

Методы обработки и анализа статистической информации

*Автор: Евгения Бородулина
группа: е6610*

Статистика - отрасль знаний, в которой излагаются общие вопросы сбора, измерения и анализа массовых статистических (количественных или качественных) данных.

Порядок преобразований первоначально полученной информации:

- 1) Данные измерений упорядочивают и группируют
- 2) После группировки составляют таблицы распределения данных
- 3) Таблицы распределения позволяют построить графики распределения данных
- 4) Составляют своего рода паспорт данных измерения.

Математической основой любой отрасли статистики, в том числе и юридической является закон больших чисел, который гласит о том, что статистические закономерности формируются и отчетливо проявляются лишь в массовом процессе и при достаточно большом числе единиц совокупности. При большом числе наблюдений экспериментальные результаты исследований становятся близкими к теоретическим.

В правовой статистике выделяют три метода

1. Метод массового статистического наблюдения заключается в изучении большого количества правонарушений, что позволяет выявить объективные закономерности в сфере преступлений, в правоприменительной деятельности судов, прокуратуры, милиции и других правоохранительных органов. Для получения объективных результатов статистическое наблюдение должно охватывать либо всю (генеральную) совокупность изучаемых явлений, либо такую ее часть, которая была бы достаточно представительной (репрезентативной) и позволяла бы сказать, что результаты, выявленные на основе неполных данных, имеют такую-то ошибку.

2. Сводка и группировка данных наблюдения - следующий метод (этап) юридической статистики. Его сущность заключается в сведении и группировке данных, полученных путем статистического наблюдения, по тем или иным признакам, для того чтобы каждая группа представляла собой определенную качественную однородность. Например, совокупность изученных преступлений группируется по объектам посягательства (против личности, экономики, государства и др.), по содержанию мотивации (корыстные, насильственные и т.д.) или по субъектам преступлений (полу, возрасту, социальному положению, прежней судимости). Сводка и группировка данных позволяет увидеть структуру изучаемых явлений, их сходства и различия. Данный метод позволяет увидеть единство количественного и качественного в той или иной совокупности.

Виды группировок

По задачам систематизации данных

Типологические

Структурные

Аналитические

По числу группировочных признаков

Простые

Сложные

Комбинационные

По используемой информации

Первичные

Вторичные

Многомерные

3. Статистический количественный анализ позволяет выявить закономерности и взаимозависимости массовых правовых, криминологических и социологических явлений. Результаты статистического исследования на этом этапе выражаются в процентах, коэффициентах, индексах и других обобщающих показателях, не включающих в себя частные, индивидуальные или случайные отклонения. В них раскрываются основные тенденции, типичные черты, корреляции, характеристики.

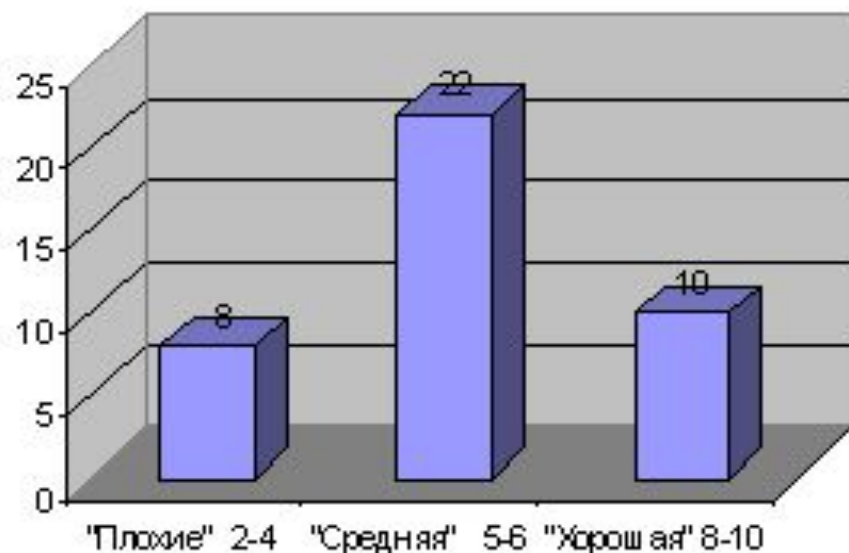
Приведем пример. Если число учтенных преступлений в разных странах в 1960 г. принять за 100% и соотнести с этими показателями данные 1990 г., то обнаружится главная закономерность развития преступности: за 30 лет зарегистрированная преступность в США увеличилась в 7,2 раза, в Англии и Уэльсе – в 6,1, во Франции – более чем в 5, в СССР – в 3,7, в ФРГ – в 2,8, а в Японии в 1,5 раза. В реальной жизни кривая динамики преступности не была строго линейной: преступность росла, сокращалась, оставалась на уровне предыдущего года, вновь росла и сокращалась.

Гистограммы.

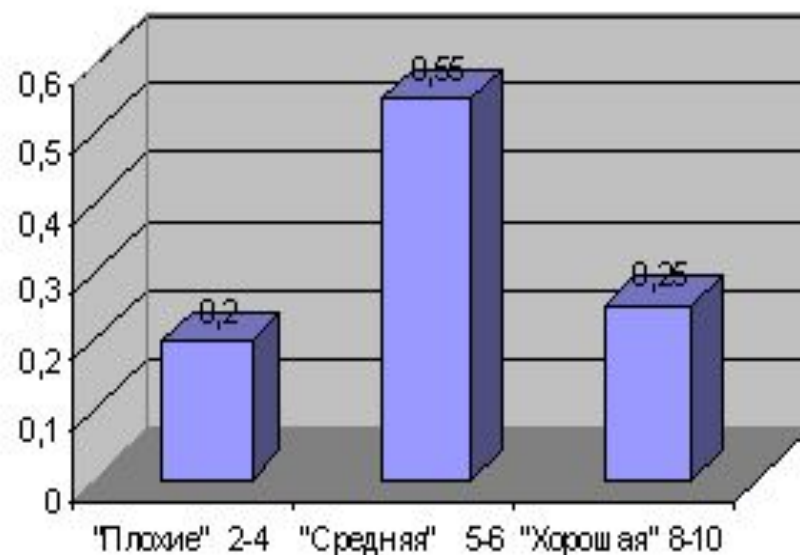
Гистограммное распределение это ступенчатая фигура.

Их строят, когда приходится первоначальную группировку данных проводить более крупными «блоками». Это когда количество вариантов измерения велико.

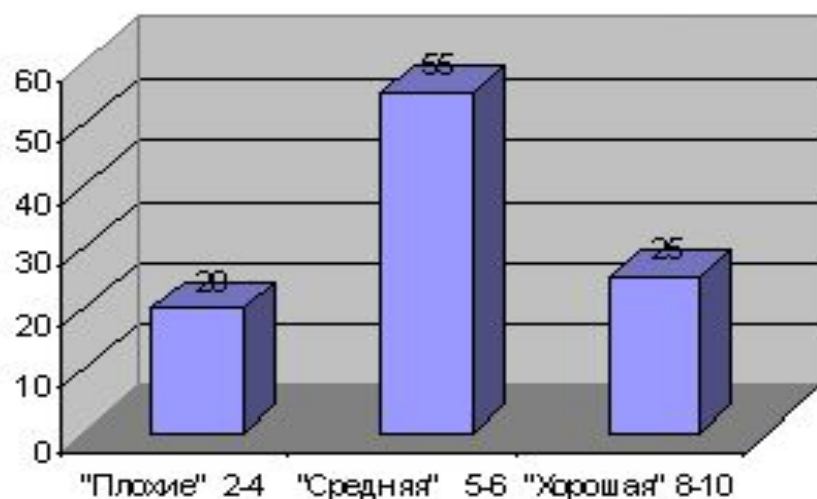
Гистограмма распределения кратностей



Гистограмма распределения частот



Гистограмма распределения частот (%)



Статистические программы:

«Statistica», «MiniTab», «Microsoft Office»,
«MatLab», «Tecplot» и др.

Варианта - каждое полученное значение
данных конкретного измерения

Кратность варианта - число,
показывающее, сколько раз эта варианта
встретилась в данном измерение

Объём измерения-общее количество данных
измерения

Частота варианты - это кратность данной варианты разделить на объём измерения

Мода ряда данных- это варианта, которая встречается в ряду чаще остальных вариантов

Размах измерения - это разность между наибольшей вариантой и наименьшей вариантой: разность $X_n - X_1$

Измерение

У 50 студентов независимо попросили назвать любую цифру. Получились следующие данные:

2 1 3 3 5 5 3 8 1 7

1 5 7 5 3 8 0 4 7 3

3 9 6 9 1 6 9 1 2 3

9 8 7 0 5 1 3 1 3 9

6 2 3 5 9 2 5 1 5 7

Шаг 1.

Упорядочивание состоит в том, что все данные выписываются последовательно в некотором порядке. Следует начать с наименьшего результата, а закончить наибольшим результатом. Таким выписыванием получается группировка информации.

Шаг 2

(Получение табличных данных)

Она состоит из двух строчек. В клетке первой строки выписывают поочерёдно все различные значения, реально полученные в измерении. Во второй строке таблицы под каждой вариантой из первой строки записывают кратность варианты. Эту таблицу называют таблицей распределения. Последний столбец приписывают для контроля. Другими словами, в этой клетке указывают объём измерения.

Шаг 3

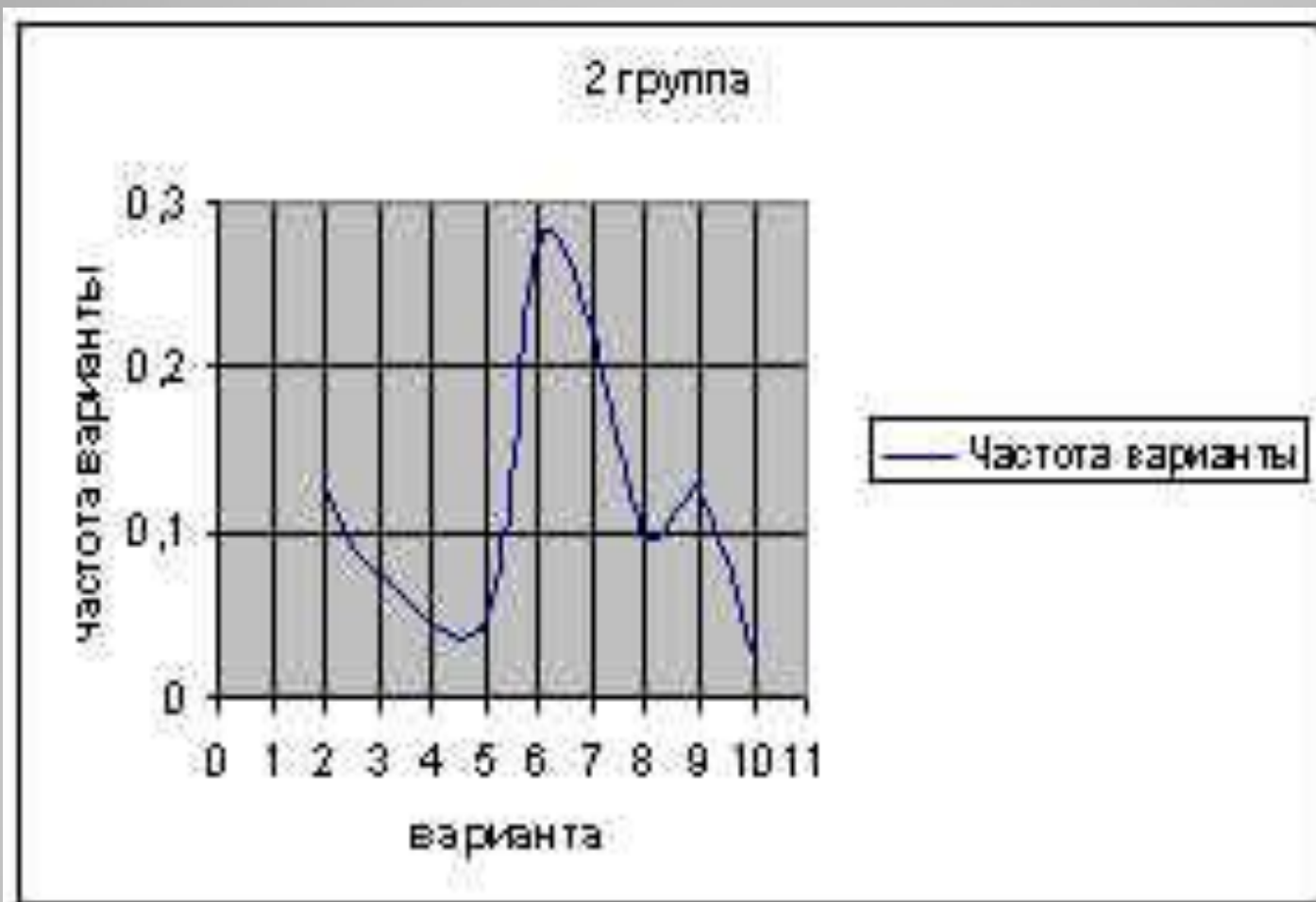
Обработки данных измерения состоит в графическом, визуальном изображении имеющейся информации. Хорошо известен табличный способ задания функций.

Аргументами этих функций являются варианты, а значениями функций – либо кратности вариантов, либо частоты, либо их процентные частоты, в зависимости от того какой график мы будем строить.

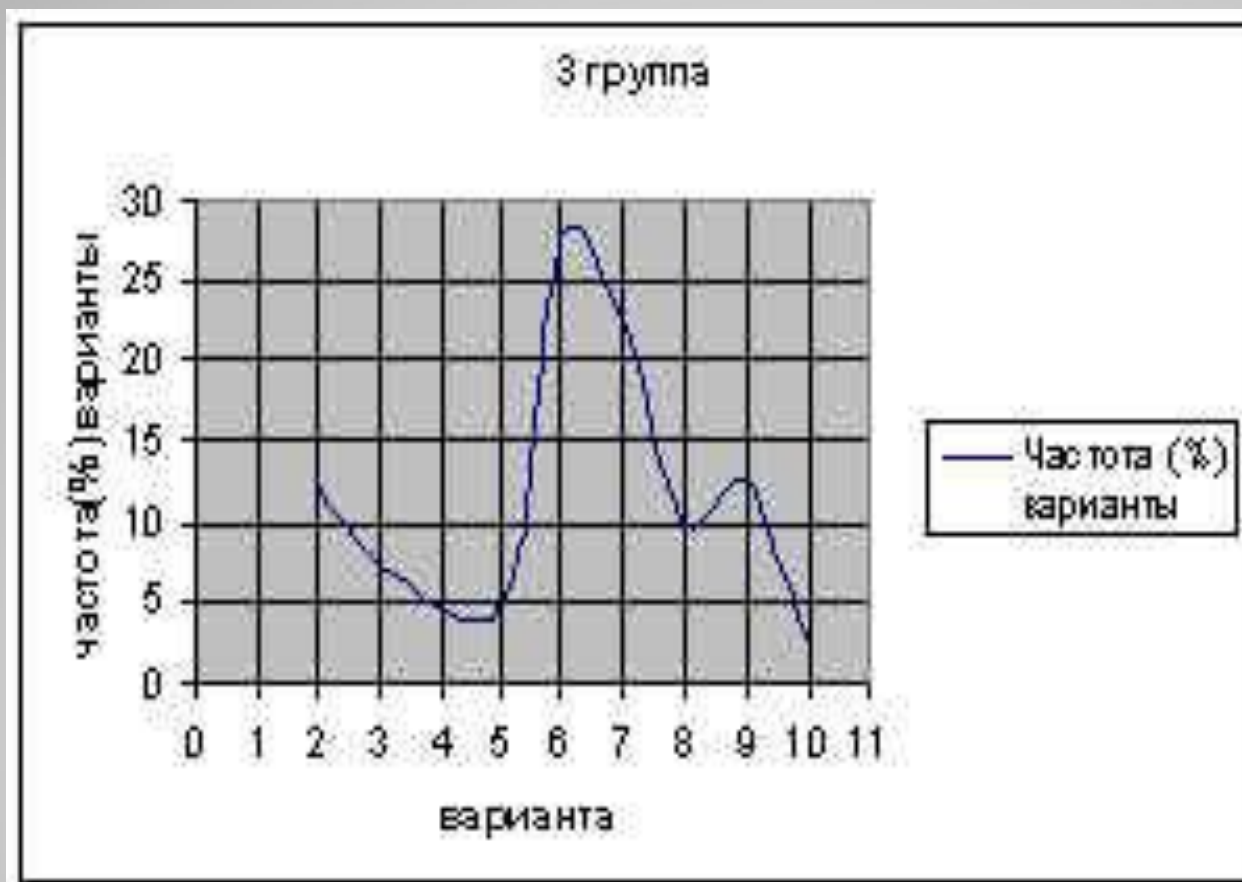
Зависимость кратности варианты от варианты.



Зависимость частоты варианты от варианты.



Зависимость частоты (%) варианты от варианты.



Спасибо за внимание.