11 Применение теории больших уклонений

В прошлый раз мы показали, что ошибки I и II рода удовлетворяют соотношениям при всех $\ln\gamma \in (\mu, \widetilde{\mu})$

$$\alpha_n \sim \frac{C(\ln\gamma)}{h_{\ln\gamma}\sqrt{n}} \exp(-\Lambda(\ln\gamma)n),\ \beta_n \sim \frac{C(\ln\gamma)}{(1-h_{\ln\gamma})\sqrt{n}} \exp(-\Lambda(\ln\gamma)n + n\ln\gamma).$$

При этом $\Lambda(\mu) = 0$, $\Lambda'(0) = h(0) > 0$, поэтому $\Lambda(\theta)$ возрастает при $\theta > \mu$. Аналогично $\Lambda(\theta) - \theta$ убывает по $\theta < \widetilde{\mu}$.

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

$$a\alpha_n + b\beta_n \leq a\alpha_n \leq C\exp(-\Lambda(\ln(1-\delta))n) \leq C\exp(-\Lambda(0)(1-\varepsilon)n)$$

при некоторых $\varepsilon, C > 0$. Аналогично при $\gamma > 1+\delta$

$$a\alpha_n + b\beta_n \leq b\eta_n \leq C\exp(-\Lambda(\ln(1+\delta))n - \ln(1+\delta)n) \leq C\exp(-\Lambda(0)(1-\varepsilon)n)$$

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). Пусть $n(\gamma_n - 1) \to u$, тогда

$$\Lambda(\ln\gamma)n = \Lambda(0) + h_0 u + o(1), \quad \Lambda(\ln\gamma)n - \ln\gamma n = \Lambda(0) + (h_0-1)u + o(1),\ n\to\infty,$$

откуда

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

Максимум будет достигаться при

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

Таким образом, наилучший асимптотический критерий имеет вид

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

11.1 Грубая асимптотика вероятностей больших уклонений

В прошлой задаче мы могли остановиться на более раннем этапе, найдя $\Lambda(\theta)$ и показав, что при любых $\varepsilon > 0$ достаточно больших n

$$\mathbf{P}_0\left(\frac{L(X_1,\ldots,X_n;\theta_1)}{L(X_1,\ldots,X_n;\theta_1)} > \gamma^n\right) \le e^{-(\Lambda(\ln\gamma)-\varepsilon)n},$$
$$\mathbf{P}_1\left(\frac{L(X_1,\ldots,X_n;\theta_1)}{L(X_1,\ldots,X_n;\theta_1)} < \gamma^n\right) \le e^{-(\Lambda(\ln\gamma)-\ln\gamma-\varepsilon)n}.$$

В этом случае мы бы смогли установить, что если γ не зависит от n, то $\gamma = 1$.

В связи с этим для первичного (грубого) анализа достаточно знать грубую асимптотику вероятностей:

$$\frac{1}{n}\ln\mathbf{P}\left(\ln\frac{L(X_1,\ldots,X_n;\theta_1)}{L(X_1,\ldots,X_n;\theta_1)} > n\theta\right) \to -\Lambda(\theta).$$

Остаток семестра мы будем заниматься именно таким, грубым, анализом вероятностей больших уклонений.

11.2 Принцип больших уклонений

Определение 1. В общем случае, назовем функцию $f(x)$ *полунепрерывной снизу (сверху)* в точке, если

$$\liminf_{x\to x_0} f(x) \ge f(x_0),\ (\limsup_{x\to x_0} f(x) \le f(x_0)).$$

Соответственно, полунепрерывность на множестве есть полунепрерывность в каждой точке множества. Это условие равносильно тому, что $f^{-1}(-\infty, a]$ — замкнутое множество при любом a.

Примером полунепрерывной сверху функции является $[x]$, полунепрерывной снизу $\{x\}$.

Определение 2. Назовем *функцией роста* $\Lambda(x)$ неотрицательную полунепрерывную снизу функцию.

Определение 3. Будем говорить, что последовательность мер $\mathbf{P}_n$ удовлетворяет ПБУ (*принципу больших уклонений*) с функцией роста Λ, если

$$-\inf_{x\in A_{int}}\Lambda(x)\le\liminf\frac{\ln\mathbf{P}_n(A)}{n}\le\limsup\frac{\ln\mathbf{P}_n(A)}{n}\le-\inf_{x\in A_{out}}\Lambda(x).$$

Что же это означает? Фактически, мы получаем оценки

$$\exp(-(1+\delta)\inf_{x\in A_{int}}\Lambda(x)n)\le\mathbf{P}_n(A)\le\exp(-(1-\delta)\inf_{x\in A_{out}}\Lambda(x)n)$$

при любом наперед взятом δ и достаточно больших n.

Нам приходится рассматривать разные множества в левой и правой частях, как и прежде это необходимая плата за то, что множество A достаточно общего вида.

Нетрудно заметить, что справедливо следующая эквивалентная формулировка:
1) Для любого замкнутого F

$$\limsup_{n\to\infty}\frac{1}{n}\ln\mathbf{P}_n(F)\le-\inf_{x\in F}\Lambda(x).$$

2) Для любого открытого G

$$\liminf_{n\to\infty}\frac{1}{n}\ln\mathbf{P}_n(G)\ge-\inf_{x\in G}\Lambda(x).$$

Действительно, тогда ПБУ будет следовать из этих утверждения для A_{int} и A_{out}.

11.3 Теорема Крамера

На следующей лекции мы докажем следующую теорему, называемую теоремой Крамера:

Теорема 1 (Крамера)**.** *Пусть S_n — случайное блуждание с $R(h)=\mathbf{E}e^{hX}$ (это матожидание всегда существует, возможно являясь бесконечным). Тогда меры $\mathbf{P}(S_n/n\in\cdot)$ удовлетворяют ПБУ с $\Lambda(x)=\sup_h(hx-\ln R(h))$.*

Отметим, что мы не накладываем никаких условий на конечность $R(h)$ или даже конечность хоть каких-нибудь моментов X. Посмотрим на то, что произойдет с формулировкой в некоторых специфических случаях:

Пример 1. 1) Пусть X_i имеют распределение Коши. Тогда

$$R(0) = 0, \quad R(h) = +\infty, \ h \neq 0.$$

Значит,

$$\Lambda(\theta) = 0 \cdot \theta - \ln R(0) = 0.$$

Таким образом, теорема говорит о том, что

$$e^{-\varepsilon n} \leq \mathbf{P}(S_n/n \in A) \leq e^{\varepsilon n}$$

при таких A, что A_{int} непусто или

$$0 \leq \mathbf{P}(S_n/n \in A) \leq e^{\varepsilon n}$$

при A с пустой внутренностью. Второе утверждение бессмысленно и не накладывает никаких ограничений на A. Первое лишь говорит, что $\mathbf{P}(S_n/n \in A)$ не может стремиться к нулю экспоненциально быстро — это довольно очевидно следует из того, что S_n/n также имеет распределение Коши и в множество с непустой внутренностью попадает с положительной вероятностью.
2) Для распределения Бернулли

$$R(h) = pe^h + (1-p), \quad \Lambda(\theta) = \sup_h(\theta h - \ln R(h)).$$

Производная $\theta h - \ln R(h)$ есть $\theta - pe^h/(pe^h + 1 - p)$. При $\theta < 0$ функция убывает по h, супремум по h достигается при $h \to \infty$ и есть $+\infty$. Аналогично при $\theta > 1$ функция возрастает по h, супремум достигается при $h \to +\infty$ и равен $+\infty$. При $h \in (0,1)$ существует единственный максимум, достигающийся при

$$h = \ln\left(\frac{(1-p)\theta}{p(1-\theta)}\right)$$

и равный

$$\Lambda(\theta) = \theta \ln \frac{\theta}{p} + (1-\theta) \ln \frac{1-\theta}{1-p}.$$

Здесь $\Lambda(p) = 0$, Λ не убывает при $\theta > p$ и не возрастает при $\theta < p$. Таким образом, если множество A не пересекается во внутренним точкам с $(0,1)$, то вероятность попасть в него будет $o(n)$ (в действительности мы видим, что она 0), если содержит среди внутренних точек точку p, то вероятность будет больше $e^{-\varepsilon n}$ при любом ε и достаточно больших n. Вполне естественно, потому что она стремится к 1.

По существу помимо тех случаев, который мы уже рассматривали в теореме Петрова, содержательных результатов мы не получим, а получим только некоторы грубые оценки.