

Опорный конспект

- 1. Решите задачи на повторение (5, 6 слайд).**
- 2. Запишите алгоритм обработки статистических данных.**
- 3. Перечислите основные понятия математической статистики и их определения.**
- 4. Какой основной метод обработки статистических данных?**
- 5. Выполнить решение задач на 19 слайде.**

Математическая статистика – раздел математики, изучающий методы сбора, систематизации и обработки результатов наблюдений массовых случайных явлений.

Задачи математической статистики:

- Сбор, хранение и обработка информации.*
- Установление и исследования различного рода зависимостей на основании экспериментальных данных.*
- Изучение вероятностных характеристик моделей реальных явлений.*
- Разработка прогнозов, оценка их достоверности.*

Повторение:

Средним арифметическим ряда чисел называется частное от деления суммы этих чисел на число слагаемых.

Размахом ряда чисел называется разность между наибольшим и наименьшим из этих чисел.

Модой ряда чисел называется число, которое встречается в данном ряду чаще других. (M_o)

Медианой упорядоченного ряда чисел с нечётным числом членов называется число, записанное посередине, а медианой упорядоченного ряда чисел с чётным числом членов называется среднее арифметическое двух чисел, записанных посередине.

Медианой произвольного ряда чисел называется медиана соответствующего упорядоченного ряда. (M_e)

Возраст учеников, которые занимаются в кружке по математике, дан ниже:

13	14	14	13	12	14	13	13	13	14	13	13	13	14	14
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Найдите:

- а) размах; б) моду; в) медиану;
г) среднее значение.*

В ряду чисел

3, 8, 15, 30, __, 24

*пропущено одно число. Найди его
если:*

- а) среднее арифметическое ряда равно 18;*
- б) размах ряда равен 40;*
- в) мода ряда равна 24.*

Статистические данные – сведения о числе объектов, какой-либо обширной совокупности, обладающих теми или иными признаками (число студентов, родившихся в 1980 г; балловые результаты ЕГЭ в городе).

Основной метод обработки статистических данных – выборочный метод.

Алгоритм обработки статистических данных:

- *Упорядочение и группировка данных;*
- *Составление таблиц распределения данных;*
- *Построение графиков распределения данных;*
- *Расчет основных числовых характеристик статистических данных.*

Основные понятия математической статистики:

Генеральная совокупность – совокупность всех исследуемых объектов.

Выборка – совокупность случайно отобранных объектов генеральной совокупности.

Объем выборки – число объектов выборки или генеральной совокупности.

Размах выборки – разность между наибольшим и наименьшим значением числовой выборки.

Выборку, представляющую собой неубывающую последовательность чисел, называют вариационным рядом.

Основные понятия математической статистики:

Вариантой называют каждое полученное значение данных конкретного измерения: $x_1; x_2; \dots; x_k$

Кратностью или частотой варианты называют число, показывающее, сколько раз варианта встретилась в данном измерении: $n_1; n_2; \dots; n_k$

Относительной частотой значений выборки называют отношение частоты варианты к объему выборки:

$$\frac{n_1}{n}; \frac{n_2}{n}; \dots; \frac{n_k}{n}$$

Статистическим рядом называют последовательность пар:

$$(x_1, n_1); (x_2, n_2); \dots; (x_k, n_k)$$

Таблицы распределения данных:

Статистическое распределение (статистический ряд):

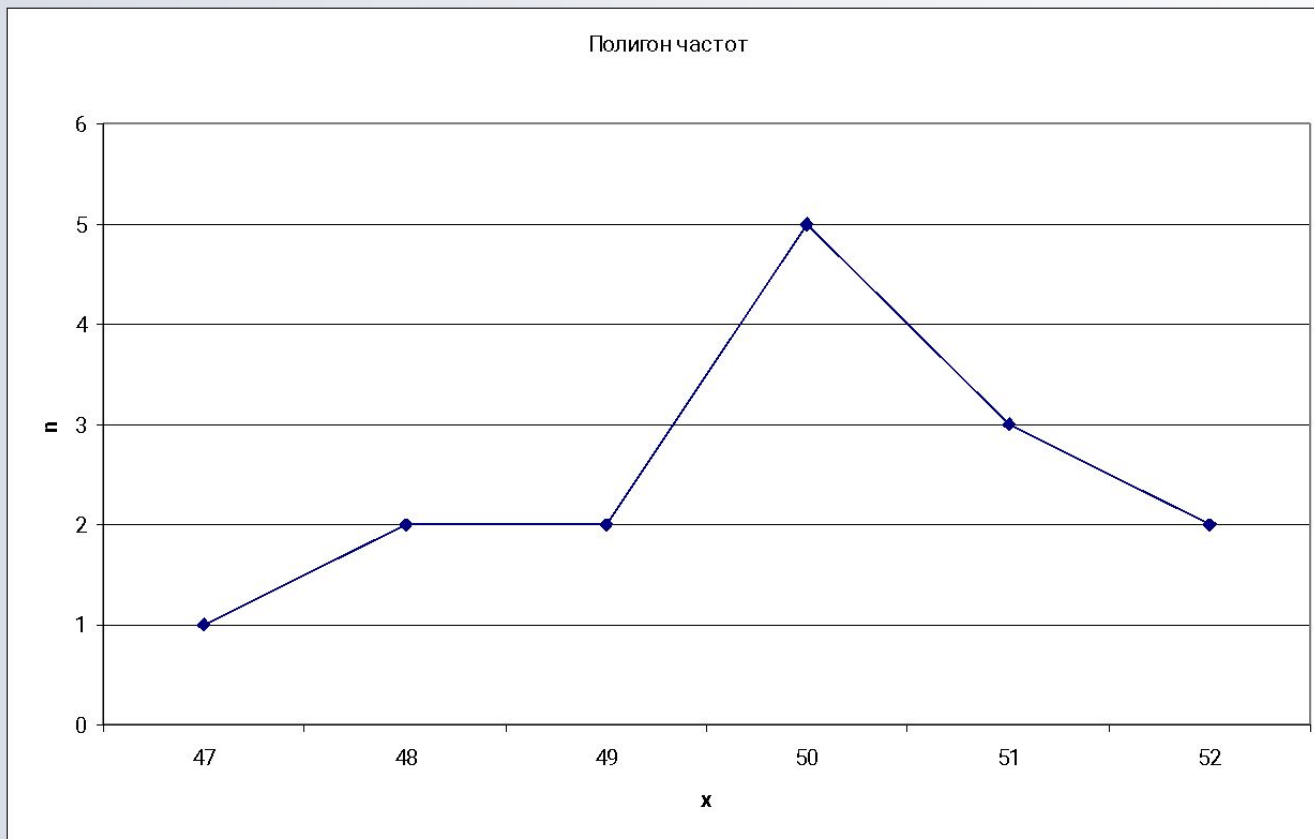
x_1	x_2	x_3	...	x_k
n_1	n_2	n_3	...	n_k

Выборочное распределение (статистический ряд):

x_1	x_2	x_3	...	x_k
$\frac{n_1}{n}$	$\frac{n_2}{n}$	$\frac{n_3}{n}$...	$\frac{n_k}{n}$

Графическое представление распределения данных:

Полигон частот:



Графическое представление распределения данных:

Гистограмма частот:



Выборочные характеристики:

Выборочным математическим ожиданием (выборочным средним) называют среднее арифметическое значение выборки:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \quad \text{ИЛИ} \quad \bar{x} = \frac{n_1 x_1 + n_2 x_2 + n_3 x_3 + \dots + n_k x_k}{n}$$

Выборочной дисперсией называют среднее арифметическое квадратов отклонений значений выборки от выборочного среднего:

$$S_0 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} \quad \text{ИЛИ}$$

$$S_0 = \frac{n_1 \cdot (x_1 - \bar{x})^2 + n_2 \cdot (x_2 - \bar{x})^2 + n_3 \cdot (x_3 - \bar{x})^2 + \dots + n_k \cdot (x_k - \bar{x})^2}{n}$$

Среднее квадратичное отклонение: $\sigma = \sqrt{S_0}$

Практическое применение средних величин (выборочных характеристик) в медицине:

- *Для оценки состояния здоровья, например параметров физического развития (средний рост, средний вес, средний объем емкости легких), соматических показателей (средний уровень сахара в крови, средний пульс).*
- *Для оценки организации работы лечебно-профилактических и санитарно-противоэпидемических учреждений, а также деятельности отдельных врачей и других медицинских работников.*
- *Для оценки состояния окружающей среды.*

Задача:

В результате измерения роста 7 - летних детей (см) получена выборка: 118, 121, 115, 125, 125, 117, 124, 120, 120, 119, 121, 119, 122, 127, 118, 120, 123, 130, 123, 116, 124, 127, 120, 122.

Представьте эти данные в виде дискретного статистического ряда распределения и постройте полигон частот. Постройте гистограмму, если число частичных промежутков равно 5. Определите среднее значение роста детей и отклонение от него.

Решение:

Составим вариационный ряд выборки: 115, 116, 117, 118, 118, 119, 119, 120, 120, 120, 120, 121, 121, 122, 122, 123, 123, 124, 124, 125, 125, 127, 127, 130.

Запишем статистическое распределение (статистический ряд):

115	116	117	118	119	120	121	122	123	124	125	127	130
1	1	1	2	2	4	2	2	2	2	2	2	1

Построим полигон частот:



Определим ширину частичного промежутка: $h = \frac{x_{\max} - x_{\min}}{5} = 3$

Число попаданий выборки в частичные промежутки соответственно равны: $[115, 118)$ -4, $[118, 121)$ -7, $[121, 124)$ - 6, $[124, 127)$ - 4, $[127, 130)$ - 3. Соответственно высоты прямоугольников равны: $4/3$; $7/3$; 2 ; $4/3$; 1 .

Построим гистограмму частот:



Вычислим среднее значение выборки (средний рост детей):

$$\bar{x} = \frac{1 \cdot 115 + 1 \cdot 116 + 1 \cdot 117 + 2 \cdot 118 + \dots + 2 \cdot 124 + 2 \cdot 125 + 2 \cdot 127 + 1 \cdot 130}{24} = 121,5 \text{ см}$$

Вычислим дисперсию:

$$S_0 = \frac{1 \cdot (115 - 121,5)^2 + 1 \cdot (116 - 121,5)^2 + \dots + 2 \cdot (127 - 121,5)^2 + 1 \cdot (130 - 121,5)^2}{24} = 13,06$$

Вычислим среднее квадратичное отклонение:

$$\sigma = \sqrt{S_0} = \sqrt{13,06} = 3,6 \text{ см}$$

Задачи для самостоятельного решения:

1. Из продукции, произведенной фармацевтической фабрикой за месяц, случайным образом отобраны 15 коробочек некоторого гомеопатического препарата, количество таблеток в которых оказалось равным соответственно 50, 51, 48, 52, 51, 50, 49, 50, 47, 50, 51, 49, 50, 52, 48. Представьте эти данные в виде дискретного статистического ряда распределения, вычислите среднее значение выборки, моду и размах.

2. Проведены измерения вязкости крови у 9 больных. Значения относительной вязкости крови у больных составили: 5, 4, 3, 2, 6, 3, 4, 8, 10. Представьте эти данные в виде дискретного статистического ряда, вычислите среднее значение относительной вязкости крови и отклонение от него.