

Алгебра та початки аналізу

Тема: Елементи математичної статистики

Урок 1

Тема уроку : Поняття про статистику. Генеральна сукупність і вибірка .  
Вибіркові характеристики



**Статистика** – це наука, яка займається збиранням, обробкою і вивченням різних даних, пов'язаних з масовими явищами, процесами та подіями.

Предметом вивчення статистики є вивчення кількісної сторони явищ.

Статистика вчить, як проаналізувати інформацію, виявити та оцінити закономірності формування, розвитку та взаємодії складових за своєю природою соціально-економічних явищ. Економічна статистика вивчає зміну цін, попиту та пропозиції на товари, прогнозує зростання та падіння виробництва й споживання. Медична статистика вивчає ефективність різних ліків і методів лікування, ймовірність виникнення деякого захворювання, прогнозує поширення епідемій . А є ще статистика демографічна, фінансова, податкова, біологічна . . .

### Що вивчає математична статистика?



**Математична статистика** – це розділ

математики, присвячений математичним методам систематизації, збору, обробки статистичних даних та їх використання для наукових і практичних висновків.

Дослідження методами математичної статистики застосовуються для прийняття рішень, прогнозування розвитку певних галузей господарства, плануванні й організації виробництва, контролі якості продукції тощо.

Серед основних задач математичної статистики можна відзначити такі :

- 1) Оцінка ймовірності.
- 2) Оцінка закону розподілу.
- 3) Оцінка числових характеристик випадкової величини.
- 4) Перевірка статистичних гіпотез .

Першим етапом будь-якого дослідження є збирання інформації, а саме, статистичне спостереження.



**Статистичне спостереження** – це спланований, науково організований збір масових даних про соціально-економічні явища та процеси. Приклади статистичних спостережень: *перепис населення; телефонне опитування та інші*. Спостереження є **суцільним**, якщо обстежують ознаки всіх одиниць сукупності. Прикладом може бути медичне обстеження населення у зв'язку з епідемією.

Спостереження є **несуцільним**, якщо обстежуються ознаки окремих одиниць сукупності. Найбільш поширеним видом несуцільного спостереження є **вибіркове спостереження**. Його застосовують тоді, коли в сукупність

входить дуже велика кількість об'єктів або спостереження пов'язане із руйнуванням об'єктів, або ж воно вимагає великих затрат. У таких випадках зі всієї сукупності вибирають обмежену кількість об'єктів і вивчають їх.

Як і в кожній науці, у статистиці використовуються свої специфічні терміни й поняття .



**Генеральна сукупність** – множина всіх можливих результатів спостереження .

**Статистична вибірка, статистичний ряд** – множина результатів , які реально одержані в даному спостереженні .

**Варіанта** – одне зі значень елементів вибірки .

**Варіаційний ряд** – упорядкована множина всіх варіант .

Кількість об'єктів сукупності (вибіркової або генеральної) називають **об'ємом сукупності**, відповідно, кількість об'єктів вибірки називають **об'ємом вибірки**. Наприклад, якщо із 800 деталей відібрано для дослідження 80 деталей, то об'єм генеральної сукупності дорівнює 800, а об'єм вибірки  $n = 80$ .

(15%)

Іноді вибірку випадкових величин або всю генеральну сукупність цих величин доводиться характеризувати одним числом. На практиці це необхідно, наприклад, для швидкого порівняння двох або більше сукупностей за загальною ознакою.

Розглянемо на конкретному прикладі **вибіркові характеристики** .

Нехай провели опитування 10 дівчат і 9 хлопців одного класу відносно кількості книг, які вони прочитали за канікули .

Результатом першого етапу статистичного дослідження є неупорядкований набір чисел, записаних дослідником у порядку їх надходження. Одержали такі ряди чисел :

Дівчата : 4,3,5,3,8,3,12,4,5,5.

Хлопці : 5,3,3,4,6,4,4,7,4.

На другому етапі статистичного дослідження, який називають **зведенням**, упорядковують і узагальнюють статистичні дані, групують їх і на цій основі дають узагальнену характеристику сукупності.

У результаті одержимо такі ряди :

Дівчата : 3,3,3,4,4,5,5,5,8,12. (1)

Хлопці : 3,3,4,4,4,4,4,5,6,7. (2)



- **Розмах вибірки ( R )** – це різниця між найбільшим і найменшим значеннями величини у вибірці :  
для ряду(1) **розмах R=12-3 = 9**, а для ряду (2) **розмах R=7-3 = 4**.
- **Мода ( Mo )** – це те значення елемента вибірки, яке зустрічається частіше за інші .  
для ряду(1) дві моди – числа 3 і 5 : **Mo<sub>1</sub>=3** і **Mo<sub>2</sub>=5** ,  
а для ряду (2) : **Mo=4** .

Зазначимо, що моди може й не бути, якщо всі значення ознаки, яка розглядається, зустрічаються однаково часто.

- **Медіана ( Me )** – це так зване середнє значення впорядкованого ряду значень, вона поділяє ряд даних на дві рівні за кількістю елементів частини.

*Якщо кількість чисел у ряді непарна, то медіана – це число, записане посередині.*

У ряді (2) кількість елементів  $n = 9$ , тому **Me=4** :

3 , 3 , 4 , 4 , **4** , 4 , 5 , 6 , 7

*Якщо кількість чисел у ряді парна, то медіана - це середнє арифметичне двох чисел, що стоять посередині .*

У ряді (1) кількість елементів  $n=10$ , тому **Me =(4+5):2= 4,5** :

3 , 3 , 3 , 4 , 4 , **4,5** , 5 , 5 , 5 , 8 , 12

- **Середнє значення вибірки ( $\bar{X}$ )** – середнє арифметичне всіх чисел ряду даних вибірки .

Якщо відомо, що в ряді даних різні значення  $x_1, x_2, \dots, x_k$  зустрічаються відповідно до частот  $m_1, m_2, \dots, m_k$  , то можна обчислити за формулою

$$\bar{X} = \frac{x_1 m_1 + x_2 m_2 + \dots + x_k m_k}{n}$$

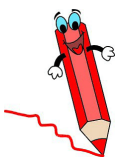
Для ряду (1) :  $\bar{X} = \frac{3 \cdot 3 + 4 \cdot 2 + 5 \cdot 3 + 8 \cdot 1 + 12 \cdot 1}{10} = \frac{52}{10} = 5,2$

Для ряду (2) :  $\bar{Y} = \frac{3 \cdot 2 + 4 \cdot 4 + 5 \cdot 1 + 6 \cdot 1 + 7 \cdot 1}{9} = \frac{40}{9} \approx 4,4$

Оскільки  $\bar{X} > \bar{Y}$  , то можна сказати, що за один і той самий проміжок часу дівчата в класі читають книг більше , ніж хлопці .

Зауважимо, що моду, медіану й середнє значення вибірки об'єднують одним терміном – **міри центральної тенденції** .

(15%)



**Вправи для самоконтролю**

1. Знайдіть середнє значення , моду і медіану вибірки :  
1,1,2,3,3,4,4,5,5,5,6.

*Відповідь* :  $M_o = 5$ ,  $M_e = 4$ ,  $\bar{X} = \frac{39}{11}$ .

(20%)

2. Знайдіть середнє значення , моду і медіану вибірки :  
2; 3; 1; 5; 4; 4; 2;1;2.

*Відповідь* :  $M_o = 2$ ,  $M_e = 2$ ,  $\bar{X} = \frac{24}{9}$ .

(20%)

3. Дано вибірка деякої випадкової величини X: 1; 2; 8; 4; 2; 3; 5; 4; 2; 3.  
Знайти середнє значення, моду й медіану, у відповідь записати суму цих величин.

*Відповідь*: 8,4 ( 3,4+2+3=8,4 )

(30%)