

Схема научно-исследовательской работы регионоведа, проводящего эксперименты

Исследователь (регионовед)

Предмет исследования (поведенческие или экономические свойства, процессы, Функции и т.д.)

Испытуемый (группа испытуемых)

Эксперимент (измерение)

Результаты эксперимента (числовые коды)

Статистическая обработка данных эксперимента

Результат статистической обработки (числовые коды)

Выводы (печатный текст, диплом, статьи и т.д.)

Получатель научной информации (заказчик, руководитель проекта и т.д.)

Исследование обычно начинается с некоторого *предположения*, требующего проверки с привлечением фактов. Это предположение — **гипотеза** — формулируется в отношении *связи явлений или свойств в некоторой совокупности объектов*.

Генеральная совокупность — это все множество объектов, в отношении которого формулируется исследовательская гипотеза.

генеральная совокупность — это хотя и не бесконечное по численности, но, как правило, недоступное для сплошного исследования множество потенциальных испытуемых.

Выборочная совокупность (Выборка) — это ограниченная по численности группа объектов (испытуемых, респондентов), специальным образом отбираемая из генеральной совокупности для изучения её свойств.

...изучение на выборке свойств генеральной совокупности называется *выборочным исследованием*.

Практически все исследования являются *выборочными*, а их выводы распространяются на *генеральные совокупности*.

1. сформулирована гипотеза

2. определены соответствующие генеральные совокупности

3. организация выборки.

Свойства выборки

Выборка должна *обосновывать генерализацию выводов*

выборочного исследования — обобщение, распространение их на генеральную совокупность.

Основные критерии обоснованности выводов исследования —

это

1. репрезентативность выборки и

2 статистическая достоверность (эмпирических) результатов.

1. Репрезентативность выборки — или её

представительность — это способность выборки

представлять изучаемые *явления* достаточно *полно*, с

точки зрения *их изменчивости* в генеральной

совокупности.

Свойства репрезентативности

- репрезентативность всегда **ограничена** в той **мере**, в какой ограничена выборка.
- репрезентативность выборки является основным критерием при определении **границ генерализации выводов исследования**.

приёмы, позволяющие получить достаточную для

исследователя репрезентативность выборки.

А.Первый (основной) приём — это простой случайный (рандомизированный) отбор.

Он предполагает обеспечение таких условий, чтобы каждый член генеральной совокупности имел равные с другими шансы попасть в выборку.

Случайный отбор обеспечивает возможность попадания в выборку самых разных представителей генеральной совокупности. При этом принимаются специальные меры, исключающие появление какой-либо закономерности отбора.

В Второй прием — стратифицированный случайный отбор,
(отбор по свойствам генеральной совокупности)

1. предварительное определение *тех качеств*, которые могут **влиять на изменчивость изучаемого свойства** (это может быть пол, уровень дохода или образования и т. д.).
2. определяется **процентное соотношение численности** различающихся по этим качествам групп (страт) в генеральной совокупности и обеспечивается идентичное процентное соотношение соответствующих групп в выборке.
3. Далее в каждую подгруппу выборки **испытуемые подбираются по принципу простого случайного отбора.**

2 Статистическая достоверность, или (*статистическая значимость*) результатов исследования определяется при помощи методов статистического вывода, которые предъявляют *определённые требования к численности*, или объёму выборки.

наиболее *общие рекомендации*:

- Наибольший объем выборки **необходим при разработке диагностической методики** – от 200 до 1000–2500 человек.
- Если **необходимо сравнивать 2 выборки**, их общая численность должна быть не менее 50 человек; численность сравниваемых выборок должна быть приблизительно одинаковой.
- Если **изучается взаимосвязь между какими-либо свойствами**, то объем выборки должен быть не меньше 30–35 человек.
- **Чем больше изменчивость изучаемого свойства, тем больше должен быть объем выборки.** Поэтому изменчивость можно уменьшить, увеличивая однородность выборки, например, по полу, возрасту и т. д. Но при этом, уменьшаются возможности генерализации выводов.

Зависимые и независимые выборки.

Обычна ситуация исследования, когда интересующее исследователя свойство изучается на двух или более выборках с целью их дальнейшего сравнения. Эти выборки могут находиться в различных соотношениях — в зависимости от процедуры их организации.

Независимые выборки характеризуются тем, что вероятность отбора любого испытуемого одной выборки не зависит от отбора любого из испытуемых другой выборки.

Зависимые выборки характеризуются тем, что каждому испытуемому одной выборки поставлен в соответствие по определённому критерию испытуемый из другой выборки.

зависимые выборки предполагают попарный подбор испытуемых в сравниваемые выборки, а независимые выборки — независимый

выделим **две парадигмы** исследования.

R-методология предполагает изучение изменчивости некоторого свойства (поведенческого) под влиянием некоторого воздействия, фактора либо другого свойства. *Выборкой является множество испытуемых.*

Q-методология, предполагает исследование изменчивости субъекта (единичного) под влиянием различных стимулов (условий, ситуаций и т. д.). Ей соответствует ситуация, когда *выборкой является множество стимулов.*

следует различать

- **объекты исследования** (в это чаще всего люди, испытуемые)
- **их свойства** (то, что интересует исследователя, составляет *предмет изучения*)
- **признаки**, отражающие в *числовой шкале* выраженность свойств.

Процесс присвоения количественных (числовых) значений, имеющейся у исследователя информации, называется **кодированием**.

кодирование это такая операция, с помощью которой экспериментальным данным придаётся форма числового сообщения (кода).

Измерение в терминах производимых исследователем операций – это *приписывание объекту числа по определённому правилу.*

Это правило устанавливает соответствие между **измеряемым свойством объекта** и **результатом измерения – признаком.**

ИЗМЕРИТЕЛЬНЫЕ ШКАЛЫ

- ◎ **НОМИНАТИВНАЯ**, номинальная или шкала наименований
- ◎ **ПОРЯДКОВАЯ**, ординарная или ранговая шкала,
- ◎ **ИНТЕРВАЛЬНАЯ** или шкала равных интервалов,
- ◎ **ШКАЛА РАВНЫХ ОТНОШЕНИЙ**, или шкала отношений

НОМИНАТИВНАЯ ШКАЛА (ШКАЛА НАИМЕНОВАНИЙ)

процедура измерения сводится к классификации свойств, группировке объектов, к объединению их в классы, группы при условии, что объекты, принадлежащие к одному классу, идентичны (или аналогичны) друг другу в отношении какого-либо признака или свойства, тогда как объекты, различающиеся по этому признаку, попадают в разные классы.

Номинативная шкала (не метрическая), или шкала наименований (номинальное измерение).

В основе лежит процедура, обычно не ассоциируемая с измерением.

Пользуясь определённым правилом, *объекты группируются по различным классам* так, чтобы внутри класса они были идентичны по измеряемому свойству. Каждому классу даётся наименование и обозначение, обычно числовое.

Затем каждому объекту присваивается соответствующее обозначение.

Ранговая, или **порядковая** шкала (**не метрическая**) (как результат ранжирования).

измерение в этой шкале предполагает *приписывание объектам чисел в зависимости от степени выраженности измеряемого свойства.*

суть : *при сравнении испытуемых друг с другом мы можем сказать, больше или меньше выражено свойство, но не можем сказать, насколько больше или насколько меньше оно выражено.*

При измерении в ранговой шкале, таким образом, из всех свойств чисел учитывается то, что **они разные**, и то, что **одно число больше, чем другое.**

Интервальная шкала (метрическая).

измерение, при котором числа *отражают насколько больше или меньше выражено свойство.*

Равным разностям между числами в этой шкале соответствуют равные разности в уровне выраженности измеренного свойства, в этой шкале предполагает возможность применения *единицы измерения (метрики).*

Объекту присваивается число единиц измерения, пропорциональное выраженности измеряемого свойства.

Важная особенность интервальной шкалы —

произвольность выбора нулевой точки: ноль вовсе не соответствует полному отсутствию измеряемого свойства.

Произвольность выбора нулевой точки отсчёта обозначает, что измерение в этой шкале не соответствует абсолютному количеству измеряемого свойства, мы можем судить,

насколько больше или насколько меньше выражено свойство при сравнении объектов, но не можем судить о

том, во сколько раз больше или меньше выражено

шкала отношений (**метрическая**, абсолютная шкала).
Измерение в этой шкале отличается от интервального только тем, что в ней *устанавливается нулевая точка*, соответствующая полному отсутствию выраженности измеряемого свойства.

шкалы полезно характеризовать по признаку их **дифференцирующей способности** (*мощности*).

шкалы по мере возрастания мощности располагаются следующим образом:

1. номинативная,
2. ранговая,
3. интервальная,
4. абсолютная.

не метрические шкалы менее мощные — они отражают меньше информации о различии объектов (испытуемых) по измеренному свойству,

метрические шкалы более мощные, они лучше дифференцируют испытуемых.

... если у исследователя есть возможность выбора, следует применить более мощную шкалу.

- Измерения, осуществляемые с помощью двух первых шкал, считаются *качественными*, а осуществляемые с помощью двух последних шкал — *количественными*.

ТАБЛИЦЫ И ГРАФИКИ ТАБЛИЦА ИСХОДНЫХ ДАННЫХ

Обычно в ходе исследования интересующий исследователя признак измеряется не у одного-двух, а у **множества объектов** (испытуемых).

Кроме того, каждый объект характеризуется не одним, а целым **рядом признаков**, измеренных в разных шкалах. Одни признаки представлены в номинативной шкале и указывают на принадлежность испытуемых к той или иной группе (пол, профессия, контрольная или экспериментальная группа и т. д.). Другие признаки могут быть представлены в порядковой или метрической шкале. Поэтому результаты измерения для дальнейшего анализа чаще всего представляют в виде **таблицы исходных данных**.

Каждая строка такой таблицы обычно соответствует одному объекту, а каждый столбец – одному измеренному признаку. Таким образом, **исходной формой представления данных** является **таблица типа «объект – признак»**.

- Процесс присвоения количественных (числовых) значений, имеющейся у исследователя информации, называется кодированием Иными словами — *кодирование* это такая операция, с помощью которой экспериментальным данным придаётся форма числового сообщения (кода).

ТАБЛИЦЫ И ГРАФИКИ РАСПРЕДЕЛЕНИЯ ЧАСТОТ

Как правило, *анализ данных* начинается с изучения того, как часто встречаются те или иные значения интересующего исследователя признака (*переменной*) в имеющемся множестве наблюдений.

Для этого строятся *таблицы и графики распределения частот*. Нередко они являются основой для получения ценных содержательных выводов исследования.

Если признак принимает всего лишь несколько возможных значений (до 10-15), то таблица распределения частот показывает *частоту встречаемости каждого значения признака*.

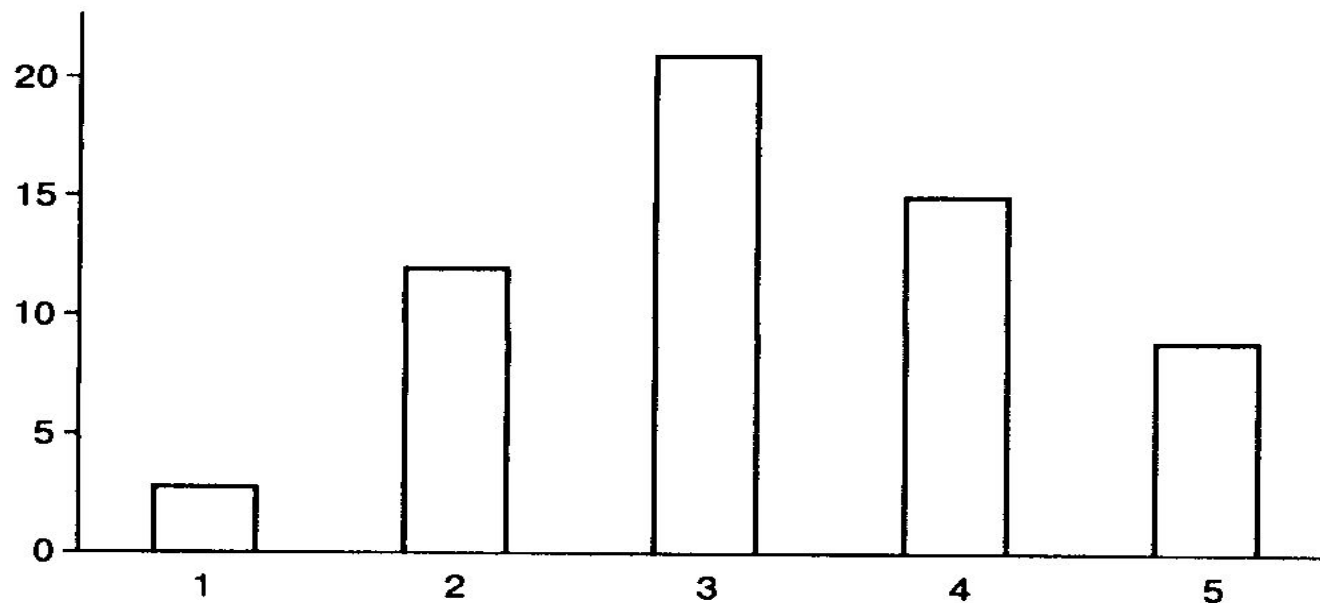
Если указывается, сколько раз встречается *каждое значение признака*, то это — *таблица абсолютных частот распределения*, если указывается доля наблюдений, приходящихся на то или иное значение признака, то говорят *об относительных частотах распределения*.

Ещё одной разновидностью таблиц распределения являются **таблицы распределения накопленных частот**.

Они показывают, как накапливаются частоты по мере возрастания значений признака. Напротив каждого значения (интервала) указывается сумма частот встречаемости всех тех наблюдений, величина признака у которых не превышает данного значения (меньше верхней границы данного интервала).

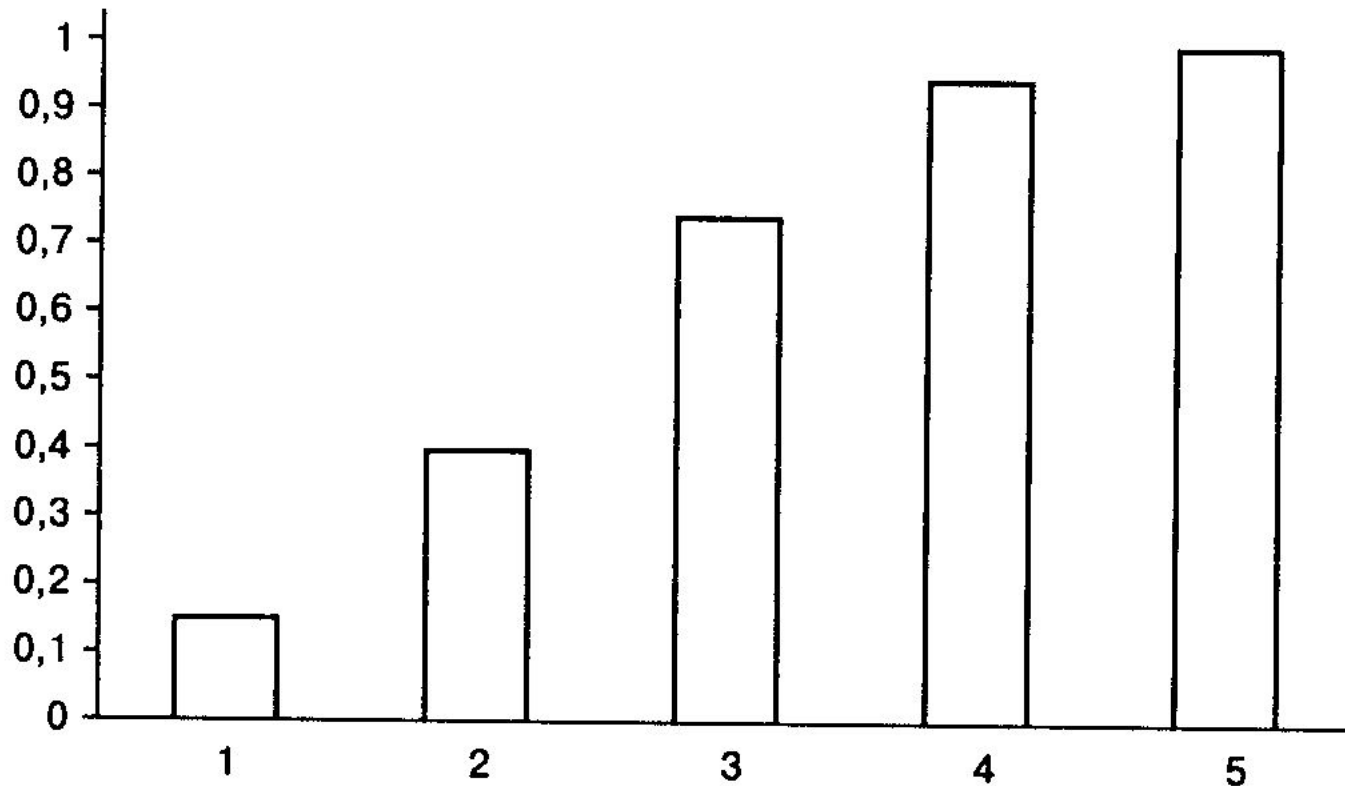
Для более наглядного представления строится **график распределения частот** или **график накопленных частот** — **гистограмма** или **сглаженная кривая распределения**.

Гистограмма распределения частот — это столбиковая диаграмма, каждый столбец которой опирается на *конкретное значение признака или разрядный интервал* (для сгруппированных частот). Высота столбика пропорциональна частоте встречаемости соответствующего значения.



Гистограмма распределения частот самооценки

Гистограмма накопленных частот отличается от гистограммы распределения тем, что *высота каждого столбика пропорциональна частоте, накопленной к данному значению (интервалу)*.

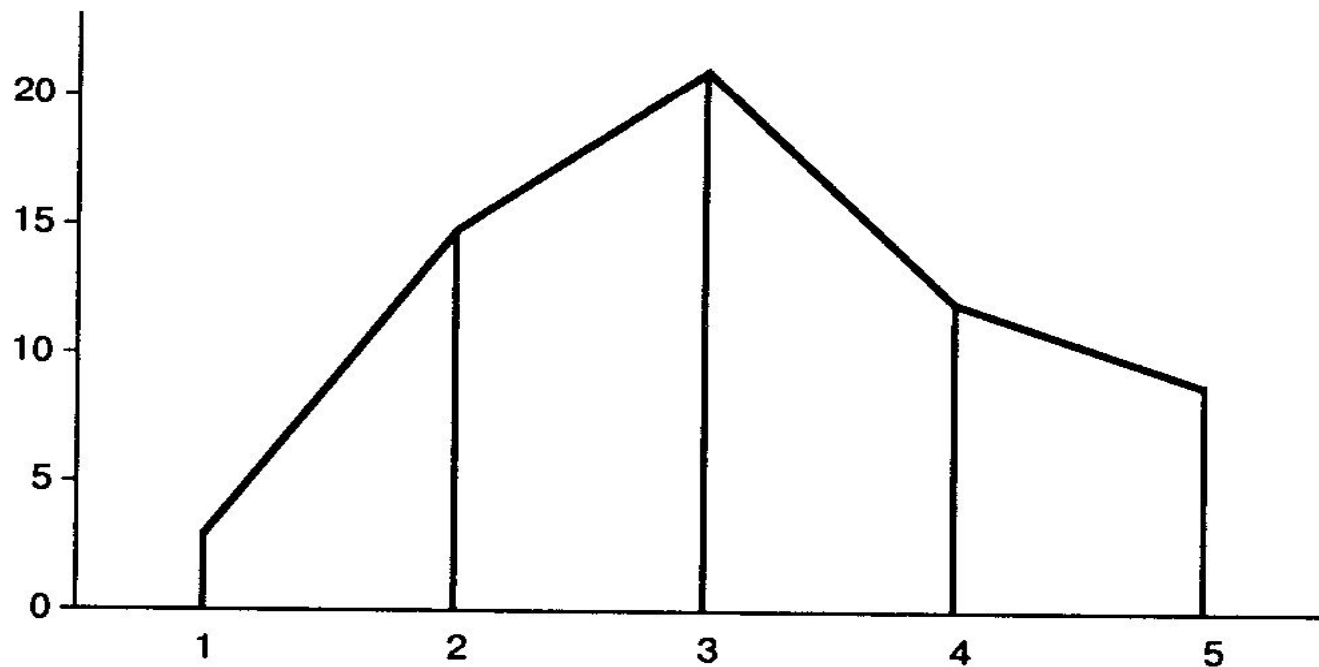


Гистограмма накопленных относительных частот самооценки

Построение **полигона распределения частот** напоминает построение гистограммы.

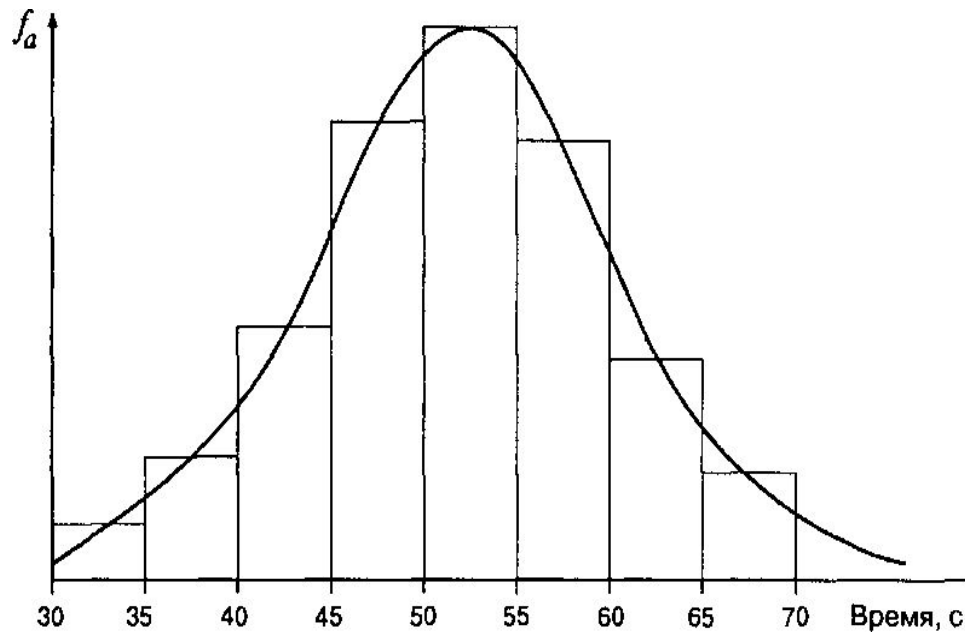
В гистограмме вершина каждого столбца, соответствующая частоте встречаемости данного значения (интервала) признака, – *отрезок прямой*.

А для полигона отмечается *точка*, соответствующая середине этого отрезка. Далее все точки соединяются *лс*



Полигон распределения частот самооценки

Вместо гистограммы или полигона часто изображают **сглаженную кривую распределения частот**



сглаженная кривая распределения частот.

Таблицы и графики распределения частот дают важную *предварительную информацию о форме распределения признака*:

о том, какие значения встречаются реже, а какие чаще, насколько выражена изменчивость признака.

выделяют следующие типичные формы распределения.

Равномерное распределение — когда все значения встречаются одинаково (или почти одинаково) часто.

Симметричное распределение — когда одинаково часто встречаются крайние значения.

Нормальное распределение — симметричное распределение, у которого крайние значения встречаются редко и частота постепенно повышается от крайних к срединным значениям признака.

Асимметричные распределения — *левосторонние* (с преобладанием частот малых значений), *правосторонние* (с преобладанием частот больших значений).

Таблицы и графики распределения признака позволяют делать некоторые содержательные выводы при сравнении групп испытуемых между собой. Сравнивая распределения, мы можем не только судить о том, какие значения встречаются чаще в той или иной группе, но и **сравнивать группы по степени выраженности индивидуальных различий – изменчивости по данному признаку.**

Таблицы и графики накопленных частот позволяют получить дополнительную информацию о том, **сколько испытуемых (или какая их доля) имеют выраженность признака не выше определённого значения.**

ТАБЛИЦЫ СОПРЯЖЕННОСТИ НОМИНАТИВНЫХ ПРИЗНАКОВ

Таблицы сопряжённости, или **кросстабуляции** — это *таблицы совместного распределения частот двух и более номинативных признаков, измеренных на одной группе объектов.*

Эти таблицы позволяют *сопоставить два или более распределения.* Столбцы такой таблицы соответствуют категориям (градациям) одного номинативного признака, а строки — категориям (градациям) другого номинативного признака.

ПЕРВИЧНЫЕ ОПИСАТЕЛЬНЫЕ СТАТИСТИКИ

К первичным **описательным статистикам** (*Descriptive Statistics*) обычно *относят числовые характеристики распределения* измеренного на выборке признака.

Каждая характеристика отражает в *одном числовом значении* свойство распределения множества результатов измерения (с точки зрения их расположения на числовой оси либо с точки зрения их изменчивости)

Основное назначение первичных описательных статистик — замена множества значений признака, измеренного на выборке, *одним числом*

Компактное описание группы при помощи первичных статистик позволяет *интерпретировать результаты измерений путём сравнения первичных статистик* разных групп.

МЕРЫ ЦЕНТРАЛЬНОЙ ТЕНДЕНЦИИ

Мера центральной тенденции (*Central Tendency*) — это число, характеризующее выборку **по уровню выраженности измеренного признака**.

Существуют **три способа** определения «центральной тенденции», каждому из которых соответствует своя мера: **мода, медиана и выборочное среднее**.

Мода (Mode) – это значение из множества измерений, которое *встречается наиболее часто*.

Моде, или *модальному интервалу признака*, соответствует наибольший подъем (вершина) графика распределения частот. Если график распределения частот имеет одну вершину, то такое распределение называется *унимодальным*.

Когда два соседних значения встречаются одинаково часто и чаще, чем любое другое значение, мода есть среднее этих двух значений.

Распределение может не иметь моды. Когда *все значения* встречаются *одинаково часто*, принято считать, что такое *распределение не имеет моды*.

мода – это **значение признака**, а не его частота.

Медиана (*Median*) – это значение признака, которое делит упорядоченное (ранжированное) множество данных пополам так, что одна половина всех значений оказывается меньше медианы, а другая – больше.

Таким образом, первым шагом при определении медианы является упорядочивание (ранжирование) всех значений по возрастанию или убыванию.

Далее медиана определяется следующим образом:

- если данные содержат **нечётное** число значений (8, 9, **10**, 13, 15), то медиана есть центральное значение, т. е. **$Md = 10$** ;
- если данные содержат **чётное** число значений (5, **8**, **9**, 11), то медиана есть точка, лежащая посередине между двумя центральными значениями, т. е. **$Md = (8+9)/2 = 8,5$** .

Среднее {Mean) (M_x — выборочное среднее, среднее арифметическое) — сумма всех значений измеренного признака, делённая на количество суммированных значений.

Если некоторый признак X измерен в группе испытуемых численностью N , мы получим значения: $x_1, x_2, \dots, x_i, \dots, x_N$ (где i — текущий номер испытуемого, от 1 до N). Тогда среднее значение M_x определяется по формуле:

$$M_x = \frac{1}{N} \sum_{i=1}^N x_i$$

Свойства среднего

Если к каждому значению переменной прибавить одно и то же число c , то среднее увеличится на это число (уменьшится на это число, если оно отрицательное):

$$M_{(x_i+c)} = \frac{1}{N} \sum_{i=1}^N (x_i + c) = M_x + c$$

А если каждое значение переменной умножить на одно и то же число c , то среднее увеличится в c раз (уменьшится в c раз, если делить на c):

$$M_{(x_i \cdot c)} = \frac{1}{N} \sum_{i=1}^N (x_i \cdot c) = M_x \cdot c$$

отклонение от среднего: $(x_i - M_x)$.

Из первого, очевидного свойства среднего следует одно важное свойство: *сумма всех отклонений от среднего равна нулю.*

$$\sum_{i=1}^N (x_i - M_x) = 0$$

Соответственно, *среднее отклонение от среднего также равно нулю.*

ВЫБОР МЕРЫ ЦЕНТРАЛЬНОЙ ТЕНДЕНЦИИ

Каждая мера центральной тенденции обладает характеристиками, которые делают её ценной в определённых условиях.

Для **номинативных данных** единственной подходящей мерой центральной тенденции является **мода**, или модальная категория — **градация номинативной переменной**, которая встречается наиболее часто.

Наиболее очевидной и часто используемой мерой центральной тенденции является **среднее значение**.

Но его использование ограничивается тем, что оно на величину среднего влияет **каждое отдельное значение**.

Если какое-нибудь значение в группе увеличится на C , то среднее увеличится на C/N .

Таким образом, **среднее значение** весьма чувствительно к «выбросам» — экстремально малым или большим значениям переменной.

На величину **моды** и **медианы** величина каждого отдельного значения не влияет.

Например, если в группе из 20 измерений переменной наибольшее значение утроится по величине, то не изменится ни мода, ни медиана. Величина среднего при этом заметно изменится.

Иначе говоря, **мода и медиана не чувствительны к «выбросам»**.

Меры центральной тенденции чаще всего используются для сравнения групп по уровню выраженности признака.

какую меру использовать?

Выборочные средние можно сравнивать, если выполняются следующие условия:

- группы достаточно большие, чтобы судить о форме распределения;
- распределения симметричны;
- отсутствуют «выбросы».

Если хотя бы одно из перечисленных условий не выполняется, то следует ограничиться модой и медианой.

Альтернативой является «сквозное» ранжирование представителей сравниваемых групп и сравнение средних, вычисленных для рангов этих групп.

Помимо **мер центральной тенденции**
широко используются **меры положения**,
которые называются
квантилями распределения

Квантиль — это **точка** на числовой оси измеренного признака, которая делит всю совокупность упорядоченных измерений на **две группы с известным соотношением их численности**.

С одним из квантилей мы уже знакомы — это **медиана**. Это значение признака, которое делит всю совокупность измерений на две группы с равной численностью. Кроме медианы часто используются **процентили и квартили**.

Процентили (*Percentiles*) – это 99 точек – значений признака (P_1, \dots, P_{99}), которые делят упорядоченное (по возрастанию) множество наблюдений на 100 частей, равных по численности.

Определение конкретного значения процентиля аналогично определению медианы. Например, при определении 10-го процентиля, P_{10} , сначала все значения признака упорядочиваются по возрастанию. Затем отсчитывается 10% испытуемых, имеющих наименьшую выраженность признака. P_{10} будет соответствовать тому значению признака, который отделяет эти 10% испытуемых от остальных 90%.

Квартили (Quartiles) – это 3 точки – значения признака (P_{25} , P_{50} , P_{75}), которые делят упорядоченное (по возрастанию) множество наблюдений на 4 равные по численности части.

Первый квартиль соответствует 25-му процентилю, второй – 50-му процентилю или медиане, третий квартиль соответствует 75-му процентилю.

Процентили и квартили используются для определения частоты встречаемости тех или иных значений (или интервалов) измеренного признака или для выделения подгрупп и отдельных испытуемых, наиболее типичных или нетипичных для данного множества наблюдений.

МЕРЫ ИЗМЕНЧИВОСТИ

Меры центральной тенденции отражают *уровень выраженности измеренного признака*.

Однако не менее важной характеристикой является *выраженность индивидуальных различий* испытуемых по измеренному признаку.

Меры изменчивости (Dispersion) применяются для численного *выражения величины межиндивидуальной вариации признака*

Наиболее простой и очевидной мерой изменчивости является *размах*, указывающий на диапазон изменчивости значений.

Размах (*Range*) – это просто разность максимального и минимального значений:

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

это очень неустойчивая мера изменчивости, на которую влияют любые возможные «выбросы».

Более устойчивыми являются разновидности размаха: *размах от 10 до 90-го перцентиля*

$(P_{90} - P_{10})$ или

междуквартильный размах $(P_{75} - P_{25})$,

применяются для описания вариации в порядковых данных

Чем больше изменчивость в данных, тем больше отклонения значений от среднего, тем больше величина дисперсии. Величина дисперсии получается при усреднении всех квадратов отклонений:

$$\sum_{i=1}^N (x_i - M_x)^2$$

Следует отличать *теоретическую (генеральную) дисперсию* – меру изменчивости бесконечного числа измерений (в генеральной совокупности, популяции в целом) и *эмпирическую, или выборочную, дисперсию* – для реально измеренного множества значений признака

$$\bar{D}_x = \frac{\sum_{i=1}^N (x_i - M_x)^2}{N}$$

для метрических данных используется **дисперсия** — величина, название которой в науке является синонимом изменчивости.

Дисперсия (Variance) — мера изменчивости для метрических данных, пропорциональная сумме квадратов отклонений измеренных значений от их арифметического среднего:

$$D_x = \frac{\sum_{i=1}^N (x_i - M_x)^2}{N - 1}$$

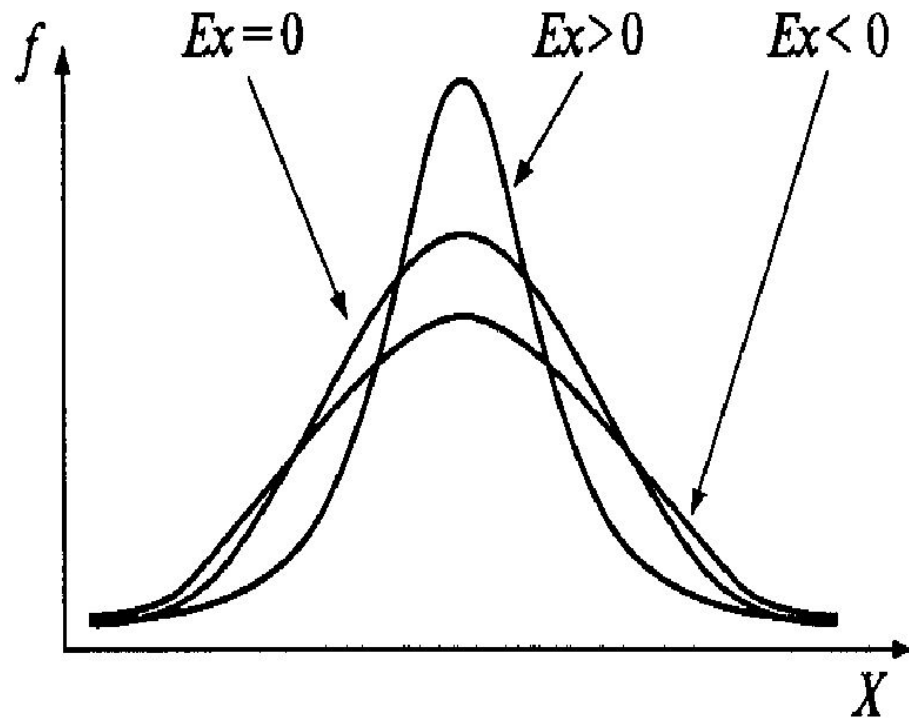
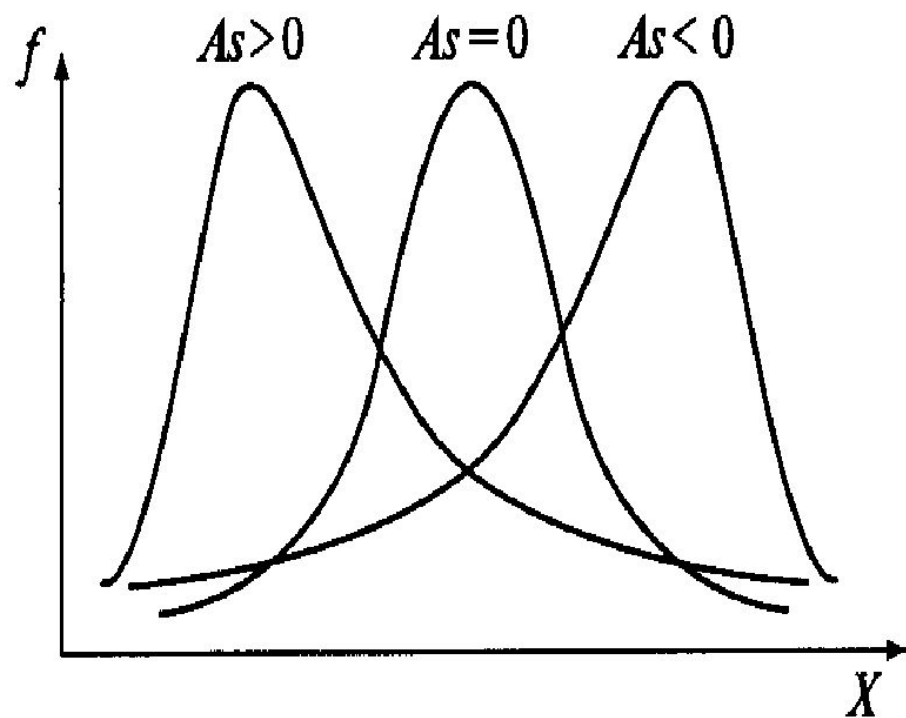
Стандартное отклонение (*Std. deviation*)
(сигма, среднеква-дратическое отклонение) —
положительное значение квадратного корня
из дисперсии:

$$\sigma_x = \sqrt{D_x} = \sqrt{\frac{\sum_i (x_i - M_x)^2}{N - 1}}$$

(На практике чаще используется именно стандартное отклонение, а не дисперсия. Это связано с тем, что сигма выражает изменчивость в исходных единицах измерения признака, а дисперсия — в квадратах исходных единиц).

Асимметрия (*Skewness*) – степень отклонения графика распределения частот от симметричного вида относительно среднего значения. Если исходные данные переведены в z-значения, показатель асимметрии вычисляется по формуле

$$As = \frac{\sum z_i^3}{N}$$



Распределения частот с разными значениями асимметрии и эксцесса

Эксцесс (*Kurtosis*) — мера плосковершинности или остроконечности графика распределения измеренного признака.

Если исходные данные переведены в **z**-значения, показатель эксцесса определяется формулой:

$$E_x = \frac{\sum_i z_i^4}{N} - 3$$

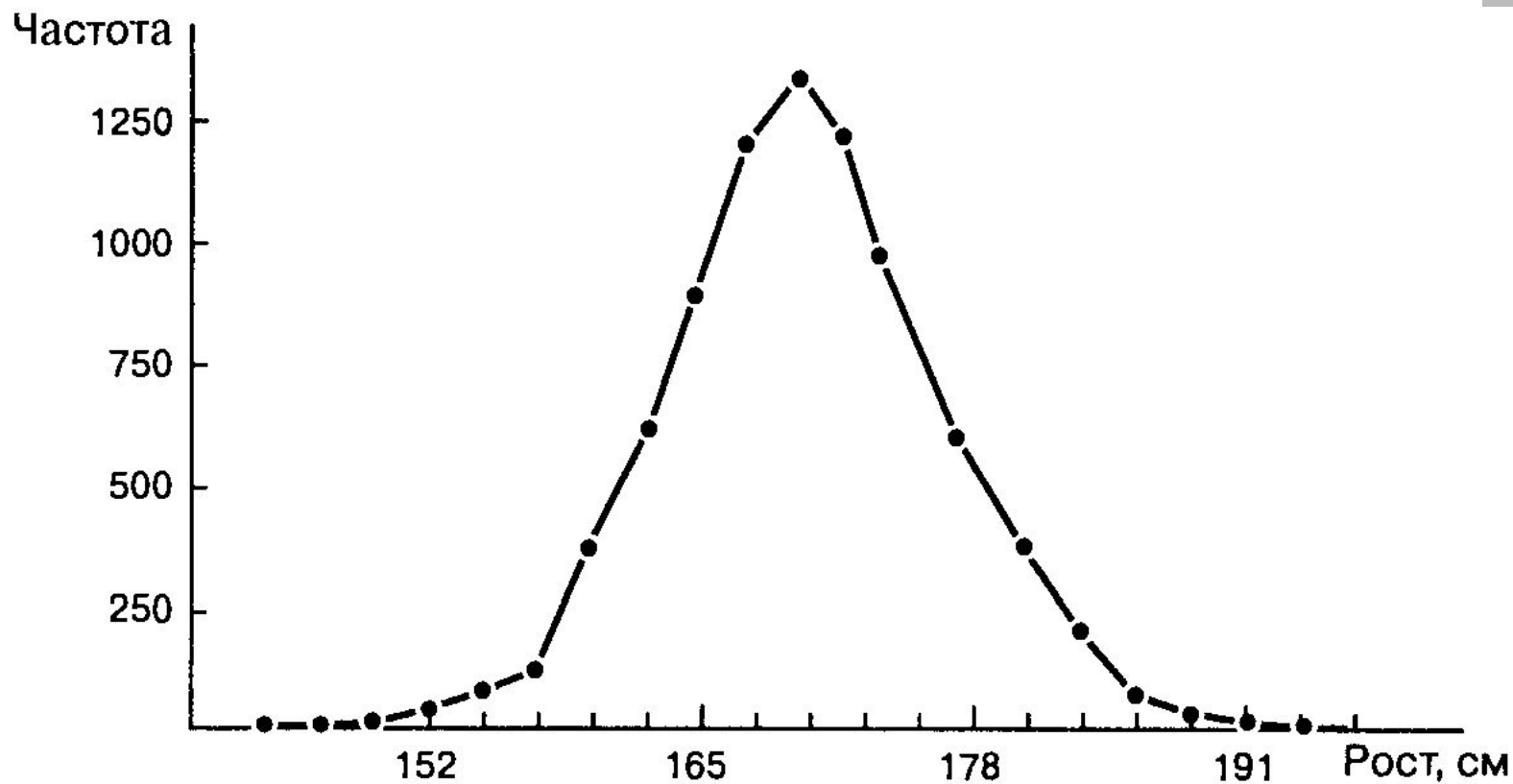
«**Острове́ршинное**» распределение характеризуется положительным эксцессом (**$E_x > 0$**).

«**Плосковершинное**» — отрицательным (**$-3 < E_x < 0$**).

«**Средневершинное**» (нормальное) распределение имеет нулевой эксцесс (**$E_x = 0$**).

НОРМАЛЬНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ И ЕГО ПРИМЕНЕНИЕ

- **Нормальный закон** распределения играет важнейшую роль в применении численных методов в психологии. Он лежит в основе измерений, разработки тестовых шкал, методов проверки гипотез.
- Если индивидуальная изменчивость некоторого свойства есть следствие действия множества причин, то распределение частот для всего многообразия проявлений этого свойства в генеральной совокупности соответствует **кривой нормального распределения**. Это и есть закон нормального распределения.
- Если при изучении некоторого свойства мы произвели его измерение на выборке испытуемых и получили отличающееся от нормального распределение, то это значит, что либо выборка нерепрезентативна генеральной совокупности, либо измерения произведены не в шкале равных интервалов.



Полигон частот для роста 8585 взрослых людей, родившихся в Англии в XIX в.

$$f(x_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_i - M)^2 / 2\sigma^2} \quad (*)$$

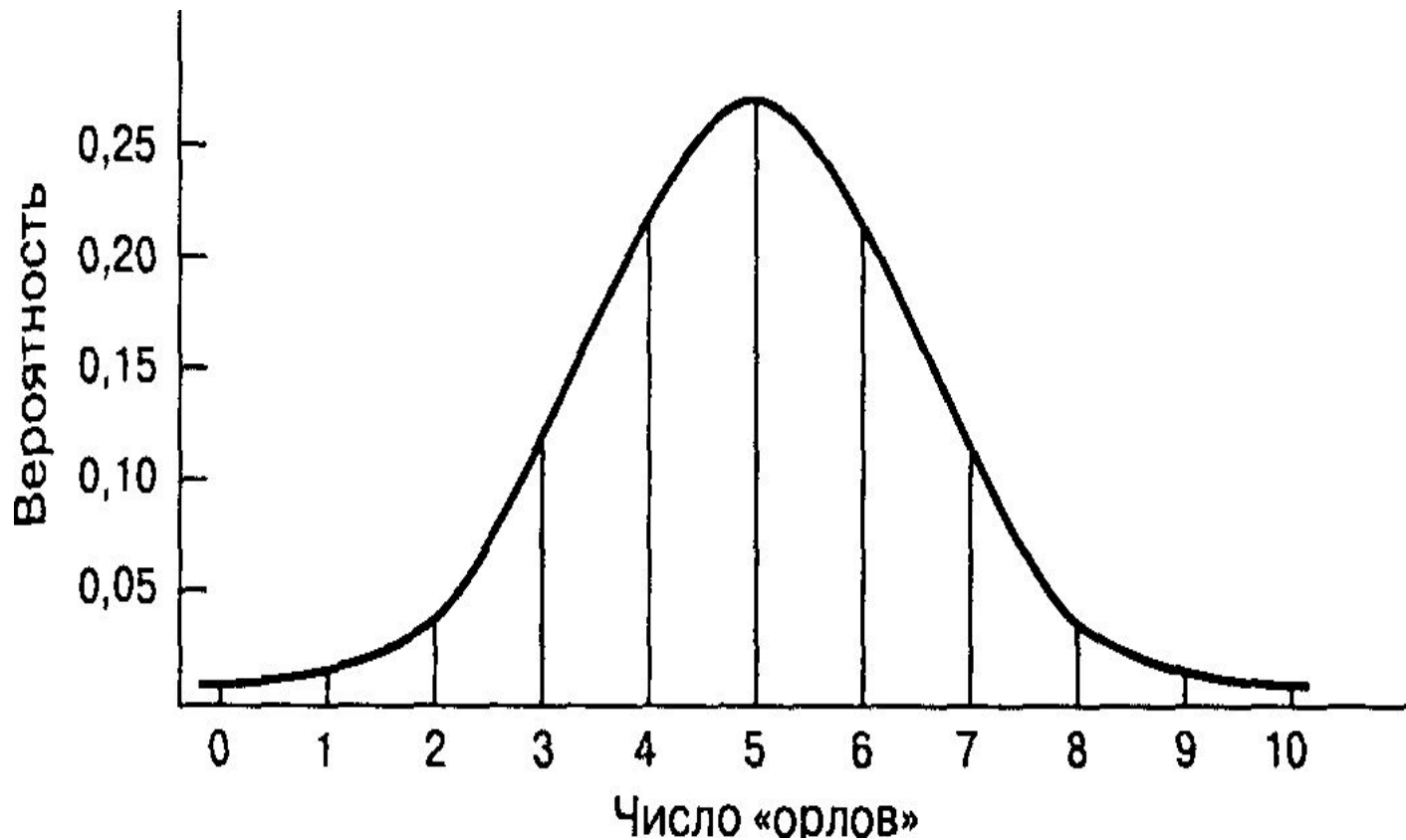
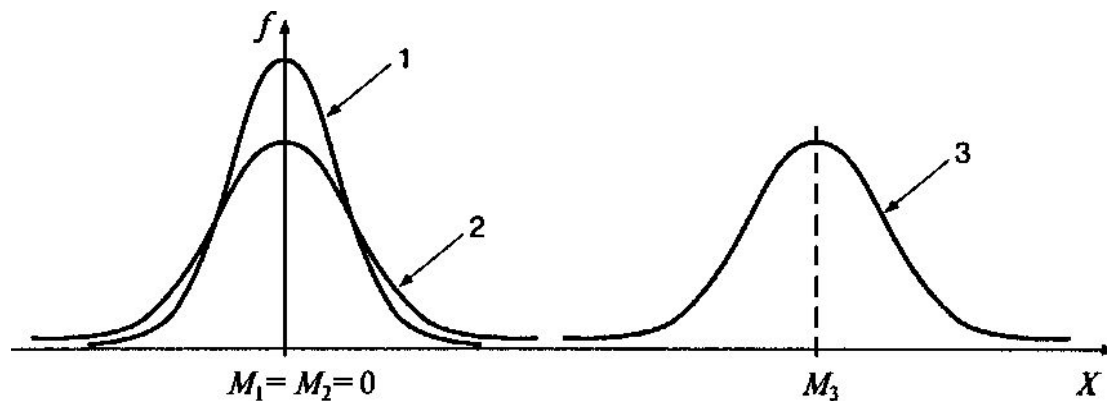


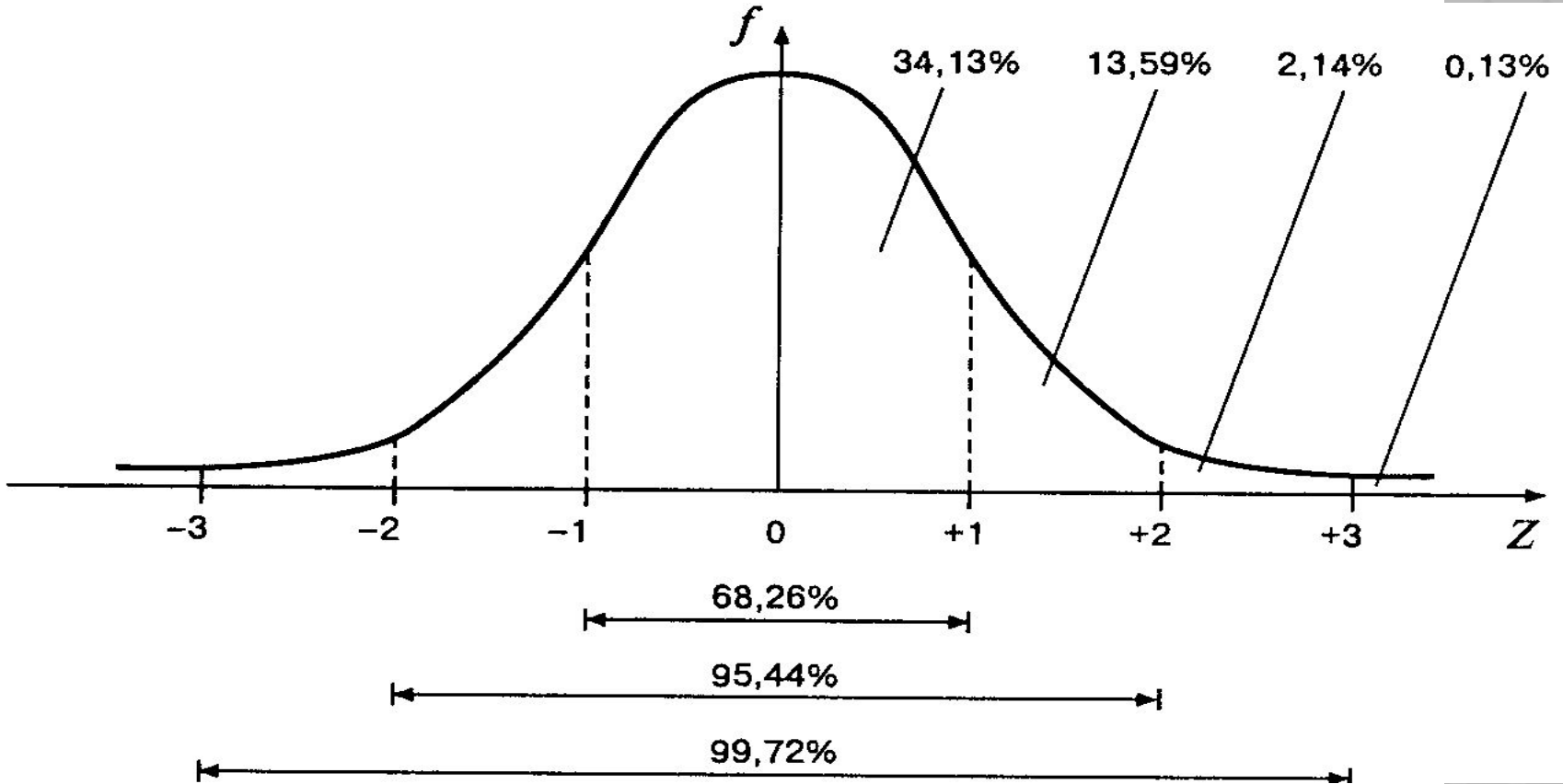
График распределения вероятностей выпадения «орлов» в игре с 10 подбрасываниями монеты и кривая нормального распределения



Семейство нормальных кривых, 1-е распределение отличается от 2-го стандартным отклонением ($\sigma_1 < \sigma_2$), 2-е от 3-го средним арифметическим ($M_2 < M_3$)

Каждому психологическому (биологическому) свойству соответствует своё распределение в генеральной совокупности. Чаще всего оно является **нормальным** и характеризуется своими параметрами: средним (M) и стандартным отклонением (σ). Только эти два значения отличают друг от друга бесконечное множество нормальных кривых, (*)
одинаковой формы, заданной уравнением Среднее M задаёт положение кривой на числовой оси и выступает как некоторая исходная, нормативная величина измерения.
Стандартное отклонение σ задаёт ширину этой кривой, зависит от единицы измерения и выступает как масштаб

Все многообразие нормальных распределений может быть сведено к одной кривой, если применить z-преоб $z_i = \frac{x_i - M_x}{\sigma_x}$ ко всем возможным измерениям свойств. Тогда каждое свойство будет иметь (среднее) $M=0$, стандартное отклонение $\sigma=1$. На рис. построен график нормального распределения для $M=0$ и $\sigma=1$. Это *единичное нормальное распределение*, КОТ



Стандартное нормальное распределение

Рассмотрим важные свойства стандартного нормального распределения

1. *Единицей измерения* единичного нормального распределения является *стандартное отклонение*.
2. Кривая приближается к оси Z *асимптотически* (никогда не коснётся её).
3. Кривая симметрична относительно $M=0$. Её *асимметрия* и *эксцесс* равны нулю.
4. Кривая имеет *характерный изгиб*: точка перегиба лежит точно на расстоянии в одну σ от M .
5. Площадь между кривой и осью Z равна 1.

Свойство (5) объясняет название единичное нормальное распределение и имеет исключительно важное значение. Благодаря этому свойству *площадь под кривой интерпретируется как вероятность, или относительная частота*.

Вся площадь под кривой соответствует вероятности того, что признак примет любое значение из всего диапазона его изменчивости (от $-\infty$ до $+\infty$). Площадь под единичной нормальной кривой слева или справа от нулевой точки равна 0,5. Это соответствует тому, что половина генеральной совокупности имеет значение признака больше 0, а половина — меньше 0. Относительная частота встречаемости в генеральной совокупности значений признака в диапазоне от z_1 до z_2 равна площади под кривой, лежащей между соответствующими

любое нормальное распределение может быть сведено к единичному нормальному распределению путем z -преобразования. Таким образом:

- если x_i - имеет нормальное распределение со средним M и стандартным отклонением σ , то $z = (x - M_x) / \sigma$ характеризуется единичным нормальным распределением со средним 0 и стандартным отклонением 1 ;
- площадь между x_1 и x_2 в нормальном распределении со средним M_x и стандартным отклонением σ равна площади между $z_1 = (x_1 - M_x) / \sigma$ и $z_2 = (x_2 - M_x) / \sigma$ в единичном нормальном распределении.
- наиболее важным общим свойством разных кривых нормального распределения является одинаковая доля площади под кривой между одними и теми же двумя значениями признака, выраженными в единицах стандартного отклонения.

Полезно помнить, что для любого нормального распределения существуют следующие соответствия между диапазонами значений и площадью под кривой:

$M \pm \sigma$ соответствует = 68% (точно – 68,26%) площади;

$M \pm 2\sigma$ соответствует = 95% (точно – 95,44%) площади;

$M \pm 3\sigma$ соответствует = 100% (точно – 99,72%) площади.

Единичное нормальное распределение устанавливает чёткую взаимосвязь стандартного отклонения и относительного количества случаев в генеральной совокупности для любого нормального распределения, зная свойства единичного нормального распределения, мы можем ответить на вопросы:

какая доля генеральной совокупности имеет выраженность свойства от -1σ до $+1\sigma$? (ответ: 68,26% всей генеральной совокупности, так как от -1 до $+1$ содержится 0,6826 площади единичного нормального распределения);

какова вероятность того, что случайно выбранный представитель генеральной совокупности будет иметь выраженность свойства, на 3σ превышающую среднее значение?(ответ:

РАЗРАБОТКА ТЕСТОВЫХ ШКАЛ

Тестовые шкалы разрабатываются для того, чтобы оценить индивидуальный результат тестирования путём сопоставления его с *тестовыми нормами*, полученными на выборке стандартизации. *Выборка стандартизации* специально формируется для разработки тестовой шкалы – она должна быть репрезентативна генеральной совокупности, для которой планируется применять данный тест, при тестировании предполагается, что и тестируемый, и выборка стандартизации принадлежат одной и той же генеральной совокупности.

Исходным принципом при разработке тестовой шкалы является предположение о том, что *измеряемое свойство распределено в генеральной совокупности в соответствии с нормальным законом*. Соответственно, измерение в тестовой шкале данного свойства на выборке стандартизации также должно обеспечивать нормальное распределение. Если это так, то тестовая шкала является *метрической – равных интервалов*.

Если это не так, то свойство удалось отразить в шкале *порядка*.

Естественно, что большинство стандартных тестовых шкал являются метрическими, что позволяет более детально интерпретировать результаты тестирования – с учётом свойств нормального распределения – и корректно применять любые методы статистического анализа.

Основная проблема стандартизации теста заключается в разработке такой шкалы, в которой распределение тестовых показателей на выборке стандартизации соответствовало бы нормальному распределению.

Исходные тестовые оценки (первичными, или «сырыми» оценками) – это ответы на вопросы теста, время или количество решённых задач и т. д.

Существует множество стандартных тестовых шкал, **основное назначение которых** – представление индивидуальных результатов тестирования в удобном для интерпретации виде. Некоторые из этих шкал представлены на рис. Общим для них является соответствие нормальному распределению, а различаются они только двумя показателями: средним значением и масштабом (стандартным отклонением $-\sigma$), определяющим широту шкалы.

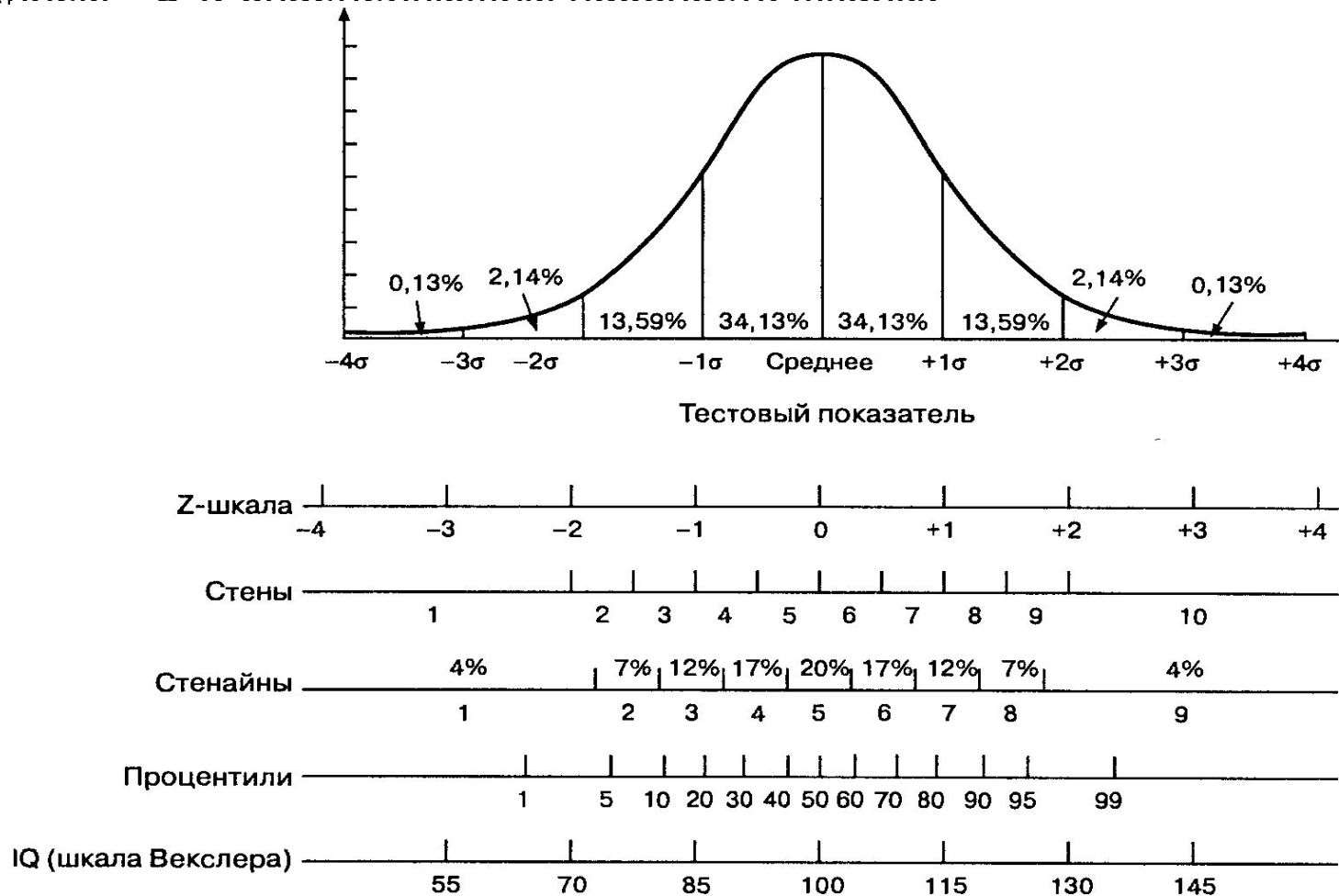


Рис. Нормальная кривая и тестовые шкалы

ПРОВЕРКА НОРМАЛЬНОСТИ РАСПРЕДЕЛЕНИЯ

Для проверки нормальности используются различные процедуры, позволяющие выяснить, отличается ли от нормального выборочное распределение измеренной переменной. Необходимость такого сопоставления возникает, когда мы сомневаемся в том, в какой шкале представлен признак – в *порядковой* или *метрической*.

Важность определения того, в какой шкале измерен признак, трудно переоценить, по крайней мере, по двум причинам. От этого зависит, во-первых, полнота учёта исходной эмпирической информации (в частности, об индивидуальных различиях), во-вторых, доступность многих методов анализа данных. Если исследователь принимает решение об измерении в порядковой шкале, то неизбежное последующее ранжирование ведёт к потере части исходной информации о различиях между испытуемыми, изучаемыми группами, о взаимосвязях между признаками и т. д. Кроме того, метрические данные позволяют использовать значительно более широкий набор методов анализа и, как следствие, сделать выводы исследования более глубокими и содержательными.

Наиболее весомым аргументом в пользу того, что признак измерен в метрической шкале, является соответствие выборочного распределения нормальному. Это является следствием *закона нормального распределения*. Если выборочное распределение не отличается от нормального, то это значит, что измеряемое свойство удалось отразить в метрической шкале (обычно – интервальной).

Графический способ (*Q-Q Plots*, *P-P Plots*).

Строят либо квантильные графики, либо графики накопленных частот. Квантильные графики (*Q-Q Plots*) строятся следующим образом. Сначала определяются эмпирические значения изучаемого признака, соответствующие 5, 10, ..., 95-перцентилю. Затем по таблице нормального распределения для каждого из этих перцентилей определяются z -значения (теоретические). Два полученных ряда чисел задают координаты точек на графике: эмпирические значения признака откладываются на оси абсцисс, а соответствующие им теоретические значения — на оси ординат. Для нормального распределения все точки будут лежать на одной прямой или рядом с ней. Чем больше расстояние от точек до прямой линии, тем меньше распределение соответствует нормальному. Графики накопленных частот (*P-P Plots*) строятся подобным образом. На оси абсцисс через равные интервалы откладываются значения накопленных относительных частот, например 0,05; 0,1; ...; 0,95. Далее определяются эмпирические значения изучаемого признака, соответствующие каждому значению накопленной частоты, которые пересчитываются в z -значения. По таблице нормального распределения определяются теоретические накопленные частоты (площадь под кривой) для каждого из вычисленных z -значений, которые откладываются на оси ординат. Если распределение соответствует нормальному, полученные на графике точки лежат на одной

Критерии асимметрии и эксцесса.

Эти критерии определяют допустимую степень отклонения эмпирических значений асимметрии и эксцесса от нулевых значений, соответствующих нормальному распределению. Допустимая степень отклонения – та, которая позволяет считать, что эти статистики существенно не отличаются от нормальных параметров. Величина допустимых отклонений определяется так называемыми стандартными ошибками асимметрии и

эксцесса.

Для формулы асимметрии $As = \frac{\sum z_i^3}{N}$

стандартная ошибка определяется:

$$As_{sd} = 3 \sqrt{\frac{6(N-1)}{(N+1)(N+3)}}$$

Для формулы эксцесса: $Ex = \frac{\sum z_i^4}{N} - 3$

стандартная ошибка эксцесса:

$$Ex_{sd} = 5 \sqrt{\frac{24N(N-2)(N-3)}{(N+1)^2(N+3)(N+5)}}$$

где N – объем выборки.

Выборочные значения асимметрии и эксцесса значительно отличаются от нуля, если не превышают значения своих стандартных ошибок. Это можно считать признаком соответствия выборочного распределения нормальному закону.

Статистический критерий нормальности Колмогорова-Смирнова

считается наиболее состоятельным для *определения степени соответствия эмпирического распределения нормальному.*

Он позволяет оценить вероятность того, что *данная выборка принадлежит генеральной совокупности с нормальным распределением.*

Если эта вероятность $p \leq 0,05$, то данное эмпирическое распределение *существенно отличается от нормального, а*

если $p > 0,05$, то делают вывод о приблизительном соответствии данного эмпирического распределения нормальному.

Причины отклонения от нормальности.

Общей причиной отклонения формы выборочного распределения признака от нормального вида чаще всего *является особенность процедуры измерения:* используемая шкала может обладать неравномерной чувствительностью к измеряемому свойству в разных частях диапазона его изменчивости.

Таким образом, такие отклонения от нормального вида, как право- или левосторонняя асимметрия или слишком большой эксцесс (больше 0), связаны с относительно низкой чувствительностью измерительной процедуры в области моды (вершины графика распределения частот).

Последствия отклонения от нормальности.

Следует отметить, что задача получения эмпирического распределения, строго соответствующего нормальному закону, нечасто встречается в практике исследования. Обычно такие случаи ограничиваются разработкой новой измерительной процедуры или тестовой шкалы, когда применяется эмпирическая или нелинейная нормализация для «исправления» эмпирического распределения.

В большинстве случаев соответствие или несоответствие нормальности является тем свойством измеренного признака, который исследователь должен учитывать при выборе статистических процедур анализа данных.

В общем случае при значительном отклонении эмпирического распределения от нормального следует отказаться от предположения о том, что признак измерен в метрической шкале. Но остаётся открытым вопрос о том, какова мера существенности этого отклонения? Кроме того, разные методы анализа данных обладают различной чувствительностью к отклонениям от нормальности. Обычно при обосновании перспективности этой проблемы приводят принцип Р. Фишера: *«Отклонения от нормального вида, если только они не слишком заметны, можно обнаружить лишь для больших выборок; сами по себе они вносят малое отличие в статистические критерии и другие вопросы»*. К примеру, при малых, но обычных для психологических исследований выборках (до 50 человек) критерий Колмогорова-Смирнова недостаточно чувствителен при определении даже весьма заметных «на глаз» отклонений от нормальности. В то же время некоторые процедуры анализа метрических данных вполне допускают отклонения от нормального распределения (одни – в большей степени, другие – в меньшей).

В дальнейшем при изложении материала мы при необходимости будем

оговаривать меру жёсткости требования нормальности.

КОЭФФИЦИЕНТЫ КОРРЕЛЯЦИИ

основные одномерные описательные статистики – меры центральной тенденции и изменчивости, которые применяются для описания одной переменной.
рассмотрим основные коэффициенты корреляции.

Коэффициент корреляции – двумерная описательная статистика, количественная мера взаимосвязи (совместной изменчивости) двух переменных.

История разработки и применения коэффициентов корреляции для исследования взаимосвязей фактически началась одновременно с возникновением измерительного подхода к исследованию индивидуальных различий – в 1870–1880 гг. Пионером в измерении способностей человека, как и автором самого термина «коэффициент корреляции», был Френсис Гальтон, а самые популярные коэффициенты корреляции были разработаны его последователем Карлом Пирсоном. С тех пор *изучение взаимосвязей с использованием коэффициентов корреляции является одним из наиболее популярных в психологии занятием.* К настоящему времени разработано великое множество различных коэффициентов корреляции, проблеме измерения взаимосвязи с их помощью посвящены сотни книг. Поэтому, не претендуя на полноту изложения, мы рассмотрим лишь самые важные, действительно незаменимые в исследованиях меры связи r - Пирсона, r -Спирмена и t -Кендалла. Их общей особенностью является то, что они *отражают взаимосвязь двух признаков, измеренных в количественной шкале – ранговой или метрической.*

Вообще говоря, любое эмпирическое исследование сосредоточено на изучении взаимосвязей двух или более переменных.

ПРИМЕРЫ

Приведем два примера исследования влияния демонстрации сцен насилия по ТВ на агрессивность подростков.

1. Изучается взаимосвязь двух переменных, измеренных в количественной (ранговой или метрической) шкале:

1) «время просмотра телепередач с насилием»; 2) «агрессивность».

2. Изучается различие в агрессивности 2-х или более групп подростков, отличающихся длительностью просмотра телепередач с демонстрацией сцен насилия.

Во втором примере изучение различий может быть представлено как исследование взаимосвязи 2-х переменных, одна из которых – номинативная (длительность просмотра телепередач). И для этой ситуации также разработаны свои коэффициенты корреляции.

ПОНЯТИЕ КОРРЕЛЯЦИИ

Взаимосвязи на языке математики обычно описываются при помощи функций, которые графически изображаются в виде линий. На рис. изображено несколько графиков функций. Если изменение одной переменной на одну единицу всегда приводит к изменению другой переменной на одну и ту же величину, функция является *линейной* (график ее представляет прямую линию); любая другая связь — *нелинейная*.

Если увеличение одной переменной связано с увеличением другой, то связь — *положительная* (прямая),

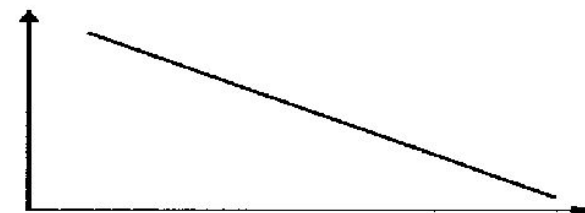
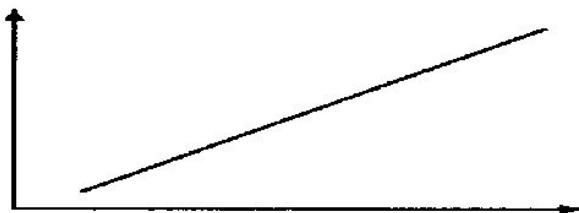
если увеличение одной переменной связано с уменьшением другой, то связь — *отрицательная* (обратная).

Если направление изменения одной переменной не меняется с возрастанием (убыванием) другой переменной, то такая функция — *монотонная*, в противном случае функцию называют *немонотонной*.

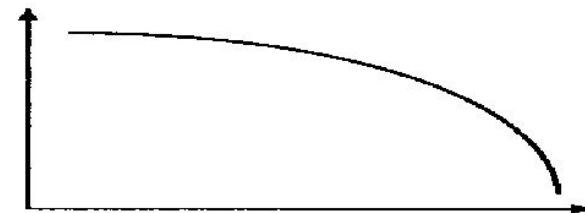
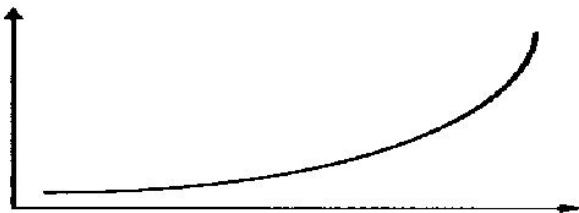
Положительная связь

Отрицательная связь

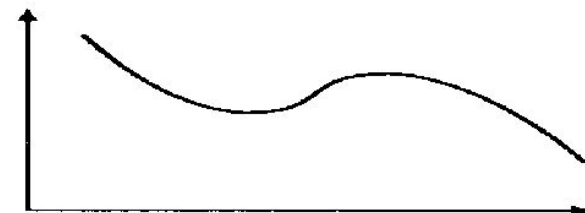
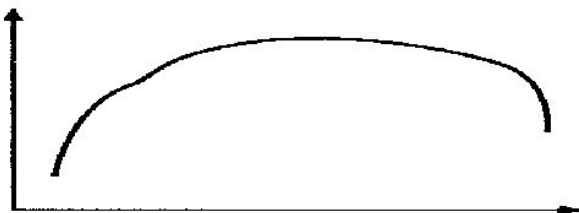
Линейная



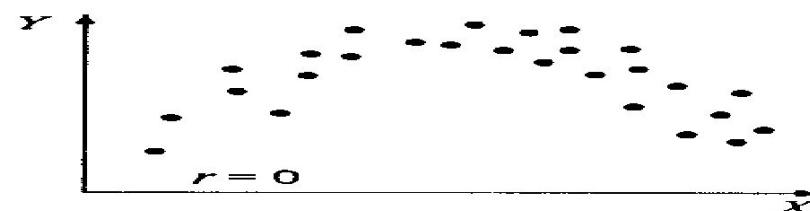
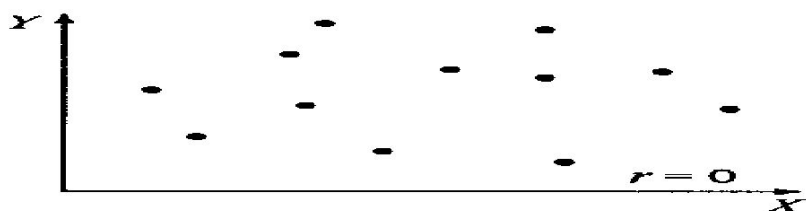
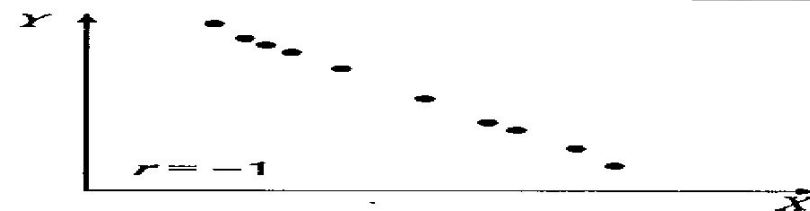
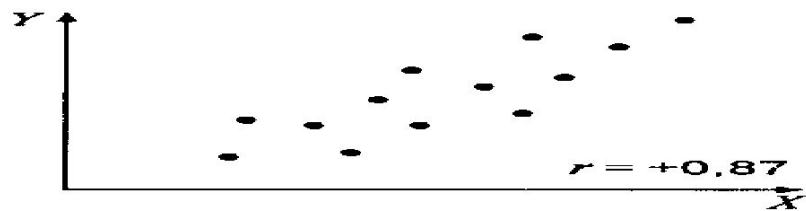
Нелинейная
монотонная



Нелинейная
немонотонная



Примеры графиков часто встречающихся функций



Примеры диаграмм рассеивания и соответствующих коэффициентов корреляции

Коэффициент корреляции — это количественная мера силы и направления вероятностной взаимосвязи двух переменных; принимает значения в диапазоне от -1 до $+1$.

Сила связи достигает максимума при условии *взаимно однозначного соответствия*: когда каждому значению одной переменной соответствует только одно значение другой переменной (и наоборот), эмпирическая взаимосвязь при этом совпадает с функциональной линейной связью.

Показателем силы связи является абсолютная (без учета знака) величина коэффициента корреляции.

Направление связи определяется прямым или обратным соотношением значений двух переменных: если возрастанию значений одной переменной соответствует возрастание значений другой переменной, то взаимосвязь называется *прямой* (положительной); если возрастанию значений одной переменной соответствует убывание значений другой переменной, то взаимосвязь является *обратной* (отрицательной).

Показателем направления связи является знак коэффициента корреляции.

КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ r -ПИРСОНА

r -Пирсона (*Pearson r*) применяется для изучения взаимосвязи двух метрических переменных, измеренных на одной и той же выборке.

Существует множество ситуаций, в которых уместно его применение. Влияет ли интеллект на успеваемость на старших курсах университета? Связан ли размер заработной платы работника с его доброжелательностью к коллегам? Влияет ли настроение школьника на успешность решения сложной арифметической задачи?

Для ответа на подобные вопросы исследователь должен измерить два интересующих его показателя у каждого члена выборки. Данные для изучения взаимосвязи затем сводятся в таблицу. Для определения формулы коэффициента корреляции, проследим логику ее возникновения, используя данные примера.

Положение каждой i -точки (испытуемого с номером i) на диаграмме рассеивания относительно остальных точек (рис.) может быть задано величинами и знаками отклонений соответствующих значений переменных от своих средних величин: $(x_i - M_x)$ и $(y_i - M_y)$.

Если знаки этих отклонений совпадают, то это свидетельствует в пользу положительной взаимосвязи (большим значениям по x соответствуют большие значения по y или меньшим значениям по x соответствуют меньшие значения по y). Таким образом, если произведение отклонений $(x_i - M_x)(y_i -$

$M_y)$ положительное, то данные i -испытуемого свидетельствуют о прямой (положительной) взаимосвязи, а если отрицательное – то об обратной (отрицательной) взаимосвязи. Соответственно, если x и y

в основном связаны прямо пропорционально, то большинство произведений

ПРИМЕР

В таблице приведен пример исходных данных измерения двух показателей интеллекта (вербального и невербального) у 20 учащихся 8-го класса.

№	Вербальный IQ (x)	Невербальный IQ (y)
1	13	12
2	9	11
3	8	8
4	9	12
5	7	9
6	9	11
7	8	9
8	13	13
9	И	9
10	12	10
11	8	9
12	9	8
13	10	10
14	10	12
15	12	10
16	10	10
17	8	11
18	9	10
19	10	11
20	11	13

Средние: **9,8** **10,4**

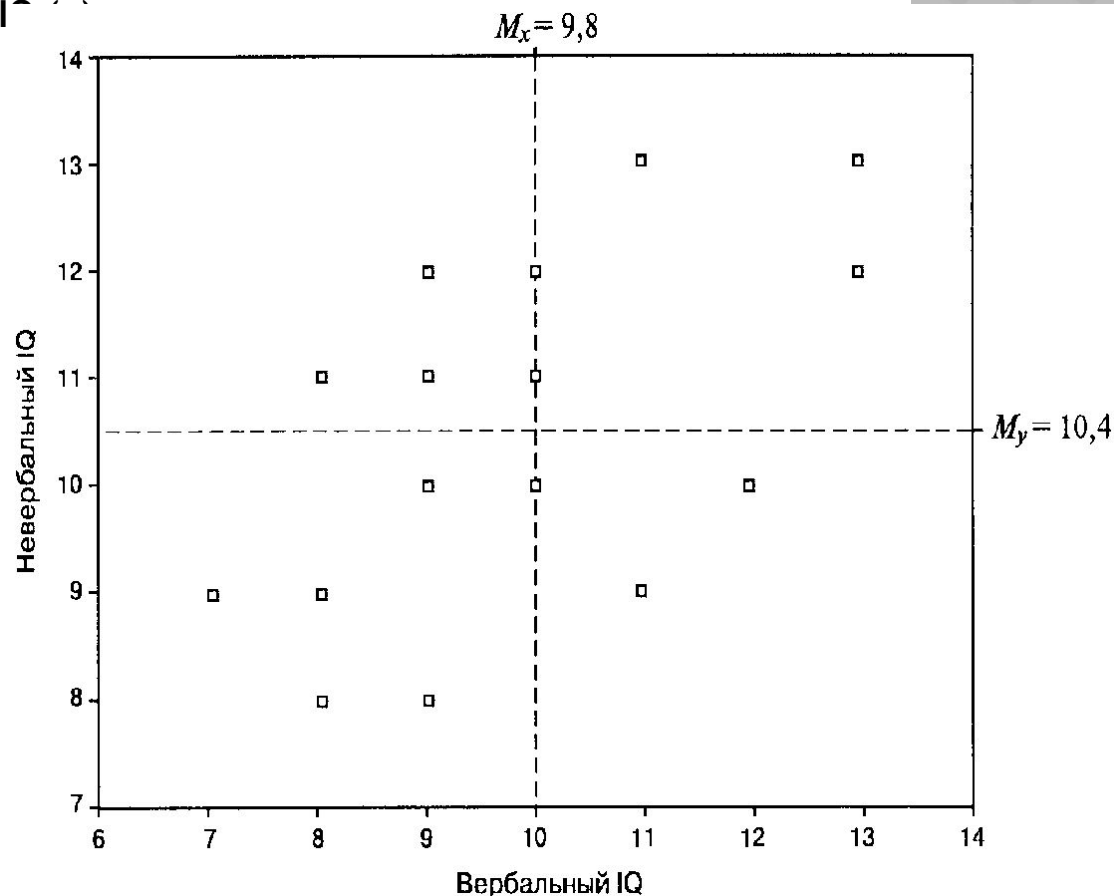


Диаграмма рассеивания для данных примера

Связь между этими переменными можно изобразить при помощи диаграммы рассеивания (см. рис.). Диаграмма показывает, что существует некоторая взаимосвязь измеренных показателей: **чем больше значения вербального интеллекта, тем (преимущественно) больше значения невербального**





















