

Элементы математической статистики

Черных А.А., преподаватель математики и информатики,
ОГБПОУ «Ангарский медицинский колледж», 2017

Статистика – наука, изучающая количественные стороны массовых явлений в конкретных условиях места и времени.

Математическая статистика – раздел математики, посвященный методам систематизации, обработки и исследования статистических данных для научных и практических выводов.

Основные понятия:

Генеральная совокупность – это совокупность всех однородных объектов, подлежащих изучению.

Выборочная совокупность (выборка) – это совокупность объектов, случайно отобранных из генеральной совокупности.

Объём совокупности (генеральной или выборочной) – это число её объектов.

Из 10 000 студентов института психологии тестируют 100 человек.

Генеральная совокупность :

студенты института психологии.

Объём генеральной совокупности :

$N = 10\ 000$

Выборочная совокупность :

тестируемые студенты.

Объём выборочной совокупности :

$N = 100$

Пусть в результате n испытаний
получены следующие значения СВ,
расположенные в неубывающем
порядке:

$$x_1, x_2, x_3, \dots, x_n,$$

$$\text{где } x_1 \leq x_2 \leq x_3 \leq \dots \leq x_n$$

Последовательность наблюдаемых значений,
записанных в неубывающем порядке,
называется **вариационным рядом**, а сами
значения x_i – **вариантами**.

Среди вариантов могут оказаться равные, тогда результат испытаний можно представить в виде таблицы:

x_i	x_1	x_2	\dots	x_k
n_i	n_1	n_2	\dots	n_k
w_i	w_1	w_2	\dots	w_k

где $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_k$

Перечень вариантов, записанных в возрастающем порядке, и соответствующих им частот (или относительных частот), называется **статистическим распределением выборки.**

$x_k - x_1$ — размах выборки

n_i — частота появления x_i

$$\sum_{i=1}^k n_i = n$$

w_i — относительная частота x_i

$$w_i = \frac{n_i}{n}$$

$$\sum_{i=1}^k w_i = 1$$

Пример 1: Ежедневное количество студентов, посещающих методический кабинет на протяжении ряда дней, следующее:
15, 17, 16, 18, 20, 21, 18, 17, 20, 15,
18, 17, 16, 19, 17, 16, 18, 19, 18, 19.
Составить статистическое распределение выборки.

Решение:

Объём выборки равен $n = 20$.

Вариационный ряд: 15, 16, 17, 18, 19, 20, 21.

Статистическое распределение выборки имеет вид:

Значение признака x_i	15	16	17	18	19	20	21
Частота n_i	2	3	4	5	3	2	1
Относительная частота $w_i = \frac{n_i}{n}$	$\frac{2}{20} = 0,1$	0,15	0,2	0,25	0,15	0,1	0,05

Статистическое распределение выборки может быть изображено графически – в виде **полигона и гистограммы:**

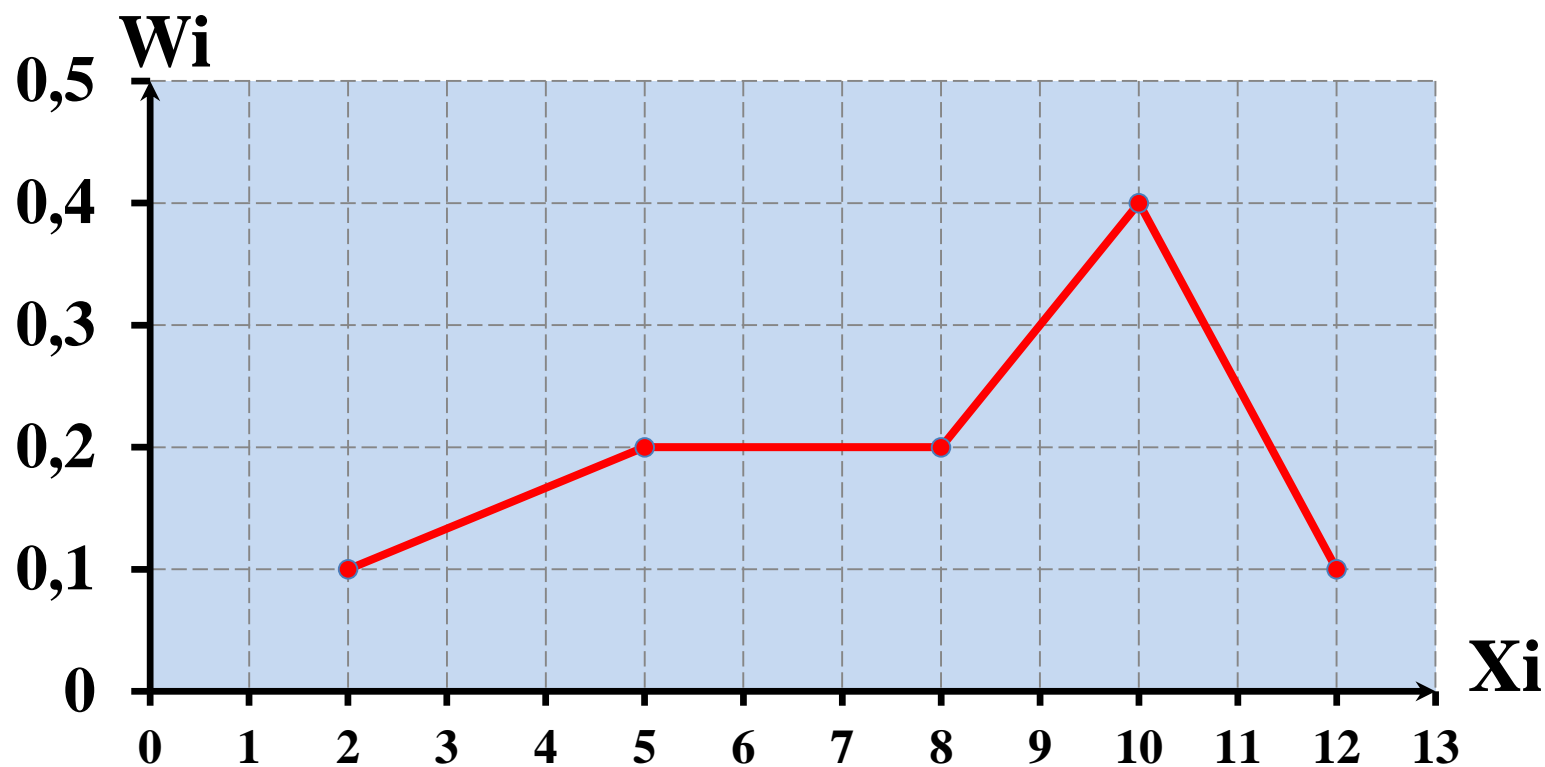
Полигоном частот называется ломаная, отрезки которой соединяют точки $(x_1; n_1), \dots, (x_k; n_k)$.

Полигоном относительных частот называется ломаная, отрезки которой соединяют точки $(x_1; w_1), \dots, (x_k; w_k)$.

Гистограммой частот называется ступенчатая фигура, состоящая из прямоугольников с основанием $h = x_i - x_{i-1}$ и высотами n_i/h (плотность частоты).

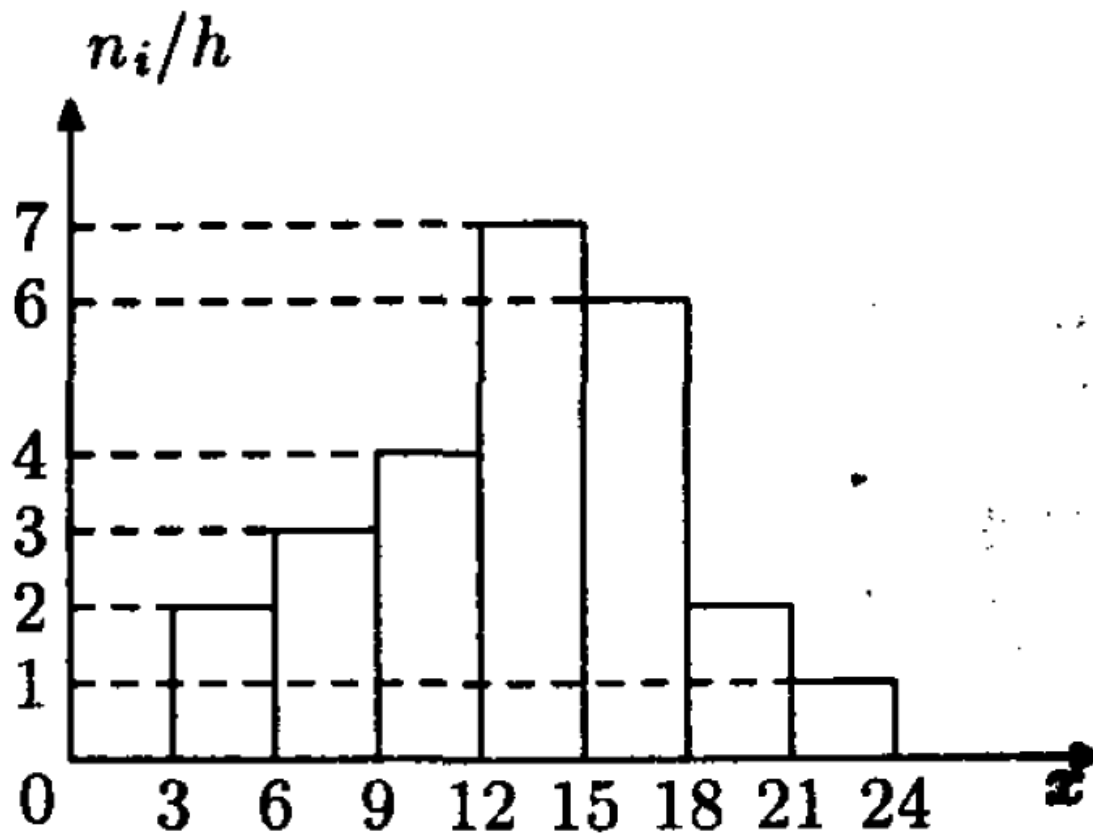
Пример 2. Построить полигон относительных частот статистического распределения:

x_i	2	5	8	10	12
w_i	0,1	0,2	0,2	0,4	0,1



Пример 3. Построить гистограмму частот распределения объёма $n = 75$.

$h = 3$	[3; 6]	[6; 9]	[9; 12]	[12; 15]	[15; 18]	[18; 21]	[21; 24]
n_i	6	9	12	21	18	6	3
$\frac{n_i}{h}$	$\frac{6}{3} = 2$	3	4	7	6	2	1



Построение полигонов и гистограмм позволяет произвести первичный анализ экспериментальных данных, а именно:

- ❖ По форме гистограммы сделать предположение о законе распределения СВ;
- ❖ Выявить наиболее часто встречающиеся значения исследуемой величины и разброс или отклонение относительно этого значения.

Характеристики положения и рассеяния статистического распределения.

Выборочная средняя :

$$\bar{x}_B = \sum_{i=1}^k x_i \cdot W_i$$

Характеристики положения и рассеяния статистического распределения.

Выборочная дисперсия:

$$D_{\text{В}} = \sum_{i=1}^k x_i^2 \cdot w_i - \bar{x}_{\text{В}}^2$$

Характеристики положения и рассеяния статистического распределения.

**Выборочное среднее квадратическое
отклонение:**

$$\sigma_{\text{в}} = \sqrt{D_{\text{в}}}$$

Пример 4. При проверке успеваемости факультета были выборочно протестированы 50 обучающихся, распределившихся по результатам тестирования следующим образом:

x_i	6	7	8	9	10
n_i	6	13	16	10	5
w_i	0,12	0,26	0,32	0,2	0,1

x_i – балл

n_i – количество обучающихся с данным баллом

Найдем выборочную среднюю:

$$\bar{x}_B = \sum_{i=1}^k x_i \cdot w_i =$$

$$= 6 \cdot 0,12 + 7 \cdot 0,26 + 8 \cdot 0,32 + 9 \cdot 0,2 + 10 \cdot 0,1 =$$

$$= 7,9$$

Найдем выборочную дисперсию:

$$D_B = \sum_{i=1}^k x_i^2 \cdot w_i - \bar{x}_B^2 =$$

$$= 6^2 \cdot 0,12 + 7^2 \cdot 0,26 + 8^2 \cdot 0,32 + 9^2 \cdot 0,2 + 10^2 \cdot 0,1 - 7,9^2 =$$

$$= 36 \cdot 0,12 + 49 \cdot 0,26 + 64 \cdot 0,32 + 81 \cdot 0,2 + 100 \cdot 0,1 - 62,41 =$$

$$= 4,32 + 12,74 + 20,48 + 16,2 + 10 - 62,41 = \mathbf{1,33}$$

**Найдем выборочное среднее
квадратическое отклонение:**

$$\sigma_B = \sqrt{D_B} = \sqrt{1,33} \approx 1,15$$

Если значения СВ заданы в виде интервалов – так называемый интервальный ряд, - то в качестве значений x_i берутся середины соответствующих интервалов.

Пример 5. Оценка проксемической характеристики – дистанции между партнерами.

Некто N собрал следующий статистический материал, касающийся дистанции при его общении с другими людьми в течение недели. Найти выборочную среднюю дистанции общения.

Вид общения	Расстояние (см)	Относительная частота
Интимное (общение близких людей)	0 – 45	0,3
Персональное (общение со знакомыми)	45 – 120	0,2
Социальное (официальное общение)	120 – 400	0,1
Публичное (выступление перед аудиторией)	400 - 750	0,4

Решение:

$$x_1 = \frac{0 + 45}{2} = 22,5$$

$$x_2 = \frac{45 + 120}{2} = 82,5$$

$$x_1 = \frac{120 + 400}{2} = 260$$

$$x_1 = \frac{400 + 750}{2} = 575$$

x_i	22,5	82,5	260	575
w_i	0,3	0,2	0,1	0,4

$$\begin{aligned}\bar{x}_B &= \sum_{i=1}^k x_i \cdot w_i = 22,5 \cdot 0,3 + 82,5 \cdot 0,2 + 260 \cdot 0,1 + 575 \cdot 0,4 = \\ &= 279,25 \text{ (см)}\end{aligned}$$

Вопросы для самоконтроля

1. Что такое статистика? Математическая статистика?
2. Перечислите основные задачи математической статистики.
3. Что такое генеральная и выборочная совокупность? Объём выборки?
4. Что такое вариационный ряд, варианта, размах выборки, частота, относительная частота.
5. Объясните понятие статистического ряда распределения.
6. Перечислите основные характеристики статистического распределения.
7. По каким формулам вычисляются основные характеристики статистического распределения?
8. Что такое полигон частот и гистограмма?

Список используемой литературы

1. Математика: учебник для студ. образоват. учреждений сред. проф. Образования/С.Г. Григорьев, С.В. Иволгина; под ред. В.А. Гусева. – 8-е изд., стер. – М.: Издательский центр «Академия», 2012. – 416 с.
2. Математика для психологов: учеб. пособие для студентов вузов/ А.В. Ганичева, В.П. Козлов. – М.: Аспект Пресс, 2005. – 239 с.