

Сумський державний університет

Кафедра

Прикладної математики та моделювання складних систем

ЗВІТ

Обов'язкове домашнє завдання

Завдання 12

Дисципліна

Теорія ймовірностей та математична статистика

Варіант 8

Виконавець:

студентка групи ПМ-81

Пороскун Олена Олегівна

Викладач:

Гончаров Олександр Андрійович

Суми, Сумська область

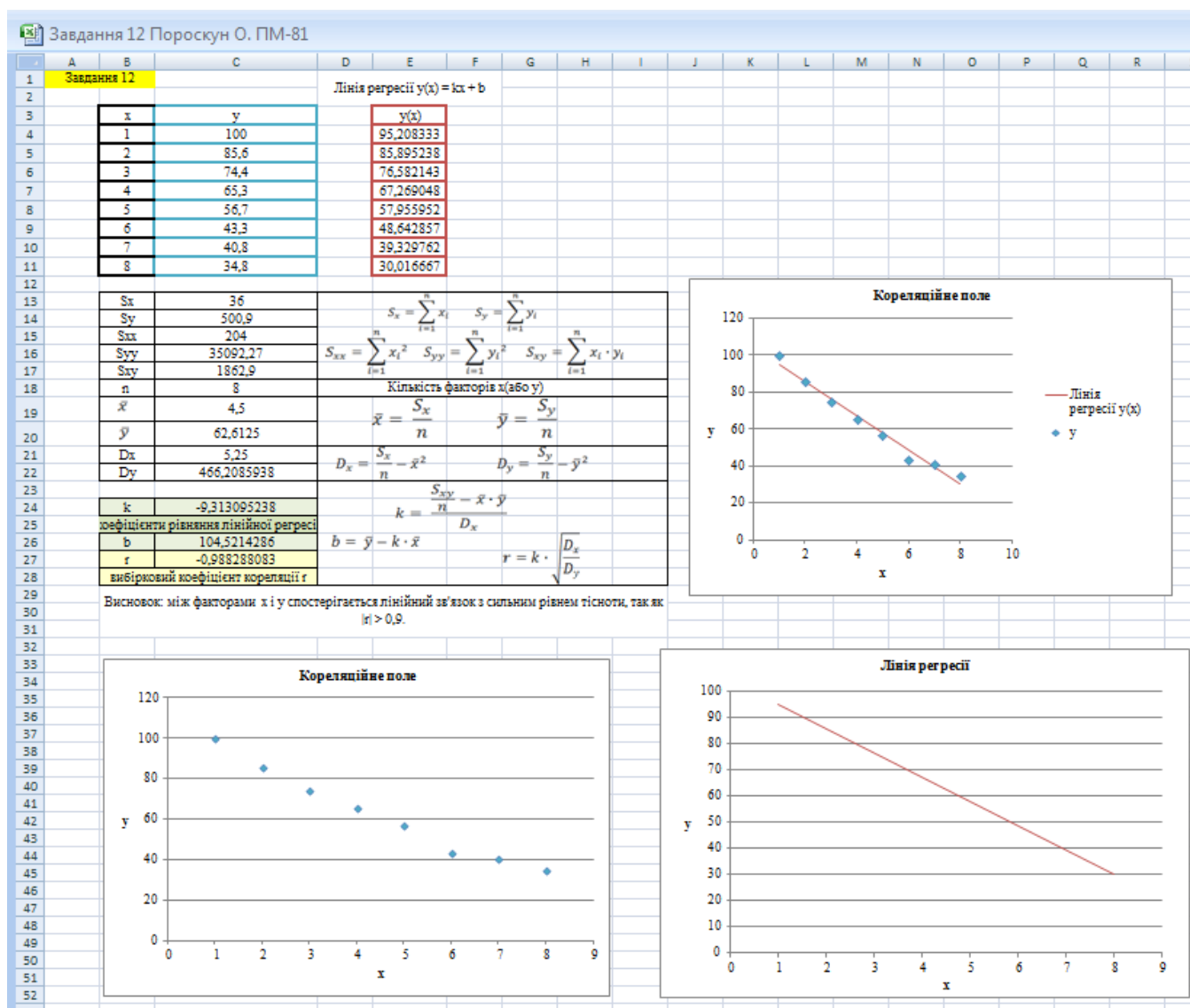
2020

12 Дані експерименту наведені в таблиці 21 в безрозмірному вигляді. Потрібно:

- побудувати кореляційне поле;
- висловити гіпотезу про вид статистичної залежності між X і Y , визначити коефіцієнт кореляції і тісноту лінійного зв'язку;
- знайти рівняння лінії регресії;
- побудувати лінію регресії.

Таблиця 21

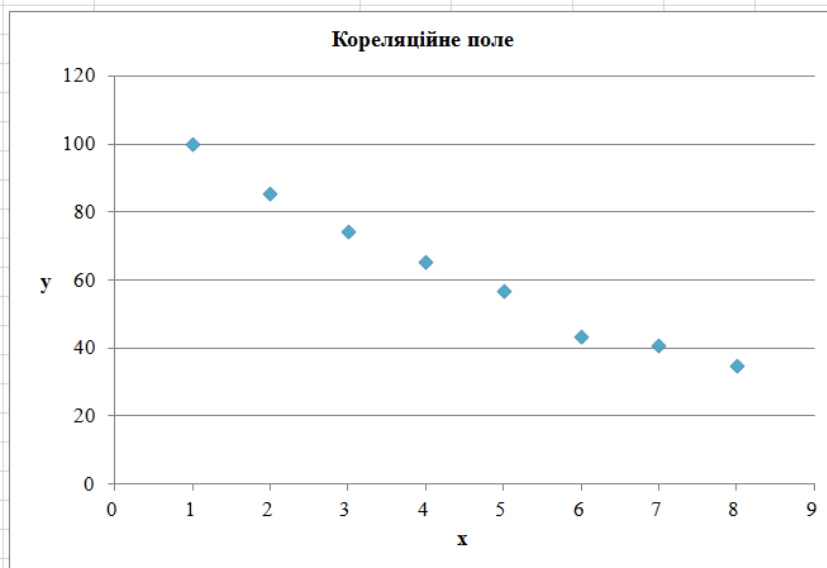
X	1	2	3	4	5	6	7	8
Y	100	85,6	74,4	65,3	56,7	43,3	40,8	34,8



	A	B	C	D	E	F	G	H	I	
1	Завдання 12			Лінія регресії $y(x) = kx + b$						
2										
3		x	y		y(x)					
4		1	100		95,2083					
5		2	85,6		85,8952					
6		3	74,4		76,5821					
7		4	65,3		67,269					
8		5	56,7		57,956					
9		6	43,3		48,6429					
10		7	40,8		39,3298					
11		8	34,8		30,0167					
12										
13		Sx	36	$S_x = \sum_{i=1}^n x_i \quad S_y = \sum_{i=1}^n y_i$ $S_{xx} = \sum_{i=1}^n x_i^2 \quad S_{yy} = \sum_{i=1}^n y_i^2 \quad S_{xy} = \sum_{i=1}^n x_i \cdot y_i$						
14		Sy	500,9							
15		Sxx	204							
16		Syy	35092,27							
17		Sxy	1862,9							
18		n	8	Кількість факторів x(або y)						
19		\bar{x}	4,5	$\bar{x} = \frac{S_x}{n}$		$\bar{y} = \frac{S_y}{n}$				
20		\bar{y}	62,6125							
21		Dx	5,25	$D_x = \frac{S_x}{n} - \bar{x}^2$		$D_y = \frac{S_y}{n} - \bar{y}^2$				
22		Dy	466,2085938							
23										
24		k	-9,313095238	$k = \frac{\frac{S_{xy}}{n} - \bar{x} \cdot \bar{y}}{D_x}$						
25		коефіцієнти рівняння лінійної регресії		$b = \bar{y} - k \cdot \bar{x}$ $r = k \cdot \sqrt{\frac{D_x}{D_y}}$						
26		b	104,5214286							
27		r	-0,988288083							
28		вибірковий коефіцієнт кореляції r								
29	Висновок: між факторами x і y спостерігається лінійний зв'язок з сильним рівнем тісноти, так як $ r $									
30	$> 0,9$.									
31										

	A	B	C	D	E	F	G	H	I
25		коефіцієнти рівняння лінійної регресії			\bar{y}	D_x			
26		b	104,5214286	$b = \bar{y} - k \cdot \bar{x}$			$r = k \cdot \sqrt{\frac{D_x}{D_y}}$		
27		r	-0,988288083						
28		вибірковий коефіцієнт кореляції r							

Висновок: між факторами x і y спостерігається лінійний зв'язок з сильним рівнем тісноти, так як $|r| > 0,9$.



Дивлячись на графік кореляційного поля можна висловити гіпотезу про лінійну залежність між X і Y.

Формули для знаходження числових характеристик:

$$S_x = \sum_{i=1}^n x_i \quad S_y = \sum_{i=1}^n y_i$$

$$S_{xx} = \sum_{i=1}^n x_i^2 \quad S_{yy} = \sum_{i=1}^n y_i^2 \quad S_{xy} = \sum_{i=1}^n x_i \cdot y_i$$

n = обсяг величин X (та Y)

$$\bar{x} = \frac{S_x}{n} \quad \bar{y} = \frac{S_y}{n}$$

$$D_x = \frac{S_x}{n} - \bar{x}^2 \quad D_y = \frac{S_y}{n} - \bar{y}^2$$

Коефіцієнти рівняння лінійної регресії:

$$k = \frac{\frac{S_{xy}}{n} - \bar{x} \cdot \bar{y}}{D_x} = -9,313095238 \quad b = \bar{y} - k \cdot \bar{x} = 104,5214286$$

Рівняння лінійної регресії: $y(x) = kx + b$

Вибірковий коефіцієнт кореляції обчислюється за формулою:

$$r = k \cdot \sqrt{\frac{D_x}{D_y}}$$

Можна оцінити тісноту лінійного зв'язку між факторами за шкалою:

$|r| < 0,6$ – слабка

$0,6 \leq |r| \leq 0,9$ – середня

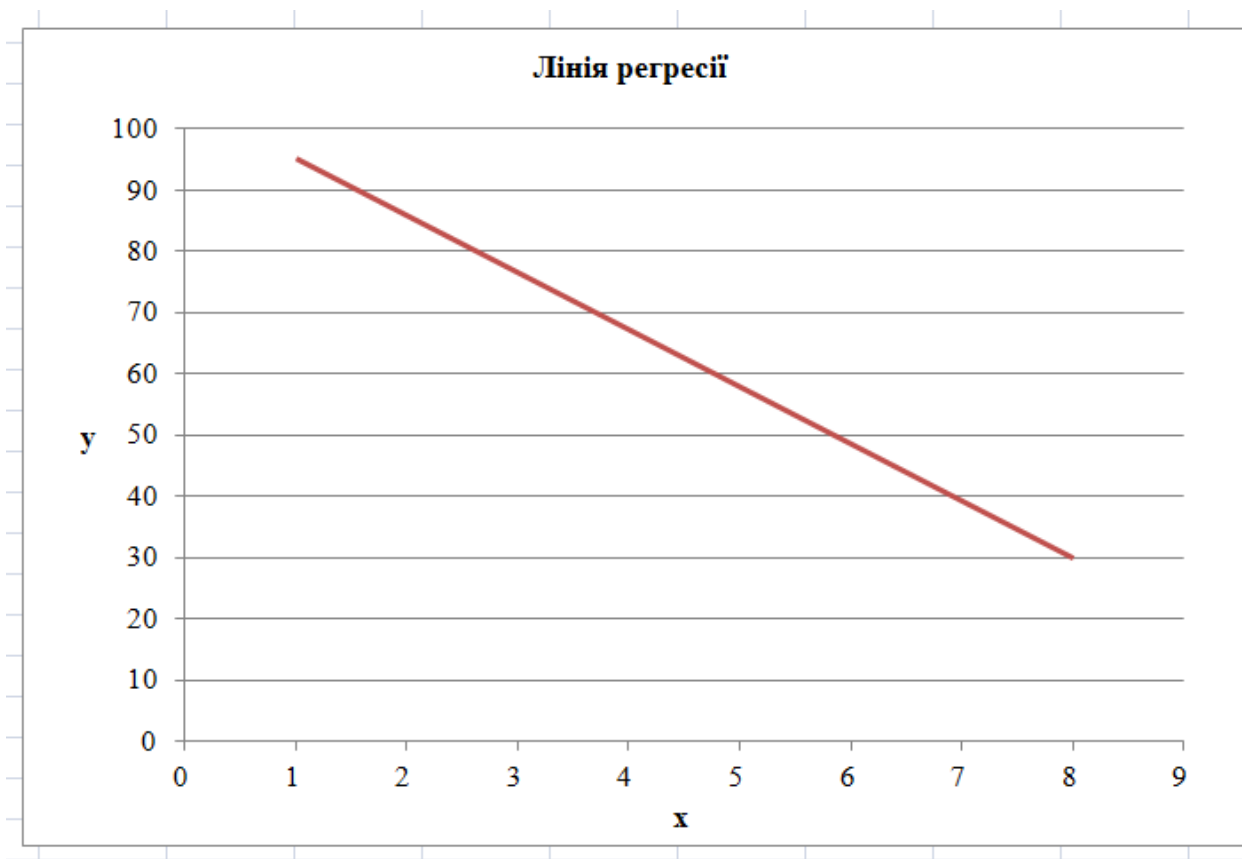
$|r| > 0,9$ – сильна.

Висновок: між факторами X і Y спостерігається лінійний зв'язок з сильним рівнем тісноти, так як $|r| = 0,988288083 > 0,9$.

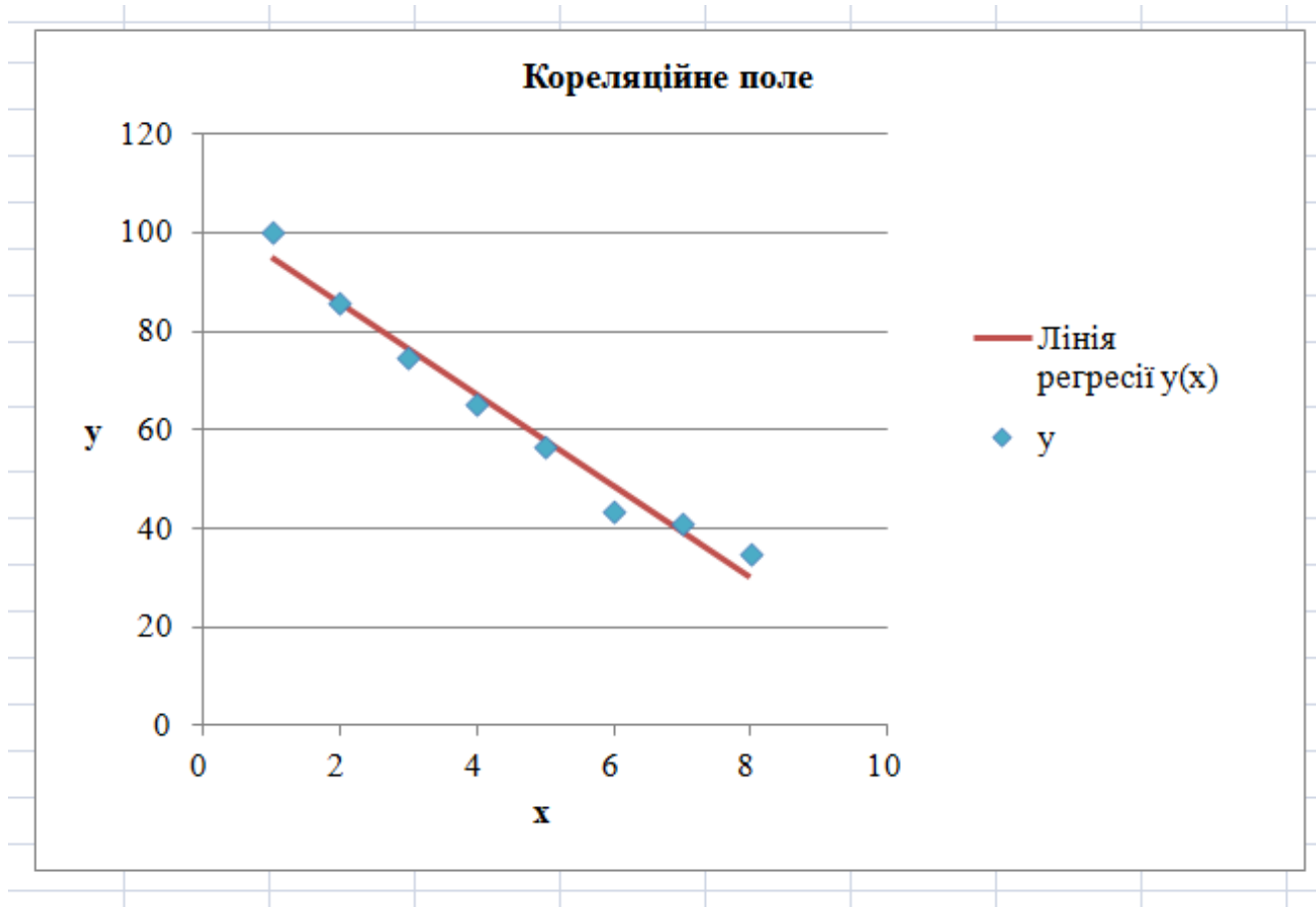
	A	B	C	D	E	F	G	H	I
1	Завдання 12				Лінія регресії $y(x) = kx + b$				
2									
3		x	y		y(x)				
4		1	100		=C\$24*\$B4+C\$26				
5		2	85,6		=C\$24*\$B5+C\$26				
6		3	74,4		=C\$24*\$B6+C\$26				
7		4	65,3		=C\$24*\$B7+C\$26				
8		5	56,7		=C\$24*\$B8+C\$26				
9		6	43,3		=C\$24*\$B9+C\$26				
10		7	40,8		=C\$24*\$B10+C\$26				
11		8	34,8		=C\$24*\$B11+C\$26				
12									
13		Sx	=СУММ(B4:B11)		$S_x = \sum_{i=1}^n x_i$	$S_y = \sum_{i=1}^n y_i$			
14		Sy	=СУММ(C4:C11)						
15		Sxx	=B4^2+B5^2+B6^2+B7^2+B8^2+B9^2+B10^2+B11^2		$S_{xx} = \sum_{i=1}^n x_i^2$	$S_{yy} = \sum_{i=1}^n y_i^2$	$S_{xy} = \sum_{i=1}^n x_i \cdot y_i$		
16		Syy	=C4^2+C5^2+C6^2+C7^2+C8^2+C9^2+C10^2+C11^2						
17		Sxy	=B4*C4+B5*C5+B6*C6+B7*C7+B8*C8+B9*C9+B10*C10+B11*C11						
18		n	=СЧЁТ(B4:B11)		Кількість факторів x(або y)				
19		\bar{x}	=C13/C18		$\bar{x} = \frac{S_x}{n}$	$\bar{y} = \frac{S_y}{n}$			
20		\bar{y}	=C14/C18						
21		Dx	=C15/C18-C19^2		$D_x = \frac{S_x}{n} - \bar{x}^2$	$D_y = \frac{S_y}{n} - \bar{y}^2$			
22		Dy	=C16/C18-C20^2						
23									
24		k	=(C17/C18-C19*C20)/C21		$k = \frac{S_{xy}}{n} - \bar{x} \cdot \bar{y}$				
25		коефіцієнти рівняння лінійної регресії			D_x				
26		b	=C20-C24*C19		$b = \bar{y} - k \cdot \bar{x}$				
27		r	=C24*КОРЕНЬ(C21/C22)				$r = k \cdot \sqrt{\frac{D_x}{D_y}}$		
28		вибірковий коефіцієнт кореляції r							
29									
30									
31									

Висновок: між факторами x і y спостерігається лінійний зв'язок з сильним рівнем тісноти, так як $|r| > 0,9$.

Лінія регресії:



Лінія регресії та кореляційне поле:



Сумський державний університет

Кафедра

Прикладної математики та моделювання складних систем

ЗВІТ

Завдання 18

Дисципліна

Теорія ймовірностей та математична статистика

Варіант 8

Виконавець:

студентка групи ПМ-81

Пороскун Олена Олегівна

Викладач:

Гончаров Олександр Андрійович

Суми, Сумська область

2020

Завдання 18. З групи А таблиці 6.5 вибрати числа, які записані в клітинах з номерами з *набору* (див. табл. 6.4) для даного варіанта. Порожня клітина означає відсутність числа. Виписані таким чином числа назвемо вибіркою А. Аналогічно за допомогою рядка В вибираємо вибірку В. Позначимо $C = A \cup B$ (об'єднана вибірка). Виконати такі завдання:

- а) знайти вибіркові середні і дисперсії для А і В;
- б) обчислити внутрішньо- і міжгрупові дисперсії;
- в) визначити середню дисперсію вибірки С за обчисленими числовими характеристиками А і В;
- г) обчислити виправлену вибіркову дисперсію і середнє квадратичне відхилення для «С».

Для розв'язання наступних трьох завдань студенту потрібно згідно зі своїм варіантом з таблиці 6.4 вибрати так званий *набір*.

Таблиця 6.4

Варіант	1	2	3	4	5	6
набір	3 4 5 6 8	1 2 3 6 8	2 3 5 7 8	2 3 4 6 8	1 2 5 6 8	1 4 5 6 7

Продовження таблиці 6.4

Варіант	7	8	9	10	11	12
набір	1 3 5 6 8	3 5 6 7 8	1 2 3 4 7	1 3 4 7 8	1 3 4 6 7	2 3 6 7 8

Продовження таблиці 6.4

Варіант	13	14	15	16	17	18
набір	2 3 4 5 6	1 4 5 7 8	1 2 3 4 5	2 5 6 7 8	1 2 4 5 8	1 2 4 6 7

Продовження таблиці 6.4

Варіант	19	20	21	22	23	24
набір	1 4 5 6 8	1 3 5 7 8	1 2 4 5 7	2 3 4 5 7	1 2 5 6 7	3 4 5 7 8

Продовження таблиці 6.4

Варіант	25	26	27	28	29	30
набір	2 3 4 5 8	1 2 3 4 6	1 2 4 6 8	1 3 4 5 7	1 4 6 7 8	1 3 4 6 8

Таблиця 6.5

Набір	1	2	3	4	5	6	7	8
Група А		-2,4	3,7		0,8	5,8	-1,1	-3,1
Група В	-2,9	8,3	1,8	-3,7	4,2		-0,5	6,1

Вибіркова середня обчислюється за формулою:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (5.1)$$

де n – об'єм вибірки;

x_i – елементи вибірки.

Вибіркова дисперсія обчислюється за формулою:

$$D = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2. \quad (5.2)$$

Якщо вибірка складається з двох (або кількох) груп, то внутрішньогрупова дисперсія обчислюється за формулою:

$$D_{\text{вн. гр.}} = \frac{D_A \cdot n_A + D_B \cdot n_B}{n_A + n_B}, \quad (5.3)$$

де D_A, D_B – дисперсії груп A і B ,

n_A, n_B – об'єми груп A і B .

Міжгрупова дисперсія:

$$D_{\text{між}} = \frac{(\bar{x}_A - \bar{x})^2 \cdot n_A + (\bar{x}_B - \bar{x})^2 \cdot n_B}{n_A + n_B}. \quad (5.4)$$

Вибіркова дисперсія дорівнює сумі внутрішньогрупової та міжгрупової дисперсій.

Сумський державний університет

Кафедра

Прикладної математики та моделювання складних систем

ЗВІТ

Завдання 19

Дисципліна

Теорія ймовірностей та математична статистика

Варіант 8

Виконавець:

студентка групи ПМ-81

Пороскун Олена Олегівна

Викладач:

Гончаров Олександр Андрійович

Суми, Сумська область

2020

Завдання 19. З рядків таблиці 6.6 згідно зі своїм набором (див. табл. 6.4) скласти вибірку, тобто початковий числовий масив. Інші потрібні дані наведені у табл. 6.7.

Пояснення. Числа набору вказують, які рядки відносяться до даного варіанта. Так, для варіанта 8 в початковий числовий масив з таблиці 6.6 треба включити елементи рядків 3, 5, 6, 7, 8; тобто отримаємо набір з $26 + 21 + 24 + 25 + 23 = 119$ елементів. Для спрощення запису в таблиці 6.6 дев'ятьма буквами: А, В, С, D, E, F, G, H, К позначені числа, що утворюють арифметичну прогресію. Два перших члени цієї прогресії вказані в таблиці 6.7. Наприклад, для варіанта 3 з таблиці 6.7 виписуємо $A = 1,0$, $B = 1,4$. Отже, крок прогресії 0,4 і наступні букви позначають числа $C = 1,8$, $D = 2,2$ і т. д.

Виконати такі завдання:

- скласти варіаційний ряд;
- знайти розмах, медіану і моду вибірки;
- побудувати полігон частот;
- побудувати гістограму вибірки. Основи прямокутників гістограми вибрати так, щоб точки, які відповідають числам «А», «В» і т. д. опинились на середині основ. Інтервали, частота яких виявиться меншою 5, об'єднати з більш показними сусідніми інтервалами;
- обчислити вибіркове середнє, дисперсію; виправлену дисперсію і середнє квадратичне відхилення;
- обчислити теоретичні частоти для нормального закону, крайні інтервали брати напівнескінченними (сума всіх частот при цьому буде дорівнювати об'єму вибірки);
- накласти теоретичну криву на гістограму;
- обчислити суму Пірсона;
- визначити кількість ступенів вільності, з рівнем значущості α перевірити гіпотезу про нормальний розподіл;
- визначити ймовірність попадання в інтервал $[a; b]$ двома способами – за відносною частотою і за теоретичною функцією розподілу;
- знайти інтервали довіри для числових параметрів «а» і « σ » нормально розподіленої генеральної сукупності.

Таблиця 6.6

Набір	Елементи вибірки	шт.
1	F E E F C E D E B E F D G D H C G B E B C C D D	24
2	B D E D G K G F D E D C H E D C G E K C	20
3	C E F F C D F G G D G C E K E H D E E F D D C G K A	26
4	E F D E D E K F E F A G E C E G G D D B D F H E G H	26
5	E F G F H E B G E G E K F B E B D G E F K	21
6	B D F F F F H D C E D A G K D E D B E F F K D G	24
7	D F H F D C D H G G E H B K E G D B B A D E G D E	25
8	H E H E E F H G E D D C K A E D B F D C D G C	23

Таблиця 6.7

Варіант	A	B	a	b	γ	α
<i>l</i>	2	3	4	5	6	7
1	0,6	1,5	3,3	4,8	0,999	0,025
2	0,8	1,2	1,2	1,7	0,999	0,025
3	1,0	1,4	1,4	1,9	0,95	0,05
4	2,8	3,6	3,6	4,9	0,99	0,05
5	0,5	1,0	1,5	2,2	0,99	0,05
6	0,5	1,0	1,0	1,7	0,999	0,05
7	1,2	1,6	2,4	2,9	0,999	0,01
8	1,9	2,7	3,5	4,8	0,999	0,05
9	0,5	1,2	1,2	2,3	0,99	0,01
10	1,7	2,3	2,9	3,8	0,95	0,05
11	1,1	1,9	2,7	4,0	0,95	0,01
12	2,1	3,0	3,9	5,4	0,99	0,02
13	1,2	1,9	3,3	4,4	0,999	0,01
14	0,7	1,2	2,2	2,9	0,99	0,05
15	0,9	1,4	1,9	2,6	0,999	0,01
16	2,9	3,6	4,3	5,4	0,95	0,025
17	2,6	3,2	4,4	5,3	0,999	0,01
18	1,2	2,1	3,9	5,4	0,99	0,05
19	1,9	2,6	4,0	5,1	0,99	0,05
20	1,7	2,6	4,4	5,9	0,99	0,025
21	1,7	2,2	3,2	3,9	0,95	0,05
22	1,5	2,2	2,9	4,0	0,99	0,05
23	1,0	1,9	2,8	4,3	0,99	0,01
24	0,7	1,2	1,7	2,4	0,95	0,05
25	0,7	1,3	1,3	2,2	0,999	0,01

Продовження таблиці 6.7

1	2	3	4	5	6	7
26	1,8	2,4	3,6	4,5	0,99	0,01
27	1,7	2,3	2,9	3,8	0,95	0,01
28	1,6	2,4	4,0	5,3	0,999	0,025
29	2,9	3,3	3,7	4,2	0,99	0,05
30	0,7	1,1	1,1	1,6	0,999	0,05

Таблиця 6.4

Варіант	1	2	3	4	5	6
набір	3 4 5 6 8	1 2 3 6 8	2 3 5 7 8	2 3 4 6 8	1 2 5 6 8	1 4 5 6 7

Продовження таблиці 6.4

Варіант	7	8	9	10	11	12
набір	1 3 5 6 8	3 5 6 7 8	1 2 3 4 7	1 3 4 7 8	1 3 4 6 7	2 3 6 7 8

Продовження таблиці 6.4

Варіант	13	14	15	16	17	18
---------	----	----	----	----	----	----

набір	2 3 4 5 6	1 4 5 7 8	1 2 3 4 5	2 5 6 7 8	1 2 4 5 8	1 2 4 6 7
-------	-----------	-----------	-----------	-----------	-----------	-----------

Продовження таблиці 6.4

Варіант	19	20	21	22	23	24
набір	1 4 5 6 8	1 3 5 7 8	1 2 4 5 7	2 3 4 5 7	1 2 5 6 7	3 4 5 7 8

Продовження таблиці 6.4

Варіант	25	26	27	28	29	30
набір	2 3 4 5 8	1 2 3 4 6	1 2 4 6 8	1 3 4 5 7	1 4 6 7 8	1 3 4 6 8

3	Набори				
4	3	5	6	7	8
5	3,5	5,1	2,7	4,3	7,5
6	5,1	5,9	4,3	5,9	5,1
7	5,9	6,7	5,9	7,5	7,5
8	5,9	5,9	5,9	5,9	5,1
9	3,5	7,5	5,9	4,3	5,1
10	4,3	5,1	5,9	3,5	5,9
11	5,9	2,7	7,5	4,3	7,5
12	6,7	6,7	4,3	7,5	6,7
13	6,7	5,1	3,5	6,7	5,1
14	4,3	6,7	5,1	6,7	4,3
15	6,7	5,1	4,3	5,1	4,3
16	3,5	9,9	1,9	7,5	3,5
17	5,1	5,9	6,7	2,7	9,9
18	9,9	2,7	9,9	9,9	1,9
19	5,1	5,1	4,3	5,1	5,1
20	7,5	2,7	5,1	6,7	4,3
21	4,3	4,3	4,3	4,3	2,7
22	5,1	6,7	2,7	2,7	5,9
23	5,1	5,1	5,1	2,7	4,3
24	5,9	5,9	5,9	1,9	3,5
25	4,3	9,9	5,9	4,3	4,3
26	4,3		9,9	5,1	6,7
27	3,5		4,3	6,7	3,5
28	6,7		6,7	4,3	
29	9,9			5,1	
30	1,9				
31					
32	A	1,9			
33	B	2,7			
34	C	3,5			
35	D	4,3			
36	E	5,1			
37	F	5,9			
38	G	6,7			
39	H	7,5			
40	K	9,9			

Виконати такі завдання:

- скласти варіаційний ряд;
- знайти розмах, медіану і моду вибірки;
- побудувати полігон частот;

	G	H	I	J	K	L	M	N	O	P
5										
6		а) скласти варіаційний ряд;								
7										
8		1,90	2,70	3,50	4,30	5,10	5,90	6,70	7,50	9,90
9		4	9	9	23	23	18	16	9	8
10										
11		б) знайти розмах, медіану і моду вибірки;								
12										
13		Обчислимо розмах вибірки $x_{\max} - x_{\min}$:								
14		x_{\max}	9,90							
15		x_{\min}	1,90							
16		$x_{\max} - x_{\min} =$	8,00							
17										
18		Медіана – це варіанта, яка ділить вибірку на дві рівні за об'ємом частини								
19										5,10
20		Мода – це варіанта ряду, що має найбільшу частоту (мод може бути декілька)								
21										4,30 5,10
22										
23		в) побудувати полігон частот;								
24										
25		Полігон частот – це ламана, що з'єднує точки з координатами (x_i , n_i).								
26										
27		Полігон частот								
28										
29										
30										
31										
32										
33										
34										
35										
36										
37										
38										
39										
40										
41										
42										

	G	H	I	J	K	L	M	N	O	P
5										
6		а) скласти варіаційний ряд;								
7										
8		1,9	2,7	3,5	4,3	5,1	5,9	6,7	7,5	9,9
9		4	9	9	23	23	18	16	9	8
10										
11		б) знайти розмах, медіану і моду вибірки;								
12										
13		Обчислимо розмах вибірки $x_{\max} - x_{\min}$:								
14		x_{\max}	=МАКС(B5:F30)							
15		x_{\min}	=МИН(B5:F30)							
16		$x_{\max} - x_{\min} =$		=I14-I15						
17										
18		Медіана – це варіанта, яка ділить вибірку на дві рівні за об'ємом частини								=МЕДИАНА(B5:F30)
19										
20		Мода – це варіанта ряду, що має найбільшу частоту (мод може бути декілька)								4,3
21										=МОДА(B5:F30)

г) побудувати гістограму вибірки. Основи прямокутників гістограми вибрати так, щоб точки, які відповідають числам «А», «В» і т. д. опинились на серединах основ. Інтервали, частота яких виявиться меншою 5, об'єднати з більш показними сусідніми інтервалами;

Для побудови гістограми потрібно вибрати на вісі ОХ основи прямокутників. У випадку варіаційного ряду з рівновіддаленими варіантами x_i , зручно розбивку $[\alpha_i : \beta_i]$ вибрати за таким правилом:

$$\alpha_i + \beta_i = 2x_i, \quad \alpha_{i+1} = \beta_i, \quad i = 1, \dots, k..$$

При побудові гістограми інтервали, які відповідають малим частотам n_i , називають непоказними і об'єднують із сусідніми. Якщо сусідніх два, то вибирають той, у якого частота n_i вища. Рівень зображуваності, взагалі кажучи, величина інтуїтивна. Прийmemo, що об'єднанню підлягають інтервали, у яких $n_i < 5$. Інколи навіть після об'єднання двох інтервалів утворюється непоказний інтервал, тоді процес об'єднання продовжується на наступний сусідній інтервал.

Об'єднання двох інтервалів описується співвідношеннями:

$$\begin{array}{ccc} [\alpha_j ; \beta_j] & \cup & [\alpha_{j+1} ; \beta_{j+1}] \rightarrow [\alpha_j ; \beta_{j+1}] \\ n_j & & n_{j+1} \quad n_j + n_{j+1} \end{array}$$

Після об'єднання інтервалів виконується їх перенумерація, щоб індекс «і» приймав всі значення підряд без пропусків. В результаті зміни інтервалів варіаційний ряд дещо видозмінюється: частоти відносяться не до чисел, а до інтервалів – показують, скільки елементів з вибірки попадає в даний інтервал. Запишемо інтервальний варіаційний ряд в таблицю 5.2 (стовпці «і», «інтервал», « δ_i » – довжина інтервалу, $\delta_i = \beta_i - \alpha_i$).

Для обчислення висот прямокутників треба вибрати масштаб, це можна зробити таким чином. Прямокутник найбільшої висоти буде відповідати моді варіаційного ряду. Тому, для того щоб рисунок гістограми ефективно

використовував площу рисунка висотою Н, масштабний множник μ повинен задовольняти співвідношенню:

$$S = H \cdot \delta = \mu \cdot n_M \qquad \mu = \frac{H \cdot \delta}{n_M}$$

де δ і S – довжина основи і площа прямокутника, який відповідає моді варіаційного ряду, n_M – частота моди (найбільша частота в варіаційному ряду). Тепер висота будь-якого прямокутника гістограми буде обчислюватись за формулою:

$$h_i = \mu \frac{n_i}{\delta_i}$$

У даному завданні, узявши висоту 80 мм, отримаємо:

$$\mu = \frac{H \cdot \delta}{n_M} = \frac{80 \cdot 0,8}{23} = 2,78$$

	Q	R	S	T	U	V	W	X	Y
4									
5		<p>г) побудувати гістограму вибірки. Основи прямокутників гістограми вибрати так, щоб точки, які відповідають числам «А», «В» і т. д. опинились на середині основ. Інтервали, частота яких виявиться меншою 5, об'єднати з більш показними сусідніми інтервалами;</p>							
6									
7									
8									
9		$S = H \cdot \delta = \mu \cdot n_M \quad \mu = \frac{H \cdot \delta}{n_M} \quad \delta_i = \beta_i - \alpha_i$							
10									
11		δ_i	0,8		n_M	23			
12		У даному прикладі, узявши висоту 80 мм, отримаємо							
13		μ	2,78						
14		де δ і S – довжина основи і площа прямокутника, який відповідає моді варіаційного ряду,							
15									
16		n_M – частота моди (найбільша частота в варіаційному ряду).							
17									
18		Тепер висота будь-якого прямокутника гістограми буде обчислюватись за формулою:					$h_i = \mu \frac{n_i}{\delta_i}$		
19									

	Q	R	S	T	U	V	W	X	Y
4									
5		<p>г) побудувати гістограму вибірки. Основи прямокутників гістограми вибрати так, щоб точки, які відповідають числам «А», «В» і т. д. опинились на середині основ. Інтервали, частота яких виявиться меншою 5, об'єднати з більш показними сусідніми інтервалами;</p>							
6									
7									
8									
9		$S = H \cdot \delta = \mu \cdot n_M \quad \mu = \frac{H \cdot \delta}{n_M} \quad \delta_i = \beta_i - \alpha_i$							
10									
11		δ_i	0,8		n_M	=МАКС(H9:P9)			
12		У даному прикладі, узявши висоту 80 мм, отримаємо							
13		μ	= (80*S11)/V11						
14		де δ і S – довжина основи і площа прямокутника, який відповідає моді варіаційного ряду,							
15									
16		n_M – частота моди (найбільша частота в варіаційному ряду).							
17									
18		Тепер висота будь-якого прямокутника гістограми буде обчислюватись за формулою:					$h_i = \mu \frac{n_i}{\delta_i}$		
19									

Результати розрахунків занесемо в стовпець «h» таблиці.

	Q	R	S	T	U	V
20						
21		i	інтервал	n_i	δ_i	h_i
22		1	1,50	13	1,6	22,61
23			3,10			
24		2	3,10	9	0,8	31,30
25			3,90			
26		3	3,90	23	0,8	80,00
27			4,70			
28		4	4,70	23	0,8	80,00
29			5,50			
30		5	5,50	18	0,8	62,61
31			6,30			
32		6	6,30	16	0,8	55,65
33			7,10			
34		7	7,10	9	0,8	31,30
35			7,90			
36		8	7,90	8	2,4	9,28
37			10,30			

	Q	R	S	T	U	V
20						
21		i	інтервал	n_i	δ_i	h_i
22		1	1,5	13	=S23-S22	=SS\$13*(\$T22/\$U22)
23			3,1			
24		2	3,1	9	=S25-S24	=SS\$13*(\$T24/\$U24)
25			3,9			
26		3	3,9	23	=S27-S26	=SS\$13*(\$T26/\$U26)
27			4,7			
28		4	4,7	23	=S29-S28	=SS\$13*(\$T28/\$U28)
29			5,5			
30		5	5,5	18	=S31-S30	=SS\$13*(\$T30/\$U30)
31			6,3			
32		6	6,3	16	=S33-S32	=SS\$13*(\$T32/\$U32)
33			7,1			
34		7	7,1	9	=S35-S34	=SS\$13*(\$T34/\$U34)
35			7,9			
36		8	7,9	8	=S37-S36	=SS\$13*(\$T36/\$U36)
37			10,3			

д) обчислити вибірове середнє, дисперсію; виправлену дисперсію і середнє квадратичне відхилення;

Сума усіх елементів вибірки:

$$S_x = \sum_{i=1}^n x_i = 639,70 \quad \text{де } n = 119 - \text{обсяг вибірки}$$

Вибіркове середнє:

$$a = \bar{x} = \frac{S_x}{n} = 5,38$$

Вибіркова дисперсія:

$$D_x = \frac{S_{xx}}{n} - \bar{x}^2 = 3,45 \quad \text{де } S_{xx} = \sum_{i=1}^n x_i^2 = 3848,95$$

Виправлена дисперсія:

$$S^2 = \frac{n}{n-1} D_x = 3,48$$

Середнє квадратичне відхилення:

$$\sigma = S = \sqrt{S^2} = 1,86$$

Z	AA	AB	AC	AD	AE	AF
4						
5	д) обчислити вибірове середнє, дисперсію; виправлену дисперсію і середнє квадратичне відхилення;					
6						
7						
8						
9	Вибіркове середнє:			$\bar{x} = \frac{S_x}{n} \quad S_x = \sum_{i=1}^n x_i$		
10						
11	n	119				
12	Sx	639,70				
13	\bar{x}	5,38				
14			$D_x = \frac{S_{xx}}{n} - \bar{x}^2 \quad S_{xx} = \sum_{i=1}^n x_i^2$			
15	Вибіркова дисперсія:					
16						
17	Sxx	3848,95				
18	Dx	3,45				
19			$S^2 = \frac{n}{n-1} D_x$			
20	Виправлена дисперсія:					
21						
22	S^2	3,48				
23					$\sigma = S = \sqrt{S^2}$	
24	Середнє квадратичне відхилення:					
25						
26	σ	1,86				

	Z	AA	AB	AC	AD	AE	AF
4							
5			д) обчислити вибіркове середнє, дисперсію; виправлену дисперсію і середнє квадратичне відхилення;				
6							
7							
8							
9			Вибіркове середнє:	$\bar{x} = \frac{S_x}{n} \quad S_x = \sum_{i=1}^n x_i$			
10							
11		n	=СЧЁТ(B5:F30)				
12		Sx	=СУММ(B5:F30)				
13		x	=AB12/AB11				
14				$D_x = \frac{S_{xx}}{n} - \bar{x}^2 \quad S_{xx} = \sum_{i=1}^n x_i^2$			
15			Вибіркова дисперсія:				
16							
17		Sxx	=СУММКВ(B5:F30)				
18		Dx	=AB17/AB11-(AB13^2)				
19				$S^2 = \frac{n}{n-1} D_x$			
20			Виправлена дисперсія:				
21							
22		S^2	=(AB11/(AB11-1))*AB18				
23				$\sigma = S = \sqrt{S^2}$			
24			Середнє квадратичне відхилення:				
25							
26		σ	=КОРЕНЬ(AB22)				

е) обчислити теоретичні частоти для нормального закону, крайні інтервали брати напівнескінченними (сума всіх частот при цьому буде дорівнювати об'єму вибірки);

Заповнимо стовпець «х», в який заносять аргументи інтегралу ймовірностей:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{t^2}{2}\right) dt,$$

на «і»-ому інтервалі $x = \frac{r-a}{s}$, де «r» – число з стовпця «інтервал».

По стовпцю «х» таблиці 5.2 можна обчислити значення інтеграла ймовірностей, використовуючи Додаток А (табл. А.2). Треба мати на увазі, що, обчислюючи ймовірності $P = \Phi(x'') - \Phi(x')$, ми віднімаємо близькі між собою числа, а це, як відомо, приводить до значного зросту відносної похибки. Тому значення функції $\Phi(x)$ треба отримувати з максимальною точністю. Аргумент «х» в Додатку А заданий з двома знаками після коми, отже, якщо в стовпці «х» таблиці 5.2 залишати три знаки після коми, то за допомогою лінійної інтерполяції можна знайти потрібні значення $\Phi(x)$ достатньо точно.

Приклад лінійної інтерполяції. Нехай треба знайти $\Phi(2,086)$. У Додатку А (табл. А.2) знаходимо $\Phi(2,08) = 0,4812$, $\Phi(2,10) = 0,4821$. Можна обчислити різницю $0,4821 - 0,4812 = 0,0009$. Отже, на одну соту частину аргументу припадає 0,0009, а на одну тисячну 0,00009. Помножимо на однозначне число 6 ($6 \cdot 0,00009 =$

0,00054) і округлимо до чотирьох знаків після коми – 0,0005. Отже $\Phi(2,086) = 0,4812 + 0,0005 = 0,4817$.

Заповнимо стовпець $\Phi(x)$ таблиці. Тепер можна обчислити теоретичні ймовірності попадання в інтервали:

$$P_i = \Phi(x_i'') - \Phi(x_i')$$

(від нижнього числа клітини стовпця $\Phi(x)$ віднімаємо верхнє число тієї ж клітини). Нарешті, останній стовпець таблиці – теоретичні частоти:

$$n'_i = P_i \cdot n \quad \text{де } n - \text{об'єм вибірки.}$$

Зауважимо, що, на відміну від фактичних, теоретичні частоти не округляються до цілого значення.

Z	AA	AB	AC	AD	AE	AF
28	е) обчислити теоретичні частоти для нормального закону, крайні інтервали брати напівнескінченними (сума всіх частот при цьому буде дорівнювати об'єму вибірки);					
29						
30						
31						
32						
33	Виходячи з вигляду гістограми (близькі до центру рисунка прямокутники високі, а до периферії знижуються), висунемо гіпотезу про нормальний закон розподілу генеральної сукупності. В якості параметрів нормального закону приймемо:					
34						
35						
36						
37						
38						
39	$\alpha = \bar{x}$	5,38		$\sigma = S$	1,86	
40						
41						
42	i	інтервал	x	Φ	p_i	n'_i
43	1	1,5	-2,086	-0,4817	0,0918	10,9242
44		3,1	-1,226	-0,3899		
45	2	3,1	-1,226	-0,3899	0,1033	12,2927
46		3,9	-0,796	-0,2866		
47	3	3,9	-0,796	-0,2866	0,1238	14,7322
48		4,7	-0,366	-0,1628		
49	4	4,7	-0,366	-0,1628	0,1887	22,4553
50		5,5	0,065	0,0259		
51	5	5,5	0,065	0,0259	0,1638	19,4922
52		6,3	0,495	0,1897		
53	6	6,3	0,495	0,1897	0,1328	15,8032
54		7,1	0,925	0,3225		
55	7	7,1	0,925	0,3225	0,0898	10,6862
56		7,9	1,355	0,4123		
57	8	7,9	1,355	0,4123	0,0836	9,9484
58		10,3	2,645	0,4959		

	Z	AA	AB	AC	AD	AE	AF
27							
28							
29							
30							
31							
32							
33							
34							
35							
36							
37							
38							
39							
40							
41							
42							
43							
44							
45							
46							
47							
48							
49							
50							
51							
52							
53							
54							
55							
56							
57							
58							

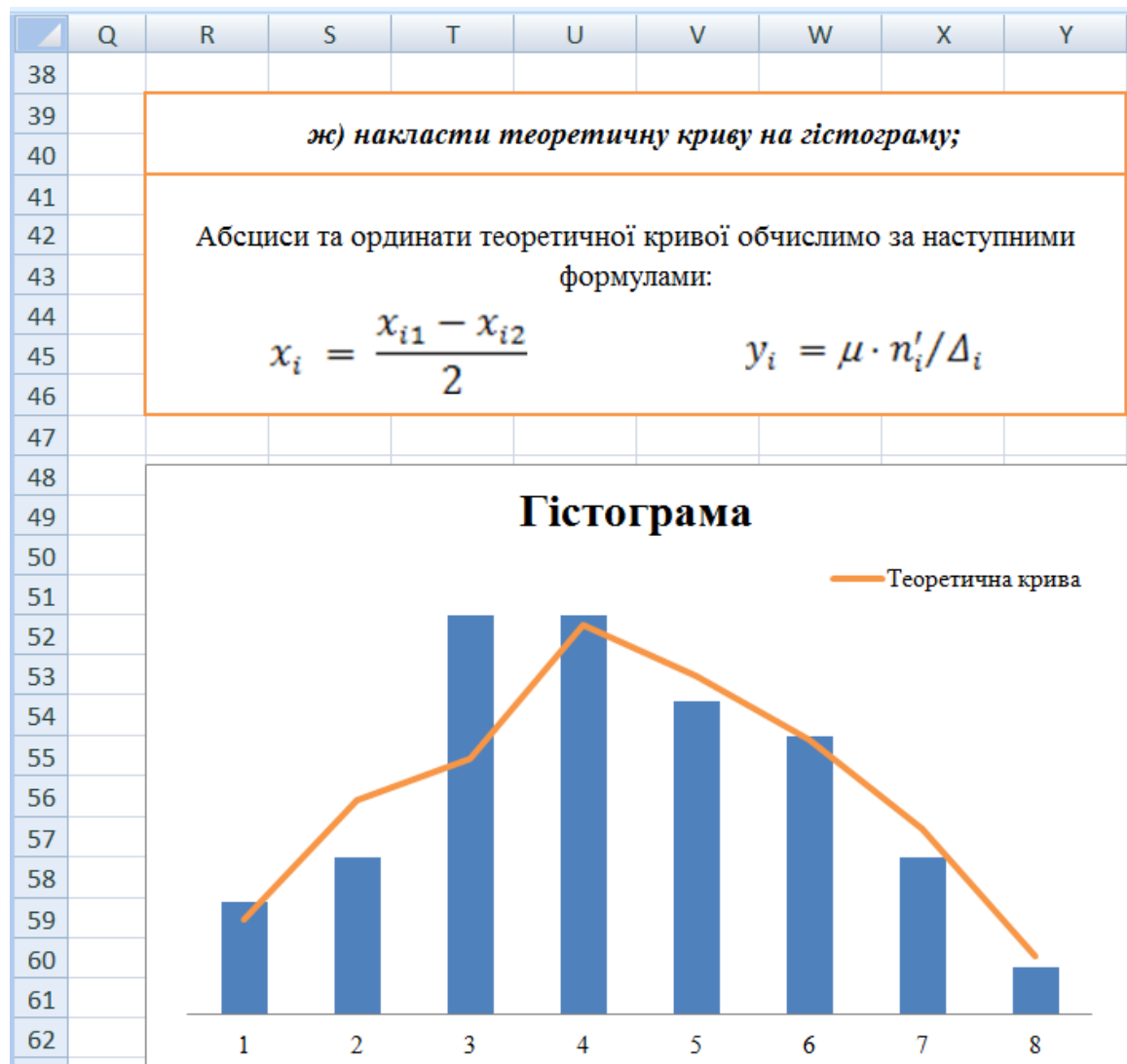
ж) накласти теоретичну криву на гістограму;

Абсциси та ординати теоретичної кривої обчислимо за наступними формулами:

$$x_i = \frac{x_{i1} - x_{i2}}{2}$$

$$y_i = \mu \cdot n'_i / \Delta_i$$

З'єднуємо точки плавною лінією (вершина кривої не зобов'язана співпадати з якою-небудь точкою), вона повинна бути симетричною відносно прямої $x = \bar{x}$.



	Q	R	S	T	U	V	W	X	Y
20									
21		i	інтервал	n_i	δ_i	h_i		x	y
22		1	1,50	13	1,6	22,61		2,3	18,9986
23			3,10						
24		2	3,10	9	0,8	31,30		3,5	42,7572
25			3,90						
26		3	3,90	23	0,8	80,00		4,3	51,2424
27			4,70						
28		4	4,70	23	0,8	80,00		5,1	78,1054
29			5,50						
30		5	5,50	18	0,8	62,61		5,9	67,799
31			6,30						
32		6	6,30	16	0,8	55,65		6,7	54,9677
33			7,10						
34		7	7,10	9	0,8	31,30		7,5	37,1694
35			7,90						
36		8	7,90	8	2,4	9,28		9,1	11,5344
37			10,30						

	Q	R	S	T	U	V	W	X	Y
20									
21		i	інтервал	n_i	δ_i	h_i		x	y
22			1,5			=\$\$13*(T22/U22)		=CP3HAЧA(S22:S23)	=\$\$13*(AF43/U22)
23		1	3,1	13	=S23-S22				
24			3,1			=\$\$13*(T24/U24)		=CP3HAЧA(S24:S25)	=\$\$13*(AF45/U24)
25		2	3,9	9	=S25-S24				
26			3,9			=\$\$13*(T26/U26)		=CP3HAЧA(S26:S27)	=\$\$13*(AF47/U26)
27		3	4,7	23	=S27-S26				
28			4,7			=\$\$13*(T28/U28)		=CP3HAЧA(S28:S29)	=\$\$13*(AF49/U28)
29		4	5,5	23	=S29-S28				
30			5,5			=\$\$13*(T30/U30)		=CP3HAЧA(S30:S31)	=\$\$13*(AF51/U30)
31		5	6,3	18	=S31-S30				
32			6,3			=\$\$13*(T32/U32)		=CP3HAЧA(S32:S33)	=\$\$13*(AF53/U32)
33		6	7,1	16	=S33-S32				
34			7,1			=\$\$13*(T34/U34)		=CP3HAЧA(S34:S35)	=\$\$13*(AF55/U34)
35		7	7,9	9	=S35-S34				
36			7,9			=\$\$13*(T36/U36)		=CP3HAЧA(S36:S37)	=\$\$13*(AF57/U36)
37		8	10,3	8	=S37-S36				

з) обчислити суму Пірсона;

$$\chi^2_{\text{спост}} = \sum_i \chi_i^2 = \sum_i \frac{(n_i - n'_i)^2}{n'_i} = 6,69$$

	AG	AH	AI	AJ	AK
4					
5		з) обчислити суму Пірсона;			
6		$\chi^2_{\text{спост}} = \sum_i \chi_i^2 = \sum_i \frac{(n_i - n'_i)^2}{n'_i}$			
7					
8					
9					
10		i	n_i	n'_i	χ_i^2
11		1	13	10,9242	0,39444
12					
13		2	9	12,2927	0,88198
14					
15		3	23	14,7322	4,63994
16					
17		4	23	22,4553	0,01321
18					
19		5	18	19,4922	0,11423
20					
21		6	16	15,8032	0,00245
22					
23		7	9	10,6862	0,26607
24					
25		8	8	9,9484	0,3816
26					
27		$\chi^2_{\text{спост}} = \sum_i \chi_i^2$			
28					

	AG	AH	AI	AJ	AK
4					
5		з) обчислити суму Пірсона;			
6		$\chi^2_{\text{спост}} = \sum_i \chi_i^2 = \sum_i \frac{(n_i - n'_i)^2}{n'_i}$			
7					
8					
9					
10		i	n_i	n'_i	χ_i^2
11					
12		1	13	=AF43	=((AI11-AJ11)^2)/AJ11
13					
14		2	9	=AF45	=((AI13-AJ13)^2)/AJ13
15					
16		3	23	=AF47	=((AI15-AJ15)^2)/AJ15
17					
18		4	23	=AF49	=((AI17-AJ17)^2)/AJ17
19					
20		5	18	=AF51	=((AI19-AJ19)^2)/AJ19
21					
22		6	16	=AF53	=((AI21-AJ21)^2)/AJ21
23					
24		7	9	=AF55	=((AI23-AJ23)^2)/AJ23
25					
26		8	8	=AF57	=((AI25-AJ25)^2)/AJ25
27		$\chi^2_{\text{спост}} = \sum_i \chi_i^2$			=СУММ(AK11:AK26)
28					

и) визначити кількість ступенів вільності, з рівнем значущості α перевірити гіпотезу про нормальний розподіл;

Для обчислення кількості ступенів вільності треба від числа інтервалів відняти кількість зв'язків, які накладає підсумовування. У випадку нормального закону зв'язків 3, це підсумовування при обчисленні \bar{x} , D, χ^2 .

Отже, ступенів вільності $8 - 3 = 5$.

Нехай заданий рівень значущості $\alpha = 0,05$. Використовуючи табл. А.5, знайдемо критичне значення розподілу χ^2 :

$$\chi^2_{\text{крит}} = 11,1$$

отже, $\chi^2_{\text{спост}} < \chi^2_{\text{крит}}$, тому немає підстав відкинути гіпотезу про нормальний розподіл.

к) визначити ймовірність попадання в інтервал $[a; b]$ двома способами – за відотною частотою і за теоретичною функцією розподілу;

Визначимо ймовірність P_{ab} попадання у інтервал $[3,5; 4,8]$ двома способами – за відотною частотою і за теоретичною функцією розподілу.

За відотною частотою:

$$P = \frac{0,5 \cdot 9 + 23}{119} = 0,23$$

(в чисельнику враховуємо частоти тих варіант варіаційного ряду, які належать інтервалу $[a; b]$, для точок, які знаходяться на границі частота зменшується вдвічі).

За теоретичним розподілом:

$$P_{ab} = P(a \leq X \leq b) = \Phi(x'') - \Phi(x')$$

$$x'' = \frac{4,8 - 5,38}{1,86} = -0,31$$

$$x' = \frac{3,5 - 5,38}{1,86} = -1,23$$

$$P_{ab} = \Phi(-0,31) - \Phi(-1,23) = -0,1217 - (-0,3907) = 0,269$$

л) знайти інтервали довіри для числових параметрів «а» і «σ» нормально розподіленої генеральної сукупності.

Знайдемо інтервали довіри для параметрів нормального розподілу; нехай їх треба обчислити з надійністю $\gamma = 0,999$. В табл. А.3 для $n = 119$ і $\gamma = 0,999$ знаходимо $t_\gamma = 3,374$ і обчислюємо відхилення:

$$\delta = \frac{t_\gamma \cdot S}{\sqrt{n}} = \frac{3,374 \cdot 1,86}{\sqrt{119}} = \frac{6,28}{10,9} = 0,5$$

$$a - \delta = 5,38 - 0,58 = 4,8$$

$$a + \delta = 5,38 + 0,58 = 5,96$$

З надійністю $\gamma = 0,999$ отримаємо $4,8 < a < 5,96$.

Додаток А (табл. А.4) дозволяє знайти $q_\gamma = 0,27$. Тому з надійністю $\gamma = 0,999$ отримаємо інтервал довіри:

$$\sigma(1 - q) < \sigma < \sigma(1 + q),$$

$$\sigma(1 - q) = 1,86(1 - 0,27) = 1,36$$

$$\sigma(1 + q) = 1,86(1 + 0,27) = 2,36$$

З надійністю $\gamma = 0,999$ отримаємо $1,36 < \sigma < 2,36$.

Додаток А

Таблиця А.2. – Таблиця значень функції $\Phi(x)$

X	$\Phi(X)$	X	$\Phi(X)$	X	$\Phi(X)$	X	$\Phi(X)$
0,00	0,0000	0,41	0,1591	0,82	0,2939	1,23	0,3907
0,01	0,0040	0,42	0,1627	0,83	0,2967	1,24	0,3925
0,02	0,0080	0,430,	0,16664	0,84	0,2995	1,25	0,3944
0,03	0,0120	0,44	0,1700	0,85	0,3023	1,26	0,3962
0,04	0,0160	0,45	0,1736	0,86	0,3051	1,27	0,3980
0,05	0,0199	0,46	0,1772	0,87	0,3079	1,28	0,3997
0,06	0,0239	0,47	0,1808	0,88	0,3106	1,29	0,4015
0,07	0,0279	0,48	0,1844	0,89	0,3133	1,30	0,4032
0,08	0,0319	0,49	0,1879	0,90	0,3159	1,31	0,4049
0,09	0,0359	0,50	0,1915	0,91	0,3186	1,35	0,4115
0,10	0,0398	0,51	0,1950	0,92	0,3212	1,36	0,4131
0,11	0,0438	0,52	0,1985	0,93	0,3238	1,37	0,4147
0,12	0,0478	0,53	0,2019	0,94	0,3264	1,38	0,4162
0,13	0,0517	0,54	0,2054	0,95	0,3289	1,39	0,4177
0,14	0,0557	0,55	0,2088	0,96	0,3315	1,40	0,4193
0,15	0,0596	0,56	0,2122	0,97	0,3340	1,41	0,4207
0,16	0,0636	0,57	0,2157	0,98	0,3365	1,42	0,4222
0,17	0,0675	0,58	0,2190	0,99	0,3389	1,43	0,4236
0,18	0,0714	0,59	0,2224	1,00	0,3414	1,44	0,4251
0,19	0,0754	0,60	0,2257	1,01	0,3438	1,45	0,4265
0,20	0,0793	0,61	0,2291	1,02	0,3461	1,46	0,4279
0,21	0,0832	0,62	0,2324	1,03	0,3485	1,47	0,4292
0,22	0,0871	0,63	0,2356	1,04	0,3508	1,48	0,4306
0,23	0,0910	0,64	0,2389	1,05	0,3531	1,49	0,4319
0,24	0,0948	0,65	0,2421	1,06	0,3554	1,50	0,4332
0,25	0,0987	0,66	0,2454	1,07	0,357	1,51	0,4345
0,26	0,1026	0,67	0,2486	1,08	0,3599	1,52	0,4347
0,27	0,1064	0,68	0,2517	1,09	0,3622	1,53	0,4370
0,28	0,1103	0,69	0,2549	1,10	0,3634	1,54	0,4382
0,29	0,1141	0,70	0,2580	1,11	0,3665	1,55	0,4394
0,30	0,1179	0,71	0,2611	1,12	0,3687	1,56	0,4406
0,31	0,1217	0,72	0,2642	1,13	0,3708	1,57	0,4418
0,32	0,1255	0,73	0,2673	1,14	0,3729	1,58	0,4429
0,33	0,1293	0,74	0,2703	1,15	0,3749	1,59	0,4441
0,34	0,1331	0,75	0,2734	1,16	0,3770	1,60	0,4452
0,35	0,1368	0,76	0,2764	1,17	0,3790	1,61	0,4463
0,36	0,1406	0,77	0,2793	1,18	0,3810	1,62	0,4474
0,37	0,1443	0,78	0,2823	1,19	0,3830	1,63	0,4485
0,38	0,1480	0,79	0,2852	1,20	0,3849	1,64	0,4495
0,39	0,1517	0,80	0,2881	1,21	0,3869	1,65	0,4505
0,40	0,1554	0,81	0,2910	1,22	0,3888	1,66	0,4515
						1,67	0,4525

Продовження таблиці А.2

X	Φ(X)	X	Φ(X)	X	Φ(X)	X	Φ(X)
1,68	0,4535	1,92	0,4726	2,32	0,4898	2,78	0,4973
1,69	0,4545	1,93	0,4732	2,34	0,4903	2,80	0,4974
1,70	0,4554	1,94	0,4738	2,36	0,4909	2,82	0,4976
1,71	0,4564	1,95	0,4744	2,38	0,4913	2,84	0,4977
1,72	0,4573	1,96	0,4750	2,40	0,4918	2,86	0,4979
1,73	0,4582	1,97	0,4756	2,42	0,4922	2,88	0,4980
1,74	0,4591	1,98	0,4761	2,44	0,4926	2,90	0,4981
1,75	0,4599	1,99	0,4767	2,46	0,4930	2,92	0,4982
1,76	0,4608	2,00	0,4772	2,48	0,4934	2,94	0,4984
1,77	0,4616	2,02	0,4783	2,50	0,4938	2,96	0,4985
1,78	0,4625	2,04	0,4793	2,52	0,4941	2,98	0,4986
1,79	0,4633	2,06	0,4803	2,54	0,4944	3,00	0,4986
1,80	0,4641	2,08	0,4812	2,56	0,4948	3,20	0,4993
1,81	0,4648	2,10	0,4821	2,58	0,4951	3,40	0,4996
1,82	0,4656	2,12	0,4830	2,60	0,4953	3,60	0,4998
1,83	0,4664	2,14	0,4838	2,62	0,4956	3,80	0,4999
1,84	0,4671	2,16	0,4846	2,64	0,4958	4,00	0,4999
1,85	0,4678	2,18	0,4854	2,66	0,4961	4,25	0,4999
1,86	0,4686	2,20	0,4861	2,68	0,4963	4,50	0,4999
1,87	0,4693	2,22	0,4868	2,70	0,4965	5,00	0,4999
1,88	0,4699	2,24	0,4874	2,72	0,4967	∞	0,5
1,89	0,4706	2,26	0,4881	2,74	0,4969		
1,90	0,4713	2,28	0,4887	2,76	0,4971		
1,91	0,4719	2,30	0,4893				

Таблиця А.3. – Таблиця значень функції $t_\gamma = t(\gamma, n)$,

$\gamma \backslash n$	0,95	0,99	0,999	$\gamma \backslash n$	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	2,633	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				

Таблиця А.4 – Таблиця значень функції $q_\gamma = q(\gamma, n)$

$\gamma \backslash n$	0,95	0,99	0,999	$\gamma \backslash n$	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

Таблиця А.5 – Критичні точки розподілу χ^2

Кількість ступенів вільності	Рівень значущості α			Кількість ступенів вільності	Рівень значущості α		
	0,01	0,025	0,05		0,01	0,025	0,05
1	6,6	5,0	3,8	16	32,0	28,8	26,3
2	9,2	7,4	6,0	17	33,4	30,2	27,6
3	11,3	9,3	7,8	18	34,8	31,5	28,9
4	13,3	11,1	9,5	19	36,2	32,9	30,1
5	15,1	12,8	11,1	20	37,6	34,2	31,4
6	16,8	14,4	12,6	21	38,9	35,5	32,7
7	18,5	16,0	14,1	22	40,3	36,8	33,9
8	20,1	17,5	15,5	23	41,6	38,1	35,2
9	21,7	19,0	16,9	24	43,0	39,4	36,4
10	23,2	20,5	18,3	25	44,3	40,6	37,7
11	24,7	21,9	19,7	26	45,6	41,9	38,9
12	26,2	23,3	21,0	27	47,0	43,2	40,1
13	27,7	24,7	22,4	28	48,3	44,5	41,3
14	29,1	26,1	23,7	29	49,6	45,7	42,6
15	30,6	27,5	25,0	30	50,9	47,0	43,8

Сумський державний університет

Кафедра

Прикладної математики та моделювання складних систем

ЗВІТ

Завдання 20

Дисципліна

Теорія ймовірностей та математична статистика

Варіант 8

Виконавець:

студентка групи ПМ-81

Пороскун Олена Олегівна

Викладач:

Гончаров Олександр Андрійович

Суми, Сумська область

2020

Завдання 20 (таблиця 6.8). Для заданої двовимірної вибірки виконати такі завдання:

а) обчислити числові характеристики

$$S_x, S_y, S_{xx}, S_{yy}, S_{xy}, \bar{x}, \bar{y}, D_x, D_y;$$

б) обчислити коефіцієнти рівняння лінійної регресії;

в) обчислити вибірковий коефіцієнт кореляції, оцінити тісноту лінійного зв'язку між факторами

г) для X_{\min} , X , X_{\max} знайти передбачення «у»;

д) для X_{\min} , X , X_{\max} знайти інтервали довіри лінії регресії з надійністю γ ;

е) побудувати на одному рисунку кореляційне поле, лінію регресії і область довіри.

В таблиці 6.8 число «k» дорівнює остачі від ділення номера варіанта на 5, а число «n» приймає значення з набору для цього варіанта (табл. 6.4). Отже, в початковий числовий масив треба включити пари чисел «х» і «у», які записані в стовпці «k» в тих клітинах (відокремлених товстими горизонтальними лініями), номер «n» яких належить набору даного варіанта.

Таблиця 6.8

k →	0		1		2		3		4	
n ↓	x	y	x	y	x	y	x	y	x	y
1	2	3	4	5	6	7	8	9	10	11
1	35	30	58	13	83	28	65	15	75	28
	31	38	29	32	86	29	86	39	77	34
	1	73	2	54	80	25	76	28	99	63
	36	28	29	34	90	34	86	37	96	58
2	6	63	31	31	61	14	86	37	65	18
	43	25	37	25	75	24	81	32	95	58
	27	42	21	39	76	23	63	15	65	21
	3	69	55	11	70	19	96	44	62	16
3	1	70	0	61	86	33	99	50	92	53
	6	63	49	23	92	38	84	37	77	33
	33	34	56	20	98	41	77	27	78	35
	46	47	26	34	79	23	62	19	96	55

Продовження таблиці 6.8

1	2	3	4	5	6	7	8	9	10	11
4	10	58	38	26	78	25	76	27	82	40
	49	16	14	40	65	18	78	29	71	23
	45	24	33	31	71	17	87	40	82	38
	46	21	21	36	64	16	72	24	75	28
5	17	52	40	27	63	16	92	46	85	40
	4	71	7	52	96	45	80	31	64	17
	17	50	4	53	66	17	85	36	92	51
	34	32	24	38	91	36	91	48	60	16
6	2	76	57	17	74	22	61	19	66	24
	4	70	43	24	95	43	60	18	96	61
	39	28	51	17	85	29	92	42	91	54
	4	70	7	49	85	31	89	43	68	24
7	9	61	55	22	70	20	66	23	88	49
	16	54	45	25	61	12	71	26	64	21
	35	28	46	21	82	26	64	18	88	45
	9	65	35	26	83	26	74	24	87	45
8	3	73	34	30	93	39	68	20	96	59
	18	52	28	31	86	29	87	42	81	33
	45	15	17	40	86	29	75	27	65	20
	3	74	16	38	88	38	73	25	69	25

Таблиця 6.4

Варіант	1	2	3	4	5	6
набір	3 4 5 6 8	1 2 3 6 8	2 3 5 7 8	2 3 4 6 8	1 2 5 6 8	1 4 5 6 7

Продовження таблиці 6.4

Варіант	7	8	9	10	11	12
набір	1 3 5 6 8	3 5 6 7 8	1 2 3 4 7	1 3 4 7 8	1 3 4 6 7	2 3 6 7 8

Продовження таблиці 6.4

Варіант	13	14	15	16	17	18
набір	2 3 4 5 6	1 4 5 7 8	1 2 3 4 5	2 5 6 7 8	1 2 4 5 8	1 2 4 6 7

Продовження таблиці 6.4

Варіант	19	20	21	22	23	24
набір	1 4 5 6 8	1 3 5 7 8	1 2 4 5 7	2 3 4 5 7	1 2 5 6 7	3 4 5 7 8

Продовження таблиці 6.4

Варіант	25	26	27	28	29	30
набір	2 3 4 5 8	1 2 3 4 6	1 2 4 6 8	1 3 4 5 7	1 4 6 7 8	1 3 4 6 8

а) Формули для знаходження числових характеристик:

$$S_x = \sum_{i=1}^n x_i \quad S_y = \sum_{i=1}^n y_i \quad S_{xx} = \sum_{i=1}^n x_i^2 \quad S_{yy} = \sum_{i=1}^n y_i^2 \quad S_{xy} = \sum_{i=1}^n x_i \cdot y_i$$

n = обсяг величин X (та Y)

$$\bar{x} = \frac{S_x}{n} \quad \bar{y} = \frac{S_y}{n}$$

$$D_x = \frac{S_{xx}}{n} - \bar{x}^2 \quad D_y = \frac{S_{yy}}{n} - \bar{y}^2$$

б) Коефіцієнти рівняння лінійної регресії:

$$k = \frac{\frac{S_{xy}}{n} - \bar{x} \cdot \bar{y}}{D_x} = 0,89 \quad b = \bar{y} - k \cdot \bar{x} = -38,3$$

Рівняння лінійної регресії: $y(x) = kx + b$

$$y = 0,89 x - 38,3$$

в) Вибірковий коефіцієнт кореляції обчислюється за формулою:

$$r = k \cdot \sqrt{\frac{D_x}{D_y}} = 0,97$$

Можна оцінити тісноту лінійного зв'язку між факторами за шкалою:

$|r| < 0,6$ – слабка

$0,6 \leq |r| \leq 0,9$ – середня

$|r| > 0,9$ – сильна.

Висновок: між факторами X і Y спостерігається лінійний зв'язок з сильним рівнем тісноти, так як $|r| = 0,97 > 0,9$.

г) для X_{min} , X , X_{max} знайти передбачення «у»

$$y(x) = kx + b$$

$$y(x_{min}) = y(60) = 15,4$$

$$y(\bar{x}) = y(77,5) = 31$$

$$y(x_{max}) = y(99) = 50,3$$

д) для X_{min} , X , X_{max} знайти інтервали довіри лінії регресії з надійністю $\gamma = 0.999$;

$$S_y = \sqrt{\frac{n}{n-1} D_y} = 10,9$$

Використовуючи таблицю А.3, знаходимо $t_\gamma = 3,883$. Тепер можна записати формулу для обчислення ширини інтервалу довіри

$$\delta(x) = \frac{t_y \cdot S_y}{\sqrt{n}} \sqrt{1 + \frac{(x - \bar{x})^2}{D}}$$
$$\frac{t_y \cdot S_y}{\sqrt{n}} = 9,5$$

$$\delta(x_{min}) = \delta(60) = 17,18$$

$$\delta(\bar{x}) = \delta(77,5) = 9,5$$

$$\delta(x_{max}) = \delta(99) = 19,9$$

Отже, отримаємо інтервали довіри для передбачень «у» в заданих точках

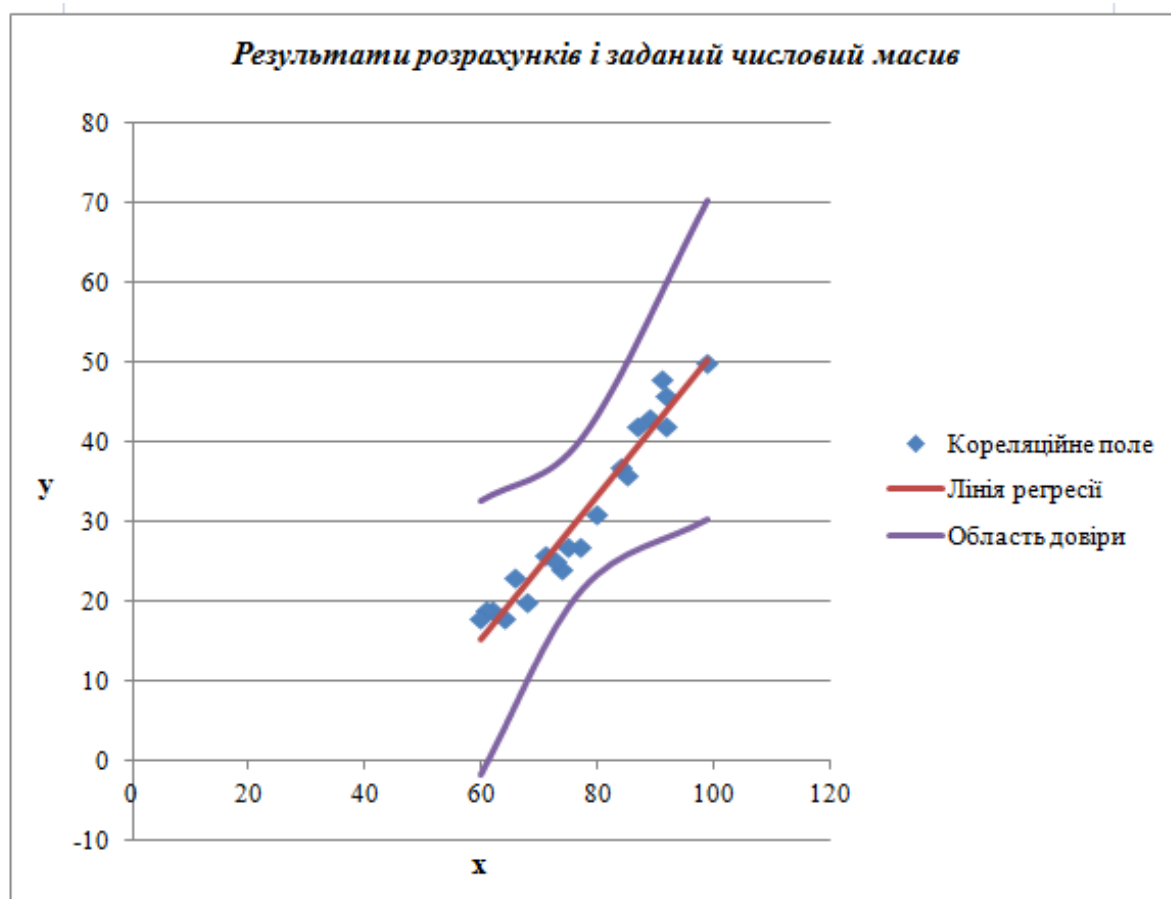
$$y_x - \delta_x \leq y \leq y_x + \delta_x;$$

$$x = 60: \quad -1,79 \leq y \leq 32,57,$$

$$x = 77,5: \quad 21,55 \leq y \leq 40,55,$$

$$x = 99: \quad 30,30 \leq y \leq 70,28.$$

е) побудувати на одному рисунку кореляційне поле, лінію регресії і область довіри



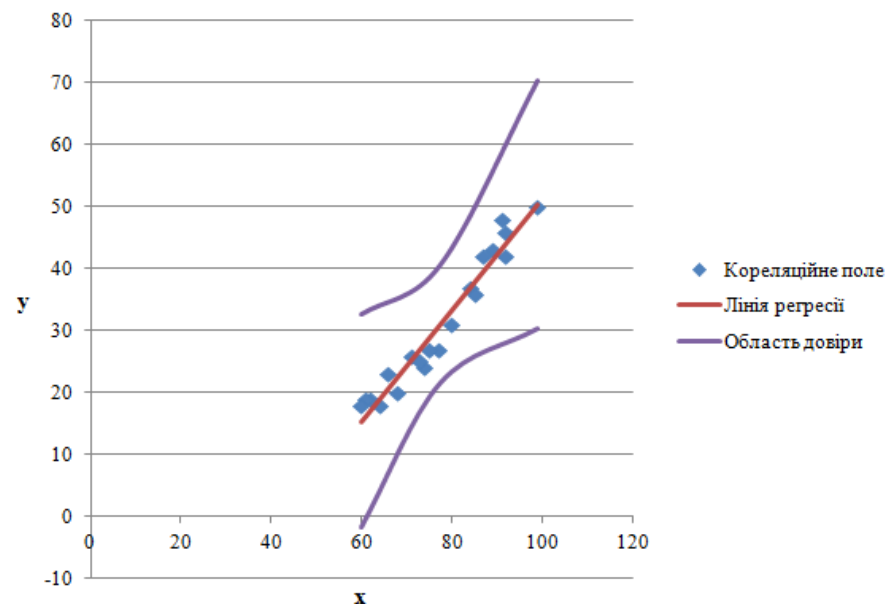
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1					Пороскун О. ПМ-81															
2																				
3																				
4		x	y		Завдання 20															
5		99	50																	
6		84	37																	
7		77	27		<p>Для заданої двовимірної вибірки виконати такі завдання:</p> <p>а) обчислити числові характеристики $S_x, S_y, S_{xx}, S_{yy}, S_{xy}, \bar{x}, \bar{y}, D_x, D_y$;</p> <p>б) обчислити коефіцієнти рівняння лінійної регресії;</p> <p>в) обчислити вибірковий коефіцієнт кореляції, оцінити тісноту лінійного зв'язку між факторами</p> <p>г) для X_{min}, X, X_{max} знайти передбачення «у»;</p> <p>д) для X_{min}, X, X_{max} знайти інтервали довіри ліній регресії з надійністю γ;</p> <p>е) побудувати на одному рисунку кореляційне поле, лінію регресії і область довіри.</p>					а) обчислити числові характеристики			б) обчислити коефіцієнти рівняння лінійної регресії							
8		62	19							S_x	1550		$y = kx + b$							
9		92	46							S_y	621		$k = 0,89$ $b = -38,30$ $y = 0,89x - 38,3$							
10		80	31							S_{xx}	122822		в) обчислити вибірковий коефіцієнт кореляції, оцінити тісноту лінійного зв'язку між факторами							
11		85	36							S_{yy}	21557									
12		91	48							S_{xy}	50541		$r = 0,97$							
13		61	19							\bar{x}	77,5									
14		60	18							\bar{y}	31,05		Висновок: між факторами X і Y спостерігається лінійний зв'язок з сильним рівнем тісноти, так як $ r = 0,97 > 0,9$.							
15		92	42							D_x	134,85									
16		89	43							D_y	113,7475									
17		66	23							n	20									
18		71	26							$S_x = \sum_{i=1}^n x_i$ $S_y = \sum_{i=1}^n y_i$ $\bar{x} = \frac{S_x}{n}$ $\bar{y} = \frac{S_y}{n}$			$D_x = \frac{S_{xx}}{n} - \bar{x}^2$			$k = \frac{\frac{S_{xy}}{n} - \bar{x} \cdot \bar{y}}{D_x}$			$r = k \cdot \sqrt{\frac{D_x}{D_y}}$	
19		64	18		$S_{xx} = \sum_{i=1}^n x_i^2$ $S_{yy} = \sum_{i=1}^n y_i^2$ $S_{xy} = \sum_{i=1}^n x_i \cdot y_i$			$D_y = \frac{S_{yy}}{n} - \bar{y}^2$			$b = \bar{y} - k \cdot \bar{x}$									
20		74	24																	
21		68	20																	
22		87	42																	
23		75	27																	
24		73	25																	
25																				
26		Лінія регресії			г) для X_{min}, X, X_{max} знайти передбачення «у»;					е) побудувати на одному рисунку кореляційне поле, лінію регресії і область довіри										
27		$y(x) = kx + b$			$y = kx + b$															
28		x	y(x)		X_{min}	60	$y(X_{min})$	15,39												
29		99	50,29		\bar{x}	77,5	$y(\bar{X})$	31,05												
30		84	36,87		X_{max}	99	$y(X_{max})$	50,28999												
31		77	30,60																	

Результати розрахунків і заданий числовий масив

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
24		73	25		$t=1$	$t=1$	$t=1$													
25																				
26		Лінія регресії			г) для X_{min} , X , X_{max} знайти передбачення «у»;						е) побудувати на одному рисунку кореляційне поле, лінію регресії і область довіри									
27		$y(x) = kx + b$			$y = kx + b$															
28		x	y(x)			X_{min}	60	$y(X_{min})$	15,39											
29		99	50,29			\bar{x}	77,5	$y(X)$	31,05											
30		84	36,87			X_{max}	99	$y(X_{max})$	50,28999											
31		77	30,60																	
32		62	17,18																	
33		92	44,03		д) для X_{min} , X , X_{max} знайти інтервали довіри лінії регресії з надійністю γ ;															
34		80	33,29																	
35		85	37,76																	
36		91	43,13		$S_y = 10,94231$			$S_y = \sqrt{\frac{n}{n-1}} D_y$												
37		61	16,28																	
38		60	15,39		Використовуючи таблицю А.3, знаходимо $t_\gamma = 3,883$. Тепер можна записати формулу для обчислення ширини інтервалу довіри:															
39		92	44,03																	
40		89	41,34																	
41		66	20,76																	
42		71	25,23																	
43		64	18,97																	
44		74	27,92																	
45		68	22,55																	
46		87	39,55																	
47		75	28,81																	
48		73	27,02																	
49																				
50																				
51					Отже, отримаємо інтервали довіри для передбачень «у» в заданих точках															
52					$y_x - \delta x \leq y \leq y_x + \delta x;$															
53																				
54																				
55		X_{min}	60					$-1,79 \leq y \leq 32,57$												
56		x	77,5					$21,55 \leq y \leq 40,55$												
57		X_{max}	99					$30,30 \leq y \leq 70,28$												
58																				

Результати розрахунків і заданий числовий масив

Результати розрахунків і заданий числовий масив



	A	B	C	D	E	F	G	H	I
1									
2					Порожун О. ПМ-81				
3									
4		x	y		Завдання 20				
5		99	50						
6		84	37						
7		77	27						
8		62	19						
9		92	46						
10		80	31						
11		85	36						
12		91	48						
13		61	19						
14		60	18						
15		92	42						
16		89	43						
17		66	23						
18		71	26						
19		64	18						
20		74	24						
21		68	20						
22		87	42						
23		75	27						
24		73	25						
25									

Для заданої двовимірної вибірки виконати такі завдання:

- а) обчислити числові характеристики S_x , S_y , S_{xx} , S_{yy} , S_{xy} , \bar{x} , \bar{y} , Dx , Dy ;
- б) обчислити коефіцієнти рівняння лінійної регресії;
- в) обчислити вибірковий коефіцієнт кореляції, оцінити тісноту лінійного зв'язку між факторами
- г) для X_{\min} , X , X_{\max} знайти передбачення «y»;
- д) для X_{\min} , X , X_{\max} знайти інтервали довіри лінії регресії з надійністю γ ;
- е) побудувати на одному рисунку кореляційне поле, лінію регресії і область довіри.

$$S_x = \sum_{i=1}^n x_i \quad S_y = \sum_{i=1}^n y_i \quad \bar{x} = \frac{S_x}{n} \quad \bar{y} = \frac{S_y}{n}$$

$$S_{xx} = \sum_{i=1}^n x_i^2 \quad S_{yy} = \sum_{i=1}^n y_i^2 \quad S_{xy} = \sum_{i=1}^n x_i \cdot y_i$$

25		
26	Лінія регресії	
27	$y(x) = kx + b$	
28	x	y(x)
29	99	=P\$11*B29+R\$11
30	84	=P\$11*B30+R\$11
31	77	=P\$11*B31+R\$11
32	62	=P\$11*B32+R\$11
33	92	=P\$11*B33+R\$11
34	80	=P\$11*B34+R\$11
35	85	=P\$11*B35+R\$11
36	91	=P\$11*B36+R\$11
37	61	=P\$11*B37+R\$11
38	60	=P\$11*B38+R\$11
39	92	=P\$11*B39+R\$11
40	89	=P\$11*B40+R\$11
41	66	=P\$11*B41+R\$11
42	71	=P\$11*B42+R\$11
43	64	=P\$11*B43+R\$11
44	74	=P\$11*B44+R\$11
45	68	=P\$11*B45+R\$11
46	87	=P\$11*B46+R\$11
47	75	=P\$11*B47+R\$11
48	73	=P\$11*B48+R\$11

г) для X_{min} , X , X_{max} знайти передбачення «у»;			
$y = kx + b$			
X_{min}	=МИН(\$B\$5:\$B\$24)	$y(X_{min})$	=P\$11*G28+R\$11
\bar{x}	=CPЗНАЧ(\$B\$5:\$B\$24)	$y(\bar{X})$	=P\$11*G29+R\$11
X_{max}	=МАКС(\$B\$5:\$B\$24)	$y(X_{max})$	=P\$11*G30+R\$11

д) для X_{min} , X , X_{max} знайти інтервали довіри лінії регресії з надійністю ;			
$S_y =$	=КОРЕНЬ((M18/(M18-1))*M17)	$S_y = \sqrt{\frac{n}{n-1} D_y}$	
Використовуючи таблицю А.3, знаходимо $t_\gamma = 3,883$. Тепер можна записати формулу для обчислення ширини інтервалу довіри:			
$\delta(x) = \frac{t_\gamma \cdot S_y}{\sqrt{n}} \sqrt{1 + \frac{(x - \bar{x})^2}{D}}$			$t_\gamma = 3,883$
$\frac{t_\gamma \cdot S_y}{\sqrt{n}} =$	=(J42*F36)/КОРЕНЬ(M18)		
X_{min}	=МИН(\$B\$5:\$B\$24)	$\delta(X_{min})$	=F\$44*КОРЕНЬ(1+(((G46-G\$47)^2))/M\$16)
\bar{x}	=CPЗНАЧ(\$B\$5:\$B\$24)	$\delta(X)$	=F\$44*КОРЕНЬ(1+(((G47-G\$47)^2))/M\$16)
X_{max}	=МАКС(\$B\$5:\$B\$24)	$\delta(X_{max})$	=F\$44*КОРЕНЬ(1+(((G48-G\$47)^2))/M\$16)

Отже, отримаємо інтервали довіри для передбачень «у» в заданих точках
 $y_x - \delta x \leq y \leq y_x + \delta x;$

X_{min}	=МИН(\$B\$5:\$B\$24)	=I28-I46	$\leq y \leq$	=I28+I46
\bar{x}	=CPЗНАЧ(\$B\$5:\$B\$24)	=I29-I47	$\leq y \leq$	=I29+I47
X_{max}	=МАКС(\$B\$5:\$B\$24)	=I30-I48	$\leq y \leq$	=I30+I48