**Вопросы “Анализ данных”.**

1. [Предметная область математической статистики. Особенности, задачи. Сравнительный анализ.…](#bookmark=kix.5mqg6i90bj66)
2. [Классификация погрешностей измерения](#bookmark=id.1fob9te)
3. [Случайные величины. Дискретные случайные величины.](#bookmark=id.3znysh7)
4. [Случайные величины. Непрерывные случайные величины.](#bookmark=id.tyjcwt)
5. [Генеральная и выборочная совокупности](#bookmark=id.3dy6vkm)
6. [Вариационный ряд: определение, графическое представление.](#bookmark=id.1t3h5sf)
7. [Алгоритм построения интервального вариационного ряда](#bookmark=id.4d34og8)
8. [Алгоритм построения дискретного вариационного ряда](#bookmark=id.2s8eyo1)
9. [Числовые характеристики вариационных рядов.](#bookmark=id.17dp8vu)
10. [Законы распределения случайных величин.](#bookmark=id.3rdcrjn)
11. [Дисперсия. Эксцесс. Показатель асимметрии.](#bookmark=id.26in1rg)
12. [Числовые характеристики случайных величин.](#bookmark=id.lnxbz9)
13. [Алгоритм вычисления числовых характеристик выборочного распределения.](#bookmark=id.35nkun2)
14. [Квантили, квартили, перцентили.](#bookmark=id.1ksv4uv)

В билете 1 теоретический вопрос и 1 практический вопрос (по выполненным лабораторным работам). На экзамен принести оформленные и распечатанные лабораторные работы.

1. **Предметная область математической статистики. Особенности, задачи. Сравнительный анализ. Примеры использования методов.**

***Математическая статистика*** - позволяет решать задачи принятия решений в условиях вероятностной неопределенности с достаточным распространением программного обеспечения.

*Основная цель математической статистики* - получение и обработка данных для статистически значимой поддержки процесса принятия решения, *например при решении задачи планирования или прогнозирования.*

Методы математической статистики можно разделить на *описательные* (дескриптивные) и *аналитические*. ***Описательные*** методы позволяют *описать реальные наблюдения* с помощью таблиц, графиков, характеристик положения (мода, медиана), характеристик рассеяния (среднее квадратичное отклонение, среднее линейное отклонение, дисперсия).

***Аналитические*** методы позволяют на основании выборочных наблюдений *сделать статистически значимые выводы* о наличии закономерностей для всей совокупности. Они основываются на соответствующих вероятностных моделях, предполагающих нормальное (или другое известное) - распределение совокупности изучаемого признака - ***методы параметрической статистики.***Другим направлением аналитических методов являются методы ***непараметрической статистики,***которые не опираются на нормальное (или любое другое) распределение и не используют его свойства.

***Основные направления методов*** математической статистики:

1. *Выборочное наблюдение* - решает задачу обобщения на всю совокупность результатов, полученных при изучении ее части *(например рейтинг политиков, анкетирование).*
2. *Проверка статистических гипотез -* позволяет ответить на вопрос о достоверности принимаемого решения *(например, обоснованность рейтинга популярности).*
3. *Дисперсионный анализ -* изучает влияние факторных признаков на результативность *(например, зависит ли производительность труда рабочего от стажа, возраста, стажа и возраста).*
4. *Корреляционно-регрессионный анализ -* позволяет выявить связи и построить модели зависимости *(например, какая зависимость существует между спросом на продукцию и курсом доллара).*
5. *Анализ временных рядов -* рассматривает последовательность чисел, зависящих от времени, и изучает их свойства *(например количество пятен на Солнце, характеризующее его активность по годам, курс доллара по дням).*
6. *Последовательный анализ* А.Вальда - один из первых разработанных методов, дающий возможность осуществлять контроль качества выпускаемой продукции.

В математической статистике предполагается, что результаты опытных данных и наблюдений являются реализацией различных случайных процессов, имеющих те или иные законы распределения, а иногда и детерминированные составляющие (регрессионный анализ). Отсюда вытекают ***основные задачи математической статистики***:

1. организация наблюдений,
2. нахождение по результатам выборочных наблюдений оценок числовых характеристик всей совокупности и исследование точности приближения (выборочный метод),
3. решение вопроса согласования результатов оценивания с опытными данными (проверка статистических гипотез),
4. оценка существенности влияния фактора признаков на результативный (дисперсионный анализ),
5. выявление аналитической зависимости между наблюдениями факторных и результативных признаков (корреляционно-регрессионный анализ).

Математическая статистика ***позволяет обосновать ответы на вопросы:***

1. случайно или закономерно изучаемое явление *(влияет ли доза внесения удобрений на урожайность?),*
2. как зависит результативный признак от факторного *(как зависит урожайность от дозы внесения удобрений при прочих равных условиях?),*
3. сколько необходимо провести наблюдений для объективного суждения об изучаемом явлении,
4. какой фактор сильнее влияет на результат *(влияет вид удобрения на урожайность?).*

**Сравнительный анализ**

***Сравнение*** - один из важных и наиболее распространенных приемов изучения взаимосвязей в развитии общественных явлений. По существу с него начинается анализ работы предприятий (хозяйств) и их внутрихозяйственных подразделений.

Непременным условием сравнения является сопоставимость показателей. Многие статистические показатели могут быть несравнимы в силу различий: их предметного содержания, календарных сроков, типа производства, форм собственности и хозяйствования, неодинаковой методологии планирования, учета и калькуляции себестоимости одинаковой продукции, единицам измерения в различных отраслях, по формам реализации и цен продукции, специализации, изменению территориальных границ, рельефно-почвенных и климатических условий. В одном случае продукция может быть валовая, в другом - товарная, а в третьем - чистая.

В экономическом анализе различают следующие виды сравнительного анализа:

1) горизонтальный - используется для определения абсолютных и относительных отклонений фактического уровня исследуемых показателей от базового;

2) вертикальный (структурный) -- используется для изучения структуры экономических явлений и процессов путем расчета:

а) удельного веса частей в общем целом (удельного веса машин и оборудования в общей стоимости основных производственных фондов);

б) соотношения частей целого между собой (основных фондов и оборотных средств);

в) выявления влияния отдельных факторов на уровень результативных показателей путем сравнения их удельного веса до и после изменения соответствующего фактора;

3) трендовый -- используется для определения относительных темпов роста и прироста показателей за ряд лет по сравнению с уровнем базисного года;

4) одномерный -- сопоставления делаются по одному или нескольким показателям одного объекта или по одному показателю сопоставляются несколько объектов;

5) многомерный -- производится сопоставление результатов деятельности нескольких предприятий (объединений) по широкому спектру показателей для комплексной оценки хозяйственной деятельности. Такая задача встает всегда, когда надо дать обобщающую сравнительную оценку результатам хозяйствования нескольких предприятий. Это делают вышестоящие органы управления, а также инвесторы для оценки степени финансового риска.

# 

1. **Классификация погрешностей измерения.**

Любые измерения, как бы точно они не выполнялись, всегда сопровождаются ***погрешностями****,* т.е отклонениями (), измеренных значений

Эти погрешности измерения часто называют *ошибками измерения.*

Существуют различные ***виды классификации ошибок измерения*:**

* По ***источникам возникновения*** погрешности измерений:
  + инструментальные,
  + внешние или погрешности среды,
  + личные,
* По ***точности*** результата измерений:
  + *равноточные* - однородные результаты, полученные при измерениях одним и тем же инструментом (или разными, но с одним классом точности), одним и тем же или равноценными методами, одинаковым числом приемов и в одинаковых условиях.
  + *неравноточные -*  это результаты измерений, когда указанные условия не соблюдаются.
* По ***закономерностям появления погрешности*** измерений:
  + *грубые погрешности* или промахи- это погрешность, существенно превышающая ожидаемую при данных условиях измерения.
  + *систематические погрешности* - это погрешности составляющие погрешности измерения, остающиеся постоянными или закономерно изменяющимися при повторных измерениях одной и той же величины. Систематические погрешности обусловлены такими факторами, как неправильная установка на нуль измерительного прибора или влияние определенных условий внешней среды.
  + *случайные погрешности* - это погрешности, составляющие погрешность измерения, изменяющиеся случайным образом при повторных измерениях одной и той же величины. Их нельзя исключить в каждом из результатов измерений.

Для оценки точности результатов измерения принято вычислять ***среднее квадратичное отклонение*** *(погрешность или ошибку)****.*** Величина среднего квадратичного отклонения S определяется в основном случайной и систематической погрешностями. Систематическая погрешность, как правило, компенсируется поправкой которая также имеет случайный характер изменения и определяется средним квадратичным отклонением

1. **Случайные величины. Дискретные случайные величины.**

***Случайная величина (СВ)*** - это величина, которая в результате опыта может принимать те или иные значения, причем до опыта нельзя сказать, какое значение она примет.

***СВ*** - это действительная функция, определенная на пространстве элементарных событий .

СВ обозначаются буквами латинского алфавита **X,Y,Z.**

СВ делятся на **три типа**:

* дискретные,
* непрерывные,
* смешанные (дискретно-непрерывные).

***Дискретная случайная величина (ДСВ)*** может принимать конечное или бесконечное счетное число значений.

Пусть Х - дискретная случайная величина, которая принимает значения: с некоторой вероятность где i = 1,2,...,n. Тогда можно говорить о вероятности того, что случайная величина Х приняла значение .

Значения и соответствующие , представляются в виде таблицы:

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | ... |  |
|  |  | ... |  |

Эта таблица является одной из форма задания ДСВ. Обычно случайные величины располагаются в возрастающем порядке. *Основное свойство таблицы заключено в том, что* ***сумма вероятностей равна 1***.

***Дискретная случайная величина*** может быть представлена в виде многоугольника распределения - фигуры, состоящей из точек с координатами (), соединенных отрезками.

# 

1. **Случайные величины. Непрерывные случайные величины.**

***Случайная величина (СВ)*** - это величина, которая в результате опыта может принимать те или иные значения, причем до опыта нельзя сказать, какое значение она примет.

***СВ*** - это действительная функция, определенная на пространстве элементарных событий .

СВ обозначаются буквами латинского алфавита **X,Y,Z.**

СВ делятся на **три типа**:

* дискретные,
* непрерывные,
* смешанные (дискретно-непрерывные).

***Непрерывная случайная величина -*** это величина, возможные значения которой непрерывно заполняют какой-то промежуток.

В общем случае случайная величина X задается функцией распределения F(x), которая выражает зависимость того, что X принимает значение, меньше, чем x: .

Функция распределения обладает следующими свойствами:

* F(x) не убывает, т.е
* Вероятность попадания случайной величины X в интервал a<x<b:
* Если все значения непрерывной случайной величины принадлежат некоторому промежутку от a до b, то:

Если функция распределения F(x) везде непрерывна и имеет производную, то случайная величина называется *непрерывной в узком смысле слова*, или просто *непрерывной*.

Вероятность того, что НСВ X попадает на участок (a,b) может быть вычислена по одной из двух формул:

1. **Генеральная и выборочная совокупности**

Понятие ***выборки*** используется, когда надо изучить какие-либо свойства совокупности объектов. *(Свойства объектов можно разделить на качественные и количественные)*

***Генеральная совокупность*** - это конечная, либо бесконечная совокупность объектов или наблюдений, все элементы которой подлежат изучению при статистическом анализе.

В математической статистике *генеральная совокупность* часто понимается *как совокупность всех мыслимых наблюдений, которые могли быть произведены при выполнение данного комплекса условий.*

Понятие генеральной совокупности аналогично понятию случайной величины (закону распределения вероятностей), так как обе они полностью определяются заданными условиями.

***Объем совокупности*** - число объектов в наблюдении этой совокупности.

***Выборочная совокупность*** *или выборка* - часть отобранных объектов из генеральной совокупности (*например, Пусть из партии 100 пачек масла для исследования выбрано 10 пачек. Тогда объем генеральной совокупности N=100, а выборка n=10).*

Сущность выборочного метода в математической статистике заключается в том, чтобы по определенной части генеральной совокупности (выборке) судить о ее свойствах в целом.

Выборочный метод является единственно возможным в случае *бесконечной* генеральной совокупности или когда исследование связано с уничтожением (гибелью) наблюдаемых объектов (*например, исследование действия вирусов на подопытных животных*). Для того, чтобы по выборке можно было адекватно судить о случайной величине, она должна быть *представительной (репрезентативной).*

***Репрезентативность выборки*** обеспечивается случайностью отбора ее элементов, так как все элементы генеральной совокупности должны иметь одинаковую вероятность попадания в выборку.

***Два способа образования выборки***:

* *повторная выборка:* когда каждый элемент, случайно отобранный и исследованный, возвращается в общую совокупность и может быть отобран повторно,
* *бесповторная выборка:* когда отобранный элемент не возвращается в общую совокупность.

1. **Вариационный ряд: определение, графическое представление.**

Пусть некоторый признак генеральной совокупности описывается случайной величиной X.   
Рассмотрим выборку объема *n* из генеральной совокупности. Элементы этой выборки представляют собой значения величины Х.

На первом этапе статистической обработки производят *ранжирование* выборки, т.е упорядочивание чисел по возрастанию.

***Варианты*** - различные элементы выборки.

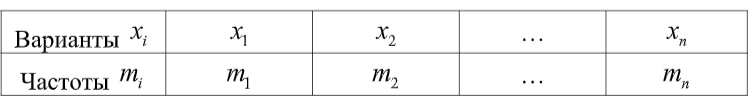
***Частота варианты***- это число показывающие, сколько раз эта варианта встречается в выборке.

***Частостью, относительной частотой*** или ***долей варианты*** называют число .

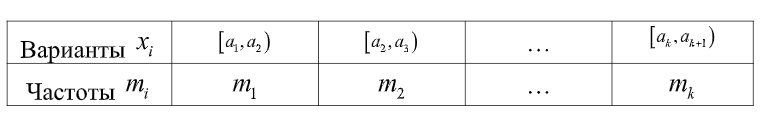
Частоты и частости называются ***весами***.

***Вариационный ряд*** - это ряд вариант расположенных в порядке возрастания значений, с соответствующими им весами. ВР подразделяются на:

* *дискретные* - вариационный ряд, представляющий выборку значений дискретной случайной величины.



* *интервальные* (непрерывные) - вариационный ряд, представляющий собой выборку непрерывной случайной величины.



***Графически вариационный ряд*** можно представить как***:***

* *полигон* - представляет собой ломаную, соединяющую точки плоскости с координатами:
  + для дискретного вариационного ряда:.
  + для интервального:
* *гистограмму* - служит только для интервальных вариационных рядом и имеет вид ступенчатой фигуры из прямоугольников с основаниями, равными длине интервалов , и высотами, равными частотам интервалов.
* *кумуляту* - представляет собой ломаную, соединяющую точки с координатами:
  + для дискретного ряда:
  + для интервального:
* *эмпирическую функцию распределения*- функция, значение которой в точке *х* равно накопленной частоте, т.е:

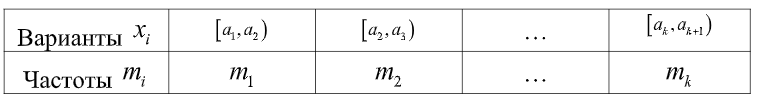
Для интервального ряда указываются не конкретные значения вариант, а только их частоты на интервалах.

* (эмпирической плотности распределения)

Пусть *х* - некоторое число. Тогда количество вариант значения которых меньше *х*, называется ***накопленной частотой:***

Отношение накопленной частоты к общему числу наблюдений *n* называется ***накопленной частостью:***

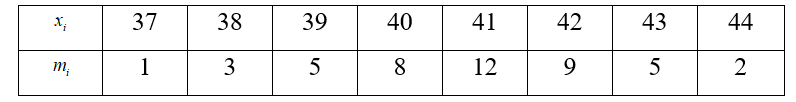
1. **Алгоритм построения интервального вариационного ряда**
2. Располагаем значения в порядке их возрастания.
3. Вычислить количество интервалов *k* по формуле Стерджерса:
4. Вычислить длину интервала:
5. Разбить множество значений вариант на полуинтервалы, т.е произвести *группировку.*
6. Подсчитать число вариант, попавших в полуинтервал . Получают значение частот
7. Считаем значение относительной частоты (частости)
8. Представить интервальный ряд таблицей вида: (в таблице не хватает графы с wi)



1. Изобразить графически

*Если варианта находится на границе интервала, то ее присоединяют к правому интервалу.*

1. **Алгоритм построения дискретного вариационного ряда**
2. Различные значения признака располагаем в порядке их возрастания
3. Под каждым из них записываем его частоту.
4. Находим относительную частоту(частость)
5. Получаем таблицу признаков и частот: (в таблице не хватает графы с wi)



1. Изобразить графически

1. **Числовые характеристики вариационных рядов.**

Основная характеристика вариационного ряда - это ***средняя арифметическая***, которая также называется ***выборочным средним****.*

Для дискретного вариационного ряда средняя арифметическая равна:

Для интервального:

Показателем вариации является ***вариационный размах R:***

Более содержательным являются ***меры рассеивания вариант вокруг своих средних величин.***

***Средним линейным отклонением*** вариационного ряда называется величина:

***Выборочной дисперсией*** называется средняя арифметическая квадратов отклонений вариант от их выборочной средней:

Для несгруппированного ряда () выборочная дисперсия равна

Дисперсию часто называют эмпирической или выборочной, подчеркивая, что она, в отличие от дисперсии случайной величины, находится опытным или статистических путем.

***Среднее квадратичное отклонение*** - мера вариации признака, равная корню из дисперсии:

В статистическом анализе рассматривается также ***коэффициент вариации,*** равные процентному отношению выборочного среднего квадратичного отклонения к выборочной средней:

1. **Законы распределения случайных величин.**

Случайная величина считается полностью заданной, если на числовой оси указаны ее возможные значения и установлен закон распределения.

***Законом распределения случайной величины***называется *соотношение*, устанавливающее связь *между возможными значениями случайной величины* и *соответствующими вероятностями*.

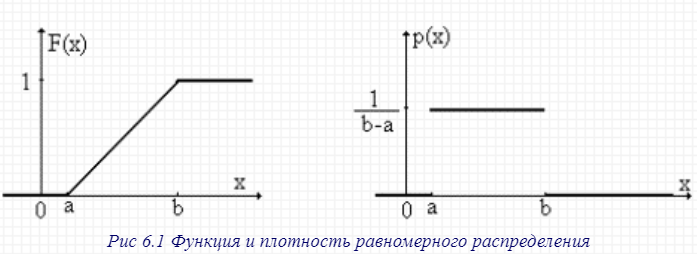
Про случайную величину говорят, что она распределена по данному закону, или подчинена данному закону распределения. В качестве законов распределения используются *ряд вероятностей, функция распределения, плотность вероятности, характеристическая функция.*

*Закон распределения дает полное вероятное описание случайной величины.* По закону распределения можно судить до опыта о том какие возможные значения случайной величины будут появляться чаще, а какие – реже.

Вид функций называют *законами распределения случайной величины.*

1. ***Равномерное распределение.***

Так называют распределение случайной величины, которая может принимать любые значения в интервале (a,b), причем вероятность попадания ее в любой отрезок внутри (a,b) пропорциональна длине отрезка и не зависит от его положения, а вероятность значений вне интервала (a,b), равна 0.

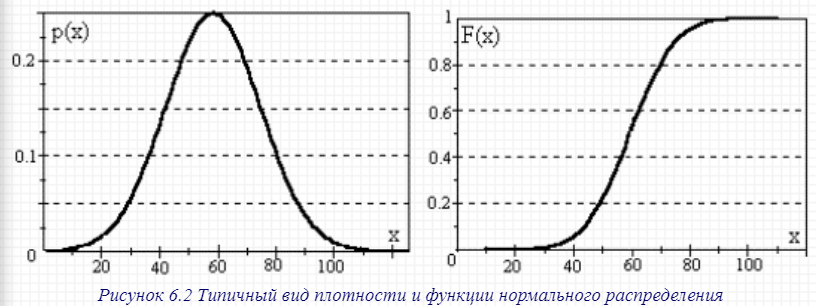


1. ***Нормальное распределение.***

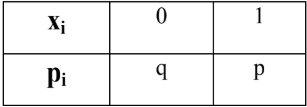
Распределение с плотностью, описываемой формулой:

, где *a=M(x)*

называется *нормальным распределением*

*.*

1. ***Закон распределения Бернулли.*** Случайная величина Х, распределенная по закону Бернулли (индикаторная случайная величина) принимает значения 1 - успех или 0 - неудача, с вероятностями p и q (p+q=1).



1. ***Биноминальный закон распределения.*** Случайная величина X принимает значения: 0, 1, 2,..., n, с вероятностью определяемой по формуле Бернулли:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | ... | k | ... | n |
|  |  |  |  | ... |  | ... |  |

1. **- *распределение Пирсона.*** Пусть одинаково распределены по нормальному закону случайной величины, являющиеся взаимозависимыми, для которых математическое ожидание равно нулю, а среднее квадратическое отклонение 1, тогда сумма квадратов этих случайных величин носит название случайной величины **-** xu-квадрат с v=n степенями свободы:

При возрастании числа степеней свободы v=n распределение медленно приближается к нормальному закону распределения. На практике обычно применяют не плотность вероятности, а квантили распределения.

*Квантилью* распределения, отвечающей заданному уровню значимости а, называется такое значение при котором вероятность того, что превысит значение равна а.

При n>30 распределение практически не отличается от нормального.

1. ***t-распределение Стьюдента***

Это распределение (критерий) используется для сравнения полученных результатов с нормальным распределением

1. ***F-распределение Фишера-Снедекора.***

F-распределение применяется для сравнение двух оценок [дисперсий](https://www.matematicus.ru/teoriya-veroyatnosti/dispersiya-sluchajnoj-velichiny) и более, для проверки гипотез в регрессионном анализе.

<https://www.matematicus.ru/category/zakony-raspredeleniya-sluchajnoj-velichiny> - сайт, куда можно посмотреть если все плохо)

1. **Дисперсия. Эксцесс. Показатель асимметрии**.

Для изучения и описания закономерностей распределения вероятностей случайных величин в математичсекой статистике используются соответствующие статистический *характеристики распределений.* Основным из них являются *средние значения и меры рассеивания.*

Средними значениями являются: ***среднее арифметическое*** (мат.ожидание), *медиана* и *мода*. К мере рассеивания относятся: ***дисперсия****,* среднеквадратичное отклонение, размах, коэффициент вариации. - Это не особо относится к вопросу, лучше начинать сразу с определения

***Дисперсия*** - характеризует степень рассеивания элементов массива вокруг среднего значения.

***Среднее арифметическое*** является *показателем среднего уровня вокруг которого колеблются значения* членов вариационного ряда, поэтому для вычисления дисперсии требуется его знать.

***Среднее арифметическое -*** это оценка теоретического среднего (математического ожидания *M)* генеральной совокупности величин и определяется по формуле:

***Момент случайной величины*** - числовая характеристика распределения данной случайной величины.

***Среднее арифметическое*** (или *мат.ожидание*) называется ***первым моментом распределения.***Он характеризует как бы центр тяжести функции плотности вероятностей *P(x).*

***Дисперсия*** является математическим ожиданием квадрата отклонения случайной величины от ее среднего значения - математического ожидания, и называется ***вторым центральным моментом*** или ***моментом инерции.***

***Третий центральный момент***относительно среднего значения характеризует ***асимметрию*** распределения относительно математического ожидания и является, следовательно, ***мерой асимметрии*** *(скошенности).*

Когда график вариационного ряда (распределение набора данных) скошен в правую сторону больше, чем в левую, то мы говорим, что распределение имеет ***правостороннюю скошенность (асимметрию),*** соответственно, скос в левую сторону дает ***левостороннюю скошенность (асимметрию).***

При симметричном распределении *мода = медиане = средней арифметической.*

Чем больше асимметричен график, тем больше расхождение между медианой, модой и средней арифметической.

Наиболее простой мерой ***скошенности вариационного ряда*** будет *разность между средней арифметической и модой*:   
Если > 0 - асимметрия правосторонняя,

Если < 0 - асимметрия левосторонняя.

Для сравнения асимметрии в нескольких рядах удобнее пользоваться относительным показателем где - среднее квадратичное отклонение.

Другой более распространенный показатель асимметрии основан на использовании центрального момента третьего порядка. В симметричных вариационных рядах нечетные центральные моменты равны нулю, при имеет место правосторонняя асимметрия,- левосторонняя. В качестве показателя асимметрии используется величина, которая называется ***нормированным моментом третьего порядка:***

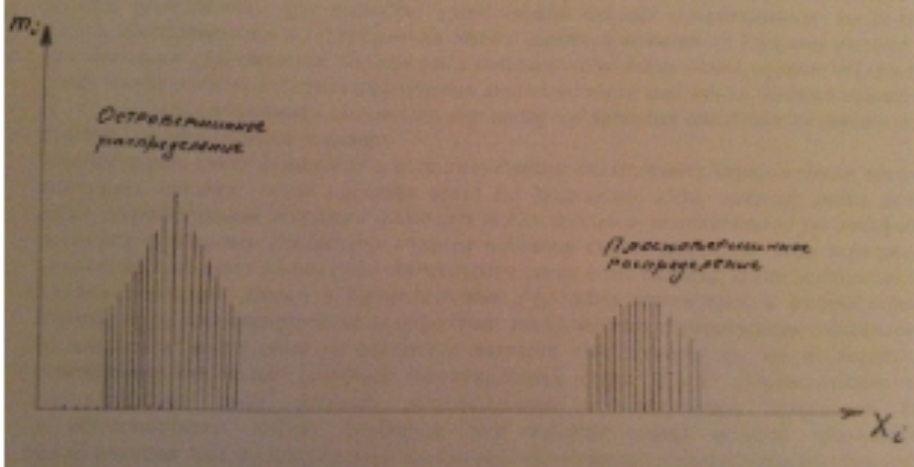
***Четвертый центральный момент*** характеризует свойство *островершинности* или *пологости* кривой плоскости вероятностей. За характеристику этого свойства принимается безразмерная величина *E,* называется ***эксцессом***.

В качестве показателя эксцесса используется величина:

- ***нормированный момент четвертого порядка.***

Если > 0, то эксцесс считают положительным (график островершинный)

Если < 0, то эксцесс считают отрицательным (график плосковершинный).

****

1. **Числовые характеристики случайных величин.**

На практике нет необходимости характеризовать величины полностью. Обычно достаточно указать только ***числовые характеристики распределения*** - т.еотдельные числовые параметры распределения***.***

*Числовые характеристики случайных величин имеют разные формы, в зависимости от того, дискретная или непрерывная случайная величина перед нами:*

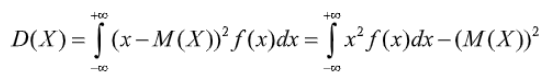
1. ***Для дискретных:***

* *Характеристики положения ряда распределения*:
  + **Математическое ожидание** M(X) - среднее значение случайной величины, т.е *сумма парных произведений случайной величины на существующую вероятность.*
  + **Медиана** - это значение случайной величины, которое делит таблицу распределения на две части таким образом, что вероятность попадания в одну из них равна 0,5. Медиана обычно не определяется для дискретных случайных величин.
  + **Мода -** это значение СВ, имеющее наиболее вероятное значение.
* *Характеристики рассеяния*:
  + **Дисперсия** D(X) - это математическое ожидание квадрата отклонения СВ от ее математического ожидания. Дисперсия служит для характеристики рассеяния СВ относительно ее математического ожидания и характеризует форму кривой распределения.
  + **Среднее квадратичное отклонение -** характеристика разброса СВ, равняется корню из дисперсии.

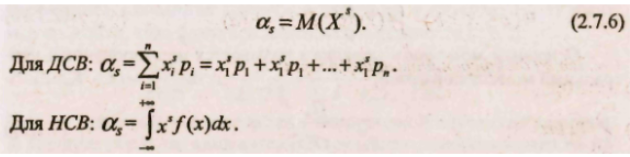
1. ***Для непрерывных:***

* ***Математическое ожидание*** - среднее значение случайной величины, Для непрерывной случайной величины X определяется по формуле:

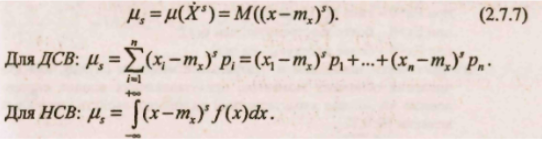
Если НСВ Х определена на интервале (a;b), то:

* ***Мода*** НСВ Х -это значение СВ, имеющее наиболее вероятное значение; будет определяться как максимум ее дифференциальной функции:
* ***Медиана*** определяется как значение случайной величины, которые делит площадь под дифференциальной функцией на две равные части:
* ***Дисперсия*** -это математическое ожидание квадрата отклонения СВ от ее математического ожидания. Дисперсия служит для характеристики рассеяния СВ относительно ее математического ожидания и характеризует форму кривой распределения.  
  

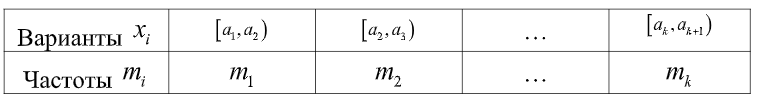
1. ***Моменты случайных величин (начальные и центральные):***Кроме характеристик положения и рассеяния существует ряд других числовых характеристик:
   * *Начальным моментом* порядка s называется математическое ожидание степени s СВ X:



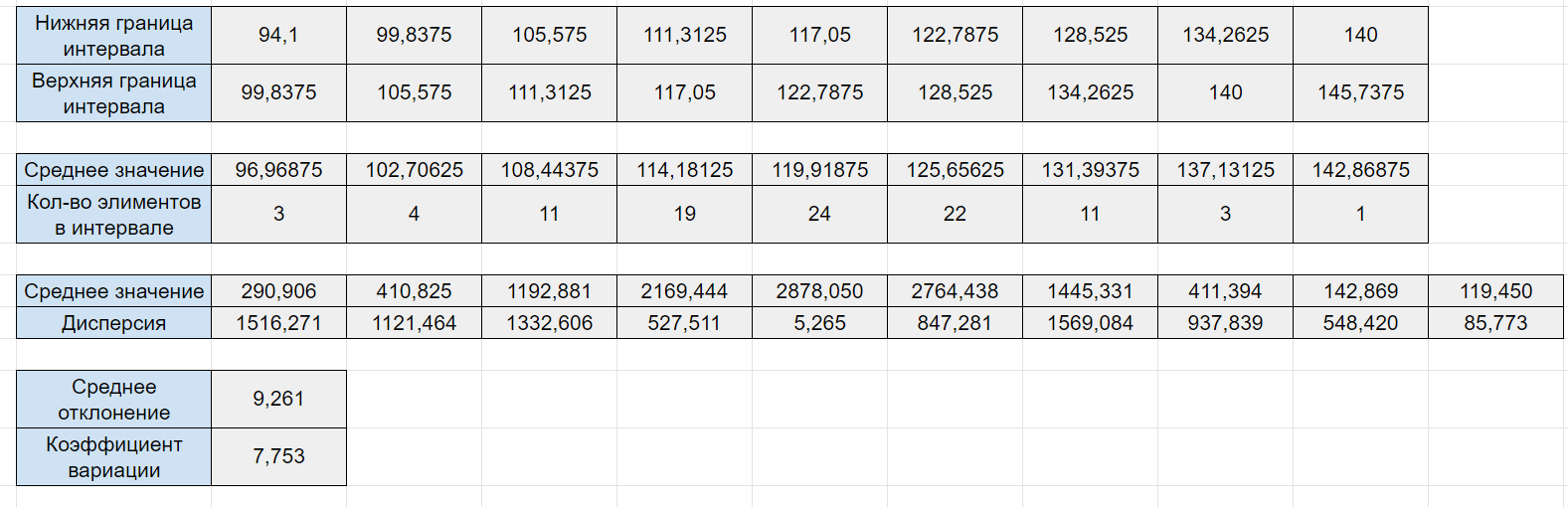
* + *Центральным моментом* порядка s СВ Х называется математическое ожидание степени s, соответствующей центрированной СВ:



1. **Алгоритм вычисления числовых характеристик выборочного распределения.**
2. Располагаем значения в порядке их возрастания.
3. Вычислить количество интервалов *k* по формуле Стерджерса:
4. Вычислить длину интервала:
5. Разбить множество значений вариант на полуинтервалы, т.е произвести *группировку.*
6. Подсчитать число вариант, попавших в полуинтервал . Получают значение частот
7. Представить интервальный ряд таблицей вида:



1. Далее находим среднее арифметическое всего ряда, путем сложения произведения середин интервалов на количество элементов в нем и деления суммы на общее количество элементов.
2. Находим дисперсию: возводим в квадрат разность мат.ожидания и среднего значения интервала, и умножаем полученное значение на количество элементов в интервале. Дисперсия распределения будет равна сумме дисперсии рядов, деленной на общее количество элементов.
3. Находим среднее отклонение, равное корню из дисперсии.
4. Находим коэффициент вариации, равный частному среднего отклонения на среднее арифметическое, умноженное на 100%.



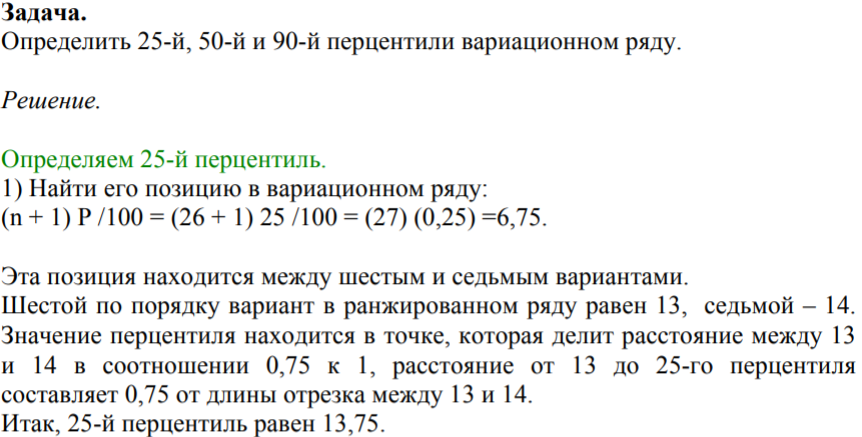
1. Находим моду и медиану, по каким то формулам, которых я не помню, но суть в том, что мода - это самый часто встречающийся элемент, а медиана - это середина ряда, так что ищем их по определениям, можно сказать

1. **Квантили, квартили, перцентили.**

***Квантили вариационного ряда*** - это варианты, занимающие определенное место в ранжированной совокупности. К числу квантилей, наиболее часто используемых в математической статистике, относят *перцентили, квартили, децили у медиан,* которые характеризуют структуру вариационного ряда.

***P-ый перцентиль вариационного ряда*** - это значение признака, слева которого лежит P% вариантов ряда. Позиция P-го перцентиля задается как *(n+1)P/100*, где *n -* число вариантов ряда.

***Перцентиль*** - это значение признака в определенной позиции ранжированного ряда, мера относительной позиции варианта в ряду.



В математической статистике наиболее часто применяются квантили, которые делят вариационный ряд на четыре равные части - ***квартили:*** *первый квартиль, второй квартиль, третий квартиль и четвертый квартиль*, которые обозначаются как

***Первый квартиль (25-й перцентиль)***,*нижний квартиль* - это значение признака в вариационном ряду, слева от которого лежит (или 25%) всех вариантов.

***Второй квартиль (50-й перцентиль)****, средний квартиль*- это медиана

***Медиана*** - значение признака ряда, относительно которого вариационный ряд делится на две равные по числу вариантов части.

***Третий квартиль (75-й перцентиль)****, верхний квартиль* - это точка, слева от которой находится или 75% вариантов ряда.

Также, применяют квантили, которые делят вариационный ряд на десять равных частей - *децили.* В дискретном вариационном ряду их значения определяются соответственно как 10,20,...,90 перцентили.

***Мода*** - это значение признака, наиболее часто встречающегося в вариационном ряду.