6 ГРАНИЧНІ ТЕОРЕМИ ТЕОРІЇ ЙМОВІРНОСТІ

Розглянемо ряд тверджень і теорем з великої групи так званих граничних теорем теорії ймовірностей, встановивши зв'язок між теоретичними та експериментальними характеристиками випадкових величин з великою кількістю випробувань на них. Вони складають основу математичної статистики. Граничні теореми умовно діляться на дві групи. Перша група теорем, так званий *законом великих чисел* (скорочено ЗВЧ), встановлює стійкість середніх значень: при великій кількості випробувань їх середній результат перестає бути випадковим і може бути передбаченим з достатньою точністю. Друга група теорем, так звана *центральна гранична теорема* (скорочено ЦГТ), встановлює умови, при яких закон розподілу суми великої кількості випадкових величин нескінченно близький до нормального.

По-перше, розглянемо нерівність Чебишева, за допомогою якого можна: а) приблизно оцінити ймовірності подій, пов'язаних з в.в., розподіл яких невідомий; б) доведення ряду теорем ЗВЧ.

6.1 Нерівність Чебишева

Теорема 6.1 Якщо в.в. X має м.о. $MX = a$ і дисперсію DX, то для будь-якого $\varepsilon > 0$ справедлива *нерівність Чебишева*

$$P\{|X - MX| \geq \varepsilon\} \leq \frac{DX}{\varepsilon^2}. \tag{6.1}$$

Доведення. Доведемо нерівність (6.1) для неперервної в.в. X з щільністю $f(x)$. Імовірність $P\{|X - a| \geq \varepsilon\}$ – це ймовірність попадання в.в. X до області поза інтервалом $[a - \varepsilon; a + \varepsilon]$. Можна записати

$$P\{|X - a| \geq \varepsilon\} = \int_{-\infty}^{a-\varepsilon} f(x)dx + \int_{a+\varepsilon}^{+\infty} f(x)dx = \int_{|x-a|\geq\varepsilon} f(x)dx =$$

$$= \int_{|x-a|\geq\varepsilon} 1 \cdot f(x)dx \leq \int_{|x-a|\geq\varepsilon} \frac{(x-a)^2}{\varepsilon^2} \cdot f(x)dx,$$

оскільки область інтегрування $|x - a| \geq \varepsilon$ можна записати як $(x - a)^2 \geq \varepsilon^2$, звідки випливає $1 \leq \frac{(x-a)^2}{\varepsilon^2}$. Тоді

$$P\{|X - a| \geq \varepsilon\} \leq \frac{1}{\varepsilon^2} \int_{|x-a|\geq\varepsilon} (x - a)^2 f(x)dx \leq \frac{1}{\varepsilon^2} \int_{-\infty}^{+\infty} (x - a)^2 f(x)dx,$$

оскільки інтеграл від невід'ємної функції може тільки зростати, коли область інтегрування розширюється. Таким чином,

$$P\{|X - a| \geq \varepsilon\} \leq \frac{1}{\varepsilon^2} \int_{|x-a|\geq\varepsilon} (x - a)^2 f(x)dx = \frac{1}{\varepsilon^2} DX,$$

тобто .$P\{|X - a| \geq \varepsilon\} \leq \frac{1}{\varepsilon^2} DX.$

Аналогічно доводиться нерівність Чебишева для дискретних в.в. X, яка приймає значення $x_1, x_2, x_3, \ldots$ з імовірностями $p_1, p_2, p_3, \ldots$; тільки інтеграли (виду $\int_{|x-a|\geq\varepsilon}$) замінюються відповідними сумами (виду $\sum_{|x_i-a|\geq\varepsilon}$).

Відзначимо, що нерівність Чебишева можна записати і в іншій формі:

$$P\{|X - MX| < \varepsilon\} \geq 1 - \frac{DX}{\varepsilon^2}. \qquad (6.2)$$

За формою (6.2) вона встановлює нижню границю ймовірності події, а за формою (6.1) встановлює верхню границю. Нерівність Чебишева *справедлива для будь-яких в.в.* Зокрема, для в.в. $X = m$, яка має *біноміальний розподіл* з математичним очікуванням $MX = a = np$ і дисперсією $DX = npq$, вона набуває вигляду

$$P\{|m - np| < \varepsilon\} \geq 1 - \frac{npq}{\varepsilon^2}; \qquad (6.3)$$

для *частоти* $\frac{m}{n}$ (або $\frac{n_A}{n}$) події в n *незалежних* випробуваннях, в кожному з яких вона може відбуватися з імовірністю $p = M\left(\frac{m}{n}\right) = a$, дисперсія якої $D\left(\frac{m}{n}\right) = \frac{pq}{n}$, нерівність Чебишева має вигляд

$$P\left\{\left|\frac{m}{n} - p\right| < \varepsilon\right\} \geq 1 - \frac{pq}{n\varepsilon^2}. \qquad (6.4)$$

Оцінка ймовірності попадання в в.в. X в інтервал $[\varepsilon; \infty)$ дає нерівність Маркова.

Теорема 6.2 *(нерівність Маркова)* Для будь-якої невід'ємної в.в. X, яка має м.о. MX і $\varepsilon > 0$, справедлива нерівність:

$$P\{X \geq \varepsilon\} \leq \frac{MX}{\varepsilon}. \qquad (6.5)$$

Доведення.

$$P\{X \geq \varepsilon\} = \int_\varepsilon^\infty f(x)dx \leq \int_\varepsilon^\infty \frac{x}{\varepsilon} f(x)dx = \frac{1}{\varepsilon}\int_\varepsilon^\infty xf(x)dx \leq \frac{1}{\varepsilon}\int_0^\infty xf(x)dx = \frac{MX}{\varepsilon}.$$

Нерівність (6,5) можна записати у вигляді

$$P\{X < \varepsilon\} \geq 1 - \frac{MX}{\varepsilon}. \qquad (6.6)$$

Приклад 1 Оцінити за допомогою нерівності Чебишева ймовірність того, що відхилення в.в. X від його м.о. буде менше трьох с.к.в., тобто менше $3\sigma_X$.

Розв'язання. Поклавши $\varepsilon = 3\sigma_X$ в формулі (6.2), отримаємо

$$P\{|X - MX| < 3\sigma_X\} \geq 1 - \frac{DX}{(3\sigma_X)^2} = 1 - \frac{1}{9} = \frac{8}{9} \approx 0{,}8889.$$

Ця оцінка, як відомо, називається *правилом трьох сигм*; для в.в. $X \sim N(a, \sigma^2)$ ця ймовірність дорівнює 0,9973.

6.2 Теорема Чебишева

Основне твердження ZBH міститься в теоремі Чебишева. Вона та інші теореми ЗБЧ використовують поняття «збіжність випадкових величин за ймовірністю».

Випадкові величини$X_1, X_2, \ldots, X_n, \ldots$ *сходяться за ймовірністю* до значення A (випадкові або невипадкові), якщо для будь-якої ймовірності подія при прагненні до одиниці, т. Е.$\varepsilon > 0\{|X_n - A| < \varepsilon\}n \to \infty$

$$\lim_{n\to\infty} P\{|X_n - A| < \varepsilon\} = 1$$

(або). Збіжність в ймовірності символічно записується наступним чином:$P\{|X_n - A| < \varepsilon\} \to 1$

$$\overset{n\to\infty}{} \quad X_n \overset{P}{\to} A.$$

Слід зазначити, що *збіжність ймовірностей* вимагає виконання нерівності$|X_n - A| < \varepsilon$ *для переважної більшості умов послідовності* (в математичному аналізі для всіх , де $n > NN$ знаходиться деяке число), і коли практично всі доданки послідовності повинні потрапляти в$n \to \infty \varepsilon$ *сусідство* A.

Теорема *(ЗБЧ у вигляді П.Л. Чебишева, 1886).* Якщо випадкові величини незалежні і існує таке число, що $X_1, X_2, \ldots, X_n, \ldots C > 0 DX_i \leq C, i = 1,2, \ldots$, то для будь-яких $\varepsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} MX_i\right| < \varepsilon\right\} = 1, \tag{6.7}$$

тобто середнє арифметичне цих випадкових величин сходиться за ймовірністю до середнього арифметичного їх m. про.:

$$\frac{1}{n}\sum_{i=1}^{n} X_i \overset{n\to\infty}{\underset{P}{\to}} \frac{1}{n}\sum_{i=1}^{n} MX_i.$$

Доведення. З $DX_i \leq C, i = 1,2, \ldots$, то

$$D\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} D\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} DX_i \leq \frac{1}{n^2}\sum_{i=1}^{n} C = \frac{1}{n^2} Cn = \frac{C}{n}.$$

Потім, звертаючись до с.в. Нерівність Чебишева ($\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$6,2), маємо

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} MX_i\right| < \varepsilon\right\} \geq 1 - \frac{D\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)}{\varepsilon^2} \geq 1 - \frac{C}{n\varepsilon^2}. \tag{6.8}$$

Переходячи до межі при і враховуючи, що ймовірність будь-якої події не перевищує 1, отримуємо$n \to \infty$

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} MX_i\right| < \varepsilon\right\} = 1.$$

Наслідок. Якщо бл. В. є незалежними і однаково розподіленими, , , то для будь-яких $X_1, X_2, \ldots, X_n, \ldots MX_i = a DX_i = \sigma^2 \varepsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - a\right| < \varepsilon\right\} = 1, \tag{6.9}$$

тобто середнє арифметичне с. сходиться за ймовірністю до очікуваного значення а:

$$\frac{1}{n}\sum_{i=1}^{n} X_i \overset{n\to\infty}{\underset{P}{\to}} a.$$

Доведення. З

$$\frac{1}{n}\sum_{i=1}^{n} MX_i = \frac{1}{n}\sum_{i=1}^{n} a = \frac{1}{n} \cdot na = a,$$

і розпорошення с.в. дорівнює числу , тобто обмежений, то, застосовуючи ЗБЧ у вигляді Чебишева ($X_i \sigma^2$6,7), отримуємо твердження (6,9).

Наслідок (6,9) теореми Чебишева обґрунтовує «принцип її носіїварифметики s. В. X_i»,, постійно використовується на практиці. Таким чином, нехай існує *n* незалежних вимірювань деякої величини, істинною величиною якої є *a* (невідомо). Результат кожного вимірювання – с. в. Згідно з дослідженням (6.9), середнє арифметичне результатів вимірювань можна прийняти за приблизне значення величини а:

$$a \approx \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}.$$

Рівність тим точніше, *чим більше р.*

Теорема Чебишева також заснована на широко використовуваному в статистиці *методі вибірки*, суть якого полягає в тому, що про якість великої кількості однорідного матеріалу можна судити по його малій вибірці.

Теорема Чебишева підтверджує зв'язок між випадковістю і необхідністю: середнє значення *випадкової величини*$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ практично не відрізняється від *невипадкової величини*$\frac{1}{n}\sum_{i=1}^{n} MX_i$.

Приклад 2 Глибина моря вимірюється приладом, який не має зміщення. Середнє квадратичне відхилення вимірювань не перевищує 15 м. Скільки самостійних вимірювань потрібно зробити так, щоб з імовірністю не менше 0,9 можна стверджувати, що середнє арифметичне цих вимірювань відрізняється від (глибини моря) за модулем менше 5 м?

Рішення. Позначимо через результати $X_i n$ незалежних вимірювань глибини моря. Нам потрібно знайти число p, яке задовольняє нерівності (6,8):

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} MX_i\right| < \varepsilon\right\} \geq 1 - \frac{C}{n\varepsilon^2},$$

де ,$MX_i = a$ *що означає, що в вимірах немає зміщення* (тобто вимірювання проводяться з однаковою точністю). За умови , , так як м.$\varepsilon = 5C = 225\sigma = \sqrt{DX} = 15$

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - a\right| < 5\right\} \geq 1 - \frac{225}{25n} \geq 0{,}9,$$

тобто $0{,}1 \geq \frac{9}{n}$, $n \geq 90$. Вимірювання повинно проводитися не менше 90 разів.

6.3 Теорема Бернуллі

Теорема Бернуллі історично є першою і найпростішою формою закону великих чисел. Він теоретично обґрунтовує стійку властивість відносної частоти.

Теорема *(ЗБЧ у вигляді Дж. Бернуллі, 1713)* Якщо ймовірність настання події A в одному тесті дорівнює p, то число настання цієї події в незалежних тестах дорівнює , то для будь-якого числа існує рівність$n_A \varepsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1, \qquad (6.10)$$

тобто відносна частота події$P^*(A)$ A сходиться за ймовірністю до ймовірності р події A: $P^*(A) \overset{n\to\infty}{\underset{P}{\to}} P(A)$.

Доведення. Давайте увійдемо з. В. наступним чином: , якщо подія А з'явилася в -му тесті $X_1, X_2, ..., X_n X_i = 1i$, а якщо воно не з'явилося, то число (кількість успіхів) можна представити у вигляді$X_i = 0\, n_A\, n_A = \sum_{i=1}^{n} X_i$. М.О. і дисперсія с. В. Рівний:X_i

$MX_i = 1 \cdot p + 0 \cdot (1-p) = p, DX_i = 0^2 \cdot (1-p) + 1^2 \cdot p - p^2 = p(1-p) = pq.$

Закон розподілу с.в. має виглядX_i

X_i	0	1
p	$1-p$	p

з будь-яким . Таким чином, с.в. незалежні, їх дисперсії обмежені однаковою кількістю $i \ X_i \frac{1}{4}$, оскільки

$$p(1-p) = p - p^2 = \frac{1}{4} - \left(p - \frac{1}{2}\right)^2 \le \frac{1}{4}.$$

Тому теорему Чебишева (6.7) можна застосувати до таких тематичних досліджень:

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} MX_i\right| < \varepsilon\right\} = 1.$$

Але, $\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{n_A}{n}$ і $\frac{1}{n}\sum_{i=1}^{n} MX_i = \frac{1}{n}\cdot np = p.$ Тому

$$\lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1.$$

Теорема Б'нуллі теоретично обґрунтовує можливість приблизного розрахунку ймовірності події за допомогою її відносної частоти. Так, наприклад, для ймовірності народження дівчинки можна взяти відносну частоту цієї події, яка, за статистичними даними, приблизно дорівнює 0, 485.

Нерівність Чебишева (6,2) для випадкових величин набуває вигляду$n_A = \frac{1}{n}\sum_{i=1}^{n} X_i$

$$P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} \ge 1 - \frac{pq}{n\varepsilon^2}. \qquad (6.11)$$

Узагальненням теореми Б'нуллі до випадку, коли ймовірність появи події Ap_i в кожному з *n* тестів різна, є *теоремою Пуассона*:

$$P^*(A) \overset{n\to\infty}{\underset{P}{\to}} \frac{1}{n}\sum_{i=1}^{n} p_i, \qquad (6.12)$$

де p_i – ймовірність події *A* в -му тесті.i

Приклад 3 Імовірність наявностіпомилки на одній сторінціпочеркознавчої руки дорівнює 0,2. Оцініть ймовірність того, що в рукописі, що містить 400 сторінок, частота появи помилки відрізняється від відповідної ймовірності за модулем менше 0,05.

Рішення: Скористаємося формулою (6.11). У цьому випадку $p = 0{,}2,$, $q = 0{,}8$ $n = 400,$. $\varepsilon = 0{,}05$

$$P\left\{\left|\frac{n_A}{n} - 0{,}2\right| < 0{,}05\right\} \ge 1 - \frac{0{,}2\cdot 0{,}8}{400\cdot 0{,}05^2} = 0{,}84,$$

тобто $p \ge 0{,}84$.

6.4 Центральна гранична теорема

Центральна гранична теорема (КПТ) – це друга група граничних теорем, які встановлюють зв'язок між законом розподілу суми с. в. і його кінцева форма, нормальний закон розподілу.

Сформулюємо СРТ для того випадку, коли члени суми мають однаковий розподіл (саме ця теорема найчастіше використовується на практиці, так як в математичній статистиці селективні випадкові величини мають однакові розподіли, так як отримані з однієї і тієї ж загальної сукупності).

Теорема. Нехай с. В. незалежні, порівну розподілені, мають скінченне очікуване очікування і дисперсію , . Тоді функція розподілу центрованої і нормованої суми цих випадкових величин прагне до функції розподілу стандартної нормальної випадкової величини:$X_1, X_2, \dots, X_n MX_i = aDX_i = \sigma^2 i = \overline{1,n} n \to \infty$

$$Z_n = \frac{\sum_{i=1}^n X_i - M(\sum_{i=1}^n X_i)}{\sqrt{D(\sum_{i=1}^n X_i)}} = \frac{\sum_{i=1}^n X_i - a}{\sigma\sqrt{n}},$$

$$F_{Z_n}(x) = P\{Z_n < x\} \xrightarrow[n\to\infty]{} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt. \qquad (6.13)$$

З відношення (6,13) випливає, що при досить великому n сума приблизно розподіляється за нормальним законом: $Z_n Z_n \sim N(0,1)$. Це означає, що сума приблизно розподіляється за нормальним законом: $S_n = \sum_{i=1}^{n} X_i \, S_n \sim N(na, \sigma\sqrt{n})$.

Кажуть, що з$n \to \infty$. В. асимптотично нормальний.$\sum_{i=1}^n X_i$

Нагадаємо, що:

1. CV X називається центрованим і нормалізованим (тобто стандартним), якщо , $MX = 0$a$DX = 1$.
2. Якщо с.в. X_i,$i = \overline{1,n}$ незалежний, $MX_i = a$,$DX_i = \sigma^2$ то
$M(\sum_{i=1}^n X_i) = \sum_{i=1}^n MX_i = na$, $D(\sum_{i=1}^n X_i) = \sum_{i=1}^n DX_i = n\sigma^2$.
3. $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$ – Функція Лапласа , де $\Phi(x) = \frac{1}{2} + \Phi_0(x)$

$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{t^2}{2}} dt$ – нормалізована функція Лапласа.

Формула (6.13) дозволяє розрахувати ймовірності різних подій, пов'язаних з сумами випадкових величин при великих величинах. в.$S_n = \sum_{i=1}^n X_i$ до стандартного с. в., отримуємо:

$$P\left\{\alpha \le \sum_{i=1}^{n} X_i \le \beta\right\} = P\left\{\frac{\alpha - na}{\sigma\sqrt{n}} \le \frac{\sum_{i=1}^n X_i - na}{\sigma\sqrt{n}} \le \frac{\beta - na}{\sigma\sqrt{n}}\right\}$$
$$\approx \Phi\left(\frac{\beta - na}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{\alpha - na}{\sigma\sqrt{n}}\right)$$

або

$$P\{\alpha \le S_n \le \beta\} \approx \Phi\left(\frac{\beta - MS_n}{\sqrt{DS_n}}\right) - \Phi\left(\frac{\alpha - MS_n}{\sqrt{DS_n}}\right) \qquad (6.14)$$

є формулою для визначення ймовірностіі що сума декількох с. с. буде в зазначених межах. Часто CPT використовується, якщо .$n > 10$

Приклад 4 Незалежні в. В. розподілені рівномірно по відрізку [0, 1X_i]. Знайти закон розподілу в. в.$Y = \sum_{i=1}^{100} X_i$, а також ймовірність того, що .$55 < Y < 70$

Рішення. УмовиКПТ дотримуються, тому в. В. має приблизну щільність розподілу,Y

$$f_Y(y) \approx \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{(y-m_y)^2}{2\sigma_y^2}}.$$

За відомими формулами для м.о. і дисперсії в разі рівномірного розподілу знаходимо:$MX_i = \frac{0+1}{2} = \frac{1}{2}$, $DX_i = \frac{(1-0)^2}{12} = \frac{1}{12}$, $\sigma X_i = \frac{1}{\sqrt{12}} = \frac{1}{2\sqrt{3}}$. Тоді

$$m_y = M\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} MX_i = 100 \cdot \frac{1}{2} = 50,$$

$$\sigma_y^2 = D\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} DX_i = 100 \cdot \frac{1}{12} = \frac{25}{3}, \qquad \sigma_y = \frac{5\sqrt{3}}{3}.$$

Тому $f_Y(y) \approx \frac{3}{5\sqrt{6\pi}} e^{-\frac{3(y-50)^2}{50}}$. Використовуючи формулу (6.14), знаходимо

$$P\{55 < Y < 70\} \approx \Phi\left(\frac{70-50}{\frac{5\sqrt{3}}{3}}\right) - \Phi\left(\frac{55-50}{\frac{5\sqrt{3}}{3}}\right) = \Phi\left(4\sqrt{3}\right) - \Phi\left(\sqrt{3}\right) \approx 0{,}04,$$

т. е. $P\{55 < Y < 70\} \approx 0{,}04$.

6.5 Інтегральна теорема Муавра—Лапласа

Наслідки КПТ є локальними та невід'ємними теоремами Муавра-Лапласа, розглянутими раніше. Ось виведення інтегральної теореми Муавра-Лапласа.

Розглянемо схему Бернуллі. Нехай n_A — число настання події A (кількість успіхів) в n незалежних тестах, в кожному з яких вона може з'явитися з імовірністю p (не з'явитися – з імовірністю). $q = 1 - p\, n_A$ В. такий, що , якщо в -му випробуванні подія $X_1, X_2, ..., X_n X_i = 1 i A$ сталася інакше$X_i = 0$, т. Е. е.$n_A = \sum_{i=1}^{n} X_i$ . . Оскільки$MX_i = p$, , $DX_i = pq$то

$$Mn_A = M(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} MX_i = np, \qquad Dn_A = D(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} DX_i = npq.$$

Потім с.в.

$$Z_n = \frac{\sum_{i=1}^{n} X_i - M(\sum_{i=1}^{n} X_i)}{\sqrt{D\left(\sum_{i=1}^{n} X_i\right)}} = \frac{n_A - np}{\sqrt{npq}}$$

також являє собою суму n незалежних, порівну розподілених випадкових величин. У цьому випадку дійсно:$Z_n \sim N(0,1)$

$$MZ_n = M\left(\frac{n_A - np}{\sqrt{npq}}\right) = \frac{np - np}{\sqrt{npq}} = 0,$$

$$DZ_n = D\left(\frac{n_A - np}{\sqrt{npq}}\right) = \frac{1}{npq} D_{n_A} = \frac{npq}{npq} = 1.$$

Отже, на підставі КПП (6.13) с.в. при великому числі $Z_n n$ має приблизно нормальний розподіл. За властивістю нормального закону пишемо

$$P\{z_1 \le Z_n \le z_2\} \approx \Phi(z_2) - \Phi(z_1)$$

($\Phi(z)$ – функція Лапласа). $z_1 = \frac{k_1 - np}{\sqrt{npq}}$, $z_2 = \frac{k_2 - np}{\sqrt{npq}}$ подвійна нерівність

$$\frac{k_1 - np}{\sqrt{npq}} \le \frac{n_A - np}{\sqrt{npq}} \le \frac{k_2 - np}{\sqrt{npq}}$$

можна переписати в еквівалентній формі$k_1 \le n_A \le k_2$. Таким чином, отримуємо

$$P\{k_1 \le n_A \le k_2\} \approx \Phi(z_2) - \Phi(z_1),$$

тобто інтегральна формула МиАвра-Лаплас.

Зауважимо, щоінтегральна теорема Пр-Еделя Муавра–Лапласа

$$P\left\{\frac{n_A - np}{\sqrt{npq}} < x\right\} \to \Phi(x) \qquad (6.15)$$

для схеми Бернуллі вонавипливає безпосередньо з КПТ (6.13) з урахуванням результатів, отриманих раніше (, $n_A = \sum_{i=1}^{n} X_i$, $MX_i = p$,$DX_i = pq$ то ,).$na = np\sigma\sqrt{n} = pq\sqrt{n} = \sqrt{npq}$

Для обчислення сум біноміальних ймовірностей можна скористатися приблизною формулою

$$\sum_{k=0}^{m} P_{n,k} \approx \Phi\left(\frac{m-np}{\sqrt{npq}}\right). \qquad (6.16)$$

Справді

$$\sum_{k=0}^{m} P_{n,k} = P\{n_A \le m\} = P\{-\infty < n_A \le m\} =$$

$$= P\left\{-\infty < \frac{n_A - n}{\sqrt{npq}} \le \frac{m-np}{\sqrt{npq}}\right\} = \Phi\left(\frac{m-np}{\sqrt{npq}}\right) - \Phi(-\infty) = \Phi\left(\frac{m-np}{\sqrt{npq}}\right).$$

де $\Phi(x)$ — функція Лапласа.

Приклад 5 Друкарці необхідно набрати текст, який містить 8000 слів з чотирьох і більше букв. Імовірність допущення помилки в будь-якому з цих слів дорівнює 0,01. Яка ймовірність того, що при наборі буде допущено не більше 90 помилок?

Рішення. Застосуємо формулу (6.16). Оскільки , , , ,$n = 8000p = 0{,}01q = 0{,}99m = 90$

$$\frac{1}{\sqrt{npq}} = \frac{1}{\sqrt{8000 \cdot 0{,}01 \cdot 0{,}99}} = \frac{1}{\sqrt{79{,}2}} \approx 0{,}112,$$

$$\frac{m-np}{\sqrt{npq}} \approx 10 \cdot 0{,}112 \approx 1{,}12,$$

$\Phi(1{,}12) = 0{,}8686$. Тому.$P\{n_A \le 90\} = \sum_{k=0}^{90} P_{n,k} \approx 0{,}8686$