Міністерство освіти і науки України
Національний університет водного господарства та природокористування
Кафедра комп’ютерних наук та прикладної математики

04-01-76М

МЕТОДИЧНІ ВКАЗІВКИ

до виконання лабораторних робіт з навчальної дисципліни
«Математична статистика та аналіз даних»
для здобувачів вищої освіти першого (бакалаврського) рівня
за освітньою програмою «Прикладна математика»
спеціальності 113 Прикладна математика
денної, заочної форм навчання
Частина 1

Рекомендовано науково-методичною радою з якості ННІ АКОТ
Протокол № 8
від «19» червня 2023 року

Рівне – 2023

Методичні вказівки до виконання лабораторних робіт з навчальної дисципліни «Математична статистика та аналіз даних» для здобувачів вищої освіти першого (бакалаврського) рівня за освітньою програмою «Прикладна математика» спеціальності 113 Прикладна математика денної, заочної форм навчання. Частина 1 [Електронне видання] / Прищепа О. В. – Рівне : НУВГП, 2023. – 58 с.

Укладач:
Прищепа О. В. – к.ф.-м.н, доцент кафедри комп'ютерних наук та прикладної математики.

Відповідальний за випуск:
Турбал Ю. В. – д.т.н., професор, завідувач кафедри комп'ютерних наук та прикладної математики.

Керівник груп забезпечення:
Прищепа О. В., к.ф.-м.н, доцент кафедри комп'ютерних наук та прикладної математики

ЗМІСТ

ВСТУП

Математична статистика та аналіз даних є важливим курсом циклу фундаментальної підготовки сучасного фахівця з прикладної математики. Студенти вивчають методи збору та групування статистичних даних, отриманих в результаті спостережень, опановують методи аналізу та інтерпретації даних з метою прийняття обґрунтованих висновків та прийняття рішень. Цей курс дає студентам знання та навички для роботи зі статистичними даними, включаючи засоби випадкового відбору, описову статистику, ймовірнісні розподіли, перевірку статистичних гіпотез, регресійний аналіз, дисперсійний аналіз та інші методи.

Лабораторна робота №1

Тема: Початкова обробка статистичних даних.

Теоретичні відомості.

Позначимо ознаку, яка вивчається ξ. **Генеральною сукупністю** називають сукупність всіх значень деякої ознаки досліджуваних об'єктів. **Вибірковою сукупністю** називають сукупність випадково відібраних об'єктів з генеральної сукупності. **Об'ємом сукупності** (вибіркової або генеральної) називають число об'єктів цієї сукупності.

Вибіркою будемо називати сукупність випадкових величин ξ_i, $i = \overline{1,n}$, що спостерігалися в експерименті. Величини ξ_i – це елементи вибірки, n – об'єм вибірки.

Числа x_1, x_2, …, x_n, які отримують на практиці при n-кратному повторенні експерименту представляють собою **реалізацію вибірки** $\xi_1, \xi_2, ..., \xi_n$.

Розташуємо величини $x_1, x_2, \ldots, x_n$ у порядку зростання:

$$x_{(1)} \le x_{(2)} \le ... \le x_{(n)}, \tag{1}$$

де $x_{(1)} = \min(x_1, ..., x_n)$, $x_{(2)}$ – друга за величиною серед $x_1, x_2, \ldots, x_n$, … , $x_{(n)} = \max(x_1, ..., x_n)$.

Впорядкована за зростанням послідовність (1) називається **варіаційним рядом вибірки**. Елемент $x_{(i)}$, $i = \overline{1,n}$ називають ще і-тим членом варіаційного ряду або варіантою**.**

Кожна варіанта може зустрічатися у вибірці певну кількість разів. Число, яке показує скільки разів зустрічалась варіанта $x_{(i)}$, називається **частотою** і позначається n_i.

Поряд з частотою використовується величина w_i – **відносна частота**, яка визначається за формулою:

$$\frac{n_i}{n} = w_i \; .$$

Відносна частота показує долю конкретної варіанти у вибірці. Іноді відносну частоту подають у відсотках.

Дискретним **статистичним розподілом вибірки** називається відповідність між варіантами та їх частотами або відносними частотами, який можна подати у вигляді таблиць:

дискретний статистичний розподіл частот

Значення	y_1	y_2	…	y_k
Частота	n_1	n_2	…	n_k

дискретний статистичний розподіл відносних частот

Значення	y_1	y_2	…	y_k
Частота	w_1	w_2	…	w_k

де y_1, y_2, …, y_k – різні по значенню дані варіаційного ряду $x_{(1)} \le x_{(2)} \le ... \le x_{(n)}$, n_i – відносна частота значення y_i, w_i – відносна частота значення y_i, $i = 1,2,...,k$, $\sum_i n_i = n$, $\sum_i w_i = 1$.

У випадку, коли обсяг дискретного варіаційного ряду досить великий, його перетворюють на інтервальний. Тобто визначають інтервал, який включає всі варіанти дискретного варіаційного ряду, ділять його на певне число проміжків $[z_0, z_1]$, $(z_1, z_2]$, …,$(z_{m-1}, z_m]$ і підраховують частоту n_i попадання варіант в кожен проміжок.

Інтервальний статистичний розподіл частот (відносних частот) вибірки можна подати у вигляді таблиць:

інтервальний статистичний розподіл частот

$(z_{i-1}, z_i]$	$[z_0, z_1]$	$(z_1, z_2]$	…	$(z_{m-1}, z_m]$
n_i	n_1	n_2	…	n_m

інтервальний статистичний розподіл відносних частот:

$(z_{i-1}, z_i]$	$[z_0, z_1]$	$(z_1, z_2]$	…	$(z_{m-1}, z_m]$
w_i	w_1	w_2	…	w_m

де $\sum_i n_i = n$, $\sum_i w_i = 1$.

Найчастіше проміжки беруть однакової ширини. Кількість проміжків та їх ширина визначається, як правило з фізичного змісту задачі. Число інтервалів можна визначити також за емпіричною формулою Стерджеса

$$m = 1 + 3.322 \cdot \lg(n). \quad (2)$$

Ширина інтервалу групування вибіркових даних h визначається залежністю:

$$h = \frac{R}{m},$$

де $R = x_{\max} - x_{\min}$, $x_{\max}$, $x_{\min}$ – відповідно найбільше і найменше значення величини ξ у вибірці.

Початок першого проміжку z_0 і кінець останнього проміжку z_m не обов'язково збігаються з $x_{\min}$ та $x_{\max}$ відповідно. Вимагається лише, щоб $z_0 \le x_{\min}$ і $x_{\max} \le z_m$. Якщо $z_0 = x_{\min}$, то вважають, що в інтервальному варіаційному ряді перший проміжок є відрізком. Початок першого інтервалу можна визначати наступним чином: $z_0 = x_{\min} - h/2$, його кінець $z_1 = z_0 + h$. Кінець першого інтервалу є початком наступного і так далі. Побудова проміжків закінчується тоді, коли найбільше значення варіанти включене в деякий проміжок. Внаслідок заокруглень, остаточна кількість інтервалів може відрізнятися від числа інтервалів, які були розраховані за формулою (2).

Накопичувальною частотою (відносною частотою) називається

$$n_j^{/} = \sum_{i=1}^{j} n_i \quad (w_j^{/} = \sum_{i=1}^{j} w_i).$$

Полігоном частот (**відносних частот**) називають ламану лінію, яка відображає залежність між варіантою і частотою (відносною частотою) та утворюється шляхом з'єднання прямими лініями точок (y_1, n_1), (y_2, n_2), …, (y_k, n_k) $\big((y_1, w_1), (y_2, w_2), ..., (y_k, w_k)\big)$.

Гістограмою частот (відносних частот) називають графічне зображення варіаційного ряду у вигляді прямокутників, основами яких є інтервали довжиною h, а висоти дорівнюють n_i / h (w_i / h).

Найбільш універсальною характеристикою випадкової величини є **функція розподілу** $F(x)$. Позначимо через $\nu_n(x)$ випадкову величину, яка дорівнює числу елементів вибірки ξ_1, $\xi_2, \ldots, \xi_n$, значення яких не перевищує x: $\nu_n(x) = \left|\{j : \xi_j < x\}\right|$, де $|\cdot|$ – число елементів скінченної множини.

Функція, яка задається рівністю

$$F_n^*(x) = \frac{\nu_n(x)}{n}$$

та визначає відносну частоту події $\{\xi < x\}$, називається **емпіричною функцією розподілу (функцією розподілу вибірки)**. Разом з цим функцію розподілу $F_\xi(x)$ випадкової величини ξ, що спостерігається, називають **теоретичною функцією розподілу.** Емпіричну функцію розподілу $F_n^*(x)$ можна використовувати як оцінку функції $F_\xi(x)$.

Якщо всі компоненти вибірки різні, функція розподілу визначається формулою:

$$F_n^*(x) = P\{\xi < x\} = \begin{cases} 0, & x \le x_{(1)} \\ \ldots & \\ \frac{i}{n}, & x_{(i)} < x \le x_{(i+1)} \\ \ldots & \\ 1, & x > x_{(n)} \end{cases}$$

Якщо деякі елементи вибірки повторюються, то функція розподілу матиме вигляд:

$$F_n^*(x) = \begin{cases} 0, & x \le y_1 \\ \ldots & \\ \frac{n_1 + \ldots + n_i}{n}, & y_i < x \le y_{i+1} \\ \ldots & \\ 1, & x > y_k \end{cases}$$

Якщо маємо в своєму розпорядженні згрупований інтервальний статистичний ряд, то можемо лише наближено побудувати емпіричну функцію розподілу $F_n^*(x)$:

$$F_n^*(x) = \begin{cases} 0, & x \le z_0; \\ F_n^{*(1)}(x), & z_0 < x \le z_1; \\ F_n^{*(2)}(x), & z_1 < x \le z_2; \\ \ldots \\ F_n^{*(m)}(x), & z_{m-1} < x \le z_m; \\ 1, & x > z_m; \end{cases}$$

де $F_n^{*(i)}(x) = \dfrac{w_i}{z_i - z_{i-1}}(x - z_{i-1}) + w_{i-1}^{/}$, $w_0^{/} = 0$, $i = \overline{1, m}$.

Приклад 1. Дана вибірка:

1	2	1	2	2	3	4	5	4	5	3	3	2	0	5	3
2	3	4	0	5	1	4	4	3	4	1	3	5	3	4	

1) Записати статистичний розподіл вибірки;
2) обчислити відносні частоти і накопичувані частоти;
3) побудувати полігон частот, полігон відносних частот;
4) знайти емпіричну функцію розподілу та побудувати її графік.

Розв’язання.

Побудуємо статистичний розподіл вибірки. Обсяг вибірки $n = 31$. Див. стовпець 1,2 табл.1. Обчислюємо відносні частоти та накопичувані відносні частоти. Всі дані заносимо до табл.1, стовпець 3, 4 :

Таблиця 1

Розрахунки для прикладу 1

1	**2**	**3**	**4**
x_i	n_i	w_i	$w_i^{/}$
0	2	0,06	0,06
1	4	0,13	0,19
2	5	0,16	0,35
3	8	0,26	0,61
4	7	0,23	0,84
5	5	0,16	1

Будуємо графіки полігону частот (рис. 1) та полігону відносних частот (рис. 2):

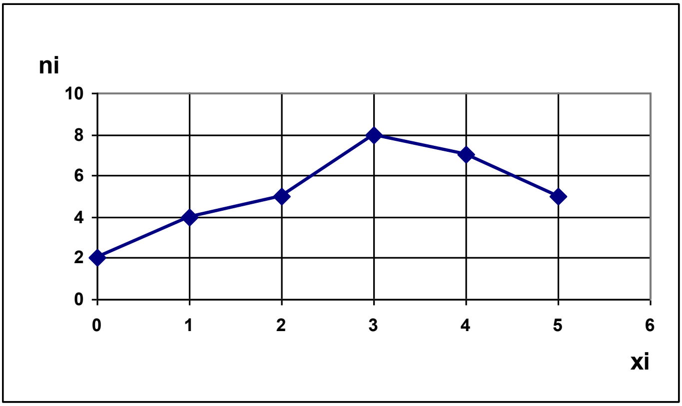

Рис. 1. Полігон частот

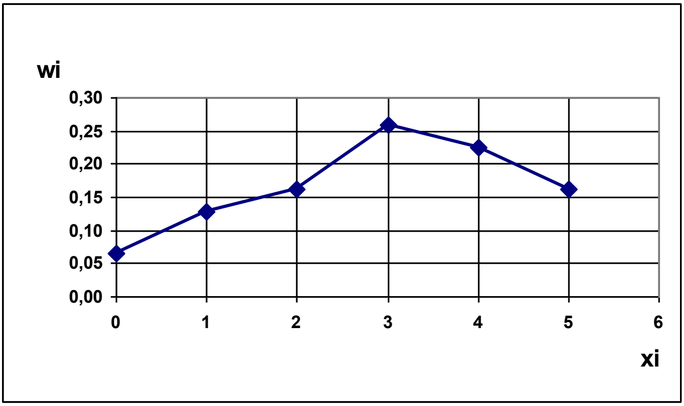

Рис. 2. Полігон відносних частот

Знаходимо емпіричну функцію розподілу $F_n^*(x)$:

$$F^*_{31} = \begin{cases} 0 & x \le 0, \\ 0{,}06 & 0 < x \le 1, \\ 0{,}19 & 1 < x \le 2, \\ 0{,}35 & 2 < x \le 3, \\ 0{,}61 & 3 < x \le 4, \\ 0{,}84 & 4 < x \le 5, \\ 1 & x > 5, \end{cases}$$

Графік емпіричної функції розподілу подана на рис. 3.

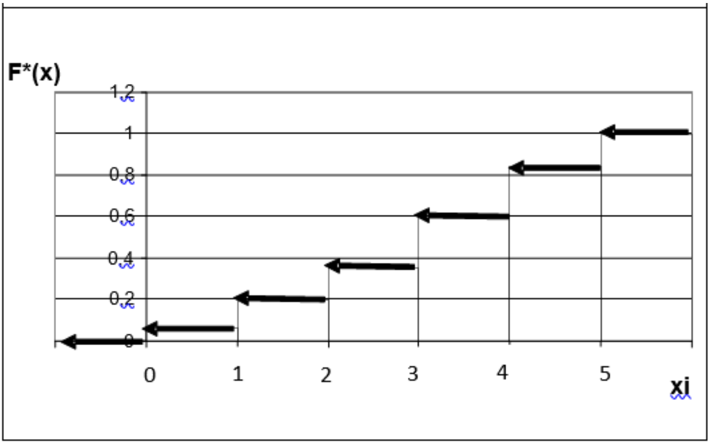

Рис. 3. Графік емпіричної функції розподілу

Приклад 2. Дана вибірка об’єму n=100.

135	132	132	137	138	134	147	145	132	147
133	134	131	133	136	131	137	133	139	134
142	142	143	135	133	146	132	132	135	139
102	126	111	122	115	121	112	109	117	122
119	107	122	118	115	107	121	119	119	133
126	114	124	126	120	112	116	116	111	120
116	118	119	120	118	126	126	121	118	117
126	130	121	130	119	117	127	123	115	130
122	117	128	118	121	119	124	122	122	123
128	123	124	126	130	126	130	126	127	126

1) Записати статистичний розподіл вибірки;
2) обчислити відносні частоти та накопичувані відносні частоти;
3) побудувати гістограму частот, гістограму відносних частот за даними розподілу вибірки;
4) знайти емпіричну функцію розподілу та побудувати її графік.

Розв'язання.

Визначимо найменше та найбільше значення вибірки: $x_{\min} = 102$, $x_{\max} = 147$. Тоді розмах вибірки становить: $R = 147 - 102 = 45$.

Число інтервалів визначаємо за емпіричною формулою Стерджеса

$$m = 1 + 3.322 \cdot \lg(100) = 7{,}644 \approx 8 .$$

За ширину інтервалу беремо $h = 45/8 \approx 6$. За початок першого інтервалу візьмемо $z_0 = 102 - 6/2 = 99$. Визначимо всі інтервали та частоту попадання варіант у інтервали. Див. стовпець 1,2,3 табл. 2. Обчислюємо відносні частоти і накопичувані відносні частоти. Дані заносимо до таблиці 2 (стовпець 4, 5).

Для побудови гістограм знаходимо середину кожного з інтервалів z_i^*, а також n_i / h *та* w_i / h, стовпці 6, 7, 8 (табл. 2) відповідно.

Таблиця 2

Розрахунки для прикладу 2

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
інтервали		n_i	w_i	$w_i^/$	z_i^*	n_i / h	w_i / h
99	105	1	0,01	0,01	102	0,17	0,002
105	111	5	0,05	0,06	108	0,83	0,008
111	117	13	0,13	0,19	114	2,17	0,022
117	123	28	0,28	0,47	120	4,67	0,047
123	129	17	0,17	0,64	126	2,83	0,028
129	135	23	0,23	0,87	132	3,83	0,038
135	141	6	0,06	0,93	138	1,00	0,010
141	147	7	0,07	1,00	144	1,17	0,012
Сума		100	1				

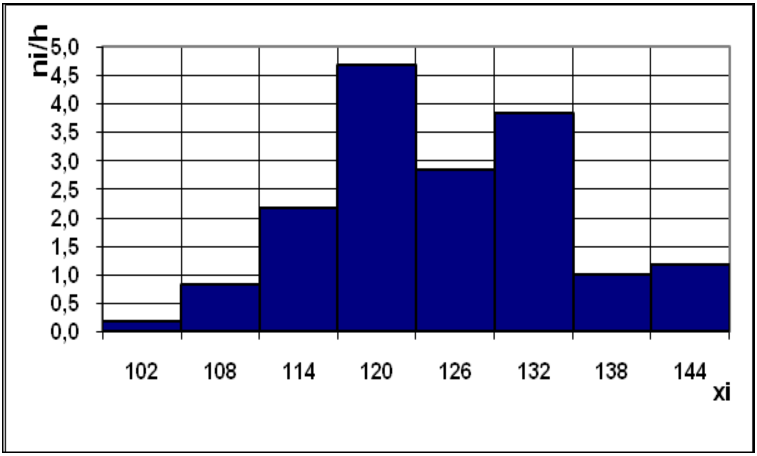

Рис. 4. Гістограма частот

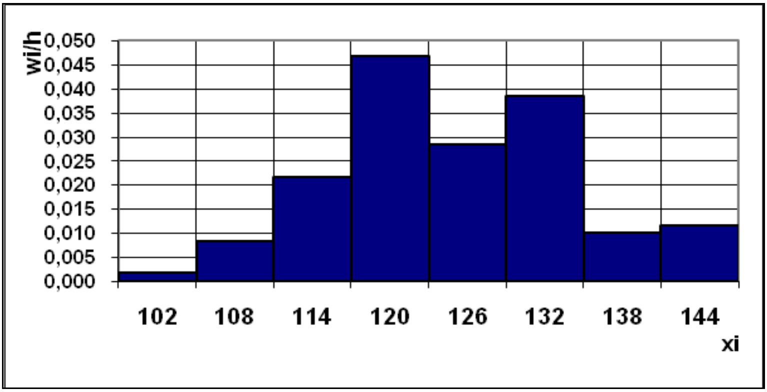

Рис. 5. Гістограма відносних частот

Знайдемо емпіричну функцію розподілу:

$$F^*_{100}(x)=\begin{cases}0, & x\le 99;\\ \dfrac{0,01}{6}(x-99), & 99<x\le 105;\\ \dfrac{0,05}{6}(x-105)+0,01, & 105<x\le 111;\\ \dfrac{0,13}{6}(x-111)+0,06, & 111<x\le 117;\\ \dfrac{0,28}{6}(x-117)+0,19, & 117\le 123;\\ \dfrac{0,17}{6}(x-123)+0,47, & 123<x\le 129;\\ \dfrac{0,23}{6}(x-129)+0,64, & 129<x\le 135;\\ \dfrac{0,06}{6}(x-135)+0,87, & 135<x\le 141;\\ \dfrac{0,07}{6}(x-141)+0,93, & 141<x\le 147;\\ 1, & x>147;\end{cases}$$

Будуємо графік емпіричної функції розподілу (рис. 6).

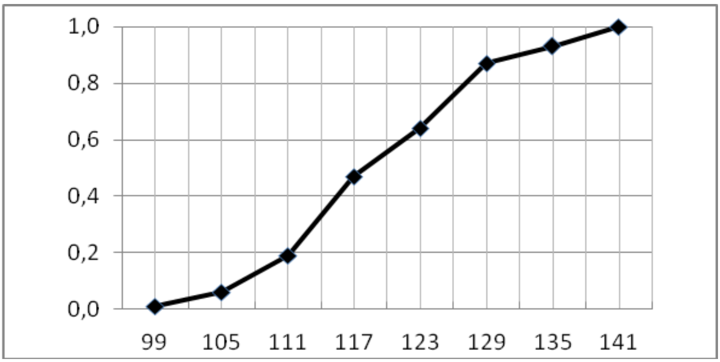

Рис. 6. Графік емпіричної функції розподілу

Завдання. Для вибірок Ai та Bi виконати наступні завдання (*i* – варіант):

1) записати статистичний розподіл частот вибірки Ai ;
2) записати інтервальний статистичний розподіл частот вибірки Bi;
3) обчислити відносні частоти та накопичувані частоти вибірок Ai та Bi;
4) побудувати полігон частот та полігон відносних частот вибірки Ai;
5) побудувати гістограму частот та гістограму відносних частот вибірки Bi;
6) знайти емпіричні функції розподілу вибірок Ai та Bi;
7) побудувати графіки емпіричних функцій розподілу вибірок Ai та Bi.

	Вибірка A_i																			
A1	0	4	2	0	5	1	1	3	0	2	2	4	3	2	3	3	0	4	5	1
	3	1	5	2	0	2	2	3	2	2	2	6	2	1	3	1	3	1	5	4
	5	5	3	2	2	0	2	1	1	3	2	3	5	3	5	2	5	2	1	1
	2	3	4	3	2	3	2	4	2											
A2	1	7	4	6	1	4	2	4	6	5	3	2	3	0	5	6	7	7	3	1
	5	5	4	2	6	2	1	5	3	3	1	5	6	4	4	3	4	1	5	5
	3	4	3	7	4	5	6	7	5	2	4	6	6	7	7	3	5	4	4	3
	5	5	7	6	6	1														
A3	0	0	2	0	1	3	0	1	5	1	2	1	3	0	0	2	1	3	2	2
	1	3	3	2	5	2	4	3	2	1	2	2	2	2	3	3	1	1	1	3
	2	1	5	1	2	1	4	4	2	3	3	6	5	2	1	2	3	2	3	1
	1	0	1	0	4	1	1	0	2	2	4	2	1	4	3	0	2	0	2	0
	3	1	2																	
A4	5	4	4	4	5	0	3	7	2	2	3	0	5	6	3	4	6	1	2	5
	3	2	3	6	6	2	3	1	7	2	3	2	2	5	2	0	2	2	6	1
	3	6	7	7	2	6	4	6	1	1	6	7	1	3	4	6	6	3	2	1
	7	2	5	4	2	3	4	5	6	6	5	3	2							
A5	4	8	4	1	7	7	5	8	8	6	7	1	6	5	8	6	7	5	6	5
	4	7	4	8	4	6	5	7	4	8	7	4	3	2	8	8	4	3	4	4
	7	5	0	4	7	6	3	5	7	2	6	8	5	8	1	7	8	5	1	6
	3	8	6	6	8	8	3	6	8	7	5	5	5	3	9	6	6	1	7	6
	5	7	7	8	3	7	3	6	5	4	4	4	7	7	4					

A6	4	10	7	6	3	7	8	7	4	7	10	7	3	9	3	5	7	8	5	7
	2	5	8	10	6	6	5	7	6	3	8	4	3	5	4	10	9	10	8	2
	10	6	6	7	8	2	7	7	4	6	7	10	4	4	10	3	6	9	6	3
	6	4	4	8	8	5	10	7	3	8	5	8	6							
A7	2	2	1	3	4	2	1	1	3	3	4	3	2	4	2	5	4	3	1	4
	4	4	2	3	4	3	7	1	3	3	3	4	3	2	1	2	3	3	1	5
	3	0	2	1	2	3	0	0	3	6	2	4	3	4	2	4	1	2	0	3
	1	0	0	2																
A8	8	4	4	7	5	5	5	3	10	2	3	6	7	6	10	6	8	4	10	8
	6	7	7	6	10	7	6	8	10	7	7	9	1	3	4	8	7	4	7	6
	7	4	4	5	4	9	6	5	9	6	6	5	6	4	7	7	6	6	3	8
	2	5	7	6	7	6	2	5	2	8	2	10	10	9	1	4	5	5	6	6
A9	2	1	2	3	1	1	0	2	2	4	3	3	0	3	0	3	2	3	2	1
	2	3	5	2	3	0	2	3	3	4	4	1	4	0	0	1	2	4	3	4
	0	0	5	2	2	3	2	1	0	0	0	3	1	0	1	2	1	2	4	2
	3	2	5	0	1	0	3	0	0	3	1	3	4	2	3	3	2	0	3	4
A10	3	5	6	8	4	5	4	7	2	7	7	3	7	4	4	7	3	3	7	3
	5	4	4	5	2	4	8	8	4	6	5	9	4	0	4	6	4	6	6	8
	4	4	9	3	3	2	1	5	2	5	5	3	4	4	7	2	8	3	8	2
	8	9	8	4	5	2	5	7	6	1	2	5	6	3	1	6	5	9	5	

A11	1	0	1	1	1	2	0	2	1	1	5	0	5	1	0	3	2	1	3	5
	0	3	1	1	0	2	0	3	3	3	2	1	3	2	1	5	0	1	0	1
	0	2	3	1	0	3	1	1	1	2	2	1	1	0	0	1	1	3	0	2
	2	1	1	0	4	2	2	1	1	1	2	0	0	1	3					
A12	6	6	5	6	9	8	7	4	4	8	3	2	3	9	7	4	7	6	9	6
	6	9	5	8	8	7	10	8	6	9	9	10	3	10	5	8	7	6	9	4
	7	6	8	9	9	3	8	4	7	4	6	9	2	8	7	8	6	8	2	3
	7	7	8	4	3	6	10	10	2	3	8									
A13	1	0	1	1	1	2	0	2	1	1	2	3	3	3	4	4	5	3	1	1
	0	2	3	3	3	4	4	5	3	1	4	1	3	2	1	0	0	1	0	2
	0	2	3	1	0	3	1	1	1	2	2	1	1	0	0	1	1	3	0	0
	2	2	3	3	3	4	2	1	1											
A14	10	6	5	6	10	8	7	10	8	6	9	9	10	3	10	5	8	7	6	9
	6	9	5	8	8	7	10	8	6	9	9	10	7	6	9	7	4	7	6	9
	7	6	8	9	9	3	8	4	2	4	6	9	2	8	7	6	6	8	2	3
	7	7	8	4	3	6	3	10	2	3	8									
A15	3	6	7	7	2	6	4	6	1	1	6	7	1	3	4	6	6	3	2	1
	3	2	3	6	6	2	3	1	7	2	3	2	2	5	2	0	2	2	6	1
	6	2	3	1	7	2	3	2	2	5	4	5	6	6	5	3	2	2	2	6
	7	2	5	4	2	3	4	5	6	6	5	3	2							

A16	7	9	6	5	4	4	4	7	7	3	4	4	8	5	1	5	7	7	8	3
	4	7	4	8	4	6	5	7	4	8	7	4	3	2	8	7	5	9	5	9
	7	5	0	4	7	6	3	5	7	2	6	6	5	8	1	6	8	6	7	9
	3	8	6	6	8	8	9	6	8	7	5	1	5	3	9	7	6	1	7	1
A17	5	7	7	8	3	7	9	6	5	2	4	8	6	2	8	6	7	5	9	5
	5	3	6	6	5	5	3	5	6	7	8	9	5	2	8	8	9	3	4	4
	4	5	6	6	3	6	5	3	4	5	3	3	7	5	1	7	8	5	2	2
	3	8	6	6	8	8	9	6	8	7	5	4	7	8	4	3	4	6		
A18	7	8	9	5	2	8	8	9	3	4	4	7	3	9	3	5	7	8	5	7
	5	4	3	7	5	1	7	8	5	6	5	5	3	5	6	7	8	9	5	2
	3	9	3	5	7	1	7	7	4	6	3	6	5	3	4	5	8	3	7	3
	6	4	4	8	8	5	2	7	3	8	5	7	6	6	7	8	1	7	7	
A19	2	3	3	3	4	4	5	3	1	4	1	4	5	5	1	7	4	3	1	4
	4	3	5	5	4	3	2	6	7	1	3	2	2	4	2	2	3	3	1	5
	2	4	3	6	1	6	4	4	4	5	2	4	5	3	5	4	1	2	0	3
	2	2	5	2	5	4	0	7	1	0	0	0	5	3	5	3	1	4		
A20	9	3	3	2	1	5	2	3	2	7	7	3	7	4	4	7	3	3	7	3
	5	4	2	5	7	6	1	2	5	6	5	9	3	0	4	6	4	3	6	8
	4	4	9	3	3	2	1	5	2	5	2	5	7	6	1	2	5	3	8	2
	8	2	3	9	7	4	5	3	6	1	2	5	6	3	1	6	5			

	Вибірка B_i											
B1	11	133	124	132	104	152	134	130	129	120	122	124
	11	123	123	129	121	122	125	131	147	124	137	112
	12	128	111	129	115	147	131	132	137	119	126	120
	11	125	123	127	132	118	133	132	132	134	131	120
	11	132	125	132	108	114	121	133	133	135	131	125
	11	115	122	131	125	132	120	126	115	117	118	118
	12	134	127	127	124	135	128	127	115	144	129	120
	12	127	125	116	132	120	117	127	118	109	127	122
	12	135	116	118	133	136	125	126	119	126	129	127
	11	124	127	132	126	131	127	130	126	124	135	127
	12	123	123	130	132	143	122	139	120	131	108	132
	13	111	123	140	137	120	125	131	118	120	120	136
	12	127	116	138	128	133	122	131	128	140	138	134
	12	126	109	137	111	115	117	130	113	126	115	124
	12	118	115	128	123	129	128	120	115	134	118	135
	12											
B2	95	96	103	89	72	105	85	85	91	101	82	91
	80	85	91	87	101	94	98	85	82	94	86	72
	89	83	100	86	85	95	95	83	87	92	92	79
	93	88	77	92	92	103	85	90	83	86	104	104
	10	85	80	95	91	93	70	83	93	95	95	78

	11	95	94	84	64	87	85	87	87	81	82	97
	10	86	89	80	88	85	93	79	95	90	107	93
	96	83	88	91	95	94	88	80	96	93	77	71
	88	97	90	86	93	91	98	95	83	84	91	99
	10	80	95	87	89	85	87	72	77	90	97	87
	95	91	88	91	81	88	78	75	80	97	95	83
	91	78	87	92	103	77	101	66	71	90	105	76
	97	75	95	88	84	96	79	89	94	100	87	100
	10	100	79	96	104	84	89	82	93	92	85	80
	10	87	90	89	89	83	84	98	81	97	86	81
	96	82	102	73	100	81	86	84	86	88	90	94
	87	99	100	81	95	88	90	87	97	90	100	94
	10	85	95	74	85	88	78	97	74			
B3	79	56	46	50	67	37	53	49	42	57	49	20
	54	43	82	54	49	27	56	49	26	46	72	41
	59	61	41	32	42	71	55	46	71	68	37	67
	39	59	39	42	45	55	45	54	87	72	68	44
	69	52	67	66	7	54	56	39	40	61	73	50
	77	54	70	45	49	50	41	39	61	70	44	42
	70	55	45	36	40	71	17	48	49	48	72	48
	64	11	66	37	12	36	15	55	39	72	61	68
	48	28	40	32	26	69	33	50	78	56	61	36

	25	67	50	60	69	63	60	58	46	11	59	53
	41	76	56	46	26	58	40	43	55	45	73	47
	86	25	27	43	26	61	64	10	50	24	42	36
	54	80	56	27	57	76	43	49	43	67	75	64
	62	88	27	37	56	67	63	56	48	28	24	42
	83	15	36	52	38	54	46	16	51	34	64	53
	29	65	45	42	44	77	76	33	39	48	43	49
	33	38	38	48	72	53	28	44	43	71	57	49
	47	63	36	72	57	33						
B4	57	61	60	63	66	68	64	72	69	59	71	62
	58	60	66	59	62	64	53	50	50	65	70	61
	66	72	71	60	74	62	49	62	76	66	64	62
	49	79	58	73	61	63	64	50	45	70	62	61
	64	42	73	62	69	60	61	69	62	67	67	72
	58	63	71	73	68	80	64	64	53	64	68	58
	54	73	59	69	60	67	57	54	69	55	70	65
	71	55	67	57	64	70	55	65	69	65	65	60
	60	54	75	62	74	63	64	76	59	71	68	55
	73	54	57	56	65	53	64	58	67	48	66	68
	58	58	62	58	52	62	65	71	64	66	65	58
	72	43	63	59	76	67	63	71	66	59	69	65
	57	69	57	61	77	70	65	60	53	65	69	67

	57	51	77	73	68	61	61	65	62	66	63	74
	68	61	57	55	77	59	66	73	73	66	50	65
	68											
B5	32	296	313	323	312	321	322	301	337	322	329	307
	32	328	312	318	327	315	319	317	309	334	323	340
	32	322	314	335	313	322	319	325	312	300	323	335
	31	326	298	298	337	322	303	314	315	310	316	321
	31	315	331	322	321	336	328	315	338	318	327	323
	32	314	297	303	322	314	317	330	318	320	312	333
	33	319	325	319	307	305	316	330	318	335	327	321
	33	288	322	334	295	318	329	305	310	304	326	319
	31	316	316	307	309	309	328	317	317	322	316	304
	30	350	309	327	345	329	338	311	316	324	310	306
	30	302	315	314	343	320	304	310	345	312	330	324
	30	326	313	320	328	309	306	306	308	324	312	309
	32	321	313	330	330	315	320	313	302	295	337	346
	32	320	307	305	323	331	345	315	318	331	322	315
	30	324	317	322	312	314	308	303	333	321	312	323
	31	288	317	327	292	316	322	319	313	328	313	309
	32	313	334	314	320	301	329	319	332	316	300	300
	30	306	314	323	318	337	325	321	322	288	313	314
	30	329	302	300	316	321	315	323	331	318	334	316

	32	294	288	312	312	315	321	332	319			
B6	61	59	60	50	58	71	57	61	55	75	68	65
	66	52	70	69	62	58	56	54	65	61	67	64
	71	60	51	54	57	56	55	57	65	56	61	49
	68	63	72	67	54	53	58	69	63	66	55	57
	55	69	54	64	54	61	63	60	63	60	62	62
	52	62	55	70	72	64	71	54	58	71	66	65
	60	64	63	61	60	64	65	68	64	66	69	53
	60	63	61	60	66	68	66	64	64	67	62	55
	65	65	56	57	72	53	62	68	63	57	55	68
	62	63	62	59	67	56	65	67	56	69	63	53
	54	68	59	63	67	57	64	68	76	64	64	62
	68	60	61	64	64	59	53	61	61	62	59	62
	67	62	60	61	63	68	58	67	68	66	68	57
	65	59	55	68	61							
B7	78	85	52	53	62	56	58	68	98	58	94	84
	61	64	62	53	89	66	54	62	57	64	66	35
	62	54	75	52	59	72	54	66	46	44	57	63
	59	54	83	53	71	64	60	48	77	47	51	54
	54	66	64	82	78	70	88	61	63	77	41	62
	64	66	80	71	53	99	58	63	43	56	51	70
	60	58	59	67	53	56	74	71	86	30	55	58

	73	86	50	63	50	74	78	60	78	68	72	65
	51	68	65	64	72	72	70	70	78	50	56	66
	66	69	64	58	71	76	61	52	67	71	61	73
	63	53	76	58	58	77	68	67	60	69	64	53
	80	53	83	51	46	63	74	45	73	70	92	79
	69	56	48	64	75	62	67	49	58	73	52	64
	70	64	75	78	59	51	86	74	72	43	53	65
	66	54	70	81	47	68	85	93	70	51	71	67
	79	46	54	49	63	96	63	61	82	61	67	85
	57	68	64	53	73	57	86	63	61	60	64	60
	73	76	73	67	76	69	87	62	72	73	67	60
	45	82	64	69	70	79	79	82	73	64	63	49
	67	57	40	53	98	64	56					
B8	56	76	65	66	76	62	89	48	62	50	47	80
	65	67	51	73	75	61	88	46	57	65	60	72
	69	68	65	34	77	63	57	61	42	85	49	41
	62	65	75	56	66	92	60	43	52	80	68	70
	42	87	81	67	65	81	90	38	58	60	79	79
	58	77	73	54	58	77	86	52	61	42	70	93
	53	64	65	76	88	59	62	67	62	90	88	69
	72	58	68	94	54	58	58	81	57	70	71	78
	57	68	70	58	72	57	62	63	87	61	91	57

	40	63	86	48	75	66	83	64	55	75	65	67
	51	86	67	58	73	71	46	86	68	79	50	58
	64	78	78	60	46	71	71	74	79	65	61	62
	83	43	61	67	50	60	83	61	83	67	67	58
	47	76	81	72	66	83	73	71	70	60	68	52
	75	61	80	51	63	62	46	62	63	80	64	70
	67	87	78	28	75	51	76	62	55	54	65	51
	61	81	65	52	93	89	57	66	68	54	70	44
	50	66	69	61	84	53	67	46	73	58	63	63
	51											
B9	71	62	43	80	70	44	42	25	48	55	58	44
	49	54	63	60	57	70	52	74	65	61	60	72
	30	62	81	56	55	38	68	55	74	50	29	35
	58	50	62	80	49	68	68	81	66	64	41	45
	66	82	76	84	47	44	72	58	58	80	61	55
	44	88	88	73	39	70	70	35	51	69	50	59
	54	66	85	63	59	52	88	64	60	61	31	64
	41	62	42	76	81	76	70	76	75	53	66	87
	73	44	61	53	46	69	71	58	63	73	56	65
	83	45	55	77	61	42	72	49	52	67	62	68
	67	53	70	76	56	62	38	59	53	50	76	52
	60	61	74	55	56	55	52	27	48	68	79	50

	48	49	69	68	47	36	69	35	43	71	34	51
	66	74	61	68	53	77	39	72	46	76	73	
B10	12	134	121	131	145	189	111	137	131	124	128	132
	12	134	110	143	140	127	132	136	134	133	121	134
	13	126	138	144	138	123	137	137	127	139	122	121
	12	127	132	129	136	135	142	132	129	128	134	138
	13	137	132	140	125	127	129	133	134	135	124	132
	12	133	141	135	141	130	130	130	131	133	133	131
	12	127	138	128	122	137	136	137	148	128	117	114
	13	137	126	137	136	130	124	130	140	132	127	125
	12	136	116	120	125	136	124	144	133	132	123	136
	13	130	130	134	132	131	125	131	130	146	132	131
	13	135	133	126	134	137	124	136	126	125	117	133
	15	130	136	132	127	131	122	125	133	133	123	136
	13	135	130	123	133	123	140	133	136	124	129	129
	12	126	128	138	127	120	144	135	126	126	144	125
	12	132	138	165	139	137	133	129	124	110	131	128
	13	130	124	142	124	129	131	143	129	127		
B11	58	49	46	53	63	64	53	46	59	64	50	55
	48	68	54	59	66	61	69	70	51	60	47	50
	40	51	46	58	63	51	65	55	55	61	45	50
	65	52	60	58	27	41	60	44	38	44	59	72

	55	53	51	33	53	70	55	60	50	51	50	55
	52	46	56	65	52	61	52	65	51	58	54	55
	52	52	47	48	55	58	72	53	69	42	41	59
	47	54	52	60	47	55	48	64	63	72	51	55
	42	63	59	60	70	54	40	58	49	66	59	55
	73	41	68	54	48	52	52	50	67	59	43	64
	54	63	63	65	60	47	72	58	55	52	53	49
	45	40	54	77	49	56	45	64	69	57	50	59
	54	57	52	63	42	41	57	60	60	52	49	46
	47	52	51	59	42	56	43	50	44	45	59	54
	50	46	42	48	59	53	64	53	54	72	55	66
	57	55	68	57	58	61	64	58	68	56	28	50
	43	62	48	50	46	58	57	71	67	59	43	45
	44	45	57	74	47	38	60	71	57	56	71	63
	43	60	50	50	65							
B12	52	40	47	64	40	54	41	74	45	45	51	76
	42	53	54	65	46	65	61	55	38	66	42	56
	43	49	77	64	53	64	58	54	56	53	43	35
	48	66	49	49	57	48	42	46	52	59	50	62
	46	53	51	50	60	30	48	56	29	74	52	60
	54	40	33	20	55	42	61	54	41	45	75	59
	54	52	62	69	65	49	48	63	52	46	44	55

	82	67	68	34	56	51	56	48	53	47	59	51
	61	42	54	33	39	47	46	47	73	63	34	44
	43	30	60	61	53	47	42	56	70	48	45	65
	40	57	56	33	44	43	45	35	35	56	59	66
	53	49	55	25	53	48	73	38	58	72	57	46
	38	53	48	68	36	53	41	55	51	50	45	50
	50	59	33	56	49	31	70	56	56	51	46	40
	58	37	40	48	48	51	56	52	44	54	55	59
	54	40	60	29	60	39	59	66	48	60	54	39
	56	34	59	41	51	45	50	55	55	44	62	23
B13	56	76	65	66	76	62	89	48	62	50	47	80
	65	67	51	73	75	61	88	46	57	65	60	72
	69	68	65	34	77	63	57	61	42	85	49	41
	62	65	75	56	66	92	60	43	52	80	68	70
	42	87	81	67	65	81	90	38	58	60	79	79
	58	77	73	54	58	77	86	52	61	42	70	93
	53	64	65	76	88	59	62	67	62	90	88	69
	72	58	68	94	54	58	58	81	57	70	71	78
	57	68	70	58	72	57	62	63	87	61	91	57
	40	63	86	48	75	66	83	64	55	75	65	67
	51	86	67	58	73	71	46	86	68	79	50	58
	64	78	78	60	46	71	71	74	79	65	61	62

	83	43	61	67	50	60	83	61	83	67	67	58
	47	76	81	72	66	83	73	71	70	60	68	52
	75	61	80	51	63	62	46	62	63	80	64	70
	67	87	78	28	75	51	76	62	55	54	65	51
	61	81	65	52	93	89	57	66	68	54	70	44
	74	71	67	73	68	68	72	68	67	70	78	74
	65	71	70	69	69	76	71	63	77	75	70	74
	71	60	69	66	71	69	73	74	80	69	73	76
	65	74	68	74	60	70	66	70	68	64	75	78
	71	75	74	72	80	72	69	69	71	70	73	65
	70	70	72	76	72	73	64	74	71	76	68	69
	64	78	66	75	72	69	68	63	70	70	78	76
	61	66	66	72	69	71	71	68	72	69	73	73
B14	70	69	74	72	69	74	70	74	72	76	71	66
	66	74	69	64	75	71	76	68	68	78	71	71
	71	81	72	68	72	71	71	71	69	61	74	66
	80	71	78	73	71	75	73	71	72	68	67	69
	61	74	67	68	69	74	69	67	74	66	74	74
	71	60	69	66	71	69	73	74	80	69	73	76
	65	74	68	74	60	70	66	70	68	64	75	78
	71	75	74	72	80	72	69	69	71	70	73	65
	70	70	72	76	72	73	64	74	71	76	68	69

B15	17	25	26	25	32	32	23	24	23	26	29	34
	21	21	19	27	17	22	31	23	23	26	37	15
	26	45	27	38	8	19	25	15	24	23	39	21
	24	21	24	31	22	23	31	24	26	19	17	26
	32	32	37	28	29	17	17	24	28	18	38	18
	18	26	27	26	20	33	24	33	33	36	22	19
	33	25	32	24	33	17	15	26	25	21	37	24
	19	25	21	23	27	32	12	25	23	28	18	41
	31	21	16	27	34	16	30	20	24	16	20	33
	17	23	14	28	37	18	29	26	26	19	28	31
	22	27	20	29	24	17	24	15	24	21	22	30
	27	23	9	25	26	12	19	33	24	35	22	26
	32	32	37	28	29	17	17	24	28	18	38	18
	18	26	27	26	20	33	24	33	33	36	22	19
	33	25	32	24	33	17	15	26	25	21	37	24
	19	25	21	23	27	32	12	25	23	28	18	41
B16	10	51	80	83	83	67	55	84	78	83	101	75
	78	99	69	99	71	67	56	74	51	78	34	67
	10	66	106	70	117	67	116	79	120	47	113	113
	69	100	78	31	68	66	91	85	64	55	83	77
	68	83	76	89	88	58	75	60	89	111	42	104

	33	96	50	42	81	78	42	64	89	60	32	46
	82	33	72	93	94	49	153	68	85	78	95	51
	76	81	67	50	75	99	114	130	108	127	110	91
	77	85	102	101	79	118	132	130	79	88	76	73
	82	75	118	50	100	70	42	79	64	78	137	83
	92	71	84	77	73	100	69	77	74	98	79	102
	83	66	59	67	87	60	91	68	91	103	73	93
	69	54	82	71	60	88	82	82	41	68	53	78
	96	97	81	86	69	52	77	66	100	119	84	102
	46	54	77	129	87	106	84	96	81	103	73	93
	83	66	59	67	87	60	91	68	91	68	53	78
	69	54	82	71	60	88	82	82	41			
B17	18	181	201	178	190	188	181	180	186	180	176	186
	19	184	187	176	189	194	196	190	193	180	186	195
	18	189	197	190	176	200	196	188	203	191	180	181
	17	185	188	173	185	180	189	178	100	175	193	184
	18	179	177	203	185	182	191	183	183	211	189	177
	17	174	200	195	189	187	193	185	184	193	181	185
	18	182	179	186	193	172	190	200	176	179	185	182
	17	191	186	182	206	181	197	197	180	193	192	200
	17	185	175	187	160	187	185	206	187	182	175	172
	18	193	178	191	197	177	175	170	174	194	188	182

	17	186	190	183	196	183	186	174	195	179	197	182
	17	184	185	172	193	175	172	179	179	184	190	183
	19	192	186	157	172	185	180	193	177	174	200	195
	18	186	185	206	192	189	189	184	183	182	179	186
	20	169	189	180	183	192	186	200	176	191	186	182
	18	184	192	179	204	197	194	182	172	185	175	187
	19	184	186	201	197	188	188	194	184	193	178	191
	81	116	75	120	75	187	181	196	204	177	178	167
	86	66	100	75	84	199	180	181	187	187	178	181
	14	95	92	105	114	189	193	191				
	11	93	110	79	63	72	75	126	73	106	75	69
	92	84	123	84	125	93	74	131	51	117	145	92
	13	137	100	102	88	67	92	130	99	94	92	76
	11	84	78	100	98	114	113	94	108	76	88	105
B18	12	96	81	116	75	120	75	62	113	109	111	78
	75	87	86	66	100	75	84	95	121	103	95	63
	10	67	148	95	92	105	114	98	102	41	76	98
	79	97	111	93	110	79	63	109	69	108	71	90
	84	136	92	84	123	84	125	102	96	72	102	100
	10	87	132	137	100	102	88	65	75	114	79	136
	12	115	90	78	86	122	119	87	115	96	137	63
	71	88	75	100	84	71	123	121	94	114	94	105

	45	94	102	109	86	45	97	93	43	48	114	118
	10	124	89	104	106	108	100	106	102	105	119	79
B19	11	115	122	128	115	118	116	124	120	127	136	129
	13	111	117	114	109	122	118	132	118	118	116	132
	11	129	123	124	122	127	126	119	115	122	131	129
	13	128	123	104	124	122	114	109	132	122	121	129
	11	120	97	123	105	135	132	133	119	137	126	102
	14	121	126	127	118	131	115	122	118	116	130	126
	13	127	116	120	119	128	104	131	115	140	115	124
	12	118	104	125	131	117	118	102	127	120	102	120
	13	128	106	132	129	131	126	110	128	134	132	124
	10	119	132	117	120	122	114	125	139	116	125	132
	11	122	120	113	123	119	104	106	108	100	106	124
	11	123	141	115	121	113	128	115	118	116	124	116
	12	126	139	114	123	125	114	109	122	118	132	100
	12	120	115	115	127	115	124	122	127	126	119	116
	10	124	112	135	117	127	104	124	122	114	109	118
	10	119	112	114	133	112	123	105	135	132	133	126
	12											
B20	33	322	329	307	300	323	322	301	307	309	309	328

	30	334	323	340	310	316	319	317	327	345	329	338
	31	313	330	330	318	327	319	325	314	343	320	304
	31	307	305	323	320	312	303	314	320	328	309	306
	30	316	330	318	335	327	328	315	338	317	322	312
	32	314	297	303	322	314	317	330	318	317	327	292
	33	319	325	319	307	305	316	330	318	334	314	320
	33	288	322	334	295	318	329	305	310	314	323	318
	31	316	316	313	302	295	315	317	317	322	316	304
	30	350	309	315	318	331	330	311	316	324	310	306
	30	302	315	303	333	321	330	310	345	312	330	324
	30	326	313	319	313	328	305	306	308	324	312	309
	32	321	313	319	332	316	320	313	302	295	337	346
	32	320	307	305	323	331	345	315	318	331	322	315
	30	324	317	322	312	314	308	303	333	321	312	323
	31	288	317	327	292	316	322	319	313	328	313	309
	32	313	334	314	320	301	329	319	332	316	300	300
	30	306	314	323	318	337	325	321	322	288	313	314
	30	329	302	300	316	321	315	323	331	318	334	316
	32	294	288	312	312	315	321					

Лабораторна робота №2

Тема: Числові характеристики вибірки.

Теоретичні відомості.

Вичерпні дані про випадкову величину ξ (генеральну сукупність) задаються при допомозі функції розподілу ймовірностей або функції щільності. Проте при практичному вивченні генеральної сукупності часто цілком достатньою виявляється менш детальна інформація, яка подана у вигляді декількох числових характеристик.

Введемо вибіркові аналоги невідомих істинних моментів розподілу.

I. Вибіркові характеристики у випадку, коли дані не повторюються:

1) вибірковий момент r-того порядку:

$$\hat{\nu}_r = \frac{1}{n}\sum_{i=1}^{n}\left(x_{(i)}\right)^r ,$$

2) вибіркове середнє:

$$\hat{\nu}_1 = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_{(i)} ,$$

3) вибіркова дисперсія:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{(i)} - \hat{\nu}_1)^2 ,$$

4) вибірковий центральний момент r-го порядку:

$$\hat{\mu}_r = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\nu}_1)^r ,$$

II. Вибіркові характеристики у випадку, коли дані повторюються:

1) вибірковий момент r-того порядку:

$$\hat{\nu}_r = \frac{1}{n}\sum_{i=1}^{k}\left(y_i\right)^r n_i ,$$

2) вибіркове середнє:

$$\hat{\nu}_1 = \frac{1}{n}\sum_{i=1}^{k} y_i n_i ,$$

3) вибіркова дисперсія:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{k} (y_i - \hat{\nu}_1)^2 n_i ,$$

4) вибірковий центральний момент r-го порядку:

$$\hat{\mu}_r = \frac{1}{n}\sum_{i=1}^{k} (y_i - \hat{\nu}_1)^r n_i .$$

III. Вибіркові характеристики у випадку згрупованих даних по інтервалах визначаються наступним чином:

1) вибірковий момент r-того порядку:

$$\hat{\nu}_r = \frac{1}{n}\sum_{i=1}^{m} \left(z_i^*\right)^r n_i ,$$

2) вибіркове середнє:

$$\hat{\nu}_1 = \frac{1}{n}\sum_{i=1}^{\mathrm{m}} z_i^* n_i ,$$

де $z_i^* = \dfrac{z_{i-1} + z_i}{2}$ – середина i-того інтервалу,

3) вибіркова дисперсія:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{m} (z_i^* - \hat{\nu}_1)^2 n_i ,$$

4) вибірковий центральний момент r-го порядку:

$$\hat{\mu}_r = \frac{1}{n}\sum_{i=1}^{m} (z_i^* - \hat{\nu}_1)^r n_i .$$

Для обчислення вибіркової дисперсії часто зручніше використовувати іншу форму

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{k} (y_i - \hat{\nu}_1)^2 n_i = \frac{1}{n}\sum_{i=1}^{k} \left(y_i\right)^2 n_i - \left(\hat{\nu}_1\right)^2$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{m}(z_i^* - \hat{\nu}_1)^2 n_i = \frac{1}{n}\sum_{i=1}^{m}\left(z_i^*\right)^2 n_i - \left(\hat{\nu}_1\right)^2 .$$

Вимірність дисперсії дорівнює квадрату вимірності значення випадкової величини, що створює незручність у дослідженні. Щоб її усунути, за характеристику розсіювання значення випадкової величини за результатами вибірки приймають вибіркове середньоквадратичне відхилення, яке визначається як

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} .$$

Як характеристику центру групування даних в математичні статистиці використовують моду (***Mo***) та медіану (***Me***).

Медіана Me генеральної сукупності визначається як середньо ймовірне значення, тобто таке, що

$$P\{x < Me\} = P\{x > Me\} .$$

Статистичною оцінкою ***медіани*** називають варіанту, яка ділить варіаційний ряд на дві частини рівні за кількістю варіант. Значення медіани може співпадати з одним із значень варіаційного ряду, іноді його шукають, інтерполюючи два сусідні значення. У випадку дискретного статистичного розподілу:

- якщо кількість варіант непарна, тобто $n = 2k + 1$, то

$$Me = x_{(k+1)} ,$$

- якщо кількість варіант парна, тобто $n = 2k$, то

$$Me = \frac{1}{2}(x_{(k)} + x_{(k+1)}) .$$

У випадку інтервального варіаційного ряду, коли дані вибірки згруповані по інтервалах, медіана визначається наступним чином:

$$Me = z_{Me} + h\frac{n/2 - n'_{Me-1}}{n_{Me}} ,$$

де z_{Me} – початок медіанного інтервалу, тобто інтервалу, у якому міститься серединний елемент; h – ширина медіанного інтервалу; n – об’єм вибірки, n'_{Me-1} – накопичена частота проміжку, що передує медіанному; n_{Me} – частота медіанного інтервалу.

Модою називають варіанту, яка має найбільшу частоту. Іншими словами, це значення вибірки, яке найчастіше зустрічається при спостереженні. У випадку інтервального варіаційного ряду, коли дані вибірки згруповані по інтервалах мода визначається наступним чином:

$$Mo = z_{Mo} + h\frac{n_{Mo} - n_{Mo-1}}{2n_{Mo} - n_{Mo-1} - n_{Mo+1}},$$

де z_{Mo} – початок модального інтервалу, тобто такого, якому відповідає перша з накопичених частот, що перевищує половину всіх спостережень; h – ширина модального інтервалу, n_{Mo} – частота модального інтервалу, n_{Mo-1} та n_{Mo+1} – частоти відповідно домодального і післямодального інтервалів.

Приклад 3. Для вибірки прикладу 1 обчислити вибіркові характеристики: вибіркове середнє, вибіркову дисперсію, вибіркове середньоквадратичне відхилення, моду, медіану.

Розв'язання.

Таблиця 3

Розрахунки для прикладу 3

1	**2**	**3**	**4**
y_i	n_i	$y_i n_i$	$(y_i - \hat{v}_1)^2 n_i$
0	2	0	17,23
1	4	4	14,98
2	5	10	4,38
3	8	24	0,03
4	7	28	7,93
5	5	25	21,31
сума	31	91	65,87

Враховуючи дані 3-го стовпця (табл. 3), отримаємо вибіркове середнє:

$$\hat{v}_1 = \frac{91}{31} = 2{,}94.$$

На основі отриманих значень 4-го стовпця (табл. 3), маємо вибіркову дисперсію:

$$\hat{\sigma}^2 = \frac{65,87}{31} = 2,12.$$

Тоді вибіркове середньоквадратичне відхилення:

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = 1,46.$$

Варіанта, що має найбільшу частоту – це мода: Mo=3, а варіанта, яка ділить варіаційний ряд на дві частини рівні за кількістю варіант – це медіана: Me=3.

Приклад 4. Для вибірки прикладу 2 обчислити вибіркові характеристики: вибіркове середнє, вибіркову дисперсію, вибіркове середньоквадратичне відхилення, моду, медіану.

Розв'язання.

Таблиця 4

Розрахунки для прикладу 4

1	**2**	**3**	**4**	**5**	**6**	**7**
інтервали		n_i	$n_i^/$	z_i^*	$z_i^* n_i$	$(z_i^* - \hat{v}_1)^2 n_i$
99	105	1	1	102	102	528,08
105	111	5	6	108	540	1441,60
111	117	13	19	114	1482	1567,29
117	123	28	47	120	3360	694,41
123	129	17	64	126	2142	17,69
129	135	23	87	132	3036	1133,45
135	141	6	93	138	828	1017,12
141	147	7	100	144	1008	2532,32
Сума		100			12498	8931,96

Враховуючи дані 6-го стовпця (табл. 4), отримаємо вибіркове середнє:

$$\hat{v}_1 = \frac{12498}{100} = 124,98.$$

На основі отриманих значень 7-го стовпця (табл. 4), маємо вибіркову дисперсію:

$$\hat{\sigma}^2 = \frac{8931{,}96}{100} = 89{,}32.$$

Тоді вибіркове середньоквадратичне відхилення: $\hat{\sigma} = \sqrt{\hat{\sigma}^2} = 9{,}45$.

Для визначення медіани за допомогою табл. 4 знайдемо спочатку:

$$z_{Me} = 123,\ h = 6,\ n = 100,\ n'_{Me-1} = 47,\ n_{Me} = 17.$$

Тоді медіана

$$Me = 123 + 6\frac{100/2 - 47}{17} \approx 124.$$

Для визначення моди за допомогою табл. 4 маємо:

$$z_{Mo} = 123,\ h = 6,\ n_{Mo} = 17,\ n_{Mo-1} = 28,\ n_{Mo+1} = 23.$$

Тоді мода

$$Mo = 123 + 6\frac{17 - 28}{2 \cdot 17 - 28 - 23} \approx 127.$$

Завдання. Для вибірок Аі, Ві лабораторної роботи №1 обчислити вибіркові характеристики: вибіркове середнє, вибіркову дисперсію, вибіркове середньоквадратичне відхилення, моду та медіану.

Лабораторна робота №3

Тема: Критерії перевірки гіпотез про розподіл випадкових величин.

Теоретичні відомості.

Статистичними гіпотезами називають різного роду твердження (припущення) про розподіл випадкових величин, які потрібно перевірити за даними вибіркової сукупності. Зазвичай ці припущення стосуються параметрів та типу розподілу випадкових величин, що розглядаються.

Будь-який критерій R для перевірки деякої гіпотези відносно розподілу випадкової величини ξ, є функцією від вибірки, що отримана в результаті спостережень. Тому вид розподілу критерію R зазвичай залежить від розподілу випадкової величини ξ.

Непараметричними називаються критерії, вид розподілу яких не залежить від розподілу генеральної сукупності ξ. Критерії перевірки гіпотез про розподіл випадкових величин називаються **критеріями згоди.**

Позначимо r – кількість параметрів розподілу, які оцінюються за вибіркою. Якщо $r = 0$, то це означає, що всі параметри розподілу відомі. Виявляється два випадки ($r = 0$ та $r > 0$) мають принципову різницю. Якщо гіпотеза повністю описує розподіл ($r = 0$), то в цьому випадку будемо говорити про перевірку гіпотези відносно (конкретного) розподілу випадкової величини і застосовується при цьому **критерій Колмогорова**. Якщо ж гіпотеза визначає розподіл лише з точністю до параметрів ($r > 0$), то будемо говорити про перевірку гіпотези відносно виду розподілу і застосовується **критерій χ^2 Пірсона**.

Критерій Колмогорова. Розглянемо гіпотезу $H_0 : F(x) = F_0(x)$ (про те, що наявна вибірка реалізована із генеральної сукупності неперервних випадкових величин з даною неперервною функцією розподілу $F_0(x)$) та альтернативу: $H_1 : F(x) \neq F_0(x)$.

Критерій **Колмогорова** базується на порівнянні теоретичної $F(x)$ та емпіричної $F^*(x)$ функцій розподілу:

$$D_n = \sup_{-\infty<x<\infty} \left| F(x) - F^*(x) \right|.$$

Розподіл D_n не залежить від розподілу $F(x)$ для будь-якого об'єму вибірки n, тобто побудований на ньому критерій буде непараметричним.

Якщо $F(x)$ є неперервною функцією, то для будь-якого значення x існує границя:

$$\lim_{n\to\infty} P\{\sqrt{n} D_n < x\} = K(x),$$

де функцію $K(x)$ можна виписати в явному вигляді:

$$K(x) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2x^2},$$

що визначає **функцію розподілу Колмогорова**.

Відмітимо, що даний розподіл не має параметрів. Значення розподілу $K(x)$, а також квантілів $K^{-1}(p)$ можна отримати, використовуючи статистичні таблиці.

Отже, в якості критерію слід брати величину:

$$R = \sqrt{n} D_n.$$

Правило перевірки гіпотези H_0: якщо $R > \lambda_\alpha$, то гіпотеза відхиляється, а при $R \le \lambda_\alpha$ – приймається, де $\lambda_\alpha = K^{-1}(1-\alpha)$.

Критичні значення λ_α для розподілу Колмогорова визначаються у табл. 5:

Таблиця 5

Критичні значення λ_α для розподілу Колмогорова

α	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001
λ_α	0,828	0,895	0,974	1,073	1,224	1,358	1,52	1,627	1,95

Приклад 5. Нехай вибірка подана у вигляді розподілу частот

$(x_{i-1}, x_i]$	[0;1]	(1;2]	(2;3]	(3;4]	(4;5]	(5;6]	(6;7]	(7;8]	(8;9]	(9;10]
n_i	24	15	21	13	27	19	35	16	16	14

Перевірити гіпотезу $H_0 : F_\xi(x) = \begin{cases} 1, x > 10 \\ \dfrac{x}{10}, x \in [0,10] \\ 0, x < 0 \end{cases}$ (про рівномірний розподіл на відрізку [0,10]), $\alpha = 0{,}05$.

Розв'язання. Нехай x_i^* – середина i-того інтервалу. Всі результати обчислень наведемо у табл. 6, об'єм вибірки становить n=200.

Таблиця 6

Розрахунки для прикладу 5

x_{i-1}	x_i	n_i	x_i^*	$F_n^*(x_i^*)$	$F(x_i^*)$	$\lvert F(x_i^*) - F_n^*(x_i^*) \rvert$
0	1	24	0,5	0,06	0,05	0,01
1	2	15	1,5	0,1575	0,15	0,0075
2	3	21	2,5	0,2475	0,25	0,0025
3	4	13	3,5	0,3325	0,35	0,0175
4	5	27	4,5	0,4325	0,45	0,0175
5	6	19	5,5	0,5475	0,55	0,0025
6	7	35	6,5	0,6825	0,65	0,0325
7	8	16	7,5	0,81	0,75	0,06
8	9	16	8,5	0,89	0,85	0,04
9	10	14	9,5	0,965	0,95	0,015

За результатами, що подані в таблиці, маємо

$$D_{200} = \sup_{-\infty < x < \infty} \left| F(x) - F^*(x) \right| = 0{,}06 .$$

Підрахуємо значення критерію:

$$R = \sqrt{200} D_{200} = \sqrt{200} \cdot 0{,}06 \approx 0{,}849 .$$

Квантіль розподілу Колмогорова, що відповідає ймовірності α $\lambda_\alpha = K^{-1}(1-\alpha) = 1{,}358$. Оскільки виконується умова $R \le \lambda_\alpha$, то гіпотеза приймається.

Критерій χ^2 Пірсона. Розглянемо випадок, коли $\mathrm{r} > \boldsymbol{0}$, тобто коли невідомі r параметрів розподілу і перевіряється гіпотеза про вид розподілу. При цьому одночасно розв'язується задача оцінки r невідомих параметрів. В даному випадку використовується критерій χ^2 Пірсона.

Правило перевірки гіпотез про вид розподілу має наступний вигляд:

1) Розбиваємо множину можливих значень випадкової величини на скінчене число m множин $\Delta_1, \Delta_2, ..., \Delta_m$, що не перетинаються, m – кількість множин (інтервалів).
2) Визначаємо $n_i = N\{\xi \in \Delta_i\}$ – кількість спостережень у вибірці значень, що належать і-тій множині Δ_i, $n = n_1 + n_2 + ... + n_m$.
3) Визначаємо число параметрів для оцінки (позначаємо r).
4) Оцінюємо невідомі параметри розподілу (якщо це необхідно).
5) Визначаємо p_i – ймовірності того, що випадкова величина, що має визначений гіпотезою розподіл, прийме значення із множини Δ_i: $p_i = P_{H_0}\{\xi \in \Delta_i\}$ або оцінюємо шукані ймовірності $\hat{p}_i$.

Наприклад, якщо множини $\{\Delta_i\}$ є інтервалами і лівий, і правий кінці і-того інтервалу Δ_i є z_{i-1} та z_i, то ймовірність p_i оцінюється за формулою:

$$\hat{p}_i = P\{z_{i-1} < \xi \le z_i\} = F_{\hat{\theta}}(z_i) - F_{\hat{\theta}}(z_{i-1}).$$

При цьому слід брати до уваги зауваження щодо обчислення ймовірності $\hat{p}_1$ та $\hat{p}_m$ ($z_0 = -\infty, z_m = +\infty$). $\sum_{i=1}^{m} p_i = 1$

6) Знаходимо np_i або $n\hat{p}_i$ для кожного інтервалу, якщо для деяких інтервалів $n\hat{p}_i < 5$, то їх об'єднуємо із сусідніми, при цьому після об'єднання визначаємо нове m.
7) За даними вибірки підраховуються значення критерію:

$$R = \sum_{i=1}^{m} \frac{(n_i - np_i)^2}{np_i} \text{ або } R = \sum_{i=1}^{m} \frac{(n_i - n\hat{p}_i)^2}{n\hat{p}_i}.$$

8) Визначаємо $\chi^2_{1-\alpha}(m-r-1)$ - квантіль χ^2-розподілу із $m-r-1$ ступенями свободи, що відповідає ймовірності $1-\alpha$.

9) Приймаємо рішення: гіпотеза відхиляється, якщо $R > \chi^2_{1-\alpha}(m-r-1)$, гіпотези приймається, якщо $R \le \chi^2_{1-\alpha}(m-r-1)$.

Приклад 6. Перевірка гіпотези про закон розподілу дискретної випадкової величини при незгрупованих даних. Використовуючи рівень значущості 0,1 перевірити гіпотезу про те, що наступні дані є вибіркою об'єму 100 з генеральної сукупності, що має розподіл Пуассона.

1	0	3	1	0	2	3	4	2	1
5	4	2	1	0	2	2	0	1	2
1	3	3	1	3	2	1	3	2	1
2	0	0	0	2	3	2	2	2	3
4	2	2	2	0	1	2	2	1	2
1	0	2	6	2	3	0	2	0	4
3	2	2	1	1	1	2	3	2	2
1	0	3	2	1	0	0	1	1	1
1	2	2	4	3	5	0	4	1	1
2	2	3	2	1	0	3	3	2	2

Розв'язання.

Об'єм вибірки становить n=100. Рівень значущості $\alpha = 0{,}1$. Кількість невідомих параметрів $r = 1$. Визначаємо незміщену оцінку для параметра λ :

$$\hat{\lambda} = \sum_{i=1}^{m} x_i n_i / n = 1{,}82 \text{ (табл. 7).}$$

Враховуючи $\hat{\lambda}$, знаходимо оцінки ймовірностей, (табл. 7, стовп. 4):

$$\hat{p}_i = \frac{\hat{\lambda}}{x_i!} e^{-\hat{\lambda}},\ i = \overline{1, m}\ .$$

Таблиця 7

Розрахунки для прикладу 6

x_i	n_i	$x_i n_i$	$\hat{p}_i$	$n\hat{p}_i$
0	16	0	0,1620	16,20
1	24	24	0,2949	29,49
2	35	70	0,2683	26,83
3	16	48	0,1628	16,28
4	6	24	0,0741	7,41
5	2	10	0,0270	2,70
6	1	6	0,0082	0,82
≥ 7	0	0	0,0027	
Сума	100	182	1	-

Для значень 5, 6, 7 умова $n\hat{p}_i \geq 5$ не виконується (табл. 7), тому об'єднуємо відповіді їм рядки із сусіднім, що відповідає $x_5 = 4$. В результаті отримаємо значення, що наведені в таблиці, та матимемо $m = 5$.

Таблиця 8

Розрахунки для прикладу 6, етап об'єднання

x_i	n_i	$n\hat{p}_i\}$	$\frac{(n_i - n\hat{p}_i)^2}{n\hat{p}_i}$
0	16	16,20	0,0025
1	24	29,49	1,0216
2	35	26,83	2,4845
3	16	16,28	0,0048
≥ 4	9	10,92	0,3425
Сума	100	-	3,856

За даними табл. 8 обчислюємо значення χ^2 - критерію:

$$R = \sum_{i=1}^{m} \frac{(n_i - n\hat{p}_i)^2}{n\hat{p}_i} = 3{,}856.$$

Оскільки за вибіркою оцінювався лише один параметр λ, то $r=1$ та кількість ступенів свободи становить $m-r-1=5-1-1=3$.

За таблицею квантілів χ^2-розподілу знаходимо $\chi^2_{0,9}(5)=6{,}2514$. Гіпотеза про розподіл Пуассона приймається, оскільки $R<6{,}2514$.

Приклад 7. Перевірка гіпотези про закон розподілу неперервної випадкової величини. Використовуючи рівень значущості 0,05 за спостереженнями, що наведені у таблиці, перевірити гіпотезу, що випадкова величина має нормальний розподіл:

Інтервал	[–4;0)	[0;2)	[2;4)	[4;6)
n_i	20	40	30	10

Розв'язання.

Обчислимо незміщені оцінки параметрів нормального розподілу за поданою вибіркою:

$$\bar{x}=\frac{-2\cdot 20+1\cdot 40+3\cdot 30+5\cdot 10}{100}=\frac{140}{100}=1{,}4\,.$$

$$\hat{S}^2=\frac{\sum\limits_{i=1}^{n}\left(z_i^*-\hat{\nu}_1\right)^2 n_i}{n-1}=$$

$$=\frac{(-2-1{,}4)^2\cdot 20+(1-1{,}4)^2\cdot 40+(3-1{,}4)^2\cdot 30+(5-1{,}4)^2\cdot 10}{100-1}=$$

$$=4{,}48$$

Середньоквадратичне відхилення**:** $\hat{S}=\sqrt{4{,}48}=2{,}1$**.**

Підрахуємо оцінки ймовірностей $\mathrm{p_i}$:

$$\hat{p}_1=P\{-\infty\le\xi<0\}=P\{-\infty<\frac{\xi-m}{S}\le\frac{0-1{,}4}{2{,}1}\}=$$

$$=\Phi(-\infty)-\Phi(-2{,}5499)=0{,}2543-0=0{,}2543;$$

$$\hat{p}_2=P\{0\le\xi<2\}=P\{\frac{0-1{,}4}{2{,}1}<\frac{\xi-m}{S}\le\frac{2-1{,}4}{2{,}1}\}=$$

$$=\Phi(0{,}2833)-\Phi(-0{,}6611)=0{,}6115-0{,}2543=0{,}3572;$$

$$\hat{p}_3 = P\{2 \le \xi < 4\} = P\{\frac{2-1,4}{2,1} < \frac{\xi - m}{S} \le \frac{4-1,4}{2,1}\} =$$

$$= \Phi(1,2277) - \Phi(0,2833) = 0,8902 - 0,6115 = 0,2787;$$

$$\hat{p}_4 = P\{4 \le \xi < +\infty\} = P\{\frac{4-1,4}{2,1} < \frac{\xi - m}{S} < +\infty\} =$$

$$= \Phi(+\infty) - \Phi(1,2277) = 1 - 0,8902 = 0,1098.$$

Таблиця 9

Розрахунки для прикладу 7

Інтервали		n_i	z_i^*	$z_i^* n_i$	$\hat{p}_i$	$n\hat{p}_i$	$\frac{(n_i - n\hat{p}_i)^2}{n\hat{p}_i}$
-4	0	20	-2	-40	0,2543	25,43	1,1587
0	2	40	1	40	0,3573	35,73	0,5115
2	4	30	3	90	0,2787	27,87	0,1630
4	6	10	5	50	0,1098	10,98	0,0871
Сума		100		140	1		1,9202

$$R = \chi^2 = \sum_{i=1}^{4} \frac{(n_i - n\hat{p}_i)^2}{n\hat{p}_i} = 1,9202.$$

Кількість ступенів свободи: $m - r - 1 = 4 - 2 - 1 = 1$.

$\chi^2_{1-\alpha}(m - r - 1) = \chi^2_{0,95}(1) = 3,8415$.

Оскільки $1,9202 < 3,8415$, то гіпотеза приймається.

Завдання.

1. Використовуючи рівень значущості α перевірити гіпотезу про те, що дані V_i є вибіркою об’єму $n = 60$ з генеральної сукупності, що має розподіл Пуассона, де $\alpha = 0,01 + k/100$, $k = остача(i/10)$, i – варіант.
2. Для зрупованих даних $\boldsymbol{Y_i}$, за допомогою критерію χ^2 Пірсона перевірити гіпотезу про нормальний розподіл при рівні значущості $\alpha = 0,01 + k/100$, $k = остача(i/10)$, i – варіант.

i	***Вибірка V_i***																			
1	3	7	10	5	5	3	6	7	8	5	4	7	7	5	7	7	7	2	5	4
	6	3	5	7	8	6	3	5	5	5	5	3	5	7	6	7	5	7	7	5
	4	4	8	6	6	5	6	6	9	5	8	4	7	1	6	7	4	10	4	7
2	2	5	6	6	3	8	3	8	9	5	5	5	4	5	7	5	2	9	5	4
	13	8	3	1	12	6	2	7	8	4	7	7	5	5	2	9	5	1	8	8
	8	6	2	10	5	6	8	2	6	6	8	4	5	5	9	4	8	9	5	3
3	5	5	4	6	8	4	4	8	3	9	6	5	5	8	7	8	4	7	4	7
	4	5	3	9	6	4	4	6	13	5	7	7	9	9	8	12	7	8	5	3
	10	7	7	9	8	6	4	6	12	5	9	6	4	3	5	5	4	7	5	6
4	8	8	1	7	2	9	4	7	7	7	8	10	1	3	6	3	6	3	8	7
	8	6	8	5	8	6	9	5	4	4	5	6	6	7	9	6	5	4	7	8
	12	3	8	10	4	7	5	7	7	7	4	9	2	9	5	6	9	3	6	9
5	6	5	3	9	7	6	5	2	8	6	5	8	4	6	4	3	7	5	9	6
	7	4	4	6	6	6	6	6	4	8	2	6	7	7	9	12	4	8	5	4
	4	4	5	2	9	0	8	6	5	4	7	3	11	5	7	2	7	12	5	7
6	5	6	6	6	5	10	8	5	4	4	8	6	5	4	6	9	1	6	8	7
	7	3	5	4	8	6	5	6	6	5	7	5	7	2	7	1	4	10	9	2
	5	9	6	6	5	6	10	7	9	1	3	6	0	9	9	7	6	5	8	4

i	***Вибірка V_i***																			
7	7	7	3	6	5	5	9	8	5	8	6	5	5	5	8	10	2	7	8	7
	2	5	9	7	9	4	4	6	4	1	7	9	5	6	10	9	5	2	9	3
	4	10	7	7	4	12	5	6	7	3	8	7	6	8	8	4	3	5	5	5
8	10	5	6	5	6	3	5	4	4	3	8	4	9	6	3	2	9	6	2	8
	2	5	4	5	5	3	9	9	4	2	7	4	7	6	8	9	7	4	8	1
	4	4	6	5	5	8	5	8	3	6	8	5	2	6	2	3	8	6	4	4
9	9	2	1	7	3	6	11	10	3	5	9	7	14	7	8	8	7	4	5	10
	4	6	11	6	6	3	11	9	4	6	4	6	10	3	3	7	6	6	9	9
	2	8	6	11	7	10	3	3	7	4	8	5	8	3	2	6	7	9	9	7
10	7	6	2	5	6	8	7	2	5	4	6	7	7	4	8	10	8	6	6	6
	11	4	6	7	9	5	7	7	7	3	4	4	8	7	7	6	9	9	6	4
	6	5	6	2	7	5	10	4	3	7	3	1	4	7	8	3	5	7	6	6
11	5	3	8	3	6	5	7	8	9	3	7	10	4	9	8	7	4	3	8	8
	5	4	7	6	5	8	3	9	8	4	2	4	7	6	7	9	6	5	3	8
	9	6	6	4	5	8	4	6	2	6	5	5	10	5	8	3	9	7	6	7

i	***Вибірка V_i***																			
12	3	5	5	9	3	5	5	2	3	4	7	8	7	4	4	5	11	4	8	5
	8	2	10	6	8	7	6	7	7	2	3	5	3	2	6	10	4	4	4	5
	5	10	2	12	10	6	3	3	3	6	5	3	4	9	6	5	9	9	6	12
13	0	5	6	6	4	7	4	7	5	7	6	4	5	11	4	6	4	9	3	5
	6	2	5	6	4	2	3	3	4	7	2	8	7	4	6	5	8	8	9	5
	6	7	9	8	5	9	7	2	5	10	4	6	6	1	7	9	5	7	9	5
14	2	7	6	9	6	7	6	3	6	5	8	7	7	7	7	11	2	7	9	2
	9	3	3	4	3	6	5	11	5	4	6	9	1	7	7	5	4	2	7	9
	1	9	7	9	9	6	6	4	5	5	6	6	6	8	6	6	6	8	7	5
15	5	7	11	1	6	10	7	4	6	4	2	8	5	5	2	2	6	8	6	5
	4	8	2	8	7	4	3	5	6	4	4	5	8	4	5	9	3	5	10	5
	7	10	5	5	5	4	10	11	5	9	8	5	12	2	3	2	8	10	10	2
16	7	11	3	2	6	6	3	5	3	4	4	6	2	5	7	5	9	9	3	5
	7	6	5	7	7	7	6	6	8	2	12	3	7	6	4	7	5	10	6	4
	4	5	6	6	6	9	6	4	10	5	5	3	3	2	4	6	7	6	4	8

i	***Вибірка V_i***																			
17	8	7	4	6	5	7	2	6	9	5	3	13	4	5	6	1	6	5	3	4
	9	3	4	2	6	6	6	5	6	6	5	6	6	6	1	4	9	4	7	3
	4	13	6	3	7	6	7	6	7	9	2	8	3	4	5	5	11	8	6	7
18	6	5	7	9	9	6	7	5	11	7	8	5	3	8	5	11	4	11	6	10
	4	5	4	14	5	5	9	4	7	7	8	6	13	5	8	5	9	7	4	12
	6	9	5	4	4	6	10	4	5	3	10	11	3	2	3	4	7	5	6	7
19	9	3	7	3	5	2	8	9	6	3	4	1	5	7	8	6	9	7	9	8
	8	5	6	7	4	3	3	3	5	5	9	3	7	6	3	11	8	4	13	7
	8	1	6	2	8	7	5	7	10	5	7	3	9	4	9	8	6	6	5	5
20	4	5	8	7	8	4	1	10	4	5	6	2	6	8	8	7	3	8	8	5
	5	9	5	3	4	4	4	12	3	7	6	9	4	5	4	10	8	4	2	5
	1	4	4	9	8	4	2	6	4	8	9	9	6	9	5	4	3	7	4	13

Вибірка Yᵢ

Варіант	**Інтервали**		**Частота**		**Варіант**	**Інтервали**		**Частота**
	z_{i-1}	x_i	n_i			z_{i-1}	x_i	n_i
1	-8	-2	11		2	0	3,5	36
	-2	4	50			3,5	7	23
	4	10	30			7	10,5	23
	10	16	10			10,5	14	100
	16	22	12			14	17,5	22
	22	28	15			17,5	21	76
3	3,0	3,6	12		4	-3	1	15
	3,6	4,2	18			1	5	75
	4,2	4,8	35			5	9	15
	4,8	5,4	43			9	13	50
	5,4	6,0	22			13	17	81
	6,0	6,6	10			17	21	54
5	0	5	15		6	-3	1	15
	5	10	75			1	5	75

Варіант	***Інтервали***		***Частота***		***Варіант***	***Інтервали***		***Частота***
	z_{i-1}	x_i	n_i			z_{i-1}	x_i	n_i
	10	15	100			5	9	15
	15	20	50			9	13	50
	20	25	12			13	17	81
	25	30	20			17	21	54
7	-3	2	13		8	-8	-6	13
	2	7	15			-6	-4	15
	7	12	24			-4	-2	24
	12	17	25			-2	0	25
	17	22	13			0	2	13
	22	27	10			2	4	10
9	-8	-3	30		10	-4	0	25
	-3	2	50			0	4	12
	2	7	12			4	8	9
	7	12	10			8	12	23
	12	17	12			12	16	18
	17	22	40			16	20	10
11	2	5	12		12	-2,5	0,5	40
	5	8	18			0,5	3,5	20

Варіант	Інтервали		Частота		Варіант	Інтервали		Частота
	z_{i-1}	x_i	n_i			z_{i-1}	x_i	n_i
	8	11	23			3,5	6,5	14
	11	14	43			6,5	9,5	9
	14	17	22			9,5	12,5	24
	17	20	10			12,5	15,5	22
13	0	4	15		14	-3	1	23
	4	8	75			1	5	17
	8	12	60			5	9	69
	12	16	50			9	13	23
	16	20	24			13	17	13
	20	24	35			17	21	56
15	-3	3	34		16	-1	3	46
	3	9	35			3	7	50
	9	15	24			7	11	16
	15	21	90			11	15	50
	21	27	13			15	19	10
	27	33	50			19	23	35
17	0	4	30		18	0	3	13

Варіант	Інтервали		Частота		Варіант	Інтервали		Частота
	z_{i-1}	x_i	n_i			z_{i-1}	x_i	n_i
	4	8	50			3	6	45
	8	12	23			6	9	23
	12	16	10			9	12	78
	16	20	12			12	15	22
	20	24	45			15	18	76
19	-1	2	50		20	1	5	23
	2	5	23			5	9	75
	5	8	23			9	13	25
	8	11	43			13	17	36
	11	14	22			17	21	13
	14	17	10			21	25	45

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

1. Бахрушин В. Є. Методи аналізу даних : навчальний посібник для студентів. Запоріжжя : КПУ, 2011. 268 с.
2. Карташов М. В. Імовірність, процеси, статистика. К. : Видавничо-поліграфічний центр «Київський університет», 2008. 494 с. URL: http://probability.univ.kiev.ua/userfiles/kmv/ VPS_Pv.pdf.
3. Лебєдєв Є. О., Лівінська Г. В., Розора І. В., Шарапов М. М. Математична статистика : навчальний посібник. К. : ВПЦ «Київський університет», 2016. 159 с.
4. Приставка П. О., Приставка П. О. Аналіз даних : навч. посіб. / МОН України. Д. : РВВ ДНУ, 2008. 92 с.
5. Турчин В. М. Теорія ймовірностей і математична статистика. Основні поняття, приклади та задачі : підручник для студентів вищих навчальних закладів. Дніпропетровськ : ІМА-прес, 2014. 556 с.
6. Introduction to Probability, Statistics and Random processes. URL: *https://www.probabilitycourse.com/*
7. Statistics and probability. Khan Academy. *URL: https://www.khanacademy.org/math/statistics-probability.*