

разложений методами информационной дифференциальной геометрии и выяснения геометрического смысла коэффициентов разложения было начато в 1980—82 гг. Амари, Барндорф-Нильсеном и их сотрудниками (см., например, [45], [44], [65], [58]). Более подробная библиография ранних работ дана в [47], [53], [156]; необходимо здесь также упомянуть пионерскую работу [115], возродившую интерес к геометрическим подходам. Можно считать установленным, что в регулярных задачах информационные количества определяют главный член асимптотики оптимального убывания погрешности оценок. При этом связь (среднего) количества информации и априорной «неопределенности» оценивания нелинейна и зависит от постановки задачи. Что же касается информационно-геометрических подходов к статистике малых выборок или больших выборок с ненормальной асимптотикой, то они только начинают обсуждаться [153]. Вопрос о существовании иных информационных количеств, связанных с парами распределений вероятностей будет затронут ниже в § 4.

§ 4. Задача различения нескольких простых гипотез

Задача проверки двух простых гипотез является классической статистической задачей, на которой отрабатывались многие принципы теоретической статистики. Однако даже здесь использование техники несимметричной информационной пифагоровой геометрии дает удобный язык для описания основных асимптотических закономерностей решения. В качестве количества информации здесь появляется в разных вариантах информационное отклонение I Кульбака (см. (0.13)), которое описывает скорость экспоненциального убывания $\exp(-NI)$ ошибки вывода для оптимального теста с ростом числа N независимых наблюдений.

Напомним основные контуры теории. Пусть по последовательности независимых одинаково распределенных исходов $\omega^1, \dots, \omega^N \in \Omega$ требуется решить, распределены ли они по мере $P_0(\cdot) \in \text{Car}(\Omega, \mathcal{A})$. Считается априори известным, что наблюдаемое явление либо описывается распределением P_0 , либо, если это не так, то распределением $P_1(\cdot) \in \text{Car}(\Omega, \mathcal{A})$. Таким образом, имеется всего два вывода: δ_0 — наблюдаем P_0 , δ_1 — наблюдаем P_1 . Каждое решающее правило Π_N вывода по последовательности $\varepsilon = (\omega^1, \dots, \omega^N)$ исходов описывается вероятностями $\Pi(\varepsilon; \delta_0) = \pi_0(\varepsilon)$ и $\Pi(\varepsilon, \delta_1) = \pi_1(\varepsilon) = 1 - \pi_0(\varepsilon)$ принять гипотезу δ_0 или ее альтернативу δ_1 . Обозначим $\alpha_N = (P_0^N \Pi_N)(\delta_1) = E_{(P_0^N)} \pi_1(\varepsilon)$ вероятность *ошибки первого рода*: ошибочно отвергнуть гипотезу $\delta_0 : P = P_0$, когда она верна. Аналогично, $\beta_N = (P_1^N \Pi_N)(\delta_0)$ обозначим вероятность *ошибки второго рода*: ошибочно принять δ_0 , в то время как распределение исходов задается распределением P_1 .