

2 Локальная предельная теорема Гнеденко

2.1 Локальная предельная теорема

Теорема 1 (Гнеденко). Пусть X_i — н.о.р. решетчатые случайные величины с a и шагом d , $EX_i = \mu$, $DX_i = \sigma^2$. Тогда

$$\mathbf{P}(S_n = k) = \frac{d}{\sqrt{2\pi n\sigma}} e^{-\frac{(k-\mu n)^2}{2n\sigma^2}} + o\left(\frac{1}{\sqrt{n}}\right),$$

при $k \in an + d\mathbb{Z}$, $n \rightarrow \infty$, причем $o()$ равномерно мало по указанным k .

Замечание 1. Заметим, что из теоремы вытекает, что при $|k - an| \leq t\sqrt{n}$

$$\mathbf{P}(S_n = k) = \frac{d + o(1)}{\sqrt{2\pi n\sigma}} e^{-\frac{(k-\mu n)^2}{2n\sigma^2}},$$

причем $o(1)$ равномерно мало по рассматриваемым k . Действительно, в силу теоремы

$$\frac{\mathbf{P}(S_n = k)}{\frac{d}{\sqrt{2\pi n\sigma}} e^{-\frac{(k-\mu n)^2}{2n\sigma^2}}} = 1 + \frac{o(1)}{e^{-\frac{(k-\mu n)^2}{2n\sigma^2}}} \rightarrow 1, \quad n \rightarrow \infty,$$

равномерность по k вытекает из ограниченности снизу знаменателя в последнем слагаемом.

Доказательство. Будем считать, что $a = 0$, $d = 1$, в противном случае перейдем от величин X_i к $(X_i - a)/d$.

Будем доказывать теорему на основе формулы обращения для дискретного случая

$$\mathbf{P}(S_n = k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itk} \psi_{X_1}^n(t) dt = \frac{1}{2\pi\sigma\sqrt{n}} \int_{-\pi\sigma\sqrt{n}}^{\pi\sigma\sqrt{n}} e^{-\frac{itk}{\sigma\sqrt{n}}} \psi_{X_1}^n\left(\frac{t}{\sigma\sqrt{n}}\right) dt.$$

Положим $x = \frac{k-an}{\sqrt{n\sigma}}$, тогда

$$\mathbf{P}(S_n = x) = \frac{1}{2\pi\sigma\sqrt{n}} \int_{-\pi\sigma\sqrt{n}}^{\pi\sigma\sqrt{n}} e^{-itx} e^{-it\frac{a}{\sigma}\sqrt{n}} \psi_{X_1}^n\left(\frac{t}{\sigma\sqrt{n}}\right) dt = \frac{1}{2\pi\sigma\sqrt{n}} \int_{-\pi\sigma\sqrt{n}}^{\pi\sigma\sqrt{n}} e^{-itx} \psi_{\frac{X_1-a}{\sigma}}^n\left(\frac{t}{\sqrt{n}}\right) dt.$$

При этом

$$\left| \int_{\delta\sigma\sqrt{n} < |t| < \pi\sigma\sqrt{n}} e^{-itx} \psi_{\frac{X_1-a}{\sigma}}^n\left(\frac{t}{\sqrt{n}}\right) dt \right| \leq \pi\sigma\sqrt{n}q^n,$$

где $\sup_{\delta < |s| < \pi} |\psi_{X_1}(s)| = q$. Заметим, что если $\psi_{X_1}(s_0) = e^{is_0b}$, то $\psi_{X_1-b}(s_0) = 1$, откуда

$$\mathbf{P}\left(X_1 \in \left\{b + \frac{2\pi k}{s_0}, k \in \mathbb{Z}\right\}\right) = 1.$$

Тогда шаг X_1 будет $\frac{2\pi}{s_0}$, что, в сочетании с решетчатостью и $d = 1$ дает $s_0 > 2\pi$. Это противоречит $\delta < s < \pi$, то есть $q < 1$ и оцениваемый интеграл экспоненциально мал по n .

В этом месте мы остановились во время лекции

При этом

$$\frac{1}{\sqrt{2\pi n\sigma}} e^{-\frac{x^2}{2}} = \frac{1}{2\pi\sigma\sqrt{n}} \int_{-\infty}^{\infty} e^{-itx} e^{-t^2/2} dt = \frac{1}{2\pi\sigma\sqrt{n}} \int_{|t| \leq \delta\sigma\sqrt{n}} e^{-itx} e^{-t^2/2} dt + \frac{o(1)}{\sqrt{n}},$$

где $o(1)$ равномерно по x стремится к 0 при $n \rightarrow \infty$, поскольку

$$\int_{|t| > \delta\sigma\sqrt{n}} e^{-t^2/2} dx = o(1).$$

Тем самым, для доказательства требуемой формулы, нам достаточно показать, что при достаточно малых δ

$$\frac{1}{2\pi\sigma\sqrt{n}} \int_{|t| < \delta\sigma\sqrt{n}} e^{-itx} \left(\psi_{\frac{X_1-a}{\sigma}}^n \left(\frac{t}{\sqrt{n}} \right) - e^{-t^2/2} \right) dt = \frac{\varepsilon}{\sqrt{n}}.$$

Для этого требуется оценить

$$\int_{|t| < \delta\sigma\sqrt{n}} e^{-t^2/2} \left| e^{n \left(\ln \psi_{\frac{X_1-a}{\sigma}} \left(\frac{t}{\sqrt{n}} \right) + \frac{t^2}{2n} \right)} - 1 \right| dt.$$

При любом $\varepsilon > 0$ и достаточно малом δ при $|s| < \delta$ верно неравенство

$$\left| \ln \psi_{\frac{X_1-a}{\sigma}}(s) + \frac{s^2}{2} \right| < \varepsilon s^2.$$

Тогда оцениваемый интеграл не превосходит

$$\int_{-\infty}^{\infty} e^{-t^2/2} \max(e^{\varepsilon t^2} - 1, 1 - e^{-\varepsilon t^2}) dt < \varepsilon_1$$

при любом $\varepsilon_1 > 0$ и достаточно малых $\varepsilon > 0$. □

2.2 Большие уклонения

Теорема Гнеденко подробно характеризует поведение S_n в $O(\sqrt{n})$ окрестности среднего, но оставляет открытым вопрос о том, как ведут себя $\mathbf{P}(S_n = k)$ при более удаленных k , давая лишь оценку $o(n^{-1/2})$.

Чтобы научиться работать с k , удаленными от μn , нам понадобится такая идея:

$$\begin{aligned} \mathbf{P}(S_n = k) &= \sum_{k_1 + \dots + k_n = k} \mathbf{P}(X_1 = k_1) \dots \mathbf{P}(X_n = k_n) = \\ &= \frac{e^{-hk}}{(\mathbf{E}e^{hX})^n} \sum_{k_1 + \dots + k_n = k} \mathbf{P}(X_1^{(h)} = k_1) \dots \mathbf{P}(X_n^{(h)} = k_n) = e^{-hk} R(h)^n \mathbf{P}(S_n^{(h)} = k), \end{aligned}$$

где $R(h) = \mathbf{E}e^{hX}$,

$$\mathbf{P}(X_1^{(h)} = k) = \frac{e^{hk}}{R(h)} \mathbf{P}(X_1 = k).$$

Такое распределение называют сопряженным. Мы хотим подобрать параметр h так, что

$$\mathbf{E}X_1^{(h)} = \frac{k}{n},$$

тогда теорема Гнеденко станет работать хорошо и

$$\mathbf{P}(S_n^{(h)} = k) = \frac{1}{\sqrt{2\pi n \mathbf{D}X_1^{(h)}}} + o\left(\frac{1}{\sqrt{n}}\right).$$

Значит,

$$\mathbf{P}(S_n = k) = \frac{1 + o(1)}{\sqrt{2\pi n \mathbf{D}X_1^{(h)}}} e^{-hk} R(h)^n.$$

Но прежде чем мы зададимся вопросами чему равны среднее и дисперсия такого распределения, отметим, что в наших рассуждениях есть неаккуратность. Поскольку мы хотим получить соотношение

$\mathbf{E}X_1^{(h)} = k/n$, параметр h будет зависеть от n . Но тогда при каждом n у X_i будет свое распределение, а в исходной теореме Гнеденко такого не позволялось. Поэтому нам понадобится равномерность о-малого в теореме Гнеденко по всем h .

Для этого отметим следующее свойство проведенных нами в доказательстве теоремы Гнеденко оценок.

Замечание 2. Если $X_i^{(h)}$, $h \in [b, c]$, семейство решетчатых случайных величин с одинаковым a, d , причем при каждом h величины $X_1^{(h)}, \dots, X_n^{(h)}, \dots$ — н.о.р. с $\mathbf{E}X_1^{(h)} = \mu(h)$, $\mathbf{D}X_i^{(h)} = \sigma^2(h)$. Предположим, что выполнены следующие условия:

1. Выполнены условия $0 < \inf_{h \in [b, c]} \sigma(h) \leq \sup_{h \in [b, c]} \sigma(h) < \infty$.
2. При любом $\delta > 0$ справедливое неравенство $\sup_{h \in [b, c]} \sup_{s \in [\delta, \pi]} |\psi_{X_1^{(h)}}(s)| = q < 1$,
3. При любом $\varepsilon > 0$ найдется такое $\delta > 0$, что при всех $h \in [b, c]$, $|s| < \delta$

$$\left| \ln \psi_{\frac{X_1^{(h)} - \mu(h)}{\sigma(h)}}(s) + \frac{s^2}{2} \right| < \varepsilon.$$

Тогда при любом t и $|k - an| \leq t\sqrt{n}$

$$\mathbf{P}(S_n = k) = \frac{1}{\sqrt{2\pi n\sigma(h)}} e^{-\frac{(k - \mu(h)n)^2}{2n\sigma(h)}} + o(n^{-1/2}),$$

причем $o(1)$ равномерно мало по рассматриваемым k и $h \in [b, c]$.