(s, t]$ auf dem Zeitintervall $(s, t]$ poissonverteilt mit dem Parameter $\lambda(t-s)$ ist:

$$\Psi(s, t] = \sum_{i=1}^{\infty} I(s < T_i \leq t), \quad P(\Psi(s, t] = d) = \frac{(\lambda(t-s))^d}{d!} e^{-\lambda(t-s)}.$$

Weiterhin folgt sofort aus der Unabhängigkeit der Wartezeiten S_i und der Gedächtnislosigkeit der Exponentialverteilung, daß die Anzahlen von Erneuerungen auf disjunkten Zeitintervallen $(t_i, t_i + s_i]$ unabhängig voneinander sind:

$$P(\Psi(t_i, t_i + s_i] = k_i, \ i = 1, 2, \dots) = \prod_{i=1}^m P(\Psi(t_i, t_i + s_i] = k_i) \\ \text{für } (t_i, t_i + s_i] \cap (t_j, t_j + s_j] = \emptyset, \ i \neq j.$$

Wir definieren nun den homogenen Poissonprozeß.

Definition 4.7 Die zufällige Folge von Punkten $(T_n)_{n \geq 1}$ heißt (homogener) Poissonprozeß mit dem Intensitätsparameter λ , wenn für den zugehörigen Zählprozeß

$$\Psi(0, t] = \sum_{i=1}^{\infty} I(T_i \leq t)$$

die folgenden beiden Eigenschaften erfüllt sind:

- (i) Der Zählprozeß $\Psi(t, t+s]$ ist für alle $t, s > 0$ poissonverteilt mit dem Parameter λs ,

(ii) Die $\Psi(t_i, t_i + s_i]$ sind unabhängig für $t_i + s_i \leq t_{i+1}$, $i = 1, 2, \dots, m$; $m = 1, 2, \dots$.

Bemerkung 1: In KÖNIG, SCHMIDT findet man den Beweis, daß diese Definition äquivalent zur Definition des homogenen Poissonprozesses als Erneuerungsprozeß mit exponentialverteilten Wartezeiten ist.

Bemerkung 2: Statt der Eigenschaft (i) kann zur Definition des Poissonprozesses auch folgende Eigenschaft verwendet werden:

$$(i') \quad P(\Psi(t, t + s] = 0) = e^{-\lambda s},$$

siehe KALLENBERG.

Nun soll eine allgemeine Klasse von Prozessen eingeführt werden. Dazu betrachten wir eine Abbildung $t \rightarrow H(t)$, wobei $H(t)$ eine wachsende Funktion mit $H(0) = 0$ ist. Wir werden, wenn es nicht anders gesagt wird, stetige Funktionen $H(\cdot)$ betrachten. Durch diese Transformation wird der Poissonprozeß mit dem Intensitätsparameter $\lambda = 1$ (andere Parameter können mit der Funktion $H(\cdot)$ erzeugt werden) in einen neuen Prozeß $(T'_n)_{n \geq 1}$ abgebildet. Der zu diesem neuen Prozeß gehörende Zählprozeß wird mit $\Psi_H(0, t]$ bezeichnet. Für ihn erhält man:

$$P(\Psi_H(0, t] = d) = \frac{H(t)^d}{d!} e^{-H(t)}.$$

Jetzt können wir Poissonprozesse allgemeiner definieren.

Definition 4.8 Eine zufällige Punktfolge $(T_n)_{n \geq 1}$ heißt (inhomogener) Poissonscher Punktprozeß mit der Erwartungswertfunktion $H(t)$, wenn

- (i) der Zählprozeß $\Psi(t, t + s] = \sum_{n=1}^{\infty} I(T_n \in (t, t + s])$ in jedem Intervall $(t, t + s]$ poissonverteilt mit dem Parameter $H(t + s) - H(t)$ ist,
- (ii) für jede Folge von disjunkten Intervallen $(t_1, t_1 + s_1], \dots, (t_m, t_m + s_m]$ mit $t_i + s_i \leq t_{i+1}$ die Zufallsgrößen $\Psi(t_i, t_i + s_i]$ $i = 1, 2, \dots, m$ unabhängig voneinander sind, $m = 1, 2, \dots$.

Bemerkung: Wenn $H(t)$ eine stetige Funktion ist, so kann auch für den inhomogenen Poissonprozeß in der Definition die Eigenschaft (i) durch folgende Eigenschaft (i') ersetzt werden:

$$(i') \quad \text{für jedes Intervall } (t, t + s] \text{ gilt} \\ P(\Psi(t, t + s] = 0) = \exp(-[H(t + s) - H(t)]).$$

Den Beweis dazu findet man ebenfalls in KALLENBERG.

Poissonsche Punktprozesse besitzen eine Reihe bemerkenswerter Eigenschaften. Um diese Eigenschaften nachzuweisen, führen wir vorher noch als Hilfsmittel den Begriff der erzeugenden Funktion ein.

Definition 4.9 Sei Ψ eine Zufallsgröße mit ganzzahligen Realisierungen und den Einzelwahrscheinlichkeiten $P(\Psi = k) = p_k$, $k = 0, 1, \dots$, $\sum_{k=0}^{\infty} p_k = 1$. Die Funktion

$$G(z) = \sum_{k=0}^{\infty} z^k p_k = E z^{\Psi}$$

heißt erzeugende Funktion der Zufallsgröße Ψ .

$G(\cdot)$ kann als Funktion einer komplexen Veränderlichen angesehen werden. Innerhalb des Kreises ($z : |z| \leq 1$) ist sie analytisch. Es besteht ein eindeutiger Zusammenhang zwischen $G(\cdot)$ und den Einzelwahrscheinlichkeiten; alle Einzelwahrscheinlichkeiten können aus $G(\cdot)$ bestimmt werden:

$$p_k = \frac{1}{k!} \left. \frac{d^k G(z)}{dz^k} \right|_{z=0}.$$

Der Nutzen von erzeugenden Funktionen zeigt sich in folgendem Lemma:

Lemma 4.1 Ψ_1, \dots, Ψ_m seien unabhängige ganzzahlige Zufallsgrößen mit den erzeugenden Funktionen $G_1(z), \dots, G_m(z)$. Dann besitzt die Zufallsgröße $\Psi = \Psi_1 + \dots + \Psi_m$ die erzeugende Funktion $G(z) = G_1(z) \cdots G_m(z)$.

Beweis: Für die erzeugende Funktion der Summe erhalten wir

$$G(z) = E z^{\Psi} = E z^{\Psi_1 + \Psi_2 + \dots + \Psi_m} = E(z^{\Psi_1} \cdots z^{\Psi_m}).$$

Aus der Unabhängigkeit von Ψ_1, \dots, Ψ_m folgt für den Erwartungswert des Produktes

$$G(z) = E z^{\Psi_1} \cdots E z^{\Psi_m} = G_1(z) \cdots G_m(z).$$

□

Folgerung 4.2 Die Summe von m unabhängigen poissonverteilten Zufallsgrößen mit den Erwartungswerten μ_i ($i = 1, \dots, m$) ist wieder poissonverteilt mit dem Erwartungswert $\mu = \mu_1 + \mu_2 + \dots + \mu_m$.

Beweis: Ψ_i ($i = 1, \dots, m$) hat die erzeugende Funktion $G_i(z) = e^{\mu_i(z-1)}$. Aus Lemma 4.1 erhält man für die erzeugende Funktion der Summe $\Psi_1 + \Psi_2 + \dots + \Psi_m$

$$G(z) = e^{\mu_1(z-1)} \cdots e^{\mu_m(z-1)} = e^{(\mu_1 + \dots + \mu_m)(z-1)}.$$

Wegen der Eineindeutigkeit der Zuordnung besitzt nur die Poissonverteilung mit dem Erwartungswert $\mu = \mu_1 + \mu_2 + \dots + \mu_m$ diese erzeugende Funktion. □

Bemerkung: Für poissonverteilte Zufallsgrößen gilt auch die umgekehrte Aussage: Ist die Summe $\Psi = \Psi_1 + \Psi_2$ unabhängiger Zufallsgrößen poissonverteilt, so sind auch die Summanden Ψ_1, Ψ_2 poissonverteilt.

Satz 4.6 Gegeben seien die unabhängigen Poissonschen Punktprozesse $(T_{1,n})_{n \geq 1}, \dots, (T_{m,n})_{n \geq 1}$ mit den Erwartungswertfunktionen $H_1(t), \dots, H_m(t)$. Dann ist die Superposition $(T_{i,n}, i = 1, \dots, m)_{n \geq 1}$ auch ein Poissonscher Punktprozeß mit der Erwartungswertfunktion $H(t) = H_1(t) + \dots + H_m(t)$.

Die Superposition ist in Bild 4.8 dargestellt.

Beweis: Es ist zu zeigen, daß für die Superposition die beiden Eigenschaften der Definition 4.8 erfüllt sind. Eigenschaft (i) folgt aus Folgerung 4.2 und Eigenschaft (ii) folgt aus der Eigenschaft (ii) für jeden der Punktprozesse $(T_{i,n})_{n \geq 1}$. \square

Wir betrachten nun ein in der Zuverlässigkeitstheorie häufig auftretendes Modell – das

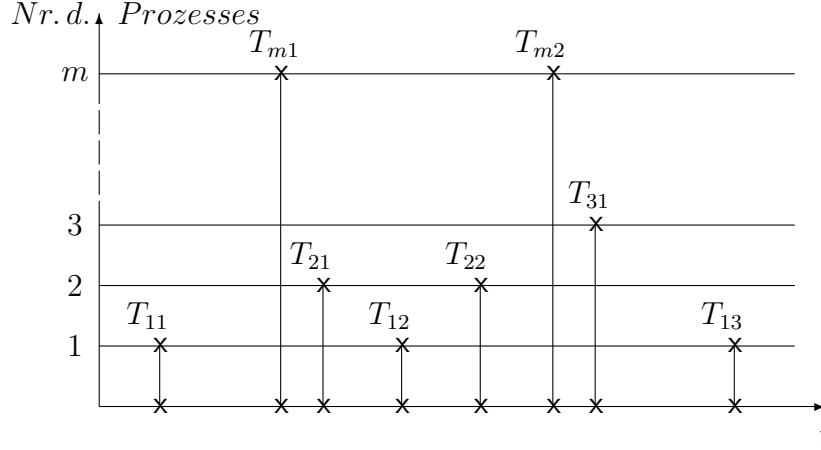


Abbildung 4.8: Die Überlagerung von Erneuerungsprozessen

Verhalten eines Bauteils bei Minimalreparatur.

Das Element besitze die Verteilungsfunktion $F(\cdot)$. Sei S die zufällige Zeit bis zum ersten Ausfall. Zum Zeitpunkt $S = s$ erfolgt ein Ausfall, das Element wird sofort repariert (wobei die Reparaturzeit vernachlässigbar klein ist) und nach der Reparatur ist der Zustand des Elementes der gleiche wie vor dem Ausfall, d.h. mit einer Wahrscheinlichkeit $\bar{F}(t)/\bar{F}(s)$ erfolgt kein Ausfall im Intervall $(s, t]$. Diese Art von Reparatur nennt man auch *Minimalinstandsetzung* und die Ausfälle *Minimalausfälle*. Wenn zur Zeit T_2 wieder ein Ausfall erfolgt, wird wiederum minimal instandgesetzt, usw.

Die Ausfallzeitpunkte des Elementes bilden eine zufällige Punktfolge $(T_n)_{n \geq 1}$.

Satz 4.7 Die Zeit bis zum Ausfall eines Elementes sei eine stetige Zufallsgröße mit der Verteilungsfunktion $F(\cdot)$. Dann ist die Folge $(T_n)_{n \geq 1}$ der Zeitpunkte aller Minimalinstandsetzungen ein Poissonscher Punktprozeß mit der Erwartungswertfunktion $H(t) = -\ln \bar{F}(t)$, d.h. der Erwartungswert der Anzahl von Ausfällen im Intervall $(0, t]$ ist gleich der Hazardfunktion an der Stelle t : $E\Psi(0, t] = H(t)$.

Beweis: Unter der Voraussetzung der Minimalinstandsetzung behält das Element nach einem Ausfall immer sein ursprüngliches Alter. Daher gilt

$$P(\Psi(t, t+s] = 0) = \frac{\bar{F}(t+s)}{\bar{F}(t)}.$$

Im Abschnitt 2.1, Gleichung (2.4) wurde bewiesen, daß für stetige Verteilungen $F(\cdot)$ die Beziehung $\bar{F}(t) = \exp(-H(t))$ gilt. Damit erhält man

$$P(\Psi(t, t+s] = 0) = \frac{\exp(-H(t+s))}{\exp(-H(t))} = \exp(-(H(t+s) - H(t)))$$

und Bedingung (i') der Definition 4.8 des Poissonschen Punktprozesses ist bewiesen. Die Bedingung (ii) folgt sofort aus der Definition der Minimalinstandsetzung. Da das Element sein ursprüngliches Alter behält, hängt ein Ausfall nicht davon ab, ob schon Ausfälle stattgefunden haben oder nicht und damit sind die Anzahlen von Instandsetzungen in disjunkten Intervallen unabhängig voneinander. \square

Als nächstes betrachten wir ein System aus m unabhängigen Elementen. Jedes Element hat eine stetige Verteilungsfunktion $F_j(\cdot)$, $j = 1, \dots, m$. Jedes Element wird bei Ausfall sofort minimalinstandgesetzt.

Satz 4.8 *Die Punktfolge aller Ausfälle in einem System mit Minimalreparaturen ist ein Poissonscher Punktprozeß mit der Erwartungswertfunktion*

$$H_s(t) = \sum_{i=1}^m H_i(t), \quad H_i(t) = -\ln \bar{F}_i(t).$$

Beweis: Aus Satz 4.7 folgt, daß die Folge $(T_{i,n})_{n \geq 1}$ von Ausfallzeitpunkten des i -ten Elementes ein Poissonscher Punktprozeß mit der Erwartungswertfunktion $H_i(t)$ ist. Die Punktprozesse der einzelnen Elemente sind unabhängig voneinander. Daher folgt aus Satz 4.6, daß die Superposition $(T_{i,n}, i = 1, \dots, m)_{n \geq 1}$ dieser Punktprozesse wieder ein Poissonscher Punktprozeß mit der Erwartungswertfunktion $H_s(t) = H_1(t) + \dots + H_m(t)$ ist. \square

Bemerkung: Wenn das Alter des Elementes nach der Reparatur Null ist, so nennt man diese Reparatur Erneuerung. Wenn die Lebensdauerverteilung des Elementes eine Exponentialverteilung ist, dann ist die Ausfallintensität des Elementes immer gleich groß, und die beiden Modelle unterscheiden sich nicht voneinander. Im allgemeinen Fall erhält man jedoch wesentliche Unterschiede.

Nach der Überlagerung oder Superposition von Poissonschen Punktprozessen betrachten wir jetzt die Verdünnung eines solchen Prozesses. Sei $(T_n)_{n \geq 1}$ ein zufälliger Punktprozeß. Jeder Zeitpunkt T_n wird entweder mit der Wahrscheinlichkeit q aus dem Prozeß gestrichen oder verbleibt im Prozeß mit der Wahrscheinlichkeit $p = 1 - q$. Dieses Vorgehen heißt *unabhängige Verdünnung* mit der Wahrscheinlichkeit q .

Wir führen nun eine Folge $\{\Delta_n\}_{n \geq 1}$ ein. Wird T_n gestrichen, so ist $\Delta_n = 1$, ansonsten ist $\Delta_n = 0$. Die Folge $\{\Delta_n\}_{n \geq 1}$ ist eine Folge unabhängiger und identisch verteilter Zufallsgrößen mit $P(\Delta_n = 0) = p$, $P(\Delta_n = 1) = q$, $q + p = 1$. Durch das Streichen von Zeitpunkten erhält man aus der Folge $(T_n)_{n \geq 1}$ eine neue Punktfolge $(T'_n)_{n \geq 1}$. $\Psi(t, t+s]$ und $\Psi'(t, t+s]$ seien wieder die zu diesen Prozessen gehörigen Zählprozesse. Dann gilt

$$\Psi'(t, t+s] = \Psi(t, t+s] - (\Delta_1 + \Delta_2 + \dots + \Delta_{\Psi(t, t+s]}) .$$

Lemma 4.2 Gegeben seien eine poissonverteilte Zufallsgröße Ψ mit dem Parameter μ und eine Folge unabhängiger und identisch verteilter Zufallsgrößen $\{\Delta_n\}_{n \geq 1}$ mit $P(\Delta_n = 1) = q$, $P(\Delta_n = 0) = p$, $p + q = 1$. Dann ist die Zufallsgröße

$$\Psi_p = \begin{cases} \Psi - \Delta_1 - \dots - \Delta_\Psi & \Psi = 1, 2, \dots \\ 0 & \Psi = 0 \end{cases}$$

ebenfalls poissonverteilt mit dem Parameter $p\mu$.

Beweis: Die Summe der Δ_i , $i = 1, \dots, r$ unterliegt einer Binomialverteilung mit den Parametern r und q , d.h.

$$P(\Delta_1 + \dots + \Delta_r = l) = \binom{r}{l} q^l p^{r-l}.$$

Aus der Formel über die totale Wahrscheinlichkeit folgt

$$\begin{aligned} P(\Psi_p = k) &= \sum_{l=0}^{\infty} P\left(\sum_{i=1}^{k+l} \Delta_i = l \mid \Psi = k+l\right) P(\Psi = k+l) = \\ &= \sum_{l=0}^{\infty} \binom{k+l}{l} q^l p^k \frac{\mu^{k+l}}{(k+l)!} e^{-\mu} = \sum_{l=0}^{\infty} \frac{(k+l)!}{l!k!} \frac{q^l p^k \mu^k \mu^l}{(k+l)!} e^{-\mu(p+q)} = \\ &= \sum_{l=0}^{\infty} \frac{(q\mu)^l (p\mu)^k}{l!k!} e^{-\mu p} e^{-\mu q} = \frac{(p\mu)^k}{k!} e^{-p\mu} \sum_{l=0}^{\infty} \frac{(q\mu)^l}{l!} e^{-\mu q} = \\ &= \frac{(p\mu)^k}{k!} e^{-p\mu}. \end{aligned}$$

□

Satz 4.9 Sei $(T_n)_{n \geq 1}$ ein Poissonscher Punktprozeß mit der Erwartungswertfunktion $H(t)$. Nach unabhängiger Verdünnung mit der Wahrscheinlichkeit q erhält man wieder einen Poissonschen Punktprozeß mit der Erwartungswertfunktion $H_p(t) = pH(t)$, $t \geq 0$.

Beweis: Es ist nun offensichtlich, daß die Bedingungen (i') und (ii) der Definition 4.8 eines Poissonschen Punktprozesses erfüllt sind:

(i') folgt aus Lemma 4.2 und (ii) folgt aus der Unabhängigkeit der Anzahl der Erneuerungen in disjunkten Intervallen des Ausgangsprozesses und der unabhängigen Verdünnung.

□

Wir wollen nun den Begriff von Punktprozessen verallgemeinern, indem wir jeden Punkt T_n mit einer Marke Y_n versehen. Die Marken Y_n seien sowohl untereinander als auch vom Prozeß $(T_n)_{n \geq 1}$ unabhängig. Wir erläutern dieses Modell an einem Beispiel:

Beispiel 4.3 Zwischen zwei Städten gibt es eine Gasfernleitung. Man registriert alle Ausfallzeiten $(T_n)_{n \geq 1}$ und die Abstände $(B_n)_{n \geq 1}$ der Ausfallpunkte der Leitung von der Stadt A. $(T_n)_{n \geq 1}$ und $(B_n)_{n \geq 1}$ sind unabhängig. In Abbildung 4.9 ist gut sichtbar, daß damit z.B. die am meisten gefährdeten Abschnitte der Gasfernleitung bestimmt werden können, z.B. ist der Abschnitt (a, b) der Leitung nicht zuverlässig. In diesem Beispiel bezeichnen die Marken den Abstand des Ausfallpunktes zur Stadt A. ◇

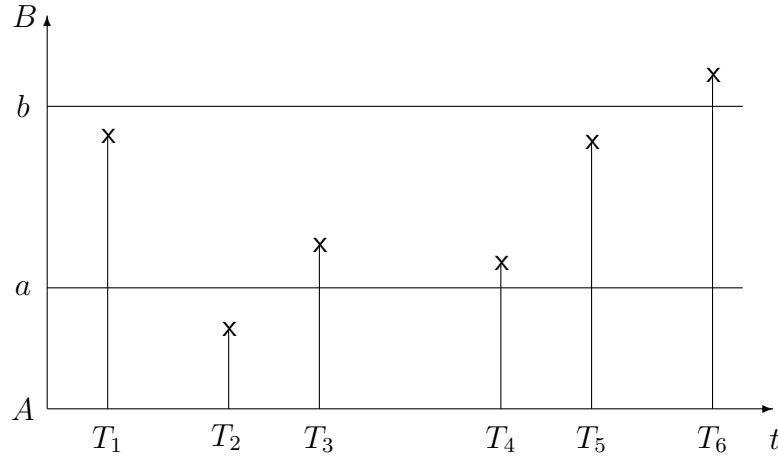


Abbildung 4.9: Ausfallzeiten und Ausfallpunkte einer Gasfernleitung

Als weiteres Beispiel aus der Zuverlässigkeitstheorie kann man annehmen, daß die Punkte T_n Ausfallzeitpunkte darstellen und zu jedem Zeitpunkt T_n die zufällige Reparaturdauer Y_n registriert wird.

Definition 4.10 *Eine zufällige Punktfolge*

$$((T_n, Y_n))_{n \geq 1} = ((T_1, Y_1), \dots, (T_n, Y_n), \dots) \quad \text{mit} \quad 0 < T_n < T_{n+1}$$

heißt markierter Punktprozeß. Jeder zufällige Zeitpunkt T_n besitzt seine zufällige Marke Y_n .

Wenn $Y_n \in \mathbb{R}^1$, besitzt der markierte Punktprozeß $((T_n, Y_n))_{n \geq 1}$ Realisierungen in der Ebene, wenn $Y_n \in \mathbb{R}^2$, besitzt er Realisierungen im \mathbb{R}^3 und wenn $Y_n \in \mathbb{R}^n$, so liegen die Realisierungen im \mathbb{R}^{n+1} .

Somit nennt man eine beliebige (nicht notwendig Poissonsche) zufällige Folge von Punkten $(\underline{X}_n)_{n \geq 1}$, wobei die \underline{X}_n zufällige Punkte im Raum \mathbb{R}^k sind, zufälligen Punktprozeß im \mathbb{R}^k .

Jedem Punktprozeß $(\underline{X}_n)_{n \geq 1}$, $\underline{X}_n \in \mathbb{R}^k$ kann ein Maß zugeordnet werden, indem die Borelsche Menge B das ganzzahlige Maß

$$\Psi(B) = \sum_{n=1}^{\infty} \mathbf{I}(\underline{X}_n \in B) \quad (4.16)$$

erhält.

Im weiteren betrachten wir nur den zweidimensionalen Fall.

Definition 4.11 *Das in (4.16) definierte Maß $\Psi : \mathfrak{B}(\mathbb{R}^2) \rightarrow (0, 1, 2, \dots)$ heißt Zählmaß und der Erwartungswert von $\Psi : \mu(B) = \mathbb{E}\Psi(B)$ heißt Momentmaß erster Ordnung des Punktprozesses \underline{X}_n .*

Verschiedene Arten der Markierung führen zu verschiedenen Punktprozessen. Wir betrachten die einfachsten, jedoch häufig anwendbaren Arten.

Definition 4.12 Man nennt eine zufällige Punktfolge $((T_n, Y_n))_{n \geq 1}$ Poissonschen Punktprozeß mit dem Momentmaß erster Ordnung $\mu(B) = E\Psi(B)$, $B \in \mathfrak{B}(\mathbb{R}^2)$, wenn

(i) für jede Menge B die zufällige Größe

$$\Psi(B) = \sum_{n=1}^{\infty} \mathbf{I}((T_n, Y_n) \in B)$$

poissonverteilt ist

(ii) für alle disjunkten Mengen B_1, B_2, \dots, B_m , $m = 2, 3, \dots$ die Größen $\Psi(B_1), \dots, \Psi(B_m)$ unabhängig voneinander sind.

Bemerkung 1: Es genügt, wenn die Eigenschaft (ii) nur für Rechtecke B_i erfüllt ist. Daraus folgt dann ihre Gültigkeit für beliebige Borelsche Mengen.

Bemerkung 2: Ein Poissonscher Punktprozeß $((T_n, Y_n))_{n \geq 1}$ hat eine charakteristische Eigenschaft, die zur Simulation solcher Prozesse genutzt werden kann. Um sie zu erläutern, betrachten wir eine beliebige Menge B mit $\mu(B) < \infty$. Dann ist die Wahrscheinlichkeit des Ereignisses $\{\Psi(B) = k\}$:

$$P(\Psi(B) = k) = \frac{(\mu(B))^k}{k!} e^{-\mu(B)}.$$

Nun teilen wir B in m disjunkte Mengen B_l , $l = 1, \dots, m$ mit $B_l \cap B_{l'} = \emptyset$, $l \neq l'$ und $B = \bigcup_{l=1}^m B_l$. Dann ist das Momentmaß 1. Ordnung $\mu(B) = \sum_{l=1}^m \mu(B_l)$. Wir betrachten die nun die bedingte Wahrscheinlichkeit

$$\begin{aligned} P(\Psi(B_l) = k_l, l = 1, \dots, m \mid \Psi(B) = k) &= \\ &= \frac{\prod_{l=1}^m \frac{\mu(B_l)^{k_l}}{k_l!} e^{-\mu(B_l)}}{\frac{\mu(B)^k}{k!} e^{-\mu(B)}} = \frac{k!}{k_1! \dots k_m!} \left(\frac{\mu(B_1)}{\mu(B)} \right)^{k_1} \dots \left(\frac{\mu(B_m)}{\mu(B)} \right)^{k_m} \end{aligned}$$

mit $k_1 + \dots + k_m = k$.

Das ist eine Polynomverteilung. Die Simulation erfolgt in zwei Stufen. Zuerst wird die Realisierung k einer poissonverteilten Zufallsgröße simuliert. Dann wird die Menge B in disjunkte Teilmengen zerlegt. Nun kann man k unabhängige „Würfe“ simulieren, wobei $\mu(B')/\mu(B)$ die Wahrscheinlichkeit ist, daß ein auf B geworfener Punkt sich in B' befindet.

Als nächstes zeigen wir, wie man durch eine unabhängige Markierung Poissonsche Punktprozesse konstruieren kann. Wir betrachten einen zufälligen Versuch mit den möglichen Ausgängen A_1, \dots, A_m . Diese Versuchsausgänge treten mit den Wahrscheinlichkeiten p_1, \dots, p_m ein. Der Versuch wird k mal unabhängig voneinander durchgeführt. Bezeichnen wir nun mit Δ_i die zufällige Anzahl des Auftretens des Ereignisses A_i in k Versuchen,

$i = 1, \dots, m$. Der Vektor $\underline{\Delta} = (\Delta_1, \dots, \Delta_m)$ besitzt eine Polynomialverteilung besitzt, d.h.

$$P(\Delta_1 = k_1, \dots, \Delta_m = k_m) = \frac{k!}{k_1! \dots k_m!} p_1^{k_1} \dots p_m^{k_m}$$

mit $k_1 + \dots + k_m = k$, $p_1 + \dots + p_m = 1$.

Vereinigt man zwei der Versuchsausgänge $B = A_i \cup A_{i+1}$, so erhält man wieder eine Polynomialverteilung mit $m - 1$ Versuchsausgängen und den Wahrscheinlichkeiten $(p_1, \dots, p_{i-1}, p_i + p_{i+1}, p_{i+2}, \dots, p_m)$.

Jetzt nehmen wir an, daß auch die Anzahl Δ der Versuche zufällig ist. Dann gilt folgendes Lemma:

Lemma 4.3 *Sei die zufällige Anzahl Δ der Versuche poissonverteilt mit dem Parameter μ . Dann sind die zufälligen Größen $\Delta_1, \dots, \Delta_m$ unabhängig und ebenfalls poissonverteilt mit den Parametern $\mu p_1, \dots, \mu p_m$, d.h.*

$$P(\Delta_1 = k_1, \dots, \Delta_m = k_m) = \prod_{i=1}^m \frac{(\mu p_i)^{k_i}}{k_i!} e^{-\mu p_i}.$$

Beweis: Aus der Formel für bedingte Wahrscheinlichkeiten erhalten wir

$$\begin{aligned} P(\Delta_1 = k_1, \dots, \Delta_m = k_m) &= \\ &= P(\Delta = \Delta_1 + \dots + \Delta_m = k, \Delta_1 = k_1, \dots, \Delta_m = k_m) = \\ &= P(\Delta_1 = k_1, \dots, \Delta_m = k_m \mid \Delta = k) P(\Delta = k) = \frac{k!}{k_1! \dots k_m!} p_1^{k_1} \dots p_m^{k_m} \frac{\mu^k}{k!} e^{-\mu}. \end{aligned}$$

Da $k_1 + \dots + k_m = k$ und $p_1 + \dots + p_m = 1$, erhält man weiter

$$\begin{aligned} P(\Delta_1 = k_1, \dots, \Delta_m = k_m) &= \frac{p_1^{k_1} \dots p_m^{k_m}}{k_1! \dots k_m!} \mu^{(k_1 + \dots + k_m)} e^{-\mu(p_1 + \dots + p_m)} = \\ &= \frac{(\mu p_1)^{k_1}}{k_1!} \dots \frac{(\mu p_m)^{k_m}}{k_m!} e^{-\mu p_1} \dots e^{-\mu p_m} \end{aligned}$$

und damit die obige Aussage. □

Satz 4.10 *Sei $(T_n)_{n \geq 1}$ ein Poissonscher Punktprozeß mit der Erwartungswertfunktion $H(t)$, $t \geq 0$ und $(Y_n)_{n \geq 1}$ eine Folge unabhängiger und identisch verteilter Zufallsgrößen mit der Verteilungsfunktion $P(Y_n \leq y) = G(y)$. Dann ist der zufällige Punktprozeß $((T_n, Y_n))_{n \geq 1}$ ein Poissonscher Punktprozeß auf dem \mathbb{R}^2 mit dem Momentenmaß erster Ordnung*

$$\mu(B) = \iint_{(u,y) \in B} dH(u) dG(y).$$

Beweis: Wir betrachten eine Menge $B = (t, t + s] \times (a, b]$. Sei $\{\Delta = \Psi(t, t + s] = k\}$ das Ereignis, daß k Punkte des Poissonschen Punktprozesses $(T_n)_{n \geq 1}$ im Intervall $(t, t + s]$ liegen. Das Ereignis $(T_i, Y_i) \in (t, t + s] \times (a, b]$ kann nun interpretiert werden als eine zufällige Anzahl Δ von Versuchen mit den beiden Ausgängen $A_1 = (Y_i \in (a, b])$ und

$A_2 = (Y_i \notin (a, b])$. Die Wahrscheinlichkeiten für diese Ausgänge sind $p = P(A_1) = P(Y_i \in (a, b]) = G(b) - G(a) = \int_a^b dG(y)$ und $q = P(A_2) = 1 - p$. Dann folgt aus Lemma 4.3, daß

$$\Psi(B) = \sum_{n=1}^{\infty} I((T_n, Y_n) \in (t, t+s] \times (a, b])$$

eine poissonverteilte Zufallsgröße mit dem Erwartungswert

$$E\Psi(B) = p[H(t+s) - H(t)] = [G(b) - G(a)][H(t+s) - H(t)] = \iint_B dH(u) dG(y)$$

ist. Somit ist die Bedingung (i) in Definition 4.12 erfüllt.

Die Bedingung (ii) wird folgendermaßen bewiesen:

Wie in Abbildung 4.10 dargestellt, werden die disjunkten Rechtecke aufgespalten. Seien $\{(t_i, a_i)\}$ alle Punkte, die Eckpunkte von Rechtecken sind. Ordnet man die t_i und a_i ,

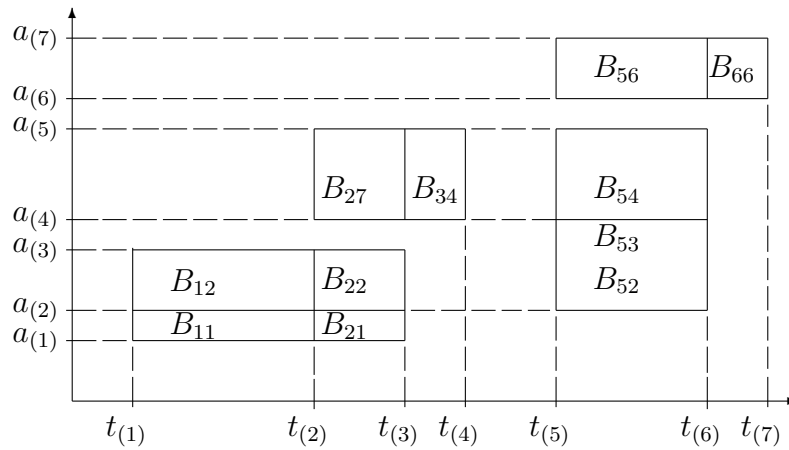


Abbildung 4.10:

so erhält man die beiden Folgen $t_{(1)} < t_{(2)} < \dots < t_{(g)}$ und $a_{(1)} < a_{(2)} < \dots < a_{(h)}$. Mit $B_{ij} = (t_{(i)}, t_{(i+1)}] \times (a_{(j)}, a_{(j+1)}]$ wird das Rechteck bezeichnet, dessen linker unterer Punkt $(t_{(i)}, a_{(j)})$ ist.

Wir betrachten nun alle Rechtecke $B_{ij_1}, \dots, B_{ij_{k_i}}$ innerhalb eines Intervalls $(t_{(i)}, t_{(i+1)}]$. Aus Lemma 4.3 folgt, daß alle $\Psi(B_{ij})$ $j = j_1, \dots, j_{k_i}$ unabhängig und poissonverteilt mit dem Parameter $[H(t_{(i+1)}) - H(t_{(i)})][G(a_{(j+1)}) - G(a_{(j)})]$ sind. Weiterhin folgt aus der Definition des Poissonprozesses $(T_n)_{n \geq 1}$, daß die Größen $\Psi(B_{ij})$ unabhängig von den Größen $\Psi(B_{i'j'})$ für $i' \neq i$ sind, da sie auf disjunkten Intervallen definiert sind. Damit sind alle $\{\Psi(B_{ij})\}$ unabhängige Ereignisse und die Bedingung (ii) ist ebenfalls erfüllt. \square

Beispiel 4.4 Sei $(T_n)_{n \geq 1}$ ein Poissonscher Punktprozeß mit der Erwartungswertfunktion $H(t) = \lambda t$ und $G(y) = \frac{y}{h}$ eine Gleichverteilung auf dem Intervall $(0, h]$. Dann erhält

man auf dem Rechteck $(0, h] \times (0, T]$ einen Poissonschen Punktprozeß mit dem Momentenmaß erster Ordnung $\mu(B) = \lambda \cdot l(B)$, wobei $l(B)$ das Lebesguesche Maß auf \mathbb{R}^2 ist. \diamond

Eine weitere wichtige Eigenschaft Poissonscher Punktprozesse ist ihre Invarianz gegenüber unabhängigen Verschiebungen. Zunächst soll so eine unabhängige Verschiebung erklärt werden. Wir betrachten einen Poissonschen Punktprozeß auf der Ebene. Der Punkt (t, y) wird jeweils in den Punkt $(t + y, y)$ abgebildet. Diese Verschiebung ist in Abbildung 4.11 dargestellt. Sie bewirkt, daß z.B. aus einem Rechteck ein Parallelogramm

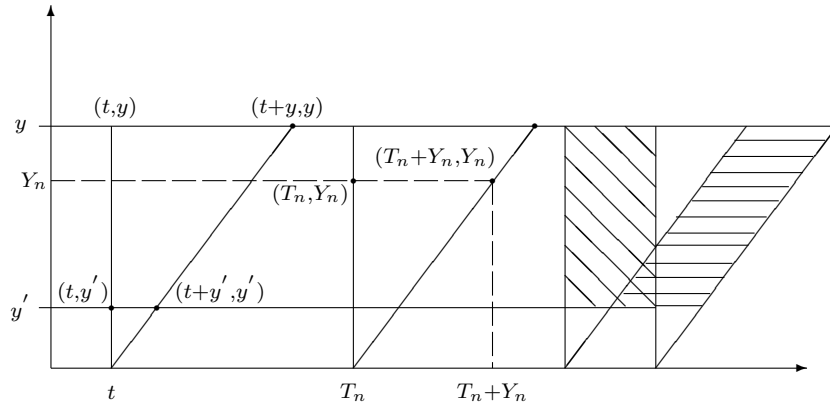


Abbildung 4.11: Eine Transformation markierter Punktprozesse

wird.

Betrachten wir einen Poissonschen Punktprozeß $(T_n)_{n \geq 1}$ und markieren wir diesen Prozeß unabhängig mit $(Y_n)_{n \geq 1}$ mit der Verteilungsfunktion $G(t)$. Das Momentmaß erster Ordnung für diesen markierten Punktprozeß $((T_n, Y_n))_{n \geq 1}$ ist wegen Satz 4.10

$$\mu(B) = \iint_B dH(t) dG(y) .$$

Betrachtet man nun einen Bereich $B = ((u, y) : u + y \leq t, u \geq 0, y \geq 0)$, so wird aus diesem Bereich bei der oben beschriebenen Transformation $B' = ((u', y) : u' \leq t)$. B' ist ein Rechteck, d.h. man wählt B so, daß man nach der Transformation ein Rechteck erhält. Damit erhält man für das Momentmaß erster Ordnung

$$\begin{aligned} \mu(B) &= \iint_{B=((u,y): u+y \leq t)} dH(u) dG(y) = \int_0^t \left(\int_0^{t-u} dG(y) \right) dH(u) = \\ &= \int_0^t G(t-u) dH(u) = G * H(t) = H * G(t) \end{aligned}$$

d.h. man erhält die Faltung der beiden Funktionen $H(t)$ und $G(t)$.

Definition 4.13 Sei $(Y_n)_{n \geq 1}$ eine Folge unabhängiger und identisch verteilter Zufallsgrößen mit $G(y) = P(Y_n \leq y)$. Der Punktprozeß $(T'_n)_{n \geq 1}$ mit

$$T'_n = \min(T_k + Y_k > T'_{n-1}, k = 1, 2, \dots) \quad (4.17)$$

heißt unabhängige Verschiebung des Punktprozesses $(T_n)_{n \geq 1}$ mit der Verteilungsfunktion G .

Die Gleichung (4.17) bedeutet hierbei, daß die Punkte nach ihrer Verschiebung wieder fortlaufend numeriert werden.

Satz 4.11 Die unabhängige Verschiebung $(T'_n)_{n \geq 1}$ eines Poissonschen Punktprozesses $(T_n)_{n \geq 1}$ mit der Erwartungswertfunktion $H(t)$ ist wieder ein Poissonscher Punktprozeß mit einer Erwartungswertfunktion $H'(t) = H * G(t)$.

Beweis: Es sind wieder die Bedingungen (i) und (ii) der Definition des Poissonschen Punktprozesses zu prüfen. Zuerst wird eine unabhängige Markierung

$$(T_n)_{n \geq 1} \rightarrow ((T_n, Y_n))_{n \geq 1}$$

durchgeführt. Aus Satz 4.10 folgt, daß $((T_n, Y_n))_{n \geq 1}$ ein Poissonscher Punktprozeß auf dem \mathbb{R}^2 mit dem Momentmaß erster Ordnung

$$\mu(B) = \iint_{(t,y) \in B} dH(t) dG(y)$$

ist. Die unabhängige Verschiebung des Punktprozesses erhält man, wenn man nun den neuen transformierten Punktprozeß auf die Gerade projiziert (siehe Abbildung 4.12). Betrachten wir als Bereich B ein Parallelogramm mit den Eckpunkten $(t - b, b)$, $(t +$

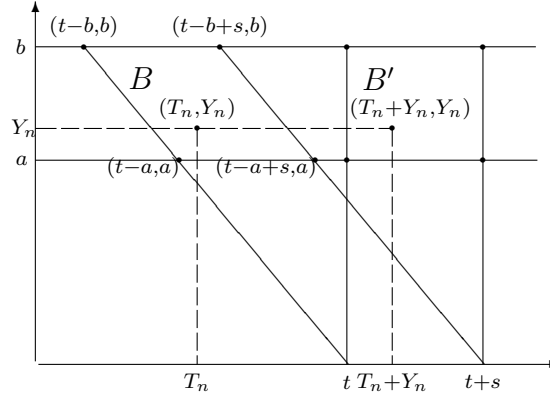


Abbildung 4.12: Eine unabhängige Verschiebung

$s - b, b)$, $(t - a, a)$ und $(t + s - a, a)$. Bei der oben beschriebenen Transformation wird aus diesem Parallelogramm ein Rechteck B' . Die Anzahl der Punkte des Prozesses in

diesem Bereich ändert sich bei der Transformation nicht, d.h. $\Psi(B) = \Psi'(B')$. Da $\Psi(B)$ poissonverteilt ist, ist auch $\Psi'(B')$ poissonverteilt und Bedingung (i') gilt.

Ebenso einfach folgt Bedingung (ii): Die disjunkten Mengen B'_1, \dots, B'_m erhalten wir aus den disjunkten Mengen B_1, \dots, B_m . Da die $\Psi_i(B_i)$ unabhängig sind, sind auch die $\Psi'_i(B'_i)$ unabhängig und (ii) gilt. Damit erhalten wir, daß die Punkte $(T_n + Y_n)_{n \geq 1}$ einen Poissonprozeß ergeben. \square

Für die hier vorgestellten Modelle gibt es sehr viele Anwendungen in der Bedienungs- und Zuverlässigkeitstheorie. Einige werden im folgenden Abschnitt vorgestellt.

4.4 Perkulationsmodelle

In diesem Abschnitt werden Modelle untersucht, bei denen Schocks die Ursache für den Ausfall darstellen. Nehmen wir an, daß ein System der Einwirkung von Schocks ausgesetzt ist. Diese Schocks erfolgen zu zufälligen Zeitpunkten T_n . Wir nehmen an, daß die Folge $(T_n)_{n \geq 1}$ ein Poissonscher Punktprozeß mit der Erwartungswertfunktion $H(t) = \lambda t$ ist. Der n -te Schock habe die Kraft Y_n . Die Folge $(Y_n)_{n \geq 1}$ ist eine Folge von unabhängigen und identisch verteilten Zufallsgrößen mit der Verteilungsfunktion $G(y) = P(Y_n \leq y)$. Damit ist $(T_n, Y_n)_{n \geq 1}$ ein markierter Punktprozeß.

Zuerst soll nun die Wahrscheinlichkeit $F(t)$ dafür bestimmt werden, daß das System trotz der Einwirkung der Schocks im Zeitintervall $(0, t]$ intakt ist. Sei $\Psi(0, t] = \sum_{k=1}^{\infty} I(T_k \leq t)$ die Anzahl der Schocks im Intervall $(0, t]$. Wenn das System nach Einwirkung eines Schocks von der Stärke Y noch intakt ist, so bezeichnen wir dieses Ereignis mit $\{\Delta(Y) = 1\}$. Sei $Q(y)$ die Wahrscheinlichkeit, daß das System nach einem Schock mit der Kraft $Y = y$ intakt ist. Aus der Formel über die totale Wahrscheinlichkeit erhalten wir

$$\begin{aligned} \bar{F}(t) &= \sum_{n=1}^{\infty} P(\Delta(Y_1) = 1, \dots, \Delta(Y_n) = 1 \mid T_n \leq t < T_{n+1}) P(T_n \leq t < T_{n+1}) = \\ &= \sum_{n=1}^{\infty} P(\Delta(Y_1) = 1) \cdots P(\Delta(Y_n) = 1) P(\Psi(0, t] = n) = \\ &= \sum_{n=1}^{\infty} P(\Delta(Y_i) = 1)^n \frac{(\lambda t)^n}{n!} e^{-\lambda t} = \sum_{n=1}^{\infty} \left(\int_0^{\infty} Q(y) dG(y) \right)^n \frac{(\lambda t)^n}{n!} e^{-\lambda t} = \\ &= e^{-\lambda t} \exp \left(\lambda t \int_0^{\infty} Q(y) dG(y) \right) = \exp \left(- \left(1 - \int_0^{\infty} Q(y) dG(y) \right) \lambda t \right). \end{aligned}$$

Damit ist folgender Satz bewiesen:

Satz 4.12 *Die Lebensdauer eines Systems unter Schockeinwirkung ist exponentialverteilt mit dem Parameter $\lambda(1 - \int_0^{\infty} Q(y) dG(y))$.*

Wir betrachten jetzt ein Schockmodell mit Schadensakkumulation, z.B. können durch Stromschwankungen Schäden an Leitungen hervorgerufen werden. Das System ist wieder der Einwirkung von Schocks ausgesetzt, die zu den Zeitpunkten $(T_n)_{n \geq 1}$ auftreten und

eine allmähliche Beschädigung des System hervorrufen. Jede Beschädigung durch einen Schock zur Zeit T_n ist eine skalare zufällige Größe $X_n \geq 0, n = 1, 2, \dots$. Die Beschädigungen akkumulieren sich, d.h. nach dem n -ten Schock ist die Beschädigung des Systems $X_1 + \dots + X_n$. Nehmen wir an es existiert ein solches R , daß das System intakt ist, wenn $X_1 + \dots + X_n \leq R$. Überschreitet das Ausmaß der Beschädigungen die Grenze R , so fällt das System aus. Tritt ein Ausfall auf Grund von Schadensakkumulation ein, so nennt man dieses Ausfallmodell Perkulationsmodell. Bezeichnen wir mit $\nu(t)$ die Nummer des Schocks, der den Ausfall hervorruft, d.h.

$$\nu = \min(n : X_1 + \dots + X_n > R) .$$

Wir berechnen wieder die Wahrscheinlichkeit $\bar{F}(t)$, daß das System während des Intervalls $(0, t]$ ausfallfrei arbeitet. Aus der Formel über die totale Wahrscheinlichkeit erhalten wir

$$\begin{aligned} \bar{F}(t) &= P(T_\nu > t) = P((T_1 > t) \cup (X_1 \leq R, T_1 \leq t < T_2) \cup \dots \\ &\dots \cup (X_1 + \dots + X_n \leq R, T_n \leq t < T_{n+1}) \cup \dots) = \\ &= \sum_{n=0}^{\infty} P(\Psi(0, t] = n) \cdot P(X_1 + \dots + X_n \leq R | T_n \leq t < T_{n+1}) . \end{aligned}$$

Nehmen wir nun zusätzlich an, daß $(T_n)_{n \geq 1}$ ein Poissonscher Punktprozeß mit der Erwartungswertfunktion $H(t) = \lambda t$ ist, daß $(X_n)_{n \geq 1}$ eine Folge unabhängiger und identisch verteilter Zufallsgrößen mit der Verteilungsfunktion $G(s)$ ist und daß die Folge $(X_n)_{n \geq 1}$ unabhängig von der Folge $(T_n)_{n \geq 1}$ ist, so erhalten wir

$$\begin{aligned} \bar{F}(t) &= \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} P(X_1 + \dots + X_n \leq R) = \\ &= \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} \cdot G^{*n}(R) . \end{aligned} \tag{4.18}$$

Hierbei ist $G^{*n}(\cdot)$ die n -fache Faltung der Verteilungsfunktion $G(\cdot)$. Wir betrachten wir ein weiteres Perkulationsmodell. Nehmen wir an, daß ein Werkstoff der Länge L und der Breite B betrachtet wird, d.h. wir beschränken uns hier auf das zweidimensionale Modell. Der Werkstoff kann in bestimmten Punkten (x, y) Defekte besitzen. Diese Defekte sind die Ursprünge von Rissen. Risse wachsen mit der Geschwindigkeit v in y -Richtung nach oben und nach unten, so daß die Grenzen des Risses mit dem Ursprung in (x, y) nach der Zeit t in den Punkten $(x, y - vt)$ und $(x, y + vt)$ liegen. Ein Ausfall liegt vor, wenn der Riß beide Seiten des Werkstoffes erreicht hat, wenn er sich also durch den gesamten Werkstoff zieht (siehe Abbildung 4.13). Modelle dieser Art heißen Perkulationsmodelle.

Wenn der Ursprung des Risses im Punkt (x, y) mit $y \leq \frac{B}{2}$ liegt, so erfolgt der Ausfall zur Zeit $\frac{B-y}{v}$ beim Erreichen der oberen Seite des Werkstoffes, ist $y > \frac{B}{2}$, so erfolgt der Ausfall zur Zeit $\frac{y}{v}$ beim Erreichen der unteren Seite. Wir nehmen an, daß die Ursprünge der Risse einen Poissonschen Punktprozeß im R^2 mit der Intensität λ_0 darstellen. Dann

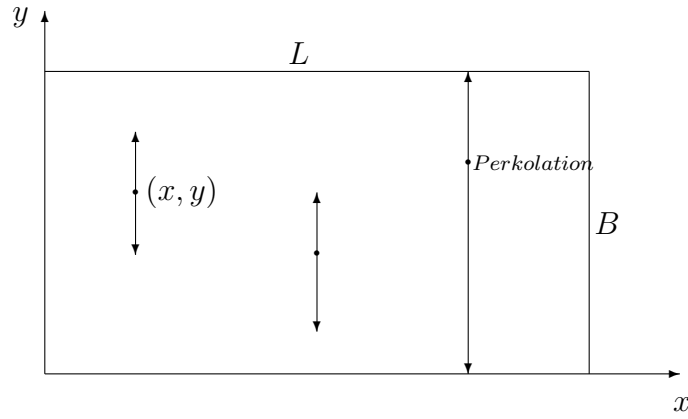


Abbildung 4.13: Rißentwicklung in Werkstoffen

ist die Wahrscheinlichkeit, daß die Anzahl Ψ der Risse gleich k ist:

$$P(\Psi = k) = \frac{(\lambda_0 BL)^k}{k!} e^{-\lambda_0 BL} . \quad (4.19)$$

Wenn $k = 0$ ist, hat man einen Werkstoff ohne Defekte. Diese Wahrscheinlichkeit ist

$$p = P(\Psi = 0) = e^{-\lambda_0 BL} .$$

Wenn $\Psi = k$ ist, dann entspricht die bedingte Verteilung für die Lage der Ursprungspunkte $(x_1, y_1), \dots, (x_k, y_k)$ der Risse einer Stichprobe von unabhängigen und identisch verteilten Größen auf dem Rechteck $[0, L] \times [0, B]$ (siehe Abschnitt 2.3). Damit besitzt die Koordinate Y_i des Punktes (X_i, Y_i) eine gleichmäßige Verteilung mit der Dichte $\frac{1}{B}$. Es folgt, daß die Wahrscheinlichkeit eines ausfallfreien Funktionierens des Werkstoffes während der Zeit s

$$\begin{aligned} \overline{G}(s) &= P\left(\frac{B - Y_i}{v} > s, Y_i \leq \frac{B}{2}\right) + P\left(\frac{Y_i}{v} > s, Y_i > \frac{B}{2}\right) = \\ &= \int_0^{\frac{B}{2}} I\left(\frac{B - y}{v} > s\right) \frac{dy}{B} + \int_{\frac{B}{2}}^B I\left(\frac{y}{v} > s\right) \frac{dy}{B} = \\ &= \int_0^{\frac{B}{2} \wedge (B - sv)} \frac{dy}{B} + \int_{\frac{B}{2} \vee sv}^B \frac{dy}{B} = \begin{cases} 1, & s \leq \frac{B}{2v}, \\ \frac{2(B - sv)}{B}, & \frac{B}{2v} < s \leq \frac{B}{v}, \\ 0, & s > \frac{B}{v} \end{cases} \end{aligned}$$

ist. Die zugehörige Dichtefunktion ist

$$g(s) = \frac{2v}{B} I\left(\frac{B}{2v} \leq s < \frac{B}{v}\right) . \quad (4.20)$$

Haben wir nun k Ursprungspunkte für gewisse Risse (Defekte), so ist die Wahrscheinlichkeit, daß bis zur Zeit s keine Perkolation erfolgt, $(\overline{G}(s))^k$. Aus der Formel über die

totale Wahrscheinlichkeit folgt aus (4.19)

$$\bar{F}_0(s) = \sum_{k=0}^{\infty} (\bar{G}(s))^k \frac{(\lambda_0 LB)^k}{k!} e^{-\lambda_0 LB} = e^{\lambda_0 LB \bar{G}(s)} \cdot e^{-\lambda_0 LB} = e^{-\lambda_0 LB G(s)}$$

und aus (4.20)

$$\begin{aligned} h_0(s) &= \frac{f_0(s)}{\bar{F}_0(s)} = \frac{e^{-\lambda_0 LB G(s)} \cdot \lambda_0 LB g(s)}{e^{-\lambda_0 LB G(s)}} = \\ &= \lambda_0 LB \cdot \frac{2v}{B} \mathbf{I}\left(\frac{B}{2v} \leq s < \frac{B}{v}\right) = 2\lambda_0 Lv \mathbf{I}\left(\frac{B}{2v} \leq s < \frac{B}{v}\right). \end{aligned}$$

Bemerkung 1: Wenn die Geschwindigkeit v der Rißausbreitung nicht konstant ist, sondern zufällig, z.B. $v = v_1$ mit p und $v = v_2$ mit $1 - p$, so erhält man unterschiedliche Verteilungen.

Bemerkung 2: Für alle $s > \frac{B}{v}$ gilt $\bar{F}(s) = e^{-\lambda_0 LB} = \text{const.}$

Ein realistischeres Modell besteht in folgendem. Wir nehmen an, daß neue Defekte nach einem markierten Poissonschen Punktprozeß $(T_n, (X_n, Y_n))_{n \geq 1}$ mit dem Intensitätsparameter λ_1 entstehen. Im Zeitpunkt T_n entsteht ein Defekt im Punkt (X_n, Y_n) , wobei die Lage des Punktes gleichmäßig über $[0, L] \times [0, B]$ verteilt ist. Wenn zur Zeit T_n ein Defekt entstanden ist, so folgt später zur Zeit $T_n + S_n$ eine Perkolation. Die Wahrscheinlichkeit $P(S_n > s) = \bar{G}(s)$ wurde schon berechnet. Alle $(S_n)_{n \geq 1}$ sind unabhängig und identisch verteilt mit der Verteilungsfunktion $G(s)$. Also erhalten wir die im Abschnitt 2.3 beschriebene unabhängige Verschiebung für den Poissonschen Punktprozeß und wenn wir die Perkolationszeitpunkte neu ordnen:

$$T'_1 = \min_{n \geq 1} (T_n + S_n), T'_2 = \min_{n \geq 1} (T_n + S_n : T_n + S_n > T'_1), \dots,$$

dann erhalten wir aufgrund von Satz 4.11 wieder einen Poissonschen Punktprozeß mit der Erwartungswertfunktion $H_1(t) = H * G(t)$, wobei in unserem Fall $H(t) = \lambda_1 t$ ist. Für diesen Fall folgt

$$\begin{aligned} H * G(t) &= \int_0^t \lambda_1(t-s) \frac{2v}{B} \mathbf{I}\left(\frac{B}{2v} \leq s < \frac{B}{v}\right) ds = \int_{\frac{B}{2v}}^{\frac{B}{v} \wedge t} \lambda_1(t-s) \frac{2v}{B} ds = \\ &= \begin{cases} 0, & t < \frac{B}{2v}, \\ \frac{v}{B} \lambda_1(t - B/2v)^2, & \frac{B}{2v} \leq t < \frac{B}{v}, \\ \lambda_1(t - 3B/4v), & t \geq \frac{B}{v}. \end{cases} \end{aligned}$$

Bezeichnen wir für dieses Modell die Ausfallintensität mit $h_1(t)$, so erhält man für diese Größe

$$h_1(t) = \frac{dH * G(t)}{dt} = \begin{cases} 0, & t < \frac{B}{2v}, \\ \lambda_1\left(\frac{2v}{B}t - 1\right), & \frac{B}{2v} \leq t < \frac{B}{v}, \\ \lambda_1, & t \geq \frac{B}{v}. \end{cases}$$

Diese Ausfallrate ist in Abbildung 4.14 dargestellt.

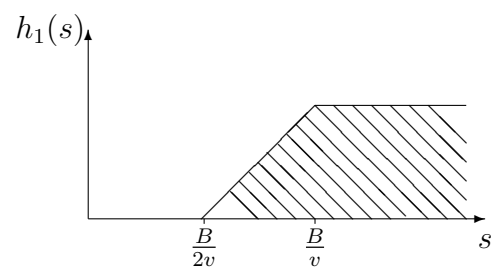


Abbildung 4.14: Die Ausfallintensität eines zweiten Perkolationsmodelles

5 Punktschätzungen für Lebensdauerverteilungen und ihre Parameter bei Zuverlässigkeitsauswertungen

5.1 Datenstrukturen in der Zuverlässigkeitsanalyse

In der Zuverlässigkeitstheorie treten häufig Datenstrukturen auf, die sich von der klassischen vollständigen Stichprobe unterscheiden. Daher ist es nötig, schon bei der Planung eines statistischen Experimentes die entstehende Datenstruktur zu berücksichtigen, um einerseits gute Verfahren auszuwählen und andererseits spezielle statistische Verfahren für diese Datenstrukturen zu entwickeln .

Der Fall der vollständigen Stichprobe ist gut bekannt. In diesem Fall beobachtet man die Daten

$$x = (S_1, \dots, S_n) \quad (5.1)$$

wobei die (S_1, S_2, \dots, S_n) unabhängig und identisch verteilte Zufallsgrößen mit der Verteilungsfunktion $F^*(s), s \geq 0$ sind. Wir interpretieren die $S_i, i = 1, \dots, n$, als Lebensdauern von Elementen. $F^*(\cdot)$ ist unbekannt. Wir nehmen jedoch an, daß man schon einige Informationen über $F^*(\cdot)$ besitzt, z. B. die Information $F^*(\cdot) \in \mathfrak{F}$, wobei \mathfrak{F} eine bekannte Familie von Verteilungsfunktionen ist.

Häufiger treten in der Praxis zensierte Lebensdauerdaten auf.

Beispiel 5.1 Es wird eine Prüfung von n Elementen e_1, \dots, e_n durchgeführt. Das Element e_i hat die zufällige Lebensdauer $S_i, i = 1, \dots, n$. Man kann jedoch bei der Prüfung nicht warten, bis alle Elemente ausgefallen sind; der Versuch wird zur Zeit T abgebrochen. Damit beobachtet man nur die zufälligen Lebensdauern, die kleiner als T sind. Von den restlichen Elementen weiß man nur, daß ihre Lebensdauer größer als T ist. Bei der Durchführung dieses Versuchs erhält man an Stelle von (5.1) die Daten

$$x = (S_1 \wedge T, \dots, S_n \wedge T) \quad \text{mit} \quad a \wedge b = \min(a, b) .$$

Dieser Versuchsplan mit fest vorgegebenem T wird mit $[n, O, T]$ bezeichnet. Das Symbol „O“ bedeutet „ohne Reparatur“, d.h. bei Ausfall ist die Beobachtung des entsprechenden Elementes beendet. Diese Art des Beobachtungsabbruches wird auch mit Typ I - Zensierung (von rechts) bezeichnet. Es ist möglich, den Versuchsplan zu verallgemeinern, indem jedes Element seinen eigenen Zensierungszeitpunkt erhält. Dann erhält man den Plan $[n, O, T_1, \dots, T_n]$. \diamond

Beispiel 5.2 Häufig legt man vor der Durchführung des Versuchs nicht die Zeit des Beobachtungsabbruches, sondern die Anzahl der ausgefallenen Elemente fest. Bezeichnen wir mit $S_{(1,n)} \leq \dots \leq S_{(n,n)}$ die geordneten Ausfallszeitpunkte, so wird bei diesem Plan der Versuch zur zufälligen Zeit $t = S_{(r,n)}$ abgebrochen. Hier erhält man die statistischen Daten

$$x = (S_1 \wedge S_{(r,n)}, \dots, S_n \wedge S_{(r,n)})$$

oder

$$x = (S_{(1,n)}, S_{(2,n)}, \dots, S_{(r,n)}) .$$

Wenn außerdem auch noch bekannt ist, welche Elemente in welcher Reihenfolge ausgefallen sind, d.h. man kennt $i_1, \dots, i_r : S_{i_1} = S_{(1,n)}, \dots, S_{i_r} = S_{(r,n)}$, dann liegen Daten der Form

$$x = (i_1, S_{i_1}, i_2, S_{i_2}, \dots, i_r, S_{i_r})$$

vor.

Erfolgt der Beobachtungsabbruch nach einer vorgegebenen Anzahl von ausgefallenen Elementen, so spricht man von Typ-II-Zensierung oder einem Versuchsplan $[n, O, r]$. Auch dieser Versuchsplan kann verallgemeinert werden:

Zum Zeitpunkt $S_{(1,n)}$ wird bei k_1 Elementen die Beobachtung abgebrochen, zum Zeitpunkt $S_{(2,n)}$ bei weiteren k_2 Elementen und zum Zeitpunkt $S_{(r,n)}$ erfolgt der Beobachtungsabbruch bei den noch verbleibenden k_r Elementen. ($r + k_1 + \dots + k_r = n$). Diesen Plan bezeichnen wir mit $[n, O, k_1 + 1, \dots, k_r + 1]$.

◇

Die Stichprobenpläne $[n, O, T]$ und $[n, O, r]$ aus den Beispielen 5.1 und 5.2 sind in Abbildung 5.1 dargestellt.

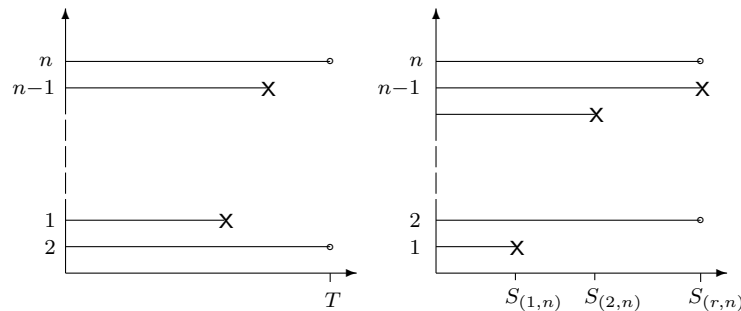


Abbildung 5.1: Die Stichprobenpläne $[n, O, T]$ und $[n, O, r]$

Beispiel 5.3 Wenn die statistischen Daten aus der Beobachtung eines Erneuerungsprozesses gewonnen werden, so nennt man den auftretenden Stichprobenplan $[N, E, T]$ mit T – Abbruchzeit der Beobachtung. Dieser Fall tritt dann auf, wenn Bauteile eines Systems hinsichtlich ihrer Lebensdauer untersucht werden. Fällt ein Bauteil aus, so wird es

im allgemeinen repariert oder ersetzt und das System setzt seine Arbeit fort. Wir setzen hier voraus, daß die Reparatur - bzw. Erneuerungsdauer gegenüber der Zeit zwischen zwei Ausfällen vernachlässigbar klein ist. Wir betrachten hier eine Verallgemeinerung des Stichprobenplanes und lassen zu, daß die Zeit des Beobachtungsabbruches bei jeder Realisierung des Erneuerungsprozesses unterschiedlich sein kann. Diese Verallgemeinerung bringt keine zusätzlichen Schwierigkeiten mit sich, ist jedoch für Anwendungen sehr wesentlich. Da der Zeitparameter oftmals nicht die real ablaufende Zeit, sondern z.B. gefahrene Kilometer oder die tatsächliche Betriebszeit sein kann, erhält man eine Transformation der Zeitachse und damit unterschiedliche Zeiten des Beobachtungsabbruches. In der Statistik für Erneuerungsprozesse unterscheidet man zwei Herangehensweisen. In der ersten Herangehensweise sind die Parameter der Verteilung der Zeit zwischen zwei Ausfällen (d.h. die Parameter der Lebensdauerverteilung) von Interesse. Im zweiten Fall spielt die zugrundeliegende Lebensdauer eine untergeordnete Rolle und man möchte Größen wie z.B. die Erneuerungsfunktion usw. schätzen. Wir wollen uns hier nur mit der Schätzung der Parameter der zugrundeliegenden Lebensdauer befassen. Bezüglich der zweiten Herangehensweise sei auf HÄRTLER verwiesen.

◇

Beispiel 5.4 Es ist auch möglich, Stichprobenpläne zu betrachten, bei denen man noch weniger Informationen aus dem Versuch erhält. Z. B. kann für jedes Element e_i eine Zeit T_i festgelegt werden, zu der das Element überprüft wird. Zu diesem Zeitpunkt wird festgestellt, ob das Element ausgefallen oder intakt ist, d.h. man hat nur die Information darüber, ob der Ausfall des Elementes e_i vor oder nach der Zeit T_i erfolgte. Verwendet man die Bezeichnung $\delta_i = 1$ für das Ereignis $\{S_i \leq T_i\}$ und $\delta_i = 0$ für $\{S_i > T_i\}$, so können die Daten in der Form

$$x = (T_1, \delta_1; \dots; T_n, \delta_n)$$

dargestellt werden.

Diese Art der Information erhält man auch bei „Dosis-Effekt“-Versuchen, auf die im Abschnitt 5.3 im Zusammenhang mit den verallgemeinerten Likelihoodprinzip noch eingegangen wird.

◇

◇ Sei

$\mathfrak{X} = (x)$ der Stichprobenraum. Teilmengen aus diesem Raum $B_i \subset \mathfrak{X}$ sind Ereignisse; im Beispiel 5.4 könnte z.B. das Ereignis $B = \{\delta_i = 0, i = 1, \dots, n\}$ beobachtet werden. Alle Ereignisse zusammen erzeugen die σ -Algebra $\mathfrak{B}(\mathfrak{X})$. Jede Verteilungsfunktion $F(\cdot)$ aus der Familie \mathfrak{F} definiert eine Wahrscheinlichkeitsverteilung auf $(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}))$. Das bedeutet, daß es im Prinzip (manchmal jedoch mit erheblichen Schwierigkeiten) möglich ist, für jedes Ereignis $B \in \mathfrak{B}$ die Wahrscheinlichkeit davon zu berechnen, daß dieses Ereignis als Ergebnis des zufälligen Versuches auftritt. Diese Wahrscheinlichkeit wird mit $P_F(B)$ bezeichnet. Da die Verteilungsfunktion jedoch nicht vollständig bekannt ist, erhält man für jedes B eine Menge von Wahrscheinlichkeiten

$$(P_F(B), F(\cdot) \in \mathfrak{F}) .$$

In dieser Menge liegt die „wahre“ Wahrscheinlichkeit $P_{F^*}(B)$, da nur eine Verteilung $F^*(\cdot) \in \mathfrak{F}$ die „wahre“ Verteilungsfunktion ist. Wir betrachten im folgenden absolut stetigen Verteilungen $F(\cdot) \in \mathfrak{F}^{(ac)}$. In diesem Fall existiert eine Dichte:

$$f(s) = \frac{dF(s)}{ds}.$$

Im folgenden sollen nun für einige Datenstrukturen die Dichte bzw. die Verteilung der Stichprobe herleiten.

Beispiel 5.5 Betrachten wir den Fall einer vollständigen Stichprobe mit den Daten

$$x = (s_1, \dots, s_n) \in \mathbb{R}^n.$$

Wenn die Beobachtungen unabhängig sind, ist die gemeinsame Dichte der Stichprobe das Produkt der Dichten der einzelnen Beobachtungen:

$$f(x) = f(s_1, \dots, s_n) = f(s_1) \cdot \dots \cdot f(s_n).$$

◇

Beispiel 5.6 Für den Stichprobenplan $[n, O, T]$ aus Beispiel 5.1 betrachten wir das Ereignis $B = \{\text{genau } d \text{ Ausfälle der Elemente } i_1, \dots, i_d \text{ fanden in den Intervallen } (s_1, s_1 + ds_1), \dots, (s_d, s_d + ds_d) \text{ statt}\}$.

Dann erhält man die Wahrscheinlichkeit

$$P_F(B) = f(s_1)ds_1 \cdots f(s_d)ds_d (\overline{F}(T))^{n-d}$$

oder die Dichte

$$f_F(s_1, \dots, s_d) = f(s_1) \cdots f(s_d) (\overline{F}(T))^{n-d}. \quad (5.2)$$

Wenn nur die Ausfallszeitpunkte bekannt sind, jedoch nicht, um welche Elemente es sich handelt, so muß man alle Möglichkeiten der Anordnung von d Ausfällen bei n Elementen betrachten, d.h. es werden $\binom{n}{d}$ Dichten der Form (5.2) summiert und man erhält

$$f_F(s_1, \dots, s_d) = \binom{n}{d} f(s_1) \cdots f(s_d) (\overline{F}(T))^{n-d} \quad (5.3)$$

Ebenfalls völlig natürlich ist es, die Dichte für die geordneten Ausfallzeitpunkte

$$s_{(1)} < s_{(2)} < \dots < s_{(d)} \quad (5.4)$$

zu betrachten.

Dann bestehen n Möglichkeiten für den ersten Ausfall, $(n-1)$ für den zweiten usw. und die Dichte der Stichprobe hat die Form

$$f_F(s_{(1)}, \dots, s_{(d)}) = n(n-1) \cdots (n-d+1) f(s_{(1)}, \dots, s_{(d)}) (\overline{F}(T))^{n-d}. \quad (5.5)$$

Die Dichten (5.2) – (5.5) sind ein Beispiel dafür, wie sich das Aussehen der Dichte der Stichprobe bei unterschiedlicher Information ändert. In diesem Beispiel handelt es sich nur um einen konstanten Faktor, der für weitere statistische Untersuchungen keine Rolle spielt. Es treten jedoch Fälle auf, bei denen wesentliche Unterschiede bei unterschiedlicher Information vorliegen.

◇

Beispiel 5.7 Die Grundgesamtheit besitze die Verteilungsfunktion $F_X(t)$ und die Verteilungsdichte $f_X(t)$. Im Falle des Stichprobenplanes $[n, E, T_1, \dots, T_n]$ werden n voneinander unabhängige Betrachtungseinheiten unter Beobachtung genommen, bei jedem Ausfall erfolgt eine sofortige vollständige Erneuerung und die Beobachtung der i -ten Betrachtungseinheit wird zur Zeit T_i abgebrochen ($i = 1, \dots, n$). Als statistische Information liegen von dieser Stichprobe folgende Daten vor:

$N_i^E = N_i^E(T_i)$ – zufällige Anzahl von Ausfällen der i -ten Betrachtungseinheit,

$X_{i1}, \dots, X_{iN_i^E}$ – Ausfallabstände der i -ten Betrachtungseinheit, $i = 1, \dots, n$.

Außerdem verfügt man noch über die Information, daß zwischen letztem Ausfall und Beobachtungsabbruch eine ausfallfreie Arbeit während der Zeit

$$R^E(T_i) = T_i - \sum_{j=1}^{N_i^E} X_{ij}$$

erfolgte.

Es ist ausreichend, die Likelihoodfunktion für eine Realisierung zu bestimmen. Da die n Betrachtungseinheiten unabhängig voneinander sind, ergibt sich die Likelihoodfunktion der Gesamtbeobachtung als Produkt der Likelihoodfunktionen $L_i(\cdot, \cdot)$, $i = 1, \dots, n$ für die Beobachtung des Erneuerungsverlaufes eines Elements.

Nach LIPTSER, SHIRYAEV ist das Wahrscheinlichkeitsmaß eines Erneuerungsprozesses P_1 absolut stetig bezüglich des Wahrscheinlichkeitsmaßes P_0 des standardisierten Poissonprozesses und es gilt

$$\frac{dP_1}{dP_0}(t, x(\cdot)) = \exp \left(\int_0^t \ln \lambda_s(x(\cdot)) dx(s) + \int_0^t (1 - \lambda_s(x(\cdot))) ds \right). \quad (5.6)$$

Im hier betrachteten Fall ist $x(\cdot)$ eine Realisierung des Zählprozesses, d.h. eine Funktion mit folgenden Eigenschaften:

1. $x(0) = 0$
2. $x(t)$ ist stückweise konstant und rechtsseitig stetig,
3. $x(t)$ ist nichtfallend und es gilt $x(t+0) - x(t) = \Delta x(t) \in \{0, 1\}$.

Diese Funktion besitzt Sprünge zu den Zeitpunkten s_j , in denen eine Erneuerung erfolgte:

$$s_{i1} = x_{i1}, s_{i2} = x_{i1} + x_{i2}, \dots, s_{id_i} = x_{i1} + \dots + x_{id_i}$$

mit d_i – Anzahl der Erneuerungen in der i -ten Realisierung und x_{ij} – Realisierungen der Zufallsgrößen X_{ij} . Die Abbruchzeit der Beobachtung ist $t = T_i$ und $\lambda_s(t)$ ist die der Realisierung $x(\cdot)$ entsprechende Ausfallrate zum Zeitpunkt s :

$$\lambda_s(x(\cdot)) := h_X(s - s_{ij}), s_{ij} \leq s < s_{i,j+1}, s_{i0} = 0, j = 0, \dots, d_i$$

mit $h_X(s) = f_X(s)/(\bar{F}_X(s))$ – Ausfallrate der Grundgesamtheit X .
Damit erhält man:

$$\begin{aligned} \int_0^t \ln \lambda_s(x(\cdot)) dx(s)_s &= \sum_{j=1}^{d_i} \ln \frac{f_X(x_{ij})}{\bar{F}_X(x_{ij})} = \\ &= \sum_{j=1}^{d_i} \ln f_X(x_{ij}) - \sum_{j=1}^{d_i} \ln(\bar{F}_X(x_{ij})) \end{aligned} \quad (5.7)$$

und

$$\begin{aligned} \int_0^t (1 - \lambda_s(x(\cdot))) ds &= \sum_{j=1}^{d_i} \int_0^{x_{ij}} (1 - h_X(u)) du + \int_0^{r^E(T_i)} (1 - h_X(u)) du = \\ &= \sum_{j=1}^{d_i} (x_{ij} + \ln(\bar{F}_X(x_{ij}))) + r^E(T_i) + \ln \bar{F}_X(r^E(T_i)) = \\ &= \sum_{j=1}^{d_i} \ln \bar{F}_X(x_{ij}) + \ln \bar{F}_X(r^E(T_i)) + T_i, \end{aligned} \quad (5.8)$$

wobei $r^E(T_i)$ die der Realisierung des Erneuerungsprozesses entsprechende Rückwärtsrekurrenzzeit ist.

Setzt man nun (5.7) und (5.8) in (5.6) ein, so erhält man für die Beobachtung einer Realisierung des Erneuerungsprozesses:

$$\frac{dP_1}{dP_0}(t, x(\cdot)) = \exp \left(\sum_{j=1}^{d_i} \ln f_X(x_{ij}) + \ln \bar{F}_X(r^E(T_i)) + T_i \right), \quad (5.9)$$

wobei der letzte Faktor $\exp(T_i)$ keinen Einfluß auf statistische Untersuchungen besitzt. \diamond

Beispiel 5.8 Wir berechnen die Wahrscheinlichkeit des Auftretens von Daten der Form

$$x = (T_1, \delta_1, \dots, T_n, \delta_n) \quad (5.10)$$

aus dem Beispiel 5.4. Für das Element e_i ist $\delta_i = 1$ mit der Wahrscheinlichkeit $F(T_i)$ und $\delta_i = 0$ mit der Wahrscheinlichkeit $\bar{F}(T_i)$. Zusammengefaßt läßt sich das Beobachtungsergebnis des i -ten Elementes als

$$(F(T_i))^{\delta_i} (\bar{F}(T_i))^{1-\delta_i}$$

aufschreiben. Somit ist die Wahrscheinlichkeit, Beobachtungsdaten der Form (5.10) zu erhalten,

$$P_F(x) = \prod_{i=1}^n (F(T_i))^{\delta_i} (\bar{F}(T_i))^{1-\delta_i}.$$

\diamond

Im einführenden Kapitel, Abschnitt 1 wurde der Begriff der Statistik definiert. Die wichtigste Unterscheidung einer Statistik von anderen Funktionen von Zufallsgrößen besteht darin, daß die Statistik keine unbekannten Parameter enthält. Das soll in den beiden folgenden Beispielen verdeutlicht werden.

Beispiel 5.9 S sei eine exponentialverteilte zufällige Größe, d.h.

$$P(S > s) = e^{-\lambda^* s}, \quad \lambda^* > 0.$$

λ^* ist der unbekannte wahre Parameter. Die Größe $\lambda^* S$ ist exponentialverteilt mit dem Parameter 1:

$$P(\lambda^* S > s) = P\left(S > \frac{s}{\lambda^*}\right) = e^{-\lambda^* \left(\frac{s}{\lambda^*}\right)} = e^{-s}.$$

Damit ist zwar die Verteilung von $\lambda^* S$ unabhängig vom unbekannten Parameter, jedoch die Größe selbst enthält λ^* und ist damit keine Statistik. \diamond

Beispiel 5.10 Wir betrachten eine nach dem Stichprobenplan $[n, O, r]$ erhaltene Stichprobe $x = (S_{(1,n)}, \dots, S_{(r,n)})$. Die Gesamtprüfzeit (total time on test) T_r in diesem Stichprobenplan ist

$$T_r = S_{(1,n)} + \dots + S_{(r-1,n)} + (n - r + 1)S_{(r,n)}.$$

T_r ist eine Statistik: $T_r : \mathfrak{X} \rightarrow \mathbb{R}_+^1$, da alle in der Definition 1.4 geforderten Eigenschaften erfüllt sind. Wenn wir annehmen, daß die Zeit bis zum Ausfall exponentialverteilt ist, kann die Verteilung von T_r bestimmt werden. Im Zeitintervall $(S_{(i,n)}, S_{(i+1,n)})$, $i = 0, \dots, n-1$, $S_{(0,n)} = 0$, werden jeweils $n - i$ Elemente geprüft. Daher gilt

$$T_r = nS_{(1,n)} + (n-1)(S_{(2,n)} - S_{(1,n)}) + \dots + (n-r+1)(S_{(r,n)} - S_{(r-1,n)}).$$

Aus den Eigenschaften der Exponentialverteilung folgt, daß der Ausfall eines von i Elementen mit der Wahrscheinlichkeit $(e^{-\lambda u})^i$ erfolgt. Dann erhält man

$$\begin{aligned} P_\lambda(nS_{(1,n)} > u_1) &= P_\lambda\left(S_{(1,n)} > \frac{u_1}{n}\right) = \left(e^{-\lambda \frac{u_1}{n}}\right)^n = e^{-\lambda u_1}, \\ P_\lambda((n-1)(S_{(2,n)} - S_{(1,n)}) > u_2) &= P_\lambda\left((S_{(2,n)} - S_{(1,n)}) > \frac{u_2}{n-1}\right) = \\ &= \left(e^{-\lambda \frac{u_2}{n-1}}\right)^{n-1} = e^{-\lambda u_2}. \end{aligned}$$

usw. Daher kann T_r als Summe von n unabhängigen und identisch exponentialverteilten zufälligen Größen betrachtet werden. Man kann zeigen, daß

$$P_\lambda(T_r > t) = \sum_{d=0}^{r-1} \frac{(\lambda t)^d}{d!} e^{-\lambda t},$$

d. h. die Statistik unterliegt einer Erlangverteilung der Stufe r mit dem Parameter λ . Wir bemerken, daß in der Verteilung der Statistik T_r wieder der „wahre“ Parameter

λ auftritt, daher wurde die Bezeichnung P_λ gewählt. Betrachten wir die Verteilung der zufälligen Größe λT_r (das ist keine Statistik), so erhalten wir

$$P_\lambda(\lambda T_r > t) = P_\lambda(T_r > \frac{t}{\lambda}) = \sum_{d=0}^r \frac{t^d}{d!} e^{-t}.$$

◇

Der wahre Parameter θ^* bzw. die wahre Verteilung P_{θ^*} , aus der die Daten stammen, ist unbekannt. Wir wollen nun aus den Daten den Parameter θ^* schätzen, d.h. wir wollen eine Statistik $T = T(x)$ so finden, daß der Abstand zwischen θ^* und T möglichst klein wird. Eine Mindestforderung an T besteht darin, daß alle Realisierungen von T in Θ liegen.

Definition 5.1 Eine Statistik $T : \mathfrak{X} \rightarrow \Theta$, die den Stichprobenraum in den Parameterraum abbildet, heißt Punktschätzung $\hat{\theta} = T(x)$, wobei x die Stichprobe ist.

Natürlich sind nicht alle Statistiken dieser Art sinnvolle Punktschätzungen. Liegen z.B. Daten aus der vollständigen Beobachtung von exponentialverteilten Lebensdauern vor, d.h. $X = (S_1, \dots, S_n)$, so ist $T_n = S_1 + \dots + S_n$ zwar eine Punktschätzung im obigen Sinne für den unbekannten Parameter λ^* , jedoch keine sinnvolle. Es ist gut bekannt, daß $1/T_n$ eine bessere Punktschätzung für λ^* ist und die korrigierte Größe

$$\hat{\lambda}_n = \frac{n-1}{T_n} = \frac{n-1}{S_1 + \dots + S_n}$$

eine erwartungstreue Punktschätzung ist. Aus diesen Ausführungen ist ersichtlich, daß die Definition einer Punktschätzung sehr allgemein ist und auch praktisch nicht sinnvolle Punktschätzungen umfaßt. Es gibt jedoch noch Fälle von Punktschätzungen, die durch diese Definition nicht erfaßt werden. Das zeigt das folgende Beispiel.

Beispiel 5.11 Wir betrachten den Fall der klassischen Stichprobe, d.h. es liegen Daten $X = (S_1, \dots, S_n)$ vor. Alle S_i ($i = 1, \dots, n$) sind unabhängig und identisch verteilt mit der Verteilungsfunktion $F(s) \in \mathfrak{F}$. Für die „wahre“ Verteilungsfunktion $F^*(s)$ soll eine Punktschätzung ermittelt werden. Der Parameterraum Θ ist in diesem Fall die Menge aller Funktionen \mathfrak{F} .

Bekanntlich ist die empirische Verteilungsfunktion eine gute Punktschätzung für die unbekannte Verteilung, d.h. wir betrachten

$$\hat{F}_n(s) = \frac{1}{n} \sum_{i=1}^n \mathbf{I}(S_i \leq s) \quad (5.11)$$

als Punktschätzung für $F^*(\cdot)$. $\hat{F}_n(\cdot)$ bildet \mathfrak{X} in \mathfrak{F} ab. Sie ist damit eine Punktschätzung im Sinne der obigen Definition, wenn \mathfrak{F} alle Verteilungsfunktionen, z.B. auch Sprungfunktionen enthält. Weiß man jedoch, daß die wahre Verteilungsfunktion stetig ist, so erhält man keine Punktschätzung im Sinne der Definition, obwohl auch eine stetige Verteilung

durch die Schätzung (5.11) beliebig genau angenähert werden kann. Um das zu beheben, muß man statt des Parameterraumes Θ die Vervollständigung $\overline{\Theta}$ betrachten. Dazu wird eine Metrik $d(\theta_1, \theta_2)$ für alle Paare $(\theta_1, \theta_2) \in \Theta$ definiert und $\overline{\Theta}$ enthält neben Θ alle die Parameterwerte, deren Abstand zu $\theta \in \Theta$ beliebig klein werden kann:

$$\overline{\Theta} = \{\bar{\theta} : \exists (\theta_n)_{n \geq 1} \quad d(\bar{\theta}, \theta_n) \rightarrow 0, \quad n \rightarrow \infty, \quad \theta_n \in \Theta\}.$$

◇

Aufgrund des Beispiels modifizieren wir die Definition 5.2:

Definition 5.2 Wenn eine geeignete Metrik d auf Θ definiert ist und eine dieser Metrik entsprechende Vervollständigung $\overline{\Theta}$ des Parameterraumes gegeben ist, so heißt eine Statistik $T : \mathfrak{X} \rightarrow \overline{\Theta}$ Punktschätzung.

Im folgenden verstehen wir unter einer Punktschätzung immer eine Punktschätzung im Sinne von Definition 5.2.

5.2 Parameterschätzungen für Lebensdauerverteilungen – das Maximum–Likelihood–Prinzip

Wir betrachten ein statistisches Modell

$$\left(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}), \mathfrak{P} = (P_\theta, \theta \in \Theta) \right).$$

Zur Bestimmung des unbekannten wahren Parameters θ aus der Menge Θ wird wegen ihrer guten asymptotischen Eigenschaften (Konsistenz, asymptotische Erwartungstreue und asymptotisch minimale Varianz) häufig die Maximum–Likelihood–Methode verwendet. Dazu benötigt man die Dichtefunktion der Beobachtung, die wir hier allgemein definieren wollen.

Zunächst betrachten wir eine dominierte Familie von Verteilungen $\mathfrak{P} = (P_\theta, \theta \in \Theta)$, für die die Dichte $p(x, \theta) = \frac{dP_\theta}{d\mu}(x)$ bezüglich eines Maßes μ auf $\mathfrak{B}(\mathfrak{X})$ existiert.

Durch die Dichtefunktion $p(\cdot, \cdot)$ ist eine Oberfläche definiert (siehe Abbildung 5.2). Wenn wir auf dieser Oberfläche θ fixieren und die Schnittlinie $p(\cdot, \theta)$ betrachten, erhalten wir eine Dichte aus obiger Familie. Fixiert man dagegen x und betrachtet diese Schnittlinie, so erhält man eine andere Funktion, die wir mit $L(\cdot, x)$ bezeichnen wollen. Natürlich gilt für jeweils fixierte x und θ

$$L(\theta, x) = p(x, \theta).$$

Die andere Bezeichnung ist nur gewählt, um die Abhängigkeit vom unbekannten Parameter zu betonen.

Definition 5.3 Die Funktion $L(\cdot, x)$ heißt Likelihoodfunktion.

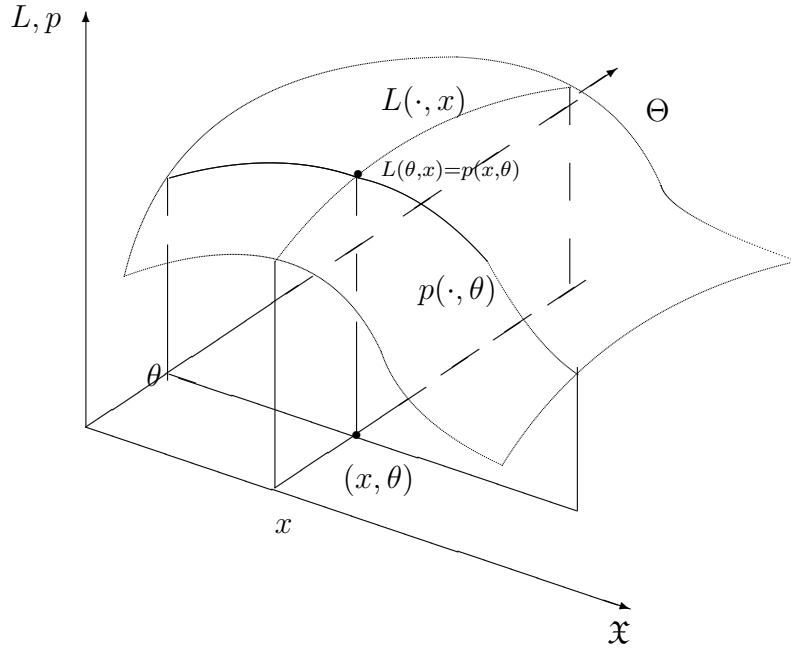


Abbildung 5.2: Die Dichte $p(\cdot, \theta)$ und die Likelihoodfunktion $L(\cdot, x)$

Beispiel 5.12 Sei $x = (s_{(1)}, \dots, s_{(r)})$ eine Stichprobe, die aufgrund des Stichprobenplanes $[n, O, r]$ erhalten wurde.

Die Lebensdauer der Grundgesamtheit sei Weibullverteilt mit der Überlebensfunktion und der Dichte

$$\bar{F}(s) = \exp(-(s/\alpha)^\beta), \quad f(s) = \frac{\beta}{\alpha} \left(\frac{s}{\alpha}\right)^{\beta-1} \exp(-(s/\alpha)^\beta).$$

Dann besitzen die statistischen Daten auf dem Konus

$$K_r = ((s_{(1)}, \dots, s_{(r)}) : 0 < s_{(1)} < \dots < s_{(r)})$$

bezüglich des Lebesgue-Maßes $d\mu = ds_{(1)} \cdots ds_{(r)}$ die Dichte

$$p(x, \theta) = n^{[r]} \left(\prod_{i=1}^r \left(\frac{\beta}{\alpha}\right) \left(\frac{s_{(i)}}{\alpha}\right)^{(\beta-1)} \exp(-(s_{(i)}/\alpha)^\beta) \right) \exp(-(n-r)(s_{(r)}/\alpha)^\beta)$$

mit $\theta = (\alpha, \beta)$, $\alpha > 0, \beta > 0, \Theta = \mathbb{R}_+^2$.

Hierbei bezeichnen wir mit $n^{[r]}$ das Produkt $n^{[r]} = n(n-1) \cdots (n-r+1)$. Für die Likelihoodfunktion erhalten wir somit

$$L(\theta, x) = n^{[r]} \beta^r \left(\prod_{i=1}^r s_{(i)}^{\beta-1} \right) \alpha^{-r\beta} \exp \left(-\alpha^{-\beta} \left(\sum_{i=1}^r s_{(i)}^\beta + (n-r)s_{(r)}^\beta \right) \right).$$

Die Likelihoodfunktion ist abhängig von x und für $x_1 \neq x_2$ erhält man im allgemeinen auch $L(\theta, x_1) \neq L(\theta, x_2)$. In Abbildung 5.3 ist die Likelihoodfunktion der Weibullverteilung für zwei verschiedene Stichproben dargestellt. Hierfür wurden jeweils $n = 50$

Realisierungen einer weibullverteilten Zufallsgröße mit den Parametern $\alpha = a = 1000$ und $\beta = b = 2$ simuliert. Abbildung 5.4 enthält die Niveaulinien der Likelihoodfunktion für diese Stichproben. Für das Modell der Weibullverteilung kann bewiesen werden, daß

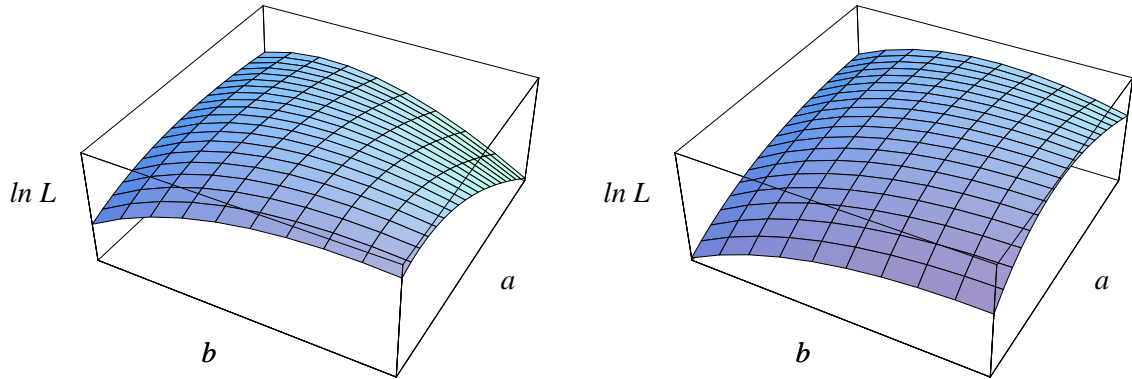


Abbildung 5.3: Die Likelihoodfunktion für zwei Stichproben x_1 und x_2

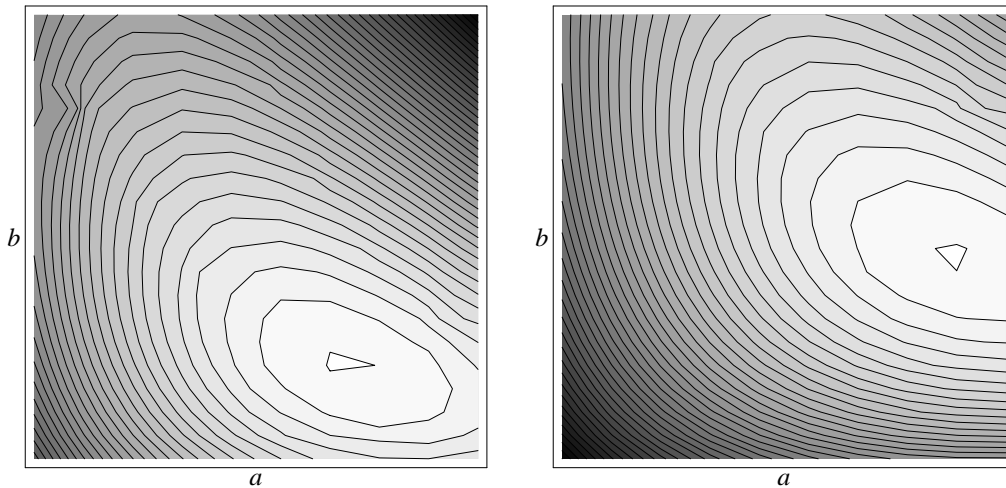


Abbildung 5.4: Niveaulinien der Likelihoodfunktion für zwei Stichproben x_1 und x_2

eine eindeutige Zuordnung zwischen x und $L(\cdot, x)$ besteht. \diamond

Es gibt jedoch auch Beispiele, in denen keine eindeutige Zuordnung zwischen Stichprobe und Likelihoodfunktion besteht.

Beispiel 5.13 Wir betrachten wieder den Stichprobenplan $[n, O, r]$, jedoch die Familie der zweiparametrischen Exponentialverteilungen

$$\mathfrak{F}_{E,2} = (F(\cdot) : \bar{F}(s) = \exp(-\lambda(s - \mu)), s > 0, \theta = (\lambda, \mu) \in \mathbb{R}_+^2) .$$

Für die Dichte der Stichprobe erhalten wir in diesem Fall

$$\begin{aligned} p(x, \theta) &= n^{[r]} \mathbf{I}(s_{(1)} \geq \mu) \prod_{i=1}^r \lambda \exp(-\lambda(s_{(i)} - \mu)) \exp(-\lambda(n-r)(s_{(r)} - \mu)) = \\ &= n^{[r]} \lambda^r \mathbf{I}(s_{(1)} \geq \mu) \exp \left(-\lambda \left(\sum_{i=1}^{r-1} s_{(i)} + (n-r+1)s_{(r)} \right) + n\lambda\mu \right). \end{aligned}$$

Betrachten wir nun zwei Stichproben $x' = (s'_{(1)}, \dots, s'_{(r)})$ und $x'' = (s''_{(1)}, \dots, s''_{(r)})$. Dann ist $p(x', \theta) \equiv p(x'', \theta)$ für beliebige $\theta \in \Theta$ genau dann, wenn

$$s'_{(1)} = s''_{(1)} \quad \text{und} \quad \sum_{i=1}^{r-1} s'_{(i)} + (n-r+1)s'_{(r)} = \sum_{i=1}^{r-1} s''_{(i)} + (n-r+1)s''_{(r)}.$$

Die Größe $\sum_{i=1}^{r-1} s_{(i)} + (n-r+1)s_{(r)}$ ist die Gesamtprüfzeit (total time on test). Wenn statistische Schlußfolgerungen auf der Grundlage der Likelihoodfunktion getroffen werden, braucht man in diesem Beispiel also nur noch die Zeit des ersten Ausfalles und die Gesamtprüfzeit zu kennen. \diamond

Bemerkung 1: Für den Stichprobenplan $[n, O, r]$ und einige andere Pläne ist es manchmal günstig, in der Likelihoodfunktion statt der Dichten die Ausfallraten zu verwenden. Mit $h(s, \theta) = f(s, \theta) / \bar{F}_\theta(s)$ erhält man für den Plan $[n, O, r]$

$$L(\theta, x) = p(x, \theta) = n^{[r]} \prod_{i=1}^r h(s_{(i)}, \theta) \prod_{i=1}^r \bar{F}_\theta(s_{(i)}) (\bar{F}_\theta(s_{(r)}))^{n-r}.$$

Bemerkung 2: Aus analytischen Gründen verwendet man häufig statt $L(\theta, x)$ die logarithmierte Likelihoodfunktion

$$l(\theta, x) := \ln L(\theta, x) \quad \forall \theta \in \Theta.$$

Bemerkung 3: $L(\cdot, x)$ bzw. $l(\cdot, x)$ können als Statistik angesehen werden. Betrachten wir die Menge aller möglichen x -Schnitte $\mathfrak{L} = (L(\cdot, x), x \in \mathfrak{X})$. Weiter sei $\mathfrak{B}(\mathfrak{L})$ die σ -Algebra, die man aus den Mengen

$$B_{(a, \theta)} = (L(\cdot, x) : L(\theta, x) \leq a) \quad \forall a \in \mathbb{R}^1, \theta \in \Theta$$

erhält. Dann ist die Likelihoodfunktion eine Abbildung $L(\cdot, \cdot) : \mathfrak{X} = (x) \rightarrow \mathfrak{L}$ und diese Abbildung ist meßbar.

Da $L(\cdot, x)$ als Funktion über alle $\theta \in \Theta$ betrachtet wird und damit nicht vom unbekannten Parameter θ^* abhängt, erfüllt sie die Forderungen an eine Statistik.

Definition 5.4 Eine Punktschätzung $\check{\theta} : \mathfrak{X} \rightarrow \Theta$ heißt Maximum-Likelihood-Schätzung des wahren unbekannten Parameters θ^* , wenn

$$L(\check{\theta}, x) = \max_{\theta \in \Theta} L(\theta, x). \quad (5.12)$$

Wir werden für beliebige Punktschätzungen das Symbol $\hat{\theta}$ und für Maximum-Likelihood-Schätzungen das Symbol $\check{\theta}$ verwenden.

Maximum-Likelihood-Schätzungen (5.12) besitzen unter bestimmten Voraussetzungen sehr gute Eigenschaften (siehe z.B. LEHMANN). Man kann jedoch auch Beispiele finden, in denen $\check{\theta}$ nicht existiert, nicht eindeutig ist bzw. keine guten Eigenschaften besitzt.

Beispiel 5.14 Die Grundgesamtheit sei auf dem Intervall $[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$ gleichmäßig verteilt, d.h. $F_{\theta}(x) = x - \theta + \frac{1}{2}, x \in [\theta - \frac{1}{2}, \theta + \frac{1}{2}]$. Dann erhält man z.B. für klassische Stichproben $[n, O, n]$

$$L(\theta, x) = 1, \quad s_i \in [\theta - \frac{1}{2}, \theta + \frac{1}{2}].$$

Die Maximum-Likelihood-Schätzung ist in diesem Fall nicht eindeutig. Alle $\check{\theta} \in [s_{(n)} - \frac{1}{2}, s_{(1)} + \frac{1}{2}]$ sind Maximum-Likelihood-Schätzungen. \diamond

Bemerkung 1: Soll eine Funktion $g(\theta) \in \mathbb{R}^1$ des unbekannten wahren Parameters θ mit der Maximum-Likelihood-Methode geschätzt werden, so ist die Schätzung der Funktion gleich der Funktion der Schätzung:

$$\check{g}(\theta) = g(\check{\theta}).$$

Bemerkung 2: Häufig kann man das Maximum der Likelihoodfunktion mittels der partiellen Ableitungen ermitteln :

$$\frac{\partial \ln L(\check{\theta}, x)}{\partial \theta_i} = 0, \quad i = 1, \dots, m.$$

Beispiel 5.15 Wir betrachten wieder die Stichprobe $x = (s_{(1)}, \dots, s_{(r)})$ für den Stichprobenplan $[n, O, r]$. Die Grundgesamtheit unterliege einer Exponentialverteilung: $\bar{F}(s) = e^{-\lambda s}$. Dann erhält man für die Likelihoodfunktion

$$L(\lambda, x) = n^{[r]} \cdot \lambda^r \cdot \exp \left(-\lambda \left(\sum_{i=1}^{r-1} s_{(i)} + (n - r + 1)s_{(r)} \right) \right).$$

Die logarithmierte Likelihoodfunktion hat die Form

$$l(\lambda, x) = \ln(n^{[r]}) + r \ln(\lambda) - \lambda \left(\sum_{i=1}^{r-1} s_{(i)} + (n - r + 1)s_{(r)} \right).$$

Das Maximum dieser Funktion kann aus der Likelihoodgleichung

$$\frac{\partial l(\lambda, x)}{\partial \lambda} = \frac{r}{\lambda} - \left(\sum_{i=1}^{r-1} s_{(i)} + (n - r + 1)s_{(r)} \right) = 0$$

ermittelt werden. Man erhält dann

$$\check{\lambda} = \frac{r}{\sum_{i=1}^{r-1} s_{(i)} + (n - r + 1)s_{(r)}}.$$

\diamond

Die Punktschätzung ist nicht erwartungstreu. Sie läßt sich jedoch explizit berechnen. Damit ist auch der Erwartungswert berechenbar und gegebenenfalls kann die Schätzung auf Erwartungstreue korrigiert werden. Wesentlich häufiger liegt der Fall vor, daß die Maximum-Likelihood-Gleichung zur Bestimmung von $\tilde{\theta}$ nichtlinear ist.

Beispiel 5.16 Die statistischen Daten werden wieder nach dem Stichprobenplan $[n, O, r]$ gewonnen, wir nehmen jedoch an, daß die Grundgesamtheit einer Weibullverteilung mit $\underline{\theta} = (\alpha, \beta)$ unterliegt.

Die logarithmierte Likelihoodfunktion ist in diesem Fall

$$l(\theta, x) = \ln n^{[r]} + r \ln \beta + (\beta - 1) \sum_{i=1}^r \ln s_{(i)} - r \beta \ln \alpha - \alpha^{-\beta} \left(\sum_{i=1}^r s_{(i)}^{\beta} + (n - r) s_{(r)}^{\beta} \right).$$

Daraus erhält man die beiden Maximum-Likelihood-Gleichungen

$$\begin{aligned} \frac{\partial l(\theta, x)}{\partial \alpha} &= -\frac{r\beta}{\alpha} + \left(\sum_{i=1}^r s_{(i)}^{\beta} + (n - r) s_{(r)}^{\beta} \right) \beta \alpha^{-\beta-1} = 0, \\ \frac{\partial l(\theta, x)}{\partial \beta} &= \frac{r}{\beta} + \sum_{i=1}^r \ln s_{(i)} - r \ln \alpha - \alpha^{-\beta} \left(\sum_{i=1}^r s_{(i)}^{\beta} \ln s_{(i)} + \right. \\ &\quad \left. + (n - r) s_{(r)}^{\beta} \ln s_{(r)} \right) + \alpha^{-\beta} \ln \alpha \left(\sum_{i=1}^r s_{(i)}^{\beta} + (n - r) s_{(r)}^{\beta} \right) = 0. \end{aligned}$$

Die erste dieser Gleichungen läßt sich nach α auflösen und liefert die Maximum-Likelihood-Schätzung für α in Abhängigkeit von β :

$$\tilde{\alpha} = \tilde{\alpha}(\beta) = \left(\frac{1}{r} \left(\sum_{i=1}^r s_{(i)}^{\beta} + (n - r) s_{(r)}^{\beta} \right) \right)^{\frac{1}{\beta}}. \quad (5.13)$$

Diese Beziehung für α kann in die zweite Gleichung eingesetzt werden und man erhält eine nichtlineare Gleichung zur Bestimmung von $\tilde{\beta}$:

$$\frac{\sum_{i=1}^r s_{(i)}^{\beta} \ln s_{(i)} + (n - r) s_{(r)}^{\beta} \ln s_{(r)}}{\sum_{i=1}^r s_{(i)}^{\beta} + (n - r) s_{(r)}^{\beta}} - \frac{1}{\beta} = \frac{1}{r} \sum_{i=1}^r \ln s_{(i)}.$$

Man kann beweisen, daß diese Gleichung eine eindeutige Lösung besitzt. \diamond

Beispiel 5.17 Wir betrachten eine Weibullverteilte Grundgesamtheit, deren Verteilungsfunktion und Verteilungsdichte folgende Form besitzen:

$$F_X(t) = 1 - e^{-(\frac{t}{\alpha})^{\beta}}, \quad f_X(t) = \beta \alpha^{-\beta} t^{\beta-1} e^{-(\frac{t}{\alpha})^{\beta}}, \quad t > 0.$$

Desweiteren betrachten wir den Stichprobenplan $[N, E, T_i]$. Die Likelihoodfunktion ist für diesen Stichprobenplan gemäß (5.9):

$$L(\alpha, \beta, x) = \prod_{i=1}^n \left(\prod_{j=1}^{d_i} \frac{\beta}{\alpha} \left(\frac{x_{ij}}{\alpha} \right)^{\beta-1} \exp \left(-\left(\frac{x_{ij}}{\alpha} \right)^{\beta} \right) \exp \left(-\left(\frac{r^E(T_i)}{\alpha} \right)^{\beta} \right) \right). \quad (5.14)$$

Zur Ermittlung von Punktschätzungen der beiden unbekannten Parameter nach der Maximum-Likelihood-Methode wird die Likelihoodfunktion (5.14) logarithmiert, nach den unbekannten Parametern differenziert, die Ableitungen werden Null gesetzt und das entstehende Gleichungssystem wird gelöst. Die Lösung hat die Form:

$$\frac{1}{\check{\beta}} + \frac{\sum_{i=1}^n \sum_{j=1}^{d_i} \ln x_{ij}}{\sum_{i=1}^n d_i} - \frac{\sum_{i=1}^n r^E(T_i)^{\check{\beta}} \ln r^E(T_i) + \sum_{i=1}^n \sum_{j=1}^{d_i} x_{ij}^{\check{\beta}} \ln x_{ij}}{\sum_{i=1}^n r^E(T_i)^{\check{\beta}} + \sum_{i=1}^n \sum_{j=1}^{d_i} x_{ij}^{\check{\beta}}} = 0, \quad (5.15)$$

$$\check{\alpha}^{\check{\beta}} = \frac{\sum_{i=1}^n r^E(T_i)^{\check{\beta}} + \sum_{i=1}^n \sum_{j=1}^{d_i} x_{ij}^{\check{\beta}}}{\sum_{i=1}^n d_i}. \quad (5.16)$$

Die Lösung von (5.15) erfordert die Verwendung numerischer Methoden, während bei bekanntem β die Schätzung für α explizit ermittelt werden kann.

◇

Diese Beispiele zeigen, daß die Lösung der Maximum-Likelihood-Gleichungen Schwierigkeiten bereiten kann. Bei zensierten Beobachtungen sind die Maximum-Likelihood-Gleichungen häufig nicht explizit lösbar, insbesondere wenn $\dim \Theta = m > 1$ ist. Eine Möglichkeit zur numerischen Bestimmung der Maximum-Likelihood-Schätzungen ist die gut bekannte Newton-Raphson-Methode. Bei dieser Methode wird mit Hilfe der Linearisierung der Likelihoodfunktion eine Folge $\theta(k), \theta(k+1), \dots$ von Näherungen für die Schätzung des unbekannten Parametervektors $\theta = (\theta_1, \dots, \theta_m) \in \Theta$ ermittelt. Sei $\theta(k)$ die k -te Näherung. Wir betrachten die Taylorentwicklung der Likelihoodgleichungen

$$\frac{\partial l(\theta, x)}{\partial \theta_i} = \frac{\partial l(\theta(k), x)}{\partial \theta_i} + \sum_{j=1}^m \frac{\partial^2 l(\theta(k), x)}{\partial \theta_i \partial \theta_j} (\theta_j - \theta_j(k)) + o(\|\theta - \theta(k)\|).$$

Da die Schätzung so bestimmt werden soll, daß $\frac{\partial l(\theta, x)}{\partial \theta_i} = 0$ für $i = 1, \dots, m$ gilt, wird die $(k+1)$ -te Näherung aus

$$0 = \frac{\partial l(\theta(k), x)}{\partial \theta_i} + \sum_{j=1}^m \frac{\partial^2 l(\theta(k), x)}{\partial \theta_i \partial \theta_j} (\theta_j(k+1) - \theta_j(k)), i = 1, \dots, m \quad (5.17)$$

ermittelt.

Damit ist in jedem Näherungsschritt ein System linearer Gleichungen der Ordnung m zu lösen. Für viele Aufgaben konvergiert die Folge $\theta(k)$ gegen die Maximum-Likelihood-Schätzung $\check{\theta}$ für $k \rightarrow \infty$.

Es gibt jedoch auch Beispiele, bei denen die Newton-Raphson-Methode versagt.

Beispiel 5.18 Wir betrachten n unabhängige und identisch verteilte Beobachtungen einer Mischverteilung der Form

$$f(x, \theta) = p \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} + (1-p)e^{-\lambda|x|}$$

mit $\theta = (p, \lambda, \sigma^2)^1$. Man kann sich leicht davon überzeugen, daß die Likelihoodfunktion in diesem Fall (wie überhaupt bei Mischverteilungen) eine Struktur aufweist, die auch die numerische Lösung der Likelihoodgleichungen fast unmöglich macht. \diamond

Für diese Art Probleme ist häufig der sogenannte EM (expectation–maximization)–Algorithmus geeignet. Eine ausführlichere Erläuterung des EM–Algorithmus sowie Bedingungen der Konvergenz der EM–Folge gegen die Maximum–Likelihood–Schätzung findet man in WU. Andere Beispiele sind in COX, OAKES angeführt.

5.3 Nichtparametrische Schätzungen von Lebensdauerverteilungen. Das verallgemeinerte Maximum–Likelihood–Prinzip

In diesem Abschnitt wenden wir uns nichtparametrischen statistischen Methoden zu. Wir werden verallgemeinerte Maximum–Likelihood–Schätzungen zuerst für den Fall vollständiger Beobachtungen herleiten und dann dieses Prinzip auf Daten der Art „Dosis–Effekt“, wie im Abschnitt 5.1 beschrieben, anwenden.

Betrachten wir das allgemeine statistische Modell

$$(\mathfrak{X} = (x), \mathfrak{B}(\mathfrak{X}), \mathfrak{P} = (P_\theta, \theta \in \Theta)) .$$

Das verallgemeinerte Likelihoodprinzip ist auch dann anwendbar, wenn $\mathfrak{P} = (P_\theta, \theta \in \Theta)$ eine nichtdominierte Verteilungsfamilie ist. Wir betrachten für jedes Paar $P_{\theta_1}, P_{\theta_2} \in \mathfrak{P}$ das Maß $\mu = P_{\theta_1} + P_{\theta_2}$. Dann sind P_{θ_1} und P_{θ_2} absolutstetig bezüglich μ , d.h. aus $\mu(B) = P_{\theta_1}(B) + P_{\theta_2}(B) = 0$ folgt, daß $P_{\theta_1}(B) = 0$ bzw. $P_{\theta_2}(B) = 0$ sind. Nach dem Satz von Radon–Nikodym existiert eine Dichte $p(x, \theta_1, \theta_2)$ für das Maß P_{θ_1} bezüglich $\mu = P_{\theta_1} + P_{\theta_2}$, d.h. $\forall B \in \mathfrak{B}(\mathfrak{X})$ gilt

$$\int_B p(x, \theta_1, \theta_2) \mu(dx) = P_{\theta_1}(B).$$

Die Dichte $p(x, \theta_1, \theta_2)$ hängt nicht nur von θ_1 , sondern auch von θ_2 ab. Analog findet man eine Dichte $p(x, \theta_2, \theta_1)$ für P_{θ_2} mit

$$\int_B p(x, \theta_2, \theta_1) \mu(dx) = P_{\theta_2}(B), \quad \forall B \in \mathfrak{B}(\mathfrak{X}) .$$

Die beiden Dichten $p(\cdot, \theta_1, \theta_2)$ und $p(\cdot, \theta_2, \theta_1)$ unterscheiden sich voneinander, d.h. es ist nicht möglich, für alle x gleiche Werte zu erhalten:

$$p(x, \theta_1, \theta_2) \not\equiv p(x, \theta_2, \theta_1) .$$

¹Diese Verteilung wurde benutzt, um Daten über den Abstand von Flugzeugen auf Start- und Landebahnen eines japanischen Flughafens auszuwerten. Dabei stand eine Stichprobe vom Umfang $n \sim 10000$ zur Verfügung (Review of RGCSP – WP/158, 13/12/88, Report of Agentia Item 1, Annex C)

Analog zur Likelihoodfunktion im vorigen Abschnitt kann man wieder eine Funktion $p(\cdot, \cdot, \cdot)$ über $\mathfrak{X} \times \Theta \times \Theta = ((x, \theta_1, \theta_2))$, $p : \mathfrak{X} \times \Theta \times \Theta \rightarrow [0, \infty)$ von 3 Veränderlichen betrachten. Wenn man θ_1, θ_2 fixiert, so erhält man als Schnitt eine Funktion von x . Wenn man dagegen x fixiert, erhält man eine Funktion $L = L(\cdot, \cdot, x)$. Dabei gilt wieder für jeden Punkt: $(x, \theta_1, \theta_2) \quad L(\theta_1, \theta_2, x) = p(x, \theta_1, \theta_2)$ (siehe auch Abbildung 5.5). $L(\cdot, \cdot, x)$

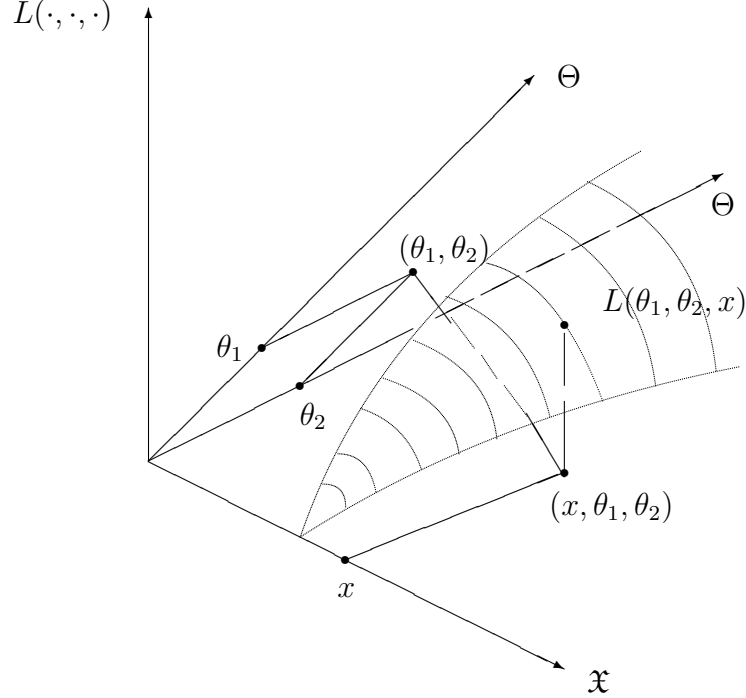


Abbildung 5.5: Die verallgemeinerte Likelihoodfunktion

ist eine Funktion auf $\Theta \times \Theta$. Diese Funktion kann man wieder als Statistik betrachten, da sie jeder Beobachtung x aus dem Stichprobenraum \mathfrak{X} einen Wert (in diesem Falle eine Funktion der beiden Parameter θ_1 und θ_2) zuordnet.

Definition 5.5 Die Funktion $L(\cdot, \cdot, x)$ heißt verallgemeinerte Likelihoodfunktion. Die Größe $L(\theta_1, \theta_2, x)$ heißt Likelihood für θ_1 bezüglich θ_2 .

Beispiel 5.19 Sei $x = (s_1, \dots, s_n)$ die vollständige Beobachtung von unabhängigen und identisch verteilten Zufallsgrößen mit der Verteilungsfunktion $F_\theta(\cdot)$ aus der Familie $\mathfrak{F}_{ac} = (F_\theta, \theta \in \Theta)$ der absolut stetigen Verteilungen. Dann existieren für beliebige $F_1, F_2 \in \mathfrak{F}_{ac}$ die entsprechenden Dichten $f(s, \theta_i) = dF_\theta(s)/ds$, $i = 1, 2$ und für die Dichte der beobachteten Stichprobe erhält man

$$p(x, \theta_i) = f(s_1, \theta_i) \cdots f(s_n, \theta_i), \quad i = 1, 2.$$

Da für jedes $B \in \mathbb{R}^n$

$$\int_B \left(\prod_{i=1}^n f(s_i, \theta_1) + \prod_{i=1}^n f(s_i, \theta_2) \right) ds_1 \cdots ds_n = P_{\theta_1}(B) + P_{\theta_2}(B) = \mu(B)$$

gilt, besitzt das Maß $P_{\theta_1} + P_{\theta_2}$ die Dichte

$$\prod_{i=1}^n f(s_i, \theta_1) + \prod_{i=1}^n f(s_i, \theta_2) .$$

Hieraus erhält man für P_{θ_1} die Dichte

$$p(x, \theta_1, \theta_2) = \frac{\prod_{i=1}^n f(s_i, \theta_1)}{\prod_{i=1}^n f(s_i, \theta_1) + \prod_{i=1}^n f(s_i, \theta_2)} . \quad (5.18)$$

Wahrscheinlichkeiten einer Menge $B \in \mathbb{R}^n$ können mit dieser Dichte folgendermaßen berechnet werden:

$$\begin{aligned} & \int_B p(x, \theta_1, \theta_2) \mu(ds_1, \dots, ds_n) = \\ &= \int_B \frac{\prod_{i=1}^n f(s_i, \theta_1)}{\prod_{i=1}^n f(s_i, \theta_1) + \prod_{i=1}^n f(s_i, \theta_2)} \left(\prod_{i=1}^n f(s_i, \theta_1) + \prod_{i=1}^n f(s_i, \theta_2) \right) ds_1 \cdots ds_n = \\ &= \int_B \prod_{i=1}^n f(s_i, \theta_1) ds_1 \cdots ds_n = P_{\theta_1}(B) . \end{aligned}$$

◇

Beispiel 5.20 Wir betrachten wieder eine vollständige Beobachtung, nehmen jedoch an, daß die beiden möglichen Verteilungsgesetze F_{θ_1} und F_{θ_2} zur Familie der reinen Sprungfunktionen gehören,

$$F_{\theta_i}(s) = \sum_{u_i \leq s} \Delta F(u_i), \quad F_{\theta_i}(\infty) = 1 .$$

In diesem Falle können natürlich nur solche s_i beobachtet werden, in denen die Verteilungsfunktion einen Sprung besitzt. Bezeichnen wir die Menge aller Sprünge mit $U(\theta_1, \theta_2) = (u : \Delta F_{\theta_1}(u) + \Delta F_{\theta_2}(u) > 0)$, so gilt $s_i \in U(\theta_1, \theta_2)$, $i = 1, \dots, n$.

Für die Wahrscheinlichkeit der Beobachtung erhält man

$$P_{\theta_i}(X = x) = \prod_{i=1}^n \Delta F_{\theta_i}(s_i), \quad x \in \mathfrak{X} .$$

Die Menge \mathfrak{X} enthält alle solche $\tilde{x} = (u_{i_1}, \dots, u_{i_n})$, für die $u_{i_k} \in U(\theta_1, \theta_2)$. \mathfrak{X} ist abzählbar. Auf \mathfrak{X} betrachten wir nun das Maß μ :

$$\mu(\{\tilde{x}\}) = \prod_{k=1}^n \Delta F_{\theta_1}(u_{i_k}) + \prod_{k=1}^n \Delta F_{\theta_2}(u_{i_k}) .$$

Dann hat der Punkt $\tilde{x} \in \mathfrak{X}$ die Dichte

$$p(\tilde{x}, \theta_1, \theta_2) = \frac{\prod_{k=1}^n \Delta F_{\theta_1}(u_{i_k})}{\prod_{k=1}^n \Delta F_{\theta_1}(u_{i_k}) + \prod_{k=1}^n \Delta F_{\theta_2}(u_{i_k})} .$$

Hiervon kann man sich leicht überzeugen, da einerseits $P_{\theta_1}(\tilde{x}) = \prod_{k=1}^n \Delta F_{\theta_1}(u_{i_k})$ und andererseits

$$\begin{aligned} P_{\theta_1}(\tilde{x}) &= \int_{\tilde{x}} p(\tilde{x}', \theta_1, \theta_2) \mu(d\tilde{x}') = p(\tilde{x}', \theta_1, \theta_2) \mu(\{\tilde{x}\}) = \\ &= \frac{\prod_{k=1}^n \Delta F_{\theta_1}(u_{i_k})}{\prod_{k=1}^n \Delta F_{\theta_1}(u_{i_k}) + \prod_{k=1}^n \Delta F_{\theta_2}(u_{i_k})} \cdot \\ &\quad \cdot \left(\prod_{k=1}^n \Delta F_{\theta_1}(u_{i_k}) + \prod_{k=1}^n \Delta F_{\theta_2}(u_{i_k}) \right) = \prod_{k=1}^n \Delta F_{\theta_1}(u_{i_k}) \end{aligned}$$

gilt. Ebenso wie im vorigen Beispiel folgt hieraus für beliebige Teilmengen $B \in \mathfrak{X}$

$$\begin{aligned} P_{\theta_1}(B) &= \sum_{\substack{\tilde{x} \in B \\ \tilde{x} \in U(\theta_1, \theta_2)}} P(\tilde{x}) = \sum_{\substack{\tilde{x} \in B \\ \tilde{x} \in U(\theta_1, \theta_2)}} \int_{\{\tilde{x}\}} p(\tilde{x}', \theta_1, \theta_2) d\mu(\tilde{x}') \\ &= \int_{\tilde{x} \in B} p(\tilde{x}', \theta_1, \theta_2) d\mu(\tilde{x}') . \end{aligned}$$

Damit erhält man im diskreten Fall

$$L(\theta_1, \theta_2, x) = \frac{\prod_{i=1}^n \Delta F_{\theta_1}(s_i)}{\prod_{i=1}^n \Delta F_{\theta_1}(s_i) + \prod_{i=1}^n \Delta F_{\theta_2}(s_i)} .$$

◇

Wir kommen nun zur Formulierung des allgemeinen Likelihood-Prinzips. Dabei werden alle statistischen Schlußfolgerungen mit Hilfe der verallgemeinerten Likelihoodfunktion $L(\cdot, \cdot, x)$ gezogen.

Definition 5.6 Eine Punktschätzung $\check{\theta} : \mathfrak{X} \rightarrow \Theta$ heißt verallgemeinerte Maximum-Likelihood-Schätzung, wenn für jedes $\theta \in \Theta$

$$L(\check{\theta}(x), \theta, x) \geq L(\theta, \check{\theta}(x), x) .$$

Im folgenden werden wir die Bezeichnung $\check{\theta}_1(x)(\geq)\check{\theta}_2(x)$ benutzen, wenn

$$L(\check{\theta}_1(x), \check{\theta}_2(x), x) \geq L(\check{\theta}_2(x), \check{\theta}_1(x), x) .$$

Damit ist $\check{\theta}(x)$ eine verallgemeinerte Maximum-Likelihood-Schätzung, wenn $\check{\theta}(x)(\geq)\theta$ für beliebige $\theta \in \Theta$.

Beispiel 5.21 Wir betrachten wieder den Fall einer Stichprobe aus vollständigen Beobachtungen $x = (s_1, \dots, s_n)$, wobei die s_i , $i = 1, \dots, n$ Realisierungen unabhängiger und identisch verteilter Zufallsgrößen S_i mit einer Verteilungsfunktion $F(\cdot) \in \mathfrak{F}$ sind. \mathfrak{F} sei die Familie aller Verteilungsfunktionen, d.h. wir treffen keine Voraussetzungen über die Stetigkeit des Verteilungsgesetzes. Sei $F_{\theta_1}(\cdot)$ eine Verteilungsfunktion, die Sprünge in den Punkten s_i besitzt. Dann gilt

$$P_{\theta_1}(x) = \prod_{i=1}^n \Delta F(s_i) .$$

Sei nun $F_{\theta_2}(\cdot)$ eine beliebige andere Verteilungsfunktion. Die verallgemeinerte Likelihood-funktion hat dann die Form

$$L(\theta_1, \theta_2, x) = \frac{\prod_{i=1}^n \Delta F_{\theta_1}(s_i)}{\prod_{i=1}^n \Delta F_{\theta_1}(s_i) + \prod_{i=1}^n \Delta F_{\theta_2}(s_i)} .$$

Ist $\Delta F_{\theta_2}(s_i) = 0$ für einen der Beobachtungswerte s_i , so gilt $\prod_{i=1}^n \Delta F_{\theta_2}(s_i) = 0$ und damit $L(\theta_1, \theta_2, x) = 1$. Andererseits ist in diesem Fall $L(\theta_2, \theta_1, x) = 0$. Damit erhält man $1 = L(\theta_1, \theta_2, x) > L(\theta_2, \theta_1, x)$ und nach dem verallgemeinerten Likelihood-Prinzip ist $F_{\theta_1}(\cdot)$ der Verteilungsfunktion $F_{\theta_2}(\cdot)$ vorzuziehen. Damit ist die Schätzung eines beliebigen Verteilungsgesetzes unter den Funktionen zu suchen, die Sprünge in den Punkten s_i , $i = 1, \dots, n$ besitzen. Um nun die verallgemeinerte Likelihood-Schätzung zu bestimmen, muß man $\check{F}_n(x)$ so finden, daß

$$\prod_{i=1}^n \Delta \check{F}_n(s_i) = \max_{\forall F(\cdot) \in \mathfrak{F}} \prod_{i=1}^n \Delta F(s_i) , \quad (5.19)$$

da dann auch

$$\frac{\prod_{i=1}^n \Delta \check{F}_n(s_i)}{\prod_{i=1}^n \Delta \check{F}_n(s_i) + \prod_{i=1}^n \Delta F_{\theta_2}(s_i)} \geq \frac{\prod_{i=1}^n \Delta F_{\theta_2}(s_i)}{\prod_{i=1}^n \Delta \check{F}_n(s_i) + \prod_{i=1}^n \Delta F_{\theta_2}(s_i)} .$$

Wie schon bemerkt, suchen wir $\check{F}_n(x)$ unter den Verteilungsfunktionen, die Sprünge in den Punkten s_i besitzen. Wir gehen zur geordneten Stichprobe $z_{(1)} < z_{(2)} < \dots < z_{(k)}$

über, wobei $z_1 = \min(s_i : i = 1, \dots, n) < z_2 = \min(s_i : s_i > z_{(1)}, i = 1, \dots, n)$ ist. Seien d_j die Anzahl der Beobachtungen $s_{(i)} : d_j = \sum_{i=1}^n \mathbf{I}(s_i = z_{(j)})$, $d_1 + \dots + d_k = n$ und p_j die Sprunghöhen $p_j = \Delta F_\theta(z_{(j)})$, $j = 1, \dots, k$. Dann folgt $\prod_{i=1}^n \Delta F_\theta(s_i) = \prod_{j=1}^k p_j^{d_j}$ mit $\sum_{j=1}^k p_j = 1$. Somit erfordert die Lösung der Gleichung (5.19) solche \check{p}_j zu finden, daß

$$\prod_{j=1}^k \check{p}_j^{d_j} = \max \left(\prod_{j=1}^k p_j^{d_j} : \sum_{j=1}^k p_j = 1, \quad 0 \leq p_j \leq 1 \right).$$

Wendet man die Lagrange-Methode an, erhält man

$$H = \prod_{j=1}^k p_j^{d_j} - \lambda \left(\sum_{j=1}^k p_j - 1 \right) \rightarrow \max$$

und hieraus

$$\frac{\partial H}{\partial p_l} = \frac{d_l}{p_l} \prod_{j=1}^k p_j^{d_j} - \lambda = 0, \quad l = 1, \dots, k, \quad \sum_{j=1}^k p_j = 1.$$

Löst man dieses Gleichungssystem, so erhält man

$$\lambda = n \prod_{j=1}^k p_j^{d_j}, \quad p_l = \frac{d_l}{n}, \quad l = 1, \dots, k$$

und damit ist die verallgemeinerte Maximum-Likelihood-Schätzung eine Funktion mit Sprüngen $\Delta \check{F}_n(z_{(j)}) = d_j/n$:

$$\check{F}_n(s) = \sum_{j: z_{(j)} \leq s} \Delta \check{F}_n(z_{(j)}) = \frac{1}{n} \sum_{i=1}^n \mathbf{I}(s_i \leq s).$$

◇

Das Ergebnis des Beispiels 5.21 kann in folgendem Satz zusammengefaßt werden:

Satz 5.1 *Für vollständige Beobachtungen ist die empirische Verteilungsfunktion*

$$\check{F}_n(s) = \frac{1}{n} \sum_{i=1}^n \mathbf{I}(s_i \leq s)$$

eine verallgemeinerte Maximum-Likelihood-Schätzung der unbekannten Verteilungsfunktion $F(s)$.

Bemerkung: Setzt man diese Schätzung in Formel (2.1) ein, so erhält man eine Maximum-Likelihood-Schätzung für die Hazardfunktion

$$\check{H}_n(s) = \sum_{i=1}^n \frac{I(s_i \leq s)}{Y(n, s_i)} \quad \text{mit} \quad Y(n, s_i) = \sum_{j=1}^n I(s_j \geq s_i) .$$

Es ist gut bekannt, daß die empirische Verteilungsfunktion viele interessante Eigenschaften besitzt. Einige dieser Eigenschaften folgen auch daraus, daß sie die eindeutig bestimmte Maximum-Likelihood-Schätzung ist. Wir beschränken uns hier auf die Betrachtung von Erwartungswert und Varianz dieser Schätzung. Für den Erwartungswert erhält man

$$E\check{F}_n(s) = E\left(\frac{1}{n} \sum_{i=1}^n I(S_i \leq s)\right) = \frac{1}{n} \sum_{i=1}^n E(I(S_i \leq s)) = \frac{1}{n} \sum_{i=1}^n F(s) = F(s) .$$

Damit ist die Schätzung erwartungstreu. Für die Varianz der Schätzung erhält man

$$\text{Var}(\check{F}_n(s)) = \frac{1}{n} F(s) \overline{F}(s) .$$

Beispiel 5.22 Wir wenden nun das verallgemeinerte Likelihood-Prinzip auf Dosis-Effekt-Daten an. Daten dieser Art werden auf folgende Weise gewonnen: Man untersucht die Elemente e_1, \dots, e_n . Auf das i -te Element wird jeweils ein Schock der Stärke t_i (Dosis) ausgeübt. Danach kann festgestellt werden, ob das Element ausgefallen ($\delta_i = 1$) oder noch intakt ist ($\delta_i = 0$) (Effekt). Wir nehmen an, daß die δ_i unabhängig voneinander sind. Sie unterliegen jedoch keiner identischen Verteilung, da der Ausfall davon abhängt, welche Kraft t_i auf das Element ausgeübt wurde. Die Funktion

$$F(t_i) = P\{\delta_i = 1\}$$

heißt Dosis-Effekt-Funktion. Sie ist in Abbildung 5.6 dargestellt. Man kann sich vor-

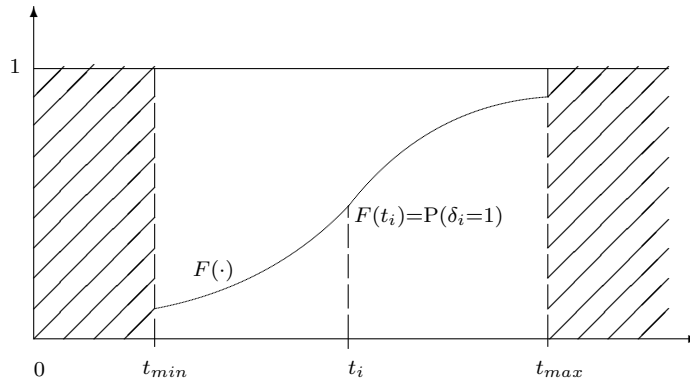


Abbildung 5.6: Die Dosis-Effekt-Funktion

stellen, daß jedes Element eine kritische Grenze S_i besitzt. Wenn $S_i \leq t_i$, dann fällt das

Element aus, d.h. $\delta_i = 1$ bzw. bei $S_i > t_i$ ist $\delta_i = 0$. Damit erhält man $\delta_i = \mathbf{I}(S_i \leq t_i)$. Die Beobachtungsdaten haben die Form

$$x = \begin{pmatrix} t_1 & \dots & t_n \\ \delta_1 & \dots & \delta_n \end{pmatrix}$$

und die Verteilung dieser Beobachtung ist durch

$$P_F(x) = \prod_{i=1}^n F(t_i)^{\delta_i} \overline{F}(t_i)^{1-\delta_i}$$

gegeben. Es soll nun eine Punktschätzung $\check{F}_n(s)$ ermittelt werden. Dazu verwenden wir wieder die verallgemeinerte Likelihoodfunktion

$$L(F_1(\cdot), F_2(\cdot), x) = \frac{\prod_{i=1}^n F_1(t_i)^{\delta_i} \overline{F}_1(t_i)^{1-\delta_i}}{\prod_{i=1}^n F_1(t_i)^{\delta_i} \overline{F}_1(t_i)^{1-\delta_i} + \prod_{i=1}^n F_2(t_i)^{\delta_i} \overline{F}_2(t_i)^{1-\delta_i}} .$$

Wie bereits früher in diesem Abschnitt muß eine Schätzung $\check{F}_n(\cdot)$ so ermittelt werden, daß

$$\prod_{i=1}^n \check{F}_n(t_i)^{\delta_i} (1 - \check{F}_n(t_i))^{1-\delta_i} = \max_{F(\cdot) \in \mathfrak{F}} \prod_{i=1}^n F(t_i)^{\delta_i} (1 - F(t_i))^{1-\delta_i} ,$$

wobei $\check{F}_n(\cdot)$ eine nichtfallende Funktion ist. Logarithmiert man diesen Ausdruck, so erhält man

$$\begin{aligned} \ln \prod_{i=1}^n F(t_i)^{\delta_i} (1 - F(t_i))^{1-\delta_i} &= \sum_{i=1}^n (\delta_i \ln F(t_i) + (1 - \delta_i) \ln(1 - F(t_i))) = \\ &= - \sum_{i=1}^n \left(\delta_i \ln \frac{1}{F(t_i)} + (1 - \delta_i) \ln \frac{1}{1 - F(t_i)} \right) . \end{aligned}$$

Bezeichnen wir mit $F_i = F(t_i)$ und

$$R(F_1, \dots, F_n) = \sum_{i=1}^n \left(\delta_i \ln \frac{1}{F_i} + (1 - \delta_i) \ln \frac{1}{1 - F_i} \right) ,$$

so ist eine Folge $0 \leq \check{F}_1 \leq \dots \leq \check{F}_n \leq 1$ so zu bestimmen, daß

$$\sum_{i=1}^n \left(\delta_i \ln \frac{1}{\check{F}_i} + (1 - \delta_i) \ln \frac{1}{1 - \check{F}_i} \right) = \min_{\forall 0 \leq \check{F}_1 \leq \dots \leq \check{F}_n \leq 1} R(F_1, \dots, F_n) .$$

Für die Funktion $R(\cdot)$ läßt sich mit Hilfe der Theorie isotonischer Schätzungen (vergleiche BELYAEV, CHEPURIN, ROBERTSON, WRIGHT, BARLOW, BARTHOLOMEW, BREMNER, BRUNK) eine sogenannte Cauchy-Funktion finden. Diese Funktion ist auf Teilmengen

$Q = (t_{i_1}, \dots, t_{i_k}) \subseteq (t_1, \dots, t_n)$ aller Punkte (t_1, \dots, t_n) definiert. Eine Funktion $K(Q)$ auf diesen Teilmengen heißt Cauchy-Funktion, wenn für zwei disjunkte Teilmengen $Q_1, Q_2 : Q_1 \cap Q_2 = \emptyset$

$$\min(K(Q_1), K(Q_2)) \leq K(Q_1 \cup Q_2) \leq \max(K(Q_1), K(Q_2))$$

gilt. Betrachten wir nun die Funktion

$$K(Q) = \sum_{i=1}^n \delta_i I(t_i \in Q) / \sum_{i=1}^n I(t_i \in Q) . \quad (5.20)$$

Diese Funktion ist eine Cauchy-Funktion. Mit dieser Cauchy-Funktion $K(Q)$ läßt sich nun das Minimum von $R(F_1, \dots, F_n)$ auf folgende Weise ermitteln:

Wir nehmen Einpunktmengen $\{t_1\}, \dots, \{t_n\}$ und berechnen die entsprechenden Cauchy-Funktionen

$$K(\{t_1\}) = \delta_1, \dots, K(\{t_n\}) = \delta_n .$$

Nun überprüfen wir die Ungleichung $K(\{t_1\}) \leq K(\{t_2\})$. Ist diese Ungleichung nicht erfüllt, so vereinigt man $\{t_1\}$ und $\{t_2\}$ zur Menge $\{t_1, t_2\}$ und definiert auf dieser Menge $K(\{t_1, t_2\}) = \frac{\delta_1 + \delta_2}{2}$.

Diese Funktion vergleicht man nun mit $K(\{t_3\})$ usw. Das Ziel des Algorithmus besteht darin, eine nichtfallende Funktion zu bestimmen. Dabei wird in jedem Schritt bei einer verletzten Ungleichung das Überprüfen abgebrochen, die entsprechenden Mengen werden vereinigt und man beginnt im nächsten Schritt die Überprüfung der Ungleichungen von vorn. Hat man solche Mengen $\{t_1, \dots, t_{\nu_1}\}, \{t_{\nu_1+1}, \dots, t_{\nu_2}\}, \dots, \{t_{\nu_{l-1}+1}, \dots, t_{\nu_n}\}$ gefunden, daß

$$K(\{t_1, \dots, t_{\nu_1}\}) \leq K(\{t_{\nu_1+1}, \dots, t_{\nu_2}\}) \leq \dots \leq K(\{t_{\nu_{l-1}+1}, \dots, t_{\nu_n}\}) ,$$

so ist die gesuchte Schätzung der Verteilungsfunktion

$$\check{F}_n(s) = \begin{cases} K(\{t_1, \dots, t_{\nu_1}\}), & 0 \leq s < t_{\nu_1} \\ K(\{t_{\nu_1+1}, \dots, t_{\nu_2}\}), & t_{\nu_1} \leq s < t_{\nu_2} \\ \dots & \dots \\ K(\{t_{\nu_{l-1}+1}, \dots, t_{\nu_n}\}), & t_{\nu_{l-1}} \leq s . \end{cases}$$

Betrachten wir dazu ein Zahlenbeispiel:

Es wurden 8 Versuche durchgeführt und dabei die Daten

$$x = \begin{pmatrix} 11 & 25 & 34 & 39 & 41 & 53 & 71 & 80 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}$$

gewonnen.

In den folgenden Schritten werden jeweils in der ersten Zeile die Teilmengen der Nummern des Versuches und in der zweiten Zeile der entsprechende Wert der Cauchy-Funktion dargestellt.

1. Schritt	$\{1\}$ 0	$\{2\}$ 1	$\{3\}$ 0	$\{4\}$	$\{5\}$	$\{6\}$	$\{7\}$	$\{8\}$
2. Schritt	$\{1\}$ 0	$\{2, 3\}$ 1/2	$\{4\}$ 1	$\{5\}$ 1	$\{6\}$ 0	$\{7\}$	$\{8\}$	
3. Schritt	$\{1\}$ 0	$\{2, 3\}$ 1/2	$\{4\}$ 1	$\{5, 6\}$ 1/2	$\{7\}$	$\{8\}$		
4. Schritt	$\{1\}$ 0	$\{2, 3\}$ 1/2	$\{4, 5, 6\}$ 2/3	$\{7\}$ 1	$\{8\}$ 1			

Damit ist der Algorithmus beendet und man hat die Schätzung

$$\check{F}_8(s) = \begin{cases} 0, & 0 \leq s < 25 \\ 1/2, & 25 \leq s < 39 \\ 2/3, & 39 \leq s < 71 \\ 1, & 71 \leq s \end{cases}$$

erhalten. ◇

Durch das verallgemeinerte Maximum-Likelihood-Prinzip ist es möglich, nichtparametrische Schätzungen zu bestimmen.

6 Tests und Konfidenzschätzungen basierend auf Zuverlässigkeitsdaten

6.1 Graphische Methoden in der Zuverlässigkeitstheorie

Graphische Methoden haben in den letzten Jahren nicht an Bedeutung verloren. Solange die Ausstattung mit Computern mangelhaft war, boten verschiedene Spezialpapiere die Möglichkeit, eine Verteilungsannahme sehr rasch auf graphischem Wege zu überprüfen. Man könnte annehmen, daß diese Hilfsmittel mit der Entwicklung von Arbeitsplatzcomputern und entsprechender Statistiksoftware überflüssig geworden wären. Gerade die Anwendung komfortabler Software birgt jedoch die Gefahr in sich, daß formal statistische Methoden auf Daten angewandt werden, für die sie nicht anwendbar sind. Graphische Methoden bieten die Möglichkeit, gewisse Eigenschaften der Stichprobe zu erkennen und eine Auswahl geeigneter Methoden zu treffen.

Natürlich wird man heute nicht mehr mit Spezialpapieren arbeiten, sondern diese „Papiere“ programmieren. Daher sollen in diesem Abschnitt einige wichtige und interessante graphische Methoden dargestellt werden.

Eine statistische Hypothese besteht im allgemeinen in der Behauptung: „Der wahre Parameter θ^* liegt in einer Teilmenge Θ_i des Parameterraumes Θ “. Wie schon im Abschnitt 5.3 erläutert, lassen sich auf diese Weise sowohl parametrische als auch nichtparametrische Aufgabenstellungen formulieren.

Wir betrachten also ein statistisches Modell

$$\left(\mathfrak{X} = (x), \mathfrak{B}(\mathfrak{X}), \mathfrak{P} = (P_\theta, \theta \in \Theta) \right)$$

und wollen auf Grundlage einer Beobachtung x entscheiden, ob die Hypothese $\theta^* \in \Theta_i$ abzulehnen ist oder nicht.

Die geometrische Idee graphischer Verfahren besteht darin, die Daten x so umzuformen, daß verschiedene Hypothesen leicht voneinander unterschieden werden können. Betrachten wir als Beispiel die 3 Annahmen

$$\begin{aligned} \Theta_1 &= \mathfrak{F}_{\text{E},1} = (F(\cdot) : \overline{F}(s) = e^{-\lambda s}) , \\ \Theta_2 &= \mathfrak{F}_{\text{IFR}} = (F(\cdot) : h(s) = \frac{f(s)}{\overline{F}(s)} \uparrow (s \uparrow)) , \\ \Theta_3 &= \mathfrak{F}_{\text{DFR}} = (F(\cdot) : h(s) = \frac{f(s)}{\overline{F}(s)} \downarrow (s \uparrow)) , \end{aligned}$$

d.h. wir wollen entscheiden, ob die Stichprobe zu einer Verteilung mit konstanter Ausfallrate, mit wachsender Ausfallrate oder mit fallender Ausfallrate gehört. Dazu versuchen wir eine solche Abbildung zu wählen, daß die transformierte Verteilungsfunktion eine lineare, konvexe oder eine konkave Funktion bildet, die dann optisch gut voneinander unterschieden werden können.

Der Einfachheit halber nehmen wir an, daß $F(\cdot)$ eine monotone Verteilungsfunktion mit der Dichte $f(s) = dF(s)/ds > 0$ und dem Erwartungswert $m_{F,1} < \infty$ ist. Dann läßt sich für jedes $p \in [0, 1]$ das Quantil s_p der Ordnung p bestimmen:

$$F(s_p) = p, \quad s_p = F^{-1}(p)$$

Das Quantil erhält man aus der Umkehrfunktion der Verteilungsfunktion, und natürlich gelten

$$F(F^{-1}(p)) = p, \quad F^{-1}(F(s)) = s.$$

Bezeichnen wir mit $p_s = F(s)$, so gilt weiterhin

$$\frac{ds_p}{dp} = \frac{dF^{-1}(p)}{dp} = \frac{1}{\frac{dF(s)}{ds} \big|_{s=s_p}} = \frac{1}{f(s_p)} > 0.$$

Wir betrachten nun die Abbildung $T_F : [0, 1] \rightarrow [0, 1]$

$$T_F(p) = \frac{1}{m_{F,1}} \int_0^{F^{-1}(p)} \bar{F}(u) du. \quad (6.1)$$

Da sich der Erwartungswert aus $m_{F,1} = \int_0^\infty \bar{F}(u) du$ (vergleiche Abschnitt 2.2) berechnen läßt, folgt aus Formel (6.1), daß $T_F(\cdot)$ nichtfallend ist und daß $T_F(0) = 0$ und $T_F(1) = 1$ gelten. Ein mögliches Aussehen dieser Funktion ist in Abbildung 6.1 dargestellt.

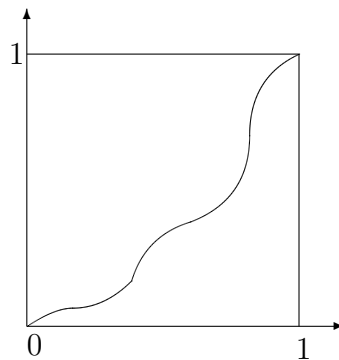


Abbildung 6.1: Die Funktion $T_F(p)$

Nun betrachten wir die Ableitung $\frac{dT_F(p)}{dp}$:

$$\begin{aligned} \frac{dT_F(p)}{dp} &= \frac{1}{m_{F,1}} \frac{d}{dp} \left(\int_0^{F^{-1}(p)} \bar{F}(u) du \right) = \frac{1}{m_{F,1}} \bar{F}(F^{-1}(p)) \frac{dF^{-1}(p)}{dp} = \\ &= \frac{1-p}{m_{F,1}} \frac{1}{f(s_p)} = \frac{1}{m_{F,1}} \frac{1}{\frac{f(s_p)}{1-F(s_p)}} = \frac{1}{m_{F,1}} \frac{1}{h(s_p)}, \end{aligned} \quad (6.2)$$

wobei $h(s) = \frac{f(s)}{1-F(s)}$ die Ausfallrate ist.

Aus dieser Herleitung lassen sich bezüglich der 3 aufgeführten Annahmen Θ_1, Θ_2 und Θ_3 drei Folgerungen ableiten.

Folgerung 6.1 Wenn $F \in \mathfrak{F}_{E,1}$, so gilt $T_F(p) \equiv p$.

In diesem Fall ist $h(s_p) = \lambda, m_{F,1} = \frac{1}{\lambda}$ und damit erhält man $\frac{dT_F(p)}{dp} = \frac{1}{\frac{1}{\lambda}} \frac{1}{\lambda} \equiv 1$.

Damit gilt $T_F(p) = \int_0^p \frac{dT_F(p')}{dp'} dp' = p$ und die Abbildung $T_F(\cdot)$ ist eine Gerade vom Punkt $(0,0)$ in den Punkt $(1,1)$. \square

Folgerung 6.2 Wenn $F \in \mathfrak{F}_{IFR} \setminus \mathfrak{F}_{E,1}$, so ist $T_F(\cdot)$ eine streng konkave Funktion, deshalb liegt sie oberhalb der Diagonalen $T_F(p) \equiv p$ aus Folgerung 6.1.

Diese Folgerung ist offensichtlich, da $h(s_p) \uparrow (s_p \uparrow)$ bedeutet, daß die Ableitung $\frac{dT_F(p)}{dp}$ eine fallende Funktion ist und damit $T_F(p)$ konkav ist. \square

Ebenso offensichtlich ist die dritte Folgerung für Verteilungen mit fallender Ausfallrate:

Folgerung 6.3 Wenn $F \in \mathfrak{F}_{DFR} \setminus \mathfrak{F}_{E,1}$, so ist $T_F(\cdot)$ eine streng konvexe Funktion, deshalb liegt sie unterhalb der Diagonalen $T_F(p) \equiv p$ aus Folgerung 6.1.

Damit haben wir Verteilungsgesetze mit konstanter, wachsender bzw. fallender Ausfallrate voneinander unterschieden. Wir betrachten nun eine Punktschätzung $\hat{F}_n(s)$ und eine Schätzung des Mittelwertes $\hat{m}_{F,1}$, die aus der Stichprobe ermittelt werden. Die Abbildung T wird auf diese Schätzungen angewandt:

$$T_{\hat{F}_n}(p) = \frac{1}{\hat{m}_{F,1}} \int_0^{\hat{F}_n^{-1}(p)} \hat{\bar{F}}_n(u) du.$$

Ein graphischer Test besteht nun darin, die Funktion $T_{\hat{F}_n}(p)$ graphisch darzustellen und anhand des Verlaufes zu entscheiden, welche der 3 Annahmen zutrifft.

In Abbildung 6.2 sind drei Funktionen $T_{\hat{F}_n}(p)$ dargestellt. Dazu wurden jeweils 50 Ausfalldaten mittels der Weibullverteilung simuliert, und zwar für das linke Bild mit einem

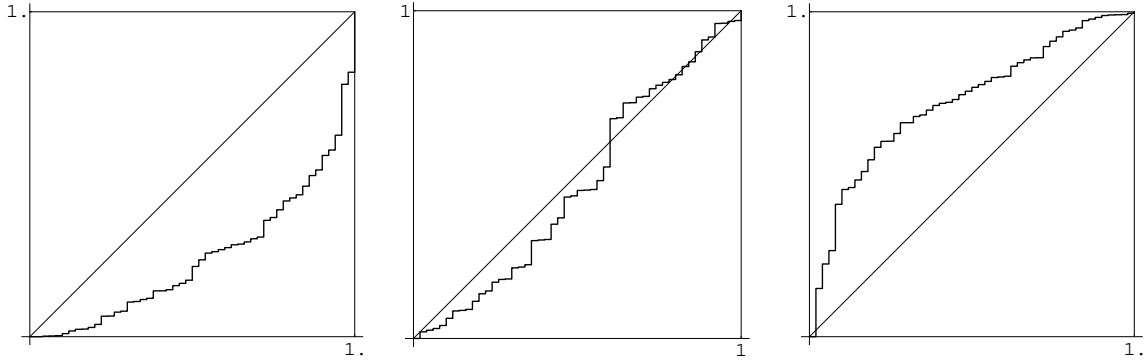


Abbildung 6.2: Drei TTT-Plots für simulierte Ausfalldaten

Formparameter $b = 0.5$ (DFR-Verteilung), für das mittlere Bild mit einem Formparameter $b = 1.0$ (Exponentialverteilung) und für das rechte Bild mit einem Formparameter $b = 2.0$ (IFR-Verteilung). Die Abbildung $F(\cdot) \rightarrow T_F(\cdot)$ nennt man auch total-time-on-test-Abbildung oder TTT-Plot. Diese Bezeichnung kommt daher, daß im Falle vollständiger Beobachtung

$$\hat{\bar{F}}_n(s) = \frac{1}{n} \sum_{i=1}^n \mathbf{I}(s_i > s)$$

und daher

$$t(s) := n \int_0^s \hat{\bar{F}}(u) du = s_{(1)} \mathbf{I}(s_{(1)} < s) + \dots + s_{d(s)} \mathbf{I}(s_{d(s)} < s) + (n - d(s))s ,$$

d.h. $n \int_0^s \hat{\bar{F}}(s) ds$ ist die Gesamtprüfzeit bis zum Zeitpunkt s . Bei vollständiger Beobachtung erhält man als Schätzung des Erwartungswertes

$$m_{F,1} = \frac{1}{n} \sum_{i=1}^n s_i$$

und damit

$$T_{\hat{F}_n}(p) = \frac{1}{\sum_{i=1}^n s_i} \int_0^{\hat{s}_p} n \hat{\bar{F}}_n(u) du = \frac{t(\hat{s}_p)}{t(\hat{s}_1)} ,$$

wobei $\hat{s}_p = \hat{F}^{-1}(p)$ und $t(\hat{s}_p)$ bzw. $t(\hat{s}_1)$ die Gesamtprüfzeiten bis zur Zeit \hat{s}_p bzw. $\hat{s}_1 = s_{(n)}$ sind.

Bei graphischen Tests wird also anhand eines Bildes optisch entschieden, welche der Annahmen getroffen werden kann. Natürlich ist diese Entscheidung keineswegs eindeutig und sehr subjektiv. Es gibt verschiedene Möglichkeiten, diese Entscheidung zu objektivieren. Wenn die Datenmenge genügend groß ist, kann man z.B. die Daten x zufällig in zwei gleiche Teilmengen x' und x'' teilen. Dann berechnet man jeweils mit diesen

Teilmengen die Punktschätzungen \hat{F}'_n und \hat{F}''_n und vergleicht die beiden dazugehörigen Graphiken $T_{\hat{F}'_n}(\cdot)$ und $T_{\hat{F}''_n}(\cdot)$. Wenn man für beide Teilmengen zu der gleichen Entscheidung kommt wie für die ungeteilten Daten $x = x' \cup x''$, so kann man relativ sicher sein, daß die Datenmenge genügend groß und die Entscheidung richtig ist.

Nicht nur im Zusammenhang mit graphischen Methoden, sondern bei jedem Test stellt sich die Frage, wie sicher die getroffene Entscheidung ist. Um diese Frage zu beantworten, müssen die Entscheidungsregeln, wann welche Hypothese abzulehnen ist, exakt formuliert werden und außerdem muß der zufällige Prozeß $T_{\hat{F}'_n}(\cdot)$ bezüglich seiner Variabilität genauer untersucht werden. Wenn die Entscheidungsregeln exakt formuliert sind, so ist es möglich, die Untersuchung des Prozesses $T_{\hat{F}'_n}(\cdot)$ mit Simulation durchzuführen.

Wir wollen hier noch eine andere Möglichkeit der Verwendung graphischer Methoden vorstellen. Diese Möglichkeit basiert auf den sogenannten Wahrscheinlichkeitspapieren.

Beispiel 6.1 Wir nehmen wieder an, daß eine vollständige Beobachtung vorliegt. Seien $X_{(n)} = (S_1, \dots, S_n)$ die Stichprobe und $S_{(1,n)}, \dots, S_{(n,n)}$ die zugehörige Variationsreihe. Die S_i , $i = 1, \dots, n$ seien unabhängig und identisch verteilt mit der Verteilungsfunktion $F(s) = \Phi(\frac{s-\mu}{\sigma})$, d.h. wir untersuchen die Hypothese, daß die Grundgesamtheit einer Normalverteilung mit den Parametern $\theta = (\mu, \sigma)$, $\sigma > 0$ unterliegt. Wir betrachten nun solche Abbildungen, in denen jedem Punkt (s, p) , $0 < p < 1$ ein Punkt (s, u_p) zugeordnet wird, wobei u_p durch die Gleichung $\Phi(u_p) = p$ definiert ist. Damit bilden wir die Verteilungsfunktion $F(s) = \Phi(\frac{s-\mu}{\sigma})$ in eine Funktion auf der Ebene \mathbb{R}^2 ab. Wenn $F(s) = \Phi(\frac{s-\mu}{\sigma}) = p$ ist, dann wird der Punkt (s, p) in den Punkt $(s, u_{p(s)})$ abgebildet. Daher wird die Funktion $F(s) = \Phi(\frac{s-\mu}{\sigma})$ oder $\frac{s-\mu}{\sigma} = u_{F(s)}$ in die Gerade

$$u_{F(s)} = \frac{s}{\sigma} - \frac{\mu}{\sigma} \quad (6.3)$$

abgebildet, d.h. jede Verteilungsfunktion der Normalverteilung bildet bei so einer Abbildung eine Gerade. Die Gerade (6.3) geht durch den Punkt $(\mu, 0)$ und besitzt den Anstiegswinkel $\tan(\varphi) = 1/\sigma$. Je kleiner die Varianz σ ist, umso steiler ist die Gerade. Wenn wir nun die empirische Verteilungsfunktion $\hat{F}_n(s)$ in dieser Abbildung darstellen, dann werden die Punkte $(s_{(i,n)}, \hat{F}_n(s_{(i,n)}))$ der Variationsreihe in die Punkte $(s_{(i,n)}, u_{\frac{i}{n}})$ abgebildet. Ist die Annahme einer Normalverteilung richtig, dann liegen diese Punkte in der Nähe der entsprechenden Punkte $(s, \frac{s}{\sigma} - \frac{\mu}{\sigma})$, d.h. sie bilden annähernd eine Gerade. Die beschriebene Abbildung ist als Spezialpapier (das sogenannte Wahrscheinlichkeitsnetz) erhältlich.

Natürlich läßt sich diese Idee auf alle Verteilungsfamilien der Form $\mathfrak{F}_{F_0(\cdot)} = \{F_0(\frac{s-\mu}{\sigma}), \theta = (\mu, \sigma) \in \Theta \subseteq \mathbb{R}^2\}$ oder auch der Form $\mathfrak{F}_{F_0(\cdot), g(\cdot)} = \{F_0(\frac{g(s)-\mu}{\sigma}), \theta = (\mu, \sigma) \in \Theta \subseteq \mathbb{R}^2\}$, wobei $g(s)$ eine monoton wachsende stetige Funktion ist, anwenden. \diamond

Als letztes wollen wir noch ein „Wahrscheinlichkeitspapier“, und zwar das für Extremwertverteilungen betrachten.

Beispiel 6.2 Sei $\mathfrak{F}_{w,2} = (F(\cdot) : \bar{F}(s) = \exp(-(\frac{s}{\alpha})^\beta), s > 0, \alpha > 0, \beta > 0)$. Die Überlebensfunktion kann in der Form

$$\bar{F}(s) = \exp(-\exp(\beta(\ln s - \ln \alpha)))$$

geschrieben werden. Ist nun $\bar{F}_0(s)$ die Funktion $\bar{F}_0(s) = \exp(-\exp(s))$, $s \in \mathbb{R}^1$ und betrachten wir die beiden Parameter $\mu = \ln \alpha$ und $\sigma = 1/\beta$, so erhalten wir die Familie

$$\mathfrak{F}_{F_0(\cdot), \ln(\cdot)} = \left(F(\cdot) : \bar{F}_0\left(\frac{\ln s - \mu}{\sigma}\right), s > 0, \alpha > 0, \beta > 0 \right).$$

Sei w_p ein Quantil, das durch $\bar{F}_0(w_p) = p$ definiert ist, d.h. $\exp(-\exp(w_p)) = p$. Dann erhält man $w_p = \ln \ln \frac{1}{p}$. Die Abbildung $(s, p) \rightarrow (\ln s, \ln \ln \frac{1}{p})$ liefert dann wieder das gewünschte Ergebnis: die Kurve $(s, \bar{F}_0(\frac{\ln s - \mu}{\sigma}))$ wird in die Gerade $(t, \frac{t - \mu}{\sigma})$ abgebildet. Diese Abbildung ist wieder lokal stetig und die Punkte der empirischen Verteilungsfunktion $(\ln s_{(i,n)}, \ln \ln(1/(i/n)))$ weichen nur wenig von der Geraden ab, wenn die Hypothese einer Weibullverteilung wahr ist. Damit hat man auch hier wieder eine Entscheidungsregel, bei der man allerdings wieder auf Intuition angewiesen ist. \diamond

In vielen Statistik-Programmpaketen findet man graphische Methoden in Form sogenannter *QQ-plots*, wobei *QQ* für „Quantil-Quantil“ steht. Die Idee dieser *QQ-plots* besteht in folgendem. Wir führen eine Folge von statistischen Versuchen durch und erhalten die Folge statistischer Daten $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n, \dots$. Jeder dieser Vektoren kann z.B. das Resultat einer Lebensdauerprüfung nach dem Stichprobenplan $[m, O, r]$ sein. Für jede der Stichproben wird der Wert einer Statistik $T(\cdot)$ berechnet, die bei Richtigkeit einer uns interessierenden Hypothese asymptotisch für $n \rightarrow \infty$ einer bekannten Verteilung unterliegt, z.B. der Standardnormalverteilung oder einer Exponentialverteilung mit bekanntem Parameter λ_0 . Die Werte der Statistik $T_i = T(x_i)$, $i = 1, \dots, n$ werden geordnet und die Punkte $(u_{k/(n+1)}, T_{(k)})$ werden im *QQ-plot* dargestellt, wobei $u_{k/(n+1)}$ das Quantil zur Ordnung $k/(n+1)$ der bekannten Grenzverteilung (im Falle der Exponentialverteilung z.B. $u_{k/(n+1)} = -\frac{1}{\lambda_0} \ln(1 - \frac{k}{n+1})$) ist. Zwei typische *QQ-plots* aus BELYAEV, LINDKVIST sind für den Fall einer asymptotischen Standardnormalverteilung in Abbildung 6.3 dargestellt.

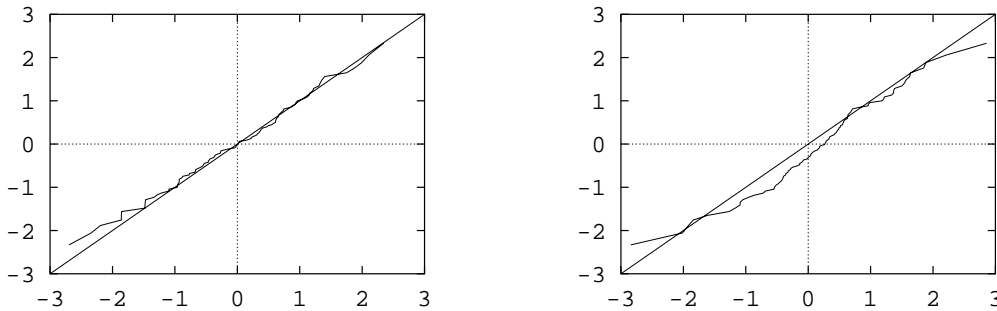


Abbildung 6.3: Zwei *QQ-plots*

Wir bemerken noch, daß diese Herangehensweise auch auf zensierte Daten anwendbar ist, wenn man für diese Daten eine konsistente Punktschätzung $\hat{F}_n(\cdot)$ findet.

6.2 Signifikanztests

Im letzten Abschnitt wurden graphische Verfahren betrachtet, mit denen Hypothesen über den Typ der Lebensdauerverteilung geprüft werden konnten. Diese Verfahren, mit denen man sich schnell einen Überblick über den Typ der Verteilung verschaffen konnte, hatten jedoch den wesentlichen Nachteil, daß man dabei auf Erfahrung und Intuition angewiesen war. Es fehlte vor allem an einer klar formulierten Entscheidungsregel, mit der man sich für oder gegen eine Hypothese entscheidet. Daher wenden wir uns in diesem Abschnitt den Signifikanztests zu, für die unter bestimmten Bedingungen klare Entscheidungsregeln existieren. Der Nachteil dieser Signifikanztests wird etwas später diskutiert.

Nehmen wir an, daß die Hypothese mathematisch exakt formuliert werden kann. Z.B. könnte die Hypothese lauten $H_0 : \theta^* \in \Theta_0$, d.h. der unbekannte wahre Parameter liegt im Bereich Θ_0 .

Wir haben bereits erwähnt, daß mit dieser Formulierung auch nichtparametrische Probleme erfaßt werden können. Nach Durchführung des statistischen Experimentes erhält man die Daten x und man möchte aufgrund dieser Daten mit großer Wahrscheinlichkeit zu einer richtigen Entscheidung kommen. Diese große Wahrscheinlichkeit bezeichnen wir mit $(1 - \alpha)$, wobei für α – die sogenannte Irrtumswahrscheinlichkeit – häufig $\alpha = 0.1$; $\alpha = 0.05$ oder $\alpha = 0.01$ verwendet wird.

Desweiteren sei T eine Statistik $T : \mathfrak{X} \rightarrow \mathfrak{T}$, deren Verteilung für $\theta^* \in \Theta_0$ (also bei Gültigkeit der zu prüfenden Hypothese) unabhängig von θ^* ist. Sei P_0 diese Verteilung der Zufallsgröße T . Nun definieren wir eine Teilmenge $B_0 = B_0(\alpha)$ so, daß

$$P_0(T \in B_0) \leq \alpha .$$

Ist die Hypothese richtig, so ist also die Wahrscheinlichkeit, daß T Werte aus B_0 annimmt, klein. Das führt dazu, daß bei $T \in B_0$ die Hypothese abgelehnt wird, da die Daten der Hypothese signifikant widersprechen. Diese Tests nennt man Signifikanztests. Dabei ist α die Wahrscheinlichkeit, eine Hypothese abzulehnen, obwohl sie richtig ist. Dieser mögliche Fehler heißt *Fehler erster Art*. Somit ist α die Wahrscheinlichkeit, einen Fehler erster Art zu begehen. Die Statistik T heißt *Testgröße* oder *Prüfgröße* und B_0 heißt *kritischer Bereich*.

Beispiel 6.3 Wir betrachten nun die Durchführung eines Tests am Beispiel des Parameters λ der Exponentialverteilung.

Die Zufallsgröße X sei exponentialverteilt mit dem Parameter λ . Es liegen Beobachtungen $T_{(1)}, \dots, T_{(r)}$ nach dem Stichprobenplan $[n, O, r]$ vor. Die Likelihoodfunktion ist für diesen Fall :

$$L(\lambda, T_{(1)}, \dots, T_{(r)}) = \binom{n}{r} \prod_{i=1}^r \lambda e^{-\lambda T_i} (e^{-\lambda T_{(r)}})^{N-r}$$

und als Punktschätzung nach der Likelihood-Methode erhält man

$$\check{\lambda} = \frac{r}{\sum_{i=1}^r T_i + (N - r)T_{(r)}} .$$

Da $S_r = \sum_{i=1}^r T_i + (N - r)T_{(r)}$ einer Erlangverteilung unterliegt, kann diese Größe als Teststatistik verwendet werden.

Ein Test wird nach folgenden Punkten durchgeführt:

1. Aufstellen der Nullhypothese $H_0 : \lambda = \lambda_0$.
2. Vorgabe einer Irrtumswahrscheinlichkeit α (z.B. $\alpha = 0.01, 0.05, 0.1$).
3. Wahl einer geeigneten Prüfgröße T .
Diese Prüfgröße ist eine Funktion der Stichprobe und ihre Verteilung ist vollständig bekannt. In unserem Fall kann gezeigt werden, daß die Größe

$$2r \frac{\lambda_0}{\lambda} = 2\lambda_0 \underbrace{\left(\sum_{i=1}^r T_i + (N - r)T_{(r)} \right)}_{\substack{\text{total} \quad \text{time} \quad \text{of} \quad \text{test}}}$$

einer χ^2 -Verteilung mit $2r$ Freiheitsgraden unterliegt.

4. Ermittlung des kritischen Bereiches B_0 aus der Beziehung $P_{\lambda_0}(T \in B_0 | H_0) = \alpha$.
Je nach ein- oder zweiseitiger Fragestellung kann ein Test aus

$$P_{\lambda_0}(2\lambda_0 S_r \in B_0) = \alpha$$

als $B_0 = (\chi_{1-\alpha}^2(2r), \infty)$ oder $B_0 = ([0, \chi_{\frac{\alpha}{2}}^2(2r)] \cup [\chi_{1-\frac{\alpha}{2}}^2(2r), \infty])$ bestimmt werden, wobei $\chi_{\alpha}^2(2r)$ das Quantil der χ^2 -Verteilung mit $2r$ Freiheitsgraden zur Ordnung α ist.

5. Als letztes wird eine Realisierung von $2\lambda_0 S_r$ berechnet und die Hypothese abgelehnt, wenn diese Realisierung in B_0 liegt.

◇

Diese Art Test ist ein Signifikanztest. Seine Durchführung ist nur möglich, wenn man eine geeignete Prüfgröße findet, deren Verteilung vollständig bekannt ist. Bisher haben wir nur über die zu prüfende Nullhypothese gesprochen. Es ist jedoch natürlich, neben dieser Nullhypothese eine Alternativhypothese H_1 zu betrachten, $H_1 : \theta^* \in \Theta_1, \Theta_1 \cap \Theta_0 = \emptyset$. Nehmen wir an, daß H_1 richtig ist. Dann ist es möglich, daß die Realisierung der Prüfgröße nicht im kritischen Bereich B_0 liegt, die Nullhypothese also nicht abgelehnt wird, obwohl H_1 richtig ist. Diesen Fehler nennt man *Fehler zweiter Art*. Die Wahrscheinlichkeit eines Fehlers zweiter Art bezeichnet man gewöhnlich mit β . Im allgemeinen hängt β von $\theta^* \in \Theta_1$ ab.

Bei den oben angeführten Punkten der Konstruktion eines Signifikanztests hatten wir keine Möglichkeit, Einfluß auf den Fehler zweiter Art zu nehmen. Im allgemeinen ist β umso größer, je kleiner α ist.

Wenn es mehrere Möglichkeiten für die Konstruktion eines Tests gibt, so wird man natürlich den verwenden, der bei gegebenem α ein kleineres β besitzt. Unter den Signifikanztests gibt es *beste (mächtigste) Tests*, d.h. Tests, deren Wahrscheinlichkeit eines Fehlers zweiter Art minimal wird. Nach dem Lemma von Neyman und Pearson (siehe z.B. RAO oder LEHMANN) besteht ein solcher bester Test für einfache Hypothesen $H_0 : \theta^* = \theta_0$ und $H_1 : \theta^* = \theta_1$ in folgendem:

Als Prüfgröße wird der Likelihoodquotient

$$L(\theta_1, X)/L(\theta_0, X)$$

gewählt. Der kritische Bereich wird aus der Beziehung

$$P_{\theta_0} \left(X : \frac{L(\theta_1, X)}{L(\theta_0, X)} \geq \delta, x \in \mathfrak{X} \right) = \alpha$$

bestimmt. Das bedeutet, man muß δ so bestimmen, daß die Wahrscheinlichkeit des Ereignisses $\{L(\theta_1, X)/L(\theta_0, X) \geq \delta\}$ gleich der vorgegebenen Irrtumswahrscheinlichkeit α ist.

Beispiel 6.4 Wir betrachten wieder die Exponentialverteilung und den Stichprobenplan $[n, O, r]$. Für den Likelihoodquotienten erhält man

$$\begin{aligned} \frac{L(\theta_1, T_{(1)}, \dots, T_{(r)})}{L(\theta_0, T_{(1)}, \dots, T_{(r)})} &= \frac{\prod_{i=1}^r \lambda_1 e^{-\lambda_1 T_i} [e^{-\lambda_1 T_{(r)}}]^{N-r}}{\prod_{i=1}^r \lambda_0 e^{-\lambda_0 T_i} [e^{-\lambda_0 T_{(r)}}]^{N-r}} \\ &= \left(\frac{\lambda_1}{\lambda_0} \right) e^{-(\lambda_1 - \lambda_0)(\sum_{i=1}^r T_i + T_{(r)}(N-r))} . \end{aligned}$$

Hier sieht man, daß der Likelihoodquotient die gleiche Teststatistik liefert, wie im Beispiel 6.3. Hieraus folgt, daß der dort vorgeschlagene Test ein bester Test ist. \diamond

An diesen Beispielen sieht man, daß die Durchführung eines Signifikanztestes vor allem zwei Schwierigkeiten bereitet : Man muß eine Prüfgröße T finden, deren Verteilung bei richtiger Nullhypothese vollständig bekannt ist und man muß einen kritischen Bereich B_0 so wählen, daß die Irrtumswahrscheinlichkeit eingehalten wird. Wir wollen diese Frage nun theoretisch etwas mehr durchleuchten.

Sei $H_1 : \theta^* \in \Theta_1$ eine Alternativhypothese. Wenn nun aus der Beobachtung die Daten $X_n = (X_1, \dots, X_n)$ vorliegen und $n \rightarrow \infty$, so wollen wir natürlich fordern, daß eine richtige Entscheidung getroffen wird. Das heißt, wenn $\Theta_1 \cap \Theta_0 = \emptyset$ und $\theta^* \in \Theta_1$, so soll die Wahrscheinlichkeit H_0 abzulehnen für $n \rightarrow \infty$ gegen 1 konvergieren.

Definition 6.1 Ein Test (T, B_0) heißt (Θ_0, Θ_1) - konsistent, wenn für $\theta^* \in \Theta_1$

$$\lim_{n \rightarrow \infty} P_{\theta}(T(x_n) \in B_0) = 1 \quad \forall \theta \in \Theta_1 .$$

Ein (Θ_0, Θ_1) - konsistenter Test (T, B_0) heißt Signifikanztest zum Signifikanzniveau $(1 - \alpha)$, wenn

$$P_{\theta}(T(x_n) \in B_0) \leq \alpha, \quad \theta \in \Theta_0 .$$

Wir betrachten nun konsistente Tests an einer Reihe von nichtparametrischen Aufgabenstellungen.

Beispiel 6.5 *Der Kolmogorov - Smirnov - Test.*

Bei diesem Test sind zwei Stichproben $Y(n_1) = (Y_1, \dots, Y_{n_1})$ und $Z(n_2) = (Z_1, \dots, Z_{n_2})$ gegeben. Es kann sich hierbei z.B. um 2 Temperaturmessungen handeln und man möchte die Frage entscheiden, ob zwischen den beiden Stichproben ein Unterschied besteht oder ob sich die unterschiedlichen Meßwerte aus Zufallsabweichungen erklären lassen.

Sei $X(n)$ die Stichprobe $(Y_1, \dots, Y_{n_1}, Z_1, \dots, Z_{n_2})$, $n = n_1 + n_2$. Wir bezeichnen mit $F_1(s) = P(Y_i \leq s)$ und $F_2(s) = P(Z_i \leq s)$ die Verteilungen der beiden Stichproben und nehmen an, daß es sich um stetige Verteilungen handelt. Der zu testende Parameter besteht aus diesen beiden Verteilungsgesetzen $\theta = (F_1(\cdot), F_2(\cdot))$. Wenn die Nullhypothese darin besteht, daß die beiden Stichproben der gleichen Verteilung unterliegen, so erhalten wir $\Theta_0 = \{(F_1(\cdot), F_2(\cdot)) : F_1(\cdot) = F_2(\cdot)\}$ und $\Theta_1 = \{(F_1(\cdot), F_2(\cdot)) : F_1(\cdot) \neq F_2(\cdot)\}$. Mit $\hat{F}_{1,n_1}(\cdot)$ und $\hat{F}_{2,n_2}(\cdot)$ bezeichnen wir die entsprechenden empirischen Verteilungsfunktionen

$$\begin{aligned}\hat{F}_{1,n_1}(s) &= \frac{1}{n_1} \sum_{i=1}^{n_1} I(Y_i \leq s) \quad \text{und} \\ \hat{F}_{2,n_2}(s) &= \frac{1}{n_2} \sum_{i=1}^{n_2} I(Z_i \leq s) .\end{aligned}$$

Wir betrachten nun die Teststatistik

$$T(x(n)) = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \sup_s |\hat{F}_{1,n_1}(s) - \hat{F}_{2,n_2}(s)| .$$

Wenn die Nullhypothese gültig ist, so ist die Verteilung dieser Größe unabhängig von $F(\cdot) = F_1(\cdot) = F_2(\cdot)$. Daher können wir den kritischen Bereich $B_0 = ((-\infty, -t_{n,\alpha}) \cup (t_{n,\alpha}, \infty))$ aus

$$P_{(F_1(\cdot), F_2(\cdot))}(|T(x(n))| > t_{n,\alpha}) = \alpha$$

bestimmen.

Da $|T(X(n))| \rightarrow \infty$ für $n \rightarrow \infty$ gilt, erhalten wir außerdem

$$\lim_{\min(n_1, n_2) \rightarrow \infty} P_{(F_1(\cdot), F_2(\cdot))}(|T(x(n))| > c) = 1 .$$

Damit erhalten wir einen konsistenten Test bei beliebiger Alternativhypothese. \diamond

Der Test aus Beispiel 6.5 kann auch auf den Stichprobenplan $[n, O, T]$ verallgemeinert werden, wenn T für beide Stichproben gleich ist.

Beispiel 6.6 In diesem Beispiel betrachten wir den Fall, daß die Nullhypothese nur eine einzige Verteilungsfunktion $H_0 : F(\cdot) = F_0(\cdot)$ enthält (einfache Nullhypothese). Wir

nehmen wieder an, daß $F_0(\cdot)$ stetig ist. Damit ist $\Theta_0 = \{\theta_0\}$, $\theta_0 = F_0(\cdot)$. Als Alternative sollen alle anderen Verteilungsfunktionen in Betracht kommen:

$$\Theta_1 = \{F(\cdot) : F(\cdot) \neq F_0(\cdot)\} .$$

Die Untersuchungen werden nach dem Stichprobenplan $[n, O, n]$ durchgeführt. Damit erhält man die statistischen Daten $X(n) = (X_1, \dots, X_n)$, $X_i \in \mathbb{R}^1$.

Die Statistik $T_n(X(n)) = \sqrt{n} \sup_s |\hat{F}_n(s) - F_0(s)|$ mit $\hat{F}_n(s) = \frac{1}{n} \sum_{i=1}^n \mathbf{I}(X_i \leq s)$ – empirische Verteilungsfunktion – ist unabhängig von F_0 , wenn die Hypothese richtig ist: $F^*(\cdot) = F_0(\cdot)$. Sie kann daher als Teststatistik verwendet werden.

Die Verteilung von $T_n(X(n))$:

$$K_n(y) = P_{F_0(\cdot)}(\sqrt{n} \sup_s |\hat{F}_n(s) - F_0(s)| \leq y)$$

heißt Kolmogorov-Verteilung. Für diese Verteilung existieren Tabellen der Quantile $k(n, p)$ (z.B. in SMIRNOV, DUNIN-BARKOVSKI):

$$K_n(k(n, p)) = p .$$

Damit kann das kritische Niveau aus

$$P_{F_0(\cdot)}(\sqrt{n} \sup_s |\hat{F}_n(s) - F_0(s)| > k(n, 1 - \alpha)) = \alpha$$

bestimmt werden.

Wenn $n \rightarrow \infty$ und $F^*(\cdot) \neq F_0(\cdot)$, so gilt

$$P_{F^*(\cdot)}(\sqrt{n} \sup_s |\hat{F}_n(s) - F_0(s)| > k(n, 1 - \alpha)) \rightarrow 1$$

und wir haben auch hier wieder einen (Θ_1, Θ_2) - konsistenten Signifikanztest zum Signifikanzniveau $(1 - \alpha)$. \diamond

Der in diesem Beispiel betrachtete Fall einer einfachen Nullhypothese ist gerade der einfachste Fall. Wesentlich interessanter (und praktisch relevanter) ist es, wenn die Nullhypothese eine Familie von Verteilungen enthält, wenn man z.B. entscheiden will, ob eine Verteilung zur Familie der Weibullverteilungen gehört. Diese Fragestellung betrachten wir im nächsten Beispiel.

Beispiel 6.7 Sei $\Theta_0 = \mathfrak{F}_{w,2}$, d.h. die Nullhypothese besteht darin, daß die betrachtete Grundgesamtheit einer zweiparametrischen Weibullverteilung unterliegt: $\bar{F}(s, a, b) = e^{-(s/a)^b}$. Die Parameter sind hier nicht weiter spezifiziert. Die Beobachtung wird wieder nach dem Stichprobenplan $[n, O, n]$ durchgeführt, d.h. wir erhalten die Daten $X(n) = (S_1, \dots, S_n)$. Mit \hat{a}_n und \hat{b}_n bezeichnen wir die Maximum-Likelihood-Punktschätzungen der beiden unbekannten Parameter. Betrachten wir nun wieder die Statistik

$$T_n(X(n)) = \sqrt{n} \sup_s |\hat{F}_n(s) - F(s, \hat{a}_n, \hat{b}_n)| .$$

Man kann beweisen, daß für $n \rightarrow \infty$ die Verteilung von $T_n(X(n))$ unabhängig von den wahren Parametern a^* und b^* ist, wenn die wahre Verteilung $F^*(s) = F(s, a^*, b^*) \in \mathfrak{F}_{w,2}$. Damit ist es auch hier möglich, asymptotisch das kritische Niveau $k(n, 1 - \alpha)$ zu bestimmen:

$$\lim_{n \rightarrow \infty} P_{F(\cdot, a^*, b^*)}(\sqrt{n} \sup_s |\hat{F}_n(s) - F(s, \check{a}_n, \check{b}_n)| > k(n, 1 - \alpha)) = \alpha.$$

Weiterhin kann man beweisen, daß für $F^*(\cdot) \notin \mathfrak{F}_{w,2}$

$$\lim_{n \rightarrow \infty} P_{F^*(\cdot)}(\sqrt{n} \sup_s |\hat{F}_n(s) - F(\check{a}_n, \check{b}_n)| > k(n, 1 - \alpha)) = 1.$$

Somit erhält man auch in diesem Fall einen $(\Theta_0 = \mathfrak{F}_{w,2}, \Theta_1 = \mathfrak{F} \setminus \mathfrak{F}_{w,2})$ -konsistenten Test. Wir bemerken jedoch, daß man hier einen kritischen Bereich nur asymptotisch bestimmen kann. In diesem Fall existieren auch keine Tabellen und man ist für die Bestimmung des kritischen Bereiches auf Simulationsuntersuchungen angewiesen. \diamond

Auf die oben gezeigte Art und Weise können verschiedene Signifikanztests für die unterschiedlichen Aufgabenstellungen konstruiert werden.

In den letzten Beispielen war es teilweise nicht möglich, die Verteilung der Testgröße exakt zu bestimmen. Einen Ausweg bieten asymptotische Methoden, die jetzt vorgestellt werden. Wir führen nun folgende Regularitätsvoraussetzungen R1 und RI ein:

- R1: (i) Θ ist eine offene Teilmenge des \mathbb{R}^m mit der gewöhnlichen Euklidischen Metrik und die Menge aller $x \in \mathfrak{X}$, für die die Likelihoodfunktion definiert ist, hängt nicht vom unbekannten Parameter $\theta \in \Theta$ ab.
- (ii) Es existieren die partiellen Ableitungen $\frac{\partial l_1(\theta, x)}{\partial \theta_i}$, $i = 1, \dots, m$ für die logarithmierte Likelihoodfunktion einer Beobachtung $l_1(\theta, x) = \ln L_1(\theta, x)$, $x \in \mathfrak{X}$.
- (iii) $E_\theta \left(\frac{\partial l_1(\theta, X_j)}{\partial \theta_i} \right)^2 < \infty \quad \forall \theta \in \Theta, i = 1, \dots, m.$
- (iv) Für eine beliebige Statistik $T_n = T_n(\underline{X}(n))$, für die $E_\theta |T_n(\underline{X}(n))|^2 < \infty$ und $E_\theta T_n(\underline{X}(n)) = g(\theta)$ gilt, ist

$$E_\theta \left(T_n(\underline{X}(n)) \frac{\partial l(n, \theta, \underline{X}(n))}{\partial \theta_i} \right) = \frac{\partial g(\theta)}{\partial \theta_i}, \quad i = 1, \dots, m,$$

wobei

$$l(n, \theta, \underline{X}(n)) = \sum_{j=1}^n l_1(\theta, X_j) \quad (6.4)$$

die logarithmierte Likelihoodfunktion für die Beobachtung aus n statistischen Experimenten ist.

RI: Die Fishersche Informationsmatrix für eine Beobachtung

$$\mathbb{I}_1(\theta) = \left(E_\theta \frac{\partial l_1(\theta, X_j)}{\partial \theta_i} \frac{\partial l_1(\theta, X_j)}{\partial \theta_k} \right), \quad i, k = 1, \dots, m$$

besitzt den vollen Rang, d.h. es gilt $\det \mathbb{I}_1(\theta) > 0$.

Bemerkung: Die Regularitätsvoraussetzung R1(iv) ersetzt die übliche Voraussetzung, beim Erwartungswert Differentiation und Integration vertauschen zu können.

Neben den Regularitätsvoraussetzungen R1, die Bedingungen an die ersten Ableitungen der Likelihoodfunktion stellen und der Voraussetzung RI, die gewährleistet, daß die Fishersche Informationsmatrix nicht entartet, führen wir nun noch Regularitätsvoraussetzungen R2 an die zweiten partiellen Ableitungen der Likelihoodfunktion ein.

R2: (i) Es existieren alle partiellen Ableitungen zweiter Ordnung

$$\frac{\partial^2 l_1(\theta, x)}{\partial \theta_i \partial \theta_k}, \quad i, k = 1, \dots, m.$$

(ii)

$$E_\theta \left| \frac{\partial^2 l_1(\theta, X_j)}{\partial \theta_i \partial \theta_k} \right| < \infty, \quad i, k = 1, \dots, m.$$

(iii)

$$E_\theta \frac{\partial l_1(\theta, X_j)}{\partial \theta_i} \frac{\partial l_1(\theta, X_j)}{\partial \theta_k} = -E_\theta \frac{\partial^2 l_1(\theta, X_j)}{\partial \theta_i \partial \theta_k}, \quad i, k = 1, \dots, m \quad \forall \theta \in \Theta.$$

(iv) Die zweiten partiellen Ableitungen sind stetig und für den lokalen Stetigkeitsmodul

$$w_{i,k,\varepsilon}(\theta, x) = \sup_{\theta' \in B_\varepsilon(\theta)} \left(\frac{\partial^2 l_1(\theta, x)}{\partial \theta_i \partial \theta_k} - \frac{\partial^2 l_1(\theta', x)}{\partial \theta_i \partial \theta_k} \right)$$

mit $x \in \mathfrak{X}$, $\theta \in \Theta$ und $B_\varepsilon(\theta) = \{\theta'' : \|\theta'' - \theta\| < \varepsilon, \theta'' \in \Theta\}$ gilt:

Für beliebige $\theta \in \Theta$ und beliebige $\delta > 0$ findet man so ein $\varepsilon = \varepsilon(\delta, \theta)$, daß

$$E_\theta w_{i,k,\varepsilon}(\theta, X_j) \leq \delta^2.$$

Satz 6.1 Sei $(\check{\theta}_n)_{n \geq 1}$ eine Folge konsistenter Maximum-Likelihood-Schätzungen für den wahren Parameter θ^* . Die Regularitätsvoraussetzungen R1, R2 und RI seien erfüllt. Dann ist die Folge $(\check{\theta}_n)_{n \geq 1}$ auch \sqrt{n} -konsistent, \sqrt{n} -asymptotisch normalverteilt und für den Abstand vom wahren Parameter θ^* gilt für $n \rightarrow \infty$

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathcal{L}(\sqrt{n}(\check{\theta}_n(\underline{X}(n)) - \theta^*)) = \\ & = \lim_{n \rightarrow \infty} \mathcal{L}\left(\mathbb{I}_1^{-1}(\theta^*) \frac{\nabla l(n, \theta^*, \underline{X}(n))}{\sqrt{n}}\right) = \Phi_{\underline{0}_m, \mathbb{I}_1^{-1}(\theta^*)}(\cdot). \end{aligned} \quad (6.5)$$

Die asymptotische Normalverteilung der Maximum-Likelihood-Schätzungen kann verwendet werden, um Parametertests zu konstruieren. Eine andere Herangehensweise ergibt sich aus dem Likelihoodquotienten.

Wenn die Verteilung einer geeigneten Prüfgröße nicht mehr bestimmt werden kann, so wird der Likelihoodquotient selbst als Prüfgröße verwendet und seine asymptotische Verteilung benutzt. Hier gilt für unabhängige und identisch verteilte Beobachtungen $X(n) = (X_1, \dots, X_n)$ unter den Regularitätsvoraussetzungen R1, RI und R2:

$$\lim_{n \rightarrow \infty} \mathcal{L}(2(\ln L(\check{\theta}_n, X(n)) - \ln L(\theta^*, X(n)))) = \chi^2(k)$$

wobei k die Dimension des interessierenden Parameters θ^* und $\check{\theta}_n$ seine Maximum-Likelihood-Schätzung sind.

Dieses Ergebnis läßt sich folgendermaßen verallgemeinern. Wenn der Parameter $\theta = (\theta_1, \dots, \theta_k, \theta_{k+1}, \dots, \theta_m) \in \Theta$ aus den k interessierenden Komponenten $\theta_1, \dots, \theta_k$ und den $(m-k)$ störenden (nuisance) Komponenten $\theta_{k+1}, \dots, \theta_m$ besteht, so entspricht der Nullhypothese H_0 eine Untermenge $\Theta_0 = ((\theta_1^0, \dots, \theta_k^0, \theta_{k+1}, \dots, \theta_m))$, in der die störenden Komponenten $\theta_{k+1}, \dots, \theta_m$ beliebige Werte so annehmen können, daß $(\theta_1^0, \dots, \theta_k^0, \theta_{k+1}, \dots, \theta_m) \in \Theta$ ist. In diesem Fall unterliegt der logarithmierte Likelihoodquotient

$$2(\ln L(\hat{\theta}_n, X(n)) - \ln L(\hat{\theta}_n^0, X(n)))$$

asymptotisch für $n \rightarrow \infty$ einer χ^2 -Verteilung mit k Freiheitsgraden, wobei $\hat{\theta}_n$ die Maximum-Likelihood-Schätzung aller Komponenten und $\hat{\theta}_n^0$ die Maximum-Likelihood-Schätzung der störenden Komponenten für $\theta^* \in \Theta_0$ sind.

Beispiel 6.8 Wir betrachten eine Weibull-Verteilung mit dem Parameter $\theta = (a, b)$ bei vollständiger Beobachtung aller Ausfallszeiten x_i , $i = 1, \dots, n$ (dieser Weg läßt sich bei zensierten Daten völlig analog durchführen). Es soll bei hinreichend großem Stichprobenumfang n die Hypothese $b = 1$ gegen $b > 1$ getestet werden. Der Parameter a ist im Modell vorhanden, jedoch für die Fragestellung nicht von Interesse.

Dann erhält man für $x(n) = (x_1, \dots, x_n)$:

$$\begin{aligned} & 2(\ln L(\check{\theta}, x(n)) - \ln L(\underline{\theta}_0, x(n))) = \\ & = 2(\ln L(\check{a}_n, \check{b}_n, x(n)) - \ln L(\check{a}_n^0, b_0, x(n))) = \\ & = 2(n \ln \check{b}_n + (\check{b}_n - 1) \sum_{i=1}^n \ln x_i - n \check{b}_n \ln \check{a}_n - \sum_{i=1}^n \left(\frac{x_i}{\check{a}_n}\right)^{\check{b}_n} - \\ & \quad - n \ln b_0 - (b_0 - 1) \sum_{i=1}^n \ln x_i + n b_0 \ln \check{a}_n^0 + \sum_{i=1}^n \left(\frac{x_i}{\check{a}_n^0}\right)^{b_0}, \end{aligned}$$

wobei $\check{a}_n^0 = \left(\frac{1}{n} \sum_{i=1}^n x_i^{b_0}\right)^{\frac{1}{b_0}}$ ist.

Ist diese Größe größer als das Quantil der χ^2 -Verteilung mit 1 Freiheitsgrad, so besteht Grund, die Hypothese $\beta = 1$ abzulehnen. \diamond

6.3 Konfidenzschätzungen

Wir wollen zunächst die asymptotische Normalverteilung bzw. die asymptotische χ^2 -Verteilung des Likelihoodquotienten anwenden, um im Fall von Beobachtungen nach dem Stichprobenplan $[n, E, T_i]$ Konfidenzschätzungen zu bestimmen. Zunächst muß für die Likelihoodfunktion (5.9) die Fishersche Information berechnet werden. Wir setzen voraus, daß die Dichte $f_X(x_{ij})$ dreimal stetig differenzierbar bezüglich aller Parameter

ist. Dann gilt für die i -te Realisierung des Erneuerungsprozesses:

$$\begin{aligned}\ln L_i(F(\cdot), x) &= \sum_{j=1}^{d_i} \ln f_X(x_{ij}) + \ln(\bar{F}_X(r^E(T_i))) , \\ \frac{\partial \ln L_i(F(\cdot), x)}{\partial \theta_k} &= \sum_{j=1}^{d_i} \frac{\partial \ln f_X(x_{ij})}{\partial \theta_k} + \frac{\partial \ln(\bar{F}_X(r^E(T_i)))}{\partial \theta_k} , \\ \frac{\partial^2 \ln L_i(F(\cdot), x)}{\partial \theta_k \partial \theta_l} &= \sum_{j=1}^{d_i} \frac{\partial^2 \ln f_X(x_{ij})}{\partial \theta_k \partial \theta_l} + \frac{\partial^2 \ln \bar{F}_X(r^E(T_i))}{\partial \theta_k \partial \theta_l} .\end{aligned}$$

Somit werden zur Berechnung der Fisherschen Information Größen der Art

$$E_1 := E \left(\sum_{j=1}^{N_i^E} l^{(2)}(X_{ij}) \right) , \quad E_2 := E (L^{(2)}(r^E(T_i))) \quad (6.6)$$

benötigt, wobei

$$l^{(2)}(X_{ij}) = -\frac{\partial^2 \ln f_X(X_{ij})}{\partial \theta_k \partial \theta_l}, \quad L^{(2)}(r^E(T_i)) = -\frac{\partial^2 \ln(\bar{F}_X(r^E(T_i)))}{\partial \theta_k \partial \theta_l} . \quad (6.7)$$

Die Berechnung dieser Größen ist eng mit der Erneuerungsfunktion verbunden. Sei

$$H^E(t) = E(N_i^E(t)) = \sum_{k=0}^{\infty} k P(N_i^E t) = k)$$

die Erneuerungsfunktion der betrachteten Grundgesamtheit mit

$$\begin{aligned}P(N_i^E t) = k) &= P \left(\sum_{j=1}^k X_{ij} \leq t < \sum_{j=1}^{k+1} X_{ij} \right) = P(S_{ik} \leq t < S_{i,k+1}) = \\ &= \int \dots \int_{s_{ik} \leq t < s_{i,k+1}} f_X(x_{i1}) \dots f_X(x_{ik+1}) dx_{i1} \dots dx_{ik+1} .\end{aligned}$$

mit $s_{ik} = \sum_{j=1}^k x_{ij}$, $k = 1, 2, \dots$, $s_0 = 0$.

Zuerst gilt für den Erwartungswert der Summe einer zufälligen Anzahl von zufälligen Größen

$$E \left(\sum_{j=1}^{N_i} l^{(2)}(X_{ij}) \right) = \sum_{k=0}^{\infty} E \left(\sum_{j=1}^k l^{(2)}(X_{ij}) I(N_i^E = k) \right) .$$

Hierbei ist $I(\cdot)$ die Indikatorfunktion und $N_i^E = N_i^E(T_i)$ ist die zufällige Anzahl der Erneuerungen bis zur Zeit T_i .

Nun gilt für einen Ausdruck der Art $E(l^{(2)}(X_{ij}) I(N_i^E = k))$:

$$E(l^{(2)}(X_{ij}) I(N_i^E = k)) = \int \dots \int_{s_{ik} \leq T_i < s_{i,k+1}} l^{(2)}(x_{ij}) f_X(x_{i1}) \dots f_X(x_{ik+1}) dx_{i1} \dots dx_{ik+1}$$

mit $s_{ik} = x_{i1} + \dots + x_{ik}$ – Sprungzeitpunkte.

Da alle X_{ij} identisch verteilt sind, kann man ohne Beschränkung der Allgemeinheit $j = 1$ betrachten. Dann erhält man:

$$\begin{aligned} & \mathbb{E}(l^{(2)}(X_{i1})\mathbb{I}(N_i^E = k)) = \\ &= \int_0^{T_i} l^{(2)}(x_{i1}) f_X(x_{i1}) \int_A \dots \int f_X(x_{i2}) \dots f_X(x_{i,k+1}) dx_{i1} \dots dx_{i,k+1} = \\ &= \int_0^{T_i} l^{(2)}(x_{i1}) \mathbb{P}(N_i^E(T_i - s_{i1}) = k - 1) f_X(x_{i1}) dx_{i1} \end{aligned}$$

mit $A = (s_{ik} - s_{i1} \leq T_i - s_{i1} < s_{i,k+1} - s_{i1})$. Damit erhält man für E_1 :

$$\begin{aligned} E_1 &= \sum_{k=1}^{\infty} k \cdot \mathbb{E}(l^{(2)}(X_{ij})\mathbb{I}(N_i^E = k)) = \\ &= \int_0^{T_i} l^{(2)}(x_{i1}) \sum_{k=1}^{\infty} k \mathbb{P}(N_i^E(T_i - s_{i1}) = k - 1) f_X(x_{i1}) dx_{i1} . \end{aligned}$$

Da

$$\begin{aligned} & \sum_{k=1}^{\infty} k \cdot \mathbb{P}(N_i^E(T_i - s_{i1}) = k - 1) = \sum_{k=1}^{\infty} (k - 1) \mathbb{P}(N_i^E(T_i - s_{i1}) = k - 1) + \\ &+ \sum_{k=1}^{\infty} \mathbb{P}(N_i^E(T_i - s_{i1}) = k - 1) = H^E(T_i - x_{i1}) + 1 \end{aligned}$$

gilt, erhält man schließlich für E_1

$$E_1 = \int_0^{T_i} l^{(2)}(x) f_X(x) dx + \int_0^{T_i} l^{(2)}(x) H^E(T_i - x) f_X(x) dx . \quad (6.8)$$

Nun wird die Größe E_2 betrachtet.

$$\begin{aligned} E_2 &= \mathbb{E}(L^{(2)}(r^E(T_i))) = \mathbb{E}(L^{(2)}(T_i - S_{iN_i^E})) = \\ &= \sum_{k=0}^{\infty} \mathbb{E}(L^{(2)}(T_i - S_{ik})\mathbb{I}(N_i^E = k)) = \\ &= \mathbb{E}(L^{(2)}(T_i)\mathbb{I}(X_{i1} > T_i)) + \sum_{k=1}^{\infty} \mathbb{E}(L^{(2)}(T_i - S_{ik})\mathbb{I}(N_i^E = k)) . \end{aligned}$$

Zur Berechnung dieser Erwartungswerte wird die gemeinsame Verteilung der Rückwärtsrekurrenzzzeit eines Erneuerungsprozesses und der Anzahl der Ausfälle benötigt. Diese

gemeinsame Verteilung ist gegeben durch:

$$P(r^E(T_i) < t, N_i^E = k) = \int_{T_i-t}^{T_i} (\bar{F}_X(T_i - z)) dF_k(z)$$

mit $r^E(T_i)$ – Rückwärtsrekurrenzzeit und $F_k(z) = P(S_{ik} < z)$.
Dann gilt

$$\begin{aligned} E_2 &= L^{(2)}(T_i)(\bar{F}_X(T_i)) + \\ &+ \sum_{k=1}^{\infty} \int_0^{T_i} L^{(2)}(T_i - s_{ik}) d_{s_{ik}} P(r^E(T_i) < s_{ik}, N_i^E(t) = k) = \\ &= L^{(2)}(T_i)(\bar{F}_X(T_i)) + \\ &+ \sum_{k=1}^{\infty} \int_0^{T_i} L^{(2)}(T_i - t)(\bar{F}_X(t)) dF_k(T_i - t) = \\ &= L^{(2)}(T_i)(\bar{F}_X(T_i)) + \\ &+ \int_0^{T_i} L^{(2)}(T_i - t)(\bar{F}_X(T_i - t)) \sum_{k=1}^{\infty} dF_k(t) = \\ &= L^{(2)}(T_i)(\bar{F}_X(T_i)) + \int_0^{T_i} L^{(2)}(T_i - t)(\bar{F}_X(T_i - t)) dH^E(t). \end{aligned} \quad (6.9)$$

Damit erhält man für die Information einer Realisierung des Erneuerungsprozesses aus (6.8) und (6.9):

$$\begin{aligned} I_{kl}^{(1)}(T_i) &= \int_0^{T_i} l^{(2)}(x) f_X(x) dx + \int_0^{T_i} l^{(2)}(x) H^E(T_i - x) f_X(x) dx + \\ &+ L^{(2)}(T_i)(\bar{F}_X(T_i)) + \int_0^{T_i} L^{(2)}(T_i - x)(\bar{F}_X(T_i - x)) dH^E(x). \end{aligned} \quad (6.10)$$

Hierbei sind $l^{(2)}(x)$ und $L^{(2)}(x)$ die für die Information benötigten negativen zweiten Ableitungen nach den unbekannten Parametern, die in (6.7) definiert wurden.

Die genauere Berechnung der Information erfordert eine erneuerungstheoretische Analyse der auftretenden Größen bei Kenntnis der konkreten Erneuerungsfunktion. Dieses wird nun am Beispiel der Weibullverteilung demonstriert.

Beispiel 6.9 Wir betrachten wieder eine Weibullverteilte Grundgesamtheit mit

$$F_X(t) = 1 - e^{-(\frac{t}{\alpha})^\beta}, \quad f_X(t) = \beta \alpha^{-\beta} t^{\beta-1} e^{-(\frac{t}{\alpha})^\beta}, \quad t > 0.$$

Die Information für eine Realisierung des Erneuerungsprozesses ist bei unterschiedlichen Abbruchzeiten T_i der Mittelwert der Informationen aus n Realisierungen und kann daher gemäß (6.10) folgendermaßen berechnet werden:

$$I_{kl}^{(1)} = \frac{1}{n} \sum_{i=1}^n I_{kl}^{(1)}(T_i) .$$

Seien die beiden Parameter $\theta_1 = \beta$ und $\theta_2 = \alpha$. Setzt man nun Verteilungsdichte, Verteilungsfunktion und Erneuerungsfunktion der Weibullverteilung in (6.10) ein und benutzt zur Vereinfachung die Beziehungen $E(\partial \ln L_i / \partial \theta_s) = 0$ ($s = 1, 2$), so erhält man folgende Fishersche Information:

$$\begin{aligned} I_{11}^{(1)} &= \frac{1}{n\beta^2} \sum_{i=1}^n \left(E(N_i^E) + E \left(\sum_{j=1}^{N_i^E} \left(\frac{X_{ij}}{\alpha} \right)^\beta \ln^2 \left(\frac{X_{ij}}{\alpha} \right)^\beta \right) + \right. \\ &\quad \left. + E \left(\left(\frac{r^E(T_i)}{\alpha} \right)^\beta \ln^2 \left(\frac{r^E(T_i)}{\alpha} \right)^\beta \right) \right) , \\ I_{12}^{(1)} &= -\frac{1}{n\alpha} \sum_{i=1}^n \left(E(N_i^E) + E \left(\sum_{j=1}^{N_i^E} \ln \left(\frac{x_{ij}}{\alpha} \right)^\beta \right) \right) , \\ I_{22}^{(1)} &= \frac{\beta^2}{n\alpha^2} \sum_{i=1}^n E(N_i^E) . \end{aligned} \quad (6.11)$$

Die weitere Berechnung der in (6.11) auftretenden Größen erfordert die numerische Bestimmung der Erneuerungsfunktion der Weibullverteilung und damit in Zusammenhang stehender Größen. Für praktische Anwendungen können jedoch alle schwer zugänglichen Werte durch ihre Schätzungen ersetzt werden (siehe Beispiel 6.10).

◇

Es soll nun gezeigt werden, wie die asymptotische Normalverteilung der Maximum-Likelihood-Schätzungen und der Likelihoodquotient benutzt werden können, um Konfidenzschätzungen für die interessierenden Parameter zu ermitteln. Sind alle oder mehrere der in der Verteilung auftretenden Parameter von Interesse, so werden gemeinsame (simultane) Konfidenzschätzungen angegeben. Dieses ist bei der Verwendung asymptotischer Methoden möglich, da die gemeinsame asymptotische Verteilung der Maximum-Likelihood-Schätzungen bekannt ist. Es ist außerdem auch sinnvoll, da für praktische Anwendungen die Konfidenzschätzungen der Parameter benutzt werden, um Konfidenzschätzungen für andere interessierende Parameter abzuleiten.

Hat man für jeden Parameter unabhängig von den anderen ein Konfidenzintervall ermittelt, so ist der gemeinsame Konfidenzbereich durch einen Quader gegeben, der durch Hyperebenen im R^p begrenzt wird. Leitet man nun aus diesem Konfidenzbereich Konfidenzschätzungen für interessierende Parameter ab, so hat man einen starken Einfluß der Ecken zu verzeichnen. „Optimistische“ bzw. „pessimistische“ Varianten liegen immer

an einem der Eckpunkte des Quaders, denen jedoch nur geringe Wahrscheinlichkeit zukommt. Außerdem stimmt bei voneinander abhängigen Parameterschätzungen die Überdeckungswahrscheinlichkeit des Quaders nicht mehr mit dem vorgegebenen Konfidenzniveau überein. Wir schlagen deshalb eine alternative Herangehensweise vor: Es werden gemeinsame Konfidenzschätzungen für die Parameter der Lebensdauerverteilung ermittelt.

Zur Konstruktion der Umrandung eines gemeinsamen Konfidenzbereiches für den Parametervektor $\underline{\theta} = (\theta_1, \dots, \theta_m)$ der Verteilung wird eine Niveaulinie der m -dimensionalen Normalverteilung so bestimmt, daß in ihr die Wahrscheinlichkeitsmasse $(1 - \alpha)$ liegt (α -Irrtumswahrscheinlichkeit). Diese Niveaulinie ist durch die quadratische Form

$$(\check{\underline{\theta}} - \underline{\theta})^T \mathbb{I}_{\underline{\theta}}^{-1} (\check{\underline{\theta}} - \underline{\theta}) < \chi_{m,1-\alpha}^2 \quad (6.12)$$

gegeben, wobei $\underline{\theta} = (\theta_1, \dots, \theta_m)$ der Parameter, $\check{\underline{\theta}}$ die Maximum-Likelihood-Schätzung des Parameters, $\mathbb{I}_{\underline{\theta}}$ die Fishersche Informationsmatrix und $\chi_{m,1-\alpha}^2$ das $(1 - \alpha)$ -Quantil der χ^2 -Verteilung mit m Freiheitsgraden sind.

Dabei sollten solche Parametervarianten gewählt werden, für die eine schnelle Konvergenz gegen die asymptotische Normalverteilung vorliegt. In diesem Fall wird das geforderte Konfidenzniveau auch schon bei relativ kleinen Stichproben erreicht. Konfidenzschätzungen für andere interessierende Größen erhält man, indem die Berandung des ermittelten Konfidenzbereiches „abgelaufen“ wird. Dieses Verfahren liefert außerdem erfahrungsgemäß kleine Konfidenzbereiche.

Auf die Frage der Parameterwahl gehen wir später noch einmal ein.

Beispiel 6.10 Wie in den vorigen Beispielen betrachten wir die Weibullverteilung und den Stichprobenplan $[n, E, T]$. Zunächst macht es sich erforderlich, die zur Berechnung der Information $I_{11}^{(1)}$, $I_{22}^{(1)}$ und $I_{12}^{(1)}$ benötigten Erwartungswerte zu bestimmen. Wie schon erwähnt, besteht eine Möglichkeit darin, die auftretenden Erwartungswerte durch ihre Schätzungen zu ersetzen. Man arbeitet dann statt mit der erwarteten Information mit der geschätzten Information:

$$\begin{aligned} J_{11}^{(1)} &= \frac{1}{n\beta^2} \left(\sum_{i=1}^n d_i + \sum_{i=1}^n \left(\sum_{j=1}^{d_i} \left(\frac{x_{ij}}{\check{\alpha}} \right)^{\check{\beta}} \ln \left(\frac{x_{ij}}{\check{\alpha}} \right)^{\check{\beta}} \right) + \right. \\ &\quad \left. + \sum_{i=1}^n \left(\left(\frac{r^E(T_i)}{\check{\alpha}} \right)^{\check{\beta}} \ln \left(\frac{r^E(T_i)}{\check{\alpha}} \right)^{\check{\beta}} \right) \right), \\ J_{12}^{(1)} &= -\frac{1}{n\alpha} \left(\sum_{i=1}^n d_i + \sum_{i=1}^n \left(\sum_{j=1}^{d_i} \left(\frac{x_{ij}}{\check{\alpha}} \right)^{\check{\beta}} \right) \right), \quad J_{22}^{(1)} = \frac{\beta^2}{n\alpha^2} \sum_{i=1}^n d_i. \end{aligned} \quad (6.13)$$

Nun kann man folgenden simultanen Konfidenzbereich bestimmen. Seien

$$\begin{aligned} C_1 &= \frac{1}{n} \left(\sum_{i=1}^n d_i + \sum_{i=1}^n \left(\sum_{j=1}^{d_i} \left(\frac{x_{ij}}{\check{\alpha}} \right)^{\check{\beta}} \ln \left(\frac{x_{ij}}{\check{\alpha}} \right)^{\check{\beta}} \right) + \right. \\ &\quad \left. + \sum_{i=1}^n \left(\left(\frac{r^E(T_i)}{\check{\alpha}} \right)^{\check{\beta}} \ln \left(\frac{r^E(T_i)}{\check{\alpha}} \right)^{\check{\beta}} \right) \right), \quad C_2 = \frac{1}{n} \sum_{i=1}^n d_i, \\ C_3 &= \frac{1}{n} \left(\sum_{i=1}^n d_i + \sum_{i=1}^n \left(\sum_{j=1}^{d_i} \left(\frac{x_{ij}}{\check{\alpha}} \right)^{\check{\beta}} \right) \right), \quad \tilde{\beta} = \frac{\beta - \check{\beta}}{\beta}, \end{aligned}$$

so erhält man obere und untere Grenzen für β :

$$\check{\beta}/(1+B) \leq \beta \leq \check{\beta}/(1-B) \quad \text{mit} \quad B = \sqrt{\frac{C_2 \chi_{2,1-\alpha}^2}{n(C_1 C_2 - C_3^2)}}.$$

Innerhalb dieser Grenzen variiert α für jedes β in

$$\frac{\check{\alpha}\beta}{\beta + \frac{C_3}{C_2}\tilde{\beta} + A} \leq \alpha \leq \frac{\check{\alpha}\beta}{\beta + \frac{C_3}{C_2}\tilde{\beta} - A} \quad \text{mit} \quad A = \sqrt{\frac{\chi_{2,1-\alpha}^2}{nC_2} - \tilde{\beta}^2 \frac{C_1 C_2 - C_3^2}{C_2^2}}. \quad (6.14)$$

Damit hat man einen simultanen Konfidenzbereich für die beiden Parameter der Weibullverteilung bestimmt.

◇

Wir betrachten nun eine alternative Herangehensweise - die Verwendung des Likelihoodquotienten. Der Likelihoodquotient ist bereits aus der Testtheorie bekannt. Dort besitzt er bekannterweise optimale Eigenschaften. Er ist der mächtigste Test zum Prüfen einer Hypothese bei einer einfachen Alternativhypothese. Um die Ergebnisse der Testtheorie auch für Aufgaben der Konfidenzschätzungen verwenden zu können, wird auf den Zusammenhang dieser Problemstellungen zurückgegriffen: Zum Konfidenzbereich gehört jeder Punkt im m -dimensionalen Parameterraum, der als Nullhypothese bei der gegebenen Punktschätzung (Prüfgröße) nicht abgelehnt wird.

Die ersten asymptotischen Entwicklungen des Likelihoodquotienten stammen von BARTLETT und wurden von LAWLEY verallgemeinert. Es wurde dabei gezeigt, daß für reguläre statistische Experimente der Likelihoodquotient asymptotisch einer χ^2 -Verteilung unterliegt. Die Konvergenzgeschwindigkeit gegen die Grenzverteilung ist $O(n^{-1})$ mit n -Stichprobenumfang, d.h. die Konvergenzordnung ist um eine Größenordnung besser als bei der Verwendung der asymptotischen Normalverteilung der Maximum-Likelihood-Schätzungen. Außerdem bietet die Entwicklung des Likelihoodquotienten die Möglichkeit, einen Korrekturterm, den sogenannten Bartlettfaktor, zu bestimmen. Auf die Berechnung des Bartlett faktors wird am Ende dieses Abschnittes eingegangen. Weitere Vorteile des Likelihoodquotienten bestehen in der Parameterinvarianz und in der Möglichkeit, Störparameter zu berücksichtigen. Schließlich ist es bei der Verwendung des Likelihoodquotienten nicht mehr nötig, die Fisher-Information zu berechnen.

Wir stellen nun die Herangehensweise zur Bestimmung von Konfidenzschätzungen bei Verwendung des Likelihoodquotienten vor.

Der m -dimensionale Parameter $\underline{\theta}$ bestehe aus den k interessierenden Parametern $\underline{\theta}^{(1)} = (\theta_1, \dots, \theta_k)$ und den $(m - k)$ Störparametern (nuisance parameters) $\underline{\theta}^{(2)} = (\theta_{k+1}, \dots, \theta_p)$ ($k \leq m$). Die Likelihoodstatistik w ist durch

$$w = 2 \left(\ln L(\check{\theta}) - \ln L(\tilde{\theta}) \right) \quad (6.15)$$

definiert. Dabei sind $\tilde{\theta}$ die Maximum-Likelihood-Schätzungen der Störparameter in Abhängigkeit von den interessierenden Parametern.

Ein Konfidenzbereich ist dann durch die Ungleichung

$$w \leq \chi_{k,1-\delta}^2 \quad (6.16)$$

gegeben.

Beispiel 6.11 Wir betrachten wieder die Weibullverteilung und den Stichprobenplan $[n, E, T]$. Wenn beide Parameter unbekannt und von Interesse sind, so gilt $p = k = 2$. Logarithmiert man die Likelihoodfunktion (5.14) aus Beispiel 5.17, so erhält man für $l = \ln L(\alpha, \beta, x)$:

$$l = nd \ln \beta - nd \ln \alpha + (\beta - 1) \sum_{i=1}^n \sum_{j=1}^{d_i} \ln x_{ij} - \sum_{i=1}^n \sum_{j=1}^{d_i} \left(\frac{x_{ij}}{\alpha} \right)^\beta - \sum_{i=1}^n \left(\frac{r^E(T_i)}{\alpha} \right)^\beta,$$

wobei $d = \frac{1}{n} \sum_{i=1}^n d_i$ die durchschnittliche Anzahl der Ausfälle pro Realisierung ist. Unter Benutzung der Likelihood-Gleichung $\frac{\partial \ln L}{\partial \alpha} \big|_{\check{\beta}, \check{\alpha}} = 0$ erhält man:

$$\begin{aligned} \ln L(\check{\beta}, \check{\alpha}, x) - \ln L(\beta, \alpha) &= nd \ln \left(\frac{\check{\beta}}{\beta} \right) - nd (\check{\beta} \ln \check{\alpha} - \beta \ln \alpha) + \\ &+ (\check{\beta} - \beta) \sum_{i=1}^n \sum_{j=1}^{d_i} \ln x_{ij} - nd + \sum_{i=1}^n \sum_{j=1}^{d_i} \left(\frac{x_{ij}}{\alpha} \right)^\beta + \sum_{i=1}^n \left(\frac{r^E(T_i)}{\alpha} \right)^\beta. \end{aligned}$$

Damit ist die Umrandung des simultanen Konfidenzbereiches durch folgende Gleichung gegeben:

$$\begin{aligned} 2 \left(nd \ln \left(\frac{\check{\beta}}{\beta} \right) - nd (\check{\beta} \ln \check{\alpha} - \beta \ln \alpha) + (\check{\beta} - \beta) \sum_{i=1}^n \sum_{j=1}^{d_i} \ln x_{ij} - \right. \\ \left. - nd + \sum_{i=1}^n \sum_{j=1}^{d_i} \left(\frac{x_{ij}}{\alpha} \right)^\beta + \sum_{i=1}^n \left(\frac{r^E(T_i)}{\alpha} \right)^\beta \right) = -2 \ln \delta, \end{aligned} \quad (6.17)$$

da $\chi_{2,1-\delta}^2 = -2 \ln \delta$ ist. Die Vorteile des Konfidenzbereiches (6.17) gegenüber den im Beispiel 6.10 hergeleiteten Konfidenzbereichen auf der Grundlage der asymptotischen

Normalverteilung der Maximum-Likelihood-Schätzungen sind die höhere Konvergenzgeschwindigkeit und damit eine höhere Genauigkeit, sowie die relativ einfache Berechnung des Konfidenzbereiches (siehe KAHLE). Gleichung (6.17) muß zwar numerisch gelöst werden, man braucht jedoch keine Approximation der Erneuerungsfunktion und damit im Zusammenhang stehender Größen. Aus Gleichung (6.17) ist außerdem ersichtlich, daß n und d den gleichen Einfluß auf die Größe des Konfidenzbereiches besitzen. Der Konfidenzbereich wird umso kleiner, je größer n bzw. d sind. \diamond

Im folgenden stellen wir die allgemeine Herangehensweise bei der Ermittlung von Konfidenzschätzungen vor.

Wir betrachten das statistische Modell

$$(\mathfrak{X}, \mathfrak{B}(\mathfrak{X}), \mathfrak{P} = (P_\theta, \theta \in \Theta)) .$$

Die Daten, die man im Resultat eines statistischen Experimentes erhält, sind Werte der Zufallsgröße X . Die Wahrscheinlichkeit des Ereignisses $X \in B$ bezeichnen wir in Abhängigkeit vom Wert des Parameters θ mit $P_\theta(B)$ oder mit $P_\theta(X \in B)$, wobei $P_\theta \in \mathfrak{P}$. Der Zufallsgröße X wird ein System $\mathfrak{C}(\Theta)$ von zufälligen Untermengen des Parameter-raumes Θ zugeordnet. Dafür geben wir eine Zahl $\gamma, 0 < \gamma < 1$ vor und ordnen jedem θ eine Untermenge $B(\theta) \in \mathfrak{B}(\mathfrak{X})$ so zu, daß

$$P_\theta(X \in B(\theta)) \geq \gamma .$$

Damit erhält man das Mengensystem $\mathfrak{B}_\gamma(\mathfrak{X}) = \{B(\theta) : B(\theta) \subset \mathfrak{X}, \theta \in \Theta\}$ aller solcher Untermengen. Mittels $\mathfrak{B}_\gamma(\mathfrak{X})$ wird jedem $x \in X$ eine Untermenge $C(x) \subset \Theta$ zugeordnet, die alle diejenigen Parameterwerte θ' enthält, für die x in $B(\theta')$ enthalten ist, also

$$C(x) := \{\theta' : x \in B(\theta')\} .$$

Sei nun $\mathfrak{C}(\Theta) = \{C(x) : x \in \mathfrak{X}\}$ das System aller solcher Untermengen. Das Ereignis $\{X \in B(\theta)\}$ tritt genau dann ein, wenn $\{\theta \in C(X)\}$. Damit ist $\{\theta \in C(X)\}$ nur eine äquivalente Schreibweise für $\{X \in B(\theta)\}$. Es gilt

$$P_\theta(\theta \in C(X)) = P_\theta(X \in B(\theta)) \geq \gamma . \quad (6.18)$$

Dieses bedeutet, daß die „zufällige“ Menge $C(X)$ den wahren Parameter θ mit einer Wahrscheinlichkeit $\geq \gamma$ enthält (siehe auch Abbildung 6.4).

Definition 6.2 Ein Mengensystem $\mathfrak{C} = (C(x) : x \in \mathfrak{X})$, für das für beliebige θ die Beziehung (6.18) gilt, heißt System von Konfidenzmengen zum Konfidenzniveau γ oder System von γ -Konfidenzmengen. Die Zahl $\gamma_0 = \inf_{\theta \in \Theta} P_\theta(\theta \in C(X))$ heißt Konfidenzniveau des Systems von γ -Konfidenzmengen, $\gamma_0 \leq \gamma$.

In Anwendungen ist es häufig erforderlich, Konfidenzschätzungen für eine reelle Funktion $g(\theta^*)$ des unbekannten wahren Parameters θ^* zu bestimmen. Für die Konfidenzgrenzen einer Funktion führen wir die folgenden Bezeichnungen ein:

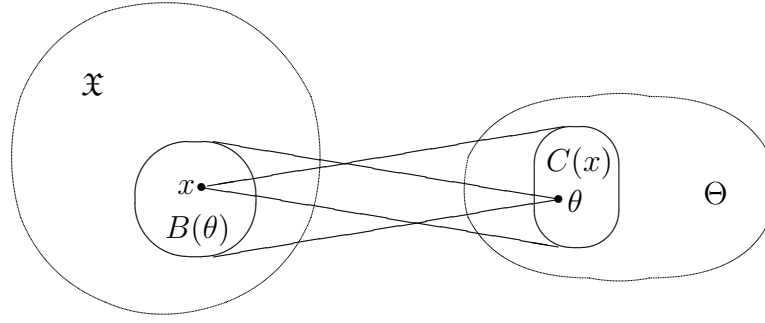


Abbildung 6.4:

Definition 6.3 Eine Zufallsgröße $g_-(X)$ bzw. $g_+(X)$ heißt untere bzw. obere γ -Konfidenzgrenze für den Wert der Funktion $g(\cdot) : \Theta \rightarrow \mathbb{R}^1$, wenn für beliebige $\theta \in \Theta$

$$P_\theta(g_-(X) \leq g(\theta)) \geq \gamma \quad \text{bzw.} \quad P_\theta(g(\theta) \leq g_+(X)) \geq \gamma .$$

Das zufällige Intervall $I(X) = [g_-(X), g_+(X)]$ heißt γ -Konfidenzintervall für den Wert der Funktion $g(\cdot)$, wenn für beliebige $\theta \in \Theta$

$$P_\theta(g_-(X) \leq g_+(X)) = 1 \quad \text{und} \quad P_\theta(g_-(X) \leq g(\theta) \leq g_+(X)) \geq \gamma .$$

Bei Verwendung eines geeigneten Systems von γ -Konfidenzmengen lassen sich γ -Konfidenzgrenzen und γ -Konfidenzintervalle bestimmen.

Satz 6.2 Sei $\mathfrak{C}(\mathfrak{X}) = (C(x) : x \in \mathfrak{X})$ ein System von γ -Konfidenzmengen der Menge Θ . Wenn $g_-(\cdot)$ und $g_+(\cdot)$ $\mathfrak{B}(\mathfrak{X})$ -meßbar sind, dann sind untere und obere γ -Konfidenzgrenzen bzw. ein γ -Konfidenzintervall für den Wert der Funktion $g(\theta^*)$ durch $g_-(X) = \inf_{\theta \in C(x)} g(\theta)$, $g_+(X) = \sup_{\theta \in C(x)} g(\theta)$ bzw. $I(X) = [g_-(X), g_+(X)]$ gegeben.

Beweis: Die Behauptung folgt unmittelbar aus den Beziehungen

$$\begin{aligned} (x : g_-(x) \leq g(\theta)) &\supseteq (x : \theta \in C(x)) = (x : x \in B(\theta)) , \\ (x : g(\theta) \leq g_+(x)) &\supseteq (x : \theta \in C(x)) = (x : x \in B(\theta)) , \\ (x : g_-(x) \leq g(\theta) \leq g_+(x)) &= (x : g_-(x) \leq g(\theta)) \cap (x : g(\theta) \leq g_+(x)) \end{aligned}$$

und den sich daraus ergebenden Ungleichungen für die Wahrscheinlichkeiten:

$$\begin{aligned} P_\theta(g_-(X) \leq g(\theta)) &\geq P_\theta(X \in B(\theta)) \geq \gamma , \\ P_\theta(g(\theta) \leq g_+(X)) &\geq P_\theta(X \in B(\theta)) \geq \gamma , \\ P(g_-(X) \leq g(\theta) \leq g_+(X)) &\geq P_\theta(X \in B(\theta)) \geq \gamma . \end{aligned} \quad \square$$

Beispiel 6.12 Wir betrachten Konfidenzschätzungen für die Wahrscheinlichkeit der ausfallfreien Arbeit eines Elementes innerhalb eines gegebenen Zeitintervalls. Die Daten sollen das Ergebnis einer sogenannten Binomialprüfung sein, d.h. die Ausfallzeiten

selbst werden nicht registriert. Wir nehmen an, daß n Elemente einer Prüfung unterzogen wurden und k Ausfälle registriert wurden. Die zufällige Anzahl der ausgefallenen Elemente unterliegt einer Binomialverteilung:

$$P(X = d) = \binom{n}{d} p^d (1-p)^{n-d}, \quad d = 0, \dots, n, \quad 0 < p < 1. \quad (6.19)$$

Sei desweitern

$$B_k^l(n, p) = P(k \leq X \leq l) = \sum_{d=k}^l \binom{n}{d} p^d (1-p)^{n-d}. \quad (6.20)$$

Zur Bestimmung von Konfidenzgrenzen geben wir $\varepsilon_1 > 0$ und $\varepsilon_2 > 0$ vor und bezeichnen mit $\gamma = 1 - (\varepsilon_1 + \varepsilon_2) > 0$ das Konfidenzniveau. Seien $d_\gamma^+ = d_\gamma^+(n, p)$ und $d_\gamma^- = d_\gamma^-(n, p)$ natürliche Zahlen, die durch die Beziehungen

$$B_0^{d_\gamma^+}(n, p) \geq 1 - \varepsilon_2 > B_0^{d_\gamma^+-1}(n, p), \quad (6.21)$$

und

$$B_0^{d_\gamma^- - 1}(n, p) < \varepsilon_1 \leq B_0^{d_\gamma^-}(n, p). \quad (6.22)$$

eindeutig bestimmt sind. Da

$$P(d_\gamma^- \leq X \leq d_\gamma^+) = B_0^{d_\gamma^+}(n, p) - B_0^{d_\gamma^-}(n, p) \geq \gamma = 1 - (\varepsilon_1 + \varepsilon_2), \quad (6.23)$$

bilden die Intervalle

$$B(p) = [d_\gamma^-(n, p), d_\gamma^+(n, p)] \quad 0 < p < 1, \quad (6.24)$$

wobei die ganzzahligen Werte $d_\gamma^- = d_\gamma^-(n, p)$ und $d_\gamma^+ = d_\gamma^+(n, p)$ den Gleichungen (6.21) und (6.22) genügen, ein System von $\mathcal{B}_\gamma(\mathfrak{X})$ -Mengen. Somit ist jede Menge $C_\gamma(d)$, $d = 1, \dots, n$, von Werten von p , die bei fixiertem d in mindestens eines der Intervalle aus $\mathcal{B}_\gamma(\mathfrak{X})$ fällt, ein γ -Konfidenzintervall. Da $d_\gamma^+(n, p)$ und $d_\gamma^-(n, p)$ nicht fallende Treppenfunktion in p sind, erhält man für $C_\gamma(d)$ Intervalle $C_\gamma(d) = [p_\gamma^-(n, d), p_\gamma^+(n, d)]$, deren Ränder aus

$$B_0^d(n, p_\gamma^+(n, d)) = 1 - \varepsilon_2, \quad B_0^{d+1}(n, p_\gamma^-(n, d)) = \varepsilon_1. \quad (6.25)$$

bestimmt werden. Die Gleichungen (6.22) heißen Clopper - Pearson - Gleichungen. \diamond

Bemerkung: Wenn eine Binomialprüfung von n Elementen im Intervall $[0, T]$ durchgeführt wird und die Lebensdauer jedes Elementes als exponentialverteilt vorausgesetzt wird: $p = 1 - e^{-\lambda T}$, $\lambda = \ln \frac{1}{1-p}$, so erhält man aus dem Konfidenzintervall für p

$$p_\gamma^-(n, d) \leq p \leq p_\gamma^+(n, d)$$

ein äquivalentes Intervall für λ :

$$\frac{1}{T} \ln \frac{1}{1 - p_\gamma^-(n, d)} \leq \lambda \leq \frac{1}{T} \ln \frac{1}{p_\gamma^+(n, d)}.$$

Beispiel 6.13 Wir führen eine Prüfung von n Elementen gemäß dem Stichprobenplan $[n, 0, r]$ durch, d.h. bei Eintritt des r -ten Ausfalls wird die Beobachtung abgebrochen. Seien die Lebensdauern der Elemente exponentialverteilt mit der Verteilungsfunktion $F(t) = 1 - e^{-\lambda t}$, $t \geq 0$. Desweiteren bezeichnen wir mit $S_1 < S_2 < \dots < S_r$ die zufälligen Ausfallszeiten. Wie im vorigen Kapitel im Beispiel 5.10 gezeigt wurde, unterliegt die Gesamtprüfzeit

$$S(r) = nS_{(1)} + (n-1)(S_{(2)} - S_{(1)}) + \dots + (n-r+1)S_{(r)} - S_{(r-1)}$$

einer Gammaverteilung mit der Dichte

$$p_\lambda(s) = \frac{\lambda^r s^{r-1}}{(r-1)!} e^{-\lambda s}, \quad s \geq 0.$$

Die Zufallsgröße $2\lambda S(r)$ unterliegt einer χ^2 -Verteilung mit $2r$ Freiheitsgraden und der Dichte $\frac{1}{2^r} \frac{s^{r-1}}{(r-1)!} e^{-\frac{1}{2}s}$. Wir bezeichnen mit $\chi_u^2(2r)$ das Quantil dieser χ^2 -Verteilung zur Ordnung u . Im Fall $r = 1$ ist dieses Quantil explizit bestimmbar

$$\chi_u^2(2) = 2 \ln \frac{1}{1 - \gamma}, \quad (6.26)$$

da die χ^2 -Verteilung mit zwei Freiheitsgraden gerade eine Exponentialverteilung ist. Wenn $\varepsilon_1 > 0$, $\varepsilon > 0$ und $\gamma = 1 - (\varepsilon_1 + \varepsilon_2) > 0$ gegeben sind, dann gilt für beliebige $\lambda > 0$:

$$\begin{aligned} & P\left(\frac{1}{2\lambda} \chi_{\varepsilon_1}^2(2r) \leq S(r) \leq \frac{1}{2\lambda} \chi_{1-\varepsilon_2}^2(2r)\right) \\ &= P\left(\chi_{\varepsilon_1}^2(2r) \leq 2\lambda S(r) \leq \chi_{1-\varepsilon_2}^2(2r)\right) = \gamma. \end{aligned} \quad (6.27)$$

Führen wir nun in Übereinstimmung mit oben die Bezeichnungen $X = S(r)$, $\mathfrak{X} = (0, \infty)$, $\Theta = (0, \infty)$, $\theta = \lambda$ ein, dann folgt aus 6.27, daß

$$B_\gamma(\lambda) = \left[\frac{1}{2\lambda} \chi_{\varepsilon_1}^2(2r), \frac{1}{2\lambda} \chi_{1-\varepsilon_2}^2(2r) \right], \quad \lambda > 0$$

ein System von $\mathcal{B}_\gamma(\mathfrak{X})$ -Mengen bildet. Zur Bestimmung der γ -Konfidenzmengen müssen die Ungleichungen $\frac{1}{2\lambda} \chi_{\varepsilon_1}^2(2r) \leq S(r) \leq \frac{1}{2\lambda} \chi_{1-\varepsilon_2}^2(2r)$ nach λ aufgelöst werden. In diesem Beispiel erhalten wir aus 6.27

$$P\left(\frac{\chi_{\varepsilon_1}^2(2r)}{2S(r)} \leq \lambda \leq \frac{\chi_{1-\varepsilon_2}^2(2r)}{2S(r)}\right) = \gamma.$$

Damit stellt das Intervall mit den zufälligen Grenzen

$$\left[\frac{\chi_{\varepsilon_1}^2(2r)}{2S(r)}, \frac{\chi_{1-\varepsilon_2}^2(2r)}{2S(r)} \right]$$

ein γ -Konfidenzintervall für den Parameter λ bei Daten nach dem Stichprobenplan $[n, 0, r]$ dar. \diamond

Die beschriebene Herangehensweise zur Bestimmung von γ -Konfidenzmengen kann auch dann verwendet werden, wenn die Lebensdauerverteilung der Elemente mehrere unbekannte Parameter $\underline{\theta} = (\theta_1, \dots, \theta_k)$ enthält und wenn es gelingt, eine Funktion $g(\underline{\theta}, X)$ zu finden, deren Verteilung nicht vom unbekannten Parameter $\underline{\theta}$ abhängt. In diesem Falle kann man wieder für beliebige Konfidenzniveaus γ solche Mengen G_γ bestimmen, daß

$$P_\theta(g(X, \underline{\theta}) \in G_\gamma) \geq \gamma. \quad (6.28)$$

Je nach Fragestellung kann G_γ durch $G_\gamma = (X : g(X, \theta) \leq g_\gamma)$ oder $G_\gamma = (X : g(X, \theta) \geq g_\gamma)$ oder, bei zweiseitige Fragestellung durch $G_\gamma = (X : g_\varepsilon \leq g(X, \underline{\theta}) \leq g_{1-\varepsilon})$ bestimmt werden, wobei g_γ , g_ε , $g_{1-\varepsilon}$ nicht vom unbekannten Parameter $\underline{\theta}$ abhängen. Wir illustrieren diese Herangehensweise am Beispiel einer Lebensdauerverteilung mit zwei unbekannten Parametern.

Beispiel 6.14 n Elemente werden einer Lebensdauerprüfung gemäß dem Stichprobenplan $[n, O, n]$ unterzogen. Mit S_1, \dots, S_n bezeichnen wir wieder die Ausfallzeitpunkte. Die Lebensdauern der Elemente seien unabhängig voneinander und identisch logarithmisch normalverteilt mit der Verteilungsfunktion $F(s) = \Phi\left(\frac{\ln s - \mu}{\sigma}\right)$. Die logarithmische Normalverteilung besitzt die Dichte

$$p(s, \underline{\theta}) = \frac{1}{\sqrt{2\pi}\sigma s} e^{-\frac{1}{2} \frac{(\ln s - \mu)^2}{\sigma^2}}, \quad s > 0, \quad \underline{\theta} = (\mu, \sigma^2)$$

und damit erhalten wir die Likelihoodfunktion

$$L(\underline{\theta}, s_1, \dots, s_n) = \frac{1}{(2\pi)^{n/2} \sigma^n} \prod_{i=1}^n \frac{1}{s_i} e^{-\frac{1}{2} \left(\frac{\ln s_i - \mu}{\sigma}\right)^2}. \quad (6.29)$$

Da die Zufallsgrößen $X_i = \ln S_i$ einer Normalverteilung $N_1(\mu, \sigma^2)$ unterliegen, erhalten wir als Maximum-Likelihood-Schätzungen für μ und σ^2

$$\hat{\mu}_n = \bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i, \quad \hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (6.30)$$

Mit den Bezeichnungen $\underline{x}_n = (x_1, \dots, x_n)$, $x_i = \ln s_i$, $i = 1, \dots, n$, $\hat{\underline{\theta}} = (\hat{\mu}_n, \hat{\sigma}_n^2)$ erhalten wir für die Likelihoodfunktion

$$\ln L(\underline{\theta}, \underline{x}_n) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^2, \quad (6.31)$$

für die Likelihoodfunktion an der Stelle $\underline{\theta}$

$$\begin{aligned} \ln L(\hat{\underline{\theta}}_n, \underline{x}_n) &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \hat{\sigma}_n^2 - \frac{1}{2} \sum_{i=1}^n \frac{(x_i - \bar{\mu})^2}{\hat{\sigma}_n^2} \\ &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 - \frac{n}{2}. \end{aligned} \quad (6.32)$$

und für den Likelihoodquotienten

$$2 \ln \frac{L(\hat{\underline{\theta}}_n, \underline{x}_n)}{L(\underline{\theta}, \underline{x}_n)} = -n \ln \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^2 - n - \sum_{i=1}^n \left(\frac{x_i - \bar{\mu}}{\sigma} \right)^2. \quad (6.33)$$

Wir betrachten nun die Zufallsgröße

$$g(\underline{\theta}, X) = 2 \ln \frac{L(\hat{\underline{\theta}}_n, \underline{X}_n)}{L(\underline{\theta}, \underline{X}_n)}, \quad (6.34)$$

die durch Gleichung 6.33 definiert ist, wobei die Beobachtungen x_i durch die entsprechenden Zufallsgrößen X_i und die Punktschätzungen durch $\hat{\mu}_n = \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$, $\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ ersetzt wurden. Die Verteilung der Zufallsgröße $g(\underline{\theta}, X)$ hängt nicht von den unbekannten Parametern ab, daher können wir den Fall $\mu = 0$, $\sigma^2 = 1$ betrachten. Hierfür erhalten wir

$$g((0, 1), \underline{X}_n) = -n \ln \left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right) + \sum_{i=1}^n x_i^2 - n. \quad (6.35)$$

Wenn g_γ ein Wert ist, für den

$$P(g((0, 1), \underline{X}_n) \leq g_\gamma) = \gamma,$$

dann gilt für die wahren Parameter μ und σ^2

$$P_{(\mu, \sigma^2)}(g(\mu, \sigma^2), \underline{X}_n) \leq g_\gamma) = \gamma.$$

Somit können γ -Konfidenzmengen $C_\gamma(x)$ für (μ, σ^2) aus dem Mengensystem $B_\gamma = (\underline{x}_n : g(\mu, \sigma^2, \underline{X}_n) \leq g_\gamma)$ ermittelt werden:

$$C_\gamma(x_n) = ((\mu, \sigma^2) : g(\mu, \sigma^2, X_n) \leq g_\gamma) \quad (6.36)$$

Zur Bestimmung von g_γ bestimmen wir die Verteilung von (6.34) für den Fall $\mu = 0$, $\sigma^2 = 1$. Dazu definieren wir die beiden Zufallsgrößen $Z_1(n) = \frac{1}{n} \sum_{i=1}^n X_i$ und $Z_2(n) = \frac{1}{\sqrt{2n}} \sum_{i=1}^n (X_i^2 - 1)$. Ihre ersten und zweiten Momente sind $EZ_1(n) = EZ_2(n) = 0$, $EZ_1^2(n) = EZ_2^2(n) = 1$, $EZ_1(n)Z_2(n) = 0$. Für das Verteilungsgesetz von $Z_1(n)$ gilt $\mathcal{L}(Z_1(n)) = N_1(0, 1)$ und die Verteilung von $Z_2(n)$ läßt sich leicht aus einer χ^2 -Verteilung bestimmen. Aus dem Zentralen Grenzwertsatz erhält man für $n \rightarrow \infty$ $\mathcal{L}(Z_2(n)) \rightarrow N_1(0, 1)$. Da außerdem $Z_1(n)$ und $Z_2(n)$ unkorreliert sind, sind sie asymptotisch unabhängig und $Z_1^2(n) + Z_2^2(n)$ unterliegt asymptotisch einer χ^2 -Verteilung mit zwei Freiheitsgraden. Somit erhält man aus (6.34)

$$g((0, 1), \underline{X}_n) = -n \ln \left(1 - \frac{1}{n} Z_1^2(n) + \sqrt{\frac{2}{n}} Z_2(n) \right) + \sqrt{2n} Z_2(n), \quad (6.37)$$

und unter Verwendung der Markov-Ungleichung

$$\begin{aligned}
& P_{(0,1)} \left(\frac{1}{n} Z_1^2(n) + \sqrt{\frac{2}{n}} |Z_2(n)| \geq \varepsilon \right) \\
& \leq \frac{1}{\varepsilon} \left(\frac{1}{n} E Z_1^2(n) + \sqrt{\frac{2}{n}} E |Z_2(n)| \right) \leq \frac{1}{\varepsilon} \left(\frac{1}{n} + \sqrt{\frac{2}{n}} (E Z_2^2(n))^{1/2} \right) \\
& = \frac{1}{\varepsilon} \left(\frac{1}{n} + \sqrt{\frac{2}{n}} \right) \rightarrow 0, \quad n \rightarrow \infty.
\end{aligned}$$

Somit folgt

$$U(n) := -\frac{1}{n} Z_1^2(n) + \sqrt{\frac{2}{n}} Z_2(n) = o_p(1), \quad n \rightarrow \infty. \quad (6.38)$$

Wiederum unter Verwendung der Markov-Ungleichung erhalten wir

$$\frac{1}{2n} Z_1^2(n) = o_p(1), \quad \sqrt{\frac{2}{n}} Z_1^2(n) Z_2(n) = o_p(1), \quad n \rightarrow \infty. \quad (6.39)$$

Es gilt fast sicher, daß $1 + U(n) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 > 0$. Daher können wir die Beziehung

$$\ln(1 + y) = 1 - \frac{y^2}{2} + w(y), \quad (6.40)$$

benutzen, wobei $|w(y)| \leq |y|^3$ wenn $|y| < 1/2$. Nun folgt aus (6.38) für $n \rightarrow \infty$

$$V(n) := g((0, 1), X_n) I(|U(n)| \geq 1/2) = o_p(1) \quad (6.41)$$

und

$$n(U(n) - \frac{1}{2} U^2(n) + w(U(n))) I(|U(n)| \geq 1/2) = o_p(1), \quad (6.42)$$

da $I(|V(n)| > 0) \leq I(|U(n)| \geq 1/2)$. Völlig analog folgt aus

$$\begin{aligned}
I(n | U(n) | I(|U(n)| \geq 1/2) > 0) & \leq I(|U(n)| \geq 1/2), \\
I(n | U^2(n) | I(|U(n)| \geq 1/2) > 0) & \leq I(|U(n)| \geq 1/2), \\
I(n | w(U(n)) | I(|U(n)| \geq 1/2) > 0) & \leq I(|U(n)| \geq 1/2),
\end{aligned}$$

daß

$$n \max(|U(n)|, |U^2(n)|, |w(U(n))|) I(|U(n)| \geq 1/2) = o_p(1). \quad (6.43)$$

Nun erhalten wir aus (6.38) bis (6.43)

$$\begin{aligned}
& g((0, 1), \underline{X}_n) = g((0, 1), \underline{X}_n) \mathbf{I}(|U(n)| < 1/2) + o_p(1) \\
& = \left[-n \left(U(n) - \frac{1}{2} U^2(n) + w(U(n)) \right) + \sqrt{2n} Z_2(n) \right] \\
& \cdot (1 - \mathbf{I}(|U(n)| \geq 1/2) + o_p(1)) \\
& = -n \left(-\frac{1}{n} Z_1^2(n) + \sqrt{\frac{2}{n}} Z_2(n) - \frac{1}{2} \left(-\frac{1}{n} Z_1^2(n) + \sqrt{\frac{2}{n}} Z_2(n) \right)^2 \right) + o_p(1) \\
& = Z_1^2(n) - \sqrt{2n} Z_2(n) - \frac{1}{2n} Z_1^2(n) + \sqrt{\frac{2}{n}} Z_1^2(n) Z_2(n) + Z_2^2(n) + \sqrt{2n} Z_2(n) \\
& = Z_1^2(n) + Z_2^2(n) + o_p(1), \quad n \rightarrow \infty.
\end{aligned}$$

Somit konvergiert die Verteilung von $g((0, 1), \underline{X}_n)$ (und damit auch von $g(\mu, \sigma^2), X_n$) gegen eine χ^2 -Verteilung mit zwei Freiheitsgraden. Das Quantil zur Ordnung γ ist, wie schon erwähnt, $2 \ln \frac{1}{1-\gamma}$. Damit erhält man eine asymptotische γ -Konfidenzmenge aus

$$-n \ln \left(\frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^2 \right) + \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^2 - n \leq 2 \ln \frac{1}{1-\gamma} \quad (6.44)$$

Die Umrandung dieses Konfidenzbereiches ist eine Linie, die mittels Niveaulinienprogrammen oder durch Modellierung ermittelt werden kann. \diamond

Der in diesen Beispielen angewendete Weg ist für viele andere statistische Modelle anwendbar. Hierfür müssen lediglich die Regularitätsvoraussetzungen des letzten Kapitels erfüllt sein. Unter diesen Voraussetzungen unterliegt der logarithmierte Likelihoodquotient $2 \ln \frac{\hat{L}}{L}$ mit $\hat{L} = L(\hat{\theta}_n, \underline{x}_n)$, $L = L(\theta, \underline{x}_n)$ und der Maximum-Likelihood-Schätzung $\hat{\theta}_n$ asymptotisch für $n \rightarrow \infty$ einer χ^2 -Verteilung mit k Freiheitsgraden. Die Konvergenzgeschwindigkeit gegen die Grenzverteilung ist hierbei $O\left(\frac{1}{n}\right)$. Somit können asymptotische γ -Konfidenzintervalle aus

$$2 \ln \frac{L(\hat{\theta}, x_n)}{L(\theta, x_n)} \leq \chi_\gamma(k)$$

ermittelt werden. Dieser Weg wurde bereits im vorigen Abschnitt beschrieben.

7 Nichtparametrische Schätzungen für zensierte Daten

7.1 Produkt-Limit-Schätzer

Wir nehmen zunächst an, daß die der Stichprobe zugrundeliegende Verteilung diskret mit Realisierungen $a_1 < a_2 < \dots$ ist. Die zugehörigen (unbekannten) Einzelwahrscheinlichkeiten seien p_j , $j = 1, 2, \dots$.

Weiterhin sei h_j , $j = 1, 2, \dots$ die diskrete Ausfallrate, das heißt die Wahrscheinlichkeit eines Ausfalls zur Zeit j unter der Bedingung, daß bisher kein Ausfall stattgefunden hat:

$$h_j = \frac{p_j}{1 - F(a_j)} = \frac{p_j}{P(X > a_j)} = \frac{p_j}{p_j + p_{j+1} + \dots}$$

Nun kann die Überlebensfunktion $\bar{F}(t)$ folgendermaßen durch die diskrete Ausfallrate ausgedrückt werden:

$$\bar{F}(t) = \prod_{j:a_j \leq t} (1 - h_j) . \quad (7.1)$$

Beweis: Sei l ein Index so, daß $a_l \leq t$ und $a_{l+1} > t$. Dann erhalten wir

$$\begin{aligned} \bar{F}(t) &= \prod_{j:a_j \leq t} (1 - h_j) = (1 - h_1)(1 - h_2) \cdot \dots \cdot (1 - h_l) \\ &= \left(1 - \frac{p_1}{p_1 + \dots}\right) \left(1 - \frac{p_2}{p_2 + \dots}\right) \cdot \dots \cdot \left(1 - \frac{p_l}{p_l + \dots}\right) \\ &= \left(\frac{p_2 + \dots}{p_1 + \dots}\right) \left(\frac{p_3 + \dots}{p_2 + \dots}\right) \cdot \dots \cdot \left(\frac{p_{l+1} + \dots}{p_l + \dots}\right) \\ &= \sum_{j:a_j > t} p_j = P(X > t) . \end{aligned}$$

□

Damit ist ein nichtparametrischer Schätzer der Überlebensfunktion:

$$\hat{\bar{F}}(t) = \prod_{j:a_j \leq t} (1 - \hat{h}_j) ,$$

wobei die Größen \hat{h}_j Maximum-Likelihood-Schätzer für h_j sind. Die Berechnung dieser

Schätzer erfolgt folgendermaßen:

$$\begin{aligned}
L &= \prod_{i=1}^n (\text{P}(X_i = a_i))^{\delta_i} (\text{P}(X_i > a_i))^{1-\delta_i} \\
&\quad \downarrow \qquad \qquad \downarrow \\
&\quad \text{Ausfall} \qquad \text{Zensierung} \\
\\
p_i &= h_i \cdot \text{P}(X_i \geq a_i) \\
\\
L &= \prod_{i=1}^n h_i^{\delta_i} \cdot \underbrace{\prod_{i=1}^n \text{P}(X_i > a_i)}_{= \prod_{i=1}^n \left(1 - \sum_{j=1}^i p_j\right)} \\
&= \prod_{i=1}^n \prod_{j=1}^{i-1} (1 - h_j) \\
\ln(L) &= \sum_{i=1}^n \delta_i \ln(h_i) + \sum_{i=1}^n \sum_{j=1}^{i-1} \ln(1 - h_j) \\
&= \sum_{i=1}^n \delta_i \ln(h_i) + \sum_{j=1}^{n-1} \underbrace{\sum_{i=j+1}^n \ln(1 - h_j)}_{= (n - j - 1) \ln(1 - h_j)} \\
&= \sum_{i=1}^n \delta_i \ln(h_i) + \sum_{i=1}^{n-1} (n - i - 1) \ln(1 - h_i) .
\end{aligned}$$

Sei d_i die Anzahl von Ausfällen bei a_i , r_i die Anzahl von verbleibenden Elementen bei a_i und $r_i - d_i$ die Anzahl von Zensierungen ($\sum_{i=1}^k r_i = n$).

$$\ln(L) = \sum_{l=1}^k r_l \ln(h_l) + (r_l - d_l) \ln(1 - h_l)$$

Dann ergibt die Maximierung

$$\hat{h}_l = \frac{d_l}{r_l} .$$

Definition 7.1 Die Stichprobenfunktion

$$\hat{F}(t) = \prod_{j: a_j \leq t} \left(1 - \frac{d_j}{r_j}\right) \tag{7.2}$$

heißt Produkt-Limit-Schätzer oder Kaplan-Meier-Schätzer.

Die gewöhnliche empirische Verteilungsfunktion ist ein Spezialfall des Produkt-Limit-Schätzer. Die Herleitung für steile Verteilungen erfolgt analog.

7.2 Actuarial Schätzung

Wir betrachten eine stetige Verteilung mit stückweise konstanter Ausfallrate.

$$h(t) = \frac{f(t)}{1 - F(t)} = \rho_j \quad (a_{j-1} \leq t < a_j \text{ und } a_0 = 0) .$$

Die Zensierung erfolgt ebenfalls nach einem zufälligen Mechanismus mit einer Ausfallrate

$$h_C = \lambda_j \quad (a_{j-1} \leq t < a_j) .$$

Von Interesse sind Schätzungen der Intensitäten ρ_j , die Zensierungsintensitäten λ_j sind nuisance Parameter. Wir betrachten das j -te Intervall, für das die Überlebenswahrscheinlichkeit geschätzt werden soll. Seien $b_j = a_j - a_{j-1}$ die Intervalllänge und r_{j-1} die Anzahl der zur Zeit a_{j-1} noch lebenden Elemente. Dann besteht die Beobachtung aus

- (i) $r_j - d_j - m_j$ unzensiert überlebenden Elementen,
- (ii) d_j Elementen, die im Intervall ausfallen,
- (iii) m_j Elementen, die im Intervall zensiert werden.

Die einzelnen Beiträge zur Likelihoodfunktion sind

- (i) $\exp(-b(\rho + \lambda))$,
- (ii)
$$\int_0^b \rho \exp(-\rho x) \exp(-\lambda x) dx = \frac{\rho}{\rho + \lambda} (1 - \exp(-b(\rho + \lambda))) ,$$
- (iii)
$$\int_0^b \lambda \exp(-\lambda x) \exp(-\rho x) dx = \frac{\lambda}{\rho + \lambda} (1 - \exp(-b(\rho + \lambda))) .$$

Für eine übersichtlichere Schreibweise wurden die Indizes j und $j - 1$ weggelassen. Die Loglikelihoodfunktion ist somit

$$-(r - d - m)b(\rho + \lambda) + d \log\left(\frac{\rho}{\rho + \lambda}\right) + m \log\left(\frac{\lambda}{\rho + \lambda}\right) + (d + m) \log(1 - \exp(-b(\rho + \lambda))) .$$

Die Maximierung dieser Likelihoodfunktion führt zu

$$\hat{\rho} = -\frac{d}{b(d + m)} \log\left(\frac{r - d - m}{r}\right) . \quad (7.3)$$

Wenn b klein ist, dann ist $(d+m)/r$ ebenfalls klein und der Logarithmus kann durch die ersten Glieder der Taylorreihe approximiert werden:

$$\log\left(\frac{r-d-m}{r}\right) = \log\left(1 - \frac{d+m}{r}\right) \approx -\frac{d+m}{r} - \frac{1}{2} \frac{(d+m)^2}{r^2}.$$

Damit erhält man für die Schätzung der Ausfallrate

$$b\hat{\rho} = \frac{d}{r} + \frac{1}{2} \frac{d(d+m)}{r^2} + O\left(\frac{(d+m)}{r}\right)^3 = \frac{d}{r - \frac{1}{2}(d+m)} (1 + O\left(\frac{(d+m)}{r}\right)^2).$$

Die Schätzung

$$\hat{\rho}_j = \frac{1}{b_j} \cdot \frac{d_j}{(r_{j-1} - \frac{1}{2}(d_j + m_j))} \quad (7.4)$$

wird gewöhnlich zur Schätzung der Ausfallrate für Intervalldaten (wie z. B. in Sterbetafeln) verwendet. Zur Interpretation kann man sich vorstellen, daß die zensierten Elemente die Hälfte der Zeit gelebt haben.

7.3 Schätzung der integrierten Hazard-Funktion

Es ist möglich den Schätzer der integrierten Hazard-Funktion aus der Schätzung der Überlebensfunktion zu ermitteln:

$$\hat{H}(t) = -\ln(\hat{F}(t)) \quad .$$

Gebräuchlicher ist jedoch der folgende Schätzer:

$$\hat{H}(t) = \sum_{j: a_j < t} \frac{d_j}{r_j} \quad \left(= \sum_{a_j < t} h_j \right) \quad .$$

Wenn keine Zensierung vorliegt und alle Beobachtungen voneinander verschieden sind, nimmt der Schätzer die folgende einfache Form an:

$$\hat{H}(a_k) = \sum_{j=1}^k \frac{1}{n+1-j}$$

wobei die a_k die geordneten Ausfallzeitpunkte sind.

Das heißt: $\hat{H}(a_k)$ ist der erwartete Wert der k -ten Order-Statistik in einer Standard-Exponentialverteilung.

8 Proportional Hazard und Accelerated Life Testing

8.1 Accelerated life testing

Wir nehmen an, daß zwei unterschiedliche Lebensbedingungen vorliegen, z. B. zwei Behandlungen, die durch eine Variable z mit den Werten $z = 0$ oder $z = 1$ gekennzeichnet sind. Die Überlebensfunktion bei $z = 0$ sei $\bar{F}_0(t)$, die bei $z = 1$

$$\bar{F}_1(t) = \bar{F}(t; 1) = \bar{F}_0(\psi t) ,$$

wobei ψ eine Konstante ist. Daraus folgt:

$$\begin{aligned} f_1(t) &= \psi f_0(\psi t) \\ h_1(t) &= \psi h_0(\psi t) \end{aligned}$$

Dieses Modell kann so interpretiert werden, daß einer Überlebenszeit t unter $z = 0$ eine Überlebenszeit $\frac{t}{\psi}$ unter $z = 1$ entspricht. Dann gilt: $T_1 = \frac{T_0}{\psi}$.

Bei mehreren möglichen Werten der erklärenden Variablen (Kovariaten) wird angenommen, daß eine Funktion $\psi(z)$ so existiert, daß folgendes gilt:

$$\begin{aligned} \bar{F}(t; z) &= \bar{F}_0(t\psi(z)) \\ f(t; z) &= f_0(t\psi(z)) \cdot \psi(z) \\ h(t; z) &= h_0(t\psi(z)) \cdot \psi(z) \end{aligned}$$

mit $\psi(0) = 1$ wenn $z = 0$ die Standardbedingung ist. Daraus ergibt sich:

$$T = \frac{T_0}{\psi(z)}$$

Zur Bestimmung der Größen $\psi(z)$ kann der folgende Ansatz dienen: Sei $\mu_0 = E(\log(T_0))$, dann gilt

$$\log(T) = \mu_0 - \log(\psi(z)) + \epsilon ,$$

wobei ϵ eine Zufallsgröße mit $E(\epsilon) = 0$ ist und die Verteilung von ϵ nicht von z abhängen soll.

Manchmal ist es nicht nötig, $\psi(z)$ näher zu spezifizieren. Wenn es jedoch nötig ist, wird eine parametrische Form von ψ

$$\psi(z; \beta)$$

gewählt. Da $\psi(z; \beta) \geq 0$ und $\psi(0; \beta) = 1$ sind, ist ein natürlicher Kandidat für $\psi(z; \beta)$

$$\psi(z; \beta) = e^{\beta^T z} \quad \text{mit} \quad \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_q \end{pmatrix}.$$

Dann können die unbekannten Parameter β mittels des lineares Regressionsmodells

$$\log(T) = \mu_0 - \beta^T z + \epsilon$$

ermittelt werden.

Zur Modellüberprüfung werden die empirischen Quantile $t_0^{(a)}$ und $t_1^{(a)}$ berechnet:

$$\begin{aligned} t_0^{(a)} &= 1 - F_0^{-1}(a) \\ t_1^{(a)} &= 1 - F_1^{-1}(a) \end{aligned}$$

und unter der Bedingung $\bar{F}_1(t) = \bar{F}_0(\psi t)$ folgt

$$t_1^{(a)} = \frac{t_0^{(a)}}{\psi}$$

d. h. ein qq-Plot ergibt eine Gerade.

Ein allgemeineres Modell umfaßt zeitabhängige erklärende Variable $Z(t)$ so, daß die Zeit transformiert wird. Damit besitzt ein Individuum, gegeben durch $Z(t)$, eine Zeitachse $t^{(z)}$, die sich relativ zu $t^{(0)}$ entwickelt:

$$\frac{dt^{(z)}}{dt^{(0)}} = \frac{1}{\psi[Z(t^{(z)})]}$$

Es folgt:

$$t^{(0)} = \int_0^{t^{(z)}} \psi(Z(u)) du = \psi(t^{(z)})$$

Damit sind die Ausfallzeiten sind durch

$$T = \psi^{-1}(T_0)$$

gegeben und es gilt

$$\begin{aligned} \bar{F}(t, Z(.)) &= \bar{F}_0(\psi(t)) \\ f(t, Z(.)) &= \psi(Z(t)) f_0(\psi(t)) \\ h(t, Z(.)) &= \psi(Z(t)) h_0(\psi(t)) \end{aligned}$$

8.2 Proportional Hazard

Die einfache Form eines proportional hazard Modells besteht in der Annahme

$$h(t; z) = \psi(z) \cdot h_0(t) \quad \text{mit} \quad \psi(0) = 1 .$$

$\psi(z)$ ist diesmal jedoch keine Zeittransformation, Verteilungsfunktion und Dichte lassen sich folgendermaßen durch die Basisgrößen ausdrücken:

$$\begin{aligned} \bar{F}(t; z) &= [\bar{F}_0]^{\psi(z)} \\ f(t; z) &= \psi(z) [\bar{F}_0(t) t]^{\psi(z)-1} f_0(t) \end{aligned}$$

Im allgemeinen wird auch hier die Funktion $\psi(z)$ in einer parametrischen Form

$$\psi(z, \beta) = e^{\beta^T z}$$

betrachtet.

Wir untersuchen jetzt den Zusammenhang zum accelerated life model. Damit die Modelle übereinstimmen, muß eine Funktion $\chi(z)$ so existieren, daß folgendes gilt:

$$\begin{aligned} (\bar{F}(t))^{\psi(z)} &= \bar{F}(t \cdot \chi(z)) \\ -\psi(z) \cdot \log(\bar{F}(t)) &= -\log(\bar{F}(t \cdot \chi(z))) \\ \log(\psi(z)) + \log(-\log(\bar{F}(t))) &= \log(-\log(\bar{F}(t \cdot \chi(z)))) \end{aligned}$$

mit $t = e^\tau$, $t \cdot \chi(z) = e^\tau \cdot e^{\chi(z)}$, $G(\tau) = \log(-\log(\tau))$ und $\lambda(z) = \log(\chi(z))$ ergibt sich somit:

$$\log(\psi(z)) + G(\tau) = G(\tau + \lambda(z))$$

Damit dies gilt, muß

$$G(\tau) = k\tau + d \quad \text{und} \quad k\lambda(z) = \log(\chi(z))$$

gelten, wobei mit d und k Konstanten sind. Somit erhalten wir

$$\begin{aligned} G(\tau) &= \log(-\log(\bar{F}(e^\tau))) \\ &= k \cdot \tau + \alpha \\ -\log(\bar{F}(e^\tau)) &= e^{k \cdot \tau} \cdot e^\alpha \\ \bar{F}(e^\tau) &= e^{-(e^\tau)^k \cdot e^\alpha} \quad \text{mit} \quad \rho = e^{\frac{\alpha}{k}} \\ \bar{F}(t) &= e^{-(\rho t)^k} \end{aligned}$$

Somit sind die beiden Modelle genau im Fall einer Weibull-Verteilung identisch.

Wir betrachten nun genauer ein parametrisiertes Modell:

$$h(t, z) = \psi(z, \beta) h_0(t) \quad ,$$

wobei sich z über die Zeit (für ein Individuum) nicht ändert.
In der Praxis werden 3 mögliche Parametrisierungen verwendet:

$$\begin{aligned}\psi(z, \beta) &= e^{\beta^T z} && \text{log-linear} \quad (\text{meist verwendete}) \\ &= 1 + \beta^T z && \text{linear} \\ &= \log(1 + \beta^T z) && \text{logistic}\end{aligned}$$

Welche der Form verwendet wird, kann durch Fitten einer übergeordneten Familie entschieden werden, zum Beispiel enthält

$$\psi(z, \beta; k) = (1 + k\beta^T z)^{\frac{1}{k}}$$

das lineare ($k = 1$) und das loglineare ($k \rightarrow 0$) Modell als Spezialfälle.

Nun soll β bei vollständig unbekanntem $h_0(t)$ geschätzt werden. Wir betrachten zunächst den Fall, daß keine Zensierung vorliegt und führen die folgenden Bezeichnungen ein:

- τ_1, \dots, τ_n seien die geordneten Ausfallzeitpunkte,
- γ_i sei die Marke des Objektes, welches zur Zeit τ_i ausfällt ($\gamma_i = l$, wenn $t_l = \tau_i$),
- $\mathcal{R}(\tau_j) = \{i : t_i \geq \tau_j\}$ sei die Menge aller Elemente unter Risiko gerade vor dem j -ten geordneten Ausfall und r_j die Länge von $\mathcal{R}(\tau_j)$.

Beispiel:

$$\begin{aligned}\mathcal{R}(\tau_1) &= \{1, 2, 3, 4\} \\ \mathcal{R}(\tau_2) &= \{1, 2, 4\} \\ \mathcal{R}(\tau_3) &= \{1, 2\} \\ \mathcal{R}(\tau_4) &= \{2\}\end{aligned}$$

Die Größen τ_j und γ_j sind gemeinsam äquivalent zur Originalstichprobe. Da h_0 unbekannt ist, können die τ_j keine Information über β geben. Daher betrachten wir die Größen γ_j . Die gemeinsame Verteilung $p(i_1, \dots, i_n)$ der γ_j über alle Permutationen von $(1, \dots, n)$ kann explizit bestimmt werden. Die bedingte Wahrscheinlichkeit $P(\gamma_j = i | \mathcal{H}_j)$, wobei $\mathcal{H}_j = \{\tau_1, \dots, \tau_j, i_1, i_2, \dots, i_{j-1}\}$ die Vergangenheit bis zum j -ten geordneten Ausfallzeitpunkt darstellt, ist die bedingte Wahrscheinlichkeit, daß i zur Zeit τ_j ausfällt, unter der Bedingung, daß ein Element des Risk-sets $\mathcal{R}(\tau_j)$ ausfällt:

$$P(\gamma_j = i | \mathcal{H}_j) = \frac{h_i(\tau_j)}{\sum_{k \in \mathcal{R}(\tau_j)} h_k(\tau_j)} = \frac{\psi(i)}{\sum_{k \in \mathcal{R}(\tau_j)} \psi(k)} \quad (\psi(k) = \psi(z_k, \beta)).$$

Man sieht, dass $P(\gamma_j = i | \mathcal{H}_j)$ unabhängig von τ_1, \dots, τ_j ist.

Daraus folgt für die gemeinsame Verteilung (Likelihoodfunktion):

$$\begin{aligned}p(i_1, \dots, i_n) &= \prod_{j=1}^n p_j(i_j | i_1, \dots, i_{j-1}) \\ &= \prod_{j=1}^n \frac{\psi(i_j)}{\sum_{k \in \mathcal{R}(\tau_j)} \psi(k)}\end{aligned}$$

Wir betrachten wieder das Beispiel:

$$\gamma_1 = 3 \quad , \quad \gamma_2 = 4 \quad , \quad \gamma_3 = 1 \quad , \quad \gamma_4 = 2$$

darausfolgt:

$$p(3, 4, 1, 2) = \frac{\psi(3)}{\psi(1) + \psi(2) + \psi(3) + \psi(4)} \cdot \frac{\psi(4)}{\psi(1) + \psi(3) + \psi(4)} \cdot \frac{\psi(1)}{\psi(1) + \psi(2)} \cdot \frac{\psi(2)}{\psi(2)}$$

Bei Zensierung kann auf gleiche Weise vorgegangen werden, wenn die Zensierung sofort nach einem Ausfall erfolgt. Da die Zensierung kaum Einfluß auf β besitzt, kann das angenommen werden.

Seien d die Anzahl der Ausfälle und $\tau_1 < \tau_2 < \dots < \tau_d$ die geordneten Ausfallzeiten. Dann läßt sich die gesamte Argumentation wiederholen, nur beinhaltet \mathcal{H}_j jetzt auch die Zensierung in $(0, \tau_j)$ und

$$L = \prod_{j=1}^d \frac{\psi(i_j)}{\sum_{k \in \mathcal{R}(\tau_j)} \psi(k)} .$$

Beispiel

$$\mathcal{R}_1 = (1, 2, 3, 4)$$

$$\mathcal{R}_2 = (1, 2)$$

$$\mathcal{R}_3 = (2)$$

A, B, C – mögliche Ausfallzeiten des zensierten Elementes

$$lik = \frac{\psi(3)}{\psi(1) + \psi(2) + \psi(3) + \psi(4)} \cdot \frac{\psi(1)}{\psi(1) + \psi(2)} \cdot \frac{\psi(2)}{\psi(2)}$$

$\hat{=}$ Summe der 3 möglichen Terme für einen Ausfall des zensierten Elementes:

$$A : \frac{\psi(3)}{\psi(1) + \psi(2) + \psi(3) + \psi(4)} \cdot \frac{\psi(4)}{\psi(1) + \psi(2) + \psi(4)} \cdot \frac{\psi(1)}{\psi(1) + \psi(2)} \cdot \frac{\psi(2)}{\psi(3)}$$

$$B : \frac{\psi(3)}{\psi(1) + \psi(2) + \psi(3) + \psi(4)} \cdot \frac{\psi(1)}{\psi(1) + \psi(2) + \psi(4)}$$

Ableiten der log-Likelihoodfunktion:

$$\begin{aligned} \frac{\partial}{\partial \beta_r} \psi(i) &= \psi_r(i) \\ \frac{\partial^2}{\partial \beta_r \partial \beta_s} \psi(i) &= \psi_{rs}(i) \end{aligned}$$

$$l = \sum_i \left[\log(\psi(i)) - \log \left(\sum_{k \in \mathcal{R}_i} \psi(k) \right) \right] = \sum l_i$$

$$\frac{\partial l_i}{\partial \beta_r} = \frac{\psi_r(i)}{\psi(i)} - \frac{\sum_{k \in \mathcal{R}_i} \psi_r(k)}{\sum_{k \in \mathcal{R}_i} \psi(k)}$$

Betrachten log-linearen Hazard:

$$\begin{aligned} \psi(z, \beta) &= e^{\beta^T z} \\ \psi_r(i) &= z_{ir} \psi(i) \\ \psi_{rs}(i) &= z_{ir} z_{is} \psi(i) \end{aligned}$$

z_{ir} - Wert der r -ten Komponente der erklärenden Variablen für das i -te Subjekt

$$\begin{aligned} \leadsto \quad \frac{\partial l_i}{\partial \beta_r} &= z_{ir} - \frac{\sum_{k \in \mathcal{R}_i} z_{kr} e^{\beta^T z_k}}{\sum_{k \in \mathcal{R}_i} z_{kr} e^{\beta^T z_k}} \\ &= z_{ir} - A_{ir}(\beta) \end{aligned}$$

z_{ir} - Wert der erklärenden Variablen des ausgefallenen Subjektes

$A_{ir}(\beta)$ - gewichtetes Mittel der selben Variablen über das gesamte Risk set

Schätzung der Baseline Hazard-Funktion

1. Wenn h_0 parametrisch spezifiziert ist:

$$h_0(t, \phi)$$

entweder gemeinsame Likelihood $l(\beta, \phi)$ wird maximiert, oder $l(\hat{\beta}, \phi)$ (conditional likelihood) wird maximiert.

2. Häufiger sind nicht parametrische Schätzungen:
einfacher Schätzer:

$$H(t) = \sum_{i=1}^n H_i(y_i) \quad - \quad \text{total cumulative hazard}$$

mit $y_i = \begin{cases} t & \text{wenn noch nicht ausgefallen} \\ c \text{ oder } t_i & \text{bei Zensierung oder Ausfall} \end{cases}$

Da für $H(t)$ gilt:

$$H(t) = \int_0^t \sum_{j \in \mathcal{R}(u)} \psi(j) h_0(u) du$$

folgt:

$$\hat{H}_o(t) = \sum_{\tau_j < t} \frac{d_j}{\sum_{l \in \mathcal{R}(\tau_j)} \hat{\psi}(l)} \quad .$$

8.3 Mischverteilungen und Frailty-Modelle

Unterschied zu Proportional Hazard: die erklärende Variable ist unbekannt und nicht meßbar!

Sei

$$\begin{aligned} h(t) &= Y \cdot h_0(t) & Y \text{ ist Zufallsvariable} \\ h_0 &= 1 & : \text{ Exponentialverteilt } (\lambda \text{ wird in } Y \text{ getreckt}) \\ h_0 &= \beta t^{\beta-1} & : \text{ Weibullverteilung.} \end{aligned}$$

Y heißt mischende Verteilung (mixture distribution)

Wenn $Y = y$ (fixiert): bedingte Verteilung (conditional distribution), beobachtet wird eine Randverteilung, wenn Y wegintegriert ist, mit marginal hazard $w(t)$.

$$S(t|Y) = \exp \left\{ -Y \int_0^t h_0(t) dt \right\}$$

Darausfolgt:

$$S(t) = ES(t|Y) \quad .$$

Wir betrachten nun einige Beispiele:

1.

$$Y = \begin{cases} y_1 & p \\ y_2 & 1-p \end{cases}$$

für $y_1 = 1$ und $y_2 = -1$:

$$h(t) = 1 - 2 \frac{1-p}{pe^{-2t} - (1-p)}$$

$h(t)$ ist fallend.

2.

$$\begin{aligned} S(t|Y) &= \exp\{-Yt^\beta\} \\ f_Y(y) &= \theta^\delta y^{\delta-1} \exp\{-\theta y\} \frac{1}{\Gamma(\delta)} \\ S(t) &= \left(\frac{\theta}{t^\beta + \theta} \right)^\delta \end{aligned}$$

Pareto fallende Ausfallrate.

3. Weibull.

Mischung beider Modelle (proportional Hazard und frailty):

$$\begin{aligned}h(t) &= h_0(t) \exp\{\beta'z + \psi'\omega\} \\&= \underbrace{Y(\omega, \psi)} h_0(t) \exp\{\beta'z\} \\&= Y, \text{ da sowieso alles unbekannt ist.}\end{aligned}$$

Darausfolgt:

$$h(t) = \frac{\lambda_0(t) \exp\{\beta'z\}}{1 + \Lambda_0 \exp\{\beta'z\} \cdot \frac{1}{\theta}}$$

(durch Ausintegrieren und weglassen des Skalenparameters. $\delta = \theta$.)

A Wahrscheinlichkeitsmodell

Exponentialverteilung¹

A.1 Einparametrische Exponentialverteilung

Wir betrachten zunächst die einparametrische Exponentialverteilung

$$F(t) = 1 - \exp(-\lambda t), \quad t \geq 0, \quad \lambda \geq 0. \quad (\text{A.1})$$

Die Beobachtungen lassen sich auf verschiedene Weise gewinnen:

1. Es werden n Elemente vom Zeitpunkt $t = 0$ an beobachtet. Sie fallen in einer zufälligen Reihenfolge aus. Die Beobachtung wird nach einer fest vorgegebenen Anzahl von Ausfällen $r = r^*$ beendet. Die Beobachtungsdauer ist zufällig.
2. Es werden n Elemente vom Zeitpunkt $t = 0$ an beobachtet. Ausgefallene Elemente werden sofort durch neue ersetzt. Das beeinträchtigt die Gültigkeit des Wahrscheinlichkeitsmodells nicht, weil die Exponentialverteilung durch die Konstanz der Ausfallrate die Eigenschaft „gebraucht so gut wie neu“ hat. Die Beobachtung wird nach einer fest vorgegebenen Anzahl von Ausfällen $r = r^*$ beendet. Die Beobachtungsdauer ist zufällig.
3. Es werden n Elemente vom Zeitpunkt $t = 0$ an beobachtet. Sie fallen in einer zufälligen Reihenfolge aus. Die Beobachtung wird nach einer fest vorgegebenen Zeit $t = t^*$ beendet. Die Anzahl der Ausfälle r ist zufällig.
4. Es werden n Elemente vom Zeitpunkt $t = 0$ an beobachtet. Ausgefallene Elemente werden sofort durch neue ersetzt. Die Beobachtung wird nach einer fest vorgegebenen Zeit $t = t^*$ beendet. Die Anzahl der Ausfälle r ist zufällig.

Diese vier Möglichkeiten der Datengewinnung sind bei der Schätzung zu unterscheiden. Sie werden getrennt beschrieben.

A.1.1 Beendigung beim r^* -ten Ausfall ohne Ersetzen der ausgefallenen Elemente

Die Beobachtung wird beendet, wenn die vorgegebene Anzahl von r^* Ausfällen erreicht ist. In diesem Fall ist $r \leq n$ eine feste Größe, und die Testdauer $t_{(r)}$ ist zufällig. Die

¹Der Anhang ist vollständig dem Buch HÄRTLER entnommen

Ausfalldaten bilden die Folge $t_{(1)} \leq t_{(2)} \leq \dots \leq t_{(r)}$. Die Likelihood-Funktion lautet in diesem Fall:

$$L(\lambda|r^*) = \frac{n!}{(n-r)!} \prod_{i=1}^r \lambda^r e^{-\lambda t_{(i)}} e^{-\lambda(n-r)t_{(r)}} = \frac{n!}{(n-r)!} \lambda^r e^{-\lambda [\sum_{i=1}^r t_{(i)} + (n-r)t_{(r)}]} \quad . \quad (\text{A.2})$$

Man erhält daraus die Maximum-Likelihood-Schätzung für λ in der Form:

$$\hat{\lambda} = \frac{r}{\sum_{i=1}^r t_{(i)} + (n-r)t_{(r)}} \quad . \quad (\text{A.3})$$

Den Nenner von A.3 bezeichnet man als summierte Lebensdauer (total time on test) S_r :

$$S_r = \sum_{i=1}^r t_{(i)} + (n-r)t_{(r)} \quad . \quad (\text{A.4})$$

Die Maximum-Likelihood-Schätzung $\hat{\lambda}$ ist in diesem Fall eine erwartungstreue Schätzung für λ und besitzt Minimalstreuung.

Das Vertrauensintervall für λ wird mit Hilfe der χ^2 -Verteilung gebildet. Epstein und Sobel zeigten, daß die Größe $2r(\lambda|\hat{\lambda})$ einer χ^2 -Verteilung mit $2r$ Freiheitsgraden folgt, wenn λ die unbekannte wahre Ausfallrate bedeutet. Somit läßt sich ein Vertrauensintervall für λ auf dem Vertrauensniveau $(1-\alpha)$ durch das oder die entsprechenden Quantile der χ^2 -Verteilung bilden. Die χ^2 -Verteilung mit $2r$ Freiheitsgraden wird symbolisch mit $\chi^2(2r)$ bezeichnet, das Quantil mit $\chi_{\gamma}^2(2r)$. Man erhält auf dieser Grundlage

1. die einseitige obere Vertrauensgrenze für λ

$$\lambda_{1-\alpha} = \frac{\chi_{1-\alpha}^2(2r)}{2S_r} \quad (\text{A.5})$$

2. die einseitige untere Vertrauensgrenze für λ

$$\lambda_{\alpha} = \frac{\chi_{\alpha}^2(2r)}{2S_r} \quad (\text{A.6})$$

3. die zweiseitige untere und obere Vertrauensgrenze für λ

$$\lambda_{\frac{\alpha}{2}} = \frac{\chi_{\frac{\alpha}{2}}^2(2r)}{2S_r} \quad , \quad \lambda_{1-\frac{\alpha}{2}} = \frac{\chi_{1-\frac{\alpha}{2}}^2(2r)}{2S_r} \quad . \quad (\text{A.7})$$

Wird in A.1 statt λ der Parameter $\vartheta = \lambda^{-1}$ verwendet, so ist die Maximum-Likelihood-Schätzung für ϑ einfach der zu A.3 reziproke Ausdruck

$$\hat{\vartheta} = \frac{1}{r} \left[\sum_{i=1}^r t_{(i)} + (n-r)t_{(r)} \right] \quad . \quad (\text{A.8})$$

Die Grenzen der Vertrauensintervalle folgen durch entsprechende Umstellungen aus den Gleichungen A.5 bis A.7:

1. einseitige untere Vertrauensgrenze für ϑ

$$\vartheta_\alpha = \frac{2S_r}{\chi_{1-\alpha}^2(2r)} \quad (\text{A.9})$$

2. einseitige obere Vertrauensgrenze für ϑ

$$\vartheta_{1-\alpha} = \frac{2S_r}{\chi_\alpha^2(2r)} \quad (\text{A.10})$$

3. zweiseitige untere und obere Vertrauensgrenze für ϑ

$$\vartheta_{\frac{\alpha}{2}} = \frac{2S_r}{\chi_{1-\frac{\alpha}{2}}^2(2r)} \quad , \quad \vartheta_{1-\frac{\alpha}{2}} = \frac{2S_r}{\chi_{\frac{\alpha}{2}}^2(2r)} \quad . \quad (\text{A.11})$$

Aus den Formeln A.5 bis A.11 ergibt sich, daß die Breite des Vertrauensintervalls entscheidend durch $r = r^*$ beeinflusst wird. Wählt man r^* groß, so wird auch S_r groß; die χ^2 -Verteilung hat eine entsprechend größere Anzahl von Freiheitsgraden, und es ergeben sich relativ engere Vertrauensgrenzen.

Für die Planung von Versuchen ist es wichtig zu wissen, in welcher Zeit etwa die r^* Ausfälle auftreten können. Diese Zeit hängt natürlich vom wahren und unbekannten Wert λ ab. Aber es interessiert auch die relative Verlängerung der Beobachtungsdauer bei einer Vergrößerung von r^* . Nach Epstein erhält man den Erwartungswert der Zeit bis zum r -ten Ausfall durch die Beziehung

$$E(t_{(r)}|n) = \vartheta \sum_{i=1}^r \frac{1}{n-i+1} \quad . \quad (\text{A.12})$$

Bild A.1 zeigt für $\vartheta = 1$ und $n = 10, 20, 50, 100$ den Quotienten $\frac{r}{n}$ über dem Erwartungswert der Zeit bis zum r -ten Ausfall $E(t_{(r)}|n)$. Für kleinere Werte von $\frac{r}{n}$ stimmen die Kurven fast überein; spürbare Unterschiede treten erst bei großen Werten $\frac{r}{n}$ auf. Für praktische Zwecke und relativ kleine Werte $\frac{r}{n}$ reicht es daher meistens, die aus dem empirischen Quantil $F(t_{(r)}) = \frac{r}{n}$ folgende Näherung

$$E(t_{(r)}|n) \approx \vartheta \ln \frac{n}{n-r} \quad (\text{A.13})$$

zu verwenden. Diese Kurve ist für $n = 100$ im Bild 4.1 gestrichelt dargestellt.

In der Praxis beschränkt sich die Parameterschätzung nur selten auf die Aufgabe, λ oder ϑ zu schätzen. Meistens sind auch davon abgeleitete Größen zu ermitteln. eine solche Größe ist das Quantil t_γ , das durch die Beziehung

$$P(t < t_\gamma) = \gamma$$

definiert ist. Für die Exponentialverteilung folgt durch Umstellung von Gleichung A.1

$$t_\gamma = \vartheta \ln \frac{1}{1-\gamma} \quad . \quad (\text{A.14})$$

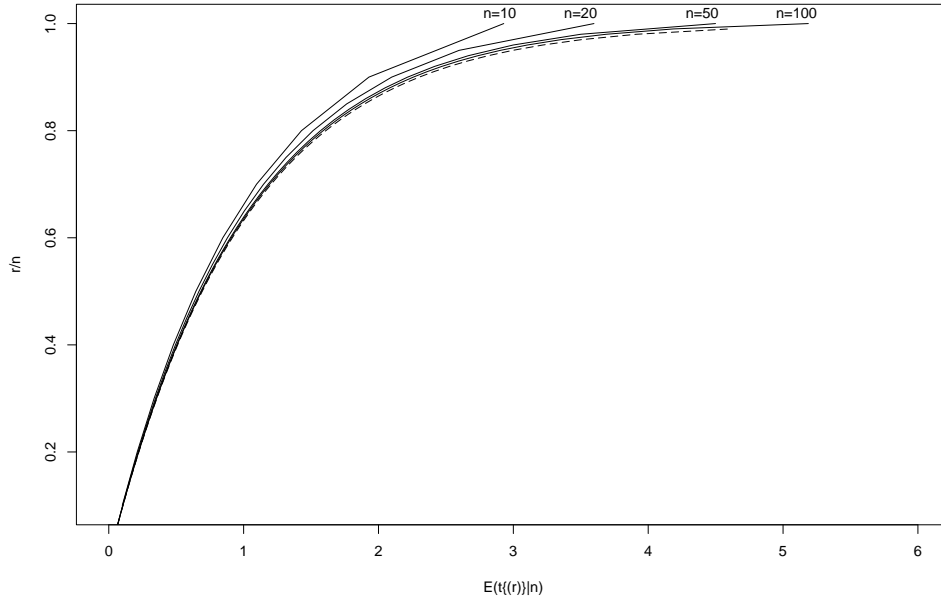


Abbildung A.1: Erwartungswert der Zeit bis zum r -tem Ausfall

Der Wert t_γ gibt die Zeit an, die mit einer vorgegebenen Wahrscheinlichkeit $(1 - \gamma)$ überlebt wird. Die Maximum-Likelihood-Schätzung von t_γ erhält man durch Einsetzen von A.8 in A.14:

$$\hat{t}_\gamma = \hat{\vartheta} \ln \frac{1}{1 - \gamma} \quad . \quad (\text{A.15})$$

In der Praxis sind auch die Vertrauensgrenzen $t_{\gamma,\alpha}, t_{\gamma,1-\alpha}$ oder $t_{\gamma,\frac{\alpha}{2}}, t_{\gamma,1-\frac{\alpha}{2}}$ von Bedeutung. Sie heißen Toleranzgrenzen und haben die Bedeutung

$$P(t_{\gamma,\alpha} \leq t_\gamma) \geq 1 - \alpha \quad . \quad (\text{A.16})$$

Je nachdem, ob man eine untere, obere oder zweiseitige Toleranzgrenze für t_γ wünscht, wird A.9, A.10 oder A.11 in A.14 eingesetzt. Die einseitige untere Toleranzgrenze $t_{\gamma,\alpha}$ auf dem Vertrauensniveau $(1 - \alpha)$ hat also die Gestalt

$$t_{\gamma,\alpha} = \frac{2S_r}{\chi_{1-\alpha}^2(2r)} \ln \frac{1}{1 - \gamma} \quad . \quad (\text{A.17})$$

Sie gibt die Zeit an, von der mit der Irrtumswahrscheinlichkeit α gesagt werden kann, daß sie von den betrachteten Elementen mit der Wahrscheinlichkeit $(1 - \gamma)$ überlebt wird.

In manchen praktischen Fällen benötigt man die Überlebenswahrscheinlichkeit zu einem festen Zeitpunkt t' . Sie wird mit

$$R(t') = e^{-\lambda t'} \quad (\text{A.18})$$

bezeichnet. Die Schätzung erhält man, wenn in A.18 für λ die Maximum-Likelihood-Schätzung $\hat{\lambda}$ (nach A.3) eingesetzt wird, d.h.

$$\hat{R}(t') = e^{-\hat{\lambda}t'} \quad . \quad (\text{A.19})$$

Eine untere, obere oder zweiseitige Vertrauensgrenze folgt entsprechend durch Einsetzen von A.5, A.6 oder A.7 in A.18. Die untere Vertrauensgrenze $R(t')_\alpha$ auf dem Vertrauensniveau $(1 - \alpha)$ hat also die Gestalt

$$R(t')_\alpha = \exp \left[- \frac{\chi_{1-\alpha}^2(2r)}{2S_r} t' \right] \quad . \quad (\text{A.20})$$

Sie gibt die Überlebenswahrscheinlichkeit zum Zeitpunkt t' an, die mit der Irrtumswahrscheinlichkeit α für die Elemente der untersuchten Grundgesamtheit gilt.

Beispiel A.1 Es sei $n = 50$ und $r^* = 15$. Im Test ergaben sich die Ausfallzeitpunkte

$$\begin{array}{lll} t_{(1)} = 125 \text{ h} & t_{(2)} = 252 \text{ h} & t_{(3)} = 319 \text{ h} \\ t_{(4)} = 401 \text{ h} & t_{(5)} = 470 \text{ h} & t_{(6)} = 691 \text{ h} \\ t_{(7)} = 965 \text{ h} & t_{(8)} = 1377 \text{ h} & t_{(9)} = 1462 \text{ h} \\ t_{(10)} = 1498 \text{ h} & t_{(11)} = 1674 \text{ h} & t_{(12)} = 1902 \text{ h} \\ t_{(13)} = 1907 \text{ h} & t_{(14)} = 2299 \text{ h} & t_{(15)} = 2312 \text{ h} \quad . \end{array}$$

Die summierte Lebensdauer beträgt nach Gleichung A.4

$$S_{15} = 17654 + 35 \cdot 2312 = 98574 \text{ h} \quad .$$

Daraus folgt der Schätzwert für λ

$$\hat{\lambda} = \frac{15}{98574} = 1,5 \cdot 10^{-4} \text{ h}^{-1} \quad .$$

Nun soll die einseitige obere Vertrauensgrenze für λ auf dem Vertrauensniveau $1 - \alpha = 0,9$ berechnet werden. Dazu wird Gleichung A.5 verwendet. Aus der Tabelle der χ^2 -Verteilung erhalten wir den Wert $\chi_{0,9}^2(30) = 40,3$; dann folgt aus A.5

$$\lambda_{0,9} = 2,0 \cdot 10^{-4} \text{ h}^{-1} \quad .$$

Der wahre Wert λ ist also mit der Irrtumswahrscheinlichkeit $\alpha = 0,1$ kleiner als $2 \cdot 10^{-4} \text{ h}^{-1}$.

Außerdem soll das 0,01-Quantil $t_{0,01}$ geschätzt werden, d.h. der Zeitpunkt, der von 99% der Elemente überlebt wird. Nach Gleichung A.15 erhält man

$$\hat{t}_{0,01} = 66 \text{ h} \quad .$$

Die untere Toleranzgrenze für $t_{0,01}$ auf dem Vertrauensniveau 0,9 folgt aus A.17 und beträgt

$$t_{0,01;0,1} = 49 \text{ h} \quad .$$

Das bedeutet: Man kann mit einer Irrtumswahrscheinlichkeit von $\alpha = 0,1$ sage, daß 99% der Elemente länger als 49 h leben werden.

Schließlich sei noch die Überlebenswahrscheinlichkeit von $t' = 1000$ h zu schätzen. Nach Gleichung A.19 gilt

$$\hat{R}(t') = 0,86 \quad .$$

Die untere Vertrauensgrenze auf dem Vertrauensniveau 0,9 ist nach Gleichung A.20

$$R(t')_{0,1} = 0,815 \quad .$$

Das bedeutet: Man kann mit einer Irrtumswahrscheinlichkeit $\alpha = 0,1$ sagen, daß mindestens 81,5% der Elemente länger als 1000 h leben werden. \diamond

A.1.2 Beendigung nach t^* Stunden ohne Ersetzen der ausgefallenen Elemente

Die Beobachtung wird zu einem vorgegebenen Zeitpunkt t^* beendet. Dabei ist die Anzahl der Ausfälle r zufällig. Es gilt stets $r \leq n$, und es kann auch der Fall $r = 0$ eintreten. Die Ausfalldaten bilden die Folge $t_{(1)} \leq t_{(2)} \leq \dots \leq t_{(r)} \leq t^*$. Die Likelihood-Funktion lautet:

$$L(\lambda|t^*) = \frac{n!}{(n-r)!} \lambda^r e^{-\lambda [\sum_{i=1}^r t_{(i)} + (n-r)t^*]} \quad . \quad (\text{A.21})$$

Daraus folgt die Maximum-Likelihood-Schätzung

$$\hat{\lambda} = \frac{r}{\sum_{i=1}^r t_{(i)} + (n-r)t^*} \quad . \quad (\text{A.22})$$

Für $r = 0$ wird auch $\hat{\lambda} = 0$. In diesem Fall hat die Likelihood-Funktion die Gestalt

$$L(\lambda|t^*, r = 0) = e^{-\lambda n t^*} \quad .$$

Das Maximum dieser Funktion liegt dann am Rande ihres Definitionsgebiets bei $\lambda = 0$. Der Nenner von A.22 ist wieder die summierte Lebensdauer

$$S_{t^*} = \sum_{i=1}^r t_{(i)} + (n-r)t^* \quad . \quad (\text{A.23})$$

Sie ist von t^* abhängig. Dadurch verändert sich die Likelihood-Funktion mit wachsendem t^* zwischen den Ausfällen kontinuierlich. Zu jedem Ausfallzeitpunkt ergibt sich aber eine sprunghafte Änderung der Likelihood-Funktion und damit auch des Schätzwerts $\hat{\lambda}$. Die Maximum-Likelihood-Schätzung ist in diesem Fall nicht erwartungstreu. Für großes n und r sind die Eigenschaften dieser Schätzung jedoch denen nach Gleichung A.3 ähnlich.

Bei der Bildung des Vertrauensintervalls für λ ist zu berücksichtigen, daß r eine Zufallsgröße ist. Für die wahre Ausfallrate λ ist die Ausfallwahrscheinlichkeit p durch die Beziehung

$$p = 1 - e^{-\lambda t^*} \quad (\text{A.24})$$

bestimmt. Die Wahrscheinlichkeit dafür, in einer n -Elemente-Stichprobe genau r Ausfälle zu erhalten, ergibt sich für bekanntes p aus der Binomialverteilung

$$p_r = \binom{n}{r} p^r (1-p)^{(n-r)} \quad , \quad r = 0, 1, \dots, n \quad . \quad (\text{A.25})$$

Nach der Beobachtung liegen n und r fest. Die Frage nach dem Vertrauensintervall bedeutet, festzustellen, welchen p - bzw. λ -Werte bei gegebenem n zu den r Ausfällen geführt haben können. Dazu wird A.25 als Betaverteilung abgesehen. Die Vertrauensgrenzen für p auf dem Vertrauensniveau $1 - \alpha$ ergeben sich durch die entsprechenden Quantile der Betaverteilung mit den Parametern $a = r + 1$ und $b = n - r + 1$. Wird mit $B_\gamma(a, b)$ das Quantil der Betaverteilung bezeichnet, so erhält man die Vertrauensgrenzen für λ durch die für p aus A.24. Es folgt somit

1. die einseitige obere Vertrauensgrenze

$$\lambda_{1-\alpha} = \frac{1}{t^*} \ln \frac{1}{1 - B_{1-\alpha}(r+1, n-r+1)} \quad (\text{A.26})$$

2. die einseitige untere Vertrauensgrenze

$$\lambda_\alpha = \frac{1}{t^*} \ln \frac{1}{1 - B_\alpha(r+1, n-r+1)} \quad (\text{A.27})$$

3. die zweiseitige untere und obere Vertrauensgrenze

$$\begin{aligned} \lambda_{\frac{\alpha}{2}} &= \frac{1}{t^*} \ln \frac{1}{1 - B_{\frac{\alpha}{2}}(r+1, n-r+1)} \\ \lambda_{1-\frac{\alpha}{2}} &= \frac{1}{t^*} \ln \frac{1}{1 - B_{1-\frac{\alpha}{2}}(r+1, n-r+1)} \quad . \end{aligned} \quad (\text{A.28})$$

Die Schätzung für ϑ erhält man als reziproken Wert zu $\hat{\lambda}$ nach Gleichung A.22, und die Vertrauensgrenzen für ϑ folgen durch entsprechende Umstellungen aus den Gleichungen A.25 bis A.28. Ähnlich verhält es sich mit der Schätzung des γ -Quantils t_γ und der Überlebenswahrscheinlichkeit $R(t')$; sie werden wie im vorigem Abschnitt aus den Schätzwerten für λ und den Vertrauensgrenzen abgeleitet.

Beispiel A.2 Die Beobachtung aus Beispiel 1 wird bis zum Zeitpunkt $t^* = 2500$ h durchgeführt. Es ist also $n = 50, r = 15$. Die Ausfallzeitpunkte sind die im Beispiel 1 angegebenen Zahlen. Aus Gleichung A.23 erhält man

$$S_{2500 \text{ h}} = 17654 + 35 \cdot 2500 = 105154 \text{ h} \quad .$$

Also ist

$$\hat{\lambda} = \frac{15}{105154} = 1,43 \cdot 10^{-4} \text{ h}^{-1} \quad .$$

Der Schätzwert für λ ist natürlich etwas kleiner als der im Beispiel 1, denn die Zeit zwischen den letzten beiden Ausfällen bei 2312 h und 2500 h verlief ohne Ausfall. Daraus folgt ein kleinerer Schätzwert der Ausfallrate.

Nun soll die einseitige obere Vertrauensgrenze auf dem Vertrauensniveau $1 - \alpha = 0,9$ berechnet werden. Dazu wird das 0,9-Quantil der Betaverteilung $B_{0,9}(16, 36)$ benötigt. In der Tabelle der Betaverteilung sind aber nur die unteren Quantile angegeben. Auf Grund der Symmetriebeziehung erhält man das gesuchte 0,9-Quantil.

Aus der Tabelle erhält man mit Hilfe einer linearen Interpolation den Wert

$$B_{0,1}(36, 16) = 0,60 \quad .$$

Daraus folgt

$$B_{0,9}(16, 36) = 0,40$$

und nach Gleichung A.26

$$\lambda_{0,9} = 2,0 \cdot 10^{-4} \text{ h}^{-1} \quad .$$

Steht keine Tabelle der Betaverteilung zur Verfügung, so wird die F -Verteilung mit $m_1 = 2a$ und $m_2 = 2b$ Freiheitsgraden verwendet. Mit Hilfe einer linearen Interpolation finden wir in der Tabelle der F -Verteilung, die in vielen Lehrbüchern über mathematische Statistik enthalten ist, den Wert $F_{0,9}(32, 72) = 1,46$. Daraus folgt der Wert $B_{0,9}(16, 36) = 0,39$ und $\lambda_{0,9} = 2,0 \cdot 10^{-4} \text{ h}^{-1}$. \diamond

Lebensdaueruntersuchungen ohne Ersetzen der ausgefallenen Elemente, die zum Zeitpunkt t^* beendet werden, treten in der Praxis häufig in solchen Fällen auf, in denen die genauen Ausfallzeitpunkte $t_{(i)}$ nicht ermittelt werden können, weil nur die Anzahl der Ausfälle zu festen Zeitpunkten erfaßt wird. Dadurch wird die exakte Berechnung der summierten Lebensdauer S_{t^*} unmöglich. Eine Schätzung von λ kann dann mit Hilfe der Maximum-Likelihood-Schätzung für p durch die Beziehung A.24 erfolgen. Die Maximum-Likelihood-Schätzung für p ist einfach

$$\hat{p} = \frac{r}{n} \quad . \quad (\text{A.29})$$

Daraus folgt der Schätzwert λ' für λ in der Form

$$\lambda' = \frac{1}{t^*} \ln \frac{n}{n-r} \quad . \quad (\text{A.30})$$

Im Beispiel 2 ist $n = 50$, $r = 15$ und $t^* = 2500 \text{ h}$. Daraus folgt $\lambda' = 1,43 \cdot 10^{-4} \text{ h}^{-1}$; der Schätzwert stimmt in diesem Fall sogar mit $\hat{\lambda}$ überein.

A.1.3 Beendigung nach r^* Ausfällen mit Ersetzen der ausgefallenen Elemente

Die Beobachtung wird beendet, wenn die vorgegebene Anzahl von r^* , $r^* > 0$, Ausfällen erreicht wurde. Der Zeitpunkt $t_{(r)}$ ist eine Zufallsgröße. Weil ausgefallene Elemente sofort

durch neue ersetzt werden, befinden sich stets n Elemente im Versuch. Die Ausfallrate für die n -Elemente-Stichprobe ist gleich $(n\lambda)$, und die Likelihood-Funktion lautet für $r = r^*$:

$$L(\lambda|r^*) = (n\lambda)^r e^{-n\lambda t_{(r)}} \quad . \quad (\text{A.31})$$

Daraus erhält man die Maximum-Likelihood-Schätzung

$$\hat{\lambda} = \frac{r}{nt_{(r)}} \quad . \quad (\text{A.32})$$

Im Nenner steht wieder die summierte Lebensdauer, die in einem Versuch mit Ersetzen der ausgefallenen Elemente einfach das Produkt aus Stichprobenumfang und Beobachtungsdauer ist. Somit ist Gleichung A.32 ein Analogon zu A.3 und hat auch dieselben Eigenschaften. Insbesondere ist $\hat{\lambda}$ eine erwartungstreue Schätzung. Den Schätzwert für $\vartheta = \lambda^{-1}$ erhält man einfach als reziproken Wert zu $\hat{\lambda}$.

Das Vertrauensintervall für λ beruht auf der χ^2 -Verteilung mit $2r$ Freiheitsgraden. Wird die Differenz zwischen zwei aufeinanderfolgenden Ausfallzeitpunkten $x_i = t_{(i)} - t_{(i-1)}$, $i = 1, \dots, r$, $x_0 = 0$, gebildet, so läßt sich der letzte Ausfallzeitpunkt $t_{(r)}$ in Form einer Summe der x_i ausdrücken:

$$t_{(r)} = \sum_{i=1}^r x_i \quad .$$

Aus den Eigenschaften des Poissonschen Prozesses folgt, daß die x_i einer Exponentialverteilung mit dem Parameter $(n\lambda)$ genügen. Die Summe von r unabhängigen Zufallsgrößen mit einer durch einen Parameter $(n\lambda)$ charakterisierten Exponentialverteilung ist eine Zufallsgröße mit einer zweiparametrigen Gammaverteilung, und zwar mit den Parametern $b = r$ und $\nu = (n\lambda)$. Die Gammaverteilung wiederum ist einer χ^2 -Verteilung mit $2r$ Freiheitsgraden folgt. Die Vertrauensintervalle stimmen im betrachteten Versuchstyp also mit denen nach Gleichung A.5 bis A.7 überein; lediglich die summierte Lebensdauer wird in der Form

$$S_{(r)} = nt_{(r)} \quad (\text{A.33})$$

berechnet.

In einem Versuch, der nach der vorgegebenen Anzahl von Ausfällen r^* beendet wird, ist die Beobachtungsdauer eine Zufallsgröße. Der Erwartungswert der Zeit bis zum r -ten Ausfall hängt natürlich von λ bzw. $\vartheta = \lambda^{-1}$ ab. Es gilt

$$E(t_{(r)}|n) = \vartheta \frac{r}{n} \quad . \quad (\text{A.34})$$

Die relative Zeiteinsparung durch die Vergrößerung von n oder durch die Verringerung von r kann mit Hilfe von A.34 beurteilt werden. Es ist in der Praxis zu beachten, daß die Breite des Vertrauensintervalls durch r stark beeinflußt wird. Deshalb sollte r^* nicht zu klein gewählt werden.

Beispiel A.3 Ein Versuch mit $n = 10$ Elementen soll bis zum dritten Ausfall ($r^* = 3$) durchgeführt werden. Ausgefallene Elemente werden sofort ersetzt (man benötigt also

insgesamt 12 Elemente). Es wurde $t_{(3)} = 750$ h ermittelt. Nach Gleichung A.32 erhält man den Schätzwert

$$\hat{\lambda} = \frac{3}{7500} = 4 \cdot 10^{-4} \text{ h}^{-1} \quad .$$

Außerdem sollen die zweiseitigen Vertrauensgrenzen für λ auf dem Vertrauensniveau $(1 - \alpha) = 0,9$ berechnet werden. Dazu wird Formel A.7 benutzt. In einer Tabelle findet man die χ^2 -Werte $\chi_{0,05}^2(6) = 1,64$, $\chi_{0,95}^2(6) = 12,6$. Durch einsetzen in A.7 erhält man $\lambda_{0,05} = 1,1 \cdot 10^{-4} \text{ h}^{-1}$, $\lambda_{0,95} = 8,4 \cdot 10^{-4} \text{ h}^{-1}$.

Wäre der Versuch statt mit 10 Elementen mit 100 Elementen durchgeführt und ebenfalls nach dem dritten Ausfall beendet worden, so hätte er sehr viel schneller zum Ziel geführt. Angenommen, der dritte Ausfall wäre entsprechend dem Erwartungswert der Zeit bis zu dem betreffenden Ausfall nach Gleichung A.34 schon zum Zeitpunkt $t_{(3)} = 75$ h aufgetreten, so hätte die Schätzung folgende Werte ergeben:

$$\begin{aligned} \hat{\lambda} &= \frac{3}{7500} = 4 \cdot 10^{-4} \text{ h}^{-1} \\ \hat{\lambda}_{0,05} &= \frac{1,64}{15000} = 1,1 \cdot 10^{-4} \text{ h}^{-1} \\ \hat{\lambda}_{0,95} &= \frac{12,6}{15000} = 8,4 \cdot 10^{-4} \text{ h}^{-1} \quad . \end{aligned}$$

Die Werte stimmen mit denen im ersten Fall überein; der Vorteil beim Einsatz von 100 Elementen liegt also nur in der Zeiteinsparung.

Wären n und r proportional vergrößert worden, d.h. $n = 100$ und $r^* = 30$, und wäre der dreißigste Ausfall entsprechend A.34 zum Zeitpunkt $t_{(30)} = 750$ h aufgetreten, so ergäbe sich

$$\hat{\lambda} = \frac{30}{75000} = 4 \cdot 10^{-4} \text{ h}^{-1} \quad .$$

Aus der Tabelle der χ^2 -Verteilung mit 60 Freiheitsgraden erhält man $\chi_{0,05}^2(60) = 43,2$, $\chi_{0,95}^2(60) = 79,1$. Daraus folgt

$$\lambda_{0,05} = \frac{43,2}{150000} = 2,9 \cdot 10^{-4} \text{ h}^{-1} \quad , \quad \lambda_{0,95} = \frac{79,1}{150000} = 5,3 \cdot 10^{-4} \text{ h}^{-1} \quad .$$

ein gegenüber dem ersten Fall engeres Vertrauensintervall. Es ist bedingt durch die umfangreichere experimentelle Information. Der Versuch hatte gegenüber dem ersten Fall den 10-fachen Elementeaufwand bei gleicher Beobachtungsdauer. \diamond

A.1.4 Beendigung nach t^* Stunden mit Ersetzen der ausgefallenen Elemente

Die Beobachtung wird nach einer vorgegebenen Zeit t^* beendet. Ausgefallene Elemente werden sofort ersetzt, so daß sich während der gesamten Beobachtungsdauer stets n

Elemente im Versuch befinden. Die Anzahl der Ausfälle $r \geq 0$ ist eine zufällige Größe. Wie im vorhergehenden Fall folgt die Zeit zwischen den Ausfällen einer Exponentialverteilung mit dem Parameter $(n\lambda)$. Die Likelihood-Funktion hat die zu A.31 analoge Form

$$L(\lambda|t^*) = (n\lambda)^r e^{-n\lambda t^*} \quad . \quad (\text{A.35})$$

Daraus erhält man als Maximum-Likelihood-Schätzung den Ausdruck

$$\hat{\lambda} = \frac{r}{nt^*} \quad . \quad (\text{A.36})$$

Im Nenner steht wieder die summierte Lebensdauer

$$S_{t^*} = nt^* \quad . \quad (\text{A.37})$$

Bei der Konstruktion von Vertrauensintervallen für λ ist die Zufälligkeit von r zu berücksichtigen. Die im Versuch mit Ersetzen der ausgefallenen Elemente auftretende Folge von Ausfällen kann durch einen Poissonschen Prozeß mit der Intensität $n\lambda$ beschrieben werden. Es gilt für die Anzahl der Ausfälle r bei festem n und λ im Zeitintervall $[0, t^*]$ die Poisson-Verteilung

$$P(X = r) = \frac{1}{r!} e^{-\lambda nt^*} (\lambda nt^*)^r \quad , \quad r = 0, 1, 2, \dots \quad (\text{A.38})$$

Nach dem Experiment steht r fest, und es sind die Grenzen für λ zu finden, die mit der vorgegebenen Wahrscheinlichkeit $1 - \alpha$ zum beobachteten r geführt haben können. Die Poisson-Verteilung wird zu diesem Zweck so betrachtet, als wäre λ die Zufallsgröße und r ein fester Wert. Es besteht zwischen der Poisson-Verteilung und einer Gammaverteilung folgende Beziehung:

$$\sum_{i=r+1}^{\infty} \frac{1}{i!} e^{-\lambda nt^*} (\lambda nt^*)^i = \frac{(nt^*)^{r+1}}{\Gamma(r+1)} \int_0^{\lambda} u^r e^{-nt^*u} du \quad . \quad (\text{A.39})$$

Dadurch kann jedem Quantil der Poisson-Verteilung ein zur gleichen Wahrscheinlichkeit gehörendes Quantil der Gammaverteilung zugeordnet und eine untere (obere) Grenze für r in eine obere (untere) Grenze für λ transformiert werden. Da r eine ganze Zahl ist, läßt sich für beliebige λ -Grenzen nicht immer eine r -Grenze mit genau der gleichen Wahrscheinlichkeit finden. In solchen Fällen verwendet man die nächstmögliche Grenze; das ist die zur ganzen Zahl r gehörende Grenze, für die das Vertrauensniveau etwas größer als $1 - \alpha$ ist. Die Gammaverteilung ist mit der χ^2 -Verteilung verwandt, mit der, weil sie in Tabellenform vorliegt, die Vertrauensgrenzen für λ bevorzugt berechnet werden. Man erhält

1. die einseitige obere Vertrauensgrenze

$$\lambda_{1-\alpha} = \frac{\chi_{1-\alpha}^2(2r+2)}{2S_{t^*}} \quad (\text{A.40})$$

2. die einseitige untere Vertrauensgrenze

$$\lambda_\alpha = \frac{\chi_\alpha^2(2r)}{2S_{t^*}} \quad (\text{A.41})$$

3. die zweiseitige untere und obere Vertrauensgrenze

$$\lambda_{\frac{\alpha}{2}} = \frac{\chi_{\frac{\alpha}{2}}^2(2r)}{2S_{t^*}} \quad , \quad \lambda_{1-\frac{\alpha}{2}} = \frac{\chi_{1-\frac{\alpha}{2}}^2(2r+2)}{2S_{t^*}} \quad . \quad (\text{A.42})$$

Die unterschiedliche Anzahl von Freiheitsgraden hat ihre Ursache in der Ganzzahligkeit von r . Die Irrtumswahrscheinlichkeit ist für die unteren Grenzen in der Regel etwas kleiner als α bzw. $\frac{1}{\alpha}$.

Beispiel A.4 Wie im Beispiel 3 wird die Beobachtung mit $n = 10$ Elementen durchgeführt. Ausgefallene Elemente werden sofort ersetzt. Der Versuch dauert bis zum Zeitpunkt $t^* = 1000$ h. Es wurden $r = 3$ Ausfälle beobachtet. Daraus folgt

$$\hat{\lambda} = \frac{3}{10000} = 3 \cdot 10^{-4} \text{ h}^{-1}$$

Die geschätzte Ausfallrate ist kleiner als im Beispiel 3; denn die Zeit zwischen 750 h und 1000 h verging ohne einen weiteren Ausfall. Es sei wieder die zweiseitige Vertrauensgrenze auf dem Vertrauensniveau $1 - \alpha = 0,9$ zu berechnen. In der χ^2 -Tabelle findet man die Werte $\chi_{0,05}^2(6) = 1,64$, $\chi_{0,95}^2(8) = 15,5$. Durch Einsetzen in Gleichung A.42 erhält man

$$\lambda_{0,05} = 8,2 \cdot 10^{-5} \text{ h}^{-1} \quad , \quad \lambda_{0,95} = 7,8 \cdot 10^{-4} \text{ h}^{-1} \quad .$$

◇

Tabelle 1: Die wichtigsten Formeln dieses Abschnittes

Funktion	Beendigungsregel			
	t^*		r^*	
	Mit Ersatz	Ohne Ersatz	Mit Ersatz	Ohne Ersatz
S	nt^*	$\sum_{i=1}^r t_{(i)} + (n-r)t^*$	$nt_{(r)}$	$\sum_{i=1}^r t_{(i)} + (n-r)t_{(r)}$
$\hat{\lambda}$	$\frac{r}{S}$	$\frac{r}{S}$	$\frac{r}{S}$	$\frac{r}{S}$
$L(\lambda)$	$(n\lambda)^r e^{-\lambda S}$	$\frac{n!}{(n-r)!} \lambda^r e^{-\lambda S}$	$(n\lambda)^r e^{-\lambda S}$	$\frac{n!}{(n-r)!} \lambda^r e^{-\lambda S}$
λ_α	$\frac{\chi_\alpha^2(2r)}{2S}$	$\frac{1}{t^*} \ln \frac{1}{1-B_\alpha(r+1, n-r+1)}$	$\frac{\chi_\alpha^2(2r)}{2S}$	$\frac{\chi_\alpha^2(2r)}{2S}$
$\lambda_{1-\alpha}$	$\frac{\chi_{1-\alpha}^2(2r)}{2S}$	$\frac{1}{t^*} \ln \frac{1}{1-B_{1-\alpha}(r+1, n-r+1)}$	$\frac{\chi_{1-\alpha}^2(2r)}{2S}$	$\frac{\chi_{1-\alpha}^2(2r)}{2S}$

A.2 Zweiparametrische Exponentialverteilung

Die Wahrscheinlichkeitsverteilung

$$F(t)1 - \exp[-\lambda(t - t_0)] \quad , \quad t \geq t_0 \geq 0 \quad , \quad \lambda > 0 \quad (\text{A.43})$$

enthält zwei unbekannte Parameter, die auf der Basis von Beobachtungswerten geschätzt werden sollen. Wie im Fall der einparametrischen Form der Exponentialverteilung können die Daten in verschiedener Weise gewonnen werden: mit und ohne Ersetzen der ausgefallenen Elemente, mit Beendigung des Versuchs nach r^* Ausfällen oder nach der vorgegebenen Beobachtungsdauer t^* .

A.2.1 Beendigung beim r^* -ten Ausfall ohne Ersetzen der ausgefallenen Elemente

Die Beobachtung wird beendet, wenn die vorgegebene Anzahl von Ausfällen r^* erreicht worden ist. Die Versuchsdauer $t_{(r)}$ ist zufällig und $r^* \leq n$ eine feste Zahl. Die Ausfalldaten bilden die Folge $t_{(1)} \leq t_{(2)} \leq \dots \leq t_{(r)}$. Die Likelihood-Funktion hat die Gestalt

$$L(\lambda, t_0 | r^*) = \frac{n!}{(n-r)!} \lambda^r \exp \left\{ -\lambda \left[\sum_{i=1}^r (t_{(i)} - t_0) + (n-r)(t_{(r)} - t_0) \right] \right\} . \quad (\text{A.44})$$

Daraus folgt als Maximum-Likelihood-Schätzung für die Parameter λ und t_0 das Paar von Gleichungen

$$\hat{t}_0 = t_{(1)} , \quad \hat{\lambda} = \frac{r}{\sum_{i=1}^r t_{(i)} + (n-r)t_{(r)} - nt_{(1)}} . \quad (\text{A.45})$$

Wird in A.43 statt λ der Parameter $\vartheta = \lambda^{-1}$ verwendet, so ist $\hat{\vartheta}$ einfach der zu $\hat{\lambda}$ reziproke Wert.

Die Maximum-Likelihood-Schätzung ist im zweiparametrischen Fall verbesserungs- fähig; denn die nach Gleichungen A.45 gewonnenen Schätzwerte sind nicht erwartungstreu und besitzen nicht die kleinstmögliche Varianz. Eine verbesserte Schätzung erhält man durch die Ausdrücke

$$\begin{aligned} t'_0 &= \frac{r}{r-1} t_{(1)} - \frac{1}{n(r-1)} \left[\sum_{i=1}^r t_{(i)} + (n-r)t_{(r)} \right] \\ \lambda' &= \frac{(r-1)}{\sum_{i=1}^r t_{(i)} + (n-r)t_{(r)} - nt_{(1)}} . \end{aligned} \quad (\text{A.46})$$

Dieses Schätzverfahren ist erwartungstreu und besitzt Minimalvarianz. Es wurde von Epstein beschrieben.

Zur Konstruktion der Vertrauensintervalle für t_0 und λ benutzt man die Eigenschaften der Zufallsgrößen

$$h = 2n\lambda(t_{(1)} - t_0) \quad (\text{A.47})$$

und

$$g = \frac{2(r-1)\lambda}{\lambda'} \quad . \quad (\text{A.48})$$

Die Größen h und g sind voneinander unabhängig; h folgt einer χ^2 -Verteilung mit zwei Freiheitsgraden und g einer χ^2 -Verteilung mit $2(r-1)$ Freiheitsgraden. Somit besitzt der Quotient

$$W = \frac{2(r-1)h}{2g} = n\lambda'(t_{(1)} - t_0) \quad (\text{A.49})$$

eine F -Verteilung mit $m_1 = 2$ und $m_2 = 2(r-1)$ Freiheitsgraden.

Die Vertrauensgrenzen für t_0 werden mit Hilfe der F -Verteilung gebildet, indem in A.49 für W das entsprechende Quantil der F -Verteilung eingesetzt wird. Ist $F_\gamma(2, 2r-2)$ das durch

$$P(W \leq F_\gamma) = \gamma \quad (\text{A.50})$$

definierte Quantil, so ergibt sich als untere Vertrauensgrenze für t_0 der Wert

$$t_{0;1-\alpha} = t_{(1)} - \frac{F_{1-\alpha}(2, 2r-2)}{n\lambda'} \quad . \quad (\text{A.51})$$

Als obere Vertrauensgrenze für t_0 verwendet man den ersten Ausfallzeitpunkt $t_{(1)}$. Das Intervall

$$\left\{ t_{(1)} - \frac{F_{1-\alpha}(2, 2r-2)}{n\lambda'}, t_{(1)} \right\} \quad (\text{A.52})$$

ist somit ein Vertrauensintervall für t_0 auf dem Vertrauensniveau $(1-\alpha)$. Es gibt mit der Irrtumswahrscheinlichkeit α an, in welchem Zahlenbereich der Parameter t_0 zu erwarten ist.

Bei der Berechnung der Vertrauensgrenzen für t_0 lässt sich eine vereinfachte Methode zur Bestimmung von F_γ anwenden. Ist F_γ das durch Gleichung A.50 definierte Quantil der F -Verteilung mit $m_1 = 2$ und $m_2 = 2(r-1)$ Freiheitsgraden, so lässt sich folgender Zusammenhang herstellen:

$$F_\gamma = \left[\left(\frac{1}{1-\gamma} \right)^{\frac{1}{(r-1)}} - 1 \right] (r-1) \quad . \quad (\text{A.53})$$

Diese Beziehung ermöglicht die Berechnung von Vertrauensgrenzen für t_0 auch ohne die Verwendung von Tabellen der F -Verteilung.

Zur Berechnung der Vertrauensgrenzen für λ' wird von der durch A.48 definierten Größe ausgegangen. Sie folgt einer χ^2 -Verteilung mit $2(r-1)$ Freiheitsgraden. Somit lassen sich folgende Vertrauensgrenzen berechnen:

1. die einseitige obere Vertrauensgrenze

$$\lambda_{1-\alpha} = \frac{\chi_{1-\alpha}^2(2r-2)}{2S_{r-1}} \quad (\text{A.54})$$

2. die einseitige untere Vertrauensgrenze

$$\lambda_\alpha = \frac{\chi_\alpha^2(2r-2)}{2S_{r-1}} \quad (\text{A.55})$$

3. die zweiseitige untere und obere Vertrauensgrenze

$$\lambda_{\frac{\alpha}{2}} = \frac{\chi_{\frac{\alpha}{2}}^2(2r-2)}{2S_{r-1}} \quad , \quad \lambda_{1-\frac{\alpha}{2}} = \frac{\chi_{1-\frac{\alpha}{2}}^2(2r-2)}{2S_{r-1}} \quad . \quad (\text{A.56})$$

Dabei ist S_{r-1} die durch die Beziehung

$$S_{r-1} = \sum_{i=1}^r t_{(i)} + (n-r)t_{(r)} - nt_{(1)} \quad (\text{A.57})$$

definierte summierte Lebensdauer, d.h. die Summe der Zeit, die alle Elemente zwischen den Zeitpunkten $t_{(1)}$ und $t_{(r)}$ gelebt haben. Die Vertrauensgrenzen für ϑ' erhält man aus den Gleichungen A.54 bis A.56 nach den üblichen Umstellungen.

Beispiel A.5 Es ist $n = 50$ und $r^* = 5$ vorgegeben. Folgende Ausfallzeitpunkte wurden beobachtet: $t_{(1)} = 625$ h, $t_{(2)} = 752$ h, $t_{(3)} = 819$ h, $t_{(4)} = 901$ h, $t_{(5)} = 970$ h. Die summierte Lebensdauer nach Gleichung A.57 beträgt $S_4 = 16467$ h. Nach Gleichung A.46 erhält man die Schätzwerte

$$\lambda' = 2,4 \cdot 10^{-4} \text{ h}^{-1} \quad , \quad t'_0 = 543 \text{ h} \quad .$$

Außerdem sollen für λ die obere Vertrauensgrenze auf dem Vertrauensniveau $1-\alpha = 0,95$ und für t_0 die untere Vertrauensgrenze auf dem Vertrauensniveau $1-\alpha = 0,95$ bestimmt werden. Aus einer Tabelle der χ^2 -Verteilung entnimmt man den Wert $\chi_{0,95}^2(8) = 15,5$. Daraus folgt

$$\lambda_{0,95} = 4,7 \cdot 10^{-4} \text{ h}^{-1} \quad .$$

Mit einer Irrtumswahrscheinlichkeit von $\alpha = 0,05$ ist die wahre Ausfallrate λ nicht größer als $4,7 \cdot 10^{-4} \text{ h}^{-1}$.

Aus einer Tabelle der F -Verteilung mit $m_1 = 2$ und $m_2 = 8$ Freiheitsgraden entnimmt man den Wert $F_{0,95}(2,8) = 4,46$, der auch nach Gleichung A.53 berechnet werden kann. Somit ergibt sich nach A.51 als untere Vertrauensgrenze auf dem Vertrauensniveau $1-\alpha = 0,95$ für t_0 der Wert

$$t_{0;0,95} = 253 \text{ h} \quad .$$

Das bedeutet, daß sich auf der Basis der Beobachtungswerte sagen läßt: Der wahre Parameter t_0 ist mit einer Irrtumswahrscheinlichkeit von 0,05 größer als 253 h. \diamond

A.2.2 Beendigung zum Zeitpunkt t^* ohne Ersetzen der ausgefallenen Elemente

Die Beobachtung wird zu einem festen Zeitpunkt beendet, und die Anzahl der Ausfälle ist zufällig. Es gilt $r \leq n$. Der Fall $r = 0$ ist möglich. Allerdings erhält man dann keine Schätzung für die beiden Parameter t_0 und λ . Die Grundlage für die Parameterschätzung bildet die Beobachtungsdauer und die Folge der Ausfallzeitpunkte $t_{(1)} \leq t_{(2)} \leq \dots \leq t_{(r)}$. Die Likelihood-Funktion lautet:

$$L(\lambda, t_0 | t^*) = \frac{n!}{(n-r)!} \lambda^r \exp \left\{ -\lambda \left[\sum_{i=1}^r (t_{(i)} - t_0) + (n-r)(t^* - t_0) \right] \right\}. \quad (\text{A.58})$$

Daraus lassen sich für die Maximum-Likelihood-Schätzung folgende Gleichungen herleiten:

$$\hat{t}_0 = t_{(1)}, \quad \hat{\lambda} = \frac{r}{\sum_{i=1}^r t_{(i)} + (n-r)t^* - nt_{(1)}}. \quad (\text{A.59})$$

Die Berechnung von Vertrauensgrenzen für t_0 und λ ist in diesem Fall nicht ohne weiteres möglich, weil die Eigenschaften der Schätzwerte vom wahren t_0 abhängig sind. Daher sind in der Praxis Experimente, die bei einer fest vorgegebenen Anzahl von Ausfällen beendet werden vorzuziehen.

A.2.3 Beendigung beim r^* -ten Ausfall mit Ersetzen der ausgefallenen Elemente

Die Beobachtung wird beendet, wenn die vorgegebene Anzahl von Ausfällen r^* erreicht worden ist. Ausgefallene Elemente werden sofort ersetzt, so daß sich stets n Elemente im Versuch befinden. Man braucht insgesamt $(n + r - 1)$ Testelemente. Die Versuchsdauer $t_{(r)}$ ist eine Zufallsgröße.

Das Modell A.43 bedeutet, daß bis zum Zeitpunkt t_0 keine Ausfälle möglich sind und das ab $t = t_0$ das Ausfallverhalten durch einen Poissonschen Prozeß mit der Intensität $(n\lambda)$ beschrieben werden kann. Die Likelihood-Funktion hat also die zu A.31 analoge Form

$$L(\lambda, t_0 | r^*) = (n\lambda)^r \exp \{ -n\lambda(t_{(r)} - t_0) \}. \quad (\text{A.60})$$

Daraus folgen die Maximum-Likelihood-Schätzungen

$$\hat{t}_0 = t_{(1)}, \quad \hat{\lambda} = \frac{r}{n(t_{(r)} - t_{(1)})}. \quad (\text{A.61})$$

Wie bei Experimenten ohne Ersetzen der ausgefallenen Elemente lassen sich die Schätzungen A.61 noch verbessern. Erwartungstreue Schätzungen mit Minimalstreuung sind gegeben durch die Gleichungen

$$t'_0 = \frac{r}{r-1}t_{(1)} - \frac{1}{r-1}t_{(r)}, \quad \lambda' = \frac{r-1}{n(t_{(r)} - t_{(1)})}. \quad (\text{A.62})$$

Die Konstruktion der Vertrauensintervalle entspricht dem Fall ohne Ersetzen der ausgefallenen Elemente (siehe Abschnitte vorher). Es werden enabhängige Hilfsgrößen h und g benutzt, wie sie durch Gleichungen A.47 und A.48 definiert sind. Daraus folgt für t_0 das Vertrauensintervall nach Gleichungen A.51 und A.52, und die Vertrauensintervalle für λ erhält man aus Gleichungen A.54 bis A.56, wobei die summierte Lebensdauer durch

$$S_{r-1} = n(t_{(r)} - t_{(1)}) \quad (\text{A.63})$$

gegeben ist.

Beispiel A.6 Ein Versuch mit $n = 10$ Elementen soll bis zum dritten Ausfall ($r^* = 3$) durchgeführt werden. Ausgefallene Elemente werden sofort ersetzt. Es wurden $t_{(1)} = 620$ h und $t_{(3)} = 1250$ h ermittelt.

Nach Gleichung A.62 erhält man die Schätzwerte $\lambda' = 3,2 \cdot 10^{-4} \text{ h}^{-1}$, $t'_0 = 305$ h. Für t_0 soll die untere Vertrauensgrenze auf dem Vertrauensniveau $1 - \alpha = 0,95$ berechnet werden. Dazu ist Gleichung A.51 anzuwenden. In einer Tabelle der F -Verteilung findet man für $F_{0,95}$ mit $m_1 = 2$ und $m_2 = 4$ Freiheitsgraden den Wert 6,94, den man auch nach Gleichung A.53 ausrechnen kann. Daraus folgt die untere Vertrauensgrenze

$$t_{0;0,95} = -1,6 \cdot 10^3 \text{ h} \quad .$$

Die Vertrauensgrenze ist negativ und daher Null zu setzen. Setzt man in A.51 den Wert $t_{0;1-\alpha} = 0$ ein, so findet man durch Umstellen von Gleichungen A.51 und A.53, daß dazu ein Vertrauensniveau von 0,75 geht. Aus den vorhandenen Daten kann man also nur mit einer Irrtumswahrscheinlichkeit von 0,25 auf einen Parameter $t_0 > 0$ schließen. Für λ sollen zweiseitige Vertrauensgrenzen auf dem Vertrauensniveau $1 - \alpha = 0,9$ bestimmt werden. Dazu sind die Gleichungen A.56 zu benutzen. Einer Tabelle der χ^2 -Verteilung mit vier Freiheitsgraden entnimmt man die Werte

$$\begin{aligned} \chi_{0,05}^2(4) &= 0,711 \\ \chi_{0,95}^2(4) &= 9,49 \quad . \end{aligned}$$

Da $S_{r-1} = 6300$ h ist, folgen die Grenzen

$$\begin{aligned} \lambda_{0,05} &= 5,6 \cdot 10^{-5} \text{ h}^{-1} \\ \lambda_{0,95} &= 7,5 \cdot 10^{-4} \text{ h}^{-1} \end{aligned}$$

◇

A.2.4 Beendigung zum Zeitpunkt t^* mit Ersetzen der ausgefallenen Elemente

Der Versuch wird zum Zeitpunkt t^* beendet; ausgefallene Elemente werden sofort ersetzt. Die Anzahl der Elemente während der Beobachtung ist stets gleich n . Die Anzahl der

Ausfälle r ist zufällig. Es kann also der Fall $r = 0$ eintreten, der keine Schätzung von t_0 und λ erlaubt. Das Modell einer zweiparametrischen Exponentialverteilung bedeutet, daß bis zum Zeitpunkt t_0 keine Ausfälle möglich sind und das Ausfallverhalten für die Zeit $t > t_0$ durch einen Poissonschen Prozeß mit der Ausfallrate $(n\lambda)$ beschrieben werden kann. Analog zu A.60 lautet die Likelihood-Funktion:

$$L(\lambda, t_0 | t^*) = (n\lambda)^r \exp \{-n\lambda(t^* - t_0)\} \quad . \quad (\text{A.64})$$

Daraus folgen die Maximum-Likelihood-Schätzungen

$$\hat{t}_0 = t_{(1)} \quad , \quad \hat{\lambda} = \frac{r}{n(t^* - t_{(1)})} \quad . \quad (\text{A.65})$$

Wie bei Experimenten mit Beendigung zum Zeitpunkt t^* ohne Ersetzen der ausgeallenen Elemente ist auch hier die Berechnung von Vertrauensintervallen für t_0 und λ nicht allgemein möglich; denn der Wert t_0 hat einen Einfluß auf die Varianz der Schätzwerte. So sind Experimente, die nach einer fest vorgegebenen Anzahl von Ausfällen r^* beendet werden, im Hinblick auf die Datenauswertung günstiger.

Tabelle 2: Die wichtigsten Formeln dieses Abschnittes

Funktion	Beendigungsregel t^*		r^*	
	Mit Ersatz	Ohne Ersatz	Mit Ersatz	Ohne Ersatz
S	$n(t^* - t_{(1)})$	$\sum_{i=1}^r t_{(i)} + (n - r)t^* - nt_{(1)}$	$n(t_{(r)} - t_{(1)})$	$\sum_{i=1}^r t_{(i)} + (n - r)t_{(r)} - nt_{(1)}$
$\hat{\lambda}$	$\frac{r}{S}$	$\frac{r}{S}$	$\frac{r}{S}$	$\frac{r}{S}$
\hat{t}_0	$t_{(1)}$	$t_{(1)}$	$t_{(1)}$	$t_{(1)}$
λ'	-	-	$\frac{r-1}{S}$	$\frac{r-1}{S}$
t'_0	-	-	$t_{(1)} - \frac{S}{n(r-1)}$	$t_{(1)} - \frac{S}{n(r-1)}$
λ_α	-	-	$\frac{\chi_\alpha^2(2r-2)}{2S}$	$\frac{\chi_\alpha^2(2r-2)}{2S}$
$\lambda_{1-\alpha}$	-	-	$\frac{\chi_{1-\alpha}^2(2r-2)}{2S}$	$\frac{\chi_{1-\alpha}^2(2r-2)}{2S}$
$t_{0;1-\alpha}$	-	-	$t_{(1)} - \frac{F_{1-\alpha}(2,2r-2)}{n\lambda'}$	$t_{(1)} - \frac{F_{1-\alpha}(2,2r-2)}{n\lambda'}$