Спецкурс "Большие уклонения"

Шкляев А.В.

2 ноября 2018 г.

Многомерная теорема Крамера

Теорема 6.1. (Крамера, в $\mathbb{R}^d$) Пусть $\vec{S}_n$ — случайное блуждание шагами $X_i \in \mathbb{R}^d$ 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. 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. Мы будем рассматривать случай $\Lambda(\vec{x}) < \infty$, оставшийся случай рассматривается аналогично тому, как это было сделано на прошлой лекции.

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

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

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

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

$h \to \infty$. 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

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

При этом

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

Тогда

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

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

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

В силу соотношения

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

имеем

$$\mathbf{P}((Y_1 + ... + Y_n)/n > \delta/2)) \leq \exp(-\widetilde{M}n).$$

при любом $\widetilde{M}$ и достаточно больших M, n. Следовательно,

$$\mathbf{P}(S_n/n \in U_\delta(\vec{x})) \geq \mathbf{P}(\tilde{S}_n/n \in U_{\delta/2}(\vec{x})) - \exp(-\widetilde{M}n) \geq \frac{1}{2}\exp\left(-(\Lambda(x) - \varepsilon)n\right)$$

при достаточно большом n. Логарифмируя, деля на n и переходя к пределу, имеем требуемое. □

0.0.1 Применения теоремы Крамера

Пример 6.1. 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

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

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.

Пример 6.2. Применим теорему Санова к задаче проверки гипотезы согласия $H_0 : p_1 = p_1^0, ..., p_k = p_k^0$, где p_i^0 заданы, с помощью критерия хи-квадрат. Критерий хи-квадрат основан на том факте, что статистика

$$T(X_1, ..., X_n) = \sum_{i=1}^{k} \frac{(N_i - np_i^0)^2}{np_i^0}$$

имеет асимптотическое распределение χ^2_{k-1} при выполнении гипотезы. Тем самым вероятность $\mathbf{P}_0(T(X_1, ..., X_n) > c)$ аппроксимируется вероятностью $1 - F_{\chi^2_{k-1}}()$. Эта аппроксимация довольно точна при c порядка константы, но, скажем, при c порядка n она является неудовлетворительной. В этом случае фактический уровень значимости критерия можно оценить с помощью теоремы Санова

$$\frac{1}{n} \ln \mathbf{P}(T(X_1, ..., X_n) > nc) = \frac{1}{n} \ln \mathbf{P}\left(\sum_{i=1}^{k} \frac{(\nu_i - p_i^0)^2}{p_i^0} > c\right) \to -\inf_{\vec{\theta}\in A} \sum_{i=1}^{k} \theta_i \ln \frac{\theta_i}{p_i^0},$$

где $A = \{\theta_i : \sum_{i=1}^{k} \theta_i = 1,\ \sum_{i=1}^{k} \frac{(\theta_i - p_i^0)^2}{p_i^0} > c\}$.
При этом A — внешность эллипсоида

$$\sum_{i=1}^{k} \frac{(\theta_i - p_i^0)^2}{p_i^0} > c,$$

пересеченная с плоскостью $\sum_{i=1}^{k} \theta_i = 1$. Минимум L_{min} выпуклой функции

$$\sum_{i=1}^{k} \theta_i \ln \frac{\theta_i}{p_i^0}$$

на A достигается на границе, то есть

$$\sum_{i=1}^{k} \frac{(\theta_i - p_i^0)^2}{p_i^0} = c.$$

Полученная задача достаточно сложна для аналитической работы, но может быть решена численно.

Решая задачу минимизации, мы находим минимальное значение L_{min} и получаем ответ: $\mathbf{P}(T(X_1, ..., X_n) > c) \approx \exp(-L_{min} n)$. Этот ответ значительно более точен, чем $1 - F_{\chi^2_{k-1}}(cn)$.