# ЗДРАВСТВУЙТЕ, УВАЖАЕМЫЕ СТУДЕНТЫ ПГС-2!

**Тема занятия «Элементы математической статистики»**

**Задание: записать теоретический материал по теме с примерами из данной презентации в рабочую тетрадь.**

**Проверка работы будет на следующей очной паре математики.**

**Если возникли вопросы – пишите teza5@mail.ru**

**Фазылова Е.Х.**

# Элементы математической статистики.

**Статистика — это точная наука, изучающая методы сбора, анализа и обработки данных, которые описывают массовые действия, явления и процессы**

**Математическая статистика – это раздел математики, изучающий методы сбора, систематизации и обработки результатов наблюдений случайных массовых явлений с целью выявления существующих закономерностей.**

## СТАТИСТИКА ИЗУЧАЕТ:

- численность отдельных групп населения страны и ее регионов;
- производство и потребление разнообразных видов продукции;
- перевозку грузов и пассажиров различными видами транспорта;
- природные ресурсы и многое другое.

**Результаты статистических исследований широко используются для практических и научных выводов.**

# ОПРЕДЕЛЕНИЕ:

**Ряд данных – это ряд результатов каких-либо измерений.**

**Например:**

**1 ) измерения роста человека**

**2) Измерения веса человека (животного)**

**3)Показания счетчика (электроэнергии, воды, тепла…)**

**4) Результаты в беге на стометровку**

**И т.д.**

***Объемом*** **ряда данных называется количество всех данных.**

**Например:** **дан ряд чисел 1; 3; 6; -4; 0 объём его будет равен 5. Почему?**

# Показатели вариации

***Размахом*** вариационного ряда называют абсолютную величину разности между максимальными и минимальными значениями (вариантами) изучаемого признака

$$R = x_{\max} - x_{\min}$$

дискретный вариационный ряд

| $x_i$ | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|
| $m_i$ | 12 | 22 | 28 | 30 | 8 |

интервальный вариационный ряд

| $x_i$ | 100-120 | 120-140 | 140-160 | 160-180 |
|---|---|---|---|---|
| $m_i$ | 9 | 16 | 11 | 4 |

# Размах вариации

- **Пример 1.**

  В ряду чисел 3,8,15,30,_,24 пропущено одно число. Найдите его, если размах ряда равен 40. В ответе указать сумму возможных вариантов

- **Решение:**

$x_{max} = 30$

$x_{min} = 3$

$R = 30 - 3 = 27 \neq 40$

$40 = x - 3 \Rightarrow x = 43$ **либо** $40 = 30 - x \Rightarrow x = -10$

$43 + (-10) = 33$

**Ответ: 33**

Пример.

Для выборки определить объем, размах, найти статистический ряд и выборочное распределение:

3, 8, -1, 3, 0, 5, 3, -1, 3, 5

Объем: $n = 10$, размах $= 8 - (-1) = 9$

*Статистический ряд:*

| | | | | | |
|---|---|---|---|---|---|
| $x_i$ | -1 | 0 | 3 | 5 | 8 |
| $n_i$ | 2 | 1 | 4 | 2 | 1 |

*Выборочное распределение:*

| | | | | | |
|---|---|---|---|---|---|
| $x_i$ | -1 | 0 | 3 | 5 | 8 |
| $\frac{n_i}{n}$ | 0,2 | 0,1 | 0,4 | 0,2 | 0,1 |

(убеждаемся 0,2 + 0,1 + 0,4 + 0,2 + 0,1 = 1)

**Модой ряда данных называется** число ряда, которое встречается в этом ряду наиболее часто.

**Медиана с нечётным числом членов** – это число, записанное посередине.

**Медиана с чётным числом членов** - это среднее арифметическое двух чисел, записанных посередине.

Например: определить медиану ряда чисел.

1) 6; -4; 5; -2; -3; 3; 3; -2; 3. Ответ: -3
2) -1; 0; 2; 1; -1; 0;2; -1. Ответ: 0

**Полигон распределения** – это зависимость абсолютной частоты варианта mi от значения варианта i x . Эту зависимость можно представить в виде таблицы.

## Рассмотрим простейшую задачу данного типа.

**Задача:**
Измерение роста детей младшей группы детского сада представлено выборкой: 92, 96, 95, 96, 94, 97, 98, 94, 95, 96.
Найдем некоторые характеристики этой выборки.

**Решение**
Размер выборки (число измерений; $N$): 10.
Наименьшее значение выборки: 92. Наибольшее значение выборки: 98.
Размах выборки: 98 – 92 = 6.
Запишем ранжированный ряд (варианты в порядке возрастания):
92, 94, 94, 95, 95, 96, 96, 96, 97, 98.

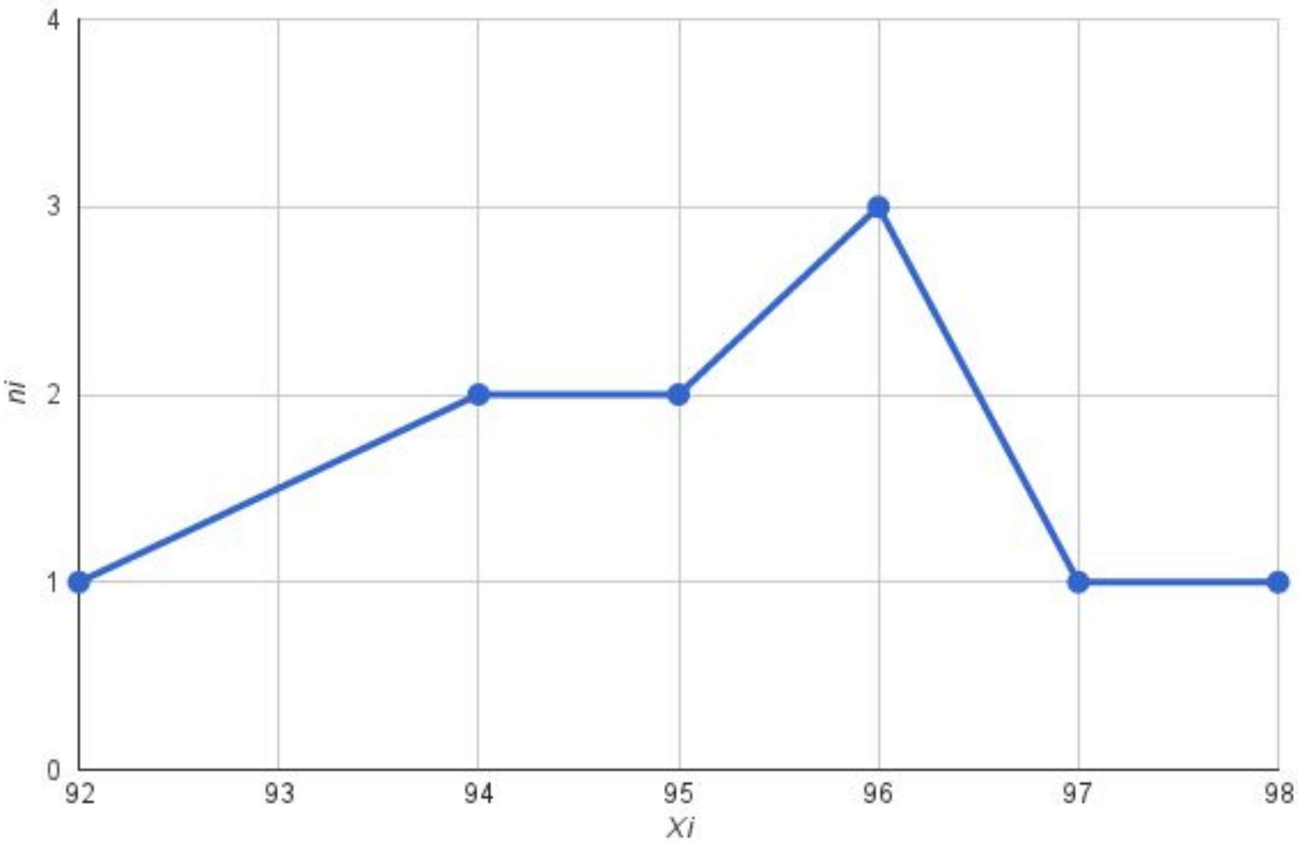
Сгруппируем ряд и запишем в таблицу (каждой варианте поставим в соответствие число ее появлений):

| $x_i$ | 92 | 94 | 95 | 96 | 97 | 98 | *N* |
|---|---|---|---|---|---|---|---|
| $n_i$ | 1 | 2 | 2 | 3 | 1 | 1 | 10 |

Вычислим относительные частоты и накопленные частоты, результат запишем в таблицу:

| $x_i$ | 92 | 94 | 95 | 96 | 97 | 98 | Итого |
|---|---|---|---|---|---|---|---|
| $n_i$ | 1 | 2 | 2 | 3 | 1 | 1 | 10 |
| | 0,1 | 0,2 | 0,2 | 0,3 | 0,1 | 0,1 | 1 |
| Накопленные частоты | 1 | 3 | 5 | 8 | 1 | 10 | |

Построим полигон частот выборки (отметим на графике варианты по оси OX, частоты по оси OY, соединим точки линией).



**Математическое ожидание** - случайной

величины XX (обозначается M(X)M(X) или реже E(X)E(X)) характеризует среднее значение случайной величины (дискретной или непрерывной). Мат. ожидание - это первый начальный момент заданной СВ.

Математическое ожидание относят к так

называемым **характеристикам положения** распределения (к которым также принадлежат мода и медиана). Эта характеристика описывает некое усредненное положение случайной величины на числовой оси. Скажем, если матожидание случайной величины - срока службы лампы, равно 100 часов, то считается, что значения срока службы сосредоточены (с обеих сторон) от этого значения (с тем или иным разбросом, о котором уже говорит дисперсия).

**Пример 1.** *Вычислить математическое ожидание дискретной случайной величины X, заданной рядом:*

| | | | | | |
|---|---|---|---|---|---|
| Xi | −1 | 2 | 5 | 10 | 20 |
| Pi | 0.1 | 0.2 | 0.3 | 0.3 | 0.1 |

Используем формулу для м.о. дискретной случайной величины:

$M(x)=\sum x_i \cdot p_i$.

Получаем:

$M(x)=\sum x_i \cdot p_i=-1\cdot 0.1+2\cdot 0.2+5\cdot 0.3+10\cdot 0.3+20\cdot 0.1=6.8$.

## Математическое ожидание дискретной случайной величины.

Говоря простым языком, это *среднеожидаемое значение* при многократном повторении испытаний. Пусть случайная величина принимает значения с вероятностями соответственно. Тогда математическое ожидание данной случайной величины равно <u>сумме произведений</u> всех её значений на соответствующие вероятности:

или в свёрнутом виде:

**<u>математическое ожидание</u>** – это уже НЕ СЛУЧАЙНАЯ величина.

**Пример:**

Мистер Х играет в европейскую рулетку по следующей системе: постоянно ставит 100 рублей на «красное». Составить закон распределения случайной величины – его выигрыша. Вычислить математическое ожидание выигрыша и округлить его до копеек. Сколько *в среднем* проигрывает игрок с каждой поставленной сотни?

***Справка:*** *европейская рулетка содержит 18 красных, 18 чёрных и 1 зелёный сектор («зеро»). В случае выпадения «красного» игроку выплачивается удвоенная ставка, в противном случае она уходит в доход казино*

***Решение:*** поскольку игрок выигрывает в 18 случаях из 37, то закон распределения его выигрыша имеет следующий вид:

| | | |
|---|---|---|
| $X$ | $0$ | $200$ |
| | $\frac{19}{37}$ | $\frac{18}{37}$ |

Вычислим математическое ожидание:

[illegible]

Таким образом, с каждой поставленной сотни игрок в среднем проигрывает 2,7 рубля.

Вычислим, например, математическое ожидание случайной величины – количества выпавших на игральном кубике очков:

[illegible] очка

В чём состоит вероятностный смысл полученного результата? Если подбросить кубик достаточно много раз, то *среднее значение* выпавших очков будет близко к 3,5 – и чем больше провести испытаний, тем ближе. Это статистическая вероятность.

Теперь вспомним нашу гипотетическую игру:

| | | | |
|---|---|---|---|
| $U$ | –5 | 2,5 | 10 |
| | 0,5 | 0,4 | 0,1 |

Возникает вопрос: а выгодно ли вообще играть в эту игру? …у кого какие впечатления? Так ведь «навскидку» и не скажешь! Но на этот вопрос можно легко ответить, вычислив математическое ожидание, по сути – *средневзвешенный* по вероятностям выигрыш:

[illegible]

таким образом, математическое ожидание данной игры **проигрышно**.

Не верь впечатлениям – верь цифрам!

**Упорядоченными рядами** данных называются ряды, в которых данные расположены по какому то правилу.

***Как упорядочить ряд чисел?***

Записать числа так, чтобы каждое последующее число было

не меньше (не больше) предыдущего; или записать

некоторые названия «по алфавиту»…

***Таблица распределения данных –*** это таблица упорядоченного ряда, в котором вместо повторений одного и того же числа записывается количество повторений.

**И наоборот,** если известна таблица распределения, то можно составить упорядоченный ряд данных.

**Номинативный ряд данных –** это НЕ ЧИСЛОВЫЕ ДАННЫЕ, а например, имена; названия; номинации…

Например:

список  финалистов чемпионатов мира по футболу с 1930

года: Аргентина, Чехословакия, Венгрия, Бразилия, Венгрия,

Швеция, Чехословакия, ФРГ, Италия, Нидерланды,

Нидерланды, ФРГ, ФРГ, Аргентина, Италия, Бразилия,

***Вероятность случайного события*** равна дроби, в знаменателе

которой содержится число всех равновероятных возможностей, из

которых состоит достоверное событие, а в числителе – число тех

возможностей, при которых рассматриваемое событие происходит.

***Процентная частота*** **- ( частота ▪ 100% )**

**Например:** если частота результата равна 5:19= 0,263157…, то

процентная частота будет равна:  0,263 · 100 = 26,3%

Часто ответы для процентных частот могут быть

***Группировка данных –*** применяется когда различных результатов измерений слишком много. Т.е их объединяют в группы.

При группировке различных данных информация становится менее точной.

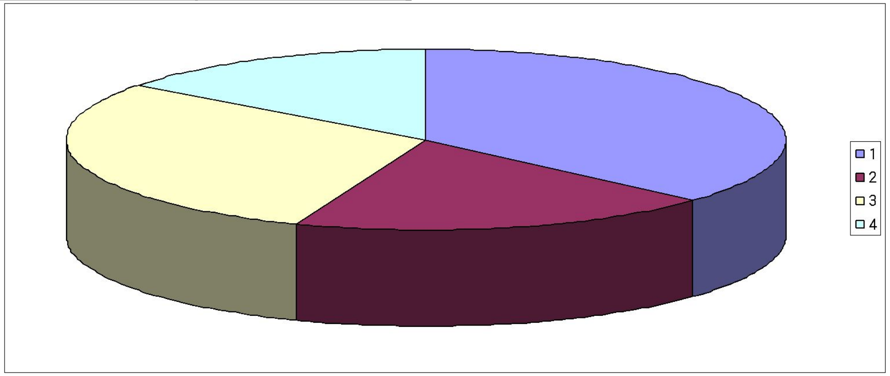
# СПОСОБЫ ОБРАБОТКИ ДАННЫХ:

**Таблица**

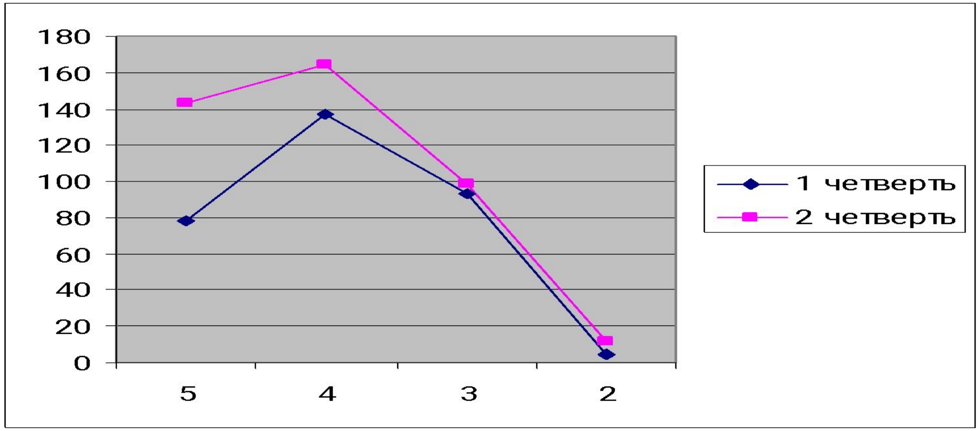
| Год обучения | 1-4 кл. | 5-9 кл | 10-11 кл |
|---|---|---|---|
| 2007-2008 | 250 | 254 | 80 |
| 2008-2009 | 253 | 248 | 78 |
| 2009-2010 | 258 | 240 | 73 |

**Диаграмма круговая**

**( каламбер )**



# График




# Гистограмма (столбчатая диаграмма)