

Статистические данные

Количественные

Непрерывные

вес
рост
суточная
производительность
.....

Дискретные

число детей
год выпуска
число проданной
продукции в день
.....

Качественные

Номинальные (классификационные)

Одно значение из
установленного числа
(пол, цвет, марка
автомобиля ...)

Ординарные (порядковые)

Упорядоченность
(тестовые баллы,
школьные оценки,
качество жизни)

Генеральная
(Множество изменений
совокупность
случайной величины X)

Должна быть
РЕПРЕЗЕНТАТИВНОЙ

Выборка

Характеристики
(статистики):
среднее
дисперсия
....

Приблизительные
характеристики
(статистики):
выборочное среднее
Выборочная
дисперсия
....

5, 3, 7, 10, 5, 5, 2, 10, 7, 2, 7, 7, 4, 2, 4.

Объем выборки $n = 15$.

Упорядочив элементы выборки по величине, получим вариационный ряд:

2, 2, 2, 3, 4, 4, 5, 5, 5, 7, 7, 7, 7, 10, 10.

Статистический ряд для примера

Элементы, x_i	2	3	4	5	7	10
Штрихи	///	/	//	///	////	//
Частота, n_i	3	1	2	3	4	2

Представить выборку из 55 наблюдений в виде таблицы частот, используя 7 интервалов группировки.

Выборка:

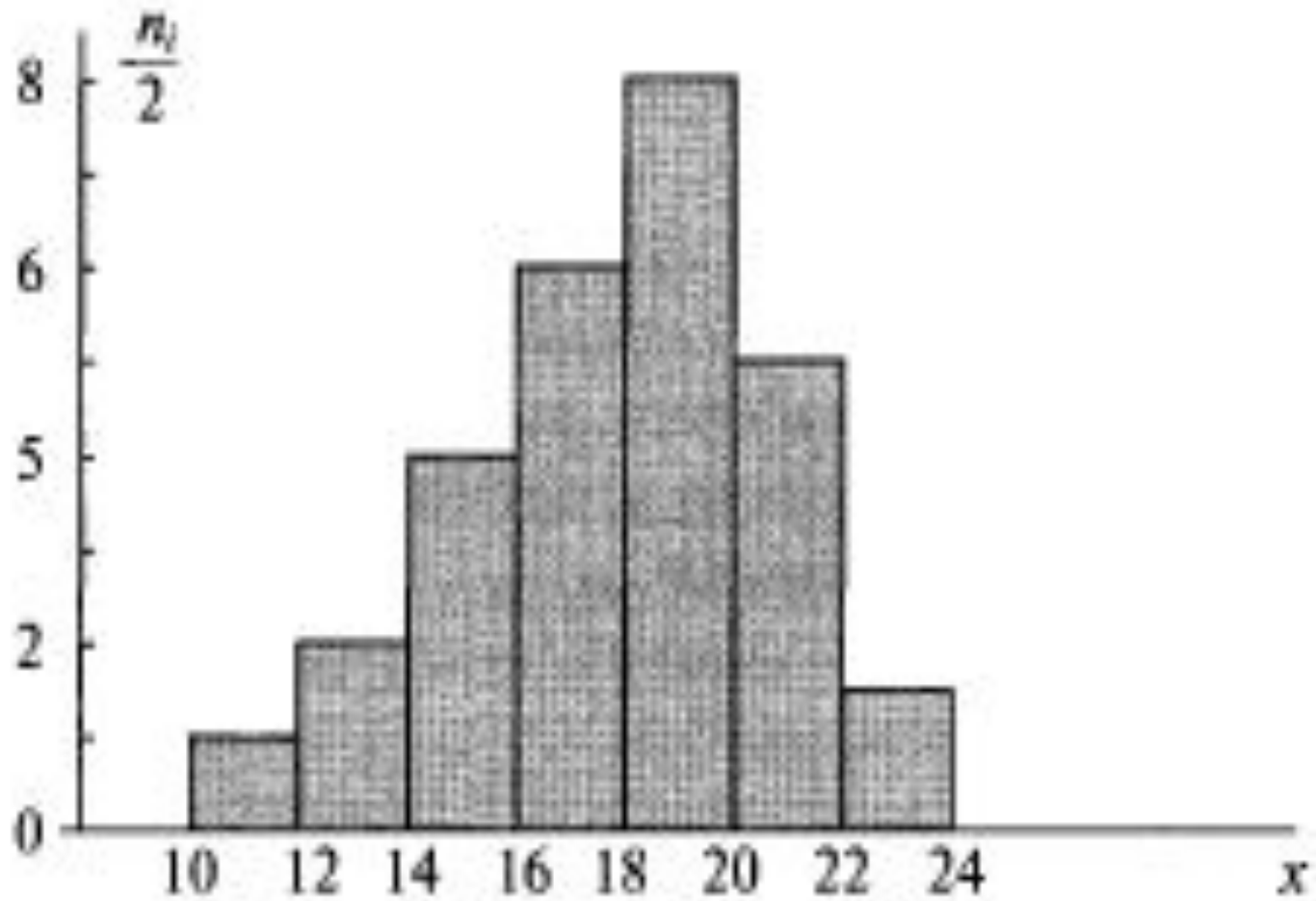
18,3	15,4	17,2	19,2	23,3	18,1	21,9
15,3	16,8	13,2	20,4	16,5	19,7	20,5
14,3	20,1	16,8	14,7	20,8	19,5	15,3
19,3	17,8	16,2	15,7	22,8	21,9	12,5
10,1	21,1	18,3	14,7	14,5	18,5	18,4
13,9	19,1	18,5	20,2	23,8	16,7	20,4
19,5	17,2	19,6	17,8	21,3	17,5	19,4
17,8	13,5	17,8	11,8	18,6	19,1	

Решение. Размах выборки $R = 23,8 - 10,1 = 13,7$. Длина интервала группировки $b = 13,7/7 \approx 2$. В качестве первого интервала удобно взять интервал 10—12.

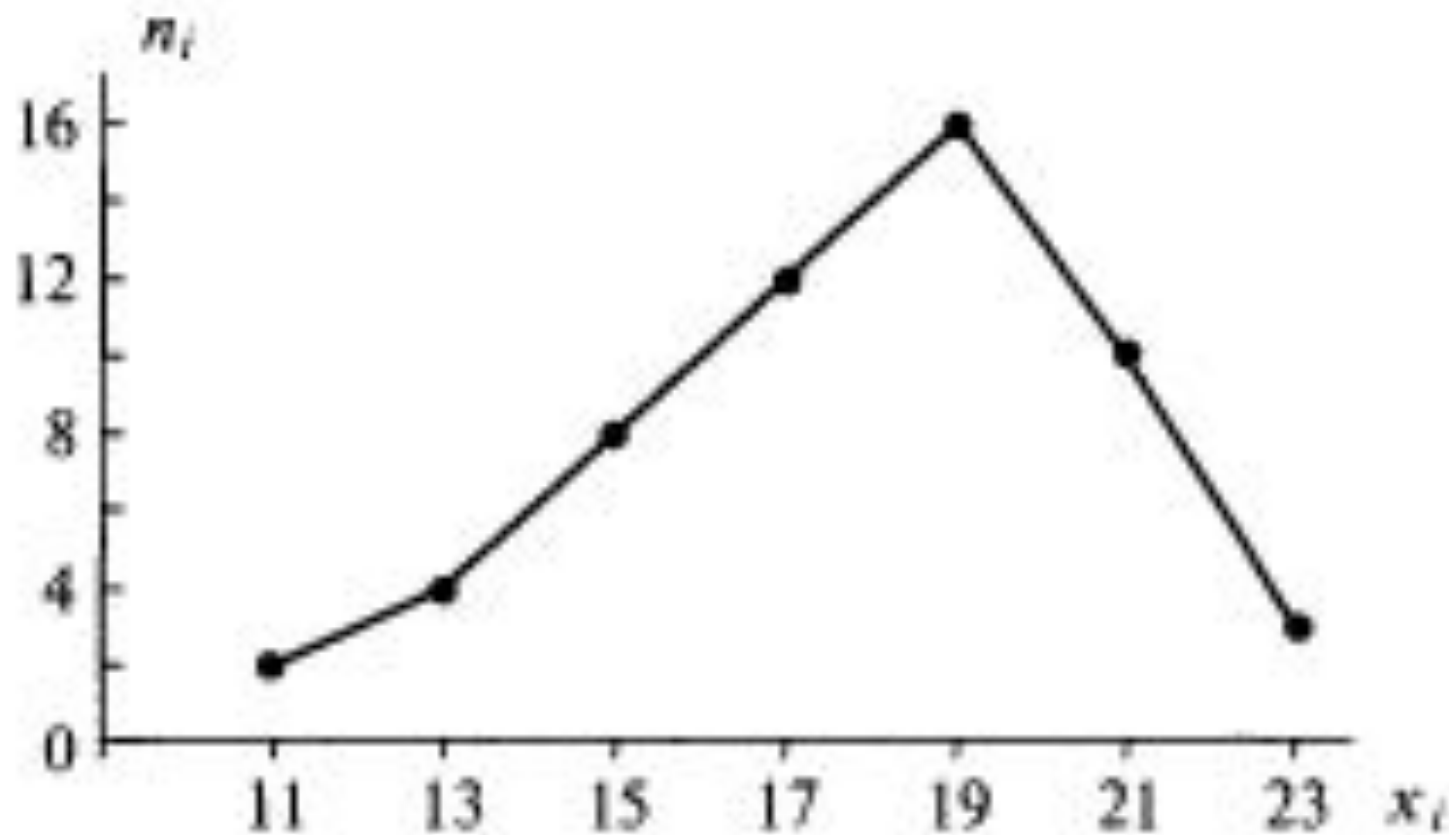
Таблица 3.2. Таблица частот для примера 3.2

Номер интервала, i	Границы интервала	Середина интервала, x_i	Частота, n_i	Накопленная частота, $\sum_{j=1}^i n_j$	Относительная частота, n_i/n	Накопленная относительная частота, $\sum_{j=1}^i n_j/n$
1	10—12	11	2	2	0,0364	0,0364
2	12—14	13	4	6	0,0727	0,1091
3	14—16	15	8	14	0,1455	0,2546
4	16—18	17	12	26	0,2182	0,4728
5	18—20	19	16	42	0,2909	0,7637
6	20—22	21	10	52	0,1818	0,9455
7	22—24	23	3	55	0,0545	1,0000

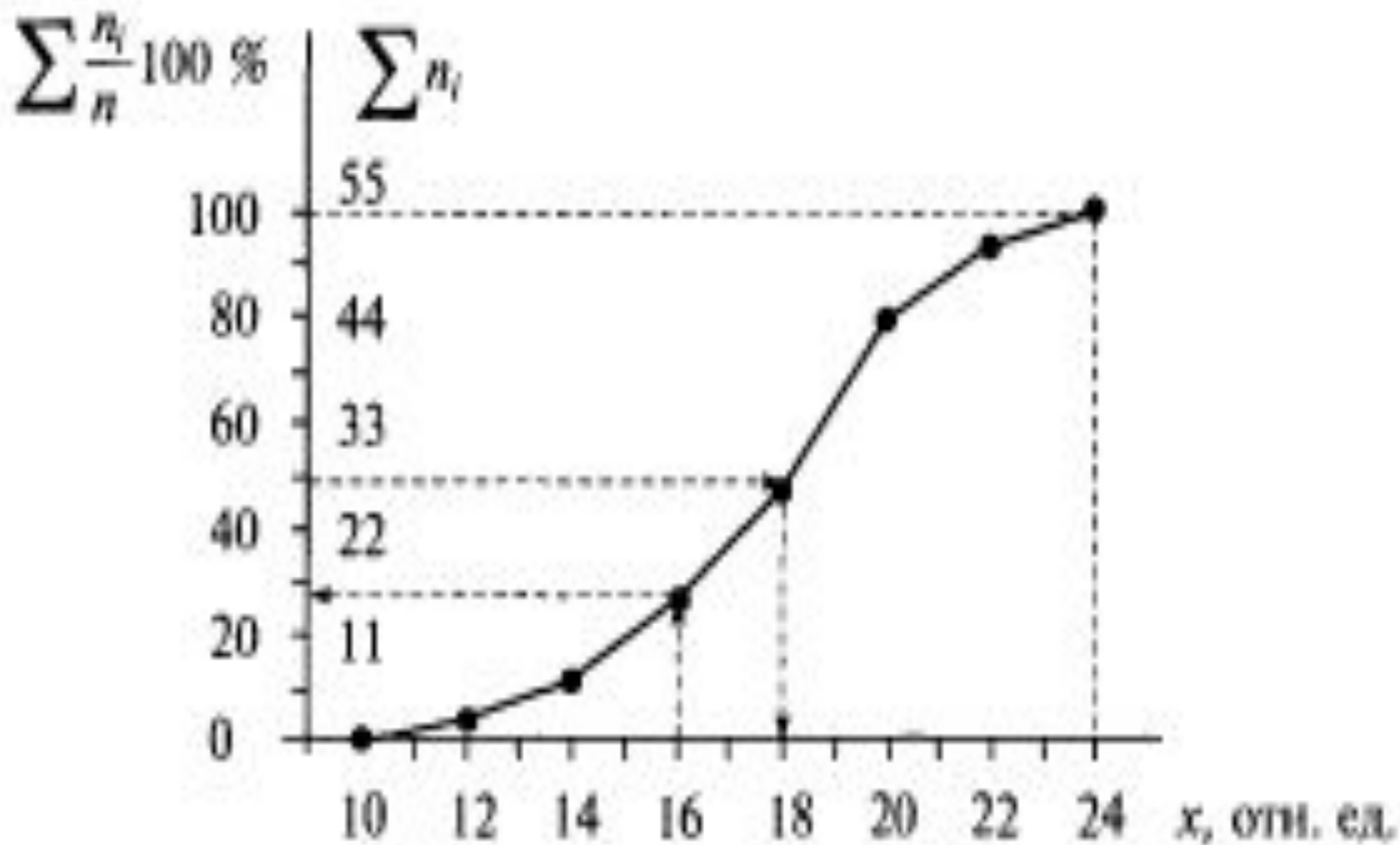
ГИСТОГРАММА



ПОЛИГОН ЧАСТОТ



ОГНИВА



ЭМПИРИЧЕСКАЯ ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ

Эмпирическая функция распределения $F_n^*(x)$ — это ступенчатая функция, имеющая скачки, кратные числу $1/n$ в точках $x^{(1)}, x^{(2)}, \dots, x^{(n)}$:

$$F_n^*(x) = \begin{cases} 0, & x \leq x^{(1)} \\ k/n, & x^{(k)} < x \leq x^{(k+1)}, \quad k = 1, 2, \dots, n-1 \\ 1, & x > x^{(n)}. \end{cases}$$

Эмпирическая функция распределения $F_n^*(x)$ является статистическим аналогом функции распределения $F(x)$ генеральной совокупности.

ВЫБОРОЧНЫЙ НАЧАЛЬНЫЙ МОМЕНТ __-ГО ПОРЯДКА

$$\alpha_r^* = \frac{1}{n} \sum_{i=1}^n x_i^r, \quad r = 1, 2, \dots \quad \bar{x} = \alpha_1^* = \frac{1}{n} \sum_{i=1}^n x_i$$

ВЫБОРОЧНЫЙ ЦЕНТРАЛЬНЫЙ МОМЕНТ __-ГО ПОРЯДКА

$$\mu_r^* = (1/n) \sum_{i=1}^n (x_i - \bar{x})^r,$$

**ДЛЯ
ВЫБОРКИ**

x_1, x_2, \dots, x_n

**КАК СТАТИСТИЧЕСКОГО
РЯДА**

$$\alpha_r^* = \frac{1}{n} \sum_{i=1}^k n_i x_i^r,$$

$$\mu_r^* = \frac{1}{n} \sum_{i=1}^k n_i (x_i - \bar{x})^r, \quad r = 1, 2, \dots$$

СВЯЗЬ НАЧАЛЬНЫХ И ЦЕНТРАЛЬНЫХ МОМЕНТОВ

$$\mu_2^* = \alpha_2^* - \bar{x}^2,$$

$$\mu_3^* = \alpha_3^* - 3\alpha_2^*\bar{x} + 2\bar{x}^3,$$

$$\mu_4^* = \alpha_4^* - 4\alpha_3^*\bar{x} + 6\alpha_2^*\bar{x}^2 - 3\bar{x}^4 \text{ и т. д.}$$

КВАНТИЛИ

Выборочная квантиль x_p порядка p , $0 < p < 1$, определяется как элемент вариационного ряда выборки $x^{(1)}, x^{(2)}, \dots, x^{(n)}$ с номером $[np] + 1$, где $[a]$ — целая часть a .

В статистической практике используют ряд квантилей, имеющих специальные названия:

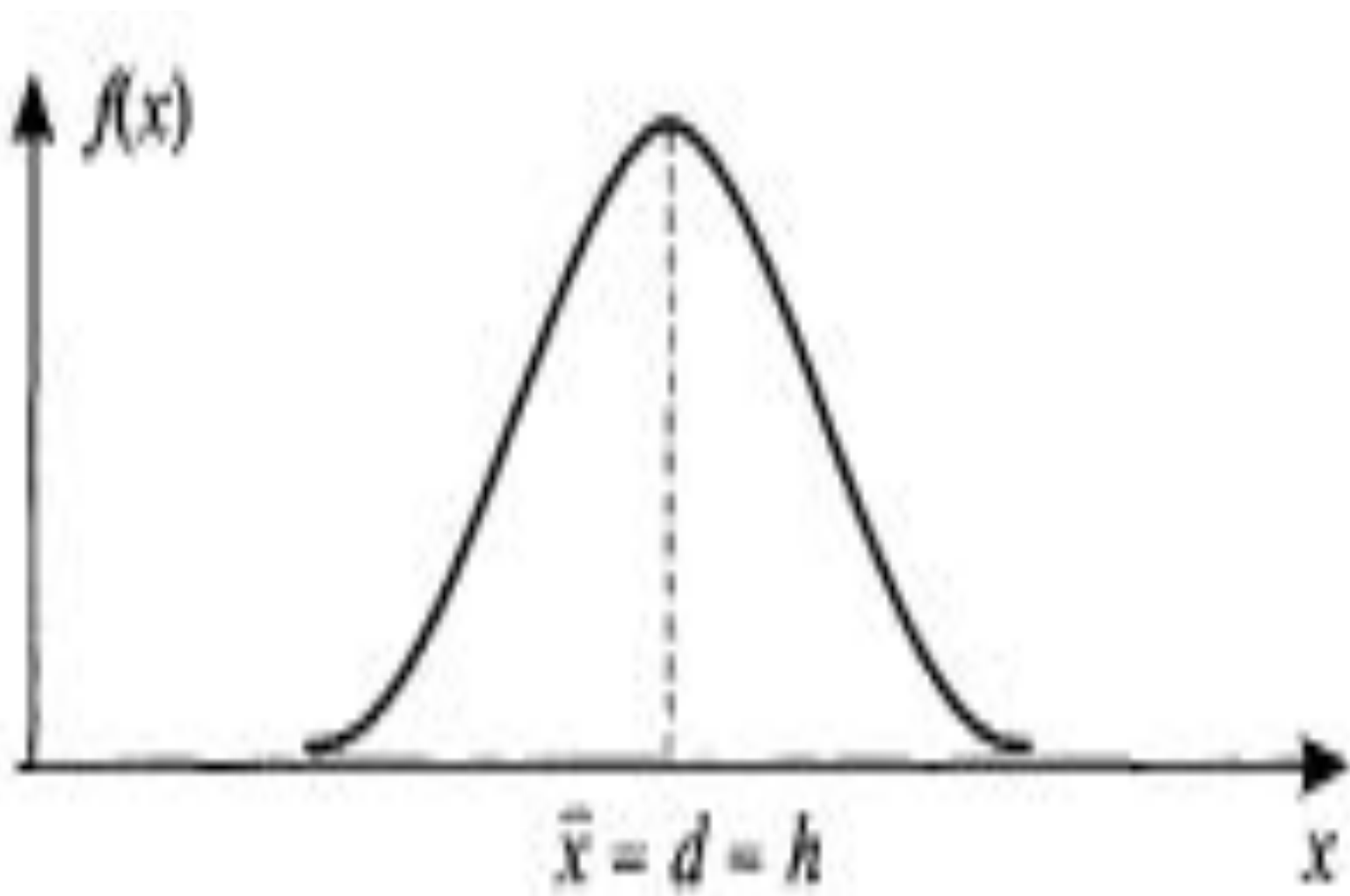
персентили: P_1, P_2, \dots, P_{99} — квантили порядков 0,01; 0,02; ...; 0,99;

децили: D_1, D_2, \dots, D_9 — квантили порядков 0,10; 0,20; ...; 0,90;

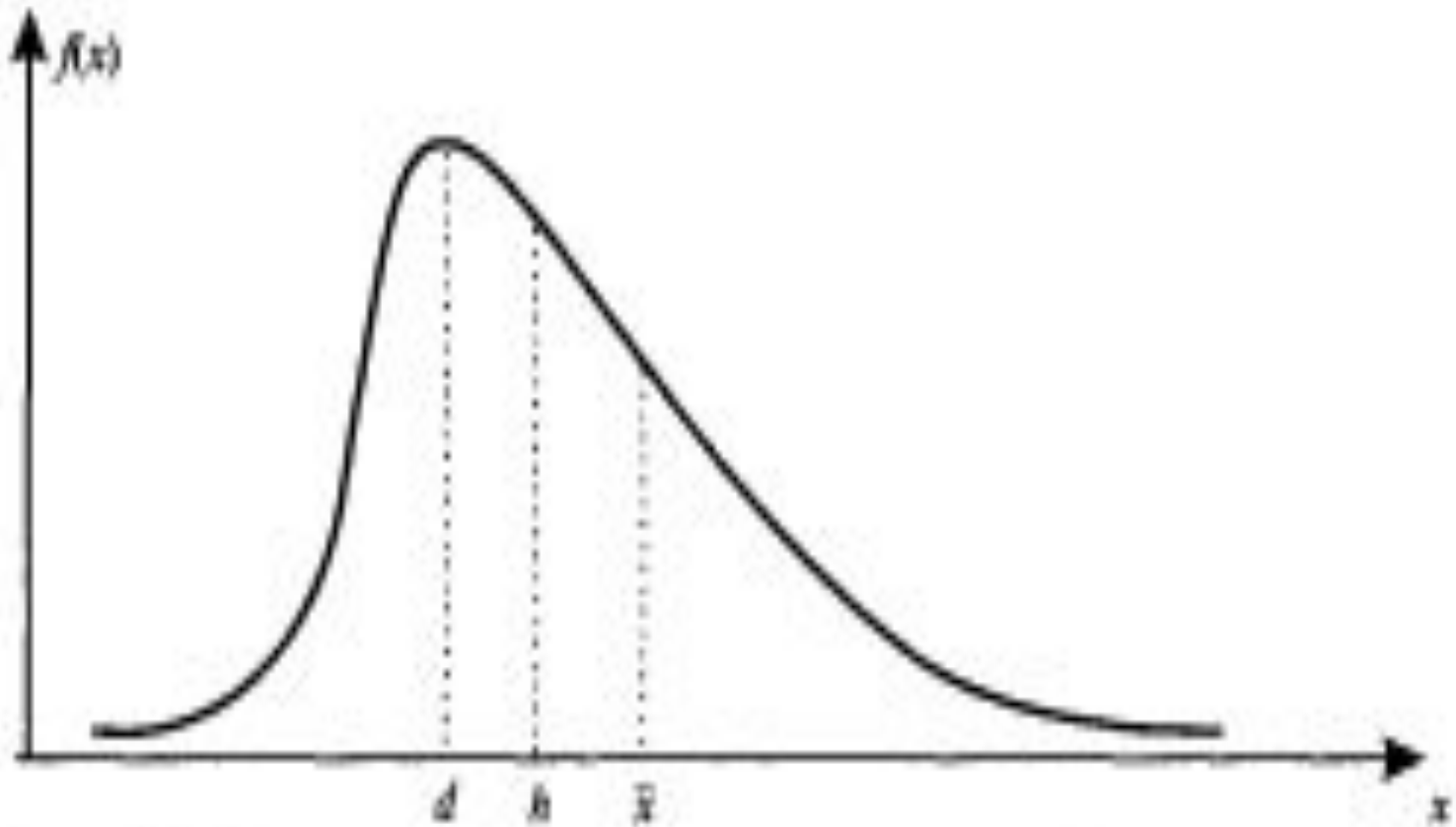
квартили: Q_1, Q_2, Q_3 — квантили порядков 0,25; 0,50; 0,75.

Для сгруппированных выборок оценки выборочных квантилей можно определить по *огиве* |

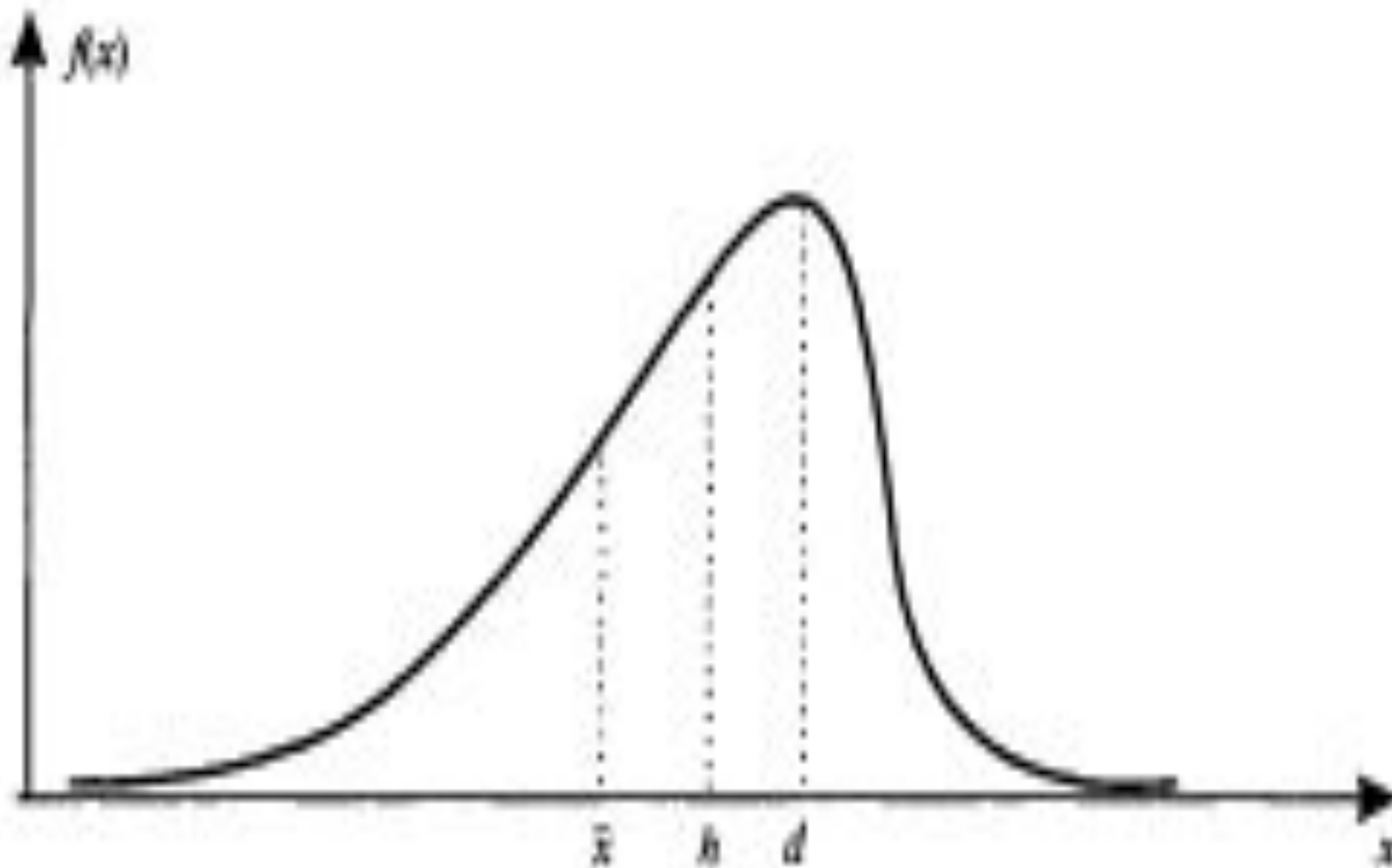
СИММЕТРИЧНОЕ УНИМОДАЛЬНОЕ РАСПРЕДЕЛЕНИЕ



**СКОШЕННОЕ СПРАВА РАСПРЕДЕЛЕНИЕ
(АССИМЕТРИЯ > 0)**



СКОШЕННОЕ СЛЕВА РАСПРЕДЕЛЕНИЕ



Для выборки объема n : x_1, x_2, \dots, x_n

СРЕДНЕЕ АРИФМЕТИЧЕСКОЕ

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

СРЕДНЕЕ ГЕОМЕТРИЧЕСКОЕ

$$x_g = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$$

СВЯЗЬ

$$\sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} \leq \frac{x_1 + x_2 + \dots + x_n}{n}$$

СРЕДНЕЕ ГАРМОНИЧЕСКОЕ x_H

$$\frac{1}{x_H} = \frac{1}{n} \sum_{i=1}^n \frac{1}{x_i}$$

СВЯЗЬ

$$\bar{x} \geq x_g \geq x_H$$

ОЦЕНКА ХАРАКТЕРИСТИК РАССЕЯНИЯ

РАЗМАХ

$$R = x_{\max} - x_{\min} = x^{(n)} - x^{(1)}$$



СРЕДНИЙ МЕЖКВАРТИЛЬНЫЙ РАЗМАХ

$$\frac{Q_3 - Q_1}{2}$$

ПЕРСЕНТИЛЬНЫЙ РАЗМАХ

$$P_{90} - P_{10} = x_{0,9} - x_{0,1}$$

ДИСПЕРСИЯ

ДЛЯ ВЫБОРКИ НЕГРУППИРОВАННЫХ

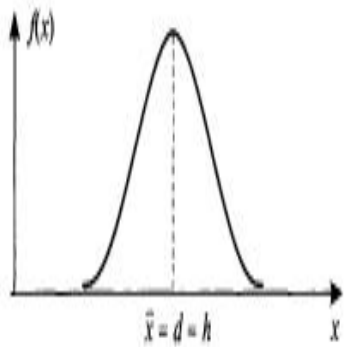
$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - n \cdot \bar{x}^2}{n-1}$$

ДЛЯ ВЫБОРКИ СГРУППИРОВАННЫХ

$$S^2 = \frac{\sum n_i (x_i - \bar{x})^2}{n-1} = \frac{\sum n_i x_i^2 - n \bar{x}^2}{n-1}$$

КОЭФФИЦИЕНТ ВАРИАЦИИ $V = \frac{S}{\bar{x}}$ или $C_V = \frac{S}{\bar{x}} \cdot 100 \%$.

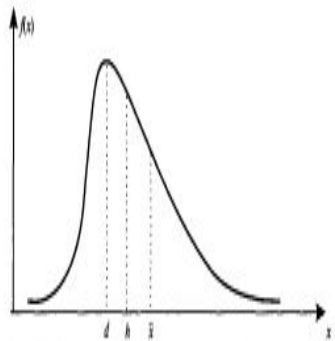
ОЦЕНКА ХАРАКТЕРИСТИК ФОРМЫ РАСПРЕДЕЛЕНИЯ



АССИММЕТРИЯ

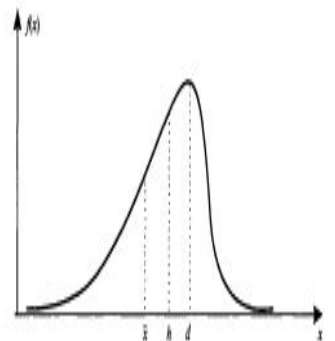
$$S_{K1} = \frac{\mu_3^*}{\sigma^3}$$

$$S_{K2} = \frac{x_{0,75} + x_{0,25} - 2 \cdot x_{0,5}}{x_{0,75} - x_{0,25}} = \frac{Q_3 + Q_1 - 2 \cdot Q_2}{Q_3 - Q_1}$$



КОЭФФИЦИЕНТ ЭКСЦЕССА

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



СПИСОК ВЫБОРОК

1	1	3	5	5	8	2	1	10	13	4	6	7	11	15	11	13	12	2	4	3	6	5	6	8	8
2	2	1	6	8	1	2	5	12	1	8	3	7	7	10	1	14	1	3	8	9	9	2	1	6	9
3	11	7	5	7	8	3	2	11	12	5	7	8	10	13	13	10	11	5	3	4	4	7	9	6	3
4	1	3	5	5	8	2	1	10	13	4	6	7	11	15	11	13	12	2	4	3	6	5	6	8	8
5	6	4	5	1	5	7	6	9	11	5	9	8	14	13	10	12	10	5	7	9	2	2	1	7	11
6	5	2	7	3	9	9	5	4	10	7	1	8	10	11	13	12	8	6	3	8	2	1	7	4	1
7	6	1	4	5	6	1	4	9	11	8	8	7	12	12	10	3	5	9	4	6	5	5	9	1	2
8	1	4	1	7	8	5	6	11	12	4	3	3	10	7	10	11	13	3	5	7	5	9	6	1	5
9	8	3	5	9	9	4	6	13	13	8	9	5	13	11	10	13	10	9	7	5	8	5	3	3	2
10	9	5	7	4	9	1	7	4	11	8	6	3	8	10	14	15	12	8	4	2	2	1	9	5	5
11	7	3	9	1	8	2	1	12	15	2	5	7	15	3	10	11	11	4	5	3	8	5	4	3	8
12	3	6	9	4	4	7	2	8	9	1	8	5	4	13	10	9	11	7	3	8	8	3	4	1	9
13	4	5	6	5	6	5	1	11	11	4	5	5	10	15	11	13	10	2	1	5	9	3	2	9	1
14	10	8	1	6	4	3	3	8	10	8	3	4	12	15	1	11	15	8	4	7	4	5	5	6	2
15	11	1	7	1	3	9	7	12	14	5	7	8	15	15	11	12	13	9	5	9	3	8	8	4	5
16	15	6	9	3	9	3	9	15	15	7	2	1	13	15	11	15	11	4	7	2	1	6	9	3	7
17	8	7	3	7	1	7	4	11	12	2	9	4	10	15	11	14	12	2	9	5	3	4	3	1	6
18	9	4	2	9	6	4	6	13	11	5	8	9	14	15	11	12	13	9	5	2	7	3	2	7	4
19	4	2	8	8	7	7	2	14	14	7	4	6	10	15	11	10	15	8	3	4	4	2	5	4	3
20	6	6	1	7	1	9	5	5	10	3	2	4	14	15	11	13	11	3	2	3	5	6	8	7	9
21	9	7	3	5	6	2	6	13	12	6	1	7	10	15	11	14	14	2	7	9	8	1	3	3	1

ЗАДАНИЕ

- а) записать в виде вариационного ряда;**
- б) записать в виде статистического ряда;**
- в) определить размах каждой выборки;**
- г) записать в виде группированного статистического ряда;**
- д) определить накопленные частоты;**
- е) определить относительные частоты;**
- ж) определить накопленные относительные частоты;**
- з) сформировать таблицу частот группированной выборки;**
- и) построить гистограмму частот;**
- к) построить полигон частот;**
- л) построить огиву частот;**
- м) определить выборочный начальный момент первого порядка;**
- н) определить моду каждой выборки;**
- о) определить медиану каждой выборки;**
- п) определить среднее геометрическое каждой выборки;**
- р) определить среднее гармоническое каждой выборки;**
- с) определить межквартильный размах каждой выборки;**
- т) определить перцентильный размах каждой выборки;**
- у) определить дисперсию каждой выборки;**
- ф) определить среднее квадратическое отклонение каждой выборки;**
- ц) определить коэффициент вариации каждой выборки;**
- ч) определить коэффициент асимметрии каждой выборки;**
- ш) определить коэффициент эксцесса каждой выборки.**