Спецкурс "Большие уклонения"

Шкляев А.В.

26 сентября 2017 г.

Задача о больших уклонениях

1 Введение

1.1 Мотивация

Классические предельные теоремы, которые мы изучаем, позволяют нам описывать "типичное" поведение случайных последовательностей и процессов. Так, например, классическая центральная предельная теорема утверждает, что

$$\mathbf{P}\left(\frac{S_n - na}{\sqrt{n}\sigma} \in [a,b]\right) \to \mathbf{P}(\xi \in [a,b]),$$

где $\xi \sim \mathcal{N}(0,1)$. Эта теорема описывает поведение только в "типичной области" $na + O(\sqrt{n})$. Мы не можем сказать насколько вероятно, что S_n окажется, скажем, правее $na + n^{2/3}$ — лишь можем сказать, что эта вероятность мала.

С другой стороны, может возникнуть вопрос — а зачем нам оценивать "малость" вероятностей?

Пример 1.1. Рассмотрим факт появления жизни на Земле. Само по себе появление жизни на случайно взятой планете очень маловероятно. С другой стороны, количество планет во вселенной огромное, скажем, 10^{20}. Тогда если вероятность появления жизни окажется 10^{-10}, то число обитаемых планет, по видимости, огромно. Если вероятность появления жизни, скажем, 10^{-20}, то вполне возможно, что мы единственная обитаемая планета. Если же она 10^{-30}, то шанс зарождения жизни крайне мал и это либо чудо, либо мы неправильно произвели наши оценки и неправильно восстановили процесс появления жизни.

Пример 1.2. Представим себе появление, что дамба выстроена так, что вероятность того, что река ее превысит равна 0.001. Тогда за 100 лет (если мы считаем независимыми уровни в разные дни) вероятность наводнения уже совсем не маленькая — порядка 10%.

Следовательно, для нас жизненно важно не только представлять, что то или иное событие редко, но и уметь оценивать его вероятность.

1.2 Неприменимость центральной предельной теоремы к редким событиям

Пример 1.3. Рассмотрим X_i с пуассоновским распределением с параметром λ, $S_n = X_1 + ... + X_n$ — пуассоновская с параметром λn. Тогда при $m = [\theta n] + 1$

$$\mathbf{P}(S_n \geq \theta n) = \sum_{k=m}^{n} \mathbf{P}(S_n = k) = \sum_{k=m+1}^{n} \frac{(\lambda n)^k}{k!} e^{-\lambda n} = (\lambda n) = \frac{(\lambda n)^m}{m!} e^{-\lambda n} \sum_{k=0}^{\infty} \frac{(\lambda n)^k}{(m+k)(m+k-1)...(m+1)}$$

Ряд в правой части не превосходит геометрической прогрессии с шагом λ/θ, а значит суммируется к некоторой величине порядка константы. Тем самым порядок нашей суммы есть

$$\frac{(\lambda n)^m}{m!} e^{-\lambda n} \sim \frac{(\lambda n)^m}{\sqrt{2\pi m} m^m} e^{m - \lambda n} = \frac{1}{\sqrt{2\pi\theta n}} e^{-m(\ln \lambda n) - \ln m) + m - \lambda n}$$

В то же время, если бы здесь была верна ЦПТ, то

$$\mathbf{P}(S_n \geq \theta n) = 1 - \Phi(\frac{(\theta-\lambda)n}{\sqrt{\lambda}\sqrt{n}}).$$

При этом

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

Для доказательства воспользуемся правилом Лопиталя:

$$\frac{1-\Phi(x)}{\frac{1}{\sqrt{2\pi}x}e^{-x^2/2}} = \frac{\frac{1}{\sqrt{2\pi}}e^{-x^2/2}}{\frac{1}{\sqrt{2\pi}}\left(e^{-x^2/2} - e^{-x^2/2}x^{-2}\right)} \to 1, \ x \to \infty.$$

Значит

$$1 - \Phi\left(\frac{(\theta-\lambda)n}{\sqrt{\lambda}\sqrt{n}}\right) \sim \frac{\sqrt{\lambda}}{\sqrt{2\pi}(\theta-\lambda)\sqrt{n}} \exp\left(-\frac{(\theta-\lambda)^2 n}{2\lambda}\right),$$

то есть величина в экспоненте не та, что требуется.

1.3 Примеры

Давайте попробуем рассмотреть нашу задачу на случайном блуждании. Пусть X_i — н.о.р. с $EX_i = \mu$, $S_n = X_1 + ... + X_n$. Рассмотрим $\theta > \mu$ и поглядим на вероятность $P(S_n \geq \theta n)$. Давайте посмотрим несколько примеров:

Пример 1.4. Пусть $X_i \sim \mathcal{N}(a, \sigma^2)$, $\theta > a$. Тогда $S_n \sim \mathcal{N}(na, n\sigma^2)$, а значит

$$P(S_n \geq \theta n) = P\left(\frac{S_n - na}{\sigma\sqrt{n}} \geq \sqrt{n}\frac{\theta - a}{\sigma}\right) = 1 - \Phi\left(\sqrt{n}\frac{\theta-a}{\sigma}\right),$$

где Φ — ф.р. $\mathcal{N}(0,1)$. Следовательно,

$$P(S_n \geq \theta n) \sim \frac{\sigma}{\sqrt{2\pi n}(\theta - a)} \exp\left(-\frac{(\theta-a)^2}{2\sigma^2} n\right).$$

Пример 1.5. Пусть X_i принимают значения 0 и 1 с вероятностями $1-p$, p, $\theta > p$. Тогда

$$P(S_n \geq \theta n) = \sum_{k=(1+\theta)n/2}^{n} C_n^k p^k (1-p)^{n-k}.$$

С помощью формулы Стирлинга можно оценить слагаемые этой суммы, откуда можно вывести, что

$$P(S_n \geq \theta n) \sim \frac{1}{\sqrt{2\pi(1-\theta)\theta}} \exp\left(-\left(\theta \ln\left(\frac{\theta}{p}\right) + (1-\theta)\ln\left(\frac{1-\theta}{1-p}\right)\right) n\right).$$

На лекции была приведена неверная формула для мультипликативной константы перед экспонентой. Как мы видим, в обоих случаях вероятность имеет вид $C(\theta)n^{-1/2}\exp(-\Lambda(\theta)n)$ при каких-то функциях $C(\theta)$, $\Lambda(\theta)$. Отсюда возникает гипотеза, что такая ситуация будет и в других случаях. Однако, так будет заведомо не всегда:

Пример 1.6. Пусть $P(X > x) \sim x^{-a}$. Тогда

$$P(S_n \geq \theta n) \geq P(X_1 > (\theta - \mu + \varepsilon)n) P(X_2 + ... + X_n > (\mu - \varepsilon)n) \sim (\theta - \mu + \varepsilon)^{-a} n^{-a}.$$

Таким образом, если X имеет степенные хвосты, то и S_n будет иметь не более чем степенные хвосты.

2 Теорема Петрова

2.1 Обозначения

Тем не менее для величин, имеющих экспоненциальные или более легкие хвосты, такой результат сформулировать можно. Введем некоторые вспомогательные обозначения: Пусть $R(h) = Ee^{hX} < \infty$ при $h^- < h < h^+$, $h^- \leq 0$, $h^+ \geq 0$. Это условие будем называть условием Крамера. Давайте положим $m(h) = (\ln R(h))' = EXe^{hX}/R(h)$, $\sigma^2(h) = m'(h)$ (чуть позже мы покажем, что $\sigma^2(h)$ положительна). Отсюда $m(h)$ монотонно возрастает на $(0, h^+)$, $m(0) = \mu = EX$, $m^- := \lim_{h\to h^- +0} m(h)$, $m^+ := \lim_{h\to h^+ -0} m(h)$. Таким образом, при всех $\theta \in (m^-, m^+)$ найдется единственное h_θ, т.ч. $m(h_\theta) = \theta$. И наконец положим $\Lambda(\theta) := h_\theta\theta - \ln R(h_\theta)$.

2.2 Теорема Петрова

Оказывается, что справедлива следующая теорема Петрова: **Теорема 1.1.** 1) Пусть величины X_i нерешетчаты (то есть $\nexists a, d:\ \mathbf{P}(X \in \{a + kd,\ k \in \mathbb{Z}\})$). Тогда

$$\mathbf{P}(S_n \geq \theta n) \sim \frac{1}{\sqrt{2\pi}\sigma(h_\theta)h_\theta} \exp(-\Lambda(\theta)n),\ n \to \infty,$$

равномерно по любому отрезку $[\theta_1, \theta_2] \subset (\mu, m^+)$. Аналогично

$$\mathbf{P}(S_n \leq \theta n) \sim \frac{1}{\sqrt{2\pi}\sigma(h_\theta)h_\theta} \exp(-\Lambda(\theta)n),\ n \to \infty,$$

равномерно по любому отрезку $[\theta_1, \theta_2] \subset (m^-, \mu)$.
2) Пусть величины X_i решетчаты с шагом d (d наибольший такой шаг). Тогда

$$\mathbf{P}(S_n \geq \theta n) \sim \frac{d}{\sqrt{2\pi}\sigma(h_\theta)\,(1 - e^{dh_\theta})} \exp(-\Lambda(\theta)n),\ n \to \infty,$$

при $\theta n \in \{an + kd,\ k \in \mathbb{Z}\}$. Аналогично

$$\mathbf{P}(S_n \leq \theta n) \sim \frac{d}{\sqrt{2\pi}\sigma(h_\theta)\,(1 - e^{dh_\theta})} \exp(-\Lambda(\theta)n),\ n \to \infty,$$

равномерно по любому отрезку $[\theta_1, \theta_2] \subset (m^-, \mu)$.
Задача 1.1. Сформулируйте теорему Петрова для пуассоновских, геометрических, экспоненциальных распределений.

2.3 Крамеровская замена мер

Опишем общую технологию, по которой доказывается теорема Петрова. Рассмотрим случай решетчатой величины, для простоты целочисленную. Тогда будем пользоваться локальной предельной теоремой Гнеденко:
Теорема 1.2. Для решетчатых величин X_i с конечной $\mathbf{E}X_i^2$ справедлива теорема

$$\mathbf{P}(S_n = k) \sim \frac{1}{\sqrt{2\pi n}\sigma} e^{-\frac{(k-\mu n)^2}{2\sigma^2 n}},\ n \to \infty,$$

причем равномерно по $|k - \mu n| \leq t\sqrt{n}$ при любом t.

Тогда при любых k, h

$$\mathbf{P}(S_n = k) = \sum_{i_1,\ldots,i_n \in \mathbb{Z}: i_1+\ldots+i_n=k} \mathbf{P}(X_1 = i_1)\ldots\mathbf{P}(X_n = i_n) = e^{-hk}R(h)^k \sum_{i_1,\ldots,i_n \in \mathbb{Z}: i_1+\ldots+i_n=k} \prod_{j=1}^{n} \frac{\mathbf{P}(X_1 = i_j)e^{hi_j}}{R(h)}$$

Введем $X^{(h)}$ — случайные величины с

$$\mathbf{P}(X^{(h)} = i) = \frac{\mathbf{P}(X = i)e^{hi}}{R(h)}.$$

Тогда

$$\mathbf{P}(S_n = k) = e^{-hk}R(h)^k\mathbf{P}(S_n^{(h)} = k).$$

При этом $\mathbf{E}X^{(h)} = \mathbf{E}Xe^{hX}/R(h) = m(h)$. Если найдется $h_{k/n}$, такое что $m(h_{k/n}) = k/n$, то в силу теоремы Гнеденко

$$\mathbf{P}(S_n^{(h_{k/n})} = k) \sim \frac{1}{\sqrt{2\pi}\sigma(h_{k/n})} e^{-\frac{(k-m(h_{k/n})n)^2}{2\sigma^2(h_{k/n})}} = \frac{1}{\sqrt{2\pi}\sigma(h_{k/n})}.$$

Здесь мы уже можем ее применять, потому что для $S_n^{(h_{k/n})}$ значение k — типичное, это математическое ожидание нашей суммы. На самом деле это не совсем аккуратно, потому что здесь распределение $X^{(h_{k/n})}$ меняется с ростом n, но это обходится тем, что теорема Гнеденко имеет некоторую равномерность по распределениям. Следовательно,

$$\mathbf{P}(S_n = k) \sim \frac{1}{\sqrt{2\pi n}\sigma(h_{k/n})} \exp(-n(h_{k/n}k/n - \ln R(h_{k/n}))).$$

Теорема Петрова получается отсюда суммированием по k. Аналогичным образом решается проблема с непрерывным случаем, только вместо теоремы Гнеденко рассматривается теорема Стоуна

$$\mathbf{P}(S_n \in [x, x+\Delta_n)) \sim \frac{\Delta_n}{\sqrt{2\pi n}\sigma} e^{-\frac{(x-\mu n)^2}{2\sigma^2 n}},\ n \to \infty,$$

где Δ_n — некоторая последовательность, стремящаяся к 0.