Спецкурс "Большие уклонения"

Шкляев А.В.

29 сентября 2016 г.

Теорема Петрова и ее доказательство

Напомним, что мы рассматриваем величины, удовлетворяющие правостороннему условию Крамеру:

$$R(h) = Ee^{hX} = \int_{\mathbb{R}} e^{hx} dF(x) = \int_{-\infty}^{0} e^{hx} dF(x) + h \int_0^{\infty} (1 - F(x)) e^{hx} dx - (1 - F(x)) e^{hx}|_0^{\infty}.$$

Соответственно, конечность этой величины равносильна тому, будет ли $(1 - F(x))e^{hx}$ интегрируема при всех $h < h^+$. Это равносильно тому, что $P(X > x)$ имеет порядок $e^{-h^+ x}$ или меньше, что и означает пресловутую экспоненциальность хвостов. При этом хвост на $-\infty$ может быть любым, лишь бы только у величин существовало математическое ожидание. На основе $R(h)$ мы определили функции $m(h)$, $\sigma(h)$, h_θ. Сопряженные случайные величины $X^{(h)}$ мы ввели как заданную

$$P_X^{(h)}(A) = \frac{1}{R(h)} \int_A e^{hx} P_X(dx).$$

Это, очевидно, вероятностная мера при любом $h \in (0, h^+)$. Рассмотрим величины $X^{(h)}$ с таким распределением. Тогда

$$EX^{(h)} = m(h),\ DX^{(h)} = \sigma^2(h).$$

Таким образом, $m(h)$ и $\sigma^2(h)$ это матожидание и дисперсия $X^{(\theta)}$, а h_θ — так подобранное h, что $X^{(h)}$ имеет м.о. θ.

Задача 2.1. Попробуйте самостоятельно найти распределение $X^{(h)}$ для $X_i \sim Bernoulli(p)$, $X_i \sim \exp(\lambda)$, $X_i \sim Geom(p)$, $X_i \sim Poiss(\lambda)$, $X_i \sim \mathcal{N}(a, 1)$, $X_i \sim R[0, 1]$.
Задача 2.2. Найти $R(h), m(h), \sigma^2(h), \Lambda(\theta)$ для $X_i \sim \mathcal{N}(a, 1)$, $X_i \sim Bernoulli(p)$, $X_i \sim \exp(\lambda)$, $X_i \sim Geom(p)$, $X_i \sim Poiss(\lambda)$.

Итак, нам осталось доказать нижнюю оценку теоремы Крамера (вернее той ее версии, которую мы назвали детской). Для этого нам нужно доказать, что

$$\liminf_{n\to\infty} \frac{\ln P(S_n \geq \theta n)}{n} \geq -\Lambda(\theta).$$

Будем использовать наше сопряженное распределение:

$$P(S_n \geq \theta n) = \int_{x_1+\ldots+x_n \geq \theta n} P_{X_1}(dx_1)\ldots P_{X_n}(dx_n) = R(h)^n \int_{x_1+\ldots+x_n \geq \theta n} e^{-h(x_1+\ldots+x_n)} P_{X_1^{(h)}}(dx_1)\ldots P_{X_n^{(h)}}(dx_n) =$$
$$R(h)^n E\left(e^{-hS_n^{(h)}}; S_n^{(h)} \geq \theta n\right) = e^{-(\theta h - \ln R(h))n} E\left(e^{-h(S_n^{(h)} - \theta n)}; S_n^{(h)} \geq \theta n\right). \tag{1}$$

Выбирая $h = h_\theta$ мы получим формулу

$$P(S_n \geq \theta n) = e^{-\Lambda(\theta)n} E\left(e^{-h_\theta\left(S_n^{(h_\theta)} - \theta n\right)}; S_n^{(h_\theta)} \geq \theta n\right).$$

Заметим, что $EX_1^{(h_\theta)} = \theta$, поэтому событие $\{S_n^{(h_\theta)} \geq \theta n\}$ перестало быть особенно редким. Оценим полученное матожидание снизу следующим образом:

$$E\left(e^{-h_\theta(S_n^{(h_\theta)} - \theta n)}; S_n^{(h_\theta)} \geq \theta n\right) \leq E\left(e^{-h_\theta(S_n^{(h_\theta)} - \theta n)}; 0 \leq S_n^{(h_\theta)} - \theta n \leq \sqrt{n}\sigma(h_\theta)\right) \geq$$
$$e^{-\sigma(h_\theta)h_\theta\sqrt{n}} P(0 \leq S_n^{(h_\theta)} - \theta n \leq \sigma(h_\theta)\sqrt{n})$$

При этом $P(0 \leq S_n^{(h_\theta)} - \theta n \leq \sigma(h_\theta)\sqrt{n}) \to \Phi(1) - \Phi(0) > 0$ в силу ЦПТ.
Применяя полученные оценки, имеем

$$\liminf_{n\to\infty} \frac{\ln P(S_n \geq \theta n)}{n} \geq -\Lambda(\theta) - \liminf_{n\to\infty} \frac{\sigma(h_\theta)}{\sqrt{n}} - \liminf_{n\to\infty} \frac{\Phi(1) - \Phi(0)}{n} = -\Lambda(\theta).$$

Теорема доказана.

При этом мы показали, что

$$e^{-\Lambda(\theta)n} e^{-\varepsilon n} \leq P(S_n \geq \theta n) \leq e^{-\Lambda(\theta)n}$$

Но погрешность нашей оценки может довольно большой, например, наша нижняя оценка могла дать погрешность $e^{-h_\theta\sigma(h_\theta)\sqrt{n}}$. Это довольно много. Хотелось бы найти асимптотическое выражение для $P(S_n \geq \theta n)$, а не для его ln.

Для этого нам понадобится понятие решетчатого распределения. Назовем распределение решетчатым, если $\exists a, d :$ $P(X \in \{a + kd, k \in \mathbb{Z}\}) = 1$. Иначе говоря, будет некоторая решетка, на которую наши шаги попадают с единичной вероятностью. Из всех таких d будем брать максимальное.

Теорема 2.1. (Петров-Бахадур-Рао) Пусть для н.о.р. X_i с $EX_1 = \mu$ выполнено правостороннее условие Крамера. Тогда при всех $\theta \in (\mu, m^+)$ выполнено соотношение

$$P(S_n \geq \theta n) = (1 + o(1))C(\theta)n^{-1/2} \exp(-\Lambda(\theta)n),\ n \to \infty, \tag{2}$$

где $o(1)$ в (2) равномерно мало по $\theta \in [\theta_1, \theta_2] \subset (0, m^+)$.
Константа $C(\theta)$ имеет вид:
а) X_i имеют нерешетчатое распределение и $C(\theta) = (\sqrt{2\pi}\sigma(h_\theta)h_\theta)^{-1}$,
б) X_i имеют решетчатое распределение, $C(\theta) = d(\sqrt{2\pi}\sigma(h_\theta)(1 - e^{-h_\theta d}))^{-1}$ и $d|(\theta - a)n$.
Замечание 2.1. Из равномерности удобно получать некоторые следствия. Выведем асимптотику $P(S_n \geq \theta n + t)$, где t — фиксированная константа.
Воспользуемся теоремой Петрова

$$P(S_n \geq \theta n + t) = P(S_n \geq \tilde{\theta} n),$$

где $\tilde{\theta} = \theta + t/n$. При этом равномерность асимптотики означает, что

$$\lim_{n\to\infty} \sup_{\theta \in [\theta_1, \theta_2]} \left| \frac{P(S_n \geq \theta n)}{\frac{C(\theta)}{\sqrt{n}} \exp(-\Lambda(\theta) n)} \right| = 0.$$

При этом $\tilde{\theta}$ при каждом $\theta \in [\theta_1, \theta_2]$ лежит в некотором $[\tilde{\theta}_1, \tilde{\theta}_2] \subset (\mu, m^+)$. Из равномерности

$$\lim_{n\to\infty} \frac{P(S_n \geq \tilde{\theta} n)}{\frac{C(\tilde{\theta})}{\sqrt{n}} \exp(-\Lambda(\tilde{\theta}) n)}$$

Остается понять, как устроены $C(\tilde{\theta})$, $\Lambda(\tilde{\theta})$. В силу непрерывности $C(\cdot)$ $C(\tilde{\theta}) \to C(\theta)$, $n \to \infty$.

$$\Lambda(\tilde{\theta}) = \Lambda(\theta) + (\theta - \tilde{\theta})\Lambda'(\theta) + o(\theta - \tilde{\theta}).$$

$\theta - \tilde{\theta} = O(n^{-1})$, откуда

$$P(S_n \geq \theta n + t) \sim C(\theta) n^{-1/2} \exp(-\Lambda(\theta) n) \exp(-\Lambda'(\theta) t).$$

Остается найти $\Lambda'(\theta)$:

$$(\theta h_\theta - \ln R(h_\theta))' = h_\theta + \theta(h_\theta)' - m(h_\theta)(h_\theta)' = h_\theta.$$

Следовательно,

$$P(S_n \geq \theta n + t) \sim C(\theta) n^{-1/2} \exp(-\Lambda(\theta) n - h_\theta t),$$

причем асимптотика равномерна по $|t| \leq a_n$ для любой $a_n = o(n^{1/2})$.
Иначе говоря, $P(\Delta_n \geq t | S_n \geq \theta n) \to \exp(-h_\theta t)$, $t > 0$, $n \to \infty$, где $\Delta_n = S_n - \theta n$. Таким образом, условное предельное распределение Δ_n при условии уклонения экспоненциальное с параметром h_θ.
Замечание 2.2. Аналогичным образом можно доказать, что для любых ограниченных множеств A_i и любого $t > 0$

$$P(S_n \geq \theta n + t, X_1 \in A_1, ..., X_k \in A_k | S_n \geq \theta n) \to e^{-h_\theta t} P(X_1^{(h_\theta)} \in A_1)...P(X_k^{(h_\theta)} \in A_k)$$

Этот результат называется теоремой Бартфаи. Для доказательства представим искомую вероятность в виде

$$(P(S_n \geq \theta n))^{-1} \int_{A_1} ... \int_{A_k} P(S_{n-k} \geq \theta n - x_1 - ... - x_k + t) P(X_1 \in dx_1)...P(X_k \in dx_k).$$

В силу Теоремы Петрова

$$P(S_{n-k} \geq \theta n - x_1 - ... - x_k + t) = P(S_{n-k} \geq \theta(n-k) + \theta k - x_1 - ... - x_k + t) \sim C(\theta) n^{-1/2} \exp(-\Lambda(\theta)(n-k)) e^{-h_\theta(\theta k - x_1 - ... - x_k + t)}.$$

Тогда искомая вероятность эквивалентна (здесь мы выносим $o(1)$ из под интеграла в силу равномерности)

$$\exp(\Lambda(\theta) k - \theta h_\theta k) e^{-h_\theta t} \int_{A_1} ... \int_{A_k} e^{h_\theta(x_1 + ... + x_k)} P(X_1 \in dx_1)...P(X_k \in dx_k) = e^{-h_\theta t} P(X_1^{(h_\theta)} \in A_1)...P(X_k^{(h_\theta)} \in A_k),$$

где мы воспользовались тем, что $\Lambda(\theta) - \theta h_\theta = -\ln R(h_\theta)$.

Доказательство Теоремы 2.1. Мы докажем теорему Петрова для нерешетчатых величин. Идея доказательства будет той же, что и Крамера, но потребует более аккуратной оценки вероятностей $P(S_n^{(h_\theta)} \geq \theta n + x)$, где $S^{(h)}$ — случайное блуждание, построенное по н.о.р. $X_i^{(h)}$. Мы будем использовать (1). Нам понадобится два вспомогательных утверждения:
Теорема (асимптотические разложения). Пусть X_i — н.о.р., нерешетчаты, $EX_1 = \mu$, $DX_1 = \sigma^2$, $E(X_1 - \mu)^3 = \mu_3$. Тогда

$$F_{\frac{S_n - \mu n}{\sqrt{n}\sigma}}(x) = \Phi(x) + \frac{Q(x) + o(1)}{\sqrt{n}}, \ n \to \infty,$$

причем $o(1)$ равномерно мало по $x \in [-t\sqrt{n}, t\sqrt{n}]$ при любом t, где $Q(x) = (1 - x^2)e^{-x^2/2}\mu_3/(6\sqrt{2\pi}\sigma^3)$.

Этот результат обобщает ЦПТ, добавляя следующий член погрешности приближения после $\Phi(x)$. Он будет доказан в рамках курса "Дополнительные главы теории вероятностей".

Лемма. Пусть F_n, G_n — функции ограниченной вариации на $[a,b]$, $f_n(x)$ — последовательность дифференцируемых функций на $[a,b]$ и соотношение $F_n(x) = G_n(x) + o(a_n)$ выполнено равномерно по $x \in [a,b]$ при $n \to \infty$, a_n — некоторая заданная последовательность. Тогда

$$\int_{[a,b]} f(x)dF(x) = \int_{[a,b]} f(x)dG(x) + o(a_n)(|f_n(a)| + |f_n(b)| + \int_a^b |f_n'(x)|dx).$$

Доказательство.
Заметим, что в силу формулы интегрирования по частям

$$\int_a^b f_n(x)dF(x) = f_n(x)F(x)|_a^b - \int_a^b F(x)f_n'(x)dx, \ \int_a^b f_n(x)dG(x) = f_n(x)G(x)|_a^b - \int_a^b G(x)f_n'(x)dx.$$

В силу соотношений

$$f_n(a)(F(a) - G(a)) = o(a_n)f_n(a), \ f_n(b)(F(b) - G(b)) = o(a_n)f_n(b), \ |\int_a^b G(x) - F(x)f_n'(x)dx| \leq a_n \int_a^b |f_n'(x)|dx$$

имеем требуемое.

На следующей лекции мы выведем из полученных результатов теорему Петрова.