М.А. МАРТИНЕНКО · О.М. НЕЩАДИМ · В.М. САФОНОВ

ТЕОРІЯ ЙМОВІРНОСТЕЙ І МАТЕМАТИЧНА СТАТИСТИКА

ЧАСТИНА II

МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ХАРЧОВИХ ТЕХНОЛОГІЙ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

М.А. МАРТИНЕНКО
О.М. НЕЩАДИМ
В.М. САФОНОВ

ТЕОРІЯ ЙМОВІРНОСТЕЙ
І
МАТЕМАТИЧНА СТАТИСТИКА

Частина II

Підручник для студентів
вищих навчальних закладів

Рекомендовано Міністерством освіти і науки,
молоді та спорту України

КИЇВ
ЦП«КОМПРИНТ»
2013

УДК 519.22/.25
ББК 22.17я73
М 29

Гриф надано Міністерством освіти і науки, молоді та спорту України
(Лист від 18.01.12 № 1 11-554)

Рецензенти:

В.В. Михайленко, д-р фіз.-мат наук, проф. кафедри вищої математики Національного авіаційного університету.

М.М. Семко, д-р фіз-мат наук, проф., завідувач кафедри вищої математики Національного університету державної податкової служби України.

В.П. Легеза, проф. кафедри вищої математики Національного університету біоресурсів і природокористування України.

Мартиненко М.А., Нещадим О.М., Сафонов В.М.
М 29 Теорія ймовірностей і математична статистика: Підручник. Ч.II — К.:ЦП «КОМПРИНТ», 2013.—278 с.
ISBN 978-617-7144-13-6

Друга частина підручника присвячена основним методам математичної статистики. Розглядаються ймовірнісно-статистичні проблеми дослідження генеральної сукупності: оцінювання параметрів, перевірка статистичних гіпотез, дисперсійного та кореляційно-регресійного аналізу. Теоретичний матеріал подається у формі лекцій і супроводжується прикладами типових задач із поясненнями.

Наприкінці кожної лекції наведено перелік основних питань для самоконтролю знань і достатню кількість завдань для проведення практичних занять. Багато задач подано із детальними розв'язками.

В посібнику вміщено завдання для індивідуальної роботи студентів та модульного контролю.

Розрахований на студентів інженерно-економічних спеціальностей вищих навчальних закладів усіх форм навчання.

УДК 519.22/.25
ББК 22.17я73

ISBN 978-617-7144-13-6

Лекція 13. Числові характеристики статистичного розподілу вибірки.

Середні вибіркові характеристики
Характеристики розсіювання
Емпіричні моменти

Статистичний розподіл містить повну інформацію про варіацію ознаки. Проте при розв'язуванні багатьох задач надмір таких числових даних затруднює їх використання. Проаналізувати одержаний матеріал можна за допомогою деяких сталих, обчислених за наявним статистичним розподілом вибірки. Такі статистичні сталі характеризують основні властивості варіаційного ряду. Для аналізу статистичних розподілів використовуються три види характеристик: 1) середні, або характеристики центральної тенденції; 2) характеристики розсіювання; 3) характеристики, які відображають додаткові особливості розподілу.

13.1 Середні вибіркові характеристики.

Середні величини характеризують значення ознаки, навколо якого концентруються варіанти статистичного розподілу (центральна тенденція розподілу). Існують різні форми середніх: арифметична, гармонічна, квадратична, геометрична тощо. Найбільш поширеною із них є середня арифметична.

Розглянемо дискретний статистичний розподіл вибірки обсягу *n*:

x_i	x_1	x_2	...	x_k
n_i	n_1	n_2	...	n_k

Величина, яка обчислюється за формулою

$$\bar{x}_в = \frac{1}{n}\sum_{i=1}^{k} x_i n_i \qquad (13.1)$$

називається *вибірковою середньою* такого розподілу (середнє арифметичне значення). Очевидно, що $\bar{x}_в = \sum_{i=1}^{k} x_i w_i$, де w_i – відносні частоти.

За вибіркове середнє неперервного статистичного розподілу приймається середнє арифметичне відповідного дискретного розподілу.

Наведемо основні властивості вибіркової середньої.

1^0. Якщо всі варіанти збільшити (зменшити) в одне і те ж число раз λ, то вибіркова середня збільшиться (зменшиться) у стільки ж разів:

$$\frac{1}{n}\sum_{i=1}^{k}(\lambda\, x_i)n_i = \lambda \overline{x}_в \text{ або } \overline{\lambda x_в} = \lambda \overline{x}_в.$$

2^0. Якщо всі варіанти збільшити (зменшити) на одне і те ж число C, то вибіркова середня збільшиться (зменшиться) на таке ж число:

$$\frac{1}{n}\sum_{i=1}^{k}(x_i - C)n_i = \overline{x}_в - C \text{ або } \overline{(x-C)}_в = \overline{x}_в - C.$$

3^0. Сума всіх відхилень значень ознаки від вибіркової середньої дорівнює нулю:

$$\sum_{i=1}^{k}(x_i - \overline{x}_i)n_i = 0 \text{ або } \overline{(x-\overline{x}_в)}_в = 0.$$

4^0. Вибіркова середня алгебраїчної суми кількох ознак дорівнює сумі вибіркових середніх цих ознак:

$$\overline{(x+y)}_в = \overline{x}_в + \overline{y}_в.$$

Обчислення вибіркової середньої безпосередньо за формулою (13.1) часто приводить до громіздких арифметичних розрахунків. Цього можна уникнути, перейшовши від варіант x_i до *умовних варіант*

$$u_i = \frac{x_i - C}{\lambda}, \tag{13.2}$$

де C, λ – відповідно підібрані сталі. Згідно властивостей 1^0–2^0 маємо формулу:

$$\overline{x}_в = \lambda \overline{u}_в + C. \tag{13.3}$$

Іноді при дослідженні сукупності її доводиться розбивати на окремі групи або, навпаки, об’єднувати декілька груп в одну сукупність.

Груповою середньою називають середнє арифметичне значень ознаки, що належать даній групі.

Нехай сукупність обсягу n складається із m груп обсягами N_j кожна, $j=\overline{1,m}$: $n=\sum_{j=1}^{m} N_j$. Якщо $\overline{x}_j$ – групові середні, то загальна вибіркова середня $\overline{x}_в$ обчислюється за формулою

$$\overline{x}_в = \frac{1}{n}\sum_{j=1}^{m}\overline{x}_j N_j. \tag{13.4}$$

Приклад 13.1. За рівнем заробітної плати робітники двох цехів підприємства розподілені так:

Зар. плата, грн.	Кількість робітників	
	Цех №1	Цех №2
250-260	–	2
260-270	3	4
270-280	6	16
280-290	7	8
290-300	4	–
Всього	20	30

Обчислити середню заробітну плату робітників цехів і підприємства.

<u>Розв'язання</u>. Маємо неперервний розподіл ознаки X – рівня заробітної плати робітників підприємства. Переходимо до дискретного розподілу.

$X=x_i$	Кількість робітників	
	№1	№2
255	0	2
265	3	4
275	6	16
285	7	8
295	4	0
Всього	20	30

При обстеженні середньої заробітної плати робітників цеху №1 скористаємось властивістю 1^0:

$$\overline{x}_1 = \frac{5}{N_1}\sum_{i=1}^{4}\frac{x_i}{5}\cdot n_i = \frac{5}{20}\left(53\cdot 3 + 55\cdot 6 + 57\cdot 7 + 59\cdot 4\right) = \frac{1124}{4} = 281$$

Для обчислення середньої заробітної плати робітників цеху №2 зменшимо варіанти на стале число C, наприклад $C=255$:

x_i	255	265	275	285
$x_i - 255$	0	10	20	30
n_i	2	4	16	8

За властивістю 2^0 маємо:

$$\overline{x}_2 = \frac{1}{N_2}\sum_{i=1}^{4}(x_i - 255)n_i + 255 = \frac{1}{30}\left(0\cdot 2 + 10\cdot 4 + 20\cdot 16 + 30\cdot 8\right) + 255 = \frac{600}{30} + 255 = 275$$

За формулою (13.4) обчислюємо середню заробітну плату робітників підприємства:

$$\overline{x}_в = \frac{1}{50}\left(281\cdot 20 + 275\cdot 30\right) = 277{,}4 \text{ (грн.)}$$

Крім вибіркової середньої, для аналізу статистичного розподілу застосовують *структурні* або *порядкові середні*. Із них найбільш поширені мода і медіана.

Означення. *Модою* $\overline{M}_0$ статистичного розподілу називається варіанта, яка має найбільшу частоту появи.

Для дискретного варіаційного ряду мода визначається безпосередньо за означенням. Для інтервального статистичного розподілу знаходиться інтервал $[\alpha_{i-1}, \alpha_i)$, який має найбільшу частоту (*модальний інтервал*). Значення моди обчислюється шляхом лінійної інтерполяції:

$$\overline{M}_0 = \alpha_{i-1} + \frac{n_i - n_{i-1}}{2n_i - n_{i-1} + n_{i+1}}\left(\alpha_i - \alpha_{i-1}\right), \qquad (13.5)$$

де n_i – частота модального інтервала;

n_{i-1} – частота домодального інтервала;

n_{i+1} – частота післямодального інтервала.

Але простіше моду можна знайти графічно за гістограмою частот: визначається абсциса точки перетину діагоналей, які з'єднують вершини прямокутника, що належить модальному інтервалу, із найближчими вершинами двох прилеглих прямокутників (рис.13.1).

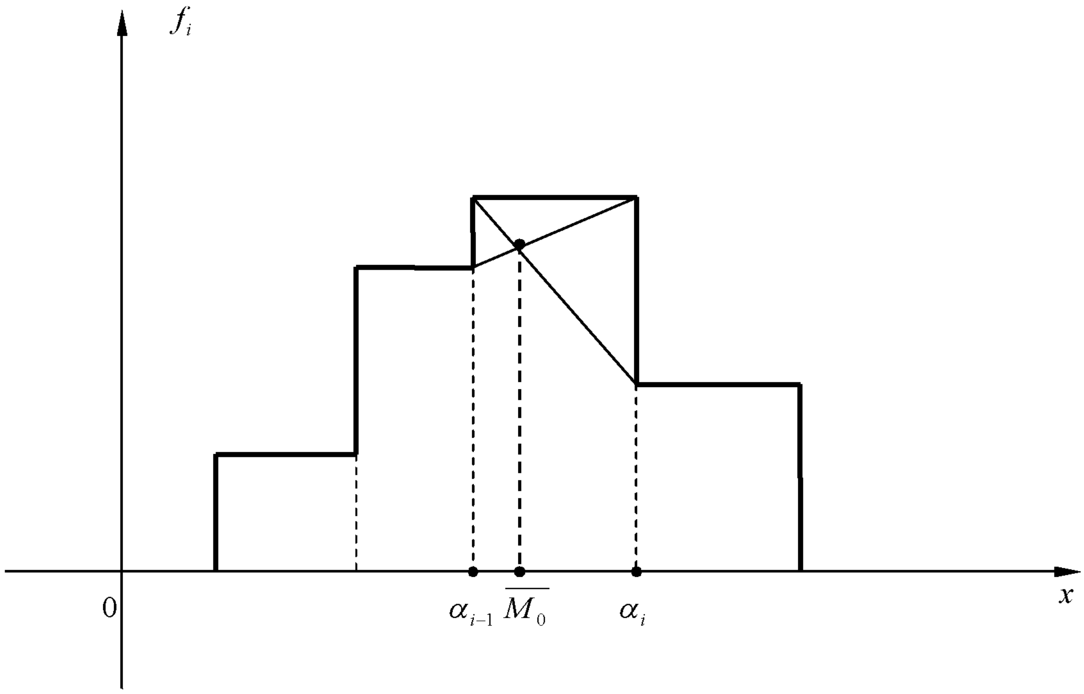

Рис.13.1.

Означення. *Медіаною* $\overline{M_e}$ статистичного розподілу називають варіанту, яка припадає на середину розподілу варіаційного ряду. Для дискретного варіаційного ряду із непарною кількістю елементів медіана дорівнює серединній варіанті, а для ряду із парною кількістю елементів – півсумі двох серединних варіант. Для інтервального статистичного розподілу визначається *медіанний інтервал* $[\alpha_{i-1}, \alpha_i)$ із умови $F_n(\alpha_{i-1}) \leq 0{,}5$ і $F_n(\alpha_i) \geq 0{,}5$. Значення медіани на цьому інтервалі обчислюється шляхом лінійної інтерполяції за формулою:

$$\overline{M}_e = \alpha_{i-1} + \frac{0{,}5 - F_n(\alpha_{i-1})}{w_i}(\alpha_i - \alpha_{i-1}), \qquad (13.6)$$

Медіану наближено можна знайти із графіка емпіричної функції розподілу як аргумент цієї функції, при якому $F_n(x) = 0{,}5$ (рис.13.2)

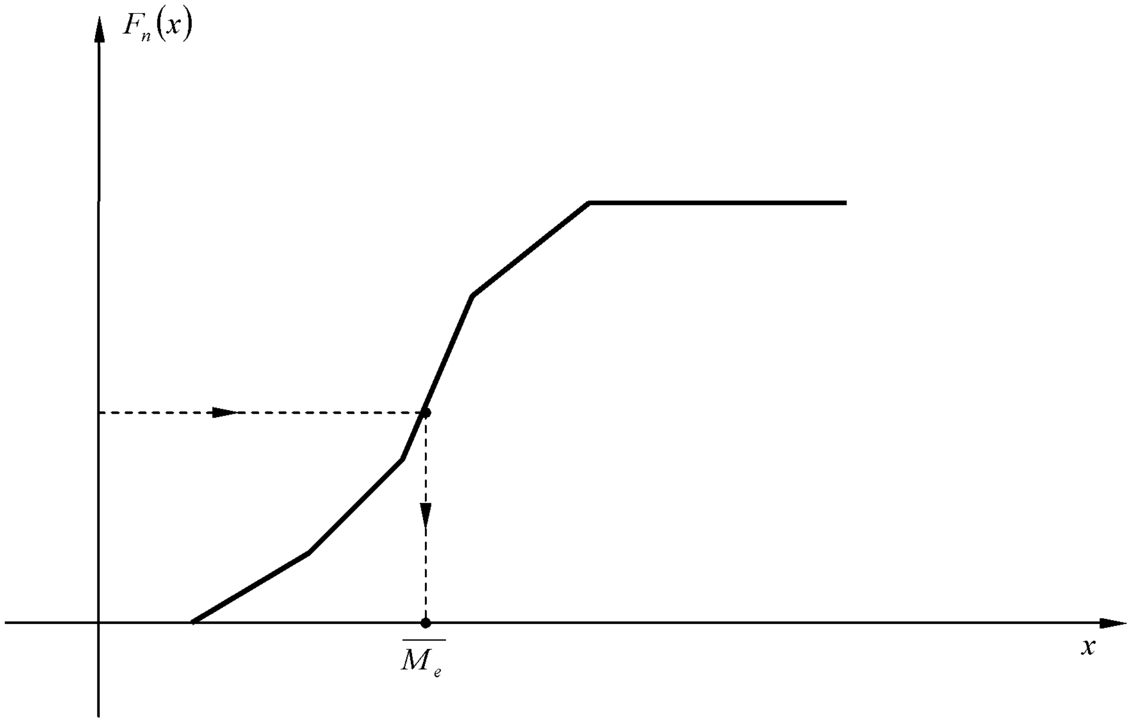

Рис.13.2.

Приклад 13.2. Наведено статистичний розподіл випадкової величини X – кількість балів, одержаних абітурієнтами на вступному іспиті з математики:

X	0-10	10-30	30-50	50-70	70-80	80-100
n_i	15	22	28	18	12	5

Знайти моду і медіану розподілу.

Розв’язання. Маємо інтервальний статистичний розподіл ознаки X ; обсяг вибірки $n = 100$. Видно, що модальним є інтервал $[30;50)$. Оскільки $\alpha_{i-1} = 30$,

$\alpha_i = 50, \quad n_i = 28, \quad n_{i-1} = 22, \quad \alpha_{i+1} = 18,$ то за формулою (13.5) обчислюємо

$$\overline{M}_0 = 30 + \frac{28-22}{2 \cdot 28 - 22 - 18}(50-30) = 30 + \frac{120}{16} = 37{,}5.$$

Будуємо емпіричну функцію розподілу $F_n(x) = W(X < x)$:

$$F_n(x) = \begin{cases} 0, \text{ при } x < 0, \\ 0{,}15, \text{ при } x = 10, \\ 0{,}37, \text{ при } x = 30, \\ 0{,}65, \text{ при } x = 50, \\ 0{,}83, \text{ при } x = 70, \\ 0{,}95, \text{ при } x = 80, \\ 1, \text{ при } x \geq 100; \end{cases}$$

її графік зображено на рис.13.3.

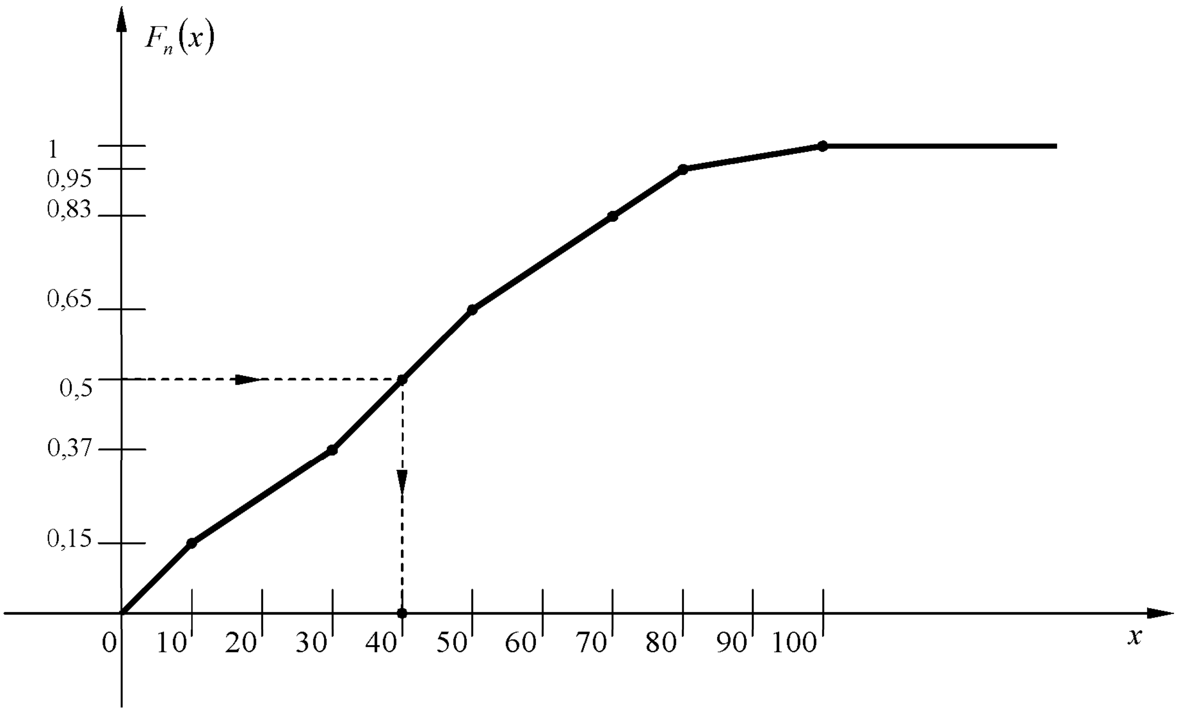

Рис.13.3.

Із графіка визнаємо медіанний інтервал $[30;50)$, якому належить аргумент функції $F_n(x)=0,5$. Враховуючи, що $\alpha_{i-1}=30$, $\alpha_i=50$, $F_n(\alpha_{i-1})=F_n(30)=0,37$, $w_i=\frac{n_i}{n}=0,28$, за формулою (13.6) обчислюємо

$$\overline{M}_e=30+\frac{0,5-0,37}{0,28}(50-30)=30+\frac{2,6}{0,28}\approx 39,3.$$

Отже, маємо значення $\overline{M}_0=37,5$, $\overline{M}_e=39,3$.

13.2Характеристики розсіювання.

Середні величини $\overline{x}_в,\overline{M}_0.\overline{M}_e$ не відображають варіації (розсіювання) значень ознаки. Найпростішою оцінкою розсіювання варіантів статистичного розподілу вибірки є *варіаційний розмах*

$$R=x_{\max}-x_{\min}, \tag{13.7}$$

де $x_{\max}$, $x_{\min}$ – найбільше і найменше значення ознаки. Видно, що цей показник залежить від двох крайніх значень ознаки і не враховує проміжних варіант і частот.

З практичної точки зору найбільш важливою є характеристика розсіювання значень варіаційного ряду навколо вибіркової середньої. Найчастіше це розсіювання оцінюється дисперсією і середнім квадратичним відхиленням.

Означення. *Дисперсією вибірки* $D_в$ називається середнє арифметичне квадратів відхилень варіант відносно $\overline{x}_в$:

$$D_в=\frac{1}{n}\sum_{i=1}^{k}\left(x_i-\overline{x}_в\right)^2 n_i \text{ або } D_в=\sum_{i=1}^{k}\left(x_i-\overline{x}_в\right)^2 w_i \tag{13.8}$$

Наведено основні властивості дисперсії вибірки.

1^0. Дисперсія сталої дорівнює нулю.

2^0. Якщо всі варіанти помножити на стале число λ, то дисперсія вибірки помножиться на число λ^2:

$$\frac{1}{n}\sum_{i=1}^{k}\left(\lambda x_i - \lambda \overline{x}_в\right)^2 n_i = \lambda^2 D_в$$

3^0. Якщо всі варіанти збільшити (зменшити) на одне і те ж число C, то дисперсія вибірки не зміниться:

$$\frac{1}{n}\sum_{i=1}^{k}\left[(x_i + C) - (\overline{x}_в + C)\right]^2 n_i = D_в$$

4^0. Дисперсія вибірки дорівнює різниці середньої арифметичної квадратів варіант і квадрату вибіркової середньої:

$$D_в = \frac{1}{n}\sum_{i=1}^{k} x_i^2 n_i - (\overline{x}_в)^2 \text{ або } D_в = \overline{\left(x^2\right)}_в - (\overline{x}_в)^2 \qquad (13.9)$$

Для спрощення обчислення дисперсії вибірки можна скористатись умовними варіантами (13.2); маємо формулу:

$$D_{вx} = \frac{\lambda^2}{n}\sum_{i=1}^{k} u_i^2 n_i - (\overline{x}_в - C)^2 \text{ або } D_{вx} = \lambda^2 D_{вu} \qquad (13.10)$$

Нехай всю вибірку обсягу *n* розбито на *m* груп. Позначимо через N_j – обсяг *j*-тої групи; $\overline{x}_j$ – групову середню; D_j – *групову дисперсію* $\left(j = \overline{1,m}\right)$; $\overline{x}_в$ – загальну середню; $D_в$ – загальну вибіркову дисперсію. Величина

$$D_m = \frac{1}{n}\sum_{j=1}^{m}\left(\overline{x}_j - \overline{x}_в\right)^2 N_j \qquad (13.11)$$

називається *міжгруповою дисперсією* і характеризує дисперсію групових середніх. Загальна дисперсія вибірки обчислюється за формулою

$$D_в = \frac{1}{n}\sum_{j=1}^{m} D_j N_j + D_m \qquad (13.12)$$

Бажано за оцінку розсіювання мати величину, яка вимірюється в тих же одиницях, що і варіанти ознаки. Такими характеристиками є:

1) *середнє квадратичне відхилення (або стандартне відхилення) вибірки* $\sigma_в = \sqrt{D_в}$;

2) *коефіцієнт варіації* $V = \dfrac{\sigma_в}{\bar{x}_в} \cdot 100\%$

13.3 Емпіричні моменти

Середнє арифметичне і дисперсія вибірки є частинними випадками більш загальних числових характеристик – моментів статистичного розподілу.

Центральним емпіричним моментом ℓ *– го* порядку називають величину v_ℓ, яка обчислюється за формулою:

$$v_\ell * = \frac{1}{n}\sum_{i=1}^{k}\left(x_i - \bar{x}_в\right)^\ell n_i \qquad (13.15)$$

Якщо $\bar{x}_в = 0$, то момент називається *початковим* і позначається через $\mu_\ell *$, тобто

$$\mu_\ell * = \frac{1}{n}\sum_{i=1}^{k} x_i^\ell n_i \qquad (13.16)$$

За допомогою емпіричних моментів розподілу можна описати всі особливості варіації ознаки (середню тенденцію, розсіювання, форму розподілу тощо).

Із формул (13.15), (13.16) видно, що $\mu_1 * = \bar{x}_в, v_1 * = 0, v_2 * = D_в$.

Центральні моменти можна виразити через початкові, наприклад $v_2 * = \mu_2 * -(\mu_1 *)^2$.

Центральні емпіричні моменти третього і четвертого порядку використовуються при дослідженні форми статистичного розподілу.

Коефіцієнтом асиметрії статистичного розподілу називається число A, яке обчислюється за формулою

$$A = \frac{\nu_3^*}{\sigma_в^3}. \tag{13.17}$$

Якщо $A=0$, то розподіл має симетричну форму (рис. 7.4), тобто варіанти, рівновіддалені від $\overline{x}_в$, мають однакову частоту.

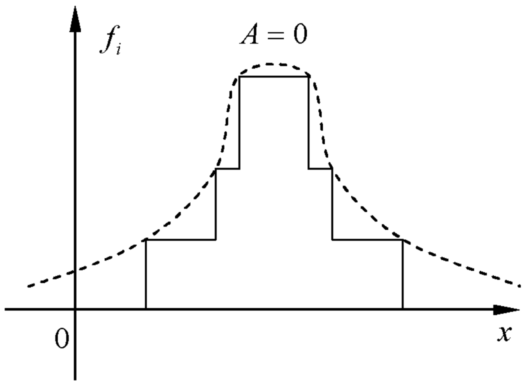

Рис. 13.4

При $A>0$ $(A<0)$ має місце правостороння (лівостороння) асиметрія (рис. 13.5 а) – б)).

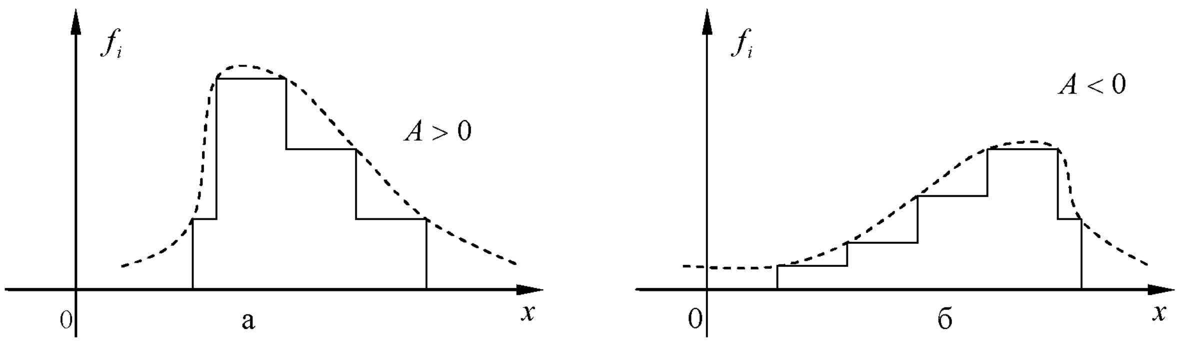

Рис.13.5

Ексцесом статистичного розподілу називається число, яке обчислюється за формулою

$$E = \frac{\nu_4^*}{\sigma_в^4} - 3. \qquad (13.18)$$

Ексцес є показником крутизни (загостреності) графіка статистичного розподілу у порівнянні з нормальним розподілом, для якого $E = 0$. Якщо $E > 0$ $(E < 0)$, то статистичний розподіл має більш загострену (пологу) вершину порівняно із нормальною кривою (рис.13.6 а) – б))

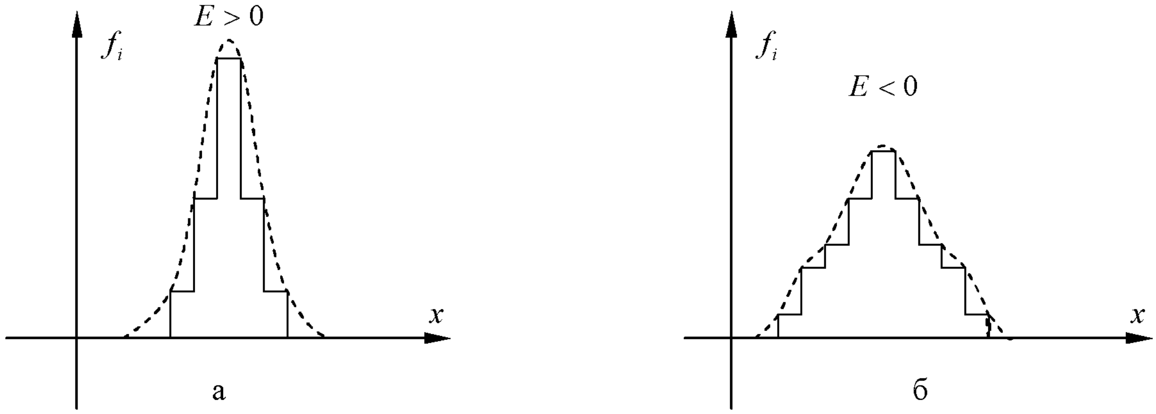

Рис.13.6

Приклад 13.3. Обчислити коефіцієнт варіації, коефіцієнт асиметрії і ексцес статистичного розподілу, наведеного у прикладі 13.2.

Розв'язання. Перейдемо від інтервального розподілу до дискретного, варіантами якого будуть $x_i = \dfrac{\alpha_{i-1} + \alpha_i}{2}$:

x_i	5	20	40	60	75	90
n_i	15	22	28	18	12	5

За формулами (13.1), (13.9), (13.13) обчислюємо $\overline{x}_в, D_в, \sigma_в$:

$$\overline{x}_в = \frac{1}{n}\sum_{i=1}^{6} x_i n_i = \frac{1}{100}(75 + 440 + 1120 + 1080 + 900 + 450) = 40,65;$$

$$D_в = \frac{1}{n}\sum_{i=1}^{6} x_i^2 n_i - (\overline{x}_в)^2 = \frac{1}{100}(375 + 8800 + 44800 + 64800 + 67500 + 40500) -$$

$$-1652,4225 = 2267,75 - 1652,4225 = 615,3275;$$

$$\sigma_в = \sqrt{D_в} = \sqrt{615,3275} \approx 24,8.$$

За формулою (13.14) знаходимо коефіцієнт варіації

$$V = \frac{\sigma_в}{x_в} \cdot 100\% = \frac{24,8}{40,65} \cdot 100\% \approx 61\%$$

Обчислюємо центральні емпіричні моменти (13.15) третього і четвертого порядків. Результати обчислень представимо у табл. 13.1.

Таблиця 13.1

x_i	n_i	$x_i - \overline{x}_в$	$(x_i - \overline{x}_в)^3 n_i$	$(x_i - \overline{x}_в)^4 n_i$
5	15	-35,65	-679625,79	24228660
20	22	-20,65	-193723,74	4000395
40	28	-0,65	-7,69	5
60	18	19,35	130411,35	2523460
75	12	34,35	486364,04	16706605
90	5	49,35	600940,5	29656413
$\sum$	100		344358,8	77115538

Отже, маємо

$$\nu_3{*} = \frac{1}{n}\sum_{i=1}^{6}(x_i - \overline{x}_в)^3 n_i = \frac{344358,8}{100} = 3443,588.$$

$$\nu_4{*} = \frac{1}{n}\sum_{i=1}^{6}(x_i - \overline{x}_в)^4 n_i = \frac{77115538}{100} = 771155,38.$$

Скориставшись формулами (13.17), (13.18), знаходимо:

$$A = \frac{v_3}{\sigma_в^3} = \frac{3443,588}{24,83^3} = \frac{3443,588}{15252,992} \approx 0,2;$$

$$E = \frac{v_4}{\sigma_в^4} - 3 = \frac{771155,38}{24,84^4} - 3 = \frac{771155,38}{378274,2} - 3 \approx -0,96.$$

Оскільки коефіцієнт $A > 0$ і порівняно близький до нуля, то статистичний розподіл ознаки X має незначну правосторонню асиметрію. Ексцес $E < 0$, тому крива даного статистичного розподілу має пологу вершину.

Запитання для самоконтролю

1. Як обчислюються $\overline{x_b}, D_b, \sigma_b$ у випадку дискретного статистичного розподілу?
2. Що таке умовні варіанти і коли їх використовують?
3. Як обчислюються $\overline{x_b}, D_b$ через групові середні та дисперсії?
4. Що таке мода, медіана дискретного статистичного розподілу?
5. Як визначити моду для інтервального статистичного розподілу?
6. Як визначається медіана інтервального статистичного розподілу?
7. За якою формулою обчислюється центральний емпіричний момент ℓ-го порядку?
8. Що таке початковий емпіричний момент ℓ-го порядку?
9. Як визначається і що характеризує коефіцієнт асиметрії статистичного розподілу?
10. Як обчислюється коефіцієнт ексцесу статистичного розподілу і що він характеризує?
11. Що таке розмах та коефіцієнт варіації вибірки?

Глосарій: варіаційний ряд, вибіркова середня, дисперсія вибірки, групова дисперсія, між групова дисперсія, ексцес, коефіцієнт асиметрії, коефіцієнт варіації, медіана мода, початковий емпіричний момент, центральний емпіричний момент.

Практичне заняття №13

Числові характеристики статистичного розподілу вибірки.

№13.1. Для обчислення середньої врожайності озимої пшениці поле площею 2000 га було поділено на 20 рівних ділянок. Фактичний урожай на кожній ділянці наведено в таблиці:

x_i $ц/га$	28	30	32	34	35
n_i	2	3	8	4	3

Потрібно: 1) обчислити $\overline{x}_b$, D_b, σ_b, R, V;

2) знайти $\overline{M_0}$, $\overline{M_e}$.

Розв'язання. Оскільки першопочаткові варіанти виявилися досить великими числами, доцільно перейти до умовних варіант $u_i = x_i - 32$. Таким чином, дістанемо розподіл умовних варіант:

u_i	–4	–2	0	2	4
n_i	2	3	8	4	3

Знайдемо вибіркову середню за формулою (13.1), скориставшись її властивістю 2^0:

$$\overline{x}_b = \frac{1}{20}\left(-4\cdot 2 - 2\cdot 3 + 0\cdot 8 + 2\cdot 4 + 4\cdot 3\right) + 32 = 0{,}3 + 32 = 32{,}3.$$

При цьому дисперсія не зміниться і, отже, її обчислимо за формулою (13.9):

$$D_b = \frac{1}{20}\left((-4)^2 \cdot 2 + (-2)^2 \cdot 3 + 0^2 \cdot 8 + 2^2 \cdot 4 + 4^2 \cdot 3\right) -$$

$$-\left[\frac{1}{20}\left(-4 \cdot 2 - 2 \cdot 3 + 0 \cdot 8 + 2 \cdot 4 + 4 \cdot 3\right)\right]^2 = 5{,}4 - 0{,}3^2 = 5{,}4 - 0{,}09 = 5{,}31.$$

Тоді $\sigma_b = \sqrt{D_b} = \sqrt{5{,}31} \approx 2{,}3$, а $V = \frac{\sigma_b}{\overline{x}_b} \cdot 100\% = \frac{2{,}3}{32{,}3} \cdot 100\% \approx 7{,}12\%$ і

$R = x_{\max} - x_{\min} = 36 - 28 = 8$.

Розглянутий розподіл є одномодальним, тому що $M_0 = 32$. Оскільки варіанта $x_3 = 32$ поділяє заданий варіаційний ряд на дві рівні за кількістю варіант частини, то $M_e = 32$.

№13.2. На кожну сотню деталей у середньому припадає дві браковані. Перевірено 10 партій по 100 деталей у кожній. Відхилення x_i кількості виявлених бракованих деталей від середнього значення наведено в таблиці:

Номер партії	1	2	3	4	5	6	7	8	9	10
x_i	–1	0	1	1	–1	1	0	–2	2	1

Побудувати дискретний статистичний розподіл та визначити коефіцієнт асиметрії A.

<u>Розв'язання.</u> За умовою дискретний статистичний розподіл має вигляд:

x_i	–2	–1	0	1	2
n_i	1	2	2	4	1

Як відомо, $\overline{x}_b = \frac{1}{10}\left(-2\cdot 1 - 1\cdot 2 + 0\cdot 2 + 1\cdot 4 + 2\cdot 1\right) = 0{,}2$,

$$D_b = \frac{1}{10}\left((-2)^2\cdot 1 + (-1)^2\cdot 2 + 0^2\cdot 2 + 1^2\cdot 4 + 2^2\cdot 1\right) - 0{,}2^2 = 1{,}4 - 0{,}04 = 1{,}36,$$

$\sigma_b = \sqrt{1{,}36} \approx 1{,}17$.

За формулою (13.15) обчислимо центральний емпіричний момент третього порядку:

$$\nu_3{}^* = \frac{1}{10}\left((-2-0{,}2)^3\cdot 1 + (-1-0{,}2)^3\cdot 2 + (0-0{,}2)^3\cdot 2 + (1-0{,}2)^3\cdot 4 + (2-0{,}2)^3\cdot 1\right) =$$

$$= \frac{1}{10}\left(-10{,}648 - 1{,}728\cdot 2 - 0{,}008\cdot 2 + 0{,}512\cdot 4 + 5{,}832\right) = -\frac{6{,}24}{10} = -0{,}624.$$

Остаточно за формулою (13.17) маємо:

$$A = \frac{-0{,}624}{1{,}17^3} \approx -0{,}39.$$

Тут коефіцієнт A<0 порівняно малий. Отже, статистичний розподіл близький до симетричного (має незначну лівосторонню асиметрію).

№13.3. Статистичний розподіл виробітку продукції за зміну наведено у таблиці:

x_i ц	7,5	8,5	9,5	10,5
n_i	2	4	3	1

Визначити коефіцієнт E.

Розв’язання. Обчислюємо значення величин:

$$\overline{x}_b = \frac{7{,}5\cdot 2 + 8{,}5\cdot 4 + 9{,}5\cdot 3 + 10{,}5\cdot 1}{2+4+3+1} = 8{,}8;$$

$$D_b = \frac{1}{10}\left(7{,}5^2\cdot 2 + 8{,}5^2\cdot 4 + 9{,}5^2\cdot 3 + 10{,}5^2\cdot 1\right) - 8{,}8^2 =$$

$$= \frac{1}{10}\left(112{,}5 + 289 + 270{,}75 + 110{,}25\right) - 77{,}44 = 0{,}81;$$

$\sigma_b = \sqrt{0{,}81} = 0{,}9.$

За формулою (13.15) дістанемо центральний емпіричний момент четвертого порядку:

$$\nu_4{}^* = \frac{1}{10}\left((7{,}5-8{,}8)^4\cdot 2 + (8{,}5-8{,}8)^4\cdot 4 + (9{,}5-8{,}8)^4\cdot 3 + (10{,}5-8{,}8)^4\cdot 1\right) =$$

$$= \frac{1}{10}(2{,}8561\cdot 2 + 0{,}0324 + 0{,}2401\cdot 3 + 8{,}3521) = 1{,}4817.$$

Отже, за формулою (13.18) маємо:

$$E = \frac{1{,}4817}{0{,}9^4} - 3 = \frac{1{,}4817}{0{,}6561} - 3 \approx -0{,}74.$$

Оскільки $E < 0,$ то полігон статистичного розподілу має більш пологу вершину, ніж крива нормального розподілу.

№13.4. Знайти вибіркові середнє та дисперсію за статистичним розподілом:

x_i	0,01	0,03	0,05
n_i	5	3	2

<u>Розв'язання.</u> Для зручності перейдемо до умовних варіант $u_i = 100x_i$. Тоді дістанемо такий розподіл умовних варіант:

u_i	1	3	5
n_i	5	3	2

Обчислюємо шукані величини: $\overline{u_b} = (1\cdot 5 + 3\cdot 3 + 5\cdot 2)/(5+3+2) = (5+9+10)/10 = 2{,}4;$

$\overline{x}_b = \frac{\overline{u_b}}{100} = 0{,}024$ (скористались властивістю 1^0 для вибіркової середньої);

$$D_{bu} = \tfrac{1}{10}(1^2\cdot 5 + 3^2\cdot 3 + 5^2\cdot 2) - \left[\tfrac{1}{10}(1\cdot 5 + 3\cdot 3 + 5\cdot 2)\right]^2 = 8{,}2 - 5{,}76 = 2{,}44$$

$D_{\text{В}} = \frac{D_{bu}}{100^2} = 0{,}000244 \approx 0{,}0002$ (за властивістю 2^0 для вибіркової дисперсії).

№13.5. Сукупність складена з двох груп, поданих в таблицях:

1)

x_i	2	4	5
n_i	1	7	2

2)

x_i	3	8
n_i	2	3

Знайти загальну середню та міжгрупову і загальну дисперсії.

Розв'язання. Знайдемо групові середні:

$$\overline{x}_1 = \frac{2 \cdot 1 + 4 \cdot 7 + 5 \cdot 2}{1 + 7 + 2} = 4;$$

$$\overline{x}_2 = \frac{3 \cdot 2 + 8 \cdot 3}{2 + 3} = 6.$$

За формулою (13.4) обчислюємо загальну середню:

$$\overline{x}_b = \frac{4 \cdot 10 + 6 \cdot 5}{10 + 5} = \frac{70}{15} = \frac{14}{3}.$$

Тепер відшукаємо групові дисперсії:

$$D_1 = \frac{(2-4)^2 \cdot 1 + (4-4)^2 \cdot 7 + (5-4)^2 \cdot 2}{10} = 0{,}6;$$

$$D_2 = \frac{(3-6)^2 \cdot 2 + (8-6)^2 \cdot 3}{5} = 6.$$

Міжгрупову дисперсію обчислюємо за формулою (13.11):

$$D_m = \frac{1}{15}\left(\left(4 - \frac{14}{3}\right)^2 \cdot 10 + \left(6 - \frac{14}{3}\right)^2 \cdot 5\right) = \frac{40 + 80}{90 \cdot 15} = \frac{8}{9}.$$

Нарешті, за формулою (13.12) маємо загальну дисперсію:

$$D_b = \frac{1}{15}(0{,}6 \cdot 10 + 6 \cdot 5) + \frac{8}{9} = \frac{36}{15} + \frac{8}{9} = \frac{148}{45}.$$

№13.6. За заданим інтервальним статистичним розподілом вибірки:

$[\alpha_{i-1};\alpha_i)$	0-2	2-4	4-6	6-8
n_i	1	3	4	2

Знайти $\overline{M_0}, \overline{M_e}, \bar{x}_b, D_b, \sigma_b, V, A, E.$

<u>Розв'язання.</u> Визначаємо модальний інтервал: $[4;6)$, якому відповідає найбільша частота $n_3 = 4$. Частоти домодального та післямодального інтервалів відповідно становлять $n_2 = 3$ і $n_4 = 2$. Беручи до уваги те, що $h = 6 - 4 = 2$ і $\alpha_2 = 4$, за формулою (13.5) дістанемо: $\overline{M_0} = 4 + \frac{4-3}{2\cdot 4 - 3 - 2}\cdot 2 = 4 + \frac{2}{3} = \frac{14}{3} \approx 4{,}67.$

Знаходимо емпіричну функцію розподілу:

$$F_n(x) = \begin{cases} 0, & при \quad x \le 0, \\ 0{,}1, & при \quad x = 2, \\ 0{,}4, & при \quad x = 4, \\ 0{,}8, & при \quad x = 6, \\ 1, & при \quad x \ge 8; \end{cases}$$

та будуємо її графік (рис. 13.7):

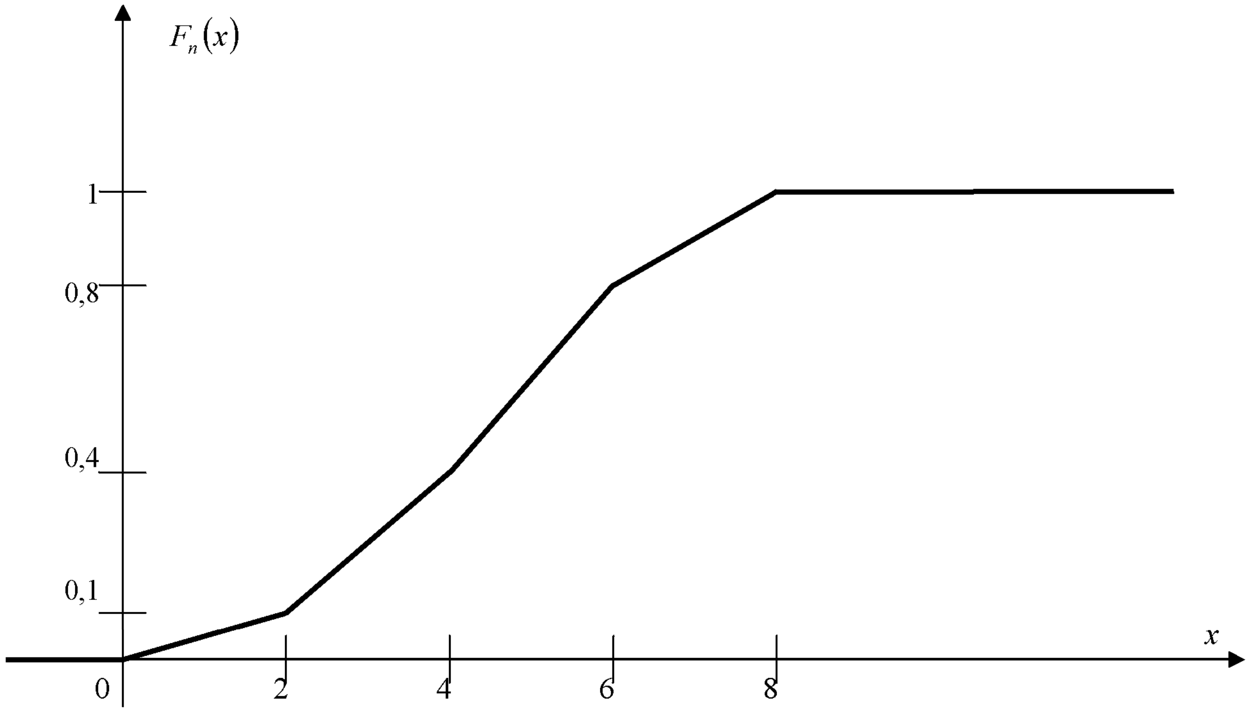

Рис. 13.7

Із графіка визначаємо медіанний інтервал $[4;6)$, який містить аргумент, для якого функція набуває значення $F_n(x) = 0{,}5$. Враховуючи, що $\alpha_{i-1} = \alpha_2 = 4$, $h = 2$, $F_n(\alpha_{i-1}) = F_n(4) = 0{,}4$ та $F_n(\alpha_i) = F_n(6) = 0{,}8$, за формулою (13.6) маємо:

$$\overline{M_e} = 4 + \frac{0{,}5 - 0{,}4}{0{,}8 - 0{,}4} \cdot 2 = 4 + \frac{0{,}2}{0{,}4} = 4{,}5.$$

Тепер перейдемо від інтервального до дискретного статистичного розподілу, варіантами якого будуть $x_i = \dfrac{\alpha_{i-1} + \alpha_i}{2}$:

x_i	1	3	5	7
n_i	1	3	4	2

Тоді, як звичайно, відшукуємо величини:

$$\overline{x}_b = \frac{1 \cdot 1 + 3 \cdot 3 + 5 \cdot 4 + 7 \cdot 2}{1 + 3 + 4 + 2} = \frac{1 + 9 + 20 + 14}{10} = 4{,}4;$$

$$D_b = \frac{1^2 \cdot 1 + 3^2 \cdot 3 + 5^2 \cdot 4 + 7^2 \cdot 2}{10} - 4{,}4^2 = \frac{1 + 27 + 100 + 98}{10} - 19{,}36 =$$
$$= 22{,}6 - 19{,}36 = 3{,}24;$$

$$\sigma_b = \sqrt{3{,}24} = 1{,}8; \ \ V = \frac{1{,}8}{4{,}4} \cdot 100\% \approx 41\%;$$

$$\nu_3{}^* = \frac{1}{10}\left((1-4{,}4)^3 \cdot 1 + (3-4{,}4)^3 \cdot 3 + (5-4{,}4)^3 \cdot 4 + (7-4{,}4)^3 \cdot 2\right) =$$
$$= \frac{1}{10}(-39{,}304 - 8{,}232 + 0{,}864 + 35{,}152) = -1{,}152;$$

$$\nu_4{}^* = \frac{1}{10}\left((1-4{,}4)^4 \cdot 1 + (3-4{,}4)^4 \cdot 3 + (5-4{,}4)^4 \cdot 4 + (7-4{,}4)^4 \cdot 2\right) =$$
$$= \frac{1}{10}(133{,}6336 + 3{,}8416 \cdot 3 + 0{,}1296 \cdot 4 + 45{,}6976 \cdot 2) = 23{,}7072;$$

$$A = \frac{-1{,}152}{1{,}8^3} \approx -0{,}2; \ \ E = \frac{23{,}7072}{1{,}8^4} - 3 \approx -0{,}74.$$

Таким чином, статистичний розподіл має незначну лівосторонню асиметрію, а його гістограма частот має пологу вершину.

№13.7. Результати опитування мешканців сільської місцевості про споживання молока і молочних продуктів (в кг на одного мешканця) наведені у вигляді статистичного ряду: 219, 222, 220, 218, 219, 221, 220, 219, 220, 222, 218, 219, 218, 220, 220, 219, 220, 221, 220, 220. Потрібно побудувати дискретний статистичний розподіл та обчислити $\bar{x}_b, D_b, \sigma_b, R, V, \overline{M_0}, \overline{M_e}$.

№13.8. Двадцять п’ять учнів класу на випускних іспитах з алгебри дістали таку кількість балів: 10, 8, 10, 9, 9, 12, 12, 8, 9, 8, 8, 9, 11, 10, 11, 10, 10, 9, 11, 10, 11, 10, 9, 10, 9. Побудувати дискретний статистичний розподіл і знайти $\bar{x}_b, D_b, \sigma_b, R, V, \overline{M_0}, \overline{M_e}$.

№13.9. Вивчались показники діяльності 100 малих підприємств харчової промисловості. Інформація про основні фонди (в млн.. грн.) подана у вигляді інтервального статистичного розподілу:

Інтервал, млн. грн.	2,6-3,0	3,0-3,4	3,4-3,8	3,8-4,2
n_i	30	40	20	10

Обчислити $\overline{M_0}, \overline{M_e}, \bar{x}_b, D_b, \sigma_b, V, A, E.$

№13.10. Число розладнань у роботі лінії розливу пива протягом місяця подано у вигляді інтервального статистичного розподілу:

Інтервал, $[\alpha_{i-1}; \alpha_i)$	0-2	2-4	4-6	6-8
n_i	2	4	3	1

Знайти $\overline{M_0}, \overline{M_e}, \bar{x}_b, D_b, \sigma_b, V, A, E.$

№13.11. Знайти вибіркові середнє і дисперсію за заданим статистичним розподілом:

x_i	0,01	0,04	0,05	0,07
n_i	5	20	15	10

№13.12. Обчислити вибіркові середнє та дисперсію згідно статистичного розподілу:

x_i	1250	1270	1290	1310
n_i	2	5	2	1

№13.13. Одержано дані про успішність студентів університету за навчальний рік:

Оцінка	Кількість студентів у %	
	зимова сесія	літня сесія
5	25	20
4	37	36
3	30	34
2	8	10

Обчислити: 1) середній бал успішності студентів за результатами зимової і літньої сесії та за навчальний рік; 2) групові дисперсії, міжгрупову і загальну дисперсії.

№13.14. Одержано дані про розподіл робітників трьох цехів підприємства за тарифними розрядами:

Розряд	Чисельність робітників цеху, чол.		
	№1	№2	№3
1	2	3	1
2	6	5	4
3	8	10	10
4	15	20	15
5	10	15	7
6	9	7	3
Разом	50	60	40

Обчислити: 1) середні тарифні розряди робітників кожного цеху і підприємства; 2) показники розсіювання (групові, міжгрупову та загальну дисперсії).

№13.15. Наведено дані про стаж роботи працівників двох підприємств концерну:

Стаж роботи, років	Кількість працівників, чол.	
	№1	№2
до 1	3	4
1 - 5	19	12
5 - 10	27	23
10 - 15	12	22
15 - 20	7	14
20 - 25	2	5
Разом	70	80

Обчислити: 1) середній стаж роботи працівників підприємств і концерну; 2) групові і загальну дисперсії.

№13.16. Обчислити коефіцієнти асиметрії та ексцесу статистичного розподілу:

x_i	1,2	1,4	1,6	1,8
n_i	2	4	3	1

№13.16. Обчислити коефіцієнти асиметрії та ексцесу статистичного розподілу:

x_i	0,1	0,3	0,5	0,7
n_i	1	2	5	2