

## Статистика.

В современном обществе статистика стала одним из важнейших инструментов управления экономикой. Принятие любого управленческого решения требует предварительного анализа имеющейся ситуации, основывается на просчете вариантов развития, сравнении этих вариантов, оценки точности прогнозов, вероятности ошибок. Методическую базу для решения этих вопросов представляет статистика.

Термин «статистика» происходит от латинского слова «статус» -определенное состояние, положение вещей.

**Опр.** *Математическая статистика* – наука, изучающая методы сбора, систематизации и обработки результатов наблюдений массовых случайных (величин) явлений с целью выявления статических закономерностей.

Задача математической статистики состоит в создании методов сбора и обработки статистических данных для получения научных и практических выводов.

Предметом математической статистики является изучение случайных величин (событий, процессов) по результатам наблюдения. Полученные в результате наблюдения (опыта, эксперимента) данные сначала надо:

- 1) упорядочить
- 2) оценить
- 3) проверить статистические гипотезы, т.е. решить вопрос согласования результатов оценивания с опытными данными.

Результаты исследования статистических данных методами математической статистики используются для принятия решения в задачах планирования, управления, прогнозирования и организации производства, при контроле качества продукции, при выборе оптимального времени настройки или замены действующей аппаратуры и т.д.

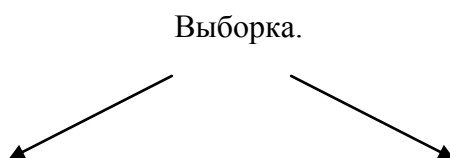
### Генеральная и выборочная совокупности.

Пусть требуется изучить данную совокупность объектов относительно некоторого признака. Например, рассматривая работу диспетчера (продавца, парикмахера и т.д.), можно исследовать: его загруженность, тип клиентов, скорость обслуживания, моменты поступления звонков и т.д. Каждый такой признак (или их комбинации) образуют случайную величину, наблюдения над которой мы и производим.

**Опр.** Вся совокупность рассматриваемых объектов называется *генеральной совокупностью*.

**Опр.** Часть генеральной совокупности называется *выборкой*.

**Опр.** *Объем выборки* (генеральной совокупности) – это число объектов этой совокупности.



**Опр.** Выборка называется *повторной*, если отобранный объект возвращается в генеральную совокупность перед извлечением следующего, в противном случае выборка называется *бесповторной*.

### Предварительная обработка статистических данных

- 1) Сначала данные *упорядочивают* и *группируют*;
- 2) Затем составляют *таблицы распределения данных*;
- 3) Строят графики распределения данных в виде *многоугольника распределения (полигона)*, *гистограммы* распределения или *круговой диаграммы*;
- 4) получение *паспорта данных* измерения, который состоит из небольшого количества основных характеристик полученной информации.

Предположим, что из генеральной совокупности извлечена выборка объемом  $n$ . Так как выборка случайна, то некоторые значения в выборке могут совпадать. Пусть значение  $x_1$  наблюдается  $n_1$  раз,  $x_2$  —  $n_2$  раз, ...,  $x_m$  —  $n_m$  раз. Значения  $x_1, x_2, \dots, x_m$  называются *вариантами*. Одно из самых простых преобразований статистических данных является их упорядочение по значениям (вариантам). Если варианты записаны в возрастающем порядке (*проранжированны*), то их называют *вариационным рядом*.

Числа  $n_1, n_2, \dots, n_m$  называются *частотами* (или кратностями), сумма частот равна объему выборки, т.е.

$$n_1 + n_2 + \dots + n_m = n$$

Пусть дан вариационный ряд

$x_1$	$x_2$	...	$x_m$
-------	-------	-----	-------

*Модой*  $M_0$  называется варианта, имеющая наибольшую частоту.

*Медианой*  $m_e$  является варианта, делящая вариационный ряд на две равные части по количеству вариант, при этом:

- если число вариант нечетно, т.е.  $n = 2k + 1$ , то

$$m_e = x_{k+1};$$

- если число вариант четно, т.е.  $n = 2k$ , то

$$m_e = \frac{x_k + x_{k+1}}{2}$$

*Среднее (или среднее арифметическое)* – это частное от деления суммы всех результатов измерения на объем измерения.

- если варианты  $x_1, x_2, \dots, x_m$  имеют частоты  $n_1, n_2, \dots, n_m$ , то  $\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_m x_m}{n}$ ;

- если все варианты различны ( $n=m$ ), то  $\bar{x} = \frac{x_1 + x_2 + \dots + x_m}{m}$

Разность между наибольшим и наименьшим значениями выборки называют *размахом выборки*  $R = x_{max} - x_{min}$ .

Отношения  $w_1 = \frac{n_1}{n}, w_2 = \frac{n_2}{n}, \dots, w_m = \frac{n_m}{n}$  называются *относительными частотами* (статистическими вероятностями или частностями) соответствующих вариантов.

Сумма относительных частот равна единице:

$$\sum_{i=1}^m w_i = \sum_{i=1}^m \frac{n_i}{n} = \frac{1}{n} \sum_{i=1}^m n_i = \frac{1}{n} \cdot n = 1.$$

Данные вариационного ряда записывают в виде статистических таблиц:

- таблицы распределения частот, в которой в первой строке записывают значения выборки (варианты), а во второй строке – соответствующие частоты:

$x_1$	$x_2$	...	$x_m$
$n_1$	$n_2$	...	$n_m$

- таблицы распределения относительных частот, в которой в первой строке записываются варианты, а во второй строке – соответствующие относительные частоты:

$x_1$	$x_2$	...	$x_m$
$w_1$	$w_2$	...	$w_m$

Решим несколько примеров.

**Пример 1.** Случайная выборка среди абитуриентов на приемных экзаменах дала следующие набранные ими баллы: 13, 12, 14, 11, 11, 12, 14, 10, 12, 13, 11, 15, 10, 13, 11, 12, 14, 12, 12, 11. Построить для данной выборки таблицу распределения частот и таблицу распределения относительных частот, определить среднее, моду, медиану и размах выборки.

Решение.

1. Составим вариационный ряд, для этого расположим данные в возрастающем порядке: 10, 10, 11, 11, 11, 11, 11, 12, 12, 12, 12, 12, 12, 13, 13, 13, 14, 14, 14, 15.
2. Выпишем варианты  $x_1=10, x_2=11, x_3=12, x_4=13, x_5=14, x_6=15$ ; их частоты равны:  $n_1=2, n_2=5, n_3=6, n_4=3, n_5=3, n_6=1$ .
3. Мода  $M_0=12$ , так как наибольшая частота, равная 6, соответствует варианту  $x_3=12$ .
4. Так как количество вариантов четно, то медиана

$$m_e = \frac{12 + 12}{2} = 12$$

5. Размах выборки равен  $x_6 - x_1 = 15 - 10 = 5$ .
6. Объем выборки  $n = \sum_{i=1}^6 n_i = 2 + 5 + 6 + 3 + 3 + 1 = 20$ .
7. Среднее:  $\bar{x} = \frac{10 \cdot 2 + 11 \cdot 5 + 12 \cdot 6 + 13 \cdot 3 + 14 \cdot 3 + 15 \cdot 1}{20} = 12,15$

8. Относительные частоты:

$$w_1 = \frac{n_1}{n} = \frac{2}{20} = 0,1; w_2 = \frac{n_2}{n} = \frac{5}{20} = 0,25; w_3 = \frac{n_3}{n} = \frac{6}{20} = 0,3;$$

$$w_4 = \frac{n_4}{n} = \frac{3}{20} = 0,15; w_5 = \frac{n_5}{n} = \frac{3}{20} = 0,15; w_6 = \frac{n_6}{n} = \frac{1}{20} = 0,05.$$

Контроль:  $\sum_{i=1}^6 w_i = 0,1 + 0,25 + 0,3 + 0,15 + 0,15 + 0,05 = 1$ .

9. Полученные данные сведем в соответствующие таблицы.

Таблица распределения частот:

$x_i$	10	11	12	13	14	15
-------	----	----	----	----	----	----

$n_i$	2	5	6	3	3	1
-------	---	---	---	---	---	---

Таблица распределения относительных частот:

$x_i$	10	11	12	13	14	15
$w_i$	0,1	0,25	0,3	0,15	0,15	0,05

**Пример 2.** Выборка задана в виде распределения частот

$x_i$	4	7	8	12
$n_i$	5	2	3	10

Построить для данной выборки таблицу распределения относительных частот, определить среднее, моду, медиану и размах выборки.

Решение.

1. Выпишем варианты  $x_1=4$ ,  $x_2=7$ ,  $x_3=8$ ,  $x_4=12$ ; их частоты равны:  $n_1=5$ ,  $n_2=2$ ,  $n_3=3$ ,  $n_4=10$ .
2. Мода  $M_0=12$ , так как наибольшая частота, равная 10, соответствует варианту  $x_4=12$ .
3. Так как количество вариант четно, то медиана

$$m_e = \frac{x_3 + x_4}{2} = \frac{8 + 12}{2} = 10$$

4. Размах выборки равен  $x_4 - x_1 = 12 - 4 = 8$ .
5. Объем выборки  $n = \sum_{i=1}^4 n_i = 5 + 2 + 3 + 10 = 20$ .
6. Среднее:  $\bar{x} = \frac{4 \cdot 5 + 7 \cdot 2 + 8 \cdot 3 + 12 \cdot 10}{20} = 8,9$

7. Относительные частоты:

$$w_1 = \frac{n_1}{n} = \frac{5}{20} = 0,25; w_2 = \frac{n_2}{n} = \frac{2}{20} = 0,1;$$

$$w_3 = \frac{n_3}{n} = \frac{3}{20} = 0,15; w_4 = \frac{n_4}{n} = \frac{10}{20} = 0,5.$$

Контроль:  $\sum_{i=1}^4 w_i = 0,25 + 0,1 + 0,15 + 0,5 = 1$ .

8. Полученные данные сведем в таблицу распределения относительных частот.

$x_i$	10	11	12	13	14	15
$w_i$	0,1	0,25	0,3	0,15	0,15	0,05

**Пример 3.** Ученик выписал из дневника свои отметки за март: 4,4,3,2,5,3,3,4,5,4,4,4,5,4,2,4,4,5,3,3.

Построить для данной выборки таблицу распределения частот и таблицу распределения относительных частот, определить среднее, моду, медиану и размах выборки.

Решение.

1. Составим вариационный ряд, для этого расположим данные в возрастающем порядке: 2,2,3,3,3,3,3,4,4,4,4,4,4,4,4,5,5,5,5.
2. Выпишем варианты  $x_1=2$ ,  $x_2=3$ ,  $x_3=4$ ,  $x_4=5$ ; их частоты равны:  $n_1=2$ ,  $n_2=5$ ,  $n_3=9$ ,  $n_4=4$ .
3. Мода  $M_0=4$ , так как наибольшая частота, равная 9, соответствует варианту  $x_3=4$ .
4. Так как количество вариант четно, то медиана

$$m_e = \frac{4 + 4}{2} = 4$$

5. Размах выборки равен  $x_4 - x_1 = 5 - 2 = 3$ .
6. Объем выборки  $n = \sum_{i=1}^4 n_i = 2 + 5 + 9 + 4 = 20$ .
7. Среднее:  $\bar{x} = \frac{2 \cdot 2 + 3 \cdot 5 + 4 \cdot 9 + 5 \cdot 4}{20} = 3,75$

8. Относительные частоты:

$$w_1 = \frac{n_1}{n} = \frac{2}{20} = 0,1; w_2 = \frac{n_2}{n} = \frac{5}{20} = 0,25; w_3 = \frac{n_3}{n} = \frac{9}{20} = 0,45; w_4 = \frac{n_4}{n} = \frac{4}{20} = 0,2.$$

Контроль:  $\sum_{i=1}^4 w_i = 0,1 + 0,25 + 0,45 + 0,2 = 1$ .

9. Полученные данные сведем в соответствующие таблицы.

Таблица распределения частот:

$x_i$	2	3	4	5
$n_i$	2	5	9	4

Таблица распределения относительных частот:

$x_i$	2	3	4	5
$w_i$	0,1	0,25	0,45	0,2

**Пример 4.** В очередном туре футбольного чемпионата состоялись 10 матчей. Вот их результаты: 3:1; 0:2; 1:1; 0:0; 0:4; 0:1; 2:2; 0:3; 1:0; 1:1. Футбольный статистик подсчитал результативность матчей (количество голов). Выписать вариационный ряд, таблицу распределения частот и таблицу распределения относительных частот, определить среднее, моду, медиану и размах выборки.

Решение.

1. Для начала выпишем несгруппированный ряд данных: 4, 2, 2, 0, 4, 1, 4, 3, 1, 2.
2. Составим вариационный ряд, для этого расположим данные в возрастающем порядке: 0, 1, 1, 2, 2, 2, 3, 4, 4, 4.
3. Выпишем варианты  $x_1=0, x_2=1, x_3=2, x_4=3, x_5=4$ ; их частоты равны:  $n_1=1, n_2=2, n_3=3, n_4=1, n_5=3$ .
4. Мода  $M_0=2$  и 4, так как вариант  $x_3$  и  $x_5$  имеют одинаковую частоту.
5. Так как количество вариантов четно, то медиана

$$m_e = \frac{2 + 2}{2} = 2$$

6. Размах выборки равен  $x_5 - x_1 = 4 - 0 = 4$ .
7. Объем выборки  $n = \sum_{i=1}^5 n_i = 1 + 2 + 3 + 1 + 3 = 10$ .
8. Среднее:  $\bar{x} = \frac{0 \cdot 1 + 1 \cdot 2 + 2 \cdot 3 + 3 \cdot 1 + 4 \cdot 3}{10} = 1,15$

9. Относительные частоты:

$$w_1 = \frac{n_1}{n} = \frac{1}{10} = 0,1; w_2 = \frac{n_2}{n} = \frac{2}{10} = 0,2; w_3 = \frac{n_3}{n} = \frac{3}{10} = 0,3;$$

$$w_4 = \frac{n_4}{n} = \frac{1}{10} = 0,1; w_5 = \frac{n_5}{n} = \frac{3}{10} = 0,3.$$

Контроль:  $\sum_{i=1}^5 w_i = 0,1 + 0,2 + 0,3 + 0,1 + 0,3 = 1$ .

10. Полученные данные сведем в соответствующие таблицы.

Таблица распределения частот:

$x_i$	0	1	2	3	4
$n_i$	1	2	3	1	3

Таблица распределения относительных частот:

$x_i$	0	1	2	3	4
$w_i$	0,1	0,2	0,3	0,1	0,3

В случае, когда число значений признака велико или признак является непрерывным, составляют генеральный статистический ряд. В первую строку таблицы статистического распределения вписывают частичные промежутки  $(x_0, x_1), (x_1, x_2), \dots, (x_{k-1}, x_k)$ , которые берут обычно одинаковыми по длине. Для определения величины интервала ( $h$ ) можно использовать формулу Стерджеса:

$$h = \frac{R}{k},$$

где  $R = x_{\max} - x_{\min}$  – размах выборки,  $k = 1 + \log_2 n$  – число интервалов ( $\log_2 n \approx 3,322 \lg n$ ).

За начало первого интервала рекомендуется брать величину  $x_{нач} = x_{\min} - \frac{h}{2}$ .

Во второй строке статистического ряда вписывают количество наблюдений  $n_i (i=1, \dots, k)$ , попавших в каждый интервал. Так как граничные значения признака могут совпадать с границами интервалов, то условимся в каждый интервал включать варианты, большие, чем нижняя граница интервала ( $x_i > a_i$ ), и меньшие или равные верхней границе ( $x_i \leq b_i$ ).

**Пример 5.** Измерили рост 30 наудачу отобранных студентов. Результаты таковы:

178, 160, 154, 183, 155, 153, 167, 186, 163, 155, 157, 175, 170, 166, 159, 173, 182, 167, 171, 169, 179, 165, 156, 179, 158, 171, 175, 173, 164, 172. Построить интервальный статистический ряд.

Решение:

1. Проранжируем полученные данные: 153, 154, 155, 155, 156, 157, 158, 159, 160, 163, 164, 165, 166, 167, 167, 169, 170, 171, 171, 172, 173, 173, 175, 175, 178, 179, 179, 182, 183, 186.  
 $X$  – рост студента – непрерывная случайная величина. При более точном измерении роста значения сл.в.  $X$  обычно не повторяются.
2. Объем выборки равен 30,  $R = 186 - 153 = 33$  – размах выборки.
3. Найдем длину интервалов:  $h = \frac{33}{1 + \log_2 30} \approx \frac{33}{5,907} \approx 5,59$ . Примем  $h = 6$ , тогда  $x_{нач} = 153 - 6 : 2 = 150$ .
4. Здесь  $k = 5,907 \approx 6$  – количество интервалов.
5. Исходные данные разбиваем на 6 интервалов: (150; 156), (156; 162), (162; 168), (168; 174), (174; 180), (180; 186).
6. Подставив число студентов, попавших в каждый из полученных промежутков, получим интервальный ряд:

Рост	(150; 156)	(156; 162)	(162; 168)	(168; 174)	(174; 180)	(180; 186)
Частота	5	4	6	7	5	3

### Полигон и гистограмма

Более наглядное представление о выборке можно получить графически, построив полигон (для дискретной случайной величины) и гистограмму (для непрерывной случайной величины).

**Опр.** Полигон – ломанная, соединяющая точки, соответствующие средним значениям интервалов группировки и частотами этих интервалов. Срединные значения откладываются по оси x, а частоты по оси y.

**Опр.** Гистограмма – ступенчатая фигура, состоящая из прямоугольников, основаниями которых являются отрезки длиной  $x_i - x_{i-1}$ , а их высоты равны  $n_i$  (если ширина всех интервалов группировки одинакова) или  $\frac{n_i}{h_i}$  ( $n_i$ - частота i-го интервала,  $h_i$ -ширина i-го интервала).

**Пример 6.** В результате тестирования группа из 24 человек набрала баллы: 4, 0, 3, 4, 1, 0, 3, 1, 0, 4, 0, 0, 3, 1, 0, 1, 1, 3, 2, 3, 1, 2, 1, 2.

Построить полигон и гистограмму.

Решение:

1. Вариационный ряд: 0,0,0,0,0,0,1,1,1,1,1,1,1,1,2,2,2,3,3,3,3,3,4,4,4.
2. Варианты:  $x_1=0$ ,  $x_2=1$ ,  $x_3=2$ ,  $x_4=3$ ,  $x_5=4$ ; их частоты равны:  $n_1=6$ ,  $n_2=7$ ,  $n_3=3$ ,  $n_4=5$ ,  $n_5=3$ .
3. Объем выборки  $n=24$ .
4. Размах выборки  $R=4-0=4$ .
5. Число интервалов:  $k = 1 + \log_2 24 = 1 + 3,3221 \cdot \lg 24 \approx 5$ .
6. Длина интервала:  $h = \frac{R}{k} = \frac{4}{5} = 0,8 \approx 1$ .
7. Полигон и гистограмма имеют вид (рис. 1):

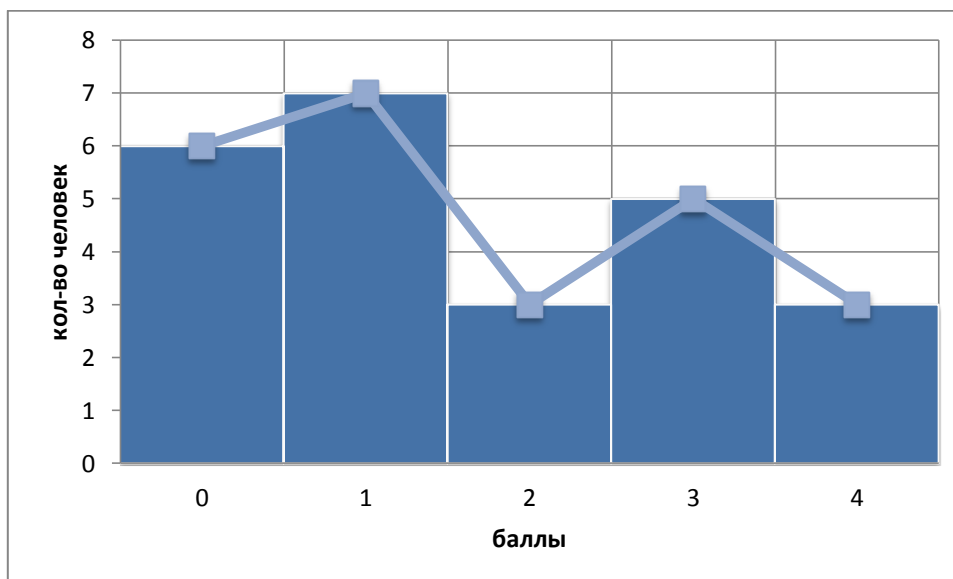


Рис. 1

Для этого примера построим еще и круговую диаграмму (рис.2). Круговые диаграммы показывают размер элементов в одном ряду данных, пропорционально сумме элементов. Значения данных в круговой диаграмме показываются как доля целого круга.



Рис.2

**Пример 7.** Рассмотрим глазодвигательную активность по числу попаданий в мишень: 15, 8, 10, 13, 17, 15, 15, 19, 22, 15, 14, 17, 15, 13, 12.

Решение:

1. Вариационный ряд: 8, 10, 12, 13, 13, 14, 15, 15, 15, 15, 15, 17, 17, 19, 22.
2. Объем выборки равен 15,  $R=22-8=14$ - размах выборки.
3. Найдем длину интервалов:  $h = \frac{14}{1+\log_2 15} \approx \frac{14}{4,8} \approx 2,9$ . Примем  $h=3$ , тогда  $x_{нач}=8-3:2=6,5 \approx 7$ .
4. Здесь  $k=4,8 \approx 5$  – количество интервалов.
5. Исходные данные разбиваем на 5 интервалов: (7; 10), (10;13), (13;16), (16;19), (19;22).
6. Составим интервальный ряд:

Число попаданий	(7; 10)	(10;13)	(13;16)	(16;19)	(19;22)
Частота	2	3	6	3	1

Полигон и гистограмма имеют вид (рис. 3-5):



Рис.3



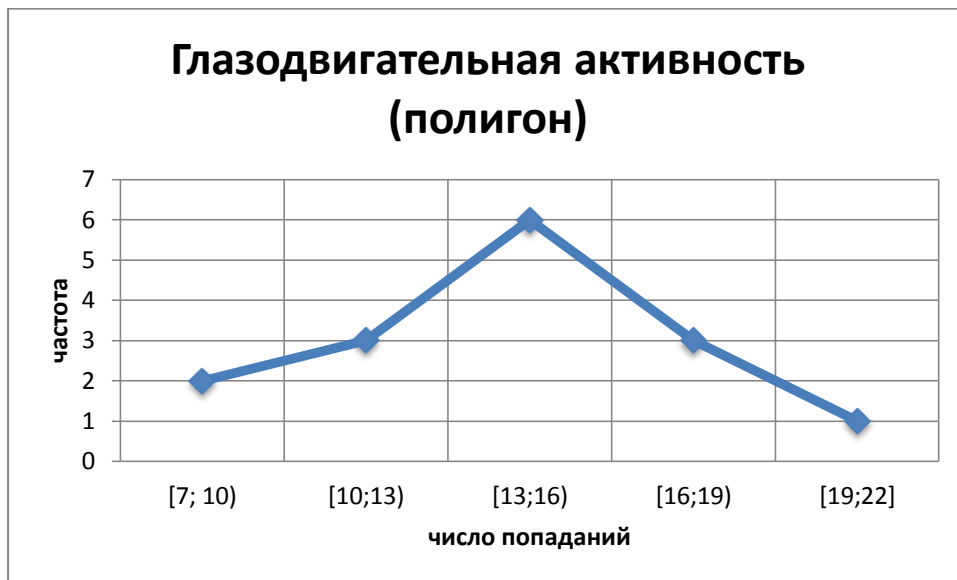


Рис.4

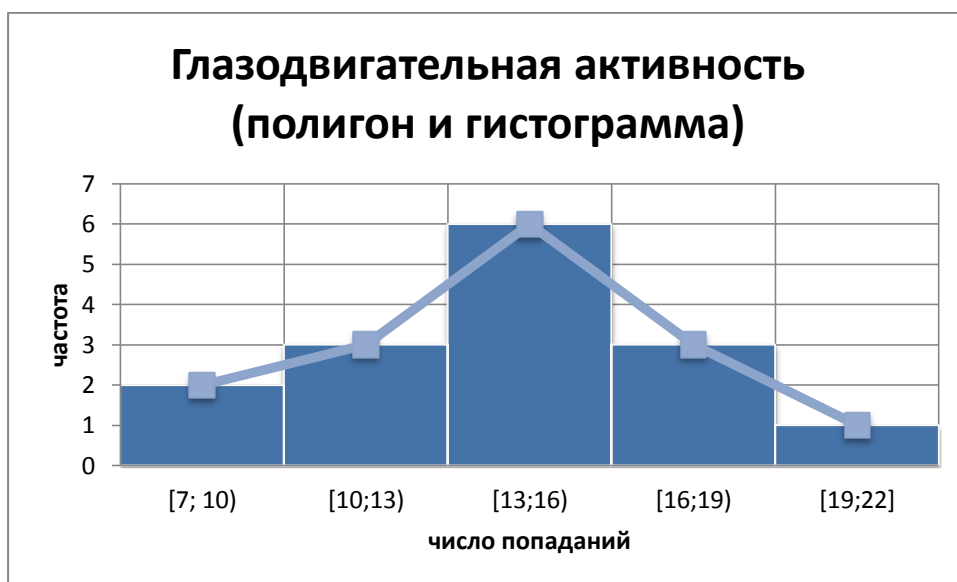


Рис. 5

**Пример 8.** Измерения напряжения электросети дали результаты:

210, 198, 215, 212, 194, 213, 199, 191, 205, 211, 189, 206, 204, 205, 201, 194, 190, 200, 202, 196, 200, 216, 214, 200, 196, 210, 206, 200, 215, 204. Построить гистограмму частот.

Решение:

1. Вариационный ряд: 189, 190, 191, 194, 194, 196, 196, 198, 199, 200, 200, 200, 200, 201, 202, 204, 204, 205, 205, 206, 206, 210, 210, 211, 212, 213, 214, 215, 215, 216.
2. Объем выборки равен 30,  $R=216-189=27$ - размах выборки.
3. Найдем длину интервалов:  $h = \frac{27}{1+\log_2 30} \approx \frac{27}{5,8} \approx 4,7$ . Примем  $h=5$ , тогда  $x_{нач}=189-5 \cdot 2 \approx 186$ .
4. Здесь  $k=5,8 \approx 6$  – количество интервалов.
5. Исходные данные разбиваем на 6 интервалов: (186; 191), (191;196), (196;201), (201;206), (206;211), (211;216).
6. Составим интервальный ряд:

Промежутки	(186; 191)	(191;196)	(196;201)	(201;206)	(206;211)	(211;216)
Частота	3	4	7	7	3	6

Гистограмма имеет вид (рис. 6) (гистограмма и полигон (рис. 7)):

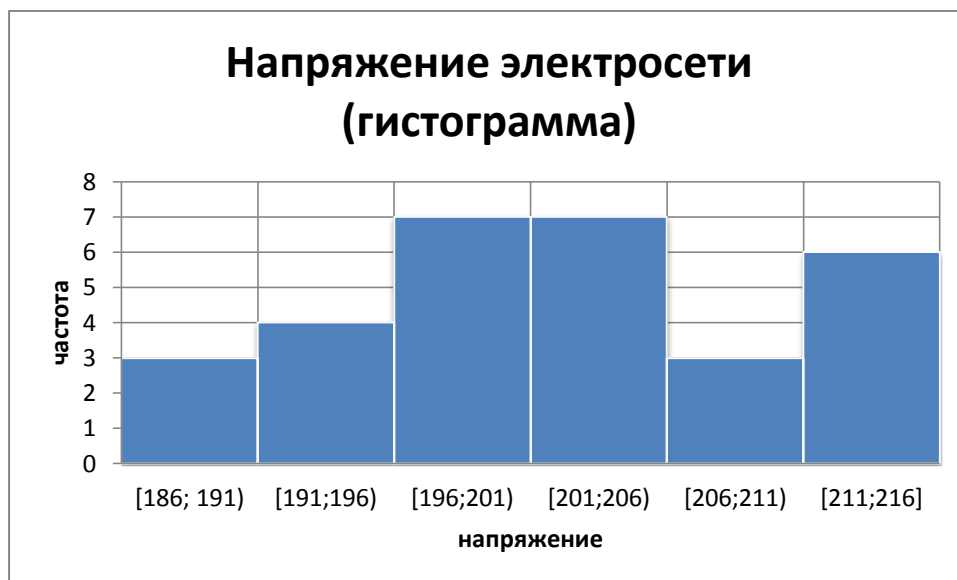


Рис.6

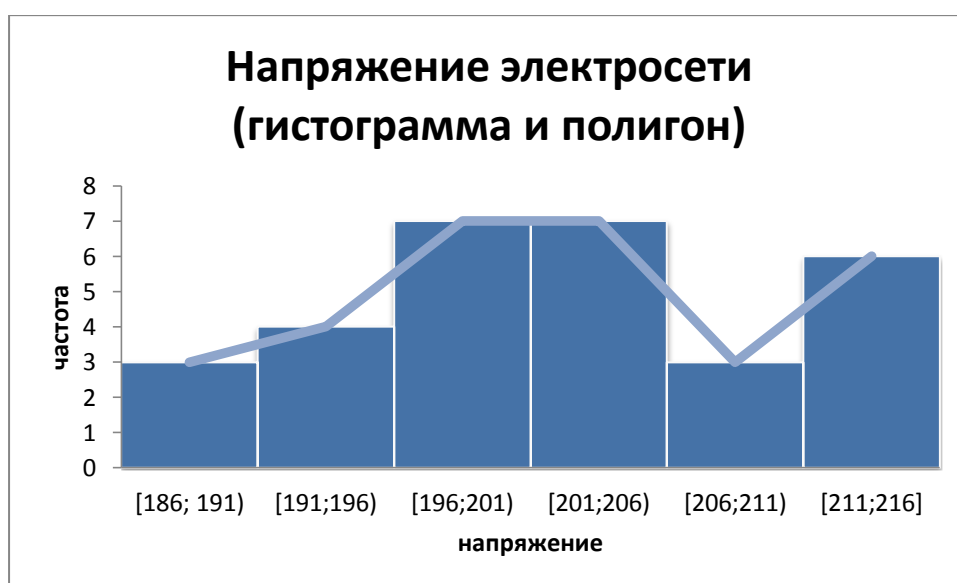


Рис.7

#### Числовые характеристики выборки

Вариационные ряды и графики эмпирических распределений дают наглядное представление о том, как варьирует признак в выборочной совокупности. Но они недостаточны для полной характеристики выборки, поскольку содержат много деталей, охватить которые невозможно без применения обобщающих числовых характеристик.

Числовые характеристики выборки дают количественное представление об эмпирических данных и позволяют сравнивать их между собой. Наибольшее практическое значение имеют характеристики положения, рассеяния и асимметрии эмпирических распределений.

Мы рассмотрим характеристики положения и рассеяния, а также практические методы их вычисления.

#### Характеристики положения

Характеристики положения определяют положение центра эмпирического распределения. Чаще всего употребляются такие характеристики положения, как среднее арифметическое, медиана и мода (определения даны выше). Иногда их называют *мерами центральной тенденции*. Если воспользоваться геометрической интерпретацией, то среднее арифметическое можно определить как точку на оси  $x$ , которая является абсциссой центра масс гистограммы.

### Характеристики рассеяния

Средние значения не дают полной информации о варьирующем признаке. Нетрудно представить себе два эмпирических распределения, у которых средние одинаковы, но при этом у одного из них значения признака рассеяны в узком диапазоне вокруг среднего, а у другого – в широком. Поэтому наряду со средними значениями вычисляют и характеристики рассеяния выборки. Рассмотрим наиболее употребительные из них: размах выборки, выборочная дисперсия и среднее квадратическое отклонение.

Размах выборки используется иногда в практических исследованиях при малых (не более 10) объемах выборки. Например, по размаху выборки легко оценить, насколько различаются лучший и худший результаты в группе спортсменов. При больших объемах выборки к его использованию надо откоситься с осторожностью.

**Опр.** *Дисперсией* называется средний квадрат отклонения значений признака от среднего арифметического. Дисперсия, вычисляемая по выборочным данным, называется *выборочной дисперсией* и обозначается  $D_B$ .

$$D_B = \frac{\sum_{i=1}^m n_i (x_i - \bar{x})^2}{n}$$

**Опр.** *Стандартным отклонением* (или *средним квадратическим отклонением*) называется корень квадратный из дисперсии:  $\sigma = \sqrt{D_B}$ .

Чем меньше дисперсия или среднее квадратическое отклонение, тем плотнее группируются данные измерения вокруг своего среднего значения.

**Пример 9.** Приводятся данные о распределении 25 работников одного из предприятий по тарифным разрядам:

4; 2; 4; 6; 5; 6; 4; 1; 3; 1; 2; 5; 2; 6; 3; 1; 2; 3; 4; 5; 4; 6; 2; 3; 4

Найти дисперсию и среднее квадратическое отклонение.

Решение:

1. В данном примере вариантами является тарифный разряд работника:  $x_1=1$ ,  $x_2=2$ ,  $x_3=3$ ,  $x_4=4$ ,  $x_5=5$ ,  $x_6=6$ .
2. Для определения частот необходимо рассчитать число работников, имеющих соответствующий тарифный разряд:  $n_1=3$ ,  $n_2=5$ ,  $n_3=4$ ,  $n_4=6$ ,  $n_5=3$ ,  $n_6=4$ .
3. Объем выборки  $n=25$ .
4. Среднее арифметическое значение выборки равно:

$$\bar{x} = \frac{1 \cdot 3 + 2 \cdot 5 + 3 \cdot 4 + 4 \cdot 6 + 5 \cdot 3 + 6 \cdot 4}{25} = \frac{98}{25} = 3,92$$

5. Вычислим дисперсию:

$$D_B = \frac{3 \cdot (1 - 3,92)^2 + 5 \cdot (2 - 3,92)^2 + 4 \cdot (3 - 3,92)^2 + 6 \cdot (4 - 3,92)^2 + 3 \cdot (5 - 3,92)^2 + 4 \cdot (6 - 3,92)^2}{25} \approx 2,7$$

6. Среднее квадратическое отклонение:  $\sigma = \sqrt{D_B} = \sqrt{2,7} \approx 1,6$ .

## Упражнения

1. На стадионе «Локомотив» была зафиксирована следующая посещаемость первых четырех футбольных матчей: 24 532, 18 711, 22 871, 24 334. Какова была средняя посещаемость (среднее арифметическое) этих матчей? Чему равен размах посещаемости?

2. В течение года Лена получала следующие отметки за контрольные по алгебре: одну «двойку», три «тройки», четыре «четверки», три «пятерки». Найдите среднее арифметическое, моду и медиану этих данных.

3. Маша, Саша, Катя, Лена, Ваня и Миша пошли в пиццерию. Ваня съел 5 кусков пиццы, Миша, Саша, Лена – по 3 куска, Катя – 2 куска, Маша - 1 кусок. 1) найдите все известные вам статистические характеристики этих данных. 2) Если бы Ваня съел не 5, а 7 кусков пиццы, как бы изменились эти величины?

4. На стройку с кирпичного завода привезли 20 упаковок кирпича. Чтобы проверить качество партии, из каждой упаковки вытащили случайным образом по кирпичу и измерили длину каждого. Ниже приведены полученные величины (в см): 20,5; 20,1; 21,3; 20,3; 19,8; 19,2; 20,1; 19,6; 20,2; 20; 20,5; 19,7; 19,9; 20,5; 19,6; 20,1; 19,4; 19,8; 19,1; 20,3.

а) Определите среднюю длину кирпича.

б) Найдите величину среднеквадратичного отклонения длинны кирпича от средней.

5. Пасечник заметил, что пчелы в двух его ульях производят мед не равномерно. Раз в 10 дней он вынимал соты из улья и заносил в таблицу массу (в кг) снятого меда, выработанного пчелами за десять дней.

а) Пчелы какого из ульев работают более стабильно? (Сделайте вывод, вычислив величину среднеквадратичного отклонения количества произведенного меда.)

б) Если в первом улье живет 100 пчел, а во втором 75 пчел, то сколько в среднем произвела каждая пчела меда за период с 19 по 28 августа каждая пчела 1 и 2 улья?

Интервалы времени	Масса меда (в кг)	
	1-й улей	2-й улей
с 20 по 30 апреля	11,4	11,9
с 1 апреля по 10 мая	12	10,8
с 11 по 20 мая	11,5	13,2
с 21 по 30 мая	11,7	12,6
с 31 мая по 9 июня	11	11,1
с 10 по 19 июня	10,6	11,4
с 20 по 29 июня	13,1	13,2
с 30 июня по 9 июля	12,8	12,9
с 10 по 19 июля	11,9	13,5
с 20 по 29 июля	13	10,9
с 30 июля по 8 августа	12,5	12,3
с 9 по 18 августа	12,9	11,7
с 19 по 28 августа	11,6	12
с 28 августа по 8 сентября	12	10,5

## Литература

1. А.Г.Мордкович. Алгебра и начала математического анализа.10-11 классы. В 2ч.Ч.1. Учебник для учащихся общеобразовательных учреждений (базовый уровень)/А.Г.Мордкович.-14-е изд., стер. – М.: Мнемозина, 2013. – 400с.: ил.
2. Алгебра и начала математического анализа.10-11 классы. В 2ч.Ч.2. Задачник для учащихся общеобразовательных учреждений (базовый уровень)/[А.Г.Мордкович и др.]; под ред. А.Г.Мордковича. -14-е изд., стер. – М.: Мнемозина, 2013. – 271с.: ил.
3. Бродский И.Л., Мешавкина О.С. Вероятность и статистика. 10-11 классы. Планирование и практикум: Пособие для учителя. – 104 с.; ил. (Школьное образование)
4. Бычков А.Г. Сборник задач по теории вероятностей, математической статистики и методам оптимизации: учебное пособие. – М: ФОРУМ, 2008. – 224 с.:ил.(Профессиональное образование)