ВСТУП

Метою освоєння дисципліни «Емпіричні методи програмної інженерії» є:

• формування у студентів професійних компетенцій, пов'язаних з використанням методів теорії ймовірності та математичної статистики в області програмної інженерії;

• отримання практичних навичок застосування статистичних та емпіричних методів комп'ютингу в області розробки і експериментального дослідження програмних за стосунків ;

• розвиток умінь, заснованих на отриманих теоретичних знаннях, що дозволяють на творчому і репродуктивному рівні застосовувати і створювати ефективне ПЗ для вирішення задач обробки інформації;

• отримання студентами навичок самостійної дослідницької роботи, яка передбачає вивчення специфічних методів математичної статистики, інструментів і засобів, необхідних для вирішення актуальної, в аспекті програмної інженерії, завдання оцінки і прогнозування результатів різних видів діяльності.

Компетенції студента, що формуються в результаті освоєння дисципліни

В результаті освоєння дисципліни студент повинен:

Знати:

- прикладні методи математичної статистики;
- методи аналізу та обробки емпіричної інформації.

Вміти:

- планувати експериментальне дослідження алгоритмічного та програмного забезпечення;
- оцінювати програмне забезпечення з використанням результатів обробки експериментальних даних.

Мати навички (набути досвіду) і володіти:

- інструментами побудови графіків, діаграм і гістограм на основі отриманих експериментальних даних;
- методами і засобами обробки результатів експерименту;
- методами і засобами прогнозування характеристик ПЗ за експериментальними результатами;
- методами розробки ефективних рішень на основі порівняльного аналізу експериментальних результатів.

В результаті вивчення дисципліни студент освоює наступні компетенції:

1. Науково-дослідна діяльність:

- здатність до формалізації у своїй предметної області з урахуванням обмежень використовуваних методів дослідження ;
- готовність до використання методів та інструментальних засобів дослідження об'єктів професійної діяльності ;
- готовність обґрунтувати прийняті проектні рішення, здійснювати постановку і виконання експериментів з перевірки їх коректності та ефективності .

2. Проектна діяльність:

- здатність оцінювати часову і ресурсну складність програмного забезпечення .

3. Технологічна діяльність:

- навички використання різних технологій розробки програмного забезпечення .

4. Виробнича діяльність:

- вміння застосовувати основні методи та інструменти розробки програмного забезпечення .

Місце дисципліни в структурі освітньої програми:

Ця дисципліна відноситься до професійного дисциплін, що забезпечують базову частину навчального плану з напряму 6.050103«Програмна інженерія».

Вивчення даної дисципліни базується на знаннях, отриманих студентами при освоєнні навчальних дисциплін:

- «Теорія ймовірностей і математична статистика»;
- «Основи алгоритмізації»;
- «Комп'ютерна дискретна математика»

і загальних знаннях математики та основ програмування.

Лекція №1.
ТЕМА 1. МЕТОДИ ПЕРЕВІРКИ СТАТИСТИЧНИХ ГІПОТЕЗ

1.1 Основні поняття теорії гіпотез

Побудова точкових і інтервальних оцінок є, як правило, попередньою формою статистичного аналізу. Основне завдання математичної статистики – вміти на основі точкових і інтервальних оцінок приймати рішення в умовах невизначеності, тобто вирішувати статистичні гіпотези. Гіпотези бувають параметричні й непараметричні. Якщо закон розподілу випадкової величини відомий, а необхідно висунути припущення тільки про деякі параметри цього розподілу, то гіпотеза називається *параметричною.* Якщо ж висловлюється припущення про невідомий закон розподілу, то гіпотеза називається *непараметричною.*

Гіпотези бувають *основні* або нуль - гіпотези (H_0) і *альтернативні* (H_1). Нульова гіпотеза завжди одна, а альтернативних може бути декілька. Якщо H_0 відкидається, то має місце протилежна їй гіпотеза H_1.

<u>Приклад.</u> Непараметрична гіпотеза: «Генеральна сукупність розподілена за законом Пуассона»; параметрична гіпотеза: «Дисперсії двох нормальних сукупностей однакові». Основна гіпотеза: «Математичне сподівання дорівнює 10»; конкуруючі: а) «Математичне сподівання менше 10», б) «Математичне сподівання більше 10», в) «Математичне сподівання не дорівнює 10».

<u>Визначення.</u> *Простою* називають гіпотезу, що містить тільки одне припущення (наприклад, H_0: $\lambda = 5$). *Складною або складовою* називають гіпотезу, що складається з кінцевого або нескінченного числа простих гіпотез (наприклад, H_0: $\lambda > 5$).

Висунута гіпотеза може бути правильною або помилковою, тобто виникає необхідність її перевірки статистичними методами. При цьому можуть бути допущені наступні помилки:

1) помилка першого роду – відкинута правильна гіпотеза. Імовірність цієї помилки $P(H_1/H_0) = \alpha$;

2) помилка другого роду – прийнята неправильна гіпотеза. Імовірність такої помилки $P(H_0/H_1) = \beta$.

Імовірність допустити помилку першого роду позначають через α - *рівень значимості.* Другою важливою характеристикою критерію є його *потужність* – величина $\gamma = 1 - \beta$.

1.2 Побудова критичної області

Для перевірки гіпотези H_0 використовують спеціально підібрану випадкову величину, розподіл якої відомий. Цю величину позначають по-різному залежно від її розподілу: *U* або *Z* при нормальному розподілі, *T* – при розподілі Стьюдента й т.і.

<u>Визначення.</u> *Статистичним критерієм* або просто *критерієм* називають випадкову величину *K*, що служить для перевірки гіпотези H_0. *Спостережуваним значенням* $K_{спос}$ називають значення критерію, обчислене за вибіркою.

Визначення. Множина всіх можливих значень критерію розбивають на дві області:

1) область прийняття гіпотези (H_0 вірна);

2) критична область (H_0 помилкова).

Критичними точками $K_{кр}$ називають точки поділу цих областей.

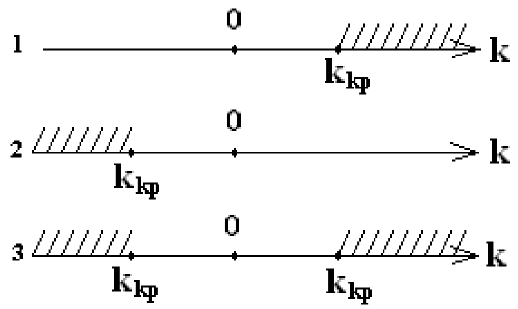

Рис.1.1 Графічна демонстрація критичних областей

Прямі 1 і 2 відповідають однобічним критичним областям - правосторонній й лівосторонній відповідно, пряма 3 - двосторонній критичній області.

Алгоритм перевірки параметричних гіпотез.

1. Висловлюється припущення про закон розподілу.
2. Відбирається вибірка.
3. Формулюються основна (H_0) і конкуруюча (H_1) гіпотези.
4. Будується статистика, що зв'язує досліджуваний параметр і вибіркові значення зі статистичним критерієм. Наприклад,

$$U = \frac{\bar{x} - \mu}{\sigma}\sqrt{n} \qquad \text{для } N(0;1) \tag{1.1}$$

$$T = \frac{\bar{x} - \mu}{S}\sqrt{n} \qquad \text{для } t_\alpha(f)$$

$$\chi^2 = \frac{S^2}{\sigma^2}(n-1) \qquad \text{для } \chi^2_\alpha(f)$$

5. Для обраної статистики будується довірчий інтервал з імовірністю 1 - α виду

$$P(-U_{1-\alpha/2} < U \le U_{1-\alpha/2}) = 1 - \alpha \tag{1.2}$$

$$P(-\infty < U \le U_{1-\alpha}) = 1 - \alpha \tag{1.3}$$

Формула (1.2) відповідає двосторонній критичній області, а формула (1.3) - односторонній.

6. Довірчий інтервал перетворюється в інтервал для досліджуваного параметра, наприклад,

$(\bar{x} - U_{1-\alpha/2}\, \sigma/\sqrt{n} \le m < \bar{x} + U_{1-\alpha/2}\, \sigma/\sqrt{n}) = 1 - \alpha$ (1.4)

7. Ухвалення рішення на основі побудованого інтервалу. Якщо $K_{спос}$ потрапляє в область прийняття гіпотези, то немає підстави відкидати основну гіпотезу H_0. Якщо ж не попадає, то приймається конкуруюча гіпотеза H_1. При цьому ймовірність помилки першого роду дорівнює рівню значимості $P(H_1/H_0) = \alpha$.

Приклад. Вимоги замовника до швидкодії програми 100 мс. З різними початковими умовами програма зпрацьовує за 98; 100; 96; 102; 98 мс. Чи відповідає програма вимогам замовника?

Нехай швидкодія є нормально розподіленою випадковою величиною з невідомим математичним сподіванням і середньо квадратичним відхиленням. Перевіримо гіпотезу H_0: m = 100 мс при конкуруючій гіпотезі H_1: $m \neq 100$. За вибіркою знаходимо середню $\bar{x}$ = 98.8 і оцінку для дисперсії S^2 = 5.2. За таблицею знаходимо квантиль на рівні значимості $\alpha = 0.05$ $t_{1-\alpha/2}(4) = 2{,}776$. За цими даними будуємо критичні точки $\bar{x} \pm t_{1-\alpha/2}\frac{S}{\sqrt{5}}$, тобто 98.8 ± 2.776*2.28/√5 і інтервал, для якого швидкодія відповідає вимогам замовника на рівні значимості 0.05, дорівнює (95.97; 101.63). Оскільки 100 потрапляє в отриманий інтервал, то немає підстав вважати, що швидкодія програми не відповідає вимогам.

1.3 Критерій згоди

У попередніх задачах закон розподілу генеральної сукупності передбачається відомим. Якщо закон розподілу невідомий, але є підстави припускати, що він має певний вигляд, то перевіряють нульову гіпотезу H_0: генеральна сукупність розподілена за законом *A*.

Визначення. *Критерієм згоди* називають критерій перевірки гіпотези про передбачуваний закон невідомого розподілу.

Побудова інтервального статистичного ряду й основних числових характеристик вибірки дозволяє попередньо обрати закон розподілу випадкової величини *X*, виходячи з таких міркувань:

1) За видом гістограми частостей статистичного ряду. Вид гістограми дає орієнтування на можливий закон розподілу. Перевага цього методу полягає в простоті й наочності, а недолік у тому, що гістограма може одночасно нагадувати декілька законів розподілу, або не нагадувати жодного.

2) Порівняння вибіркових оцінок випадкової величини. Наприклад, для нормального закону розподілу порівнюють початкові моменти: $M(x) = \mu_e = \mu_0$. Перевага - простота й наочність, недолік - висновки можуть бути дуже наближені, в яких не відбивається ступінь симетрії кривої розподілу.

3) За емпіричним коефіцієнтом варіації $V = S/\bar{x}$. Відомо, що кожному закону розподілу відповідає певний діапазон значень коефіцієнта варіації:

Перевага методу - простота, недолік - коефіцієнт варіації не відбиває ступінь симетричності емпіричної кривої розподілу, велика неоднозначність вибору.

Таблиця 1.1 Вибір закону за емпіричним коефіцієнтом

Закон розподілу	Границі	Середнє значення
Нормальний	(0.08;0.4)	0.25
Вейбулла	(0.36;0.63) (0.4;0.85)	0.44, 0.71
Логарифмічно-нормальний	(0.35;0.8)	0.68
Експонентний	(0.6;1.3)	0.92

4) За вибірковими коефіцієнтами асиметрії й ексцесу.

$$\hat{A} = \frac{\sum (x_j - \bar{x})^3 m_j}{nS^3} \tag{1.5}$$

$$\hat{E} = \frac{\sum (x_j - \bar{x})^4 m_j}{nS^4} - 3 \tag{1.6}$$

Емпірична функція вважається відповідною гіпотетичній функції, якщо вибіркові коефіцієнти асиметрії й ексцесу відрізняються за абсолютною величиною від своїх математичних сподівань не більше ніж на потроєні середні квадратичні відхилення, тобто якщо:

$$|A - M(A)| < 3S_A \tag{1.7}$$

$$|E - M(E)| < 3S_E \tag{1.8}$$

то обраний закон розподілу погоджується з експериментальними даними. Наприклад, для нормального закону розподілу $M(A) = M(Э) = 0$, тому умови узгодження набувають такого вигляду:

$$|A| < 3S_A\,,\ |E| < 3S_E \tag{1.9}$$

де

$$S_A^2 = \frac{6(n-1)}{(n+1)(n+3)}\ ,\quad S_E^2 = \frac{24n(n-2)(n-3)}{(n-1)^2(n+3)(n+5)}$$

Достоїнство методу - урахування симетрії й крутості, тобто форми кривої, недолік - немає строгої кількісної оцінки припустимої розбіжності між вибірковими коефіцієнтами асиметрії й ексцесу та їхніх математичних сподівань, оскільки використовується правило «трьох сигм», що не дуже точне.

5) За допомогою чисел Вестергарда. Використовується тільки для перевірки відповідності емпіричного закону розподілу до нормального. Числа Вестергарда: 0.3, 0.7, 1.1, 3. Емпіричний розподіл близький до нормального

закону розподілу, якщо в інтервалі ($\bar{x}$ - 0.3S; $\bar{x}$ + 0.3S) розташована чверть всієї сукупності даних, в інтервалі ($\bar{x}$ - 0.7S; $\bar{x}$ + 0.7S) – половина всієї сукупності даних, в інтервалі ($\bar{x}$ - 1.1S; $\bar{x}$ + 1.1S) – три чверті, а в інтервалі ($\bar{x}$ - 3S; $\bar{x}$ + 3S) – 0.998 відсотків всієї сукупності даних.

Приклад обробки емпіричних даних. Представимо вибірку з 55 спостережень у вигляді таблиці частот, використовуючи 7 інтервалів угрупування. Вибірка:

17	19	23	18	21	15	16	13	20	18	15
20	14	20	16	14	20	19	15	19	16	19
15	22	21	12	10	21	18	14	14	17	16
13	19	18	20	24	16	20	19	17	18	18
21	17	19	17	13	17	11	18	19	19	17

Розмах вибірки $\omega = 24 - 10 = 14$. Довжина інтервалу b = 14/7 = 2, результати угруповання зведені в таблицю 13.2

Таблиця 1.2

№ інтервалу i	Границі інтервалу	Середина інтервалу x_i^*	Частота n_i^*	Накопичена частота $\sum_{j=1}^{i} n_j^*$	Відносна частота n_i^* / n	Накопичена відносна частота $\sum_{j=1}^{i} n_j^* / n$
1	10-12	11	2	2	0,0364	0,0364
2	12-14	13	4	6	0,0727	0,1091
3	14-16	15	8	14	0,1455	0,2546
	16-18	17	12	26	0,2182	0,4728
5	18-20	19	16	42	0,2909	0,7637
6	20-22	21	10	52	0,1818	0,9455
7	22-24	23	3	55	0,0545	1,0000

Далі обчислимо середнє, дисперсію, коефіцієнти асиметрії й ексцесу для нашої вибірки. Довжина інтервалу угруповання b = 2. Значення x_i^*, що зустрічається з найбільшою частотою, $\tilde{d}_X = 19$. Тому нормування має вигляд:

$$u_i = \frac{x_i^* - 19}{2}, \qquad i = 1,2,\ldots,7.$$

Обчислення оформимо у вигляді таблиці 1.3

Таблиця 1.3

i	x_i^*	u_i	n_i^*	$n_i^* u_i$	u_i^2	$n_i^* u_i^2$	u_i^3	$n_i^* u_i^3$	u_i^4	$n_i^* u_i^4$	u_i+1	$(u_i+1)^4$	$n_i^* (u_i+1)^4$
1	11	-4	2	-8	16	32	-64	-128	256	512	-3	81	162
2	13	-3	4	-12	9	36	-27	-108	81	324	-2	16	64

3	15	-2	8	-16	4	32	-8	-64	16	128	-1	1	8
4	17	-1	12	-12	1	12	-1	-12	1	12	0	0	0
5	19	0	16	0	0	0	0	0	0	0	1	1	16
6	21	1	10	10	1	10	1	10	1	10	2	16	160
7	23	2	3	6	4	12	8	24	16	48	3	81	243
$\sum$	-	-	55	-32	-	134	-	-278	-	1034	-	-	653
$\sum \frac{n_i^* u_i^l}{n}$			-	-0,582	-	2,436	-	-5,05	-	18,8	-	-	-

Контроль обчислень за тотожністю дає наступне
55+4 (-32)+6 134+4 (-278)-1034=653

Нарешті, скориставшись формулами, знаходимо параметри розподілу:

$$\bar{x} = 19 + 2 \cdot \frac{-32}{55} \approx 17{,}8,$$

$$s_u^2 = \frac{134 - (-32)^2 / 55}{54} \approx 2{,}14, \quad s_u \approx 1{,}46, \quad s_x^2 \approx 2^2 \cdot 2{,}14 \approx 8{,}56,$$

$$\tilde{a}_X = \frac{1}{1{,}46^3}\left[-5{,}054 - 3 \cdot (-0{,}582) \cdot 2{,}436 + 2 \cdot (-0{,}582)^3\right] \approx -0{,}393,$$

$$\tilde{e}_X = \frac{1}{1{,}46^4}\left[18{,}8 - 4 \cdot (-0{,}582) \cdot (-5{,}054) + 6 \cdot (-0{,}582)^2 \cdot 2{,}436 - 3 \cdot (-0{,}582)^4\right] -$$
$$- 3 \approx -0{,}303.$$

1.4 Критерій Пірсона (χ^2)

Нехай $x_1, x_2, \ldots, x_n$ – вибірка спостережень випадкової величини X. Перевіряється гіпотеза H_0, яка стверджує, що випадкова величина X має закон розподілу $F(X)$. Процедура застосування критерію χ^2 для перевірки гіпотези H_0 складається з наступних етапів:

1) За вибіркою спостережень випадкової величини X будують групований (інтервальний) статистичний ряд і знаходять оцінки невідомих параметрів передбачуваного закону розподілу (тобто $\tilde{M}(X)$ *та* $\tilde{Д}(X)$).

2) Використовуючи закон розподілу, що перевіряють, обчислюють імовірності p_k, з якими випадкова величина X потрапляє в кожний інтервал. Наприклад, для нормального закону:

$$p_j = P\{x_{j-1} < x \le x_j\} = \Phi(U_j) - \Phi(U_{j-1}) \qquad (1.10)$$

де $U_j = (x_j - \bar{x})/S$.

3) Визначають теоретичні частоти n^*p_j для кожного часткового інтервалу.

4) Обчислюють вибіркову статистику (критерій) за формулою:

$$\chi_B^2 = \sum_{j=1}^{k} \frac{\left(m_j - np_j\right)^2}{np_j} \tag{1.11}$$

5) Приймають статистичний розв’язок: гіпотеза H_0 не суперечить вибірці спостережень на заданому рівні значимості α, якщо:

$$\chi_B^2 < \chi_{1-\alpha}^2 (r - \ell - 1) \tag{1.12}$$

де r – число інтервалів, ℓ - число параметрів розподілу, які оцінюються за вибіркою. У тому випадку, коли $\chi_B^2 \geq \chi_{1-\alpha}^2 (r - \ell - 1)$, нульова гіпотеза відхиляється.

<u>Зауваження</u>. Необхідно, щоб для всіх інтервалів виконувалася умова $np_j \geq 5$. Якщо для деяких інтервалів ця умова не виконується, то їх варто об'єднати із сусідніми.

<u>Приклад</u>. Перевірити гіпотезу про нормальний розподіл групованих даних, рівень значимості прийняти рівним 0.1. Об'єм вибірки n = 55, середнє значення $\bar{x}$ = 17.84, оцінка для середньо квадратичного відхилення дорівнює S = 2.92.

Таблиця 1.4 Перевірка гіпотези H_0 за критерієм Пірсона

x_j-x_{j+1}	m_j	$U_j;U_{j+1}$	$P_j=\Phi(U_j)$ $-\Phi(U_{j+1})$	np_j		m_j-np_j	$\frac{\left(m_j - np_j\right)^2}{np_j}$
10-12	2	-2.68; -2	0. 0191	1. 0505	14.333	-0. 333	0. 0077
12-14	4	-2; -1.32	0. 0714	3. 927			
14-16	8	-1.32; -0.63	0. 1701	9. 3555			
16-18	12	-0.63; 0.05	0. 2556	14. 058		-2. 058	0. 3013
18-20	16	0.05; 0.74	0. 2504	13. 772		2. 228	0. 3604
20-22	10	0.74; 1.43	0. 1526	8. 393	11.671	1. 329	0. 1513
22-24	3	1.43; 2.11	0. 0596	3. 278			
χ_B^2							0. 8207

Для нормального закону число оцінюваних за вибіркою параметрів дорівнює двом (тобто ℓ = 2), число ступенів волі f = 4 – 2 – 1 = 1. З таблиці χ^2 знаходимо: $\chi_{0.92}$ (1) = 2.71. Оскільки χ_B^2 = 0.919 < $\chi_{0.92}$ (1), то гіпотеза про нормальний закон розподілу не суперечить результатам спостережень.

1.5 Критерій Колмогорова (λ)

Критерій згоди χ^2 дозволяє перевірити гіпотезу про згоду даної вибірки з конкретним теоретичним законом розподілу для будь-якої випадкової величини (як безперервної, так і дискретної). Критерій λ застосовується для перевірки гіпотез про закони розподілу тільки безперервних випадкових величин. Його відмінність полягає в тому, що порівнюються не емпіричні й теоретичні частоти, а емпірична й гіпотетична функції розподілу. Передбачається, що теоретичні значення параметрів функції розподілу генеральної сукупності відомі. Перевірку гіпотези H_0 проводять за наступною схемою:

1) Будують інтервальний статистичний ряд і знаходять емпіричну функцію розподілу:

$$F_n^{\wedge}(X) = \frac{m_x}{n} \tag{1.47}$$

2) Знаходять нормовані інтервали (для перевірки нормального закону розподілу):

$$U = \frac{x - \bar{x}}{S} \tag{1.13}$$

3) Визначають значення теоретичної функції розподілу, які відповідають спостережуваним значенням випадкової величин X.

4) Для кожного x_j (або інтервалу) знаходять модуль різниці між емпіричною й теоретичною функціями розподілу: $\left|F_n^{\wedge}(X) - F(X)\right|$.

5) Обчислюють спостережуване значення вибіркової статистики Колмогорова:

$$\lambda = \max\left|F_n^{\wedge}(X) - F(X)\right| \cdot \sqrt{n} \tag{1.14}$$

6) Порівнюють спостережуване значення вибіркової статистики λ із критичним значенням λ_α, яке визначається за таблицею квантилів розподілу Колмогорова на заданому рівні значимості α. При цьому, якщо $\lambda < \lambda_\alpha$, то немає підстав для відхилення нульової гіпотези, якщо ж $\lambda \geq \lambda_\alpha$, то приймають альтернативну гіпотезу.

Таблиця 1.5 Квантилі розподілу Колмогорова

α	0.20	0.10	0.05	0.02	0.01	0. 001
λ_α	1. 078	1. 224	1. 358	1. 520	1. 827	1. 950

Приклад. Перевірити гіпотезу про нормальний розподіл групованих даних, рівень значимості прийняти 0.1. Об'єм вибірки n = 55, середнє значення $\bar{x}$ = 17.84, оцінка середньо квадратичного відхилення дорівнює S =2.92.

Таблиця 1.6 Перевірка гіпотези H_0 за критерієм Колмогорова

x_j-x_{j+1}	m_j	$U_j;U_{j+1}$	$F_n(X)$ *для* U_j	$F(X)$ *для* U_j	$\left\|F_n^{\wedge}(X) - F(X)\right\|$
10-12	2	-2.68; -2	0. 0364	0. 0465	0. 0101
12-14	4	-2; -1.32	0. 1091	0. 0934	0. 0157
14-16	8	-1.32; -0.63	0. 2545	0. 2643	0. 0098
16-18	12	-0.63; 0.05	0. 4727	0. 5199	0. 0472
18-20	16	0.05; 0.74	0. 7636	0. 7703	0. 0067
20-22	10	0.74; 1.43	0. 9455	0. 9236	0. 0219
22-24	3	1.43; 2.11	1	0. 9826	0. 0174

Для обчислення значень теоретичної функції нормального розподілу використовують функцію Лапласа:

$$F(x)=\begin{cases}0.5-\Phi(x) & x\le 0\\ 0.5+\Phi(x) & x>0\end{cases}$$

$$\lambda = 0.0472\cdot\sqrt{55}=0.35$$

Оскільки $\lambda < 1.224$, то немає підстав відкидати гіпотезу H_0.

1.6 Критерій знаків

У практиці обробки результатів розподіл генеральної сукупності часто невідомий або відрізняється від нормального розподілу випадкової величини. У цих випадках застосовують методи, що не залежать від розподілу, або, так звані *непараметричні*. Ці методи використовують не самі чисельні значення елементів вибірки, а *структурні властивості вибірки* (наприклад, порядок між елементами). Тому потужність таких методів нижча, ніж потужність параметричних гіпотез. Перевагою цих методів є простота й можливість побудови загальних суджень. Більша група непараметричних критеріїв використовується для перевірки гіпотези про приналежності двох вибірок однієї й тієї ж генеральної сукупності, Тобто про рівність функцій розподілу: $F(x)\equiv F(y)|_{y=x}$. Такі генеральні сукупності називаються *однорідними*.

Простішим критерієм такого роду є критерій знаків. Він служить для перевірки гіпотези H_0 про однорідність генеральних сукупностей за попарно зв'язаними вибірками (задача порівняння двох вимірювальних приладів).

Критерій знаків. Використовуються n об'єктів і над кожним проводять по одному виміру за допомогою обох приладів. Якщо вибірки однорідні, то значення x_i і y_i взаємозамінні й імовірності появи того, що x_i-y_i>0 і x_i-y_i<0 рівно.

Статистикою критерію знаків є число знаків «+» у послідовності знаків різностей, які не дорівнюють нулю. За умови, що H_0 вірна, а пари спостережень (x_i; y_i) незалежні, число знаків має біноміальний розподіл з параметрами p=1/2 і n=L (L – число ненульових знаків). Таким чином, нульова гіпотеза має вигляд: H_0: p=1/2, а можливі конкуруючі гіпотези: $H_1^{(1)}$:p>1/2, $H_1^{(2)}$: p<1/2, $H_1^{(3)}$: $p\neq$1/2.

Для перевірки гіпотези H_0 використовують статистику Фішера – Снедекора. Критичні області:

для $H_1^{(1)}$

$$F_B=\frac{R}{L-R+1}\ge F_{1-\alpha}(k_1;k_2) \tag{1.15}$$

де k_1=2(L-R+1), k_2=2R;

для $H_1^{(2)}$:

$$F_B=\frac{L-R}{R+1}\ge F_{1-\alpha}(k_1;k_2) \tag{1.16}$$

де k_1=2(R+1), k_2=2(L–R);

для $H_1^{(3)}$: об'єднання перших двох випадків для рівня значимості $\alpha/2$. Тут R – число знаків «+», k_1, k_2 – ступеня волі.

Приклад. Перевірялася швидкість десяти автомобілів за допомогою двох приладів. Результати перевірки представлені у вигляді наступної таблиці:

V_1	70	85	63	54	65	80	75	95	52	55
V_2	72	86	62	55	63	80	78	90	53	57

Чи можна сказати, що другий прилад дає завищені значення? Рівень значимості прийняти 0,1.

У припущенні, що швидкості руху автомобілів не залежать одна від одної, задачу можна розв'язати, застосовуючи критерій знаків. Складемо послідовність знаків різниці швидкостей автомобілів отриманих на першому й другому приладах:

$v_1 - v_2$:   –, –, +, –, +, 0, –, +, –, –.

Число нульових різностей $L = 9$, число позитивних різностей $R = 3$. Перевіримо гіпотезу про те, що розбіжності в показниках приладів зумовлені випадковими помилками, тобто гіпотезу $H_0 : p = 1/2$. Альтернативна гіпотеза припускає, що показники другого приладу вищі, тоді знаків «+» повинно бути менше, а ймовірність, отже, повинна бути менше 1/2. Отже, альтернативну гіпотезу приймаємо $H_1^{(2)}$: p<1/2 і використовуємо формулу (2.13)

k_1=2(3+1)=8, k_2=2(9–3)=12, F_B=(9–3)/(3+1)=1,5

За таблицею Фішера $F_{0,9}(8;12)$ =2,24, а тому що $F_B < F_{0.9}$ (8;12) немає підстави відкидати гіпотезу H_0. Варто вважати, що розбіжності в показниках приладів викликані випадковими помилками.

1.7 Критерій Вілкоксона

Критерій Вілкоксона застосовується для порівняння двох незалежних вибірок об'єму n_1 і n_2 і перевіряє гіпотезу H_0, що підтверджує, що вибірки отримані з однорідних генеральних сукупностей, і, зокрема, мають рівні середні й медіани.

Статистика W будується в такий спосіб. Розташуємо n_1+ n_2 значень об'єднаної вибірки в порядку зростання (у вигляді варіаційного ряду). Для кожного елемента ряду поставимо його відповідно номер у ряді – *ранг*. Якщо кілька елементів збігається за величиною, то кожному з них привласнюється ранг, що дорівнює середньому арифметичному їхніх номерів. Останній елемент у ранжируваній об'єднаної вибірки повинен мати ранг n_1+ n_2. Далі нехай R_1 – сума рангів першої вибірки, R_2 – другий. Обчислимо такі величини:

$$\omega_1 = n_1 \cdot n_2 + \frac{n_1(n_1+1)}{2} - R_1 \qquad (1.17)$$

$$\omega_2 = n_1 \cdot n_2 + \frac{n_2(n_2+1)}{2} - R_2 \qquad (1.18)$$

Правильність обчислень перевіряють за формулою: $\omega_1 + \omega_2 = n_1 \cdot n_2$. Вибіркове значення ω_B є мінімум з ω_1 і ω_2.

У таблиці розподілу Вілкоксона наводяться ймовірності того, що $W<\omega_У$ за умови, що H_0 вірна, тобто p=P{ $W<\omega_В$ / H_0} для вибірок об'єму n_1 і n_2 ($n_1 \geq n_2$). При однобічній (двосторонній) альтернативній гіпотезі H_1 гіпотеза H_0 відхиляється, якщо $p \leq \alpha$ ($p \leq \alpha/2$), де α – рівень значимості. У протилежному випадку приймається нульова гіпотеза.

Якщо об'єм кожної з вибірок більше восьми, то перевірку гіпотези H_0 можна проводити, використовуючи статистику:

$$Z = \frac{W - n_1 n_2 / 2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}} \tag{1.19}$$

Z має для гіпотези H_0 нормальний розподіл N (0;1). Критичні точки перебувають у таблиці розподілу Лапласа залежно від виду альтернативної гіпотези:

1) двостороння $\Phi(Z_{kp})=(1-\alpha)/2$. Критична область: $|Z_В|>Z_{kp}$;

2) однобічна $\Phi(Z_{kp})=(1-2\alpha)/2$. Критична область: $Z_В<-Z_{kp}$;

або $Z_В>Z_{kp}$ для лівосторонньої й правобічної відповідно.

Приклад. Вимірялася напруга пробою в діодів двох партій, результати вимірів представлені в таблиці:

1 партія	39	50	61	67	40	40	54	-
2 партія	60	53	42	41	40	54	63	69

Чи можна вважати, що для другої партії напруга пробою вища при рівні значимості 0,1?

Для критерію Вілкоксона складемо варіаційний ряд, відзначаючи приналежність до першої партії рискою вгорі:

Елемент	$\overline{39}$	$\overline{40}$	$\overline{40}$	40	41	42	$\overline{50}$	53	$\overline{54}$	54	60	$\overline{61}$	63	$\overline{67}$	69
Ранг	1	3	3	3	5	6	7	8	9.5	9.5	11	12	13	14	15

Знайдемо суми рангів першої й другої вибірок: R_1=49.5, R_2=70.5. Використовуючи формули (14.1) і (14.2) визначимо величини ω_1 і ω_2:

ω_1=7·8+7(7+1)/2-49.5=34.5; ω_2=7·8+8(8+1)/2-70.5=21.5.

Проведемо перевірку: 34.5+21.5=7(8=56.

$\omega_В$ = min(ω_1;ω_2) =21.5.

За таблицею приблизно знаходимо p=P[W<21.5]=0.25. Оскільки альтернативна гіпотеза однобічна, а $p>\alpha$=0.1, то приймаємо гіпотезу H_0.

Приклад. В умовах попереднього прикладу отримані результати нової серії вимірів напруги пробою в діодів (у вольтах):

1-я партія	50	41	48	60	46	60	51	42	62	54	42	46

2-я партія	38	40	47	51	63	50	63	57	59	51	-	-

Чи є підстави стверджувати, що напруга пробою в діодів обох партій різна?

Розв'язати приклад, використовуючи:

а) критерій Вілкоксона;

б) перевірку гіпотези про рівність середніх (у припущенні, що обидві вибірки отримані з нормального розподілених генеральних сукупностей). Прийняти $\alpha = 0{,}1$.

а) упорядкуємо результати вимірів і визначимо ранги кожного результату. Маємо

Елемент	38	40	41	42	42	46	46	47	48	50	50
Ранг	1	2	3	4,5	4,5	6,5	6,5	8	9	10,5	10,5

Елемент	51	51	51	54	57	59	60	60	62	63	63
Ранг	13	13	13	15	16	17	18,5	18,5	20	21,5	21,5

Знайдемо суми рангів $R_1 = 129{,}5$, $R_2 = 123{,}5$.

Якщо $n_1 = 12$, $n_2 = 10$, знаходимо

$$\omega_1 = 12 \cdot 10 + \frac{12 \cdot (12+1)}{2} - 129{,}5 = 68{,}5,$$

$$\omega_2 = 12 \cdot 10 + \frac{10 \cdot (10+1)}{2} - 123{,}5 = 51{,}5.$$

Вибіркове значення ω_B статистики критерію таке:

$$\omega_B = 51{,}5.$$

Якщо $n_1 > 8$ й $n_2 > 8$, то для перевірки гіпотези H_0 використаємо статистику Z. Вибіркове значення цієї статистики визначається за формулою (1.16):

$$z_B = \frac{51{,}5 - \frac{1}{2} \cdot 12 \cdot 10}{\sqrt{\frac{1}{12} \cdot 12 \cdot 10 \cdot (12+10+1)}} \approx -0{,}56.$$

Припущення, що перевіряє, відповідає двосторонній альтернативній гіпотезі, отже, $|z_B|$ значення урівнюється із $u_{1-\frac{\alpha}{2}}$ квантіллю, що визначається за таблицею Д 2: $u_{0,95} = 1{,}645$.

Таким образом, твердження про те, що напруга пробою в діодів обох партій є різною, варто відхилити.

б) Перевіримо гіпотезу про рівність середніх. За результатами спостережень обчислимо оцінки середніх і дисперсій для кожної вибірки:

$$\bar{x}_1 \approx 50{,}17, \quad s_1^2 \approx 55{,}06, \quad \bar{x}_2 = 51{,}90, \quad s_2^2 \approx 76{,}32.$$

Попередньо варто перевірити гіпотезу про рівність дисперсій H_0: $\sigma_1^2 = \sigma_2^2$ при альтернативній гіпотезі H_1: $\sigma_1^2 \neq \sigma_2^2$. Для цього знайдемо відношення вибіркових дисперсій і порівняємо його із квантіллю $F_{1-\frac{\alpha}{2}}(n_2 - 1,\ n_1 - 1)$:

$$\frac{s_2^2}{s_1^2} \approx 1{,}39; \quad F_{0{,}95}(9,\,11) = 2{,}90.$$

Якщо 1,39 < 2,90, то гіпотеза про рівність дисперсій приймається, отже, перевірку гіпотези про рівність середніх можна проводити на основі статистики Стьюдента. Вибіркове значення цієї статистики дорівнює

$$t_B = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} =$$

$$= \frac{50{,}17 - 51{,}90}{\sqrt{\dfrac{(12-1)\cdot 55{,}06 + (10-1)\cdot 76{,}32}{12+10-2}} \cdot \sqrt{\dfrac{1}{12} + \dfrac{1}{10}}} \approx -0{,}5$$

За умови, що значення квантіллі $t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2)$ за таблицею Д6 дорівнює $t_{0{,}95}(20) = 1{,}725$ й $|t_B| < 1{,}725$, гіпотеза про рівність середніх приймається. Таким чином, твердження про те, що напруга пробою в діодів обох партій по-різному, варто відхилити. Це збігається з результатом, отриманим в а).

1.8 Критерій серій

Критерій серій застосовується для перевірки гіпотези H_0, що підтверджує, що елементи вибірки отримані випадковим чином і незалежні. Нехай $x_1, x_2, ..., x_n$ – вибірка результатів спостережень, а $\tilde{h}_X$ – вибіркова медіана, визначена за цими даними. Кожному елементу вибірки поставимо відповідно знак «+» або «–» залежно від того, більше або менше медіани його значення (нульові значення не враховуються). Тим самим всій вибірці поставлений відповідно набір знаків. Позначимо n_1 число знаків «+», а n_2 – число знаків «–» в отриманому наборі знаків. Серією в цьому наборі називається всяка послідовність, що складається з однакових знаків і обмежена протилежними знаками, або перебуває на початку або вкінці набору. Наприклад, у наборі

$+ - + + + - - - - - + +$

утримується 5 серій: (+), (–), (+ + +), (− − − − −), (+ +), $n_1 = 6$, $n_2 = 6$.

Статистикою критерію серій є число серій *N*. Критична область визначається нерівностями $N \le N_1$ й $N \ge N_2$. Значення границь критичної області N_1 й N_2 для рівня значимості α = 0,05 наведені в таблиці Д 9.

<u>Приклад</u>. Швидкості автомобілів у деякій точці траси утворили наступний ряд (км/год):

31, 39, 40, 45, 27, 28, 35, 55, 21, 33, 42, 36.

Чи можна вважати отримані значення випадковими? Прийняти $\alpha = 0{,}05$.

Знайдемо оцінку медіани цих даних. Для цього представимо дані у вигляді варіаційного ряду:

21, 27, 28, 31, 33, 35, 36, 39, 40, 42, 45, 55.

Оцінка медіани $\tilde{h} = \dfrac{35+36}{2} = 35{,}5$. Вихідному ряду спостережень відповідає наступна послідовність знаків:

$- + + + - - - + - - + +,$

де $n_1 = 6$, $n_2 = 6$, число серій $N = 6$. За таблицею Д 9 при $\alpha = 0{,}05$ знаходимо $N_1 = 3$, $N_2 = 11$. Таким чином, гіпотеза H_0 приймається: отримані значення швидкості можна вважати випадковими.

При більших об'ємах вибірки, коли або n_1, або n_2, або обидва значення більше 20, для перевірки гіпотези H_0 можна використати статистику *Z*, вибіркове значення z_B якої обчислюється за формулою

$$z_B = \frac{\left(N - \dfrac{2n_1 n_2}{n_1+n_2} - 1\right) - \dfrac{1}{2}}{\sqrt{\dfrac{2n_1 n_2[2n_1 n_2 - (n_1+n_2)]}{(n_1+n_2)^2(n_1+n_2-1)}}} \tag{1.20}$$

За умови, що вірна гіпотезу H_0, статистика *Z* має приблизно нормальний розподіл *N (0,1)*. У цьому випадку критична область визначається нерівностями

$z_B \le u_{\alpha/2}$ або $z_B \ge u_{1-\alpha/2.}$

<u>Зауваження.</u> Критерій серій застосовується для перевірки випадковості будь – якої вибірки, елементами якої є два різних символи, наприклад: 1 і 0, А і В, + і –. Статистикою критерію є число серій у цій вибірці.

<u>Приклад</u>. Чи можна вважати, що послідовність

110010001010100111011010010000101000

отримана із сукупності випадкових послідовностей? Прийняти $\alpha = 0{,}01$.

У цій послідовності число нулів n_1 = 21, а число одиниць n_2 = 15. Число серій N = 23.

Якщо $n_1 > 20$, для перевірки гіпотези H_0, що підтверджує, що дана послідовність отримана із сукупності випадкових послідовностей, скористаємося статистикою Z. За формулою (1.6) вибіркове значення z_B цієї статистики дорівнює

$$z_B = \frac{(23 - \frac{2 \cdot 21 \cdot 15}{21+15} - 1) - \frac{1}{2}}{\sqrt{\frac{2 \cdot 21 \cdot 15 \cdot [2 \cdot 15 \cdot 21 - (21+15)]}{(21+15)^2 \cdot (21+15-1)}}} \approx 1{,}741.$$

Оскільки $u_{1-\alpha/2.} = u_{0,995} = 2{,}576$, гіпотеза H_0 приймається: можна вважати, що дана послідовність отримана із сукупності випадкових послідовностей.

1.9 Питання для самоперевірки теми 1

1. Які гіпотези називаються параметричними (непараметричними)?
2. Яку гіпотезу називають нульовою (альтернативною)?
3. Що називається рівнем значимості?
4. Що таке помилка першого (другого) роду?
5. Що називається статистичним критерієм?
6. Сформулюйте основний принцип перевірки статистичної гіпотези.
7. Які ви знаєте критичні області?
8. Що називається потужністю критерію?
9. Сформулюйте загальний підхід до статистичної перевірки гіпотези.
10. Наведіть схему перевірки гіпотези про числове значення дисперсії й МС при відомій (невідомій) дисперсії.
11. Наведіть схему перевірки гіпотези про рівність дисперсій при невідомих середніх.
12. Наведіть схему перевірки гіпотези про рівність МС при відомих (невідомих і однакових) дисперсіях.
13. Наведіть схему перевірки гіпотези про рівність МС з невідомими дисперсіями (залежні вибірки).
14. Наведіть схему перевірки гіпотези про числове значення ймовірності події.
15. Наведіть схему перевірки гіпотези про значимість коефіцієнта кореляції.
16. Наведіть схему перевірки гіпотези про рівність дисперсій декількох нормальних генеральних сукупностей за критеріями Бартлета й Кочрена.
17. Що називається критерієм згоди?
18. Назвіть основні критерії попереднього вибору закону розподілу.
19. Наведіть алгоритм перевірки гіпотези про закон розподілу за критерієм згоди Пірсона.

20. Які обмеження накладаються на теоретичні частоти при перевірці гіпотези за критерієм Пірсона?
21. Наведіть алгоритм перевірки гіпотези про закон розподілу за критерієм Колмогорова.
22. У чому відмінність критеріїв Пірсона й Колмогорова?
23. Які методи математичної статистики називають непараметричними?
24. Які генеральні сукупності називаються однорідними?
25. Для перевірки якої гіпотези служить критерій знаків?
26. Що є статистикою критерію знаків?
27. Яку гіпотезу перевіряє критерій Вілкоксона?
28. Як будується статистика за критерієм Вілкоксона?
29. Дайте визначення серії.
30. Для перевірки якої гіпотези застосовується критерій серій?
31. Що є статистикою критерію серій?
32. Як визначається критична область за критерієм серій?
33. Що характеризує ранговий коефіцієнт кореляції?