

Тема: Статистика та її методи

Посилання

на

підручник:

<https://lib.imzo.gov.ua/wa-data/public/site/books2/pidruchnyky-11-klas-2019/13-matematyka-11-klas/merzlyak-ag-matematyka-algebra-i-poch-analizu-ta-geometriya-riven-standartu-11-kl.pdf>

Матеріали до теми:

Математична статистика вивчає методи збирання, обробки та інтерпретації (*інтерпретація* – роз'яснення, тлумачення наукових і літературних текстів) різноманітних даних.

- ☐ Розглянемо на прикладі *основні елементи математичної статистики*

- **Вибірка**

№1

Опитавши 19 дітей, що прийшли на підготовку до ЗНО, скільки часу вони витратили на шлях до аудиторії, отримали наступні результати: 5, 15, 40, 70, 10, 30, 60, 15, 40, 90, 40, 30, 10, 40, 60, 30, 60, 30, і 15 хв

- ☐ Отримавши відповіді на запитання маємо ряд чисел, такий ряд чисел називається *вибіркою*
- ☐ Якщо ми впорядкуємо отриманий ряд чисел за зростанням – отримаємо *варіаційний ряд* (або ранжований ряд, відповідно операцію розташування випадкових величин називають ранжуванням)
- ☐ Кількість елементів, з яких складається вибірка називається *об'ємом вибірки*
- ☐ Кожен елемент вибірки називається її *варіантою*
- ☐ Чи завжди можна дослідити множину всіх об'єктів, що підлягають дослідженню?
(Учні висловлюють власну думку)

Генеральна сукупність – це сукупність усіх об'єктів, що підлягають дослідженню.

- Наведіть приклад генеральної сукупності та вибірки з неї
(Учні наводять власні приклади. Бувають випадки коли можна дослідити кожен елемент генеральної вибірки, наприклад, проведення зовнішнього незалежного оцінювання з української мови (і математики, якщо на цей момент ЗНО з математики вже є обов'язковим))

• Характеристики варіаційних рядів

1. **Середнє значення вибірки** $\bar{x}$ – це середнє арифметичне усіх її варіант:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Середнє значення для нашої вибірки:

$$\bar{x} = \frac{5+2 \cdot 10+3 \cdot 15+4 \cdot 30+4 \cdot 40+3 \cdot 60+70+90}{19} \approx 36,3$$

2. **Медіана вибірки** Me



- Спробуйте самостійно сформулювати означення медіани вибірки.
(Серединне значення варіаційного ряду)

Медіана вибірки Me – це число, що ділить відповідний варіаційний ряд навпіл

- Як знайти медіану вибірки, якщо варіаційний ряд складається з парної кількості елементів?

{3, 15, 27, **42, 44**, 44, 87, 88}

В такому випадку, медіаною вибірки вважається середнє арифметичне двох чисел, що розташовані посередині цього варіаційного ряду:

$$Me = \frac{42+44}{2} = 43$$

- Як знайти медіану вибірки, якщо елементами варіаційного ряду є не числа і маємо парну кількість елементів ряду?
(В такому випадку медіаною вибірки називається будь-яке з двох елементів, що розташовані посередині. Наприклад, маємо ряд оцінок за шкалою ECTS: F, E, E, E, E, C, C, B, B, A, отже медіаною цієї вибірки є оцінки E і C)

3. **Мода вибірки** Mo

- Спробуйте самостійно сформулювати означення моди вибірки
(Варіанта вибірки з найбільшою частотою)
- Як можна зручніше подати інформацію варіаційного ряду?
(Можна скласти **статистичну таблицю** із відповідними варіантами варіаційного ряду та їх частотою у цьому ряді)

5	10	15	30	40	60	70	90
1	2	3	4	4	3	1	1

- ☐ Що є модою цього варіаційного ряду?

$$M_{O_1} = 30 \text{ хв}$$

$$M_{O_2} = 40 \text{ хв}$$

4. Розмах вибірки R – це різниця між найбільшим і найменшим значенням вибірки.

- ☐ Обчисліть розмах цієї вибірки

5	10	15	30	40	60	70	90
1	2	3	4	4	3	1	1

$$R = 90 - 5 = 85$$

- ☐ Досліджуючи популярність артиста – опитали всіх людей, що прийшли до нього на концерт. Чи будуть отримані висновки об'єктивними?

(Такі висновки не будуть об'єктивними, так як люди, що прийшли на концерт – прийшли на нього, бо їм подобається цей артист)

- ☐ Отже, нам потрібна *репрезентативна* вибірка, так як збирання даних має ґрунтуватися не тільки на масовості.

Вибірка називається *репрезентативною*, якщо в ній всі основні ознаки *генеральної сукупності* присутні в тій самій пропорції і з тією самою відносною частотою, з якою дана ознака виступає в заданій генеральній сукупності.

• Статистична та математична ймовірності

- ☐ Якою є математична ймовірність народження хлопчика?

$$P = \frac{1}{2} = 0,5$$

- ☐ Якщо поглянемо на таблицю народжуваності в Києві за 2005 рік, то можемо помітити, що статистична ймовірність відрізняється від математичної. Статистична ймовірність народжуваності хлопчиків трішки більша ніж дівчаток і мало відрізняється від відомого числа в *демографії* (наука, що вивчає склад і рух населення та закономірності його розвитку) 0,518.

Місяць	Загальна кількість	Хлопчики	Частота народжуваності	Дівчатка	Частота народжуваності
1	1396	711	0,509312	685	0,490688
2	1254	659	0,525518	595	0,474482

3	1361	702	0,515797	659	0,484203
4	1301	664	0,510377	637	0,489623
5	1440	758	0,526389	682	0,473611
6	1462	762	0,5121204	700	0,478796
7	1587	811	0,511027	776	0,488973
8	1617	849	0,525046	768	0,474954
9	1321	699	0,529145	622	0,470855
10	1528	771	0,504581	757	0,495419
11	1465	748	0,51058	717	0,48942
12	1332	698	0,524024	634	0,475976
Усього	17064	8832	0,517581	8232	0,482419

• Способи подання даних

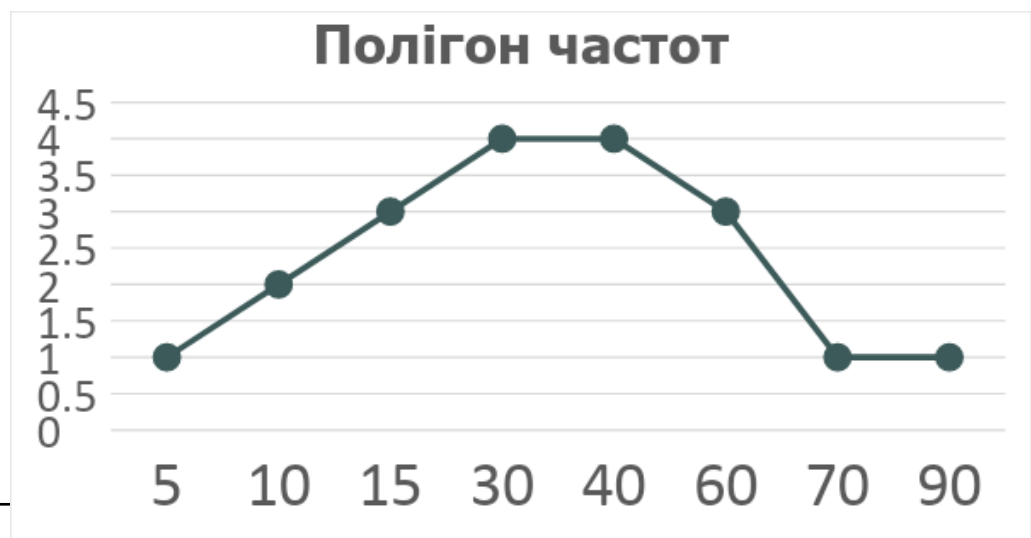
- Для кращого сприйняття інформації про ту чи іншу вибірку – отримані дані можна подати графічно.

Розглянемо на прикладі нашої задачі найпоширеніші графічні способи подачі інформації:

1. Стовпчаста діаграма



2. Полігон частот



3. Секторна діаграма



● Розв'язування задач

№1 (Середнє арифметичне і медіана)

Нехай заробітна плата в області знаходиться від 3 000 – 250 000 грн (3k; 3.2k; 4k; 7,5k; 10k; 3k; 3,3k; 250 000k; 4.5k). Знайдіть середню заробітну плату.

Можемо знайти **середнє значення вибірки**, тоді середня заробітна плата пересічного жителя області приблизно 32 000 грн:

$$\bar{x} = \frac{2 \cdot 3k + 3,2k + 4k + 7,5k + 10k + 3,3k + 250k + 4,5k}{9} \approx 32k$$

Можемо знайти **медіану варіаційного ряду**, тоді середня заробітна плата пересічного жителя області приблизно 4 000 грн:

{3k; 3k; 3.2k; 3,3k; **4k**; 4.5k; 7,5k; 10k; 250k; }

- ☐ Що краще відображає середню заробітну плату пересічного жителя області – середнє значення чи медіана варіаційного ряду?

№2

На новорічні свята ви купили багато цукерок: 2 кг по 980 грн, 4 кг по 450 грн і 8 кг по 350 грн. За якою середньою ціною Ви купили всі цукерки?

Знайдемо **середнє значення вибірки**:

$$\bar{x} = \frac{2 \cdot 980 + 4 \cdot 450 + 8 \cdot 350}{14} \approx 469 \text{ грн}$$

№1

Результатом роботи комп'ютерної програми, що моделює статистичне дослідження, є деяке ціле число в діапазоні від -128 до 128 . Після п'яти послідовних запусків програма видала такі результати: $62, -15, 31, 103, -22$. Що в даному статистичному дослідженні є генеральною сукупністю? Що є вибіркою? Знайдіть розмах вибірки.

Розв'язок:

Генеральна сукупність:

Усі цілі числа від -128 до 128

Вибірка:

$\{62, -15, 31, 103, -22\}$

Розмах вибірки:

$103 - (-22) = 125$

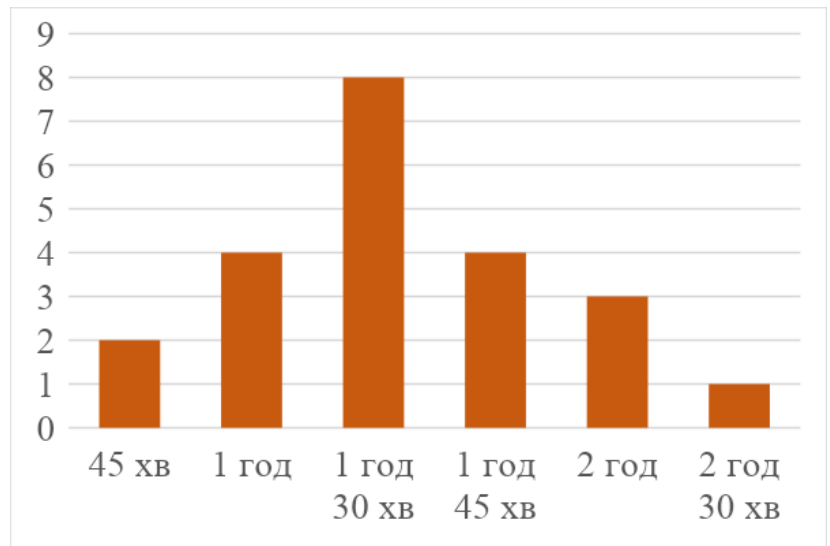
№2

На замовлення підприємств легкої промисловості проведено дослідження, результатами якого є розміри одягу в міжнародному форматі (символи: XS, S, M, L, XL, XXL, XXXL). Які статистичні показники (розмах, середнє значення, медіана, мода) можна визначити для зібраних даних?

Відповідь: медіану і моду;

№3

Серед учнів і учениць 10 класу провели опитування: скільки часу вони щодня перебувають на свіжому повітрі. Результати опитування подано у вигляді діаграми, зображеної на рисунку. Знайдіть розмах, середнє значення та моду даної вибірки.



Розв'язок:

Розмах:

$R = 150 - 45 = 105$

Середнє значення:

$\bar{x} = \frac{2 \cdot 45 + 4 \cdot 60 + 8 \cdot 90 + 4 \cdot 105 + 3 \cdot 120 + 1 \cdot 150}{22} = \frac{1980}{22} = 90 \text{ хв}$

Мода:
 $Mo = 1 \text{ год } 30 \text{ хв}$

№4

Визначте середнє значення та медіану вибірки 1, 4, 7, 3, 9, 6, 7, 2, 4, 9

Розв'язок:

$$\bar{x} = \frac{1+3+2\cdot4+6+2\cdot7+2\cdot9}{10} = \frac{50}{10} = 5$$

{1; 3; 4; 4; 6; 7; 7; 9; 9}

$Me = 6$

Відповідь: $\bar{x} = 5$; $Me = 6$

№5

У чемпіонаті України з футболу 2017-2018 рр. команда «Шахтар», що стала чемпіоном України, зіграла 32 матчі, у яких двічі забила 5 голів, 3 рази – 4 голи, 9 разів – 3 голи, 8 разів – 2 голи, 6 разів – один гол і в 4 матчах не забила жодного гола. Обчисліть середню кількість м'ячів, яку команда «Шахтар» забивала в одному матчі.

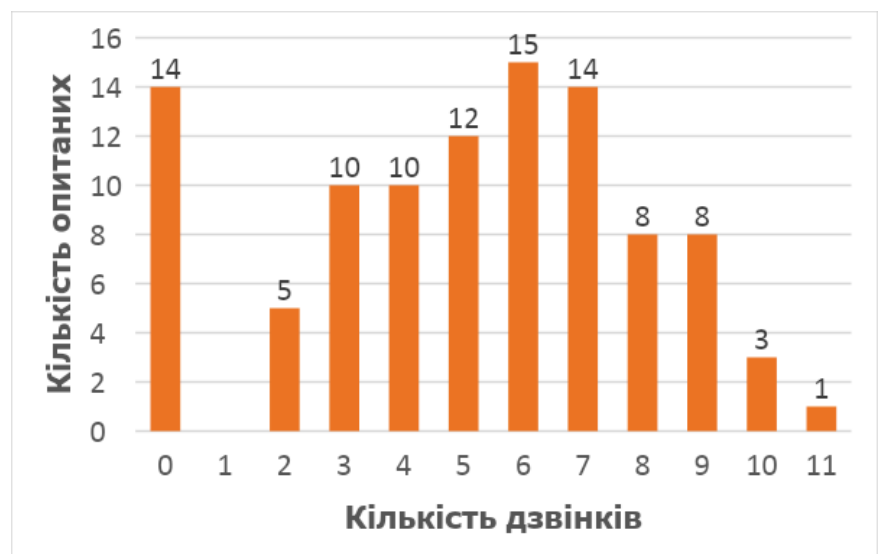
Розв'язок:

$$\bar{x} = \frac{2\cdot5+3\cdot4+9\cdot3+8\cdot2+6\cdot1+4\cdot0}{32} = \frac{71}{32} \approx 2,2$$

Відповідь: $\approx 2,2$ м'яча за гру

№6

Телефонна компанія хоче дізнатися про кількість телефонних дзвінків, які робить людина протягом доби. Дані щодо 100 людей подано на діаграмі. Обчисліть розмах, середнє значення, медіану та моду цієї вибірки.



Розв'язок:

$$R = 11 - 0 = 11$$

$$\bar{x} = \frac{14 \cdot 0 + 0 \cdot 1 + 5 \cdot 2 + 10 \cdot 3 + 10 \cdot 4 + 12 \cdot 5 + 15 \cdot 6 + 14 \cdot 7 + 8 \cdot 8 + 8 \cdot 9 + 3 \cdot 10 + 1 \cdot 11}{100} =$$

$$= \frac{10 + 30 + 40 + 60 + 90 + 98 + 64 + 72 + 30 + 11}{100} = \frac{505}{100} = 5,05$$

$$Mo = 6$$

$$Me = 5$$

Відповідь: $R = 11$; $Mo = 6$; $Me = 5$;

Завдання:

1. Опрацювати теоретичний матеріал §3, п.15.
2. Законспектувати означення.
3. Виконати усно вправи: 15.1-15.4; письмово вправи: 15.5, 15.7, 15.8, 15.9.

ЗВЕРНІТЬ УВАГУ!!! Роботу виконувати у робочому або окремому зошиті (якщо робочий залишився у гуртожитку), фотографувати і надсилати на електронну адресу valentinatalavera@ukr.net, у темі листа вказувати – ПІБ, предмет, номер групи.

Можна підготувати мультимедійну презентацію з теми і надіслати на електронну адресу valentinatalavera@ukr.net.