Ряд статистичних даних, отриманий у результаті їх зведення та групування за певною кількісною чи якісною ознакою, називають *рядом розподілу*.

Розглянемо принципи формування груп. У кожному конкретному дослідженні вирішуються три питання:

1) вибір ознаки групування;
2) визначення числа груп і визначення інтервалу зміни ознаки для кожної групи;
3) визначення показників, що повинні характеризувати групу.

Ознака, за якою здійснюється групування, може бути кількісною або атрибутивною. *Атрибутивна ознака* виражається словом (стать, сімейний стан, форма власності тощо). *Кількісна* або *варіаційна ознака* виражається числом (вік, заробітна плата, обсяг виробництва).

Якщо групування проводять за варіаційною ознакою, постає питання щодо кількості груп і величини інтервалів групування (інтервалів зміни ознаки). Чим більшою є різниця між максимальним та мінімальним значенням ознаки (цю різницю називають варіаційним розмахом), тим на більшу кількість груп поділяють генеральну сукупність. Кількість груп залежить також від чисельності досліджуваної сукупності. Чим більше створено груп, тим менший інтервал групування, і навпаки.

Застосовують рівні та нерівні інтервали групування. Величину рівних інтервалів групування визначають за формулою

$$h = \frac{x_{\max} - x_{\min}}{n}, \qquad (1.3)$$

де h – довжина інтервалу, $x_{\max}$ – найбільше значення ознаки, $x_{\min}$ – її найменше значення, n – число груп.

Приклад 1.7. Величина щоденної виручки магазинів торгівельної мережі коливається від 750 до 950 тис. грн. Необхідно провести групування магазинів за розмірами виручки, утворивши 5 груп з рівними інтервалами.

Знайдемо довжину одного інтервалу. $x_{\max}$ = 950 тис. грн., $x_{\min}$ = 750 тис. грн., n = 5.

$$h = \frac{950 - 750}{5} = 40 \text{ } тис.грн.$$

Отже, отримуємо наступні групи магазинів зі значеннями виручки, що знаходяться у інтервалах (тис. грн.): 750–790, 790–830, 830–870, 870–910, 910–950.

Для визначення кількості рівних інтервалів, на які розбивають сукупність спостережень, застосовують *формулу,* запропоновану американським вченим Гербертом *Стерджесом*:

$$n = 1 + 3{,}322 \cdot \lg N, \tag{1.4}$$

де N – кількість спостережень.
Число інтервалів, яке відповідає обсягу сукупності N, розраховане за формулою Стерджеса, наведене у таблиці 1.1.

Таблиця 1.1. Визначення кількості рівних інтервалів за формулою Стерджеса

Обсяг сукупності, N	15-24	25-44	45-89	90-179	180-359	360-719
Число інтервалів, n	5	6	7	8	9	10

На практиці часто застосовують також нерівні інтервали.

Приклад 1.8. За даними таблиці 1.2 побудувати інтервальний ряд розподілу статистичних даних щодо випуску продукції підприємствами регіону у 2020 р.

Таблиця 1.2. Випуск продукції підприємствами регіону у 2020 р.

№ підприємства	Валова продукція, млн. грн.		Виконання плану,%
	План	Факт	
1	6,5	7,1	109,2
2	3,6	2,9	80,6
3	13,5	14,0	103,7
4	7,2	4,8	66,7
5	14,3	15,2	109,8
6	10,2	11,8	115,7
7	16,0	16,1	100,6
8	13,4	16,6	123,9
9	11,2	10,2	91,1
10	0,6	0,6	100,0
11	0,8	0,9	112,5
12	2,2	2,6	118,2
13	7,0	5,5	78,6
14	4,0	4,1	102,5
15	4,5	4,9	108,9
Разом	115,0	117,8	102,4

$h=\frac{x_{max}-x_{min}}{n}$

Виконаємо складне зведення цих статистичних даних (таблиця 1.3).

Таблиця 1.3. Випуск продукції підприємствами регіону (зведені дані)

Рівень виконання плану, %	Кількість підприємств	Випуск валової продукції, млн. грн.		Виконання плану, %	Вироблено понад план (+) чи недовиконано (-), млн. грн.
		план	факт		
До 100	4	29	23,4	80,7	–5,6
100-110	7	59,4	62,5	105,2	+3,1
Понад 110	4	26,6	31,9	119,9	+5,3
Разом	15	115	117,8	102,4	+2,8

Для побудови інтервального ряду використаємо формулу Стерджеса. З таблиці 1.2 знаходимо: $x_{\max} = 16,6$, $x_{\min} = 0,6$, $R = x_{\max} - x_{\min} = 16,6 - 0,6 = 16,0$ (млн. грн.)

Кількість інтервалів дорівнює:

$$n = 1 + 3,322 \cdot \lg 15 = 1 + 3,322 \cdot 1,176 \approx 5.$$

Зауважимо, що те ж значення $n = 5$ отримуємо для обсягу сукупності $N = 15$, використавши таблицю 1.1.

Знайдемо довжину інтервалу:

$$h = \frac{x_{\max} - x_{\min}}{n} = \frac{16}{5} = 3,2 \text{ (млн. грн.)}$$

Перший інтервал починається з $x_{\min} = 0,6$ млн. грн. Додаючи $h = 3,2$ млн. грн., знаходимо верхню межу початкового інтервалу: 0,6+3,2=3,8 (млн. грн.). Продовжуючи послідовно цей процес, знаходимо всі інші інтервали і заповнюємо таблицю 1.4 згрупованих даних.

Таблиця 1.4. Випуск продукції підприємствами регіону (згруповані дані).

Валова продукція, факт, млн. грн.	Кількість підприємств	Виконання плану, %
0,6 – 3,8	4	102,9

3,8 – 7,0	4	85,0
7,0 – 10,2	1	109,2
10,2 – 13,4	2	102,8
13,4 – 16,8	4	109,1
Разом	15	102,4

1.3 Види рядів розподілу

Розрізняють *атрибутивні* (утворені за якісною ознакою) та *варіаційні* (утворені за кількісною ознакою) ряди розподілу. Таблиця 1.5 є прикладом атрибутивного, таблиця 1.6 – варіаційного ряду розподілу.

Таблиця 1.5. Розподіл підприємств і обсягу виробленої ними промислової продукції за формами власності

Форма власності підприємства	Кількість підприємств	Обсяг продукції, %
Державна	1367	18,3
Комунальна	2093	2,2
Приватна	11170	1,5
Колективна	32498	77,1
Міжнародних організацій	220	0,9
Всього	47348	100

Варіаційні ряди поділяються на *інтервальні* та *дискретні* (таблиці 1.6 та 1.7). Інтервальні ряди поділяються на ряди з однаковими та неоднаковими інтервалами.

Варіант – це окреме значення групувальної ознаки, частота – це кількість появ варіанта у ряді. Для рядів з неоднаковими інтервалами використовують додаткову характеристику – *щільність розподілу*, тобто частоту, що припадає на одиницю довжини інтервалу (таблиця 1.8).

Накопичена частота показує, для скількох одиниць сукупності значення ознаки не перевищує відповідного варіанту (для дискретного ряду) або верхньої межі відповідного інтервалу (для інтервального ряду).

Позначимо x – значення ознаки, f – частота, ρ – щільність, d_f – *відносна частота (частка)*, s_f – накопичена частота, s_d – накопичена відносна частота (частка), k – порядковий номер варіанту (інтервалу), h – довжина інтервалу. З урахуванням цих позначень отримуємо:

$$d_f = \frac{f}{N}, \quad (1.5)$$

$$s_f^k = s_f^{k-1} + f^k,\ k = 1,...,n,\ s_f^0 = 0, \tag{1.6}$$

$$s_d^k = s_d^{k-1} + d^k,\ k = 1,...,n,\ s_d^0 = 0, \tag{1.7}$$

$$\rho = \frac{f}{h}. \tag{1.8}$$

Наведемо приклад варіаційного дискретного ряду (таблиця 1.6) з розрахованими за формулами (1.5) – (1.7) для кожного варіанту показниками відносної частоти, накопиченої частоти та накопиченої відносної частоти.

Таблиця 1.6. Розподіл робітників підприємства за рівнем кваліфікації

Значення ознаки, x (кваліфікаційний розряд)	Частота, f (кількість робітників, чоловік)	Відносна частота, d_f (питома вага робітників, %)	Накопичена частота, s_f, чол.	Накопичена відносна частота, s_d, %
1	3	5,2	3	5,2
2	6	10,4	9	15,6
3	8	13,8	17	29,4
4	13	22,4	30	51,8
5	22	37,9	52	89,7
6	6	10,3	58	100,0
Разом	58	100	–	–

Таблиця 1.7 є прикладом інтервального варіаційного ряду з рівними інтервалами.

Таблиця 1.7. Розподіл підприємств регіону за рівнем рентабельності

Значення ознаки, x (рентабельність, %)	Частота, f (кількість підприємств)	Відносна частота, d_f (питома вага підприємств, %)	Накопичена частота, s_f	Накопичена відносна частота, s_d, %
До 5	2	6,2	2	6,2
5–10	3	9,4	5	15,6
10–15	10	31,2	15	46,8
15–20	8	25,0	23	71,8
20–25	6	18,8	29	90,6
Понад 25	3	9,4	32	100,0
Разом	32	100	–	–

Таблиця 1.8 відображає варіаційний ряд з нерівними інтервалами, для яких розраховано щільність розподілу.

Таблиця 1.8. Розподіл підприємств регіону за рівнем рентабельності

Значення ознаки, x (рентабельність, %)	Частота, f (кількість підприємств)	Відносна частота, d_f (питома вага підприємств, %)	Накопичена частота, s_f	Накопичена відносна частота, s_d, %	Щільність, ρ
До 5	2	6,2	2	6,2	0,4
5–20	21	65,6	23	71,8	1,4
20–30	10	28,2	32	100,0	1,9
Разом	32	100	–	–	–

Для графічного зображення рядів розподілу використовують:

– *гістограму* – для інтервального варіаційного ряду;

– *полігон* (ламану лінію, що сполучає сукупність ізольованих точок на площині) – для дискретного варіаційного ряду;

– *кумулятивну криву* або *кумуляту* (криву накопичених частот або часток) для графічного зображення дискретного та інтервального варіаційних рядів.

2. Статистичні характеристики рядів розподілу

2.1 Середні величини

Середньою величиною у статистиці називають узагальнюючий показник, що характеризує типовий рівень змінної ознаки у розрахунку на одиницю однорідної сукупності. Середні величини використовують з метою виявлення характерних закономірних рис об'єктів дослідження у конкретних умовах місця і часу.

При обчисленні середньої величини необхідно чітко усвідомити її логічну формулу. На підставі цієї формули вибирають вид середньої величини. Чисельник логічної формули дорівнює обсягу значень змінної ознаки, а знаменник – обсягу сукупності. Наведемо приклади логічних формул середніх величин.

1) $\text{середня місячна заробітна плата} = \dfrac{\text{фонд заробітної плати}}{\text{кількість працюючих}}$;

2) $\text{середня ціна} = \dfrac{\text{виручка}}{\text{кількість товару}}$;

3) $\text{середня урожайність} = \dfrac{\text{валовий збір}}{\text{площа}}$;

У кожному конкретному випадку для реалізації логічної формули використовується певний вид середньої. Розглянемо ці види.

Середня арифметична використовується для осереднення прямих значень ознаки шляхом їх підсумовування. Якщо дані не згруповані, використовується *простa середня арифметична*:

$$\bar{x} = \frac{\sum x}{N} = \frac{x_1 + x_2 + \ldots + x_N}{N}. \quad (2.1)$$

Приклад 2.1. За місяць відділення комерційного банку залучило депозити на суму, тис. грн.: 50, 40, 35, 20, 22. 26. Середня величина залученого депозиту становить:

$$\bar{x} = \frac{50 + 40 + 35 + 20 + 22 + 26}{6} \approx 32{,}2(\text{тис.грн.})$$

Якщо дані згруповані, використовують *середню арифметичну, зважену за частотою*:

$$\bar{x} = \frac{\sum x \cdot f}{\sum f} = \frac{x_1 f_1 + x_2 f_2 + \ldots + x_n f_n}{f_1 + f_2 + \ldots + f_n}. \quad (2.2)$$

Використовують також *середню арифметичну, зважену за відносною частотою (часткою):*

$$\bar{x} = x_1 d_1 + x_2 d_2 + \ldots + x_n d_n. \quad (2.3)$$

Приклад 2.2. За даними таблиці 2.1 обчислити середню банківську процентну ставку за депозитами.

Таблиця 2.1. Дані щодо процентних ставок за депозитами, залученими комерційним банком

Депозитна ставка, %	x	23	10	13	6	7	11	Разом
Кількість депозитів	f	10	15	12	6	3	4	50

Частка депозитів, %	d	20	30	24	12	6	8	100

Обчислимо середню депозитну ставку, зважену за частотою:

$$\bar{x} = \frac{\sum x \cdot f}{\sum f} = \frac{23 \cdot 10 + 10 \cdot 15 + 13 \cdot 12 + 6 \cdot 6 + 7 \cdot 3 + 11 \cdot 4}{50} = 12,7(\%)$$

Середня депозитна ставка, зважена за часткою:

$$\bar{x} = \frac{\sum x \cdot f}{\sum f} = 23 \cdot 0,2 + 10 \cdot 0,3 + 13 \cdot 0,24 + 6 \cdot 0,12 + 7 \cdot 0,06 +$$
$$+ 11 \cdot 0,08 = 0,127 = 12,7\%$$

Для розрахунку середньої арифметичної для інтервального варіаційного ряду перейдемо до дискретного ряду, утвореного з серединних (центральних) значень ознаки у кожному з інтервалів. Серединне значення ознаки у замкненому інтервалі розраховують за формулою

$$\bar{x}_i = x_* + \frac{h}{2} \ ,$$

де x_* – нижня межа інтервалу, h – його довжина. Умовою замкненості інтервалу можна знехтувати, якщо ознака є неперервною.

Розглянемо основні властивості середньої арифметичної.

1. Сума відхилень окремих варіант ознаки від середньої арифметичної дорівнює нулю: $\sum(x - \bar{x}) \cdot f = 0$, тобто у середній взаємно компенсуються додатні та від'ємні відхилення окремих варіант від середнього.
2. Сума квадратів відхилень окремих варіант ознаки від середньої менша, ніж від будь-якої іншої величини: $\sum(x - \bar{x})^2 \cdot f = min$.
3. Якщо всі варіанти збільшити або зменшити на одну і ту ж саму величину С (збільшити або зменшити в С разів), той середня арифметична величина зміниться аналогічно.

Для осереднення обернених показників використовують *середню гармонічну*. Якщо дані не згруповані, то використовується *проста середня гармонічна*:

$$\bar{x} = \frac{N}{\sum \frac{1}{x}} = \frac{N}{\frac{1}{x_1} + \frac{}{x_2} + \ldots + \frac{1}{x_N}}. \qquad (2.4)$$

Приклад 2.3. Витрати часу на виготовлення однієї деталі у першого робітника становлять 0,25 годин, у другого – 0,20 годин. Визначити середні витрати часу на виготовлення однієї деталі одним робітником.

Нехай x_1 та x_2 – витрати часу кожного робітника на виготовлення однієї деталі. Тоді $1/x_1$ кількість деталей, виготовлена першим робітником, $1/x_2$ – другим робітником. Середні витрати часу знаходимо за формулою (2.4):

$$\bar{x} = \frac{2}{\frac{1}{0,25} + \frac{1}{0,20}} = \frac{2}{4+5} \approx 0,22(\text{год.})$$

Для згрупованих даних використовують зважену середню гармонічну:

$$\bar{x} = \frac{\sum \omega}{\sum \frac{\omega}{x}} = \frac{\omega_1 + \omega_2 + \ldots + \omega_n}{\frac{\omega_1}{x_1} + \frac{\omega_2}{x_2} + \ldots + \frac{\omega_n}{x_n}}, \tag{2.5}$$

де $\omega = x \cdot f$ – обсяг значень ознаки.

Приклад 2.4. Маємо дані про показники заробітної плати в банку, наведені у таблиці 2.2.

Таблиця 2.2 Дані про заробітну плату у банку

Відділи	Середня заробітна плата одного працівника за місяць $(\bar{x})$, г.о.	Фонд заробітної плати за місяць $(\omega = x \cdot f)$, г.о.
Кредитний	600	4200
Валютний	700	3500

Визначити середню заробітну плату за місяць одного працівника по двох відділах у цілому.

Побудуємо логічну формулу середньої заробітної плати:

$$\textit{Середня місячна заробітна плата} = \frac{\textit{Фонд заробітної плати}}{\textit{Кількість працівників}}$$

Фонд заробітної плати загалом по двох відділах складає $\omega_1 + \omega_2 = 4200 + 3500 = 7700$ г.о. Кількість працівників:

$$\frac{\omega_1}{x_1}+\frac{\omega_2}{x_2}=\frac{4200}{600}+\frac{3500}{700}=12.$$

Середня заробітна плата становить:

$$\bar{x}=\frac{\omega_1+\omega_2}{\frac{\omega_1}{x_1}+\frac{\omega_2}{x_2}}=\frac{4200+3500}{\frac{4200}{600}+\frac{3500}{700}}=641,7(г.о.)$$

<u>*Середня*</u> <u>*хронологічна*</u> використовується для осереднення моментних показників (тих, що відносяться до певних моментів часу). Якщо є два моментних показники (на початок та кінець періоду), то середня розраховується як півсума цих значень за формулою:

$$\bar{x}=\frac{x_1+x_n}{2}, \tag{2.6}$$

де x_1 – значення показника на початок періоду, x_n – на кінець періоду.

Якщо моментів часу більше, ніж два, а інтервали часу між ними рівні, то у чисельнику до півсуми крайніх значень додають усі проміжні, а знаменником є число, яке на одиницю менше від числа значень ознаки:

$$\bar{x}=\frac{\frac{x_1+x_n}{2}+x_2+\ldots+x_{n-1}}{n-1}\quad. \tag{2.7}$$

Приклад 2.5. Залишки коштів на рахунку клієнта у банку на початок кожного місяця становили, тис. грн.: липень – 50, серпень – 60, вересень – 62, жовтень – 80. Знайти середньомісячний залишок коштів за третій квартал.

Використаємо формулу середньої хронологічної (2.7):

$$\bar{x}=\frac{\frac{50+80}{2}+60+62}{4-1}\approx 62,3(тис.грн.)$$

<u>*Середня*</u> <u>*геометрична*</u> використовується для обчислення середніх значень ланцюгових відносних показників динаміки, наприклад, середнього темпу зростання деякого показника. Вона обчислюється за формулою:

$$\overline{x} = \sqrt[n]{x_1 \cdot x_2 ... x_n}\ , \tag{2.8}$$

де $x_1, x_2, ..., x_n$ – ланцюгові відносні показники динаміки.

Приклад 2.6. За даними таблиці 2.3 розрахувати середній темп зміни обсягу реалізованої продукції.

Таблиця 2.3. Розподіл обсягу реалізованої продукції по кварталах

Квартал	Обсяг реалізації продукції, тис. г.о.	Темп зміни обсягу реалізованої продукції, у % до попереднього кварталу
1	2560,3	–
2	1978,6	77,3
3	2380,4	120,3
4	2876,3	120,8
Разом	9795,6	–

Середній темп зміни обсягу реалізованої продукції розрахуємо за формулою (2.8) середнього геометричного:

$$\overline{x} = \sqrt[3]{77,3 \cdot 120,3 \cdot 120,8} \approx 103,47\% .$$

<u>Середня</u> *<u>квадратична</u>* використовується у статистиці при розрахунку показників варіації. Формула для її розрахунку має вигляд:

$$\overline{x} = \sqrt{\frac{\sum x^2}{n}}\ . \tag{2.9}$$

Застосування середньої квадратичної буде розглянуто у подальшому при вивченні показників варіації статистичного розподілу.

2.2 Характеристики центру розподілу

Аналіз варіаційного ряду розподілу полягає у виявленні закономірностей зміни частот залежно від зміни кількісної ознаки, покладеної в основу групування. При аналізі варіаційних рядів найважливішими є такі групи показників: характеристики центру розподілу, характеристики величини варіації, характеристики форми розподілу.

<u>Центром</u> <u>розподілу</u> називають значення змінної ознаки, навколо якого групуються інші варіанти. До характеристик центру розподілу відносять середню величину, моду та медіану.

Середні величини були розглянуті вище. Вони характеризують типовий рівень ознаки у сукупності. Крім середніх величин, важливу роль при дослідженні статистичних розподілів відіграють мода та медіана, а також квартилі та децилі.

<u>Модою</u> називають значення ознаки, яке найчастіше зустрічається у сукупності. Для дискретного варіаційного ряду модою є варіант, що має найбільшу частоту (відносну частоту). Для інтервального варіаційного ряду з однаковими інтервалами моду обчислюють за *<u>формулою</u> <u>Орженцького</u>*:

$$M_0 = \underline{x}_{M_o} + h \cdot \frac{f_{M_o} - f_{M_o-1}}{\left(f_{M_o} - f_{M_o-1}\right)\left(f_{M_o} - f_{M_o+1}\right)}. \qquad (2.10)$$

У формулі (2.10) M_0 – мода варіаційного ряду, $\underline{x}_{M_o}$ – нижня межа модального інтервалу (інтервалу з найбільшою частотою), f_{M_o}, f_{M_o-1}, f_{M_o+1} – відповідно частоти модального, передмодального та післямодального інтервалів, h – довжина інтервалу.

Для інтервального варіаційного ряду з нерівними інтервалами при рівномірному розподілі щільності у межах кожного інтервалу моду розраховують після зведення ряду до ряду з рівними інтервалами.

Варіаційний ряд може мати одну або декілька мод. Наявність кількох мод свідчить про неоднорідність сукупності, тобто об’єднання у одній сукупності різноякісних одиниць.

Оскільки при графічному зображенні варіаційного ряду у вигляді полігону частот по осі абсцис відкладається варіант (значення ознаки), а по осі ординат – частота, то визначення моди для дискретного варіаційного ряду зводиться до знаходження абсциси точки з найбільшим значенням частоти на полігоні.

Для графічного визначення моди у інтервальному ряді з рівними інтервалами використовують гістограму розподілу. Праву верхню вершину модального прямокутника з’єднують з правою верхньою вершиною попереднього прямокутника, а ліва верхня вершина модального прямокутника – з лівою верхньою вершиною післямодального прямокутника. Абсциса точки перетину цих прямих є модою розподілу.

Приклад 2.7. У таблиці 2.4 наведено дані щодо розподілу працівників підприємства за розміром заробітної платні. Знайти середнє значення заробітної платні, а також її моду (найпоширенішу на підприємстві величину заробітної платні).

Оскільки варіаційний ряд є інтервальним, то для розрахунку середньої заробітної платні на підприємстві замінимо інтервальний ряд на дискретний, який утворюється з серединних значень ознаки (заробітної платні) для кожного з інтервалів.

Таблиця 2.4. Розподіл працівників за розміром заробітної платні

Заробітна платня, г.о.	Серединне значення заробітної платні, $\overline{x}_i$, г.о.	Частота, кількість працівників, f чол.	Частка, d, %	Накопичена	
				частота, чол.	частка, %
До 120	104,5	8	12,3	8	12,3
120-150	134,5	10	15,4	18	27,7
150-180	164,5	15	23,1	33	50,8
180-210	194,5	18	27,7	51	78,5
210-240	224,5	9	13,8	60	92,3
Понад 240	254,5	5	7,7	65	100,0
Разом	–	65	100,0	–	_–

Будемо вважати, що розмір заробітної платні установлюється з точністю до цілих значень. Варіаційний ряд є інтервальним, проте значення ознаки (величини заробітної платні) є дискретними. Для встановлення середнього розміру заробітної платні знайдемо довжини інтервалів і перейдемо від інтервального ряду до ряду, що утворюється серединними значеннями $\overline{x}_i$ інтервалів, взятими з відповідними частотами.

У даному прикладі не визначено чітко межі першого та останнього інтервалів. У такому разі на них поширюються довжини наступного (для початкового інтервалу) та попереднього (для кінцевого інтервалу) проміжків. У таблиці для інтервального варіаційного ряду верхня та нижня межі сусідніх інтервалів збігаються. У таких випадках прийнято вважати, що інтервали є напіввідкритими, тобто замкненими зліва і відкритими справа. Знайдемо довжини інтервалів:

$$h = 149 - 120 = 209 - 180 = 179 - 150 = 239 - 210 = 29(г.о.)$$

Отже, маємо ряд з рівними інтервалами довжиною 29 г.о. З урахуванням цього знаходимо середні у кожному з інтервалів:

$$\overline{x}_1 = 90 + \frac{29}{2} = 104,5; \overline{x}_2 = 120 + \frac{29}{2} = 134,5; \overline{x}_3 = 150 + \frac{29}{2} = 164,5;$$

$$\overline{x}_4 = 180 + \frac{29}{2} = 194,5;\ \overline{x}_5 = 210 + \frac{29}{2} = 224,5;\ \overline{x}_6 = 240 + \frac{29}{2} = 254,5.$$

Середній розмір заробітної платні:

$$\overline{x} = \frac{104,5 \cdot 8 + 134,5 \cdot 10 + 164,5 \cdot 15 + 194,5 \cdot 18 + 224,5 \cdot 9 + 254,5 \cdot 5}{65} \approx 176(г.о.)$$

Найбільша частота ($f_{\max} = 18$ чоловік) відповідає працівникам із заробітною платнею у межах від 180 до 210 одиниць, тобто цей інтервал є модальним. Передмодальним є інтервал (150; 180), післямодальним – інтервал (210; 240).Цим інтервалам відповідають наступні частоти: $f_{M_o} = 18$, $f_{M_o - 1} = 15$, $f_{M_o + 1} = 9$. Підставивши ці дані у формулу (2.10), отримаємо:

$$M_o = 180 + 30 \cdot \frac{18 - 15}{(18 - 15) + (18 - 9)} = 187,5(г.о.)$$

Отже, найбільш поширеною на підприємстві є заробітна платня працівника у розмірі 187,5 г.о.

<u>*Медіаною*</u> (M_e) називають варіант, розташований посередині впорядкованого варіаційного ряду. Медіана ділить ряд на дві рівні частини за чисельністю одиниць сукупності.

Для дискретного впорядкованого варіаційного ряду з непарним числом елементів медіану знаходять як варіант x з порядковим номером $\frac{n+1}{2}$, для ряду з парним числом членів медіану визначають як середню арифметичну двох варіантів з порядковими номерами $\frac{n}{2}$ та $\frac{n}{2} + 1$. Отже, медіану впорядкованого дискретного варіаційного ряду можна визначити за формулою:

$$M_e = \begin{cases} x_{\frac{n+1}{2}}, \quad n = 2k + 1, \\ \dfrac{x_{\frac{n}{2}} + x_{\frac{n+1}{2}}}{2}, n = 2k, k \in \mathbb{N}. \end{cases} \tag{2.11}$$

Знаходження медіани у інтервальному варіаційному ряді потребує попереднього визначення медіанного інтервалу, тобто інтервалу, для якого накопичена частота (накопичена частка) дорівнює півсумі всіх частот (часток) ряду, або перевищує її. Значення медіани для інтервального варіаційного ряду обчислюють за *формулою Фехнера*:

$$M_e = \underline{x}_{M_e} + h_{M_e} \cdot \frac{\frac{1}{2}\sum f - S_{M_e - 1}}{f_{M_e}}, \tag{2.12}$$

де $\underline{x}_{M_e}$ – нижня межа медіанного інтервалу, $\sum f$ – сума частот, S_{M_e-1} – сума частот, накопичених перед медіанним інтервалом, f_{M_e} – частота медіанного інтервалу.

У формулі (2.12) замість частот f можна використовувати частки d.

Медіана є центром розподілу кількості одиниць сукупності, на відміну від середньої арифметичної, яка є центром розподілу відхилень ознаки від її середнього значення $\bar{x}$. Для неоднорідної сукупності вона краще характеризує типовий рівень ознаки, ніж середня величина. Для якісно однорідної сукупності середня, мода і медіана незначно відрізняються між собою.

Водночас з модою та медіаною для повнішої характеристики сукупності використовують варіанти, що займають у впорядкованому ряді цілком визначене місце. До таких варіантів належать чверті (квартилі) та десяті частини (децилі). *Квартилі* ділять ряд за сумою частот на чотири рівні частини, *децилі* – на десять рівних частин.

Налічують 3 квартилі: Q_1 – нижню, $Q_2\left(Me\right)$ – серединну, Q_3 – верхню. Для початкової четвертої частини одиниць сукупності значення ознаки не перевищує Q_1, серединна чверть Q_2 – для половини одиниць сукупності значення ознаки не перевищує Q_2, Q_3 – для трьох четвертих одиниць сукупності.

При великому значенні обсягу сукупності, крім квартилів, можуть використовуватися і децилі (десяті частини), зміст і техніка обчислень яких аналогічні до обчислення квартилів.

Приклад 2.8. За даними таблиці 2.4 знайти медіану та чверті інтервального варіаційного ряду.

Обсяг сукупності становить 65 одиниць, тому посередині варіаційного ряду розміщений варіант з порядковим номером 33: $x_{\frac{65+1}{2}} = x_{33}$. Значення накопиченої частоти $S_f = 33$ свідчить, що 33-й член ряду міститься у інтервалі $\left|150;180\right)$. Нижня межа інтервалу $\underline{x}_{Me} = 150$. Сума частот, накопичених до медіанного інтервалу, $S_{M_e-1} = 18$. Частота медіанного інтервалу $f_{M_e} = 13$.

Довжина медіанного інтервалу $h_{M_e} = 30$. З урахуванням цих значень за формулою (2.12) отримуємо:

$$M_e = 150 + 30 \cdot \frac{\frac{1}{2} \cdot 65 - 18}{15} = 179,0.$$

Отже, для половини працівників розмір заробітної платні не перевищує 179 г.о.

Для знаходження першої квартилі використаємо формулу:

$$Q_1 = \underline{x}_{Q_1} + h_{Q_1} \cdot \frac{\frac{1}{4}\sum f - S_{Q_1 - 1}}{f_{Q_1}}, \qquad (2.13)$$

де $\underline{x}_{Q_1}$ – нижня межа чвертного інтервалу, h_{Q_1} – довжина цього інтервалу, $S_{Q_1 - 1}$ – сума частот, накопичених до першого чвертного інтервалу, f_{Q_1} – частота чвертного інтервалу.

У даному прикладі $\sum f = n = 65$, тому в першу чверть входять члени ряду з номерами від 1 до $\left[\frac{65}{4}\right] + 1 = 17$. Першим чвертним інтервалом є $[120; 150)$, оскільки для нього $S_f = 18 > \frac{65}{4}$. З урахуванням цього знаходимо першу квартиль:

$$Q_1 = 120 + 30 \cdot \frac{\frac{1}{4} \cdot 65 - 8}{10} = 144,75.$$

Знайдене значення Q_1 свідчить, що для 25% працівників заробітна платня не перевищує 144,75 г.о.

Друга квартиль $Q_2 = M_e = 179,0$ г.о.

Третю квартиль Q_3 розраховуємо аналогічно до Q_1, при цьому формула (2.13) набуває вигляду:

$$Q_3 = \underline{x}_{Q_3} + h_{Q_3} \cdot \frac{\frac{3}{4}\sum f - S_{Q_3 - 1}}{f_{Q_3}}. \qquad (2.14)$$