

**МІНІСТЕРСТВО ВНУТРІШНІХ СПРАВ УКРАЇНИ**

**Харківський національний університет внутрішніх справ**

**Навчально-науковий інститут підготовки фахівців для підрозділів  
кримінальної міліції**

**Кафедра інформаційної та економічної безпеки**

**Навчально-методичні вказівки  
до самостійної роботи  
з дисципліни  
"Математична статистика"**

Галузь знань	<b>0301 - „Соціально-політичні науки ”</b>
Напрямок підготовки	<b>6.030102 - „ Психологія ”</b>
Спеціалізація	<b>кримінальна міліція у справах дітей</b>
Ступень вищої освіти	<b>бакалавр</b>
Форма навчання	<b>денна</b>

**Харків 2015**

## **Передмова**

Робоча навчальна програма з дисципліни „Математична статистика” для курсантів за напрямом підготовки 6.030103 - „Психологія ”

### **СХВАЛЕНО**

Науково-методичною радою  
Харківського національного  
університету внутрішніх справ  
\_25.08.2015 Протокол № 8  
(дата, місяць, рік)

### **ЗАТВЕРДЖЕНО**

Вченою радою навчально-наукового  
інституту кримінальної міліції  
\_21.08.2015 Протокол № 1/15-16\_  
(дата, місяць, рік)  
\_\_\_\_\_  
(підпис) Корнієнко Д.М.  
(П.І.Б.)

### **ПОГОДЖЕНО**

Секцією Науково-методичної ради  
ХНУВС з технічних дисциплін  
\_25.08.2015 Протокол № 8  
(дата, місяць, рік)  
\_\_\_\_\_  
(підпис) \_ Сезонова І.К. \_\_\_\_\_  
(П.І.Б.)

### **ЗАТВЕРДЖЕНО**

На засіданні кафедри інформаційної та  
економічної безпеки  
20.08.2015 Протокол № \_12\_  
(дата, місяць, рік)  
\_\_\_\_\_  
(підпис) Сезонова І.К.  
(П.І.Б.)

### **Рецензенти:**

Гнусов Ю.В., кандидат технічних наук, доцент кафедри захисту інформації  
ФПФПБКТЛ ХНУВС

Чуб І.А., доктор технічних наук, професор, начальник кафедри пожежної  
профілактики Національного університету цивільного захисту України

**Розробник:** Соколовська О.Г. - Харків, Харківський національний  
університет внутрішніх справ, Навчально-науковий інститут підготовки  
фахівців для підрозділів кримінальної міліції ХНУВС, 2015

## 1. Загальні методичні вказівки

Математична статистика – розділ математики, присвячений математичним методам систематизації, обробки й використанні статистичних даних для наукових і практичних висновків.

*Прийоми й способи наукового аналізу даних, що ставляться до масових явищ, з метою визначення деяких узагальнюючі ці дані характеристик і виявлення статистичних закономірностей, становлять предмет математичної статистики.*

Відповідно до програми курсу «Математична статистика» курсанти вивчають основні поняття в математичній статистиці, а саме статистичні сукупності, дискретний статистичний розподіл вибірки та її числові характеристики, інтервальний статистичний розподіл вибірки та його числові характеристики, варіаційні ряди та їх характеристики, перевірка статистичних гіпотез, основи кореляційного та регресійного аналізів. Ці розділи є теоретичною основою більшості спеціальних дисциплін і займають важливе місце у формуванні фахівця високої кваліфікації.

Навчальним планом передбачене виконання самостійних робіт. Кожна робота включає кілька завдань, номери варіантів яких визначаються відповідно до наведеної нижче таблиці.

При оформленні роботи задачі варто розташовувати в порядку зростання їхніх номерів, зберігаючи нумерацію. Перед вирішенням кожної задачі необхідно повністю переписати її умову.

## 2. Розподіл часу навчальної дисципліни за темами

Номер та назва навчальної теми	Кількість годин, відведених на вивчення навчальної дисципліни					Література	Вид контролю	
	Всього	з них:						
		лекції	Семінарські заняття	Практичні заняття	Лабораторні заняття			Самостійна робота
Семестр № 4								
Тема № 1 Основні поняття в математичній статистиці.	14	4			6	4	1,2,4,6,8	залік
Тема № 2 Варіаційні ряди та їх характеристики. Перевірка статистичних гіпотез.	24	6			8	10	1,2,5,6,8	
Тема № 3 Основи кореляційного та регресійного аналізів..	16	2			8	6	1,2,5,6,8	
Всього за семестр № 4:	54	12			22	20		

### 3. Методичні вказівки до самостійної роботи

#### Тема № 1: Основні поняття в математичній статистиці.

**Приклад. 1.** За заданим дискретним статистичним розподілом вибірки

$X = x_i$	-6	-4	-2	2	4	6
$n_i$	5	10	15	20	40	10
$W_i$	0,05	0,1	0,15	0,2	0,4	0,1

потрібно:

1. Побудувати  $F^*(x)$  і зобразити її графічно;
2. Накреслити полігони частот і відносних частот.

**Розв'язання.** Згідно з означенням та властивостями  $F^*(x)$  має такий вигляд:

$$F^*(x) = W(X < x) = \frac{n_x}{n} = \begin{cases} 0 & x \leq -6, \\ 0,05 & -6 < x \leq -4, \\ 0,15 & -4 < x \leq -2, \\ 0,3 & -2 < x \leq 2, \\ 0,5 & 2 < x \leq 4, \\ 0,9 & 4 < x \leq 6, \\ 1, & x > 6. \end{cases}$$

Графічне зображення  $F^*(x)$  подано на рис. 106.

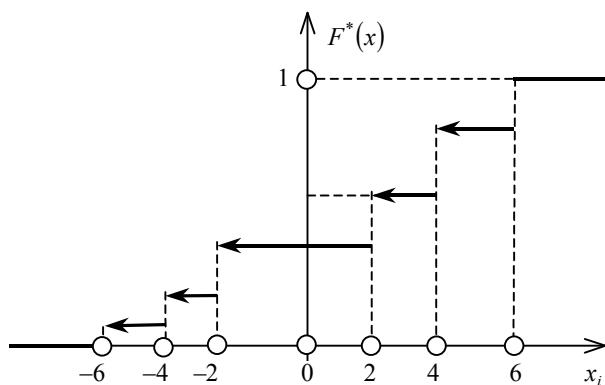


Рис. 1.1

Полігони частот та відносних частот зображено на рис.1.1, 1.3.

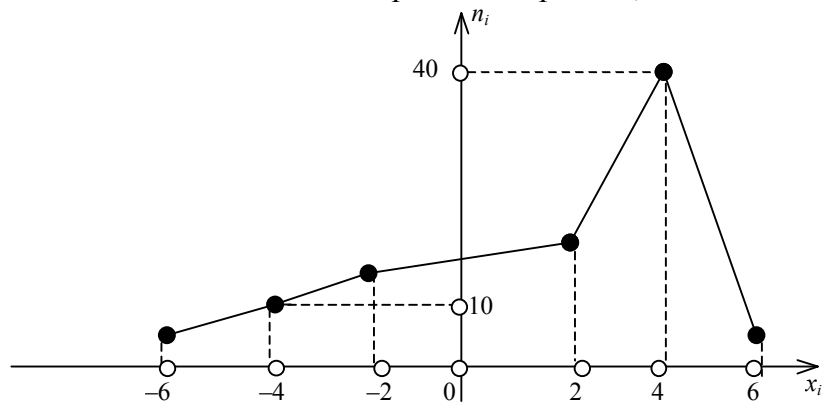


Рис. 1.2

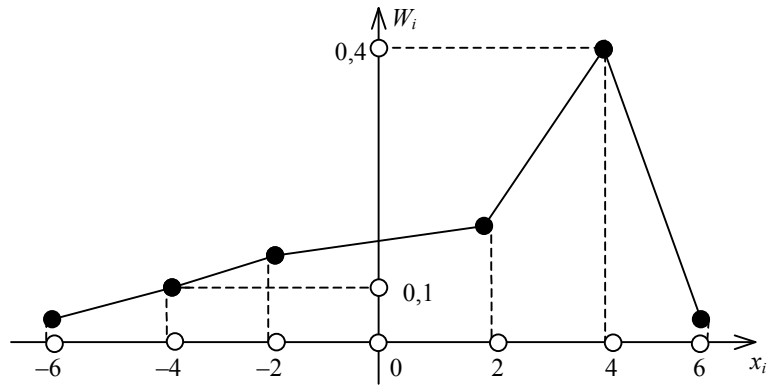


Рис. 1.3

**Приклад. 2.** За заданим статистичним розподілом вибірки

$X = x_i$	2,5	4,5	6,5	8,5	10,5
$n_i$	10	20	30	30	10

потрібно:

- 1) обчислити  $\bar{x}_B$ ,  $D_B$ ,  $\sigma_B$ ;
- 2) знайти  $Mo^*$ ,  $Me^*$ ;
- 3) обчислити  $R$ ,  $V$ .

**Розв'язання.** Оскільки  $n = \sum n_i = 100$ , то згідно з формулами (354), (357), (358) дістанемо:

$$\bar{x}_B = \frac{\sum x_i n_i}{n} = \frac{2,5 \cdot 10 + 4,5 \cdot 20 + 6,5 \cdot 30 + 8,5 \cdot 30 + 10,5 \cdot 10}{100} = 6,7;$$

$$\bar{x}_B = 6,7.$$

Для обчислення  $D_B$  визначається

$$\frac{\sum x_i^2 n_i}{n} = \frac{(2,5)^2 \cdot 10 + (4,5)^2 \cdot 20 + (6,5)^2 \cdot 30 + (8,5)^2 \cdot 30 + (10,5)^2 \cdot 10}{100} = 50,05.$$

Тоді

$$D_B = \frac{\sum x_i^2 n_i}{n} - (\bar{x}_B)^2 = 50,05 - (6,7)^2 = 50,05 - 44,89 = 5,16.$$

$$D_B = 5,16.$$

$$\sigma_B = \sqrt{D_B} = \sqrt{5,16} \approx 2,27.$$

$$\sigma_B = 2,27.$$

$$Mo^* = 6,5; 8,5.$$

Отже, наведений статистичний розподіл вибірки буде двомодальним.  $Me^* = 6,5$ , оскільки варіанта  $x = 6,5$  поділяє варіаційний ряд 2,5; 4,5; **6,5**; 8,5; 10,5 на дві частини: 2,5; 4,5 і 8,5; 10,5, які мають однакову кількість варіант.

$$R = x_{\max} - x_{\min} = 10,5 - 2,5 = 8.$$

$$V = \frac{\sigma_B}{\bar{x}_B} 100\% = \frac{2,27}{6,7} 100\% = 33,88\%.$$

**Задачі до теми**

1. При вивченні випадкової величини  $X$  у результаті 40 незалежних спостережень дістали вибірку:

10, 13, 10, 9, 9, 12, 12, 6, 7, 9, 8, 9, 11, 9, 14, 13, 9, 8, 8, 7, 10, 10, 11, 11, 11, 12, 8, 7, 9, 10, 14, 13, 8, 8, 9, 10, 11, 11, 12, 12.

Потрібно:

1. Побудувати дискретний статистичний розподіл для цієї вибірки, а також полігон частот і  $F^*(x)$ .

2. Обчислити  $\bar{x}_B$ ,  $\sigma_B$ ,  $R$ ,  $V$ .

3. Знайти  $Mo^*$ ,  $Me^*$ .

*Відповідь.*  $\bar{x}_B = 10$ ,  $\sigma_B = 2$ .

2. П'ятдесят абітурієнтів на вступних іспитах з інформатики дістали таку кількість балів:

12, 14, 19, 15, 14, 18, 13, 16, 17, 12,  
20, 17, 15, 13, 17, 16, 20, 14, 14, 13,  
17, 16, 15, 19, 16, 15, 18, 17, 15, 14,  
16, 15, 15, 18, 15, 15, 19, 14, 16, 18,  
18, 15, 15, 17, 15, 16, 16, 14, 14, 17.

Потрібно:

1. Побудувати дискретний статистичний розподіл. Полігон частот,  $F^*(x)$ .

2. Обчислити  $\bar{x}_B$ ,  $\sigma_B$ ,  $R$ ,  $V$ .

3. Знайти  $Mo^*$ ,  $Me^*$ .

*Відповідь.*  $\bar{x}_B = 15,78$ ,  $\sigma_B = 1,93$ .

3. На кожну сотню деталей, що їх виготовляє цех, у середньому припадає дві браковані. Було перевірено 10 партій по 100 деталей у кожній. Відхилення кількості виявлених бракованих деталей від середнього  $x_i$  наведено в таблиці:

Номер партії	1	2	3	4	5	6	7	8	9	10
$x_i$	-1	0	1	1	-1	1	0	-2	2	1

Потрібно:

1. Побудувати дискретний статистичний розподіл, полігон частот і  $F^*(x)$ .

2. Обчислити  $\bar{x}_B$ ,  $\sigma_B$ ,  $R$ ,  $V$ .

3. Знайти  $Mo^*$ ,  $Me^*$ .

*Відповідь.*  $\bar{x}_B = 0,2$ ,  $\sigma_B = 1,233$ .

4. Результати вимірювання максимальної місткості  $x_i$  20-ти конденсаторів у пікофарадах наведено в таблиці:

Номер конденсатора	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$x_i$ пФ	4,4	4,3 <sub>1</sub>	4,4	4,4	4,6 <sub>5</sub>	4,5 <sub>6</sub>	4,7 <sub>1</sub>	4,5 <sub>4</sub>	4,3 <sub>6</sub>	4,5 <sub>6</sub>	4,3 <sub>1</sub>	4,4 <sub>2</sub>	4,6	4,3 <sub>5</sub>	4,5	4,4	4,4 <sub>3</sub>	4,4 <sub>8</sub>	4,4 <sub>2</sub>	4,4 <sub>5</sub>

Потрібно:

1. Побудувати дискретний статистичний розподіл, полігон відносних частот і  $F^*(x)$ .

2. Обчислити  $\bar{x}_B$ ,  $\sigma_B$ ,  $R$ ,  $V$ .

3. Знайти  $Mo^*$ ,  $Me^*$ .

*Відповідь.*  $\bar{x}_B = 4,47$  пФ,  $\sigma_B = 1,108$  пФ.

5. 200 однотипних деталей  $n_i$  після шліфування були піддані контрольним вимірюванням, результати яких наведено в таблиці  $x_i$ :

$x_i$ , мм	3,7	3,8	3,9	4	4,1	4,2	4,3	4,4
$n_i$	1	22	40	79	27	26	4	1

Потрібно:

1. Побудувати полігон відносних частот, комуляту,  $F^*(x)$ .

2. Обчислити  $\bar{x}_B$ ,  $\sigma_B$ ,  $R$ ,  $V$ ;

3. Знайти  $Mo^*$ ,  $Me^*$ .

*Відповідь.*  $\bar{x}_B = 4,34$ ,  $\sigma_B = 0,039$ .

6. Із партії сталевих кілець, виготовлених заводом, була здійснена вибірка обсягом 200, і результати вимірювання товщини цих кілець  $x_i$  наведено у вигляді дискретного статистичного розподілу:

$x_i$ , мм	14,41	14,43	14,45	14,47	14,49	14,51	14,53	14,55	14,57	14,59	14,61	14,63
$n_i$	2	2	8	9	9	14	41	76	21	11	4	3

Потрібно:

1. Побудувати полігон відносних частот і комуляту.
2. Обчислити  $\bar{x}_B$ ,  $\sigma_B$ ,  $R$ ,  $V$ .
3. Знайти  $Mo^*$ ,  $Me^*$ .

Відповідь.  $\bar{x}_B = 14,34$ ;  $\sigma_B = 0,039$ .

7. Із партії однотипних сталевих болтів, виготовлених заводом, була здійснена вибірка обсягом 200 шт. і результати вимірювання їх діаметрів  $x_i$  наведено у вигляді інтервального статистичного розподілу:

$x_i$ , мм $h = 2$ мм	14,40— 14,42	14,42— 14,44	14,44— 14,46	14,46— 14,48	14,48— 14,50	14,50— 14,52	14,52— 14,54	14,54— 14,56	14,56— 14,58	14,58— 14,60	14,60— 14,62	14,62— 14,64
$n_i$	2	2	8	9	9	14	41	76	21	11	4	3

Потрібно:

1. Побудувати гістограму частот і  $F^*(x)$ .
2. Обчислити  $\bar{x}_B$ ,  $\sigma_B$ ,  $E_s^*$ ,  $A_s^*$ ,  $Mo^*$ ,  $Me^*$ .

Відповідь.  $\bar{x}_B = 14,34$  мм;  $\sigma_B = 0,039$  мм;  $A_s^* = 0,311$  мм;  $E_s^* = 1,549$  мм,  $Mo^* = 15,34$  мм,  $Me^* = 15,39$  мм.

8. У 100 осіб було виміряно зріст  $x_i$ . Результати вимірювання наведено як інтервальний статистичний розподіл:

$x_i$ , см $h = 4$ см	168— 172	172— 176	176— 180	180— 184	184— 188	192— 196	196— 200
$n_i$	10	20	30	25	10	3	2

Потрібно:

1. Побудувати гістограму частот і  $F^*(x)$ .
2. Обчислити  $\bar{x}_B$ ,  $\sigma_B$ ,  $E_s^*$ ,  $A_s^*$ ,  $Mo^*$ ,  $Me^*$ .

Відповідь.  $\bar{x}_B = 178,88$  см,  $\sigma_B = 98,87$  см,  $A_s^* = 0,0063$  см,  $E_s^* = -2,9999$  см,  $Mo^* = 178,7$  см,  $Me^* = 178,6$  см.

9. Для дослідження розподілу маси новонароджених  $x_i$  була зібрана інформація про 100 дітей. Ця інформація подана як інтервальний статистичний розподіл, що має такий вигляд у табличній формі:

$x_i$ , кг $h = 0,5$	1,0— 1,5	1,5— 2,0	2,0— 2,5	2,5— 3,0	3,0— 3,5	3,5— 4,0	4,0— 4,5	4,5— 5,0
$n_i$	2	8	10	30	40	6	3	1

Потрібно:

1. Побудувати гістограму відносних частот і  $F^*(x)$ .
2. Обчислити  $\bar{x}_B$ ,  $\sigma_B$ ,  $E_s^*$ ,  $A_s^*$ ,  $Mo^*$ ,  $Me^*$ .

Відповідь.  $\bar{x}_B = 2,915$  кг,  $\sigma_B = 0,625$  кг,  $A_s^* = -0,26$ ,  $E_s^* = 0,73$ ,  $Mo^* = 3,11$  кг,  $Me^* = 3$  кг.

10. Залежність річної заробітної плати  $Y$  від загального виробітку  $X$  показано у вигляді двовимірного статистичного розподілу:

$Y$	$X$					
	1,5	2,5	3,5	4,5	5,5	$n_{yi}$
0,82	1	3	—	—	—	
0,86	—	3	2	1	—	
0,9	—	2	5	9	3	
0,94	—	—	—	6	4	
0,98	—	—	—	—	2	
$n_{xj}$						

Обчислити  $r_B$ ,  $\bar{y}_x = 4,5$ ,  $\bar{x}_y = 0,80$ .

Відповідь.  $r_B = 0,783$ ;  $\bar{y}_{x=4,5} = 0,913$ ;  $\bar{x}_{y=0,86} = 3,17$ .

11. Зі старших класів ліцею було відібрано групу учнів. Дані про їх середньорічні оцінки з математики  $x_i$  та решти дисциплін  $n_i$  (за стобальною системою) наведено в таблиці:

$n_i$	45	30	48	50	52	54	51	60	62	63	65	70	71	74	76	68	79	85
$x_i$	30	35	40	44	48	55	52	65	69	72	78	82	84	86	90	91	92	95

Обчислити  $K_{xy}$ ,  $r_B$ .

Відповідь.  $K_{xy} = 252,62$ ;  $r_B = 0,903$ .

12. При аналізі руди дістали такі дані про відсотковий вміст у ній свинцю та срібла. Результати аналізу наведено в таблиці:

$Y = y_j$	$X = x_j$									
	2,5	7,5	12, 5	17, 5	22, 5	27, 5	32, 5	37, 5	42, 5	$n_{yi}$
2	11 9	9	—	—	—	—	—	—	—	
6	9	59	7	—	—	—	—	—	—	
10	1	4	28	3	—	—	—	—	—	
14	—	—	8	12	4	—	—	—	—	
18	—	—	1	6	7	1	1	—	—	
22	—	—	—	1	1	8	3	—	—	
26	—	—	—	—	—	2	1	—	—	
30	—	—	—	—	—	—	3	2	1	
34	—	—	—	—	—	—	—	—	—	
38	—	—	—	—	—	—	—	—	1	
$n_{xj}$										

Обчислити  $r_B$ ,  $\bar{y}_{x=12,5}$ ,  $\bar{x}_{y=14}$ .

Відповідь.  $r_B = 0,865$ ;  $\bar{y}_{x=12,5} = 3,32\%$ ;  $\bar{x}_{y=14} = 50\%$ .

## Тема № 2. Варіаційні ряди та їх характеристики. Перевірка статистичних гіпотез.

### 1.1. Завдання:

За даними вибірових сукупностей  $A$  і  $B$ :

1. Побудувати варіаційні ряди розподілу частот, відносних частот, накопичених частот, накопичених відносних частот.



2. Обчислити числові характеристики (сумарні статистики): середню арифметичну, моду, медіану; розмах варіації, середнє лінійне відхилення, дисперсію, стандартне відхилення, коефіцієнт варіації; коефіцієнт асиметрії, коефіцієнт ексцесу.

3. Побудувати графіки варіаційного ряду (полігон і гістограму).

1.2. Порядок виконання

### Побудова дискретних варіаційних рядів по вибірковій сукупності *A*.

У таблиці 1 наведені значення елементів вибіркової сукупності *A*, об'єм якої дорівнює  $n = 50$  одиниць.

**Таблиця 1**

Вибіркова сукупність <i>A</i>									
16	13	14	12	15	15	16	15	14	15
16	12	15	12	17	17	14	13	15	16
11	15	16	15	13	13	15	16	14	15
17	15	14	16	14	13	15	16	14	13
17	16	14	14	11	15	16	13	16	15

Дискретні варіаційні ряди розподілу частот, відносних частот, накопичених частот, накопичених відносних частот побудовані в таблиці 2.

**Таблиця 2**

Дискретні варіаційні ряди					
№	Варіанти	Частота	Відносна частота	Накопичена частота	Накопичена відносна частота
1	11	2	0,04	2	0,04
2	12	3	0,06	5	0,1
3	13	7	0,14	12	0,24
4	14	9	0,18	21	0,42
5	15	14	0,28	35	0,7
6	16	11	0,22	46	0,92
7	17	4	0,08	<b>50</b>	<b>1</b>
	Сума	<b>50</b>	<b>1</b>		

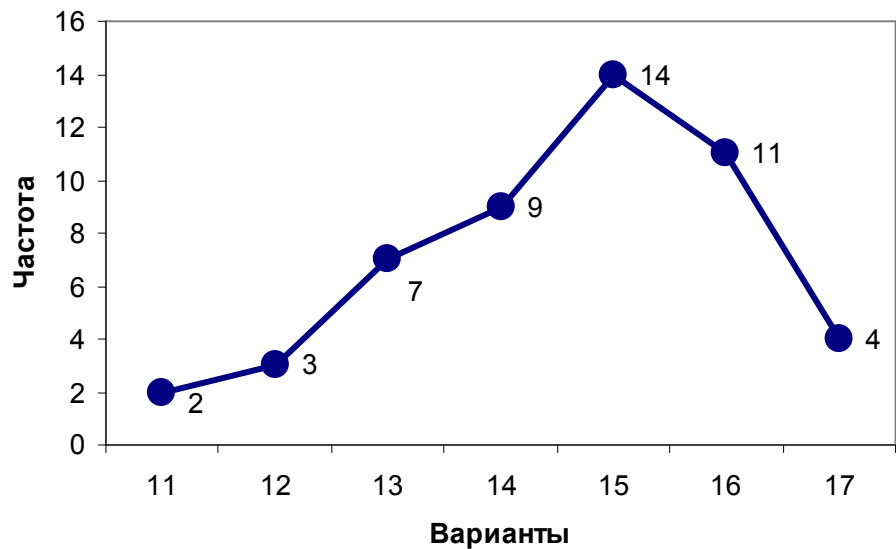
Результати обчислення числових характеристик (сумарних статистик) вибіркової сукупності *A* наведені в таблиці 3.

**Таблиця 3**

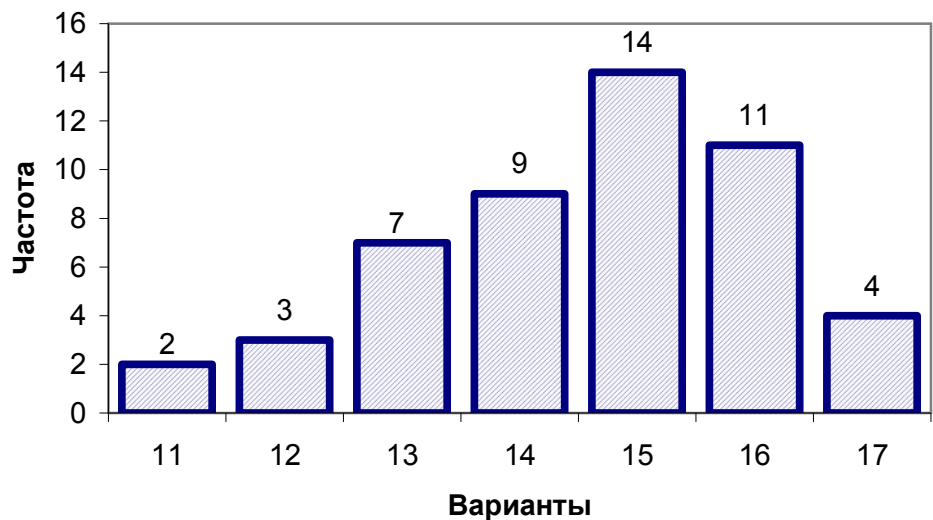
Сумарні статистики вибіркової сукупності <i>A</i>	
Числова характеристика	Значення характеристики
Об'єм вибірки	50
Середня арифметична	14,58
Мода	15
Медіана	15
Мінімум	11
Максимум	17
Розмах варіації	6

Числова характеристика	Значення характеристики
Середнє лінійне відхилення	1,25
Дисперсія вибірки	2,33
Стандартне відхилення	1,53
Коефіцієнт варіації (%)	15,98
Коефіцієнт асиметрії	-0,50
Коефіцієнт ексцесу	-0,27

На рисунках 1, 2 наведений полігон і гістограма частот вибіркової сукупності *A*.



**Рис. 1. Полігон частот вибіркової сукупності *A***



**Рис. 2. Гістограма частот вибіркової сукупності *A***

Побудова варіаційних рядів, обчислення числових характеристик і графічне зображення статистичних даних можна здійснити з використанням електронної таблиці Microsoft Excel.

Допустимо, що дані, по яких необхідно побудувати дискретні варіаційні ряди перебувають в комірках від A2 по J6 (рис. 3).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	<b>Выборочная совокупность A:</b>																
2	16	13	14	12	15	15	16	15	14	15							
3	16	12	15	12	17	17	14	13	15	16							
4	11	15	16	15	13	13	15	16	14	15							
5	17	15	14	16	14	13	15	16	14	13							
6	17	16	14	14	11	15	16	13	16	15							
7																	
8	<b>Суммарные статистики:</b>																
9																	
10	Объем выборки					50											
11	Средняя арифметическая					14,6											
12	Мода					15											
13	Медиана					15											
14	Минимум					11											
15	Максимум					17											
16	Размах вариации					6											
17	Среднее линейное отклонение					1,25											
18	Дисперсия выборки					2,33											
19	Стандартное отклонение					1,53											
20	Коэффициент вариации (%)					15,98											
21	Коэффициент асимметрии					-0,50											
22	Коэффициент эксцесса					-0,27											

**Дискретные вариационные ряды**

№	Варианты	Частота	Относительная частота	Накопленная частота	Накопленная относительная частота
1	11	2	0,04	2	0,04
2	12	3	0,06	5	0,1
3	13	7	0,14	12	0,24
4	14	9	0,18	21	0,42
5	15	14	0,28	35	0,7
6	16	11	0,22	46	0,92
7	17	4	0,08	50	1
	Сумма	50	1		

**Рис. 3. Вихідні дані, варіаційні ряди та сумарні статистики вибіркової сукупності A**

У таблицях 4 і 5 наведені формули та функції Microsoft Excel, які використовуються при побудові дискретних варіаційних рядів і обчисленні числових характеристик

**Таблиця 4**

**Дискретні варіаційні ряди**

№	Варианти	Частота	Відносна частота	Накопичена частота	Накопичена відносна частота
1	11	=СЧЁТЕСЛИ(\$A\$2:\$J\$6;"11")	=N10/\$F\$10	=N10	=P10/\$F\$10
2	12	=СЧЁТЕСЛИ(\$A\$2:\$J\$6;"12")	=N11/\$F\$10	=P10+N11	=P11/\$F\$10
3	13	=СЧЁТЕСЛИ(\$A\$2:\$J\$6;"13")	=N12/\$F\$10	=P11+N12	=P12/\$F\$10
4	14	=СЧЁТЕСЛИ(\$A\$2:\$J\$6;"14")	=N13/\$F\$10	=P12+N13	=P13/\$F\$10
5	15	=СЧЁТЕСЛИ(\$A\$2:\$J\$6;"15")	=N14/\$F\$10	=P13+N14	=P14/\$F\$10
6	16	=СЧЁТЕСЛИ(\$A\$2:\$J\$6;"16")	=N15/\$F\$10	=P14+N15	=P15/\$F\$10
7	17	=СЧЁТЕСЛИ(\$A\$2:\$J\$6;"17")	=N16/\$F\$10	=P15+N16	=P16/\$F\$10
	Сума	=СУММ(N10:N16)	=СУММ(O10:O16)		

**Таблиця 5**

**Сумарні статистики вибіркової сукупності A**

Числова характеристика	Формула Microsoft Excel
Об'єм вибірки	=СЧЁТ(A2:J6)

Числова характеристика	Формула Microsoft Excel
Середня арифметична	=СРЗНАЧ(A2:J6)
Мода	=МОДА(A2:J6)
Медіана	=МЕДИАНА(A2:J6)
Мінімум	=МИН(A2:J6)
Максимум	=МАКС(A2:J6)
Розмах варіації	=F15-F14
Середнє лінійне відхилення	=СРОТКЛ(A2:J6)
Дисперсія вибірки	=ДИСП(A2:J6)
Стандартне відхилення	=КОРЕНЬ(F18)
Коефіцієнт варіації (%)	=(F18/F11)*100
Коефіцієнт асиметрії	=СКОС(A2:J6)
Коефіцієнт ексцесу	=ЭКСЦЕСС(A2:J6)

Діаграми можна створити на окремому аркуші або помістити як впроваджений об'єкт на аркуш із даними. Щоб створити діаграму варто скористатися майстром діаграм для покрокового створення діаграми, при якому вибираються її тип і різні параметри.

### Побудова інтервальних варіаційних рядів по вибірковій сукупності **B**.

У таблиці 6 наведені значення елементів вибіркової сукупності **B**, об'єм якої дорівнює  $n = 100$  одиниць.

**Таблиця 6**

Вибіркова сукупність <b>B</b>									
67	70	64	71	71	73	69	73	72	75
73	80	69	66	66	81	75	62	65	73
69	70	69	73	74	75	71	74	69	72
71	75	71	67	73	75	65	78	72	75
74	76	67	80	71	70	70	70	64	62
71	65	69	70	70	65	77	69	77	74
71	80	71	65	67	70	72	79	67	60
73	71	71	71	66	74	69	68	69	64
68	72	69	67	68	80	70	75	68	62
68	72	66	69	73	69	66	69	75	64

Інтервальні варіаційні ряди розподілу частот, відносних частот, накопичених частот, накопичених відносних частот побудовані в таблиці 7.

Таблиця 7

## Інтервальні варіаційні ряди

№	Нижня границя інтервалу	Верхня границя інтервалу	Середина інтервалу	Частота	Відносна частота	Накопичена частота	Накопичена відносна частота
1	59	62	60,5	4	0,04	4	0,04
2	62	65	63,5	9	0,09	13	0,13
3	65	68	66,5	16	0,16	29	0,29
4	68	71	69,5	34	0,34	63	0,63
5	71	74	72,5	19	0,19	82	0,82
6	74	77	75,5	11	0,11	93	0,93
7	77	80	78,5	6	0,06	99	0,99
8	80	83	81,5	1	0,01	<b>100</b>	<b>1</b>
Сума				<b>100</b>	<b>1</b>		

Результати обчислення числових характеристик (сумарних статистик) вибіркової сукупності **B** по варіаційному ряді та за вихідним даними наведені, відповідно, у таблицях 8, 9.

Таблиця 8

Сумарні статистики (по варіаційному ряді) вибіркової сукупності **B**

Числова характеристика	Значення характеристики
Об'єм вибірки	100
Середня арифметична	70,01
Мода	69,64
Медіана	69,85
Мінімум	60
Максимум	81
Розмах варіації	21
Середнє лінійне відхилення	3,40
Дисперсія вибірки	19,83
Стандартне відхилення	4,45
Коефіцієнт варіації (%)	6,36
Коефіцієнт асиметрії	0,08
Коефіцієнт ексцесу	-0,19

Таблиця 9

Сумарні статистики (за вихідним даними) вибіркової сукупності **B**

Числова характеристика	Значення характеристики
Об'єм вибірки	100
Середня арифметична	70,52
Мода	69

Числова характеристика	Значення характеристики
Медіана	70
Мінімум	60
Максимум	81
Розмах варіації	21
Середнє лінійне відхилення	3,41
Дисперсія вибірки	19,08
Стандартне відхилення	4,37
Коефіцієнт варіації (%)	6,19
Коефіцієнт асиметрії	0,19
Коефіцієнт ексцесу	-0,01

Розбіжності в значеннях числових характеристик вибіркової сукупності **B** обчислених по варіаційному ряді та за вихідним даними пояснюються тим, що значення ознаки усередині кожного інтервалу приймаються рівними серединам інтервалів.

На рисунках 4, 5 наведений полігон і гістограма частот вибіркової сукупності **B**.

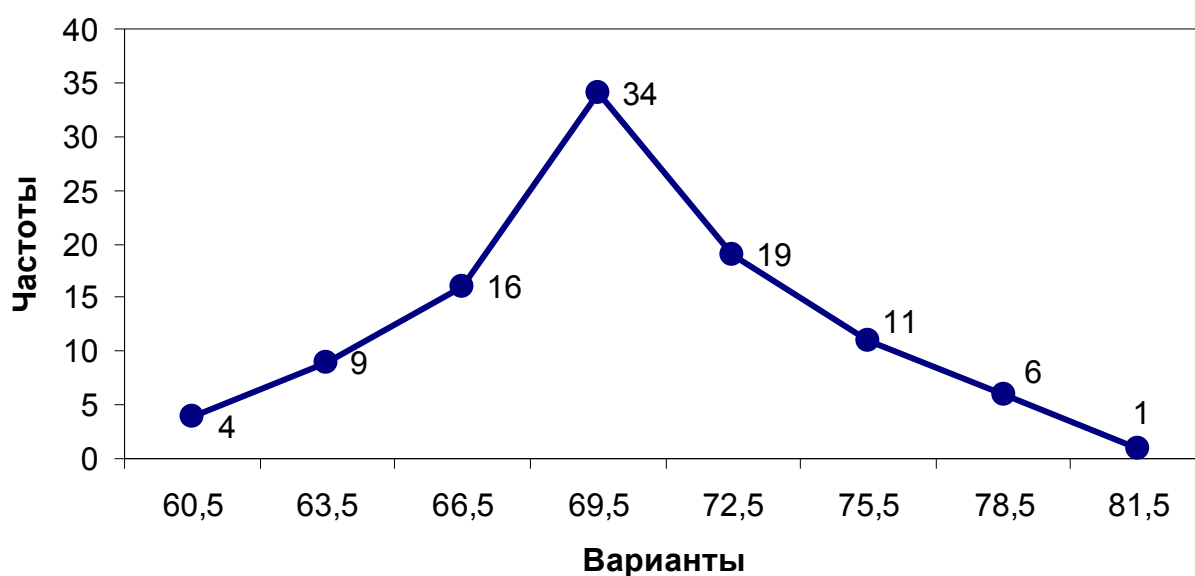


Рис. 4. Полігон частот вибіркової сукупності **B**



27			
28	Суммарные статистики (по исходным данным):		Вспомогательные вычисления
29	Число наблюдений	100	№
30	Средняя арифметическая	70,52	$ v_i - \bar{x}  \cdot f_i$
31	Мода	69	$(v_i - \bar{x})^2 \cdot f_i$
32	Медиана	70	$(v_i - \bar{x})^3 \cdot f_i$
33	Минимум	60	$(v_i - \bar{x})^4 \cdot f_i$
34	Максимум	81	
35	Размах вариации	21	
36	Среднее линейное отклонение	3,41	
37	Дисперсия выборки	19,08	
38	Стандартное отклонение	4,37	
39	Кoeffициент вариации (%)	6,19	
40	Кoeffициент асимметрии	0,19	
41	Кoeffициент эксцесса	-0,01	

№	$ v_i - \bar{x}  \cdot f_i$	$(v_i - \bar{x})^2 \cdot f_i$	$(v_i - \bar{x})^3 \cdot f_i$	$(v_i - \bar{x})^4 \cdot f_i$
1	38,04	361,76	-3440,3	32717,6
2	58,59	381,42	-2483,1	16164,7
3	56,16	197,12	-691,9	2428,6
4	17,34	8,84	-4,5	2,3
5	47,31	117,80	293,3	730,4
6	60,39	331,54	1820,2	9992,7
7	50,94	432,48	3671,8	31173,2
8	11,49	132,02	1516,9	17429,3
Сумма	340,26	1962,99	682,36	110638,78

Рис. 6. Вихідні дані, варіаційні ряди та сумарні статистики вибіркової сукупностіВ  
Таблиця 10

#### Інтервальні варіаційні ряди

№	Нижня границя інтервалу	Верхня границя інтервалу	Середина інтервалу	Частота	Відносна частота	Накопичена частота	Накопичена відносна частота
1	59	62	=(F15+G15)/2	{=ЧАСТОТА(B2:K11;G15:G22)}	=I15/\$C\$14	=I15	=K15/\$C\$14
2	62	65	=(F16+G16)/2	{=ЧАСТОТА(B2:K11;G15:G22)}	=I16/\$C\$14	=K15+I16	=K16/\$C\$14
3	65	68	=(F17+G17)/2	{=ЧАСТОТА(B2:K11;G15:G22)}	=I17/\$C\$14	=K16+I17	=K17/\$C\$14
4	68	71	=(F18+G18)/2	{=ЧАСТОТА(B2:K11;G15:G22)}	=I18/\$C\$14	=K17+I18	=K18/\$C\$14
5	71	74	=(F19+G19)/2	{=ЧАСТОТА(B2:K11;G15:G22)}	=I19/\$C\$14	=K18+I19	=K19/\$C\$14
6	74	77	=(F20+G20)/2	{=ЧАСТОТА(B2:K11;G15:G22)}	=I20/\$C\$14	=K19+I20	=K20/\$C\$14
7	77	80	=(F21+G21)/2	{=ЧАСТОТА(B2:K11;G15:G22)}	=I21/\$C\$14	=K20+I21	=K21/\$C\$14
8	80	83	=(F22+G22)/2	{=ЧАСТОТА(B2:K11;G15:G22)}	=I22/\$C\$14	=K21+I22	=K22/\$C\$14
	Сума			=СУМ(I15:I22)	=СУМ(J15:J22)		

■ Функція ЧАСТОТА обчислює частоту появи значень в інтервалі значень і повертає масив цифр. ЧАСТОТА вводиться як **формула масиву** після **виділення** інтервалу суміжних комірок, у які потрібно повернути отриманий масив розподілу.



Формула масиву обробляє кілька наборів значень, названих аргументами масиву. Кожний аргумент масиву повинен включати однакове число рядків і стовпців. Формула масиву створюється так само, як і інші формули, з тією різницею, що для уведення такої формули виділите діапазон, починаючи з комірки, що містить формулу. Натисніть клавішу F2, а потім натисніть клавіші CTRL+SHIFT+ENTER. **Якщо формула не буде уведена як формула масиву, єдине значення буде дорівнює 1.**

У таблиці 11 наведені формули Microsoft Excel для обчислення числових характеристик по варіаційному ряді.

Таблиця 11

**Обчислення числових характеристик по варіаційному ряді**

Числова характеристика	Формула Microsoft Excel
Об'єм вибірки	=СЧЁТ(B2:K11)
Середня арифметична	=СУММПРОИЗВ(H15:H22;I15:I22)/C14
Мода	=F18+(G18-F18)*(I18-I17)/((I18-I17)+(I18-I19))
Медіана	=F18+(G18-F18)*(C14/2-СУММ(I15:I17))/I18
Мінімум	=МИН(B2:K11)
Максимум	=МАКС(B2:K11)
Розмах варіації	=C19-C18
Середнє лінійне відхилення	=F38/C14
Дисперсія вибірки	=G38/(C14-1)
Стандартне відхилення	=КОРЕНЬ(C22)
Коефіцієнт варіації (%)	=(C23/C15)*100
Коефіцієнт асиметрії	=(H38/C14)/C23^3
Коефіцієнт ексцесу	=(I38/C14)/C23^4-3

У таблиці 12 наведені формули для допоміжних обчислень сум  $\sum_{i=1}^k |v_i - \bar{x}| \cdot f_i$ ,  $\sum_{i=1}^k (v_i - \bar{x})^2 \cdot f_i$ ,  $\sum_{i=1}^k (v_i - \bar{x})^3 \cdot f_i$ ,  $\frac{1}{n} \sum_{i=1}^k (v_i - \bar{x})^4 \cdot f_i$  (комірки E30:I39), які необхідні для визначення характеристик  $d = \frac{1}{n} \sum_{i=1}^k |v_i - \bar{x}| \cdot f_i$ ,  $S^2 = \frac{1}{n-1} \sum_{i=1}^k (v_i - \bar{x})^2 \cdot f_i$ ,  $As = \frac{\frac{1}{n} \sum_{i=1}^k (v_i - \bar{x})^3 \cdot f_i}{S^3}$ ,  $Ek = \frac{\frac{1}{n} \sum_{i=1}^k (v_i - \bar{x})^4 \cdot f_i}{S^4} - 3$ , відповідно.

Таблиця 12

**Допоміжні обчислення**

№	$ v_i - \bar{x}  \cdot f_i$	$(v_i - \bar{x})^2 \cdot f_i$	$(v_i - \bar{x})^3 \cdot f_i$	$(v_i - \bar{x})^4 \cdot f_i$
1	=ABS(H15-\$C\$15)*I15	=(H15-\$C\$15)^2*I15	=(H15-\$C\$15)^3*I15	=(H15-\$C\$15)^4*I15
2	=ABS(H16-\$C\$15)*I16	=(H16-\$C\$15)^2*I16	=(H16-\$C\$15)^3*I16	=(H16-\$C\$15)^4*I16
3	=ABS(H17-\$C\$15)*I17	=(H17-\$C\$15)^2*I17	=(H17-\$C\$15)^3*I17	=(H17-\$C\$15)^4*I17
4	=ABS(H18-\$C\$15)*I18	=(H18-\$C\$15)^2*I18	=(H18-\$C\$15)^3*I18	=(H18-\$C\$15)^4*I18
5	=ABS(H19-\$C\$15)*I19	=(H19-\$C\$15)^2*I19	=(H19-\$C\$15)^3*I19	=(H19-\$C\$15)^4*I19
6	=ABS(H20-\$C\$15)*I20	=(H20-\$C\$15)^2*I20	=(H20-\$C\$15)^3*I20	=(H20-\$C\$15)^4*I20
7	=ABS(H21-\$C\$15)*I21	=(H21-\$C\$15)^2*I21	=(H21-\$C\$15)^3*I21	=(H21-\$C\$15)^4*I21
8	=ABS(H22-\$C\$15)*I22	=(H22-\$C\$15)^2*I22	=(H22-\$C\$15)^3*I22	=(H22-\$C\$15)^4*I22
Сума	=СУМ(F30:F37)	=СУМ(G30:G37)	=СУМ(H30:H37)	=СУМ(I30:I37)

У таблиці 13 наведені формули та функції Microsoft Excel, які використовуються при обчисленні числових характеристик вибіркової сукупності безпосередньо за вихідним даними, тобто без побудови варіаційного ряду.

**Таблиця 13**

**Обчислення числових характеристик за вихідним даними**

Числова характеристика	Формула Microsoft Excel
Об'єм вибірки	=СЧЁТ(B2:K11)
Середня арифметична	=СРЗНАЧ(B2:K11)
Мода	=МОДА(B2:K11)
Медіана	=МЕДИАНА(B2:K11)
Мінімум	=МИН(B2:K11)
Максимум	=МАКС(B2:K11)
Середнє лінійне відхилення	=C34-C33
Розмах варіації	=СРОТКЛ(B2:K11)
Дисперсія вибірки	=ДИСП(B2:K11)
Стандартне відхилення	=СТАНДОТКЛОН(B2:K11)
Коефіцієнт варіації (%)	=(C38/C30)*100
Коефіцієнт асиметрії	=СКОС(B2:K11)
Коефіцієнт ексцесу	=ЭКЦЕСС(B2:K11)

**Тема №3. Основи кореляційного та регресійного аналізу.**

1.В таблиці 9.1 відображені відсоток непрацюючих осіб та кількість крадіжок приватного майна на 10 тисяч населення в десяти районах. Визначити наявність або відсутність взаємозв'язку між двома рядами розподілу методом співставлення паралельних рядів.

Таблиця 9.1

Район	Відсоток непрацюючих осіб	Кількість крадіжок на 10000 населення
1	22	38
2	18	32
3	30	46
4	16	36
5	24	39
6	20	34
7	26	36
8	21	30
9	19	28
10	28	38

2. В таблиці 9.2 відображені споживання спиртних напоїв на душу населення в літрах та кількість засуджених за хуліганство на 100 тисяч населення в десяти районах. Визначить наявність або відсутність взаємозв'язку між цими показниками методом співставлення паралельних рядів.

Таблиця 9.2

Район	Споживання спиртних напоїв на душу населення	Кількість засуджених за хуліганство на 100 тис. населення
1	3,5	76
2	4	85
3	4,3	94
4	3,2	80
5	5	102
6	3,8	74
7	6	108
8	5,2	104
9	4,8	98
10	7,4	112

3. В таблиці 9.3 відображені середньомісячна кількість зареєстрованих в районі адміністративних правопорушень та злочинів, вчинених неповнолітніми особами. Визначить наявність або відсутність взаємозв'язку між цими показниками методом співставлення паралельних рядів.

Таблиця 9.3

Рік	1998	1999	2000	2001	2002	2003	2004	2005
Кількість адміністративних правопорушень	10	12	14	11	13	9	15	16
Кількість злочинів	4	5	7	3	6	2	8	7

4. В таблиці 9.4 відображені кількість зареєстрованих в районі злочинів та кількість злочинів, що залишились нерозкритими. Визначить наявність або відсутність взаємозв'язку між двома рядами розподілу методом співставлення паралельних рядів.

Таблиця 9.4

Рік	1998	1999	2000	2001	2002	2003	2004	2005
-----	------	------	------	------	------	------	------	------

Кількість зареєстрованих злочинів	324	316	310	308	311	302	296	303
Кількість злочинів, що залишились нерозкритими	12	18	10	11	8	12	14	13

5. За даними таблиці 9.1 визначить наявність або відсутність взаємозв'язку між показниками графічним шляхом за допомогою поля кореляції.

6. За даними таблиці 9.2 визначить наявність або відсутність взаємозв'язку між показниками графічним шляхом за допомогою поля кореляції.

7. За даними таблиці 9.3 визначить наявність або відсутність взаємозв'язку між показниками графічним шляхом за допомогою поля кореляції.

8. За даними таблиці 9.4 визначить наявність або відсутність взаємозв'язку між показниками графічним шляхом за допомогою поля кореляції.

9. В таблиці 9.5 відображені відсоток непрацюючих осіб та кількість грабежів на 10 тисяч населення в десяти районах. Використовуючи метод статистичного групування визначить наявність або відсутність зв'язку між кількістю непрацюючих осіб та коефіцієнтом злочинної інтенсивності грабежів.

Таблиця 9.5

Показники	Р а й о н и									
	1	2	3	4	5	6	7	8	9	10
Відсоток непрацюючих осіб	12	15	10	16	20	32	14	24	26	18
Кількість грабежів на 10000 населення	6	8	9	5	10	18	4	14	12	7

Статистична модель зв'язку соціально-правових явищ (рівняння регресії) виражається математичною функцією:

$$\bar{y}_x = f(x_1, x_2, \dots, x_n)$$

При дослідженні парної регресії, що характеризує зв'язок між двома ознаками – факторною та результативною, аналітичне вираження зв'язку описується наступними рівняннями:

$$\text{прямої - } \bar{y}_x = a + bx \quad \text{гіперболи - } \bar{y}_x = a + b \frac{1}{x} \quad \text{параболи - } \bar{y}_x = a + bx + cx^2$$

У цих рівняннях  $a, b$  і  $c$  – параметри рівнянь регресії

Встановити вид рівняння можна лише досліджуючи зв'язок графічно.

Оцінка параметрів регресії ( $a, b$ ) здійснюється *методом найменших квадратів*, сутність якого полягає в знаходженні параметрів  $a$  і  $b$ , при яких мінімізується сума квадратів відхилень *емпіричних (фактичних)* значень результативної ознаки від *теоретичних*, отриманих за обраним рівнянням регресії:

$$\sum (y - \bar{y}_x)^2 = \min$$

Для лінійної залежності це виглядає так:

$$\sum (y - a - bx)^2 = \min$$

Проводячи математичне перетворення (диференціювання) одержуємо систему нормальних рівнянь:

$$\begin{cases} na + b \sum y = \sum y \\ a \sum x + b \sum x^2 = \sum xy \end{cases}$$

де:  $n$  – обсяг досліджуваної сукупності (кількість одиниць спостереження)

Вирішуючи цю систему рівнянь можна визначити параметри лінійного рівняння:

$$a = \frac{\sum x^2 \sum y - \sum x \sum xy}{n \sum x^2 - \sum x \sum x} \quad b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - \sum x \sum x}$$

**Приклад.** Побудуємо рівняння регресії для зв'язку між ступенем незайнятості та інтенсивністю злочинності неповнолітніх осіб у п'яти районах (таблиця 9.6).

Таблиця 9.6

Район	Кількість н/л осіб, які не працюють і не навчаються на 1000 н/л (x)	Кількість злочинів, вчинених н/л на 10000 населення (y)	$x^2$	xy	$\bar{y}_x$
1	7	4	49	28	3,25
2	9	3	81	27	4,02
3	12	6	144	72	5,2
4	14	5	196	70	5,98
5	18	8	324	144	7,54

$$\sum x = 60 \quad \sum y = 26 \quad \sum x^2 = 794 \quad \sum xy = 341$$

Система нормальних рівнянь для цього прикладу має вигляд:

$$\begin{cases} 5a + b60 = 26 \\ a60 + b794 = 341 \end{cases}$$

Звідси:  $a = 0,52$ ;  $b = 0,39$ ; при цьому  $\bar{y}_x = 0,52 + 0,39X$

Підставляючи значення параметрів  $a$  і  $b$  у рівняння прямої, знаходимо теоретичні вирівняні значення  $\bar{y}_x$  (табл.9.6, гр.6).

Якщо в результаті графічного дослідження встановлена криволінійна залежність, яка має форму параболи другого порядку, зв'язок виражається рівнянням кривої  $\bar{y}_x = a + bx + cx^2$ . Необхідно знайти параметри  $a$ ,  $b$ ,  $c$ .

Для цього необхідно розв'язати систему нормальних рівнянь:

$$\begin{cases} na + b\sum x + c\sum x^2 = \sum y \\ a\sum x + b\sum x^2 + c\sum x^3 = \sum xy \\ a\sum x^2 + b\sum x^3 + c\sum x^4 = \sum x^2 y \end{cases}$$

Для гіперболи:  $\bar{Y} = a + b\frac{1}{x}$

$$\begin{cases} na + b\sum \frac{1}{x} = \sum y \\ a\sum \frac{1}{x} + b\sum \frac{1}{x^2} = \sum \frac{1}{x} y \end{cases}$$

Параметри гіперболи мають такий вигляд:

$$a = \frac{\sum \frac{1}{x^2} \sum y - \sum \frac{1}{x} \sum \frac{y}{x}}{n \sum \frac{1}{x^2} - \sum \frac{1}{x} \sum \frac{1}{x}} \quad b = \frac{n \sum \frac{y}{x} - \sum \frac{1}{x} \sum y}{n \sum \frac{1}{x^2} - \sum \frac{1}{x} \sum \frac{1}{x}}$$

Розв'язавши систему цих рівнянь знайдемо параметри  $a$ ,  $b$ ,  $c$  і отримаємо рівняння регресії.

## Завдання

1. За даними таблиці 9.1 визначить аналітичне вираження взаємозв'язку між кількістю непрацюючих осіб та кількістю крадіжок приватного майна на 10 тисяч населення (рівняння регресії).

2. За даними таблиці 9.2 визначить аналітичне вираження взаємозв'язку між споживанням спиртних напоїв на душу населення та кількістю засуджених за хуліганство на 100 тисяч населення (рівняння регресії).

3. За даними таблиці 9.3 визначить аналітичне вираження взаємозв'язку адміністративними правопорушеннями та злочинами, вчинених неповнолітніми особами (рівняння регресії).

4. За даними таблиці 9.7 визначить аналітичне вираження взаємозв'язку між середньою заробітною платою та коефіцієнтом злочинної інтенсивності в семи районах (рівняння регресії).

**Таблиця 9.7**

Район	Середня заробітна плата (тис. грн.)	Коефіцієнт злочинної інтенсивності (на 10 тис. населення)
1	1,08	70
2	1,63	62
3	1,02	75
4	1,5	50
5	0,85	80
6	0,73	86
7	0,61	94

5. За даними таблиці 9.8 визначить аналітичне вираження взаємозв'язку між віком та кількістю осіб, які вчинили злочини (рівняння регресії).

**Таблиця 9.8**

Вік	14-17	18-24	25-29	30-39	40-49	50-59
Кількість осіб, які вчинили злочини	12	26	34	42	28	8

6. За даними таблиці 9.9 визначить аналітичне вираження взаємозв'язку між віком робітників та їх середньою продуктивністю праці (рівняння регресії).

**Таблиця 9.9**

Вік	20	25	30	35	40	45	50
Середня кількість виготовлених деталей	24	28	32	38	46	34	26

**Приклад.** Визначимо тісноту зв'язку між рівнем незайнятості та інтенсивністю злочинної діяльності неповнолітніх осіб у п'яти районах (таблиця 9.10). Щоб усунути вплив інших факторів в якості ряду злочинів будемо використовувати згладжений (теоретичний) ряд.

**Таблиця 9.10**

Райони	Кількість н/л, які не працюють і не навчаються, на 1000 н/л ( $x$ )	Кількість злочинів, вчинених н/л на 10000 населення ( $y$ )	$dx$	$dy$	$dx^2$	$dy^2$
1	7	4	-5	-1,95	25	3,8
2	9	3	-3	-1,18	9	1,39
3	12	6	0	0	0	0
4	14	5	2	0,78	4	0,6
5	18	8	6	2,34	36	5,47

	$\bar{x}=12$	$\bar{y}=5,2$			$\Sigma dx^2=74$	$\Sigma dy^2=14,8$
--	--------------	---------------	--	--	------------------	--------------------

$$R = \frac{(-5) \cdot (-1,95) + (-3) \cdot (-1,18) + 0 \cdot 0 + 2 \cdot (0,78) + 6 \cdot 2,34}{\sqrt{74 \cdot 11,26}} = 0,9$$

Значення лінійного коефіцієнта кореляції  $0,9$  свідчить про наявність сильного та прямого зв'язку між рівнем незайнятості та інтенсивністю злочинності неповнолітніх осіб.

*Коефіцієнт рангової кореляції Спірмена* розраховується за формулою:

$$\rho = 1 - \frac{6 \cdot \sum d^2}{n \cdot (n^2 - 1)}$$

де:  $d$  – різниця рангів,  $n$  – кількість рангів

*Ранг* – це порядковий номер значень ознаки, розташованих в порядку зростання або зменшення їх величин.

**Приклад.** Ряд  $x$  у таблиці 9.11 проранжируємо (визначимо ранги або номери місць) від 1 до 5. Так як значення  $x$  вже розміщені в порядку зростання, то значення рангів збігаються зі значеннями графі *райони*. Проранжируємо ряд  $y$  (графа 5).

Таблиця 9.11

Райони	Кількість неповнолітніх, які не працюють і не навчаються, на 1000 н/л ( $x$ )	Кількість злочинів, вчинених неповнолітніми на 1000 н/л ( $y$ )	Ранги $x$	Ранги $y$	Різниця рангів ( $d$ )	$d^2$
1	7	4	1	2	1	1
2	9	3	2	1	1	1
3	12	6	3	4	1	1
4	14	5	4	3	1	1
5	18	8	5	5	0	0

Визначимо різниці рангів ( $d$ ), зведемо їх у квадрат ( $d^2$ ) і просумуємо ( $\sum d^2 = 4$ ).  
*Коефіцієнт Спірмена* розраховуємо за такою формулою:

$$\rho = 1 - \frac{6 \cdot \sum d^2}{n \cdot (n^2 - 1)} = 1 - \frac{6 \cdot 4}{5 \cdot (25 - 1)} = 0,8$$

Цей коефіцієнт змінюється від  $+1$  до  $-1$ . Таким чином, значення коефіцієнта Спірмена  $0,8$  свідчить про наявність сильного та прямого зв'язку між рівнем незайнятості та інтенсивністю злочинності неповнолітніх осіб.

Для визначення тісноти зв'язку двох *якісних* ознак використовується *коефіцієнт контингенції Пірсона*.

При цьому кожна з ознак складається з двох альтернативних значень (наприклад, злочин *розкритий*, злочин *не розкритий*).

Числовий матеріал розташовують у вигляді таблиці спряженості, яка називається «*таблицею чотирьох полів*» (таблиця 9.12).

Таблиця 9.12

$\begin{matrix} \diagdown \\ X \end{matrix} \begin{matrix} Y \\ \diagup \end{matrix}$	В	не В	сума
А	a	b	a + b
не А	c	d	c + d
сума	a + c	b + d	a + b + c + d

*Коефіцієнт контингенції* визначається за формулою:

$$Kk = \frac{ad - bc}{\sqrt{(a + b) \cdot (b + d) \cdot (a + c) \cdot (c + d)}}$$

Зв'язок вважається підтвердженим, якщо  $Ka \geq 0,3$ .

**Приклад.** Досліджено 100 осіб, які виховувалися в повній або благополучній родині, і 100 осіб, які виховувалися в неповній або неблагополучній родині, щодо вчинення ними злочинів. Результати дослідження відображені в таблиці 9.13. Визначимо ступінь щільності зв'язку між цими ознаками.

Таблиця 9.13

Показники	Кількість осіб, які вчинили злочини	Кількість осіб, які не вчинили злочини	Сума
Виховання в неповній або неблагополучній родині	38	62	100
Виховання в повній або благополучній родині	11	89	100
Сума	49	151	200

Коефіцієнт контингенції буде дорівнювати:

$$Kk = \frac{38 \cdot 89 - 11 \cdot 62}{\sqrt{100 \cdot 151 \cdot 49 \cdot 100}} = \frac{2700}{8602} = 0,31$$

Виходячи зі значення коефіцієнта можна зробити висновок, що зв'язок між вихованням дітей в неповній або неблагополучній родині і злочинністю цілком підтверджується.

Якщо по кожній з взаємозалежних ознак виділяється число груп більше двох, то для подібного роду таблиць щільність зв'язку між якісними ознаками може бути вимірена за допомогою коефіцієнта взаємної спряженості Чупрова.

Коефіцієнт взаємної спряженості Чупрова визначається за формулою:

$$Kc = \sqrt{\frac{\varphi^2}{\sqrt{(K_1 - 1)(K_2 - 1)}}}$$

де  $\varphi^2$  – сума відносин квадратів частот кожної клітини таблиці до добутку підсумкових частот відповідної граfi і рядка без одиниці:

$$\varphi^2 = \sum \frac{f_{i,j}^2}{F_i \cdot F_j} - 1$$

де  $K_1$  – число рядків у таблиці;  $K_2$  – число граfi у таблиці.

**Приклад.** В таблиці 9.14 відображена кількість розкритих та не розкритих злочинів за окремими розділами Особливої частини Кримінального кодексу України. Визначимо, чи є зв'язок між видами злочинів та їх розкриттям.

Таблиця 9.14

Види злочинів ( $f_i$ )	Кількість злочинів ( $f_j$ )		Сума ( $F_j$ )
	Розкритих	Не розкритих	
Злочини проти життя та здоров'я особи	1042	336	1378
Злочини проти власності	17333	5263	22596
Злочини в сфері господарської діяльності	683	232	915
Злочини проти громадської безпеки	566	46	612
Злочини проти безпеки руху та експлуатації транспорту	765	770	1535
Сума ( $F_i$ )	20389	6647	27036



$$\phi^2 = \sum \frac{f_{i \cdot j}^2}{F_{i \cdot} \cdot F_{\cdot j}} - 1 = \frac{1042^2}{20389 \cdot 1378} + \frac{17333^2}{20389 \cdot 22596} + \frac{683^2}{20389 \cdot 915} + \frac{566^2}{20389 \cdot 612} + \frac{765^2}{20389 \cdot 1535} + \frac{336^2}{6647 \cdot 1378} + \frac{5263^2}{6647 \cdot 22596} + \frac{232^2}{6647 \cdot 915} + \frac{46^2}{6647 \cdot 612} + \frac{770^2}{6647 \cdot 1535} - 1 = 0,1132$$

Коефіцієнт взаємної спряженості Чупрова буде дорівнювати:

$$K_{\text{ч}} = \sqrt{\frac{0,1132}{\sqrt{(5-1)(2-1)}}} = 0,238$$

Цей коефіцієнт змінюється від 0 до 1, його значення не може бути негативним. Зв'язок вважається *істотним* при величині  $K_{\text{ч}}=0,3$ . Чим ближче значення коефіцієнта до одиниці, тим щільніше зв'язок.

Розрахований  $K_{\text{ч}}=0,238$  свідчить про наявність зв'язку між видами злочинів та їх розкриттям.

### Завдання до самостійної роботи

1. За даними таблиці 9.1 визначить ступінь тісноти зв'язку між кількістю непрацюючих осіб та кількістю крадіжок приватного майна на 10 тисяч населення шляхом розрахунку лінійного коефіцієнта кореляції.

2. За даними таблиці 9.2 визначить ступінь тісноти зв'язку між показниками шляхом розрахунку лінійного коефіцієнта кореляції.

3. За даними таблиці 9.3 визначить ступінь тісноти зв'язку між споживанням спиртних напоїв на душу населення та кількістю засуджених за хуліганство на 100 тисяч населення шляхом розрахунку коефіцієнта рангової кореляції Спірмена.

4. За даними таблиці 9.7 визначить ступінь тісноти зв'язку між середньою заробітною платою та коефіцієнтом злочинної інтенсивності в семи районах шляхом розрахунку коефіцієнта рангової кореляції Спірмена.

5. В таблиці 9.15 відображені результати опитування працюючих на підприємствах різної форми власності щодо задоволеності рівнем життя. Розрахуйте коефіцієнт контингенції, сформулюйте висновки, що витікають з аналізу отриманого коефіцієнту.

Таблиця 9.15

Форма власності підприємства	Задоволеність рівнем життя		Всього
	Задоволені	Не задоволені	
Державне	30	45	75
Приватне	50	25	75
Всього	80	70	150

7. В таблиці 9.16 наведені результати дослідження осіб, які вживають або не вживають наркотичні речовини, щодо їх сімейного положення. Розрахуйте коефіцієнт контингенції. Зробіть висновки, чи залежить вживання наркотичних речовин від сімейного положення.

Таблиця 9.16

Відношення до наркотиків	Сімейний стан		Всього
	Наявність сім'ї	Відсутність сім'ї	
Споживають	26	72	100
Не споживають	64	36	100
Всього	92	108	200

8. Використовуючи дані таблиці 9.17 встановіть, чи є взаємозв'язок між видами злочинів та їх розкриттям.

Таблиця 9.17

Види злочинів	Розкрито злочинів	Не розкрито злочинів	Всього
Умисне вбивство	38 (95,0%)	2 (5,0%)	40 (100%)
Розбій	31 (91,2%)	3 (8,8%)	34 (100%)
Грабіж	151 (72,9%)	56 (27,1%)	207 (100%)
Порушення правил безпеки руху	46 (63,9%)	26 (36,1%)	72 (100%)
Всього	266	87	353

#### 4. Інформаційне та методичне забезпечення

##### Базова

1. Ермалаев О.Ю. Математическая статистика для психологов. –М.: Московский психолого-социальный институт: Флинта, 2004. – 336 с
2. Сидоренко Е.В. Методы математической обработки в психологии. СПб.: ООО «Речь», 2001. -350с.
3. Гласс Дж., Стенли Дж. Статистические методы в педагогике и психологии. - М.: Изд-во «Прогресс», 1976. – 496 с.
4. Кремер Н.Ш. Теория вероятностей и математическая статистика. -М.: Юнити-Дана, 2001. – 543 с.
5. Гмурман В.Е. Теория вероятностей и математическая статистика. -М.: Высш. шк., 2001. – 479 с.
6. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. -М.: Высш. шк. 2001. – 400 с.
7. Лапач С.Н., Чубенко А.В., Бабич П.Н. Статистические методы в медико-биологических исследованиях с использованием Excel. – К.: МОРИОН, 2001.– 408 с.
8. Колде Я.К. Практикум по теории вероятностей и математической статистике. -М.: Высш. шк. 1991. – 157 с.

##### Допоміжна

1. Вайнберг Д., Шуменер Д. Статистика. – М.: Статистика, 1979. – 390 с.
2. Афифи А., Эйзен С. Статистический анализ. – М.: Мир 1982. – 450 с.
3. Калинина В.Н., Панкин В.Ф. Математическая статистика. Учебное пособие для вузов. – М.: Высш. Школа, 1994. – 336 с.
4. Макаров Ф.А., Тюрин Ю.Н. Анализ данных на компьютере. – М.: Финансы и статистика, 1995. – 400 с.
5. Артемьева Е.Ю., Мартянов Е.М. Вероятностные методы в психологии. – Изд-во Московского университета. – М., 1975. – 206 с.
6. Дружинин Н.К. Выборочное наблюдение и эксперимент. – М.: Статистика, 1977. – 176 с.
7. Енаков И.С. Методы, алгоритмы, программы многомерного статистического анализа. – М.: Финансы и статистика, 1986. – 250 с.