Екзамен з Теорії Ймовірності

Захарова Дмитра Олеговича, МП-31

17 червня 2024 р.

Білет №5

Вміст

1 Теоретичне питання

Умова. Теорема Хінчина (закон великих чисел).
Відповідь. Отже, сформулюємо саму теорему.

Theorem 1.1. Нехай маємо послідовність незалежних однаково розподілених випадкових величин $\{\xi_n\}_{n\in\mathbb{N}}$, що визначені на однаковому ймовірністному просторі $(\Omega, \mathcal{F}, \Pr)$, причому їх математичне сподівання скінченне і $\mu = \mathbb{E}[\xi_n]$. Тоді, справедливий **закон великих чисел у формі Хінчина**:

$$\forall \varepsilon > 0 : \lim_{n\to\infty} \Pr\left[\left|\frac{1}{n}\sum_{k=1}^{n} \xi_k - \mu\right| \geq \varepsilon\right] = 0 \tag{1}$$

Remark. Практично, теорема означає наступне: середнє арифметичне великої кількості незалежних однаково розподілених випадкових величин з великою ймовірністю зближується до їхнього спільного математичного сподівання. На відміну від теореми Чебишева, тут не вимагається існування дисперсії.

Доведення. Спочатку зафіксуємо *ідею доведення*. Зафіксуємо $\delta > 0$ і введемо наступні випадкові величини:

$$\eta_j(\omega) = \xi_j(\omega) \cdot \mathbb{1}(|\xi_j(\omega)| < \delta n),\ \zeta_j(\omega) = \xi_j(\omega) \cdot \mathbb{1}(|\xi_j(\omega)| \geq \delta n) \tag{2}$$

За побудовою видно, що $\xi_j \equiv \eta_j + \zeta_j$. Далі, ми окремо розглянемо величини $\{\eta_j\}_{j=1}^n$ та $\{\zeta_j\}_{j=1}^n$, для яких знайдемо оцінки, які врешті-решт допоможуть нам оцінити і ξ_j.

Випадкові величини $\{\eta_j\}$. Нехай тепер $\Phi_\xi(x)$ – функція розподілу кожної з випадкових величин ξ_j. Помітимо, що для величин η_j існує як математичне сподівання, так і дисперсія. Дійсно,

$$\mu_n := \mathbb{E}[\eta_j] = \int_{(-\delta n, +\delta n)} x d\Phi_\xi(x) \leq \int_{\mathbb{R}} x d\Phi_\xi(x) = \mu, \tag{3}$$

$$\mathrm{Var}[\eta_j] = \int_{(-\delta n, +\delta n)} x^2 d\Phi_\xi(x) - \mu_n^2 \leq \delta n \int_{(-\delta n, +\delta n)} |x| d\Phi_\xi(x) \leq \delta\beta n, \tag{4}$$

де ми позначили $\beta := \int_{\mathbb{R}} |x| d\Phi_\xi(x) < \infty$. Отже, звідси $\mu_n \xrightarrow[n\to\infty]{} \mu$, тобто

$$(\forall \varepsilon > 0)\ (\exists n_\varepsilon \in \mathbb{N})\ (\forall n \geq n_\varepsilon) : \{|\mu_n - \mu| < \varepsilon\} \tag{5}$$

Тепер, введемо наступну борельову функцію $f_n(x) = x \cdot \mathbb{1}(|x| < \delta n)$. Оскільки $\{\xi_j\}_{j=1}^n$ є незалежними, то і група величин $\{f_n(\xi_j)\}_{j=1}^n$ теж є незалежними.

Також видно, що $\eta_j := f_n(\xi_j)$. Застосуємо другу нерівність Чебушева для випадкової величини $M_n := \frac{1}{n}\sum_{j=1}^n \eta_j$:

$$\Pr\left[|M_n - \mu n| \geq \varepsilon\right] \leq \frac{\mathrm{Var}\left[\sum_{j=1}^n \eta_j\right]}{n^2\varepsilon^2} = \frac{\sum_{j=1}^n \mathrm{Var}\left[\eta_j\right]}{n^2\varepsilon^2} \leq \frac{\delta\beta}{\varepsilon^2} \tag{6}$$

Тепер помітимо дуже важливий факт:

$$\begin{aligned}\{\omega \in \Omega : |M_n - \mu| \geq 2\varepsilon\} &= \{\omega \in \Omega : |M_n - \mu_n + \mu_n - \mu| \geq 2\varepsilon\} \\ &\subset \{\omega \in \Omega : |M_n - \mu| \geq \varepsilon\} \cup \{\omega \in \Omega : |\mu_n - \mu| \geq \varepsilon\}\end{aligned} \tag{7}$$

Тоді, з півадитивності і монотонності міри випливає

$$\Pr[|M_n - \mu| \geq 2\varepsilon] \leq \Pr[|M_n - \mu_n| \geq \varepsilon] + \Pr[|\mu_n - \mu| \geq \varepsilon] \tag{8}$$

Помітимо, що для всіх $n \geq n_\varepsilon$ подія $|\mu_n - \mu| \geq \varepsilon$ не виконується, а тому

$$(\forall n \geq n_\varepsilon) : \{\Pr[|M_n - \mu| \geq 2\varepsilon] \leq \frac{\beta\delta}{\varepsilon^2}\} \tag{9}$$

Випадкові величини $\{\zeta_j\}$. Тепер перейдемо до величин ζ_j. Вони однаково розподілені, а тому $\Pr[\zeta_j \neq 0] = \Pr[\zeta_n \neq 0]$. Окрім того,

$$\Pr[\zeta_j \neq 0] = \Pr[\zeta_n \neq 0] = \Pr[|\xi_n| \geq n\delta] = \int_{|x|\geq\delta n} d\Phi_\xi(x) \leq \frac{1}{\delta n}\int_{|x|\geq\delta n} |x| d\Phi_\xi(x) \tag{10}$$

Оскільки $\mathbb{E}[|\xi_n|] = \beta = \int_{\mathbb{R}} |x| d\Phi_\xi(x) < \infty$, то

$$\lim_{n\to\infty} \int_{|x|\geq\delta n} |x| d\Phi_\xi(x) = 0 \tag{11}$$

Звідси випливає

$$(\forall\delta > 0)\ (\exists n_\delta \in \mathbb{N})\ (\forall n \geq n_\delta) : \left\{\int_{|x|\geq\delta n} |x| d\Phi_\xi(x) < \delta^2\right\} \tag{12}$$

Звідси для всіх $n \geq n_\delta$ справедливо $\Pr[\zeta_j \neq 0] = \Pr[\zeta_n \neq 0] \leq \frac{\delta}{n}$. Далі, оскільки

$$\bigcap_{j=1}^n \{\omega \in \Omega : \zeta_j(\omega) = 0\} \subset \{\omega \in \Omega : \sum_{j=1}^n \zeta_j(\omega) = 0\}, \tag{13}$$

то $\{\omega \in \Omega : \sum_{j=1}^n \zeta_j(\omega) \neq 0\} \subset \bigcup_{j=1}^n \{\omega \in \Omega : \zeta_j(\omega) \neq 0\}$. Тому,

$$\Pr\left[\sum_{j=1}^n \zeta_j \neq 0\right] \leq \sum_{j=1}^n \Pr[\zeta_j \neq 0] \leq \delta \tag{14}$$

Оцінка $\{\xi_j\}$. Тому, остаточно

$$\Pr\left[\left|\frac{1}{n}\sum_{j=1}^{n}\xi_j-\mu\right|\geq 4\varepsilon\right]=\Pr\left[\left|\frac{1}{n}\sum_{j=1}^{n}(\eta_j+\zeta_j)-\mu\right|\geq 4\varepsilon\right]$$
$$\leq \Pr\left[\left|\frac{1}{n}\sum_{j=1}^{n}\eta_j-\mu\right|\geq 2\varepsilon\right]+\Pr\left[\left|\frac{1}{n}\sum_{j=1}^{n}\zeta_j\right|\geq 2\varepsilon\right]$$
$$\Pr\left[\left|\frac{1}{n}\sum_{j=1}^{n}\eta_j-\mu\right|\geq 2\varepsilon\right]+\Pr\left[\sum_{j=1}^{n}\zeta_j\neq 0\right] \tag{15}$$

Тепер, зафіксуємо $n'=\max\{n_\varepsilon, n_\delta\}$, тоді для всіх $n\geq n'$:

$$\Pr\left[\left|\frac{1}{n}\sum_{j=1}^{n}\xi_j-\mu\right|\geq 4\varepsilon\right]\leq\frac{\beta\delta}{\varepsilon^2}+\delta \tag{16}$$

Оскільки $\varepsilon,\delta>0$ – довільні додатні числа, то остаточно

$$(\forall\varepsilon>0)\ \exists\lim_{n\to\infty}\Pr\left[\left|\frac{1}{n}\sum_{j=1}^{n}\xi_j-\mu\right|\geq 4\varepsilon\right]=0 \tag{17}$$

Теорему доведено.

2 Практичне питання №1

Умова. В жовтій коробці 4 шоколадних цукерки і 6 карамельок, а в блакитній коробці – 7 шоколадних цукерок і 3 карамельки. Із навмання обраної коробки дитина взяла 2 цукерки. Яка ймовірність, що вони взяті із жовтій коробки, якщо обидві цукерки виявились шоколадними ?

Відповідь. Розглянемо три події: E – обидві цукерки виявились шоколадними, Y – обидві цукерки взяті з жовтої коробки, B – з блакитної. Тоді, за умовою, нам треба знайти ймовірність події Y при умові E, тобто $\Pr[Y|E]$. За **формулою Баєса**, маємо

$$\Pr[Y|E] = \frac{\Pr[E \mid Y]\Pr[Y]}{\Pr[E]} = \frac{\Pr[E \mid Y]\Pr[Y]}{\Pr[E \mid Y]\Pr[Y] + \Pr[E \mid B]\Pr[B]} \tag{18}$$

(тут Y, B утворюють повну групу подій). Легко бачити, що оскільки коробка обирається навмання, то $\Pr[Y] = \Pr[B] = \frac{1}{2}$. Отже, залишається знайти лише $\Pr[E \mid Y]$ та $\Pr[E \mid B]$.

Якщо випала жовта коробка, то маємо C_4^2 способів вибрати шоколадну цукерку та C_{10}^2 вибрати 2 цукерки з 10. Тому, за класичним означенням ймовірності, ймовірність дістати дві шоколадні цукерки дорівнює $C_4^2 \Big/ C_{10}^2$. Аналогічно, якщо випала синя коробка, то $C_7^2 \Big/ C_{10}^2$. Таким чином,

$$\Pr[E \mid Y] = \frac{C_4^2}{C_{10}^2} = \frac{2}{15},\ \Pr[E \mid B] = \frac{C_7^2}{C_{10}^2} = \frac{7}{15} \tag{19}$$

Отже, остаточно:

$$\Pr[Y|E] = \frac{\Pr[E \mid Y]\Pr[Y]}{\Pr[E \mid Y]\Pr[Y] + \Pr[E \mid B]\Pr[B]} = \frac{\frac{2}{15}\cdot\frac{1}{2}}{\frac{2}{15}\cdot\frac{1}{2} + \frac{7}{15}\cdot\frac{1}{2}} = \frac{2}{9} \tag{20}$$

Відповідь. $\frac{2}{9}$.

3 Практичне питання №2

Умова. Ймовірність того, що деяка банкнота виявиться фальшивою дорівнює $p = 0.0001$. Знайти ймовірність того, що з $N = 4000$ перевірених купюр буде знайдено більш ніж $n = 2$ фальшиві.

Відповідь. Введемо випадкову величину ξ – кількість фальшивих купюр. Тоді, цю величину можна вважати випадковою величиною, що має біноміальний розподіл з параметрами $N = 4000$ та $p = 0.0001$, де в якості експерименту вважаємо "чи є банкнота фальшивою", де подія "банкнота фальшива" відбувається з ймовірністю p. В такому разі, оскільки $\xi \sim \text{Bin}(N, p)$, то розподіл має вигляд

$$\text{Pr}[\xi = k] = C_N^k p^k (1-p)^{N-k}, k \in \{0, 1, \ldots, N\}. \tag{21}$$

Подія "більш ніж $n = 2$ фальшиві купюри" відповідає $\xi > 2$, тобто шукана подія має вигляд $\text{Pr}[\xi > n]$. В цілому, на цьому етапі можна знайти ймовірність "в лоб":

$$\text{Pr}[\xi > n] = \sum_{k=n+1}^{N} C_N^k p^k (1-p)^{N-k}, \tag{22}$$

проте її дуже складно обрахувати. Можна трошки спростити життя і замість події $\xi > 2$ розглядати подію $\xi \leq 2$, тоді

$$\text{Pr}[\xi > 2] = 1 - \text{Pr}[\xi \leq 2] = 1 - \sum_{k=0}^{2} C_N^k p^k (1-p)^{N-k}. \tag{23}$$

В цьому випадку, нам лише залишиться порахувати три наступних вирази: $C_N^0(1-p)^N, C_N^1 p(1-p)^{N-1}, C_N^2 p^2(1-p)^{N-2}$, проте це все ще достатньо складно, оскільки треба рахувати значення накшталт $(0.0001)^{4000}$.

Тому, можемо скористатися апроксимацією біноміального розподілу за розподілом Пуасона. Згідно цієї апроксимації, якщо N велике, а Np мале, то наближено можна вважати, що $\xi \sim \text{Poiss}(\lambda)$, де $\lambda = Np = 0.4$. В такому разі, можемо знайти ймовірність події $\xi \leq 2$ за формулою

$$\text{Pr}[\xi \leq 2] = \sum_{k=0}^{2} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-0.4} \left(1 + 0.4 + \frac{0.4^2}{2}\right) \approx 0.9921. \tag{24}$$

Отже, шукана ймовірність дорівнює

$$\text{Pr}[\xi > 2] = 1 - \text{Pr}[\xi \leq 2] \approx 1 - 0.9921 = 0.0079. \tag{25}$$

Відповідь. Ймовірність того, що з $N = 4000$ перевірених купюр буде знайдено більш ніж $n = 2$ фальшиві наближено дорівнює 0.0079.

4 Практичне питання №3

Умова. Випадкова величина ξ має стандартний нормальний розподіл, а випадкова величина η має показниковий розподіл із математичним сподіванням 1.5. Знайти щільність розподілу випадкового вектору (ξ, η), якщо його компоненти є незалежними випадковими величинами

Відповідь. Згідно означенню, $\xi \sim \mathcal{N}(0,1)$, тобто щільність розподілу ξ має вигляд

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \; x \in \mathbb{R}. \tag{26}$$

Оскільки $\eta \sim \mathrm{Exp}(\lambda)$, то її щільність розподілу має вигляд

$$f_\eta(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases} \tag{27}$$

За умовою ми знаємо, що $\mathbb{E}[\xi] = \frac{3}{2}$, звідси можна знайти λ. Дійсно,

$$\mathbb{E}[\xi] = \int_{\mathbb{R}} x f_\eta(x) dx = \lambda \int_0^{+\infty} x e^{-\lambda x} dx \tag{28}$$

Далі можна проінтегрувати частинами, якщо візьмемо $u = x, du = dx$ та $dv = e^{-\lambda x} dx$:

$$\mathbb{E}[\xi] = \lambda \left[-\frac{1}{\lambda} x e^{-\lambda x} \Big|_{x \to 0}^{x \to +\infty} + \frac{1}{\lambda} \int_0^{+\infty} e^{-\lambda x} dx \right] = \int_0^{+\infty} e^{-\lambda x} dx = \frac{1}{\lambda} \tag{29}$$

Отже, $\lambda = \frac{2}{3}$. Як відомо, якщо дві випадкові величини є незалежними, то їхній спільний розподіл є добутком їхніх щільностей розподілу. Тобто щільність розподілу випадкового вектору (ξ, η) має вигляд

$$f_{(\xi,\eta)}(x, y) = f_\xi(x) f_\eta(y) = \begin{cases} \frac{2}{3\sqrt{2\pi}} e^{-\frac{x^2}{2}} e^{-\frac{2y}{3}}, & y \geq 0 \\ 0, & y < 0 \end{cases}. \tag{30}$$

Отже, маємо нульову щільність розподілу випадкового вектору (ξ, η) при $y < 0$, а при $y \geq 0$ маємо

$$f_{(\xi,\eta)}(x, y) = \frac{2}{3\sqrt{2\pi}} \exp\left\{ -\frac{x^2}{2} - \frac{2y}{3} \right\}. \tag{31}$$

Відповідь. $f_{(\xi,\eta)}(x, y) = \frac{2}{3\sqrt{2\pi}} e^{-\frac{x^2}{2} - \frac{2y}{3}} \cdot \mathbb{1}_{[0,+\infty)}(y)$