Лекція 4. **Закон великих чисел у формі Хінчина**

Нерівність Чебишова дозволяла доводити для випадкових величин закон великих чисел при умові існування дисперсій. Базуючись на апараті характеристичних функцій можна поглибити аналіз і відкинути цю умову.

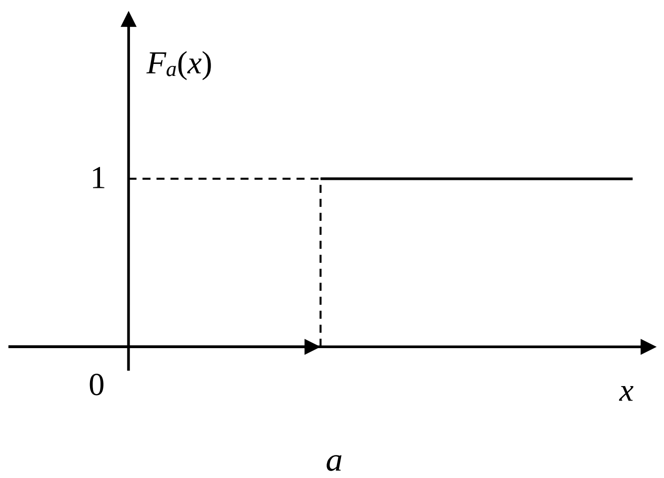
Теорема 1. Нехай $\xi_1, \xi_2, ..., \xi_n, ...$ – послідовність незалежних, однаково розподілених випадкових величин, які мають $M\xi_1 = a$ і нехай $S_n = \xi_1 + ... + \xi_n$. Тоді $\frac{S_n}{n} \underset{n\to\infty}{\overset{P}{\Rightarrow}} a$ при $n \to \infty$.

Доведення. Оскільки

$$P\left\{\left|\frac{S_n}{n} - a\right| \ge \varepsilon\right\} = P\left\{\frac{S_n}{n} \ge a + \varepsilon\right\} + P\left\{\frac{S_n}{n} \le a - \varepsilon\right\} =$$

$$= F_{\frac{S_n}{n}}(a-\varepsilon) + \left(1 - F_{\frac{S_n}{n}}\left((a+\varepsilon) - 0\right)\right) \le F_{\frac{S_n}{n}}(a-\varepsilon) + \left(1 - F_{\frac{S_n}{n}}\left(a + \frac{\varepsilon}{2}\right)\right) \to 0$$.Позаяк,

$$F_{\frac{S_n}{n}}(a-\varepsilon) \to F_a(a-\varepsilon) = 0, \text{ а } F_{\frac{S_n}{n}}\left(a + \frac{\varepsilon}{2}\right) \to F_a\left(a + \frac{\varepsilon}{2}\right) = 1$$



Отже, достатньо показати, що $\frac{S_n}{n} \underset{n\to\infty}{\overset{сл}{\Rightarrow}} a$, тобто, що для кожного фіксованого t

$$f_{\frac{s_n}{n}}(t) \underset{n\to\infty}{\to} e^{iat}.$$

Характеристична функція випадкової величини ξ_1 в деякому околі точки ”0” задовольняє нерівності

$$|f(t) - 1| < \frac{1}{2}.$$

Тому для таких t можна визначити функцію $l(t) = \ln f(t)$ - береться головне значення логарифму (Функція $Ln(z)$ багатозначна ($Ln(z) = \ln \rho e^{i(\varphi + 2\pi k)} = \ln \rho + i(\varphi + 2\pi k), k = 0, \pm 1, ...$) Значення $\ln z = \ln \rho + i\varphi$, $-\pi < \varphi \le \pi$ називають головним значення логарифму.

Позаяк, існує математичне сподівання ξ_1, то існує

$$l'(0)=\frac{f'(0)}{f(0)}=ia .$$

При кожному фіксованому t і достатньо великому n значення $l\left(\frac{t}{n}\right)$ визначене і $f_{\frac{s_n}{n}}(t)=f^n\left(\frac{t}{n}\right)=e^{nl\left(\frac{t}{n}\right)}$. Так як $l(0)=0$, то при $n\to\infty$

$$e^{nl\left(\frac{t}{n}\right)}=\exp\left\{t\frac{l\left(\frac{t}{n}\right)-l(0)}{\frac{t}{n}}\right\}\to e^{l'(0)t}=e^{iat} .$$

Теорему доведено.

Доведений результат будемо називати законом великих чисел у формі Хінчина.

Центральна гранична теорема

Розглянемо проблему апроксимації сум незалежних випадкових величин. Нехай $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x}e^{-\frac{u^2}{2}}du$ – функція розподілу нормального закону. Наступний результат є одним з варіантів центральної граничної теореми – теореми про апроксимацію розподілу сум величин нормальним розподілом.

Теорема 2. Якщо випадкові величини $\xi_1,\xi_2,...$ – незалежні, однаково розподілені і мають скінченні $M\xi_1=a$ і $D\xi_1=\sigma^2>0$, то для будь-якого $x\in(-\infty,+\infty)$

$$\lim_{n\to\infty}P\left\{\frac{\xi_1+...+\xi_n-na}{\sigma\sqrt{n}}\le x\right\}=\Phi(x) .$$

Доведення. Позначимо $S_n=\frac{\xi_1+...+\xi_n-na}{\sigma\sqrt{n}}$. Тоді необхідно показати, що

$$\lim_{n\to\infty}f_{S_n}(t)=e^{-\frac{t^2}{2}} .$$

Нехай $\tilde{\xi}_1=\xi_1-a$. З властивості 6) для характеристичних функцій маємо

$$f_{\tilde{\xi}_1}(t)=1-\frac{t^2\sigma^2}{2}+o\left(t^2\right) \text{ при } t\to 0 .$$

Звідси випливає, що для довільного фіксованого t

$$f_{S_n}(t)=\left[f_{\tilde{\xi}_1}\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n=\left(1-\frac{t^2}{2n}+o\left(\frac{t^2}{n}\right)\right)^n\underset{n\to\infty}{\to}e^{-\frac{t^2}{2}} .$$

Теорему доведено.

Застосування:

$\xi_1+\ldots+\xi_n\approx\sqrt{n}\sigma\xi_0+na$, де $\xi_0 \sim N(0,1)$ ($p_{\xi_1+\xi_2}(x)=\int\limits_{-\infty}^{+\infty}p_{\xi_1}(x-u)\,p_{\xi}(u)\,du$)

****

Нехай $\xi_1, \xi_2, ..., \xi_n, ...$ - послідовність незалежних, однаково розподілених випадкових величин, причому

$$\xi_1 = \begin{cases} 1, & \text{з імовірністю } p, \\ 0, & \text{з імовірністю } 1 - p = q. \end{cases}$$

Тоді $M\xi_1 = p, \quad D\xi_1 = p - p^2 = pq, \quad \mu_n = \sum_{k=1}^{n} \xi_k$ - число успіхів у n випробовуваннях Бернуллі. Тепер інтегральна теорема Муавра-Лапласа для μ_n є наслідком теореми 2.

****

Розглянемо ще два застосування центральної граничної теореми.

1) При вимірюванні деякої невипадкової величини a ми отримуємо наближене значення ξ. Похибка $\delta = \xi - a$, яку ми зробили, може бути подана у вигляді суми двох похибок $\delta = (\xi - M\xi) + (M\xi - a)$, де $\xi - M\xi$ - випадкова похибка, $M\xi - a$ - систематична похибка. Ефективні методи вимірювань не мають систематичної похибки, тому будемо вважати, що $M\xi = a, \quad M\delta = 0, \quad D\delta = \sigma^2$.

Для зменшення похибки роблять n незалежних вимірювань $\xi_1, \xi_2, ..., \xi_n$. За оцінку величини a приймають середнє $\hat{a} = \frac{1}{n}(\xi_1 + ... + \xi_n)$. Центральна гранична теорема дозволяє проаналізувати похибку $\hat{\delta} = \hat{a} - a$:

$$P\{|\hat{a} - a| \le \varepsilon\} = P\left\{\left|\frac{\xi_1 + ... + \xi_n}{n} - a\right| \le \varepsilon\right\} =$$

$$= P\left\{\left|\frac{\xi_1 + ... + \xi_n - na}{\sigma\sqrt{n}}\right| \le \frac{\varepsilon\sqrt{n}}{\sigma}\right\} \approx \frac{1}{\sqrt{2\pi}} \int_{-\varepsilon\sqrt{n}/\sigma}^{\varepsilon\sqrt{n}/\sigma} e^{-\frac{u^2}{2}} du \underset{n\to\infty}{\to} 1.$$

****

2) Доведемо за допомогою центральної граничної теореми співвідношення

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$

Нехай ζ_n - випадкова величина, розподілена за законом Пуассона з параметром n. Тоді $P\{\zeta_n \le n\} = e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!}$. Величину ζ_n подамо у вигляді $\xi_1 + ... + \xi_n$, де ξ_n - незалежні і мають пуассонівський розподіл з параметром $\lambda = 1$, $M\xi_1 = \lambda = 1, \quad D\xi_1 = \lambda = 1$.

Отже

$$\lim_{n\to\infty} P\{\zeta_n \le n\} = \lim_{n\to\infty} P\left\{\frac{\xi_1 + ... + \xi_n - na}{\sigma\sqrt{n}} \le 0\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{u^2}{2}} du = \frac{1}{2}.$$

****

Розширимо тепер умови, при яких можлива апроксимація розподілу сум випадкових величин нормальним розподілом.

Будемо розглядати серії довільних незалежних випадкових величин $\xi_{1,n}, \xi_{2,n}, ..., \xi_{n,n}$ і їх суми $S_n = \sum_{k=1}^{n} \xi_{k,n}$.

$$\begin{matrix} \xi_{11} & 0 & \dots & 0 \\ \xi_{12} & \xi_{22} & 0 & \dots \\ \xi_{13} & \xi_{23} & \xi_{33} & 0 \\ \dots & \dots & \dots & \dots \end{matrix}$$

Припустимо, що $\xi_{k,n}$ мають скінченні другі моменти $\sigma_{k,n}^2 = D\xi_{k,n} < \infty$ і нехай

$$M\xi_{k,n} = 0, \quad DS_n = \sum_{k=1}^{n} \sigma_{k,n}^2 = 1. \tag{1}$$

Будемо припускати також, що виконується умова Ліндеберга: для будь-якого $\varepsilon > 0$

$$\sum_{k=1}^{n} M\left(\xi_{k,n}^2, \left|\xi_{k,n}\right| > \varepsilon\right) = \sum_{k=1}^{n} \int_{|x|>\varepsilon} x^2 dF_{kn}(x) \underset{n\to\infty}{\to} 0. \tag{2}$$

Лема 1 З умови Ліндеберга випливає властивість рівномірної малості $\xi_{k,n}$, тобто для будь-якого $\varepsilon > 0$

$$\max_{1\le k\le n} P\left(\left|\xi_{k,n}\right| > \varepsilon\right) \underset{n\to\infty}{\to} 0. \tag{3}$$

Доведення. Так як

$$\max_{1\le k\le n} D\xi_{kn} = \max_{1\le k\le n} \int_{-\infty}^{+\infty} x^2 dF_{kn}(x) \le \max_{1\le k\le n} \left\{ \int_{|x|\le\varepsilon} x^2 dF_{kn}(x) + \int_{|x|>\varepsilon} x^2 dF_{kn}(x) \right\} \le$$

$$\le \varepsilon^2 + \max_{1\le k\le n} \int_{|x|>\varepsilon} x^2 dF_{kn}(x) \le \varepsilon + \sum_{k=1}^{n} \int_{|x|>\varepsilon} x^2 dF_{kn}(x),$$

то $\max\limits_{1\le k\le n} D\xi_{kn} \underset{n\to\infty}{\to} 0$. Для завершення доведення треба застосувати нерівність Чебишова $\max\limits_{1\le k\le n} P\left(\left|\xi_{k,n}\right| > \varepsilon\right) \le \dfrac{1}{\varepsilon^2} \max\limits_{1\le k\le n} D\xi_{kn} \underset{n\to\infty}{\to} 0$.

Лему доведено.

****

Застосовуючи метод характеристичних функцій, результат теореми 2 можна узагальнити на суми S_n, складені з серії незалежних випадкових величин.

Теорема 3. Якщо послідовність серій незалежних випадкових величин $\left\{\xi_{k,n}\right\}_{k=1}^{n}$, $n = 1,2...$ задовольняє умовам (1), (2), то

$$P\left\{S_n \le x\right\} \underset{n\to\infty}{\to} \Phi(x).$$

Доведення. Нехай $\varphi_{k,n}(t)$ - характеристичні функції $\xi_{k,n}$. Тоді $\varphi_{S_n}(t) = \prod\limits_{k=1}^{n} \varphi_{k,n}(t)$. Нам достатньо довести, що

$$\lim_{n\to\infty} \varphi_{S_n}(t) = \exp\left\{-t^2/2\right\}.$$

Оскільки $M\xi_{k,n} = 0$, то

$$\left|\varphi_{k,n}(t) - 1\right| = \left|M\left(e^{it\xi_{k,n}} - 1 - it\xi_{k,n}\right)\right| \le \frac{t^2}{2} M\xi_{k,n}^2 = \frac{t^2}{2} D\xi_{k,n}.$$

Звідси випливає

$$\max_{k\le n} \left|\varphi_{k,n}(t) - 1\right| \to 0 \text{ при } n \to \infty. \tag{4}$$

Неважко перевірити, що при $|\alpha| \le \dfrac{1}{2}$

$$\left|\ln(1+\alpha) - \alpha\right| \le |\alpha|^2. \tag{5}$$

З (4) маємо, що при достатньо великих n і всіх $k \le n$ виконується нерівність

$$\left|\varphi_{k,n}(t)-1\right|<\frac{1}{2},$$

що дає можливість на основі (5) при $\alpha=\varphi_{k,n}(t)-1$ отримати

$$\left|\ln\varphi_{k,n}(t)-\left(\varphi_{k,n}(t)-1\right)\right|\le\left|\varphi_{k,n}(t)-1\right|^2. \quad (6)$$

Розглянемо тепер характеристичну функцію суми S_n. В силу незалежності $\xi_{k,n}$ і нерівності (4.30)

$$\ln\varphi_{S_n}(t)=\sum_{k=1}^{n}\ln\varphi_{k,n}(t)=\sum_{k=1}^{n}\left(\varphi_{k,n}(t)-1\right)+R_n\,, \quad (7)$$

де

$$\left|R_n\right|\le\sum_{k=1}^{n}\left|\varphi_{k,n}(t)-1\right|^2\le\max_{k\le n}\left|\varphi_{k,n}(t)-1\right|\sum_{k=1}^{n}\left|\varphi_{k,n}(t)-1\right|,$$

$$\sum_{k=1}^{n}\left|\varphi_{k,n}(t)-1\right|\le\sum_{k=1}^{n}\frac{t^2}{2}D\xi_{k,n}=\frac{t^2}{2}.$$

Таким чином $R_n\to0$ і головним доданком в правій частині (7) є сума

$$\sum_{k=1}^{n}\left(\varphi_{k,n}(t)-1\right)=\sum_{k=1}^{n}M\left(e^{it\xi_{k,n}}-1-it\xi_{k,n}\right)==\int_{-\infty}^{+\infty}\frac{e^{itx}-1-itx}{x^2}\sum_{k=1}^{n}x^2dF_{\xi_{k,n}}(x)=\int_{-\infty}^{+\infty}f(t,x)dG_n(x),$$

де $f(t,x)=\dfrac{e^{itx}-1-itx}{x^2}$, $\quad G_n(x)=\sum_{k=1}^{n}M\left(\xi_{k,n}^2;\xi_{k,n}\le x\right)=\sum_{k=1}^{n}\int_{-\infty}^{x}u^2dF_{\xi_{k,n}}(u).$ Умова Ліндеберга означає, що

$$G_n(x)\underset{n\to\infty}{\overset{сл}{\Rightarrow}}E(x),$$

де $E(x)$ – функція виродженого розподілу, зосередженого в точці 0.

Використовуючи розклад e^{itx} у ряд, можна переконатись, що функція $f(t,x)$ при кожному t обмежена і неперервна, $f(t,0)=-\dfrac{t^2}{2}$. Тому на основі другої теореми Хеллі маємо

$$\int_{-\infty}^{+\infty}f(t,x)dG_n(x)\underset{n\to\infty}{\longrightarrow}\int_{-\infty}^{+\infty}f(t,x)dE(x)=f(t,0)=-\frac{t^2}{2}.$$

Таким чином $\ln\varphi_{S_n}(t)=-\dfrac{t^2}{2}+o(1)$, $\quad\varphi_{S_n}(t)\to e^{-t^2/2}$ при $n\to\infty$.

Теорему доведено.

****

Умова Ліндеберга (2) не є необхідною для збіжності розподілу S_n до нормального. В зв'язку з цим розглянемо приклад:
$\xi_{1,n}=\eta,\quad\xi_{2,n}=...=\xi_{n,n}=0$, де η – випадкова величина з стандартним нормальним розподілом. Очевидно, умова (1) виконується, $P\{S_n\le x\}=\Phi(x)$, але доданки $\xi_{k,n}$ не є рівномірно малими, а значить (2) не виконується.

Однак, якщо разом із збіжністю $P(S_n\le x)\to\Phi(x)$ вимагати, щоб $\xi_{k,n}$ були рівномірно малими, то умова Ліндеберга стає необхідною.

****

Теорема 4. Якщо послідовність серій незалежних випадкових величин $\left\{\xi_{k,n}\right\}_{k=1}^{n}$, $\quad n=1,2...$ задовольняє умовам (1), (3) і

$$P\{S_n\le x\}\underset{n\to\infty}{\to}\Phi(x),$$

то виконується умова Ліндеберга.

****

Розглянемо послідовність незалежних, але не обов'язково однаково розподілених випадкових величин $\xi_1, \xi_2, ..., \xi_n, ...$

$$M\xi_k = a_k, \quad D\xi_k = \sigma_k^2, \quad \sigma^2(n) = \sum_{k=1}^{n} \sigma_k^2,$$

$$S_n = \frac{\xi_1 + ... + \xi_n - \sum_{k=1}^{n} a_k}{\sigma(n)} = \sum_{k=1}^{n} \xi_{k,n}, \text{ де } \xi_{k,n} = \frac{\xi_k - a_k}{\sigma(n)}.$$

Очевидно, що $M\xi_{k,n} = 0, \quad \sum_{k=1}^{n} D\xi_{k,n} = 1$. Умова Ліндеберга приймає вид

$$\sum_{k=1}^{n} M\left(\xi_{k,n}^2, |\xi_{k,n}| > \varepsilon\right) = \frac{1}{\sigma^2(n)} \sum_{k=1}^{n} M\left\{(\xi_k - a_k)^2; |\xi_k - a_k| > \varepsilon\sigma(n)\right\} =$$

$$= \frac{1}{\sigma^2(n)} \sum_{k=1}^{n} \int_{|x-a_k|>\varepsilon\sigma(n)} (x - a_k)^2 \, dF_{\xi_k}(x) \underset{n\to\infty}{\to} 0. \quad (8)$$

Наслідок 1. Якщо випадкові величини $\xi_1, \xi_2, ..., \xi_n, ...$ – незалежні, мають математичні сподівання і дисперсії, для них виконується умова Ліндеберга (8), то

$$P\{S_n \le x\} \underset{n\to\infty}{\to} \Phi(x).$$

Останній результат, а також теорема 4 представляють собою варіанти центральної граничної теореми, які мають широке застосування при аналізі похибок вимірювань, в актуарній математиці та інших областях.