Спецкурс "Большие уклонения"

Шкляев А.В.

31 октября 2017 г.

Применения теорем Крамера

Пример 5.1. 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

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

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 $\alpha_n = \mathbf{P}_0((X_1, ..., X_n) \in D)$, 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, $R(1) = 1$.
Следовательно, $\mathbf{E}_0 Y_1 = \mu < \infty$ (возможно $-\infty$). При $\gamma < \mu$ α_n сходится к 1. При $\gamma > \mu$

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

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

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

Тогда

$$\tilde{R}(h) = \mathbf{E}_1 e^{-hY_1} = \int_{\mathbb{R}} f_1^{1-h}(x) f_0^h(x) dx = R(1-h),$$

$\tilde{R}(1) = 1$, $\tilde{\mu} = -\mathbf{E}_1 Y_1 < \infty$. 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 5.2. 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

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

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

Пример 5.3. Рассмотрим следующую задачу. Пусть мы рассматриваем эксперимент с k возможными исходами, имеющими вероятности $p_1,...,p_k$, $\vec{N} = (N_1,...,N_k)$ — число исходов каждого типа за n испытаний, $\vec{\nu} = \vec{N}/n$ — частоты каждого исхода. Тогда мы можем задать вопросом насколько вероятно то, что ν попадет в то или иное множество F.
Для результата о больших уклонениях η нам понадобятся векторы $\vec{X}_i \in \mathbb{R}^k$, принимающие одно из значений $(1,0,...,0)$, $(0,1,...,0)$,..., $(0,0,...,1)$ с вероятностями $p_1,...,p_k$. Тогда $\vec{S}_n = \vec{N}$.
Функция $R(\vec{h}) = \sum_{i=1}^k p_i e^{h_i}$. При этом

$$grad \ln R(\vec{h}) = \left(\frac{p_i e^{h_i}}{R(\vec{h})},\ i \le k\right).$$

Для применения многомерной теоремы Крамера, нам понадобится такое $\vec{h}$, что $grad \ln R(\vec{h}) = \vec{\theta}$, $\sum_{i=1}^k \theta_i = 1$. Тогда

$$\Lambda(\vec{\theta}) = (\vec{\theta}, \vec{h}) - \ln R(\vec{h}) = \sum_{i=1}^k \theta_i h_i - \ln R(\vec{h}).$$

Но $h_i = \ln R(\vec{h}) + \ln(\theta_i/p_i)$, откуда

$$\Lambda(\vec{\theta}) = \sum_{i=1}^k \theta_i h_i - \ln R(\vec{h}) = \sum_{i=1}^k \theta_i \ln \frac{\theta_i}{p_i}.$$

Отсюда в силу теоремы Крамера имеем

$$-\inf_{A^{int}} \Lambda(\vec{\theta}) \le \lim \inf_{n\to\infty} \frac{1}{n} \ln \mathbf{P}(\nu \in A) \le \lim \sup_{n\to\infty} \frac{1}{n} \ln \mathbf{P}(\nu \in A) \le -\inf_{A^{out}} \Lambda(\vec{\theta}),$$

где A — какое-либо множество векторов в $\mathbb{R}^k$ с неотрицательными компонентами, в сумме дающими 1. Этот результат называется теоремой Санова.