Семинар 13. Законы больших чисел и предельные теоремы.

Начнем с того, что сформулируем несколько полезных неравенств.

Неравенство Маркова. При $X \geq 0$ п.н., $x > 0$ верно неравенство

$$\mathbf{P}(X \geq x) \leq \frac{\mathbf{E}X}{x}.$$

Неравенство Чебышева. При $\mathbf{D}X < \infty$, $x > 0$ верно неравенство

$$\mathbf{P}(|X - \mathbf{E}X| \geq x) \leq \frac{\mathbf{D}X}{x^2}.$$

Неравенство Колмогорова. При X_i — независимых, $\mathbf{E}X_i = 0$, $\mathbf{D}X_i < \infty$ справедливо неравенство

$$\mathbf{P}(\max_{i \leq n} |S_i| \geq x) \leq \frac{\mathbf{D}S_n}{x^2}.$$

Говорят, что последовательность случайных величин $X_1, ...$ удовлетворяет *закону больших чисел* (ЗБЧ), если существует константа C, т.ч.

$$\frac{X_1 + ... + X_n}{n} \xrightarrow{\mathbf{P}} C, n \to \infty.$$

Говорят, что последовательность случайных величин $X_1, ...$ удовлетворяет *усиленному закону больших чисел* (УЗБЧ), если существует константа C, т.ч.

$$\frac{X_1 + ... + X_n}{n} \to C\text{п.н. }, n \to \infty.$$

Пример 1. Простейший пример ЗБЧ дает нам схема Бернулли. Пусть X_i — результаты броска монеты с вероятностью единицы (орла) p. Тогда $\hat{p} = (X_1 + ... + X_n)/n$ — число орлов, деленное на число бросков монеты, должно сходиться к p. В случае УЗБЧ мы утверждаем, что эта сходимость будет иметь место при увеличении числа бросков (кроме заранее выделенного множества нулевой меры). В случае ЗБЧ — что при больших n вероятность отличия $\hat{p}$ и p более чем на эпсилон будет мала.

Теорема 1 (ЗБЧ Чебышева). Пусть X_i — н.о.р. величины $\mathbf{E}X_i = \mu$, $\mathbf{D}X_i = \sigma^2 < \infty$. Тогда они удовлетворяют ЗБЧ с $= \mu$.

Доказательство этой теоремы крайне просто

$$\mathbf{P}\left(\left|\frac{X_1 + ... + X_n}{n} - \mu\right| > \varepsilon\right) \leq \frac{\mathbf{D}\left(\frac{X_1 + ... + X_n}{n}\right)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0,\ n \to \infty$$

Как видно из доказательства, нам не важна независимость, а только некореллированность X_i, а также не требуется одинаково распределенность, достаточно

$$\frac{1}{n}\sum_{i=1}^{n} \mu_i \to C,\ \frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2 \to 0.$$

Кроме того, можно заметить, что мы доказываем сходимость по вероятности с помощью сходимости в L^2, поэтому сходимость в среднем порядка 2 мы тоже доказали.

Теорема 1 в случае н.о.р. может быть усилена

Теорема 2 (ЗБЧ Хинчина). Пусть X_i — н.о.р. величины, $\mathbf{E}X_i = \mu$. Тогда они удовлетворяют ЗБЧ с $= \mu$.

Доказательство опять-таки несложно. Воспользуемся тем, что сходимость по вероятности к константе равносильна сходимости по распределению, а значит сходимости хар.функций. Тогда

$$\psi_{\frac{X_1+...+X_n}{n}}(t) = \psi^n_{X_1}(t/n) = \exp(n \ln \psi_{X_1}(t/n)).$$

Остается заметить, что

$$(\ln \psi_{X_1})'(1) = \frac{\psi'_{X_1}(1)}{\psi_{X_1}(1)} = i\mathbf{E}X_1 = i\mu,$$

откуда $n \ln \psi_{X_1}(t/n) \to e^{it\mu}$, что и требовалось доказать.

Оказывается, что для н.о.р. величин верны и более сильные теоремы

Теорема 3 (УЗБЧ Колмогорова). Пусть X_i — н.о.р. величины. Тогда они удовлетворяют УЗБЧ тогда и только тогда, когда существует $\mathbf{E}X_i = C$.

Теорема 4 Пусть X_i — н.о.р. величины. Тогда они удовлетворяют ЗБЧ тогда и только тогда, когда $n\mathbf{P}(|X_1| > n) \to 0$, $n \to \infty$ и $\mathbf{E}(X; |X| \leq n) \to C$.

Пример 2. Рассмотрим величины X_i с симметричным распределением, т.ч. $\mathbf{P}(|X| > x) \sim (x \ln x)^{-1}$, $x \to \infty$. Тогда $n\mathbf{P}(|X_1| > n) \to 0$, $\mathbf{E}(|X|; |X| \leq n) = 0$, но $\mathbf{E}X$ не определено. Следовательно, X_i удовлетворяют ЗБЧ, но не УЗБЧ.

Пример 3. УЗБЧ сам по себе довольно удивительный факт. Рассмотрим, например, случайное равномерное число из $[0, 1)$. Последовательность его двоичных цифр $X_1, X_2, ...$. есть последовательность н.о.р. сл.в., равномерно распределенных на $\{0, 1\}$. Тогда $(X_1 + ... + X_n)/n$ — доля единиц среди первых n цифр, почти при всех разыгранных числах будет стремится к 0.5.

Если $\frac{X_1+...+X_n-\mu n}{n} \to 0$, то можно ожидать, что при делении на что-то меньшее n мы получим какой-то конечный предел. Для н.о.р. оказывается верна следующая теорема, называемая центральная предельная теорема:

Теорема 5. Пусть X_i — н.о.р., $\mathbf{E}X_i = \mu$, $0 < \mathbf{D}X_i = \sigma^2 < \infty$. Тогда

$$\frac{X_1 + ... + X_n - \mu n}{\sigma\sqrt{n}} \xrightarrow{d} Z \sim \mathcal{N}(0, 1).$$

Доказательство теоремы похоже на доказательство теоремы 2, только разложение функции $\ln \psi$ придется делать до второго члена.

Следствие.

$$\mathbf{P}\left(\frac{X_1 + ... + X_n - \mu n}{\sigma\sqrt{n}} \in [a, b]\right) \to \Phi(b) - \Phi(a), \ n \to \infty,$$

где Φ — ф.р. $\mathcal{N}(0, 1)$.

Пример 4. Посмотрите какой удивительный факт. Как бы не были устроены величины

X_i с заданными средними и дисперсиями, вероятности попадания $S_n = X_1 + ... + X_n$ в различные множества будут при больших n одинаково устроены.
Скажем, рассмотрим $X_i \sim \mathcal{N}(0,1)$, $Y_i = \pm 1$ равновероятно. Тогда $\mathbf{E}X_1 = 0$, $\mathbf{D}X_1 = 1$, $\mathbf{E}Y_1 = 0$, $\mathbf{D}Y_1 = 1$.
Следовательно,

$$\mathbf{P}\left(\frac{X_1 + ... + X_n}{\sqrt{n}} \in [a,b]\right) \approx \mathbf{P}\left(\frac{Y_1 + ... + Y_n}{\sqrt{n}} \in [a,b]\right) \approx \Phi(b) - \Phi(a).$$

При этом $X_1 + ... + X_n \sim \mathcal{N}(0,n)$, а $Y_1 + ... + Y_n$ имеют дискретное распределение с $\mathbf{P}(Y_1 + ... + Y_n = k) = C_n^{(k+n)/n} 2^{-n}$, k,n одной четности и 0 иначе.
Пример 5. Какая вероятность, что при 1600 подбрасываниях монеты выпадет менее 700 орлов?

$$\mathbf{P}(X_1 + ... + X_n \le 700) = \mathbf{P}\left(\frac{X_1 + ... + X_n - n\mathbf{E}X_1}{\sqrt{n\mathbf{D}X_1}} \le \frac{700 - 800}{\sqrt{\frac{1}{4}1600}}\right) = \Phi(-5) \approx 0,$$

откуда вероятность довольно мала.
При существовании $\mathbf{E}X^3$ можно получить оценку

$$\left|\mathbf{P}\left(\frac{X_1 + ... + X_n - \mu n}{\sigma\sqrt{n}} \le x\right) - \Phi(x)\right| \le C\frac{\mathbf{E}(X - \mathbf{E}X)^3}{(\mathbf{D}X)^{3/2}\sqrt{n}},$$

где C – некоторая константа, не зависящая от распределения X_i, $C \le 0.5$.

Можно обобщать полученные результаты на разнораспределенные величины. Сформулируем одну из таких теорем — теорему Ляпунова
Теорема 6. Пусть X_i независимы, $\mathbf{E}X_i = \mu_i$, $\mathbf{D}X_i = \sigma_i^2$, $\sigma^2(n) = \sigma_1^2 + ...\sigma_n^2$, $\mathbf{E}|X_i|^3 < \infty$, причем

$$\frac{1}{\sigma(n)^{3/2}} \sum_{i=1}^{n} \mathbf{E}|X_i|^3 \to 0,\ n \to \infty.$$

Тогда

$$\frac{1}{\sigma(n)} \sum_{i=1}^{n} (X_i - \mu_i) \xrightarrow{d} Z \sim \mathcal{N}(0,1).$$

13.1.0. Из поселка с 2500 жителей раз в сутки ходит поезд. Жители независимо от других в среднем 6 раз в месяц ездят на поезде в случайные дни. Сколько мест сделать в поезде, чтобы вероятность переполнения была не больше 0.01.
13.2.0. Театр, вмещающий 1000 человек, имеет два разных входа с одинаковыми гардеробами. Сколько сделать мест в гардеробах, чтобы в среднем в 99 случаях из 100 все зрители могли раздеться в гардеробе входа, через который они вошли? Зрители приходят парами, равновероятно выбирая один из входов
13.3.0. Игральная кость подбрасывается до тех пор, пока общая сумма выпавших очков не превысит 700. Оценить вероятность того, что для этого потребуется более 210 бросаний.

13.1.1. ξ_i — н.о.р., $\mathbf{E}\xi_i = a$, $\eta_n = (\xi_1 + ... + \xi_n)/n$. Показать, что $\sqrt[n]{\eta_1 ... \eta_n} \to a$ п.н.
13.2.1. Пусть X_i - число исходов i среди n испытаний с N исходами, вероятность i-ого исхода - p_i. $\eta_n = X_1 a_1 + ... + X_n a_n$, где a_i — заданы. Оценить $\mathbf{P}(\eta_n \in (a, b))$ из ЦПТ.
13.3.1. Верен ли ЗБЧ для Y_i, где X_i — независимы, а) $(1 - \mathbf{P}(X_i = 0))/2 = \mathbf{P}(X_i = 1) = \mathbf{P}(X_i = -1) = 1/(2i)^2$, $Y_i = iX_i$. б) $\mathbf{P}(Y_i = 2^i) = \mathbf{P}(Y_i = -2^i) = 1/2$.
13.1.2 Найти предел интеграла $\int_{[0,1]^n} \frac{x_1^2 + ... + x_n^2}{x_1 + ... + x_n} dx_1 ... dx_n$ при $n \to \infty$.
13.2.2 X_1,...,X_n - н.о.р. с $\mathbf{E}X = 0$, $\mathbf{D}X = 1$. Доказать, что $\mathbf{P}(\frac{X_1 + ... + X_n}{\sqrt{X_1^2 + ... + X_n^2}} \le x) \to \Phi(x)$.
13.3.2 Привести пример последовательности нез. сл.в., принимающих три значения, имеющих м.о. 0 и дисперсию 1, для которых не выполнена ЦПТ.
13.1.3. Пусть X_1,...,X_n - н.о.р с $\mathbf{E}X = 0$, $\mathbf{D}X = 1$. Найти $\lim_{n\to\infty} \mathbf{P}((X_1 + ... + X_{\nu_n})(\nu_n)^{-1/2} \le x)$, где сл.в. $\nu_n \in \mathbb{N}$ не зависит от X_i и $\mathbf{P}(\nu_n \le M) \to 0$ при $n \to \infty$.
13.2.3. Пусть $X_1 \sim R[0, 1]$, $X_i = \{10^i X_1\}$. Док-ть, что X_i удовлетворяют а) ЗБЧ б) ЦПТ
13.3.3. Θ_i — сл.в. из $(0, 1)$, $1 - P(X_i = 0|\Theta_i = \theta) = P(X_i = 1|\Theta_i = \theta) = \theta$. Выполнена ли для $S_n = X_1 + ... + X_n$ ЦПТ, если а) Θ_i — н.о.р., б) $\Theta_i = \Theta$ — одна и та же величина.