

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Київський національний університет будівництва і архітектури

Методичні вказівки
до самостійної роботи
з математичної статистики
для здобувачів першого (бакалаврського) рівня вищої освіти
за спеціальностями 073 «Менеджмент»,
051 «Економіка», 071 «Облік і оподаткування»

Київ 2024

УДК 517.521

Б81

Укладачі: З.І.Наголкіна, канд. фіз.-мат. наук, доцент
Л.В.Соколова, канд. фіз.-мат. наук, доцент
Ю.П. Філонов, канд. фіз.-мат. наук, доцент

Рецензент В.В. Отрашевська, канд. фіз.-мат. наук, доцент

Відповідальний за випуск Н.В. Бондаренко, канд. фіз.-мат. наук,
доцент

*Затверджено на засіданні кафедри вищої математики, протокол
№ 11 від 21 березня 2024 року.*

В авторській редакції

Методичні вказівки до самостійної роботи з математичної
Б81 статистики / уклад. : З.І. Наголкіна та ін. – Київ : КНУБА, 2024. – с.
44

Викладено розділ математичної статистики, який включає в себе обробку і аналіз статистичних даних одновимірної і двовимірної вибірки. Містить стислі теоретичні відомості і приклади обробки вибірок, а також набори вибірок для самостійної роботи.

Призначено для здобувачів першого (бакалаврського) рівня вищої освіти за спеціальностями 073 «Менеджмент», 051 «Економіка», 071 «Облік і оподаткування».

Зміст

Вступ	
Тема 1. Статистичний аналіз і дослідження одновимірної вибірки ...	
1.1. Первинна обробка одновимірної вибірки.....	
1.2. Вибіркові числові характеристики.....	
1.3. Геометрична інтерпретація вибірки.....	
Тема 2. Перевірка статистичної гіпотези про розподіл методом узгодження Пірсона.	
2.1. Загальні відомості	
2.2. Критерій згоди Пірсона	
2.3. Перевірка гіпотези про нормальний розподіл, за критерієм Пірсона	
Вибірки до завдань 1-2	
Тема 3. Елементи теорії кореляції	
3.1. Первинна обробка двовимірної вибірки.....	
3.2. Рівняння лінійної регресії.....	
Вибірки для завдань 3-4	
Додаток1	
Додаток 2	
Список літератури.....	

Вступ

Математична статистика – це розділ математики, в якому на основі дослідних даних вивчаються ймовірнісні закономірності випадкових масових явищ. Математична статистика застосовує апарат теорії ймовірностей для дослідження і побудови математичних моделей. Важливою задачею є методи відбору і групування статистичних даних залежно від мети дослідження випадкових вибірок. Такий аналіз стає можливим на основі збору, систематизації і обробки статистичних даних, що одержані в результаті експериментів та спостережень. Задачами математичної статистики є статистична перевірка гіпотез, оцінка розподілу статистичних ймовірностей, вивчення статистичної залежності, визначення основних числових характеристик випадкових вибірок, якими є таке: вибіркове середнє, вибірккові дисперсії, стандартне відхилення, коефіцієнт кореляції. Методи математичної статистики широко застосовуються в організації виробництва, економіці, геодезії, будівництві, соціології.

У методичній роботі докладно розібрані задачі обробки й аналізу необхідної інформації за наявності одновимірної і двовимірної вибірки статистичних даних. Це можуть бути числові дані, отримані в результаті різних економічних досліджень, аудиту, соціологічних опитувань, кількісних характеристик параметрів вимірювань в технологічних і фізичних процесах. Маючи матеріал у вигляді великих таблиць чисел, виникає запитання, як його систематизувати і яку інформацію можна з цього матеріалу отримати. Відповідь на ці запитання продемонстрована на прикладах розв'язаних статистичних задач.

Мета навчального видання – надати студентам фундаментальні знання і практичні навички з основних методів математичної статистики, які необхідні для ефективного аналізу експериментальних чи спостережних даних.

Тема 1. Статистичний аналіз і дослідження одновимірної вибірки

Математична статистика теоретично базується на основних поняттях і методах теорії ймовірностей. Всі необхідні поняття з теорії ймовірностей наведені в навчальному посібнику [3], укладеному авторами. Основні поняття з математичної статистики, що необхідні для виконання індивідуальної роботи, наведені в підручнику [4].

Завдання 1. Нехай задана вибірка з генеральної сукупності X (табл. 1), яка містить $n = 30$ елементів. За даними вибірки визначити вибіркове математичне сподівання, вибіркову дисперсію, вибіркові моду, медіану, коефіцієнти асиметрії і ексцесу. А також побудувати гістограму і кумулятивну (накопичувальну криву розподілу).

Таблиця 1

73,34	67,21	68,98	74,45	71,81	68,59	66,25	68,07	71,45	70,04
72,81	70,37	68,77	69,32	70,56	70,71	68,99	67,99	68,56	69,82
72,43	72,65	68,51	67,94	66,71	69,97	68,42	68,93	70,45	70,93

Розв'язання.

1.1. Первинна обробка одновимірної вибірки

Перше завдання з математичної статистики полягає в первинній обробці інформації. Залежно від предмету дослідження це можуть бути дані з аудиту певного підприємства або результати вимірів деяких технічних характеристик, або похибки під час виміру відстаней, або кутів. Розглядають велику кількість таких даних. Вони можуть бути внесені в журнали спостережень або задані електронними таблицями. Всі ці дані називаються генеральною сукупністю. Але з них випадковим чином, за законами математичної статистики, відбирають обмежену вибірку. Ця вибірка повинна бути репрезентативною – тобто адекватно представляти властивості генеральної сукупності.

Зміст першого завдання полягає в тому, щоб визначити, яку інформацію ми можемо отримати, досліджуючи вибірку і як таку інформацію можна буде розповсюдити на всю генеральну сукупність.

Різні можливі значення x_i генеральної сукупності X , які потрапили до вибірки, називаються *варіантами*. Ряд варіант, x_1, x_2, \dots, x_k , розташованих в порядку зростання їх значень, називається *варіаційним статистичним рядом*. Число спостережень кожної варіанти називається *частотою* і позначається n_i ($i = 1, 2, \dots, k$). Відношення частоти до всього об'єму вибірки $w_i = \frac{n_i}{n}$ називають *відносною частотою*. В класичній теорії ймовірностей вона має зміст статистичної ймовірності $w_i = p_i$.

Зрозуміло, що сума відносних частот має бути 1. Сукупність варіант і відповідних ймовірностей називається *рядом розподілу*.

Якщо кількість варіант достатньо велика, то для зручності дослідження їх треба згрупувати. Вибірку доцільно розбити на інтервали з певним кроком. Далі обчислюють частоту (кількість) варіант, що попали в певний інтервал. Такий ряд називається *інтервальним варіаційним рядом*. Для зручності досліджень ряд, як правило, розбивають на 6–10 інтервалів.

1.2. Побудова варіаційного ряду

Спочатку потрібно побудувати варіаційний ряд, впорядкувавши вибірку (табл. 2).

Варіаційний ряд

Таблиця 2

66,25	66,71	67,21	67,94	67,99	68,07	68,42	68,51	68,56	68,59
68,77	68,93	68,98	68,99	69,32	69,82	69,97	70,04	70,37	70,45
70,56	70,71	70,93	71,45	71,81	72,43	72,65	72,81	73,34	74,45

Побудуємо інтервальний варіаційний ряд. Для цього з вибірки визначаємо мінімальне та максимальне значення $x_{min} = 66,25$; $x_{max} = 74,45$. У деяких випадках можна округлити значення на кінцях всього інтервалу вибірки, але обов'язково крайні точки вибірки мають знаходитись в інтервалі $[66,25; 74,45] \subseteq [66; 75]$.

Розмах вибірки $\omega = 75 - 66 = 9$. Розіб'єм з кроком h вибірку на k інтервалів. У даному прикладі візьмемо $h = 1$, а $k = 9$.

Інтервальний варіаційний ряд оформлюється у вигляді табл. 3.

Таблиця 3

$[x_{i-1}, x_i]$	66,67	67,68	68,69	69,70	70,71	71,72	72,73	73,74	74,75
n_i	2	3	9	3	6	2	3	1	1

Зауваження. Якщо групується більша кількість елементів для вибірки обсягом n , то можна взяти інший крок h . Це визначає точність відповідних обчислень.

Примітка. У випадку побудови інтервального варіаційного ряду, якщо об'єм вибірки достатньо великий, необов'язково спочатку впорядковувати варіанти. Достатньо розбити всю вибірку формально на інтервали і обчислювати кількість варіант, що попала в кожний інтервал. Далі обчислюється за даними вибірки кількість елементів, яка попадає в кожний інтервал. Причому якщо варіанта попадає на границю двох інтервалів, то можна домовитись, щоб зараховувати її в правий інтервал. Умовна частота, або ймовірність, визначається за формулою:

$$p_i = \frac{n_i}{n}, \quad \sum_{i=1}^k p_i = 1, \quad \sum_{i=1}^k n_i = n = 30.$$

Представником кожного інтервалу є його середина

$$x_i^* = \frac{x_i + x_{i+1}}{2}.$$

Далі знаходимо вибіркові числові характеристики. Спочатку знаходимо вибіркове математичне сподівання, яке позначається x_B (x_B – середнє вибіркове), і обчислюється за наступною формулою:

$$M_B = x_B = \sum_{i=1}^{k=9} x_i^* p_i.$$

У деяких задачах середнє значення інтерпретують як прогнозований, або очікуваний, результат, а дисперсія характеризує відхилення від прогнозу.

Вибіркова дисперсія:

$$D_B = \sum_{i=1}^{k=9} (x_i^* - x_B)^2 p_i.$$

Вибіркове середнє квадратичне відхилення:

$$\sigma_B^2 = D_B.$$

Також обчислюються відповідні центральні моменти m_3 і m_4 , які надалі можуть бути використані для дослідження таких характеристик вибірки, як асиметрія і ексцес:

$$m_3 = \sum_{i=1}^{k=9} (x_i^* - x_g)^3 p_i, \quad m_4 = \sum_{i=1}^{k=9} (x_i^* - x_g)^4 p_i.$$

Треба відмітити, що, оскільки вибірка з генеральної сукупності взята випадковим чином, всі числові характеристики є також випадковими величинами. Це означає, що, коли взяти іншу вибірку з генеральної сукупності, відповідно будуть інші числові характеристики.

Складемо табл. 4 для обчислення числових характеристик.

Таблиця складена в Excel, але можна обчислити, користуючись калькулятором.

1-й рядок – це інтервали $[x_i, x_{i+1}]$, на які розбили задану вибірку.

2-й рядок – n_i відповідні частоти.

3-й – відносні частоти, або ймовірності.

Представником кожного інтервалу є його середина. Надалі знаходять числові характеристики випадкової величини, заданої розподілом (x_i^*, p_i) .

Розподіл, пов'язаний з вибіркою, називається емпіричним. Треба

переконатись, що $\sum_{i=1}^{k=9} p_i = 1$. Всі числові характеристики, які отримані з

вибірки, називаються *емпіричними*.

Сума в 5 рядку – математичне сподівання, а в 8, 9, 10 відповідно дисперсія, третій і четвертий центральні моменти.

Таблиця 4

$x_i - x_{i+1}$	66-67	67-68	68-69	69-70	70-71	71-72	72-73	73-74	74-75	сума
n_i	2	3	9	3	6	2	3	1	1	30
p_i	0,07	0,10	0,30	0,10	0,20	0,07	0,10	0,03	0,03	1,00
x_i^*	66,50	67,50	68,50	69,50	70,50	71,50	72,50	73,50	74,50	
$x_i^* p_i$	4,43	6,75	20,55	6,95	14,10	4,77	7,25	2,45	2,48	69,73
$x_i^* - x_{\theta}$	-3,23	-2,23	-1,23	-0,23	0,77	1,77	2,77	3,77	4,77	6,90
$(x_i^* - x_{\theta})^2$	10,45	4,99	1,52	0,05	0,59	3,12	7,65	14,19	22,72	
$(x_i^* - x_{\theta})^2 p_i$	0,70	0,50	0,46	0,01	0,12	0,21	0,77	0,47	0,76	3,98
$(x_i^* - x_{\theta})^3 p_i$	-2,25	-1,11	-0,56	0,00	0,09	0,37	2,12	1,78	3,61	4,04
$(x_i^* - x_{\theta})^4 p_i$	7,29	2,49	0,69	0,00	0,07	0,65	5,86	6,71	17,21	8,01

1.3. Вибіркові числові характеристики

Для даної вибірки (таблиця 1) обчислені відповідні числові характеристики. Вибіркове математичне сподівання – сума в 5-ому рядку:

$$x_B = 69,73.$$

Вибіркова дисперсія – сума в 8-му рядку:

$$D_e = \sigma_e^2 = 3,98.$$

Вибіркове середньоквадратичне відхилення:

$$\sigma_e = \sqrt{D_e} = 1,99; \sigma_e^3 = 1,99 \times 3,98 = 7,9; \sigma_e^4 = 3,98 \times 3,98 = 15,08.$$

Вибірковий коефіцієнт асиметрії у випадку інтервального ряду:

$$a_{sB} = \frac{m_3}{\sigma_B^3} = \frac{4,04}{7,9} = 0,5.$$

Якщо $a_{sB} > 0$, то коефіцієнт асиметрії вказує на розтягненість емпіричного розподілу праворуч відносно математичного сподівання, якщо $a_{sB} < 0$, то емпіричний розподіл витягнутий ліворуч від математичного сподівання.

Вибірковий коефіцієнт ексцесу e_{kB} у випадку інтервального ряду:

$$e_{kB} = \frac{m_4}{\sigma_B^4} - 3 = \frac{8,04}{15,08} - 3 = -2,4.$$

Він визначає відхилення від нормального розподілу. Для нормального розподілу $e_{kB} = 0$. Якщо $e_{kB} > 0$, то крива емпіричного розподілу має більш круту і гостру вершину, ніж нормальний. Якщо $e_{kB} < 0$, то вершина емпіричного розподілу більш полого, ніж для нормального.

Мода $M = 68,5$ – варіанта, що відповідає найбільшій частоті.

Медіана $Me = 70,5$ – ділить варіаційний ряд навпіл.

В таблиці 4 наведено приклад обчислення числових характеристик в програмі Excel. Далі будуюмо гістограму. Це геометричне зображення емпіричного розподілу. Гістограма представляє собою ступінчасту фігуру, складену з прямокутників, основа яких є інтервал, довжиною h , а висоти – відповідні відносні частоти, або ймовірності, поділені на h :

$$\frac{n_i}{nh} = \frac{p_i}{h}.$$

Таким чином площа гістограми є:

$$S = \sum_{i=1}^{k=9} (x_{i+1} - x_i) \frac{p_i}{h} = \sum_{i=1}^{k=9} \frac{\omega}{9} \frac{n_i}{hn} = \sum_{i=1}^{k=9} \frac{\omega}{9h} \frac{n_i}{n} = \frac{1}{n} \sum_{i=1}^9 n_i = 1.$$

Згідно з означенням кроку розбиття k , $h = 1$. В даному прикладі $9h = \omega$.

1.4. Геометрична інтерпретація вибірки

В якості геометричного зображення розподілу для варіаційного ряду використовують багатокутник розподілу. Для інтервального варіаційного ряду використовують гістограму (рис. 1).

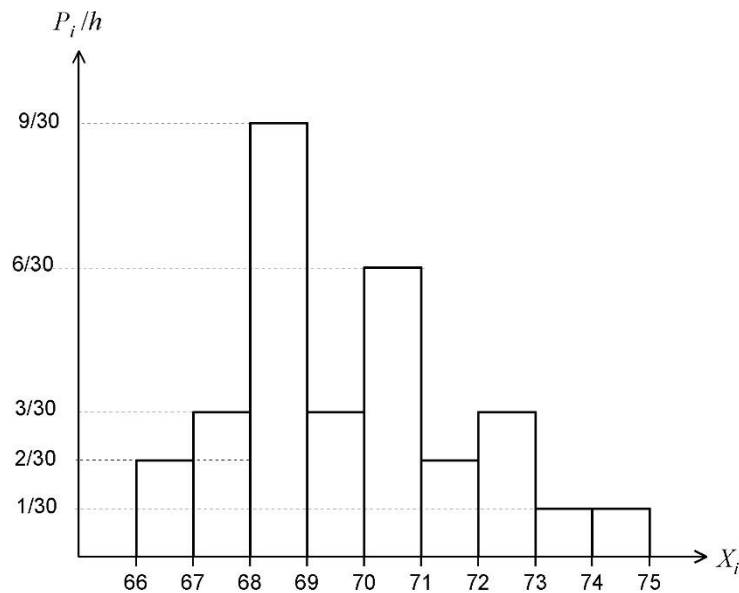


Рис. 1. Гістограма частот (до задачі № 29)

Проте і для інтервального ряду можна побудувати багатокутник емпіричного розподілу, з'єднавши точки з координатами (x_i^*, p_i) .

Кумулятивна крива розподілу (накопичувальна крива) – ламана з вершинами в точках:

$$\left(x_i^* + \frac{h}{2}, \frac{1}{n} \sum_{i=1}^k n_i \right), i = 1, \dots, k.$$

У деяких задачах математичної статистики важливо оцінити накопичувальний ефект. Наприклад, ймовірна загроза життю людини при

накопиченні деяких шкідливих речовин або економічний фінансовий кумулятивний ефект, що досягається поступовим накопиченням фондів або ресурсів. Цей накопичувальний ефект характеризується так званою кумулятивною кривою (рис. 2). В даному випадку це ламана, яка з'єднує точки з координатами:

$$A_1(66,5 + 0,5; 0,07) = A_1(67; 0,07)$$

$$A_2(67,5 + 0,5; 0,07 + 0,1) = A_2(68; 0,17)$$

$$A_3(69; 0,17 + 0,3) = A_3(69; 0,47)$$

$$A_4(70; 0,47 + 0,1) = A_4(70; 0,57)$$

$$A_5(71; 0,57 + 0,2) = A_5(71; 0,77)$$

$$A_6(72; 0,77 + 0,07) = A_6(72; 0,84)$$

$$A_7(73; 0,84 + 0,1) = A_7(73; 0,94)$$

$$A_8(74; 0,94 + 0,03) = A_8(74; 0,97)$$

$$A_9(75; 0,97 + 0,03) = A_9(75; 1)$$

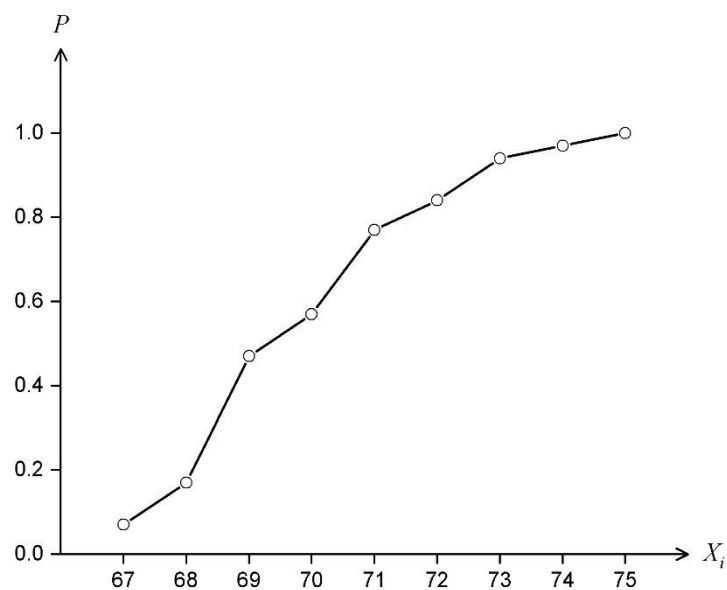


Рис. 2. Кумулятивна крива, або крива накопичувальних частот

Зауважимо, що кумулятивна крива завжди зростаюча (борги накопичуються).

Тема 2. Перевірка статистичної гіпотези про розподіл методом узгодження Пірсона

2.1. Загальні відомості

Зважаючи на геометричний характер розподілу, а саме з вигляду гістограми, або з багатокутника розподілу, можна зробити припущення про тип емпіричного розподілу. Це припущення називається статистичною гіпотезою.

Взагалі статистичною називають гіпотезу про характер розподілу, або про значення параметрів розподілу.

Позначимо гіпотезу про вигляд розподілу H_0 і назвемо цю гіпотезу основною. Альтернативною називають протилежну гіпотезу, або таку, що логічно заперечує основну. Нехай альтернативна гіпотеза – H_1 .

У результаті перевірки гіпотези може бути прийняте одне з двох правильних рішень:

- гіпотеза приймається і вона правильна;
- гіпотеза відхиляється і вона хибна.

При цьому, в результаті перевірки можуть бути допущені такі помилки:

- помилки 1 роду: відхиляється правильна гіпотеза;
- помилки 2 роду: приймається хибна гіпотеза.

Виявляється, що помилки 1 роду більш суттєві, ніж 2 роду. І тому, щоб мінімізувати наслідки помилок 1 роду – вводять рівень значущості $\alpha = 0,01; 0,05; 0,001$. Тобто ймовірність зробити помилку 1 роду дорівнює рівню значущості. Ймовірність P (відхилити правильну гіпотезу) $= \alpha$. В розпорядженні є дані вибірки. І тому перевірку гіпотез проводять за даними вибірки. Це і є статистична перевірка гіпотези.

Для перевірки гіпотези H_0 вводять числову характеристику, яку обчислюють на основі вибірки і за допомогою якої вирішують прийняти основну гіпотезу або альтернативну. Зрозуміло, що ця характеристика залежить від конкретної вибірки, тобто вона є випадковою величиною. Така випадкова величина називається критерієм $K = K(X)$. K обирають таким, щоб його закон розподілу був відомий. Значення критерію, обчислені за даними вибірки, називаються k_{cn} . При одних значеннях k_{cn}

гіпотеза приймається, при інших відхиляється. Сукупність тих значень критерію, при яких гіпотеза відхиляється, називається критичною. Область, при яких гіпотеза приймається, називається областю прийняття рішень.

Оскільки критерій є одновимірною випадковою величиною, то всі його можливі значення належать інтервалам, тому критична область і область прийняття рішень – інтервали. Точка, що відділяє ці дві області, називається $k_{кр}$.

Розглядають правосторонню, лівосторонню і двосторонню критичні області. В цій роботі ми обмежимося вивченням правосторонньої критичної області.

Якщо $k_{cn} \in (k_{кр}, \infty)$ – правостороння критична область ($k_{кр} > 0$) і $P(k_{cn} > k_{кр}) = \alpha$, то H_0 – гіпотеза відхиляється, якщо ж $k_{cn} < k_{кр}$, тоді немає підстав відхилити гіпотезу H_0 .

Для знаходження критичної області необхідно визначити критичні точки. Для однієї гіпотези можна побудувати різні критерії і тому вводять поняття потужності критерію. Потужністю критерію називають ймовірність потрапляння критерію в критичну область, за умови, що альтернативна гіпотеза є істиною. Побудова правостороннього критерію $P(K > k_{кр}) = \alpha$ – є складною математичною задачею.

Існують різні критерії для перевірки статистичних гіпотез: Пірсона, Стьюдента, Фішера, Колмогорова. Критерії – це випадкові величини, які мають певний розподіл. Для кожного критерію є таблиці, за якими знаходять критичну точку.

2.2. Критерій згоди Пірсона

Критерієм згоди називають статистичний критерій про перевірку закону розподілу випадкової величини, яка представлена вибіркою з генеральної сукупності. Існує кілька критеріїв згоди.

Ми розглянемо критерій згоди Пірсона, який базується на порівнянні емпіричних і теоретичних частот. Припустимо, є певна вибірка з генеральної сукупності $X(x_1, x_2, \dots, x_n)$. Гіпотеза H_0 полягає в тому, що

ця випадкова величина має певний розподіл A . Складемо з вибірки інтервальний варіаційний ряд:

$[x_0, x_1]$	$[x_1, x_2]$...	$[x_{k-2}, x_{k-1}]$	$[x_{k-1}, x_k]$
n_1	n_2		n_{k-1}	n_k

$$\sum_{i=0}^k n_i = n.$$

Вважаючи, що теоретичний розподіл A має вигляд $F(x) = P(X \leq x)$, можна знайти ймовірність попадання в інтервал $x_{i-1} \leq X \leq x_i$ за загальною формулою:

$$p_i = P(x_{i-1} \leq X \leq x_i) = F(x_i) - F(x_{i-1}).$$

Зазначимо, що під час обчислення ймовірностей треба покласти $x_0 = -\infty$ і $x_k = \infty$. Це необхідно для того, щоб функцію розподілу A визначити на всій області. Тоді $\sum_{i=0}^k p_i = 1$. Відповідні частоти для

теоретичного (гіпотетичного) розподілу будуть $n'_i = np_i$. Складемо таблицю.

Таблиця 5

$[x_0, x_1]$	$[x_1, x_2]$...	$[x_{k-2}, x_{k-1}]$	$[x_{k-1}, x_k]$
n_1	n_2	...	n_{k-1}	n_k
$n'_1 = np_1$	$n'_2 = np_2$...	$n'_{k-1} = np_{k-1}$	$n'_k = np_k$

Критерій згоди Пірсона позначається χ^2 (читається «хі-квадрат») і має вигляд:

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - n'_i)^2}{n'_i},$$

де k – кількість інтервалів, на які розбито вибірку, n_i – емпіричні частоти, $n'_i = np_i$ – теоретичні частоти, отримані з урахуванням теоретичного розподілу.

Теоретичні і емпіричні частоти наведені в табл. 5.

З формули критерію видно, що коли теоретичні та емпіричні частоти співпадають, то $\chi^2 = 0$, тобто чим ближче χ^2 до 0, тим краще

узгоджуються вибіркові і теоретичні дані. χ^2 – це випадкова величина. Пірсон дослідив цю випадкову величину. Вона позначається $\chi^2(\alpha, m)$ і існують таблиці її розподілу. Це двопараметричний закон. Він залежить від ступеня вільності m і рівня значущості α .

Ступенем вільності називають число $m = k - r - 1$, де k – кількість інтервалів розбиття, r – кількість параметрів теоретичного розподілу A . Наприклад, для нормального розподілу $r = 2$, а для показникового $r = 1$.

Для критерію χ^2 будують правосторонню критичну область за правилом $P[\chi^2 > \chi_{кр}^2] = \alpha$. За даним рівнем значущості α і за ступенем вільності $m = k - r - 1$ в таблиці критичних точок значень розподілу Пірсона знаходять $k_{кр} = \chi_{кр}^2(\alpha, m)$.

За даними вибірки обчислюють емпіричне значення критерію Пірсона $\chi_{сп}^2 = \sum_{i=1}^k \frac{(n_i - n'_i)^2}{n'_i}$. Далі порівнюють значення $\chi_{сп}^2$ і табличне значення $k_{кр}$. Якщо $\chi_{сп}^2 \geq k_{кр}$, то гіпотезу H_0 відхиляють. Якщо $\chi_{сп}^2 \leq k_{кр}$, то немає підстав відхилити дану гіпотезу.

2.3. Перевірка гіпотези про нормальний розподіл, за критерієм Пірсона

Завдання 2. Перевірити статистичну гіпотезу про нормальний розподіл для вибірки, заданої табл. 1.

Розв'язання.

Розглянемо вибірку, задану табл. 1. Візьмемо до уваги відповідний інтервальний варіаційний ряд (табл. 3) і вибіркові числові характеристики, обчислені в табл. 4. Як відомо, щільність нормального закону розподілу має вигляд

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}},$$

де $a = MX$, $\sigma^2 = DX$ параметри розподілу.

Зробимо припущення (гіпотезу), що вибіркова статистика № 29 має нормальний розподіл. Тоді цей емпіричний розподіл, який отримано з вибірки і геометрично зображений гістограмою, має такі параметри: $a = x_B = 69,73$; $\sigma^2 = \sigma_B^2 = 3,98$; $\sigma_B = 1,99$.

З урахуванням цього

$$f(x) = \frac{1}{1,99\sqrt{2\pi}} e^{-\frac{(x-69,73)^2}{2 \times 3,98}}.$$

Відомо, ймовірність попадання в інтервал (α, β) нормально розподіленої випадкової величини визначається за допомогою функції Лапласа за формулою:

$$P(\alpha \leq X \leq \beta) = \Phi\left(\frac{\beta - MX}{\sqrt{DX}}\right) - \Phi\left(\frac{\alpha - MX}{\sqrt{DX}}\right),$$

$$a = MX = x_B, \sigma_B = \sqrt{DX} = \sigma.$$

Таким чином, знаходимо ймовірність попадання в частковий інтервал (x_i, x_{i+1}) :

$$P(x_i \leq X \leq x_{i+1}) = \Phi\left(\frac{x_{i+1} - MX}{\sqrt{DX}}\right) - \Phi\left(\frac{x_i - MX}{\sqrt{DX}}\right) = \Phi(z_{i+1}) - \Phi(z_i).$$

Знайдемо емпіричні ймовірності попадання в частковий інтервал при гіпотетичному розподілі з параметрами $a = x_B = 69,73$, $\sigma^2 = \sigma_B^2 = 3,98$.

Скориставшись попередньою формулою, отримуємо:

$$p'_i = P(x_i \leq X \leq x_{i+1}) =$$

$$= \Phi\left(\frac{x_{i+1} - 69,73}{1,99}\right) - \Phi\left(\frac{x_i - 69,73}{1,99}\right) = \Phi(z_{i+1}) - \Phi(z_i).$$

Позначимо $z_i = \frac{x_i - 69,73}{1,99}$. Для отримання гіпотетичних ймовірностей і відповідно гіпотетичних частот складемо табл. 6 відносно нових змінних. При цьому треба врахувати, що $z_0 = -\infty$, $z_9 = \infty$. Гіпотетичні частоти і ймовірності ще називають теоретичними.

Для визначення $\Phi(z_i)$ треба скористатися таблицями функцій

Лапласа (дод. 1).

Таблиця 6

	z_i	z_{i+1}	$\Phi(z_i)$	$\Phi(z_{i+1})$	p'_i	$p'_i n = n'_i$
1	$-\infty$	-1,371	-0,5	-0,4147	0,0853	2,56
2	-1,371	-0,86	-0,4147	-0,3051	0,1096	3,289
3	-0,86	-0,367	-0,3051	-0,1443	0,1608	4,824
4	-0,367	0,135	-0,1443	0,0537	0,198	5,94
5	0,135	0,638	0,0537	0,2389	0,1852	5,556
6	0,638	1,141	0,2389	0,3749	0,136	4,08
7	1,141	1,642	0,3749	0,4495	0,0747	2,241
8	1,642	2,145	0,4495	0,4838	0,0353	1,059
9	2,145	∞	0,4838	0,5	0,0152	0,456

Перевірка:

$$\sum_{i=1}^9 p'_i = 1,0001 \approx 1; \quad \sum_{i=1}^9 p'_i n = \sum_{i=1}^9 n'_i = 30,005 \approx 30.$$

Частоти, що отримані з вибірки n_i , називають емпіричними частотами, а частоти $n'_i = np_i$ – теоретичними або гіпотетичними.

В табл. 6 обчислені теоретичні ймовірності і відповідні теоретичні частоти

$$p'_i \approx \frac{n'_i}{n}.$$

Критерій згоди Пірсона визначається за формулою:

$$\chi_{cn}^2 = \sum_{i=1}^9 \frac{(n_i - n'_i)^2}{n'_i}.$$

Знайдемо всі елементи цієї формули за табл. 7.

Критерій Пірсона правосторонній. Тобто, якщо $\chi_{cn}^2 > \chi_{кр}^2$, то гіпотеза про вигляд розподілу відхиляється.

В табл. 7 обчислено χ_{cn}^2 :

$$\chi_{cn}^2 = \sum_{i=1}^9 \frac{(n_i - n'_i)^2}{n'_i} = 7,21.$$

Таблиця 7.

	n_i	n_i'	$n_i - n_i'$	$(n_i - n_i')^2$	$\frac{(n_i - n_i')^2}{n_i'}$
1	2	2,56	-0,56	0,313	0,122
2	3	3,289	-0,289	0,083	0,025
3	9	4,824	4,176	17,44	3,61
4	3	5,94	-2,94	8,64	1,45
5	6	5,556	0,444	0,197	0,035
6	2	4,08	-2,08	4,32	1,06
7	3	2,241	0,759	0,576	0,257
8	1	1,059	0,059	0,0035	0,0033
9	1	0,456	0,544	0,295	0,646

При рівні значущості $\alpha = 0,05$ по таблицях Пірсона при числі ступенів вільності $m = 6$: $\chi_{кр}^2 = 12,592$. Значення критерію $\chi_{сн}^2$, що спостерігаються за даними вибірки: $\chi_{сн}^2 = 7,27$.

Висновок: якщо $\chi_{сн}^2 \leq \chi_{кр}^2$ при даному $\alpha = 0,05$ нема підстав відхилити гіпотезу про нормальний розподіл.

Кількість інтервалів, на які розбито вибірку – 9, кількість зв'язків – $2 + 1 = 3$. Кількість ступенів вільностей $m = 9 - 3 = 6$.

Примітка: для показникового закону $F(x) = 1 - e^{-\lambda x}$, $\lambda = \frac{1}{a}$ – один параметр, і тому кількість зв'язків $1 + 1 = 2$, кількість ступенів вільності було б $9 - 2 = 7$. При $\alpha = 0,05$ значення $\chi_{кр}^2$ при різних ступенях вільності наведені в табл. 8.

Таблиця 8

$m = 4$	$\chi_{кр}^2 = 9,48$
$m = 5$	$\chi_{кр}^2 = 11,1$
$m = 6$	$\chi_{кр}^2 = 12,6$

Табл. 8 є фрагментом повної таблиці. Повна таблиця для критичних значень критерія Пірсона наведена в дод. 2. Слід відмітити, що критерій узгодження Пірсона може бути використано для перевірки статистичних

гіпотез про інші вигляди розподілів.

Вибірки до завдань 1–2.

Варіант 1.

72,01	68,70	68,92	70,77	69,42	69,55	74,72	70,98	71,62	68,15
70,08	69,01	73,31	70,83	70,99	67,56	69,25	69,26	68,74	71,31
72,43	68,85	70,46	72,19	68,60	70,69	73,75	69,10	68,18	72,47

Варіант 2.

67,89	69,90	68,93	67,20	74,16	68,78	69,49	72,616	71,49	66,87
67,83	68,78	70,01	71,03	72,89	66,51	72,19	73,85	68,36	69,56
73,68	74,09	71,55	67,99	71,55	71,58	70,84	69,42	73,11	75,35

Варіант 3.

70,92	68,88	73,29	68,27	72,86	72,13	71,32	68,14	71,17	67,36
74,73	71,87	67,99	67,08	67,57	70,15	70,56	70,18	68,51	71,84
70,25	70,33	67,61	71,58	69,52	68,29	72,51	71,75	70,62	69,60

Варіант 4.

70,92	68,88	73,29	68,27	72,86	72,13	71,32	68,14	71,17	67,36
74,73	71,87	67,99	67,08	67,57	70,15	70,56	70,18	68,51	71,84
70,25	70,33	67,61	71,58	69,52	68,29	72,52	71,75	70,62	69,60

Варіант 5.

65,46	68,21	67,69	72,59	69,92	70,92	70,31	71,14	68,10	70,78
72,93	70,40	71,53	73,98	69,16	69,26	71,73	72,55	71,29	72,16
70,61	69,30	70,75	71,56	72,43	70,22	71,01	71,52	69,99	72,19

Варіант 6.

71,78	66,46	69,38	69,28	69,08	70,89	73,89	72,17	68,63	73,65
71,35	68,78	69,70	67,29	67,61	69,47	68,30	69,79	74,98	72,31
66,86	71,89	70,34	68,75	69,76	70,04	69,78	70,51	68,75	71,61

Варіант 7.

68,37	68,40	71,24	69,96	66,62	73,76	70,53	69,44	71,10	75,39
68,52	69,64	71,98	71,24	68,23	71,55	72,31	65,96	73,61	70,49

70,70	71,97	69,20	68,59	70,51	67,88	72,47	70,21	69,00	71,63
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Варіант 8.

65,44	69,38	68,40	69,51	69,66	66,65	67,82	71,40	71,36	65,95
68,62	73,28	67,03	69,45	68,80	71,33	70,43	70,92	70,96	71,14
67,53	71,29	71,05	71,87	73,97	69,08	68,31	70,64	69,94	70,57

Варіант 9.

65,97	70,78	72,03	70,98	66,05	68,39	69,55	70,15	71,75	66,38
71,74	68,31	71,76	68,92	69,48	72,87	67,87	68,84	67,73	70,88
68,37	70,79	71,70	70,91	69,37	69,05	70,91	69,02	70,37	68,93

Варіант 10.

71,40	70,19	72,75	68,80	72,29	68,66	70,00	71,37	72,87	65,60
72,44	71,42	67,92	68,66	74,30	70,86	67,06	73,41	68,28	70,97
69,45	66,78	71,57	69,57	71,89	72,76	70,86	70,46	69,58	71,55

Варіант 11.

71,55	71,12	70,38	70,25	69,93	72,32	72,39	68,39	73,04	70,54
69,03	71,62	69,55	70,75	72,59	70,63	68,34	69,91	69,60	69,42
71,13	64,86	68,76	66,04	71,72	71,88	69,02	68,40	71,20	68,03

Варіант 12.

70,13	70,17	69,35	68,61	70,87	67,29	69,81	70,31	68,23	73,68
72,01	71,73	74,87	70,28	67,73	69,08	68,14	67,30	65,09	68,80
70,99	72,61	69,09	68,71	67,15	73,15	72,39	69,77	69,64	72,56

Варіант 13.

72,04	69,81	66,42	71,09	68,73	65,69	68,48	72,74	68,53	70,30
70,48	66,26	66,90	70,02	71,61	70,22	71,51	68,17	66,03	71,97
68,58	71,39	70,17	70,86	73,17	70,73	75,06	69,17	74,25	68,00

Варіант 14.

67,40	71,42	70,13	70,91	70,45	73,42	71,44	72,84	69,41	71,23
70,08	72,12	71,02	68,52	70,08	70,70	69,49	67,94	69,42	68,64
72,01	67,95	72,23	71,48	71,64	70,52	65,81	68,85	71,42	69,44

Варіант 15.

72,16	69,89	67,46	72,73	72,37	71,91	72,73	68,22	65,19	70,13
70,67	69,03	72,48	67,17	69,30	71,22	70,20	71,17	72,50	71,18
67,98	66,74	68,05	70,14	66,47	68,82	69,72	69,04	69,68	70,25

Варіант 16.

69,96	72,24	66,41	68,56	69,61	66,83	68,57	70,48	69,49	66,72
69,72	69,67	70,87	69,52	64,81	71,51	67,76	72,07	73,54	70,44
71,36	70,61	70,51	71,83	73,13	74,08	65,30	71,83	72,00	71,52

Варіант 17.

67,91	68,65	70,40	69,21	71,40	72,28	70,67	70,86	67,74	69,98
69,03	66,81	67,31	67,06	74,65	67,26	68,82	71,80	69,79	71,60
70,13	72,62	68,32	71,04	70,93	66,19	71,44	70,07	72,29	69,93

Варіант 18.

70,99	71,18	66,10	71,50	69,56	69,41	71,49	70,25	70,31	71,26
71,67	70,73	71,93	69,44	69,20	70,41	67,84	68,91	72,47	68,95
68,91	67,94	69,86	68,84	71,94	68,13	72,16	70,63	67,40	69,45

Варіант 19.

70,18	71,58	69,12	67,15	70,37	66,72	70,93	69,86	67,01	69,88
68,86	68,53	68,86	68,11	70,28	74,78	71,75	68,90	67,48	71,28
68,73	66,88	72,30	70,84	67,87	67,23	68,01	72,67	69,35	68,86

Варіант 20.

68,24	68,87	67,66	70,56	68,29	69,82	67,87	72,02	68,68	71,21
69,72	70,80	72,73	70,5	68,98	69,70	68,18	68,49	68,65	68,88
67,20	67,70	72,01	69,01	71,53	71,62	72,46	72,57	69,36	69,73

Варіант 21.

66,72	68,75	72,18	73,07	70,80	66,74	69,33	71,09	68,97	71,62
71,31	69,15	70,65	67,25	72,76	67,89	68,95	70,07	70,36	73,50
72,99	70,14	69,77	67,68	69,38	72,42	67,17	72,47	71,26	71,76

Варіант 22.

71,94	70,45	69,02	71,39	68,89	68,17	70,60	69,90	71,36	68,65
68,81	66,91	68,51	70,06	70,88	69,92	70,74	69,49	71,27	71,08
67,19	67,98	70,58	69,74	68,46	68,91	70,65	72,32	69,56	66,71

Варіант 23.

66,27	67,81	73,21	66,70	71,08	69,80	67,99	69,10	70,87	73,13
66,95	70,01	71,28	70,16	69,39	69,74	70,06	70,21	71,41	71,28
67,28	69,48	72,84	71,60	68,48	69,38	66,83	71,70	68,16	70,62

Варіант 24.

71,94	66,27	68,27	70,12	66,54	70,47	68,66	69,10	66,52	69,91
69,65	68,72	71,71	71,50	69,29	69,46	70,43	68,16	72,93	69,75
71,76	67,20	69,86	70,40	72,69	68,83	71,38	72,58	70,70	70,87

Варіант 25.

69,70	68,13	69,17	71,08	64,73	70,21	71,96	68,12	68,97	69,41
68,16	72,93	68,10	69,29	70,09	67,52	72,10	72,17	69,68	66,69
68,94	70,48	72,66	70,26	70,94	73,22	70,83	70,09	68,61	66,64

Варіант 26.

73,30	71,27	70,40	68,48	69,80	69,93	71,09	69,84	70,86	69,56
71,75	67,48	71,84	70,34	71,23	71,74	69,12	68,28	68,12	70,69
71,47	68,75	73,07	73,06	70,03	68,77	70,02	71,78	69,13	71,42

Варіант 27.

68,69	68,98	69,21	67,56	70,33	68,49	71,72	73,96	65,72	68,93
68,84	66,57	67,55	68,74	68,31	70,38	68,11	71,06	71,64	69,14
73,44	69,18	71,94	70,84	70,16	69,81	67,62	66,95	70,81	69,51

Варіант 28.

68,03	67,97	68,28	69,12	70,29	69,35	67,77	69,46	68,95	70,02
69,78	71,97	69,89	73,60	70,10	70,85	68,16	67,72	69,80	67,24
70,23	64,10	70,59	73,36	69,12	68,57	72,52	70,70	71,75	71,36

Тема 3. Елементи теорії кореляції

3.1. Первинна обробка двовимірної вибірки

В багатьох прикладних задачах необхідно виявити залежність або між двома властивостями одного об'єкта, або між певними ознаками різних об'єктів. Наприклад, під час аудиту підприємства можна визначити зв'язок між заробітною платою і якістю виробленої продукції, або зв'язок між розміром податків і розмірами кредитів для клієнтів певних установ. Якщо певні ознаки допускають кількісне вимірювання, то за статистичними даними можна визначити міру зв'язку між різними ознаками. Зміст обробки багатовимірних вибірок і полягає у встановленні цього зв'язку. Нехай є двовимірна вибірка X, Y . Зв'язок між X і Y може бути функціональний, коли кожному значенню однієї величини відповідає одне значення іншої $Y = f(X)$, а може бути статистичним, коли зміна однієї величини змінює розподіл іншої. Якщо є випадкові величини, то зв'язок між ними має випадковий характер.

Розділ теорії ймовірностей, який вивчає зв'язок між випадковими величинами, називається кореляційним аналізом. Статистичні зв'язки між випадковими величинами також називаються кореляцією.

Нехай є двовимірна вибірка з генеральної сукупності. Ставиться задача – яку інформацію можна отримати з цієї вибірки як про кожну випадкову величину окремо, так і про зв'язок між ними. Двовимірну випадкову величину геометрично слід інтерпретувати як точку площини з випадковими координатами або як випадковий вектор. Якщо представляти результати вимірювання пари X, Y на координатній площині, то це будуть точки з координатами (x_i, y_j) , де $i, j: 1, 2, \dots, n$.

Приклад обробки інформації, яка представлена двовимірною вибіркою, розглянуто в завданні 3.

Завдання 3. За даними двовимірної вибірки визначити числові характеристики X і Y , кореляційний момент і коефіцієнт кореляції. У табл. нижче задано 30 пар статистичних даних (X, Y) .

Варіант 29.

№	1	2	3	4	5	6	7	8	9	10
X	0,58	0,82	3,19	5,72	6,30	5,35	5,89	6,57	8,99	9,93
Y	1,49	4,48	14,58	25,37	27,84	23,80	26,09	29,01	39,33	43,33
№	11	12	13	14	15	16	17	18	19	20
X	10,30	11,76	11,89	12,88	13,90	15,68	13,97	17,79	17,78	20,48
Y	44,90	51,15	51,68	55,92	60,30	67,85	60,59	76,86	76,82	88,36
№	21	22	23	24	25	26	27	28	29	30
X	21,05	22,28	22,33	23,27	25,08	26,49	27,88	27,61	28,67	28,66
Y	90,79	96,04	96,24	100,26	107,99	114,01	119,94	118,77	123,28	123,25

Розв'язання.

Законом розподілу двовимірної випадкової величини називають перелік всіх можливих пар значень випадкових величин $(x_i; y_i)$, а також ймовірностей, з якими приймаються ці значення $p_{ij} = P(X = x_i; Y = y_j)$ $i = 1, 2, \dots, n; j = 1, 2, \dots, m$. Цей закон розподілу задають таблицею з подвійним входом. Така таблиця називається кореляційною.

За даними таблиці (В29) неможливо визначити розподіл або ступінь їх залежності. Для впорядкування даних потрібно скласти кореляційну таблицю. Зробимо її за аналогією завдання 1, у якому за одновимірною вибіркою складають варіаційний статистичний ряд, а в двовимірному випадку треба скласти кореляційну таблицю. Якщо кількість пар невелика, то кореляційну таблицю можна скласти одразу. Якщо кількість пар достатньо велика ($n > 20$), то для побудови кореляційної таблиці їх треба згрупувати. В даному випадку кореляційна таблиця буде двовимірним аналогом інтервального варіаційного ряду розподілу. В першому рядку стоять всі значення $X(x_1, x_2, \dots, x_k)$, в першому стовпчику всі значення $Y(y_1, y_2, \dots, y_m)$, а на перетині – сумісна частота n_{ij} . Відносна частота n_{ij}/n має зміст статистичної ймовірності. Надалі будемо позначати її p_{ij} і називати сумісною ймовірністю.

Ступінь залежності випадкових величин X, Y визначається кореляційним моментом

$$\mu_{XY} = \mu(X, Y) = M(X - MX)(Y - MY)$$

і коефіцієнтом кореляції

$$r_{XY} = \frac{\mu_{XY}}{\sigma_X \sigma_Y}.$$

Для цього на основі кореляційної таблиці треба знайти відповідно розподіли окремо для X і Y , а потім їхні числові характеристики.

Побудова кореляційної таблиці. В даній задачі кількість пар $n = 30$. Утворимо кореляційну таблицю для заданої вибірки. Для згрупування даних визначимо $x_{min} = -0,58$, $x_{max} = 28,67$ і розмах вибірки для X :

$$\Delta = 28,67 - 0,58 = 29,25.$$

Розбиваємо вибірку X на $7 \div 10$ інтервалів, зручніше зробити це так, щоб крок був парним. Візьмемо найближчий зручний відрізок, який покриває задану вибірку $[-0,58; 29,25] \subset [-1; 31]$. Тоді розіб'ємо вибірку X на 8 інтервалів. А саме:

$$[-1,3][3,7][7;11][11;15][15;19][19;23][23;27][27;31].$$

Аналогічно згрупуємо вибірку для Y :

$$y_{min} = -1,49; y_{max} = 123,25; \Delta = 123,25 - (-1,49) = 124,74.$$

Найбільшим цілим числом, що покриває довжину інтервалу $[-1.49, 123.25] \subset [-2, 124]$, є 126. Аналогічно попередньому знайдемо таку найближчу довжину відрізка покриття, яка дасть можливість розбити на кількість інтервалів $7 \div 10$ з парним кроком. Це буде $126 = 7 \times 18$:

$$[-2;16][16;34][34;52][52;70][70;88][88;106][106;124].$$

Складемо прямокутну кореляційну табл. 9 згрупованих даних: перший рядок – інтервали змінної X , x^* – середина відповідного інтервалу. Перший стовпчик – інтервали змінної Y , y^* – середина відповідного інтервалу.

Таблиця 9

Y/X	$[-1; 3]$ $x^* = 1$	$[3; 7]$ $x^* = 5$	$[7; 11]$ $x^* = 9$	$[11; 15]$ $x^* = 13$	$[15; 19]$ $x^* = 17$	$[19; 23]$ $x^* = 21$	$[23; 27]$ $x^* = 25$	$[27; 31]$ $x^* = 29$
$[-2; 16] y^* = 7$.. 2	.1						
$[16; 34] y^* = 25$	 5						
$[34; 52] y^* = 43$... 3	.. 2				
$[52; 70] y^* = 61$... 3	. 1			
$[70; 88] y^* = 79$.. 2			
$[88; 106] y^* = 97$					 4	. 1	
$[106; 124] y^* = 115$.. 2 4

Кількість точок в кожній клітинці визначає кількість пар (n_{ij} – частота), що потрапляють в даний прямокутний інтервал. Це і називають кореляційним полем. Наприклад, перший інтервал – $X \in [-1;3], Y \in [-2;16]$. Перевіряємо пари вибірки на приналежність до кожного інтервалу. Таким чином перевіримо всі пари. Їхня кількість має бути $\sum_{i=1}^8 \sum_{j=1}^7 n_{ij} = 30$. Якщо кожен частоту поділити на 30, то отримаємо

сумісну ймовірність $p_{ij} = n_{ij}/n$. Представником кожного розряду для x і y будуть середні арифметичні $x_i^* = \frac{x_{i-1} + x_i}{2}, y_j^* = \frac{y_{j-1} + y_j}{2}$.

Примітка. Якщо значення X або Y потрапляє на границю інтервалу, то попередньо необхідно визначити приналежність цього значення до попереднього або наступного проміжку.

Складемо розподіли відповідно для X і Y . У випадку завдання двовимірної випадкової величини елементами кореляційного поля є сумісні ймовірності

$$p_{ij} = P(X = x_i, Y = y_j), \frac{n_{ij}}{n} = p_{ij} = P(X = x_i, Y = y_j),$$

$$\sum_{i=1}^k \sum_{j=1}^m n_{ij} = n, \text{ а } \sum_{i=1}^k \sum_{j=1}^m p_{ij} = 1.$$

З кореляційної таблиці можна визначити розподіл для кожної випадкової величини X і Y окремо. В якості представників інтервалів візьмем їхні середини. Це будуть відповідно варіанти X і Y в табл. 10 і 11. Частоти – є суми відповідних частот для кожної варіанти. Відповідна ймовірність є:

$$p(X = x_i^*) = p_i = \frac{n_i}{n}, p(Y = y_j^*) = p_j = \frac{n_j}{n}, i:1-8; j:1-7.$$

Розподіл X

Таблиця 10

x_i^*	1	5	9	13	17	21	25	29
n_i	2	6	3	5	3	4	3	4

$$n_i = \sum_{j=1}^7 n_{ij}, \sum_{i=1}^8 n_i = 30.$$

Таблиця 11

y_j^*	7	25	43	61	79	97	115
n_j	3	5	5	4	2	5	6

$$n_j = \sum_{i=1}^8 n_{ij}, \quad \sum_{j=1}^7 n_j = 30.$$

Користуючись таблицями розподілів для кожної варіанти, знайдемо відповідні числові характеристики:

$$MX = \frac{1}{n} \sum_{i=1}^8 x_i^* n_i, \quad DX = \frac{1}{n} \sum_{i=1}^8 (x_i^* - MX)^2 n_i = \frac{1}{n} \sum_{i=1}^8 (x_i^*)^2 n_i - MX^2.$$

$$MY = \frac{1}{n} \sum_{j=1}^7 y_j^* n_j, \quad DY = \frac{1}{n} \sum_{j=1}^7 (y_j^* - MY)^2 n_j = \frac{1}{n} \sum_{j=1}^7 (y_j^*)^2 n_j - MY^2.$$

Як відомо, дисперсія може бути обчислена за двома формулами. Обираємо ту, яка зручніша в даному випадку. За відповідними таблицями 10, 11 знайдемо

$$MX = \frac{1}{30} (1 \times 2 + 5 \times 6 + 9 \times 3 + 13 \times 5 + 17 \times 3 + 21 \times 4 + 23 \times 3 + 25 \times 4) = 14,3;$$

$$MY = \frac{1}{30} (7 \times 3 + 25 \times 5 + 43 \times 5 + 61 \times 4 + 79 \times 2 + 97 \times 5 + 115 \times 6) = 64,6;$$

$$DX = \frac{\sum_{i=1}^8 x_i^2 n_i}{30} - MX^2;$$

$$DX = \frac{1}{30} (1^2 \times 2 + 5^2 \times 6 + 9^2 \times 3 + 13^2 \times 5 + 17^2 \times 3 + 21^2 \times 4 + 23^2 \times 3 + 25^2 \times 4) - 14,3^2 = 114,49;$$

$$DY = \frac{1}{30} (7^2 \times 3 + 25^2 \times 5 + 43^2 \times 5 + 61^2 \times 4 + 79^2 \times 2 + 97^2 \times 5 + 115^2 \times 6) - 64,6^2 = 1369,4;$$

$$\sigma(X) = \sqrt{DX} \approx 10,68; \quad \sigma(Y) = \sqrt{DY} \approx 37.$$

Кореляційний момент визначає зв'язок між двома випадковими величинами і позначається:

$$\mu_{XY} = \mu(X, Y) = M(X - MX)(Y - MY).$$

Якщо випадкові величини X та Y незалежні, то $\mu_{XY} = 0$. Дійсно, якщо X та Y – незалежні, то:

$$M((X - MX) \times (Y - MY)) = M(X - MX) \times M(Y - MY) = 0$$

Користуючись кореляційною табл. 9 і відповідно числовими характеристиками для X і Y знайдемо кореляційний момент за формулою:

$$\mu(X, Y) = \sum_{i=1}^n \sum_{j=1}^m x_i y_j p_{ij} - MX \times MY = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m x_i y_j n_{ij} - MX \times MY,$$

$$p_{ij} = \frac{n_{ij}}{n},$$

де n_{ij} – сумісна частота, тобто кількість пар, що попали в даний прямокутник. Наприклад, за табл. $n_{23} = 5$, $n_{87} = 4$. Обчислимо кореляційний момент за даними табл.:

$$\begin{aligned} \mu_{XY} = \frac{1}{30} & (1 \times 7 \times 2 + 5 \times 7 \times 1 + 5 \times 25 \times 5 + 9 \times 43 \times 3 + 13 \times 43 \times 2 + \\ & + 13 \times 61 \times 3 + 17 \times 61 \times 1 + 17 \times 79 \times 2 + 21 \times 97 \times 4 + 25 \times 97 \times 1 + \\ & + 25 \times 115 \times 2 + 29 \times 115 \times 4) - MX \times MY = 366,82. \end{aligned}$$

У даному випадку кореляційний момент є вибірковою, тобто таким, що залежить від вибірки і її групування. Кореляційний момент має розмірність добутку розмірностей величин X і Y .

Коефіцієнт кореляції визначається $r_{XY} = \frac{\mu_{XY}}{\sigma_X \sigma_Y}$ – безрозмірна

величина, що характеризує зв'язок між X і Y . Причому $|r_{XY}| \leq 1$.

В даному випадку

$$r_{XY} = \frac{\mu_{XY}}{\sigma(X) \sigma(y)} = \frac{366,82}{10,68 \times 37} = 0,92 \leq 1.$$

Якщо $r_{XY} = 1$, то між величинами існує точна лінійна залежність.

Залежно від значення коефіцієнта кореляції можна зробити висновки про ступінь зв'язку: якщо $0 < |r_{XY}| \leq 0,5$ – зв'язок слабкий, $0,5 < |r_{XY}| \leq 0,7$, – середній і, якщо $0,7 < |r_{XY}| \leq 1$, – сильний. Найбільш тісним зв'язком вважають лінійну залежність.

Зауваження. Визначення коефіцієнта кореляції важливе в задачах надійності окремих конструкцій і споруд в будівельній галузі, а також при менеджменті будівельної галузі. Коефіцієнт кореляції входить в значення характеристики безпеки β_s [5]:

$$\beta_s = \frac{\bar{R} - \bar{F}}{\sqrt{\sigma_R^2 + \sigma_F^2 - 2r_{RF}\sigma_R\sigma_F}},$$

де \bar{R} – середнє значення несучої здатності конструкції, \bar{F} – середнє зовнішнє навантаження. Ймовірність неруйнування конструкції визначається:

$$P(R, F) = \frac{1}{2} + \Phi(\beta_s),$$

де $\Phi(x)$ – функція Лапласа (завдання 2). Якщо існує статистичний матеріал (R, F) , то можна визначити коефіцієнт надійності певної споруди або конструкції.

3.2. Рівняння лінійної регресії

Завдання 4.

Визначити рівняння лінійної регресії методом найменших квадратів для вибірки **Завдання 3**.

Розв'язання.

Вигляд функціональної залежності між X і Y , який встановлений за двовимірною вибіркою, називають емпіричною формулою. Регресія характеризує форму зв'язку між випадковими величинами. Розрізняють прямолінійну, криволінійну та інші типи регресії. В даній роботі треба скласти рівняння лінійної регресії.

Якщо побудувати графік емпіричної залежності між X і Y на кореляційному полі, то він не буде обов'язково проходити через всі точки (x_i, y_i) вибірки. Але він може бути найкращим наближенням, тобто $Y \approx kX + b$, або $y_i \approx kx_i + b$. Для визначення невідомих коефіцієнтів k і b застосовують метод найменших квадратів. Цей метод детально розглянуто в [4]. Рівняння лінійної регресії Y на X по згрупованим даним має вигляд:

$$(y - y_e) = r_{XY} \frac{\sigma(Y)}{\sigma(X)} (x - x_e).$$

В даному випадку: $y_e = MY$, $x_e = MX$, $\sigma(Y) = \sigma_B(Y)$, $r_{XY} = r_{BXY}$ відповідно вибіркові числові характеристики (отримані з даних кореляційної таблиці).

Як відомо $y = kx + b$ – рівняння прямої з кутовим коефіцієнтом, а $(y - y_0) = k(x - x_0)$ – рівняння прямої, що проходить через задану точку з кутовим коефіцієнтом k . З урахуванням числових характеристик даної вибірки маємо

$$y - 64,6 = k(x - 14,3);$$

$$k = r_{XY} \frac{\sigma(Y)}{\sigma(X)} = 0,92 \frac{37}{10,68} = 3,18;$$

$$y - 64,6 = 3,18(x - 14,3).$$

Зауваження. Для даної двовимірної вибірки коефіцієнт кореляції близький до 1 і тому доцільно розглядати рівняння лінійної регресії. Якщо представлення вибірки в кореляційному полі має нелінійний характер, то залежність може бути представлена нелінійною функцією.

Вибірки для завдань 3–4.

Варіант 1.

№	1	2	3	4	5	6	7	8	9	10
X	1,00	1,34	2,67	3,27	4,56	4,60	6,58	8,56	9,94	11,89
Y	6,12	6,76	9,27	10,41	12,83	12,91	16,65	20,38	23,00	26,67
№	11	12	13	14	15	16	17	18	19	20
X	10,41	11,33	11,60	13,57	16,94	16,89	16,46	19,44	18,56	20,25
Y	23,88	25,61	26,12	29,84	36,21	36,11	35,29	40,92	39,33	42,45
№	21	22	23	24	25	26	27	28	29	30
X	20,00	22,43	23,71	24,37	25,36	26,68	26,44	27,87	28,56	29,96
Y	41,97	46,58	48,98	50,22	52,10	54,59	54,13	56,83	58,15	60,79

Варіант 2.

№	1	2	3	4	5	6	7	8	9	10
X	1,37	3,06	2,42	3,23	4,74	7,80	6,92	9,22	8,34	10,61
Y	5,31	8,01	7,00	8,29	10,69	15,57	14,18	17,84	16,44	20,05
№	11	12	13	14	15	16	17	18	19	20
X	11,58	11,84	13,46	16,10	15,18	16,29	16,47	19,32	18,14	20,26
Y	21,60	22,02	24,60	28,82	27,35	29,11	29,40	33,94	32,07	35,45
№	21	22	23	24	25	26	27	28	29	30
X	21,94	22,66	21,99	23,90	24,01	27,72	26,73	29,90	30,39	29,37
Y	38,12	39,28	38,21	41,25	41,44	47,34	45,77	50,83	51,61	49,98

Варіант 3.

№	1	2	3	4	5	6	7	8	9	10
X	1,58	2,07	5,12	4,28	5,44	8,26	6,30	8,02	8,81	10,30
Y	5,72	6,40	10,68	9,49	11,13	15,08	12,32	14,73	15,84	17,93
№	11	12	13	14	15	16	17	18	19	20
X	10,41	12,02	15,12	15,03	13,95	15,53	18,59	19,03	20,55	19,80
Y	18,08	20,34	24,68	24,56	23,04	25,26	29,55	30,16	32,28	31,24
№	21	22	23	24	25	26	27	28	29	30
X	18,78	22,43	22,62	23,06	24,94	24,88	26,18	28,30	29,26	28,17
Y	29,81	34,92	35,19	35,80	38,44	38,36	40,18	43,14	44,49	42,95

Варіант 4.

№	1	2	3	4	5	6	7	8	9	10
X	0,48	3,04	2,66	3,96	6,23	5,09	8,71	8,09	7,88	9,86
Y	3,21	12,43	11,06	15,74	23,94	19,83	32,87	30,63	29,87	37,03
№	11	12	13	14	15	16	17	18	19	20
X	10,36	11,57	12,60	14,54	14,76	17,10	16,03	18,09	18,88	19,85
Y	38,82	43,19	46,91	53,89	54,66	63,12	59,27	66,68	69,54	73,02
№	21	22	23	24	25	26	27	28	29	30
X	24,35	24,81	23,58	23,56	24,22	25,85	26,99	27,57	28,78	30,52
Y	89,25	90,90	86,46	86,38	88,75	94,64	98,76	100,82	105,21	111,49

Варіант 5.

№	1	2	3	4	5	6	7	8	9	10
X	2,70	2,34	2,08	3,71	4,54	5,76	5,36	7,94	7,82	11,19
Y	12,11	11,10	10,41	14,91	17,20	20,58	19,48	26,61	26,29	35,61
№	11	12	13	14	15	16	17	18	19	20
X	9,61	10,83	12,52	14,76	13,44	15,80	16,11	18,96	18,94	19,70
Y	31,22	34,60	39,28	45,49	41,84	48,36	49,21	57,11	57,05	59,15
№	21	22	23	24	25	26	27	28	29	30
X	22,63	21,85	23,75	25,52	26,35	26,05	27,58	28,85	27,69	32,32
Y	67,28	65,09	70,36	75,24	77,53	76,71	80,95	84,46	81,25	94,06

Варіант 6.

№	1	2	3	4	5	6	7	8	9	10
X	1,45	1,02	2,64	4,07	2,37	6,21	7,39	7,06	10,12	7,73
Y	7,11	6,22	9,56	12,50	9,02	16,90	19,34	18,66	24,96	20,04
№	11	12	13	14	15	16	17	18	19	20
X	10,64	12,41	13,39	12,56	14,98	17,18	16,55	18,80	18,15	22,33
Y	26,03	29,66	31,68	29,97	34,96	39,47	38,18	42,82	41,47	50,07
№	21	22	23	24	25	26	27	28	29	30
X	20,46	21,36	24,76	23,76	25,33	26,58	29,21	28,82	29,30	30,81
Y	46,23	48,07	55,07	53,03	56,25	58,83	64,23	63,43	64,42	67,52

Варіант 7.

№	1	2	3	4	5	6	7	8	9	10
X	0,03	1,55	0,90	4,29	3,49	6,61	7,41	7,60	6,60	10,18
Y	6,80	9,01	8,06	12,97	11,83	16,35	17,51	17,78	16,34	21,54
№	11	12	13	14	15	16	17	18	19	20
X	11,05	13,61	13,59	12,40	15,12	15,02	19,07	18,61	18,73	21,17
Y	22,80	26,51	26,48	24,75	28,71	28,56	34,44	33,77	33,94	37,48
№	21	22	23	24	25	26	27	28	29	30
X	21,61	22,64	23,63	25,03	24,75	25,96	27,32	27,56	28,58	27,91
Y	38,12	39,62	41,06	43,09	42,68	44,44	46,41	46,76	48,24	47,27

Варіант 8.

№	1	2	3	4	5	6	7	8	9	10
X	-0,19	2,47	1,65	4,52	4,82	5,09	7,81	7,28	8,84	10,77
Y	2,76	10,69	8,25	16,78	17,67	18,49	26,60	25,02	29,66	35,41
№	11	12	13	14	15	16	17	18	19	20
X	10,21	11,58	13,42	14,58	16,71	16,34	16,59	16,65	19,15	18,92
Y	33,74	37,82	43,32	46,78	53,11	52,03	52,76	52,94	60,38	59,71
№	21	22	23	24	25	26	27	28	29	30
X	21,27	20,64	23,84	22,06	25,27	27,81	27,00	28,96	28,23	31,22
Y	66,70	64,81	74,36	69,06	78,63	86,21	83,77	89,63	87,44	96,34

Варіант 9.

№	1	2	3	4	5	6	7	8	9	10
X	-0,05	2,38	2,37	2,61	5,51	6,28	5,53	9,20	10,34	8,28
Y	5,90	8,86	8,84	9,13	12,66	13,61	1270	17,16	18,55	16,04
№	11	12	13	14	15	16	17	18	19	20
X	10,89	11,70	13,75	15,83	15,06	15,73	18,55	17,89	20,28	22,05
Y	19,22	20,21	22,70	25,23	24,29	25,11	28,54	27,75	30,65	32,81
№	21	22	23	24	25	26	27	28	29	30
X	20,89	21,39	22,35	23,90	24,69	25,66	26,94	26,92	30,23	28,80
Y	31,40	32,00	33,17	35,06	36,02	37,20	38,76	38,74	42,77	41,04

Варіант 10.

№	1	2	3	4	5	6	7	8	9	10
X	0,76	2,20	3,8197	4,94	5,45	5,98	6,81	8,28	7,16	10,06
Y	11,26	13,99	17,0649	19,20	20,17	21,18	22,74	25,54	23,41	28,93
№	11	12	13	14	15	16	17	18	19	20
X	11,03	12,25	13,96	13,72	14,57	13,96	16,06	18,11	17,92	19,90
Y	30,77	33,09	36,34	35,89	37,51	36,35	40,34	44,23	43,87	47,64
№	21	22	23	24	25	26	27	28	29	30
X	20,34	21,15	23,48	26,69	23,48	24,11	27,41	27,48	29,07	32,29
Y	48,48	50,01	54,44	60,54	54,43	55,64	61,92	62,04	65,07	71,20

Варіант 11.

№	1	2	3	4	5	6	7	8	9	10
X	0,16	1,04	4,14	3,73	3,71	6,16	5,35	8,97	8,98	10,41
Y	7,56	10,13	19,16	17,97	17,92	25,07	22,71	33,28	33,29	37,46
№	11	12	13	14	15	16	17	18	19	20
X	9,80	13,07	14,83	13,51	13,28	16,04	18,29	17,26	18,09	19,12
Y	35,68	45,23	50,36	46,51	45,84	53,90	60,47	57,45	59,88	62,89
№	21	22	23	24	25	26	27	28	29	30
X	21,53	21,13	24,63	23,86	24,65	26,85	25,83	27,41	28,22	30,40
Y	69,93	68,75	78,96	76,71	79,04	85,45	82,47	87,09	89,46	95,82

Варіант 12.

№	1	2	3	4	5	6	7	8	9	10
X	-0,01	2,85	3,72	4,86	5,73	5,78	5,81	6,82	8,08	8,95
Y	7,99	18,97	22,30	26,69	30,03	30,23	30,35	34,21	39,038	42,40
№	11	12	13	14	15	16	17	18	19	20
X	10,15	11,92	12,86	13,03	16,44	15,25	17,11	17,60	20,40	19,15
Y	46,99	53,78	57,42	58,07	71,14	66,58	73,73	75,59	86,37	81,58
№	21	22	23	24	25	26	27	28	29	30
X	21,00	21,08	23,14	25,56	24,39	25,64	26,68	28,67	29,73	30,33
Y	88,66	89,00	96,88	106,20	101,70	106,50	110,49	118,14	122,20	124,52

Варіант 13.

№	1	2	3	4	5	6	7	8	9	10
X	0,48	1,78	4,05	3,78	5,16	5,87	7,12	9,92	9,97	9,95
Y	11,38	13,34	16,80	16,39	18,49	19,56	21,47	25,72	25,80	25,76
№	11	12	13	14	15	16	17	18	19	20
X	10,67	11,57	13,25	13,74	14,21	15,27	18,41	16,85	19,81	19,32
Y	26,87	28,24	30,78	31,53	32,25	33,85	38,64	36,26	40,76	40,01
№	21	22	23	24	25	26	27	28	29	30
X	20,30	22,92	23,11	22,63	24,66	27,84	27,43	27,84	28,51	31,044
Y	41,51	45,49	45,78	45,06	48,14	52,97	52,35	52,97	53,99	57,84

Варіант 14.

№	1	2	3	4	5	6	7	8	9	10
X	0,30	2,31	3,42	6,10	4,33	5,17	6,96	8,43	6,98	11,14
Y	2,70	7,35	9,90	16,08	12,00	13,92	18,06	21,43	18,11	27,67
№	11	12	13	14	15	16	17	18	19	20
X	10,87	12,25	12,19	14,17	13,28	15,87	16,58	17,57	19,70	20,18
Y	27,06	30,24	30,09	34,67	32,60	38,58	40,21	42,48	47,40	48,49
№	21	22	23	24	25	26	27	28	29	30
X	20,023	20,02	22,00	25,35	24,60	27,06	27,01	25,99	30,98	30,52
Y	48,14	48,12	52,69	60,40	58,69	64,33	64,23	61,89	73,38	72,33

Варіант 15.

№	1	2	3	4	5	6	7	8	9	10
X	3,14	2,47	5,59	3,72	5,12	6,08	6,49	6,52	6,19	10,06
Y	8,41	7,48	11,79	9,22	11,14	12,48	13,04	13,09	12,63	17,99
№	11	12	13	14	15	16	17	18	19	20
X	11,06	11,63	14,13	13,57	16,86	14,09	16,41	18,18	20,14	20,43
Y	19,38	20,16	23,62	22,84	27,39	23,57	26,78	29,23	31,93	32,34
№	21	22	23	24	25	26	27	28	29	30
X	20,06	20,42	22,05	23,95	23,69	23,71	26,44	28,53	29,05	29,10
Y	31,82	32,32	34,58	37,21	36,85	36,88	40,66	43,55	44,27	44,34

Варіант 16.

№	1	2	3	4	5	6	7	8	9	10
X	0,68	2,16	3,90	3,38	5,93	4,75	7,44	10,07	8,18	9,30
Y	3,84	7,73	12,30	10,94	17,67	14,56	21,62	28,56	23,57	26,54
№	11	12	13	14	15	16	17	18	19	20
X	10,50	12,02	14,07	13,98	13,19	13,91	15,83	17,31	18,74	20,38
Y	29,70	33,70	39,10	38,86	36,79	38,67	43,74	47,63	51,40	55,74
№	21	22	23	24	25	26	27	28	29	30
X	21,51	20,56	23,32	22,04	25,53	26,80	28,46	29,46	28,88	29,48
Y	58,70	56,21	63,48	60,09	69,29	72,64	77,02	79,65	78,11	79,69

Варіант 17.

№	1	2	3	4	5	6	7	8	9	10
X	-0,87	2,76	3,06	4,56	2,48	6,32	5,88	7,24	8,97	10,38
Y	5,52	7,39	7,54	8,32	7,25	9,22	8,90	9,70	10,59	11,32
№	11	12	13	14	15	16	17	18	19	20
X	11,60	11,59	13,99	12,01	13,68	17,59	17,77	18,39	19,13	20,63
Y	11,95	11,94	13,18	12,15	13,02	15,03	15,13	15,45	15,82	16,60
№	21	22	23	24	25	26	27	28	29	30
X	21,23	21,55	23,93	25,27	26,16	27,06	27,45	28,58	30,76	30,41
Y	16,91	17,07	18,30	18,99	19,45	19,92	20,12	20,70	21,82	21,64

Варіант 18.

№	1	2	3	4	5	6	7	8	9	10
X	0,02	2,14	3,06	3,42	6,98	6,69	5,62	8,18	8,59	9,60
Y	5,87	3,28	2,16	1,72	-2,63	-2,27	-0,97	-4,08	-4,58	-5,82
№	11	12	13	14	15	16	17	18	19	20
X	11,12	12,97	12,39	15,34	14,40	16,29	15,56	18,72	19,19	20,68
Y	-7,67	-9,93	-9,22	-12,82	-11,68	-13,98	-13,09	-16,95	-17,53	-19,33
№	21	22	23	24	25	26	27	28	29	30
X	21,30	21,36	23,15	24,84	24,37	25,44	28,87	27,74	27,87	28,47
Y	20,10	20,17	22,36	24,41	23,84	25,14	29,33	27,95	28,11	28,85

Варіант 19.

№	1	2	3	4	5	6	7	8	9	10
X	1,51	1,37	2,60	4,24	5,00	5,05	5,65	7,72	7,95	9,99
Y	17,82	16,53	28,20	43,84	51,05	51,58	57,32	77,04	79,22	98,67
№	11	12	13	14	15	16	17	18	19	20
X	11,20	11,50	13,26	15,29	16,57	14,58	17,94	21,09	19,89	20,40
Y	110,20	113,03	129,85	149,21	161,37	142,38	174,40	204,50	193,01	197,84
№	21	22	23	24	25	26	27	28	29	30
X	21,57	19,70	21,72	22,67	26,43	24,30	27,93	31,25	31,05	30,42
Y	208,99	191,22	210,47	219,50	255,34	235,06	269,69	301,26	299,40	293,35

Варіант 20.

№	1	2	3	4	5	6	7	8	9	10
X	-0,85	2,30	3,54	5,91	5,90	7,21	6,83	7,70	8,63	10,68
Y	-8,48	29,01	43,68	71,86	71,84	87,37	82,81	93,24	104,34	128,64
№	11	12	13	14	15	16	17	18	19	20
X	10,30	10,52	12,82	14,29	14,14	17,42	17,85	17,80	21,43	20,83
Y	124,17	126,75	154,10	171,61	169,79	208,90	214,04	213,43	256,59	249,46
№	21	22	23	24	25	26	27	28	29	30
X	21,24	23,56	22,04	24,41	24,88	27,31	29,14	30,50	26,36	28,93
Y	254,32	281,95	263,79	292,07	297,58	326,48	348,27	364,44	315,25	345,83

Варіант 21.

№	1	2	3	4	5	6	7	8	9	10
X	-0,13	2,77	2,12	4,92	5,12	6,84	6,99	10,10	8,40	8,90
Y	3,29	19,12	15,55	30,79	31,88	41,27	42,07	59,06	49,76	52,49
№	11	12	13	14	15	16	17	18	19	20
X	13,55	11,94	12,25	13,49	16,33	15,29	19,40	17,53	18,80	20,58
Y	77,84	69,06	70,74	77,49	92,97	87,34	109,71	99,53	106,47	116,13
№	21	22	23	24	25	26	27	28	29	30
X	20,49	21,89	23,25	23,51	24,76	25,77	25,98	27,30	31,05	31,59
Y	115,66	123,26	130,71	132,09	138,90	144,41	145,55	152,78	173,20	176,11

Варіант 22.

№	1	2	3	4	5	6	7	8	9	10
X	-0,86	0,57	2,70	2,76	5,29	6,85	7,22	8,34	9,18	8,70
Y	-1,90	3,82	12,39	12,62	22,77	29,06	30,52	35,05	38,43	36,48
№	11	12	13	14	15	16	17	18	19	20
X	11,63	10,77	12,97	12,44	16,32	15,85	18,21	18,51	18,80	19,00
Y	48,23	44,77	53,61	51,49	67,07	65,18	74,67	75,88	77,05	77,84
№	21	22	23	24	25	26	27	28	29	30
X	20,73	23,69	22,90	24,71	24,75	26,74	28,20	28,48	28,95	30,97
Y	84,77	96,68	93,49	100,75	100,92	108,92	114,78	115,91	117,79	125,91

Варіант 23.

№	1	2	3	4	5	6	7	8	9	10
X	0,96	2,42	3,69	4,46	6,04	7,39	8,94	8,45	7,83	8,81
Y	12,08	24,57	35,46	41,98	55,58	67,10	80,41	76,21	70,86	79,28
№	11	12	13	14	15	16	17	18	19	20
X	10,31	12,92	12,37	15,66	13,95	16,93	17,34	17,70	19,49	19,26
Y	92,13	114,52	109,73	137,94	123,27	148,85	152,31	155,44	170,72	168,81
№	21	22	23	24	25	26	27	28	29	30
X	20,66	22,19	24,37	23,61	24,56	26,75	27,84	28,63	29,48	30,14
Y	180,74	193,85	212,53	206,03	214,22	232,94	242,25	249,00	256,28	261,99

Варіант 24.

№	1	2	3	4	5	6	7	8	9	10
X	1,83	1,77	3,49	5,84	5,27	6,85	5,08	7,48	8,14	9,06
Y	8,19	8,04	12,97	19,70	18,08	22,61	17,53	24,41	26,32	28,94
№	11	12	13	14	15	16	17	18	19	20
X	9,14	13,02	14,27	16,35	15,23	16,64	16,29	18,37	18,84	18,93
Y	29,19	40,31	43,89	49,86	46,64	50,70	49,70	55,65	57,00	57,28
№	21	22	23	24	25	26	27	28	29	30
X	20,18	23,14	21,90	23,05	26,34	28,38	26,44	27,28	27,95	29,53
Y	60,88	69,35	65,80	69,09	78,52	84,39	78,81	81,23	83,16	87,69

Варіант 25.

№	1	2	3	4	5	6	7	8	9	10
X	0,84	2,79	3,07	4,64	4,15	5,03	7,25	8,77	8,78	8,65
Y	1,44	3,78	4,11	6,00	5,40	6,46	9,11	10,95	10,95	10,79
№	11	12	13	14	15	16	17	18	19	20
X	11,09	12,36	15,27	15,03	14,29	14,22	18,96	19,591	19,54	18,12
Y	13,72	15,24	18,74	18,44	17,56	17,48	23,15	23,91	23,85	22,15
№	21	22	23	24	25	26	27	28	29	30
X	20,35	21,91	23,27	23,51	24,68	27,20	27,26	28,55	27,98	29,68
Y	24,82	26,68	28,32	28,60	30,00	33,02	33,10	34,64	33,96	36,00

Варіант 26.

№	1	2	3	4	5	6	7	8	9	10
X	0,40	2,66	2,22	4,42	5,10	3,90	6,03	7,95	11,21	9,78
Y	4,80	8,85	8,08	12,01	13,23	11,07	14,89	18,33	24,19	21,61
№	11	12	13	14	15	16	17	18	19	20
X	9,20	12,11	14,69	12,98	14,75	16,34	16,04	17,54	18,44	19,32
Y	20,58	25,80	30,41	27,35	30,53	33,38	32,83	35,53	37,15	38,73
№	21	22	23	24	25	26	27	28	29	30
X	22,18	21,50	22,90	21,84	24,42	26,82	27,44	27,52	29,07	30,92
Y	43,85	42,64	45,15	43,24	47,87	52,18	53,28	53,43	56,20	59,52

Варіант 27.

№	1	2	3	4	5	6	7	8	9	10
X	0,35	0,80	3,26	3,35	4,93	5,78	5,97	7,37	8,026	9,19
Y	3,89	5,64	15,33	15,70	21,90	25,26	26,03	31,54	34,11	38,71
№	11	12	13	14	15	16	17	18	19	20
X	11,80	12,33	13,13	12,95	16,40	14,16	18,05	16,41	19,08	21,14
Y	48,99	51,05	54,20	53,50	67,11	58,26	73,59	67,12	77,65	85,78
№	21	22	23	24	25	26	27	28	29	30
X	20,83	21,49	22,95	24,68	25,90	25,19	26,12	28,21	30,37	30,97
Y	84,55	87,13	92,87	99,70	104,49	101,72	105,38	113,59	122,10	124,47

Варіант 28.

№	1	2	3	4	5	6	7	8	9	10
X	2,81	1,72	2,67	3,10	5,66	6,98	8,22	8,97	10,56	10,29
Y	17,96	13,04	17,30	19,23	30,84	36,78	42,41	45,78	52,96	51,75
№	11	12	13	14	15	16	17	18	19	20
X	11,42	10,90	12,61	14,19	15,71	15,96	17,59	17,44	18,13	20,89
Y	56,88	54,51	62,27	69,38	76,28	77,41	84,77	84,09	87,20	99,69
№	21	22	23	24	25	26	27	28	29	30
X	22,09	22,86	23,40	23,13	25,55	25,40	25,49	27,92	29,52	30,34
Y	105,12	108,59	111,02	109,80	120,73	120,07	120,49	131,45	138,70	142,39

$$\text{Значення функції Лапласа } \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$$

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
0,00	0,0000	0,50	0,1915	1,00	0,3413	1,50	0,4332	2,00	0,4772	3,00	0,49865
0,01	0,0040	0,51	0,1950	1,01	0,3438	1,51	0,4345	2,02	0,4783	3,20	0,49931
0,02	0,0080	0,52	0,1985	1,02	0,3461	1,52	0,4357	2,04	0,4793	3,40	0,49966
0,03	0,0120	0,53	0,2019	1,03	0,3485	1,53	0,4370	2,06	0,4803	3,60	0,499841
0,04	0,0160	0,54	0,2054	1,04	0,3508	1,54	0,4382	2,08	0,4812	3,80	0,499928
0,05	0,0199	0,55	0,2088	1,05	0,3531	1,55	0,4394	2,10	0,4821	4,00	0,499968
0,06	0,0239	0,56	0,2123	1,06	0,3554	1,56	0,4406	2,12	0,4830	4,50	0,499997
0,07	0,0279	0,57	0,2157	1,07	0,3577	1,57	0,4418	2,14	0,4838	5,00	0,499997
0,08	0,0319	0,58	0,2190	1,08	0,3599	1,58	0,4429	2,16	0,4846		
0,09	0,0359	0,59	0,2224	1,09	0,3621	1,59	0,4441	2,18	0,4854		
0,10	0,0398	0,60	0,2257	1,10	0,3643	1,60	0,4452	2,20	0,4861		
0,11	0,0438	0,61	0,2291	1,11	0,3665	1,61	0,4463	2,22	0,4868		
0,12	0,0478	0,62	0,2324	1,12	0,3686	1,62	0,4474	2,24	0,4875		
0,13	0,0517	0,63	0,2357	1,13	0,3708	1,63	0,4484	2,26	0,4881		
0,14	0,0557	0,64	0,2389	1,14	0,3729	1,64	0,4495	2,28	0,4887		
0,15	0,0596	0,65	0,2422	1,15	0,3749	1,65	0,4505	2,30	0,4893		
0,16	0,0636	0,66	0,2454	1,16	0,3770	1,66	0,4515	2,32	0,4898		
0,17	0,0675	0,67	0,2486	1,17	0,3790	1,67	0,4525	2,34	0,4904		
0,18	0,0714	0,68	0,2517	1,18	0,3810	1,68	0,4535	2,36	0,4909		
0,19	0,0753	0,69	0,2549	1,19	0,3830	1,69	0,4545	2,38	0,4913		
0,20	0,0793	0,70	0,2580	1,20	0,3849	1,70	0,4554	2,40	0,4918		
0,21	0,0832	0,71	0,2611	1,21	0,3869	1,71	0,4564	2,42	0,4922		
0,22	0,0871	0,72	0,2642	1,22	0,3883	1,72	0,4573	2,44	0,4927		
0,23	0,0910	0,73	0,2673	1,23	0,3907	1,73	0,4582	2,46	0,4931		
0,24	0,0948	0,74	0,2703	1,24	0,3925	1,74	0,4591	2,48	0,4934		
0,25	0,0987	0,75	0,2734	1,25	0,3944	1,75	0,4599	2,50	0,4938		
0,26	0,1026	0,76	0,2764	1,26	0,3962	1,76	0,4608	2,52	0,4941		
0,27	0,1064	0,77	0,2794	1,27	0,3980	1,77	0,4616	2,54	0,4945		
0,28	0,1103	0,78	0,2823	1,28	0,3997	1,78	0,4625	2,56	0,4948		
0,29	0,1141	0,79	0,2852	1,29	0,4015	1,79	0,4633	2,58	0,4951		
0,30	0,1179	0,80	0,2881	1,30	0,4032	1,80	0,4641	2,60	0,4953		
0,31	0,1217	0,81	0,2910	1,31	0,4049	1,81	0,4649	2,62	0,4956		
0,32	0,1255	0,82	0,2939	1,32	0,4066	1,82	0,4656	2,64	0,4959		
0,33	0,1293	0,83	0,2967	1,33	0,4082	1,83	0,4664	2,66	0,4961		
0,34	0,1331	0,84	0,2995	1,34	0,4099	1,84	0,4671	2,68	0,4963		
0,35	0,1368	0,85	0,3023	1,35	0,4115	1,85	0,4678	2,70	0,4965		
0,36	0,1406	0,86	0,3051	1,36	0,4131	1,86	0,4686	2,72	0,4967		
0,37	0,1443	0,87	0,3078	1,37	0,4147	1,87	0,4693	2,74	0,4969		
0,38	0,1480	0,88	0,3106	1,38	0,4162	1,88	0,4699	2,76	0,4971		
0,39	0,1517	0,89	0,3133	1,39	0,4177	1,89	0,4706	2,78	0,4973		
0,40	0,1554	0,90	0,3159	1,40	0,4192	1,90	0,4713	2,80	0,4974		
0,41	0,1591	0,91	0,3186	1,41	0,4207	1,91	0,4719	2,82	0,4976		

Закінчення дод. 1

0,42	0,1628	0,92	0,3212	1,42	0,4222	1,92	0,4726	2,84	0,4977		
0,43	0,1664	0,93	0,3238	1,43	0,4236	1,93	0,4732	2,86	0,4979		
0,44	0,1700	0,94	0,3264	1,44	0,4251	1,94	0,4738	2,88	0,4980		
0,45	0,1736	0,95	0,3289	1,45	0,4265	1,95	0,4744	2,90	0,4981		
0,46	0,1772	0,96	0,3315	1,46	0,4279	1,96	0,4750	2,92	0,4982		
0,47	0,1808	0,97	0,3340	1,47	0,4292	1,97	0,4756	2,94	0,4984		
0,48	0,1844	0,98	0,3365	1,48	0,4306	1,98	0,4761	2,96	0,4985		
0,49	0,1879	0,99	0,3389	1,49	0,4319	1,99	0,4767	2,98	0,4986		

Додаток 2

Таблиця Пірсона. α – рівень значущості, m – кількість ступенів вільності

m/α	0,99	0,98	0,95	0,90	0,80	0,20	0,10	0,05	0,02	0,01	0,001
1	0,00016	0,00063	0,00393	0,0158	0,0642	1,642	2,706	3,841	5,412	6,635	10,827
2	0,0201	0,0404	0,103	0,211	0,446	3,219	4,605	5,991	7,824	9,210	13,815
3	0,115	0,185	0,352	0,584	1,005	4,642	6,251	7,815	9,837	11,341	16,268
4	0,297	0,429	0,711	1,064	1,649	5,989	7,779	9,488	11,668	13,277	18,465
5	0,554	0,752	1,145	1,610	2,343	7,289	9,236	11,070	13,388	15,086	20,517
6	0,872	1,134	1,635	2,204	3,070	8,558	10,645	12,592	15,033	16,812	22,457
7	1,239	1,564	2,167	2,833	3,822	9,803	12,017	14,067	16,622	18,475	24,322
8	1,646	2,032	2,733	3,490	4,594	11,030	13,362	15,507	18,679	20,090	26,125
9	2,088	2,532	3,325	4,168	5,380	12,242	14,684	16,919	19,679	21,666	27,877
10	2,588	3,059	3,940	4,865	6,179	13,442	15,987	18,307	21,161	23,209	29,588

Список літератури

1. *Барковський В.В.* Теорія ймовірностей та математична статистика. 5-те видання. / В.В. Барковський, Н.В. Барковська, О.К. Лопатін. – Київ : Центр учбової літератури, 2010. – 424 с.
2. *Гмурман В.Е.* Теория вероятностей и математическая статистика : Учеб. Пособие. 12-е изд., перераб. – Москва : Высшее образование, 2006. – 479 с.
3. *Гмурман В.Е.* Руководство к решению задач по теории вероятностей и математической статистике : Учеб. пособие 11-е изд. перераб. – Москва : Высшее образование, 2006 – 404 с.
4. *Наголкіна З.І.* Практикум з математичної статистики. / З.І. Наголкіна, С.Г. Розе, В.С. Шитюк. – Київ : КНУБА, 2013.
5. *Бондаренко Н.В.* Теорія ймовірностей : Навчальний посібник. / Н.В. Бондаренко, З.І. Наголкіна, М.С. Пастухова. – Київ : КНУБА, 2016.
6. *Шаповалов О.М.* Забезпечення надійності експлуатації будівельних конструкцій : методичні вказівки. / О.М. Шаповалов, Н.О. Псурцева. – Харків : ХНУМО ім. Бекетова, 2017.

Навчально-методичне видання

**Методичні вказівки до самостійної роботи
з математичної статистики**

для здобувачів першого (бакалаврського) рівня вищої освіти
за спеціальностями 073 «Менеджмент»,
051 «Економіка», 071 «Облік і оподаткування»

Укладачі : **НАГОЛКІНА** Зоя Іванівна,
СОКОЛОВА Людмила Віталіївна,
ФІЛОНОВ Юрій Петрович

Випусковий редактор *Л. С. Тавлуй*

Комп'ютерне верстання *К. А. Мавроді*

Підписано до друку 17.01.2025. Формат 60 x 84_{1/16}

Ум. друк. арк. 2,56. Обл.-вид. арк. 2,75.

Електронний документ. Вид. № 142/III-24

Видавець і виготовлювач:

Київський національний університет будівництва і архітектури

Проспект Повітряних Сил, 31, Київ, Україна, 03037

Свідоцтво про внесення до Державного реєстру суб'єктів
видавничої справи ДК № 808 від 13.02.2002

