

Спецкурс "Большие уклонения"

Шкляев А.В.

10 ноября 2018 г.

Применения теорем Крамера

Пример 9.1. 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:

1) Отвергать гипотезу H_0 , если выборка попала в множество D

$$D = \left\{ \frac{f_1(x_1) \dots f_1(x_n)}{f_0(x_1) \dots f_0(x_n)} > \gamma_n \right\},$$

2) Принимать ее в противном случае.

γ здесь — некоторое заданное число. Предположим, что размер выборки растет, а $\gamma_n = \gamma^n$, γ — фиксированно.

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

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

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

$$R(h) = \mathbf{E}_0 e^{h \ln \frac{f_1(X_1)}{f_0(X_1)}} = \int_{\mathbb{R}} e^{h \ln \frac{f_1(x)}{f_0(x)}} f_0(x) dx = \int_{\mathbb{R}} f_1^h(x) f_0^{1-h}(x) dx.$$

В частности, $R(1) = 1$.

Следовательно, $\mathbf{E}_0 Y_1 = \mu < \infty$ (возможно $-\infty$). При $\gamma < \mu$ α_n сходится к 1. При $\gamma > \mu$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln \alpha_n = - \inf_{\theta \geq \ln \gamma} \Lambda(\theta) = -\Lambda(\gamma).$$

Если же $\mu > -\infty$, а $\mu < \ln \gamma < m^+$, то в силу теоремы Петрова

$$\alpha_n \sim C(\ln \gamma) n^{-1/2} \exp(-\Lambda(\ln \gamma) n).$$

Аналогично для ошибки второго рода

$$\beta_n = \mathbf{P}_1((X_1, \dots, X_n) \notin D) = \mathbf{P}_1 \left(\sum_{i=1}^n \ln \frac{f_0(X_i)}{f_1(X_i)} > -n \ln \gamma \right).$$

Тогда

$$\tilde{R}(h) = \mathbf{E}_1 e^{-h Y_1} = \int_{\mathbb{R}} f_1^{1-h}(x) f_0^h(x) dx = R(1-h),$$

$\tilde{R}(1) = 1$, $\tilde{\mu} = -\mathbf{E}_1 Y_1 < \infty$. Тогда аналогично

$$\beta_n \sim \tilde{C}(-\ln \gamma) n^{-1/2} \exp(-\tilde{\Lambda}(-\ln \gamma) n)$$

при $\ln \gamma > -\tilde{\mu}$. При этом $\tilde{C}(-\ln \gamma) = C(\ln \gamma)$.

$$\tilde{\Lambda}(-\theta) = \inf_h (-\theta h - \ln R(1-h)) = \inf_h (\theta(1-h) - \ln R(1-h)) - \theta = \Lambda(\theta) - \theta.$$

Пример 9.2. Давайте зададимся таким статистическим вопросом для задачи из предыдущего примера. Пусть α_n , β_n — вероятности ошибок I, II родов, a , b — положительные константы, $a + b = 1$. Тогда какой минимальной может быть байесовская ошибка $\Delta_n = a\alpha_n + b\beta_n$? Мы знаем, что наилучшим критерием будет критерий Неймана-Пирсона и наименьшую Δ_n будет давать он.

Рассмотрим критерий Неймана-Пирсона с $\gamma = 1$. Рассмотрим любой другой критерий T_n Неймана-Пирсона с $\gamma_n^{(T)}$. При $\gamma_n^{(T)} \leq 0$ $\alpha_n^{(0)} \leq \alpha_n^{(T)}$, откуда

$$\ln \min_{T_n} \Delta_n \geq \ln(a\alpha_n^{(0)}) = \ln a + \ln \alpha_n^{(0)},$$

при $\gamma_n^{(T)} > 0$

$$\ln \min_{T_n} \Delta_n \geq \ln(b\beta_n^{(0)}) = \ln b + \ln \beta_n^{(0)},$$

Следовательно,

$$\liminf \frac{1}{n} \ln \min_{T_n} \Delta_n \geq -\Lambda(0).$$

При этом на критерии Н-П $\lim \frac{1}{n} \ln \Delta_n = -\Lambda(0)$. Отсюда имеем утверждение — наилучшая возможная байесовская ошибка критерия будет Δ_n с

$$\lim \frac{1}{n} \ln \Delta_n = -\Lambda(0) = \inf_h \ln R(h).$$

Критерий при этом принимает вид

$$\{x_1, \dots, x_n : f_1(x_1) \dots f_1(x_n) > f_0(x_1) \dots f_0(x_n)\},$$

иначе говоря, критерий выбирает ту гипотезу, при которой правдоподобие больше.

Для гипотезы $H_0 : \theta = \theta_0$ с общей альтернативой $H_1 : \theta \neq \theta_0$ естественно использовать критерий отношения правдоподобий с критерием

$$D_c = \left\{ x_1, \dots, x_n : \frac{\sup_{\theta} L(x_1, \dots, x_n; \theta)}{L(x_1, \dots, x_n; \theta_0)} > c \right\}.$$

В частном случае, если выборка принимает конечное множество значений с неизвестными вероятностями p_1, \dots, p_k , правдоподобие имеет вид

$$L(x_1, \dots, x_n; p_1, \dots, p_k) = p_1^{\nu_1} \dots p_k^{\nu_k},$$

где ν_i — количество исходов i . ОМП в этом случае имеет вид $p_i = \nu_i/n$, откуда

$$D_c = \left\{ \vec{x} : \frac{\sup_{\theta} L(x_1, \dots, x_n; \theta)}{L(x_1, \dots, x_n; \theta_0)} > c \right\} = \left\{ \vec{x} : \frac{\left(\frac{\nu_1}{n}\right)^{\nu_1} \dots \left(\frac{\nu_k}{n}\right)^{\nu_k}}{p_{1,0}^{\nu_1} \dots p_{k,0}^{\nu_k}} > c \right\} = \left\{ \vec{x} : \sum_{i=1}^k \frac{\nu_i}{n} \ln \left(\frac{\nu_i}{p_{i,0} n} \right) > \tilde{c} \right\},$$

где $p_{i,0}$ — набор вероятностей при простой гипотезе. Функцию

$$H(\vec{x}; \vec{p}) = \sum_{i=1}^k x_i \ln \left(\frac{x_i}{p_{i,0}} \right)$$

назовем энтропией. На прошлой лекции мы видели, что эта величина является функцией уклонений в теореме Санова.

Назовем критерий оптимальным, если среди критериев таким что

$$\limsup \frac{1}{n} \ln \alpha_n \leq -\alpha, \quad (1)$$

где α_n — вероятности ошибок первого рода, $\limsup(\ln \beta_n)/n$ минимально, где β_n — вероятности ошибки 2 рода

Критерий отношений правдоподобий в силу теоремы Санова имеет ошибку первого рода, удовлетворяющую соотношению

$$\limsup \frac{1}{n} \ln \mathbf{P}_{\vec{p}_0} \left((\nu_1, \dots, \nu_k) : H \left(\frac{\nu_1}{n}, \dots, \frac{\nu_k}{n}; \vec{p}_0 \right) > \tilde{c} \right) = - \inf_{\vec{x} \in A} H(\vec{x}; \vec{p}_0),$$

где

$$A = \{ \vec{x} : H(\vec{x}; \vec{p}_0) > \tilde{c} \}.$$

Следовательно, критерий с $\tilde{c} = \alpha$ будет критерием, удовлетворяющим (1).