

СТАТИСТИКА.



Описательная статистика.

Лекция 3. Показатели формы распределения.

Автор: Равичев Л.В.

РХТУ им. Д.И.Менделеева

Кафедра управления технологическими инновациями

Москва - 2013

Показатели формы распределения

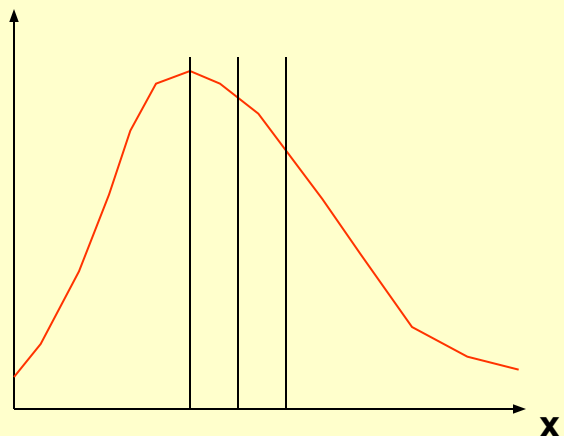
Показатель асимметрии

Для сравнительного анализа степени асимметрии нескольких распределений рассчитывается **относительный показатель асимметрии (коэффициент асимметрии)**:

$$A_x = \frac{\bar{x} - M_o}{\sigma}$$

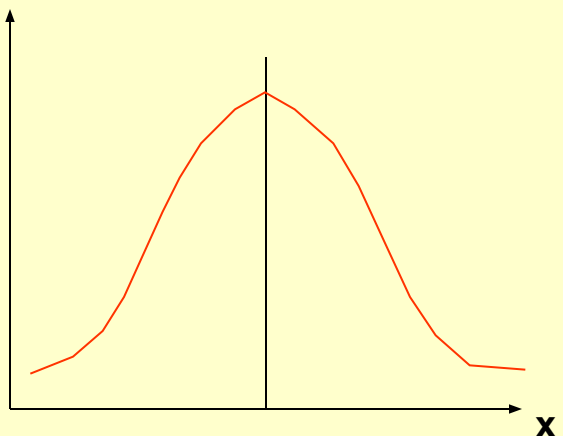
**Правосторонняя
асимметрия:**

$$A_x > 0 \quad \bar{x} > M_e > M_o$$



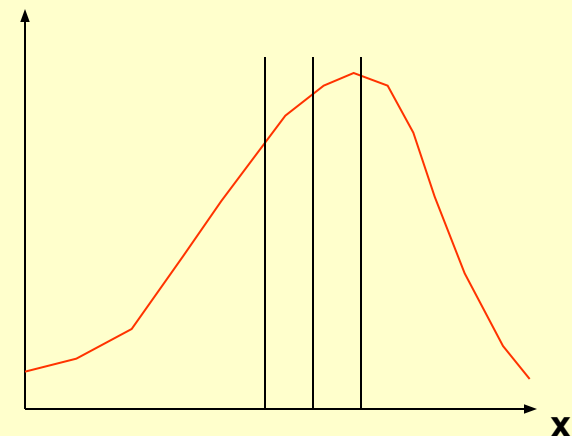
**Симметричная
кривая:**

$$A_x = 0 \quad \bar{x} = M_e = M_o$$



**Левосторонняя
асимметрия:**

$$A_x < 0 \quad \bar{x} < M_e < M_o$$



Показатели формы распределения

Показатель асимметрии

Наиболее распространенным является показатель асимметрии, вычисляемый по формуле:

$$A_x = \frac{\mu_3}{\sigma^3}$$

где μ_3 - центральный момент третьего порядка:

$$\mu_3 = \frac{\sum_{i=1}^n (x_i - \bar{x})^3 f_i}{\sum_{i=1}^n f_i}$$

Показатели формы распределения

Показатель эксцесса

Для симметричных распределений рассчитывается показатель эксцесса (островершинности) :

$$E_x = \frac{\mu_4}{\sigma^4} - 3$$

где μ_4 - центральный момент четвертого порядка:

$$\mu_4 = \frac{\sum_{i=1}^n (x_i - \bar{x})^4 f_i}{\sum_{i=1}^n f_i}$$

Построение дискретного ряда распределения

Пример №1. Имеются следующие данные о квалификации рабочих цеха: тарифные разряды 24 рабочих – 4; 3; 6; 4; 4; 2; 3; 5; 4; 4; 5; 2; 3; 4; 4; 5; 2; 3; 6; 5; 4; 2; 4; 3. Требуется: 1) построить дискретный ряд распределения; 2) дать графическое изображение ряда; 3) вычислить показатели центра распределения, показатели вариации и формы распределения.

Решение.

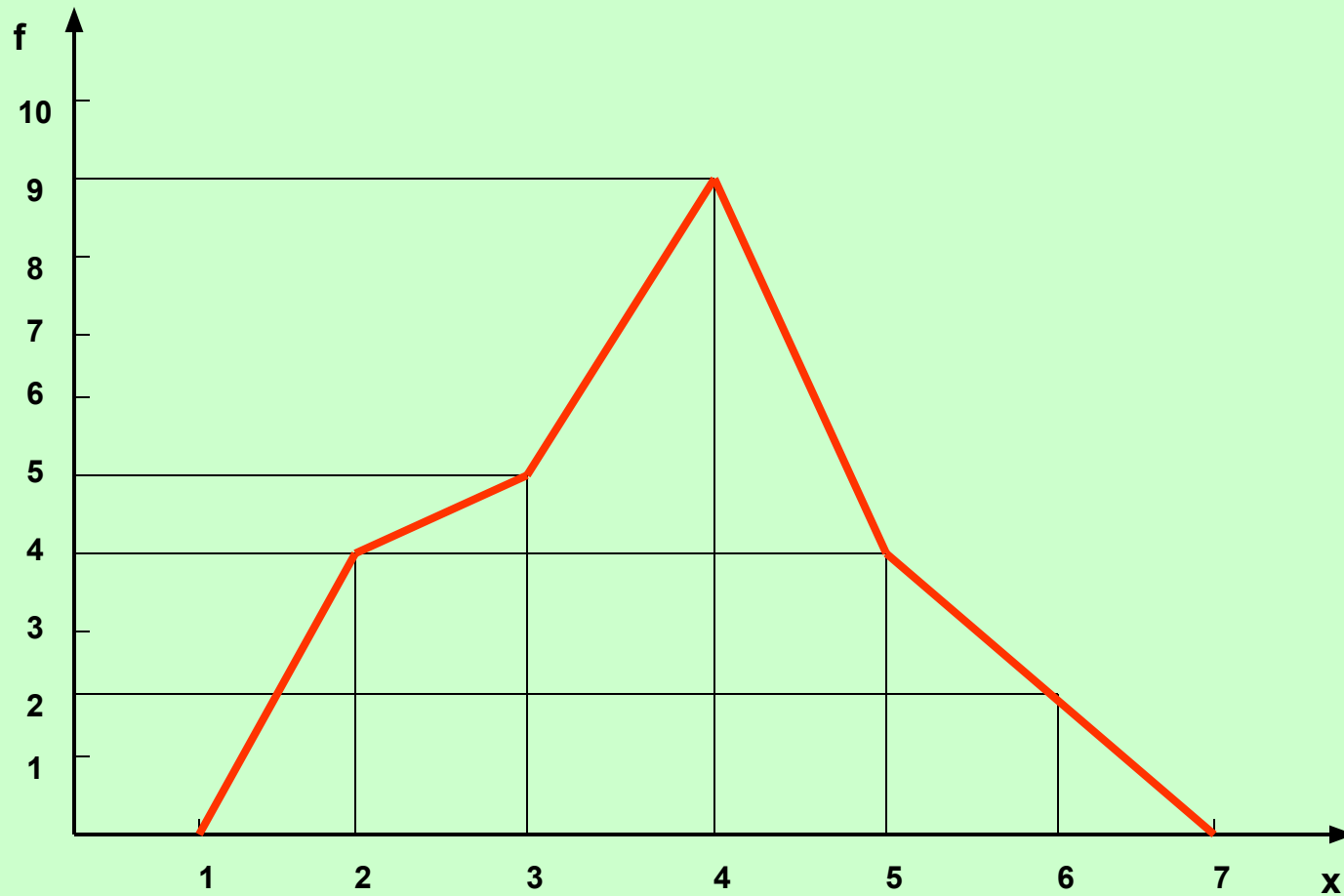
1). Дискретный ряд распределения имеет вид:

Тарифный разряд, x	Число рабочих, f	Накопленная частота
2	4	4
3	5	9
4	9	18
5	4	22
6	2	24
Итого	24	-

Дискретный ряд распределения

Полигон частот

2). Дискретный вариационный ряд в виде полигона частот:



Дискретный ряд распределения

Показатели центра распределения

3). Расчет показателей центра распределения.

3.1). Средняя арифметическая:

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i} = \frac{2 \cdot 4 + 3 \cdot 5 + 4 \cdot 9 + 5 \cdot 4 + 6 \cdot 2}{4 + 5 + 9 + 4 + 2} = \frac{91}{24} = 3,8 \text{ разр}$$

3.2). Мода:

$$M_o = 4 \text{ разр}$$

3.3). Медиана:

$$N_{M_e} = \frac{24+1}{2} = 12,5; \quad M_e = 4 \text{ разр}$$

Дискретный ряд распределения

Показатели вариации

4). Расчет показателей вариации.

4.1). Среднее линейное отклонение:

$$d = \pm \frac{\sum_{i=1}^n |x_i - \bar{x}| f_i}{\sum_{i=1}^n f_i} = \pm \frac{22,2}{24} = \pm 0,9 \text{ разр}$$

4.2). Среднее квадратическое отклонение:

$$\sigma = \pm \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 f_i}{\sum_{i=1}^n f_i}} = \pm \sqrt{\frac{31,96}{24}} = \pm 1,2 \text{ разр}$$

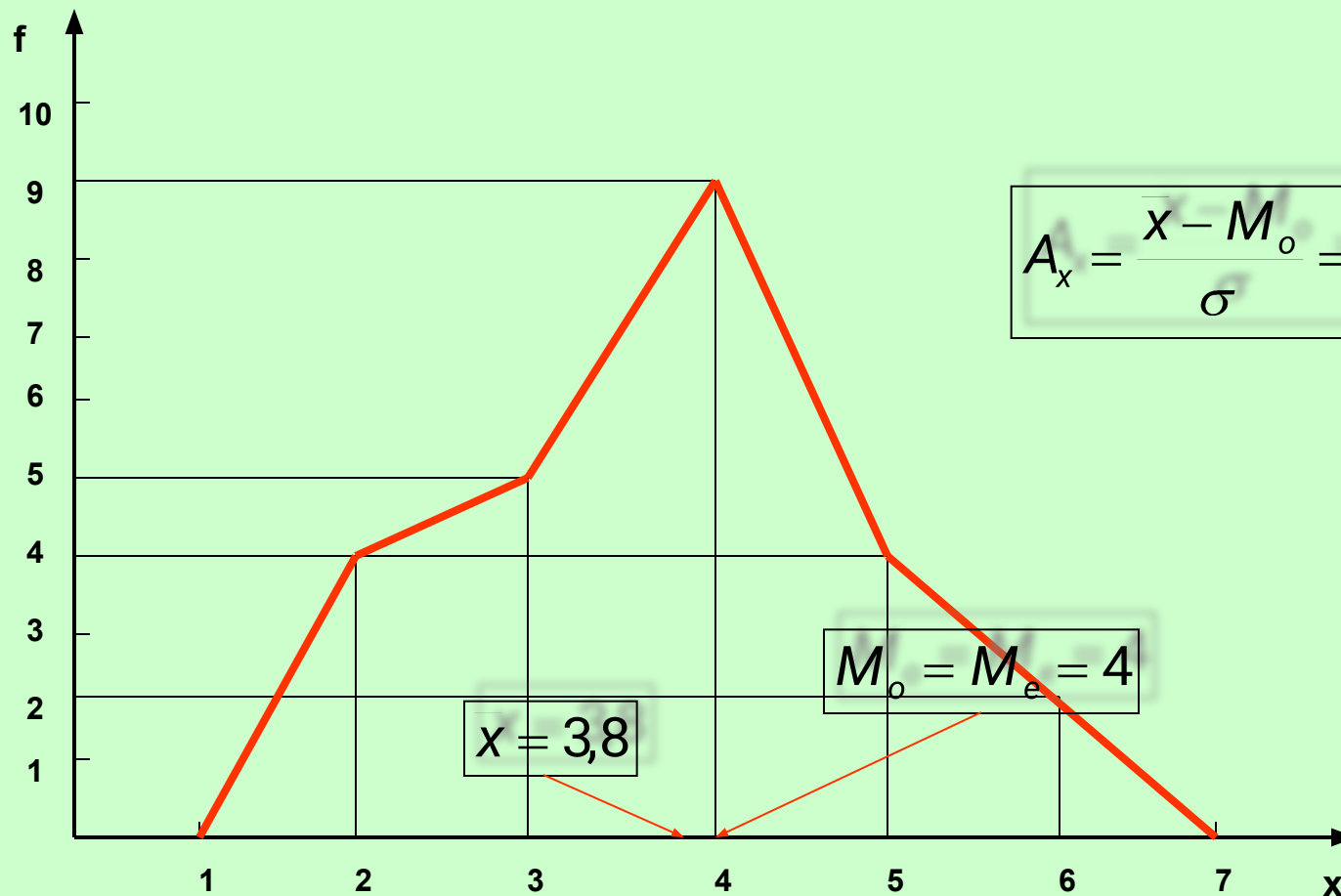
4.3). Коэффициент вариации:

$$V = \frac{\sigma}{\bar{x}} \cdot 100 = \frac{1,2}{3,8} \cdot 100 = 31,6\%$$

Дискретный ряд распределения

Показатель асимметрии

5). Показатель асимметрии:



$$A_x = \frac{x - M_o}{\sigma} = \frac{3,8 - 4}{1,2} = -0,17$$

$$M_o = M_e = 4$$

$$x = 3,8$$

Построение интервального ряда распределения

Пример №2. Имеются следующие данные о возрастном составе рабочих цеха (лет): 18; 38; 28; 29; 26; 38; 34; 22; 28; 30; 22; 23; 35; 33; 27; 24; 30; 32; 28; 25; 29; 26; 31; 24; 29; 27; 32; 25; 29; 29. Требуется: 1) построить интервальный ряд распределения; 2) дать графическое изображение ряда; 3) вычислить показатели центра распределения, показатели вариации и формы распределения.

Решение.

1). Интервальный ряд распределения.

Величина интервала группировки:

$$i = \frac{R}{m}$$

где m – число групп, приближенно определяется по формуле Стерджеса:

$$m = 1 + 3,322 \cdot \lg(n)$$

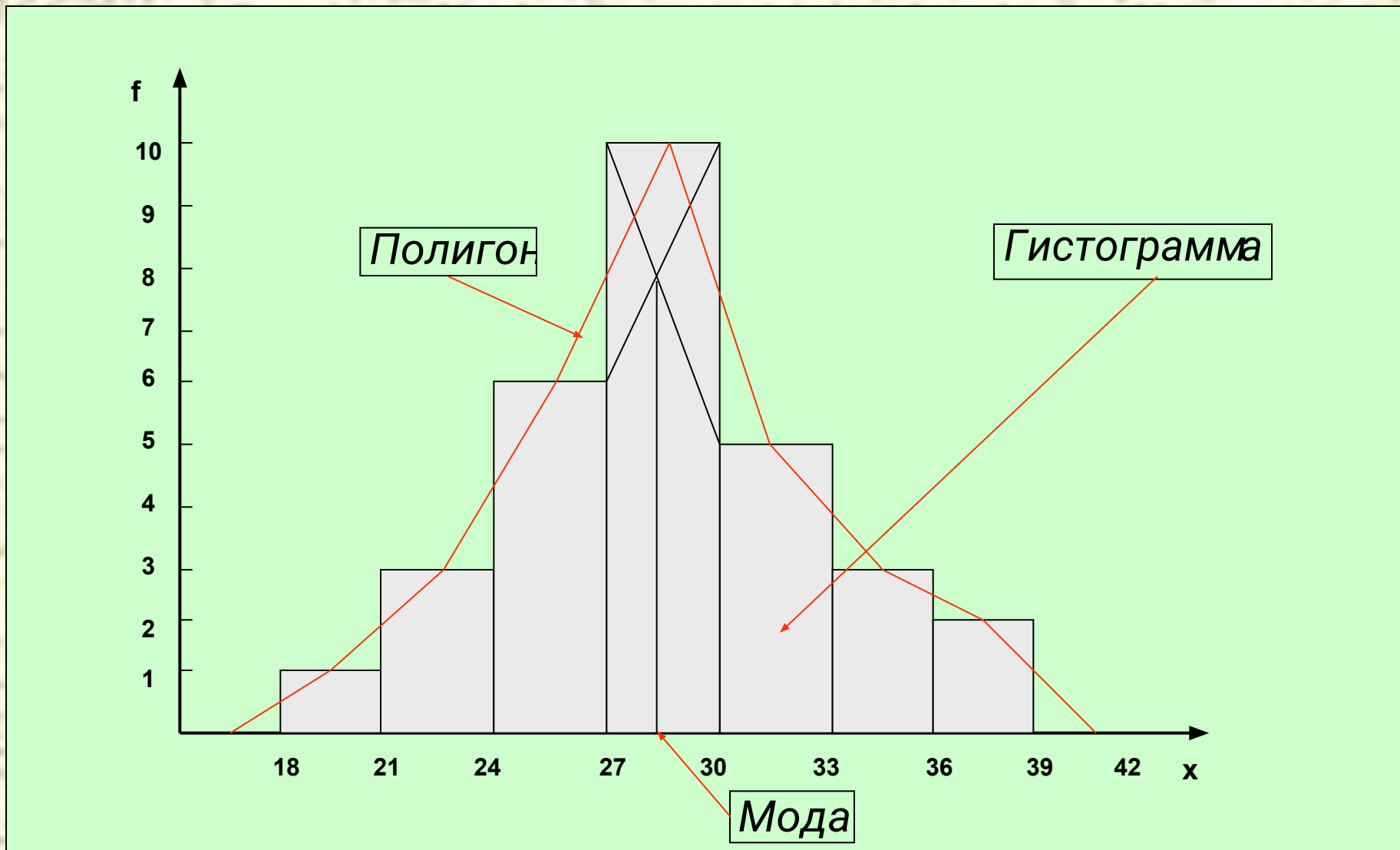
$$m = 1 + 3,322 \cdot \lg(30) \cong 6$$

$$i = \frac{38 - 18}{6} \cong 3 \text{ года}$$

Группы рабочих по возрасту, x	Число рабочих, f	Накопленная частота
18-21	1	1
21-24	3	4
24-27	6	10
27-30	10	20
30-33	5	25
33-36	3	28
36-39	2	30
Итого	30	-

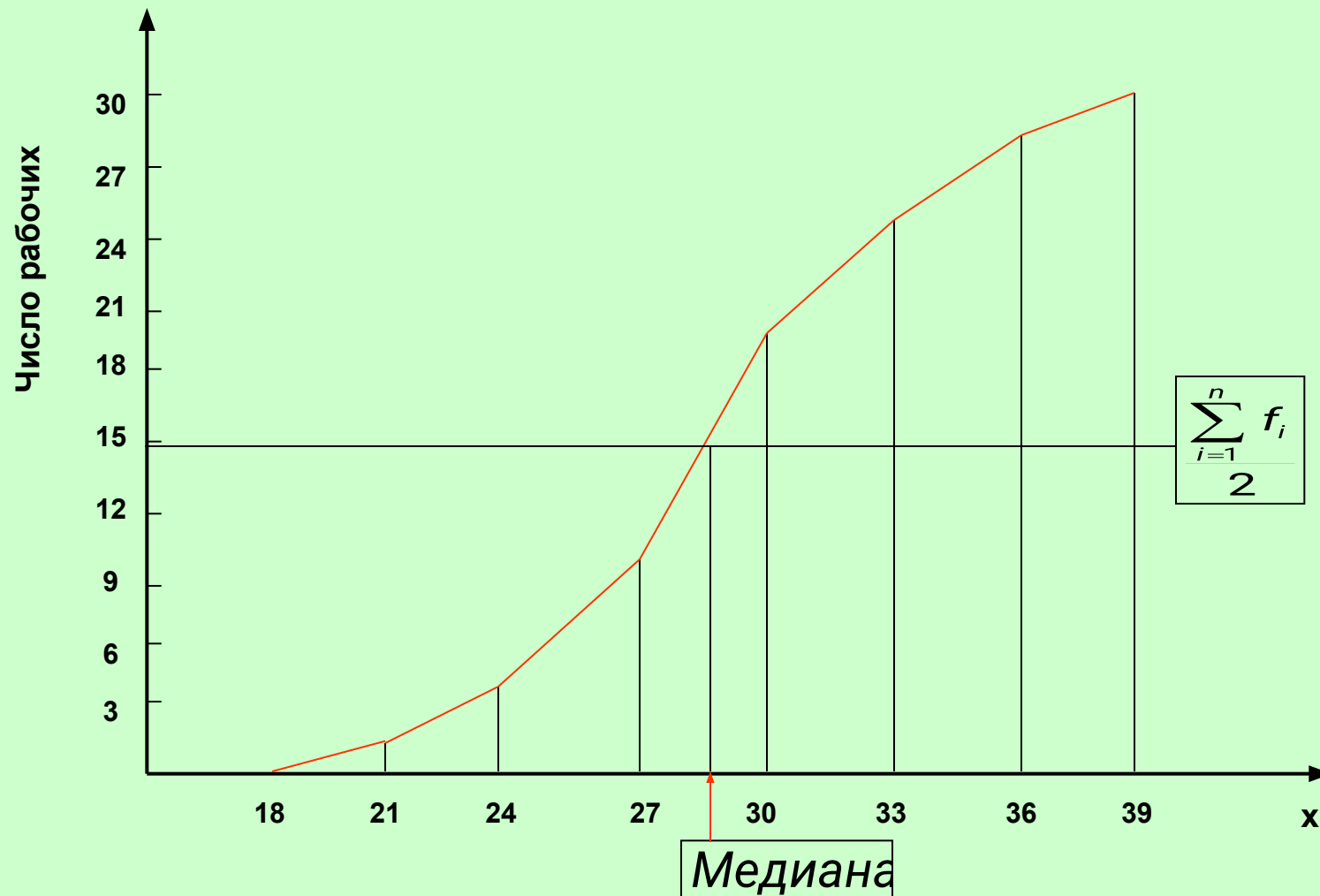
Интервальный ряд распределения

Построение гистограммы и полигона



Интервальный ряд распределения

Построение кумуляты



Интервальный ряд распределения

Показатели центра распределения

3). Расчет показателей центра распределения.

3.1). Средняя арифметическая:

$$\bar{x} = \frac{\sum_{i=1}^n x_i' f_i}{\sum_{i=1}^n f_i} = \frac{19,5 \cdot 1 + 22,5 \cdot 3 + 25,5 \cdot 6 + 28,5 \cdot 10 + 31,5 \cdot 5 + 34,5 \cdot 3 + 37,5 \cdot 2}{4 + 5 + 9 + 4 + 2} = \frac{861}{30} = 28,7 \text{ года}$$

3.2). Мода:

$$M_o = x_{mo} + i \frac{f_2 - f_1}{(f_2 - f_1) + (f_2 - f_3)}$$

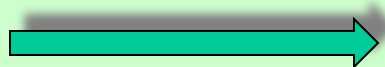
$$M_o = 27 + 3 \frac{10 - 6}{(10 - 6) + (10 - 5)} = 28,33 \text{ года}$$

Интервальный ряд распределения

Показатели центра распределения

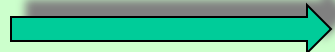
3.3). Медиана:

$$N_{M_e} = \frac{n+1}{2}$$



$$N_{M_e} = \frac{30+1}{2} = 15,5$$

$$M_e = x_o + i \frac{\frac{\sum_{i=1}^n f_i}{2} - S_{m-1}}{f_m}$$



$$M_e = 27 + 3 \frac{\frac{30}{2} - 10}{10} = 28,5 \text{ года}$$

Интервальный ряд распределения

Показатели вариации

Вспомогательная таблица для расчета показателей вариации.

Группы рабочих по возрасту	Центр интервала x'	f	$x' f$	d	$ d f$	d^2	$d^2 f$
18-21	19,5	1	19,5	-9,2	9,2	84,64	84,64
21-24	22,5	3	67,5	-6,2	18,6	38,44	115,32
24-27	25,5	6	153,0	-3,2	19,2	10,24	61,44
27-30	28,5	10	285,0	-0,2	20,0	0,04	0,40
30-33	31,5	5	157,5	2,8	14,0	7,84	39,20
33-36	34,5	3	103,5	5,8	17,4	33,64	100,92
36-39	37,5	2	75,0	8,8	17,6	77,44	154,88
Сумма	-	30	861,0	-	116,0	-	556,80

Интервальный ряд распределения

Показатели вариации

4). Расчет показателей вариации.

4.1). Среднее линейное отклонение:

$$d = \pm \frac{\sum_{i=1}^n |x_i' - \bar{x}| f_i}{\sum_{i=1}^n f_i} = \pm \frac{1160}{30} = \pm 3.87 \text{ года}$$

4.2). Среднее квадратическое отклонение:

$$\sigma = \pm \sqrt{\frac{\sum_{i=1}^n (x_i' - \bar{x})^2 f_i}{\sum_{i=1}^n f_i}} = \pm \sqrt{\frac{5568}{30}} = \pm 4,31 \text{ года}$$

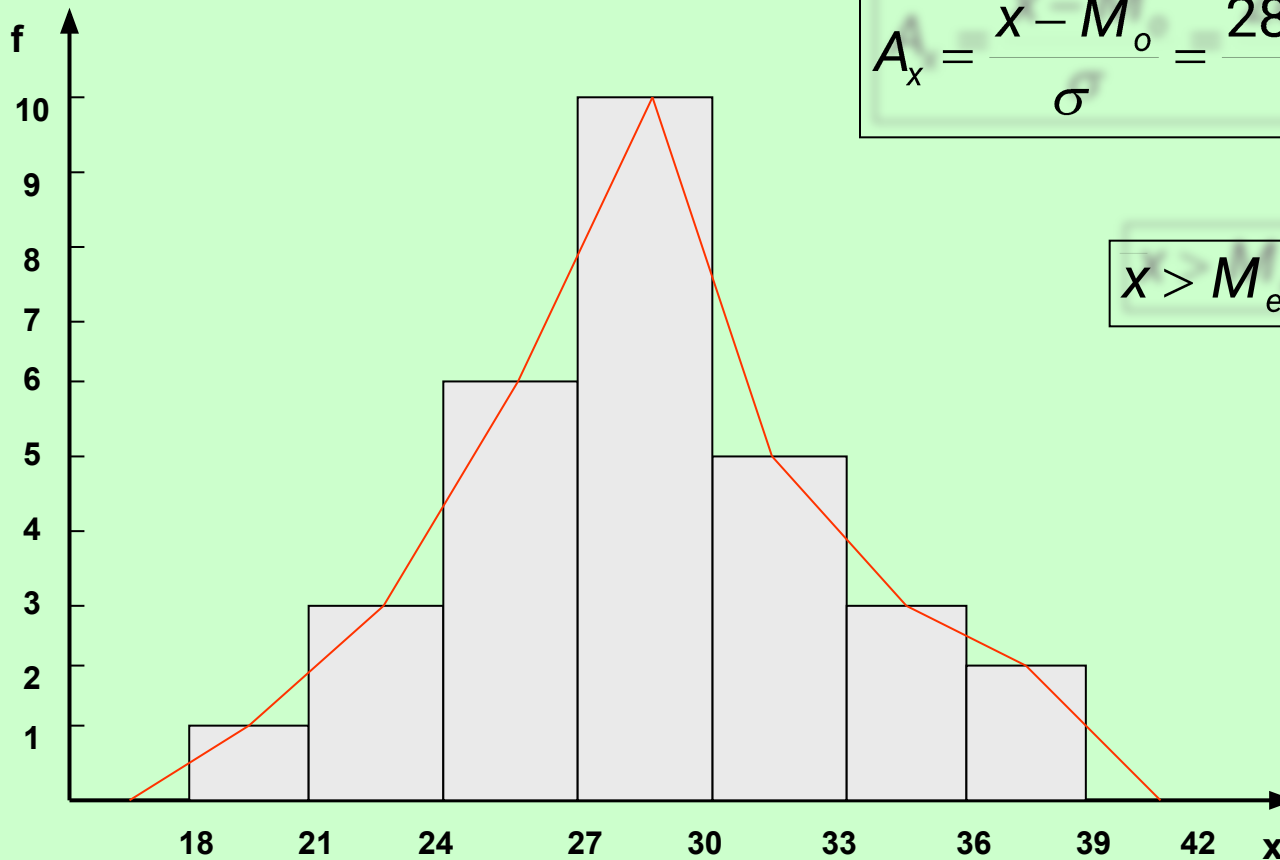
4.3). Коэффициент вариации:

$$V = \frac{\sigma}{\bar{x}} \cdot 100 = \frac{4,31}{28,7} \cdot 100 = 15,0\%$$

Интервальный ряд распределения

Показатель асимметрии

5). Показатель асимметрии:



$$A_x = \frac{\bar{x} - M_o}{\sigma} = \frac{28,7 - 28,3}{4,31} = +0,09$$

$$x > M_e > M_o$$

Интервальный ряд распределения

Показатель эксцесса

б). Показатель эксцесса:

$$E_x = \frac{\mu_4}{\sigma^4} - 3 \longrightarrow E_x = \frac{9307}{4,31^4} - 3 = -0,3$$

$$\mu_4 = \frac{\sum_{i=1}^n (x_i' - \bar{x})^4 f_i}{\sum_{i=1}^n f_i} \longrightarrow \mu_4 = 9307$$