МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Київський національний університет будівництва і архітектури

МАТЕМАТИЧНА СТАТИСТИКА ТА ВИПАДКОВІ ПРОЦЕСИ

Методичні вказівки та завдання
до проведення практичних занять № 1 – 12
для здобувачів першого (бакалаврського) рівня вищої освіти
за спеціальностями 122 «Комп’ютерні науки»,
126 «Інформаційні системи та технології»

Київ 2025

УДК 519.21
М33

Укладач О.В. Горда, канд. техн. наук, доцент

Рецензент Ю.В. Рябчун, д-р філософії, доцент

Відповідальний за випуск Т.А. Гончаренко, д-р техн. наук, доцент

Затверджено на засіданні кафедри інформаційних технологій, протокол № 5 від 9 грудня 2024 року.

В авторській редакції.

М33 **Математична** статистика та випадкові процеси : методичні вказівки та завдання до проведення практичних занять №1 – 12 / уклад. О.В. Горда. – Київ : КНУБА, 2025. – 128 с.

Містять зміст, необхідні теоретичні відомості, порядок виконання практичних робіт та завдання для самостійного виконання.

Призначено для здобувачів першого (бакалаврського) рівня вищої освіти за спеціальностями 122 «Комп'ютерні науки» та 126 «Інформаційні системи та технології».

Зміст

ЗАГАЛЬНІ ПОЛОЖЕННЯ

Дисципліна «Математична статистика та випадкові процеси» входить до навчальних планів усіх комп'ютерних, а також ряду природничо-наукових та економічних спеціальностей. Це викликано вимогами науково-технічного прогресу, в процесі якого людство йде по шляхом використання дедалі складніших технічних та інформаційних засобів, а також використання складних природних об'єктів або систем.

Підбір матеріалу для навчального посібника зорієнтований на подальше використання в таких напрямах досліджень, як аналіз та моделювання систем, науковий експеримент, інтелектуальний аналіз даних (*Data mining*), прийняття рішень в умовах ризиків та невизначеності, інтелектуальні методи оптимізації складних систем, системи штучного інтелекту, теорія надійності.

Методичні вказівки містять рекомендації для вирішення задач, які безпосередньо в практичному зрізі відображають якість засвоєння здобувачами вивчення дисципліни, а саме:

– принципи та методи збору інформації та її оцінки на основі статистичних оцінок;

– дослідження залежності між випадковими величинами та побудова математичної моделі у вигляді лінії регресії, прогнозування та рангова кореляція;

– перевірка статистичних гіпотез;

– визначення ступеня впливу фактора на систему;

– поняття, класифікація та основні характеристики випадкових процесів;

– системи масового обслуговування, їх дослідження та імітаційне моделювання.

Завершивши вивчення дисципліни, здобувач повинен здійснювати логічні математичні міркування із чітким зазначенням припущень та висновків, демонструвати здатність до ймовірнісного мислення, вміти застосовувати інструментарій математичної статистики у практичних ситуаціях в міждисциплінарному контексті, володіти практичними навичками моделювання і аналізу випадкових процесів.

Для кращого розуміння та полегшення засвоєння матеріалу в методичних вказівках наведені типові приклади вирішення завдань.

Практична робота № 1.
Побудова та аналіз варіаційного ряду

Задано вибірку з 50 значень:

1) побудувати варіаційний ряд (частот та відносних частот) з 8 значень;
2) побудувати функцію розподілу та її графік;
3) побудувати кумуляту та знайти квартилі;
4) побудувати полігон та гістограму частот;
5) знайти точкові оцінки вибірки та генеральної сукупності: середнє, моду, медіану, дисперсію, середньоквадратичне відхилення, розмах, коефіцієнт варіації;
6) сформулювати гіпотезу про закон розподілу та перевірити її на основі критерію Пірсона ($\alpha = 0{,}05$);
7) методом моментів оцінити параметри функції розподілу.

Для виконання розрахунків та побудови графіків доцільно застосовувати програмний додаток *Mathcad*.

Теоретичні відомості

Варіаційний ряд – це пара, яка складається із значення випадкової величини (варіанти) та відповідної частоти або відносної частоти, причому значення випадкової величини впорядковані за зростанням:

Частота – кількість об’єктів з конкретним числовим значенням ознаки: n_i.

Відносна частота – частка з даними значенням ознаки ($w_i=n_i/N$), де N – об’єм сукупності. Відносна частота є статистичною оцінкою теоретичної ймовірності.

Емпіричною функцією розподілу, що відповідає вибірці $(x_1, x_2, ..., x_n)$, називається функція $F^*(x)$, яка для кожного значення x визначає відносну частоту події $X < x$: $F^*(x) = \frac{n_x}{N}$, де n_x – кількість спостережень $X < x$, N – об’єм вибірки.

Кумулятою називають графік накопичення частот, де на осі OX відкладаються значення випадкової величини, а на осі OY – накопичені частоти.

Квантиль (***процентиль***) – одна з числових характеристик випадкових величин, це таке число x_p, що значення p-ї частини сукупності менше або рівне x_p. Наприклад, квантиль 0,25 (також називається 25-м процентилем або нижнім квартилем) – це таке значення (x_p), що 25 % (p) значень змінної потрапляють нижче даного значення. Як правило, останній квантиль не обчислюється.

Розрахунок квантилів для емпіричного закону розподілу можна виконувати за наступними схемами:

✓ схема 1:

– складаємо варіаційний ряд $x_0 < x_1 < \ldots < x_{N-1}$ і покладаємо $x_N = x_{N-1}$;

– знаходимо величину $k = p(N-1)$, N – об'єм вибірки;

– порівнюємо k і pN:

якщо $k+1 < pN$, то покладаємо $x_p = x_{k+1}$;

якщо $k+1 = pN$, то покладаємо $x_p = (x_{k+1} + x_k)/2$;

якщо $k+1 > pN$, то покладаємо $x_p = x_k$;

✓ схема 2:

– складаємо варіаційний ряд і додаємо рядок накопиченої частоти (приклад наведений в табл. 1.1) і знаходимо, скільки значень попадає у відповідний процентиль: $n_p = p \cdot N$.

Для наведеного прикладу, якщо $p = 0{,}25$, $n_p = 0{,}25 \cdot 30 = 7{,}5$

Таблиця 1.1

Варіаційний ряд та накопичені частоти

<table>
<tr><td>Значення x_i</td><td>35</td><td>36</td><td>37</td><td>38</td><td>39</td><td>40</td><td>41</td></tr>
<tr><td rowspan="2">Частота n_i</td><td rowspan="2">2</td><td>4</td><td>5</td><td rowspan="2">6</td><td rowspan="2">7</td><td rowspan="2">5</td><td rowspan="2">1</td></tr>
<tr><td colspan="2">інтервал, що відповідає $p = 0{,}25$</td></tr>
<tr><td>Накопичена частота Nn</td><td>2</td><td>6</td><td>11</td><td>17</td><td>24</td><td>29</td><td>30</td></tr>
</table>

– в рядку накопичених частот знайдемо нижню границю величини n_p і обчислимо L – середину відповідного інтервалу та $rn = n_p - n_{нижня\ границя}$ (для нашого прикладу $n_p = 6$ і йому відповідає значення випадкової величини $x_2 = 26$, звідси $L = 36{,}5$, $rn = 7{,}5 - 6 = 1{,}5$);

– знаходимо відповідний процентиль $P_p = L + rn / n_{верхняграниця}$, де $n_{верхняграниця}$

– частота верхньої границі інтервалу (5) і додаємо до L ($P_4 = 36{,}5 + 1{,}5 / 5 = 36{,}8$).

Полігоном частот (відносних частот) називають ламану, відрізки якої з'єднують точки з координатами (x_i, n_i) / $\left(x_i, {}^{n_i}\!/\!_{N} \right)$.

Гістограмою частот (умовних частот) називають ступінчасту фігуру, що складається з прямокутників, основи яких розташовані на осі x і довжини їх дорівнюють довжинам часткових інтервалів h, а висоти дорівнюють відношенню: ${}^{n_i}\!/\!_{h}$ – для частот ${}^{n_i}\!/\!_{Nh}$ – для умовних частот.

Математичним сподіванням величини X є середньо зважена величина:

$$m_x = \bar{x} = \frac{\sum_{i=1}^{N} n_i x_i}{N},$$

де n_i – частоти відповідних значень випадкової величини x_i; N– об'єм сукупності.

Диспéрсія – це міра розсіювання значень випадкової величини відносно середнього значення розподілу $D = \dfrac{\sum_{i=1}^{n} n_i (x_i - \bar{x})^2}{N}$.

Середньоквадратичне відхилення: $\sigma = \sqrt{D}$.

Модою M_0 називають варіанту, яка має найбільшу частоту.

Особливості застосування моди:

1) якщо всі значення варіаційного ряду мають однакову частоту, то кажуть, що цей варіаційний ряд не має моди;

2) якщо дві сусідні варіанти мають однакову домінуючу частоту, то мода обчислюється як середнє арифметичне цих варіант;

3) якщо дві не сусідні варіанти мають однакову домінуючу частоту, то такий варіаційний ряд називається бімодальним;

4) якщо таких варіант більше двох, то ряд полімодальний.

Медіаною m_0 називають варіанту, яка ділить варіаційний ряд на дві частини, рівні за кількістю варіант. Якщо число варіант непарне, то $m_O = x_{k+1}$, якщо парне, то $m_O = (x_k + x_{k+1}) / 2$.

Розмахом варіювання R називають різницю між найбільшою і найменшою варіантами: $R = x_{\max} - x_{\min}$.

Коефіцієнтом варіації V називають виражене у % відношення вибіркового середнього квадратичного відхилення до вибіркової середньої: $V = \sigma_в / \bar{x}_в \cdot 100\%$.

Перевірка закону розподілу за критерієм Пірсона $\left(\chi^2\right)$**.**

Загальна формула для обчислення значення критерію, що спостерігається: $\chi^2_{спост.} = \dfrac{\left(n_i - n_i'\right)^2}{n_i'}$.

1. Перевірка розподілу за рівномірним законом $R(a,b)$**.**

Функція щільності рівномірного розподілу:

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in (a,b) \\ 0, & x \notin (a,b), \end{cases}$$

1) оцінити параметри a та b – кінці інтервалу (a,b):

$$a^* = \bar{x}_в - \sqrt{3}\sigma_в, \quad b^* = \bar{x}_в + \sqrt{3}\sigma_в;$$

2) знайти щільність передбачуваного розподілу: $f(x) = \dfrac{1}{b^* - a^*}$;

3) знайти теоретичні частоти (N – об'єм вибірки, x_i – границі інтервалів):

$$n_1' = NP = N \cdot f(x) \cdot (x_1 - a^*) = \frac{N(x_1 - a^*)}{b^* - a^*},$$

$$n_n' = NP = N \cdot f(x) \cdot (x_1 - a^*) = \frac{N(b^* - x_n)}{b^* - a^*},$$

$$n_2' = n_3' = \ldots = n_{n-1}' = \frac{N(x_i - x_{i-1})}{b^* - a^*}.$$

4) За допомогою критерію Пірсона порівняти теоретичні та емпіричні частоти за кількість степенів свободи $k = n - 3$, де n – кількість груп або частинних інтервалів).

2. Перевірка розподілу за показниковим (експоненціальним) законом E(λ ***)***

Функція щільності показникового розподілу:

$$p(k) = \begin{cases} \lambda \cdot e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}.$$

1) знайти середнє вибіркове $\bar{x}_в$ і в якості оцінки параметра взяти:

$$\lambda^* = \frac{1}{\bar{x}_в};$$

2) знайти ймовірності попадання випадкової величини в частковий інтервал (x_i, x_{i+1}): $P_i = e^{-\lambda x_i} - e^{-\lambda x_{i+1}}$;

3) знайти теоретичні частоти: $n_i' = N \cdot P_i$ (N – об'єм вибірки);

4) за допомогою критерію Пірсона порівняти теоретичні та емпіричні частоти за кількість степенів свободи $k = n - 2$, де n – кількість інтервалів.

Зауваження. Для перевірки експоненціального закону розподілу необхідно перейти до інтервального варіаційного ряду.

3. Перевірка розподілу за законом Пуассона P(λ)

Закон Пуассона $P_n(i) = \lambda^i \dfrac{e^{-\lambda}}{i!}$.

1) знайти середнє вибіркове $\bar{x}_в$ і в якості оцінки параметра взяти: $\lambda^* = \bar{x}_в$,

2) поклавши $i = x_i$, знайти ймовірності $P_n(i)$, $i \geq 0$ – ціле число;

3) знайти теоретичні частоти: $n_i' = N \cdot P_i$ (N – об'єм вибірки);

4) за допомогою критерію Пірсона порівняти теоретичні та емпіричні частоти за кількість степенів свободи $k = n - 2$, де n – кількість інтервалів.

4. Перевірка розподілу за нормальним законом $N(a, \sigma)$**.**

1) знайти середнє вибіркове $\bar{x}_в$ та вибіркове середньоквадратичне відхилення $\sigma_в$;

2) обчислити теоретичні частоти:

$$n_i' = \frac{N \cdot h}{\sigma_в} \varphi(u_i) = \Phi\left(\frac{x_{i+1} - \bar{x}_в}{\sigma_в}\right) - \Phi\left(\frac{x_i - \bar{x}_в}{\sigma_в}\right),$$

де N – об'єм вибірки,

h – довжина інтервалу,

$u_i = \dfrac{x_i - \bar{x}_в}{\sigma_в}$, $\varphi(u_i) = \dfrac{1}{\sqrt{2\pi}} e^{-u^2/2}$,

$\Phi(\ )$ – функція Лапласа (дод. 2),

(x_i, x_{i+1}) – частковий інтервал.

3) за допомогою критерію Пірсона порівняти теоретичні та емпіричні частоти за кількість степенів свободи $k = n - 3$, де n – кількість інтервалів.

Приклад 1. Нехай задано варіаційний ряд (табл. 1.2):

Таблиця 1.1

Варіаційний ряд (результати спостережень)

№ спостереження	1	2	3	4	5	6
x_i – варіанти (інтервали)	2–5	5–8	8–11	11–14	14–17	17–20
x_i (значення)	3,5	6,5	9,5	12,5	15,5	18,5
n_i – частоти	6	8	15	22	14	5

Об’єм вибірки – $N = 70$, середнє вибіркове – $\bar{x}_в = 11{,}43$, середньоквадратичне відхилення $\sigma_в = 4{,}03$, виправлене середньо вибіркове – $s = 4{,}05$.

1) Перевірка рівномірного закону розподілу.

Знайдемо оцінку функції щільності розподілу:

$$a^* = 11{,}43 - \sqrt{3} \cdot 4{,}03 = 4{,}45,\ b^* = 11{,}43 + \sqrt{3} \cdot 4{,}03 = 18{,}41,$$

$$f(x) = 1/(b^* - a^*) = 0{,}072.$$

Знайдемо вирівнюючи частоти:

$$n_1' = 70 \cdot 0{,}072 \cdot (5 - 4{,}45) = 2{,}77,$$

$$n_2' = n_3' = n_4' = n_5' = 70 \cdot 0{,}072 \cdot 3 = 15{,}12,$$

$$n_6' = 70 \cdot 0{,}072 \cdot (18{,}41 - 17) = 7{,}1.$$

Обчислимо значення критерію, що спостерігалося:

$$\chi^2_{сп} = \frac{(6 - 2{,}77)^2}{2{,}77} + \ldots + \frac{(5 - 7{,}1)^2}{7{,}1} = 10{,}95.$$

З таблиці знайдемо критичне значення критерію: $\chi^2_{кр}(0{,}05;3) = 7{,}8.$

$\chi^2_{сп} > \chi^2_{кр}$, отже гіпотеза про рівномірний закон відхиляється.

2) ***Перевірка показникового закону розподілу.***

Знайдемо оцінку параметра $\lambda^* = 1/\bar{x}_в = 1/11{,}43 = 0{,}087.$

Знайдемо вирівнюючи частоти $n_i' = N \cdot \left(e^{-\lambda x_i} - e^{-\lambda x_{i+1}}\right)$:

$$n_1' = 70 \cdot \left(e^{-0{,}087\cdot 2} - e^{-0{,}087\cdot 5}\right) = 13{,}44; \qquad \text{аналогічно} \qquad n_2' = 10{,}37;$$

$n_3' = 8{,}05;\ n_4' = 6{,}23;\ n_5' = 4{,}76;\ n_6' = 3{,}64.$

Обчислимо значення критерію, що спостерігалося:

$$\chi^2_{сп} = \frac{(6-13{,}44)^2}{13{,}44} + \ldots + \frac{(5-3{,}64)^2}{3{,}64} = 69{,}02.$$

З таблиці знайдемо критичне значення критерію: $\chi^2_{кр}(0{,}05;4) = 9{,}5.$

$\chi^2_{сп} > \chi^2_{кр}$, отже гіпотеза про рівномірний закон відхиляється.

3) Перевірка нормального закону розподілу

Знайдемо вирівнюючи частоти:

$$n_1' = 70 \cdot \left(\Phi\left(\frac{5-11{,}43}{4{,}05}\right) - \Phi\left(\frac{2-11{,}43}{4{,}05}\right)\right) = 3{,}2,$$

$$n_2' = 9{,}9;\quad n_3' = 18{,}2;\quad n_4' = 19{,}6;\quad n_5' = 12{,}5;\quad n_6' = 4{,}7.$$

Обчислимо значення критерію, що спостерігалося:

$$\chi^2_{сп} = \frac{(6-3{,}2)^2}{3{,}2} + \ldots + \frac{(5-4{,}7)^2}{4{,}7} = 3{,}87.$$

З таблиці знайдемо критичне значення критерію: $\chi^2_{кр}(0{,}05;4) = 9{,}5.$

$\chi^2_{сп} < \chi^2_{кр}$, отже гіпотеза про рівномірний закон приймається.

Приклад 2. Нехай задано варіаційний ряд (табл. 1.3):

Таблиця 1.3

Варіаційний ряд (результати спостережень)

№ спостереження	1	2	3	4	5	6
x_i (значення)	0	1	2	3	4	5
n_i – частоти	270	166	49	10	3	2

Об'єм вибірки – $N = 500$, середнє вибіркове – $\bar{x}_в = 0{,}632$.

Знайдемо оцінку параметра $\lambda^* = \bar{x}_в = 0{,}632.$

Значення частот в останніх 2-х стовпчиках таблиці малі, тому ми їх об'єднаємо. Знайдемо вирівнюючи частоти $n_i' = N \cdot \dfrac{\bar{x}_в^i}{i!} e^{-\bar{x}_в}$:

$$n_1' = 265{,}76;\ n_2' = 167{,}96;\ n_3' = 53{,}08;\ n_4' = 11{,}18;\ n_5' = 1{,}99.$$

Обчислимо значення критерію, що спостерігалося:

$$\chi^2_{сп} = \frac{(270 - 265{,}764)^2}{270} + \ldots + \frac{(5 - 1{,}99)^2}{1{,}99} = 5{,}081.$$

З таблиці (дод. 5) знайдемо критичне значення критерію: $\chi^2_{кр}(0{,}05;4) = 9{,}5$. $\chi^2_{сп} < \chi^2_{кр}$, отже гіпотеза про закон розподілу Пуасона приймається.

Метод моментів полягає у прирівнюванні параметрам закону розподілу відповідних теоретичних моментів.

Приклад виконання роботи

1. ***Побудова варіаційного ряду. Занесемо початкові дані у*** *Mathcad*. Для цього їх можна скопіювати та вставити у документ Блокноту, розташувати у вигляді стовпчика і записати у текстовий файл (рис. 1.1).

Зчитати дані у вектор стовпчик документа Mathcad можна за допомогою функції READPRN:

Ім’я змінної вектора стовпчика = READPRN(«шлях до файлу»).

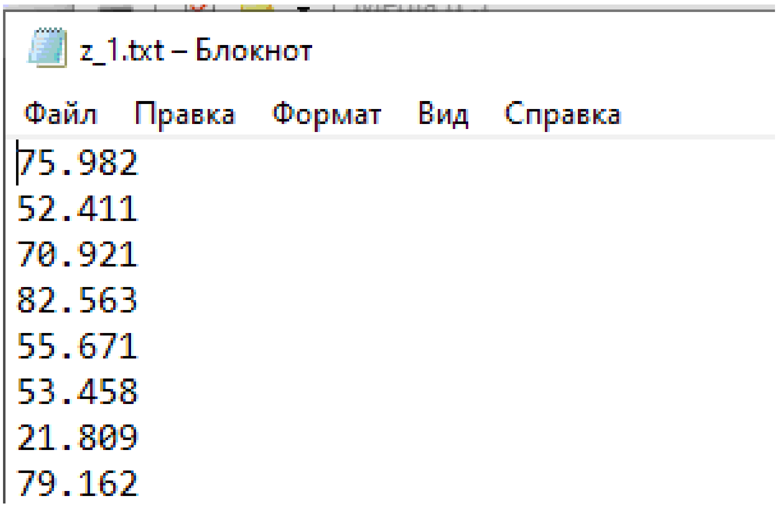


Рис. 1.1. Фрагмент файлу даних

Приклад. Зчитування файлу з даними *data1.txt*, який розташований на диску *C* в папці *My_file*:

x:= READPRN(“C:My_file\data1.txt”).

Застосовуючи стандартну функцію сортування у порядку зростання *sort*(), відсортуємо дані: x = sort(x).

Побудуємо варіаційний ряд. Застосовуючи стандартні функції, знайдемо максимальне та мінімальне значення та весь діапазон можливих значень R, розіб'ємо його на 8 частинних відрізків і знайдемо довжину частинного відрізка h. В якості значення візьмемо значення середини відрізка X_i (рис. 1.2).

Розділимо відрізок на 7 рівних частин та знайдемо значення середин частинних відрізків (рис. 1.2).

$$x := \mathrm{READPRN}(\text{"F:dat\_1.txt"}) \qquad x := \mathrm{sort}(x)$$

$$R := \max(x) - \min(x) = 110.378 \qquad h := \frac{R}{8} = 13.797$$

$$X_0 := x_0 + 0.5 \cdot h \quad i := 0..7 \quad X_i := X_0 + h \cdot i \qquad N := \mathrm{last}(x) = 49$$

Рис. 1.2. Зчитування даних у вектор *x*, сортування та визначення довжини частинних відрізків *h* та середніх точок відрізків X_i

Для знаходження частот складемо програмний блок, використовуючи інструменти панелі Програмування, що складається з двох циклічних процесів:

– 1-й цикл – відлік частинних інтервалів;

– 2-й цикл – обчислення кількості значень, що попали у відповідний інтервал.

Вхідні дані: N – об'єм сукупності; X – значення варіант; h – довжина частинних інтервалів. Вихідні дані: n – вектор стовпчик частот.

Результати обчислення програмного блоку – кількість значень, менших за кінцеве значення відповідного відрізку, запишемо у вектор *nv* і знайдемо кількість значень на кожному відрізку (рис. 1.3).

```
var(N,X,h) := | n_0 ← 0
              | for i ∈ 0..7
              |   | for j ∈ n_i..N
              |   |   | n_i ← n_i + 1  if  x_j ≤ X_i + 0.5·h
              |   |   | break  otherwise
              |   | n_(i+1) ← 0
              | n
                                     nv := var(N,X,h)
```

$$i := 0..6 \qquad n_0 := nv_0 \qquad n_{i+1} := nv_{i+1} - nv_i \qquad k := 7$$

Рис. 1.3. Програмний блок обчислення частот

Знайдемо відносні частоти і виведемо значення варіаційного ряду на екран (рис. 1.4).

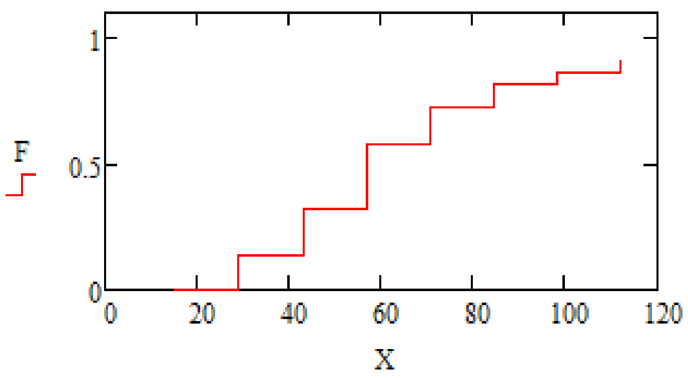
$$N := \sum_{i=0}^{k} n_i = 50 \qquad i := 0..k \qquad w_i := \frac{n_i}{N}$$

$X_i =$	$n_i =$	$w_i =$
15.335	4	0.08
29.132	7	0.14
42.929	9	0.18
56.726	13	0.26
70.524	7	0.14
84.321	5	0.1
98.118	2	0.04
111.915	3	0.06

Рис. 1.4. Значення варіаційного ряду

2. Знаходження функції розподілу. Функцію розподілу можна обчислити однією формулою та задати крайні значення. Побудуємо графік функції розподілу. Для отримання ступінчастої фігури у закладці Формат – Трасування виберемо вигляд функції – ступінчастий (рис. 1.5).

$$i := 1..7 \qquad F_0 := 0 \qquad F_i := F_{i-1} + w_i \qquad F_8 := 1$$



$$F = \begin{pmatrix} 0 \\ 0.14 \\ 0.32 \\ 0.58 \\ 0.72 \\ 0.82 \\ 0.86 \\ 0.92 \\ 1 \end{pmatrix}$$

Рис.1.5. Обчислення та графік функції розподілу

3. Побудуємо огинаючу криву функції розподілу. На осі *OY* відрізок від 0 до 1 розіб’ємо на чотири рівні частини, проведемо паралельні прямі до осі *OX* і знайдемо точки їх перетину з кумулятою з яких опустимо перпендикуляри на вісь *OX*. Основи перпендикулярів будуть відмічати значення квартилів (рис. 1.6).

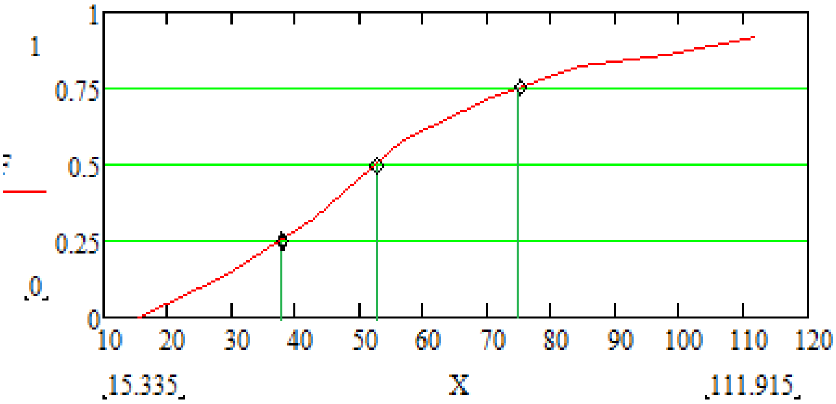


Рис. 1.6. Кумулята та визначення квартилів

З графіка знайдемо наближені значення квартилів:

$$Q_1 \approx 39,\ Q_2 \approx 52,\ Q_3 \approx 74{,}5.$$

Розрахуємо квартилі за схемою для емпіричного закону розподілу:

$$-\ p = 0{,}25;\ 0{,}5;\ 0{,}75,\ \ N = 50;$$

$$-\ k_i = p_1(N-1):$$

$$k_1 = 0{,}25 \cdot 49 = 12{,}25,\ k_2 = 0{,}5 \cdot 49 = 24{,}5,\ k_3 = 0{,}75 \cdot 49 = 36{,}75.$$

Порівняємо $k_i + 1$ з відповідними значеннями $p_i \cdot N$ і згідно з правилами визначимо відповідні значення варіаційного ряду (необхідно пам'ятати, що k – номер варіанти – натуральне число). Отримаємо:

$$Q_1 \approx x_{13},\ Q_2 \approx x_{25},\ Q_2 \approx x_{38}.$$

4. Для побудови полігону та гістограми частот також скористуємось графічними можливостями Mathcad. Для побудови полігону виберемо тип графіка – лінії, на осі OX відкладемо значення елементів варіаційного ряду, а OY – частоти n (рис. 1.7).

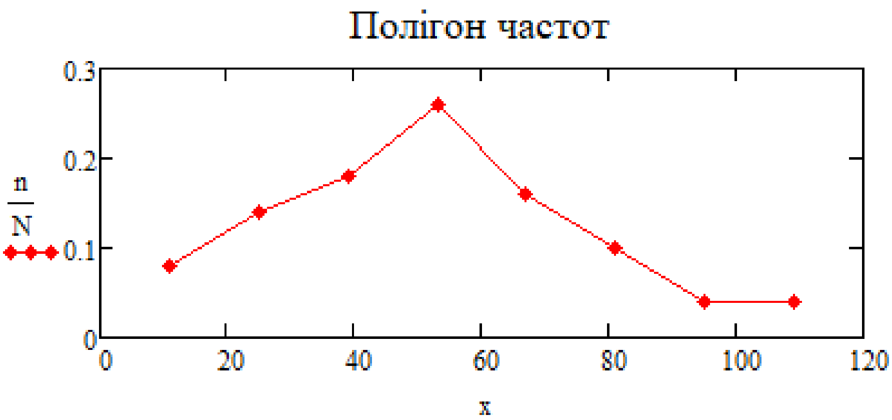


Рис. 1.7. Полігон частот

Для побудови гістограми виберемо тип графіка – суцільні стовпчики, на осі OX відкладемо значення X, а OY – n/h (рис. 1.8).

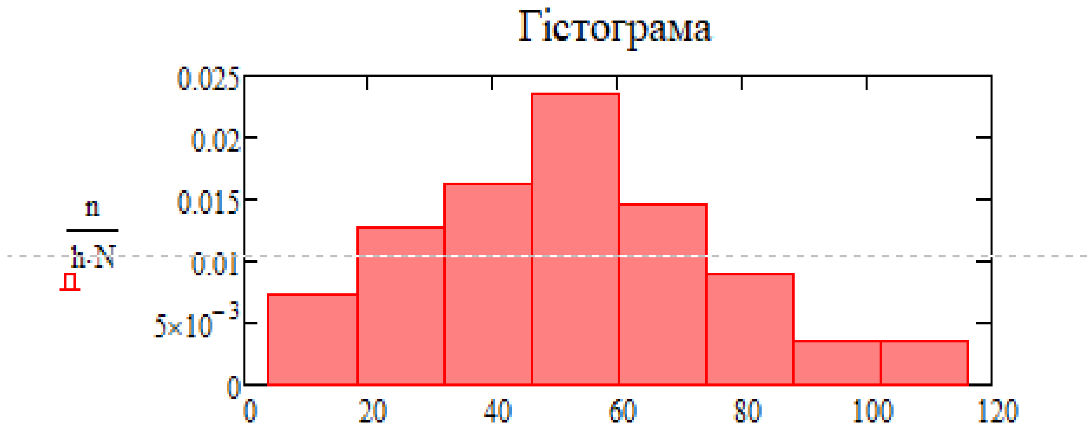


Рис. 1.8. Гістограма частот

5. Обчислення точкових оцінок (рис. 1.9). Під час обчислення точкових оцінок вибірки та генеральної сукупності слід пам'ятати, що дисперсія є зміщеною оцінкою, і необхідно робити поправку на кількість степенів свободи $((N-1))$.

$$xv := \frac{\sum_{i=0}^{7} (X_i \cdot n_i)}{N} = 56.726 \qquad Dv := \frac{\sum_{i=0}^{7} (X_i - xv)^2}{N} = 167.52 \qquad \sigma v := \sqrt{Dv} = 12.943$$

середнє вибіркове | вибіркова дисперсія | вибіркове середньоквадратичне відхилення

$$xg := xv = 56.726 \qquad D := \frac{N \cdot Dv}{N-1} = 170.939 \qquad \sigma := \sqrt{D} = 13.074$$

середнє генеральне | генеральна дисперсія | генеральне середньоквадратичне відхилення

$$Mo := X_3 = 56.726 \qquad Me := X_3 = 56.726 \qquad R := \max(X) - \min(X) = 96.581 \qquad V := \frac{\sigma}{xg} = 0.23$$

мода | медіана | розмах | коефіцієнт варіації

Рис. 1.9. Точкові оцінки варіаційного ряду

6. З графіків полігону та гістограми, а також оцінок середнього вибіркового, моди та медіани можна висунути гіпотезу, що вибірка має нормальний закон розподілу. Обчислимо значення χ^2-критерію, що спостерігається, для нормального закону розподілу (рис. 1.10).

$$i := 0..7$$

$$f(u) := \frac{1}{\sqrt{2\cdot\pi}}\cdot\exp\left(\frac{-u^2}{2}\right) \qquad np_i := \frac{N\cdot h}{\sigma}\cdot f\left(\frac{X_i - xv}{\sigma}\right) \qquad \chi2 := \sum_{i=0}^{7}\frac{\left(n_i - np_i\right)^2}{np_i} = 2894.836$$

Рис. 1.10. Обчислення χ^2-критерію

З таблиці дод. 5 за заданим рівнем значущості $\alpha = 0{,}05$ та кількістю степенів свободи $k = 7 - 3 = 4$ знайдемо критичне значення χ^2-критерію: $\chi^2_{кр} = 9{,}5$. Оскільки значення, що спостерігається, більше критичного, то гіпотеза про нормальний закон розподілу відхиляється.

7. Якщо б гіпотеза про нормальний закон розподілу підтвердилась, то можна було б оцінити його параметри. Нормальний закон розподілу залежить від 2-х параметрів (a, σ) (див. дод. 1), де a – математичне сподівання, а σ – середньоквадратичне відхилення, тоді $a = xg$, $\sigma = \sigma$.

Варіанти для самостійного виконання

Варіант 1 (нормальний)

64.39	79.96	62.92	70.56	65.99	60.88	79.23	74.53	62.54	71.22
70.50	64.99	70.21	60.30	74.86	60.75	71.90	75.43	70.68	69.46
62.11	72.41	73.59	65.17	68.75	73.17	71.29	68.42	74.69	57.43
62.76	68.53	77.15	63.11	75.78	68.85	80.32	71.40	86.81	67.15
69.96	77.26	80.04	85.28	65.03	68.73	64.77	62.42	65.67	58.33

Варіант 2 (експоненціальний)

1.38	8.01	14.87	3.72	0.23	14.83	0.56	0.84	2.84	8.57
7.01	2.55	8.84	0.02	32.25	5.50	10.23	3.27	6.28	12.56
3.83	2.23	0.54	15.29	1.53	8.09	7.99	3.59	4.01	10.16
10.79	1.74	11.29	4.89	1.21	1.34	11.56	11.59	5.47	8.56
3.05	4.78	1.15	13.43	0.84	5.04	11.97	0.95	15.46	5.24

Варіант 3 (рівномірний)

3	2	8	10	5	4	6	10	10	5
7	7	7	1	7	1	9	6	9	10
5	3	10	3	5	5	5	10	10	1
5	6	8	10	4	4	5	8	7	2
2	10	7	8	3	10	9	6	4	5

Варіант 4 (рівномірний)

3.14	1.57	-3.01	-1.94	8.50	2.04	8.06	1.87	8.41	7.97
-4.10	-0.80	3.17	3.17	-1.73	1.21	4.04	5.83	6.67	1.30
4.87	9.78	7.42	8.12	-3.84	2.54	-1.02	0.08	-3.96	-1.40
8.34	4.13	7.55	-3.18	2.11	-3.12	7.97	1.02	-0.82	3.96
-3.36	-1.19	7.50	7.85	7.53	-3.02	-4.13	2.90	0.69	2.19

Варіант 5 (Пуассона)

11	14	15	8	12	12	15	14	8	11
21	9	11	5	13	9	11	15	6	12
7	8	7	5	2	2	8	9	12	14
8	5	17	9	10	7	16	4	9	7
11	12	12	15	13	11	11	8	9	11

Варіант 6 (експоненціальний)

3.56	19.19	16.77	0.74	19.74	1.05	15.19	1.40	20.08	37.11
11.23	39.29	3.27	38.60	2.75	5.61	13.46	23.67	8.18	8.61
2.82	4.29	9.29	5.17	0.07	8.47	14.97	16.24	1.95	9.26
16.98	2.77	11.54	21.73	6.47	14.48	11.38	8.32	41.97	3.96
0.36	1.21	2.96	13.78	5.27	6.82	10.29	10.79	12.28	21.84

Варіант 6 (нормальний)

30.95	29.27	29.36	28.12	32.07	28.23	31.93	31.92	30.32	35.36
29.54	28.28	33.66	30.83	36.57	29.06	27.46	33.52	32.80	28.96
34.45	34.60	32.12	27.21	32.56	33.78	32.98	25.46	32.24	30.32
30.98	28.36	30.52	27.48	26.57	31.21	25.68	29.35	32.09	28.15
28.28	24.86	35.71	28.90	29.34	33.17	32.56	30.06	30.57	29.80

Варіант 7 (рівномірний)

18	4	4	20	13	12	8	10	13	6
6	3	3	18	8	16	19	2	6	3
16	24	12	12	20	16	11	15	24	2
9	4	17	18	6	4	21	19	13	21
23	4	21	24	6	4	24	9	14	20

Варіант 8 (експоненціальний)

8.16	16.95	2.08	2.41	4.87	4.42	2.43	9.08	0.83	23.03
0.55	4.29	14.75	7.71	6.68	7.62	1.25	1.134	2.70	1.06
1.90	3.57	2.93	2.19	7.11	0.76	0.23	10.87	2.67	2.48

2.10	2.69	1.70	4.83	2.37	0.75	0.42	9.20	0.38	3.67
0.05	12.43	3.59	11.33	3.24	6.29	4.77	5.17	2.77	6.57

Варіант 9 (рівномірний)

26.24	19.80	41.98	48.79	22.53	41.12	41.79	36.27	34.40	43.91
34.18	25.16	18.34	32.59	23.08	27.07	11.55	19.28	25.35	30.18
13.79	20.81	45.85	18.71	43.60	21.20	39.07	34.88	11.22	10.32
23.44	18.62	39.83	44.31	29.70	23.34	44.91	13.05	44.30	46.76
15.79	35.35	31.45	44.48	12.08	24.68	21.43	48.67	34.14	26.43

Варіант 10 (Пуассона)

52	43	42	42	55	47	41	53	43	49
42	62	50	45	49	52	45	34	44	55
44	44	50	41	44	60	48	46	48	56
44	38	54	60	55	53	54	41	47	56
46	35	52	46	47	54	49	46	57	40

Практична робота № 2
Перевірка та оцінки нормального закону розподілу

Задано вибірку із 100 значень:

1) побудувати варіаційний ряд;
2) побудувати полігон та гістограму відносних частот;
3) знайти точкові оцінки: середнє, моду, медіану, дисперсію, середньоквадратичне відхилення, асиметрію, ексцес, розмах, коефіцієнт варіації;
4) перевірити, що вибірка має нормальний закон розподілу за критерієм Пірсона та графічно;
5) знайти інтервальні оцінки: математичного сподівання та дисперсії з достовірністю $\gamma = 0{,}95$;
6) оцінити об’єм необхідної вибірки, щоб похибка не перевищувала 5%.

За необхідності, у розрахунках виконати округлення і перейти до умовних варіант.

Для виконання розрахунків та побудови графіків доцільно застосовувати програмний додаток *Mathcad*.

Теоретичні відомості

Нормальний закон розподілу, чи закон Гауса, відіграє у статистиці і займає особливе становище серед інших законів. Нормальний розподіл природним чином виникає практично скрізь, де йдеться про вимір з помилками. Наприклад, координати точки влучення снаряду, зростання, вага людини мають нормальний закон розподілу. Більш того, центральна гранична теорема взагалі стверджує, що сума великої кількості доданків сходить до нормальної випадкової величини, незалежно від того, який був вихідний розподіл у вибірки. Таким чином, дана теорема встановлює умови, за яких виникає нормальний розподіл та порушення яких веде до розподілу, відмінного від нормального.

Тестування даних на нормальність є досить частим етапом первинного аналізу даних, оскільки велика кількість статистичних методів використовує те, що дані розподілені нормально.

Перші 3 пункти аналогічні до 1,4 і 5 пунктів завдання 1. Різниця полягає в тому, що для нормального закону важливими є ще дві характеристики: асиметрія та ексцес.

Коефіцієнт асиметрії – в теорії ймовірностей це величина, що характеризує асиметрію розподілу даної випадкової величини (рис. 2.1).

Якщо $As < 0$, то це означає, що переважають дані з великими значеннями (правостороння асиметрія), а якщо $As > 0$, то більше даних з меншими значеннями, ніж середнє арифметичне (лівостороння асиметрія). Якщо $As = 0$, то розподіл є симетричним.

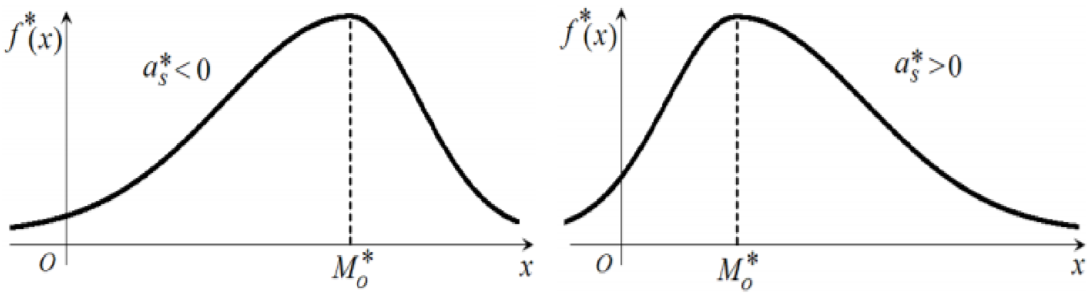


Рис. 2.1. Право- та лівостороння асиметрія

Для оцінки ступеня асиметрії застосовують моментні та структурні коефіцієнти асиметрії.

Найчастіше застосовують відносний показник, структурний коефіцієнт асиметрії, запропонований англійським статистиком К.Пірсоном: $As = \dfrac{\bar{x} - m_o}{\sigma}$.

Досить точним є показник, що ґрунтується на визначенні центрального моменту третього порядку – моментний коефіцієнт асиметрії: $As = \dfrac{\mu_3}{\sigma^3}$, де $\mu_3 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{n}$ – третій центральний момент.

Оцінка суттєвості As проводиться на основі середньої квадратичної помилки коефіцієнта асиметрії, яка залежить від числа спостережень (n):

$$\sigma_{As} = \sqrt{\frac{6(n-1)}{(n+1)(n+3)}}.$$

Якщо $\dfrac{|As|}{\sigma_{As}} > 3$, асиметрія істотна і розподіл ознаки у генеральній сукупності несиметричний. Якщо $\dfrac{|Ex|}{\sigma_{As}} < 3$, асиметрія несуттєва, та її наявність пояснюється впливом різних випадкових обставин.

Коефіцієнт ексцесу (коефіцієнт гостроти) в теорії ймовірностей – міра гостроти піку розподілу випадкової величини (рис. 2.2).

Найбільш точним є показник, який використовує центральний момент четвертого порядку:

$$Es = \frac{\mu_4}{\sigma^4} - 3, \text{ де } \mu_4 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{n} \text{ – четвертий центральний момент.}$$

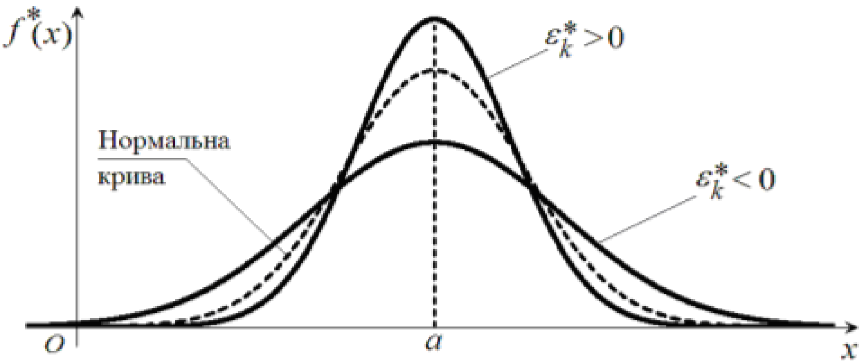


Рис. 2.2. Ексцес

Середньоквадратична похибка ексцесу розраховується за формулою:

$$\sigma_{Ex} = \sqrt{\frac{24n(n-2)(n-3)}{(n-1)^2(n+3)(n+5)}}.$$

Якщо $\frac{|Ex|}{\sigma_{Ex}} > 3$ – відхилення від нормального вважається суттєвим, якщо $\frac{|Ex|}{\sigma_{Ex}} < 3$ – не суттєвим.

В наближених обчисленнях, коли величина вимірюється багаторазово, в якості істинного значення береться середнє значення, асиметрія та ексцес визначають тенденції прямування до середнього значення.

Якщо середнє значення, мода та медіана майже співпадають, то закон розподілу можна вважати нормальним.

Для перевірки нормального закону розподілу застосовуються різні критерії: Пірсона, Колмогорова – Смірнова, Романовського $n < 20$(), Шапіро – Уїлка, Д'Агостино та ін., а також графічні методи.

Критерій Пірсона χ^2 – найчастіше вживаний критерій згоди (див. завдання 1).

Зауваження. Необхідні умови застосування критерію Пірсона:

1) обсяг вибірки має бути досить великий, принаймні щонайменше 50 спостережень;

2) кожен частковий інтервал повинен містити щонайменше п'ять спостережень. Якщо ця кількість в окремих інтервалах мала, то є сенс об'єднати деякі інтервали, підсумовуючи частоти.

Графічна перевірка нормального закону розподілу. Метод спрямлених діаграм. Нехай варіаційних ряд задано за інтервальною шкалою. Для перевірки закону розподілу складається розрахункова таблиця (табл. 2.1) для знаходження квантилів. Якщо задана ймовірність p, то p-квантилем випадкової величини X називають таке значення аргументу z_p функції розподілу $F(x)$ для якого ймовірність події $X < z_p$ дорівнює заданому значенню p.

Таблиця 2.1

Розрахункова таблиця квантилів

№ ін.-лу	Правий кінець ін.-лу	Часто-та	Накопичена частота	Відносна накопичена частота	Відносна накопичена частота %	Квантилі
i	x_i	n_i	$\sum_{k=1}^{i} n_k$	$p_1 = \frac{\sum n_k}{N}$	$p_i \cdot 100\%$	z_{pi}

Квантилі для нормального закону розподілу знаходимо з таблиці дод. 8.

В прямокутній системі координат XOZ наносимо точки (x_i, z_i). Якщо точки розташовані поблизу деякої прямої (рис. 2.3), то немає підстав спростовувати гіпотезу про нормальний закон розподілу, в протилежному випадку гіпотеза відхиляється. Відповідну пряму можна побудувати за двома точками: координата точки перетину прямої і осі OX відповідає математичному сподіванню, а друга точка має координати $(a-\sigma;-1)$, де a – математичне сподівання, σ – середньоквадратичне відхилення.

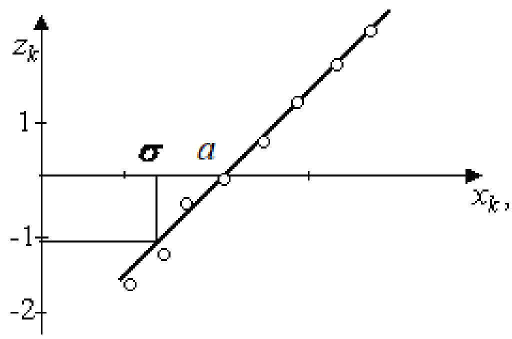


Рис. 2.3. Відхилення спостережень від прямої $z = \dfrac{x-a}{\sigma}$

Якщо точки несуттєво відхиляються рівномірно по обидві сторони лінії, то закон можна вважати нормальним. Якщо точки переважно відхиляються по одну із сторін, то це може свідчити про систематичну похибку.

<u>*Зауваження*</u>. Початкові та кінцеві точки можуть відхилятися від прямої.

Інтервальні оцінки. Побудова довірчих інтервалів.

Точність оцінки характеризується позитивним числом, яке становить величину розбіжності між оцінками вибірки і генеральної сукупності: $|\theta - \theta^*| < \delta$, $\delta > 0$.

Надійністю (довірчою ймовірністю) оцінки називають ймовірність, з якої здійснюється нерівність $|\theta - \theta^*| < \delta$: $P(|\theta - \theta^*| < \delta) = \gamma$.

Для надійності найбільш часто використовують величини, близькі до одиниці: 0,95; 0,99 і 0,999. Це значення помножене на 100 означає відсоток надійності оцінки.

Довірчим називають інтервал $(\theta^* - \delta, \theta^* + \delta)$, який покриває невідомий параметр із заданою надійністю γ.

Довірчий інтервал для математичного сподівання при невідомій дисперсії: $\bar{x} - t_\gamma s/\sqrt{n} < a < \bar{x} + t_\gamma s/\sqrt{n}$, де t_γ – величина з розподілом Стьюдента (знаходиться з таблиці розподілу Стьюдента (дод. 3) за заданими значеннями γ та n) s – виправлене середньо квадратичне відхилення.

Оцінку використовують для визначення найменшого об'єму вибірки (яку найменшу кількість вимірювань чи іспитів необхідно провести), щоб отримати значення із заданою точністю δ: $n = (t_\gamma s/\delta)^2$. Необхідно пам'ятати, що n натуральне число, тому результат округлюється до найближчого більшого цілого.

Довірчий інтервал для оцінки середньо квадратичного відхилення σ нормального розподілу. Для отримання оцінки будемо вимагати виконання умови: $P(|\sigma - s| < \delta) = \gamma$. Нехай $\delta/s = q$, тоді отримаємо: $s(1-q) < \sigma < s(1+q)$, де q за заданими значеннями γ та n знаходиться з таблиці χ^2 розподілу (дод. 4).

Цю оцінку використовують в якості оцінки точності вимірювань.

Приклад виконання роботи

Пункти **1** – **3** виконуються аналогічно до завдання 1. Оскільки в даному завданні ми маємо справу з нормальним законом розподілу, то для пункту **3** додається дві точкові оцінки – асиметрія та ексцес (рис. 2.4).

$$As := \frac{\sum_{i=0}^{7}(x_i - xv)^3}{N\cdot\sigma^3} = 1.663 \qquad \sigma As := \sqrt{\frac{6\cdot(N-1)}{(N+1)\cdot(N+3)}} = 0.33 \qquad \frac{|As|}{\sigma As} = 5.041$$

Асиметрія є істотною

$$Ex := \frac{\sum_{i=0}^{7}(x_i - xv)^4}{N\cdot\sigma^4} - 3 = 8.352 \qquad \sigma Ex := \sqrt{\frac{24\cdot N\cdot(N-2)\cdot(N-3)}{(N-1)^2\cdot(N+3)\cdot(N+5)}} = 0.622 \qquad \frac{|Ex|}{\sigma Ex} = 13.429$$

Ексцес є суттєвим

Рис. 2.4. Оцінка асиметрії та ексцесу

4. *Перевірка нормального закону розподілу*. Перевірку за *критерієм Пірсона* доцільно виконати в середовищі Mathcad, що дозволить повністю автоматизувати обчислення і не використовувати статистичні таблиці для обчислення вирівнюючих частот (рис. 2.5).

$$f(u) := \frac{1}{\sqrt{2\cdot\pi}}\cdot\exp\left(\frac{-u^2}{2}\right) \qquad i := 0..7 \qquad N = 50 \qquad h := 11$$

$$z_i := \frac{x_i - xv}{\sigma} \qquad np_i := \frac{h\cdot N}{\sigma}\cdot f(z_i)$$

$$\chi 2 := \sum_{i=0}^{7}\frac{(n_i - np_i)^2}{np_i} = 11.226 \qquad \chi 2kr := 14.1$$

Рис. 2.5. Перевірка нормального закону розподілу за критерієм Пірсона

Для перевірки закону розподілу за *методом спрямлених діаграм* створимо розрахункову таблицю для квантилів (рис. 2.6).

$$i := 1..7 \quad Xk_i := x_i + 0.5\cdot h \qquad Nn_0 := n_0 \qquad Nn_i := n_i + Nn_{i-1} \qquad i := 0..7 \qquad P_i := \frac{Nn_i}{N} \qquad Pv_i := P_i\cdot 100$$

№	Кінцева точка інтервалу	Частота	Накопичена частота	Відносна зн.	Накопичена частота %	Квантилі
$i =$	$Xk_i =$	$n_i =$	$Nn_i =$	$P_i =$	$Pv_i =$	$z :=$
0	0	4	4	0.08	8	-1.405
1	30.019	7	11	0.22	22	-0.772
2	44.019	9	20	0.4	40	-0.253
3	58.019	13	33	0.66	66	0.412
4	72.019	8	41	0.82	82	0.915
5	86.019	5	46	0.92	92	1.405
6	100.019	2	48	0.96	96	1.751
7	114.019	2	50	1	100	2

Рис. 2.6. Розрахункова таблиця для знаходження квантилів

Значення квантилів беремо з таблиці дод. 8. Нанесемо на графік точки $(x_i; z_i)$ та побудуємо лінію за двома точками, спираючись на оцінки математичного сподівання та середньоквадратичного відхилення (рис. 2.7)

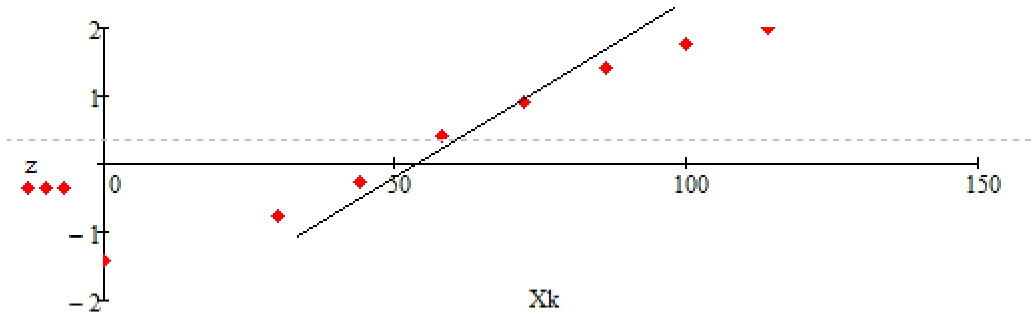


Рис. 2.7. Розташування точок емпіричного розподілу

Отримані оцінки досить грубі, але можна вважати, що закон розподілу нормальний.

***5.** Знаходження інтервальних оцінок.*

Знайдемо інтервальну оцінку для математичного сподівання. Оскільки дисперсія невідома, скористаємось законом розподілу Стьюдента. З дод. 3 за заданою надійність $\gamma = 0{,}95$ та об'ємом вибірки $N = 100$ знайдемо значення $t_\gamma = 1{,}984$, оцінимо похибку δ та знайдемо довірчий інтервал $(a1; a2)$ (рис. 2.8).

$$t := 1.984 \qquad \delta := \frac{t \cdot \sigma}{\sqrt{N}} = 1.492 \qquad a1 := xg - \sigma = 44.64 \qquad a2 := xg + \sigma = 59.68$$

Рис. 2.8. Інтервальна оцінка математичного сподівання

За заданими значеннями γ та N знаходиться з таблиці χ^2 розподілу (дод. 5) знайдемо величину $q = 0{,}143$ і знайдемо довірчий інтервал для дисперсії $(s1; s2)$ (рис. 2.9).

$$q := 0.143 \qquad s1 := \sigma \cdot (1 - q) = 6.445 \qquad s2 := \sigma \cdot (1 + q) = 8.595$$

Рис. 2.9. Довірчий інтервал дисперсії

***6.** Визначення оптимального об'єму вибірки.*

Для визначення оптимального об'єму вибірки (яку найменшу кількість вимірювань чи іспитів необхідно провести), щоб отримати значення із заданою точністю δ, використовують оцінку точності, яка обчислювалась для знаходження довірчого інтервалу математичного сподівання:

$$N = \left(t_\gamma s / \delta\right)^2 = 89038{,}741.$$

Так як N – натуральне число, то в якості остаточного значення візьмемо $N = 89039$.

Варіанти для самостійного виконання

Варіант 1

0.897 1.517 1.719 0.985 1.766 1.043 0.299 1.348 1.094 0.700
0.945 1.363 1.163 0.918 0.615 0.254 0.289 1.417 0.959 1.245
1.745 0.848 0.623 1.314 1.186 0.629 1.245 0.879 0.034 1.370
1.705 1.147 1.685 1.547 0.887 0.469 0.911 1.108 0.781 1.856
1.709 0.606 0.144 1.555 1.559 2.175 0.902 0.417 0.103 0.903
1.336 1.444 0.949 0.568 0.456 0.692 1.710 0.426 1.420 0.069
0.396 0.427 0.879 1.039 1.016 1.374 1.146 1.052 0.556 0.580
1.359 0.466 1.160 0.393 1.276 0.904 1.099 1.361 1.050 1.677
1.815 0.595 1.156 0.443 1.550 1.444 1.794 2.293 0.728 0.464
1.244 0.472 0.568 0.997 1.772 0.618 0.598 0.667 1.152 1.480

Варіант 2

11.459 9.992 11.124 7.377 10.507 12.138 12.671 14.845 14.056 12.906
16.316 15.064 6.065 6.721 8.356 6.349 10.518 6.865 16.650 11.967
3.745 9.972 15.172 9.026 16.001 7.798 11.6780 15.664 12.728
7.516 10.268 13.596 12.296 3.590 10.168 10.347 9.238 13.443 10.224
10.399 6.531 3.408 9.894 14.568 8.821 13.928 14.189 10.749 13.313
10.460 21.760 9.691 11.085 12.924 9.913 6.072 5.962 10.278
11.678 9.032 14.053 7.953 10.031 10.926 7.863 2.380 12.684 10.195
7.190 17.224 16.102 5.165 4.524 10.871 6.856 5.638 7.860 9.878
14.512 7.760 7.085 12.880 9.943 8.047 20.3462 12.234 9.761
11.085 12.136 9.210 9.266 12.043 7.924 14.769 17.125 12.448
13.045 11.190 11.500 8.366

Варіант 3

5.129 8.777 9.397 3.175 9.818 2.649 6.278 12.199 10.646 7.544
10.039 10.220 11.062 10.577 21.119 13.358 4.391 7.548 1.431 13.247
9.630 12.380 4.146 19.021 -0.292 4.936 15.624 7.742 13.906 11.503
18.650 9.694 8.715 11.492 22.147 11.331 11.000 16.961 0.038 15.130
9.755 13.217 17.212 11.207 12.688 14.262 10.725 2.941 0.596 13.681
5.679 12.960 5.667 7.329 16.000 12.635 15.556 10.878 11.401 15.489
19.017 4.747 15.669 10.846 2.679 16.413 13.973 1.256 6.258 10.340
17.173 6.260 12.752 12.162 8.050 16.030 12.751 16.103 13.058 10.264
9.852 7.295 10.855 13.132 9.610 7.746 17.251 15.309 14.385 16.039
3.469 8.461 11.912 9.136 13.114 12.285 15.649 11.006 15.289 0.382

Варіант 4

12.982 16.601 12.757 21.305 12.582 16.235 11.381 22.512 14.160
15.439 9.918 11.280 23.711 13.076 15.413 14.065 13.847 18.763
8.205 19.515 13.470 14.293 19.554 16.310 17.253 10.772 16.400
18.149 17.430 13.840 17.595 6.472 5.013 17.659 15.454 13.863
7.657 11.496 14.529 20.046
18.303 19.581 16.765 15.341 11.381 14.653 19.144 16.280 17.797
16.902 10.940 12.484 9.407 18.524 13.129 9.122 24.698 12.767
16.079 19.696 13.116 10.185 13.980 16.293 14.500 15.769 18.838
13.754 16.977 15.508 15.548 13.984 15.658 11.863 20.916 11.711
13.737 12.720 9.068 12.373
13.832 9.285 17.991 7.778 11.557 14.623 16.715 10.897 10.919
9.074 16.207 14.917 13.908 22.434 18.139 16.345 10.856 11.365
13.212 15.622

Варіант 5

31.548 25.875 14.384 37.130 16.050 19.211 28.302 23.416 20.132
28.520 22.659 39.512 13.472 24.739 19.989 20.254 17.104 27.621
19.104 25.813 20.674 27.496 28.215 38.088 26.996 28.212 31.076
23.094 11.001 47.299 22.531 39.436 36.762 21.599 17.056 32.537
19.742 26.837 32.284 15.667
16.227 19.215 22.386 29.715 32.800 27.404 20.169 28.520 31.937
10.166 21.056 29.212 31.499 24.498 19.874 22.015 26.416 20.065
24.361 15.875 23.554 22.918 29.364 8.824 39.471 31.524 22.540

27.199 32.188 16.253 25.367 29.801 16.587 17.143 16.361 31.391
23.945 29.809 26.469 21.531
24.490 29.751 28.180 29.900 26.594 25.962 29.763 25.738 27.326
23.652 29.891 7.512 18.985 24.561 12.832 29.570 33.375 38.839
25.904 23.252

Варіант 6

30.955 25.390 10.228 31.923 19.152 32.138 11.392 15.993 34.724
19.393 29.282 49.499 13.646 22.657 35.698 21.039 0.872 30.717
26.983 24.969 33.264 24.239 39.643 15.477 19.255 32.765 21.332
24.844 31.059 35.119 44.561 19.419 20.392 13.526 26.942 40.574
32.849 12.289 32.513 26.648
38.454 27.594 23.456 22.029 35.678 7.350 28.492 24.895 37.877
40.454 10.884 27.704 34.898 20.180 20.446 36.317 29.550 17.412
30.235 36.375 0.618 25.771 21.699 20.505 20.179 41.038 18.652
17.666 40.535 28.207 34.921 18.306 36.323 35.819 47.037 25.522
26.584 34.506 25.249 16.445
4.808 11.552 21.204 21.725 35.461 44.209 6.975 26.248 24.473
31.234 25.734 28.482 36.141 37.581 20.248 26.709 16.639 29.382
43.715 9.791

Варіант 7

26.222 22.918 20.302 25.551 16.863 26.589 30.317 22.953 18.074
28.450 29.042 31.123 24.812 28.935 25.832 25.690 30.785 24.196
26.553 27.779 26.065 24.782 15.519 24.989 26.881 21.446 25.265
27.047 23.753 19.399 29.398 27.912 14.360 25.466 23.865 28.885
18.558 20.237 27.519 17.336
35.194 19.967 19.115 23.109 19.255 28.112 23.145 26.587 20.537
19.511 29.620 25.323 20.047 17.587 35.122 28.237 21.211 25.390
34.543 17.921 26.335 28.002 19.135 24.781 13.202 22.872 22.180
31.622 25.611 25.298 28.208 18.192 16.373 29.804 22.450 30.243
27.776 23.934 30.235 22.944
27.127 26.738 26.441 33.691 18.392 28.304 22.216 24.328 23.865
23.160 18.426 24.091 17.029 22.849 21.819 37.544 20.524 19.143
24.188 18.195

Варіант 8

44.355 28.821 17.883 35.364 27.888 46.645 16.394 68.469 39.814
39.944 40.273 24.243 22.827 45.237 29.362 28.177 37.059 43.731
46.156 39.866 33.924 20.561 32.441 39.851 45.636 44.720 34.254
25.723 15.730 30.635 47.254 47.454 31.096 26.593 18.377 37.079
49.388 45.040 42.938 27.809
24.512 24.849 58.333 15.434 11.519 28.933 44.620 21.460 60.662
27.925 39.976 32.925 28.139 52.520 40.048 20.680 47.640 17.906
41.494 45.243 39.181 20.496 32.000 59.601 39.809 42.764 26.014
43.609 16.509 12.764 39.191 31.435 16.168 36.446 36.142 30.756
23.764 25.665 32.562 32.512
26.249 3.785 29.271 23.121 40.960 35.557 19.771 38.792 29.000
38.244 38.922 21.956 18.944 49.373 47.987 25.485 40.976 51.878
39.596 27.167

Варіант 9

44.545 24.129 45.817 26.460 9.065 10.412 5.661 1.133 22.503 14.263
33.574 16.762 9.749 47.526 23.575 29.581 2.484 26.012 11.626
17.858 16.234 48.054 57.208 29.266 39.280 17.001 4.295 33.314
42.852 31.603 38.227 20.315 15.209 31.053 53.847 29.286 87.104
4.840 61.037 31.167
29.636 45.813 1.577 43.112 36.874 25.752 54.447 51.825 23.127
17.684 26.803 54.517 62.690 6.613 12.554 26.804 40.140 26.707
43.727 38.613 20.774 31.863 33.745 20.411 41.123 24.929 15.515
73.244 24.913 60.331 36.229 40.556 43.978 57.947 11.553 59.666
12.073 27.182 37.042 29.977
1.923 47.790 27.735 46.957 15.781 47.206 45.953 43.184 58.925
30.909 27.033 33.380 14.588 9.374 21.925 36.181 66.302 12.151
37.406 9.030

Варіант 10

11.765 57.958 41.441 46.674 87.340 61.508 43.632 51.529 80.318
11.657 6.798 58.348 95.323 0.320 0.184 48.078 35.097 43.826
39.920 3.0932 23.413 9.409 35.767 36.192 106.322 61.902 45.968
11.665 26.517 51.589 50.768 51.211 44.249 21.482 80.782 50.142
138.800 3.052 69.565 22.297

80.698 32.254 6.853 38.277 40.055 22.973 31.627 18.052 0.605
45.848 88.955 12.192 85.227 27.483 25.964 19.585 -11.698 22.312
55.211 24.443 31.493 14.192 38.750 29.303 14.040 58.891 92.453
3.407 14.927 43.086 25.394 73.444 50.903 4.013 59.983 14.864
53.295 54.433 22.990 41.934
59.525 10.708 6.438 25.301 14.162 70.600 15.563 25.471 14.846
48.976 49.705 2.105 60.621 57.996 21.144 58.721 113.520 88.070
52.269 90.607

Практична робота № 3.
Знаходження статистичних оцінок для згрупованих даних

Спостереження проводились за декількома групами:

1) знайти середні оцінки для кожної групи та оцінити середнє вибірки;

2) знайти групову, міжгрупову, внутрішньогрупову і загальну дисперсію;

3) обчислити коефіцієнт варіації.

Теоретичні відомості

Середні вибіркові для кожної групи: $\bar{x}_j = \dfrac{\sum\limits_{i=1}^{N_j} n_{ji} x_{ji}}{N_j}$,

де j – номер групи, N_j – об’єм j-ї групи , n_{ji} – частоти відповідної групи, x_{ji} – значення варіант відповідної групи.

Оскільки математичне сподівання – оцінка незміщена, то ***загальне середнє вибіркове*** можна оцінити середнє арифметичне групових середніх: $\bar{x}_в = \sum\limits_{j=1}^{m} N_j \cdot \bar{x}_j \Big/ N$, де m – кількість груп, де $N = \sum\limits_{j=1}^{m} N_j$ – загальний об’єм сукупності.

Оскільки дисперсія є зміщеною оцінкою, то розрізняють групову, міжгрупову, внутрішньогрупову і загальну дисперсію.

Групова дисперсія – це дисперсія значень в групі, а саме відхилення значень варіант в групі від середнього групи:

$$D_{jгг} = \frac{\sum_{i=1}^{N_j} n_{ji} \cdot \left(x_{ji} - \bar{x}_j\right)^2}{N_j}.$$

Внутрішньогруповою дисперсією називають середню арифметичну дисперсію, зважену за об'мами груп:

$$D_{внгр} = \frac{\sum_{j=1}^{m} N_j \cdot D_{jгг}}{N}.$$

Міжгруповою дисперсією називають дисперсію групових середніх щодо загальної середньої:

$$D_{міжгр} = \frac{\sum_{j=1}^{m} N_j \cdot \left(\bar{x}_j - \bar{x}\right)^2}{N}.$$

Загальною дисперсією називають дисперсію значень ознаки всієї сукупності щодо загальної середньої:

$$D_{заг} = \frac{\sum_{j=1}^{m}\sum_{i=1}^{N_j} n_{ji} \cdot \left(x_{ji} - \bar{x}\right)^2}{N}.$$

Зауваження: $D_{заг} = D_{внгр} + D_{міжгр}$.

Приклад виконання роботи

Нехай сукупність складається з двох груп, заданих своїми варіаційними рядами (табл. 3.1):

Таблиця 3.1

Вибірки двох груп

1-а група							2-а група			
x_i		2	4	5			x_i	3	8	
n_i		1	7	2	$N = 10$		n_i	2	5	$N = 5$

Знайти точкові статистичні оцінки за згрупованими даними.

Розв'язок. Знайдемо середні вибіркові для кожної групи:

$$\bar{x}_1 = \frac{2 \cdot 1 + 4 \cdot 7 + 5 \cdot 2}{10} = 4, \quad \bar{x}_2 = \frac{3 \cdot 2 + 8 \cdot 5}{5} = 6.$$

Знайдемо загальне середнє: $\bar{x} = \dfrac{4 \cdot 10 + 6 \cdot 5}{15} = \dfrac{14}{3}$.

Знайдемо групові дисперсії:

$$D_{1гр} = \frac{1 \cdot (2-4)^2 + 7 \cdot (4-4)^2 + 2 \cdot (5-4)^2}{10} = 0{,}6,$$

$$D_{2гр} = \frac{2 \cdot (3-6)^2 + 3 \cdot (8-6)^2}{5} = 6.$$

Знайдемо внутрішньогрупові дисперсії: $D_{внгр} = \dfrac{10 \cdot 0{,}6 + 5 \cdot 6}{15} = \dfrac{36}{15}$

Знайдемо міжгрупову дисперсію:

$$D_{міжгр} = \frac{10 \cdot (4 - 14/3)^2 + 5 \cdot (6 - 14/3)^2}{15} = \frac{8}{9}.$$

Знайдемо загальну дисперсію:

$$D_{заг} = \frac{1 \cdot (2 - 14/3)^2 + 7 \cdot (4 - 14/3)^2 + 2 \cdot (5 - 14/3)^2}{15} +$$

$$+ \frac{2 \cdot (3 - 14/3)^2 + 3 \cdot (8 - 14/3)^2}{15} = \frac{148}{45}.$$

Виконаємо перевірку; $\dfrac{36}{15} + \dfrac{8}{9} = \dfrac{148}{45}$.

Приклад 2. Нехай проведено обстеження 3-х незалежних груп стосовно ознаки X. Результати обстеження задані табл. 3.2.

Таблиця 3.2

Результати обстеження 3-х груп

1-а група	2-а група	3-я група
4,7	5	4,9
4,2	4,8	4,7
3,6	3,5	4,4
3,4	3,3	4
3,1	3,3	3,2

Кожне значення з’явилось один раз. Знайти оцінки загального середнього та загальної дисперсії.

Будемо виконувати завдання, використовуючи середовище *Mathlab*.

Дані для кожної групи запишемо в окремий вектор стовпчик та визначимо об’єми груп (рис. 3.1).

Вхідні дані

$$n1 := \begin{pmatrix} 4.7 \\ 4.2 \\ 3.6 \\ 3.4 \\ 3.1 \end{pmatrix} \qquad n2 := \begin{pmatrix} 5 \\ 4.8 \\ 3.5 \\ 3.3 \\ 3.3 \end{pmatrix} \qquad n3 := \begin{pmatrix} 4.9 \\ 4.7 \\ 4.4 \\ 4 \\ 3.2 \end{pmatrix}$$

$$N_1 := 5 \qquad N_2 := 5 \qquad N_3 := 5$$

Рис. 3.1. Задання початкових даних

Скориставшись інструментом знаходження суми (панель Аналіз), знайдемо групові середні та загальне середнє (рис. 3.2). При знаходженні загального середнього врахуємо, що групи мають однаковий об’єм.

Середні групові та загальне

$$x1 := \frac{\sum_{i=0}^{4} \left(n1_i\right)}{N_1} = 3.8 \qquad x2 := \frac{\sum_{i=0}^{4} \left(n2_i\right)}{N_2} = 3.98 \qquad x3 := \frac{\sum_{i=0}^{4} \left(n3_i\right)}{N_3} = 4.24$$

загальне середнє $xv := \frac{x1 + x2 + x3}{3} = 4.007$

Рис. 3.2. Обчислення середніх значень

Знайдемо дисперсії для кожної групи (рис. 3.3).

Обчислення дисперсії

$$D_1 := \frac{\sum_{i=0}^{4} \left(n1_i - x1\right)^2}{N_1} = 0.332 \qquad D_2 := \frac{\sum_{i=0}^{4} \left(n2_i - x2\right)^2}{N_2} = 0.574 \qquad D_3 := \frac{\sum_{i=0}^{4} \left(n3_i - x3\right)^2}{N_3} = 0.362$$

Рис. 3.3. Обчислення групових дисперсій

Далі обчислимо міжгрупову, внутрішньогрупову дисперсії і знайдемо загальну дисперсію (рис. 3.4).

Внутрішньогрупова д

$$Dv := \frac{\sum_{i=1}^{3} (N_i \cdot D_i)}{N_1 + N_2 + N_3} = 0.423$$

Міжгрупова дисперсія

$$Dm := \frac{(x1 - xv)^2 \cdot N_1 + (x2 - xv)^2 \cdot N_2 + (x3 - xv)^2 \cdot N_3}{\sum_{i=1}^{3} N_i} = 0.033$$

Загальна дисперсія

$$D := Dv + Dm = 0.455$$

Середньоквадратичне відхилення

$$\sigma := \sqrt{D} = 0.675$$

Коефіцієнт варіації

$$V := \frac{\sigma}{xv} \cdot 100 = 16.841$$

Рис. 3.4. Знаходження загальної дисперсії та коефіцієнту варіації

Варіанти для самостійного виконання

Варіант 1

Є дані стосовно кількості дітей у сім'ях різних підрозділів корпорації:

Кількість дітей	Кількість сімей співробітників по підрозділах		
	перший	***другий***	***третій***
0	4	7	5
1	6	10	13
2	3	3	3
3	2	1	-

Варіант 2

Для дослідження успішності з трьох груп студентів першого курсу відібрали по п'ять студентів і отримали наступні результати середніх балів:

№ студ.	1-а група	2-а група	3-я група
1	4,7	5	4,9
2	4,2	4,8	4,7
3	3,6	3,5	4,4
4	3,4	3,3	4
5	3,1	3,3	3,2

Варіант 3

Для контролю якості виробів з декількох верстатів відібрали партії продукції. Середні значення розмірів наведені у таблиці.

№ верстата	Розміри							
1	19,65	19,7	19,7	20,05	20,6	21,4	21,5	22
2	19,6	19,8	19,85	20	20	21,5	21,8	-
3	19,8	20	20	20,3	20,4	20,4	21,3	22,2

Варіант 4

Якість проєкту оцінило три групи експертів за шкалою від 0 д 1. Результати оцінювання наведені у таблиці

№ групи	Оцінки							
1	0,65	0,7	0,25	0,5	0,6	0,4	0,5	0,3
2	0,6	0,8	0,5	0,4	0,9	0,5	0,8	0,2
3	0,8	0,5	0,4	0,3	0,7	0,4	0,3	0,2

Варіант 5

Для характеристики виробничого стажу працівників проведено обстеження декількох груп різних категорій працівників. Результати обстеження наведені у таблиці.

Стаж	Кількість працівників		
	робітники	технологи	конструктори
До 2 р.	7	1	-
2 – 4	15	10	3
4 – 6	20	22	20
6 – 8	30	20	10
8 – 10	10	23	32
10 – 12	8	7	20
12 – 14	2	6	10
більше 14	8	11	5

Варіант 6

В таблиці наведено дані випуску продукції підприємствами за тиждень. Визначити необхідні характеристики та визначити, яке підприємство працювало більш ритмічно.

Підприємство	Тиждень						
	I	II	III	IV	V	VI	VII
1	11	12	16	17	24	40	32
2	19	21	18	22	28	30	26

Варіант 7

В таблиці наведено дані кількості бракованих деталей на 100 одиниць продукції, які виробляються декількома працівниками за кожну декаду півріччя. Визначити необхідні характеристики та визначити, який працівник працює краще.

Працівник	Декада					
1	6	3	2	5	6	4
2	2	5	5	7	9	5
3	4	-	7	3	7	4
4	5	3	9	2	-	7

Варіант 8

В таблиці наведено дані кількості бракованих деталей на 100 одиниць продукції, які виробляються декількома бригадами за кожну декаду півріччя. Визначити необхідні характеристики та визначити, яка бригада працює краще.

Бригада	К-ть працівників	Декада					
1	10	9	6	7	11	6	12
2	11	11	7	7	10	9	13
3	16	13	9	10	15	7	17

Варіант 9

В таблиці наведено дані кількості зерен будяків у пробах зерна з декількох партій. Визначити необхідні характеристики та визначити, яка парія краща.

№ партії	К-ть зерен	К-ть зерен будяків в пробі					
1	200	19	16	17	21	16	32
2	250	11	17	27	30	19	33
3	300	23	29	30	35	17	37

Варіант 10

На декількох ринках зареєстрована на протязі місяця наступна ціна на картоплю. Визначити необхідні характеристики.

№ ринку	Ціна за 1 кг												
1	10	15	20	17	12	15	15	15	20	20	17	15	20	17
2	20	25	25	20	18	17	18	20	15	-	-	-	-	-
3	15	15	20	17	17	20	20	15	18	18	17	25	20	15

Практична робота № 4.
Лінійна кореляція. Побудова лінії регресії

1. За даними, наведеними у кореляційній таблиці, знайти вибіркове рівняння прямої лінії регресії ***У*** на ***Х***.

2. Оцінити тісноту зв'язку за вибірковим коефіцієнтом кореляції (табл. 4.1). Провести аналіз значущості вибіркового коефіцієнта кореляції при рівні значущості $\alpha = 0{,}05$.

3. Побудувати графік та оцінити похибку.

Таблиця 4.1

Тіснота зв'язку і величина коефіцієнта кореляції

Коефіцієнт кореляції r_{xy}	Тіснота зв'язку
0,91 –1	дуже сильний
0,81 – 0,90	досить сильний
0,65 – 0,80	сильний
0,45 – 0,64	помірний
0,25 – 0,44	слабкий
До 0,25	дуже слабкий

Для спрощення обчислень та побудови графіків завдання доцільно виконувати в середовищі Mathcad.

Теоретичні відомості

Статистичною називають залежність, за якої зміна однієї з величин тягне за собою зміну розподілу іншої. Зокрема, статистична залежність проявляється в тому, що у разі зміни однієї з величин змінюється середнє значення іншої; в цьому випадку статистичну залежність називають ***кореляційною***.

Кореляція (пізньолатинське *correlatio* – співвідношення) – взаємозв`язок, взаємозалежність, взаємовідповідність; співвідношення понять, функцій, предметів тощо.

Лінійна регресія – це метод моделювання залежності між скалярною змінною ***Y*** та векторною (у загальному випадку) змінною ***X***:

$$\overline{y}_x - \overline{y} = r_в \frac{\overline{\sigma}_y}{\overline{\sigma}_x}\left(x - \overline{x}\right),$$

де $\overline{x}$, $\overline{y}$ – середні вибіркові, $\overline{\sigma}_x$, $\overline{\sigma}_y$ – відповідні середньоквадратичні відхилення, $r_в$ – ***вибірковий коефіцієнт кореляції,*** який відображає тісноту зв'язку між випадковими величинами. Якщо $r_в = 1$, то зв'язок строго лінійний.

Лінійна регресія повертає розподіл умовної імовірності Y залежно від ***X***.

Під час розрахунків параметрів моделі лінійної регресії зазвичай застосовується метод найменших квадратів (МНК).

Порядок побудови ліній регресії

1. До заданої кореляційної таблиці додати ще один стовпчик та рядок, де в стовпчику визначити частоти n_y – суми частот за рядками, а в нижньому рядку n_x – суми частот за стовпчиками. Знайти загальний об'єм:

$$N = \sum_{i=1}^{n} n_{yi} = \sum_{j=1}^{m} n_{xj}.$$

2. Знайти середні значення для x та y:

$$\overline{x} = \frac{\sum_{j=1}^{m} n_{xj} \cdot x_j}{N}, \quad \overline{y} = \frac{\sum_{i=1}^{n} n_{yi} \cdot y_i}{N}.$$

3. Для обчислення середньоквадратичного відхилення скористаємось спрощеною формулою. Для цього знайдемо середнє квадратів:

$$\overline{x}^2 = \frac{\sum_{j=1}^{m} n_{xj} \cdot x_j^2}{N}, \quad \overline{y} = \frac{\sum_{i=1}^{n} n_{yi} \cdot y_i^2}{N}.$$

Обчислимо середньоквадратичні відхилення:

$$\sigma_x = \sqrt{\overline{x}^2 - (\overline{x})^2}, \ \sigma_y = \sqrt{\overline{y}^2 - (\overline{y})^2}.$$

4. Обчислимо сумісну появу значень x та y – $\sum n_{xy}xy$. Для цього складемо розрахункову таблицю:

– у верхньому куточку кожної клітинки кореляційної таблиці запишемо добуток $X = x_j \cdot n_{ij}$;

– у верхньому куточку кожної клітинки кореляційної таблиці запишемо добуток $Y = y_i \cdot n_{ij}$.

Допишемо до кореляційної таблиці ще один стовпчик і один рядок, де в стовпчик занесемо значення суми значень, що стоять у верхньому куточку на відповідне значення:

$$y_i - \left(\sum_{j=1}^{m} x_j \cdot n_{ij} \right) \cdot y_i,$$

а у рядок – суми значень, що стоять у нижньому куточку, на відповідне значення:

$$x_j - \left(\sum_{i=1}^{n} y_i \cdot n_{ij} \right) \cdot x_j.$$

Знайдемо суми значень додаткового стовпчика та додаткового рядка, це і буде шукана сума сумісної появи випадкових величин:

$$\sum_{i=1}^{n} \left(\sum_{j=1}^{m} x_j \cdot n_{ij} \right) \cdot y_i = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} y_i \cdot n_{ij} \right) \cdot x_j. = \sum n_{xy}xy.$$

Значення цих сум повинно співпадати (рис. 4.1).

v \ u	-2	-1	0	1	2	$U = \sum n_{uv} \cdot u$	$v \cdot U$
-2	-8; 4; -8	-6; 6; -12				-14	28
-1	—	-8; 8; -8	0; 10; -10			-8	8
0	—	—	0; 32; 0	3; 3; 0	18; 9; 0	21	0
1	—	—	0; 4; 4	12; 12; 12	12; 6; 6	24	24
2	—	—	—	1; 1; 2	10; 5; 10	11	22
$V = \sum n_{uv} \cdot v$	-8	-20	-6	14	16		$\sum_v v \cdot U = 82$ ↑
$u \cdot V$	16	20	0	14	32	$\sum_u u \cdot V = 82$ ←	—Контроль

Рис. 4.1. Приклад обчислення сумісної появи випадкових величин

5. Ступінь зв'язку між випадковими величинами X та Y визначається такими числовими характеристиками їх сумісного розподілу, як коваріація та кореляція:

$$\operatorname{cov}(X,Y)=M(XY)-M(X)M(Y),\ cor(X,Y)=\frac{\operatorname{cov}(X,Y)}{\sigma_X\sigma_Y}.$$

Коефіцієнт коваріації випадкових величин X та Y визначається як:

$$m_{XY}=\frac{1}{N-1}\sum(x_i-\bar{x})(y_i-\bar{y}).$$

Вибірковий коефіцієнт кореляції обчислюється за формулою:

$$r_в=\frac{\sum n_{xy}xy-N\cdot\bar{x}\cdot\bar{y}}{N\cdot\sigma_x\cdot\sigma_y}.$$

За табл. 4.1 оцінимо тісноту зв'язку між випадковими величинами.

6. Побудуємо вибіркову лінію регресії:

$$f_y(x)=r_в\cdot\frac{\sigma_y}{\sigma_x}(x-\bar{x})+\bar{y}.$$

7. Побудуємо графік знайденої лінійної регресії. На осі OX відкладемо значення аргумента у діапазоні значень випадкової величини X, на осі OY – відповідні значення Y, знайдені з отриманого рівняння. Для наочного представлення похибки експериментальних даних кореляційної таблиці та теоретичних значень, отриманих з рівняння, для кожного значення величини X знайдемо середнє значення $\bar{Y}$ і нанесемо на графік пари точок $(x_j,\bar{y}_j)$.

8. Знайдемо середньоквадратичну похибку:

$$\varepsilon=\sqrt{\frac{\sum_{j=1}^{m}\left(f_y(x_j)-\bar{y}_j\right)^2}{m}}.$$

9. Оскільки $r_в$ – випадкова величина, яка визначається за вибіркою, то її значення залежить від об'єму вибірки, і можуть бути ситуації, коли $r_в\neq 0$, а залежності між випадковими величинами немає, і навпаки. Отже для перевірки наявності зв'язку між випадковими величинами необхідно перевірити гіпотезу про значущість величини вибіркового коефіцієнта кореляції: $H_0:r_в=0$. Якщо гіпотеза не підтверджується, то коефіцієнт кореляції є значущим. За наявності невеликих обсягів

вибірки $(N \le 100)$ для перевірки гіпотези в якості критерію застосовують випадкову величину з розподілом Стьюдента: $t_{сд} = \sqrt{\frac{r_в^2}{1-r_в^2}(n-2)}$.

За заданим рівнем значущості α та кількістю степенів свободи $k = N - 1$ з таблиці додатку знайти критичне значення критерію $t_{кр}$. Якщо $t_{сп} > t_{кр}$, то коефіцієнт кореляції вважається значущим і робиться висновок про існування статистичного зв'язку між випадковими величинами.

Приклад виконання роботи

Нехай задана кореляційна таблиця:

У	*X*						
	5	10	15	20	25	30	35
100	-	-	-	-	-	6	1
120	-	-	-	-	-	4	2
140	-	-	8	10	5	-	-
160	3	4	3	-	-	-	-
180	2	1	-	1	-	-	-

Запишемо значення випадкових величин у вигляді відповідних векторів стовпчиків, а частоти представимо у вигляді матриці (рис. 4.2).

$$x := \begin{pmatrix} 5 \\ 10 \\ 15 \\ 20 \\ 25 \\ 30 \\ 35 \end{pmatrix} \qquad y := \begin{pmatrix} 100 \\ 120 \\ 140 \\ 160 \\ 180 \end{pmatrix} \qquad n := \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 6 & 1 \\ 0 & 0 & 0 & 0 & 0 & 4 & 2 \\ 0 & 0 & 8 & 10 & 5 & 0 & 0 \\ 3 & 4 & 3 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Рис. 4.2. Задання значень кореляційної таблиці

1. Побудуємо лінію регресії. Знайдемо n_y – суми частот за рядками, а в нижному рядку n_x – суми частот за стовпчиками та загальний об'єм сукупності (рис. 4.3).

$$i := 0..4 \qquad j := 0..6$$

$$ny_i := \sum_{j=0}^{6} n_{i,j} \qquad ny = \begin{pmatrix} 7 \\ 6 \\ 23 \\ 10 \\ 4 \end{pmatrix} \qquad nx_j := \sum_{i=0}^{4} n_{i,j} \qquad nx = \begin{pmatrix} 5 \\ 5 \\ 11 \\ 11 \\ 5 \\ 10 \\ 3 \end{pmatrix} \qquad N := \sum_{i=0}^{4} \sum_{j=0}^{6} n_{i,j} = 50$$

Рис. 4.3. Обчислення частот n_x, n_y та об’єму вибірки

Обчислимо середні значення та середньоквадратичні відхилення (рис. 4.4).

$$xs := \frac{\sum_{i=0}^{6} (nx_i \cdot x_i)}{N} = 19.8 \qquad ys := \frac{\sum_{i=0}^{4} (ny_i \cdot y_i)}{N} = 139.2$$

$$xs2 := \frac{\sum_{i=0}^{6} \left[nx_i \cdot (x_i)^2\right]}{N} = 466 \qquad ys2 := \frac{\sum_{i=0}^{4} \left[ny_i \cdot (y_i)^2\right]}{N} = 19856$$

$$\sigma x := \sqrt{xs2 - xs^2} = 8.6 \qquad \sigma y := \sqrt{ys2 - ys^2} = 21.894$$

Рис. 4.4. Обчислення точкових оцінок

Обчислимо сумісну появу випадкових величин (рис. 4.5).

$$u_i := \sum_{j=0}^{6} (n_{i,j} \cdot x_j) \qquad v_j := \sum_{i=0}^{4} (n_{i,j} \cdot y_i)$$

$$\boxed{nxy := \sum_{i=0}^{4} (u_i \cdot y_i) = 129800} \qquad \boxed{nxx := \sum_{i=0}^{6} (v_i \cdot x_i) = 129800}$$

Рис. 4.5. Сумісна поява випадкових величин

Знайдемо вибірковий коефіцієнт кореляції та побудуємо функцію лінії регресії (рис. 4.6). За табл. 4.1 оцінимо тісноту статистичного зв’язку між випадковими величинами. Значення вибіркового коефіцієнту кореляції за модулем свідчить, що зв’язок досить сильний.

$$r := \frac{nxy - N \cdot xs \cdot ys}{N \cdot \sigma x \cdot \sigma y} = -0.851$$

$$Y(X) := r \cdot \frac{\sigma y}{\sigma x} \cdot (X - xs) + ys \;\left|\begin{array}{l} \text{float}, 3 \\ \text{simplify} \end{array}\right. \rightarrow -2.16 \cdot X + 182.0$$

Рис. 4.6. Обчислення вибіркового коефіцієнту кореляції та побудова лінії регресії

Для побудови лінії регресії скористаємось панеллю інструментів Символьні. Значення коефіцієнтів будемо відображати з точністю до 3-го знаку (*float*,3), *simplify* – спрощення розкриє дужки та приведе подібні.

2. Оцінимо тісноту зв’язку. $|r_в| = 0{,}851$, отже між випадковими величинами існує досить сильний статистичний зв’язок.

Перевіримо гіпотезу про значущість коефіцієнта кореляції. Обчислимо значення критерію, що спостерігається:

$$t_{cg} = \sqrt{\frac{r_в^2}{1-r_в^2}(n-2)} = 11{,}227.$$

З таблиці розподілу Стьюдента (дод. 6) для $\alpha = 0{,}05$ і $k = 49$ знайдемо критичне значення $t_{кр} = 2{,}01$. Оскільки $t_{сп} > t_{кр}$, то статистичний зв’язок між випадковими величинами існує і зв’язок є значущим.

3. Знайдемо для кожного значення випадкової величини X середнє значення Y. Задамо значення аргументу знайденої функції в діапазоні від 5 до 50 (трохи збільшимо діапазон з метою прогнозування протікання процесу у майбутньому).

Побудуємо графік функції та нанесемо точки $(x_j, \bar{y}_j)$ (рис. 4.7).

$$X := 5..50 \qquad yz_j := \frac{\sum_{i=0}^{4} \left(n_{i,j} \cdot y_i\right)}{nx_j}$$

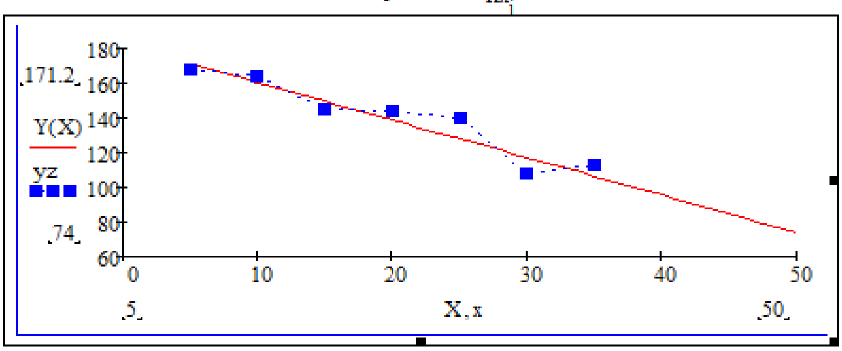


Рис. 4.7. Графік лінії регресії

Оцінимо точність наближення (рис. 4.8).

$$\varepsilon := \frac{\sqrt{\sum_{i=0}^{6} \left(Y(x_i) - yz_i\right)^2}}{7} = 2.636$$

Рис. 4.8. Середньоквадратична похибка наближення

Варіанти для самостійного виконання

Варіант 1

У	*X*							
	5	10	15	20	25	30	35	40
100	2	1	-	-	-	-	-	-
120	3	4	3	-	-	-	-	-
140	-	-	5	10	8	-	-	-
160	-	-	-	1	-	6	1	1
180	-	-	-	-	-	-	4	1

Варіант 2

У	*X*							
	18	23	28	33	38	43	48	53
125	-	1	-	-	-	-	-	-
150	1	2	5	-	-	-	-	-
175	-	3	2	12	3	-	-	-
200	-	-	1	8	7	2	1	-
250	-	-	-	-	-	-	4	1

Варіант 3

У	*X*						
	5	10	15	20	25	30	35
100	-	-	-	-	-	6	1
120	-	-	-	-	-	4	2
140	-	-	8	10	5	-	-
160	3	4	3	-	-	-	-
180	2	1	-	1	-	-	-

Варіант 4

У	*X*						
	20	25	30	35	40	45	50
10	4	6	-	-	-	-	-
20	-	7	8	-	-	-	-
30	-	-	29	3	8	-	-
40	-	-	4	12	6	3	-
50	-	-	-	1	5	2	2

Варіант 5

У	*X*						
	5	10	15	20	25	30	35
100	5	2	-	-	-	-	-
120	3	4	-	-	-	-	-
140	-	-	8	8	4	-	-
160	-	-	2	4	3	1	-
180	-	-	-	-	1	2	2

Варіант 6

У	*X*						
	12	22	32	42	52	62	72
16	3	6	-	-	-	-	-
26	-	7	8	-	-	-	-
36	-	-	30	5	7	-	-
46	-	-	4	10	7	2	-
56	-	-	-	1	5	2	3

Варіант 7

У	*X*						
	2	7	13	18	23	28	33
15			-	-	-	4	5
20	4	-	-	-	-	8	7
25	-	-	-	7	30	5	-
30	-	-	3	10	7	2	-
35	1	1	5	2	3	-	-

Варіант 8

У	*X*						
	22	27	32	37	42	47	52
36	2	3	-	-	-	-	-
46	-	5	8	-	-	-	-
56	-	-	2	7	6	-	-
66	-	-	1	2	9	4	-
76	-	-	-	1	-	4	6

Варіант 9

У	*Х*						
	5	10	15	20	25	35	40
48			-	-	-	2	1
58	-	-	-	-	3	4	3
68	-	-	5	10	8	-	-
78	-	1	6	1	1	-	-
88	4	1	-	-	1	-	-

Варіант 10

У	*Х*						
	17	22	27	32	37	42	47
110	3	3	-	-	-	-	-
120	-	5	7	-	-	-	-
130	-	-	3	6	5	-	-
140	-	-	1	3	7	4	-
150	-	-	-	1	-	6	6

Практична робота № 5.
Криволінійна кореляція

За даними, наведеними у кореляційній таблиці:

1. Знайти вибіркове рівняння квадратичної лінії регресії ***У*** на ***Х***.
2. Побудувати графік.
3. Оцінити тісноту зв’язку.

Для виконання розрахунків та побудови графіків доцільно застосовувати програмний додаток *Mathcad*.

Теоретичні відомості

Якщо графік регресії – крива лінія, то кореляцію називають ***криволинійною***.

Зокрема, у випадку параболічної кореляції другого порядку вибіркове рівняння регресії *Y* на *X* має вигляд: $Y(X) = aX^2 + bX + c.$

Для побудови криволінійної лінії регресії, також, застосовують метод найменших квадратів. Коефіцієнти рівняння квадратичної залежності знаходять з системи лінійних рівнянь:

$$\begin{gathered} a \cdot \sum n_x x^4 + b \cdot \sum n_x x^3 + c \cdot \sum n_x x^2 = \sum n_x \bar{y}_x x^2 \\ a \cdot \sum n_x x^3 + b \cdot \sum n_x x^2 + c \cdot \sum n_x x = \sum n_x \bar{y}_x x \\ a \cdot \sum n_x x^2 + b \cdot \sum n_x x + c \cdot n = \sum n_x \bar{y}_x . \end{gathered} \tag{5.1}$$

Для складання рівнянь заповнюють розрахункову таблицю (табл. 5.1), де в нижньому рядку запишемо суми значень за стовпчиками.

Таблиця 5.1

Розрахункова таблиця

x	n_x	$\bar{y}_x$	$n_x x$	$n_x x^2$	$n_x x^3$	$n_x x^4$	$n_x \bar{y}_x$	$n_x \bar{y}_x x$	$n_x \bar{y}_x x^2$
Σ									

Для оцінки тісноти зв'язку криволінійної кореляції застосовується вибіркове кореляційне відношення: $\eta_{yx} = \sigma_{міжгр} / \sigma_{заг} = \sigma_{\bar{y}_x} / \sigma_y$,

де $\sigma_{\bar{y}_x} = \sqrt{D_{міжгр}} = \sqrt{\dfrac{\sum n_x (\bar{y}_x - \bar{y})^2}{n}}$, $\sigma_y = \sqrt{D_{заг}} = \sqrt{\dfrac{\sum n_y (y - \bar{y})^2}{n}}$.

Приклад виконання роботи

Нехай задана кореляційна таблиця:

Y	**X**				
	0	1	2	3	4
0	18	1	1	-	-
3	1	20	-	-	-
5	3	5	10	2	-
10	-	-	7	12	-
17	-	-	-	-	20

Запишемо початкові дані X та Y у вигляді векторів стовпчиків, а частоти – у вигляді матриці (рис. 5.1).

$$x := \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \qquad y := \begin{pmatrix} 0 \\ 3 \\ 5 \\ 10 \\ 17 \end{pmatrix} \qquad n := \begin{pmatrix} 18 & 1 & 1 & 0 & 0 \\ 1 & 20 & 0 & 0 & 0 \\ 3 & 5 & 10 & 2 & 0 \\ 0 & 0 & 7 & 12 & 0 \\ 0 & 0 & 0 & 0 & 20 \end{pmatrix}$$

Рис. 5.1. Початкові дані

Знайдемо відповідні частоти по X та Y і для кожного X середнє значення Y (рис. 5.2).

$$i := 0..4$$

$$nx_i := \sum_{j=0}^{4} n_{j,i} \qquad ny_i := \sum_{j=0}^{4} n_{i,j} \qquad ys_i := \frac{\sum_{j=0}^{4} \left(n_{j,i} \cdot y_j\right)}{nx_i}$$

Рис. 5.2. Обчислення частот та середнього значення Y

Обчислимо коефіцієнти системи лінійних рівнянь 5.1 (рис. 5.3)

$$x1 := \sum_{i=0}^{4} \left(nx_i \cdot x_i\right) = 184 \qquad x2 := \sum_{i=0}^{4} \left[nx_i \cdot \left(x_i\right)^2\right] = 544 \qquad x3 := \sum_{i=0}^{4} \left[nx_i \cdot \left(x_i\right)^3\right] = 1828$$

$$x4 := \sum_{i=0}^{4} \left[nx_i \cdot \left(x_i\right)^4\right] = 6568 \qquad N := \sum_{i=0}^{4} nx_i = 100$$

$$b1 := \sum_{i=0}^{4} \left(nx_i \cdot ys_i\right) = 693 \qquad b2 := \sum_{i=0}^{4} \left(nx_i \cdot ys_i \cdot x_i\right) = 2075 \qquad b3 := \sum_{i=0}^{4} \left[nx_i \cdot ys_i \cdot \left(x_i\right)^2\right] = 7175$$

Рис. 5.3. Обчислення коефіцієнтів рівняння регресії

Складемо систему лінійних рівнянь 5.1 і розв'яжемо її відносно невідомих параметрів a, b, c. Для розв'язання системи лінійних рівнянь застосуємо стандартний блок *Given – Find*. На початку необхідно задати початкові значення невідомих. В якості початкових значень візьмемо нулі. Далі записується ключове слово *Given* – дано, після якого описується вигляд системи рівнянь. Необхідно звернути увагу, що ліва частина кожного рівняння прирівнюється правій, тому між ними ставиться знак логічної рівності (жирне =). Цей знак знаходиться на панелі Булева алгебра або його можна встановити комбінацією клавіш *Ctrl*+=. Розв'язок знаходиться за допомогою стандартної функції *Find* (функції – категорія Розв'язання рівнянь), параметрами якої є шукані змінні. Результатом виконання функції є вектор стовпчик невідомих параметрів, з якого виділимо окремі значення кожного з них (рис. 5.4).

$$A := 0 \qquad B := 0 \qquad C := 0$$

Given

$$A \cdot x4 + B \cdot x3 + C \cdot x2 = b3$$

$$A \cdot x3 + B \cdot x2 + C \cdot x1 = b2$$

$$A \cdot x2 + B \cdot x1 + C \cdot N = b1$$

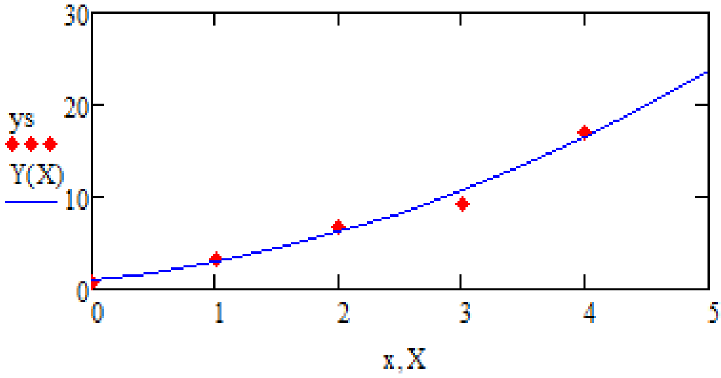
$$r := \mathrm{Find}(A, B, C) \qquad r = \begin{pmatrix} 0.663 \\ 1.226 \\ 1.069 \end{pmatrix} \qquad A := r_0 \qquad B := r_1 \qquad C := r_2$$

Рис. 5.4. Розв'язання системи лінійних рівнянь

Після знаходження значень параметрів, запишемо рівняння регресії.

4. Задамо значення аргументу X у відповідному діапазоні і побудуємо графік функції. Нанесемо на графік відповідні точки $(x_i, \bar{y}_i)$ і оцінимо точність наближення (рис. 5.5).

$$Y(X) := A \cdot X^2 + B \cdot X + C \qquad X := 0, 0.5 .. 5$$



$$\varepsilon := \frac{\sqrt{\sum_{i=0}^{4} \left(y_i - Y(x_i)\right)^2}}{5} = 0.358$$

Рис. 5.5. Графік квадратичної лінії регресії

Знайдемо вибіркове кореляційне відношення (рис. 5.6)

$$yv := \frac{\sum_{i=0}^{4} \left(ny_i \cdot y_i\right)}{N} = 6.93 \qquad \sigma z := \sqrt{\frac{\sum_{i=0}^{4} \left[ny_i \cdot \left(y_i - yv\right)^2\right]}{N}} = 5.972$$

$$\sigma m := \sqrt{\frac{\sum_{i=0}^{4} \left[nx_i \cdot \left(ys_i - yv\right)^2\right]}{N}} = 5.725 \qquad \eta := \frac{\sigma m}{\sigma z} = 0.959$$

Рис. 5.6. Обчислення вибіркового кореляційного відношення

За табл. 4.1 (завдання 4) оцінимо тісноту зв’язку. Як видно, зв’язок між випадковими величинами X та Y дуже сильний.

Варіанти для самостійного виконання

Варіант 1

У	*X*				
	0	1	2	3	4
0	18	1	1	-	-
3	1	20	-	-	-
5	3	5	10	2	-
10	-	-	7	12	-
17	-	-	-	-	20

Варіант 2

У	*X*				
	0	4	6	7	10
7	19	1	1	-	-
13	2	14	-	-	-
40	-	3	22	2	-
80	-	-	-	15	-
200	-	-	-	-	21

Варіант 3

У	*X*				
	0	4	5	7	8
1	50	1	1	-	-
35	-	44	-	-	-
50	-	5	45	2	-
80	-	-	7	15	-
210	-	-	-	2	20

Варіант 4

У	*X*				
	0	1	2	3	4
10	20	5	-	-	-
11	7	15	3	1	-
20	-	3	17	4	-
35	-	-	8	13	7
50	-	-	-	5	42

Варіант 5

У	*X*					
	15	20	25	30	35	40
100	-	-	-	-	2	2
120	-	-	4	3	10	3
140	-	2	50	7	10	-
160	1	4	-	3	-	
180	1	1	-	-	-	

Варіант 6

У	*X*				
	1	2	3	4	5
-1	-	-	-	6	4
0	-	-	1	4	6
1	-	5	9	5	-
2	3	7	-	-	-
3,5	4	1	-	-	-

Варіант 7

У	*X*				
	70	80	90	100	150
2	1	5	-	-	-
3,5	-	4	10	-	-
3	-	1	9	5	-
4	-	2	2	12	-
5	-	-	1	1	2

Варіант 8

У	*X*				
	165	170	175	180	185
5	12	-	13	-	-
9	11	11	-	13	11
20	-	11	12	15	11
45	-	-	11	11	11
90	-	-	11	-	11

Варіант 9

У	*X*					
	15	20	30	40	50	60
2	5	7	-	-	-	-
4	-	20	23	-	-	-
15	-	-	30	47	2	-
140	-	-	10	11	20	6
280	-	-	-	9	7	3

Варіант 10

У	*X*					
	1	2	3	4	5	6
2	3	7	9	-	-	-
3	-	20	23	-	-	-
8	-	-	30	47	2	-
44	-	-	10	11	20	6
180	-	-	-	-	5	7

Практична робота № 6.
Рангова кореляція

Знайти вибірковий коефіцієнт рангової кореляції за критерієм Спірмена та Кендалла за результатами оцінок, отриманих двома групами експертів або в результаті порівняння двох об'єктів чи предметів. Перевірити гіпотезу про значущість коефіцієнта рангової кореляції за рівня значущості $\alpha = 0{,}05$. Правильно представити табличні значення.

Якісну характеристику тісноти зв'язку коефіцієнта рангової кореляції можна оцінити за шкалою Чеддока (табл. 6.1).

Таблиця 6.1

Шкала оцінки тісноти зв'язку коефіцієнта рангової кореляції

Кількісна міра тісноти зв'язку	Якісна характеристика сили зв'язку
0,1 – 0,3	слабка
0,3 – 0,5	помірна
0,5 – 0,7	суттєва
0,7 – 0,9	висока
0,9 – 1	дуже висока

Для виконання розрахунків доцільно застосовувати програмний додаток *Mathcad*.

Теоретичні відомості

Кореляція рангу – це будь-яка із кількох статистик, які вимірюють **порядкове-з'єднання** – відношення між ранжируванням різних порядкових змінних або різних кореляцій однієї і тієї ж змінної, де «ранжируванням» називається наданням позначок «перший», «другий», «третій» і т.д. різним спостереженням конкретної змінної.

Коефіцієнт кореляції рангу вимірює ступінь подібності між двома ранжируваннями і може бути використаний для оцінки статистичної значущості співвідношення між ними.

Для оцінки ступеня зв'язку ознак X і Y служать, зокрема, коефіцієнти рангової кореляції Спірмена й Кендалла. Для визначення коефіцієнта рангової кореляції обидві послідовності необхідно відсортувати у однаковому порядку, наприклад, за зростанням із збереженням відповідних рангів.

Вибірковий коефіцієнт рангової кореляції Спірмена знаходять за формулою:

$$\rho_в = 1 - \frac{6\sum_i d_i^2}{n^3 - n},$$

де n – кількість елементів послідовностей, різниці між відповідними рангами $d = R_X - R_Y$, $|\rho_в| \le 1$.

Тіснота зв'язку перевіряється на основі гіпотези про значущість вибіркового коефіцієнта рангової кореляції, де в якості критерію береться випадкова величина з розподілом Стьюдента. Критичне значення критерію обчислюється за формулою:

$$T_{кр} = t_{кр}(\alpha, k)\sqrt{\frac{1-\rho_в^2}{n-2}},$$

де $t_{кр}(\alpha, k)$ – критична точка, значення якої знаходиться з таблиці розподілу Стьюдента за рівнем значущості α та кількістю степенів свободи $k = n - 2$. Якщо $|\rho_в| < T_{кр}$, то кореляційний зв'язок між ознаками незначущий.

Вибірковий коефіцієнта рангової кореляції за Кендаллом знаходять за формулою:

$$\tau_в = \frac{2R}{n(n-1)};$$

де n – кількість елементів послідовностей, $R = \sum_i (R_i^+ - R_i^-)$ – сума різниць рангів, де R_i^+ – сума рангів більших за поточне значення, R_i^- – сума рангів менших за поточне значення.

Тіснота зв'язку перевіряється на основі гіпотези про значущість вибіркового коефіцієнта рангової кореляції, де в якості критерію береться випадкова величина з нормальним розподілом. Критичне значення критерію обчислюється за формулою:

$$T_{кр} = z_{кр}\sqrt{\frac{2(2n+5)}{9n(n-1)}},$$

де $z_{кр}$ – критична точка, значення якої знаходиться з таблиці функції Лапласа з рівняння $\Phi\left[z_{кр}\right] = (1-\alpha)/2$. Якщо $|z_в| < T_{кр}$, то кореляційний зв’язок між ознаками незначущий.

Приклад виконання роботи

Знання 10 студентів перевірялись за двома дисциплінами *A* та *B*. Оцінки за тестами наведені у табл. 6.21.

Таблиця 6.2

Оцінки за двома дисциплінами

A	95	90	86	84	75	70	62	60	57	50
B	92	93	83	80	55	60	45	72	62	70

За допомогою коефіцієнта рангової кореляції оцінити, чи суттєво відрізняються знання студентів за різними дисциплінами?

Для спрощення виконання операцій застосуємо середовище *Mathcad*.

Як видно, оцінки за першою дисципліною вже впорядковані у порядку спадання, припишемо кожному студенту за першою дисципліною ранг і впорядкуємо оцінки за другою дисципліною також у порядку спадання за значенням оцінок із збереженням рангу студента. Для виконання цієї процедури оцінки та відповідні ранги запишемо у матриці розмірності 10 х 2, де перший стовпчик відповідає значенню оцінки, а другий – рангу (рис. 6.1).

$$A := \begin{pmatrix} 95 & 1 \\ 90 & 2 \\ 86 & 3 \\ 84 & 4 \\ 75 & 5 \\ 70 & 6 \\ 62 & 7 \\ 60 & 8 \\ 57 & 9 \\ 50 & 10 \end{pmatrix} \qquad B := \begin{pmatrix} 92 & 1 \\ 93 & 2 \\ 83 & 3 \\ 80 & 4 \\ 55 & 5 \\ 60 & 6 \\ 45 & 7 \\ 72 & 8 \\ 62 & 9 \\ 70 & 10 \end{pmatrix}$$

Рис. 6.1. Задання початкових даних

Для впорядкування оцінок за другою дисципліною застосуємо стандартну функцію *csort* (*a1*,*a2*), де перший параметр – ім'я матриці, другий – номер стовпчика, за яким необхідно виконати сортування, і розташуємо значення у зворотному порядку (функція *reverse*(), рис. 6.2).

$B1 := \mathrm{csort}(B,0)$ $\quad B1 := \mathrm{reverse}(B1)$ $\quad i := 0..9$ $\quad n := 10$

Критерій Спірмена

$d_i := A_{(i,1)} - B1_{i,1}$

A =

	0	1
0	95	1
1	90	2
2	86	3
3	84	4
4	75	5
5	70	6
6	62	7
7	60	8
8	57	9
9	50	10

B1 =

	0	1
0	93	2
1	92	1
2	83	3
3	80	4
4	72	8
5	70	10
6	62	9
7	60	6
8	55	5
9	45	7

d =

	0
0	-1
1	1
2	0
3	0
4	-3
5	-4
6	-2
7	2
8	4
9	3

Рис.6.2. Знаходження різниці між рангами

Як видно з отриманих результатів, що, наприклад, перший студент, який має найвищий ранг за першою дисципліною, за другою – вже має ранг два, четвертий – за другою дисципліною має ранг 8.

1. Знайдемо вибірковий коефіцієнт кореляції за Спірменом.

Знайдемо d_i – різниці між рангами оцінок для кожного студента (рис. 6.2). Обчислимо коефіцієнт рангової кореляції за Спірменом. З таблиці розподілу Стьюдента за рівнем значущості $\alpha = 0{,}05$ знаходимо значення $t_{кр}$ і обчислюємо критичне значення критерію (рис. 6.3).

$$\rho := 1 - \frac{6 \cdot \sum_{i=0}^{9} \left(d_i\right)^2}{n^3 - n} = 0.636 \qquad k := n - 2 = 8 \qquad tk := 2.31$$

$$T := tk \cdot \sqrt{\frac{1-\rho^2}{k}} = 0.63$$

Рис. 6.3. Обчислення коефіцієнта рангової кореляції

Значення критерію, отримане за результатами спостережень, переважає критичне, то можна вважати, зв'язок між оцінками значущий. За таблицею Чеддока, зв'язок можна вважати сильним.

2) Обчислення вибіркового коефіцієнта рангової кореляції за Кедаллом. Для кожного студента визначимо значення:

– R_i^+ – кількість рангів за другою дисципліною, що стоять правіше (або нижче) і перевищують поточний i-й ранг. Отже для першого студента $R_1^+ = 8$, бо ранг другого студента дорівнює 1 (менший за значенням), $R_2^+ = 8$, $R_3^+ = 7$, $R_4^+ = 6$, $R_5^+ = 2$, $R_6^+ = R_7^+ = 0$, $R_8^+ = R_9^+ = 1$;

– R_i^- – кількість рангів за другою дисципліною, що стоять правіше (або нижче) і менші за поточний ранг, $R_1^- = 1$, $R_2^- = 0$ і т.д.

Для підрахунку значень R_i^+ та R_i^- можна скласти програмний блок (рис. 6.4)

```
b := B1⟨1⟩   K(b,n) := | for i ∈ 0..n − 2
                       |    | Rp_i ← 0
                       |    | Rn_i ← 0
                       |    | for j ∈ i + 1..n − 1
                       |    |    | Rp_i ← Rp_i + 1  if  b_i < b_j
                       |    |    | Rn_i ← Rn_i + 1  if  b_i > b_j
                       | R⟨0⟩ ← Rp
                       | R⟨1⟩ ← Rn
                       | R
```

$R := K(b,n)$ $\quad Rp := R^{\langle 0\rangle}$ $\quad Rn := R^{\langle 1\rangle}$ $\quad Rp = \begin{pmatrix} 8 \\ 8 \\ 7 \\ 6 \\ 2 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$ $\quad Rn = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 3 \\ 4 \\ 3 \\ 1 \\ 0 \end{pmatrix}$

Рис. 6.4. Обчислення різниць рангів

Обчислимо значення вибіркового коефіцієнта рангової кореляції за Кедаллом та перевіримо гіпотезу про рівень значущості. Для цього за заданим значенням α з таблиці функції Лапласа знайдемо значення $z_в$ та обчислимо критичне значення критерію Tk (рис. 6.5).

$$\tau := \frac{2\cdot\left[\left[\sum_{i=0}^{n-2}\left(Rp_i - Rn_i\right)\right]\right]}{n\cdot(n-1)} = 0.467 \qquad \frac{(1-\alpha)}{2} = 0.475$$

$$zk := 1.96 \qquad Tk := zk\cdot\sqrt{\frac{2\cdot(2\cdot n+5)}{9\cdot n\cdot(n-1)}} = 0.487$$

Рис 6.5. Обчислення вибіркового коефіцієнта рангової кореляції за Кедаллом

З отриманих даних видно, що зв’язок між оцінками існує, але не суттєвий.

Варіанти для самостійного виконання

Варіант 1

A	95	90	86	84	75	70	62	60	57	50
B	92	93	83	80	55	60	45	72	62	70

Варіант 2

A	98	88	80	76	70	63	61	60	58	56	51
B	91	93	74	78	65	64	66	52	53	48	62

Варіант 3

A	1	2	3	4	5	6	7	8	9	10	11	12	13
B	6	3	4	2	1	10	7	8	9	5	11	13	12

Варіант 4

A	1	11	3	4	5	7	6	8	9	10	2
B	1	10	3	5	4	8	9	11	6	7	2

Варіант 5

A	1	2	3	4	5	6	7	8	9	10	11
B	3	10	7	2	8	5	6	11	9	1	4

Варіант 6

A	1	2	3	4	5	6	7	8	9	10	11
B	1	2	4	3	6	5	7	11	10	9	8

Варіант 7

A	1	2	3	4	5	6	7	8	9	10	11	12
B	4	3	6	2	1	7	10	8	9	5	12	11

Варіант 8

A	1	2	3	4	5	6	7	8	9	10	11	12
B	1	3	2	6	4	7	5	8	9	12	12	11

Варіант 9

A	1	2	3	4	5	6	7	8	9	10	11	12
B	220	10	1000	606	780	790	900	544		1070	1500	11

Варіант 10

A	20	10	50	40	30	60	70	80	90	100	110	120
B	220	10	1000	606	780	790	900	544		1070	1500	1100

Практичне заняття № 7.
Перевірка статистичних гіпотез для параметрів нормального закону розподілу

За даними вибірок ***X*, *У*** сукупностей, що мають нормальний закон розподілу перевірити статистичні гіпотези:

1. Про рівність їх дисперсій та математичних сподівань за заданого рівня значущості $\alpha = 0{,}05$ та конкуруючої гіпотези $H_1 : x \neq y$, де x, y – відповідні оцінки першої та другої вибірки.

2. Для першої вибірки перевірити рівність вибіркового математичного сподівання гіпотетичному, для другої вибірки перевірити рівність дисперсії гіпотетичній при $\alpha = 0{,}1$ та конкуруючій гіпотезі $H_1 : x \neq x_0$, де x – відповідні оцінки першої та другої вибірки, а x_0 – відповідні гіпотетичні значення. В якості гіпотетичних значень взяти відповідні статистичні оцінки, збільшені на 10%.

Теоретичні відомості

Статистичною називають гіпотезу про вид невідомого розподілу, або про параметри відомих розподілів.

Нульовою (основною) називають висунуту гіпотезу H_0.

Конкуруючою (альтернативною) називають гіпотезу H_1, яка суперечить нульовий.

Статистичним критерієм Kr називають випадкову величину, яка застосовується для перевірки гіпотези.

Критичною областю називають сукупність значень критерію, за яких гіпотеза відхиляється. Розрізняють лівосторонню, правосторонню та двосторонню критичні області (рис. 7.1).

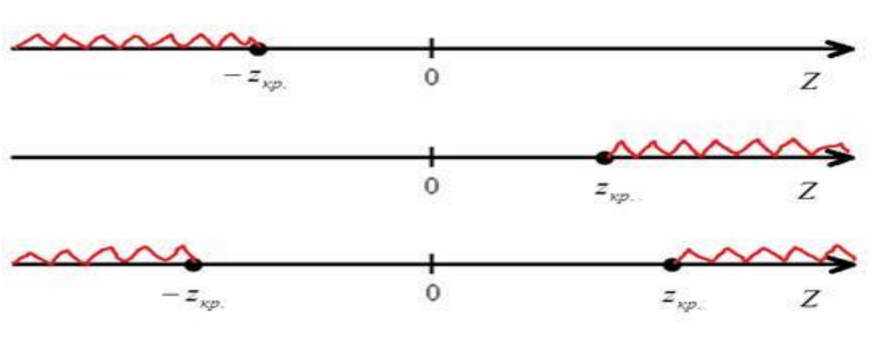


Рис. 7.1. Критичні області

Критичними точками називають точки, які відокремлюють критичну область від області прийняття гіпотези.

Якщо основна гіпотеза полягає у рівності двох параметрів, наприклад, $H_0 : \bar{x} = \bar{y}$, то можуть існувати наступні варіанти конкуруючих гіпотез: $H_1 : \bar{x} < \bar{y}$, $H_1 : \bar{x} > \bar{y}$ чи $H_1 : \bar{x} \neq \bar{y}$. У першому випадку будується лівостороння критична область, у другому – правостороння та у третьому – двостороння.

Гіпотеза про рівність генеральних середніх двох розподілів.

Постановка задачі. З двох генеральних сукупностей вилучено вибірки, об'ємом N_1 та N_2, знайдено їх вибіркові середні: $\bar{x}$ і $\bar{y}$ відповідно. Потрібно на рівні значущості α перевірити гіпотезу $H_0 : \bar{x} = \bar{y}$ про рівність генеральних середніх проти однієї з наступних конкуруючих гіпотез:

Якщо значення критерію Kr, що спостерігається, не попадає у критичну область, то гіпотеза H_0 приймається:

– $|Kr_{сп}| < kr_{кр}$ для $H_1 : \bar{x} \neq \bar{y}$;

– $Kr_{сп} < kr_{кр}$ для $H_1 : \bar{x} > \bar{y}$;

– $Kr_{сп} > -kr_{кр}$ для $H_1 : \bar{x} < \bar{y}$.

При цьому можливі наступні варіанти задачі.

1. Вибірки незалежні, генеральні сукупності розподілені нормально та відомі їх дисперсії σ_X, σ_Y. Тоді для перевірки нульової гіпотези використовують статистичний критерій:

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_X^2 / N_1 + \sigma_Y^2 / N_2}}.$$

Критичне значення критерію $z_{кр}$ шукається із співвідношення $\Phi(z_{кр}) = \frac{1-2\alpha}{2}$ для односторонньої і $\Phi(z_{кр}) = \frac{1-\alpha}{2}$ для двосторонньої критичної області, де α – заданий рівень значущості $\Phi(\ )$ – функція Лапласа (дод. 2).

Бажано, щоб N_1, N_2 були більше 30.

2. Незалежні вибірки досить великі $(N_1, N_2 > 30)$, генеральні дисперсії невідомі, причому генеральні сукупності можуть мати й інший розподіл (не нормальний). В цьому випадку застосовують схожий, але наближений критерій:

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{D_в(X)/ N_1 + D_в(Y)/ N_2}},$$

де $D_в(X)$, $D_в(Y)$ – відповідні вибіркові дисперсії. Критичне значення критерію $z_{кр}$ шукається аналогічно до випадку 1.

3. Незалежні вибірки малі, генеральні сукупності розподілені нормально та дисперсії їх не відомі. Для перевірки гіпотези застосовується критерій Стьюдента:

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{(N_1 - 1)s_x^2 + (N_2 - 1)s_Y^2}} \cdot \sqrt{\frac{N_1 \cdot N_2 (N_1 + N_2 - 2)}{N_1 + N_2}}.$$

Критичне значення за заданим рівнем значущості α та кількістю степенів свободи $k = N_1 + N_2 - 2$ знаходиться з таблиці критичних точок розподілу Стьюдента (дод. 6).

Гіпотеза про рівність середнього вибіркового з гіпотетичним середнім генеральної сукупності.

Постановка задачі. З генеральної сукупності вилучена вибірка X з нормальним законом розподілу, об'ємом N. За заданого рівня значущості α необхідно перевірити гіпотезу: $H_0 : a = a_0$, де a – генеральне середнє сукупності, a_0 – гіпотетичне середнє.

1. Дисперсія σ^2 відома. Обчислюється значення критерію, що спостерігається: $U_{сп} = \frac{(\bar{x} - a_0)\sqrt{N}}{\sigma}$. Можливі наступні варіанти конкуруючої гіпотези:

– $H_1 : a \neq a_0$. з таблиці функції Лапласа знаходиться критичне значення $\Phi(u_{кр}) = (1-\alpha)/2$ (дод. 2). Якщо $|U_{сп}| < u_{кр}$, то гіпотеза H_0 приймається;

– $H_1 : a > a_0$. з таблиці функції Лапласа знаходиться критичне значення $\Phi(u_{кр}) = (1-2\alpha)/2$;

– $H_1 : a < a_0$. з таблиці функції Лапласа знаходиться критичне значення $\Phi(u_{кр}) = (1-2\alpha)/2$.

2. Дисперсія σ^2 невідома (наприклад, у випадку малих вибірок). В якості критерію перевірки гіпотези H_0 застосовують критерій Стьюдента:

$$T_{сп} = \frac{(\bar{x} - a_0)\sqrt{N}}{s}, \text{ де } s = \sqrt{\frac{\sum_i n_i x_i^2 - \left[\sum_i n_i x_i\right]^2 / N}{N-1}}$$

з $k = N-1$ свободи.

Можливі наступні варіанти конкуруючої гіпотези:

– $H_1 : a \neq a_0$. $t_{кр}(\alpha, k)$знаходять з таблиці критичних точок таблиці Стьюдента (дод. 6);

– $H_1 : a > a_0$. $t_{кр\ правост}(\alpha, k)$ – знаходиться у нижньому рядку таблиці критичних точок таблиці Стьюдента (дод. 6).

– $H_1 : a < a_0$. Знаходять $t_{кр\ правост}(\alpha, k)$ (див. випадок $H_1 : a > a_0$).

Гіпотеза про рівність генеральних дисперсій двох нормальних розподілів.

Постановка задачі. Із двох нормальних генеральних сукупностей вилучено незалежні вибірки X та Y об'ємом N_1 та N_2 та знайдено їх виправлені дисперсії: s_X^2 і s_Y^2 відповідно. Виникає питання: значущо чи незначно вони відрізняються? Для цього перевіряється гіпотеза $H_0 : D(X) = D(Y)$. Якщо вона буде прийнята, то різницю між вибірковими значеннями можна пояснити випадковими факторами.

Для перевірки гіпотези застосовують критерій Фішера-Снедекора:

$$F_{сп} = s_Б^2 / s_М^2 ,$$

відношення більшого значення виправленої дисперсії до меншого. Критичне значення критерію $F_{кр}(\alpha, k_1, k_2)$ знаходиться з таблиці критичних точок розподілу Фішера-Снедекора (дод. 7), де k_1 – кількість степенів свободи вибірки з більшою дисперсією, а k_1 – з меншою. Для двосторонньої критичної області $F_{кр}(\alpha/2, k_1, k_2)$.

Гіпотеза про рівність вибіркової дисперсії та гіпотетичної генеральної дисперсії нормальної сукупності.

Постановка задачі. За результатами вибірки обсягу N знайдено виправлену вибіркову дисперсію s^2 і виникає питання: вона значущо відрізняється від гіпотетичної (заданої) дисперсії σ_0^2: $H_0 : s^2 = \sigma_0^2$.

Для перевірки гіпотези застосовується критерій: $\chi_{сп}^2 = \dfrac{(N-1)s^2}{\sigma_0^2}$.

Критичне значення знаходиться з таблиці дод. 5 за заданим рівнем значущості α та степенями свободи $k = N - 1$:

– $\chi_{кр}^2(\alpha, k)$ для $H_0 : s^2 > \sigma_0^2$;

– $\chi_{кр}^2(1 - \alpha/2, k)$ для $H_0 : s^2 \neq \sigma_0^2$;

– $\chi_{кр}^2(1 - \alpha, k)$ для $H_0 : s^2 < \sigma_0^2$.

Приклад виконання роботи

Нехай з двох генеральних сукупностей з нормальним законом розподілу вилучено дві вибірки:

X	41	49	56		64	65	72	73	77	82	87
У	39	53	57		64	70	74	81	89		

Необхідно перевірити статистичні гіпотези про середні значення та дисперсії.

Розв'язок. Для спрощення розрахунків застосуємо програмне середовище *Mathcad*. Запишемо дані у вигляді векторів стовпчиків. За допомогою функції $last()$, яка повертає останній номер елемента у векторі, визначимо $n1 = N_1 - 1$, $n2 = N_2 - 1$, оскільки нумерація елементів починається з 0 (рис. 7.2).

$$x := \begin{pmatrix} 41 \\ 49 \\ 56 \\ 64 \\ 65 \\ 72 \\ 73 \\ 77 \\ 82 \\ 87 \end{pmatrix} \qquad y := \begin{pmatrix} 39 \\ 53 \\ 57 \\ 64 \\ 70 \\ 74 \\ 81 \\ 89 \end{pmatrix} \qquad \begin{array}{l} n1 := \mathrm{last}(x) = 9 \\ n2 := \mathrm{last}(y) = 7 \end{array}$$

Рис. 7.2. Введення початкових даних

Обчислимо середні вибіркові та виправлені дисперсії (рис. 7.3).

$$xv := \frac{\sum_{i=0}^{n1} x_i}{n1 + 1} = 66.6 \qquad yv := \frac{\sum_{i=0}^{n2} y_i}{n2 + 1} = 65.875$$

$$s1 := \frac{\sum_{i=0}^{n1} \left(x_i - xv\right)^2}{n1} = 213.156 \qquad s2 := \frac{\sum_{i=0}^{n2} \left(y_i - yv\right)^2}{n2} = 259.554$$

Рис. 7.3. Середні вибіркові та виправлені дисперсії

1. Перевіримо гіпотезу про рівність середніх: $H_0 : \bar{x} = \bar{y}$, за конкуруючої гіпотези: $H_1 : \bar{x} \neq \bar{y}$. Оскільки задані вибірки малі і дисперсія невідома, для перевірки гіпотези застосуємо критерій Стьюдента, де $N1 = n1 + 1$, $N2 = n2 + 1$ (рис. 7.4).

$$T := \frac{xv - yv}{\sqrt{n1 \cdot s1 + n2 \cdot s2}} \cdot \sqrt{\frac{N1 \cdot N2 \cdot (N1 + N2 - 2)}{N1 + N2}} = 0.1$$

Рис. 7.4. Обчислення значення критерію Стьюдента, що спостерігається

З таблиці дод. 6 за рівнем значущості $\alpha = 0{,}05$ та кількістю степенів свободи $k = n1 + n2 = 16$ знайдемо критичне значення критерію: $t_{кр}(\alpha, k) = 2{,}12$. Оскільки $|T| < t_{кр}$, то гіпотеза H_0 про рівність середніх приймається.

Перевіримо гіпотезу про рівність дисперсій: $H_0 : D(X) = D(y)$, за конкуруючої гіпотези: $H_1 : D(X) \neq D(Y)$. Для перевірки гіпотези застосуємо критерій Фішера-Снедекора: $F_{сп} = s2 / s1 = 1{,}218$, де у чисельнику буде виправлена дисперсія вибірки Y, бо вона є більшою. За рівнем значущості $\alpha = 0{,}05$ та кількістю степенів свободи $k_1 = 7$ (відповідає вибірці з більшою дисперсією), $k_1 = 9$ (відповідає вибірці з меншою дисперсією) з таблиці дод. 7 знайдемо критичне значення критерію: $F_{кр}(\alpha, k1, k2) = 3{,}29$. Оскільки $|F_{сп}| < F_{кр}$, то гіпотеза H_0 про рівність дисперсій приймається.

2. Перевіримо гіпотезу про рівність середнього вибіркового X гіпотетичному середньому генеральному: $H_0 : \bar{x} = a_0$, за конкуруючої гіпотези $H_1 : \bar{x} \neq a_0$. В якості гіпотетичного значення візьмемо середнє вибіркове, збільшене на 10%: $a_0 = \bar{x} + 0{,}1\bar{x} = 73{,}26$. Оскільки дисперсія невідома і вибірки малі, то для перевірки гіпотези застосуємо критерій Стьюдента:

$$T_{сп} = \frac{(xv - a_0)\sqrt{N}}{s1} = -0{,}01.$$

З таблиці дод. 6 знайдемо критичне значення критерію: $t_{кр}(\alpha, k) = 1{,}83$. Оскільки $|T| < t_{кр}$, то немає підстав відхиляти основну гіпотезу.

Перевіримо гіпотезу про рівність вибіркової дисперсії та гіпотетичної генеральної дисперсії: $H_0 : s^2 = \sigma_0^2$ за $H_1 : s^2 \neq \sigma_0^2$. Обчислимо значення гіпотетичної дисперсії: $\sigma_0^2 = s2 + 0{,}1s2 = 285{,}51$. Для цього застосуємо χ^2-критерій: $\chi^2_{сп} = \dfrac{n2 \cdot s2}{\sigma_0^2} = 6{,}36$.

З таблиці дод. 5 за заданим рівнем значущості $\alpha = 0{,}1$ та кількістю степенів свободи $n2$ знайдемо критичне значення критерію: $\chi^2_{кр}(\alpha, n2) = 18{,}5$. Оскільки $|\chi^2_{сп}| < \chi^2_{кр}$, то немає підстав відхиляти основну гіпотезу.

Варіанти для самостійного виконання

Варіант 1

X	41	49	56	64	65	72	73	77	82	87
У	39	53	57	64	70	74	81	89		

Варіант 2

X	12	14	13	16	11	9	13	15	15	18	14
У	13	9	11	10	7	6	8	10	11		

Варіант 3

X	20	15	11	16	13	16	19	15	9	18
У	19	22	17	14	21	25	26	24	15	16

Варіант 4

X	90	29	39	79	88	53	34	40	75	79
У	41	49	56	64	72	65	63	87	77	62

Варіант 5

X	90	29	39	79	88	53	34	40	75	79
У	41	49	56	64	72	65	63	87	77	62

Варіант 6

X	2	2	2	2	2	3	3	3	3	3	2	2	3	3	3
У	2	3	3	4	3	2	3	4	4	3	4	3	2	4	4

Варіант 7

X	1,08	1,10	1,12	1,14	1,15	1,25	1,36	1,38	1,40	1,42
У	1,11	1,12	1,18	1,22	1,33	1,35	1,36	1,38		

Варіант 8

X	12,3	12,5	12,8	13,0	13,5		***У***	12,2	12,3	13,0
n_i	1	2	4	2	1		m_i	6	8	2

Варіант 9

X	20	15	11	16	13	16	16	19	15	9
У	19	22	17	24	21	25	26	24	15	15

Варіант 10

X	90	29	39	79	88	53	34	40	75	79
У	41	49	56	64	72	65	63	87	77	62

Практичне заняття № 8
Однофакторний дисперсійний аналіз

Результати спостережень занесені у таблицю, де стовпчики – рівні фактору, рядки – номер спостереження. Визначити суттєвість впливу фактора за рівня значущості $\alpha = 0{,}05$.

Теоретичні відомості

Факторний аналіз – статистичний метод перевірки гіпотез про вплив різних чинників на досліджувану випадкову величину.

Постановка задачі. Нехай генеральні сукупності X_1, $X_2, \ldots,$ X_p, які розподілені нормально і мають однакову, хоча і невідому дисперсію. Математичні сподівання також невідомі, але можуть бути різні. Потрібно за заданого рівня значущості за вибірковими середніми перевірити нульову гіпотезу про рівність всіх математичних сподівань:

$$H_0 : M(X_1) = M(X_2) = \ldots = M(X_p).$$

Іншими словами, потрібно встановити, значимо чи не значимо розрізняються вибіркові середні.

Для порівняння декількох середніх на практиці використовують інший метод, який заснований на порівнянні дисперсій, і тому названий дисперсійним аналізом.

Порядок перевірки гіпотези. Нехай ознака X розподілена за нормальним законом. На неї впливає фактор F, котрий має p постійних рівнів. Будемо припускати, що число спостережень на кожному рівні постійне і дорівнює q. Таким чином, спостерігається $n = p \cdot q$ значень фактора, яким відповідають значення ознаки x_{ij}, де i – номер випробування, $i = \overline{1, q}$, j – номер рівня фактора, $j = \overline{1, p}$.

Результати спостережень зведені в таблицю (табл. 8.1).

Введемо в розгляд величини:

– $S_{заг} = \sum_{j=1}^{p} \sum_{i=1}^{q} \left(x_{ij} - \bar{x} \right)^2$ – загальну суму квадратів відхилень спостережуваних значень від загальної середньої $\bar{x}$;

– $S_{факт} = q\sum_{j=1}^{p}\left(\overline{x}_{грj} - \overline{x}\right)^2$ – факторну суму квадратів відхилень групових середніх від загальної середньої $\overline{x}$. Вона характеризує розсіювання «між групами»;

– $S_{зал} = \sum_{i=1}^{q}\left(x_{i1} - \overline{x}_{гр1}\right)^2 + \sum_{i=1}^{q}\left(x_{i2} - \overline{x}_{гр2}\right)^2 + ... + \sum_{i=1}^{q}\left(x_{ip} - \overline{x}_{грp}\right)^2$ – залишкову суму квадратів відхилень спостережуваних значень групи від своєї групової середньої, яка характеризує розсіювання «всередині групи».

Таблиця 8.1

Результати спостережень на різних рівнях фактора

Номер спостереження	Рівні фактора			
	F_1	F_2	…	F_p
1	x_{11}	x_{12}	…	x_{1p}
2	x_{21}	x_{22}	…	x_{2p}
….	….	….	….	….
q	x_{q1}	x_{q2}	….	x_{qp}
Групові середні	$x_{гр1}$	$x_{гр2}$	….	$x_{грp}$

Ці суми пов’язані рівністю: $S_{зал} = S_{заг} - S_{факт}$.

Розділивши суми квадратів відхилень на відповідне число ступенів свободи, отримаємо загальну, факторну і залишкову дисперсії:

$$s_{заг}^2 = \frac{S_{заг}}{pq-1},\ s_{факт}^2 = \frac{S_{факт}}{p-1},\ s_{зал}^2 = \frac{S_{зал}}{p(q-1)},$$

де p – число спостережень на кожному рівні, q – число рівнів, $pq-1$ – число степенів свободи загальної дисперсії, $p-1$ – число степенів свободи факторной дисперсії, $p(q-1)$ – число степенів свободи залишкової дисперсії.

Якщо нульова гіпотеза про рівність середніх справедлива, то всі ці дисперсії є незміщеними оцінками генеральної дисперсії. Тоді вирішення поставленої задачі зводиться до порівняння факторної і залишкової дисперсій за критерієм Фішера: $F_{кр} = s_{факт} / s_{зал}$.

Якщо нульова гіпотеза про рівність декількох середніх (будемо називати їх груповими) справедлива, то, очевидно, критерій вкаже, що нульову гіпотезу про рівність факторної і залишкової дисперсій немає підстав відкинути, а отже вплив фактора несуттєвий. Якщо гіпотеза про рівність групових середніх помилкова, то вплив фактора значний.

Зауваження. Якщо кількість спостережень на рівнях фактора є неоднаковою: q_1 на F_1, q_2 на F_2 і т.д., то факторну суму квадратів відхилень знаходять за формулою:

$$S_{факт} = \frac{T_1^2}{q_1} + \frac{T_2^2}{q_2} + \ldots + \frac{T_p^2}{q_p}, \text{ де } T_j = \sum_{i=1}^{q_i} x_{ij} .$$

Приклад виконання роботи

В результаті спостережень впливу фактора F на випадкову величну X були отримані дані, представлені у табл. 8.2.

Таблиця 8.2

Результати спостереджень

Номер іспиту	Рівні фактора			
	F_1	F_2	F_3	F_4
1	3	7	2	5
2	9	5	6	5
3	4	5	2	8
4	2	3	5	4
5	6	9	5	8

Кількість рівнів фактора $p = 4$, кількість спостережень на кожному рівні однакова $q = 5$.

Необхідно визначити, чи суттєвим є вплив даного фактора.

Розв'язок. Для спрощення обчислень застосуємо програмне середовище *Mathcad*. Занесемо дані у таблицю $t1$. Для цього скористаємось стандартними можливостями: Вставка – Дані – Таблиця (рис. 8.1).

$t1 :=$

	0	1	2	3
0	3	7	2	5
1	9	5	6	5
2	9	5	2	8
3	2	3	5	4
4	6	9	5	8

$p := 4$

$q := 5$

Рис. 8.1. Початкові дані

З таблицею можна працювати як з матрицею. Скориставшись цим, знайдемо середні значення для кожного рівня фактора $x_{грі}$ (рис. 8.2).

$$x_1 := \frac{\sum_{i=0}^{4} t1_{i,0}}{q} = 5.8 \quad x_2 := \frac{\sum_{i=0}^{4} t1_{i,1}}{q} = 5.8 \quad x_3 := \frac{\sum_{i=0}^{4} t1_{i,2}}{q} = 4 \quad x_4 := \frac{\sum_{i=0}^{4} t1_{i,3}}{q} = 6$$

Рис. 8.2. Обчислення середніх для кожної групи

Знайдемо загальне середнє: $xv = \left(\sum_{i=1}^{p} x_i\right) / p$. Обчислимо загальну, факторну і залишкову суми квадратів (рис. 8.3).

$$Szag := \sum_{i=0}^{p-1} \sum_{j=0}^{q-1} \left(t1_{j,i} - xv\right)^2 = 104.8 \quad Sf := q \cdot \sum_{i=1}^{p} \left(x_i - xv\right)^2 = 13.2 \quad Szal := Szag - Sf = 91.6$$

Рис. 8.3. Суми квадратів

Зробимо поправки на кількість степенів свободи і знайдемо факторну $s2f$ та залишкову s2z виправлені дисперсії (рис. 8.4)

$$s2f := \frac{Sf}{p-1} = 4.4 \qquad s2z := \frac{Szal}{p \cdot (q-1)} = 5.725$$

Рис.8.4. Факторна та залишкова дисперсії

В нашому прикладі залишкова дисперсія більша, ніж факторна, тому зрозуміло, що вплив фактора несуттєвий і можна не виконувати перевірку гіпотези на основі критерію Фішера.

Варіанти для самостійного виконання

Варіант 1

Номер іспиту	Рівні фактора			
	F_1	F_2	F_3	F_4
1	3	7	2	5
2	9	5	6	5
3	4	5	2	8
4	2	3	5	4
5	6	9	5	8

Варіант 2

Номер іспиту	Рівні фактора			
	F_1	F_2	F_3	F_4
1	36	56	52	39
2	47	61	57	57
3	50	64	59	63
4	58	66	58	61
5	67	66	79	65

Варіант 3

Номер іспиту	Рівні фактора				
	F_1	F_2	F_3	F_4	F_5
1	100	92	74	68	64
2	101	102	87	80	83
3	126	104	88	83	83
4	133	115	93	87	84
5	141	119	94	96	90

Варіант 4

Номер іспиту	Рівні фактора			
	F_1	F_2	F_3	F_4
1	51	52	56	54
2	59	58	56	58
3	53	66	58	62
4	59	69	58	64
5	63	70	70	66
6	69	72	74	67

Варіант 5

Номер іспиту	Рівні фактора		
	F_1	F_2	F_3
1	8	7	4
2	7	8	5
3	9	5	3
4	5	4	6
5	6	6	2
6	8	7	4

Варіант 6

Номер іспиту	Рівні фактора		
	F_1	F_2	F_3
1	10	8	5
2	9	8	7
3	10	7	6
4	8	9	4
5	9	6	5
6	8	4	2
7	9	5	3

Варіант 6 (продовження)

8	7	5	2
9	7	6	1
10	6	4	2

Варіант 7

Номер іспиту	Рівні фактора		
	F_1	F_2	F_3
1	3,13	1,39	5,47
2	3,25	5,38	5,6
3	3,64	4,07	6,88
4	3,4	3,87	6,4
5	2,59	4,37	3,02
6	1,97	3,79	6,18
7	3,16	3,33	5,52
8	4,22	5,39	4,15
9	1,36	3,37	2,07
10	3,47	4,74	4,68

Варіант 8

Номер іспиту	Рівні фактора			
	F_1	F_2	F_3	F_4
1	1,38	1,41	1,32	1,31
2	1,38	1,42	1,33	1,33
3	1,42	1,44	1,34	1,32
4	1,42	1,45		
5	1,43	1,45		

Варіант 9

Номер іспиту	Рівні фактора		
	F_1	F_2	F_3
1	7	9	65
2	5	8	7
3	6	10	6
4	4	8	6
5	6	7	9
6	7	10	5
7	8	10	76
8	6	9	8
9	5	7	7
10	7	6	8

Варіант 10

Номер іспиту	Рівні фактора			
	F_1	F_2	F_3	F_4
1	1600	1580	1460	1510
2	1610	1640	1550	1520
3	1650	1640	1600	1530
4	1680	1700	1620	1570
5	1700	1750	1640	1600
6	1700		1660	1680
7	1800		1740	
8			1820	

Практичне завдання № 9.
Статистичні оцінки випадкових процесів

Задано випадковий процес (процеси). Оцінити основні статистичні характеристики (завдання 1 – 6). Для оцінки параметрів випадкових величин див. дод. 1.

Теоретичні відомості

Визначення 1. *Математичним сподіванням* випадкової функції (ВФ) $X(t)$ називається невипадкова функція $m_X(t)$, яка для кожного значення аргументу *t* дорівнює математичному сподіванню відповідного розрізу випадкової функції, тобто $m_X(t) = M[X(t)]$.

Властивості математичного сподівання:

1. Математичне сподівання невипадкової функції, $\varphi(t)$ дорівнює самій функції, тобто $M[\varphi(t)] = \varphi(t)$.

2. Невипадковий множник $\varphi(t)$ можна виносити за знак математичного сподівання: $M[\varphi(t)\cdot X(t)] = \varphi(t)\cdot M[X(t)] = \varphi(t)\cdot m_X(t)$.

3. Математичне сподівання суми двох випадкових функцій дорівнює сумі математичних сподівань доданків:

$$M[X(t)+Y(t)] = M[X(t)] + M[Y(t)] = m_X(t) + m_Y(t).$$

Визначення 2. *Дисперсією* випадкової функції X(t) називається невипадкова функція $D_X(t)$, значення якої для кожного аргументу *t* дорівнює дисперсії відповідного розрізу ВФ: $D_X(t) = D[X(t)]$, середнє квадратичне відхилення ВФ: $\sigma_X(t) = \sqrt{D_X(t)}$

Властивості дисперсії:

1. Дисперсія невипадкової функції дорівнює 0: $D[\varphi(t)] = 0$.

2. Дисперсія суми випадкової функції $X(t)$ і невипадкової функції $\varphi(t)$ дорівнює дисперсії випадкової функції: $D[X(t)+\varphi(t)] = D_X(t)$.

3. Дисперсія добутку випадкової функції $X(t)$ і невипадкової функції $\varphi(t)$ дорівнює добутку квадрату невипадкового множника і дисперсії випадкової функції, тобто $D[X(t)\cdot\varphi(t)] = \varphi^2(t)\cdot D_X(t)$.

Визначення 3. Кореляційною функцією ВФ $X(t)$ називається невипадкова функція двох аргументів $K_X(t_1,t_2)$, яка для кожної пари значень t_1 й t_2 дорівнює кореляційному моменту відповідних розрізів:

$$K_X(t,t')=M\left[\overset{\circ}{X}(t_1)\cdot\overset{\circ}{X}(t_2)\right],\text{ де }\overset{\circ}{X}(t)=X(t)-m_X(t).$$

Якщо покласти $t=t'$, то $K_X(t,t)=D_X(t)$

Властивості:

1. Від додавання невипадкового доданка кореляційна функція ВФ не змінюється.

2. Під час множення ВФ на невипадковий множник $\varphi(t)$ її кореляційна функція множиться на $\varphi(t_1)\cdot\varphi(t_2)$.

3. Кореляційна функція $K_X(t_1,t_2)$ симетрична відносно своїх аргументів: $K_X(t_1,t_2)=K_X(t_2,t_1)$.

4. Абсолютна величина кореляційної функції не перевищує середнього геометричного дисперсій відповідних розрізів, а саме: $K_X(t_1,t_2)\le D_X(t_1)\cdot D_X(t_2)$.

Нормована кореляційна функція ВФ є коефіцієнтом кореляції відповідних розрізів випадкової функції:

$$r_X(t_1,t_2)=\frac{K_X(t_1,t_2)}{\sigma_X(t_1)\cdot\sigma_X(t_2)},\ |r_X(t_1,t_2)|\le 1.$$

Визначення 4. *Взаємна кореляційна функція* визначає залежність між декількома випадковими процесами.

Взаємною кореляційною функцією двох випадкових функцій $X(t)$ і $Y(t)$ називають невипадкову функцію $R_{XY}(t_1,t_2)$ двох незалежних аргументів t_1 та t_2, значення якої за кожної пари фіксованих значень аргументів дорівнює кореляційному моменту перерізів цих функцій, що відповідають цим значенням аргументів:

$$R_{xy}(t_1,t_2)=M\left[\hat{X}(t_1)\hat{Y}(t_2)\right]=R_{yx}(t_2,t_1).$$

1. Некорельованими називають дві випадкові функції, якщо їх взаємна кореляційна функція тотожно дорівнює нулю.

2. Додавання до випадкових функції $X(t)$ і $Y(t)$ невипадкових доданків, відповідно $\varphi(t)$ та $\psi(t)$, не змінює їх взаємної кореляційної функції. Тобто, якщо $X1(t)=X(t)+\varphi(t)$ і $Y1(t)=Y(t)+\psi(t)$, то

$$R_{X1Y1}(t_1,t_2)=R_{XY}(t_1,t_2).$$

3. У процесі множення випадкових функції $X(t)$ і $Y(t)$ на невипадкові множники, відповідно $\varphi(t)$ та $\psi(t)$, їх взаємна кореляційна функція множиться на добуток $\varphi(t)\cdot\psi(t)$. Тобто, якщо $X1(t)=X(t)\cdot\varphi(t)$ і

$Y1(t)=Y(t)\cdot\psi(t)$, то $R_{X1Y1}(t_1,t_2)=R_{XY}(t_1,t_2)\varphi(t_1)\psi(t_2)$

4. Абсолютна величина взаємної кореляційної функції двох випадкових функцій не перевищує середнього геометричного їх дисперсій: $|R_{XY}(t_1,t_2)|\le\sqrt{D_X(t_1)D_Y(t_2)}$.

Нормованою взаємною кореляційною функцією двох випадкових функцій $X(t)$ і $Y(t)$ називають невипадкову функцію двох незалежних аргументів t_1 та t_2:

$$\rho_{XY}(t_1,t_2)=\frac{R_{XY}(t_1,t_2)}{\sqrt{K_X(t_1,t_1)}\cdot\sqrt{K_Y(t_2,t_2)}}=\frac{R_{XY}(t_1,t_2)}{\sqrt{D_X(t_1)}\cdot\sqrt{D_Y(t_2)}}.$$

Кореляційна функція суми декількох випадкових функцій:

$$Z(t)=X(t)+Y(t)$$

$$K_Z(t_1,t_2)=K_X(t_1,t_2)+K_Y(t_1,t_2)+R_{XY}(t_1,t_2)+R_{XY}(t_2,t_1).$$

Якщо випадкові функції $X(t)$ і $Y(t)$ незалежні, то

$$K_Z(t_1,t_2)=K_X(t_1,t_2)+K_Y(t_1,t_2).$$

Визначення 5. Похідною випадкової функції $X(t)$ називають границю у середньому квадратичному відношення приросту випадкової функції до приросту аргументу Δt, коли $\Delta t\to 0$:

$$X'(t)=\lim_{\Delta t\to 0}\frac{X(t+\Delta t)-X(t)}{\Delta t}/$$

Математичне сподівання похідної $X'(t)$ від випадкової функції $X(t)$ дорівнює похідній від її математичного сподівання: $m_{X'}(t)=m'_X(t)$.

Кореляційна функція похідної $X'(t)$ від випадкової функції $X(t)$ дорівнює другій мішаній похідній від її кореляційної функції:

$$K_{X''}(t_1,t_2)=\frac{\partial^2 K_X(t_1,t_2)}{\partial t_1\partial t_2}.$$

Взаємна кореляційна функція випадкової функції $X(t)$ та її похідної $X'(t)$ дорівнює частинній похідній від її кореляційної функції по відповідному аргументу: $R_{XX'}(t_1,t_2)=\dfrac{\partial K_X(t_1,t_2)}{\partial t_1}$ або

$$R_{XX'}(t_1,t_2)=\frac{\partial K_X(t_1,t_2)}{\partial t_2}.$$

Визначення 6. Інтегралом від випадкової функції $X(t)$ на відрізку $[0;t]$ називають границю в середньому квадратичному інтегральної суми при прямуванні до нуля максимальної довжини часткового інтервалу Δs_i:

$$Y(t)=\lim_{\max\Delta s_i\to 0}\sum_i X(s_i)\cdot\Delta s_i=\int_0^t X(s)ds.$$

Математичне сподівання інтеграла $Y(t)=\int_0^t X(s)ds$ від випадкової функції $X(t)$ дорівнює інтегралу її математичного сподівання

$$m_Y(t)=\int_0^t m_X(s)ds.$$

Кореляційна функція інтеграла $Y(t)=\int_0^t X(s)ds$ від випадкової функції $X(t)$ дорівнює подвійному інтегралу від її кореляційної функції:

$$K_Y(t_1,t_2)=\int_0^{t_1}\int_0^{t_2}K_X(s_1,s_2)ds_1ds_2.$$

Взаємна кореляційна функція випадкової функції $X(t)$ та інтеграла $Y(t)=\int_0^t X(s)ds$ дорівнює інтегралу від кореляційної функції випадкової функції:

$$R_{XY}(t_1,t_2)=\int_0^{t_2}K(t_1,s)ds,\ R_{YX}(t_1,t_2)=\int_0^{t_1}K(s,t_2)ds$$

Визначення 7. Стаціонарною називають випадкову функцію $X(t)$, математичне сподівання якої є сталим при всіх значеннях аргументу t і кореляційна функція якої залежить тільки від різниці аргументів t_2-t_1.

З цього визначення випливає, що:

1) кореляційна функція стаціонарної випадкової функції є функцією одного аргументу $\tau=t_2-t_1$: $K_X(t_1,t_2)=k_X(t_2-t_1)=k_X(\tau)$;

2) дисперсія стаціонарної випадкової функції стала при всіх значеннях аргументу t і дорівнює значенням її кореляційної функції на початку координат ($\tau=0$): $D_X(t)=K_X(t,t)=k_X(t-t)=k_X(0)$.

Достатня умова ергодичності стаціонарної випадкової функції $X(t)$ відносно кореляційної функції полягає в тому, що кореляційна функція $k_Y(\tau)\to 0$ при $\tau\to\infty$: $\lim\limits_{\tau\to\infty} k_Y(\tau)=0$, де $Y(t,\tau)=\overset{\circ}{X}(t)\cdot \overset{\circ}{X}(t+\tau)$.

Приклад виконання роботи

Завдання 1. Знайти математичне сподівання $m_X(t)$, кореляційну функцію $K_X(t_1,t_2)$ та дисперсію $D_X(t)$ випадкового процесу $X(t)=t^2\cdot U-V\cdot ch(t)-t^2$, де U, V – некорельовані випадкові величини: $U\in E(0{,}1)$, $V\in B(20;0{,}2)$.

Розв'язок. Запишемо математичне сподівання та дисперсію випадкових величин U та V:

$$M(U)=10,\ M(V)=4,\ D(U)=100,\ D(V)=3{,}2.$$

1.1. Знайдемо математичного сподівання випадкового процесу, користуючись його властивостями:

$$m_X(t)=M[X(t)]=M\left[t^2\cdot U-V\cdot ch(t)-t^2\right]=$$

$$=t^2M(U)-ch(t)M(V)-t^2=10t^2-4ch(t)-t^2.$$

1.2. Для знаходження кореляційної функції запишемо значення:

$$\overset{\circ}{X}(t)=X(t)-m_X(t)=t^2\cdot U-V\cdot ch(t)-t^2-10t^2+4ch(t)+t^2=t^2\cdot(U-10)-ch(t)\cdot(V-4)$$

Враховуючи, що випадкові величини U, V – некорельовані (їх добуток дорівнює нулю), знайдемо кореляційну функцію:

$$K_X(t_1,t_2)=M\left[\overset{\circ}{X}(t_1)\cdot\overset{\circ}{X}(t_2)\right]=$$

$$=M\left[\left(t_1^2\cdot(U-10)-ch(t_1)\cdot(V-4)\right)\cdot\left(t_2^2\cdot(U-10)-ch(t_2)\cdot(V-4)\right)\right]=$$

$$=M\left[t_1^2\cdot t_2^2(U-10)^2+ch(t_1)\cdot ch(t_2)(V-4)^2\right]=$$

$$=t_1^2\cdot t_2^2M(U-10)^2+ch(t_1)\cdot ch(t_2)M(V-4)^2=$$

$$=t_1^2\cdot t_2^2D(U)+ch(t_1)\cdot ch(t_2)D(V)=100t_1^2\cdot t_2^2+3{,}2ch(t_1)\cdot ch(t_2)$$

Знайдемо дисперсію: $D_X(t)=K_X(t,t)=100t^4+3{,}2ch^2(t)$.

Завдання 2. Знайти кореляційну функцію $K_Z(t_1,t_2)$ та дисперсію $D_Z(t)$, випадкового процесу $Z(t)=X(t)\sin(t)-t^2Y(t)+e^t$, якщо $X(t)$, $Y(t)$ – некорельовані випадкові процеси і задані їх кореляційні функції,

$$K_X(t_1,t_2)=\cos(t_2-t_1),\ K_Y(t_1,t_2)=1/\left(t_1^2t_2^2\right).$$

Розв'язок. Оскільки $X(t)$, $Y(t)$ – некорельовані випадкові процеси, і користуючись властивостями кореляційної функції від суми випадкових процесів, запишемо:

$$K_Z(t_1,t_2)=K_X(t_1,t_2)+K_Y(t_1,t_2)=$$
$$=\sin(t_1)\sin(t_2)\cos(t_2-t_1)+\frac{t_1^2t_2^2}{t_1^2t_2^2}=\sin(t_1)\sin(t_2)\cos(t_2-t_1)+1.$$
$$D_Z(t)=\sin^2(t)+1.$$

Завдання 3. $Z(t)\ =\ t^2\ +\ g(t)\,X(t)\ -h(t)\,Y(t)$, де $g(t)=t$, $h(t)=e^{-t}$ – невипадкові функції, $X(t)$, $Y(t)$ – центровані випадкові процеси з кореляційними функціями $K_X(t_1,t_2)=4(1+t_1\cdot t_2)$, $K_Y(t_1,t_2)=9(1+t_1\cdot t_2)$ та взаємною кореляційною функцією $K_{XY}(t_1,t_2)=6(1+t_1\cdot t_2)$. Знайти математичне сподівання $m_Z(t)$ кореляційну функцію $K_Z(t_1,t_2)$, дисперсію $D_Z(t)$, нормовану кореляційну функцію $\rho_Z(t_1,t_2)$ випадкового процесу Z.

Розв'язок. Запишемо випадкову функцію, підставивши значення невипадкових множників: $Z(t)\ =\ t^2\ +\ t\cdot\ X(t)\ -e^{-t}\cdot Y(t)$. Користуючись властивостями кореляційної функції від суми випадкових процесів, запишемо:

$$K_Z(t_1,t_2)=t_1t_2\cdot K_X(t_1,t_2)+e^{-(t_1+t_2)}\cdot K_Y(t_1,t_2)+t_1e^{-t_2}\cdot R_{XY}(t_1,t_2)+t_2e^{-t_1}\cdot R_{XY}(t_2,t_1)=$$
$$=4t_1t_2\cdot(1+t_1\cdot t_2)+9e^{-(t_1+t_2)}\cdot(1+t_1\cdot t_2)+6t_1e^{-t_2}\cdot(1+t_1\cdot t_2)+6t_2e^{-t_1}\cdot(1+t_1\cdot t_2)$$

Знайдемо дисперсію:

$$D_Z(t)=4t^2(1+t^2)+9e^{-2t}(1+t^2)+12te^{-t}(1+t^2)=(1+t^2)(2t+3e^{-t})^2.$$

Знайдемо нормовану кореляційну функцію:

$$\rho_{XY}(t_1,t_2)=\frac{R_{XY}(t_1,t_2)}{\sqrt{D_X(t_1)}\cdot\sqrt{D_Y(t_2)}}=\frac{6(1+t_1\cdot t_2)}{(2t_1+3e^{-t_1})(2t_2+3e^{-t_2})\sqrt{(1+t_1^2)(1+t_2^2)}}.$$

Завдання 4. Знайти математичне сподівання $m_Y(t)$, кореляційну функцію $K_Y(t_1,t_2)$ та дисперсію $D_Y(t)$ випадкового процесу $Y(t)=X'(t)$, не виконуючи диференціювання $X(t)$, де $X(t)=-U\cdot t^3-\cos(t)$, $U\in R(-1;3)$ – випадкова величина. Знайти взаємну кореляційну функцію $K_{XY}(t_1,t_2)$.

Розв'язок. Знайдемо математичне сподівання випадкової величини U та випадкового процесу $X(t)$: $M(U)=1$, $m_X(t)=-t^3-\cos(t)$. Математичне сподівання від похідної: $m_Y(t)=m'_X(t)=-3t^2+\sin(t)$.

Знайдемо дисперсію випадкової величини U та випадкового процесу $X(t)$: $D(U)=4/3$, $K_X(t_1,t_2)=\frac{4}{3}t_1^3t_2^3$. Кореляційну функцію від похідної:

$$K_Y(t_1,t_2)=\frac{\partial^2 K_X(t_1,t_2)}{\partial t_1\partial t_2}=12(t_1t_2)^2 .$$

Знайдемо взаємну кореляційну функцію

$$K_{XY}(t_1,t_2)=\frac{\partial K_X(t_1,t_2)}{\partial t_2}=4t_1^3t_2^2 .$$

Завдання 5. $X(t)=f(t)U$, $f(t)=\sin(2t)$ – невипадкова функція $U\in N(-1;3)$ – випадкова величина, $Z(t)=\int_0^t X(s)ds$. Знайти математичне сподівання $m_Z(t)$, кореляційну функцію $K_Z(t_1,t_2)$, дисперсію $D_Z(t)$, взаємні кореляційні функції $K_{ZX}(t_1,t_2)$, $K_{XZ}(t_1,t_2)$, не інтегруючи $X(t)$.

Розв'язок. Знайдемо математичне сподівання та дисперсію випадкової величини U: $M(U)=-1$, $D(U)=6$. Математичне сподівання, кореляційна функція та дисперсія випадкового процесу $X(t)$ $m_X(t)=-\sin(2t)$, $K_X(t_1,t_2)=6\sin(2t_1)\sin(2t_2)$, $D_X(t)=6\sin^2(2t)$.

Знайдемо відповідні характеристики для випадкового процесу

$$Z(t)=\int_0^t X(s)ds:\ m_Z(t)=-\int_0^t\sin(2s)ds=\frac{\cos(2s)}{2}\Big|_0^t=\frac{1}{2}(\cos(2t)-1),$$

$$K_Z(t_1,t_2)=\int_0^{t_1}\int_0^{t_2}6\sin(2s_1)\sin(2s_2)ds_1ds_2=6\sin^2(t_1)\sin^2(t2),$$

$$D_Z(t)=6\sin^4(t).$$

Знайдемо взаємні кореляційні функції випадкового процесу та його інтегралу:

$$K_{ZX}(t_1,t_2)=\int_0^{t1}6\sin(2s)\sin(2t_2)ds=2\sin^2(t1)\sin(2t_2),$$

$$K_{XZ}(t_1,t_2)=\int_0^{t2}6\sin(2t_1)\sin(2s)ds=2\sin(2t1)\sin^2(2t_2).$$

Завдання 6. Довести, що випадковий процес $X(t)=(U-1)\cos(20t)-V\sin(20t)$ є стаціонарним у широкому сенсі. Перевірити властивість ергодичності для математичного сподівання та кореляційної функції. Знайти дисперсію випадкового процесу, $U \in E(1)$, $V \in N(0{,}1)$ − некорельовані випадкові величини.

Розв'язок. Знайдемо математичне сподівання та дисперсію випадкових величин U, V: $M(U)=1$, $D(U)=1$, $M(V)=0$, $D(V)=1$.

Математичне сподівання: $m_X(t) = \cos(20t) - \cos(20t) = 0$.

Оскільки математичне сподівання дорівнює 0, то процес є стаціонарним у вузькому сенсі. Обчислимо кореляційну функцію:

$$K_X(t_1,t_2) = M\left[\left((U-1)\cos(20t_1) - V\sin(20t_1)\right)\left((U-1)\cos(20t_2) - V\sin(20t_2)\right)\right] =$$
$$= \cos(20t_1)\cos(20t_2)M\left[(U-1)^2\right] + \cos(20t_1)\cos(20t_2)M\left[(V-0)^2\right] =$$
$$= 2\cos(20t_1)\cos(20t_2) = \left(\cos(20(t1-t2)) + \cos(20(t1+t2))\right)$$

Процес не є ергодичним, оскільки кореляційна функція залежить від 2-х аргументів, а не від $\tau = t_2 - t_1$.

Дисперсія випадкового процесу: $D_X(t) = 1 + \cos(40t)$.

Варіанти для самостійного виконання

Завдання 1. Знайти математичне сподівання $m_X(t)$, кореляційну функцію $K_X(t_1,t_2)$ та дисперсію $D_X(t)$ випадкового процесу $X(t)$, U, V – некорельовані випадкові величини.

№ варіанта	Завдання
1	$X(t) = t^2 \cdot U + V \cdot \cos(t) - \sin(t)$. U ∈ N(3; 2), V ∈ E(0.5).
2	$X(t) = t \cdot U - 3e^{-3t}V + \cos(t)$. U ∈ R(0; 6), V∈ B(10; 0.5).
3	$X(t) = e^t \cdot U - V \cdot ch(t) + 3$. U ∈ P(0.2), V ∈ R(–2; 2).
4	$X(t) = U \cdot \sin(t) - t \cdot V + t^5$. U ∈ N(1; 2), V ∈ P(2).
5	$X(t) = t^3 \cdot U + V \cdot \cos(t) - 2$. U ∈ R(–1; 3), V ∈ E(0.4).
6	$X(t) = 3 \cdot U \cdot sh(t) - e^{3t} \cdot V + \cos(t)$. U ∈ E(0.25), V ∈ R(2; 4).
7	$X(t) = U \cdot \sin(2t) - 4t \cdot V + 3$. U ∈ B(10; 0.3), V ∈ P(3).
8	$X(t) = U \cdot \cos(3t) - V \cdot \sin(t) - t$. U ∈ R(–3; 1), V ∈ N(–1; 0.5).
9	$X(t) = t^2 \cdot U - V \cdot ch(t) - t^2$. U ∈ E(0.1), V ∈ B(20; 0.2).
10	$X(t) = e^t \cdot U - V \cdot \sin(t) + t$. U ∈ N(–2; 2), V ∈ E(4).

Завдання 2. Знайти кореляційну функцію $K_Z(t_1,t_2)$ та дисперсію $D_Z(t)$, якщо $X(t)$, $Y(t)$ – некорельовані випадкові процеси і задані їх кореляційні функції $K_X(t_1,t_2)$, $K_Y(t_1,t_2)$.

№ варіанта	Завдання
1	$Z(t)=X(t)\sin(t)-Y(t)(t^2+1)+e^t$, $K_X(t_1,t_2)=1/(1+\lvert t_2-t_1\rvert)$, $K_Y(t_1,t_2)=t_1t_2+1$.
2	$Z(t)=X(t)e^t-Y(t)\cos(t)+e^{2t}$, $K_X(t_1,t_2)=K_Y(t_1,t_2)=1/(1+(t_2-t)^2)$.
3	$Z(t)=X(t)t-Y(t)\cos(t)+\sin(t)$, $K_X(t_1,t_2)=1/(1+\lvert t_2-t_1\rvert)$, $K_Y(t_1,t_2)=\cos(t_2-t_1)$.
4	$Z(t)=X(t)\sin(t)-t^2Y(t)+e^t$, $K_X(t_1,t_2)=\cos(t_2-t_1)$, $K_Y(t_1,t_2)=1/(t_1^2t_2^2)$.
5	$Z(t)=X(t)\cos(3t)-Y(t)(t+3)+\sin(t)$, $K_X(t_1,t_2)=K_Y(t_1,t_2)=\exp(-\lvert t_2-t_1\rvert)$.
6	$Z(t)=X(t)e^{-3t}-Y(t)\sin(t)-t$, $K_X(t_1,t_2)=1+\cos(t_2-t_1)$, $K_Y(t_1,t_2)=\sin(t_2)\cdot\sin(t_1)$.
7	$Z(t)=X(t)t-Y(t)e^{2t}+sh(t)$, $K_X(t_1,t_2)=K_Y(t_1,t_2)=\exp(-2\lvert t_2-t_1\rvert)$.
8	$Z(t)=X(t)\cos(t)-(3t^2+1)Y(t)+\sin(t)$, $K_X(t_1,t_2)=t_1^2t_2^2$, $K_Y(t_1,t_2)=\cos(t_2-t_1)$.
9	$Z(t)=X(t)ch(t)-3tY(t)+e^t$, $K_X(t_1,t_2)=2+\cos(t_2-t_1)$, $K_Y(t_1,t_2)=t_1\cdot t_2+1$.
10	$Z(t)=X(t)t^4-Y(t)ch(t)+ch(t)$, $K_X(t_1,t_2)=t_1\cdot t_2+2$, $K_Y(t_1,t_2)=\exp(-4\lvert t_2-t_1\rvert)$.

Завдання 3. $Z(t)=t^2+g(t)X(t)-h(t)Y(t)$, де $g(t)$, $h(t)$ – невипадкові функції, $X(t)$, $Y(t)$ – центровані випадкові процеси з кореляційними функціями $K_X(t_1,t_2)$, $K_Y(t_1,t_2)$ та взаємною кореляційною функцією $K_{XY}(t_1,t_2)$. Знайти математичне сподівання $m_Z(t)$, кореляційну функцію $K_Z(t_1,t_2)$, дисперсію $D_Z(t)$, нормовану кореляційну функцію $\rho_Z(t_1,t_2)$ випадкового процесу Z.

№ варіанта	Завдання
1	$g(t)=t^2$, $h(t)=e^t$, $K_X(t_1,t_2)=\exp(-t_2-t_1)$, $K_Y(t_1,t_2)=16\exp(-t_2-t_1)$, $K_{XY}(t_1,t_2)=4\exp(-t_2-t_1)$.
2	$g(t)=e^{-2t}$, $h(t)=\sin(4t)$, $K_X(t_1,t_2)=4\cos(t_1-t_2)$, $K_Y(t_1,t_2)=36\cos(t_1-t_2)$, $K_{XY}(t_1,t_2)=12\cos(t_1-t_2)$.
3	$g(t)=t$, $h(t)=e^{-t}$, $K_X(t_1,t_2)=4(1+t_1\cdot t_2)$, $K_Y(t_1,t_2)=9(1+t_1\cdot t_2)$ $K_{XY}(t_1,t_2)=6(1+t_1\cdot t_2)$.
4	$g(t)=\sin(t)$, $h(t)=t$, $K_X(t_1,t_2)=9\cos(t_1)\cos(t_2)$, $K_Y(t_1,t_2)=25\cos(t_1)\cos(t_2)$, $K_{XY}(t_1,t_2)=15\cos(t_1)\cos(t_2)$.
5	$g(t)=\cos(t)$, $h(t)=\sin(t)$, $K_X(t_1,t_2)=9t_1\cdot t_2$, $K_Y(t_1,t_2)=36t_1\cdot t_2$, $K_{XY}(t_1,t_2)=18t_1\cdot t_2$.
6	$g(t)=t^2$, $h(t)=\cos(4t)$, $K_X(t_1,t_2)=25(2+\lvert t_2-t_1\rvert)-1$, $K_Y(t_1,t_2)=(2+\lvert t_2-t_1\rvert)-1$, $K_{XY}(t_1,t_2)=5(2+\lvert t_2-t_1\rvert)-1$.
7	$g(t)=e^{-3t}$, $h(t)=3t$, $K_X(t_1,t_2)=4\exp(-\lvert t_2-t_1\rvert)$, $K_Y(t_1,t_2)=9\exp(-\lvert t_2-t_1\rvert)$, $K_{XY}(t_1,t_2)=6\exp(-\lvert t_2-t_1\rvert)$.
8	$g(t)=\sin(6t)$, $h(t)=e^t$, $K_X(t_1,t_2)=4t_1\cdot t_2$, $K_Y(t_1,t_2)=49t_1\cdot t_2$, $K_{XY}(t_1,t_2)=14t_1\cdot t_2$.
9	$g(t)=t^3$, $h(t)=\sin(2t)$, $K_X(t_1,t_2)=4\sin(t_1)\sin(t_2)$, $K_Y(t_1,t_2)=16\sin(t_1)\sin(t_2)$, $K_{XY}(t_1,t_2)=8\sin(t_1)\sin(t_2)$.
10	$g(t)=e^{-2t}$, $h(t)=\cos(4t)$, $K_X(t_1,t_2)=(t_1\cdot t_2)^2$, $K_Y(t_1,t_2)=16(t_1\cdot t_2)^2$, $K_{XY}(t_1,t_2)=4(t_1\cdot t_2)^2$.

Завдання 4. Знайти математичне сподівання $m_Y(t)$, кореляційну функцію $K_Y(t_1,t_2)$ та дисперсію $D_Y(t)$ випадкового процесу $Y(t)=X'(t)$, не виконуючи диференціювання $X(t)$. Знайти взаємну кореляційну функцію $K_{XY}(t_1,t_2)$, U – випадкова величина.

№ варіанта	Завдання
1	$X(t)=t^2-Ue^{-3t}$, U ∈ N(2; 0,7).
2	$X(t)=-U\cdot t^2-\sin(t)$, U ∈ B(10;0.5).
3	$X(t)=U\cdot t-4t^2$, U ∈ R(3;6).
4	$X(t)=-U\cdot t^3-\sin(t)$, U ∈ P(4).
5	$X(t)=U\cos(3t)-3$, U ∈ P(5).
6	$X(t)-Ue^{-2t}-t$, U ∈ P(2).
7	$X(t)=3t^2+Ue^{-2t}$, U ∈ E(0.2).
8	$X(t)=U\sin(t)+t$, U ∈ N(1;2).
9	$X(t)=5t^2-U\sin(t)$, U ∈ B(10; 0.1).
10	$X(t)=-U\cdot t^3-\cos(t)$, U ∈ R(–1;3).

Завдання 5. $X(t)=f(t)U$, $f(t)$ – невипадкова функція, U – випадкова величина $Z(t)=\int_0^t X(s)ds$. Знайти математичне сподівання $m_Z(t)$, кореляційну функцію $K_Z(t_1,t_2)$, дисперсію $D_Z(t)$, взаємні кореляційні функції $K_{ZX}(t_1,t_2)$, $K_{XZ}(t_1,t_2)$, не інтегруючи $X(t)$.

№ варіанта	Завдання
1	$f(t)=\cos(6t)$, U ∈ E(0.2).
2	$f(t)=1/(1+t)^2$, U ∈ E(0.5).
3	$f(t)=\sin(3t)$, U ∈ B(10;0.3).
4	$f(t)=1+e^{-2t}$, U ∈ B(20;0.4).
5	$f(t)=e^{-3t}$, U ∈ P(2).
6	$f(t)=e^{-3t}$, U ∈ P(5).
7	$f(t)=\sin(2t)$, U ∈ N(–1;3).
8	$f(t)=1/(2t+1)$, U ∈ R(–2;4).
9	$f(t)=sh(2t)$, U ∈ R(–2;2).
10	$f(t)=1/(t^2+1)$, U ∈ N(–3;2).

Завдання 6. Довести, що випадковий процес $X(t)$ є стаціонарним у широкому сенсі. Перевірити властивість ергодичності для математичного сподівання та кореляційної функції. Знайти дисперсію випадкового процесу, U, V − некорельовані випадкові величини.

№ варіанта	Завдання
1	$X(t)=(U-2)\cos(3t)-V\sin(3t)$, U ∈ R(0;4), V∈N(32;0).
2	$X(t)=(U+2)\cos(2t)-V\sin(2t)$, U ∈ N(−2;2), V∈R(− 32;32).
3	$X(t)=U\cos(5t)-(V-5)\sin(5t)$, U ∈ R(−4;4), V ∈N(34;5).
4	$X(t)=(U-4)\cos(8t)-V\sin(8t)$, U ∈ P(4), V∈ N(0;2).
5	$X(t)=(U-1)\cos(20t)-V\sin(20t)$, U ∈E(1), V ∈ N(0;1).
6	$X(t)=(U-2)\cos(11t)-(V-8)\sin(11t)$, U ∈B(10;0.2), V ∈ B(10;0.8).
7	$X(t)=(U-1)\cos(6t)-(V-4/3)\sin(6t)$, U ∈ R(−1;3), V ∈P(4/3).
8	$X(t)=U\cos(21t)-V\sin(21t)$, U ∈ R(−1;1), V ∈E (3).
9	$X(t)=(U-8)\cos(15t)-V\sin(15t)$, U ∈ B(20;0.4), V ∈N(8.4;0).
10	$X(t)=U\cos(5t)-(V-5)\sin(5t)$, U ∈ R(−2;2), V ∈N(32;5).

Практичне заняття № 10.
Ланцюги Маркова

Задано матриці перехідних станів та вектор початкових станів дискретної системи:

1. Доповнити матриці перехідних станів, поставивши замість зірочок необхідні значення. Побудувати граф станів Марківського ланцюга.

2. Знайти матрицю перехідних ймовірностей за 2, 3, 4, 5 кроків.

3. Визначити розподіл ймовірностей станів системи за 1, 2, 3 кроків за початкового стану системи $S_0 = s_1$ ($p_2 = 1$, $p_i = 0, i \neq 2$).

4. Скласти балансну систему рівнянь для стаціонарного стану системи. Визначити фінальні ймовірності.

Теоретичні відомості

Процес Маркова – випадковий процес, в якому ймовірність значень випадкової величини у наступні моменти часу $t+\tau$ не залежить від того, які значення вона приймала у попередні моменти.

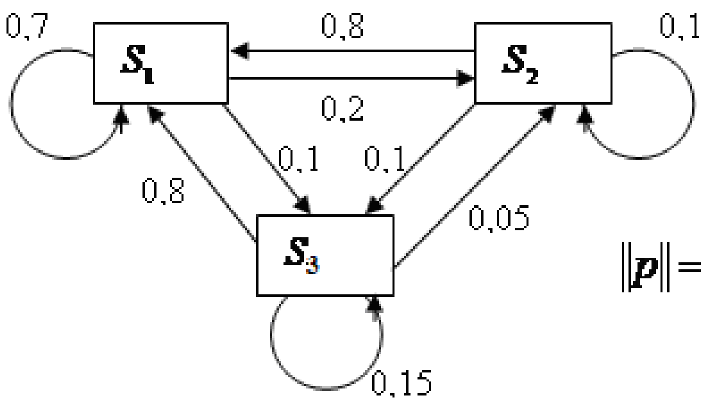
Ланцю́г Ма́ркова – випадковий процес, частковий випадок Марківського процесу, у якого часова множина та/або фазовий простір дискретні.

Ланцюг Маркова можна задавати двома способами: аналітично та графічно.

Аналітично ланцюг Маркова задається ***матрицею переходу*** P розмірності $n \times n$, елементами якої є перехідні імовірності p_{ij}, а також вектором ймовірностей усіх станів системи у початковий момент часу $t=0$, які називаються ***початковими ймовірностями*** станів системи:

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}, \quad P_0 = \left(p_1^{(0)}, p_2^{(0)}, \ldots, p_n^{(0)}\right),$$

де n – кількіть станів системи, всі ймовірності $p_i^{(0)}$ крім однієї, дорівнюють 0, а ймовірність, в якій перебуває система, дорівнює 1. Кожний рядок матриці перехідних ймовірностей повинен складати повну групу подій: $\forall k, \sum_{i=1}^{n} p_{ik} = 1$.



$$\|p\| = \begin{pmatrix} 0{,}7 & 0{,}2 & 0{,}1 \\ 0{,}8 & 0{,}1 & 0{,}1 \\ 0{,}8 & 0{,}05 & 0{,}15 \end{pmatrix}$$

Рис. 10.1. Приклад графу станів системи та відповідної матриці перехідних ймовірностей

Графічно ланцюг Маркова задається графом, вершинами якого є стани системи, дугами – ймовірності переходу системи з одного стану в інший (рис. 10.1).

Властивість Чепмена-Колмогорова: $p_{ij}^{(k)} = \sum_{m=1}^{n} p_{im}^{(k)} p_{mj}^{(k)}$ – матриця перехідних ймовірностей за один крок.

Матрицю перехідних ймовірностей на *k*-му кроці можна визначити як $\left\| p_{ij}^{(k)} \right\| = P^k$.

Ймовірності перебування системи в j-му стані на *k*-му кроці визначаються за рекурентною формулою:

$$p_j^{(k)} = \sum_{i=1}^{n} p_i^{(k-1)} p_{ij}. \qquad (10.1)$$

Ергодична властивість. За деяких умов у ланцюгу Маркова із зростанням номера кроку встановлюється стаціонарний режим, в якому система S продовжує блукати по станах, проте ймовірності цих станів вже не залежать ні від номера кроку, ні від початкового розподілу ймовірностей. Такі ймовірності називаються ***граничними*** (або ***фінальними***).

У випадку систем з дискретними станами та дискретним часом для визначення граничних ймовірностей роботи системи у стаціонарному режимі складається система балансних рівнянь за наступними правилами: в кожному рівнянні ліворуч стоїть гранична ймовірність i-го стану p_i помножена на сумарну інтенсивність вихідних потоків, а праворуч – сума добутків інтенсивностей вхідних потоків на ймовірності тих станів, з яких ці потоки виходять. Також додається рівняння, яке визначає, що ймовірності становлять повну групу подій.

Так для системи, представленої на рис. 10.1, система балансних рівнянь може бути записана наступним чином:

$$\begin{cases} p_1(p_{12} + p_{13}) = p_2 \cdot p_{21} + p_3 \cdot p_{31} \\ p_2(p_{21} + p_{31}) = p_1 \cdot p_{13} + p_2 \cdot p_{23} \\ p_1 + p_2 + p_3 = 0. \end{cases}$$

Оскільки система в загальному вигляді складається з 4-х рівнянь з 3-ма невідомими, то для того, що б вона не була перевизначеною і мала єдиний розв’язок, одне з рівнянь можна виключити, але рівняння, що визначає суму всіх шуканих ймовірностей як одиницю, обов’язково лишається.

Приклад виконання роботи

Нехай система задана матрицею перехідних ймовірностей P. Оскільки кожний рядок матриці повинен складати повну групу подій (сума ймовірностей дорівнює 1), то замінимо зірочки відповідними значеннями:

$$P=\begin{pmatrix}0 & * & 1/2 & 1/3\\ 0 & 1/3 & * & 0\\ 0 & * & 1/6 & 1/6\\ 1 & 0 & 0 & 0\end{pmatrix}\Rightarrow P=\begin{pmatrix}0 & 1/6 & 1/2 & 1/3\\ 0 & 1/3 & 2/3 & 0\\ 0 & 2/3 & 1/6 & 1/6\\ 1 & 0 & 0 & 0\end{pmatrix}$$

1. Система має чотири стани, на основі матриці перехідних ймовірностей складемо її граф (рис. 10.2).

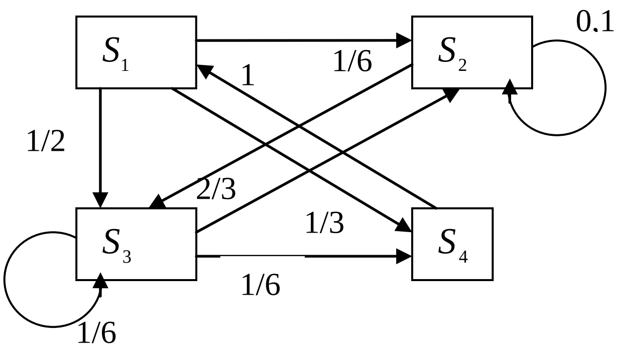


Рис. 10.2. Граф станів системи

2. Знайдемо матриці перехідних ймовірностей за 2, 3, 4, 5 кроків. Для обчислення застосуємо програмне середовище *Mathcad*. Задамо матрицю перехідних ймовірностей P і піднесемо її у відповідні степені (рис. 10.3).

$$P := \begin{pmatrix} 0 & \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & \frac{2}{3} & \frac{1}{6} & \frac{1}{6} \\ 1 & 0 & 0 & 0 \end{pmatrix} \qquad k := 2 \quad P2 := P^2 = \begin{pmatrix} \frac{1}{3} & \frac{7}{18} & \frac{7}{36} & \frac{1}{12} \\ 0 & \frac{5}{9} & \frac{1}{3} & \frac{1}{9} \\ \frac{1}{6} & \frac{1}{3} & \frac{17}{36} & \frac{1}{36} \\ 0 & \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \end{pmatrix}$$

$$k := 3 \quad P3 := P^3 = \begin{pmatrix} \frac{1}{12} & \frac{17}{54} & \frac{11}{24} & \frac{31}{216} \\ \frac{1}{9} & \frac{11}{27} & \frac{23}{54} & \frac{1}{18} \\ \frac{1}{36} & \frac{49}{108} & \frac{83}{216} & \frac{29}{216} \\ \frac{1}{3} & \frac{7}{18} & \frac{7}{36} & \frac{1}{12} \end{pmatrix}$$

$$k := 4 \quad P4 := P^4 = \begin{pmatrix} \frac{31}{216} & \frac{275}{648} & \frac{425}{1296} & \frac{5}{48} \\ \frac{1}{18} & \frac{71}{162} & \frac{43}{108} & \frac{35}{324} \\ \frac{29}{216} & \frac{89}{216} & \frac{493}{1296} & \frac{95}{1296} \\ \frac{1}{12} & \frac{17}{54} & \frac{11}{24} & \frac{31}{216} \end{pmatrix} \qquad k := 5 \quad P5 := P^5 = \begin{pmatrix} \frac{5}{48} & \frac{1493}{3888} & \frac{1061}{2592} & \frac{797}{7776} \\ \frac{35}{324} & \frac{409}{972} & \frac{751}{1944} & \frac{55}{648} \\ \frac{95}{1296} & \frac{1607}{3888} & \frac{3151}{7776} & \frac{841}{7776} \\ \frac{31}{216} & \frac{275}{648} & \frac{425}{1296} & \frac{5}{48} \end{pmatrix}$$

Рис. 10.3. Початкова матриця перехідних ймовірностей та перехідні ймовірності за 2, 3, 4, 5 кроків

3. Визначимо розподіл ймовірностей станів системи за 1, 2, 3 кроки за початкового стану системи $P_0 = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}$. Для цього скористаємось рекурентною формулою (10.1) (рис. 10.4).

$$p0 := \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}$$

$$k := 1 \quad p1 := p0 \cdot P = \begin{pmatrix} 0 & \frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix} \qquad k := 2 \quad p2 := p1 \cdot P = \begin{pmatrix} 0 & \frac{5}{9} & \frac{1}{3} & \frac{1}{9} \end{pmatrix}$$

$$k := 3 \quad p3 := p2 \cdot P = \begin{pmatrix} \frac{1}{9} & \frac{11}{27} & \frac{23}{54} & \frac{1}{18} \end{pmatrix}$$

Рис. 10.4. Розподіл ймовірностей станів системи за 1, 2, 3 кроки

4. Складемо балансну систему рівнянь:

$$\begin{cases} p_1 (p_{12} + p_{13} + p_{14}) = p_4 \\ p_2 \cdot p_{23} = p_1 \cdot p_{21} + p_3 \cdot p_{32} \\ p_3 (p_{32} + p_{34}) = p1 \cdot p_{13} + p_2 \cdot p_{23} \\ p_4 = p_1 \cdot p_{14} + p_3 \cdot p_{34} \\ p_1 + p_2 + p_3 + p_4 = 1. \end{cases}$$

Виключимо з розгляду четверте рівняння і розв'яжемо систему, застосувавши блок *Given – Find* (рис. 10.5).

$$p1 := 0 \qquad p2 := 0 \qquad p3 := 0 \qquad p4 := 0$$

Given

$$p1 + p2 + p3 + p4 = 1$$

$$p1 = p4$$

$$p2 \cdot \frac{2}{3} = \frac{1}{6} \cdot p1 + \frac{2}{3} \cdot p3$$

$$p3 \cdot \left(\frac{2}{3} + \frac{1}{6}\right) = p1 \cdot \frac{1}{2} + p2 \cdot \frac{2}{3}$$

$$r := \mathrm{Find}(p1, p2, p3, p4) \qquad p1 := r_0 = \frac{4}{41} \qquad p2 := r_1 = \frac{17}{41}$$

$$p3 := r_2 = \frac{16}{41} \qquad p4 := r_3 = \frac{4}{41}$$

Рис. 10.5. Розв’язання системи балансних рівнянь

Отже фінальні ймовірності становлять: $P_f = \left(\frac{4}{41}, \frac{17}{41}, \frac{16}{41}, \frac{4}{41}\right)$.

Варіанти для самостійного виконання

Варіант 1

$$\begin{pmatrix} 1/3 & 1/3 & \ast & 0 \\ 1/2 & 1/2 & 0 & \ast \\ 1/4 & 1/4 & \ast & 1/2 \\ 0 & 1/2 & 0 & \ast \end{pmatrix};$$

Варіант 2

$$\begin{pmatrix} \ast & 2/3 & 1/3 \\ 1/2 & \ast & 1/2 \\ 1/3 & 2/3 & \ast \end{pmatrix};$$

Варіант 3

$$\begin{pmatrix} 0 & 1/6 & 1/2 & 1/3 \\ 0 & 1/3 & 2/3 & 0 \\ 0 & 2/3 & 1/6 & 1/6 \\ 1 & 0 & 0 & 0 \end{pmatrix};$$

Варіант 4

$$\begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix};$$

Варіант 5

$$\begin{pmatrix} \frac{1}{2} & \ast & 0 & \frac{1}{4} \\ 0 & 0 & 1 & 0 \\ \frac{1}{3} & 0 & \ast & \frac{2}{3} \\ 0 & \frac{1}{2} & \ast & \frac{1}{2} \end{pmatrix};$$

Варіант 6

$$\begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \ast \\ \frac{1}{4} & \frac{1}{2} & \ast & \frac{1}{4} \\ \frac{1}{3} & \ast & 0 & 0 \\ \ast & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix};$$

Варіант 7

$$\begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \ast & 0 & 0 \\ 0 & \frac{1}{4} & \frac{3}{4} & \ast & 0 \\ 0 & \ast & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{2}{3} & \ast \\ \ast & 0 & 0 & 0 & \frac{3}{4} \end{pmatrix}$$

Варіант 8

$$\begin{pmatrix} \ast & 0{,}4 & 0 & 0{,}5 & 0 \\ 0{,}3 & \ast & 0 & 0{,}3 & 0 \\ 0{,}4 & 0{,}5 & \ast & 0 & 0{,}1 \\ 0{,}2 & 0{,}6 & 0 & \ast & 0 \\ 0 & 0{,}1 & 0{,}4 & 0{,}2 & 0{,}3 \end{pmatrix};$$

Варіант 9

$$\begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 & \ast \\ \frac{1}{2} & 0 & 0 & \ast & 0 \\ 0 & 1 & \ast & 0 & 0 \\ 0 & \ast & \frac{1}{2} & 0 & \frac{1}{2} \\ \ast & 1 & 0 & 0 & 0 \end{pmatrix};$$

Варіант 10

$$\begin{pmatrix} \frac{1}{3} & \ast & 0 & \frac{1}{6} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} & 0 & 0 & \ast \\ \frac{1}{3} & \ast & 0 & \frac{1}{3} & 0 \\ \frac{1}{4} & \frac{3}{8} & 0 & \ast & 0 \\ 0 & 0 & 0 & \ast & 0 \end{pmatrix}$$

Практичне заняття № 11
Системи масового обслуговування

1. Визначити тип системи та записати його в позначеннях Кендалла-Лі.

2. Побудувати граф системи.

3. Розв’язати задачі аналітично.

Теоретичні відомості

Система масового обслуговування (СМО) – математичний (абстрактний) об’єкт, що складається з одного або декількох приладів (каналів) обслуговування замовлень, що надходять у систему через накопичувач або без накопичувача безпосередньо на обслуговування

Структура СМО визначається заданим потоком замовлень, кількістю обслуговуючих приладів, тривалістю обслуговування, кількістю місць очікування на обслуговування.

Функціонування СМО визначається наступними основними факторами:

– a – ймовірнісним розподілом моментів надходження замовлень (вхідний потік);

– b – ймовірнісним розподілом часу обслуговування (вихідний потік);

– c – кількістю та продуктивністю каналів обслуговування та їх конфігурацією (паралельна, послідовна, комбінована);

– e – максимальною ємністю системи;

– d – дисципліною черги, довжина черги визначається як різниця $e– c$;

– f – потужністю генератора замовлень.

Для скороченого позначення типу СМО застосовується символіка, введена Кендаллом-Лі:

$(a/ b/ c)(d/ e/ f)$,

де використовуються наступні стандартні позначення:

– для вхідного та вихідного потоку (a, b):

- M – Марківський процес (експоненціальний розподіл);
- D – детермінований потік;
- E_k – потік Ерланга, або гама розподіл (суми незалежних випадкових величин мають експоненціальний розподіл);
- GI – довільний тип розподілу моментів надходження замовлень;

- G – довільний тип розподілу часу обслуговування;

– для дисципліни черги:

- FCFS – першим прийшов – першим вийшов (FIFO);
- LCFS – останнім прийшов – першим вийшов (LIFO);
- SIRO – випадковий відбір замовлень;
- GD – довільний тип дисципліни.

Значення параметрів *c, e, f* – натуральні числа, включаючи нескінченність.

Характеристики СМО. Нехай: λ – інтенсивність вхідного потоку, μ – інтенсивність вихідного потоку.

До основних характеристик. що визначають процес функціонування системи масового обслуговування відносять:

– *завантаженість,* або *коефіцієнт використання системи* ρ – це відношення інтенсивності вхідного потоку до інтенсивності обслуговування $\rho = \lambda / N\mu$ (для одноканальної системи $\rho = \lambda / \mu$). Якщо система з відмовами (не всі замовлення можуть потрапити до системи), то $\rho = \lambda_{еф} / \mu$, де $\lambda_{еф}$ – ефективна інтенсивність надходження клієнтів у систему $(\lambda_{еф} < \lambda)$. Це відношення показує, наскільки задіяні її ресурси. Значення коефіцієнта завантаження визначає необхідну і достатню умови існування стаціонарного режиму, а саме $\rho \le 1$ або $\lambda \le \mu$. В протилежному випадку система буде працювати у режимі перевантаження. Завантаженість ρ характеризує:

- середня кількість замовлень *m*, що надходять до системи, за середній час обслуговування одного замовлення τ ($\rho = m \cdot \tau$);
- частку від інтервалу часу, на протязі якого сервіс зайнятий;
- ймовірність зайнятості сервісу (ймовірність простою сервісу буде складати: $1-\rho$);
- середню кількість замовлень, що знаходиться в обслуговуючому пристрої (оскільки сервіс може обслуговувати лише одне замовлення з ймовірністю ρ і простоювати з ймовірністю $1-\rho$, то середня кількість замовлень в обслуговуючому пристрої: $1 \cdot \rho + 0 \cdot (1-\rho) = \rho$).

– *відносна пропускна здатність системи Q* – відносне середнє число замовлень (табл. 11.1).

Таблиця 11.1

Відносна пропускна здатність СМО

Тип СМО	З відмовами	З чергою	З необмеженою чергою
Одноканальна	$Q = \mu / (\mu + \lambda)$	$Q = 1 - \rho^{m+1} \cdot p_0$	$Q = 1$
Багатоканальна	$Q = 1 - p_{відмови}$	$Q = 1 - p_{відмови}$	$Q = 1$

p_0 – ймовірність того, що канал вільний, $p_{відмови}$ – ймовірність відмови.

– *абсолютна пропускна здатність системи* A – середнє число замовлень, яке може бути оброблене за одиницю часу (табл. 11. 2).

Таблиця 11.2

Абсолютна пропускна здатність СМО

Тип СМО	З відмовами	З чергою	З необмеженою чергою
Одноканальна	$A = Q \cdot \lambda = p_0 \cdot \lambda$	$A = Q \cdot \lambda$	$A = \lambda$
Багатоканальна	$A = Q \cdot \lambda$	$A = Q \cdot \lambda$	$A = \lambda$

p_0 – ймовірність того, що канал вільний, Q – відносна пропускна здатність.

– *ймовірність відмови* p_0 – ймовірність, що замовлення не потрапить у систему (гранична ймовірність того, що всі канали зайняті):

- $p_{відмови} = \dfrac{\lambda}{\lambda + \mu}$ – для одноканальної СМО;
- $p_{відмови} = \dfrac{\rho^n}{n!} \cdot p_0$ – для багатоканальної СМО.

– *середня кількість зайнятих сервісів* k (для багатоканальних СМО):

$$k = \sum_{i=0}^{n} i p_i = A / \mu = \rho\left(1 - \frac{\rho^n}{n!} p_0\right),$$

n – кількість сервісів, p_i – граничні ймовірності станів.

– *ймовірності станів системи* – найбільш важлива характеристика в тому сенсі, що, знаючи ймовірності, можна визначити всі інші характеристики. При цьому під станом системи розуміють кількість замовлень, що знаходиться у системі. Ймовірність стану системи, коли в ній знаходиться *k* замовлень, позначається як p_k (розрахункові формули наведені нижче);

– *час очікування* W_q – середня величина випадкового часу, який замовлення перебуває у черзі в стані очікування W_q;

– *час перебування у системі* W_s – середній час перебування замовлення у системі з моменту надходження до завершення обслуговування:

$$W_s = W_q + \frac{1}{\mu}$$

(сума середнього часу перебування у черзі та середнього часу обслуговування);

– *середня кількість замовлень у черзі* L_q – середня довжина черги

$$L_q = \sum_{n=m+1}^{\infty}(n-m)p_n = \lambda_{еф} \cdot \mathrm{W_q}$$

– *середня кількість замовлень* m, що знаходяться *у системі* L_s:

$$L_s = \sum_{n=1}^{\infty} n \cdot p_n = \lambda_{еф} \cdot \mathrm{W_s}$$

Залежності між середньою кількістю замовлень в системі L_s і черзі L_q та середнім часом перебування відповідно у системі W_S та черзі W_q називаються *формулами Літтла*.

Середню кількість зайнятих сервісів можна визначити як:

$$m = \frac{\lambda_{еф}}{\mu} = L_s - L_q .$$

Одноканальна система з відмовами (без черги):

$$p_0 = \frac{\mu}{\mu+\lambda}, \quad p_1 = \frac{\lambda}{\mu+\lambda},$$

$$Q = \frac{\mu}{\mu+\lambda} = 1 - p_1, \quad p_{відмови} = \frac{\lambda}{\mu+\lambda} \quad A = \frac{\mu \cdot \lambda}{\mu+\lambda} = \lambda Q .$$

Багатоканальна система з відмовами:

$$p_0 = \left(1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2!\mu^2} + \ldots + \frac{\lambda^i}{i!\mu^i} + \ldots + \frac{\lambda^n}{n!\mu^n}\right)^{-1} = \left(1 + \rho + \frac{\rho^2}{2!} + \ldots + \frac{\rho^i}{i!} + \ldots + \frac{\rho^n}{n!}\right)^{-1}$$

$$p_1 = \rho \cdot p_0,\ p_2 = \frac{\rho^2}{2!} \cdot p_0, \ldots,\ p_k = \frac{\rho^k}{k!} \cdot p_0, \ldots,\ p_n = \frac{\rho^n}{n!} \cdot p_0,\ p_{відмови} = \frac{\rho^n}{n!} p_0;$$

$$Q = 1 - p_{відмови} = 1 - \frac{\rho^n}{n!} p_0, \quad A = \lambda \cdot Q = \lambda \cdot \left(1 - \frac{\rho^n}{n!} p_0\right).$$

$$k_{зайнятих} = \sum_{i=0} (i \cdot p_i) = \frac{A}{\mu} = \rho \cdot \left(1 - \frac{\rho^n}{n!} p_0\right), \quad k_{вільних} = \sum_{i=0}^{n} (n - i) \cdot p_i$$

Одноканальна СМО з необмеженою чергою (без відмов):

$$p_0 = 1 - \rho,\ p_1 = \rho \cdot p_0,\ p_2 = \rho^2 \cdot p_0, \ldots,\ p_i = \rho^i \cdot p_0,\ (\rho < 1);$$

$$L_{сист} = \frac{\rho}{1 - \rho}, \quad L_{черги} = L_{сист} - L_{обсл},\ L_{обсл} = \rho,\ L_{очікування} = \frac{\rho^2}{1 - \rho};$$

$$\mathrm{W}_{сист} = \frac{1}{\lambda} L_{сист}, \quad \mathrm{W}_{очік} = \frac{1}{\lambda} L_{очік}.$$

$$\mathrm{W}_{сиис} = \frac{1}{\lambda} \cdot \frac{\rho}{1 - \rho}, \quad \mathrm{W}_{очік} = \frac{1}{\lambda} \cdot \frac{\rho^2}{1 - \rho}$$ – для одного замовлення.

Багатоканальна СМО з необмеженою чергою (без відмов):

$$p_0 = \left(1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \ldots + \frac{\rho^n}{n!} + \frac{\rho^{n+1}}{n!(n - \rho)}\right)^{-1},\ p_1 = \frac{\rho}{1!} p_0, \ldots,\ p_k = \frac{\rho^k}{k!} p_0, \ldots,$$

$$p_n = \frac{\rho^n}{n!} p_0,\ p_{n+1} = \frac{\rho^{n+1}}{n \cdot n!} p_0, \ldots,\ p_{n+r} = \frac{\rho^{n+r}}{n^r \cdot n!} p_0,\ P_{черга} = \frac{\rho^{n+1}}{n!(n - \rho)} p_0;$$

$$k = \frac{\lambda}{\mu} = \rho,\ L_{черга} = \frac{\rho^{n+1}}{n \cdot n! \left(1 - \rho/n\right)^2} p_0, \quad L_{сист} = L_{черга} + \rho.$$

***СМО з обмеженою чергою (з відмовами)*:**

Одноканальна СМО	Багатоканальна СМО
Граничні ймовірності	
$p_0=\dfrac{1-\rho}{1-\rho^{m+2}}$, $p_1=\rho\cdot p_0,\ p_2=\dfrac{\rho^2}{2!}\cdot p_0,\dots,$ $p_k=\dfrac{\rho^k}{k!}\cdot p_0$	$p_0=\left(1+\dfrac{\rho}{1!}+\dots+\dfrac{\rho^n}{n!}+\dots+\dfrac{\rho^{n+1}\left(1-\left(\rho/n\right)^m\right)}{n\cdot n!\left(1-\rho/n\right)}\right)$ $p_1=\rho\cdot p_0,\dots,\ p_n=\dfrac{\rho^n}{n!}\cdot p_0,$ $p_{n+1}=\dfrac{\rho^{n+1}}{n\cdot n!}\cdot p_0,\dots,\ p_{n+r}=\dfrac{\rho^{n+r}}{n^r\cdot n!}\cdot p_0,\ r=\overline{1,m}$
Ймовірність відмови	
$p_{відмови}=p_{m+1}=\rho^{m+1}p_0$	$p_{відмови}=p_{m+n}=\dfrac{\rho^{m+n}}{n^m\cdot n!}p_0$
Абсолютна пропускна здатність	
$A=\lambda Q=\lambda\left(1-\rho^{m+1}p_0\right)$	$A=\lambda Q=\lambda\left(1-\dfrac{\rho^{m+n}}{n^m\cdot n!}p_0\right)$
Відносна пропускна здатність	
$Q=1-p_{відмови}=1-\rho^{m+1}p_0$	$Q=1-p_{відмови}=1-\dfrac{\rho^{m+n}}{n^m\cdot n!}p_0$
Середня кількість замовлень у черзі	
$L_{черги}=\rho^2\dfrac{1-\rho^m(m+1-m\rho)}{\left(1-\rho^{m+2}\right)(1-\rho)}$	$L_{черги}=\rho^{n+1}p_0\left(1-\left(m+1-m\dfrac{\rho}{n}\right)\left(\dfrac{\rho}{n}\right)^m\right)$
Середня кількість зайнятих сервісів	
$L_{обсл}=1-p_0$	$k=\rho\left(1-\dfrac{\rho^{n+m}}{n^m n!}p_0\right)$
Середня кількість замовлень у системі	
$L_{сист}=L_{очікування}+L_{обсл}$	$L_{сист}=L_{очікування}+k$
Середній час перебування у системі	
$W_{сист}=\dfrac{L_{систми}}{\lambda}$	$W_{сист}=\dfrac{L_{систми}}{\lambda}$
Середній час перебування у черзі	
$W_{очікування}=\dfrac{L_{очікування}}{\lambda}$	$W_{очікування}=\dfrac{L_{очікування}}{\lambda}$

Більш детальна інформація викладена у посібнику [3].

Приклад виконання роботи

Приклад 1. СМО з відмовами.

Служба ремонту має три телефонні лінії для виклику майстра. Якщо всі телефони зайняті, клієнт не може додзвонитись, отже отримує відмову. Дзвінки надходять з інтенсивністю $\lambda = 60 \,{}^{\text{дзвінків}}\!/_{\text{год.}}$ Час на прийом замовлення (обслуговування) є випадковою величиною з експоненціальним законом розподілу, середня тривалість розмови $t_{обсл.} = 3хв.$ Визначити показники ефективності роботи системи.

Розв'язок. Система відноситься до типу багатоканальних СМО з відмовами, граф станів якої наведений на рис. 11.1.

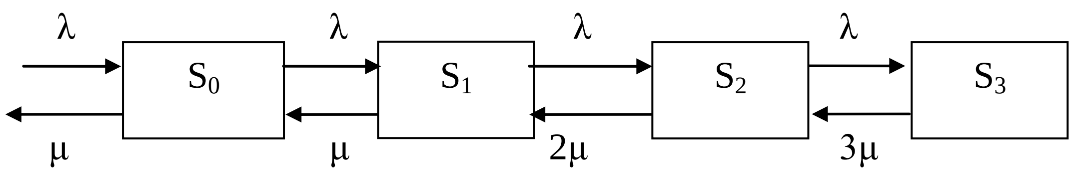


Рис. 11.1. Приклад триканальної СМО з відмовами

Приведемо всі параметри до одних одиниць вимірювання часу (годин): $t_{обсл.} = 3хв = 3/60 = 0{,}05 \, год.$

Тоді інтенсивність обслуговування:

$\mu = {}^{1}\!/_{t_{обсл.}} = {}^{1}\!/_{0{,}05} = 20 \, дзвінків$, приведена інтенсивність потоку:

$$\rho = {}^{\lambda}\!/_{\mu} = {}^{60}\!/_{20} = 3.$$

Обчислимо граничні ймовірності станів:

$$p_0 = \left(1 + \rho + \frac{\rho^2}{2!} + \frac{\rho^3}{3!}\right)^{-1} = \left(1 + 3 + \frac{9}{2} + \frac{27}{6}\right)^{-1} = \frac{1}{13} = 0.077$$

$$p_1 = \rho \cdot p_0 = 0{,}231, \quad p_2 = \frac{\rho^2}{2!} \cdot p_0 = 0{,}346, \quad p_3 = \frac{\rho^3}{3!} \cdot p_0 = 0{,}346$$

Ймовірність відмови буде становити: $p_{відмови} = 0.346$.

Обчислимо відносну та абсолютну пропускні здатності:

$$Q = 1 - p_{\text{відмови}} = 1 - 0{,}346 = 0.654, \quad A = \lambda \cdot Q = 60 \cdot 0{,}654 = 39{,}24.$$

Середнє число зайнятих (вільних) каналів:

$$k_{зайнятих} = \frac{A}{\mu} = \frac{39,24}{20} = 1,962\,, \qquad k_{вільних} = 1,038\,.$$

Відповідні коефіцієнти зайнятості будуть становити:

$$K_{зайнятих} = \frac{k_{зайнятих}}{n} = \frac{1,962}{3} = 0,654 \qquad K_{простою} = \frac{k_{простою}}{n} = \frac{1,038}{3} = 0,094\,.$$

Приклад 2. СМО з очікуванням.

На деяку базу в середньому через 30 хв прибувають автомашини з продукцією. Середній час розвантаження однієї машини становить 1,5 години. Розвантаження виконує дві бригади вантажників. На території бази можуть знаходитися в черзі та очікувати розвантаження не більше 4 автомашин. Визначити показники роботи СМО.

Розв'язок. Визначимо основні показники СМО:

– інтенсивність замовлень $\lambda = 2$;

– інтенсивність обслуговування $\mu = 1/t_{обсл} = 2/3$;

– кількість сервісів $n = 2$;

– довжина черги $m = 3$;

– відносне навантаження на систему $\rho = \lambda / \mu = 3$;

– приведена інтенсивність потоку $\rho / n = 1,5$

Знайдемо ймовірність відсутності клієнтів у системі:

$$p_0 = \left(1 + \frac{3}{1!} + \frac{3^2}{2!} + \frac{3^3}{2 \cdot 2!} \cdot \frac{\left(1 - \left(\frac{3}{2}\right)^4\right)}{\left(1 - \frac{3}{2}\right)} \right)^{-1} = 0,0158$$

Ймовірність відмови: $p_{відм} = \frac{\rho^{n+m}}{n^m \cdot n!} \cdot p_0 = \frac{3^6}{2^4 \cdot 2!} \cdot 0,0158 = 0,36$

Відносна пропускна здатність: $Q = 1 - p_{відмови} = 1 - 0,36 = 0,64$.

Абсолютна пропускна здатність: $A = \lambda Q = 2 \cdot 0,64 = 1,28 \left(авто/год \right)$.

Середня кількість зайнятих бригад: $k_{зайн.} = \frac{A}{\mu} = \rho Q = 3 \cdot 0,64 = 1,92$

Середня довжина черги:

$$L_{черги}=\frac{3^3}{2\cdot 2!}\cdot\frac{1-(1{,}5)^4\left(4+1-\frac{4}{2}\cdot 3\right)}{(1-1{,}5)^2}\cdot 0{,}0158\approx 2{,}6(авто).$$

Середній час перебування клієнта у черзі: $W_{очікування}=\frac{2{,}6}{2}=1{,}3(год)$.

Середня кількість клієнтів у системі: $L_{сист.}=2{,}6+1{,}92=4{,}52$

Середній час перебування клієнта у системі:

$$W_{сист.}=\frac{4{,}52}{2}=2{,}26(год).$$

Варіанти для самостійного виконання

Варіант 1

1.1. Автозаправна станція має 4 бензоколонки. Середній час заправки 2 хв. Вхідний потік автомашин – найпростіший з інтенсивністю 1,5 авт./хв. Якщо всі колонки зайняті, то машина на станцію не заїжджає. Визначте ймовірність відмови і середнє число зайнятих колонок.

1.2. Покупці магазину утворюють найпростіший потік вимог з інтенсивністю 150 осіб/год. Визначте найменше число продавців, коли середня кількість покупців, які очікують обслуговування, не перевищуватиме 3 особи, час обслуговування становитиме 4 хв.

Варіант 2

2.1. На телефонну лінію надходить простий потік дзвінків з інтенсивністю $\lambda=0{,}4$ дзвінки/хв. Середня тривалість розмови становить 3 хв. Час розмови має показниковий закон розподілу. Знайти граничні ймовірності станів системи та характеристики обслуговування.

2.2. У нафтоналивному порту 4 причали для заправки танкерів, які приходять в середньому через 18 год, а час завантаження становить в середньому дві доби. У черзі можуть стояти не більше 2 танкерів. Визначте пропускну здатність і холостий хід порту.

Варіант 3

3.1. Потік бажаючих оформити виклик лікаря додому – найпростіший. В середньому абоненти телефонують через кожні 10 с. Час прийому виклику розподілено за показниковим законом із середнім значенням 12 с. Визначте найменше число телефонів у реєстратурі, коли виклик приймається не менш, ніж від 90% абонентів. Вважається, що в разі невдачі абонент не робить більше спроб додзвонитися.

3.2. Автоматична мийка може прийняти на обслуговування одночасно 4 автомашини. В середньому машини прибувають через 2 хв, а середня тривалість мийки – 10 хв. У черзі можуть перебувати не більше 6 машин. Визначте ймовірність того, що в системі знаходиться хоча б одна машина і завантаженість однієї установки для миття машин.

Варіант 4

4.1. У магазині є 3 довідкових телефони. В середньому звертаються за довідками 40 осіб/год. Середня тривалість розмови 3 хв. Витрати, пов’язані з роботою одного телефона, 0,5 грн/хв. Визначте мінімальну вартість однієї хвилини розмови по телефону, за якої система не збиткова.

4.2. У перукарні працюють два майстри. Час обслуговування розподілено за показниковим законом із середнім 12 хв. Очікувати обслуговування можуть не більше трьох осіб. Потік клієнтів – найпростіший з інтенсивністю 10 клієнтів/год. Знайдіть найважливіші операційні характеристики цієї системи.

Варіант 5

5.1. Платна стоянка для легкових машин має 7 місць. Знайдіть ймовірність того, що прибула машина знайде вільне місце, якщо машини в середньому прибувають через 10 хв, а займають місце на стоянці в середньому на 1 год.

5.2. Система автоматичної посадки літаків одночасно може зберігати дані тільки про 6 літаків, що знаходяться в повітрі. Літаки, що підлітають до аеродрому, утворюють найпростіший потік з інтенсивністю 6 літаків/год. Якщо в момент запиту посадки система заповнена, то літак відлітає до запасного аеродрому. Аеродром має 3 посадкові смуги, літак займає смугу в середньому 20 хв. Знайдіть пропускну здатність СМО, завантаженість однієї смуги, середнє число зайнятих смуг, середній час очікування початку посадки після запиту.

Варіант 6

6.1. Потік деталей, що сходять з конвеєра, найпростіший з інтенсивністю 2 дет./хв. Час перевірки деталі контролером має показниковий закон розподілу із середнім 2 хв/дет. Визначте частку неперевірених деталей.

6.2. На склад в середньому прибуває 3 машини за годину.

Розвантаження здійснюють 3 бригади вантажників. Середній час розвантаження машини – 1 год. В черзі на розвантаження може перебувати не більше 4-х машин. Дати оцінку роботи СМО.

Варіант 7

7.1. Місто обслуговують 4 машини швидкої допомоги. Виклики надходять в середньому через 4 год. Імовірність того, що хоча б одна машина зайнята, дорівнює 0,25. Визначте середнє число зайнятих машин і середню частку простою машин.

7.2. Розглядається робота автозаправної станції (АЗС), на якій є 2 заправні колонки. Припустимо, що вона описується процесом розмноження і загибелі в стаціонарному режимі. Заправка кожної машини триває в середньому 3 хвилини. В середньому на АЗС кожні дві хвилини прибуває машина. Число місць в черзі необмежене. Всі машини, що стали на заправку, терпляче чекають своєї черги. Визначте імовірність того, що на заправці знаходиться 5 машин; імовірність того, що знову прибулій машині доведеться чекати обслуговування.

Варіант 8

8.1. У річковому порту один причал, інтенсивність вхідного потоку 5 суден/день. Інтенсивність вантажно-розвантажувальних робіт – 6 суден/день. Маючи на увазі стаціонарний режим роботи, визначте всі середні характеристики системи.

8.2. У автоматизовану інформаційну систему надходять запити з інтенсивністю $\lambda = 2$ зап./хв. Середня тривалість відповіді може бути відповідно: 10, 15, 20 сек. Вбудований буфер дозволяє зберігати інформацію про два запити. Якщо в буфері вже зберігаються два запити, черговий запит отримує відмову. Визначте основні ймовірнісні характеристики ефективності системи.

Варіант 9

9.1. Яке оптимальне число каналів обслуговування повинна мати СМО, якщо інтенсивність потоку заявок дорівнює 3, середнє число заявок, що були обслужені в одиницю часу, дорівнює 2, штраф за кожну відмову дорівнює 5, а вартість простою одній лінії дорівнює 2?

9.2. Магазин відвідує в середньому 90 осіб на годину, які обслуговуються одним касиром. Час на обслуговування однієї особи в середньому становить 1 хв. Черга в зал обслуговування обмежена 5 покупцями. Оцінити ефективність роботи СМО.

Варіант 10

10.1. Визначте число злітно-посадочних смуг для літаків з урахуванням вимоги, що ймовірність очікування повинна бути менша, ніж 0,05. При цьому інтенсивність вхідного потоку 27 літаків на добу, а інтенсивність їх обслуговування – 30 літаків на добу.

10.2. Інтенсивність потоку відвідувачів їдальні становить 150 осіб на годину. Є 3 касири, кожен з яких обслуговує в середньому 1 відвідувача за хвилину. Знайти характеристики системи масового обслуговування, черга необмежена.

Практичне заняття № 12.
Імітаційне моделювання систем масового обслуговування

До задач, наведених у завданні 11, побудувати імітаційну модель засобами *Matlab – Simulink – SimEvents*.

Теоретичні відомості

Matlab – середовище імітаційного моделювання з вбудованою графічною мовою ***Simulink***, де реалізовано багато спеціалізованих бібліотек. Для моделювання СМО використовується бібліотека ***SimEvents***.

Розглянемо послідовність дій для створення імітаційної моделі.

1. Після завантаження програмного середовища *Matlab* на панелі головного меню необхідно вибрати кнопку запуску мови *Simulink*:

Відкриється початкова сторінка додатка, де необхідно витрати створення вікна для моделі: *Blank Model*. На екрані з’явиться вікно, де буде створюватись модель. На верхній панелі меню необхідно вибрати кнопку *Library Browser*, за допомогою якої завантажується панель графічних бібліотек, де в правому вікні – перелік бібліотек, в лівому – графічні компоненти відповідної бібліотеки (рис. 12.1).

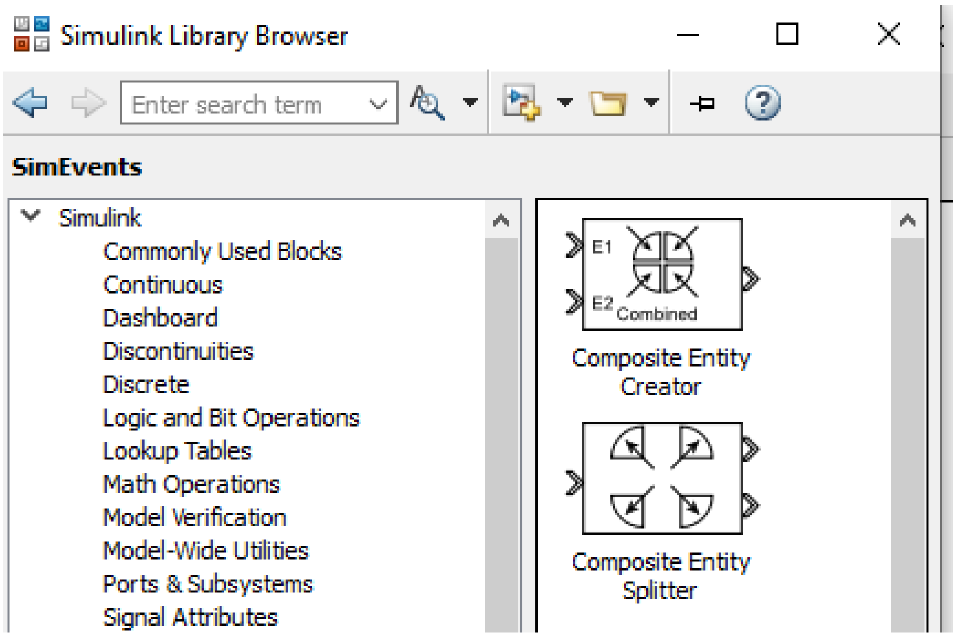


Рис. 12.1. Браузер графічних бібліотек *Simulink*

Назви бібліотек розташовані в алфавітному порядку. Необхідно завантажити бібліотеку *SimEvents*. В правому вікні з'являться графічні компоненти, які будемо використовувати для створення нашої імітаційної моделі (рис. 12.2). Встановлювати необхідні компоненти у модель можна шляхом перетягування або, натиснувши на компонент правою кнопкою «миші», вибрати дію Add block to model.

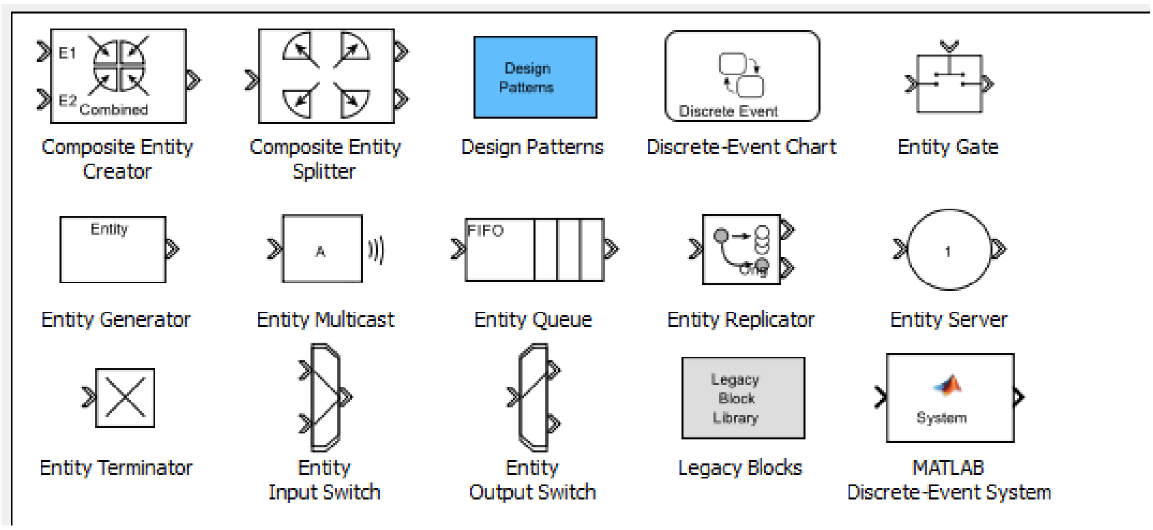


Рис. 12.2. Фрагмент графічних компонентів бібліотеки *SimEvents*

2. Побудову моделі будемо виконувати, спираючись на узагальнену структуру СМО (рис. 12.3).

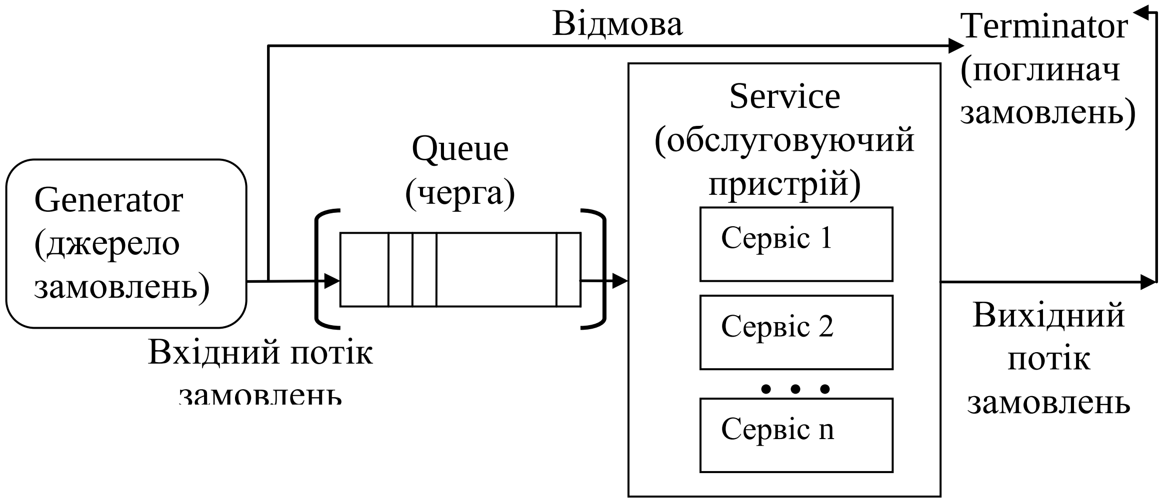


Рис. 12.3. Узагальнена структурна схема СМО

Оскільки модель буде простою, то будемо використовувати лише необхідні основні блоки:

– Generator (джерело замовлень);

– Queue (черга);

– Service (обслуговуючий пристрій);

– Terminator (поглинач замовлень);

– Засоби візуалізації результатів моделювання.

Розглянемо основні блоки моделі та їх налаштування. Для налаштування блоків необхідно встановити курсор на поле блоку та двічі натиснути правою кнопкою «миші». Відкриється вікно налаштування, в якому треба встановити необхідні параметри для моделі та натиснути кнопку ОК. Блоки з’єднуються протягуванням курсору від виходу з попереднього блоку до входу наступного, тримаючи натиснутою ліву клавішу «миші».

2.1 ***Генератор замовлень*** (рис. 12.4). У вікні налаштування генератора є декілька закладок, розглянемо основні налаштування.

На головній закладці ***Entity Generation*** необхідно налаштувати наступні пункти:

– *Generation method* (метод генерації):

– Time-based (на основі часу) – генерування замовлень, використовуючи час між генераціями з вхідного сигналу або статистичного розподілу;

– Event-based (на основі події) для зовнішньої події, щоб визначити час взаємодії замовлення.

Для наших моделей оберемо метод Time-based (на основі часу).

– Time source (джерело часу):

– Dialog – фіксований період часу між створеннями замовлень задається у діалоговому вікні;

– Signal port – створення сутностей на основі вхідного сигналу;

– MATLAB action – створення сценарію для визначення часу між генераціями. Цей параметр відображається, якщо для методу генерації встановлено значення Time-based. Отже оберемо цей параметр.

– Period – визначення періоду між генерацією замовлень. Параметр відображається, якщо метод генерації має значення Time-based.

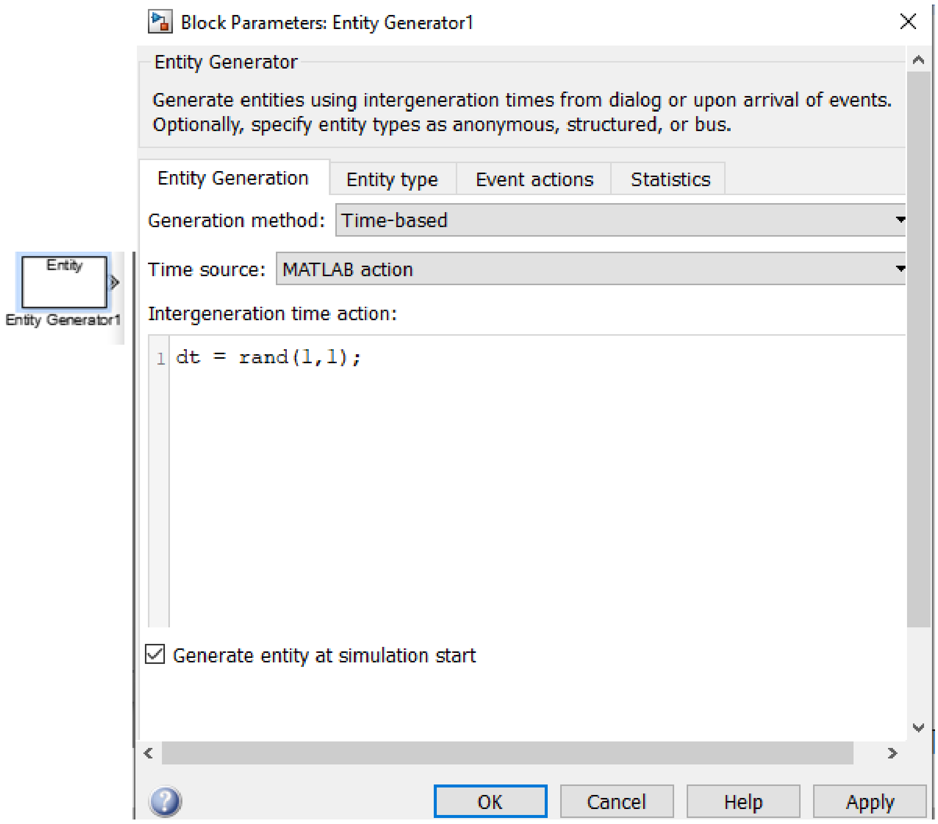


Рис. 12.4. Генератор замовлень

Якщо при цьому джерело часу визначається як MATLAB action, то з’являється вікно Intergeneration Time action (рис. 12.4), у якому необхідно визначити проміжок часу Δt. Оскільки всі наші задачі мають вхідний потік з експоненціальним законом розподілу, то зробимо наступне налаштування:

– задамо значення математичного сподівання $m = 1/\lambda$;

– задамо крок Δt через генератор випадкових чисел на проміжку $(0,1)$ (функція rand()) з поправкою на експоненціальний закон розподілу (рис. 12.5).

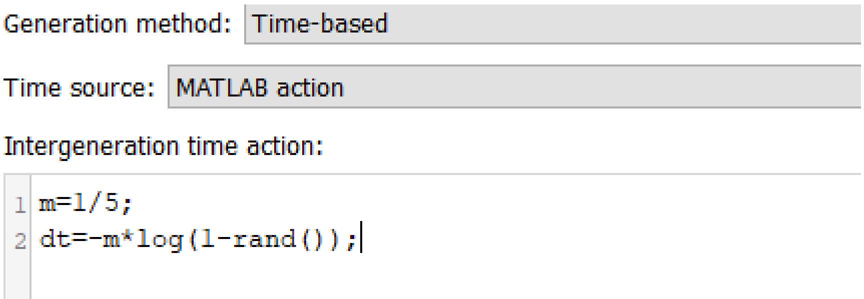


Рис. 12.5. Задання проміжку часу між надходженням замовлень за експоненціальним законом розподілу

– Generate entity at simulation start (створити сутність на початку симуляції) – поставте прапорець, щоб створити об’єкт на початку симуляції.

Наступна закладка *Entity priority* (пріоритет замовлення) – визначення пріоритету генерованого замовлення. Менше значення означає вищий пріоритет.

Закладка *Event actions* (подійні дії) – визначення поведінки замовлення під час певних подій.

В наших задачах всі замовлення рівнозначні, не пов’язані з іншими подіями, тому в цих закладках нічого не будемо змінювати.

Закладка *Statistics* (статистика) дає доступ до статистичних даних блоку:

– Number of entities departed, d – виводить кількість замовлень, які залишили блок;

– Pending entity present in block, pe – вказує, чи є незавершені сутності, які ще не залишили блок.

– Average intergeneration time, w – виводить середній час між створенням замовлень.

Щоб отримати доступ до необхідної статистики, необхідно поставити галочку проти відповідного значення.

2.2. ***Черга*** (рис. 12.6).

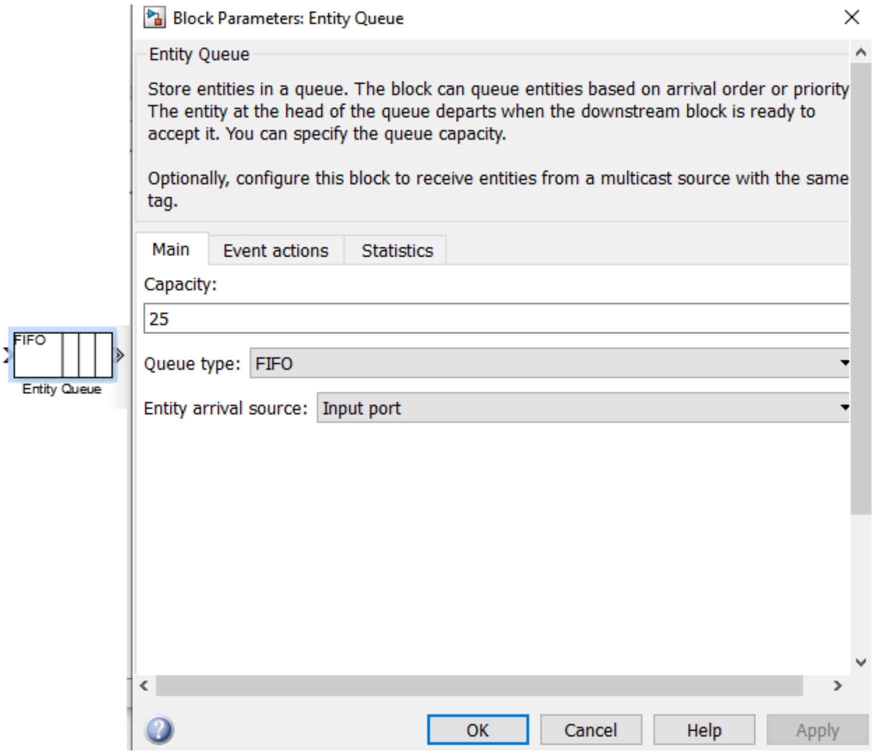


Рис. 12.6. Черга

На головній вкладці необхідно налаштувати наступні параметри:

– Capacity (ємність) – довжина черги, кількість місць. Для системи без відмов довжина (черга нескінченна) встановлюється значення Inf;

– Queue type (тип черги) – FIFO для черги «першим прийшов – першим вийшов», LIFO для черги «останній прийшов – перший вийшов», Priority для черги з пріоритетом.

В наших задачах ми маємо справу із звичайною чергою, тому оберемо тип FIFO.

– Entity arrival source (джерело надходження замовлень). Наші замовлення надходять з Input port.

Оскільки замовлення не пов’язані з подіями, закладку *Event actions* не чіпаємо.

У закладці *Statistics* (статистика) можна зробити налаштування доступу до статистичних даних:

– Number of entities departed, d – кількість замовлень, що вийшли з блоку;

– Number of entities in block, n – кількість замовлень, присутніх у блоці;

– Average wait – середній час очікування замовлення у блоці;

– Average queue length – середня довжина черги.

2.3. Сервіс (рис. 12.7) – обслуговуючий пристрій.

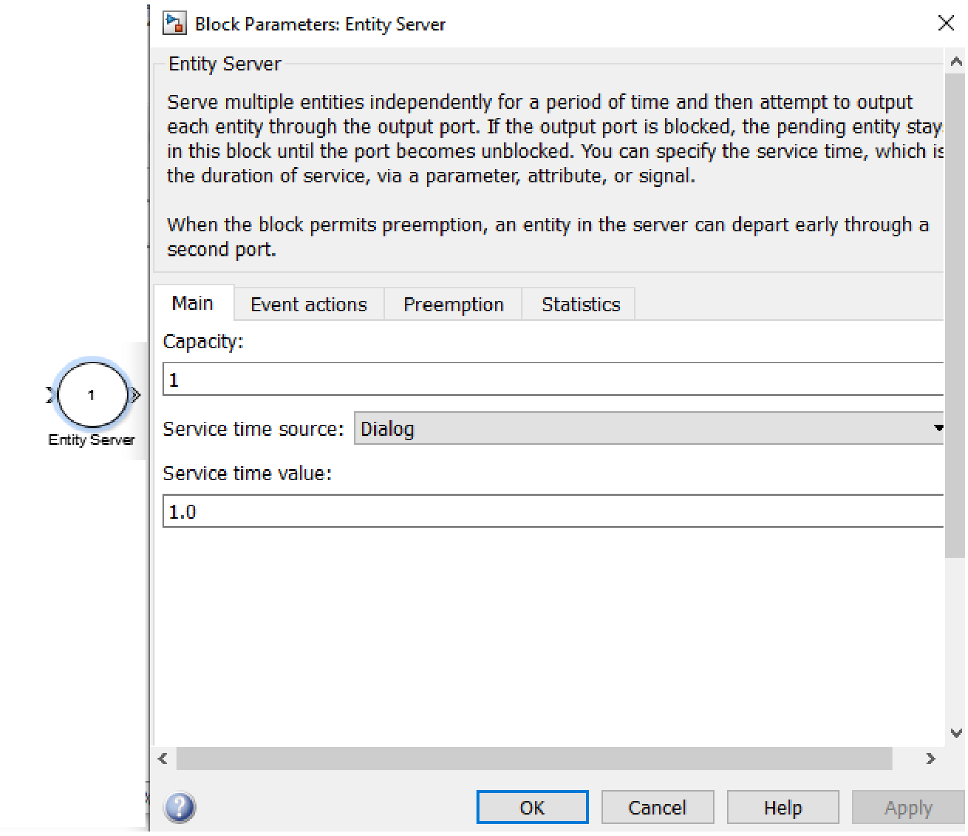


Рис. 12.7. Сервер – блок обслуговування

На головній вкладці налаштовуються наступні параметри:

– Capacity – кількість паралельних обслуговуючих пристроїв;

– Service time source – джерело часу обслуговування. Передбачається декілька варіантів (рис. 12.8).

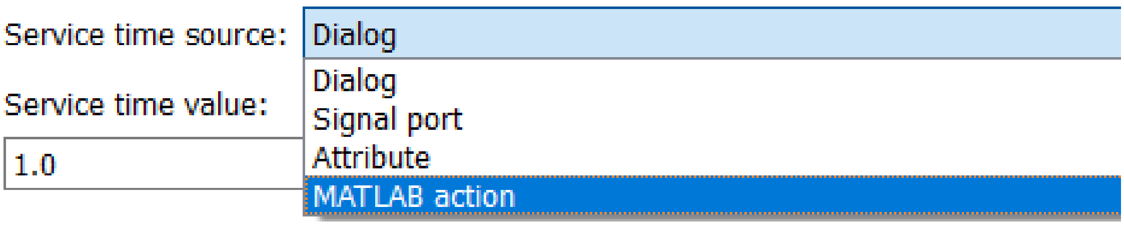


Рис. 12.8. Варіанти визначення часу обслуговування

Якщо встановлено значення Dialog, то у вікні Service time value необхідно задати час обслуговування одного замовлення.

Якщо встановлено значення Attribute, то необхідно вказати назву атрибуту.

В нашому випадку встановимо в якості джерела часу обслуговування значення MATLAB action і пропишемо сценарій обробки замовлень за експоненціальним законом (рис. 12.9).

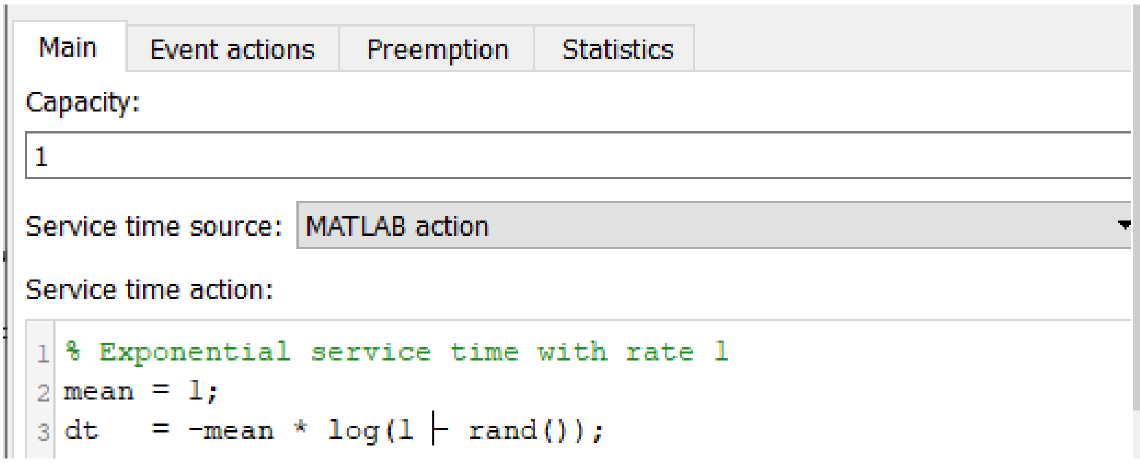


Рис. 12.9. Сценарій обробки замовлення для експоненціального закону

Наступні дві закладки Event actions, Preemption пропустимо, бо замовлення не пов’язані з подіями і не мають пріоритетів.

В закладці Snatistics можна налаштувати доступ до наступних даних:

– Number of entities departed, d – кількість оброблених замовлень;

– Number of entities in block, n – кількість замовлень, що обробляються;

– Pending entiity present in block, pe – вказує, чи є в блоці невідправлені замовлення;

– Number of pending entities, np – кількість недоопрацьованих замовлень;

– Average wait – середній час очікування для замовлень у блоці;

– Utilization – утилізація.

2.4. Термінатор – блок прийому та знищення замовлень (рис. 12.10).

Закладка Event actions дозволяє визначити поведінку в параметрі Event (подія).

Закладка Snatistics містить єдиний пункт:

– Number of entities arrived, a – кількість сутностей, які прибули до блоку.

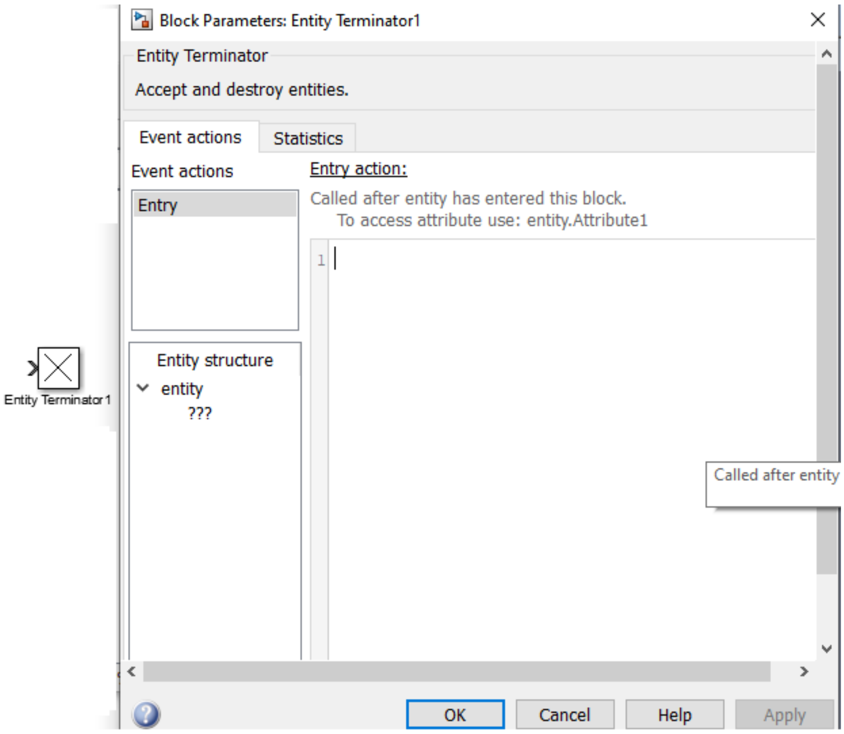


Рис. 12.10. Блок прийому та знищення замовлень

2.5. Додаткові блоки для виводу даних та встановлення часу. Для виводу результатів імітації застосуємо блоки бібліотеки виводу ***Sinks***. Дані можна вивести у віконце Display (після завершення імітації буде відображено останнє значення) або на графік Scope – виводить генеровані дані на протязі імітації (рис. 12.11). Блок Scope має розвинені можливості налаштування, зокрема встановлення заголовків графіків, підпис координатних вісей, можливість виводу декількох графіків в одному вікні тощо.

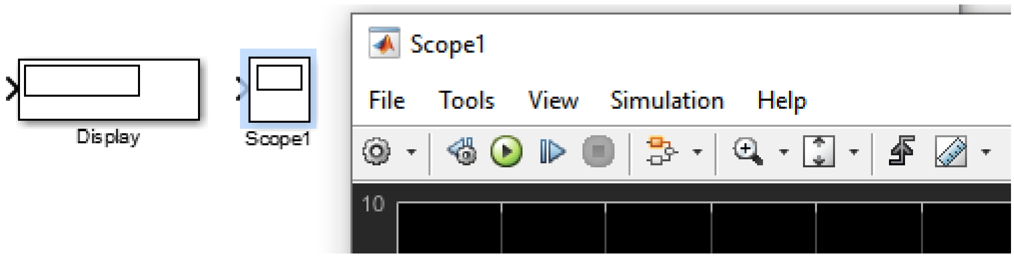


Рис. 12.11. Блоки виводу результатів

Перед запуском моделі на виконання у віконці на верхній панелі меню необхідно встановити час виконання імітації. Час встановлюється у тактах, тому повинен дорівнювати кількості часових одиниць, в яких визначались параметри системи.

Приклад виконання роботи

Інтенсивність потоку замовлень $\lambda = 5$ зам./хв. Система без відмов (має нескінченну чергу). Блок обслуговування складається з двох сервісів, середній час обробки замовлень 2 хв. Скільки замовлень система може обробити за 10 хв? Який середній час перебування замовлення у черзі?

Система має тип: $M/M/2$.

Імітаційна модель СМО, що відповідає умові поставленій задачі, та результати її роботи наведені на рис. 12.12.

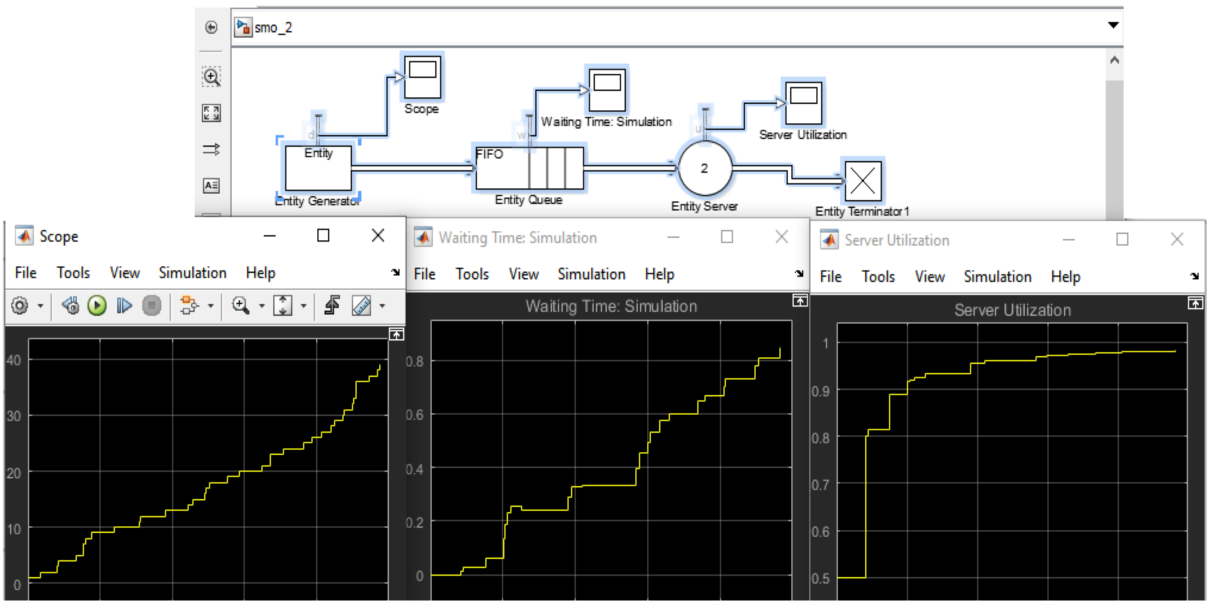


Рис. 12.12. Приклад імітаційної моделі СМО

Щоб покращити роботу системи, доцільно збільшити кількість серверів.

Список літератури

1. Лебєдєв Є.О. Математична статистика : навч. посіб. / Є.О. Лебєдєв, Г.В. Лівінська, І.В. Розора, М.М. Шарапов. – Київ : ВПЦ «Київський університет», 2016. – 160 с.
2. Швець В.Т. Теорія ймовірностей, математична статистика та випадкові процеси : навч. посіб. / В. Т. Швець. – Одеса, 2021. – Електрон. текст. дані: 234 с. Режим доступу https://card-file.ontu.edu.ua/handle/123456789/17874. – Назва з екрана.
3. Горда О.В. Математичне та імітаційне моделювання систем масового обслуговування : навч. посіб. / О.В. Горда, В.М. Міхайленко. – Київ : КНУБА, 2019. – 196 с.
4. Новицький І.В. Випадкові процеси : навч. посіб. / І.В. Новицький, С.А. Ус. – Дніпро : НГУ, 2014. – 132 с.
5. Robert V. Hogg, Joseph W. McKean Introduction to Mathematical Statistics. Pearson Education, Inc. 2019. – 762 p.
6. Bosq D. Mathematical Statistics and Stochastic Processes. Wiley-ISTE. 2012. – 304 p.
7. Leonid Koralov, Yakov G. Sinai Theory of Probability and Random Processes. Princeton University, New Jersey, 2007. – 252 p.

Додаток 1

Параметри законів розподілу

Закон розподілу	Щільність розподілу ймовірність	Мат. сподівання	Дисперсія	Позн.	к-ть п.
Рівномірний	$f(x)=\begin{cases}\dfrac{1}{b-a}, x\in[a,b]\\ 0 \quad\quad , x\notin[a,b]\end{cases}$	$M(x)=\dfrac{a+b}{2}$	$D(x)=\dfrac{(b-a)^2}{12}$	$R(a,b)$	2
Нормальний	$f(x)=\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$	$M(x)=a$	$D(x)=.\sigma^2$	$N(a,\sigma)$	2
Експоненціальний	$f(x)=\begin{cases}\lambda e^{-\lambda x} & x\ge 0\\ 0 & x<0\end{cases}$	$M(x)=1/\lambda$	$D(x)=1/\lambda^2$	$E(\lambda)$	1
Пуассона	$p(k)=\dfrac{\lambda^k}{k!}e^{-\lambda}$	$M(x)=\lambda$	$D(x)=\lambda$	$P(\lambda)$	1
Біноміальний	$p_n(k)=C_n^k p^k q^{n-k}$	$M(x)=np$	$D(x)=npq$	$B(n,p)$	2

Оцінка параметрів методом моментів полягає у прирівнюванні теоретичним параметрам функції розподілу відповідних емпіричних моментів того ж порядку.

Якщо гіпотеза про закон розподілу відхиляється, то він невідомий і що оцінювати, також невідомо. Якщо закон розподілу експоненціальний або Пуассона, то вони залежать від одного параметру λ, який оцінюється через середнє вибіркове.

Для біноміального розподілу кількість спостережень в одному іспиті відома, тож, як правило, оцінюється параметр p – ймовірність: $p=\bar{x}/n$.

Для нормального закону розподілу $a=\bar{x}_в$, $\sigma=\sigma_в$ (іноді в якості оцінки береться виправлене середньоквадратичне відхилення $\sigma=s_в$).

Для рівномірного розподілу параметрів a та b беруться наступні оцінки: $a^*=\bar{x}_в-\sqrt{3}\sigma_в$, $b^*=\bar{x}_в+\sqrt{3}\sigma_в$.

Додаток 2

Таблиця значень функції $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0	0	0,18	0,0714	0,36	0,1406	0,5	0,1915	x	$\Phi(x)$	x	$\Phi(x)$
0,01	0,004	0,19	0,0753	0,37	0,1443	0,51	0,195	0,68	0,2517	0,85	0,3023
0,02	0,008	0,2	0,0793	0,38	0,148	0,52	0,1985	0,69	0,2549	0,86	0,3051
0,03	0,012	0,21	0,0832	0,39	0,1517	0,53	0,2019	0,7	0,258	0,87	0,3078
0,04	0,016	0,22	0,0871	0,4	0,1554	0,54	0,2054	0,71	0,2611	0,88	0,3106
0,05	0,0199	0,23	0,091	0,41	0,1591	0,55	0,2088	0,72	0,2642	0,89	0,3133
0,06	0,0239	0,24	0,0948	0,42	0,1628	0,56	0,2123	0,73	0,2673	0,9	0,3159
0,07	0,0279	0,25	0,0987	0,43	0,1664	0,57	0,2157	0,74	0,2703	0,91	0,3186
0,08	0,0319	0,26	0,1026	0,44	0,17	0,58	0,219	0,75	0,2734	0,92	0,3212
0,09	0,0359	0,27	0,1064	0,45	0,1736	0,59	0,2224	0,76	0,2764	0,93	0,3238
0,1	0,0398	0,28	0,1103	0,46	0,1772	0,6	0,2257	0,77	0,2794	0,94	0,3264
0,11	0,0438	0,29	0,1141	0,47	0,1808	0,61	0,2291	0,78	0,2823	0,95	0,3289
0,12	0,0478	0,3	0,1179	0,48	0,1844	0,62	0,2324	0,79	0,2852	0,96	0,3315
0,13	0,0517	0,31	0,1217	0,49	0,1879	0,63	0,2357	0,8	0,2881	0,97	0,334
0,14	0,0557	0,32	0,1255	0,36	0,1406	0,64	0,2389	0,81	0,291	0,98	0,3365
0,15	0,0596	0,33	0,1293	0,37	0,1443	0,65	0,2422	0,82	0,2939	0,99	0,3389
0,16	0,0636	0,34	0,1331	0,38	0,148	0,66	0,2454	0,83	0,2967	1	0,3413
0,17	0,0675	0,35	0,1368	0,39	0,1517	0,67	0,2486	0,84	0,2995	1,01	0,3438
1,02	0,3461	1,22	0,3883	1,42	0,4222	1,62	0,4474	1,83	0,4664	2,06	0,4803

Продовження табл. дод. 2

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,03	0,3485	1,23	0,3907	1,43	0,4236	1,63	0,4484	1,84	0,4671	2,08	0,4812
1,04	0,3508	1,24	0,3925	1,44	0,4251	1,64	0,4495	1,85	0,4678	2,1	0,4821
1,05	0,3531	1,25	0,3944	1,45	0,4265	1,65	0,4505	1,86	0,468	2,12	0,483
1,06	0,3554	1,26	0,3962	1,46	0,4279	1,66	0,4515	1,87	0,4693	2,14	0,4838
1,07	0,3577	1,27	0,398	1,47	0,4292	1,67	0,4525	1,88	0,4699	2,16	0,4846
1,08	0,3599	1,28	0,3997	1,48	0,4306	1,68	0,4535	1,89	0,4706	2,18	0,4854
1,09	0,3621	1,29	0,4015	1,49	0,4319	1,69	0,4545	1,9	0,4713	2,2	0,4861
1,1	0,3643	1,3	0,4032	1,5	0,4332	1,7	0,4554	1,91	0,4719	2,22	0,4868
1,11	0,3665	1,31	0,4049	1,51	0,4345	1,71	0,4564	1,92	0,4726	2,24	0,4875
1,12	0,3686	1,32	0,4066	1,52	0,4357	1,72	0,4573	1,93	0,4732	2,26	0,4881
1,13	0,3708	1,33	0,4082	1,53	0,437	1,73	0,4582	1,94	0,4738	2,28	0,4887
1,14	0,3729	1,34	0,4099	1,54	0,4382	1,74	0,4591	1,95	0,4744	2,3	0,4893
1,15	0,3749	1,35	0,4115	1,55	0,4394	1,75	0,4599	1,96	0,475	2,32	0,4898
1,16	0,377	1,36	0,4131	1,56	0,4406	1,76	0,4608	1,97	0,4756	2,34	0,4904
1,17	0,379	1,37	0,4147	1,57	0,4418	1,78	0,4625	1,98	0,4761	2,36	0,4909
1,18	0,381	1,38	0,4162	1,58	0,4429	1,79	0,4633	1,99	0,4767	2,38	0,4913
1,19	0,383	1,39	0,4177	1,59	0,4441	1,8	0,4641	2	0,4772	2,4	0,4918
1,2	0,3849	1,4	0,4192	1,6	0,4452	1,81	0,4649	2,02	0,4783	2,42	0,4922
1,21	0,3869	1,41	0,4207	1,61	0,4463	1,82	0,4656	2,04	0,4793	2,44	0,4927
2,46	0,4931	2,58	0,4951	2,7	0,4965	2,82	0,4976	2,94	0,4984	3,6	0,499841
2,48	0,4934	2,6	0,4953	2,72	0,4967	2,84	0,4977	2,96	0,4985	3,8	0,499928

Закінчення табл. дод. 2

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2,5	0,4938	2,62	0,4956	2,74	0,4969	2,86	0,4979	2,98	0,4986	4	0,499968
2,52	0,4941	2,64	0,4959	2,76	0,4971	2,88	0,498	3	0,49865	4,5	0,499997
2,54	0,4945	2,66	0,4961	2,78	0,4973	2,9	0,4981	3,2	0,49931	5	0,499997
2,56	0,4948	2,68	0,4963	2,8	0,4974	2,92	0,4982	3,4	0,49966	$x > 5$	0,5

Під час знаходження значень необхідно враховувати, що $\Phi(-x) = -\Phi(x)$ – непарна функція.

Додаток 3

Таблиця значень $t_\gamma = t(\gamma, n)$

n / γ	0,95	0,99	0,999	n / γ	0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2.064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3.464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,663	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

Додаток 4

Таблиця значень $q = q(\gamma, n)$

n / γ	0,95	0,99	0,999	n / γ	0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0,32	0,49	0,73
7	0,92	1,62	2,98	30	0,28	0,43	0,63
8	0,80	1,38	2,42	35	0,26	0,38	0,56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1,08	1,80	45	0,22	0,32	0,46
11	0,59	0.98	1,60	50	0,21	0,30	0,43
12	0,55	0,90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0,78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,198	0,27
17	0,42	0,66	1,01	150	0,115	0,160	0,211
18	0,40	0,63	0,96	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

Додаток 5

Критичні точки розподілу χ^2

k /α	0,01	0,025	0,05	0,95	0,975	0,99
1	6,63490	5,02389	3,84146	0,00393	0,00098	0,00016
2	9,21034	7,37776	5,99146	0,10259	0,05064	0,02010
3	11,34487	9,34840	7,81473	0,35185	0,21580	0,11483
4	13,2767	11,14329	9,48773	0,71072	0,48442	0,29711
5	15,08627	12,8325	11,0705	1,14548	0,83121	0,55430
6	16,81189	14,44938	12,59159	1,63538	1,23734	0,87209
7	18,47531	16,01276	14,06714	2,16735	1,68987	1,23904
8	20,09024	17,53455	15,50731	2,73264	2,17973	1,64650
9	21,66599	19,02277	16,91898	3,32511	2,70039	2,08790
10	23,20925	20,48318	18,30704	3,94030	3,24697	2,55821
11	24,72497	21,92005	19,67514	4,57481	3,81575	3,05348
12	26,21697	23,33666	21,02607	5,22603	4,40379	3,57057
13	27,68825	24,7356	22,36203	5,89186	5,00875	4,10692
14	29,14124	26,11895	23,68479	6,57063	5,62873	4,66043
15	30,57791	27,48839	24,99579	7,26094	6,26214	5,22935
16	31,99993	28,84535	26,29623	7,96165	6,90766	5,81221
17	33,40866	30,19101	27,58711	8,67176	7,56419	6,40776
18	34,80531	31,52638	28,86930	9,39046	8,23075	7,01491
19	36,19087	32,85233	30,14353	10,11701	8,90652	7,63273
20	37,56623	34,16961	31,41043	10,85081	9,59078	8,26040
21	38,93217	35,47888	32,67057	11,59131	10,2829	8,89720
22	40,28936	36,78071	33,92444	12,33801	10,98232	9,54249
23	41,63840	38,07563	35,17246	13,09051	11,68855	10,19572
24	42,97982	39,36408	36,41503	13,84843	12,40115	10,85636
25	44,31410	40,64647	37,65248	14,61141	13,11972	11,52398
26	45,64168	41,92317	38,88514	15,37916	13,84391	12,19815
27	46,96294	43,19451	40,11327	16,15140	14,57338	12,87850
28	48,27824	44,46079	41,33714	16,92788	15,30786	13,56471

Закінчення табл. дод. 5

29	49,58788	45,72229	42,55697	17,70837	16,04707	14,25645
30	50,89218	46,97924	43,77297	18,49266	16,79077	14,95346
31	52,19139	48,23189	44,98534	19,28057	17,53874	15,65546
32	53,48577	49,48044	46,19426	20,07191	18,29076	16,36222
33	54,77554	50,72508	47,39988	20,86653	19,04666	17,07351
34	56,06091	51,96600	48,60237	21,66428	19,80625	17,78915
35	57,34207	53,20335	49,80185	22,46502	20,56938	18,50893
36	58,61921	54,43729	50,99846	23,26861	21,33588	19,23268
37	59,89250	55,66797	52,19232	24,07494	22,10563	19,96023
38	61,16209	56,89552	53,38354	24,8839	22,87848	20,69144
39	62,42812	58,12006	54,57223	25,69539	23,65432	21,42616
40	63,69074	59,34171	55,75848	26,5093	24,43304	22,16426
41	64,95007	60,56057	56,94239	27,32555	25,21452	22,90561
42	66,20624	61,77676	58,12404	28,14405	25,99866	23,65009
43	67,45935	62,99036	59,30351	28,96472	26,78537	24,39760
44	68,70951	64,20146	60,48089	29,78748	27,57457	25,14803
45	69,95683	65,41016	61,65623	30,61226	28,36615	25,90127
46	71,20140	66,61653	62,82962	31,43900	29,16005	26,65724
47	72,44331	67,82065	64,00111	32,26762	29,95620	27,41585
48	73,68264	69,02259	65,17077	33,09808	30,75451	28,17701
49	74,91947	70,22241	66,33865	33,93031	31,55492	28,94065
50	76,15389	71,42020	67,50481	34,76425	32,35736	29,70668

Додаток 6

Критичні точки розподілу Стьюдента

k / α	значення α для двосторонньої критичної області					
	0,1	0,05	0,02	0,01	0,002	0.001
1	6.31	12.7	31.82	63.7	318.3	637.0
2	2.92	4.30	6.97	9.92	22.33	31.6
3	2.35	3.18	4.54	5.84	10.22	12.9
4	2.13	2.78	3.75	4.60	7.17	8.61
5	2.01	2.57	3.37	4.03	5.89	6.86
6	1.94	2.45	3.14	3.71	5.21	5.96
7	1.89	2.36	3.00	3.50	4.79	5.40
8	1.86	2.31	2.90	3.36	4.50	5.04
9	1.83	2.26	2.82	3.25	4.30	4.78
10	1.81	2.23	2.76	3.17	4.14	4.59
11	1.80	2.20	2.72	3.11	4.03	4.44
12	1.78	2.18	2.68	3.05	3.93	4.32
13	1.77	2.16	2.65	3.01	3.85	4.22
14	1.76	2.14	2.62	2.98	3.79	4.14
15	1.75	2.13	2.60	2.95	3.73	4.07
16	1.75	2.12	2.58	2.92	3.69	4.01
17	1.74	2.11	2.57	2.90	3.65	3.95
18	1.73	2.10	2.55	2.88	3.61	3.92
19	1.73	2.09	2.54	2.86	3.58	3.88
20	1.73	2.09	2.53	2.85	3.55	3.85
11	1.80	2.20	2.72	3.11	4.03	4.44
21	1.72	2.08	2.52	2.83	3.53	3.82
22	1.72	2.07	2.51	2.82	3.51	3.79
23	1.71	2.07	2.50	2.81	3.59	3.77
24	1.71	2.06	2.49	2.80	3.47	3.74
25	1.71	2.06	2.49	2.79	3.45	3.72
26	1.71	2.06	2.48	2.78	3.44	3.71
27	1.71	2.05	2.47	2.77	3.42	3.69
28	1.70	2.05	2.46	2.76	3.40	3.66
29	1.70	2.05	2.46	2.76	3.40	3.66
30	1.70	2.04	2.46	2.75	3.39	3.65
40	1.68	2.02	2.42	2.70	3.31	3.55
60	1.67	2.00	2.39	2.66	3.23	3.46
120	1.66	1.98	2.36	2.62	3.17	3.37
∞	1.64	1.96	2.33	2.58	3.09	3.29
	0,05	0,025	0,01	0,005	0,001	0,0005
	значення α для односторонньої критичної області					

Додаток 7

Критичні точки розподілу Фішера – Снедекора
k_1 (k_2)– кількість степенів свободи вибірки з більшою (меншою) дисперсією

k_1 / k_2	$\alpha = 0{,}01$											
	1	2	3	4	5	6	7	8	9	10	11	12
1	10,5	4999	5403	5625	5764	5889	5928	5981	6022	6056	6082	6106
2	98,49	99,01	90,17	99,25	99,30	99,33	99,34	99,36	99,38	99,40	99,41	99,42
3	34,12	30,81	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,13	27,05
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,54	14,45	14,37
5	16,26	13,27	12,06	11,39	10,97	10,67	10,45	10,27	10,15	10,05	9,96	9,89
6	13,74	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72
7	12,25	9,55	8,45	7,85	7,46	7,46	7,00	6,84	6,71	6,62	6,54	6,47
8	11,26	8,65	7,59	7,01	6,63	6,37	6,19	6,03	5,91	5,82	5,74	5,67
9	10,56	8,02	6,99	6,42	6,06	5,80	5,62	5,47	5,35	5,26	5,18	5,11
10	10,04	7,56	6,55	5,99	5,64	5,39	5,21	5,06	4,95	4,85	4,78	4,71
11	9,85	7,20	6,22	5,67	5,32	5,07	4,88	4,74	4,63	4,54	4,46	4,40
12	9,33	6,93	5,95	5,41	5,06	4,82	4,65	4,50	4,39	4,30	4,22	4,16
13	9,07	6,70	5,74	5,20	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96
14	8,86	6,51	5,56	5,03	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67
16	8,53	6,23	5,29	4,77	4,44	2 40	4,03	3,89	3,78	3,69	3,61	3,55
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52	3,45
k_1 / k_2	$\alpha = 0{,}05$											
	1	2	3	4	5	6	7	8	9	10	11	12
1	161	200	216	225	230	234	237	239	241	242	243	244
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,40	19,41
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,76	8,74
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70	4,68
6	5,99	5,14	4,76	4,35	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60	3,57
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31	3,28
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10	3,07
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94	2,91
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82	2,79
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72	2,69
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,63	2,60
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56	2,53
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,51	2,48
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,41	2,38

Додаток 8

Квантилі нормального розподілу u_p

(в цій таблиці перед всіма значеннями необхідно поставити знак мінус)

%	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0	-	3,090	2,878	2,748	2,652	2,576	2,512	2,457	2,409	2,366
1	2,326	2,290	2,257	2,226	2.197	2,170	2,144	2,120	2.097	2,075
2	2,054	2,034	2,014	1,995	1,977	1,960	1,943	1,927	1,911	1,896
3	1,881	1,866	1,852	1,838	1,825	1,812	1,799	1,787	1,774	1,762
4	1,751	1,739	1,728	1,717	1,706	1,695	1,685	1,675	1,665	1,655
5	1,645	1,635	1,626	1,616	1,607	1,598	1,589	1,580	1,572	1,563
6	1,555	1,546	1,538	1,530	1,522	1,514	1,506	1,499	1,491	1,483
7	1,476	1,468	1,461	1,454	1,447	1,440	1,433	1,426	1,419	1,412
8	1,405	1,398	1,392	1,385	1,379	1,372	1,366	1,359	1,353	1,347
9	1,341	1,335	1,329	1,323	1,317	1,311	1,305	1,299	1,293	1,287
10	1,282	1,276	1,270	1,265	1,259	1,254	1,248	1,243	1,237	1,232
11	1,227	1,221	1,216	1,211	1,206	1,200	1,195	1,190	1,185	1,180
12	1,175	1,170	1,165	1,160	1,155	1,150	1,146	1,141	1,136	1,131
13	1,126	1,122	1,117	1,112	1,108	1,103	1,098	1,094	1,089	1,085
14	1,080	1,076	1,071	1,067	1,063	1,058	1,054	1,049	1,045	1,041
15	1,036	1,032	1,028	1,024	1,019	1,015	1,011	1,007	1,003	0,999
16	0,994	0,990	0,986	0,982	0,978	0,974	0,970	0,966	0,962	0,958
17	0,954	0,950	0,946	0,942	0,938	0,935	0,931	0,927	0,923	0,919
18	0,915	0,912	0,908	0,904	0,900	0,896	0,893	0,889	0,885	0,882
19	0,878	0,874	0,871	0,867	0,863	0,860	0,856	0.852	0,849	0,845
20	0,842	0,838	0,834	0,831	0,827	0,824	0,820	0,817	0,813	0,810
21	0,806	0,803	0,800	0,796	0,793	0,789	0,786	0,782	0,779	0,776
22	0,772	0,769	0,765	0,762	0,759	0,755	0,752	0,749	0,745	0,742
23	0,739	0,736	0,732	0,729	0,726	0,722	0,719	0,716	0,713	0,710
24	0,706	0,703	0,700	0,697	0,693	0,690	0,687	0,684	0,681	0,678
25	0,674	0,671	0,668	0,665	0,662	0,659	0,656	0,653	0,650	0,646
26	0,643	0,640	0,637	0,634	0,631	0,628	0,625	0,622	0,619	0,616
27	0,613	0,610	0,607	0,604	0,601	0,598	0,595	0,592	0,589	0,586
28	0,583	0,580	0,577	0,574	0,571	0,568	0,565	0,562	0,559	0,556
29	0,553	0,550	0,548	0,545	0,542	0,539	0,536	0,533	0,530	0,527
30	0,524	0,522	0,519	0,516	0,513	0,510	0,507	0,504	0,502	0,499
31	0,496	0,493	0,490	0,487	0,485	0,482	0,479	0,476	0,473	0,470
32	0,468	0,465	0,462	0,459	0,457	0,454	0,451	0,448	0,445	0,443
33	0,440	0,437	0,434	0,432	0,429	0,426	0,423	0,421	0,418	0,415
34	0,412	0,410	0,407	0,404	0,402	0,399	0,396	0,393	0,391	0,388
35	0,385	0,383	0,380	0,377	0,375	0,372	0,369	0,366	0,364	0,361
36	0,358	0,356	0,353	0,350	0,348	0,345	0,342	0,340	0,337	0,335
37	0,332	0,329	0,327	0,324	0,321	0,319	0,316	0,313	0,311	0,308
38	0,305	0,303	0,300	0,298	0,295	0,292	0,290	0,287	0,285	0,282
39	0,279	0,277	0,274	0,272	0,269	0,266	0,264	0,261	0,259	0,256
40	0,253	0,251	0,248	0,246	0,243	0,240	0,238	0,235	0,233	0,230
41	0,228	0,225	0,222	0,220	0,217	0,215	0,212	0,210	0,207	0,204
42	0,202	0,199	0,197	0,194	0,192	0,189	0,187	0,184	0,181	0,179
43	0,176	0,174	0,171	0,169	0,166	0,164	0,161	0,159	0,156	0,154
44	0,151	0,148	0,146	0,143	0,141	0,138	0,136	0,133	0,131	0,128
45	0,126	0,123	0,121	0,118	0,116	0,113	0,111	0,108	0,105	0,103

Продовження табл. дод. 8

46	0,100	0,098	0,095	0,093	0,090	0,088	0,085	0,083	0,080	0,079
47	0,075	0,073	0,070	0,068	0,065	0,063	0,060	0,058	0,055	0,053
48	0,050	0,048	0,045	0,043	0,040	0,038	0,035	0,033	0,030	0,028
49	0,025	0,023	0,020	0,018	0,015	0,013	0,010	0,008	0,005	0,003
(в цій таблиці всі значення додатні)										
%	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
50	-	0,003	0,005	0,008	0,010	0,013	0,015	0,010	0,013	0,015
51	0,025	0,028	0,030	0,033	0,035	0,038	0,040	0,035	0,038	0,040
52	0,050	0,053	0,055	0,058	0,060	0,063	0,065	0,060	0,063	0,065
53	0,075	0,079	0,080	0,083	0,085	0,088	0,090	0,085	0,088	0,090
54	0,100	0,103	0,105	0,108	0,111	0,113	0,116	0,111	0,113	0,116
55	0,126	0,128	0,131	0,133	0,136	0,138	0,141	0,136	0,138	0,141
56	0,151	0,154	0,156	0,159	0,161	0,164	0,166	0,161	0,164	0,166
57	0,176	0,179	0,181	0,184	0,187	0,189	0,192	0,187	0,189	0,192
58	0,202	0,204	0,207	0,210	0,212	0,215	0,217	0,212	0,215	0,217
59	0,228	0,230	0,233	0,235	0,238	0,240	0,243	0,238	0,240	0,243
60	0,253	0,256	0,259	0,261	0,264	0,266	0,269	0,264	0,266	0,269
61	0,279	0,282	0,285	0,287	0,290	0,292	0,295	0,290	0,292	0,295
62	0,305	0,308	0,311	0,313	0,316	0,319	0,321	0,316	0,319	0,321
63	0,332	0,335	0,337	0,340	0,342	0,345	0,348	0,342	0,345	0,348
64	0,358	0,361	0,364	0,366	0,369	0,372	0,375	0,369	0,372	0,375
65	0,385	0,388	0,391	0,393	0,395	0,399	0,402	0,395	0,399	0,402
66	0,412	0,415	0,418	0,421	0,423	0,426	0,429	0,423	0,426	0,429
67	0,440	0,443	0,445	0,448	0,451	0,454	0,457	0,451	0,454	0,457
68	0,468	0,470	0,473	0,476	0,479	0,482	0,485	0,479	0,482	0,485
69	0,496	0,499	0,502	0,504	0,507	0,510	0,513	0,507	0,510	0,513
70	0,524	0,527	0,530	0,533	0,536	0,539	0,542	0,536	0,539	0,542
71	0,553	0,556	0,559	0,562	0,565	0,568	0,571	0,565	0,568	0,571
72	0,583	0,586	0,589	0,592	0,595	0,598	0,601	0,595	0,598	0,601
73	0,613	0,616	0,619	0,622	0,625	0,628	0,631	0,625	0,628	0,631
74	0,643	0,646	0,650	0,653	0,656	0,659	0,662	0,656	0,659	0,662
75	0,674	0,678	0,681	0,684	0,687	0,690	0,693	0,687	0,690	0,693
76	0,706	0,710	0,713	0,716	0,719	0,722	0,726	0,719	0,722	0,726
77	0,739	0,742	0,745	0,749	0,752	0,755	0,759	0,762	0,765	0,769
78	0,772	0,776	0,779	0,782	0,786	0,789	0,793	0,796	0,800	0,803
79	0,806	0,810	0,813	0,817	0,820	0,824	0,827	0,831	0,834	0,838
80	0,842	0,845	0,849	0,852	0,856	0,860	0,863	0,867	0,871	0,874
81	0,878	0,882	0,885	0,889	0,893	0,896	0,900	0,904	0,908	0,913
82	0,915	0,919	0,923	0,927	0,931	0,935	0,938	0,942	0,946	0,950
83	0,954	0,958	0,962	0,966	0,970	0,974	0,978	0,982	0,986	0,990
84	0,994	0,999	1,003	1,007	1,011	1,015	1,019	1,024	1,028	1,032
85	1,036	1,041	1,045	1,049	1,054	1,058	1,063	1,067	1,071	1,076
86	1,080	1,085	1,089	1,094	1,098	1,103	1,108	1,112	1,117	1,122
87	1,126	1,131	1,136	1,141	1,146	1,150	1,155	1,160	1,165	1,170
88	1,175	1,180	1,185	1,190	1,195	1,200	1,206	1,211	1,216	1,221
89	1,227	1,232	1,237	1,243	1,248	1,254	1,259	1,265	1,270	1,276
90	1,282	1,287	1,293	1,299	1,305	1,311	1,317	1,323	1,329	1,335
91	1,341	1,347	1,353	1,359	1,366	1,372	1,379	1,385	1,392	1,398
92	1,405	1,412	1,419	1,426	1,433	1,440	1,447	1,454	1,461	1,468
93	1,476	1,483	1,491	1,499	1,506	1,514	1,522	1,530	1,538	1,546

94	1,555	1,563	1,572	1,580	1,589	1,589	1,607	1,616	1,626	1,635
95	1,645	1,655	1,665	1,675	1,685	1,695	1,706	1,717	1,728	1,739
96	1,751	1,762	1,774	1,787	1,799	1,812	1,825	1,838	1,852	1,866
97	1,881	1,896	1,911	1,927	1,943	1,960	1,977	1,995	2,014	2,034
98	2,054	2,075	2,097	2,120	2,144	2,170	2,197	2,226	22,257	2,290
99	2,326	2,366	2,409	2,457	2,512	2,576	2,652	2,748	2,878	3,090

ДЛЯ НОТАТОК

ДЛЯ НОТАТОК

Навчально-методичне видання

МАТЕМАТИЧНА СТАТИСТИКА ТА ВИПАДКОВІ ПРОЦЕСИ

Методичні вказівки та завдання
до проведення практичних занять № 1 – 12
для здобувачів першого (бакалаврського) рівня вищої освіти
за спеціальностями 122 «Комп'ютерні науки»,
126 «Інформаційні системи та технології»

Укладач **Горда** Олена Володимирівна

Випусковий редактор *Л. С. Тавлуй*
Комп'ютерне верстання *К. А. Мавроді*

Підписано до друку 16.06.2025. Формат 60 х $84_{1/16}$
Ум. друк. арк. 7,44. Обл.-вид. арк. 8,0.
Електронний документ. Вид. № 42/ІІІ-25

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