**ПРАКТИЧНЕ ЗАНЯТТЯ 3**

**Повторення матеріалу**

1. Дайте визначення моди. Наведіть приклади.

2. Дайте визначення медіани. Наведіть приклади.

Як визначається номер члена ряду, який є медіаною, у випадку непарної кількості членів ряду?

Як визначається номер члена ряду, який є медіаною, у випадку парної кількості членів ряду?

3. За якою формулою визначається середнє арифметичне?

4. Визначення дисперсії. Як за відомою дисперсією знайти стандартне відхилення?

**План**

1. Побудова інтервального розподілу в Excel.

2. Побудова гістограми інтервального розподілу.

3. Побудова полігону дискретного варіаційного ряду.

4. Розв’язання задач.

**Теоретичні відомості**

Встановлення статистичних закономірностей, що до масових випадкових явищ, основане на вивченні статистичних даних – відомостей про те, які значення прийняла окрема ознака (випадкова величина Х), у наслідок проведення досліду.

Різні значення ознаки (випадкової величини Х) носять назву **варіант** (позначають х).

 Розгляд та розуміння даних, що були отримані за результатами вибіркового дослідження (особливо при великій кількості спостережень n) провести досить важко, та за ними практично неможливо уявити характер розподілу ознаки. Першим кроком до розуміння статистичного матеріалу, що маємо унаслідок проведення деякого дослідження є упорядковування, розташування варіантів за зростанням (зменшенням), тобто **ранжування** ряду. В такому вигляді вивчати дані також не дуже зручно. Це пов’язано з великою кількістю варіант, тому їх розбивають на окремі інтервали, тобто проводять **групування***.* Кількість інтервалів k треба обирати не дуже великою, бо постає та ж проблема великої кількості варіант, але й не зовсім малим, щоб не втратити особливостей розподілу ознаки. У деякій літературі кількість інтервалів *k* визначається як корінь квадратний з об’єму вибірки , проте на практиці майже завжди використовується формула *Стерджеса*, де ,

 тоді величина інтервалу *l* визначається за формулою:



де xmax – xmin – різниця між найбільшим та найменшим значенням ознаки.

Числа, які показують скільки разів зустрічаються варіанти із даного інтервалу, називають **частотами** (позначають ni), а їх відношення до загальної кількості спостережень – **частостями**, або відносними частотами, тобто

.

Частоти та відносні частоти називають **вагами**.

**Приклад 1.1**

Розглянемо дискретний ряд оцінок з іспиту 3-го курсу:

2, 3, 3, 5, 4, 3, 4, 4, 5, 2, 5, 4, 3, 3, 5, 4, 3, 3, 3, 3

Ранжуємо його за зростанням:

2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5

Оскільки варіант лише 20, (та вони дискретні) розбивати на інтервали не будемо, а побудуємо одразу ж таблицю.

**Таблиця 1.1**

|  |  |  |  |
| --- | --- | --- | --- |
| I | **Значення варіант хі (оцінка)** | **Частота ni** | **Частістьwi=ni/n** |
| 1 | 2 | 2 | 0,1 |
| 2 | 3 | 9 | 0,45 |
| 3 | 4 | 5 | 0,25 |
| 4 | 5 | 4 | 0,2 |
| **Всього** |   | 20 | 1 |

**Статистичним рядом** називають ранжований за зростанням (зменшенням) ряд варіант з вагами, що їм відповідають.

Для задання статистичного ряду достатньо вказати варіанти та частоти.

Статистичний ряд називають **дискретним**, якщо будь-які його варіанти різняться постійною величиною, та – **неперервним** (*інтервальним*), якщо варіанти можуть різнитися між собою, будь-якою маленькою величиною.

Для графічного зображення статистичних рядів найчастіше використовують гістограму, полігон, та кумулятивну криву (кумуляту).

**Приклад побудови варіаційного ряду та гістограми**

В групі з 40 осіб було виконано заміри часу розв’язання тестового завдання. Були отримані результати: 35, 32, 41, 45, 48, 50, 54, 51, 47, 55, 44, 51, 47, 56, 54, 52, 57, 43, 55, 59, 54, 59, 56, 60, 62, 42, 58, 53, 49, 38, 64, 46, 65, 67, 46, 53, 63, 48, 54, 41. Отже, максимальний час склав 67 секунд, мінімальний – 32.

Побудувати варіаційний ряд та гістограму.

1. Побудуємо таблицю розподілу частот.

Визначимо розмах:

R = Xmax – Xmin; R = 67 – 32 = 35.

2. Для вибору кількості інтервалів (k) можна використати формулу Стержеса

k=$1+3,322lgn\_{}$=$1+3,322lg40\_{}=$1+3,32·1,6≈6,3≈7

k≈$\sqrt{n}$=$\sqrt{40}$≈6,3≈7

h=$\frac{R}{k}$=$\frac{35}{7}=5$

3. Ліву межу першого часткового інтервалу вибираємо рівною Хmin = 32, всі наступні межі отримуємо з попередньої, додатком ширини інтервалу групування h=5

4. Підрахунок частот в Excel

**Теоретичні відомості**

**Полігон**, як правило, використовують для зображення дискретного статистичного ряду, він являє собою ламану, для якої кінці відрізків мають координати (xi,ni).

**Побудувати полігон частот**

У групі з 27 осіб провели опитування, одним із запитань якого було: «Як часто Ви відвідували кінотеатр минулого року?». За результатами опитування отримали таку таблицю:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Варіанта xi | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Частота ni | 0 | 5 | 1 | 8 | 3 | 2 | 2 | 6 |

Побудувати за цим розподілом полігон частот та гістограму частот.

**Гістограму** використовують для зображення лише неперервних статистичних рядів (вона являє собою ступінчасту фігуру із прямокутників з основами, що дорівнюють значенням ознаки *ki=xi+1–xi*, та висотами, що дорівнюють частотам (частостям) nі (wi) інтервалів). Якщо з’єднати середини верхніх сторін відрізками, то отримаємо полігон того ж самого розподілу.

**Кумулята**– крива нагромаджених частот (частостей).

**Емпіричною функцією розподілу** Fn(x) називається відносна частота (частість) того, що ознака (випадкова величина Х) прийме значення, менше ніж заданий х, тобто

*Fn(x)=w(X<x)*

Можна помітити, що статистичний ряд – це статистичний аналог розподілу ознаки (випадкової величини *Х*). В цьому сенсі полігон (гістограма) – аналог кривої розподілу, а емпірична функція розподілу - функції розподілу випадкової величини *Х.*

За даними прикладу 1.1 побудуємо полігон (рис. 1.1) та емпіричну функцію розподілу (рис. 1.2).

Розрахуємо значення емпіричної функції розподілу

F(2)=0; F(3)=0,1; F(4)=0,55; F(5)=0,8; F(x>5)=1;

та побудуємо її графік (мал. 1.2)

Рис. 1.1



Функція розподілу



Полігон

**Завдання до практичного заняття 3**

**Задача 1.**

За виконання індивідуального завдання студент може отримати максимально 15 балів. У групі 35 студентів. Нехай вони отримали таку кількість балів:10, 10, 11, 9, 15, 12, 9, 12, 13, 9, 8, 11, 14, 13, 12, 9, 10, 14, 10, 7, 8, 7, 9, 11, 15, 12, 7, 10, 7, 7, 8, 13, 13, 14, 10. Побудувати варіаційний ряд та полігон розподілу.

35, 32, 41, 45, 48, 50, 54, 51, 47, 55, 44, 51, 47, 56, 54, 52, 57, 43, 55, 59, 54, 59, 56, 60, 62, 42, 58, 53, 49, 38, 64, 46, 65, 67, 46, 53, 63, 48, 54, 41