Спецкурс "Большие уклонения"

Шкляев А.В.

24 октября 2017 г.

Принцип больших уклонений

Закончим доказательство теоремы Крамера

Теорема 5.1. (Крамера, в $\mathbb{R}^d$) Пусть $\vec{S}_n$ — случайное блуждание шагами $\vec{X}_i \in \mathbb{R}^d$ с $R(\vec{h}) = \mathbf{E}e^{(\vec{h},\vec{X})} < \infty$, $\vec{h} \in \mathbb{R}^d$. Тогда меры $\mathbf{P}(\vec{S}_n/n \in \cdot)$ удовлетворяют ПБУ с $\Lambda(\vec{x}) = \sup_{\vec{h}}((\vec{h},\vec{x}) - \ln R(\vec{h}))$.

Доказательство. Обсудим еще один вопрос, связанный с доказательством верхнего принципа больших уклонений. А что если $\inf_{\vec{\theta}\in F}\Lambda(\vec{\theta}) = \infty$? В этом случае стоит действовать слегка иначе. Мы должны доказать, что при любом T и достаточно больших n

$$\mathbf{P}\left(\frac{1}{n}\vec{S_n} \in F\right) \le e^{-Tn}$$

При этом, мы знаем, что

$$\liminf_{n\to\infty}\frac{1}{n}\ln\mathbf{P}\left(\frac{1}{n}\vec{S_n} \in F\right) \le -\inf_{\vec{\theta}\in F_M}\Lambda(\vec{\theta})$$

При этом

$$\inf_{\vec{\theta}\in F_M}\Lambda(\vec{\theta}) \to \inf_{\vec{\theta}\in F}\Lambda(\vec{\theta}) = \infty$$

при $M\to\infty$, а значит

$$\mathbf{P}\left(\frac{1}{n}\vec{S_n} \in F\right) \le \sum_{i=1}^d \mathbf{P}(S_{n,i} \ge Mn) + \sum_{i=1}^d \mathbf{P}(S_{n,i} \le -Mn) + e^{-2Tn}$$

при достаточно больших M. В силу неравенства Маркова

$$\mathbf{P}(S_{n,i} \ge Mn) \le \frac{R_i(1)^n}{e^{Mn}}, \mathbf{P}(S_{n,i} \le -Mn) \le \frac{R_i(-1)^n}{e^{Mn}}.$$

Выбирая M достаточно большим, чтобы $R_i(1)e^{-M} < e^{-2T}$, $R_i(-1)e^{-M} < e^{-2T}$, имеем

$$\mathbf{P}\left(\frac{1}{n}\vec{S_n} \in F\right) \le (2d+1)e^{-2Tn} \le e^{-Tn}.$$

2) Докажем нижний принцип больших уклонений: для любого открытого G

$$\liminf_{n\to\infty}\frac{1}{n}\ln\mathbf{P}\left(\frac{\vec{S}_n}{n} \in G\right) \ge -\inf_{\vec{\theta}\in G}\Lambda(\vec{\theta}).$$

Аналогично одномерному случаю достаточно доказать, что при любом $\vec{x} \in D_\Lambda$

$$\mathbf{P}(\vec{S_n}/n \in B_{\vec{x},\delta}) \ge -\Lambda(\vec{x}).$$

Предположим, что $\vec{x} = grad \ln R(\vec{h})$ при некотором $\vec{h} \in \mathbb{R}^d$. Рассмотрим меру

$$\mathbf{P}(\vec{S_n}^{(h)} \in A) = R(\vec{h})^{-n}\int_A e^{(\vec{h},\vec{y})}\mathbf{P}(\vec{S_n} \in d\vec{y}).$$

Тогда

$$\frac{1}{n}\ln\mathbf{P}(\vec{S_n}/n \in B_{\vec{x},\delta}) = \frac{1}{n}\ln\left(R(\vec{h})^n\int_{B_{\vec{x},\delta}} e^{-n(\vec{h},\vec{y})}\mathbf{P}\left(\vec{S_n}^{(\vec{h})}/n \in d\vec{y}\right)\right) \ge \ln R(\vec{h}) - (\vec{h},\vec{x}) - \frac{1}{n}\ln\left(e^{-n\delta|\vec{h}|}\mathbf{P}\left(\vec{S_n}^{(\vec{h})}/n \in d\vec{y}\right)\right).$$

В силу ЗБЧ последнее слагаемое сходится к $-\delta|\vec{h}|$ при $n\to\infty$, т.к. $\mathbf{E}\vec{X}^{(\vec{h})} = grad \ln R(\vec{h}) = \vec{x}$, откуда имеем требуемую оценку.

Итак, мы доказываем нижнюю оценку, причем уже знаем, что будет, если есть такое $\vec{h}$, что $\Lambda(\vec{x}) = (\vec{x},\vec{h}) - \ln R(\vec{h})$. Заметим, что если в $\Lambda(\vec{\theta}) = \sup_y((\vec{y},\vec{\theta}) - \ln R(\vec{y}))$ супремум достигается в конечной точке $\vec{y}$, то в силу необходимого условия экстремума такое $\vec{h}$ найдется и оно в точности равно $\vec{y}$. Поэтому сложности могут быть в том случае, когда супремум $\Lambda(\vec{x})$ не достижим. Рассмотрим $\vec{Z}_i = \vec{X}_i + \vec{Y}_i$, $\tilde{S}_n = \vec{Z}_1 + ... + \vec{Z}_n$, $\vec{Y}_i \sim \mathcal{N}(\vec{0}, E/M)$ н.о.р., $M > 0$.
Тогда

$$\ln R_Z(\vec{h}) = \ln R(\vec{h}) + \frac{1}{2M}|\vec{h}|^2 \ge \ln R(\vec{h}),\ \Lambda_Z(x) \le \Lambda(x).$$

При этом $\ln R(\vec{h}) \ge (\vec{h},\vec{\mu})$, откуда

$$(\vec{h},\vec{x}) - \ln R_Z(\vec{h}) \le (\vec{h},(\vec{x}-\vec{\mu})) - \frac{1}{2M}|\vec{h}|^2 \to -\infty,$$

$h \to \infty$. Следовательно, супремум в $\Lambda_M(\vec{\theta})$ достижим в конкретной точке $\vec{h}$, удовлетворяющей условию $\vec{x} = grad \ln R_Z(\vec{h})$. Пользуясь доказанным соотношением

$$\liminf_{n\to\infty} \frac{1}{n} \ln P(\tilde{S}_n/n \in B_{\delta/2}(\vec{x}) \geq -\Lambda_Z(\vec{x}) \geq -\Lambda(x).$$

При этом

$$P(S_n/n \in B_\delta(\vec{x})) \geq P(\tilde{S}_n/n \in B_{\delta/2}(\vec{x})) - P(|\vec{S}_n^{(Y)}/n| > \delta/2).$$

Тогда

$$-P(S_n/n \in B_\delta(\vec{x}) \leq -P(\tilde{S}_n \in B_{\delta/2}(\vec{x}) + P(|\vec{S}_n^{(Y)}/n| > \delta/2).$$

Модуль вектора больше $\delta/2$ только если одна из координат больше $\delta/(2\sqrt{d})$. Значит

$$P(|\vec{S}_n^{(Y)}/n| > \delta/2) \leq d2 \left(1 - \Phi\left(\frac{\delta\sqrt{nM}}{2d}\right)\right).$$

В силу соотношения

$$\Phi(x) \sim \frac{1}{\sqrt{2\pi x}} e^{-x^2/2},$$

имеем $\limsup_{n\to\infty} \frac{1}{n} \ln P((Y_1 + ... + Y_n)/n > \delta/2)) \leq -M\delta^2/(2d)$. Отсюда при достаточно больших M посл-ь $P((Y_1 + ... + Y_n)/n > \delta/2)$ ЭО со сколь угодно малым t. Поскольку $-P(\tilde{S}_n/n \in B_{\delta/2}(\vec{x})$ ЭО с $-\Lambda(x)$, имеем $-P(S_n/n \in B_\delta(\vec{x})$ ЭО с $\Lambda(x)$. Что и требовалось доказать. □

Рассмотрим несколько примеров применения теорем Крамера.

Пример 5.1. Рассмотрим задачу оценивания среднего $\vec{\mu}$ по выборке $\vec{X}_i$, $i = 1...n$, $E\vec{X}_i = \vec{\mu}$. Допустим при большой выборке $\vec{X}_i$ мы заметили, что $\overline{X}$ и $\vec{\mu}$ отличаются более чем на ε, например, в $\mathbb{R}^d$. Насколько уверенно мы можем сказать, что $\vec{X}$ имеет другое среднее?

Обычно мы используем для этого ЦПТ

$$P\left(\left|\left|\frac{\overline{X} - \vec{\mu}}{\sqrt{n}}\right|\right| \geq x\right) \to P(||\vec{Y}|| \geq x),$$

где $\vec{Y} \sim \mathcal{N}(0, \Sigma^2)$, Σ^2 — матрица ковариации $\vec{X}$.

Но, вообще говоря, нельзя использовать это как оценку для $P(||\overline{X} - \vec{\mu}|| > \varepsilon)$, поскольку ЦПТ верна при $\varepsilon = O(n^{-1/2})$ но не при $O(1)$. Таким образом, видя большое отклонение $\overline{X}$ и μ, мы не можем оценить вероятность этого исхода из ЦПТ. Зато можем с помощью теоремы Крамера:

$$\frac{1}{n} \ln P(||\overline{X} - \vec{\mu}|| \geq \varepsilon) \to - \inf_{\vec{\theta} \notin \mathbb{R}^d \setminus B_\varepsilon(\vec{\mu})} \Lambda(\vec{\theta}).$$

Здесь мы воспользовались тем, что $\Lambda(\vec{\theta})$ на внешности открытого и замкнутого шаров имеет один и тот же инфимум. Теорема Крамера, таким образом, позволит нам оценить фактический уровень значимости нашей гипотезы более правильным образом.