РОЗДІЛ ІІ. МАТЕМАТИЧНА СТАТИСТИКА

ТЕМА 3. СТАТИСТИЧНІ РОЗПОДІЛИ ВИБІРОК ТА ЇХ ЧИСЛОВІ ХАРАКТЕРИСТИКИ

Лекція 17, 18. Основні поняття математичної статистики. Генеральна і вибіркова сукупності.

Основним змістом математичної статистики є систематизація, обробка і використання статистичної інформації для виявлення статистичних закономірностей ознаки або ознак певної сукупності елементів.

Оскільки суцільна обробка всіх елементів сукупності практично неможлива, то, як правило, застосовується вибірковий метод. Отже, розрізняють генеральну і вибіркову сукупності.

Множина Ω однотипних елементів, яким притаманні певні кількісні ознаки (розміри, вага, маса тощо), утворює генеральну сукупність. Кількість усіх елементів генеральної сукупності називають її обсягом і позначають символом N, значення якого здебільшого невідоме.

Кожна непорожня підмножина A множини Ω ($A \subset \Omega$) випадково вибраних елементів із генеральної сукупності називається вибіркою. Кількість усіх елементів вибірки називають її обсягом і позначають символом n. Його значення відоме, причому воно набагато менше за обсяг генеральної сукупності ($n << N$).

Математична статистика розв’язує дві категорії задач:

1) статистичне оцінювання (точкове, інтервальне) параметрів генеральної сукупності;

2) перевірка правдивості статистичних гіпотез про значення параметрів генеральної сукупності або про закон розподілу ознаки генеральної сукупності на підставі обробки результатів вибірки.

Кількісні ознаки елементів генеральної сукупності можуть бути одновимірними і багатовимірними, дискретними і неперервними.

Коли реалізується вибірка, кількісна ознака, наприклад X, набуває конкретних числових значень ($X = x_i$), які називають *варіантою*.

Зростаючий числовий ряд варіант називають *варіаційним*.

Кожна варіанта вибірки може бути спостереженою n_i раз ($n_i \geq 1$), число n_i називають *частотою варіанти* x_i. При цьому

$$n = \sum_{i=1}^{k} n_i \tag{18.1}$$

де k — кількість варіант, що різняться числовим значенням;
n — обсяг вибірки.

Відношення частоти n_i варіанти x_i до обсягу вибірки n називають її *відносною частотою* і позначають через W_i , тобто

$$W_i = \frac{n_i}{n} \tag{18.2}$$

Для кожної вибірки виконується рівність

Якщо досліджується ознака генеральної сукупності X, яка є неперервною, то варіант буде багато. У цьому разі варіаційний ряд — це певна кількість рівних або нерівних частинних інтервалів чи груп варіант зі своїми частотами.

Такі частинні інтервали варіант, які розміщені у зростаючій послідовності, утворюють *інтервальний варіаційний ряд*.

На практиці для зручності, як правило, розглядають інтервальні варіаційні ряди, у котрих інтервали є рівними між собою.

1. Статистичним розподілом вибірки

Перелік варіант варіаційного ряду і відповідних їм частот, або відносних частот, називають *дискретним статистичним розподілом вибірки*.

У табличній формі він має такий вигляд:

$X = x_i$	x_1	x_2	x_3	...	x_k
n_i	n_1	n_2	n_3	...	n_k
W_i	W_1	W_2	W_3	...	W_k

Дискретний статистичний розподіл вибірки можна подати емпіричною функцією $F^*(x)$.

2. Емпірична функція розподілу

Емпірична функція $F^*(x)$ та її властивості. Функція аргументу x, що визначає відносну частоту події $X < x$, тобто

$$F^*(x) = W(X < x) = \frac{n_x}{n} \qquad (19.1)$$

називається *емпіричною*, або *комулятою*. Тут n — обсяг вибірки; n_x — кількість варіант статистичного розподілу вибірки, значення яких менше за фіксовану варіанту x; $F^*(x)$ — називають ще *функцією нагромадження відносних частот*.

Властивості $F^*(x)$:

1) $0 \le F^*(x) \le 1$;
2) $F(x_{\min}) = 0$, де $x_{\min}$ є найменшою варіантою варіаційного ряду;
3) $F(x)\big|_{x > x_{\max}} = 1$, де $x_{\max}$ є найбільшою варіантою варіаційного ряду;
4) $F(x)$ є неспадною функцією аргументу x, а саме: $F(x_2) \ge F(x_1)$ при $x_2 \ge x_1$.

3. Полігон частот і відносних частот.

Дискретний статистичний розподіл вибірки можна зобразити графічно у вигляді ламаної лінії, відрізки якої сполучають координати точок $(x_i; n_i)$, або $(x_i; W_i)$.

У першому випадку ламану лінію називають *полігоном частот*, у другому — *полігоном відносних частот*.

Приклад. За заданим дискретним статистичним розподілом вибірки

$X = x_i$	–6	–4	–2	2	4	6

n_i	5	10	15	20	40	10
W_i	0,05	0,1	0,15	0,2	0,4	0,1

потрібно:

1. Побудувати $F^*(x)$ і зобразити її графічно;
2. Накреслити полігони частот і відносних частот.

Розв'язання. Згідно з означенням та властивостями $F^*(x)$ має такий вигляд:

$$F^*(x) = W(X < x) = \frac{n_x}{n} = \begin{cases} 0 & x \le -6, \\ 0{,}05 & -6 < x \le -4, \\ 0{,}15 & -4 < x \le -2, \\ 0{,}3 & -2 < x \le 2, \\ 0{,}5 & 2 < x \le 4, \\ 0{,}9 & 4 < x \le 6, \\ 1, & x > 6. \end{cases}$$

Графічне зображення $F^*(x)$ подано на рис. 14.

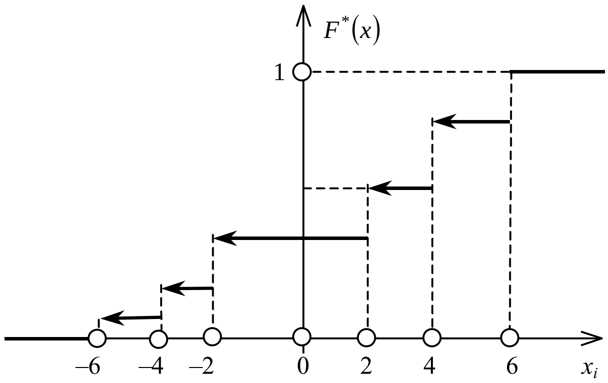

Рис. 14

Полігони частот та відносних частот зображено на рис.15, 16.

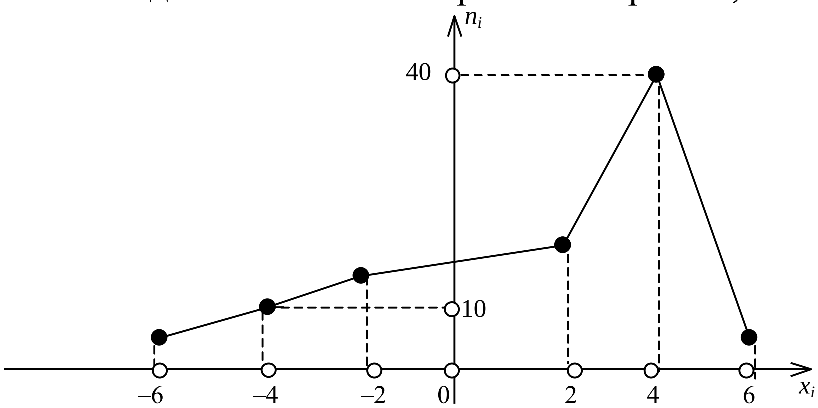

Рис. 15

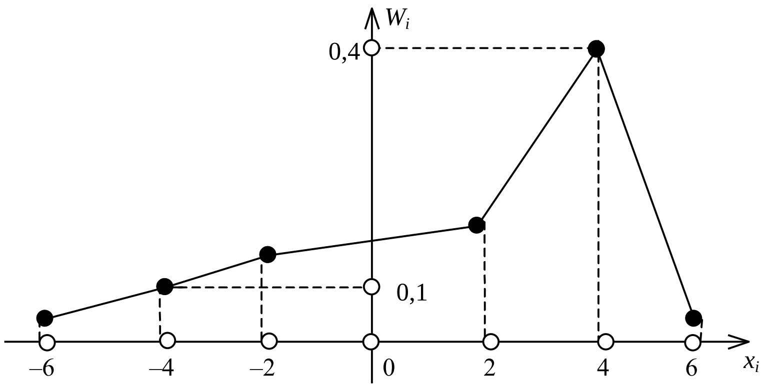

Рис. 16