

Возможны две постановки задачи оптимального различения. Если ошибка первого рода приводит к существенно более тяжелым последствиям, чем ошибка второго рода, то стараются минимизировать вероятность α_N , удерживая величину β_N достаточно малой, т. е. минимизируя α_N в классе $B(b, N)$ решающих правил P_N с вероятностью ошибки $\beta_N \leq b < 1$. Если же ошибки примерно равноценны по последствиям, то стараются минимизировать $\max\{\alpha_N, \beta_N\}$.

Мы будем предполагать, что распределения вероятностей P_0 и P_1 взаимно абсолютно непрерывны. Именно такие гипотезы наиболее плохо тестируются (невозможно достоверное различение), и как раз для них приложим геометрический язык. В этой ситуации по фундаментальной лемме Неймана—Пирсона [171] достаточной статистикой является отношение правдоподобия — производная Радона—Никодима

$$(dP_1^N/dP_0^N)(\varepsilon) = [p_1(\omega^1)/p_0(\omega^1)] \cdot \dots \cdot [p_1(\omega^N)/p_0(\omega^N)], \quad (4.1)$$

или, что эквивалентно, *нормированная логарифмическая функция правдоподобия* $N^{-1} \ln (dP_1^N/dP_0^N)(\varepsilon) = f_N(\varepsilon)$,

$$f_N(\varepsilon) = N^{-1} [q(\omega^1) + \dots + q(\omega^N)], \quad q(\omega) = \ln p_1(\omega) - \ln p_0(\omega), \quad (4.2)$$

где $p_1(\omega)$ и $p_0(\omega)$ — плотности P_1 и P_0 по любой доминирующей мере. Так что всякий оптимальный тест имеет вид

$$\pi_0(\varepsilon) = 1 \text{ при } f_N(\varepsilon) < k_N, \quad \pi_0(\varepsilon) = 0 \text{ при } f_N(\varepsilon) > k_N,$$

где граница k_N критической области $K_N = \{\varepsilon : f_N(\varepsilon) > k_N\}$

определяется постановкой задачи оптимизации. Если при этом $P_{0,1}^N \{\varepsilon : f_N(\varepsilon) = k_N\} > 0$, то решение при таких значениях может быть принято путем рандомизации.

Теорема 4.1. Пусть распределения P_0 и P_1 взаимно абсолютно непрерывны. Тогда, для любого уровня $b < 1$ вероятности ошибки второго рода, вероятность α_N ошибки первого рода у оптимального теста убывает как $\exp[-NI(P_0 : P_1)]$, точнее

$$\lim_{N \rightarrow \infty} \inf_{P \in B(b, N)} N^{-1} \ln \alpha_N(P) = -I(P_0 : P_1). \quad (4.3)$$

Доказательство этой теоремы восходит к неопубликованной работе Стейна [193] и диссертации Джоши, см. [103], [155], [34]. Напомним его идею при $I(P_0 : P_1) < \infty$. Из условия $P(N) \in B(b, N)$ следует, что $\beta_N = E_1^{(N)} \pi_0(\varepsilon) \leq b$, т. е. $E_1^{(N)} \pi_1(\varepsilon) \geq 1 - b$. Далее, если исходы $\varepsilon = (\omega^1, \dots, \omega^N)$ распределены по P_1^N , то $f_N(\varepsilon) \Rightarrow -I(P_0 : P_1)$ в силу закона больших чисел Хинчина. Другими словами, если $U(\eta, N) = \{\varepsilon : f_N(\varepsilon) + I(P_0 : P_1) \geq \eta\}$,