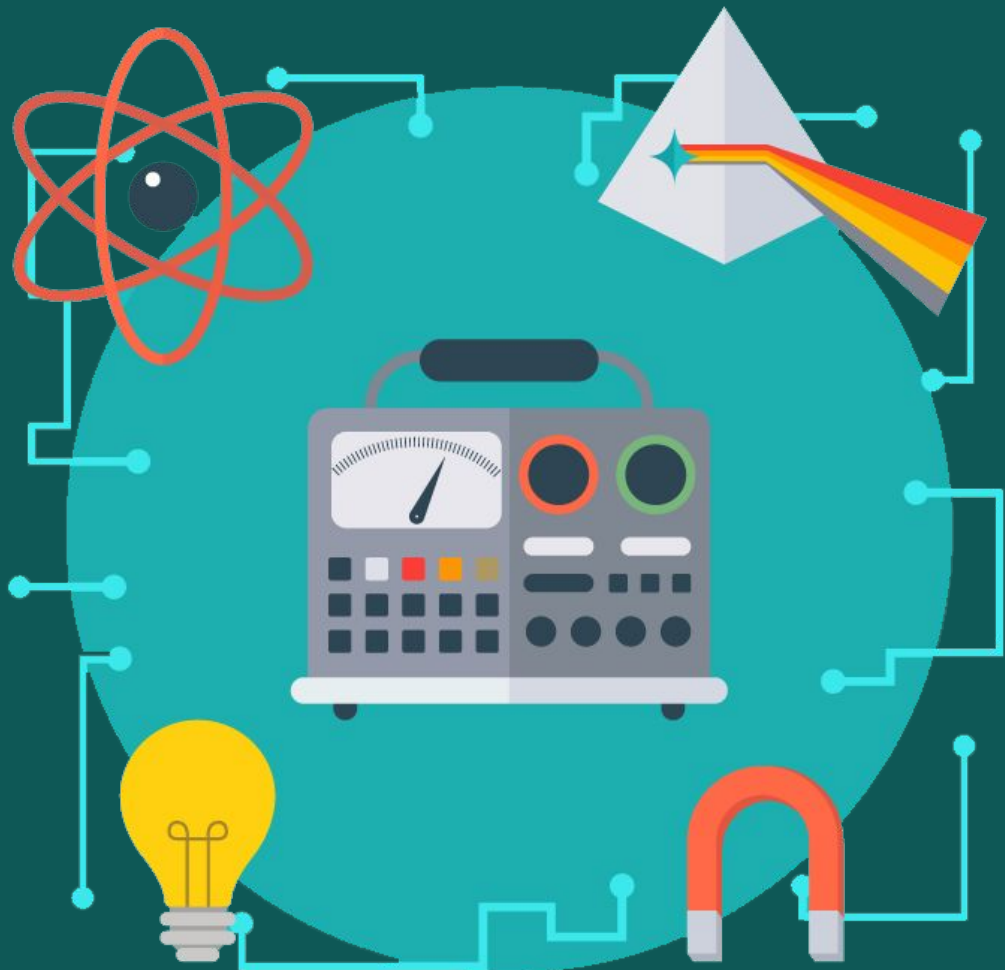


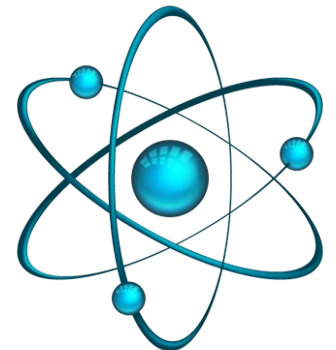
ПРАКТИЧЕСКАЯ ФИЗИКА



Доцент кафедры
экспериментальной
физики
Ерина Марина Васильевна

Лекция № 10

Гистограммы. Построение и анализ.

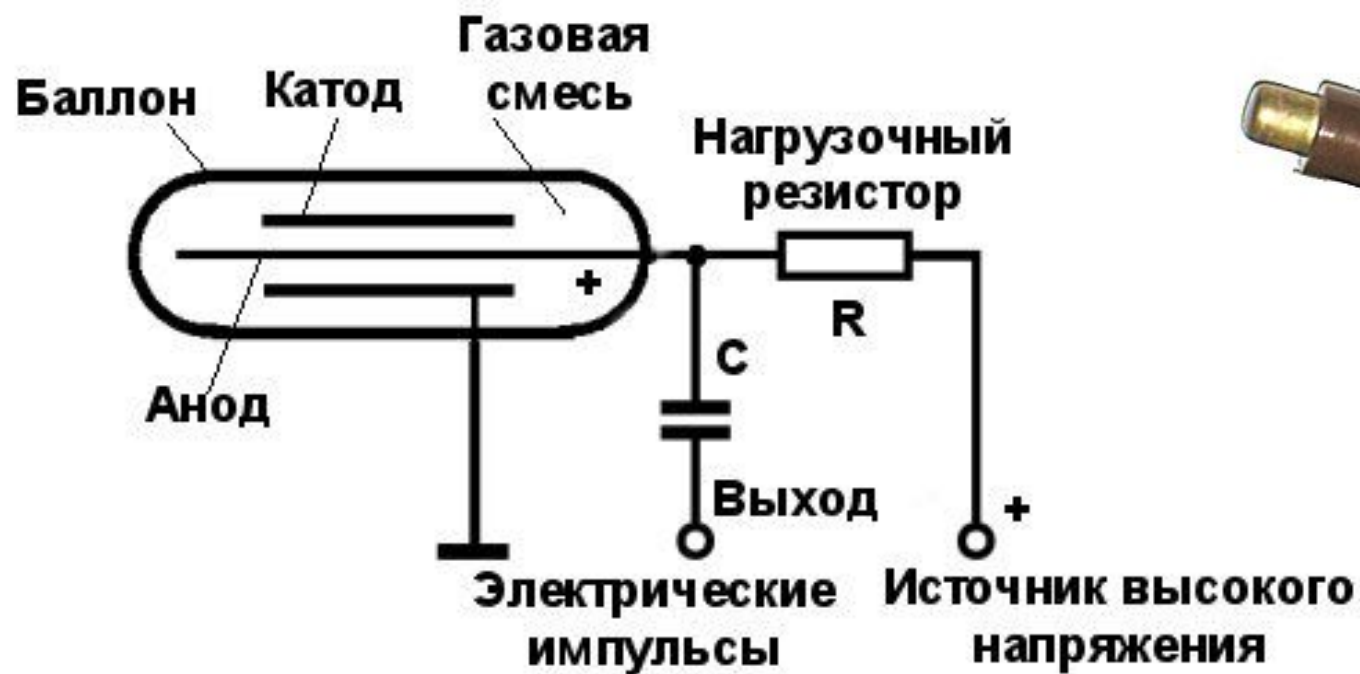


Гистограммы

Статистический анализ случайных погрешностей подразумевает проведение многократных измерений. Если число измерений превышает 15-20, то их представление в виде таблицы или списка не удобно и не наглядно. В таких случаях результаты наглядно представить в виде гистограммы.

Допустим, получен ряд результатов измерений количества α - частиц, попавших в счетчик Гейгера за 1 мин. (26; 24; 26; 28; 23; 24; 25; 24; 26; 25). Если в серии измерений один и тот же результат встречается несколько раз, то можно подсчитать число реализаций каждого результата.

Счетчик Гейгера



Устройство, схема включения и внешний вид счетчика Гейгера

Таблица
1.

x_i	23	24	25	26	27	28
n_i	1	3	2	3	0	1

Среднее значение:

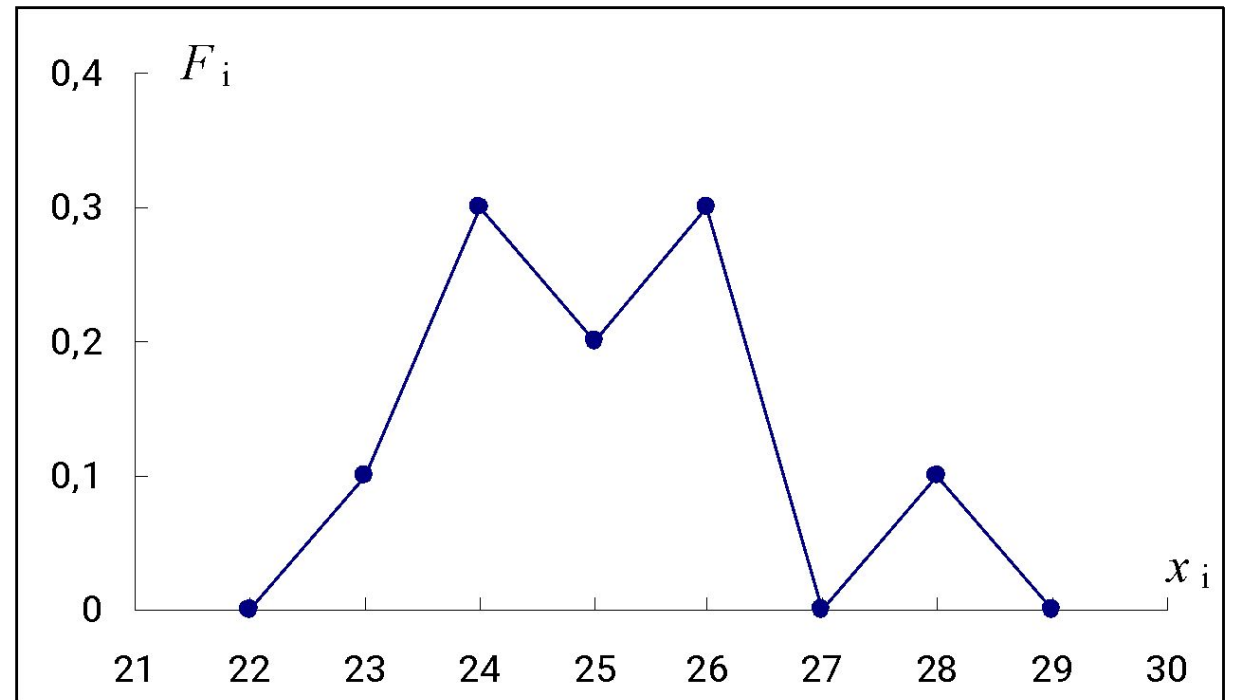
$$\bar{x} = \frac{23 \cdot 1 + 24 \cdot 3 + 25 \cdot 2 + 26 \cdot 3 + 27 \cdot 0 + 28 \cdot 1}{10} = 25,1$$

$$\bar{x} = \frac{\sum_i x_i n_i}{N}$$

$$N = \sum_i n_i$$

Можно ввести величину, показывающую сколько раз в N экспериментах встречается величина x_i .

$$F_i = \frac{n_i}{N} \quad \bar{x} = \sum_i F_i x_i \quad 1 = \sum_i F_i$$



В большинстве случаев в физике приходится иметь дело с непрерывными физическими величинами.

Пример: пусть при помощи штангенциркуля произведены измерения длины и получены следующие результаты (в мм): (26,4; 23,9; 25,1; 24,6; 22,7; 23,8; 25,1; 23,9; 25,3; 25,4). Результаты практически не повторяются, поэтому для построения гистограммы необходимо разбить диапазон результатов измерений на равные интервалы-промежутки и подсчитать число попаданий в каждый интервал. Каждый такой интервал называется **бином (карманом)**.

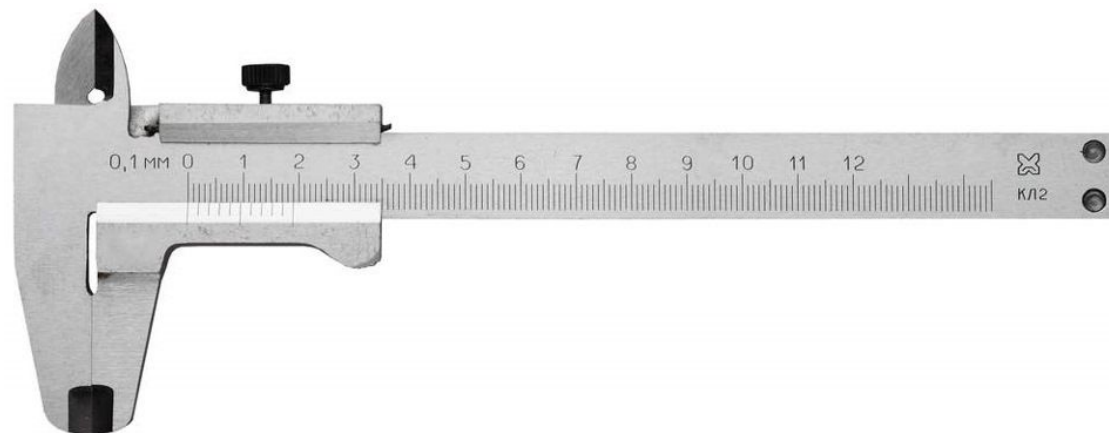
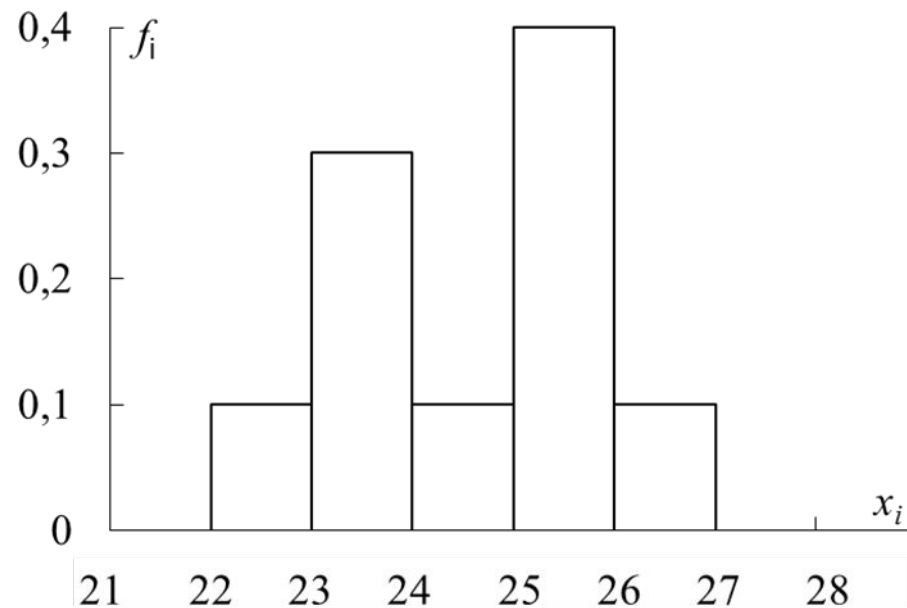


Таблица
2.

бин	22-23	23-24	24-25	25-26	26-27	27-28
число попаданий n_i	1	3	1	4	1	0

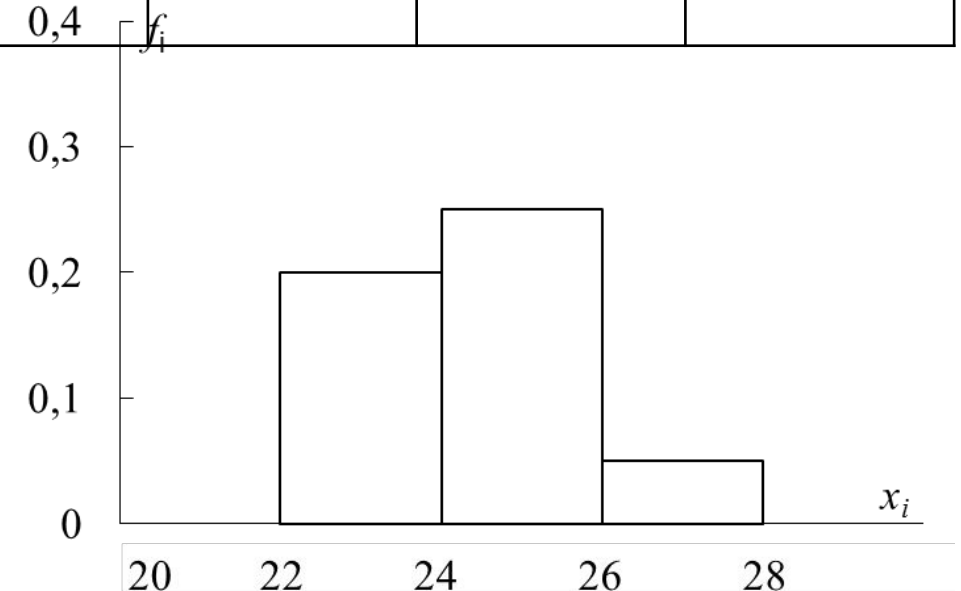
Построим
гистограмму
зависимости $f_i(x_i)$, где
 $F_i = f_i(x_i)\Delta x_i$



Построим гистограмму для другой ширины
бинов

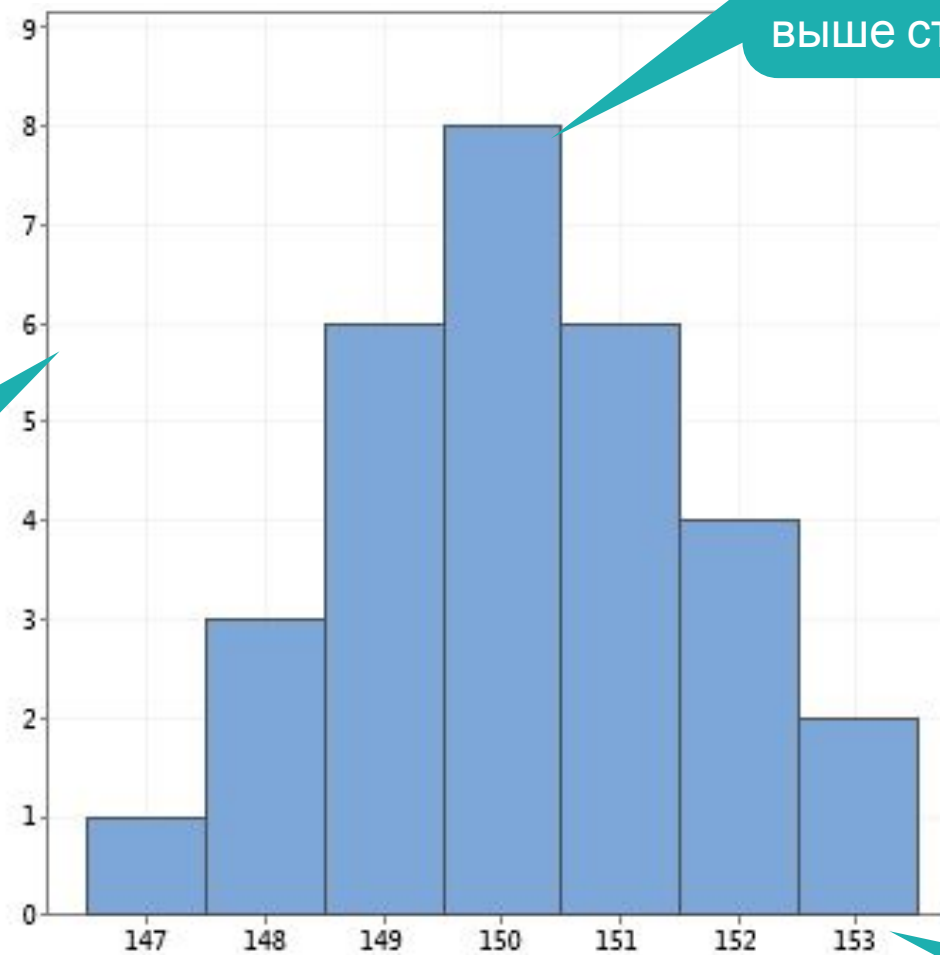
Таблица
3.

бин	22-24	24-26	26-28
число попаданий n_i	4	5	1



Выводы

По вертикальной
оси – частота
появления
переменной



Чем больше
наблюдений, тем
выше столбик

По горизонтальной
оси – числовые
интервалы

Порядок построения гистограммы:

- Собрать данные, выявить максимальное и минимальное значения и определить диапазон (размах) гистограммы.
- Полученный диапазон разделить на интервалы, предварительно определив их число (обычно 5-20 в зависимости от числа показателей) и определить ширину интервала.
- Все данные распределить по интервалам в порядке возрастания: левая граница первого интервала должна быть меньше наименьшего из имеющихся значений.
- Подсчитать частоту попаданий результатов в каждый интервал.
- Вычислить относительную частоту попадания данных в каждый из интервалов.
- По полученным данным построить гистограмму - столбчатую диаграмму, высота столбиков которой соответствует частоте или относительной частоте попадания данных в каждый из интервалов:
 - наносится горизонтальная ось, выбирается масштаб и откладываются соответствующие интервалы;
 - затем строится вертикальная ось, на которой также выбирается масштаб в соответствии с максимальным значением частот.

Правило Стёрджеса - эмпирическое правило определения оптимального количества интервалов, на которые разбивается наблюдаемый диапазон изменения случайной величины при построении гистограммы плотности её распределения.

Количество интервалов n определяется как:

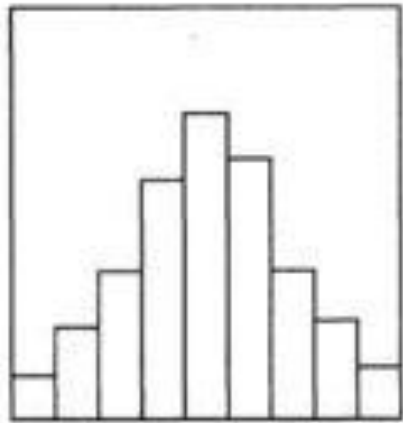
$$n = 1 + \lceil \log_2 N \rceil \quad \text{или} \quad n = 1 + \lceil 3.322 \ln N \rceil$$

где N – общее число наблюдений величины, \log_2 – логарифм по основанию 2, \ln – натуральный логарифм, $\lceil \quad \rceil$ – обозначает целую часть числа x .

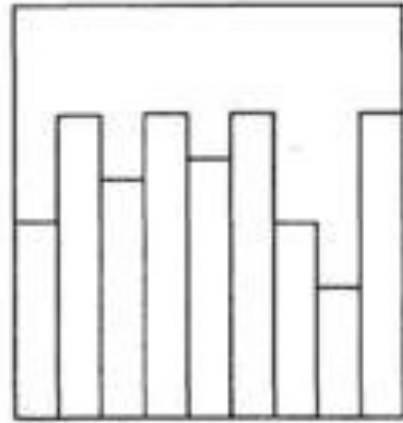
К выбору рекомендуемого числа интервалов на гистограмме:

Количество данных в выборке	Число интервалов
23 – 45	6
46 – 90	7
91 – 1 80	8
181 – 361	9
362 – 723	10
724 – 1447	11
1448 – 2885	12

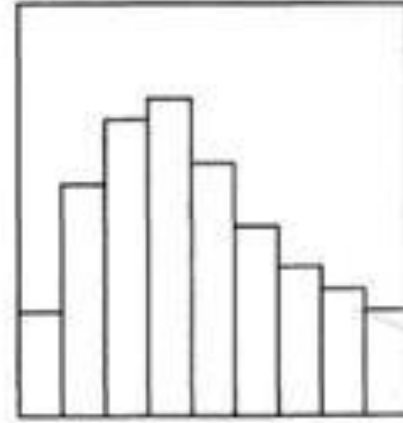
Примеры гистограмм



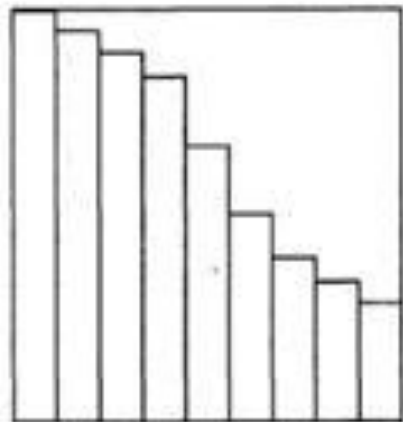
а



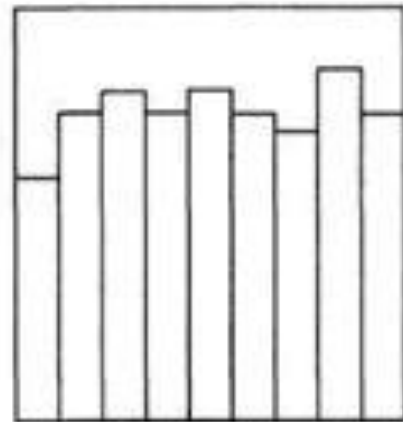
б



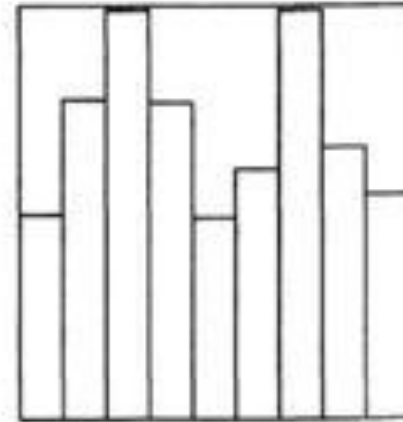
в



г



д

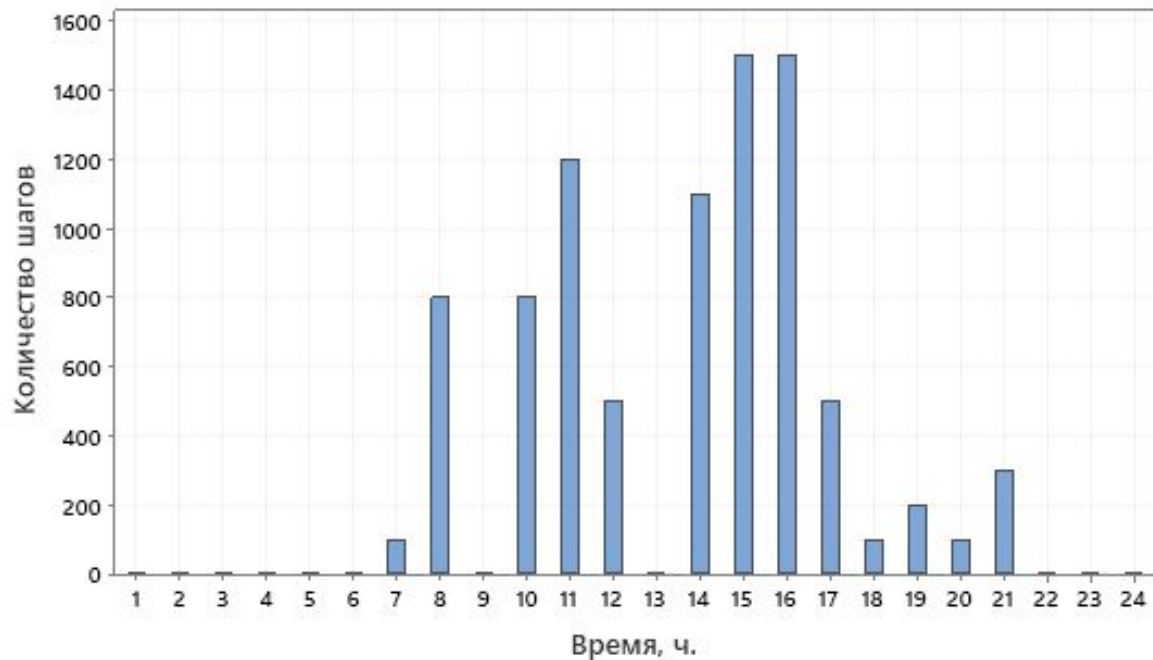


е

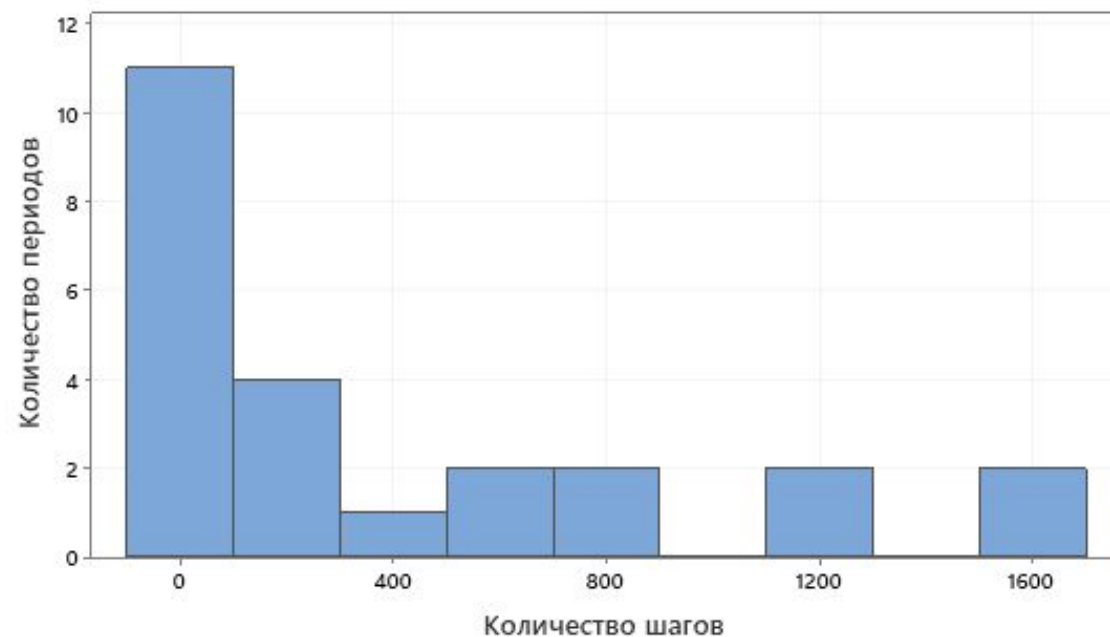
- а – симметричная, или колоколообразная;
- б – гребенка;
- в – положительно скошенное распределение;
- г – распределение с обрывом справа;
- д – равномерное распределение (плато);
- е – двухпиковая (бимодальная) форма.

Диаграмма временного ряда - гистограмма

Количество шагов за день
Диаграмма временного ряда (столбчатая)



Распределение количества шагов в час
Гистограмма



Как проводить анализ гистограмм?

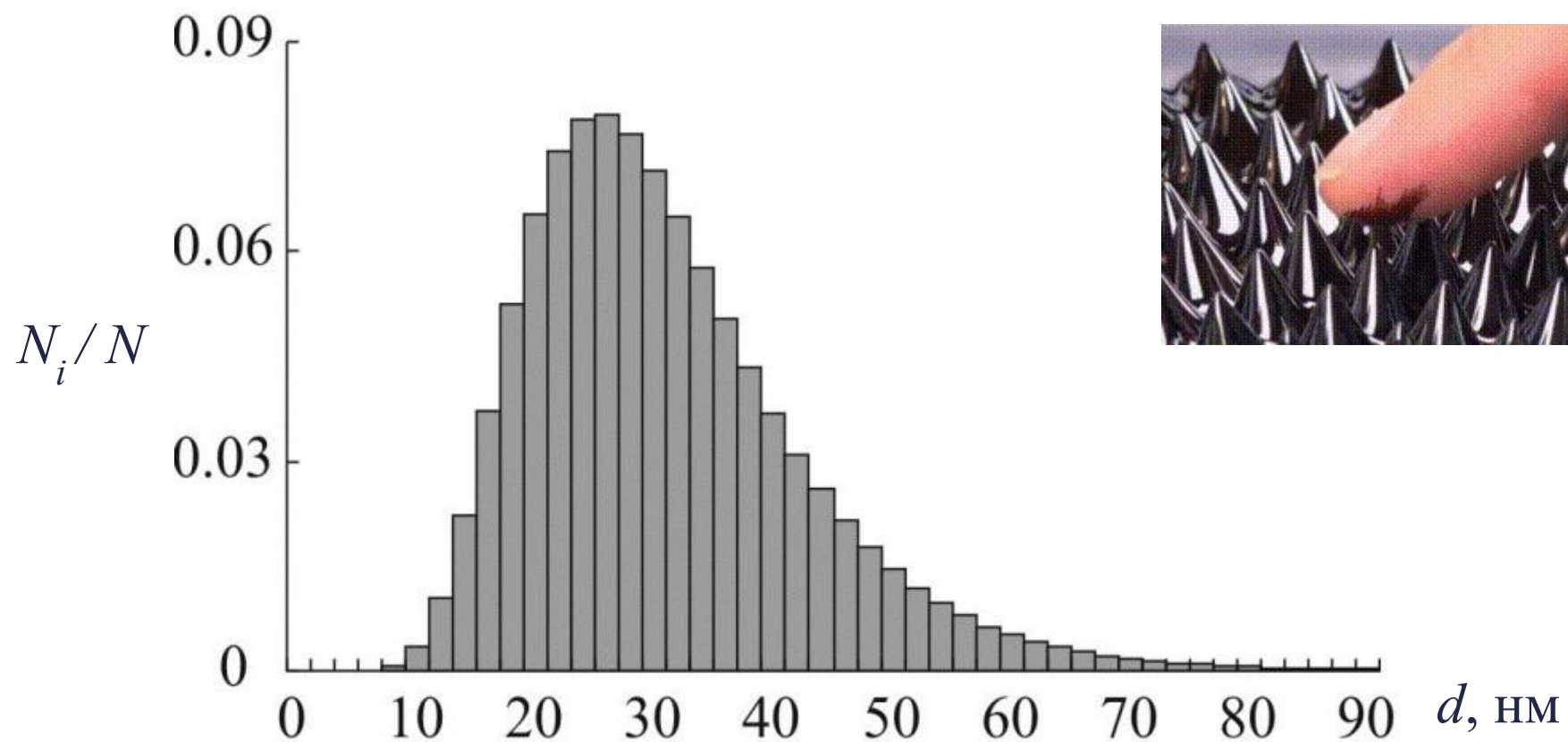
Гистограммы нужны для того, чтобы наглядно представить распределение наблюдений. Но что еще мы можем увидеть, рассматривая прямоугольники, составляющие этот график? Какую информацию можем «прочитать»? На что обратить внимание и как это лучше сделать?

Визуальная оценка гистограмм позволяет воспринять ряд статистических показателей:

- распределение наблюдений (distribution);
- наибольшую концентрацию данных – моду (mode);
- минимальное и максимальное значения (min и max);
- размах (range);
- степень асимметрии – скос (skewness);
- эксцесс (kurtosis);
- наличие явных выбросов (outliers);
- возможное присутствие нескольких распределений (популяций);
- ширину интервалов – дистанцию между правым и левым краями частотной ячейки по оси X;
- количество интервалов – общее (в том числе и нулевые значения) количество частотных ячеек гистограммы.

Пример

Распределение частиц магнитной жидкости по размерам



СПАСИБО!

$$S = \frac{(v - v_0)}{2a}$$

$$\Delta U = A + Q$$

$$F = \frac{q_1 q_2}{R^2}$$

$$Q = \lambda m$$

$$X = X_{\max} \cdot \cos \omega t$$

$$N = N_0 \cdot 2^{-t/T}$$

$$A = FS \cos \alpha$$

$$P = \frac{F}{S}$$

$$\Delta d = \frac{(2k+1)\lambda}{2}$$

$$\phi = \frac{P}{P_0 \cdot 100\%}$$

$$Ft = \Delta p$$

$$F = mg$$

$$v_2 = \frac{(v_1 + v)}{1 + v_1 v / c^2}$$

$$t = \frac{t_1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$\lambda = vT$$

$$T = 2\pi \sqrt{LC}$$

$$P = IU$$

СПАСИБО!

$$Z = \sqrt{(X_C - X_L)^2 + R^2}$$

$$E = \frac{mv^2}{2}$$

$$\eta = \frac{(Q_1 - Q_2)}{Q_1}$$

$$E = 2\pi k \sigma$$

$$F = \rho g V$$

$$Q = C(T_2 - T_1)$$

$$P = m(g+a)$$

$$\frac{v}{T} = \text{const}$$

$$\rho = \frac{m}{V}$$

$$p = mc = \frac{h}{\lambda} = \frac{E}{c}$$

$$T = \frac{2\pi\sqrt{l}}{g}$$

$$F = \frac{kq_1 q_2}{r^2}$$

$$F_{\text{упр}} = -kx$$