

Урок 7. Обчислення основних статистичних характеристик вибірки

Вивчення нового матеріалу

Слайд № 1

В курсі алгебри в 9 класі ви вже ознайомилися з основними статистичними характеристиками вибірки.

Математична статистика — розділ математики, в якому за допомогою математичних методів опрацьовуються та систематизуються кількісні результати досліджень для отримання наукових і практичних висновків.

У математичній статистиці вивчають числові характеристики вибірок, які узагальнюють об'єкти вибірки за певною ознакою.

Наприклад, в оцінюванні PISA значеннями варіант є суми балів, набраних кожним учасником тестування за 1000-бальною шкалою, а узагальненою характеристикою для кожної країни-учасниці є середнє арифметичне значення сум балів учасників.



Найважливіше, що статистичні характеристики вибірки дають змогу робити висновки щодо характеристик всієї генеральної сукупності!

Слайд № 2

На заводі протягом кожної години семигодинного робочого дня робітник виготовляв 10, 8, 11, 12, 11, 9, 7 деталей.

Знайти моду, медіану, середнє значення.

Мода — це елемент вибірки, що зустрічається найчастіше.

В нашому ряді мода — це число 11 (число 11 зустрічається двічі). **Мода = 11**

Медіана — число, що є серединою варіаційного ряду, тобто набору чисел, записаних в порядку зростання. Розташуємо наш варіаційний ряд в порядку зростання: 7, 8, 9, 10, 11, 12. Середина ряду — число 10. **Медіана = 10**

Середнє значення — середнє арифметичне усіх варіант вибірки: $(10+8+11+12+11+9+7)/7=9,7$.

Середнє значення = 9,7

Увага!

Якщо кількість членів ряду парна, то медіана дорівнює середньому арифметичному двох членів, що містяться посередині ряду. Наприклад, медіаною ряду **2, 3, 3, 5, 7, 10** буде **4**.



Розглянемо приклад.

Слайд № 3

Дисперсією вибірки (від лат. dispersion – «розкид») є величина, що обчислюється за формулою

$$D = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2, \text{ де } n - \text{об'єм вибірки, } (i = 1 \dots n) - \text{варіанти, } \bar{x} - \text{вибіркове середнє.}$$

Стандартне відхилення – це параметр, який характеризує ступінь розкиду елементів вибірки відносно середнього значення: $\sigma = \sqrt{D}$

Асиметрія – величина, яка характеризує ступінь несиметричності розподілу відносно його середнього значення.

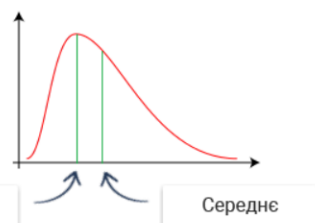
- ♦ Додатна асиметрія вказує на скошеність розподілу в бік малих значень.
- ♦ Від'ємна асиметрія вказує на скошеність розподілу в бік великих значень.
- ♦ Якщо асиметрія дорівнює 0, то варіанти розподілені симетрично.



Розглянемо ще деякі нові статистичні характеристики.

Слайд № 4

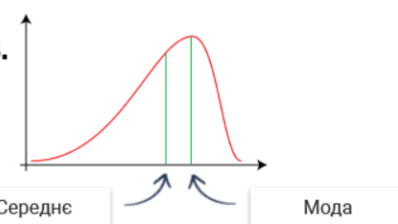
Вибірка А.



Розподіл скошено в бік малих значень – **асиметрія додатна.**

Мода менша за середнє значення.

Вибірка В.



Розподіл скошено в бік великих значень – **асиметрія від'ємна.**

Мода більша за середнє значення.

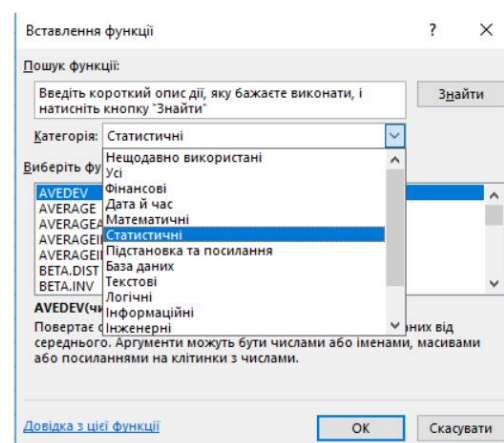


Зауважте, що у вибірці А великий «правий хвіст» тягне середнє вправо, а у вибірці В – навпаки.

Слайд № 5

Розв'язування завдань статистичного аналізу даних значно спрощується завдяки статистичним функціям, реалізованим в табличному процесорі.

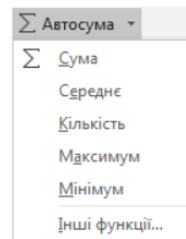
Табличний процесор надає ряд вбудованих функцій для обчислення основних статистичних характеристик. Вони належать до категорії **Статистичні.**



Електронні таблиці – це чудовий засіб статистичного дослідження!

Слайд № 6

Найбільш уживані функції (сума, середнє, максимальне, мінімальне значення) можна знайти в меню кнопки **Автосума** на вкладці **Основне**.



Аргументами статистичних функцій є адреси діапазонів або набори чисел.



Якщо аргументом статистичної функції буде текст, виникне помилка.

Вправа

Вправа № 1 Крок № 1



Вправа 1. Обчислити статистичні характеристики (медіана, мода, середнє, розмах, дисперсія, ст. відхилення, асиметрія) для ряду **2, 3, 3, 3, 5, 5, 7, 7, 7, 10, 10**

Крок 1. Запустіть табличний процесор і створіть таблицю за зразком.

Зауважте, що між модою і середнім один рядок відступу зроблено спеціально; він нам знадобиться згодом.

Ряд (x _i)	Статистичні характеристики
2	Медіана
3	Мода
3	
3	Середнє
5	Розмах
5	Дисперсія
7	Ст. відхилення
7	Асиметрія
7	
10	
10	



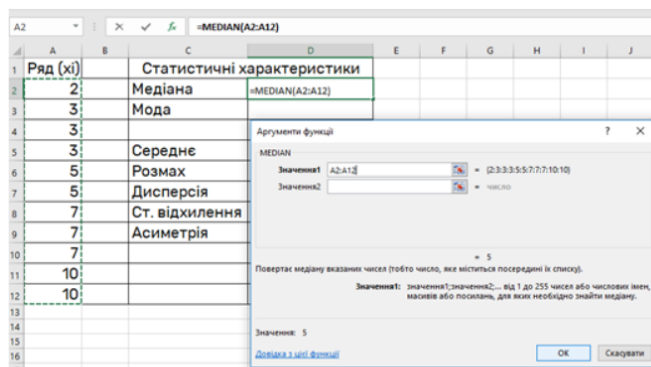
Я покажу, як розв'язувати цю задачу крок за кроком.

Вправа № 1
Крок № 2



Крок 2. Обчисліть медіану.

- виділіть клітинку **D2**
- натисніть кнопку f_x
- знайдіть категорію **Статистичні**
- знайдіть функцію **MEDIAN**
- виділіть діапазон **A2:A12**
- натисніть **OK**



Введіть формулу **=MEDIAN(A2:A12)** в клітинку **D2**.

Вправа № 1
Крок № 3



Крок 3. Аналогічним способом визначте і всі інші статистичні характеристики за допомогою функцій табличного процесора.

	A	B	C	D
1	Ряд (xi)		Статистичні характеристики	
2	2		Медіана	=MEDIAN(A2:A12)
3	3		Мода	=MODE.MULT(A2:A12)
4	3			
5	3		Середнє	=AVERAGE(A2:A12)
6	5		Розмах	=MAX(A2:A12)-MIN(A2:A12)
7	5		Дисперсія	=VAR.S(A2:A12)
8	7		Ст. відхилення	=STDEV.S(A2:A12)
9	7		Асиметрія	=SKEW(A2:A12)
10	7			
11	10			
12	10			



Скористайтеся зразком формул у клітинках діапазону **D2:D10**.

Слайд № 7

Отже, ми отримали такі результати:

	A	B	C	D
1	Ряд (x_i)		Статистичні характеристики	
2	2		Медіана	5,0
3	3		Мода	3,0
4	3			
5	3		Середнє	5,6
6	5		Розмах	8,0
7	5		Дисперсія	7,9
8	7		Ст. відхилення	2,8
9	7		Асиметрія	0,4
10	7			
11	10			
12	10			

Але якщо уважно подивитися на ряд, то можна помітити, що найбільшу кількість разів (тричі) повторюються числа **3** та **7**, тобто обидва ці числа є **модами** даного ряду.



А в нашій таблиці відображується тільки одне число 3.

Вправа

Вправа № 1
Крок № 4

Крок 4. Виправите електронну таблицю так, щоб було обчислено обидва значення моди.

	A	B	C	D
1	Ряд (x_i)		Статистичні характеристики	
2	2		Медіана	5,0
3	3		Мода	3,0
4	3			7,0
5	3		Середнє	5,6
6	5		Розмах	8,0
7	5		Дисперсія	7,9
8	7		Ст. відхилення	2,8
9	7		Асиметрія	0,4
10	7			
11	10			
12	10			



Не забудьте перед введенням формули виділити діапазон клітинок **D3:D4**.

Вправа № 2



Вправа 2. У магазині провели рекламну кампанію.

Зробіть висновок щодо її ефективності на основі даних про обсяги продажу товару кожного дня протягом двох тижнів до кампанії і двох тижнів після неї.

До кампанії	60	51	61	52	56	53	68	57	60	70	72	68	70	78
Після кампанії	58	56	64	57	60	62	70	64	75	82	66	64	76	73



Завантажте з електронного підручника файл-заготовку [рекламна кампанія.xlsx](#).

Вивчення нового матеріалу

Слайд № 8 Висновки із вправи № 2

Отже, порівняння статистичних характеристик свідчить про таке:

Статистичні характеристики		
	До кампанії	Після кампанії
Середнє:	62,57	66,21
Медіана:	60,50	64,00
Ст. відхилення:	8,43	7,90
Асиметрія:	0,23	0,58

- Середнє значення і медіана вибірки після кампанії перевищують відповідні значення до кампанії, а отже, кампанія дала позитивний ефект.
- Значення стандартного відхилення зменшилось, тому обсяги продажів після кампанії стали більш однорідними.
- Вдвічі збільшилося значення асиметрії, що свідчить про те, що найбільші значення продажу більш суттєво відрізняються від середнього значення, ніж найменші значення, у порівнянні з тим, як це було до рекламної кампанії.



Результати статистичного аналізу дають змогу зробити висновок про позитивний вплив рекламної кампанії на обсяги продажів товару.