

## **Лабораторна робота № 8.**

### **Тема: Використання однофакторного дисперсійного аналізу**

**Мета:** навчитись за допомогою калькулятора Excel застосовувати дисперсійний аналіз для дослідження впливу як одного фактора, так і взаємодії кількох, на результативну ознаку.

#### **Теоретичні відомості**

Дисперсійний аналіз, запропонований Р. Фішером, є статистичним методом для виявлення впливу окремих факторів на результати проведеного експерименту. Цей метод ґрунтується на припущенні про те, що якщо на об'єкт (групу досліджуваних) впливає декілька незалежних факторів і їх вплив накладається, то загальну дисперсію результативної ознаки можна розкласти на суму дисперсій, що виникає внаслідок дії кожного окремого фактора та дисперсії, зумовленої впливом випадкових чинників (залишкова дисперсія). Порівняння дисперсій, зумовлених впливом різних факторів із випадковою (залишковою) дисперсією, дає змогу оцінити значущість внеску кожного з факторів.

Основою дисперсійного аналізу є припущення про те, що одні змінні можуть розглядатися як причини, а інші – як наслідки. В психологічних дослідженнях саме змінні, що розглядаються як причини, вважаються незалежними факторами (змінними). Суть дисперсійного аналізу полягає в розчленуванні загальної дисперсії залежної змінної на окремі компоненти, зумовлені впливом конкретних факторів, і в перевірці гіпотез про значущість впливу цих факторів на досліджувану ознаку. Порівнюючи компоненти дисперсії одна з одною за допомогою критерію Фішера, можна визначити, яка частка загальної варіації результативної ознаки зумовлена впливом факторів. За кількістю факторів дисперсійний аналіз може бути однофакторним (вивчається вплив одного фактора), двофакторним і багатофакторним.

Нехай потрібно дослідити вплив на ознаку  $X$  певного одного фактора. Результати експерименту ділять на певну кількість груп, які відрізняються між собою ступенем дії фактора. Для зручності проведення обчислень результати експерименту заносимо у таку таблицю:

Ступінь впливу фактора	1	2	3	...	p
Спостережувані значення ознаки X	$x_{11}, x_{21}, x_{31}, \dots, x_{n11}$	$x_{12}, x_{22}, x_{32}, \dots, x_{n22}$	$x_{13}, x_{23}, x_{33}, \dots, x_{n33}$		$x_{1p}, x_{2p}, x_{3p}, \dots, x_{npp}$
Групові середні	$\bar{x}_1 = \frac{1}{n1} \sum_{i=1}^{n1} x_{i1}$	$\bar{x}_2 = \frac{1}{n2} \sum_{i=1}^{n2} x_{i2}$	$\bar{x}_3 = \frac{1}{n3} \sum_{i=1}^{n3} x_{i3}$		$\bar{x}_p = \frac{1}{np} \sum_{i=1}^{np} x_{ip}$
Загальна середня	$\bar{x} = \frac{1}{N} \sum_{j=1}^p \sum_{i=1}^{n_j} x_{ij}, \quad N = \sum_{j=1}^p n_j$				
Вид дисперсій	Сума квадратів відхилень		Число ступенів свободи	Статистичні оцінки дисперсій	
Внутрішньо-групова	$D_{внутр} = \sum_{i=1}^{n1} \sum_{j=1}^p (x_{ij} - \bar{x}_j)^2$		$N-p$	$SS_{внутр} = \frac{D_{внутр}}{N-p}$	
Міжгрупова	$D_{міжгруп} = \sum_{j=1}^p n_j \cdot (\bar{x}_j - \bar{x})^2$		$p-1$	$SS_{міжгруп} = \frac{D_{міжгруп}}{p-1}$	
Загальна	$D_{заг} = \sum_{i=1}^{n1} \sum_{j=1}^p (x_{ij} - \bar{x})^2$		$N-1$	$SS_{заг} = \frac{D_{заг}}{N-1}$	

Формулюємо такі гіпотези:

- $H_0$ : фактор не впливає на досліджувану ознаку .
- $H_1$ : фактор впливає на досліджувану ознаку.

Критичну точку знаходимо з таблиці значень критерію Фішера, критична область є правобічною. Емпіричне значення критерію обчислюємо за такою формулою

$$F_{емп} = \frac{SS_{міжгруп}}{SS_{внутр}}. \quad (8.1)$$

## Завдання

Три різні мобільні додатки відтворювали з різною швидкістю (низькою, середньою, високою) інформацію. Довести (або спростувати) припущення про те, що фактор швидкості пред'явлення інформації впливає на показники їх відтворення. Емпіричні дані наведені у таблиці.

№ п/п	Швидкість пред'явлення		
	Низька	Середня	Висока
1	7	5	5
2	8	5	4
3	7	6	5
4	6	5	3
5	5	4	4
6	7	6	5
7		4	4
8			3

## Хід роботи

Сформулюємо такі гіпотези.

- $H_0$ : відмінності в обсязі відтворення інформації не є вираженішими, ніж випадкові відмінності в середині групи (фактор «швидкість пред'явлення» не впливає на показник відтворення).
  - $H_1$ : відмінності в обсязі відтворення інформації вираженіші, ніж випадкові відмінності всередині групи.
1. Формуємо таблицю для обчислень. Для цього об'єднуємо комірки **A1:A2**, куди вводим назву «№» та комірки **B1:D1**, куди вводим назву «Швидкість пред'явлення». В комірку **B2** вводим назву «Низька», в комірку **C2** вводим назву «Середня», в комірку **D2** вводим назву «Висока». В сформовану таблицю вносимо дані задачі.
  2. Вводим в комірку **A11** назву “**ni**=”, а в комірки **B11**, **C11**, **D11** – числа, що відповідають обсягам відповідних груп, які обчислюємо за формулами:  
**B11: =СЧЁТ(В3:В8);**  
**C11: =СЧЁТ(С3:С9);**  
**D11: =СЧЁТ(Д3:Д10).**
  3. У комірку **A12** вводим назву “**1/ni**=”. У відповідних комірках проводимо обчислення:  
**B12: =1/В11;**  
**C12: =1/С11;**  
**D12: =1/Д11.**

4. У комірку **A13** вводимо назву “**Суми**”. Проводимо потрібні обчислення у відповідних комітках:
  - **B13:** =СУММ(B3:B8);
  - **C13:** =СУММ(C3:C9);
  - **D13:** =СУММ(D3:D10).
5. В комірку **A14** вводимо назву «**Середні**». Проводимо потрібні обчислення у відповідних комітках:
  - **B14:** = СРЗНАЧ(B3:B8);
  - **C14:** =СРЗНАЧ(C3:C9);
  - **D14:** =СРЗНАЧ(D3:D10).
6. Об’єднуємо комірки **A15:B15**, вводимо назву “**Загальний обсяг**”. В комірці **C15** вводимо формулу **C15:** =B11+C11+D11.
7. Об’єднуємо комірки **A16:B16**, вводимо назву “**Загальна середня**”. В комірці **C16** вводимо формулу **C16:**=СРЗНАЧ(B3:B8;C3:C9;D3:D10), як показано на рис. 8.1.

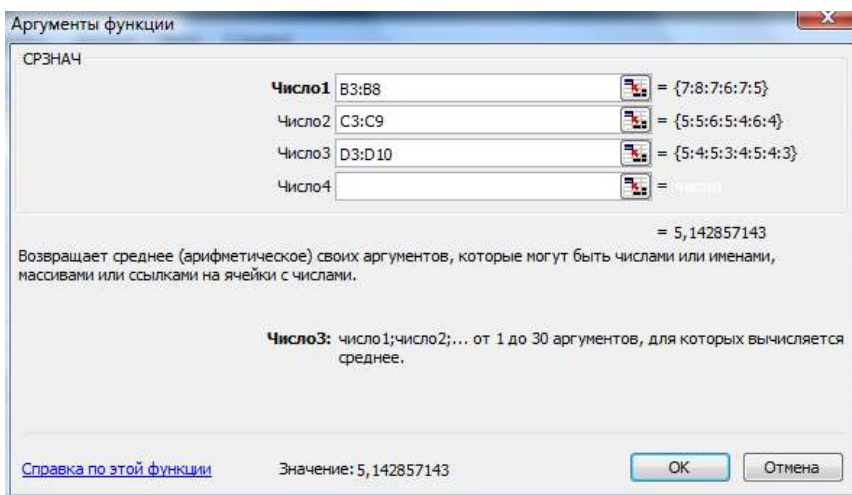


Рис. 8.1. Обчислення середнього значення

8. Об’єднуємо комірки **E1:G1**, вводимо назву “**Квадрати різниць по групах**”. У комірки **E2, F2, G2** вводимо назви «**Низька**», «**Середня**» і «**Висока**».
9. Об’єднуємо комірки **H1:J1**, вводимо назву “**Квадрати різниць із заг. сер.**”. У комірки **H2, I2, J2** вводимо назви «**Низька**», «**Середня**» і «**Висока**».

10. У комірці **E3** вводимо формулу  $=(B3-B\$14)^2$ . Поширюємо її на стовпець **E4: E8**.
11. У комірці **F3** вводимо формулу  $=(C3-C\$14)^2$ . Поширюємо її на стовпець **F4: F9**.
12. У комірці **G3** вводимо формулу  $=(D3-D\$14)^2$ . Поширюємо її на стовпець **G4: G10**.
13. У комірці **H3** вводимо формулу  $=(B3-C\$16)^2$ . Поширюємо її на стовпець **H4: H8**.
14. У комірці **I3** вводимо формулу  $=(C3-C\$16)^2$ . Поширюємо її на стовпець **I4: I9**.
15. У комірці **J3** вводимо формулу  $=(D3-C\$16)^2$ . Поширюємо її на стовпець **J4: J10**.

Результати обчислень показано на рис. 8.2.

	A	B	C	D	E	F	G	H	I	J
1	номери	Швидкість пред'явлення			Квадрати різниць по групах			Квадрати різниць із заг. сер.		
2		Низька	Середня	Висока	Низька	Середня	Висока	Низька	Середня	Висока
3	1	7	5	5	0,111111	0,000000	0,765625	3,44898	0,020408	0,020408
4	2	8	5	4	1,777778	0,000000	0,015625	8,163265	0,020408	1,306122
5	3	7	6	5	0,111111	1,000000	0,765625	3,44898	0,734694	0,020408
6	4	6	5	3	0,444444	0,000000	1,265625	0,734694	0,020408	4,591837
7	5	7	4	4	0,111111	1,000000	0,015625	3,44898	1,306122	1,306122
8	6	5	6	5	2,777778	1,000000	0,765625	0,020408	0,734694	0,020408
9	7		4	4		1,000000	0,015625		1,306122	1,306122
10	8			3			1,265625			4,591837
11	ni=	6	7	8						
12	1/ni=	0,166667	0,142857	0,125						
13	Суми	40	35	33						
14	Середні	6,666667	5	4,125						
15	Загальний обсяг		21							
16	Загальна середня	5,142857								

Рис. 8.2. Результати обчислень квадратів різниць

16. Об'єднуємо комірки **A18:B18**; **A19:B19**; **A20:B20**; **A21:B21**. Вводимо назви: «Вид дисперсій», «Внутрішньогрупова», «Міжгрупова», «Загальна».
17. Об'єднуємо комірки **C18:E18**, **F18:G18**, **H18:I18**. Вводимо назви: «Сума квадратів відхилень», «Ступені свободи», «Статистичні оцінки дисперсій».
18. Об'єднуємо комірки **D19:E19**, **D20:E20**, **D21:E21**, **F19:G19**, **F20:G20**, **F21:G21**, **I19:J19**, **I20:J20**, **I21:J21**.
19. У комірки **C19**, **C20**, **C21** вводимо назви: «Двугрупа», «Дві групи», «Два», «Дві». В комірки **H19**, **H20**, **H21** вводимо назви «SSвнутр», «SSміжгруп», «SSзаг».
20. Об'єднуємо комірки **A22:B22** і вводимо назву «Критерій Фішера». У комірки **A22**, **A23**, **A24** вводимо назви « $F_{\text{емп}}$ », « $F_{0,01}$ », « $F_{0,05}$ ».

21. У комірку **DE19** вводимо формулу **=СУММ(E3:E8;F3:F9;G3:G10)**. Введення даних в діалоговому вікні здійснюється, як у п. 7.
22. Виділяємо блок комірок **B11:D11**, копіюємо їх вміст за допомогою клавіш **Ctrl+C** та здійснюємо спеціальну вставку.
23. У комірку **DE20** вводимо формулу **=B11\*(B14-C16)^2+C11\*(C14-C16)^2+D11\*(D14-C16)^2**.
24. У комірку **DE21** вводимо формулу **=СУММ(H3:H8;I3:I9;J3:J10)**, аналогічно, як у п. 7.
25. У комірки **FG19, FG20, FG21** вводимо формули
  - **FG19: =C15-3;**
  - **FG20: =3-1;**
  - **FG21: =C15-1.**
26. У комірки **IJ19, IJ20, IJ21** вводимо формули:
  - **IJ19: =D19/F19;**
  - **IJ20: =D20/F20;**
  - **IJ21: =D21/F21.**
27. У комірки **B23, B24, B25** вводимо формули:
  - **B23: =I20/I19;**
  - **B24: =ФРАСПОБР(0,05;F20;F19);**
  - **B25: =ФРАСПОБР(0,01;F20;F19).**
28. Робимо висновки на підставі отриманих числових результатів. Якщо емпіричне значення критерію є більшим за критичну точку, то нульову гіпотезу відхиляють (правобічна критична область). Результати обчислень наведено на рис. 8.3.

18	Вид дисперсії	Сума квадратів відхилень	Ступені свободи	Статистичні оцінки дисперсій
19	Внутрішньогрупова	$D_{внутр} = 14,208333$	18	$SS_{внутр} = 0,789351852$
20	Міжгрупова	$D_{міжгруп} = 22,36309524$	2	$SS_{міжгруп} = 11,18154762$
21	Загальна	$D_{заг} = 36,57142857$	20	$SS_{заг} = 1,828571429$
22	Критерій Фішера			
23	$F_{емп} =$	14,16548		
24	$F_{0,01} =$	3,554557		
25	$F_{0,05} =$	6,012905		
26				

Рис. 8.3. Результати дисперсійного аналізу

29. Задачу можна виконати значно швидше, застосувавши пакет «Анализ данных», розділ «Однофакторный дисперсионный анализ». Для цього необхідно внести дані в діалогове вікно так, як показано на рис. 8.4.

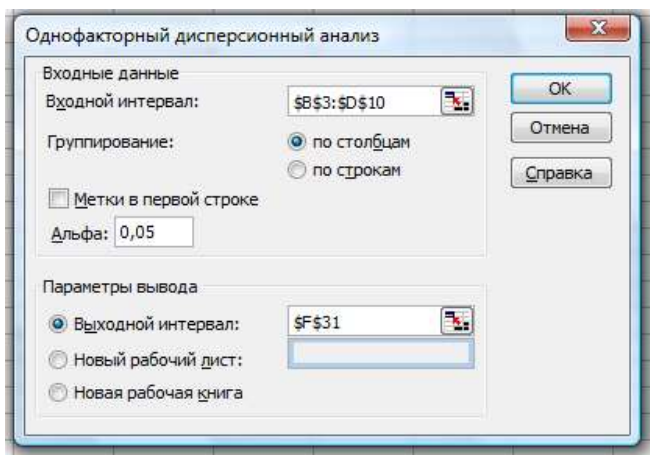


Рис. 8. 4. Ввід даних для проведення дисперсійного аналізу

Отримані результати наведені на рис.8. 5.

Однофакторный дисперсионный анализ						
ИТОГИ						
Группы	Счет	Сумма	Среднее	Дисперсия		
Низька	6	40	6,666667	1,066667		
Середня	7	35	5	0,666667		
Висока	8	33	4,125	0,696429		
Дисперсионный анализ						
Источник вариации	SS	df	MS	F	P-Значение	F критическое
Между группами	22,3631	2	11,18155	14,16548	0,0002017	3,554557146
Внутри групп	14,20833	18	0,789352			

Рис. 8.5. Результати дисперсійного аналізу

### Рекомендації щодо оформлення звіту

Звіт повинен містити:

- титульний аркуш;
- найменування і мету роботи;
- відомості щодо виконання завдання;
- висновки по роботі.

## Індивідуальні завдання № 8

При рівні значущості  $\alpha=0,01$  перевірити гіпотези про те, чи існує вплив частоти центрального процесора (фактор А) на швидкість завантаження програмного додатку.

1.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	9	8	10	12
A <sub>2</sub>	10	12	11	11
A <sub>3</sub>	8	16	10	18
A <sub>4</sub>	9	18	10	8

2.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	60	80	75	85
A <sub>2</sub>	75	66	85	80
A <sub>3</sub>	60	80	65	60
A <sub>4</sub>	95	85	100	80

3.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	25	28	20	22
A <sub>2</sub>	29	22	21	18
A <sub>3</sub>	19	25	30	22
A <sub>4</sub>	18	30	24	20

4.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	6	8	3	2
A <sub>2</sub>	5	4	10	11
A <sub>3</sub>	5	4	13	12
A <sub>4</sub>	18	16	21	20



5.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	28	26	21	25
A <sub>2</sub>	24	28	32	28
A <sub>3</sub>	23	24	20	24
A <sub>4</sub>	29	28	27	28

6.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	90	85	105	110
A <sub>2</sub>	80	110	115	90
A <sub>3</sub>	90	75	120	110
A <sub>4</sub>	75	120	110	85

7.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	10	7	8	6
A <sub>2</sub>	8	14	6	10
A <sub>3</sub>	15	12	11	9
A <sub>4</sub>	12	13	6	8

8.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	1	3	2	4
A <sub>2</sub>	1	2	5	3
A <sub>3</sub>	3	4	6	4
A <sub>4</sub>	2	4	4	5

9.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	2,2	2,6	2,8	3,1
A <sub>2</sub>	3,1	3,3	3,5	3,7
A <sub>3</sub>	2,4	2,6	2,8	2,6
A <sub>4</sub>	3,3	2,7	2,5	3,9

10.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	1,9	1,5	1,7	1,9
A <sub>2</sub>	2,1	2,2	2,5	1,8
A <sub>3</sub>	3	3,1	2,9	2,8
A <sub>4</sub>	3,1	3,3	2,9	2,8

11.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	10	12	15	11
A <sub>2</sub>	9	7	8	6
A <sub>3</sub>	11	9	8	10
A <sub>4</sub>	8	8	9	11

12.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	22	24	26	25
A <sub>2</sub>	27	24	26	23
A <sub>3</sub>	31	30	29	27
A <sub>4</sub>	29	29	25	30

13.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	0,9	0,8	0,5	1,1
A <sub>2</sub>	1,5	1,7	1,2	1,6
A <sub>3</sub>	2,2	1,9	2,1	2,2
A <sub>4</sub>	1,9	1,7	1,4	1,6

14.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	12	15	14	11
A <sub>2</sub>	18	19	21	23
A <sub>3</sub>	19	17	22	19
A <sub>4</sub>	21	23	22	23

15.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	5,5	5,7	6,1	6,5
A <sub>2</sub>	7,4	7,6	7,2	7,1
A <sub>3</sub>	6,3	6,9	6,5	6,6
A <sub>4</sub>	7,9	8,1	7,8	8,1

16.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	5,3	6,7	6,6	7,5
A <sub>2</sub>	6,4	5,6	6,2	7,4
A <sub>3</sub>	6,3	6,8	6,9	6,1
A <sub>4</sub>	7,7	7,1	7,5	7,1

17.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	3,9	3,8	3,5	4,1
A <sub>2</sub>	5,5	5,7	5,2	5,6
A <sub>3</sub>	6,2	6,9	6,1	6,2
A <sub>4</sub>	5,9	5,7	5,4	5,6

18.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	2	5	4	1
A <sub>2</sub>	8	9	11	13
A <sub>3</sub>	9	7	12	9
A <sub>4</sub>	11	13	12	13

19.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	1,5	1,7	1,1	1,5
A <sub>2</sub>	3,4	3,6	3,2	3,1
A <sub>3</sub>	5,3	5,9	5,5	5,6
A <sub>4</sub>	7,9	7,1	7,8	8,1

20.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	3	5	6	5
A <sub>2</sub>	5	6	2	4
A <sub>3</sub>	6	8	9	7
A <sub>4</sub>	7	8	5	6

21.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	7,9	7,8	7,5	7,1
A <sub>2</sub>	5,5	5,7	5,2	5,6
A <sub>3</sub>	3,2	3,9	3,1	3,2
A <sub>4</sub>	1,9	1,7	1,4	1,6

22.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	12	15	14	11
A <sub>2</sub>	16	17	19	21
A <sub>3</sub>	19	17	22	19
A <sub>4</sub>	22	24	23	24

23.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	5,5	5,7	6,1	6,5
A <sub>2</sub>	7,0	7,2	6,8	6,6
A <sub>3</sub>	5,3	5,9	5,5	5,6
A <sub>4</sub>	7,9	8,1	7,8	8,1

24.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	5,3	6,7	6,6	7,5
A <sub>2</sub>	7,4	6,6	7,2	6,4
A <sub>3</sub>	8,3	8,8	8,9	8,1
A <sub>4</sub>	7,7	7,1	7,5	7,1

25.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	3,9	3,8	3,5	3,1
A <sub>2</sub>	4,5	4,7	4,2	4,6
A <sub>3</sub>	3,2	3,9	3,1	3,2
A <sub>4</sub>	1,9	1,7	1,4	1,6

26.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	12	15	14	11
A <sub>2</sub>	13	14	16	18
A <sub>3</sub>	18	16	21	18
A <sub>4</sub>	21	23	22	23

27.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	5,1	5,2	6,0	6,1
A <sub>2</sub>	7,1	7,2	7,0	7,0
A <sub>3</sub>	6,1	6,2	6,0	6,3
A <sub>4</sub>	7,1	8,1	7,0	8,1

28.

Рівень фактора	Спостережувані значення			
A <sub>1</sub>	5,3	6,7	6,6	7,5
A <sub>2</sub>	6,0	5,0	6,1	7,0
A <sub>3</sub>	6,1	6,2	6,3	6,4
A <sub>4</sub>	7,7	7,6	7,5	7,4