6 Применение теории больших уклонений

6.1 Критерий Неймана-Пирсона

Рассмотрим критерий Неймана-Пирсона проверки простой гипотезы H_0 (X_i имеют плотность $f_0(x)$) с простой альтернативой H_1 (X_i имеют плотность $f_1(x)$). Допустим у нас есть выборка размера n, тогда критерий Неймана-Пирсона предлагает действовать следующим образом: отвергать гипотезу H_0, если выборка попала в множество D

$$D = \left\{ \frac{f_1(x_1)...f_1(x_n)}{f_0(x_1)....f_0(x_n)} > \gamma^n \right\}$$

и принимать ее в противном случае, при некотором $\gamma \in (0,1)$.

Рассмотрим ошибку первого рода $\alpha_n = \mathbf{P}_0((X_1,...,X_n) \in D)$, т.е. вероятность того, что гипотеза была верна, а мы ее отвергли. Как ведет себя α_n при $n \to \infty$?

$$\alpha_n = \mathbf{P}_0((X_1,...,X_n) \in D) = \mathbf{P}_0\left(\sum_{i=1}^{n} \ln\frac{f_1(X_i)}{f_0(X_i)} \geq n\ln\gamma\right).$$

Для $Y_1 = \ln\frac{f_1(X_1)}{f_0(X_1)}$ при гипотезе H_0

$$R(h) = \mathbf{E}_0 e^{h\ln\frac{f_1(X_i)}{f_0(X_i)}} = \int_{\mathbb{R}} e^{h\ln\frac{f_1(X)}{f_0(x)}} f_0(x)dx = \int_{\mathbb{R}} f_1^h(x) f_0^{1-h}(x)dx.$$

В частности, $R(1) = 1$.

При этом

$$\mathbf{E}_0 Y_1 = \int_{\mathbb{R}} \ln\frac{f_1(x)}{f_0(x)} f_0(x)dx \leq \ln\left(\int_{\mathbb{R}} \frac{f_1(x)}{f_0(x)} f_0(x)dx\right) = 0.$$

Аналогичным образом $\mathbf{E}_1 Y_1 \geq 0$. При этом оба неравенства строгие, если только Y_0 не является вырожденной величиной по мере одной из мер. Будем считать, что это не так.

Следовательно, $\mathbf{E}_0 Y_1 = \mu < 0$, $\mathbf{E}_1 Y_1 = \widetilde{\mu} > 0$ (будем считать, что $\mu > -\infty$, $\widetilde{\mu} < \infty$). При $\gamma < \mu$ α_n сходится к 1. При $\gamma > \mu$, $\mu < \ln\gamma < \widetilde{\mu}$ в силу теоремы Петрова

$$\alpha_n \sim \frac{C(\ln\gamma)}{h_{\ln\gamma}\sqrt{n}} \exp(-\Lambda(\ln\gamma)n).$$

Здесь

$$C(\theta) = \frac{1}{\sqrt{2\pi}\sigma(h_\theta)}, \ \Lambda(\theta) = \theta h_\theta - \ln R(h_\theta),$$

где
$$h_\theta : \int_{\mathbb{R}} f_1^{h_\theta}(x) f_0^{1-h_\theta}(x)(\ln f_1(x) - \ln f_0(x) - \theta)dx = 0.$$

Аналогично для ошибки второго рода
$$\beta_n = \mathbf{P}_1((X_1, ..., X_n) \notin D) = \mathbf{P}_1\left(\sum_{i=1}^n \ln \frac{f_1(X_i)}{f_0(X_i)} < n\ln\gamma\right).$$

Тогда при $-1 \le h < 0$
$$\widetilde{R}(h) = \mathbf{E}_1 e^{hY_1} = \int_{\mathbb{R}} f_1^{1+h}(x) f_0^h(x)dx = R(1+h),$$

$\widetilde{R}(-1) = 1$, $\widetilde{\mu} = \mathbf{E}_1 Y_1 < \infty$. При этом
$$\widetilde{m}(h) = m(1+h),\ \widetilde{\sigma}^2(h) = \sigma^2(1+h),\ \widetilde{h}_\theta : h_\theta - 1$$

при $\theta \in (\mu, \widetilde{\mu})$, откуда
$$\widetilde{\Lambda}(\theta) = \Lambda(\theta) - \theta.$$

Аналогично предыдущему
$$\beta_n \sim \frac{\widetilde{C}(\ln\gamma)}{(1-h_{\ln\gamma})\sqrt{n}} \exp(-\widetilde{\Lambda}(\ln\gamma)n)$$

при $\mu < \ln\gamma < \widetilde{\mu}$. Таким образом, при всех $\ln\gamma \in (\mu, \widetilde{\mu})$
$$\alpha_n \sim \frac{C(\ln\gamma)}{h_{\ln\gamma}\sqrt{n}} \exp(-\Lambda(\ln\gamma)n),\ \ \beta_n \sim \frac{C(\ln\gamma)}{(1-h_{\ln\gamma})\sqrt{n}} \exp(-\Lambda(\ln\gamma)n + n\ln\gamma).$$

6.2 Наилучший критерий с заданной функцией потерь

Пусть мы платим цену a за ошибку I рода, ценку b за ошибку второго и хотим минимизировать $a\alpha_n + b\beta_n$.

При $\gamma < 1-\delta$, $\delta > 0$, мы получим $\exp(-\Lambda(\ln\gamma)) > \exp(-\Lambda(0))$. Следовательно, при любых a, b
$$a\alpha_n + b\beta_n \le a\alpha_n \le C\exp(-\Lambda(\ln(1-\delta))n) \le C\exp(-\Lambda(0)(1-\varepsilon)n)$$

при некоторых $\varepsilon, C > 0$. Аналогично при $\gamma > 1+\delta$
$$a\alpha_n + b\beta_n \le b\eta_n \le C\exp(-\Lambda(\ln(1+\delta))n - \ln(1+\delta)n) \le C\exp(-\Lambda(0)(1-\varepsilon)n)$$

при некоторых $\varepsilon, C > 0$.

При этом при $\gamma = 1$ имеем

$$a\alpha_n + b\beta_n \sim De^{-\Lambda(0)n}.$$

Следовательно, при $\gamma = 1$ наши потери асимптотически будут меньше, чем при любом $\gamma \leq 1 - \delta$ или $\gamma \geq 1 + \delta$. Аналогичные рассуждения проходит при $\delta = \delta_n$, если только $n\delta_n \to \infty$. В этом случае $\Lambda(\ln(1 - \delta_n)n) = \Lambda(0)n - \delta_n n h_0 + o(\delta_n n)$, откуда мы получим

$$a\alpha_n + b\beta_n \leq a\alpha_n \leq C\exp(-\Lambda(0)n - \delta_n n h_0/2),$$

то есть ошибка опять же будет иметь больший порядок чем $De^{-\Lambda(0)n}$. Аналогичным образом рассматривается $\Lambda(\ln(1 + \delta_n)n - \ln(1 + \delta_n)n$. Таким образом, наилучшую потерю будет давать случай $\gamma_n = 1 + O(1/n)$, то есть γ^n порядка O(1).

Итак, можно считать, что $\ln\gamma = o(1/\sqrt{n})$. Тогда

$$\Lambda(\ln\gamma)n = \Lambda(0) + h_0\ln\gamma + o(1/n),\ n \to \infty, (\Lambda(\ln\gamma) - \ln\gamma)n = \Lambda(0) + (h_1 - 1)\ln\gamma + o(1/n), n \to \infty.$$

Таким образом,

$$a\alpha_n + b\beta_n \sim \frac{C(\gamma)}{\sqrt{n}} e^{-\Lambda(0)n}\left(\frac{a}{h_0}e^{-h_0 n\ln\gamma} + \frac{b}{(1-h_0)}e^{(1-h_0)n\ln\gamma}\right).$$

Максимум будет достигаться при $n\gamma = u$, где

$$be^{(1-h_0)u} = ae^{-h_0 u},\ u = \ln\frac{a}{b}.$$

Таким образом, наилучший асимптотический критерий имеет вид

$$\frac{L(x_1, ..., x_n; \theta_1)}{L(x_1, ..., x_n; \theta_0)} > \left(1 + \frac{1}{n}\ln\frac{a}{b}\right)^n \sim \frac{a}{b}.$$