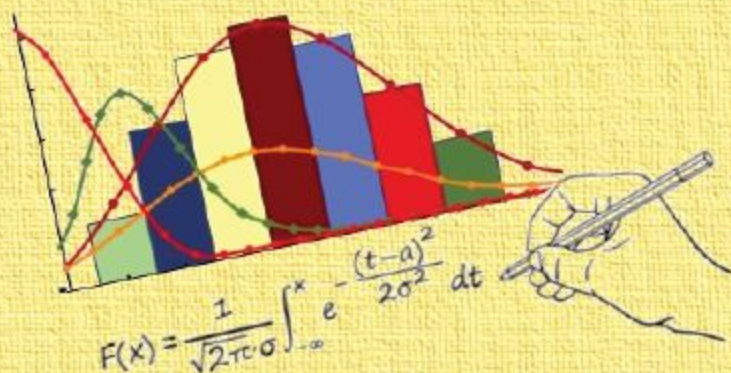


О. В. Дьячкова

ТЕОРІЯ ЙМОВІРНОСТЕЙ МАТЕМАТИЧНА СТАТИСТИКА

Опорний конспект



МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
імені В. Н. КАРАЗІНА

О. В. Дьячкова

ТЕОРІЯ ЙМОВІРНОСТЕЙ І МАТЕМАТИЧНА СТАТИСТИКА

Опорний конспект

Харків – 2019

УДК 519.2 (075.8)

Д 93

Рецензенти:

Є. В. Свищова – канд. фіз.-мат. наук, доцент кафедри інформаційних технологій і математики ХГУ «НУА»;

А. Ю. Петрова – канд. фіз.-мат. наук, доцент кафедри економічної кібернетики та прикладної економіки ХНУ імені В. Н. Каразіна.

*Затверджено до друку рішенням Науково-методичної ради
Харківського національного університету імені В. Н. Каразіна
(протокол № 6 від 29 травня 2019 року)*

Дьячкова О. В.

Д 93

Теорія ймовірностей і математична статистика : опорний конспект /
О. В. Дьячкова. — Х. : ХНУ імені В. Н. Каразіна, 2019. — 92 с.

Видання призначено для студентів усіх форм навчання, які вивчають базовий курс теорії ймовірностей і математичної статистики.

Розглянуто основні поняття та найважливіші відомості цього курсу. Зокрема, висвітлено події та операції з ними, імовірності та операції з ними, випадкові величини та їхні розподіли, граничні теореми. Також розглянуто роботу з вибірками даних, основу оцінювання параметрів, перевірки гіпотез та елементи кореляційного і регресійного аналізу.

Це дозволяє використовувати видання як довідник студентам, аспірантам, викладачам, науковим співробітникам та всім бажаючим, які опановують імовірнісні та статистичні методи.

УДК 519.2 (075.8)

© Харківський національний університет
імені В. Н. Каразіна, 2019

© О. В. Дьячкова, 2019

ЗМІСТ

ТЕОРІЯ ЙМОВІРНОСТЕЙ	5
1. Випадкові події.....	5
1.1. Різновиди подій.....	5
1.2. Алгебра подій.....	6
2. Ймовірність події.....	8
2.1. Статистична ймовірність	8
2.2. Класичне визначення ймовірності	8
2.3. Властивості ймовірності	9
2.4. Геометрична ймовірність	9
2.5. Аксиоматичне визначення ймовірності	10
3. Додавання і множення ймовірностей	11
3.1. Умовна ймовірність	11
3.2. Теорема множення ймовірностей	12
3.3. Ймовірність появи хоча б однієї події	13
3.4. Теорема додавання ймовірностей сумісних подій	14
3.5. Формула повної ймовірності	14
3.6. Формула Бейєса	14
4. Повторні випробування.....	15
4.1. Формула Бернуллі	15
4.2. Асимптотичні наближення	16
5. Випадкова величина	18
5.1. Поняття випадкової величини	18
5.2. Властивості функції розподілу ймовірностей	19
5.3. Класифікація випадкових величин	20
5.4. Система декількох випадкових величин	22
5.5. Операції над дискретними випадковими величинами	25
5.6. Числові характеристики дискретних випадкових величин.....	28
6. Основні закони розподілу дискретних випадкових величин	32
6.1. Біноміальний розподіл	32
6.2. Розподіл Пуассона	33
6.3. Дискретний рівномірний розподіл	34
7. Неперервні випадкові величини.....	34
7.1. Функція розподілу неперервної випадкової величини	34
7.2. Щільність розподілу неперервної випадкової величини	35
7.3. Числові характеристики неперервних випадкових величин	36

8. Основні закони розподілу неперервних випадкових величин.....	38
8.1. Рівномірний розподіл	38
8.2. Показовий (експоненціальний) розподіл.....	39
8.3. Нормальний розподіл.....	40
9. Граничні теореми теорії ймовірностей	45
9.1. Закон великих чисел	45
9.2. Нерівності Маркова і Чебишова	45
9.3. Теореми закону великих чисел.....	46
9.4. Центральна гранична теорема.....	49
9.5. Інтегральна теорема Муавра – Лапласа.....	51
МАТЕМАТИЧНА СТАТИСТИКА	52
10. Основи статистичних методів.....	52
10.1. Генеральна і вибіркова сукупності	52
10.2. Варіаційні ряди	53
10.3. Емпірична функція розподілу	55
10.4. Операція «рису».....	56
10.5. Числові характеристики вибірки.....	57
11. Оцінки параметрів розподілу	58
11.1. Поняття про оцінку параметрів. Види оцінок.....	58
11.2. Точкові оцінки параметрів.....	59
11.3. Інтервальні оцінки параметрів.....	64
12. Взаємозв'язки між випадковими величинами	67
12.1. Основи теорії кореляції.....	68
12.2. Елементи кореляційного аналізу.....	71
12.3. Елементи регресійного аналізу	72
13. Основи перевірки статистичних гіпотез.....	75
13.1. Статистичні гіпотези.....	75
13.2. Критерії узгодження	81
Додатки	85
Список літератури	88
Предметний покажчик.....	89

ТЕОРІЯ ЙМОВІРНОСТЕЙ

♣ Явища, наслідок яких заздалегідь не можна передбачити точно, називають **випадковими**.

♣ **Детерміновані** явища – ті, що можна точно передбачити, не випадкові.

♣ **Теорія ймовірностей** – розділ математики, що вивчає спільні закономірності, властиві масовим випадковим явищам.

1. Випадкові події

♣ **Випробування** (дослід, експеримент) – здійснення деякого комплексу умов, при якому може відбутися явище (випадкове або детерміноване), що вивчається.

♣ **Подія** – усякий наслідок (результат) випробування.

1.1. Різновиди подій

♣ **Достовірна (вірогідна) подія** – та, яка обов'язково відбудеться при певній сукупності умов.

♣ **Неможлива подія** $\emptyset$ – та, яка завідомо не відбудеться при будь-якій сукупності умов.

♣ **Випадкова подія** – та, яка при здійсненні сукупності умов може або відбутися, або не відбутися, причому до проведення випробування невідомо, чи відбудеться вона, чи ні.

Випадкові події позначають прописними літерами латинського алфавіту $A, B, C, \dots$



♣ Події називають **рівноможливими**, якщо немає об'єктивних причин вважати одну подію більш можливою, ніж іншу.

♣ **Елементарні події** $\omega_1, \omega_2, \dots$ – усі рівноможливі наслідки випробування, взаємовиключні один до іншого, у випадку коли наслідків скінченна або зліченна множина.

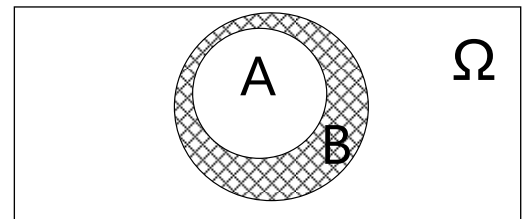
♣ **Простір (або множина) елементарних подій** Ω – сукупність усіх елементарних подій: $\Omega = \{\omega_1, \dots, \omega_n\}$.

Подія – це будь-яка сукупність елементарних наслідків.

1.2. Алгебра подій

Зобразимо множину всіляких подій Ω як прямокутник, а події – колами в ньому.

♣ $A \subset B$: подія A **приводить до** події B , якщо з настання події A випливає настання події B .

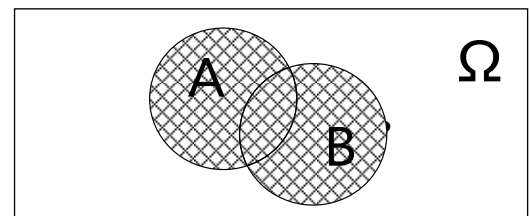


$$A \subset B$$

♣ $A \subseteq B$: поява A приводить до появи події B , але не навпаки.

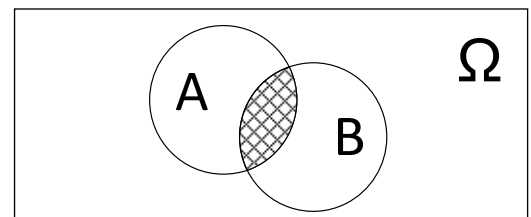
♣ $A = B$: події A і B **рівносильні (еквівалентні)**, якщо $A \subset B$ і $B \subset A$.

♣ $A + B$: **сума** подій A і B – подія C , що полягає в появі або події A , або події B , або обох цих подій.



$$C = A + B$$

♣ $A \cdot B$: **добуток** подій A і B – подія C , що полягає в сумісній появі цих подій.



$$C = A \cdot B$$

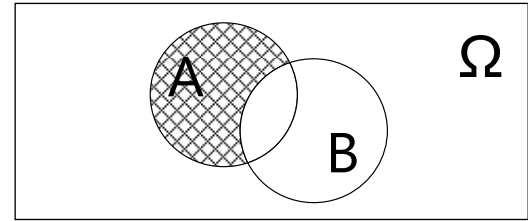
Для суми і добутку подій виконується:

$$\begin{array}{lll} A + A = A & A + \Omega = \Omega & A + \emptyset = A \\ A \cdot A = A & A \cdot \Omega = A & A \cdot \emptyset = \emptyset \end{array}$$

$\sum_{k=1}^n A_k$ – сума будь-якого скінченного числа подій A_k .

$\prod_{k=1}^n A_k$ – добуток будь-якого скінченного числа подій A_k .

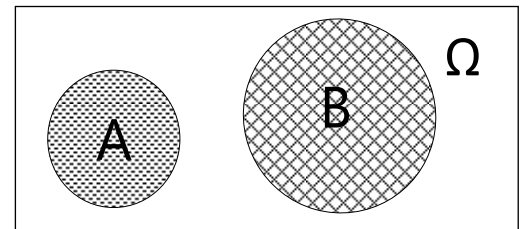
♣ $A \setminus B$: **різниця** подій A і B – подія C , яка відбувається тоді і лише тоді, коли відбувається A і не відбувається B .



$$C = A \setminus B$$

♣ **Складні** події – ті, що отримані внаслідок застосування скінченного або зліченного числа операцій додавання, множення або різниці подій.

♣ $A \cdot B = \emptyset$: події A і B **несумісні**, якщо поява однієї з них виключає появу іншої в одному і тому ж випробуванні. Інакше події **сумісні**.



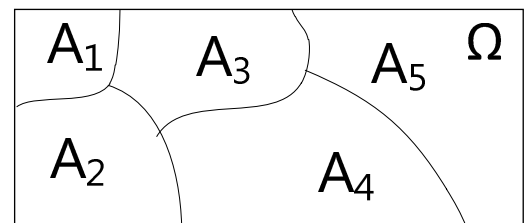
$$A \cdot B = \emptyset$$

♣ Події $A_1, A_2, \dots, A_n$ **попарно несумісні**, якщо будь-які дві з них несумісні: $A_k \cdot A_j = \emptyset$ (при усіх $k \neq j$).

♣ Події $A_1, A_2, \dots, A_n$ утворюють **повну групу**, якщо їхня сума – достовірна подія і вони попарно несумісні:

$$A_1 + A_2 + \dots + A_n = \Omega,$$

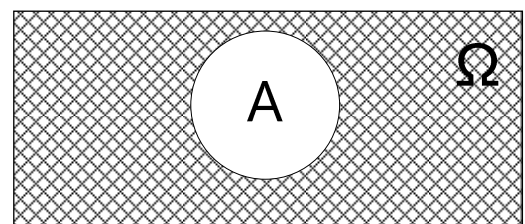
$$A_i \cdot A_j = \emptyset \text{ при } i \neq j.$$



$$A_1 + A_2 + A_3 + A_4 + A_5 = \Omega$$

♣ $\bar{A}$ – подія, **протилежна** події A , якщо A і $\bar{A}$ несумісні і внаслідок випробування одна з цих подій обов'язково відбудеться:

$$A \cdot \bar{A} = \emptyset; \quad A + \bar{A} = \Omega.$$



$$\bar{A}$$

♣ **Алгебра подій** U – така сукупність подій, що для будь-яких двох із них їхня сума, різниця і добуток також входять до U , тобто:

якщо $A, B \in U$, то: $A+B \in U$, $A \setminus B \in U$, $A \cdot B \in U$.

Деякі властивості операцій додавання і добутку подій:

$$A + B = B + A$$

$$A \cdot B = B \cdot A$$

$$A \cdot B \subset A$$

$$A \setminus B = A \cdot \bar{B}$$

$$\overline{A + B} = \bar{A} \cdot \bar{B}$$

$$\overline{A \cdot B} = \bar{A} + \bar{B}$$

$$\bar{\bar{A}} = A$$

$$A + (B + C) = (A + B) + C$$

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C$$

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C$$

$$(A + B) \cdot C = A \cdot C + B \cdot C$$

$$A + \bar{A} = \Omega$$

$$A \cdot \bar{A} = \emptyset$$

$$A + A = A$$

$$A \cdot A = A$$

$$A + \emptyset = A$$

$$A \cdot \emptyset = \emptyset$$

$$A + \Omega = \Omega$$

$$A \cdot \Omega = A$$

2. Імовірність події

2.1. Статистична ймовірність

Нехай проводять n однакових незалежних один від одного дослідів, і нехай в m із них подія A відбулася.

♣ **Відносна частота події A** у проведеній серії експериментів – це відношення:

$$\frac{m}{n} \xrightarrow{n \rightarrow \infty} P^*(A). \quad [2.1]$$

♣ **Статистична ймовірність $P^*(A)$ події A** – це її відносна частота, якщо при збільшенні числа дослідів вона коливається біля деякого фіксованого значення.

2.2. Класичне визначення ймовірності

При класичному підході розглядають лише скінченне число рівноможливих подій.

♣ **Елементарні події** (чи **елементарні наслідки випробування**) – попарно несумісні, рівноможливі події, сума яких є достовірною подією.

Множину усіх елементарних подій позначають $\Omega = \{\omega_1, \dots, \omega_n\}$.

Елементарними є рівноможливі події, що утворюють повну групу.

Наслідки, що сприяють події, – ті елементарні наслідки, в яких настає ця подія.

♣ **Імовірність** $P(A)$ події A – відношення числа наслідків, що сприяють цій події, до загального числа усіх елементарних наслідків випробування.

$$P(A) = \frac{m}{n}, \quad [2.2]$$

де m – число елементарних наслідків, що сприяють A ;

n – число усіх можливих елементарних наслідків випробування.

2.3. Властивості ймовірності

1. Імовірність достовірної події дорівнює 1:

$$P(\Omega) = 1 \quad (\text{оскільки } m = n).$$

2. Імовірність неможливої події дорівнює 0:

$$P(\emptyset) = 0 \quad (\text{оскільки } m = 0).$$

3. Імовірність події набуває значень від 0 до 1:

$$0 \leq P(A) \leq 1 \quad (\text{оскільки } 0 \leq m \leq n).$$

4. Імовірність протилежної події:

$$P(\bar{A}) = 1 - P(A) \quad (\text{оскільки } P(\bar{A}) = \frac{n-m}{n} = 1 - \frac{m}{n} = 1 - P(A)).$$

5. Імовірність суми подій, що утворюють повну групу, дорівнює 1:

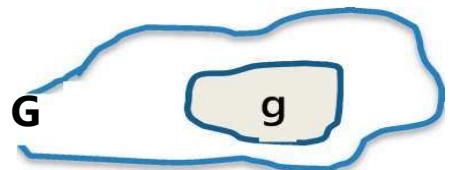
$$P\left(\sum_{k=1}^n A_k\right) = 1 \quad \left(\text{оскільки для повної групи подій } \sum_{k=1}^n A_k = \Omega\right)$$

2.4. Геометрична ймовірність

♣ **Геометрична ймовірність** події A – відношення міри (*mes*) тій області g , що сприяє появі події, до міри усієї області G можливих наслідків події:

$$P(A) = \frac{\text{mes}(g)}{\text{mes}(G)}$$

де *mes* – міра області.



Мірами можуть бути, наприклад:

- довжини відрізків: $P(A) = \frac{|I|}{|L|}$ (якщо розглядають імовірність попадання точки на відрізок I , що є частиною відрізка L);
- площі плоских фігур: $P(A) = \frac{S(g)}{S(G)}$;
- об'єми тривимірних фігур: $P(A) = \frac{V(g)}{V(G)}$ і так далі.

2.5. Аксиоматичне визначення ймовірності

Поставимо у відповідність кожній випадковій події A з алгебри подій U певне число $P(A)$ і назвемо його ймовірністю, якщо виконана низка умов.

♣ **Імовірністю** $P(A)$ називають числову функцію подій, що задана на алгебрі подій і задовольняє аксіомам:

Аксиома 1. Імовірність набуває невід'ємних значень:

$$P(A) \geq 0$$

Аксиома 2. Імовірність достовірної події Ω дорівнює 1:

$$P(\Omega) = 1$$

Аксиома 3 (аксіома додавання ймовірностей, аксіома адитивності).

Якщо A і B – несумісні події, то ймовірність того, що відбудеться хоча б одна з цих двох подій, дорівнює сумі їхніх ймовірностей:

$$P(A+B) = P(A) + P(B), \quad \text{якщо } A \cdot B = \emptyset \quad [2.3]$$

$$(A, B \in U)$$

У загальній математичній теорії ймовірностей досліджують і нескінченне число випадкових подій. При цьому передбачають, що виконується ще одна аксіома – безперервності.

Аксиома 4 (аксіома безперервності). Для будь-якої спадної послідовності подій

$$A_1 \supset A_2 \supset \dots \supset A_k \supset \dots, \text{ такий що } \bigcap_k A_k = \emptyset, \text{ справедливо:}$$

$$\lim_{k \rightarrow \infty} P(A_k) = 0$$

Якщо число подій скінченне, то аксіома 4 впливає з перших трьох.

Якщо прийнято аксіоматику Колмогорова, то тепер деяку величину можна назвати ймовірністю, лише якщо усі ці аксіоми виконано (тобто слід перевірити виконання кожної з них).

Класична, статистична і геометрична ймовірності – це окремі випадки аксіоматики Колмогорова.

Наслідок 1. Якщо події $A_1, A_2, \dots, A_n$ – попарно несумісні, то ймовірність появи їхньої суми дорівнює сумі їхніх ймовірностей:

$$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

$$\left(\text{або } P\left(\sum_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k) \right).$$

Наслідок 2. Сума ймовірностей подій $A_1, A_2, \dots, A_n$, що утворюють повну групу подій, дорівнює 1:

$$P(A_1) + P(A_2) + \dots + P(A_n) = 1$$

$$\left(\text{або } \sum_{k=1}^n P(A_k) = 1 \right).$$

Наслідок 3. Сума ймовірностей протилежних подій дорівнює 1:

$$P(A) + P(\bar{A}) = 1$$

Звідси: $P(\bar{A}) = 1 - P(A)$.

Наслідок 4. Ймовірність будь-якої події менше або дорівнює 1:

$$P(A) \leq 1$$

Наслідок 5. Ймовірність неможливої події дорівнює нулю:

$$P(\emptyset) = 0$$

3. Додавання і множення ймовірностей

3.1. Умовна ймовірність

♣ **Умовна ймовірність** $P(B/A)$ – ймовірність події B , обчислена в припущенні, що подія A вже відбулася.

♣ Подія B **незалежна** від події A з $P(A) \neq 0$, якщо ймовірність події B не залежить від того, відбулася подія A чи ні.

♣ Подія B **залежна** від події A , якщо ймовірність події B змінюється залежно від того, відбулася подія A чи ні.

Умова незалежності події B від події A :

$$P(B/A) = P(B). \quad [3.1]$$

♣ Декілька подій називають **незалежними у сукупності**, якщо ймовірність настання кожної з них не залежить від результатів будь-якого числа інших подій, що вже відбулися.

3.2. Теорема множення ймовірностей

§ **Теорема (множення).** Ймовірність сумісного настання двох залежних подій дорівнює добутку ймовірностей однієї з них на умовну ймовірність іншої, обчислену в припущенні, що перша подія вже відбулася.

$$P(A \cdot B) = P(A) \cdot P(B/A). \quad [3.2]$$

Зауваження. Немає значення, яку з подій A і B вважати першою, а яку другою, тобто виконується також:

$$P(A \cdot B) = P(B) \cdot P(A/B) \quad [3.3]$$

Формула **обчислення умовної ймовірності** $P(B/A)$ події B за умови, що відбулася подія A із ненульовою ймовірністю $P(A) \neq 0$:

$$P(B/A) = \frac{P(A \cdot B)}{P(A)}, \quad P(A) > 0. \quad [3.4]$$

Наслідок 1. Якщо подія A не залежить від події B , то і подія B не залежить від події A .

Наслідок 2. Якщо події A і B незалежні, то ймовірність добутку цих подій дорівнює:

$$P(A \cdot B) = P(A) \cdot P(B). \quad [3.5]$$

(добутку ймовірностей цих подій).

Наслідок 3. Імовірність добутку скінченного числа незалежних у сукупності подій дорівнює:

$$P(A_1 \cdot A_2 \cdot \dots \cdot A_n) = P(A_1) \cdot P(A_2) \cdot \dots \cdot P(A_n)$$

(добутку ймовірностей цих подій)

або
$$P\left(\prod_{k=1}^n A_k\right) = \prod_{k=1}^n P(A_k).$$

Наслідок 4. Імовірність добутку скінченного числа залежних подій дорівнює:

$$P(A_1 \cdot A_2 \cdot \dots \cdot A_n) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 A_2) \cdot \dots \cdot P(A_n/A_1 A_2 \dots A_{n-1})$$

(добутку ймовірності однієї з них на умовну ймовірність усіх інших, обчислених у припущенні, що усі попередні події вже відбулися).

3.3. Імовірність появи хоча б однієї події

Нехай унаслідок випробування може з'явитися хоча б одна (чи декілька) із n незалежних в сукупності подій, імовірності яких відомі.

 **Теорема.** Імовірність появи хоча б однієї з незалежних у сукупності подій $A_1, A_2, \dots, A_n$ дорівнює:

$$P(A_1 + A_2 + \dots + A_n) = 1 - P(\overline{A_1}) \cdot P(\overline{A_2}) \cdot \dots \cdot P(\overline{A_n})$$

(різниці між 1 і добутком ймовірностей протилежних подій $\overline{A_1}, \overline{A_2}, \dots, \overline{A_n}$)

або
$$P\left(\sum_{k=1}^n A_k\right) = 1 - \prod_{k=1}^n P(\overline{A_k}).$$

Якщо позначити ймовірність кожної з подій $\overline{A_1}, \overline{A_2}, \dots, \overline{A_n}$ через $q_1, q_2, \dots, q_n$, то попередню формулу можна записати коротше:

$$P(A) = 1 - q_1 \cdot q_2 \cdot \dots \cdot q_n.$$

Зауваження. Якщо події $A_1, A_2, \dots, A_n$ мають однакову ймовірність p , то ймовірність появи хоча б однієї з них:

$$P(A) = 1 - q^n \quad (\text{де } q = 1 - p).$$

3.4. Теорема додавання ймовірностей сумісних подій



Теорема. Ймовірність суми двох сумісних подій дорівнює:

$$P(A+B) = P(A) + P(B) - P(A \cdot B) \quad [3.6]$$

(сумі ймовірностей цих подій без ймовірності їх сумісної появи).

3.5. Формула повної ймовірності

♣ Нехай подія A може відбутися лише одночасно з однією з подій $H_1, H_2, \dots, H_n$, що утворюють повну групу. Події $H_1, H_2, \dots, H_n$ називають **гіпотезами** (оскільки заздалегідь невідомо, яка з них настане).



Теорема (формула повної ймовірності). Ймовірність події A дорівнює:

$$P(A) = P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) + \dots + P(H_n) \cdot P(A/H_n) \quad [3.7]$$

(сумі попарних добутків ймовірностей гіпотез і відповідних умовних ймовірностей події A)

або
$$P(A) = \sum_{j=1}^n P(H_j) \cdot P(A/H_j).$$

3.6. Формула Бейєса

Нехай подія A може відбутися лише водночас з однією із подій (гіпотез) $H_1, H_2, \dots, H_n$, що утворюють повну групу. Ймовірності цих гіпотез відомі до досвіду (априорні ймовірності) і дорівнюють $P(H_1), P(H_2), \dots, P(H_n)$. І нехай відомо, що подія A таки відбулася, але невідомо, з якою саме гіпотезою.

Формула Бейєса дозволяє переобчислити умовні ймовірності $P(H_j/A)$ гіпотез (*апостеріорні ймовірності*) в цих нових умовах – за умови, що відбулася подія A :

$$P(H_j/A) = \frac{P(H_j) \cdot P(A/H_j)}{P(A)} \quad (j=\overline{1, n}) \quad [3.8]$$

4. Повторні випробування

4.1. Формула Бернуллі

♣ **Схема випробувань Бернуллі** – це послідовність n однакових незалежних випробувань, що задовольняють умовам:

- 1) кожне випробування має два наслідки: успіх (поява події, що позначають 1) і невдача (непоява події, позначають 0); вони взаємно несумісні і протилежні (випробування з *бінарними* результатами);
- 2) ймовірність успіху p у кожному випробуванні стала. Ймовірність невдачі $q=1-p$.

🌀 **Теорема.** Ймовірність того, що в серії з n випробувань за схемою Бернуллі подія A з'явиться точно k разів ($0 \leq k \leq n$), може бути обчислена за **формулою Бернуллі**:

$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k}. \quad [4.1]$$

Наслідок 1. Ймовірність, що подія A відбудеться не більше ніж k_1 разів:

$$P_n(k \leq k_1) = \sum_{k=0}^{k_1} C_n^k \cdot p^k \cdot q^{n-k}.$$

Наслідок 2. Ймовірність, що подія A відбудеться не менше ніж k_2 разів:

$$P_n(k \geq k_2) = \sum_{k=k_2}^n C_n^k \cdot p^k \cdot q^{n-k}.$$

Наслідок 3. Ймовірність, що подія A відбудеться не менше ніж k_1 і не більше ніж k_2 разів:

$$P_n(k_1 \leq k \leq k_2) = \sum_{k=k_1}^{k_2} C_n^k \cdot p^k \cdot q^{n-k}.$$

♣ **Найімовірнішим числом** появи події в n незалежних випробуваннях називають число k_0 , що задовольняє умові $P_n(k_0) \geq P_n(k)$ для усіх $k = \overline{0, n}$, що не дорівнюють k_0 .

Твердження. Найімовірніше число k_0 появи події A в серії із n незалежних випробувань задовольняє нерівності

$$np - q \leq k_0 \leq np + p. \quad [4.2]$$

Якщо $np - q$ не є цілим числом, то значення k_0 єдине. Якщо $np - q$ і $np + p$ – цілі числа, то найімовірніших значень два: $k_1 = np - q$; $k_2 = np + p$.

4.2. Асимптотичні наближення

При великих значеннях n (наприклад, $n > 50$) і k обчислення за формулою Бернуллі стає громіздким. Тому застосовують асимптотичні формули, що дають наближені розрахунки ймовірностей.

4.2.1. Локальна теорема Муавра – Лапласа

Теорема (локальна Муавра – Лапласа). Якщо в кожному з n незалежних випробувань за схемою Бернуллі подія A відбувається з однією і тією ж сталою ймовірністю p , відмінною від 0 і 1, то ймовірність появи події A точно k разів для великих значень n приблизно дорівнює:

$$P_n(k) \approx \frac{\varphi(x)}{\sqrt{n \cdot p \cdot q}}, \quad [4.3]$$

де $x = \frac{k - np}{\sqrt{npq}}$ – обмежена величина, $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$ – функція Гаусса, $p \neq 0, q \neq 0, n \rightarrow \infty$.

Значення функції Гаусса $\varphi(x)$ протабульовано і зведено в таблицю для невід'ємних значень x . Оскільки ця функція парна, то для від'ємних значень x : $\varphi(-x) = \varphi(x)$. Для $|x| \geq 4$ вважають $\varphi(x) = 0$.

4.2.2. Інтегральна теорема Муавра – Лапласа

Теорема (інтегральна Муавра – Лапласа). Якщо в кожному з n незалежних випробувань за схемою Бернуллі подія A відбувається з однією і тією ж сталою ймовірністю p , відмінною від 0 і 1 , то ймовірність того, що подія A відбудеться від k_1 до k_2 разів, для великих значень n приблизно дорівнює:

$$P_n(k_1 \leq k \leq k_2) \approx \Phi(x_2) - \Phi(x_1), \quad [4.4]$$

де $x_1 = \frac{k_1 - np}{\sqrt{npq}}$, $x_2 = \frac{k_2 - np}{\sqrt{npq}}$, $p \neq 0$, $q \neq 0$, $n \rightarrow \infty$,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt \quad \text{— функція Лапласа (або інтеграл імовірностей).}$$

Оскільки інтеграл $\int_0^x e^{-\frac{t^2}{2}} dt$ не можна виразити через елементарні функції, то значення функції Лапласа протабульовано і зведено в таблицю — для невід’ємних значень x . Для від’ємних значень x (враховуючи, що функція непарна): $\Phi(-x) = -\Phi(x)$. Для $x \geq 4$ з великою точністю вважають $\Phi(x) \approx 0,5$.

4.2.3. Формула Пуассона для малоймовірних випадкових подій

Якщо ймовірність p настання події в окремому випробуванні близька до нуля, то навіть при великому числі випробувань n , але при невеликій величині добутку np , значення ймовірностей $P_n(k)$, отримані з локальної теореми Лапласа, виявляються недостатньо точними, і виникає потреба в іншій наближеній формулі.

У випадку, коли $n \rightarrow \infty$, а $p \rightarrow 0$, але так, що $\lim_{n \rightarrow \infty} np = \lambda > 0$ (де $p = \frac{\lambda}{n} + o(\frac{1}{n})$), із формули Бернуллі можна отримати іншу формулу — Пуассона.

Теорема Пуассона. Якщо ймовірність p настання події A в кожному випробуванні стала і мала, число незалежних випробувань n достатньо велике, але значення добутку $np = \lambda$ залишається невеликим, то ймовірність того, що в цих випробуваннях подія A відбудеться k разів, приблизно дорівнює:

формула Пуассона:

$$P_n(k) \approx \frac{\lambda^k}{k!} e^{-\lambda}. \quad [4.5]$$

Значення функції Пуассона протабульовано і зведено в таблицю.

5. Випадкова величина

♣ Величину, що набуває внаслідок випробування можливого значення (одного або декількох), наперед невідомого і залежного від випадкових причин, які заздалегідь не можуть бути враховані, називають **випадковою величиною** (ВВ).

5.1. Поняття випадкової величини

♣ У теорії ймовірностей **випадкову величину** визначають як числову функцію $\xi(\omega)$ від елементарної події $\omega \in \Omega$, причому сукупність усіх подій ω , для яких $\xi(\omega) < x$, належить алгебрі подій для будь-якого $x \in \mathbb{R}$:

$$\Omega \xrightarrow{\xi(\omega)} \mathbb{R}, \quad \text{при цьому } \{\omega: \xi(\omega) < x\} \in U \quad (U - \text{алгебра подій, } x \in \mathbb{R}).$$

Випадкові величини позначають часто грецькими літерами ξ, η, ζ або прописними латинськими X, Y, Z , а можливі конкретні значення випадкових величин – рядковими x, y, z .

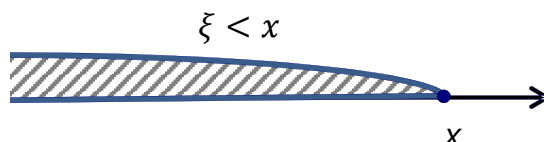
Випадкова величина (тобто функція $\xi(\omega)$) – це абстрактна функція, яку можна вивчати лише побічно. Щоб описати ВВ, потрібно визначити ймовірності подій, пов'язаних з нею. Якнайповнішим чином характеризує випадкову величину *функція розподілу ймовірностей* – вона відображає, як ці ймовірності розподілені по всіх значеннях ВВ.

♣ **Функція розподілу ймовірностей** $F(x)$ – це ймовірність тих подій, в яких ВВ набуває значення менше ніж x (тобто $\xi < x$):

$$F(x) = P(\xi < x) \quad (x \in \mathbb{R})$$
 – функція розподілу ймовірностей.

Функція розподілу визначена для усіх подій.

Якщо значення ВВ ξ – точки на числовій осі, то геометрично функцію розподілу інтерпретують як ймовірність того, що ВВ потрапляє лівіше заданої точки x .



Функція розподілу – основна характеристика випадкової величини. Знаючи функцію розподілу ВВ, можна обчислити ймовірності будь-яких подій, пов'язаних з нею, а також інші випадкові характеристики.

Одна і та ж функція розподілу може описувати скільки завгодно різних випадкових величин, тобто $F_\xi(x) = F_\eta(x) \not\Rightarrow \xi(\omega) \equiv \eta(\omega)$.

5.2. Властивості функції розподілу ймовірностей

1. Значення функції розподілу належать відрізку $[0; 1]$:

$$\diamond 0 \leq F(x) \leq 1.$$

2. На мінус нескінченності функція розподілу дорівнює 0, на плюс нескінченності дорівнює 1:

$$\diamond F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0; \quad F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1.$$

3. Функція розподілу – неспадна на всій числовій осі:

$$\diamond \text{Якщо } x_1 < x_2, \text{ то } F(x_1) \leq F(x_2).$$

4. Ймовірність влучення ВВ в інтервал $[a, b)$ (включаючи a) дорівнює приросту функції її розподілу на цьому інтервалі:

$$\diamond P(a \leq \xi < b) = F(b) - F(a).$$

Це одна з основних формул теорії ймовірностей.

5. Функція розподілу неперервна зліва:

$$\diamond \lim_{x \rightarrow x_0 - 0} F(x) = F(x_0).$$

Графіком функції розподілу $F(x)$ у загальному випадку є графік неспадної функції, значення якої починаються від 0, доходять до 1, причому в окремих точках функція може мати скачки (розриви) і точки перегину.

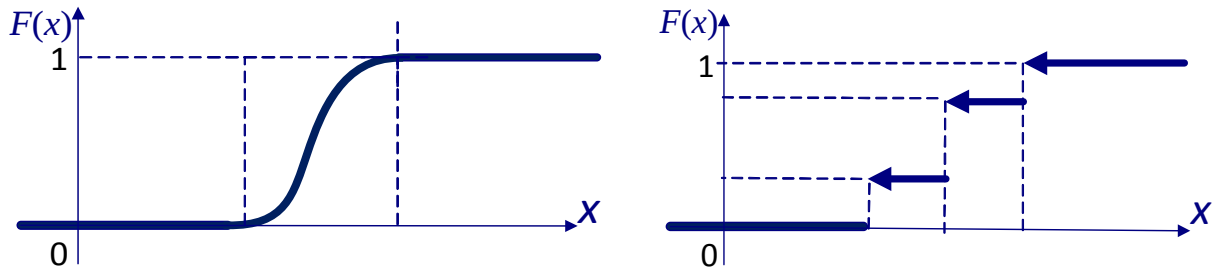


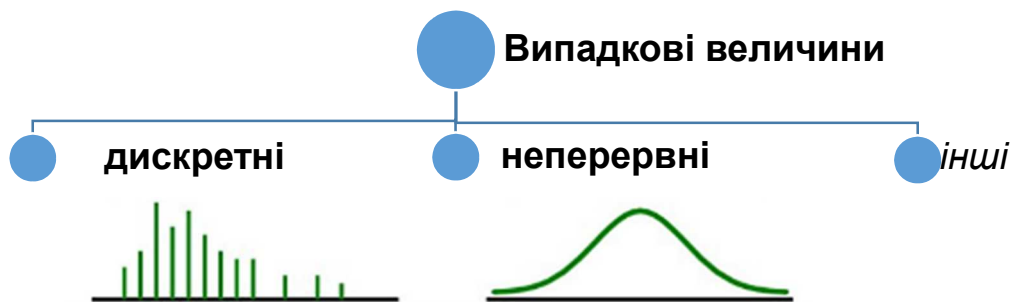
Рис. 5.1. Приклади функцій розподілу

5.3. Класифікація випадкових величин

5.3.1. Дискретні і неперервні ВВ

♣ ВВ називають **дискретною**, якщо вона приймає скінченну або зліченну кількість окремих значень.

♣ ВВ називають **неперервною**, якщо її функція розподілу неперервна у будь-якій точці і диференційована усюди, окрім, можливо, окремих точок.



5.3.2. Закон розподілу і ряд розподілу дискретної ВВ

♣ Задано розподіл імовірностей дискретної випадкової величини ξ , якщо визначено ймовірності всіляких подій, у яких ξ набуває значень $x = \xi(\omega)$.

♣ Будь-яке правило, що дозволяє знаходити ймовірності усіх можливих значень ВВ, називають **законом розподілу ймовірностей** (=законом розподілу) цієї випадкової величини.

Закон розподілу ВВ може бути задано таблично, графічно або аналітично.

Для табличного опису дискретної ВВ необхідно задати її значення $x_1, x_2, \dots, x_n$ ($n \leq \infty$) та їхні ймовірності $p_k = P(\xi = x_k)$. Цю інформацію зручно зображувати у вигляді таблиці:

ξ	x_1	x_2	$\dots$	x_n
P	p_1	p_2	$\dots$	p_n

[5.1]

♣ Табличну відповідність між можливими значеннями ВВ та їхніми ймовірностями називають **рядом розподілу**.

При цьому розглядатимемо такі чотири обмеження: [5.2]

- $p_k > 0$;
- $x_j \neq x_k$ при $j \neq k$;
- значення ВВ x_k розташовують за збільшенням:
 $x_1 < x_2 < \dots < x_n$;
- сума ймовірностей подій $\xi(\omega) = x_k$ дорівнює одиниці:

$$\sum_{k=1}^n p_k = 1 \quad (\text{умова нормування}).$$

♣ Ряд розподілу [5.1] із властивостями (5.2) є **законом розподілу ймовірностей дискретної ВВ**.

Для графічного опису закону розподілу дискретної ВВ використовують багатокутник розподілу (рис. 5.2): по осі абсцис відкладають можливі значення ВВ, а по осі ординат – їхні ймовірності. Отримані точки з'єднують відрізками прямих.

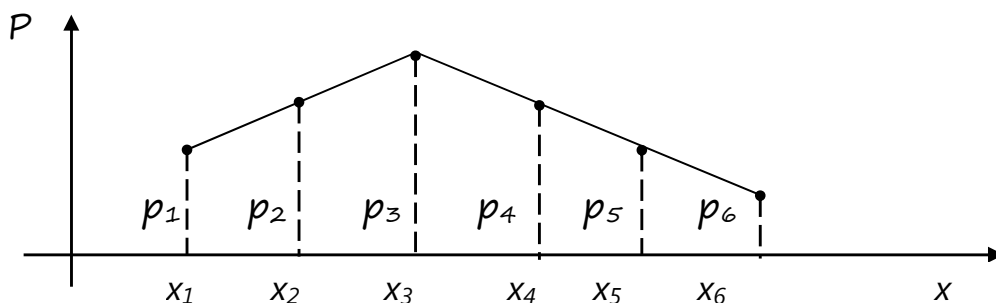


Рис. 5.2. Багатокутник розподілу

5.3.3. Загальний вигляд функції розподілу дискретної ВВ

Загальний вигляд функції розподілу для дискретної ВВ:

$$F(x) = \begin{cases} 0, & x \leq x_1 \\ p_1, & x_1 < x \leq x_2 \\ p_1 + p_2, & x_2 < x \leq x_3 \\ \dots & \dots \\ p_1 + \dots + p_{n-1}, & x_{n-1} < x \leq x_n \\ 1, & x > x_n \end{cases}$$

Графік функції розподілу $F(x)$ будь-якої дискретної ВВ ξ має ступінчастий вигляд. Стрибки відбуваються у точках, що відповідають можливим значенням ВВ. Величини стрибків дорівнюють імовірностям, з якими ВВ набуває цих значень. Сума усіх стрибків дорівнює 1.

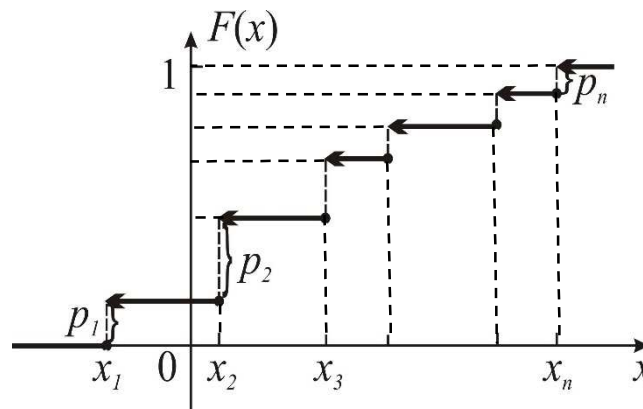


Рис. 5.3. Функція розподілу дискретної ВВ

І ряд розподілу, і функція розподілу повністю описують випадкову величину та є різними формами закону її розподілу. При цьому функція розподілу – найбільш загальна його форма, придатна для опису поведінки як дискретних, так і неперервних ВВ.

5.4. Система декількох випадкових величин

5.4.1. Багатовимірні випадкові величини

Випадкові величини, значеннями яких є число, називають одновимірними. ВВ, значення яких визначаються двома, трьома, ... n числами, називають відповідно двовимірними, тривимірними, ... n -вимірними ВВ.

Можна розглядати n -вимірну ВВ як систему n одновимірних випадкових величин. Наприклад, розглядаючи одночасно дві ВВ ξ і η , можна вважати їх системою двох ВВ або ж двовимірною випадковою величиною – її позначають тоді (ξ, η) .

Властивості системи декількох ВВ визначаються не лише властивостями самих цих ВВ, а також і зв'язками (залежностями) між ними.

♣ Упорядкований набір $(\xi_1, \xi_2, \dots, \xi_n)$ n випадкових величин ξ_k , заданих на одному і тому ж просторі елементарних подій Ω , називають **n -вимірною випадковою величиною, або системою n випадкових величин.**

Закон розподілу такої системи може бути задано різними способами – графічно, таблично, аналітично.

Наприклад, дискретна двовимірна ВВ (ξ, η) може бути задана формулою ймовірності одночасного прийняття¹ значень:

$$p_{ij} = P(\xi = x_i, \eta = y_j) \quad (i=1, \dots, n, j=1, \dots, m)$$

або таблицею:

$\xi \backslash \eta$	y_1	y_2	$\dots$	y_m
x_1	p_{11}	p_{12}	$\dots$	p_{1m}
x_2	p_{21}	p_{22}	$\dots$	p_{2m}
$\dots$				
x_n	p_{n1}	p_{n2}	$\dots$	p_{nm}

При цьому сума усіх ймовірностей дорівнює 1 (як сума ймовірностей повної групи несумісних подій):

$$\sum_{i=1}^n \sum_{j=1}^m p_{ij} = 1.$$

Універсальною формою закону розподілу багатовимірної ВВ є функція сумісного розподілу ймовірностей.

♣ Функція розподілу двовимірної ВВ (ξ, η) – це функція $F(x, y)$, яка для кожної пари чисел (x, y) визначає ймовірність одночасного виконання нерівностей $\xi < x$ і $\eta < y$:

$$F(x, y) = P(\xi < x, \eta < y).$$

Тобто $F(x, y) = P(AB)$, де $A = \{\omega: \xi < x\}$, $B = \{\omega: \eta < y\}$.

Аналогічно визначають функцію розподілу системи n ВВ:

$$F(x_1, x_2, \dots, x_n) = P(\xi_1 < x_1, \xi_2 < x_2, \dots, \xi_n < x_n).$$

Знаючи функцію сумісного розподілу, можна знайти розподіл кожної зі складових ВВ. Розв'язання зворотної задачі у загальному випадку не існує. Але можна відновити функцію сумісного розподілу по розподілах окремих ВВ, якщо вони незалежні.

¹ Запис $(\xi = x_i, \eta = y_j)$ позначає добуток подій $(\xi = x_i)$ і $(\eta = y_j)$.

5.4.2. Залежні і незалежні випадкові величини

У теорії ймовірностей йдеться не про повну (функціональну) залежність – коли, знаючи значення однієї величини, можна точно вказати значення іншої. Імовірнісна (стохастична) залежність – така, коли знаючи значення однієї величини, можна лише вказати залежний від нього закон розподілу іншої величини.

Дві випадкові величини вважають незалежними, якщо закон розподілу однієї з них не дає жодної інформації про розподіл іншої ВВ:

♣ **Дві випадкові величини ξ і η називають незалежними**, якщо незалежні події $A=\{\omega: \xi < x\}$ і $B=\{\omega: \eta < y\}$ для будь-яких можливих $x, y \in R$. Інакше величини називають **залежними**.

Як відомо, для незалежних подій A і B : $P(A \cdot B) = P(A) \cdot P(B)$. [5.3]

Цю рівність можна записати за допомогою функцій розподілу (враховуючи їхнє визначення $F_\xi(x)=P(\xi < x)$). Таким чином, для двох незалежних ВВ, що утворюють систему, виконується рівність:

$$F_{\xi, \eta}(x, y) = F_\xi(x) \cdot F_\eta(y) \quad [5.4]$$

(сумісна функція розподілу двох ВВ дорівнює добутку функцій розподілу кожної з них).

Умова незалежності двох дискретних ВВ, що утворюють систему, також може бути записана у вигляді (5.3), якщо через A і B позначити події: $A=\{\omega: \xi=x_j\}$ і $B=\{\omega: \eta=y_k\}$. Тоді (5.3) набере вигляду:

$$P(\xi=x_j, \eta=y_k) = P(\xi=x_j) \cdot P(\eta=y_k).$$

♣ **Дві дискретні ВВ ξ і η називають незалежними**, якщо незалежні події $\{\omega: \xi = x_j\}$ і $\{\omega: \eta = y_k\}$ при будь-яких можливих x_j і y_k . Інакше величини називають **залежними**.

Залежність або незалежність випадкових величин завжди взаємні: якщо ξ не залежить від η , то і η не залежить від ξ .

Для набору n випадкових величин $(\xi_1, \xi_2, \dots, \xi_n)$ можна говорити про попарну незалежність ВВ і про незалежність у сукупності.

♣ **Випадкові величини** називають **попарно незалежними**, якщо незалежні будь-які дві з них.

♣ **Випадкові величини** $\xi_1, \xi_2, \dots, \xi_n$ називають **незалежними в сукупності** (=взаємно незалежними), якщо для будь-якого їх піднабору функція їхнього сумісного розподілу дорівнює добутку функцій розподілу кожної ВВ, тобто:

для усіх комбінацій індексів $1 \leq k_1 < \dots < k_j \leq n$ (де $j=2, \dots, n$) виконується:

$$F_{\xi_{k_1}, \dots, \xi_{k_j}}(x_{k_1}, \dots, x_{k_j}) = F_{\xi_{k_1}}(x_{k_1}) \cdot \dots \cdot F_{\xi_{k_j}}(x_{k_j}).$$

Так само як і для подій, для випадкових величин із незалежності у сукупності впливає попарна їх незалежність. Зворотне неправильно.

5.5. Операції над дискретними випадковими величинами

У загальному випадку для дискретної ВВ ξ , що задана законом розподілу:

ξ	x_1	x_2	...	x_n
$P(\xi=x_k)$	p_1	p_2	...	p_n

результат застосування до неї деякої операції f задає випадкову величину $\eta=f(\xi)$, що описується законом розподілу:

$\eta=f(\xi)$	$f(x_1)$	$f(x_2)$	...	$f(x_n)$
$P(\xi=x_k)$	p_1	p_2	...	p_n

якщо усі $f(x_k)$ різні. При цьому якщо внаслідок застосування операції в ряді розподілу ВВ η з'являться значення, що повторюються, їх включають у ряд лише один раз, складаючи їхні ймовірності.

Операції над однією випадковою величиною.

♣ **Добуток випадкової величини ξ на сталу величину C** – випадкова величина $\eta = C \cdot \xi$, усі значення якої є помноженими на C значеннями ξ з тими ж ймовірностями.

$$\eta = C \cdot \xi$$

$\eta=C\xi$	Cx_1	Cx_2	...	Cx_n
P	p_1	p_2	...	p_n

♣ **m -й степінь випадкової величини ξ** – випадкова величина $\eta = \xi^m$, усі значення якої дорівнюють m -му степеню значень ξ з тими ж імовірностями.

$$\eta = \xi^m$$

$\eta = \xi^m$	x_1^m	x_2^m	...	x_n^m
P	p_1	p_2	...	p_n

– якщо m не є цілим парним числом або дробовим із парним числівником.

♣ **Сума випадкової величини ξ і сталої величини C** – випадкова величина $\eta = \xi + C$, усі значення якої є сумами значень ξ і сталої C з тими ж імовірностями.

$$\eta = \xi + C$$

$\eta = \xi + C$	$x_1 + C$	$x_2 + C$	...	$x_n + C$
P	p_1	p_2	...	p_n

Операції над двома випадковими величинами.

Нехай задано закони розподілу ВВ ξ і η :

ξ	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

η	y_1	y_2	...	y_m
P	$\hat{p}_1$	$\hat{p}_2$	...	$\hat{p}_m$

де $p_j = P(\xi = x_j)$, $\hat{p}_k = P(\eta = y_k)$.

Застосуємо деяку (узагальнену) операцію f до обох цих ВВ. Результатом буде нова ВВ $\zeta = f(\xi, \eta)$, яка набуває значень

$$\zeta = z_{jk} = f(x_j, y_k).$$

Значення ζ розраховують, заповнюючи спочатку прямокутну матрицю вигляду:

$\eta \backslash \xi$	x_1	x_2	...	x_n
y_1	$f(x_1, y_1)$	$f(x_2, y_1)$		$f(x_n, y_1)$
y_2	$f(x_1, y_2)$	$f(x_2, y_2)$		$f(x_n, y_2)$
...				
y_m	$f(x_1, y_m)$	$f(x_2, y_m)$		$f(x_n, y_m)$

Імовірності цих можливих значень ζ також розраховують та записують у відповідні клітинки тієї ж матриці через рису. Інший спосіб –

формують аналогічну прямокутну матрицю вигляду:

$\eta \backslash \xi$	x_1	x_2	$\dots$	x_n
y_1	p_{11}	p_{12}	$\dots$	p_{1n}
y_2	p_{21}	p_{22}	$\dots$	p_{2n}
$\dots$				
y_m	p_{m1}	p_{m2}	$\dots$	p_{mn}

де $p_{jk} = P(\xi = x_j, \eta = y_k)$ – імовірності появи подій $A = (\omega: \xi = x_j)$ і $B = (\omega: \eta = y_k)$.
Тобто $p_{jk} = P(A \cdot B)$.

Якщо внаслідок застосування операції в таблиці значень ζ з'явилися значення, що повторюються, їх включають до закону розподілу лише один раз, складаючи відповідні їм імовірності (як імовірності несумісних подій).

Умова нормування:
$$\sum_{j,k}^{n,m} p_{jk} = 1.$$

♣ **Сума** двох дискретних випадкових величин ξ і η – випадкова величина $\zeta = \xi + \eta$, що набуває значень сум $x_j + y_k$ з імовірностями p_{jk} , коли $\xi = x_j, \eta = y_k$.

При цьому якщо дискретні ВВ ξ і η незалежні, то $p_{jk} = p_j \hat{p}_k$ для будь-яких j та k .

Якщо ж ВВ ξ і η не є незалежними, то для ймовірнісного опису їхньої суми ζ потрібно задавати ймовірності $p_{jk} = P(\xi = x_j, \eta = y_k)$ – їх не можна отримати із законів розподілу ξ і η .

♣ **Різниця** двох випадкових величин ξ і η – сума двох випадкових величин ξ і $-\eta$:

$$\xi - \eta = \xi + (-\eta).$$

♣ **Добуток** двох випадкових величин ξ і η – випадкова величина $\zeta = \xi \cdot \eta$, що набуває значень всіляких добутків $x_j \cdot y_k$ з імовірностями p_{jk} .

Таким чином, так само як і для суми ВВ, для ймовірнісного опису добутку ζ потрібно задавати ймовірності $p_{jk} = P(\xi = x_j, \eta = y_k)$, якщо ВВ ξ і η не є незалежними. Тільки для незалежних ВВ ξ і η : $p_{jk} = p_j \hat{p}_k$ для будь-яких j та k .

5.6. Числові характеристики дискретних випадкових величин

Числові характеристики випадкової величини – деякі числові її параметри, що характеризують найбільш суттєві риси її розподілу або поведінку. Серед них: середнє значення ВВ; рівень розкиду значень ВВ відносно її середнього значення, найбільш імовірне значення ВВ тощо.

Найбільш значущими і часто використовуваними характеристиками ВВ є її математичне очікування і дисперсія.

5.6.1. Математичне очікування

♣ **Математичне очікування** дискретної ВВ – сума добутків усіх її можливих значень на їхні імовірності.

$$M\xi = \sum_{k=1}^n x_k p_k.$$

У разі зліченного числа значень ВВ математичне очікування представляє собою суму ряду

$$M\xi = \sum_{k=1}^{\infty} x_k p_k$$

за умови $\sum_{k=1}^{\infty} |x_k| p_k < \infty$.

Математичне очікування – одне з найфундаментальніших понять теорії ймовірностей. За його допомоги можуть бути визначені усі інші числові і функціональні характеристики ВВ.

Математичне очікування – не випадкова величина, а стала (незалежна від випадку).

При великому числі випробувань середнє арифметичне спостережуваних значень ВВ наближається до її мат. очікування. Таким чином, мат. очікування характеризує середнє значення випадкової величини. Навколо значення мат. очікування групуються усі можливі її значення.

5.6.2. Властивості математичного очікування

Властивості математичного очікування для ВВ ξ і η , що задані законами розподілу:

ξ	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

η	y_1	y_2	...	y_m
P	$\hat{p}_1$	$\hat{p}_2$	...	$\hat{p}_m$

1. Математичне очікування сталої величини дорівнює цій сталій:

$$\diamond Mc = c.$$

2. Сталій множник можна виносити за знак математичного очікування:

$$\diamond M(c\xi) = c \cdot M\xi.$$

3. Математичне очікування ВВ $f(\xi)$ (де f – не випадкова функція):

$$\diamond Mf(\xi) = \sum_{k=1}^n f(x_k) \cdot p_k.$$

4. Математичне очікування суми двох ВВ дорівнює сумі їхніх мат. очікувань:

$$\diamond M(\xi + \eta) = M(\xi) + M(\eta).$$

Наслідок 1. Мат. очікування різниці двох ВВ дорівнює різниці їхніх мат. очікувань:

$$\diamond M(\xi - \eta) = M(\xi) - M(\eta).$$

Наслідок 2. Мат. очікування суми декількох ВВ дорівнює сумі їхніх мат. очікувань:

$$\diamond M\left(\sum_{k=1}^n \xi_k\right) = \sum_{k=1}^n M\xi_k.$$

5. Математичне очікування добутку двох ВВ у загальному випадку **не** дорівнює добутку їхніх мат. очікувань:

$$\diamond M(\xi \cdot \eta) \neq M\xi \cdot M\eta.$$

Проте для незалежних ВВ виконується рівність:

$$\diamond M(\xi \cdot \eta) = M\xi \cdot M\eta.$$

Наслідок. Математичне очікування добутку декількох незалежних у сукупності ВВ дорівнює добутку їхніх мат. очікувань:

$$\diamond M\left(\prod_{k=1}^n \xi_k\right) = \prod_{k=1}^n M\xi_k.$$

5.6.3. Відхилення і дисперсія

Існують різні показники розсіювання навколо мат. очікування (тобто наскільки в середньому значення ВВ зосереджені навколо нього).

♣ **Відхилення** $\xi - M\xi$ – різниця між випадковою величиною та її математичним очікуванням.

Відхилення є випадковою величиною. Його мат. очікування завжди дорівнює 0. Тому зазвичай використовують не відхилення $\xi - M\xi$, а його квадрат $(\xi - M\xi)^2$.

♣ **Дисперсія** $D\xi$ випадкової величини ξ – математичне очікування квадрата відхилення ξ від її математичного очікування:

$$D\xi = M(\xi - M\xi)^2. \quad [5.5]$$

Так само як і мат. очікування, дисперсія – не випадкова (стала) величина. Формула розрахунку дисперсії:

$$D\xi = \sum_{k=1}^n (x_k - M\xi)^2 \cdot p_k.$$

Часто для розрахунку дисперсії використовують іншу формулу.

 **Теорема.** Дисперсія ВВ ξ :

$$D\xi = M\xi^2 - (M\xi)^2 \quad [5.6]$$

(дорівнює різниці між математичним очікуванням квадрата ВВ і квадратом її математичного очікування)

5.6.4. Властивості дисперсії

1. Дисперсія сталої величини дорівнює нулю:

$$\diamond Dc = 0.$$

2. Сталій множник можна винести за знак дисперсії, підносячи його до квадрата:

$$\diamond D(c\xi) = c^2 D\xi.$$

3. Дисперсія – величина невід'ємна:

$$\diamond D\xi \geq 0.$$

У загальному випадку дисперсія суми двох ВВ **не** дорівнює сумі їхніх дисперсій:

$$\diamond D(\xi + \eta) \neq D\xi + D\eta.$$

Проте для незалежних ВВ виконується така властивість.

4. Дисперсія суми двох незалежних ВВ дорівнює сумі їхніх дисперсій:

$$\diamond D(\xi + \eta) = D\xi + D\eta.$$

Наслідок 1. Дисперсія суми декількох взаємно незалежних ВВ дорівнює сумі їхніх дисперсій:

$$\diamond D\left(\sum_{k=1}^n \xi_k\right) = \sum_{k=1}^n D\xi_k.$$

Наслідок 2. Дисперсія ВВ не зміниться, якщо до неї додати сталу:

$$\diamond D(\xi + c) = D\xi.$$

Наслідок 3. Дисперсія різниці двох незалежних ВВ дорівнює сумі їхніх дисперсій:

$$\diamond D(\xi - \eta) = D\xi + D\eta.$$

Зауваження. Якщо для дискретної ВВ $D\xi = 0$, то $\xi = \text{const}$, тобто ξ – не випадкова величина.

5.6.5. Середнє квадратичне відхилення

♣ **Середнє квадратичне відхилення** σ випадкової величини ξ – це арифметичне значення квадратного кореня з її дисперсії:

$$\sigma(\xi) = \sqrt{D\xi}. \quad [5.7]$$

Середнє квадратичне відхилення (=стандартне відхилення) також, як і дисперсія, характеризує міру відхилення ВВ від її математичного очікування, але має розмірність ВВ.

 **Теорема.** Середнє квадратичне відхилення суми скінченного числа взаємно незалежних ВВ дорівнює квадратному кореню з суми квадратів середніх квадратичних відхилень цих величин:

$$\diamond \sigma\left(\sum_{k=1}^n \xi_k\right) = \sqrt{\sum_{k=1}^n \sigma^2(\xi_k)}.$$

6. Основні закони розподілу дискретних випадкових величин

6.1. Біноміальний розподіл

♣ Дискретна ВВ ξ має **біноміальний розподіл** (=розподілена за **біноміальним законом розподілу**), якщо вона набуває значень $0, 1, 2, \dots, n$ з імовірностями (формула Бернуллі):

$$P(\xi=k) = P_n(k) = C_n^k \cdot p^k \cdot q^{n-k} \quad [6.1]$$

$(k=0, 1, 2, \dots, n; \quad 0 < p < 1; \quad q=1-p).$

Біноміальний розподіл описує ВВ ξ – число появ події A в серії з n випробувань за схемою Бернуллі: проводять n незалежних випробувань, в кожному з яких імовірність p появи події A незмінна і не залежить від результатів інших випробувань. Вочевидь, ВВ ξ може набувати усіх цілих значень від 0 до n .

Біноміальний закон розподілу випадкової величини:

ξ	0	1	2	...	k	...	n
P	q^n	$C_n^1 \cdot p^1 \cdot q^{n-1}$	$C_n^2 \cdot p^2 \cdot q^{n-2}$	...	$C_n^k \cdot p^k \cdot q^{n-k}$	...	p^n

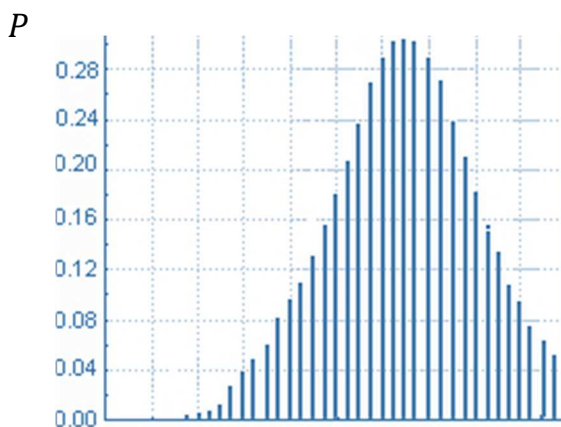
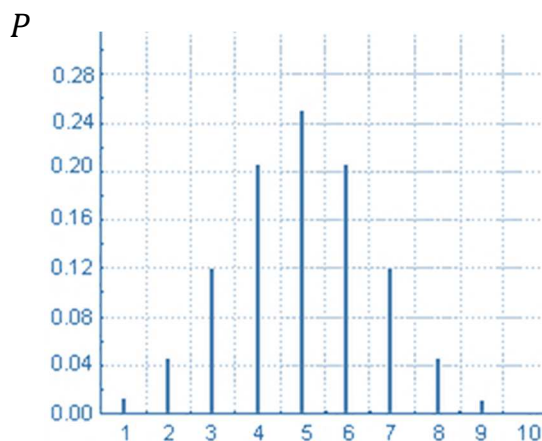


Рис. 6.1. Діаграми біноміального розподілу для: а) $n=10, p=0,5$; б) $n=100, p=0,7$

Теорема. Математичне очікування і дисперсію ВВ ξ , розподіленої за біноміальним законом, розраховують за формулами:

$$M\xi = np, \quad D\xi = npq. \quad [6.2]$$

6.2. Розподіл Пуассона

♣ Дискретна ВВ ξ розподілена за **законом Пуассона** з параметром $\lambda > 0$, якщо її можливі значення дорівнюють $0, 1, 2, \dots, k, \dots$, а відповідні ймовірності обчислюються за формулою:

$$P(\xi=k) = F(k,\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \lambda > 0. \quad [6.3]$$

Розподіл Пуассона описує ВВ ξ – число появ події A в схемі випробувань: проводять n незалежних випробувань, в кожному з яких ймовірність появи події A дорівнює p (ймовірність неяви $q=1-p$).

Закон розподілу Пуассона:

ξ	0	1	2	...	k	...
P	$e^{-\lambda}$	$e^{-\lambda} \cdot \frac{\lambda}{1!}$	$e^{-\lambda} \cdot \frac{\lambda^2}{2!}$	...	$e^{-\lambda} \cdot \frac{\lambda^k}{k!}$	...

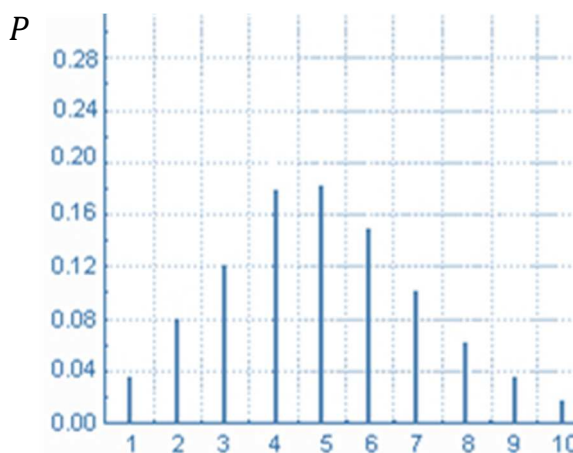


Рис. 6.3. Діаграма розподілу Пуассона для $n=100$, $\lambda=5$

Теорема. Математичне очікування і дисперсія ВВ ξ , розподіленої за законом Пуассона, дорівнюють параметру Пуассона λ :

$$M\xi = \lambda, \quad D\xi = \lambda. \quad [6.4]$$

Чисельно математичне очікування і дисперсія цієї ВВ співпадають, проте мають різні розмірності.

6.3. Дискретний рівномірний розподіл

♣ Дискретна ВВ, що приймає скінченне число значень із рівними ймовірностями, розподілена за **дискретним рівномірним законом**:

$$P(\xi=k) = \frac{1}{n} \quad (k=1,2,\dots,n). \quad [6.5]$$

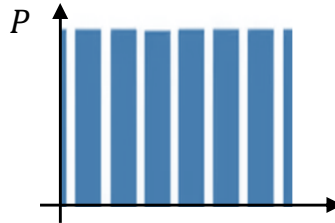


Рис. 6.4. Діаграма рівномірного розподілу

Математичне очікування і дисперсія рівномірно розподіленої дискретної випадкової величини ξ :

$$M\xi = \frac{n+1}{2}, \quad D\xi = \frac{n^2-1}{12}. \quad [6.6]$$

7. Неперервні випадкові величини

7.1. Функція розподілу неперервної випадкової величини

♣ **Функція розподілу** $F(x)$ неперервної ВВ ξ – функція, що визначає ймовірність подій, у яких величина може набути значення менше x .

$$F(x) = P(\omega: \xi < x) \quad (x \in \mathbb{R}). \quad [7.1]$$

Для функцій розподілу неперервних ВВ виконуються усі властивості дискретних ВВ, а також справедливі додаткові.

Властивості функції розподілу неперервної ВВ,
спільні з дискретною ВВ

1. Значення функції розподілу належать відріzkу $[0; 1]$.

$$\diamond 0 \leq F(x) \leq 1$$

2. Якщо можливі значення ВВ розташовані на всій числовій осі, то

$$\diamond F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$$

$$\diamond F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1$$

Якщо можливі значення ВВ належать інтервалу $[a; b]$, то

$$\diamond F(x) = 0 \text{ при } x \leq a$$

$$\diamond F(x) = 1 \text{ при } x \geq b$$

3. Функція розподілу – неспадна на всій числовій осі.

$$\diamond \text{Якщо } x_1 < x_2, \text{ то } F(x_1) \leq F(x_2)$$

4. Імовірність влучення значення неперервної ВВ у деякий інтервал дорівнює приросту функції розподілу на цьому інтервалі.

$$\diamond P(a \leq \xi < b) = F(b) - F(a)$$

Властивості саме неперервної ВВ

5. Імовірність того, що неперервна ВВ набуде одного певного значення, дорівнює нулю.

$$\diamond P(\xi = x_0) = 0$$

6. Для неперервної ВВ справедливо:

$$\diamond P(a \leq \xi \leq b) = P(a < \xi < b) = P(a < \xi \leq b) = F(b) - F(a)$$

7.2. Щільність розподілу неперервної випадкової величини

Зручно закон розподілу неперервних ВВ описувати за допомогою *щільності ймовірностей (щільності розподілу)*.

♣ **Щільність розподілу $p(x)$** неперервної випадкової величини – перша похідна від функції розподілу:

$$p(x) = F'(x).$$

[7.2]

Звідси: функція розподілу ВВ є первісною щільності розподілу цієї величини.

Як і функція розподілу $F(x)$, щільність імовірності $p(x)$ є однією з форм закону розподілу. На відміну від функції розподілу, щільність розподілу існує лише для неперервних ВВ.

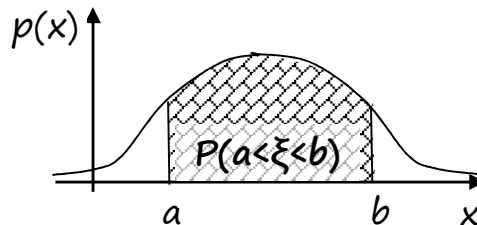
Властивості щільності розподілу неперервної ВВ

1. Щільність розподілу невід'ємна.

$$\diamond p(x) \geq 0 \text{ для усіх } x.$$

2. Імовірність влучення неперервної ВВ в інтервал (a, b) :

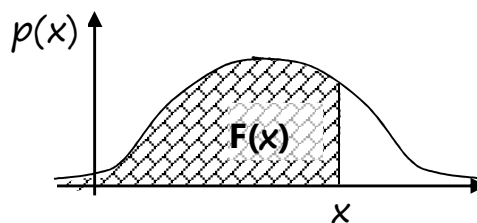
$$\diamond P(a < \xi < b) = \int_a^b p(x) dx$$



[7.3]

3. Функція розподілу дорівнює невластному інтегралу від щільності розподілу.

$$\diamond F(x) = \int_{-\infty}^x p(u) du$$



[7.4]

4. (Умова нормування). Невласний інтеграл із нескінченними межами від функції розподілу дорівнює 1.

$$\diamond \int_{-\infty}^{\infty} p(x) dx = 1$$

Якщо ВВ набуває значень лише з інтервалу (a, b) , то

$$\diamond \int_a^b p(x) dx = 1$$

♣ Графік щільності розподілу називають **кривою розподілу**.

7.3. Числові характеристики неперервних випадкових величин

7.3.1. Математичне очікування

♣ **Математичне очікування** $M\xi$ неперервної ВВ ξ , можливі значення якої належать відрізку $[a; b]$, — це визначений інтеграл

$$M\xi = \int_a^b x \cdot p(x) dx.$$

[7.5]

Якщо можливі значення належать усій осі Ox , то

$$M\xi = \int_{-\infty}^{\infty} x \cdot p(x) dx \quad [7.6]$$

(якщо цей інтеграл збігається абсолютно, тобто існує $\lim_{N \rightarrow \infty} \int_{-N}^N |x| \cdot p(x) dx$).

Так само як і для дискретної ВВ, математичне очікування неперервної ВВ – невипадкова (стала) величина.

7.3.2. Дисперсія, середнє квадратичне відхилення

♣ **Дисперсія $D\xi$** неперервної ВВ ξ – математичне очікування квадрата її відхилення.

Якщо можливі значення належать відрітку $[a; b]$:

$$D\xi = \int_a^b (x - M\xi)^2 \cdot p(x) dx. \quad [7.7]$$

Якщо можливі значення належать усій осі Ox :

$$D\xi = \int_{-\infty}^{\infty} (x - M\xi)^2 \cdot p(x) dx. \quad [7.8]$$

Для обчислень часто застосовують формули:

для $[a; b]$ –
$$D\xi = \int_a^b x^2 \cdot p(x) dx - (M\xi)^2, \quad [7.9]$$

для Ox –
$$D\xi = \int_{-\infty}^{\infty} x^2 \cdot p(x) dx - (M\xi)^2. \quad [7.10]$$

Усі властивості математичного очікування і дисперсії для дискретних ВВ справедливі і для неперервних величин.

♣ **Середнє квадратичне відхилення $\sigma(\xi)$** неперервної ВВ ξ – квадратний корінь із дисперсії:

$$\sigma(\xi) = \sqrt{D\xi}. \quad [7.11]$$

8. Основні закони розподілу неперервних випадкових величин

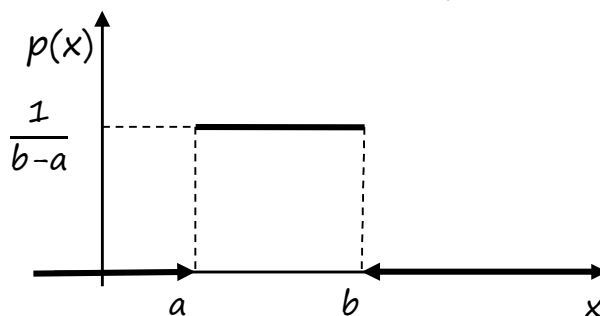
Класифікацію неперервних ВВ проводять за видом функції розподілу або щільності розподілу ймовірностей. Щільність розподілу неперервної ВВ називають також законом розподілу.

8.1. Рівномірний розподіл

♣ **Рівномірним** називають розподіл неперервної ВВ, усі значення якої належать деякому скінченному інтервалу $-\infty < a \leq \xi(\omega) \leq b < +\infty$, а щільність імовірності $p(x)$ стала на цьому інтервалі:

$$p(x) = \begin{cases} 0, & x < a \\ \frac{1}{b-a}, & a \leq x \leq b \\ 0, & x > b \end{cases} \quad [8.1]$$

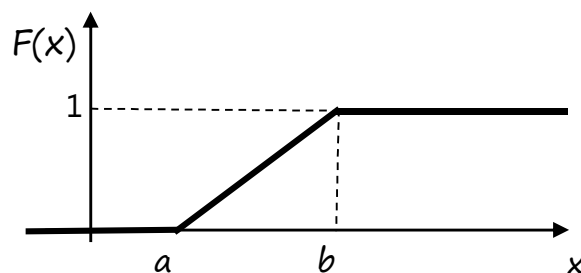
Графік щільності рівномірного розподілу має вигляд:



Функція розподілу ймовірності:

$$F(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x \geq b \end{cases} \quad [8.2]$$

Графік функції розподілу:



Математичне очікування ВВ, що має рівномірний розподіл:

$$M\xi = \frac{a+b}{2}. \quad [8.3]$$

Дисперсія ВВ, що має рівномірний розподіл:

$$D\xi = \frac{(b-a)^2}{12}. \quad [8.4]$$

Середнє квадратичне відхилення:

$$\sigma(\xi) = \frac{b-a}{2\sqrt{3}}. \quad [8.5]$$

Імовірність влучення ВВ ξ в інтервал (x_1, x_2) , що лежить усередині $[a, b]$:

$$P(x_1 < \xi < x_2) = \frac{x_2 - x_1}{b - a}.$$

8.2. Показовий (експоненціальний) розподіл

♣ Показовим (експоненціальним) називають розподіл неперервної випадкової величини, який описується щільністю:

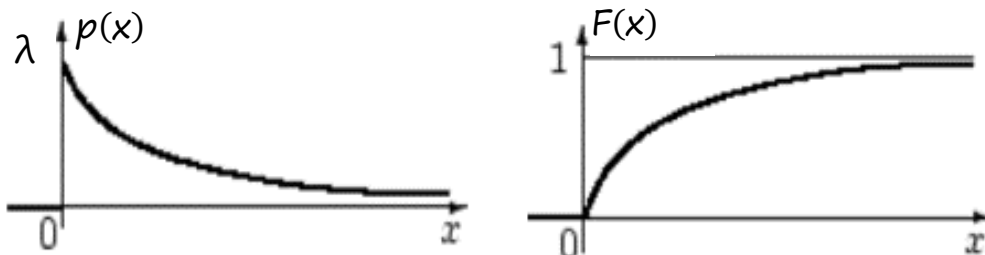
$$p(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0 \end{cases} \quad [8.6]$$

де $\lambda > 0$ – стала величина.

Функція розподілу показового закону має вигляд:

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0 \end{cases} \quad [8.7]$$

Графіки щільності і функції показового розподілу:



Математичне очікування показово розподіленої ВВ:

$$M\xi = \frac{1}{\lambda}. \quad [8.8]$$

Дисперсія ВВ, що має показовий розподіл:

$$D\xi = \frac{1}{\lambda^2}. \quad [8.9]$$

Середнє квадратичне відхилення:

$$\sigma = \frac{1}{\lambda}. \quad [8.10]$$

Імовірність влучення ВВ ξ в інтервал (x_1, x_2) :

$$P(x_1 < \xi < x_2) = e^{-\lambda x_1} - e^{-\lambda x_2}.$$

8.3. Нормальний розподіл

Нормальний розподіл грає надзвичайно важливу роль в теорії ймовірностей. До нього приблизно збігаються більшість інших розподілів за цілком певних умов. Із нього отримуються інші розподіли, широко використовувані в математичній статистиці.

Нормальний закон розподілу найчастіше зустрічається на практиці. Він проявляється завжди, коли ВВ є результатом дії великого числа різних незалежних випадкових чинників, кожен з яких окремо має малий вплив на значення ВВ.

♣ **Нормальним** називають **розподіл** імовірностей неперервної ВВ, що описується щільністю

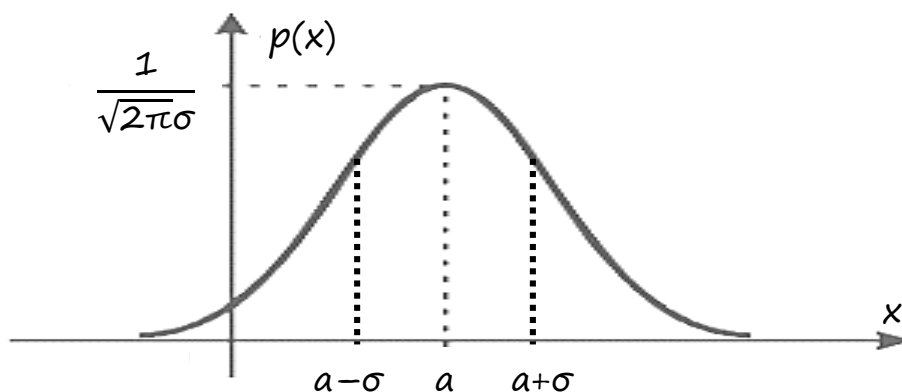
$$p(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}} \quad \forall x \in \mathbb{R}, a \in \mathbb{R}, \sigma > 0. \quad [8.11]$$

Нормальний закон розподілу також називають законом Гаусса.

8.3.1. Крива Гаусса

♣ Графік щільності нормального закону називають **нормальною кривою**, або **кривою Гаусса**.

Графік функції $p(x)$ (крива Гаусса) має вигляд:



Параметр a визначає положення на графіку центру щільності розподілу, а σ – розсіювання відносно центру.

8.3.2. Характеристики нормального розподілу

Випадкову величину, розподілену за нормальним законом, називають нормальною ВВ з параметрами a і σ . Достатньо знати значення цих параметрів, щоб задати нормальний розподіл.

Імовірнісний сенс параметрів: a – математичне очікування, σ – середнє квадратичне відхилення.

Математичне очікування нормального розподілу:

$$M\xi = a.$$

[8.12]

Дисперсія нормального розподілу:

$$D\xi = \sigma^2.$$

[8.13]

Середнє квадратичне відхилення нормального розподілу:

$$\sigma(\xi) = \sqrt{\sigma^2} = \sigma.$$

[8.14]

Нормальний розподіл із мат. очікуванням a і середнім квадратичним відхиленням σ позначають² $N(a, \sigma)$. Той факт, що ВВ ξ має такий нормальний розподіл, позначають $\xi \sim N(a, \sigma)$.

♣ Якщо $a=0$ і $\sigma=1$, то нормальний розподіл $N(0, 1)$ (і його криву Гаусса) називають **стандартним (нормованим)**. Щільність нормованого розподілу, яку називають **функцією Гаусса**:

$$p(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}. \quad [8.15]$$

8.3.3. Функція розподілу нормальної ВВ

Щоб знайти функцію розподілу нормальної ВВ $N(a, \sigma)$, підставимо її щільність у формулу властивості (7.4):

$$F(x) = \int_{-\infty}^x p(u) du = \frac{1}{\sqrt{2\pi} \cdot \sigma} \int_{-\infty}^x e^{-\frac{(u-a)^2}{2\sigma^2}} du = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^{\frac{x-a}{\sigma}} e^{-\frac{y^2}{2}} dy.$$

Отриманий інтеграл не «береться» (тобто не виражається через елементарні функції). Проте його можна розрахувати через спеціальну функцію Лапласа, для якої складені розрахункові таблиці.

♣ **Функція Лапласа, або інтеграл імовірностей:**

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{y^2}{2}} dy.$$

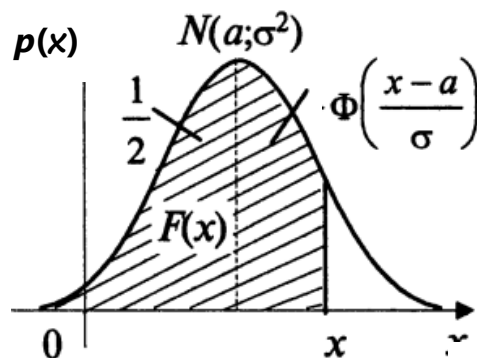
Таким чином,

♣ **функція розподілу нормальної ВВ $N(a, \sigma)$:**

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right). \quad [8.16]$$

Значення функції Лапласа $\Phi(z)$ лежать у межах $-1/2 \leq \Phi(z) \leq 1/2$. Це відповідає тому, що функція розподілу (як імовірність) має значення від 0 до 1. Функція $\Phi(z)$ непарна: $\Phi(-z) = -\Phi(z)$. Окрім того, $\Phi(z \geq 4) \approx 1/2$.

² Іноді в літературі зустрічається також позначення $N(a, \sigma^2)$.



Для стандартної нормальної випадкової величини $N(0, 1)$ функція розподілу:

$$F(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{y^2}{2}} dy = \frac{1}{2} + \Phi(x).$$

8.3.4. Імовірність влучення нормальної ВВ в заданий інтервал

Загальна формула ймовірності влучення ВВ в інтервал (x_1, x_2) : $P(x_1 \leq \xi \leq x_2) = F(x_2) - F(x_1)$.

Врахуємо у цій формулі вигляд функції $F(x)$ для нормальної ВВ.

Теорема. Імовірність влучення нормально розподіленої ВВ в інтервал $[x_1, x_2]$:

$$P(x_1 \leq \xi(\omega) \leq x_2) = \Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right). \quad [8.17]$$

Наслідок. Імовірність відхилення нормальної ВВ ξ з параметрами a і σ від її мат. очікування a :

$$P(|\xi - a| < \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right) \quad [8.18]$$

для будь-якого числа $\varepsilon > 0$.

Для стандартної нормальної ВВ формули (8.17) і (8.18) мають вигляд:

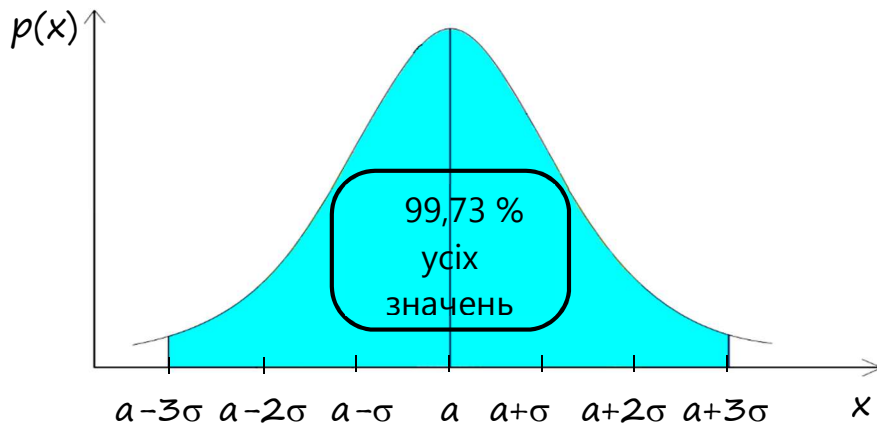
$$P(x_1 \leq \xi \leq x_2) = \Phi(x_2) - \Phi(x_1), \quad P(|\xi| < \varepsilon) = 2\Phi(\varepsilon).$$

8.3.5. Правило трьох сигм

Імовірність влучення нормальної ВВ в інтервал $a \pm 3\sigma$ за (8.17):

$$P(a - 3\sigma \leq \xi \leq a + 3\sigma) = 2 \cdot \Phi(3) \approx 2 \cdot 0,49865 = 0,9973.$$

Тобто інтервалу $a \pm 3\sigma$ відповідає 99,73 % площі під нормальною кривою.



Таким чином, імовірність події $A = \{\omega: |\xi - a| \leq 3\sigma\}$ близька до одиниці, тобто ця подія майже достовірна.

Імовірність протилежної події $P(|\xi - a| > 3\sigma)$ дуже мала ($1 - 0,9973 = 0,0027$). Це означає, що лише в 0,27 % випадків відхилення може бути більше ніж 3σ , тобто такі події практично неможливі. Це і є правило трьох сигм.

Правило трьох сигм. Для нормально розподіленої ВВ ξ імовірність події $|\xi - a| \leq 3\sigma$ близька до одиниці:

$$P(|\xi - a| \leq 3\sigma) \approx 1 - \varepsilon$$

[8.19]

(практично достовірно, що відхилення значень ВВ від її мат. очікування не перевищує потрібного середнього квадратичного відхилення).

На практиці це правило застосовують таким чином. Якщо розподіл деякої ВВ ξ невідомий, але виконана умова правила трьох сигм $|\xi - a| \leq 3\sigma$, то можна припустити, що ВВ розподілена нормально.

9. Граничні теореми теорії ймовірностей

9.1. Закон великих чисел

Якщо в однорідних умовах відбувається достатньо велике число випадкових подій, ряд їхніх характеристик практично втрачають випадковий характер. Їхній середній результат або їхню сукупну дію можна передбачати – з великою мірою надійності.

Це доводить ціла серія граничних теорем під загальною назвою *закон великих чисел*.

♣ **Закон великих чисел (ЗВЧ)** – низка математичних теорем, які загалом стверджують:

при великому числі випробувань їхні середні характеристики перестають бути випадковими і наближаються до деяких НЕвипадкових, сталих величин.

Кожна з теорем ЗВЧ встановлює, що середнє арифметичне n випадкових величин для достатньо великих n наближається до невинпадкової величини – за певних умов, тобто теореми різняться умовами обмежень на ВВ.

Для доведення цих теорем використовують нерівності Маркова і Чебишова.

9.2. Нерівності Маркова і Чебишова

9.2.1. Нерівність Маркова

Якщо існує $M|\xi|$, то

$$P(|\xi| \geq t) \leq \frac{M|\xi|}{t}. \quad [9.1]$$

9.2.2. Нерівність Чебишова

 **Нерівність Чебишова.** Якщо ВВ ξ має обмежені мат. очікування і дисперсію, то

$$P(|\xi - M\xi| \geq \varepsilon) \leq \frac{D\xi}{\varepsilon^2}. \quad [9.2]$$

Тобто ймовірність, що ξ відхилиться від її мат. очікування $M\xi$ за абсолютною величиною не менше ніж на довільне число $\varepsilon > 0$, не перевищує $\frac{D\xi}{\varepsilon^2}$.

Існує еквівалентна форма нерівності Чебишова – для протилежної події:

$$P(|\xi - M\xi| < \varepsilon) \geq 1 - \frac{D\xi}{\varepsilon^2}. \quad [9.3]$$

Якщо число ε достатньо мале, то (9.3) оцінює ймовірність того, що ξ набуває значень, близьких до свого математичного очікування.

Нерівність Чебишова має обмежене практичне застосування, оскільки часто дає грубу оцінку, але має дуже велике теоретичне значення.

9.3. Теореми закону великих чисел

Теореми закону великих чисел указують різні умови, при яких для послідовності ВВ $\xi_1, \xi_2, \dots, \xi_n, \dots$ (де n достатньо велике) ймовірність нерівності

$$\left| \frac{\xi_1 + \dots + \xi_n}{n} - \frac{M\xi_1 + \dots + M\xi_n}{n} \right| < \varepsilon$$

буде як завгодно близька до 1 для будь-якого малого числа $\varepsilon > 0$.

Таким чином, (при виконанні певних умов) як би сильно кожна ВВ не відхилялася від свого математичного очікування, середнє великого числа випадкових величин практично не відрізнятиметься від середнього їхніх математичних очікувань.

Умови виконання закону великих чисел:

- у формі *теореми Маркова*: якщо дисперсія суми n доданків зростає не занадто швидко зі зростанням n ;
- у формі *теореми Чебишова*: якщо дисперсії незалежних ВВ (однаково або по-різному розподілених) обмежені в сукупності;
- у формі *теореми Бернуллі*: якщо ВВ розподілена за біноміальним законом розподілу та інші.

9.3.1. Теорема Маркова

Теорема Маркова (закон великих чисел).

Якщо послідовність ВВ $\xi_1, \xi_2, \dots, \xi_n, \dots$ має скінченні математичні очікування і дисперсії та виконується яка-небудь одна з двох умов:

$$\text{або а) } \lim_{n \rightarrow \infty} \frac{D(\xi_1 + \dots + \xi_n)}{n^2} = 0, \quad [9.4]$$

$$\text{або б) } \xi_1, \xi_2, \dots, \xi_n \text{ незалежні і } \lim_{n \rightarrow \infty} \frac{D\xi_1 + \dots + D\xi_n}{n^2} = 0, \quad [9.5]$$

то для будь-якого числа $\varepsilon > 0$ виконується

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{\xi_1 + \dots + \xi_n}{n} - \frac{M\xi_1 + \dots + M\xi_n}{n}\right| < \varepsilon\right) = 1. \quad [9.6]$$

Теорема Маркова припускає, що ВВ $\xi_1, \xi_2, \dots, \xi_n, \dots$ можуть мати різні розподіли (і навіть за певних умов бути залежними).

9.3.2. Теорема Чебишова

Часто замість умови (9.5) простіше перевірити обмеженість дисперсії. Ця умова досліджена в теоремі Чебишова. Вона може бути доведена на підставі теореми Маркова, тобто є її наслідком.

Наслідок 1. Закон великих чисел Чебишова.

Якщо попарно незалежні ВВ $\xi_1, \xi_2, \dots, \xi_n, \dots$ мають скінченні математичні очікування й обмежені в сукупності дисперсії (тобто $D\xi_k = \sigma_k^2 \leq C < \infty$), то:

для будь-якого числа $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{\xi_1 + \dots + \xi_n}{n} - \frac{M\xi_1 + \dots + M\xi_n}{n}\right| < \varepsilon\right) = 1. \quad [9.7]$$

Таким чином, можна з великою часткою впевненості стверджувати, що при великих значеннях n величина $\frac{\xi_1 + \dots + \xi_n}{n}$ вже не є випадковою,

а дорівнює $\frac{M\xi_1 + \dots + M\xi_n}{n}$. А ймовірність тих подій, коли це невірно, практично дорівнює O .

На практиці часто ВВ мають один і той же розподіл, а отже, і однакові математичні очікування. Тоді теорема Чебишова набуває такого вигляду.

Наслідок 2. Окремий випадок закону великих чисел Чебишова.

Якщо $\xi_1, \xi_2, \dots, \xi_n, \dots$ – попарно незалежні, однаково розподілені ВВ з $M\xi_k = a$ і $D\xi_k = \sigma^2 < \infty$, то:
для будь-якого числа $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{\xi_1 + \dots + \xi_n}{n} - a\right| < \varepsilon\right) = 1. \quad [9.8]$$

9.3.3. Теорема Бернуллі

У схемі Бернуллі проводять n незалежних випробувань, у кожному з яких може відбутися або не відбутися подія A з однією і тією ж імовірністю p .

При достатньо великому n практично достовірно, що відносна частота події збігатиметься з імовірністю її появи в одному випробуванні. У цьому полягає закон великих чисел у вигляді теореми Бернуллі.

Наслідок 3. Теорема Бернуллі.

З імовірністю, скільки завгодно близькою до 1, можна стверджувати, що при необмеженому збільшенні числа випробувань n відносна частота події скільки завгодно мало відрізняється від її ймовірності.

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right) = 1, \quad [9.9]$$

де μ_n – число появи подій A в n випробуваннях;

p – імовірність появи події A в одному випробуванні;

ε – довільне додатне число.

Теорема Бернуллі надає спосіб з'ясувати ймовірність появи події A в одному випробуванні – якщо вона невідома. Для цього слід провести серію з n випробувань і розрахувати відносну частоту появи цієї події.

В умовах теореми Бернуллі нерівність Чебишова (9.2) стосовно відносної частоти $\frac{\mu_n}{n}$ має вигляд:

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

(для будь-якого числа $\varepsilon > 0$).

Закон великих чисел є одним з найважливіших застосувань теорії ймовірностей, оскільки «гарантує» стабільні результати для середніх значень деяких випадкових подій. Крім того, він дає гарантію, що можна розглядати не точно мат. очікування, а його оцінку. Це грає величезну роль у математичній статистиці.

9.4. Центральна гранична теорема

♣ **Центральна гранична теорема** – загальна назва групи теорем, що доводять твердження:

якщо випадкова величина η є сумою дуже великого числа взаємно незалежних ВВ, вплив кожної з яких на усю суму нікчемно малий, то η має розподіл, близький до нормального.

Центральна гранична теорема має порівняно високу точність навіть при $n \geq 30$. Вона виконується навіть при деяких (не дуже значних) відхиленнях від її умов. Тому вона широко застосовується в найрізноманітніших областях.

Різні форми центральної граничної теореми (ЦГТ) розглядають різні умови виконання цього твердження.



Теорема Ляпунова (центральна гранична теорема Ляпунова)

вказує певні умови, за якими для незалежних ВВ (довільних розподілів) $\xi_1, \xi_2, \dots, \xi_n$ буде виконуватися

$$\lim_{n \rightarrow \infty} P(\eta_n < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du, \quad [9.10]$$

$$\text{де } \eta_n = \frac{(\xi_1 + \dots + \xi_n) - (M\xi_1 + \dots + M\xi_n)}{\sqrt{\sigma_1^2 + \dots + \sigma_n^2}}.$$

Умови теореми Ляпунова гарантують, що усі складові суми η_n роблять рівномірно малий внесок у загальну суму.

На практиці дуже часто доводиться мати справу з випадковими величинами, які є сумами інших незалежних ВВ. На більшість реальних процесів у житті впливає безліч незалежних чинників. І хоча кожен з доданків може не підпорядковуватися нормальному закону, їхня сума – за деяких вельми загальних умов – має розподіл, близький до нормального. Сенс цих умов полягає в тому, що дія кожного доданка (випадкової величини) невелика в порівнянні з сумарною дією їх усіх.



Наслідок із теореми Ляпунова (теорема Гаусса). Для незалежних, однаково розподілених ВВ $\xi_1, \xi_2, \dots, \xi_n$, для яких $M\xi_k = a$, $D\xi_k = \sigma^2 < \infty$, виконується:

$$\lim_{n \rightarrow \infty} P(\eta_n < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du = \frac{1}{2} + \Phi(x), \quad [9.11]$$

$$\text{де } \eta_n = \frac{(\xi_1 + \dots + \xi_n)/n - a}{\sigma/\sqrt{n}},$$

$\Phi(x)$ – функція Лапласа (або інтеграл імовірностей).

Таким чином, граничним розподілом для η_n є нормальний розподіл із нульовим мат. очікуванням та одиничною дисперсією $N(0, 1)$.

9.5. Інтегральна теорема Муавра – Лапласа

Інтегральна теорема Муавра – Лапласа є наслідком центральної граничної теореми для повторних випробувань.

Нехай кожна з дискретних ВВ $\xi_1, \xi_2, \dots, \xi_n$ набуває одного з двох можливих значень 1 і 0 з імовірностями відповідно p і q . Сума цих ВВ, випадкова величина η_n , – число появи події в n випробуваннях.

Числові характеристики η_n : $M\eta_n = np$, $D\eta_n = npq$, $\sigma(\eta_n) = \sqrt{npq}$.

На підставі ЦГТ при $n \rightarrow \infty$ розподіл η_n наближається до нормального. Тоді за формулою (8.17) (при $n \rightarrow \infty$):

$$P(x_1 \leq \eta_n \leq x_2) = \Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right).$$

Підставивши до неї характеристики η_n , отримаємо формулювання інтегральної теореми Муавра – Лапласа:

$$P_n(k_1 \leq \eta_n \leq k_2) \approx \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right). \quad [9.12]$$

МАТЕМАТИЧНА СТАТИСТИКА

Математична статистика розробляє методи отримання висновків зі статистичних даних (результатів спостережень або вимірювань) на підставі методів теорії ймовірності. Мета – виявити закономірності явищ і їхній взаємозв'язок.

Також мат. статистика займається побудовою ймовірнісних моделей і перевіркою їхньої адекватності.

Основні задачі математичної статистики:

- вказати способи систематизації даних (їхнього збору, угруповання);
- визначити закон розподілу випадкової величини;
- оцінити невідомі параметри розподілу;
- оцінити залежність випадкової величини від однієї або декількох інших випадкових величин;
- перевірити правдоподібність гіпотез – про закон розподілу, про значення оцінюваного параметра, про форму зв'язку між випадковими величинами та ін.

10. Основи статистичних методів

10.1. Генеральна і вибіркова сукупності

♣ Множину однорідних об'єктів, які вивчають з точки зору їхнього розподілу за деякою ознакою, називають **генеральною сукупністю**.

♣ **Вибіркова сукупність**, або **вибірка**, – сукупність об'єктів, випадково відібраних із генеральної сукупності.

♣ **Обсяг вибірки** – число елементів у вибірковій сукупності.

Обсяг генеральної сукупності позначають N , вибірки – n . Об'єкти вибірки зазвичай позначають $x_1, x_2, \dots, x_n$.

Можна розглядати генеральну сукупність як деяку випадкову величину X . При цьому випробуванням буде – випадковий вибір об'єкта із сукупності, а значенням $x_1, x_2, \dots, x_n$ цієї ВВ – значення досліджуваної ознаки.

10.2. Варіаційні ряди

♣ Спостережувані (числові) значення $x_1, x_2, \dots, x_n$ (елементи вибірки) називають **варіантами**.

♣ Послідовність варіантів, записаних у зростаючому порядку, називається простим **варіаційним рядом**.

♣ **Частотою** n_k деякого варіанта називають кількість його повторень у вибірці.

♣ Відношення частоти n_k до обсягу вибірки n називають **відносною частотою** (позначають n_k^*): $n_k^* = \frac{n_k}{n}$.

♣ Сукупність впорядкованих значень x_k і відповідних їм частот n_k у вигляді таблиці (матриці) називають **дискретним варіаційним рядом**:

значення X	x_1	x_2	$\dots$	x_m
частоти n_k	n_1	n_2	$\dots$	n_m

Якщо обсяг вибірки n великий, а кожен варіант зустрічається рідко, то увесь діапазон значень вибірки розбивають на інтервали, що не перетинаються. Так само роблять, якщо випадкова величина X неперервна. Тоді одержують інтервальний (неперервний) варіаційний ряд.

♣ **Інтервальний варіаційний ряд** – таблиця, де кожному інтервалу значень вибірки $[\alpha_{k-1}, \alpha_k)$ ставиться у відповідність кількість n_k значень, що потрапили до нього:

інтервали $[\alpha_{k-1}, \alpha_k)$	$[\alpha_0, \alpha_1)$	$[\alpha_1, \alpha_2)$	$\dots$	$[\alpha_{m-1}, \alpha_m)$
частоти n_k	n_1	n_2	$\dots$	n_m

Зазвичай інтервали мають однакову довжину.

Для побудови інтервального варіаційного ряду за вибіркою:

1) визначають **розмах варіації** – різницю між максимальним і мінімальним значеннями вибірки: $R = x_{\max} - x_{\min}$;

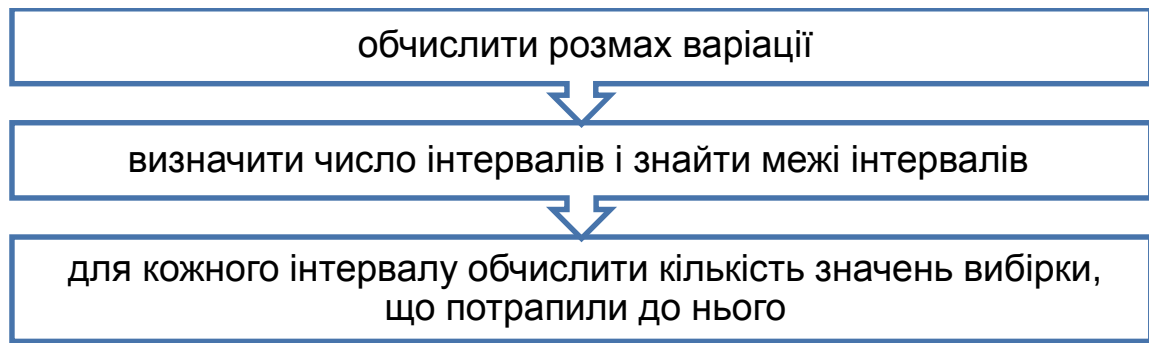
2) задають число інтервалів (наприклад, $m = 1 + \log_2 n = 1 + 3,322 \lg n$).

Довжина інтервалів дорівнює $h = \frac{R}{m}$;

3) обчислюють початок першого інтервалу: $\alpha_0 = x_{\min} - \frac{h}{2}$ і наступних: $\alpha_k = \alpha_{k-1} + h$ ($k = \overline{1, m}$);

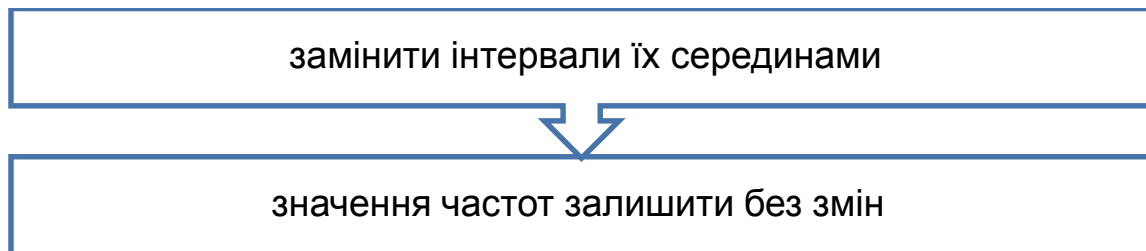
4) для кожного інтервалу підраховують кількість значень вибірки, що потрапили до нього. Якщо значення потрапило на межу двох інтервалів, його відносять до попереднього інтервалу.

Алгоритм побудови інтервального ряду за вибіркою



Для переходу від інтервального варіаційного ряду до дискретного потрібно замінити інтервали числами – серединами цих інтервалів $x_k = \frac{\alpha_{k-1} + \alpha_k}{2}$ (а значення частот залишити без змін).

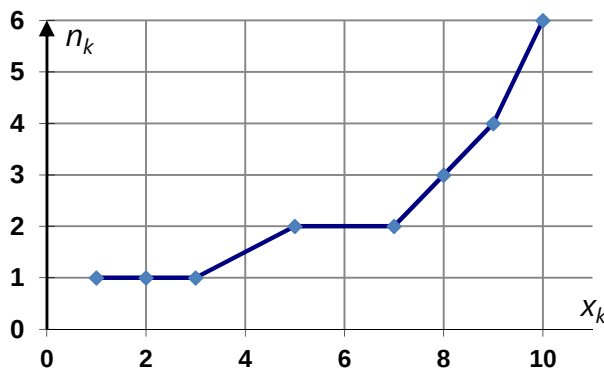
Алгоритм побудови дискретного ряду за інтервальним



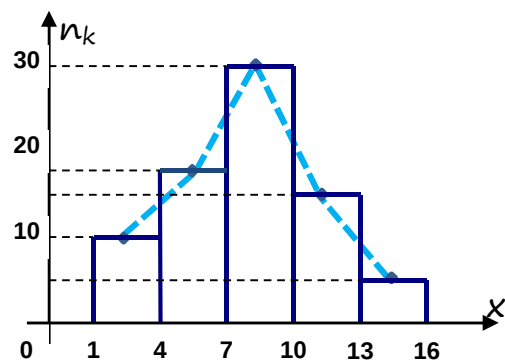
Зазвичай аналіз починають з побудови полігону (багатокутника) частот або гістограми.

♣ **Полігон частот** (*багатокутник частот*) – ламана, сполучаюча точки $(x_1, n_1), (x_2, n_2), \dots, (x_m, n_m)$.

♣ **Гістограма** – ступінчаста фігура, складена із прямокутників, основою яких є інтервали $[\alpha_{k-1}, \alpha_k)$, а висоти дорівнюють відповідним частотам n_k .



а)



б)

Рис. 10.1. а) полігон дискретного варіаційного ряду;
б) полігон і гістограма інтервального варіаційного ряду

Гістограму застосовують для відображення інтервального ряду. Полігон – і для дискретного, і для інтервального варіаційних рядів (для інтервального ряду лініями сполучають точки, значення $x_1, x_2, \dots, x_n$ яких – середини інтервалів).

Графічне зображення варіаційного ряду дозволяє уявити приблизно закон розподілу випадкової величини.

10.3. Емпірична функція розподілу

♣ **Емпірична функція розподілу** – функція $F_n(x)$, що показує для кожного x відносну частоту тих спостережень, у яких значення вибірки менше x .

$$F_n(x) = \frac{n_x^{(H)}}{n}, \quad \text{де } n - \text{обсяг вибірки, } n_x^{(H)} = \sum_{x_k < x} n_k - \text{накопичена частота}$$

(сума частот значень вибірки, менших ніж x).

Для дискретного варіаційного ряду

X	x_1	x_2	...	x_m
n	n_1	n_2	...	n_m

емпірична функція розподілу:

$$F_n(x) = \begin{cases} 0 & \text{при } x \leq x_1 \\ \frac{n_1}{n} & \text{при } x_1 < x \leq x_2 \\ \frac{n_1 + n_2}{n} & \text{при } x_2 < x \leq x_3 \\ \frac{n_1 + n_2 + n_3}{n} & \text{при } x_3 < x \leq x_4 \\ \dots\dots\dots & \\ \frac{n_1 + n_2 + \dots + n_m}{n} = 1 & \text{при } x > x_m \end{cases}$$

Для інтервального ряду обчислюють значення $F_n(x)$ на кінцях інтервалів. Щоб побудувати графік, відображають ці точки (накопичені частоти відповідають верхнім межах інтервалів) і з'єднують їх прямими (іноді кривими) лініями.

♣ Функцію розподілу $F(x)$ генеральної сукупності, що визначає ймовірність події $\xi < x$, називають **теоретичною функцією розподілу**.

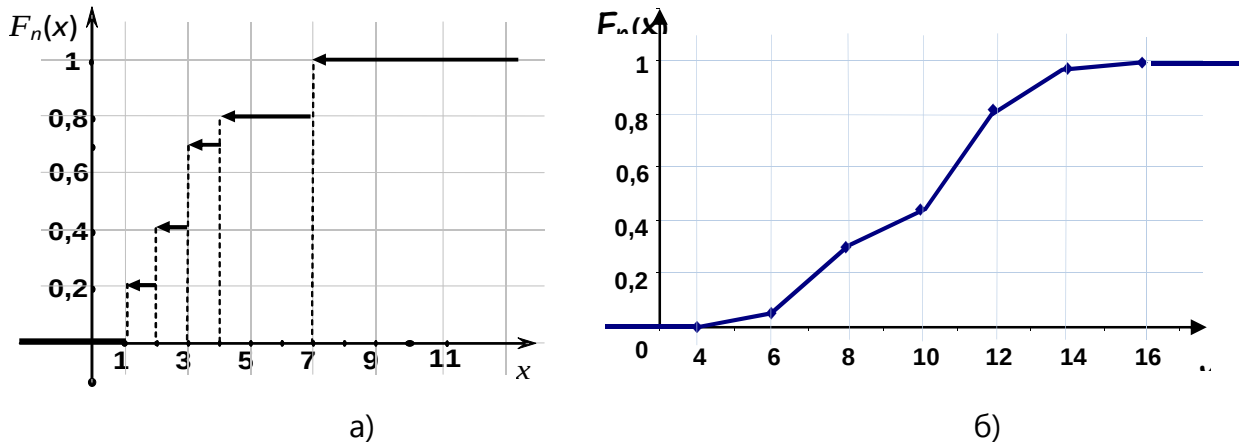


Рис. 10.2. Емпірична функція розподілу: а) дискретного, б) інтервального ряду

За теоремою Бернуллі, можна використовувати емпіричну функцію розподілу $F_n(x)$ стандартної вибірки для наближеного подання теоретичної функції $F(x)$.

Властивості емпіричної функції $F_n(x)$ співпадають із властивостями теоретичної функції розподілу $F(x)$:

1. Значення $F_n(x)$ належать інтервалу $[0; 1]$: $0 \leq F_n(x) \leq 1$.
2. $F_n(x)$ – неспадна функція, тобто з $x_1 \leq x_2$ випливає, що $F_n(x_1) \leq F_n(x_2)$.
3. Якщо x_1 – найменший варіант, то $F_n(x) = 0$ при $x \leq x_1$.

Якщо x_m – найбільший варіант, то $F_n(x) = 1$ при $x > x_m$.

4. $F_n(x)$ неперервна зліва: $\lim_{x \rightarrow x_0 - 0} F_n(x) = F_n(x_0)$.

10.4. Операція «риса»

Операція «риса»:

$$\bar{x} = \frac{1}{n} \sum_{k=1}^n x_k$$

За її допомогою можна скоротити запис виразів:

$$\overline{x+y} = \frac{1}{n} \sum_{k=1}^n (x_k + y_k),$$

$$\overline{x^2} = \frac{1}{n} \sum_{k=1}^n x_k^2,$$

$$\overline{xy} = \frac{1}{n} \sum_{k=1}^n x_k y_k,$$

$$\overline{(x-\bar{x})^2} = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2.$$

Властивості операції «риса».

$$1) \overline{C} = C, \text{ де } C - \text{ стала.}$$

$$2) \overline{Cx} = C\bar{x}.$$

$$3) \overline{\bar{x}} = \bar{x}.$$

$$4) \overline{x+y} = \bar{x} + \bar{y}.$$

$$5) |\overline{xy} - \bar{x}\bar{y}| \leq \sqrt{\overline{(x-\bar{x})^2} \cdot \overline{(y-\bar{y})^2}}.$$

10.5. Числові характеристики вибірки

Основні числові характеристики вибірки: вибіркове середнє, вибіркова дисперсія, вибіркове середнє квадратичне відхилення.

♣ **Вибіркове середнє** $\bar{x}$ – середнє арифметичне варіант із урахуванням їхніх частот.

Якщо дані у вигляді дискретного варіаційного ряду $x_1, x_2, \dots, x_n$ і усі значення різні, то вибіркове середнє:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{k=1}^n x_k. \quad [10.1]$$

Якщо значення $x_1, x_2, \dots, x_m$ мають частоти $n_1, n_2, \dots, n_m$, то вибіркове середнє:

$$\bar{x} = \frac{x_1 n_1 + x_2 n_2 + \dots + x_m n_m}{n} = \frac{1}{n} \sum_{k=1}^m x_k n_k. \quad [10.2]$$

Якщо дані у вигляді інтервального варіаційного ряду, то для обчислення вибіркового середнього або дисперсії необхідно спочатку перейти до дискретного варіаційного ряду.

♣ **Вибіркова дисперсія** (дисперсія варіаційного ряду) – середнє арифметичне квадратів відхилень варіант від їхньої вибіркової середньої $\bar{x}$:

$$S_x^2 = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2. \quad [10.3]$$

Якщо значення $x_1, x_2, \dots, x_m$ мають частоти $n_1, n_2, \dots, n_m$, то вибіркова дисперсія:

$$S_x^2 = \frac{1}{n} \sum_{k=1}^m (x_k - \bar{x})^2 n_k. \quad [10.4]$$

На практиці для обчислення S_x^2 використовують таку формулу:

$$S_x^2 = \frac{1}{n} \sum_{k=1}^n x_k^2 - \left(\frac{1}{n} \sum_{k=1}^n x_k \right)^2 = \overline{x^2} - (\bar{x})^2. \quad [10.5]$$

♣ **Вибіркове середнє квадратичне відхилення** S_x – арифметичне значення квадратного кореня з вибіркової дисперсії:

$$S_x = \sqrt{S_x^2}.$$

Вибіркові середнє, дисперсія, середнє квадратичне відхилення залежать від значень вибірки, тому вони самі є випадковими величинами.

11. Оцінки параметрів розподілу

Одна з типових задач мат. статистики – на підставі вибіркової сукупності отримати наближені значення параметрів її закону розподілу, загальний вигляд якого відомий (з яких-небудь міркувань).

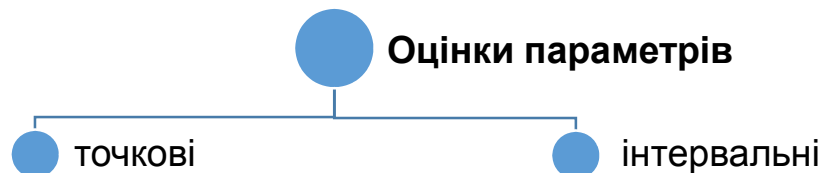
Оскільки для багатьох законів розподілу ВВ їхні параметри виражаються через числові характеристики цієї ВВ (математичне очікування, дисперсія, середнє квадратичне відхилення та ін.), то проводять оцінку цих числових характеристик.

11.1. Поняття про оцінку параметрів. Види оцінок

Нехай для функції $F(x)$ розподілу генеральної сукупності відомий її вигляд, але невідомі параметри $\theta_1, \theta_2, \dots, \theta_m$, що її визначають: $F(x, \theta_1, \theta_2, \dots, \theta_m)$. Тоді виникає задача оцінки цих невідомих параметрів за вибіркою.

♣ **Оцінка $\hat{\theta}$ параметра θ** – його наближене значення, залежне від вибірових даних $x_1, x_2, \dots, x_n$.

Розрізняють два види статистичних оцінок параметрів – точкові та інтервальні.



♣ **Точковою** називають **оцінку**, яка визначається одним числом $\hat{\theta}$.

Знайти оцінку $\hat{\theta}$ – означає знайти функцію $\hat{\theta}_n(x_1, x_2, \dots, x_n)$, яка при підстановці до неї результатів спостереження $x_1, x_2, \dots, x_n$ і дає наближене значення θ .

Такі функції (залежні від результатів спостережень) називають статистичними оцінками, або *статистиками*. Оцінка $\hat{\theta}_n$ є випадковою величиною, оскільки залежить від випадкових значень.

Якщо вибірка малого обсягу, точкова оцінка може суттєво відрізнятися від оцінюваного параметра. Тоді доцільно використовувати інтервальні оцінки.

♣ **Інтервальною** називають **оцінку**, яка визначена двома числами θ_1 та θ_2 – кінцями інтервалу, що містить із заданою ймовірністю γ оцінюваний параметр θ .

♣ Інтервал (θ_1, θ_2) , що містить справжнє значення параметра θ з заданою ймовірністю γ , називають **довірчим інтервалом** для параметра θ : $P(\theta_1 < \theta < \theta_2) = \gamma$.

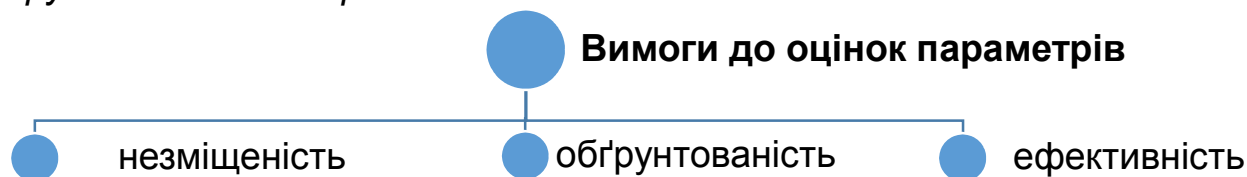
♣ Ймовірність γ називають **довірчою ймовірністю** (рівнем довіри, надійністю оцінки), а число $\alpha = 1 - \gamma$ – **рівнем значущості**.

Зазвичай довірчу ймовірність γ задають заздалегідь числом, близьким до одиниці, – найчастіше використовують значення 0,90; 0,95; 0,99; 0,999.

11.2. Точкові оцінки параметрів

Часто існує декілька статистик для оцінки одного і того ж параметра. Задача – вибрати таку функцію $\hat{\theta}_n$, щоб отримувані оцінки були найбільш близькі до значення θ , а зі збільшенням обсягу вибірки точність оцінки зростала. Тому для вибору найкращої оцінки висувають низку вимог.

Основні властивості, що характеризують якість оцінок: *незміщеність*, *обґрунтованість* й *ефективність*.



♣ Оцінку $\hat{\theta}$ називають **незміщеною**, якщо її мат. очікування дорівнює оцінюваному параметру θ при будь-якому обсязі вибірки:

$$M\hat{\theta} = \theta. \quad [11.1]$$

Якщо рівність не виконується, оцінку називають **зміщеною**, а різницю $M\hat{\theta} - \theta$ називають **зміщенням** або **систематичною похибкою**.

♣ Оцінку $\hat{\theta}$ називають **обґрунтованою**, якщо при необмеженому збільшенні числа спостережень n вона наближається до параметра θ з ймовірністю, скільки завгодно близькою до одиниці:

$$\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| < \varepsilon) = 1, \quad \forall \varepsilon > 0.$$

♣ Оцінку $\hat{\theta}$ називають **ефективною**, якщо вона має найменшу дисперсію серед усіх інших незміщених оцінок параметра θ , обчислених за вибіркою того ж обсягу n .

При здобутті оцінок параметрів прагнуть до того, щоб вони були незміщеними, обґрунтованими й ефективними. Але на практиці не завжди вдається задовольнити усі ці три умови одночасно.

11.2.1. Оцінка математичного очікування і дисперсії за вибіркою

З Твердження 1. Вибіркове середнє

$$\bar{x} = \frac{1}{n} \sum_{k=1}^n x_k \quad [11.2]$$

– незміщена, обґрунтована оцінка математичного очікування: $\widehat{MX} = \bar{x}$.

Ефективність оцінки $\bar{x}$ залежить від вибору закону розподілу.

З Твердження 2. Якщо ВВ має нормальний розподіл, то оцінка $\bar{x}$ є ефективною: вона має серед усіх незміщених оцінок мат. очікування мінімальну дисперсію, що дорівнює $\frac{\sigma^2}{n}$.

З Твердження 3. Вибіркова дисперсія

$$S_x^2 = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2 \quad [11.3]$$

– обґрунтована, але зміщена оцінка дисперсії: $\widehat{DX} = S_x^2$.

Проте при $n \rightarrow \infty$ оцінка стає незміщеною: $\lim_{n \rightarrow \infty} MS_x^2 = \sigma^2$. Такі оцінки називають *асимптотично незміщеними*.

Щоб усунути зміщення, можна скоректувати оцінку, запровадивши поправку $\frac{n}{n-1}$. Тоді отримаємо виправлену оцінку $\widetilde{S}_x^2$, яку називають **виправленою дисперсією**:

$$\widetilde{S}_x^2 \stackrel{\text{def}}{=} \frac{n}{n-1} S_x^2 = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2. \quad [11.4]$$

Виправлену дисперсію використовують при невеликих вибірках. При $n > 50$ практично немає різниці між S_x^2 і $\widetilde{S}_x^2$.

І вибірка дисперсія, і виправлена дисперсія – обґрунтовані оцінки. У загальному випадку вони не є ефективними, проте є асимптотично ефективними.

Окрім розглянутих, існує безліч інших оцінок математичного очікування, дисперсії та інших параметрів, а також різні методи знаходження цих оцінок.

11.2.2. Метод максимальної правдоподібності

Ідея методу максимальної правдоподібності – набути за оцінку $\hat{\Theta}$ параметра Θ таке його значення, при якому ймовірність отримати вже наявну вибірку максимальна.

Для визначення параметрів знаходять максимум деякої функції, що задає цю ймовірність.

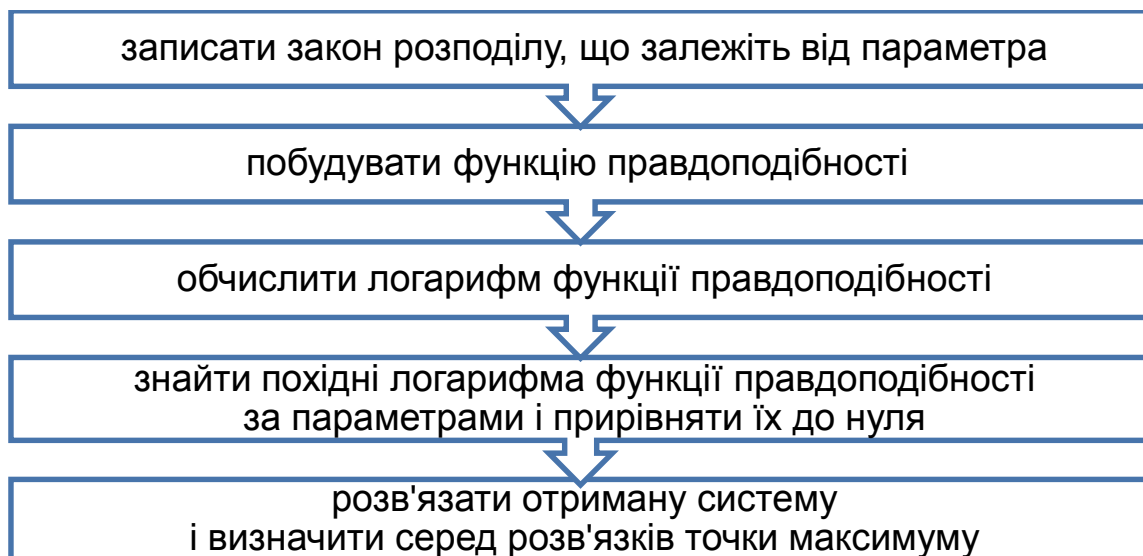
♣ **Функція правдоподібності** – функція, що визначає ймовірність отримання в незалежних випробуваннях саме значень вибірки $x_1, x_2, \dots, x_n$:
 $L(\Theta, x_1, x_2, \dots, x_n) = P_{x_1}(\Theta) \cdot P_{x_2}(\Theta) \cdot \dots \cdot P_{x_n}(\Theta)$ – для дискретної ВВ;
 $L(\Theta, x_1, x_2, \dots, x_n) = p(x_1, \Theta) \cdot p(x_2, \Theta) \cdot \dots \cdot p(x_n, \Theta)$ – для неперервної ВВ.
 [11.5]

Знаходять значення Θ , при якому функція правдоподібності досягає максимуму, – це значення і приймають за оцінку $\hat{\Theta}$. Для цього прирівнюють до нуля похідну функції, розв'язують це рівняння і досліджують отримані критичні точки на екстремум:

$$\max_{\Theta} L(\Theta, x_1, x_2, \dots, x_n): \quad \frac{dL}{d\Theta} = 0 \quad \Rightarrow \quad \hat{\Theta} = \hat{\Theta}(x_1, x_2, \dots, x_n).$$

Часто для спрощення розрахунків знаходять максимум не самої функції, а її логарифма $\ln L$ (оскільки вони досягають максимуму в одній точці).

Метод максимальної правдоподібності



Оцінки, отримані методом максимальної правдоподібності, є обґрунтованими й ефективними (або асимптотично ефективними), але можуть бути як зміщеними, так і незміщеними.

11.2.3. Метод моментів

Згідно з методом моментів, теоретичні числові характеристики ВВ виражають через шуканий параметр Θ , а потім прирівнюють їх до відповідних емпіричних характеристик (залежних від вибірових даних). Розв'язання отриманої системи рівнянь і будуть оцінками параметрів розподілу $\hat{\Theta}$.

Як такі числові характеристики використовують *моменти*.

Теоретичні моменти

♣ **Початковий момент** v_m порядку m випадкової величини ξ – це мат. очікування величини ξ^m :

$$v_m = M\xi^m \quad m=1, 2, \dots \quad [11.6]$$

За визначенням мат. очікування, формули початкових моментів:

для дискретної ВВ:

$$v_m = \sum_{k=1}^n x_k^m \cdot p_k. \quad [11.7]$$

для неперервної ВВ:

$$v_m = \int_{-\infty}^{\infty} x^m \cdot p(x) dx. \quad [11.8]$$

Початковий момент 1-го порядку ($m=1$) збігається з мат. очікуванням ВВ: $v_1 = M\xi$.

♣ **Центральний момент** μ_m порядку m випадкової величини ξ – це мат. очікування величини $(\xi - M\xi)^m$:

$$\mu_m = M(\xi - M\xi)^m, \quad m=1, 2, \dots$$

Центральний момент 2-го порядку збігається з дисперсією ВВ:

$$\mu_2 = M(\xi - M\xi)^2 = D\xi.$$

Для будь-якої ВВ центральний момент 1-го порядку дорівнює нулю:

$$\mu_1 = M(\xi - M\xi) = 0.$$

Центральні моменти можна виразити через початкові. Наприклад:

$$\mu_2 = v_2 - v_1^2.$$

Емпіричні моменти

Емпіричні моменти як функції від стандартної вибірки $x_1, x_2, \dots, x_n$.

♣ **Початковий емпіричний момент** порядку m :

$$v_m^* = \overline{x^m} = \frac{1}{n} \sum_{k=1}^n x_k^m. \quad [11.9]$$

При $m=1$ початковий емпіричний момент збігається з вибіркоvim середнім: $v_1^* = \bar{x}$.

Через початкові емпіричні моменти можна виразити вибіркoву дисперсію: $S_x^2 = v_2^* - (v_1^*)^2$.

♣ **Центральний емпіричний момент порядку m :**

$$\mu_m^* = \overline{(x - \bar{x})^m} = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^m. \quad [11.10]$$

При $m=1$ центральний емпіричний момент дорівнює нулю: $\mu_1^* = \overline{(x - \bar{x})} = 0$, а при $m=2$ збігається з вибірковою дисперсією: $S_x^2 = \mu_2^* = \overline{(x - \bar{x})^2}$.

Опис методу моментів

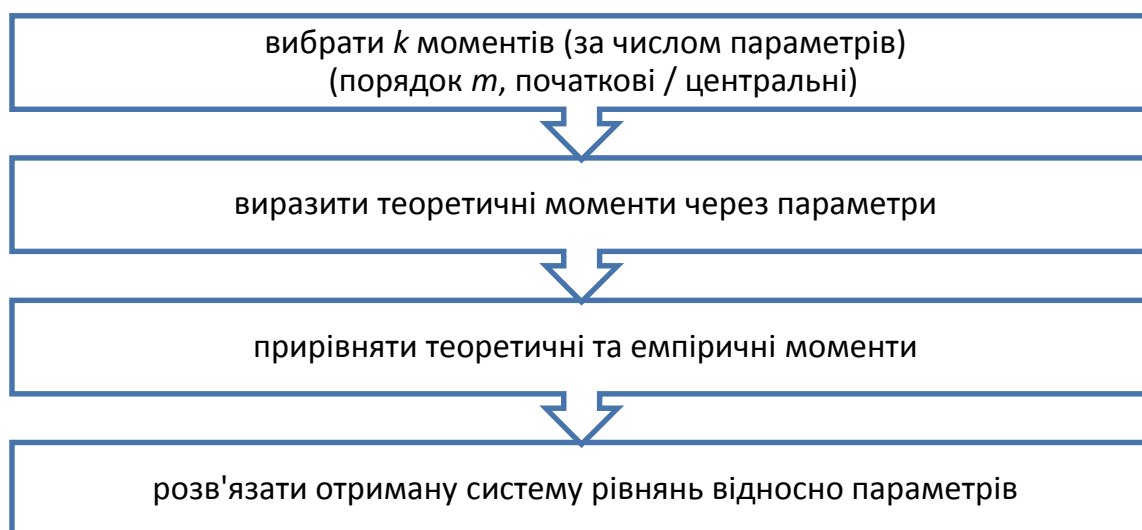
Прирівняємо k емпіричних моментів до відповідних теоретичних. Отримаємо систему k рівнянь із k невідомими – параметрами $\theta_1, \dots, \theta_k$:

$$\begin{cases} v_1(\theta_1, \dots, \theta_k) = v_1^*(x_1, \dots, x_n), \\ \dots \\ v_k(\theta_1, \dots, \theta_k) = v_k^*(x_1, \dots, x_n). \end{cases} \quad [11.11]$$

Число рівнянь має дорівнювати числу параметрів розподілу. Які саме моменти включати до системи, залежить від складності функцій $v_k(\theta_1, \dots, \theta_k)$ і поставлених задач.

Розв'язавши систему (11.11), отримаємо оцінки $\hat{\theta}_1, \dots, \hat{\theta}_k$ як функції від значень вибірки $x_1, \dots, x_n$.

Метод моментів



У методі моментів часто розглядають лише початкові моменти, іноді використовують центральні або сукупність початкових і центральних.

Оцінки, отримані методом моментів, обґрунтовані. Незміщеність цих оцінок не гарантована (зазвичай вони лише асимптотично незміщені). Найчастіше ці оцінки не є ефективними і навіть асимптотично ефективними.

11.3. Інтервальні оцінки параметрів

11.3.1. Довірчі інтервали

Точкові оцінки є випадковими величинами, тому виникає питання про їхню точність і надійність.

Оцінка $\hat{\theta}$ тим ближче до визначуваного параметра θ , чим менша їхня абсолютна різниця $|\hat{\theta} - \theta|$. Таким чином, якщо $|\hat{\theta} - \theta| < \delta$, то чим менше $\delta > 0$, тим оцінка $\hat{\theta}$ точніша.

Ця нерівність виконуватиметься лише з деякою ймовірністю, її вважають надійністю оцінки.

♣ **Довірча ймовірність** (надійність оцінки) $\hat{\theta}$ параметра θ – ймовірність

$$\gamma = P(|\hat{\theta} - \theta| < \delta) \quad [11.12]$$

Довірчу ймовірність γ задають заздалегідь, зазвичай близькою до одиниці: 0,9, 0,95, 0,99, 0,999. Тоді великі (за абсолютною величиною) помилки з'являтимуться лише з малою ймовірністю $\alpha = 1 - \gamma$.

♣ Число $\alpha = 1 - \gamma$ називають **рівнем значущості**.

Довірчу ймовірність можна записати у вигляді ймовірності потрапляння в інтервал $(\hat{\theta} - \delta, \hat{\theta} + \delta)$.

♣ Інтервал (θ_1, θ_2) називають **довірчим інтервалом** для параметра θ , якщо він містить цей параметр із заздалегідь заданою довірчою ймовірністю γ :

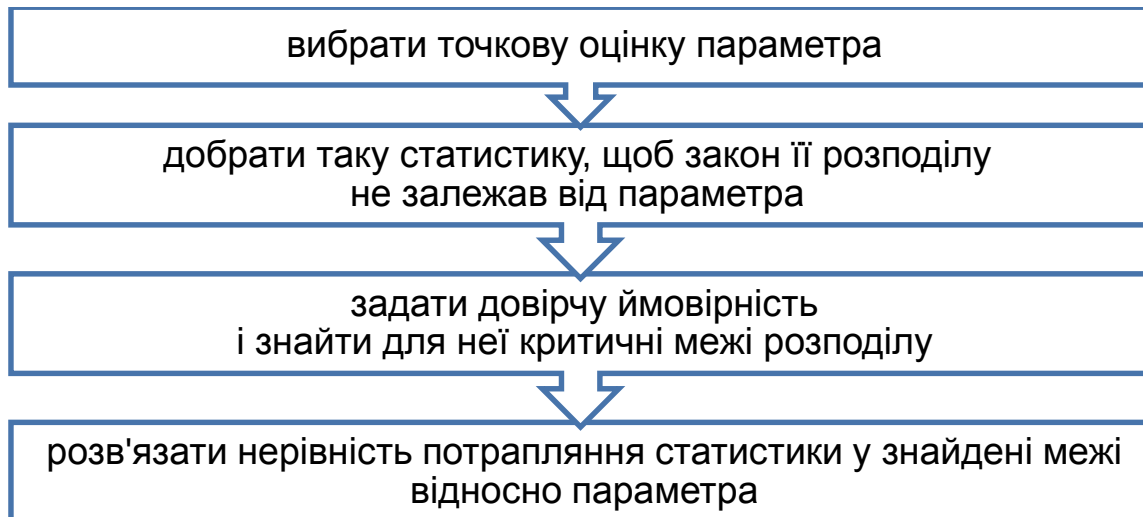
$$P(\theta_1 < \theta < \theta_2) = \gamma. \quad [11.13]$$

Один із підходів до побудови довірчих інтервалів:

- 1) вибирають точкову оцінку $\hat{\theta}(x_1, x_2, \dots, x_n)$ параметра θ ;
- 2) добирають деяку неперервну випадкову величину (*статистику*) – функцію $Y(\hat{\theta}, \theta)$ так, щоб закон її розподілу $p(x)$ не залежав від θ ;

- 3) задавшись довірчою ймовірністю γ , знаходять для цього розподілу *критичні межі* a і b (значення, для яких функція розподілу набуває значень відповідно $\frac{1}{2} - \frac{\gamma}{2}$ і $\frac{1}{2} + \frac{\gamma}{2}$);
- 4) розв'язують нерівність $a < Y(\hat{\theta}, \theta) < b$ відносно θ (тобто виражають θ з $Y(\hat{\theta}, \theta)$) і тим самим отримують довірчий інтервал $\theta_1 < \theta < \theta_2$, для якого $P(\theta_1 < \theta < \theta_2) = \gamma$.

Пошук довірчого інтервалу



Довірчий інтервал визначається однозначно у разі вибірки з нормального закону розподілу.

11.3.2. Інтервальна оцінка математичного очікування при відомій дисперсії

Нехай $x_1, x_2, \dots, x_n$ – стандартна вибірка з ВВ X , розподіленої за нормальним законом із невідомим математичним очікуванням і відомою дисперсією σ^2 .

Для знаходження інтервалу для мат. очікування використовують:

- вибіркове середнє $\bar{x}$ – як точкову оцінку мат. очікування;
- статистику $\frac{\bar{x} - a}{\sigma/\sqrt{n}}$ – як величину зі стандартним нормальним розподілом, що не залежить від параметра a (мат. очікування 0 і дисперсія 1).

Довірчий інтервал, який містить невідоме математичне очікування $MX = a$ з довірчою ймовірністю γ :

$$\left(\bar{x} - \varepsilon \frac{\sigma}{\sqrt{n}}, \bar{x} + \varepsilon \frac{\sigma}{\sqrt{n}} \right).$$

Значення ε знаходять з умови $\Phi(\varepsilon) = \gamma/2$ за таблицями функції Лапласа чи за допомогою комп'ютера.

11.3.3. Інтервальна оцінка математичного очікування при невідомій дисперсії

Нехай $x_1, x_2, \dots, x_n$ – стандартна вибірка з ВВ X , розподіленої за нормальним законом із невідомими математичним очікуванням і дисперсією.

Для знаходження інтервалу для мат. очікування використовують:

- вибіркове середнє $\bar{x}$ – як точкову оцінку мат. очікування;
- виправлену дисперсію $\widetilde{S}_x^2$ – як точкову оцінку дисперсії;
- статистику $\frac{\bar{x}-a}{\widetilde{S}_x/\sqrt{n}}$ – як величину, що має розподіл Стюдента $t(n-1)$, щільність якого $p_t(x)$ не залежить від параметра a .

Довірчий інтервал, який містить невідоме математичне очікування $MX=a$ з довірчою ймовірністю γ :

$$\left(\bar{x} - \varepsilon \frac{\widetilde{S}_x}{\sqrt{n}}, \bar{x} + \varepsilon \frac{\widetilde{S}_x}{\sqrt{n}} \right).$$

Значення ε знаходять з умови $\int_0^\varepsilon p_t(x) dx = \frac{\gamma}{2}$ у таблицях розподілу

Стюдента (з $n-1$ ступенями вільності й довірчою ймовірністю γ) чи за допомогою комп'ютера.

11.3.4. Інтервальна оцінка дисперсії

Нехай $x_1, x_2, \dots, x_n$ – стандартна вибірка з ВВ X , розподіленої за нормальним законом із невідомими математичним очікуванням і дисперсією.

Для знаходження інтервалу для дисперсії використовують:

- вибіркове середнє $\bar{x}$ – як точкову оцінку мат. очікування;
- виправлену дисперсію $\widetilde{S}_x^2$ – як точкову оцінку дисперсії;
- статистику $\frac{n-1}{\sigma^2} \widetilde{S}_x^2$ – як величину, що має розподіл хі-квадрат $\chi^2(n-1)$, який не залежить від дисперсії σ^2 .

Довірчий інтервал, який містить невідомий параметр – дисперсію σ^2 – з довірчою ймовірністю γ :

$$\left(\frac{n-1}{\chi_2^2} \widetilde{S}_x^2, \frac{n-1}{\chi_1^2} \widetilde{S}_x^2 \right).$$

Значення χ_1^2 і χ_2^2 знаходять у таблицях розподілу хі-квадрат χ^2 (для $n-1$ ступеня вільності й рівнів значущості $\frac{1}{2} - \frac{\gamma}{2}$ і $\frac{1}{2} + \frac{\gamma}{2}$ відповідно) чи за допомогою комп'ютера.

Якщо $x_1, x_2, \dots, x_n$ – стандартна вибірка з ВВ X , розподіленої за нормальним законом із відомим математичним очікуванням a і невідомою дисперсією, використовують іншу точкову оцінку дисперсії.

Для знаходження інтервалу для дисперсії використовують:

- вибіркове середнє $\bar{x}$ – як точкову оцінку мат. очікування;
- ефективну оцінку $S_o^2 = \frac{1}{n} \sum_{k=1}^n (x_k - a)^2$ – як точкову оцінку дисперсії;
- статистику $\frac{n}{\sigma^2} S_o^2$ – як величину, що має розподіл хі-квадрат $\chi^2(n)$, який не залежить від дисперсії σ^2 .

Довірчий інтервал, який містить невідомий параметр – дисперсію σ^2 – з довірчою ймовірністю γ :

$$\left(\frac{nS_o^2}{\chi_2^2}, \frac{nS_o^2}{\chi_1^2} \right).$$

Значення χ_1^2 і χ_2^2 знаходять у таблицях розподілу хі-квадрат χ^2 (для n ступенів вільності й рівнів значущості $\frac{1}{2} - \frac{\gamma}{2}$ і $\frac{1}{2} + \frac{\gamma}{2}$ відповідно) чи за допомогою комп'ютера.

12. Взаємозв'язки між випадковими величинами

♣ Між величинами існує **функціональний зв'язок**, якщо кожному значенню однієї з них відповідає одне або декілька значень іншої.

♣ Зв'язок між випадковими величинами, при якому зі зміною значень однієї ВВ в іншій змінюється закон розподілу або його характеристики, називають **стохастичним (статистичним, імовірнісним) зв'язком**.

Наявність статистичного зв'язку між випадковими величинами не означає причинно-наслідкової залежності між ними. Математична статистика не встановлює й не досліджує причини явищ і наслідків.

Існують три підходи в математичній статистиці для виявлення й дослідження зв'язків між ВВ – кореляційний, регресійний і дисперсійний аналіз.

Кореляційний аналіз встановлює факт статистичної взаємозалежності між випадковими величинами та оцінює її рівень.

Регресійний аналіз дозволяє побудувати конкретну функціональну математичну модель і оцінити її адекватність.

Дисперсійний аналіз досліджує, як зміни випадкових чинників впливають на результат, і дозволяє відібрати найбільш значущі з них.

12.1. Основи теорії кореляції

♣ Якщо при зміні однієї випадкової величини змінюється середнє значення іншої, то в цьому випадку статистичну залежність називають **кореляційною залежністю**.

♣ Показником наявності кореляційного зв'язку між ВВ є **коефіцієнт кореляції** – із певною ймовірністю він показує, наскільки цей зв'язок близький до лінійного.

Множинний коефіцієнт кореляції враховує взаємозв'язки трьох і більше ВВ, парний коефіцієнт кореляції – для двох ВВ.

12.1.1. Коваріація і коефіцієнт кореляції

♣ **Коваріація** двох ВВ ξ і η – математичне очікування добутку їхніх відхилень:

$$\text{cov}(\xi, \eta) = M((\xi - M\xi) \cdot (\eta - M\eta)). \quad [12.1]$$

Властивості коваріації

Властивість 1. Коваріацію можна подати у вигляді:

$$\text{cov}(\xi, \eta) = M(\xi \cdot \eta) - M\xi \cdot M\eta.$$

Властивість 2. Для незалежних величин ξ і η коваріація дорівнює нулю:

$$\text{cov}(\xi, \eta) = 0.$$

Зворотне, у загальному випадку, неправильно.

Властивість 3. Коваріація ВВ із собою дорівнює дисперсії:

$$\text{cov}(\xi, \xi) = D\xi.$$

Властивість 4. Дисперсія суми двох довільних ВВ:

$$D(\xi + \eta) = D\xi + D\eta + 2 \text{cov}(\xi, \eta). \quad [12.2]$$

Властивість 5. Для коваріації виконується:

$$\begin{aligned} \text{cov}(K \cdot \xi, \eta) &= K \cdot \text{cov}(\xi, \eta) \\ \text{cov}(\xi, K \cdot \eta) &= K \cdot \text{cov}(\xi, \eta). \end{aligned}$$

Властивість 6.

$$|\text{cov}(\xi, \eta)| \leq \sigma_\xi \cdot \sigma_\eta.$$

Коваріація показує, чи є залежність між двома ВВ (див. 12.3).

Значення коваріації залежить від одиниць вимірювання значень ВВ. Щоб отримати безрозмірний коефіцієнт, коваріацію ділять на добуток середніх квадратичних відхилень – і отримують коефіцієнт кореляції Пірсона.

♣ **Коефіцієнт (парної) кореляції Пірсона** r випадкових величин ξ і η – відношення

$$r = \frac{\text{cov}(\xi, \eta)}{\sigma_{\xi} \cdot \sigma_{\eta}} = \frac{M((\xi - M\xi) \cdot (\eta - M\eta))}{\sigma_{\xi} \cdot \sigma_{\eta}}. \quad [12.3]$$

♣ Випадкові величини ξ і η , що мають ненульову кореляцію $r \neq 0$, називають **корельованими**. Якщо ж $r = 0$, то говорять, що ξ і η не корелюють.

Властивості коефіцієнта кореляції.

Властивість 1. Значення коефіцієнта кореляції задовольняють нерівності $|r| \leq 1$.

Властивість 2. Якщо ВВ ξ і η незалежні, то коефіцієнт кореляції дорівнює нулю:

$$r = 0.$$

Властивість 3. 1) Якщо між ВВ ξ і η наявний лінійний зв'язок, то коефіцієнт кореляції ξ і η за модулем дорівнює одиниці:

$$\xi = A \cdot \eta + B \Rightarrow |r| = 1.$$

2) І навпаки: якщо коефіцієнт кореляції ВВ ξ і η за модулем дорівнює одиниці, то з імовірністю 1 між ξ і η є лінійний зв'язок:

$$|r| = 1 \Rightarrow P(\xi = A \cdot \eta + B) = 1.$$

12.1.2. Кореляція і залежність випадкових величин

Дві корельовані величини завжди залежні. Зворотного висновку зробити не можна: залежні ВВ можуть бути як корельованими, так і некорельованими.

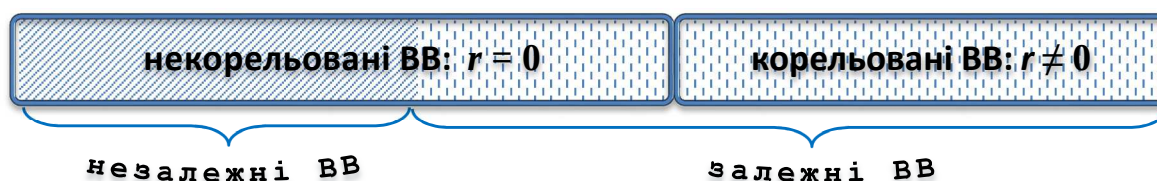


Рис. 12.1. Зв'язок між кореляцією й залежністю

Значення коефіцієнта кореляції (у межах від -1 до $+1$) показує силу зв'язку: чим ближче за абсолютною величиною до 1 , тим вона тісніша. Знак цього значення вказує напрям зв'язку:

$r > 0$ відповідає **прямому зв'язку** (коли зростання значень однієї ВВ пов'язане зі збільшенням значень іншої);

$r < 0$ відповідає **зворотному зв'язку** (коли зростання значень однієї ВВ пов'язане зі зменшенням значень іншої).

12.1.3. Множинна кореляція

♣ **Множинна (лінійна) кореляція** вивчає (лінійні) зв'язки між декількома випадковими величинами.

Для аналізу множинної кореляції складається матриця частинних коефіцієнтів кореляції для кожної пари величин, коли виключено вплив усіх інших ВВ (пам'ятаючи при цьому, що $r_{\xi\eta} = r_{\eta\xi}$ і що $r_{\xi\xi} = 1$):

$$\Delta = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & r_{nn} \end{pmatrix}$$

♣ Така матриця називається **кореляційною матрицею** системи n випадкових величин $\xi_1, \dots, \xi_n$.

Міру сумісного впливу на випадкову величину ξ_1 усіх інших ВВ $\xi_2, \dots, \xi_n$ відображає коефіцієнт множинної кореляції:

$$R_{\xi_1 | \xi_2, \dots, \xi_n} = \sqrt{1 - \frac{|\Delta|}{\Delta_{11}}}.$$

де $|\Delta|$ – визначник кореляційної матриці;

Δ_{11} – алгебраїчне доповнення її першого елемента r_{11} :

$$\Delta_{11} = \begin{vmatrix} r_{22} & \dots & r_{2n} \\ \dots & \dots & \dots \\ r_{n2} & \dots & r_{nn} \end{vmatrix}, \text{ тобто } \Delta_{11} \text{ є кореляційною матрицею випадкових величин } \xi_2, \dots, \xi_n.$$

Отже, величина множинного коефіцієнта кореляції залежить і від кореляції досліджуваної ВВ з кожною з інших величин, і від кореляції між самими цими величинами.

12.2. Елементи кореляційного аналізу

♣ **Кореляційний аналіз** – це розділ математичної статистики, що вивчає кореляційні зв'язки між випадковими величинами.

Основні завдання кореляційного аналізу: • виявити, як змінюються в середньому значення однієї ВВ при зміні значень іншої; • виміряти тісноту зв'язку між ВВ.

12.2.1. Емпіричний коефіцієнт кореляції

♣ **Емпіричний (вибірковий) коефіцієнт кореляції** – величина, яку обчислюють за формулою:

$$\rho = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{S_x \cdot S_y}, \quad [12.4]$$

де $\bar{x} = \frac{1}{n} \sum_{k=1}^n x_k$, $\bar{y} = \frac{1}{n} \sum_{k=1}^n y_k$.

Ця формула отримана з формули (12.3) заміною мат. очікування й середнього квадратичного відхилення на їхні обґрунтовані оцінки.

Емпіричний коефіцієнт кореляції ρ за абсолютною величиною не перевищує одиниці:

$$-1 \leq \rho \leq 1.$$

Якщо досліджувані величини X і Y мають сумісний нормальний розподіл, то емпіричний коефіцієнт кореляції ρ є оцінкою (обґрунтованою й асимптотично ефективною) коефіцієнта кореляції r :

$$\rho = \hat{r}.$$

У різних прикладних галузях прийняті різні оцінки міри статистичної залежності – наприклад, така:



Рис. 12.2. Градація значень коефіцієнта кореляції

При значеннях ρ , більших 0,99, залежність зазвичай вважають повною, функціональною.

Емпіричний коефіцієнт ρ – величина випадкова. І лише за його величиною не можна судити про достовірність зв'язку: можливо, його значення – лише випадкова обставина. Тому необхідно перевіряти гіпотезу про значущість емпіричного коефіцієнта кореляції ρ .

12.2.2. Діаграма розсіювання

♣ Графічне представлення кореляційного зв'язку між двома величинами X і Y називають **діаграмою розсіювання** (кореляційним полем).

При наближенні значення коефіцієнта кореляції до $+1$ або -1 сумісний розподіл (X , Y) має тенденцію концентруватися поблизу деякої прямої лінії.

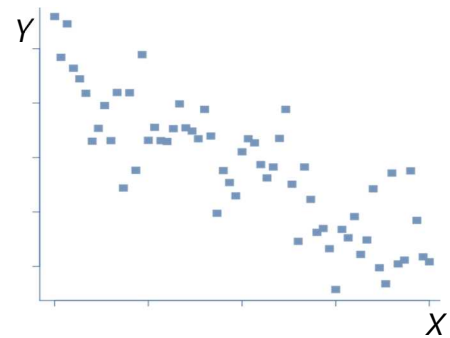


Рис. 12.3. Діаграма розсіювання

Чим ближче абсолютна величина коефіцієнта кореляції до 1 , тим вужче еліпс, обмежуючий точки (x, y) на діаграмі.

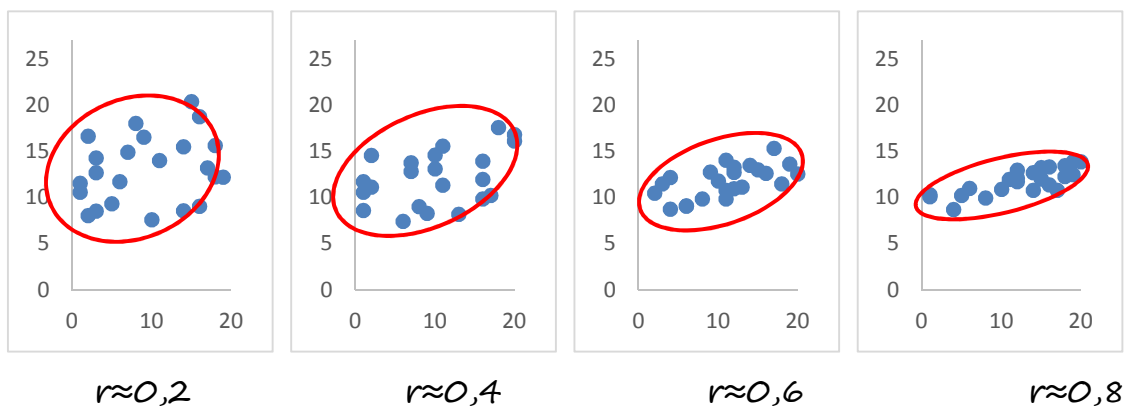


Рис. 12.4. Діаграма розсіювання для різних значень коефіцієнта кореляції

Коли значення r близькі до 0 , лінійну тенденцію на діаграмі не видно.

Аналіз діаграми може привести до висновку про відсутність лінійної тенденції, але наявності іншого, нелінійного взаємозв'язку даних або яких-небудь їхніх особливостей.

12.3. Елементи регресійного аналізу

12.3.1. Моделі регресії

Зазвичай розглядають закономірність зміни *середнього значення* однієї випадкової величини від значень іншої – тобто умовного математичного очікування.

♣ Для двовимірної ВВ (ξ, η) **умовним математичним очікуванням** ВВ η відносно ВВ ξ називають функцію $M_{\xi}(\eta)$, яка для кожного значення ξ дорівнює математичному очікуванню η , знайденому за умови, що ξ набула цього значення.

♣ Залежності між умовним математичним очікуванням однієї величини і значеннями іншої (чи декількох інших) називають **функцією регресії**:

$M_{\xi}(\eta) = f(\xi)$ – рівняння регресії η на ξ ,

$M_{\eta}(\xi) = \varphi(\eta)$ – рівняння регресії ξ на η .

(У цих формулах $f(\xi)$ – функція регресії η на ξ , а $\varphi(\eta)$ – функція регресії ξ на η .)

Графік функції регресії називають *лінією регресії*.

У математичній статистиці умовному математичному очікуванню відповідає **умовне середнє**.

12.3.2. Начала регресійного аналізу

За результатами експерименту можна отримати наближене вираження (оцінку) функцій регресії (тобто *вибіркові рівняння регресії*) вигляду:

$$\overline{y}_x = \hat{f}(x), \quad \overline{x}_y = \hat{\varphi}(y), \quad [12.5]$$

де $\overline{y}_x$ – умовне середнє змінної Y при фіксованому значенні змінної X ,

$\overline{x}_y$ – умовне середнє змінної X при фіксованому значенні змінної Y .

Кореляційна залежність між величинами X і Y може бути подана як функціональна залежність між $\overline{y}_x$ і X або між $\overline{x}_y$ і Y .

Методи знаходження таких залежностей і оцінки їхніх статистичних властивостей складають зміст регресійного аналізу.

12.3.3. Метод найменших квадратів. Парна лінійна регресія

Завдання: за заданими значеннями $(x_1, y_1), \dots, (x_n, y_n)$ і відповідними частотами цих значень n_{kj} , а також за умовними середніми $\overline{x}_{y_j}$ і $\overline{y}_{x_k}$ знайти функції регресії $f(x)$ і $\varphi(y)$.

Найпростішою формою кореляційно-регресійного зв'язку є лінійний зв'язок між двома величинами – **парна лінійна регресія**. Її рівняння:

$$\overline{y}_x = ax + b \quad \text{або} \quad \overline{x}_y = cy + d. \quad [12.6]$$

У кожному з цих рівнянь по дві невідомі величини – коефіцієнти лінійної регресії. Так, у рівнянні $\overline{y}_x = ax + b$ це параметри a, b . Однозначно визначити їх за вибіркою неможливо – через вплив випадкових чинників.

Вплив цих випадкових *відхилень* (*погрешностей, помилок, збурень, шуму*) на спостережувані значення подають так:

$$\overline{y}_{x_k} = ax_k + b + \varepsilon_k \quad (k=1, \dots, n),$$

де ε_k – випадкова змінна, вона характеризує відхилення значень вибірки від теоретичної регресії.

Треба знайти такі значення невідомих a і b , щоб випадкові відхилення ε_k у сукупності були близькі до нуля.

Суть **методу найменших квадратів**: знайти такі значення параметрів a і b , щоб була як можна меншою зважена сума квадратів відхилень ε_k :

$$\sum_k \varepsilon_k^2 \cdot m_{xk} \rightarrow \min.$$

Тоді буде знайдено рівняння прямої, найменш віддаленої від усіх точок $(x_k, \overline{y}_{x_k})$ з урахуванням їхніх вагів. (Підсумовуються саме квадрати величин ε_k^2 , щоб запобігти зменшенню суми самих величин ε_k з різними знаками.)

Цю суму розглядають як функцію від двох невідомих параметрів a і b :

$$\sum_k \varepsilon_k^2 \cdot m_{xk} = \sum_k (ax_k + b - \overline{y}_{x_k})^2 \cdot m_{xk}.$$

Записавши необхідну умову існування її мінімуму (рівність нулю двох частинних похідних), отримують лінійну систему двох рівнянь із двома змінними a і b . Знайшовши рішення системи і перевіривши, що у знайденій точці досягається саме мінімум, отримують вирази для параметрів a і b :

$$a = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - (\bar{x})^2} \quad \text{або} \quad a = \rho \frac{S_y}{S_x}; \quad [12.7]$$

$$b = \bar{y} - a\bar{x}.$$

Таким чином, рівняння лінійної парної регресії одержує вигляд:

$$y_x = \rho \frac{S_y}{S_x} x + (\bar{y} - a\bar{x})$$

або

$$y_x - \bar{y} = \rho \frac{S_y}{S_x} (x - \bar{x}).$$

[12.8]

Для лінійної регресії X на Y отримують аналогічні формули.

Наприклад:
$$c = \rho \frac{S_x}{S_y}; \quad d = \bar{x} - c\bar{y}, \quad [12.9]$$

$$x_y - \bar{x} = \rho \frac{S_x}{S_y} (y - \bar{y}). \quad [12.10]$$

Зазвичай на графіці зображують обидві прямі регресії $y_x = ax + b$ і $y_x = cy + d$ та точки вибірки (x_k, y_k) . Вочевидь, вони обоє проходять через точку $(\bar{x}, \bar{y})$ – центр розподілу.

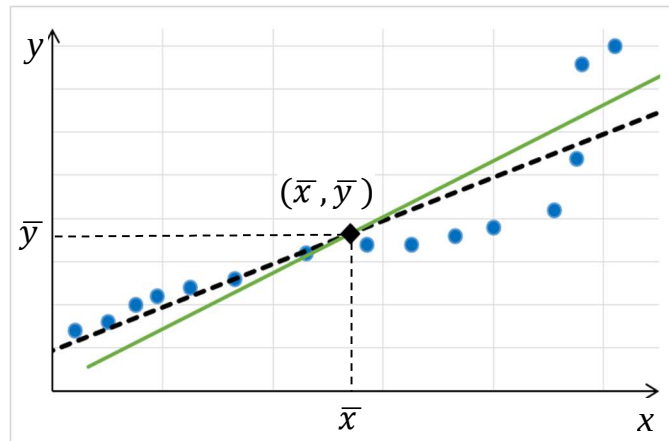


Рис. 12.5. Прямі регресії на кореляційному полі

Прямі регресії Y на X і X на Y співпадають лише тоді, коли $|\rho| = 1$, тобто у разі функціональної лінійної залежності між величинами X і Y .

13. Основи перевірки статистичних гіпотез

13.1. Статистичні гіпотези

13.1.1. Гіпотези і критерії. Критичні області і точки

Перевірка правдоподібності статистичних гіпотез є однією з основних задач математичної статистики.

♣ **Статистичною гіпотезою** називають будь-які статистичні припущення про властивості генеральної сукупності, що перевіряються за значеннями вибірки.

Перевірка гіпотез – це процедура перевірки, чи відповідає висунена гіпотеза вибіровим даним, та вирішення, чи допустима ця гіпотеза.

♣ Правило, за яким вирішують, прийняти або відхилити гіпотезу, називають **критерієм**.

Ця процедура не доводить правильність або помилковість гіпотези. Вона лише перевіряє, чи узгоджені вибіркові дані з гіпотезою.

♣ Гіпотезу, що перевіряється (основну), зазвичай позначають H_0 і називають **нульовою гіпотезою**. Гіпотезу, що суперечить їй, називають **альтернативною гіпотезою** й позначають H_1 .

Для перевірки гіпотези розраховують за значеннями вибірки деяку випадкову величину, яку називають **статистикою критерію** (*статистичним критерієм*). Її добирають так, щоб був відомий її розподіл за умови справедливості гіпотези. Обчислюють ймовірність появи цього розрахованого значення і за нею роблять висновок – відхиляти гіпотезу або ні.

Таким чином, усі можливі значення статистики розбивають на дві підмножини: область відхилення гіпотези й область прийняття гіпотези.

♣ Область відхилення гіпотези називають **критичною областю**, а область прийняття гіпотези – **областю прийняття гіпотези** (*областю допустимих значень*).

Якщо розраховане за вибіркою значення статистики критерію потрапляє до критичної області, то гіпотезу відкидають.

♣ Точки, що відокремлюють критичну область від області прийняття гіпотези, називають **критичними точками**. Позначають $K_{кр}$.

Для кожної статистики існують таблиці, за якими знаходять її критичні точки.

13.1.2. Рівень значущості критерію

Перевірку гіпотез здійснюють на підставі випадкової вибірки. Отже, існує деяка ймовірність прийняти помилкове розв'язання.

Розрізняють помилки першого й другого роду. **Помилка 1 роду** полягає в тому, що буде знехтувана правильна гіпотеза. **Помилка 2 роду** полягає в тому, що буде прийнята неправильна гіпотеза.

гіпотеза H_0	приймається	відкидається
• правильна	+	помилка 1 роду
• неправильна	помилка 2 роду	+

Імовірність помилки 1 роду позначають α . Імовірність помилки 2 роду позначають β .

♣ Імовірність α припуститися помилки 1 роду (тобто відкинути нульову гіпотезу, якщо вона істинна) називають **рівнем значущості критерію**. Зазвичай на практиці використовують рівні значущості 0,05; 0,01.

♣ Імовірність $1 - \beta$ не припуститися помилки 2 роду (тобто ймовірність відкинути нульову гіпотезу, якщо вона помилкова) називають **потужністю критерію**.

Величину $1 - \alpha$ називають **достовірністю**.

13.1.3. Загальна схема перевірки гіпотези

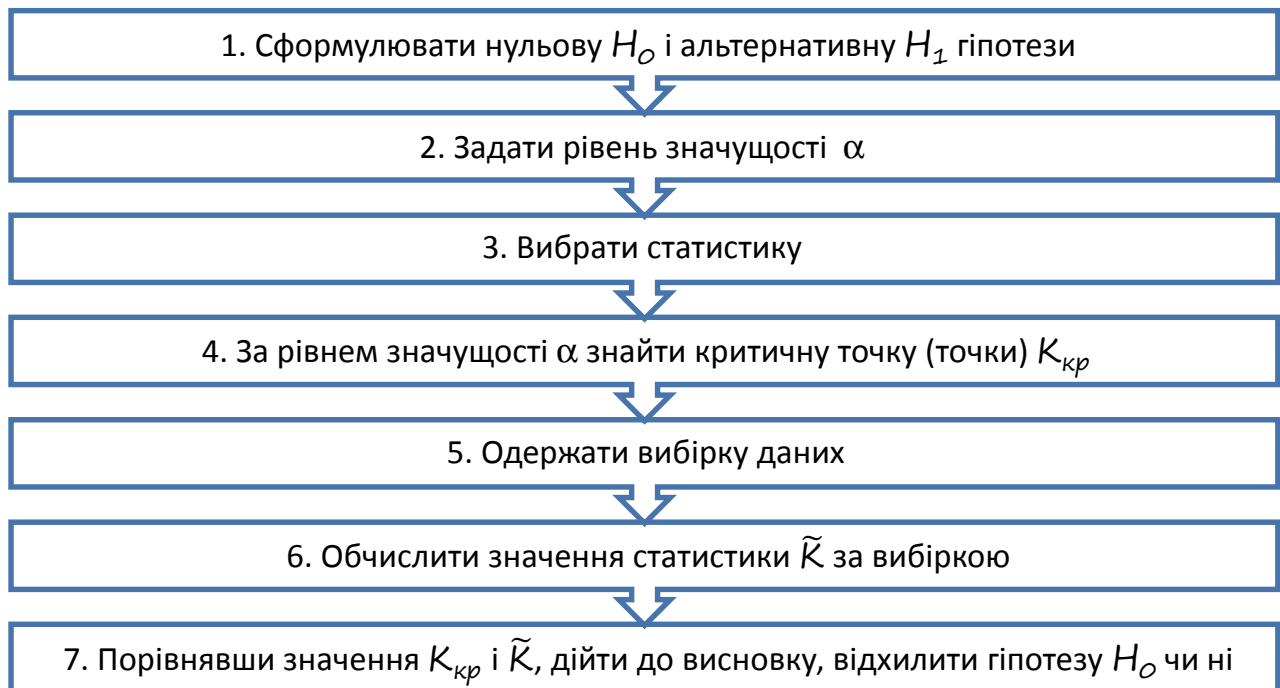
Висунувши гіпотезу та задавши рівень значущості, вибирають статистику для її перевірки і знаходять для неї критичну точку (точки). Обчислюють значення статистики за вибірковими даними й порівнюють із критичною точкою. Якщо розраховане значення потрапляє в критичну область, гіпотезу відкидають.

Загальна схема перевірки статистичної гіпотези:

[13.1]

1. Формулюють нульову гіпотезу H_0 і альтернативну їй гіпотезу H_1 .
2. Задають рівень значущості α (межу малої ймовірності), найчастіше $\alpha = 0,05$.
3. Вибирають статистику K – таку випадкову величину, щоб закон її розподілу при справедливості H_0 був відомий.
4. Знаходять критичну точку (точки) $K_{кр}$ статистики згідно з вибраним рівнем значущості α .
5. Отримують вибіркові дані.
6. Розраховують за ними значення статистики $\tilde{K}$.
7. Якщо значення $\tilde{K}$ потрапляє в критичну область, нульову гіпотезу H_0 відхиляють і приймають альтернативну гіпотезу H_1 .

Якщо значення $\tilde{K}$ не потрапляє до критичної області, це не означає, що доведена істинність нульової гіпотези H_0 . Це лише значить, що вибіркові дані не суперечать цій гіпотезі; говорять: «результати спостережень з імовірністю $1 - \alpha$ не суперечать висуненій гіпотезі H_0 ».

Алгоритм перевірки статистичної гіпотези

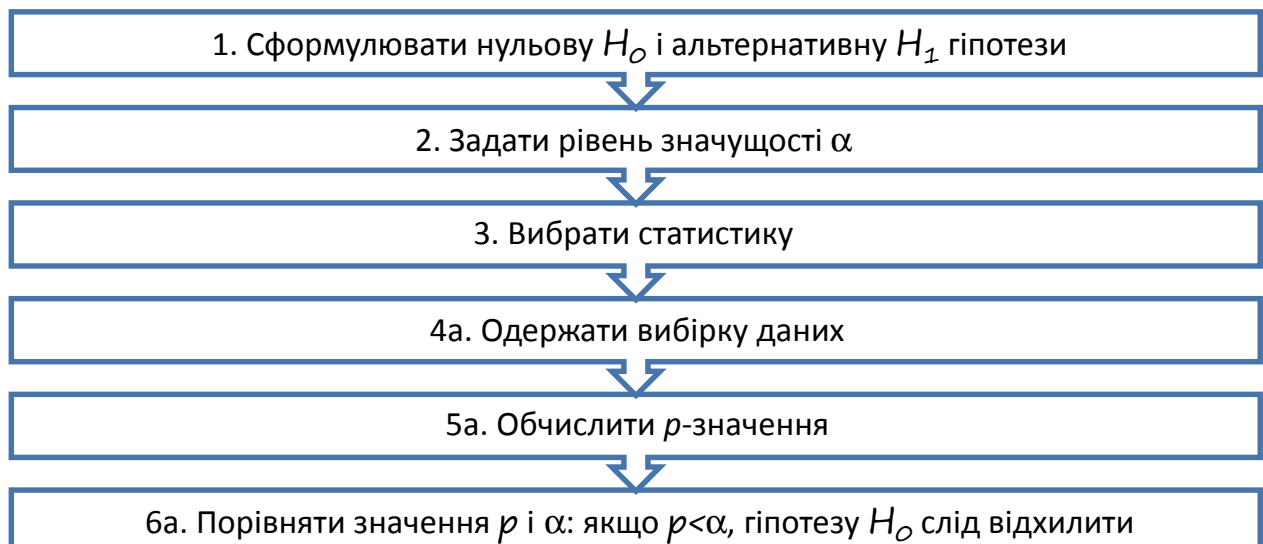
З появою комп'ютерів замість пошуку критичних точок (п. 4 схеми) і порівняння з ними $\tilde{K}$ (пп. 6, 7) зазвичай відразу обчислюють так зване p -значення – спостережуваний рівень значущості.

♣ Спостережуваний рівень значущості p -значення – це ймовірність отримати значення статистики $\tilde{K}$ (або ще більше), якщо гіпотеза H_0 була б істинна.

У програмах відразу обчислюють p -значення. Якщо воно менше вибраного рівня значущості ($p < \alpha$), гіпотезу H_0 відхиляють, інакше – ні.

Алгоритм перевірки статистичної гіпотези з використанням p -значення

[13.2]



13.1.4. Перевірка гіпотези про рівність двох дисперсій нормальних генеральних сукупностей

Нехай X і Y – нормально розподілені генеральні сукупності, дисперсії яких невідомі. Із них отримані дві незалежні виборки обсягів n_X і n_Y відповідно, за якими знайдені виправлені вибіркові дисперсії $\widetilde{S}_X^2$ і $\widetilde{S}_Y^2$. Необхідно перевірити з рівнем значущості α нульову гіпотезу H_0 , що дисперсії генеральних сукупностей однакові: $\sigma_X^2 = \sigma_Y^2$.

Як критерій перевірки використовують відношення більшої виправленої дисперсії до меншої, тобто випадкову величину $F = \frac{\widetilde{S}_{\max}^2}{\widetilde{S}_{\min}^2}$

(де $\widetilde{S}_{\max}^2 = \max(\widetilde{S}_X^2, \widetilde{S}_Y^2)$, $\widetilde{S}_{\min}^2 = \min(\widetilde{S}_X^2, \widetilde{S}_Y^2)$).

Якщо гіпотеза H_0 істинна, то випадкова величина F має F -розподіл Фішера – Снедекора із ступенями вільності $k_1 = n_1 - 1$ і $k_2 = n_2 - 1$ (де n_1 і n_2 – обсяги вибірок, за якими обчислені відповідно більша та менша виправлені дисперсії).

Обчислене за вибіркою значення F -статистики $\widetilde{F}$ порівнюють із критичним значенням $F_{\text{критич}}$, знайденим у таблицях (чи за допомогою комп'ютера) для заданого рівня значущості і ступенів вільності. Якщо $\widetilde{F} > F_{\text{критич}}$, то скоріш за все гіпотезу H_0 слід відхилити.

Схема перевірки гіпотези:

1. Нульова гіпотеза $H_0: \sigma_X^2 = \sigma_Y^2$; альтернативна $H_1: \sigma_X^2 > \sigma_Y^2$ (чи $\sigma_X^2 < \sigma_Y^2$).
2. Отримують вибіркові дані. Обчислюють виправлені вибіркові дисперсії

$$\widetilde{S}_X^2 = \frac{1}{n_X - 1} \sum_{k=1}^{n_X} (x_k - \bar{x})^2, \quad \widetilde{S}_Y^2 = \frac{1}{n_Y - 1} \sum_{k=1}^{n_Y} (y_k - \bar{y})^2,$$
 знаходять із них найбільшу $\widetilde{S}_{\max}^2$ та найменшу $\widetilde{S}_{\min}^2$.

3. Розраховують за ними значення статистики $\widetilde{F} = \frac{\widetilde{S}_{\max}^2}{\widetilde{S}_{\min}^2}$.
4. За заданим рівнем значущості α і ступенями вільності $k_1 = n_1 - 1$ і $k_2 = n_2 - 1$ знаходять критичну точку $F_{\text{критич}} = F(\alpha, k_1, k_2)$.
5. Якщо $\widetilde{F} \leq F_{\text{критич}}$, нема підстав відхиляти нульову гіпотезу H_0 .
Якщо $\widetilde{F} > F_{\text{критич}}$, нульову гіпотезу H_0 відхиляють і приймають гіпотезу H_1 .

Якщо альтернативна гіпотеза H_1 має вигляд $\sigma_X^2 \neq \sigma_Y^2$, то п. 4 схеми набуває вигляду:

4'. За заданим рівнем значущості $\frac{\alpha}{2}$ (удвічі меншим заданого) і ступенями вільності $k_1=n_1-1$ і $k_2=n_2-1$ знаходять критичну точку $F_{\text{критич}} = F\left(\frac{\alpha}{2}, k_1, k_2\right)$.

Зауваження. Якщо відомо математичне очікування якої-небудь з генеральних сукупностей X чи Y , тоді замість виправленої вибіркової дисперсії використовують відповідну її оцінку $S_X^2 = \frac{1}{n_X} \sum_{k=1}^{n_X} (x_k - MX)^2$ чи

$$S_Y^2 = \frac{1}{n_Y} \sum_{k=1}^{n_Y} (y_k - MY)^2 \text{ та відповідні ступені вільності } k_1=n_1 \text{ або } k_2=n_2.$$

13.1.5. Перевірка гіпотези про рівність математичних очікувань нормальних генеральних сукупностей

Нехай X і Y – нормально розподілені генеральні сукупності з невідомими однаковими дисперсіями. Із них отримані дві незалежні виборки обсягів n_X і n_Y відповідно, для них визначені вибіркові середні $\bar{x}$ і $\bar{y}$ та виправлені вибіркові дисперсії $\widetilde{S}_X^2$ і $\widetilde{S}_Y^2$. Необхідно перевірити з рівнем значущості α нульову гіпотезу $H_0: MX=MY$ при альтернативній гіпотезі $H_1: MX \neq MY$.

Як критерій перевірки використовують випадкову величину $t = \frac{\bar{x} - \bar{y}}{\sqrt{D(X-Y)}}$. Якщо гіпотеза H_0 істинна, то величина t має t -розподіл

Стюдента із ступенем вільності $k=n_X+n_Y-2$. Для обчислення t підста-

$$\text{вимо оцінки дисперсії } \widetilde{S}_X^2 \text{ і } \widetilde{S}_Y^2: \quad t = \frac{(\bar{x} - \bar{y}) \cdot \sqrt{n_X + n_Y - 2}}{\sqrt{(n_X - 1)\widetilde{S}_X^2 + (n_Y - 1)\widetilde{S}_Y^2}} \cdot \sqrt{\frac{n_X \cdot n_Y}{n_X + n_Y}}.$$

Обчислене за вибіркою значення t -статистики $\tilde{t}$ порівнюють із критичним значенням $t_{\text{критич}}$, знайденим у таблицях (чи за допомогою комп'ютера) для заданого рівня значущості і ступеню вільності. Якщо $|\tilde{t}| > t_{\text{критич}}$, то скоріш за все гіпотезу H_0 слід відхилити.

Схема перевірки гіпотези:

1. Нульова гіпотеза $H_0: MX=MY$; альтернативна $H_1: MX \neq MY$.
2. Отримують вибірові дані. Обчислюють виправлені вибірові дисперсії $\widetilde{S}_X^2$ і $\widetilde{S}_Y^2$.
3. Розраховують за ними значення статистики
$$t = \frac{(\bar{x} - \bar{y}) \cdot \sqrt{n_X + n_Y - 2}}{\sqrt{(n_X - 1)\widetilde{S}_X^2 + (n_Y - 1)\widetilde{S}_Y^2}} \cdot \sqrt{\frac{n_X \cdot n_Y}{n_X + n_Y}}.$$
4. За заданим рівнем значущості α і ступенем вільності $k = n_X + n_Y - 2$ знаходять критичну точку $t_{\text{критич}} = t(\alpha, k)$.
5. Якщо $|\widetilde{t}| \leq t_{\text{критич}}$, нема підстав відхиляти нульову гіпотезу H_0 .
Якщо $|\widetilde{t}| > t_{\text{критич}}$, нульову гіпотезу H_0 відхиляють і приймають гіпотезу H_1 .

13.2. Критерії узгодження

Критерії узгодження – це критерії перевірки, чи відповідає нульова гіпотеза вибіровим даним. Під альтернативною гіпотезою в цьому випадку розуміють, що не виконана нульова гіпотеза H_0 .

Передусім критерії узгодження застосовують при перевірці, що теоретичний розподіл має певний вигляд. Найчастіше використовують критерій узгодження Пірсона χ^2 (хі-квадрат).

13.2.1. Критерій узгодження Пірсона χ^2

Висунена гіпотеза H_0 : «Вибірка $x_1, x_2, \dots, x_n$ отримана з генеральної сукупності, що має деякий певний розподіл $F(x)$ ». При цьому параметри розподілу можуть бути як відомі (проста гіпотеза), так і невідомі (і тоді оцінені за вибіркою).

Для перевірки гіпотези діапазон значень вибірки розбивають на m інтервалів $[z_{k-1}, z_k)$ (де $k=1, 2, \dots, m$) і для кожного з них розраховують частоту n_k і відносну частоту $p_k^* = \frac{n_k}{n}$ потрапляння до нього значень вибірки. Також розраховують теоретичні ймовірності p_k потрапляння значень до кожного інтервалу за передбаченим розподілом $F(x)$: $p_k = F(z_k) - F(z_{k-1})$.

Як міру розбіжності даних приймають величину K :

$$K = \sum_{k=1}^m c_k (p_k^* - p_k)^2.$$

Якщо $c_k = \frac{n}{p_k}$, то закон розподілу K при великих n практично не залежить від функції розподілу $F(x)$ і числа випробувань n , а лише від числа інтервалів m і збігається до розподілу хі-квадрат χ^2 . Тому таку величину K

також позначають χ^2 :
$$K = \chi^2 = \sum_{k=1}^m \frac{(n_k - np_k)^2}{np_k}. \quad [13.3]$$

Розподіл χ^2 залежить від параметра ν – числа ступенів вільності. Для критерію узгодження Пірсона він пов'язаний із числом інтервалів m :

$$\nu = m - u - 1,$$

де m – число інтервалів, u – число невідомих параметрів теоретичного розподілу, розрахованих за вибіркою.

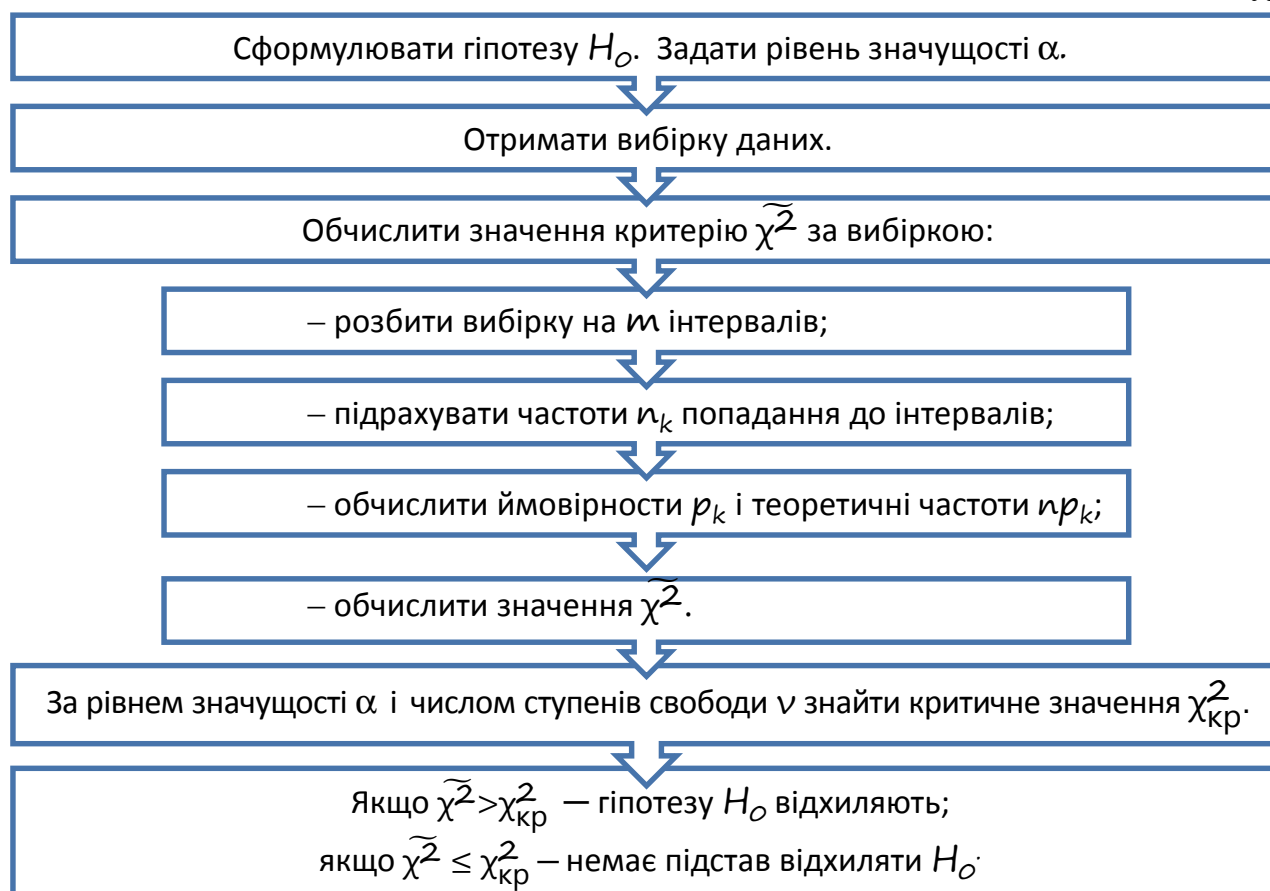
Для перевірки гіпотези обчислене за вибіркою значення $\tilde{K} = \tilde{\chi}^2$ порівнюють за схемою (13.1) з критичним значенням $\chi_{\text{критич}}^2$, знайденим у таблицях (чи за допомогою комп'ютера) для заданого рівня значущості α і числа ступенів вільності ν . Якщо $\tilde{\chi}^2 > \chi_{\text{критич}}^2$, то скоріш за все гіпотезу H_0 слід відхилити.

Схема (13.1) набере вигляду (13.4).

Схема перевірки за критерієм Пірсона χ^2 [13.4]
гіпотези про закон розподілу неперервної ВВ:

- 1–3. Формулюють нульову гіпотезу H_0 . Задають рівень значущості α . Як статистику критерію використовують $K = \chi^2$.
4. Отримують вибіркові дані.
5. Розраховують за ними значення статистики $\tilde{\chi}^2$:
 - а) увесь діапазон значень розбивають на m груп (інтервалів) і підраховують частоти n_k потрапляння значень до кожного інтервалу;
 - б) для кожного інтервалу розраховують (за формулою теоретичного розподілу гіпотези H_0) імовірності p_k потрапляння до нього значень і теоретичні частоти np_k . Якщо точні значення u параметрів розподілу невідомі, їх заміняють на їхні точкові оцінки;
 - в) розраховують значення $\tilde{\chi}^2$.
6. За таблицями або за допомогою комп'ютера визначають критичне значення $\chi_{\text{критич}}^2$ для заданого рівня значущості α і числа ступенів вільності $\nu = m - u - 1$.
7. Якщо $\tilde{\chi}^2 > \chi_{\text{критич}}^2$, то на рівні значущості α гіпотезу H_0 відхиляють як таку, що не узгоджується з результатами спостережень.
Якщо ж $\tilde{\chi}^2 \leq \chi_{\text{критич}}^2$, то результати спостережень із імовірністю α не суперечать висуненій гіпотезі H_0 і немає підстав її відкидати.

Алгоритм застосування критерію Пірсона χ^2



13.2.2. Критерій Колмогорова

Критерій узгодження Колмогорова призначений для перевірки гіпотези про приналежність вибірки деякому закону розподілу. Він придатний лише для неперервних ВВ і тільки якщо для передбачуваного закону розподілу відомі усі параметри.

Нехай $F(x)$ – теоретична функція розподілу, параметри якої відомі, $F_n(x)$ – емпірична функція розподілу. Висувається проста гіпотеза H_0 : генеральна сукупність має розподіл $F(x)$.

Як міру розбіжності розглядають величину D_n – статистику Колмогорова:

$$D_n = \max_x |F_n(x) - F(x)|. \quad [13.5]$$

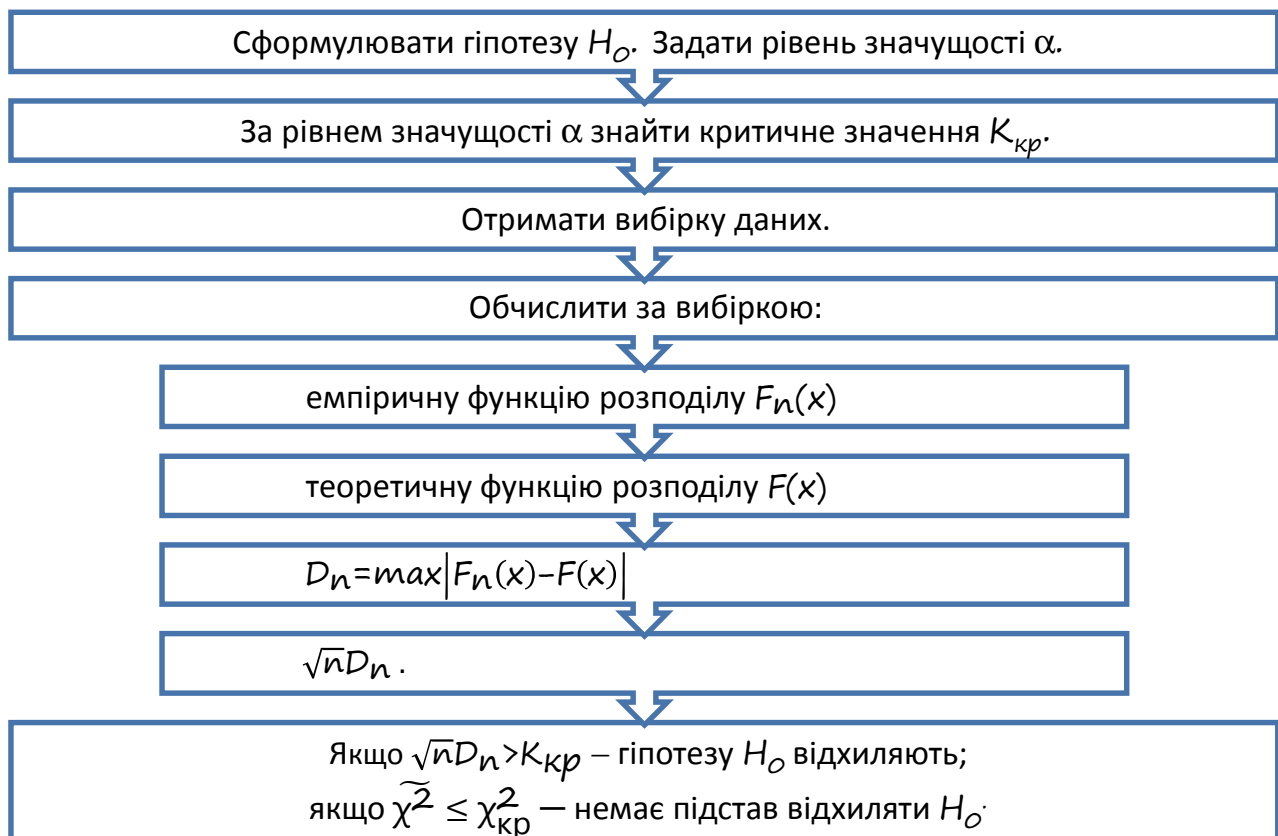
Якщо гіпотеза H_0 правильна, то для будь-якої функції розподілу $F(x)$ статистика D_n у певному сенсі збігається до нуля.

Доведено, що розподіл величини $\sqrt{n}D_n$ при $n \rightarrow \infty$ збігається до функції $K(x)$, для якої складено таблицю критичних значень. Задавши рівень значущості, за цими таблицями знаходять критичне значення $K_{\text{критич}}$, і якщо виявиться, що $\sqrt{n}D_n > K_{\text{критич}}$, то гіпотезу H_0 відкидають.

Схема застосування критерію Колмогорова:

[13.6]

- 1–2. Формулюють нульову гіпотезу H_0 . Задають рівень значущості α .
- 3–4. Як статистику критерію використовують $K=K(x)$. Для вибраного α знаходять критичне значення $K_{\text{критич}}$.
5. Отримують вибіркові дані.
6. Обчислюють:
 - а) емпіричну функцію розподілу $F_n(x)$ за вибіркою;
 - б) теоретичну функцію розподілу $F(x)$ за вибіркою;
 - в) величину $D_n = \max_x |F_n(x) - F(x)|$;
 - г) значення $\sqrt{n}D_n$.
7. Якщо $\sqrt{n}D_n > K_{\text{критич}}$, то на рівні значущості α гіпотезу H_0 відхиляють. Якщо ж $\sqrt{n}D_n \leq K_{\text{критич}}$, то результати спостережень з імовірністю α не суперечать висуненій гіпотезі H_0 і немає підстав її відкидати.

Алгоритм застосування критерію Колмогорова

Перевага критерію Колмогорова – у його простоті і вільності від розподілу: неважливо, з яким саме законом перевіряється узгодженість. Але бажано, щоб дані були негрупованими, і потрібно, щоб були відомі усі параметри передбачуваного закону. Якщо ж параметри невідомі, виявляється перевага критерію Пірсона. Але він має недоліки через вимоги до інтервалів угруповання даних.

Додатки

Додаток 1

Основні позначення і скорочення

Позначення:

$A, B, C \dots$ — випадкові події

$\overline{A}$ — подія, протилежна A

$\emptyset$ — неможлива подія

Ω — достовірна подія

$\omega_1, \omega_2 \dots$ — елементарні події

$\Omega = \{\omega_1, \omega_2, \dots\}$ — простір елементарних подій

U — алгебра подій

$P(A)$ — імовірність події A

$P(B/A)$ — умовна ймовірність події B (імовірність події B за умови, що відбулася подія A)

$H_1, H_2, H_3 \dots$ — гіпотези (події, що утворюють повну групу, з однією з яких може відбутися подія A)

$\xi, \eta, \zeta, \xi(\omega), \eta(\omega) \dots$ — випадкові величини

$M\xi$ — математичне очікування випадкової величини ξ

$D\xi$ — дисперсія випадкової величини ξ

$\sigma(\xi)$ — середнє квадратичне (стандартне) відхилення випадкової величини ξ

ν_m — початковий момент порядку m

μ_m — центральний момент порядку m

$F_\xi(x), F(x)$ — функція розподілу випадкової величини ξ

$p_\xi(x), p(x)$ — щільність розподілу випадкової величини ξ

(ξ, η) — двовимірна випадкова величина (система двох випадкових величин)

$F_{\xi, \eta}(x, y)$ — функція сумісного розподілу (ξ, η)

$\text{cov}(\xi, \eta)$ — коваріація випадкових величин ξ і η

$r_{\xi, \eta}$ — коефіцієнт кореляції випадкових величин ξ і η

$\Phi(x)$ — функція Лапласа (інтеграл імовірностей)

χ^2 — розподіл Пірсона

$\chi^2(n)$ — випадкова величина, що має розподіл χ^2 з n ступенями вільності

C_n^k — число сполук із n елементів по k елементів

A_n^k — число розміщень із n елементів по k елементів

P_n — число перестановок із n елементів

$n!$ — факторіал числа n

$F_n(x)$ — емпірична функція розподілу

$N(a, \sigma)$ — нормальний розподіл з параметрами a (математичне очікування) і σ (середнє квадратичне відхилення)

$N(0,1)$ — стандартний (нормований) нормальний розподіл (з математичним очікуванням 0 і середнім квадратичним відхиленням 1)

$\xi \sim N(a, \sigma)$ — випадкова величина ξ має нормальний розподіл з параметрами a і σ

$t(n)$ — випадкова величина, що має t -розподіл Стюдента з n ступенями вільності

$\bar{x}$ — вибіркове середнє

S_x^2 — вибіркова дисперсія

$\widetilde{S}_x^2$ — виправлена (незміщена) вибіркова дисперсія

S_x — вибіркове середнє квадратичне відхилення

$\hat{\Theta}$ — оцінка параметра Θ

γ — довірна ймовірність (надійність оцінки)

α — рівень значущості

$L(\Theta, x_1, x_2, \dots, x_n)$ — функція правдоподібності

v_m^* — початковий емпіричний момент порядку m

μ_m^* — центральний емпіричний момент порядку m

ρ — емпіричний коефіцієнт кореляції

$\overline{y_x}$ — умовне середнє змінної Y при фіксованому значенні змінної X

ε_k — відхилення (залишки, погрішності, помилки, обурення, шум)

H_0 — нульова гіпотеза

H_1 — альтернативна гіпотеза

Скорочення:

ТЙ — теорія ймовірностей

МС — математична статистика

ВВ — випадкова величина

ЦГТ — центральна гранична теорема

ЗВЧ — закон великих чисел

Додаток 2

Грецький алфавіт

Α	α	альфа	Ι	ι	йота	Ρ	ρ	ро
Β	β	бета	Κ	κ	каппа	Σ	σ	сигма
Γ	γ	гамма	Λ	λ	лямбда	Τ	τ	тау
Δ	δ	дельта	Μ	μ	мю	Υ	υ	юпсilon (іпсilon)
Ε	ε	епсілон	Ν	ν	ню (ні)	Φ	φ	фі
Ζ	ζ	дзета	Ξ	ξ	ксі	Χ	χ	хі
Η	η	ета	Ο	ο	омікрон	Ψ	ψ	псі
Θ	θ, ϑ	тета	Π	π	пі	Ω	ω	омега

Основні формули комбінаторики

Правило суми. Число способів виконати будь-яку (одну) з двох взаємовиключних дій дорівнює сумі способів виконати кожен з них.

Правило множення. Число способів виконати одну за іншою декілька дій дорівнює добутку способів виконати кожен з них.

Правило включення – виключення. Нехай m об'єктів мають властивість A , а n об'єктів мають властивість B , причому k об'єктів мають одночасно властивості A і B . Тоді кількість елементів, що мають властивості A або B , дорівнює $m+n-k$.

Усі три правила поширюються на будь-яке скінченне число дій або властивостей.

* * *

Перестановки P_n з n елементів – всілякі сукупності з n елементів, що відрізняються одна від одної лише порядком розташування елементів.

Розміщення A_n^k з n елементів по k в кожному – всілякі впорядковані сукупності k елементів із n , які відрізняються: або складом елементів, або порядком їх розташування, або й тим й іншим.

Сполука C_n^k з n елементів по k в кожному – всілякі довільні (невпорядковані) сукупності з k елементів, які відрізняються одна від одної лише складом елементів.

Число:

♣ перестановок $P_n = n!$

♣ розміщень $A_n^k = n \cdot (n-1) \cdot (n-2) \dots (n-k+1)$

або $A_n^k = \frac{n!}{(n-k)!} \quad (0 \leq k \leq n)$

♣ сполук $C_n^k = \frac{A_n^k}{P_n} = \frac{n!}{k! \cdot (n-k)!}$

(де $n! = 1 \cdot 2 \cdot \dots \cdot n$ – факторіал числа. При цьому $0! = 1$.)

Властивості сполук:

1. $C_n^0 = C_n^n = 1$.

3. $C_n^0 + C_n^1 + \dots + C_n^n = 2^n$.

2. $C_n^1 = n$.

4. $C_n^k = C_n^{n-k}$ (властивість симетрії).

5. $C_{n+1}^{k+1} = C_n^k + C_n^{k+1}$ (властивість Паскаля) $(0 < k < n)$.

Список літератури

1. Барковський В. В. Теорія ймовірностей та математична статистика : навч. посіб. / В. В. Барковський, Н. В. Барковська, О. К. Лопатін. – 5-те вид., випр. та доп. – К. : Центр учбової літератури, 2016.
2. Валєєв К. Г. Теорія ймовірностей та математична статистика : навч. посібник / К. Г. Валєєв, І. А. Джалладова. – К. : КНЕУ, 2009. – 342 с.
3. Бобик О. І. Теорія ймовірностей і математична статистика : підручник / О. І. Бобик, Г. І. Берегова, Б. І. Копитко. – К. : ВД «Професіонал», 2007.
4. Кремер Н. Ш. Теория вероятностей и математическая статистика : учебник для студ. вузов, обуч. по экон. специальностям / Н. Ш. Кремер. – 3-е изд., перераб. и доп. – М. : Юнити-Дана, 2012. – 551 с.
5. Жлуктенко В. І. Теорія ймовірностей і математична статистика : у 2 ч. Ч. І. Теорія ймовірностей / В. І. Жлуктенко, С. І. Наконечний. – К. : КНЕУ, 2000.
6. Жлуктенко В. І. Теорія ймовірностей і математична статистика : у 2 ч. Ч. 2. Математична статистика / В. І. Жлуктенко, С. І. Наконечний. — К. : КНЕУ, 2000.
7. Михайленко С. В. Теория вероятностей и математическая статистика : пособие для самостоят. изучения дисциплины / С. В. Михайленко, Е. В. Свищева, А. А. Янцевич ; Нар. укр. акад. – Изд. 2-е, испр. – Х. : Изд-во НУА, 2010. – 215 с.
8. Колемаев В. А. Теория вероятностей и математическая статистика / В. А. Колемаев, В. Н. Калинина. – М. : Юнити-Дана, 2012.
9. Гмурман В. Е. Теория вероятностей и математическая статистика / В. Е. Гмурман. – М. : ЮРАЙТ, 2013.

Предметний покажчик

- аксіома
 - адитивності (додавання ймовірностей), 10
 - безперервності, 10
- алгебра подій, 7
- багатовимірна випадкова величина, 22
- багатокутник розподілу, 21
- багатокутник частот, 54
- варіаційний ряд
 - дискретний, 53
 - інтервальний, 53
- вибірка, вибіркова сукупність, 52
- вибіркова дисперсія, 57
- вибіркове середнє, 57
- випадкова величина, 18
- виправлена дисперсія, 60
- випробування, 5
- відносна частота, 8, 53
- генеральна сукупність, 52
- гіпотеза статистична, 75
 - альтернативна, 76
 - нульова, 76
 - схема перевірки, 77, 79, 81
- гіпотези, 14
- гістограма, 54
- дискретна випадкова величина, 20
- дисперсійний аналіз, 67
- дисперсія
 - дискретної випадкової величини, 30
 - неперервної випадкової величини, 37
- діаграма розсіяння, 72
- довірча ймовірність, 59, 64
- довірчий інтервал, 59, 64
- емпіричний (вибірковий) коефіцієнт кореляції, 71
- емпіричний момент
 - початковий, 62
 - центральний, 63
- закон великих чисел, 45, 46
 - теорема Бернуллі, 48
 - теорема Маркова, 47
- теорема Чебишова, 47
 - окремий випадок, 48
- закон розподілу
 - біноміальний, 32
 - дискретної випадкової величини, 21
 - ймовірностей, 21
 - неперервної випадкової величини, 35
 - нормальний (Гаусса), 40
 - показовий (експоненціальний), 39
 - Пуассона, 33
 - рівномірний
 - дискретної випадкової величини, 34
 - неперервної випадкової величини, 38
- залежні випадкові величини, 69
- ймовірність
 - апостеріорна, 15
 - априорна, 14
- імовірнісна залежність, 24
- імовірнісний зв'язок, 67
- імовірність
 - аксіоматична, 10
 - геометрична, 9
 - добуток, 12
 - класична, 9
 - появи хоча б однієї події, 13
 - статистична, 8
 - сумісних подій, 14
- імовірності
 - додавання, 10, 14
- інтеграл ймовірностей, 17, 42
- коваріація, 68
- коефіцієнт кореляції (Пірсона), 68, 69
- корельовані величини, 69
- корельовані випадкові величини, 69
- кореляційна залежність, 68
- кореляційна матриця, 70
- кореляційне поле, 72
- кореляційний аналіз, 67, 71
- кореляційний зв'язок
 - зворотний, 70
 - прямий, 70

- кореляція, 68
- крива нормальна (Гаусса), 41
- крива розподілу, 36
- критерій статистичний, 76
 - достовірність, 77
 - критична область, 76
 - критична точка, 76
 - область допустимих значень, 76
 - область прийняття гіпотези, 76
 - помилка 1 роду, 77
 - помилка 2 роду, 77
 - потужність, 77
 - рівень значущості, 77
 - статистика критерію, 76
- критерій узгодження
 - Колмогорова, 83
 - схема застосування, 84
 - Пірсона χ^2 , 81
 - схема застосування, 82
- математичне очікування
 - дискретної випадкової величини, 28
 - неперервної випадкової величини, 36
- метод моментів, 62
- метод найменших квадратів, 73
- множина елементарних подій, 6
- множинна кореляція, 70
- моделі регресії, 72
- незалежні випадкові величини
 - взаємно незалежні, 25
 - у сукупності, 25
- неперервна випадкова величина, 20
- нерівність
 - Маркова, 45
 - Чебишова, 45
- обсяг вибірки, 52
- операція «риса», 56
- оцінка (параметра), 58
 - ефективна, 59
 - інтервальна, 58, 64
 - дисперсії нормальної ВВ, 66
 - мат. очікування нормальної ВВ, 65, 66, 67
 - незміщена, 59
 - обґрунтована, 59
- точкова, 58
 - математичного очікування, 60
- парна лінійна регресія, 73
- події
 - алгебра подій, 7
 - добуток, 6
 - елементарні, 8
 - залежні, 12
 - незалежні, 11
 - незалежні у сукупності, 12
 - несумісні, 7
 - повна група, 7
 - попарно несумісні, 7
 - протилежні, 7
 - рівноможливі, 5
 - рівносильні (еквівалентні), 6
 - різниця, 7
 - складні, 7
 - сума, 6
 - сумісні, 7
- подія, 5
 - випадкова, 5
 - достовірна, 5
 - імовірність, 9
 - найімовірніше число появи, 16
 - неможлива, 5
- полігон (багатокутник) частот, 54
- початковий момент, 62
- правило трьох сигм, 44
- причинно-наслідкова залежність, 67
- регресійний аналіз, 67, 73
- рівень значущості, 59, 64
- рівняння лінійної парної регресії, 74
- розмах варіації, 53
- розподіл
 - біноміальний, 32
 - нормальний, 40
 - показовий (експоненціальний), 39
 - Пуассона, 33
 - рівномірний
 - дискретної випадкової величини, 34
 - неперервної випадкової величини, 38
- розподіл імовірностей, 20
- ряд розподілу, 21

середнє квадратичне (стандартне)
 відхилення, 31, 37
 стандартний (нормований) нормальний
 розподіл, 42
 статистика критерію, 76
 статистичний зв'язок, 67
 стохастичний зв'язок, 67
 схема випробувань Бернуллі, 15
 теорема
 Бернуллі (закон великих чисел), 48
 Гаусса (центральна гранична), 50
 додавання ймовірностей сумісних
 подій, 14
 інтегральна Муавра – Лапласа, 17
 локальна Муавра – Лапласа, 16
 Ляпунова (центральна гранична), 50
 Маркова (закон великих чисел), 47
 множення ймовірностей, 12
 Пуассона, 18
 формула Бернуллі, 15
 формула повної ймовірності, 14
 Чебишова (закон великих чисел), 47
 теорія ймовірностей, 5
 умова нормування, 36

умовне математичне очікування, 72
 умовне середнє, 73
 формула
 Бернуллі, 15
 повної ймовірності, 14
 Пуассона, 18
 функціональна залежність, 24
 функціональний зв'язок, 67
 функція Лапласа, 17, 42
 функція правдоподібності, 61
 функція регресії, 73
 функція розподілу
 дискретної випадкової величини, 22
 емпірична, 55
 ймовірностей, 19
 неперервної випадкової величини, 34
 нормальної випадкової величини, 42
 теоретична, 55
 центральна гранична теорема, 49
 Гаусса, 50
 Ляпунова, 50
 центральний момент, 62
 частота, 53
 щільність розподілу, 35

Навчальне видання

Дьячкова Ольга Володимирівна

ТЕОРІЯ ЙМОВІРНОСТЕЙ І МАТЕМАТИЧНА СТАТИСТИКА

Опорний конспект

Коректор *О. В. Анцибора*

Комп'ютерний набір і верстання *О. В. Дьячкова*

Макет обкладинки *О. В. Дьячкова*

Формат 60×84/16. Умов. друк. арк. 5,37. Наклад 100 пр. Зам. № 121/19.

Видавець і виготовлювач

Харківський національний університет імені В. Н. Каразіна,
61022, м. Харків, майдан Свободи, 4.

Свідоцтво суб'єкта видавничої справи ДК № 3367 від 13.01.2009.

Видавництво ХНУ імені В. Н. Каразіна

Тел. 705-24-32