

Министерство образования Республики Беларусь  
Учреждение образования «Витебский государственный  
университет имени П.М. Машерова»  
Кафедра инклюзивного образования

**Н.В. Кухтова**

**МЕТОДЫ  
СТАТИСТИЧЕСКОГО  
АНАЛИЗА ДАННЫХ**

*Курс лекций*

*Витебск  
ВГУ имени П.М. Машерова  
2025*

УДК 159.9:311:519.25(075.8)  
ББК 88в631я73  
К95

Печатается по решению научно-методического совета учреждения образования «Витебский государственный университет имени П.М. Машерова». Протокол № 1 от 08.09.2025.

Автор: заведующий кафедрой психологии ВГУ имени П.М. Машерова,  
кандидат психологических наук, доцент **Н.В. Кухтова**

Р е ц е н з е н т :  
доцент кафедры среднего общего образования  
и социального проектирования Псковского государственного университета,  
кандидат педагогических наук, доцент *С.Ю. Буренина*

**Кухтова, Н.В.**

**К95**      Методы статистического анализа данных : курс лекций /  
Н.В. Кухтова. – Витебск : ВГУ имени П.М. Машерова, 2025. – 57 с.

Курс лекций по учебной дисциплине «Статистические методы в психологии» содержит описание основных понятий и процедур, определяющих стратегию использования статистических методов в результате обработки данных эмпирического исследования.

Предназначен для специальности 7-07-0114-01 Специальное и инклюзивное образование и будет полезен преподавателям психологических и педагогических дисциплин.

УДК 159.9:311:519.25(075.8)  
ББК 88в631я73

© Кухтова Н.В., 2025  
© ВГУ имени П.М. Машерова, 2025

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ .....	5
МОДУЛЬ 1. Общее представление о статистических методах .....	7
Тема 1. Общие положения методов прикладной статистики .....	7
1.1. Определение статистики .....	7
1.2. Статистическая совокупность .....	8
1.3. Признаки и переменные. Классификация признаков .....	8
1.4. Шкалы измерения признаков .....	10
1.5. Методы статистической обработки .....	13
1.6. Природа статистических данных .....	13
Тема 2. Статистические наблюдения и измерения .....	14
2.1. Определение статистического наблюдения .....	14
2.2. Классификация видов статистического наблюдения .....	15
2.3. Типы вопросов в статистическом бланке .....	15
2.4. Ошибки статистического наблюдения .....	16
2.5. Сводка и группировка статистического материала .....	16
2.6. Вариационный ряд в статистическом анализе .....	18
Тема 3. Статистические показатели .....	19
3.1. Обобщающие статистические показатели .....	19
3.2. Атрибуты статистических показателей .....	19
3.3. Виды статистических показателей: абсолютные, относительные, средние	20
Тема 4. Методы описательной статистики .....	21
4.1. Меры центральной тенденции в описательной статистике .....	21
4.2. Меры изменчивости в описательной статистике .....	23
4.3. Меры и положения разброса в описательной статистике .....	24
Тема 5. Генеральная совокупность в прикладной статистике .....	25
5.1. Генеральная совокупность и выборка .....	25
5.2. Репрезентативная выборка .....	26
5.3. Зависимые и независимые случайные выборки .....	28
Тема 6. Понятие нормального распределения в прикладной статистике .....	29
6.1. Понятие нормального распределения .....	29
6.2. Свойства нормального распределения .....	30
6.3. Проверка нормальности распределения .....	31

Тема 7. Понятие статистической гипотезы. Процедура проверки гипотезы .....	32
7.1. Гипотеза и ее характеристики .....	32
7.2. Проверка статистических гипотез .....	34
7.3. Уровень значимости и уровень достоверности .....	34
7.4. Виды ошибок, возможные при проверке статистических гипотез .....	35
Тема 8. Способы представления статистических данных .....	37
8.1. Графические способы представления статистических данных .....	37
8.2. Табличный способ представления статистических данных .....	39
Тема 9. Показатели корреляционной связи в статистическом ряду .....	41
9.1. Общее понятие о корреляционном анализе и коэффициенте корреляции	41
9.2. Характеристики коэффициентов корреляции .....	43
9.3. Содержательный анализ корреляционных матриц .....	47
МОДУЛЬ 2. Методы сравнения выборок .....	49
Тема 10. Параметрические методы сравнения выборок .....	49
10.1. Анализ различий. Т-тест (критерий Стьюдента) .....	49
10.2. Сравнение долей признака: угловое преобразование $\rho$ -Фишера .....	52
Тема 11. Непараметрические методы сравнения выборок .....	53
11.1. Сравнение двух независимых совокупностей (критерии Манна – Уитни) .....	53
11.2. Сравнение трех и более независимых совокупностей. Критерий Крускала – Уоллиса как непараметрический аналог дисперсионного анализа для независимых совокупностей .....	54
СПИСОК ЛИТЕРАТУРЫ .....	56

## **ВВЕДЕНИЕ**

В условиях интенсивной информатизации общества значение компьютерных технологий обработки эмпирических данных постоянно возрастает. Рост их популярности обусловлен развитием аппаратного и программного обеспечения, совершенствованием методов численной обработки в статистике, анализе данных, методах оптимизации и других отраслях математики. Знакомство с научными подходами, принципами и методами статистики, формирование навыков использования методов эмпирического анализа информации позволяют сформировать необходимые компетенции для принятия решений. В настоящее время методы статистики широко используются, поэтому овладение знаниями и умениями при проведении эмпирических исследований является важным для формирования профессиональной компетентности студентов.

Цель учебной дисциплины – обеспечение специальных педагогов знаниями основ современных методов прикладной статистики, применяемых в научных исследованиях.

Задачи учебной дисциплины:

- повышение статистической грамотности студентов;
- ознакомление студентов с основными положениями теории вероятностей и их применением в математической статистике;
- углубление знаний студентов о различных методах шкалирования, используемых в эмпирических исследованиях, об их связи с методами статистической обработки информации;
- формирование умения квалифицированного использования статистических методов анализа;
- формирование навыков использования полученных знаний в профессиональной деятельности;
- рассмотрение основных параметрических и непараметрических методов обработки количественных данных;
- формирование навыков выбора статистического метода в соответствии с типом данных и гипотезой использования методов при помощи программного обеспечения;
- ознакомление со способами интерпретации результатов применения статистических методов и их презентации в отчетах.

Учебная дисциплина «Методы статистического анализа данных» относится к государственному компоненту, к модулю «Методология и методы научного исследования».

Содержание учебной дисциплины «Методы статистического анализа данных» позволяет углубить и расширить знания в исследовательской деятельности, сформировать практические умения и навыки по применению методов математической статистики для анализа экспериментальных данных исследования. Особенностью этой учебной дисциплины является ее интегративный характер, что предполагает реализацию междисциплинарных связей с учебными дисциплинами модуля «Современные технологии и средства коммуникации в образовании», в частности «Информационные технологии в образовании», и модуля «Методология и методы научного исследования», дисциплины «Методология научного исследования».

Требования к освоению учебной дисциплины «Методы статистического анализа данных» определены образовательным стандартом высшего образования, который разработан с учетом компетентностного подхода.

Освоение учебной дисциплины должно обеспечить формирование следующих компетенций:

*Информационно-коммуникационные компетенции:*

– владеть методологией поиска нового (информации, идей и т.п.), методикой анализа и адаптации знаний к своим профессиональным потребностям,

– уметь преобразовывать информацию в специальные знания,

– формировать информационную культуру обучающихся;

*проектные компетенции*

– осознанно и самостоятельно планировать профессиональную деятельность,

– осмысленно строить профессиональную карьеру,

– находить оптимальные решения инновационного характера,

– быть способными воплотить свои замыслы в инновационный проект и реализовать его,

– системно совершенствовать образовательный процесс,

– формировать у обучающихся способность к построению собственной образовательной траектории.

В рамках образовательного процесса по учебной дисциплине «Методы статистического анализа данных» студент должен приобрести не только теоретические и практические знания, умения и навыки по специальности, но и развить свой ценностно-личностный, духовный потенциал, сформировать качества патриота и гражданина, готового к активному участию в экономической, производственной, социально-культурной и общественной жизни страны.

**В результате изучения учебной дисциплины студент должен:**

**знать:**

– виды измерительных шкал;

– меры центральной тенденции и изменчивости;

– параметрические и непараметрические критерии;

**уметь:**

– проверять эмпирические распределения на нормальность;

– применять статистические расчеты для различных видов измерительных шкал;

– выявлять различия в уровне исследуемого признака;

– проводить оценку достоверности сдвига исследуемого признака;

– использовать методы корреляционного анализа;

**владеть:**

– навыками формулирования статистических гипотез для различных исследовательских задач;

– алгоритмом выбора соответствующего метода проверки статистических гипотез;

– техниками получения и интерпретации описательной, аналитической и дифференциальной статистики;

– правилами презентации результатов статистической обработки данных в устных и письменных отчетах.

# МОДУЛЬ 1

## ОБЩЕЕ ПРЕДСТАВЛЕНИЕ

### О СТАТИСТИЧЕСКИХ МЕТОДАХ

#### Тема 1. Общие положения методов прикладной статистики

- 1.1. *Определение статистики.*
- 1.2. *Статистическая совокупность.*
- 1.3. *Признаки и переменные. Классификация признаков.*
- 1.4. *Шкалы измерения признаков.*
- 1.5. *Методы статистической обработки.*
- 1.6. *Природа статистических данных.*

#### 1.1. *Определение статистики*

**С**татистика – наука, изучающая количественную сторону массовых общественных явлений с целью установления закономерностей в неразрывной связи с их качественной стороной в конкретных условиях места и времени в их взаимосвязи и взаимозависимости.

В настоящее время данный термин употребляется в трех значениях:

- 1) под статистикой понимают отрасль практической деятельности, которая имеет своей целью сбор, обработку, анализ и публикацию массовых данных о самых различных явлениях общественной жизни (в этом смысле «статистика» выступает как синоним словосочетания «статистический учет»);
- 2) статистикой называют цифровой материал, служащий для характеристики какой-либо области общественных явлений или территориального распределения какого-то показателя;
- 3) статистикой называется отрасль знания, особая научная дисциплина и соответственно учебный предмет в высших и средних специальных учебных заведениях.

Предмет статистики – изучение и анализ массовых общественных явлений и совокупностей.

В свою очередь, математическая статистика (от итальянского «*statio*» – государство) – раздел математики, который изучает методы сбора, систематизации, обработки и использования статистических данных для получения научных и практических выводов.

Первая задача математической статистики – определить способы сбора и группировки статистических данных.

Вторая задача математической статистики состоит в разработке методов анализа статистических данных в зависимости от целей исследования.

Статистические данные – это совокупность чисел, полученных эмпирическим путем, представляющая собой количественные признаки изучаемых объектов. Эти данные получают в результате специально организованных исследований, в том числе психологических.

Математическая статистика включает три раздела:

- 1) описательная (дискриптивная) статистика занимается описанием, систематизацией, графическим представлением и табулированием данных, полученных в ходе исследования, а также выявлением центральных тенденций распределения и оценки разброса данных;
- 2) индуктивная (аналитическая) статистика или теория статистического вывода занимается проверкой того, можно ли результаты, полученные на ограниченной сово-

купности объектов, распространить на всю совокупность объектов данного вида. Базируется на описательной статистике;

3) планирование и анализ эксперимента – статистические методы, разработанные для обнаружения и проверки причинной связи между изучаемыми показателями (переменными).

## **1.2. Статистическая совокупность**

**С**овокупность представляет собой однородную по какому-либо признаку группу, которая состоит из ядра и окружающих его явлений («слоев»). Ядро – концентрированное выражение всех специфических свойств данной группы, отличающих одну совокупность от других. «Слой» – единицы с неполным набором специфических свойств, которые принадлежат к данной совокупности с определенной вероятностью.

Статистическая совокупность – это множество объективно существующих единиц изучаемого явления, объединенных единой качественной основой, общей связью, но отличающихся друг от друга отдельными признаками.

## **1.3. Признаки и переменные. Классификация признаков**

**П**ризнаки и переменные – это измеряемые психологические явления. Такими явлениями могут быть время решения задачи, количество допущенных ошибок, уровень тревожности, показатель интеллектуальной лабильности, интенсивность агрессивных реакций, угол поворота корпуса в беседе, показатель социометрического статуса и множество других переменных.

Понятия признака и переменной могут использоваться как взаимозаменяемые. Они являются наиболее общими. Иногда вместо них используются понятия показателя или уровня, например, уровень настойчивости, показатель вербального интеллекта и др. Понятия показателя и уровня указывают на то, что признак может быть измерен количественно, так как к ним применимы определения «высокий» или «низкий», например, высокий уровень интеллекта, низкие показатели тревожности и др.

**П**ризнак – это качественная особенность единицы совокупности.

Признак может быть классифицирован по разным основаниям:

1. По характеру выражения:

1.1. Описательные (качественные) – признаки, выражаются словесно (например, форма собственности предприятия, вид покупаемой продукции, направление подготовки и т. д.). Описательные признаки подразделяют на номинальные, которые нельзя упорядочить, ранжировать (например, национальность, отраслевая принадлежность предприятия и др.), и порядковые, которые можно ранжировать (например, балл успеваемости студента, тарифный разряд, рейтинги компаний и др.).

1.2. Количественные – признаки, отдельные значения которых имеют числовое выражение (например, площадь территории области, стоимость фондов предприятия, цена товара и т. д.).

2. По способу измерения:

2.1. Первичные признаки – это признаки, которые характеризуют единицу совокупности в целом. Они могут быть измерены, сосчитаны, взвешены и существуют сами по себе, независимо от их статистического изучения (например, численность обучающихся, валовой сбор картофеля, сумма страховых выплат и др.).

2.2. Вторичные признаки получают расчетным путем через соотношение первичных признаков. Вторичные признаки являются продуктами человеческого сознания,

результатами познания изучаемого объекта (например, удовлетворенность браком, уровень успеваемости и др.).

3. По отношению к характеризуемому объекту:

3.1. Прямые признаки – свойства, присущие тому объекту, который ими характеризуется (например, студенческая стипендия, доход предприятия и т. п.).

3.2. Косвенные признаки – свойства, присущие не самому изучаемому объекту, а другим совокупностям, относящимся к объекту (например, оплата представителей определенной профессии, средняя заработка и т. п.).

4. По характеру вариации:

4.1. Альтернативные признаки – те, которые принимают только одно из имеющихся значений (пол человека, место проживания; например, мужчины – женщины, город – село), признаки обладания или необладания чем-то (например, состоит в браке – не состоит в браке).

4.2. Дискретные – это признаки, которые имеют только целочисленные значения (например, количество домашних животных в семье).

4.3. Непрерывные – признаки, способные принимать любые значения, как целые, так и дробные. К непрерывным признакам относятся все вторичные признаки.

5. По отношению ко времени:

5.1. Моментные признаки – характеристики состояния, наличия чего-либо на определенный момент времени (например, численность обучающихся на 1 сентября).

5.2. Интервальные признаки – характеристики процесса за определенный промежуток времени: год, полугодие, квартал, месяц, сутки и т. д. (например, количество пропусков учащимися занятий за учебный год).

**Переменные** – параметр реальности, который может изменяться и / или изменяется в экспериментальном исследовании. Различают независимые переменные, зависимые, внешние (побочные), латентные, дополнительные и т.д. [3; С. 20].

Психологические переменные являются случайными величинами, поскольку заранее неизвестно, какое именно значение они примут.

Переменная может быть символом, используемым для обозначения неуказанного заранее элемента множества

Пример: если выборка данных состоит из студентов нескольких уровней (лет) обучения, то уровень обучения является переменной.

Зависимая переменная:

- описывается в терминах других переменных;
- «предсказываемая»;
- выходная, критериальная или переменная отклика на некоторые изменения; измеряется.

Независимая переменная:

- описывает или предсказывает некоторую переменную отклика;
- изменяемая в процессе манипуляций или действий человека или естественным образом во времени и/или «пространстве»;
- стимул, фактор или входная переменная.

Следует отметить, что статистическая обработка – это оперирование со значениями признака, полученными у испытуемых в психологическом исследовании. Такие индивидуальные результаты называют также «наблюдениями», «наблюдаемыми значениями», «вариантами», «датами», «индивидуальными показателями» и др.

Значения признака определяются при помощи специальных шкал измерения.

#### **1.4. Шкалы измерения признаков**

**И**змерение – приписывание значений признакам объекта в соответствии с определенными правилами или шкалой измерения. С. Стивенс выделил четыре способа измерения, которые назвал шкалами. Первые два способа – неметрические, это номинальная и ранговая шкалы. К метрическим шкалам относятся шкала интервалов и шкала отношений. Так, тип шкалы определяется исследователем исходя из операционных определений переменных. Выбор шкалы, наряду с поставленными в исследовании задачами, предопределяет подходящий статистический анализ.

**Наиболее часто употребляются следующие шкалы измерения:**

1. Номинальная (наименований)
2. Порядковая (порядка, ранга)
3. Интервальная (интервалов)
4. Относительная (равных отношений)

**Ш**кала наименований (номинальная, номинативная, шкала классификации – неметрическая) предполагает простое обозначение числом или символом определенного класса объектов, объединенных общим признаком. В этом случае число выполняет чисто номинальную функцию. Каждому классу дается наименование и обозначение, обычно числовое и присваивается соответствующее наименование. Объекты, отнесенные к одному и тому же классу, получают одни и те же обозначения. После наименования объектов могут быть подсчитаны их количество в каждом классе и коэффициент ассоциации, характеризующий степень совпадения классов. Допустима замена одного числа-наименования другим при условии, что различные классы будут иметь неодинаковые наименования. Таким образом, объекты группируются по разным классам так, чтобы внутри класса они были идентичны по измеряемому свойству.

Выделяют несколько вариантов обозначения шкалы наименований:

– *категоризованная*. В том случае, если количество классов шкалы известно, а также известны правила отнесения к ним объекта (например: пол м – ж);  
– *дихотомическая*. Состоит только из двух классов (например: любит – не любит).

Для номинальной шкалы арифметические операции не имеют смысла, так как классифицируются исходные объекты на классы, и можно перейти от наименований к числам, подсчитав количество наблюдений в каждом из классов. Такая величина называется частотой. Учитывается только одно свойство чисел – то, что это разные символы. Остальные свойства чисел не учитываются. Привычные операции с числами — упорядочивание, сложение-вычитание, деление – при измерении в номинативной шкале теряют смысл. При сравнении объектов можно сделать вывод только о том, принадлежат они к одному или разным классам, тождественны или нет по измеренному свойству.

Номинальные переменные могут принимать дискретные, не связанные друг с другом значения. Вопросы анкеты, кодируемые номинальными переменными, могут быть как закрытыми (с вариантами ответов), так и открытыми (с текстовым полем вместо прямого указания вариантов ответа).

**Р**анговая (порядковая, ординальная – неметрическая). Ранговая шкала основана на ранжировании объектов в порядке нарастания или убывания величины определенного, общего для объектов признака. Измерение в этой шкале предполагает приписывание объектам чисел в зависимости от степени выраженности измеряемого свойства. При сравнении испытуемых друг с другом можно сказать, больше или меньше выражено свойство, но нельзя сказать, насколько больше или насколько меньше оно выражено. В порядковой шкале должно быть не менее 3-х классов. Например, отметки по результатам экзамена.

Чем больше классов в шкале, тем больше возможности для математической обработки полученных данных. В общих случаях числа в порядковой шкале не отражают количества свойства, которыми обладают исследуемые объекты. Поэтому для этой шкалы арифметические операции также чаще всего не имеют смысла. Фактически в качестве единицы измерения в порядковой шкале используется расстояние в 1 ранг, но при этом расстояние между соседними рангами может быть различным. При измерении в ранговой шкале, таким образом, из всех свойств чисел учитывается то, что они разные, и то, что одно число больше, чем другое.

При ранжировании «вручную», а не при помощи компьютера, существует два обстоятельства:

1. Следует установить для себя и запомнить порядок ранжирования. Так можно ранжировать испытуемых по их «месту в группе»: ранг 1 присваивается тому, у которого наименьшая выраженность признака, и далее – увеличение ранга по мере увеличения уровня признака. Или можно ранг 1 присваивать тому, у которого 1-е место по выраженности данного признака (например, «самый быстрый»). Строгих правил выбора здесь нет, но важно помнить, в каком направлении производилось ранжирование.

2. Следует соблюдать правило ранжирования для связанных рангов, когда двое или более испытуемых имеют одинаковую выраженность измеряемого свойства. В этом случае таким испытуемым присваивается один и тот же, средний ранг. Например, если ранжировать испытуемых по «месту в группе» и двое имеют одинаковые самые высокие исходные оценки, то обоим присваивается средний ранг 1,5:  $(1+2)/2 = 1,5$ . Следующему за этой парой испытуемому присваивается ранг 3 и т. д. Это правило основано на соглашении соблюдения одинаковой суммы рангов для связанных и несвязанных рангов.

Так, ранжирование позволяет определить объект, относящийся к середине ряда (медиану), и подсчитать процент объектов с меньшей или большей величиной признака. Допустима замена чисел, характеризующих величину признака, любыми другими, не нарушающими ранговый порядок объектов.

В том числе, порядковые переменные кодируют такие закрытые вопросы, варианты ответа на которые подчиняются логическому числовому порядку. То есть варианты ответа на такие вопросы представляют собой связанные между собой группы значений.

**Интервальная шкала (метрическая)** позволяет классифицировать и упорядочивать объекты, а также количественно описать различия между свойствами объектов. *Интервальная шкала* основана на сравнении различия между объектами по величине измеряемого признака или свойства. Различия оцениваются определенным эталоном, интервалом (отсюда и название данной шкалы).

Для интервальной шкалы характерно отсутствие абсолютного нуля. За нуль может быть принята любая точка шкалы. Отсутствие абсолютного нуля не позволяет сравнивать отношения величин. Для задания такой шкалы устанавливают единицу измерения и произвольную точку отсчета (например, календарное время). Для этой шкалы арифметические операции имеют смысл.

Интервальная шкала позволяет подсчитывать кроме параметров, характерных для номинальной и ординальной шкал, средние арифметические, средние квадратичные отклонения и коэффициенты корреляции. Данная шкала имеет очень важное с точки зрения конструирования тестов свойство: возможность перевода величин из одной измерительной шкалы в другую с использованием линейного преобразования  $Y = aX + b$ . В большинстве психологических опросников (Айзенка, Кеттелла и др.) шкала интервалов реализована как итоговая оценочная шкала. Применяя эту шкалу, можно судить, насколько больше или насколько меньше выражено свойство при сравнении объектов, но нельзя судить о том, во сколько раз больше или меньше выражено свойство.

Измерение в этой шкале предполагает возможность применения *единицы измерения (метрики)*. Объекту присваивается число единиц измерения, пропорциональное выраженности измеряемого свойства.

Так, интервальными являются переменные, не имеющие выделенных категорий. Они содержат числовые данные (например, номер анкеты в базе данных) и кодируют чаще всего открытые вопросы. Интервальные переменные (или другие типы переменных, приводимые к интервальному виду) используются практически во всех статистических процедурах.

**Шкала отношений (абсолютная шкала)** характеризуется наличием абсолютного нуля, что делает допустимым сравнение отношения двух величин. Шкала отношений отличается от интервальной шкалы только тем, что в ней задано абсолютное начало отсчета. Например, рост в см – абсолютное начало 0. В шкале отношений можно определить не только на сколько одно измерение превосходит другое, но и во сколько раз. (В.В. Сечко).

Для этой шкалы допустимо только преобразование вида  $Y = aX$ , которое не изменяет отношения между величинами. В дополнение к статистикам предыдущих шкал можно вычислять коэффициент вариации. С помощью шкалы отношений производятся измерения в психофизиологических исследованиях, где единицы измерения — физические величины, при этом предполагается равенство отношения степеней выраженности психических свойств у двух объектов отношению двух чисел, приписанных этим объектам по данному свойству. Отношение: порядок и интервал + включает «нуль» и может быть смысловым, то есть иметь единицу измерений. Числа могут сравниваться в процентном отношении.

Таблица 1 – Шкалы измерения признаков (по А.Д. Наследову)

Название шкалы	Основная операция	Допустимое преобразование	Допустимые статистические величины	Примеры
Номинальная	Приписывание одинаковых чисел-наименований объектам, имеющим общий признак	Замена одного, числа-наименования другим	Число объектов в классе. Коэффициент ассоциации	Классификация испытуемых по полу
Ранговая	Ранжирование объектов по выраженности определенного признака	Замена на любую монотонно возрастающую функцию	Медиана. Процентили	Шкала ценностных ориентаций
Интервальная	Определение величины различий между объектами	$Y = aX + b$	Средняя арифметическая. Стандартное квадратичное отклонение. Коэффициент корреляции	Итоговая оценочная шкала опросника Айзенка
Отношений	Определение равенства отношений величин	$Y = aX$	Коэффициент вариации	Измерение психофизиологических параметров

Важно понимать, какая именно шкала используется, так как каждый способ обработки экспериментальных данных рассчитан на определенный тип шкал. Применение математических методов, неадекватных данным, приводит к странным, а часто и ложным результатам.

Таким образом, неметрические шкалы менее мощные и отражают меньше информации о различии объектов (испытуемых) по измеренному свойству, и, напротив, метрические шкалы более мощные, они лучше дифференцируют испытуемых. Поэтому следует применить более мощную шкалу. При отсутствии выбора приходится использовать доступную измерительную шкалу.

### **1.5. Методы статистической обработки**

**М**етоды статистической обработки – это способы количественных расчетов, математические формулы и приемы, которые позволяют обобщать эмпирические данные, выявляя скрытые в них закономерности. Выделяют следующие методы:

1. *Первичные методы*. С помощью них получают показатели, непосредственно отражающие результаты эмпирических исследований: вычисление мер центральной тенденции (среднего значения, моды, медианы), мер изменчивости (размаха, дисперсии, стандартного отклонения), наглядное представление данных в виде графиков и диаграмм.

2. *Вторичные методы*. Используя первичные данные, позволяют выявить скрытые статистические закономерности, произвести качественный анализ данных: выдвижение статистических гипотез, подготовка данных для применения статистических методов (например, проверка нормальности распределения, уравновешивание дисперсионных комплексов и др.), проверка гипотез с помощью выбранных статистических критериев, формулирование выводов, имеющих определенную доверительную вероятность.

Использование математико-статистического анализа эмпирических данных позволяет:

- 1) более четко и лаконично описывать изучаемые объекты, обобщать данные исследования;
- 2) выявлять наличие существенных различий между группами, количественно сравнивая исследуемые признаки;
- 3) устанавливать скрытые причины и суть психологических явлений;
- 4) повышать доказательность выводов, сопроводив их статистическим подтверждением.

### **1.6. Природа статистических данных**

**П**рирода статистических данных обусловлена тремя основными свойствами: 1. Неопределенность статистических данных.  
2. Абстрактность статистических данных.

3. Вероятностный характер статистических данных (признак может принять это значение, а может и не принять).

Рассмотрим на примере изучения успеваемости студентов 1 курса. Уровень успеваемости студентов – это соотношение объемов двух понятий, которые без конкретизации обладают неопределенностью, абстрактностью, вероятностным характером.

Неопределенность «уровня успеваемости студентов» может проявляться в таких аспектах, как:

1) неопределенность объемов для соотношения количества «успевающих» и «неуспевающих» студентов или количества «успевающих» студентов и общего количества студентов первого курса;

2) неопределенность самого понятия «студент»: в расчет принимаются только очные студенты, или только заочные студенты, или и те и другие;

3) неопределенность понятия «первый курс» – учитываются студенты бакалавриата/магистратуры/аспирантуры все вместе или по отдельности.

Абстрактность «уровня успеваемости студентов» заключается в том, что необходимо осуществить ряд уточнений:

1) «успевающий студент» – это студент со всеми пятерками или без троек;

2) «успевающий студент» – это «успевающий» по всем предметам или только по части предметов;

3) может ли студент, который не ходит на занятия (текущая успеваемость «0»), но сдает все экзамены на «отлично», быть отнесен к категории «успевающий»?

Вероятностный характер «уровня успеваемости студентов» может проявляться в том, что:

1) результаты сдачи студентом сессий носят вероятностный характер: даже если при сдаче первой сессии студент «неуспевающий», то это не значит, что и по результатам сдачи второй (и последующих) сессии он останется «неуспевающим»;

2) уровень успеваемости может варьироваться в зависимости от специальности, изучаемых в семестре курсов;

3) по результатам промежуточных тестирований студент – «успевающий», а по результатам итогового экзамена – «неуспевающий».

## **Тема 2. Статистические наблюдения и измерения**

### **2.1. Определение статистического наблюдения.**

### **2.2. Классификация видов статистического наблюдения.**

### **2.3. Типы вопросов в статистическом бланке.**

### **2.4. Ошибки статистического наблюдения.**

### **2.5. Сводка и группировка статистического материала.**

### **2.6. Вариационный ряд в статистическом анализе.**

### **2.1. Определение статистического наблюдения**

**Статистическое наблюдение** – планомерная регистрация признаков у элементов статистической совокупности.

Статистическое исследование состоит из нескольких этапов:

1. Проведение мероприятий по подготовке исследования.

2. Сбор первичных данных (непосредственное статистическое наблюдение).

3. Сводка, то есть систематизация и группировка статистических данных, что позволяет описать отдельные группы и исследуемый объект в целом.

4. Анализ, обобщение полученных данных, обнаружение закономерностей.

Результаты представлены в форме аналитических отчетов, таблиц, графиков, научных статей.

В зависимости от организации выделяют три формы статистического наблюдения:

1. Отчетность – форма статистического наблюдения, при которой сведения поступают в статистические органы от отдельных предприятий, организаций, учреждений и т. д. в порядке предоставления обязательной информации об их деятельности, по заранее установленным программам и в заранее оговоренные сроки (например, отчетность загсов, финансовая отчетность).

2. Специально организованное статистическое наблюдение – наблюдение, проводимое с определенной целью, на определенную дату, для получения информации, которая не может быть собрана посредством отчетности (например, переписи, бюджетные исследования, исследование проституции, изучение брачно-семейных ориентаций студентов и др.).

3. Регистровые наблюдения (регистры) – форма непрерывного статистического наблюдения за долговременными процессами, имеющими фиксированное начало, стадию развития и фиксированный конец (например, регистрационные больничные карты, студенческие зачетки, карточки, библиотечные формуляры, личный листок сотрудника).

## **2.2. Классификация видов статистического наблюдения**

В зависимости от сочетания различных условий выделяют следующие разновидности статистического наблюдения.

**Классификация видов** статистического наблюдения:

1. По объему охватываемых единиц:
  - сплошное;
  - несплошное: монографическое, основной массив, выборочное.
2. По моменту регистрации по времени:
  - текущее;
  - прерывное (периодическое, единовременное).
3. По источнику получаемой информации:
  - непосредственное наблюдение;
  - опрос;
  - документальное.
4. По способу сбора информации:
  - экспедиционное;
  - корреспондентское;
  - явочное;
  - саморегистрация.

## **2.3. Типы вопросов в статистическом бланке**

Статистический бланк исследования (форма, анкета, опросник) – методический документ, на основе которого проходит исследование статистической совокупности.

**T**ипы вопросов в статистическом бланке:

1. По структуре вопроса: открытые, закрытые, полузакрытые.
2. По форме вопроса: прямые, косвенные.
3. По функциям: программно-тематические, служебные, контрольные, контактные, вопрос-фильтры.
4. По оформлению: текстовые, вопросы-рисунки, графики, табличные вопросы.

## **2.4. Ошибки статистического наблюдения**

**Ошибки статистического наблюдения** – расхождение полученных данных с реальной действительностью. Ошибки могут возникнуть на разных этапах статистического наблюдения.

Виды ошибок:

1. Ошибки выборки – ошибки, связанные с неправильным отбором единиц исследования.
2. Ошибки наблюдения (регистрации) – ошибки связанные с нарушением процедуры регистрации исследуемого признака:
  - 2.1. преднамеренные ошибки – ошибки, тенденциозно искажающие информацию: завышение или занижение тенденций;
  - 2.2. непреднамеренные ошибки:
    - 2.2.1. случайные ошибки – описки, ошибки, оговорки (нивелируются посредством закона больших чисел);
    - 2.2.2. систематические ошибки – ошибки аккумуляции, ошибки округления количественных параметров.
3. Ошибки исчисления – ошибки, связанные с вводом статистических данных и расчетом показателей и коэффициентов.

## **2.5. Сводка и группировка статистического материала**

На основе всестороннего теоретического анализа сущности и содержания изучаемых явлений и процессов проводится статистическая сводка. Программой и планом проведения статистической сводки обеспечивается достоверность и обоснованность ее результатов.

**Сводка** – процесс обобщения статистических данных, сведение единичных фактов воедино. Цель сводки – получение обобщающих статистических показателей.

Статистическая сводка также рассматривается как переход от единичных данных к сведениям о группах единиц и совокупности в целом.

Проведение сводки включает три этапа:

- 1) предварительный контроль – это проверка данных;
- 2) группировка данных по заданным признакам – это определение производных показателей;
- 3) оформление результатов сводки в виде статистических таблиц, они являются удобной формой для восприятия полученной информации.

*Разновидности сводки в статистическом наблюдении:*

1. В зависимости от способа организации:
  - 1.1. Централизованная – сводка, при которой данные, полученные в результате наблюдения, поступают в единый центр для их последующей обработки.
  - 1.2. Децентрализованная – сводка, при которой результаты наблюдения проходят обработку в несколько этапов, отличающихся между собой по степени обобщения материала (любые виды отчетности, результаты переписей).
2. В зависимости от техники обработки:
  - 2.1. Ручная сводка – результаты обрабатываются вручную – методом «а» («конверт», «домик»).
  - 2.2. Механизированная сводка – информация с бланка наблюдения сначала набивается на перфокарты или перфоленты, а потом обрабатывается (сейчас не используется).

2.3. Автоматизированная сводка – обработка с помощью компьютерных программ обработки: SPSS (статистический пакет социальной информации) и др.

3. В зависимости от глубины преобразований:

3.1. Простая сводка – сведение данных в статистические таблицы, подведение общих итогов по совокупности в целом и дает возможность определить число единиц изучаемой совокупности и объем изучаемых признаков, но тем самым простая сводка не дает нам представления о целостности состава изучаемой совокупности.

3.2. Сложная сводка – сведение данных с использованием метода группировки по определенной программе, позволяет нам изучить состав совокупности и выявить влияние одних признаков на другие, т. е. раскрыть свойственные данной совокупности закономерности.

Программа статистической сводки содержит группы, на которые может быть разбита или разбивается совокупность единиц статистического наблюдения, а также систему показателей, характеризующих изучаемую совокупность явлений и процессов как в целом, так и отдельных ее частей. От целей и задач исследования зависит программа статистической сводки. Разработка программы включает следующие этапы:

- 1) выбирается группировочный признак для образования однородных групп;
- 2) определяется порядок формирования и число групп;
- 3) разрабатывается система статистических показателей для характеристики групп и объекта в целом;
- 4) создаются макеты статистических таблиц для предоставления результатов сводки.

**Группировка** – деление совокупности на группы, одно родные по одному или нескольким признакам. Например, «социологии», «политологи» – группы, однородные по признаку «направление подготовки». Признак – это измеряемые свойства или особенности единиц совокупности. Все признаки делятся на качественные и количественные. Группировочный признак – признак, положенный в основу группировки.

Все группировки можно разбить на три вида в зависимости от целей и содержания группировки:

1. Типологическая группировка – предполагает выделение и изучение крупных социальных или экономических типов. На пример, группировка по полу, по профессии.
2. Структурная группировка – предполагает изучение структуры явления (принципы организации явления). Элементы совокупности, значения признака строго упорядочены между собой, выстроены в порядке возрастания или убывания исследуемого признака. Например, группировка по уровню образования.
3. Аналитическая группировка – предполагает выявление связи, зависимости между исследуемыми признаками. Например, влияние дохода на оценку деятельности российского правительства.

В аналитической группировке всегда присутствует два признака:

- факторный признак – положен в основу группировки и обуславливает изменение, вариацию другого признака (доход);
- результативный признак – тот признак, который изменяется под воздействием факторного признака.

По итогу сводки и группировки можно получить статистические ряды распределения.

## **2.6. Вариационный ряд в статистическом анализе**

Вариация – это свойство статистической совокупности, отражающее способность к изменению, обусловленное как внешними, так и внутренними факторами, как связанными с сущностью исследуемого объекта, так и не связанными с ней.

**Вариационный (статистический) ряд** – таблица, первая строка которой содержит в порядке возрастания элементы  $x_i$ , а вторая – их частоты  $m_i$  ( относительные частоты  $f_i$ ). В строке  $x_i$  расположены в возрастающем порядке баллы, полученные испытуемыми по шкале реактивной тревожности; в строке  $m_i$  указана частота каждого балла, полученного испытуемыми.

Статистический ряд – результаты сводки, где статистическая совокупность и элементы, которые ее составляют (единицы наблюдения), упорядочены по какому-либо принципу.

Определение вида статистического ряда – важная процедура статистического анализа. Она позволяет определить алгоритм расчета (выбрать необходимую формулу) статистического показателя.

Для определения вида статистического ряда необходимо:

1) внимательно посмотреть, что содержательно характеризует материал таблицы или рисунка. И тогда:

– если он описывает распределение признака на некоторой территории, то это территориальный ряд;

– если в таблице или на рисунке представлено изменение признака во времени, то это ряд динамики;

– если материал таблицы или рисунка содержательно характеризует статистическую совокупность и ее элементы упорядочены по какому-либо другому принципу (не территории и времени), то это ряд распределения;

2) далее работаем только с рядами распределения:

– если признак в таблице или на рисунке качественный (выражен словесно), то это атрибутивный ряд, а если количественный, то вариационный ряд. Важно понимать, что иногда качественные признаки могут быть представлены в виде чисел, а количественные – слов. На пример, для сохранения конфиденциальности или для сокращения объема текста качественный признак – название фирмы – может быть выражен в виде ее порядкового номера. Или количественный признак – количество детей – в виде слов: один, два, три и т. д.;

3) в вариационном ряду определяем его подвид и на какие элементы делится признак. Если элементы признака представлены в виде целого числа, то это дискретный ряд, а если в виде интервала чисел, то интервальный.

## **Тема 3. Статистические показатели**

- 3.1. Обобщающие статистические показатели.*
- 3.2. Атрибуты статистических показателей.*
- 3.3 Виды статистических показателей: абсолютные, относительные, средние.*

### ***3.1. Обобщающие статистические показатели***

**Л**юбое статистическое исследование связано с расчетом и анализом статистических показателей. **Статистический показатель** представляет собой количественную характеристику явлений и процессов. Как правило, изучаемые процессы и явления сложны, их сущность нельзя отразить одним показателем. Поэтому возникает необходимость в применении системы статистических показателей.

**Статистические показатели** – это количественные характеристики совокупности, а также ее частей.

**Система статистических показателей** – это совокупность взаимосвязанных показателей, предназначенных для решения конкретной задачи.

**Обобщающие статистические показатели** – показатели, полученные в результате сводки путем перехода от индивидуальных значений признаков совокупности к характеристике всей совокупности.

Все статистические показатели по охвату единиц совокупности разделяются на индивидуальные и сводные, а по форме выражения – на абсолютные, относительные и средние. Индивидуальные показатели характеризуют отдельный объект или единицу совокупности. Сводные показатели в отличие от индивидуальных характеризуют группу единиц, представляют часть или всю совокупность. Эти показатели, в свою очередь, подразделяются на объемные и расчетные.

Объемные показатели получают путем сложения значений единиц совокупности. Полученная величина может выступать в качестве абсолютного показателя или относительного. Расчетные показатели, вычисляемые по формулам, служат для решения задач анализа. Они делятся на абсолютные, относительные, средние. В зависимости от принадлежности к одному или двум объектам различают однообъектные и межобъектные показатели. Первые характеризуют только один объект, а вторые получают в результате сопоставления двух величин.

### ***3.2. Атрибуты статистических показателей***

Все статистические показатели имеют характеристики – **атрибуты**:

- 1. Качественная сторона статистического показателя: объект и его свойство.

Она отражает сущность изучаемого свойства статистической совокупности без указания места, времени и возможности определения числового значения. Определяется она понятиями, входящими в наименование показателя и связана с функцией, которую выполняет показатель. Основные функции статистического показателя:

- плановая (показатель в плане, норматив),
- отчетная (показатель в отчете),
- прогностическая (оценочная) функция.

2. Количественная сторона статистического показателя: число и единицы измерения. Отражает методологию расчета (формулу), число и единицу измерения.

3. Пространственные границы: территориальные, отраслевые и иные границы объекта (статистического показателя).

#### 4. Интервал или момент времени измерения.

Для получения правильных выводов при сравнении, анализе, обобщении показателей необходимо, чтобы показатели были сопоставимы друг с другом. Сопоставляемые показатели не должны отличаться более, чем одним из перечисленных атрибутов.

### **3.3. Виды статистических показателей: абсолютные, относительные, средние**

В зависимости от методов расчета могут быть выделены следующие **виды статистических показателей**: абсолютные, относительные, средние.

**Абсолютные показатели** – величины, которые характеризуют абсолютный размер (уровень) социально-экономического явления.

Виды абсолютных величин:

*1. По форме выражения выделяют:*

1.1. Натуральные абсолютные показатели – величины, предназначенные для характеристики физических свойств объекта (кг, м, км, граммы и т. д.):

- простые натуральные величины (кг, м, км);
- сложные натуральные величины (м/с, км/ч);
- условные натуральные величины (лошадиные силы).

1.2. Стоимостные абсолютные показатели – величины, предназначенные для характеристики стоимости (₽, \$, €).

1.3. Трудовые абсолютные показатели – величины, предназначенные для характеристики трудозатрат, трудоресурсов (человекодень, человеко-час).

*2. По уровню обобщения:*

2.1. Индивидуальные показатели – отражают характеристику конкретного элемента исследования (персональный доход студента 1 курса).

2.2. Групповые показатели – отражают итоговые, суммарные выражения величины характеристики группы (доход первокурсников).

2.3. Обобщающие показатели – характеризуют всю совокупность исследуемых элементов (доход студентов).

**Относительные показатели** – величины, которые отражают относительный размер явления (т. е. соотношение статистических показателей).

Виды относительных показателей:

*1. В зависимости от содержания:*

1.1. Относительные показатели динамики (ОПД) – это отношение показателя, достигнутого на данный период времени, к показателю за предшествующий период времени или к любому другому, взятыму за базу: ОПД = достигнутый уровень (текущий)/базисный.

1.2. Относительные показатели структуры (ОПСт) – это показатели соотношения размеров частей и целого: ОПСт = часть/целое.

1.3. Относительные показатели координации (ОПК) – это соотношение частей целого между собой: ОПК = часть 1/часть 2.

1.4. Относительные показатели сравнения (ОПСр) – это соотношение одноименных величин, характеризующих разные объекты или территории: ОПСр = отрасль (территория 1)/отрасль (территория 2).

1.5. Относительные показатели интенсивности (ОПИ) – это соотношение разноименных показателей, относящихся к одному объекту/территории: ОПИ = численность 1 (объект)/численность 2 (объект).

1.6. Относительные показатели плана (ОПП) – это отношение плана в текущий данный период времени к показателю, взятому за базу: ОПП = по плану в текущий период/базисный.

1.7. Относительный показатель выполнения плана (ОПВП) – это отношение фактически достигнутого плана к запланированному уровню: ОПВП = фактически достигнутый уровень плана/по плану.

2. В зависимости от того, что принимают за базу:

2.1. В виде кратного соотношения, доли ( $a > b$ , выражается в виде целого числа).

2.2. В процентах (база = 100, выражается в %).

2.3. В промиллях (база = 1 000, выражается в ‰).

2.4. В проценциимиллях (база = 10 000, выражается в 0 000).

**Средние показатели** – величины, которые дают характеристики средней тенденции в развитии явления, они могут быть рассчитаны только по количественному признаку. Все средние показатели делятся на два типа: структурные средние (мода, медиана) и степенные средние.

Виды степенных средних показателей:

1. Средняя арифметическая – это такое среднее значение признака, при получении которого сохраняется неизменным общий объем признака в совокупности.

2. Средняя гармоническая – эту среднюю называют обратной средней арифметической, поскольку эта величина используется при  $m = -1$ .

3. Средняя геометрическая – чаще всего находит свое применение при определении средних темпов роста (средних коэффициентов роста), когда индивидуальные значения признака представлены в виде относительных величин. Она используется также, если необходимо найти среднюю между минимальным и максимальным значениями признака (например, между 100 и 1 000 000).

4. Средняя квадратическая – основной сферой ее применения является измерение вариации признака в совокупности (расчет среднего квадратического отклонения).

## Тема 4. Методы описательной статистики

4.1. Меры центральной тенденции в описательной статистике.

4.2. Меры изменчивости в описательной статистике.

4.3. Меры положения и разброса в описательной статистике.

### 4.1. Меры центральной тенденции в описательной статистике

**Мера центральной тенденции** – это число, характеризующее выборку по уровню выраженности измеренного признака и характеристики совокупности переменных (признаков), указывающие на наиболее типичный, репрезентативный для изучаемой выборки результат. Не менее важной характеристикой является выраженность индивидуальных различий испытуемых по измеренному признаку.

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

**М**ода (*Mode*) – это такое значение из множества измерений, которое встречается наиболее часто. Мода, 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 иметь и не одну моду. Когда все значения встречаются одинаково часто, принято считать, что такое распределение не имеет моды.

*Бимодальное распределение* имеет на графике распределения две вершины, даже если частоты для двух вершин не строго равны. В последнем случае выделяют большую и меньшую моду. Во всей группе может быть и несколько локальных вершин распределения частот. Тогда выделяют *наибольшую моду и локальные моды*.

Если распределение имеет несколько мод, то говорят, что оно мультимодально или многомодально (имеет два или более «пика»). Мультимодальность распределения дает важную информацию о природе исследуемой переменной. Мультимодальность также служит индикатором того, что выборка не является однородной и наблюдения, возможно, порождены двумя или более «наложенными» распределениями.

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

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

Так, медиана разбивает выборку на две равные части и дает общее представление о том, где сосредоточены значения переменной, иными словами, где находится ее центр.

Медиана используется когда:

- 1) распределение асимметрично;
- 2) есть опасность перекоса из-за экстремальных значений. Медиана не чувствительна к экстремальным значениям, в то время как среднее очень чувствительно;
- 3) медиану можно вычислять для данных шкалы порядка и выше.

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

\*если данные содержат нечетное число значений (8, 9, 10, 13, 15), то медиана есть центральное значение, т. е.  $Md = 10$ ;

\*если данные содержат четное число значений (5, 8, 9, 11), то медиана есть точка, лежащая посередине между двумя центральными значениями, т. е.  $Md = (8+9)/2 = 8,5$ .

**С**реднее арифметическое – сумма всех значений измеренного признака, деленная на количество суммированных значений, а также центр распределения, вокруг которого группируются все варианты статистической совокупности; сумма всех членов совокупности, деленная на их общее число. Так, среднее – это сумма значений переменной, деленная на  $N$  (число значений переменной). Обычно обозначается  $\bar{x}$ .

**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). 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  $c$ ). Следует отметить, что среднее отклонение от среднего также равно 0.

При обработке исходных данных желательно использовать все меры центральной тенденции, поскольку использование одной из них может приводить к не совсем достоверным выводам.

1. В небольших выборках мода может быть совершенно нестабильной.
2. На медиану не влияют величины самых больших и самых малых значений элементов выборки (т.е. «выбросы»).
3. На величину среднего значения оказывает влияние каждый элемент выборки. Если какой-либо элемент выборки изменится на величину  $c$ , то среднее значение изменится в том же направлении на величину  $c - \bar{x}$ , где  $n$  – объем выборки.
4. Некоторые выборки вообще нельзя охарактеризовать с помощью мер центральной тенденции. Особенно это справедливо для выборок, имеющих более чем одну моду.
5. Если выборка является унимодальной, т.е. имеет одну моду, и гистограмма такой выборки является симметричной, то в этом случае мода, медиана и среднее значение совпадают.
6. Среднее значение представляет вариационный ряд, если одновременно выполняются следующие условия:
  - объем выборки достаточно велик;
  - гистограмма распределения симметрична;
  - отсутствуют выбросы (т.е. очень большие или очень малые значения). Если хотя бы одно из этих условий не выполняется, то при анализе данных следует ограничиться модой и медианой.
7. Мода и медиана являются более устойчивыми характеристиками выборки, чем среднее значение.

#### **4.2. Меры изменчивости в описательной статистике**

**М**ера изменчивости – численное выражение величины межиндивидуальной вариации признака и статистический показатель вариации (разброса) признака (переменной) относительно среднего значения, степени индивидуальных отклонений от центральной тенденции распределения. *Меры изменчивости* применяются для численного выражения величины межиндивидуальной вариации признака.

**C**тандартное (среднее квадратическое) отклонение ( $\sigma$ ,  $s_x$ ,  $S_x$ ) – это положительное значение квадратного корня из дисперсии. В том числе, рассматривается как квадратный корень из суммы квадратов отклонений индивидуальных значений признака от среднего, то есть дисперсии.

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.

**R**азмах – наиболее простая и очевидная мера изменчивости, указывающая на диапазон изменчивости значений. *Размах* – это просто разность максимального и минимального значений. Более устойчивыми являются разновидности размаха: *размах от 10 до 90-го процентиля* ( $P_{90} - P_{10}$ ) или *междуквартильный размах* ( $P_{75} - P_{25}$ ). Последние две меры изменчивости находят свое применение для описания вариации в порядковых данных. А для метрических данных используется дисперсия – величина, название которой в науке является синонимом изменчивости.

**D**исперсия (*Variance*) – мера изменчивости для метрических данных, пропорциональная сумме квадратов отклонений измеренных значений от их арифметического среднего:

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

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. Выборочное значение в статистике используется для оценки дисперсии в генеральной совокупности. Выше указана формула для генеральной (теоретической) дисперсии ( $A$ ), которая, понятно, не вычисляется. Для вычислений используется формула выборочной (эмпирической) дисперсии ( $D_x$ ), отличающаяся знаменателем:

#### **Свойства дисперсии:**

1. Если значения измеренного признака не отличаются друг от друга (равны между собой) – дисперсия равна нулю. Это соответствует отсутствию изменчивости в данных.
2. Прибавление одного и того же числа к каждому значению переменной не меняет дисперсию:

### **4.3. Меры положения и разброса в описательной статистике**

В описательной статистике 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. Одним из квантилей является 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 ( $P_u \dots, P_{99}$ ), 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%. Так, перцентили (центили, процентили) отделяют от совокупности по 0,01 части.

Процентиль какого-либо значения, таким образом, представляет собой процент случаев, которые имеют то же самое или меньшее значение.

Процентиль = (накопленная частота/ $N$ )\*100.

Процентиль всегда выражает положение значения по отношению к какой-либо выборке.

**Квартили** (*Quartiles*) – это 3 точки – значения признака, которые делят упорядоченное (по возрастанию) множество наблюдений на 4 равные по численности части.

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. Так, квартили – это значения, которые делят две половины выборки (разбитые медианой) еще раз пополам.

Таким образом, медиана и квартили делят диапазон значений переменной на четыре равные части. Первый quartиль соответствует 25-му процентилю, второй – 50-му процентилю или медиане, третий quartиль соответствует 75-му процентилю. Другими словами, существует:

– *Верхний квартиль* ( $Q_3$ ) делит пополам верхнюю часть выборки (значения переменной больше медианы). Верхний квартиль часто обозначают символом 75%, это означает, что 75% значений переменной меньше верхнего квартиля.

– *Нижний квартиль* ( $Q_1$ ) делит пополам нижнюю часть выборки (значения переменной меньше медианы). Нижний квартиль часто обозначают символом 25%, это означает, что 25% значений переменной меньше нижнего квартиля.

– Внутриквартильный (квартильный) размах =  $Q_3 - Q_1$ .

**Квантили** – структурные характеристики вариационного ряда, отсекающие в пределах ряда определенную часть его членов. Также **квантиль** рассматривается как точка на числовой оси, на которой откладываются результаты наблюдений. Эта точка делит всю совокупность наблюдений на части (группы) с определенными пропорциями между ними.

**Квинтили** делят значения наблюдений на 5 частей, их 4 ( $K_1, K_2, K_3, K_4$ ).

**Децили** делят совокупность на 10 частей, их 9 ( $D_1, \dots, D_9$ ) (рис. 1).

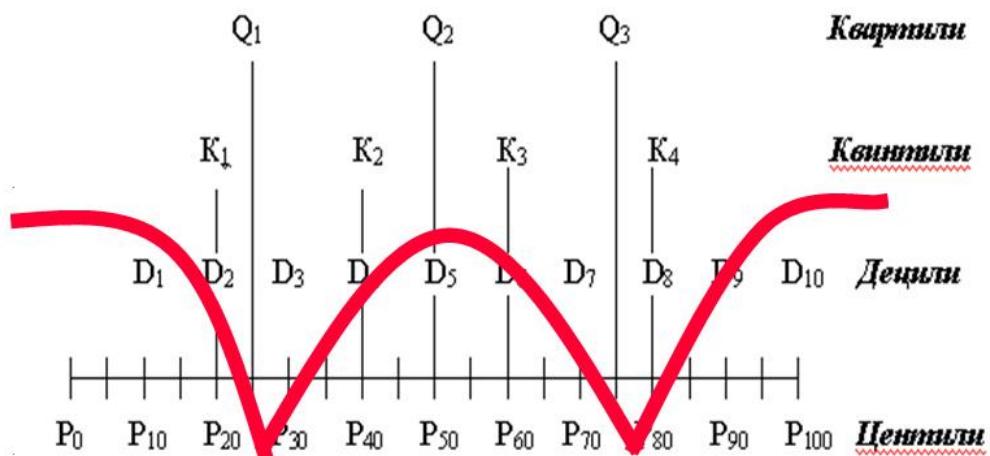


Рисунок 1 – Распределение мер положения в описательной статистике

## Тема 5. Генеральная совокупность в прикладной статистике

5.1. Генеральная совокупность и выборка.

5.2. Репрезентативная выборка.

5.3. Зависимые и независимые случайные выборки.

### 5.1. Генеральная совокупность и выборка

**Генеральная совокупность** – это совокупность всех мысленных наблюдений (или всех мыслимо возможных объектов), которые могут быть проведены при данном реальном комплексе условий и в отношении которых формулируется исследовательская гипотеза. В том числе генеральной совокупностью называют совокупность объектов, из которых производится выборка.

Понятие генеральной совокупности – это абстрактное математическое понятие и может быть конечной или бесконечной.

Выборка из данной генеральной совокупности представляет собой результат ограниченного ряда наблюдений интересующего нас показателя (признака, переменной). Генеральной совокупности всегда больше, чем выборка. В статистике выборка обозначается  $x_1, x_2, \dots, x_n$  количество наблюдений  $n$ . Количество наблюдений – « $n$ » называется объемом выборки.

Сущность статистических методов – чтобы по некоторой части генеральной совокупности, т.е. по выборке, выносить суждения о свойствах генеральной совокупности в целом.

**В**ыборочной совокупностью или просто *выборкой* называют совокупность случайно отобранных объектов или результаты ограниченного ряда наблюдений. Выборку рассматривают как некий эмпирический аналог генеральной совокупности, с которым чаще всего на практике имеют дело, поскольку обследование всей генеральной совокупности бывает либо слишком трудоемко, либо невозможно. Иными словами, выборка – это ограниченная по численности группа объектов (испытуемых, респондентов), специально отбираемая из генеральной совокупности для изучения ее свойств. Соответственно, изучение на выборке свойств генеральной совокупности называется *выборочным исследованием*. Практически все исследования являются выборочными, а их выводы распространяются на генеральные совокупности.

В зависимости от количества наблюдений выборочные совокупности подразделяются на три группы: малые – до 30 наблюдений; средние – от 30 до 200 наблюдений; большие – от 200 наблюдений и выше.

Объемом совокупности (выборочной или генеральной) называют число наблюдений этой совокупности.

*Распределением* признака (переменной) называют закономерность встречаемости различных его значений. *Параметры распределения* (описательные статистики) – это его числовые характеристики, описывающие распределение. Они указывают, где «в среднем» располагаются значения признака, насколько эти значения изменчивы и наблюдается ли преимущественное появление определенных значений признака. В реальных исследованиях оперируем не параметрами, а их приближенными значениями, так называемыми оценками параметров. Это объясняется ограниченностью обследованных выборок. Чем больше выборка, тем ближе могут быть оценки параметра к его истинному значению. Когда речь идет о параметрах, имеется в виду их оценки. Параметры определяют то, что называется средней тенденцией, разброс значений и форму распределения.

Таким образом, генеральная совокупность – это хотя и не бесконечное по численности, но, как правило, недоступное для сплошного исследования множество потенциальных испытуемых. В свою очередь, выборка – ограниченная по численности группа объектов, специально отбираемая из генеральной совокупности для изучения ее свойств.

Таким образом, после того, как сформулирована гипотеза и определены соответствующие генеральные совокупности, перед исследователем возникает проблема организации выборки. Выборка должна быть такой, чтобы была обоснована генерализация выводов выборочного исследования – обобщение, распространение их на генеральную совокупность. Основные критерии обоснованности выводов исследования – это репрезентативность выборки и статистическая достоверность (эмпирических) результатов.

## 5.2. Репрезентативная выборка

**Р**епрезентативность выборки (ее представительность) – это способность выборки представлять изучаемые явления достаточно полно – с точки зрения их изменчивости в генеральной совокупности; соответствие свойств

исследуемой выборки свойствам генеральной совокупности. Репрезентативной является выборка, элементы которой правильно представляют пропорции генеральной совокупности. В этом случае выборка будет представлять изучаемое явление достаточно полно с точки зрения его изменчивости в генеральной совокупности.

Необходимо учитывать, что для большинства переменных в генеральной совокупности значения распределяются следующим образом:

- большинство значений отражает средний уровень выраженности переменной;
- меньшинство значений отражает низкий или ниже среднего уровень выраженности переменной;
- меньшинство значений отражает высокий или выше среднего уровень выраженности переменной.

Полное представление об изучаемом явлении, во всем его диапазоне и нюансах изменчивости, может дать только генеральная совокупность. Поэтому репрезентативность всегда ограничена в той мере, в какой ограничена выборка. И именно репрезентативность выборки является основным критерием при определении границ генерализации выводов исследования.

При отборе объектов из генеральной совокупности для получения репрезентативности выборки используются следующие способы:

1. Простой случайный отбор (рандомизированный). 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 при отборе. И это позволяет надеяться на то, что в конечном итоге в выборке изучаемое свойство будет представлено если и не во всем, то в максимально возможном его многообразии.

Объекты генеральной совокупности, имеющей объем  $N$ , нумеруют от 1 до  $N$ . Затем, используя таблицу случайных чисел или процедуру жеребьевки (например, корзину с пронумерованными карточками), отбирают  $n$  объектов выборки.

2. Простой отбор с помощью регулярной, но не существенной для изучаемого явления процедуры (например, отбор испытуемых по их номеру в списке).

3. Стратифицированный (расслоенный). Он предполагает предварительное определение тех качеств, которые могут влиять на изменчивость изучаемого свойства (это может быть пол, уровень дохода или образования и т. д.). Затем определяется процентное соотношение численности различающихся по этих качествам групп (страт) в генеральной совокупности и обеспечивается идентичное процентное соотношение соответствующих групп в выборке. Далее в каждую подгруппу выборки испытуемые подбираются по принципу простого случайного отбора. В этом случае генеральная совокупность объема  $N$  разделяется на непрессекающиеся подсовокупности (страты, слои)  $N_1, N_2, \dots, N_k$ . Из каждого слоя извлекается простая случайная выборка, имеющая объем  $n_1, n_2, \dots, n_k$  соответственно, причем  $n_1 + n_2 + \dots + n_k = n$ . Стратифицированный отбор применяется, когда объекты внутри каждого слоя являются однородными по изучаемому свойству.

4. Серийный (гнездовой) отбор. Он применяется, если удобно исследовать не отдельные элементы генеральной совокупности, а целые блоки или серии таких элементов. Например, исследуются все ученики одного класса или все семьи в одном доме.

5. Комбинированный (ступенчатый). Объединяет несколько вышеперечисленных способов отбора, которые составляют различные ступени выборочного исследования.

Существуют определенные требования к численности, или *объему выборки*.

К сожалению, строгих рекомендаций по предварительному определению требуемого объема выборки не существует. Более того, ответ на вопрос о необходимой

и достаточной ее численности исследователь обычно получает слишком поздно – только после анализа данных уже обследованной выборки. Тем не менее, можно сформулировать наиболее общие рекомендации:

1. Наибольший объем выборки необходим при разработке диагностической методики – от 200 до 1000 – 2500 человек.
2. Если необходимо сравнивать 2 выборки, их общая численность должна быть не менее 50 человек; численность сравниваемых выборок должна быть приблизительно одинаковой.
3. Если изучается взаимосвязь между какими-либо свойствами, то объем выборки должен быть не меньше 30 – 35 человек.
4. 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. При этом, естественно, уменьшаются возможности генерализации выводов.

### **5.3. Зависимые и независимые случайные 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. Напротив, **зависимые выборки** характеризуются тем, что каждому испытуемому одной выборки поставлен в соответствие по определенному критерию испытуемый из другой выборки.

В общем случае зависимые выборки предполагают попарный подбор испытуемых в сравниваемые выборки, а независимые выборки – независимый отбор испытуемых.

Следует отметить, что случаи «частично зависимых» (или «частично независимых») выборок недопустимы: это непредсказуемым образом нарушает их представительность.

Выделяют две парадигмы исследования. Так называемая **R-методология** предполагает изучение изменчивости некоторого свойства под влиянием некоторого воздействия, фактора либо другого свойства. Выборкой является множество испытуемых. Другой подход, **Q-методология**, предполагает исследование изменчивости субъекта (единичного) под влиянием различных стимулов (условий, ситуаций и т. д.). Ей соответствует ситуация, когда выборкой является множество стимулов.

Таким образом, независимые выборки:

1. Выборки, которые характеризуются тем, что вероятность отбора любого испытуемого одной выборки не зависит от отбора любого из испытуемых другой выборки.
2. Допущение независимости предполагает, что представители двух выборок не составляют пары коррелирующих значений признака.

Зависимые выборки.

1. Выборки, характеризующиеся тем, что каждому испытуемому одной выборки поставлен в соответствие по определенному критерию испытуемый из другой выборки.
2. Допущение зависимости чаще всего означает, что признак измерен на одной и той же выборке дважды, например, до воздействия и после него.

## **Тема 6. Понятие нормального распределения в прикладной статистике**

- 6.1. Понятие нормального распределения.*
- 6.2. Свойства нормального распределения.*
- 6.3. Проверка нормальности распределения.*

### ***6.1. Понятие нормального распределения***

**Нормальный закон распределения** играет важнейшую роль в применении численных методов. Он лежит в основе измерений, разработки тестовых шкал, методов проверки гипотез.

Начиная со второй половины XIX столетия измерительные и вычислительные методы разрабатываются на основе следующего принципа: 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, которое используется для приблизительного описания многих случайных явлений, в которых на интересующий нас признак оказывает влияние большое количество независимых случайных величин, среди которых нет сильно выделяющихся.

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

Необходимо отметить, что, если при изучении некоторого свойства произведено 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.

Нормальное распределение:  $f(x) = (1/\sqrt{2\pi}) \exp\{-(x-\mu)^2/2\sigma^2\}$ ,

где среднее значение  $\mu$ ; дисперсия  $\sigma^2$ ; асимметрия  $A = 0$ ; эксцесс  $E = 3$ .

Так, нормальное распределение – вероятность отклонения любой варианты  $x_i$  от центра распределения  $\mu$ , где  $x_i - \mu = 0$ , определяется функцией нормированного отклонения  $t$ . Графически эта функция выражается в виде кривой вероятности, называемой нормальной кривой. Форма и положение этой кривой определяются только двумя параметрами:  $\mu$  и  $\sigma$ . При изменении величины  $\mu$  форма нормальной кривой не меняется, лишь график ее смещается вправо или влево. Изменение же величины  $\sigma$  влечет за собой изменение только ширины кривой: при уменьшении  $\sigma$  кривая делается более узкой за счет меньшего рассеяния вариант вокруг средней, а при увеличении  $\sigma$  кривая расширяется. Симметричное распределение, у которого крайние значения встречаются редко и частота постепенно повышается от крайних к средним значениям признака. Стандартное нормальное распределение имеет нулевое среднее и единичную дисперсию.

## 6.2. Свойства нормального распределения

**С**войства нормального распределения:

- Единицей измерения единичного нормального распределения является стандартное отклонение.
- Кривая приближается к оси  $Z$  по краям асимптотически – никогда не касаясь ее.
- Кривая симметрична относительно  $A=0$ . Ее асимметрия и эксцесс равны нулю.
- Кривая имеет характерный изгиб.
- Площадь между кривой и осью  $Z$  равна 1 и объясняет название *единичное* нормальное распределение и имеет исключительно важное значение. Благодаря этому свойству *площадь под кривой интерпретируется как вероятность, или относительная частота*. Действительно, 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. Следует отметить, что любое нормальное распределение может быть сведено к единичному нормальному распределению путем преобразования.

– Для нормального распределения характерно совпадение моды, медианы и среднего значения.

– Форма и положение графика нормального распределения определяются средним значением и стандартным отклонением.

При этом: а) если  $\sigma$  постоянно, а изменяется  $x$ , то форма кривой неизменна, а график смещается по оси абсцисс вправо (при увеличении  $x$ ) или влево (при уменьшении  $x$ );

б) если  $x$  постоянно, то изменение  $\sigma$  влечет изменение ширины кривой: при уменьшении  $\sigma$  кривая делается более узкой и поднимается вверх, а при увеличении  $\sigma$  кривая расширяется и опускается вниз.

Однако во всех случаях нормальная кривая строго симметрична относительно среднего значения и сохраняет правильную колоколообразную форму.

– Чем больше величина признака отклоняется от среднего значения, тем меньше частота его встречаемости (или вероятность – для генеральной совокупности) в распределении;

Таким образом:

\*если  $x$ , имеет нормальное распределение со средним  $M$  и стандартным отклонением  $\sigma$ , 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  $\sigma$  равна площади между  $z_1 = (x_1 - M_x)/\sigma$  и  $z_2 = (x_2 - M_x)/\sigma$  в единичном нормальном распределении.

Итак, 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.

Важно 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 четкую взаимосвязь стандартного отклонения и относительного количества случаев в генеральной совокупности для любого нормального распределения.

Исходя из мнения, в соответствии с которым свойства в генеральной совокупности имеют нормальное распределение, реальные данные, полученные на выборке, нечасто распределены нормально. Более того, разработано множество методов, позволяющих анализировать данные без всякого предположения о характере их распределения как в выборке, так и в генеральной совокупности. Эти обстоятельства иногда приводят к ложному убеждению, что нормальное распределение – пустая математическая абстракция, не имеющая отношения к психологии. Тем не менее, можно указать по крайней мере на три важных аспекта применения нормального распределения:

1. Разработка тестовых шкал.
2. Проверка нормальности выборочного распределения для принятия решения о том, в какой шкале измерен признак – в метрической или порядковой.
3. Статистическая проверка гипотез, в частности – при определении риска принятия неверного решения.

### **6.3. 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. А сомнения такие возникают очень часто, так как заранее нам, как правило, не известно, в какой шкале удастся измерить изучаемое свойство (исключая, конечно, случаи явно номинативного измерения).

Важность определения того, в какой шкале измерен признак, трудно переоценить, по крайней мере, по двум причинам. От этого зависит, *во-первых*, полнота учета 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 ведет 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).

Существует множество различных способов проверки нормальности.

**Графический способ (*Q-Q Plots, P-P Plots*).** Странят 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-процентилю. Затем по таблице нормального распределения для каждого из этих процентиелей 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 значений, которые откладываются на оси ординат. Если распределение соответствует нормальному, полученные на графике точки лежат на одной прямой.

**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.

Выборочные значения асимметрии и эксцесса значительно отличаются от нуля, если не превышают значения своих стандартных ошибок. Это можно считать признаком соответствия выборочного распределенияциальному закону. Следует отметить, что компьютерные программы вычисляют показатели асимметрии, эксцесса и соответствующие им стандартные ошибки по другим, более сложным формулам.

**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  $p > 0,05$ , то данное эмпирическое распределение существенно отличается от нормального, а если  $p < 0,05$ , то делают вывод о приблизительном соответствии данного эмпирического распределения нормальному.

**Критерий  $\lambda$  Колмогорова – Смирнова:** критерий, определяющий, отличается ли данное эмпирическое распределение от теоретического распределения (нормального). Здесь первый раз возникла необходимость расчета значения статистического критерия и уровня статистической достоверности (значимости).

## **Тема 7. Понятие статистической гипотезы. Процедура проверки гипотезы**

- 7.1. Гипотеза и ее характеристики.*
- 7.2. Проверка статистических гипотез.*
- 7.3. Уровень значимости и уровень достоверности.*
- 7.4. Виды ошибок, возможные при проверке статистических гипотез.*

### **7.1. Гипотеза и ее характеристики**

**О**дним из основных понятий, используемых в научных исследованиях, является **гипотеза** (от греч. ὑπόθεσις – основание, предположение), которая выдвигается для того, чтобы дать предварительное, но вероятностное объяснение реальному обстоятельству, которое более или менее исследовано.

Выделяют следующие виды гипотез: теоретические и статистические.

Теоретическая (научная) гипотеза – это научно обоснованное высказывание вероятностного характера относительно сущности, взаимосвязей и причин явлений объективной действительности.

Признаки правильно сформулированной научной гипотезы были предложены П. Фрессом:

- гипотеза должна быть адекватной поставленной проблеме;
- гипотеза должна учитывать уже имеющиеся в науке знания по исследуемой проблеме;
- гипотеза должна быть доступной для проверки (верифицируемой).

Различают два основных вида теоретических гипотез:

- a) описательные – направлены на изучение структуры исследуемого объекта (характера связей в нем, типичного набора факторов) и его функций (предположение о тесноте связей);
- б) объяснительные – представляют собой предположения о причинно-следственных связях в изучаемом объекте.

После того как научная гипотеза сформулирована, проводится поиск показателей (понятий, входящих в ее состав), которые могут быть измерены в ходе эмпирического исследования, а затем обработаны с помощью методов математической статистики и проинтерпретированы. Поскольку в психологических исследованиях мы имеем дело с данными, в которых интересующие исследователя закономерности искажены различными случайными факторами, большинство статистических выводов сопровождается проверкой некоторых гипотез об источнике этих данных.

Статистическая гипотеза – это предположение о распределении вероятностей признака, которое мы хотим проверить по имеющимся данным.

Выделяют нулевую (основную) и альтернативную (конкурирующую) статистические гипотезы.

**Н**улевая гипотеза ( $H_0$ ) – это основное проверяемое предположение об отсутствии статистического эффекта. Оно формулируется как отсутствие различий, отсутствие связи, отсутствие влияния фактора, равенство нулю значений выборочных характеристик и т.п. Например, в случае установления наличия (отсутствия) различий, нулевая гипотеза говорит о том, что исследуемые выборки принадлежат к одной совокупности и различия между ними признаны случайными (недостоверными).

**А**льтернативная гипотеза ( $H_1$ ) – это рабочая гипотеза исследования, предположение о наличии статистического эффекта. Она говорит о том, что между выборками существуют достоверные различия (взаимосвязь, влияние фактора и т.д.), которые обусловлены влиянием независимой переменной.

Альтернативная гипотеза – это предположение, которое хочет доказать исследователь, поэтому ее иногда называют экспериментальной гипотезой.

Нулевая и альтернативная гипотезы могут быть **направленными и ненаправленными**. Ненаправленные гипотезы говорят только о наличии или отсутствии различий (связи и т.д.). Направленные гипотезы говорят не только о наличии или отсутствии различий (связи и т.д.), но и указывают их направление.

В статистических гипотезах речь идет не об арифметических (т.е. числовых), а о статистически значимых различиях, т.е. с одинаковой ли частотой встречаются различные значения признака в обоих эмпирических распределениях. Нулевая гипотеза всегда проверяется по отношению к альтернативной, поскольку они образуют полную группу несовместных событий: если верна одна из них, то другая является ложной, и наоборот, отклонение одной из них неизбежно влечет принятие другой. Поэтому альтернативная гипотеза  $H_1$  принимается (а не доказывается), если с помощью рассуждений мы можем отвергнуть нулевую гипотезу  $H_0$ , т.е. альтернативная гипотеза подтверждается не прямо, а косвенно.

## **7.2. Проверка статистических гипотез**

**О**бщая схема проверки статистических гипотез состоит из следующих этапов:

- 1 этап – выдвигаются две статистические гипотезы:
  - 1) основная нулевая  $H_0$  и 2) альтернативная (конкурирующая)  $H_1$ .

2 этап – задаемся уровнем значимости. Статистический вывод никогда не может быть сделан со стопроцентной уверенностью. Всегда допускается риск принятия неправильного решения. При проверке статистических гипотез мерой такого риска и выступает уровень значимости, который представляет собой долю и процент ошибок, которые можно позволить при статистических выводах.

3 этап – по исходным данным, то есть по выборке вычисляется наблюдаемое значение статистики критерия. Для этого используются статистические таблицы. Выбор необходимой статистической таблицы осуществляется в зависимости от распределения статистики критерия. При проверке статистических гипотез статистика критерия выбирается (статистиками).

4 этап – путем сравнения найденных наблюдаемых критических значений делаем вывод о правильности этой или иной гипотезы.

Существует еще вариант принятия решения об отклонении (принятия) статистической гипотезы, который разбивается на семь этапов:

- 1) формулировка нулевой и альтернативной гипотез;

- 2) выбор соответствующего уровня значимости (как правило,  $p = 0,05$ );

- 3) выбор статистического критерия, который определяется типом решаемой психологической задачи;

- 4) вычисление по экспериментальным данным эмпирического значения вы бранного статистического критерия (ЭМП);

- 5) нахождение по статистическим таблицам критических значений (КР) для выбранного статистического метода, соответствующих уровню значимости  $p = 0,05$  и  $p = 0,01$  при заданном объеме выборки (числе степеней свободы);

- 6) сравнение вычисленного эмпирического и найденного критического значений;

- 7) выбор соответствующей статистической гипотезы.

При выборе гипотезы используется правило отклонения  $H_0$  (принятия  $H_1$ ):

- если эмпирическое значение критерия равняется критическому значению или превышает его ( $\text{ЭМП} \geq \text{КР}$ ) при уровне значимости  $p \leq 0,05$ , то  $H_0$  отклоняется;

- если эмпирическое значение критерия равняется критическому значению или превышает его ( $\text{ЭМП} \geq \text{КР}$ ) при уровне значимости  $p \leq 0,01$ , то  $H_0$  отклоняется и принимается  $H_1$ . Следует помнить, что у этого правила есть исключения, к которым относятся G-критерий знаков, Т-критерий Вилкоксона и U-критерий Манна – Уитни. При их проверке устанавливаются обратные соотношения между эмпирическим и критическим значениями. Для облегчения процесса принятия решения об отвержении (принятии) гипотез можно использовать ось значимости.

## **7.3. Уровень значимости и уровень достоверности**

**Уровень значимости** – это вероятность того, что мы сочли различия существенными, а они на самом деле случайны. Уровень статистической значимости  $p$  представляет собой, таким образом, вероятность неправильного отверждения нуль-гипотезы.

Уровень статистической значимости, установленный исследователями для заключения о действии неслучайных факторов часто называется уровнем  $\alpha$  (в более новых книгах он обычно обозначается латинской буквой  $p$ ). Чтобы определить, стоит ли

объяснять какое-либо явление действием некоторого неслучайного фактора, надо найти вероятность того, что это явление произойдет случайно и сравнить с выбранным уровнем статистической значимости. Следует отметить, что приемлемый уровень статистической значимости должен быть определен до проведения исследования.

В социальных науках исследователи согласились, что следующие два значения будут основанием для допущения действия неслучайного фактора:

1) Если некоторое событие происходит случайно в 5% случаев или еще реже, то предполагается, что это происходит благодаря действию некоторых неслучайных факторов. Это значение называется 5%-м уровнем статистической значимости или уровнем статистической значимости, равным 0,05. Когда указывается, что различия достоверны на 5%-ом уровне значимости, или при  $p<0,05$ , то имеется в виду, что вероятность того, что они все-таки недостоверны, составляет 0,05.

2) Если некоторое событие происходит случайно в 1% случаев или еще реже, то предполагается, что это происходит благодаря действию некоторых неслучайных факторов. Это значение называется 1%-м уровнем статистической значимости или уровнем статистической значимости, равным 0,01. Когда указывается, что различия достоверны на 1 %-ом уровне значимости, или при  $p<0,01$ , то мы имеем в виду, что вероятность того, что они все-таки недостоверны, составляет 0,01.

Если перевести все это на более формализованный язык, то уровень значимости – это вероятность отклонения нулевой гипотезы, в то время как она верна. Так, в отношении научной гипотезы уровень статистической значимости – это количественный показатель степени недоверия к выводу о наличии связи, вычисленный по результатам выборочной, эмпирической проверки этой гипотезы. Чем меньше значение р-уровня, тем выше статистическая значимость результата исследования, подтверждающего научную гипотезу.

Уровень значимости при прочих равных условиях выше (значение р-уровня меньше), если:

- величина связи (различия) больше;
- изменчивость признака (признаков) меньше;
- объем выборки (выборок) больше.

#### **7.4. Виды ошибок, возможные при проверке статистических гипотез**

Ошибки, допускаемые при проверке гипотез, удобно разделить на два типа:

- 1) отклонение гипотезы  $H_0$ , когда она верна, – ошибка первого рода;
- 2) принятие гипотезы  $H_0$ , когда в действительности верна какая-то другая гипотеза, – ошибка второго рода.

Вероятность ошибки первого рода обозначается  $\alpha$ . Величина  $\alpha$  называется уровнем значимости критерия, по которому проверяется справедливость гипотезы  $H_0$ .

Вероятность ошибки второго рода обозначается  $\beta$ . Ее величина зависит от альтернативной гипотезы  $H_1$ .

Наглядно все варианты принятия и отвержения статистической гипотезы  $H_0$  и их вероятность можно представить в виде табл. 2.

Таблица 2 – Варианты принимаемых решений при проверке статистических гипотез

Решение	Нуль-гипотеза – верна	Альтернативная гипотеза – верна
Отвержение нуль-гипотезы	Ошибка I рода (вероятность равна $\alpha$ )	Нет ошибки Правильное решение (вероятность равна $1 - \beta$ )
Принятие нуль-гипотезы	Нет ошибки Правильное решение (вероятность равна $1 - \alpha$ )	Ошибка II рода (вероятность равна $\beta$ )

Ошибка, состоящая в том, что отклонена нуль-гипотеза, в то время как она верна, называется **ошибкой I рода**. Вероятность такой ошибки обозначается  $\alpha$  (или  $p$ ).

Ошибка, состоящая в том, что принята нуль-гипотеза, в то время как она неверна, называется **ошибкой II рода**. Вероятность такой ошибки обозначается  $\beta$ .

Следует помнить, что критерии различаются по мощности. Мощность критерия – это его способность не допустить ошибку II рода. Поэтому мощность =  $1 - \beta$ . Мощность критерия определяется эмпирическим путем.

Основной проблемой статистического вывода является установление оптимальной величины вероятности  $\alpha$ , удовлетворяющей двум противоречивым требованиям:

1. величина  $\alpha$  должна быть достаточно мала, т.к. чем меньше  $\alpha$ , тем больше вероятность правильного решения;
2. величина  $\alpha$  должна быть достаточно велика, чтобы не допустить ошибки второго рода, т.к. при увеличении значения  $\alpha$  вероятность  $\beta$  уменьшается.

В ряде случаев используются так называемые стандартные (пороговые или критические) величины ошибки первого рода  $\alpha = 0,1; 0,05; 0,01; 0,001$ . Традиционно при интерпретации различных уровней значимости, полученных на основе выборочных показаний об ошибочности статистической гипотезы ( $p$  – уровень значимости), исходят из величины  $\alpha = 0,05$ , реже  $\alpha = 0,01$  (табл. 3).

Таблица 3 – Интерпретация уровней значимости

Уровень значимости	Решение относительно нулевой гипотезы	Возможный статистический вывод
$p > 0,1$	принимается $H_0$	результаты статистически незначимы
$0,05 \leq p \leq 0,01$	сомнение в истинности $H_0$	результаты значимы на уровне тенденции
$p \leq 0,05$	значимость, отклонение $H_0$	результаты статистически значимы
$p \leq 0,01$	высокая значимость, отклонение $H_0$	результаты на уровне высокой статистической значимости

$p$ -уровень значимости позволяет установить, с какой вероятностью (или в каком проценте случаев) все же возможна ошибка в тех выводах, которые делаются в оценке достоверности показателей или различий между какими-то величинами, полученными эмпирическим путем.

Когда говорят, что различия (или взаимосвязь) достоверны при  $p \leq 0,05$  (или на 5%-ном уровне значимости), это означает, что имеется 5%-ная вероятность того, что они все-таки недостоверны, т.е. установленные различия (взаимосвязь) являются случайными, характерными только для данной выборки. Аналогично, если различия достоверны при  $p \leq 0,01$  (или на 1%-ном уровне значимости), то имеется 1%-ная веро-

ятность их недостоверности, т.е. различия являются случайными, присущими только исследуемой выборке.

В современных статистических пакетах уровни значимости вычисляются непосредственно в процессе обработки эмпирических данных. Эти уровни, обозначаемые в разных пакетах различной аббревиатурой, могут иметь разное числовое выражение в интервале от 0 до 1.

## Тема 8. Способы представления статистических данных

8.1. Графические способы представления статистических данных.

8.2. Табличный способ представления статистических данных.

### 8.1. Графические способы представления статистических данных

Графическая представленность полученных результатов возможна посредством, как минимум, гистограмм, диаграмм, графиков, полигонов частот.

**Гистограмма** – это тип диаграммы, в которой каждая точка данных выводится в виде отдельного вертикального столбика, высота которого отображает значение. Шкала значений выводится на вертикальной оси, располагаемой обычно на левой стороне диаграммы.

Гистограмма в статистике называется изображения статистических данных в виде различных геометрических образов: точек, линий, фигур и т.п. Главное достоинство – наглядность, ознакомление масс со статистическими данными; позволяют оживить таблицу, делая ее более доступной.

Гистограммы широко используются для обобщения и анализа статистических данных, легче уяснить закономерности развития, распределения и размещения явлений.

В гистограммах используются контрольные значения, которые позволяют выявлять различного рода ошибки и неточности и являются контролером точности расчётов и вычислений.

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

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 ломаной линией.

Вместо гистограммы или полигона часто изображают сглаженную кривую распределения частот.

**Круговая диаграмма.** Круг берется за 100% и разделяется на число секторов, равное числу групп, и угловая величина каждого сектора берется пропорциональной частоты группы.

Графики распределения частот дают важную предварительную информацию о *форме распределения признака*: о том, какие значения встречаются реже, а какие чаще, насколько выражена изменчивость признака. Обычно 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 серединным значениям признака. *Асимметричные распределения* – *левосторонние* (с преобладанием частот малых значений), *правосторонние* (с преобладанием частот больших значений).

В том числе, для графического представления статистических данных используются самые разнообразные виды графиков (рис. 2–3). Их можно классифицировать по разным признакам: характеру графического образа, способу построения и назначению (содержанию). По характеру графического образа различают графики объемные, линейные и плоскостные (рис. 2). По способу построения графики можно разделить на диаграммы и статистические карты (рис. 3).

Диаграмма представляет собой чертеж, показывающий соотношение статистических данных при помощи разнообразных геометрических и изобразительных средств.

Статистические карты предназначены для графического изображения одноименных показателей, относящихся к разным территориям. Для этого в основу изображения берется географическая карта. Изображение на карте статистических данных называется картограммой или картодиаграммой.

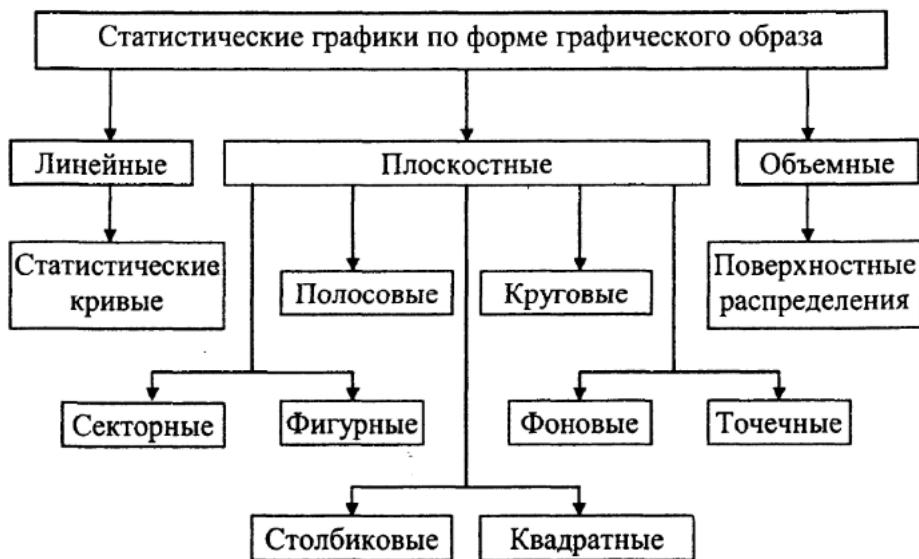


Рисунок 2 – Классификация статистических графиков по форме графического образа

По содержанию или назначению можно выделить графики сравнения в пространстве, графики относительных величин (структуры, динамики и т.п.), графики вариационных рядов, графики взаимосвязанных показателей и графики размещения по территории.



Рисунок 3 – Классификация статистических графиков по способу построения и содержанию изображаемых данных

Несмотря на большое разнообразие статистических графиков, существуют общие правила их построения. При построении графика важно найти такие способы изображения, которые наилучшим образом отвечают содержанию и логической природе изображаемых показателей.

Каждый график состоит из графического образа и вспомогательных элементов.

Графический образ (основа графика) – это геометрические знаки, то есть совокупность точек, линий, фигур, с помощью которых изображаются статистические показатели. Важно правильно выбрать графический образ, который должен соответствовать цели графика и способствовать выразительности изображаемых статистических данных.

Вспомогательные элементы делают возможным чтение графика, его понимание и использование. К ним относятся: 1) экспликация графика; 2) пространственные ориентиры; 3) масштабные ориентиры; 4) поле графика.

## 8.2. Табличный способ представления статистических данных

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 (табл. 4).

Таблица 4 – Пример таблицы распределения частот

Значение	Абсолютная частота	Относительная частота	Накопленная частота
5	3	0,05	1,00
4	12	0,20	0,95
3	21	0,35	0,75
2	15	0,25	0,40
1	9	0,15	0,15
(сумма)	60	1	—

Нередко относительная частота применяется для оценки вероятности встречаемости значения. Во многих случаях признак может принимать множество различных значений и тогда о распределении признака позволяет судить *таблица сгруппированных частот*, в которых частоты группируются по разрядам или интервалам значений признака.

Еще 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).

Сами по себе таблицы распределения признака позволяют делать некоторые содержательные выводы при сравнении групп испытуемых между собой. Сравнивая распределения, можно 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 накопленных частот позволяют быстро получить дополнительную информацию о том, сколько испытуемых (или какая их доля) имеют выраженность признака не выше определенного значения. Следует отметить, что для сравнения групп разной численности следует использовать таблицы и графики относительных частот.

Табличным способом представления статистических данных являются **таблицы сопряженности**, или кросstabуляции – 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. Если номинативные признаки внесены в электронную таблицу исходных данных, то таблицу сопряженности можно построить, воспользовавшись функцией «Кросstabуляция» одного из стандартных статистических пакетов (например, Crosstabs – в SPSS). Таблицы сопряженности могут включать номинативные признаки, имеющие и более двух градаций.

## **Тема 9. Показатели корреляционной связи в статистическом ряду**

*9.1. Общее понятие о корреляционном анализе и коэффициенте корреляции.*

*9.2. Характеристики коэффициентов корреляции.*

*9.3. Содержательный анализ корреляционных матриц.*

***9.1. Общее понятие о корреляционном анализе и коэффициенте корреляции***

**Корреляция** (от лат. *correlatio* – соотношение) – это мера степени и направления связи между значениями двух переменных. В том числе – это статистический показатель вероятности связи между двумя переменными, измеренными в количественной шкале. Корреляционный анализ позволяет проверить гипотезу о связях между переменными с использованием коэффициентов корреляции.

При определении понятия «корреляция» используют термины «корреляционная связь» и «корреляционная зависимость», «взаимосвязь». При этом в математической статистике сложилось два под хода. В соответствии с первым подходом корреляционная связь рассматривается как согласованные изменения двух или нескольких признаков. Это означает, что изменчивость одного признака находится в некотором соответствии с изменчивостью другого, которое может быть обусловлено множеством разных причин, т.е. эти два признака зависят от некоторого третьего признака или сочетания признаков, не рассматриваемых в данном исследовании.

При изучении корреляции стараются установить, существует ли связь между двумя признаками, измеренными на одной выборке (например, уровнем интеллекта и успеваемостью школьников), или между признаками, измеренными на разных выборках (например, между уровнем интеллекта детей и образованием родителей). Если эта связь существует, то стремятся выяснить, как изменение одного признака связано с изменением другого. Если при увеличении (уменьшении) одного признака X увеличивается (уменьшается) другой признак Y, то говорят о положительной, или прямой, корреляции ( $X \uparrow \uparrow Y$  или  $X \downarrow \downarrow Y$ ). Если при увеличении одного признака X уменьшается другой признак Y, то говорят об отрицательной или обратной корреляции ( $X \uparrow \downarrow Y$  или  $X \downarrow \uparrow Y$ ).

В случае качественных признаков положительная корреляция обозначает, что присутствие одного признака совпадает с присутствием другого, а отрицательная корреляция – что присутствие одного признака совпадает с отсутствием другого. Степень взаимосвязи между признаками и ее направление характеризует коэффициент корреляции, который обозначают  $r$ . Так, *коэффициент корреляции* – двумерная описательная статистика, количественная мера взаимосвязи (совместной изменчивости) двух переменных. *Другими словами* – это показатель степени связи между двумя переменными или измерениями.

Коэффициент корреляции – это величина, которая может варьироваться в пределах от  $-1$  до  $+1$ , т.е.  $-1 \leq r \leq +1$ . Корреляционные связи можно классифицировать по различным основаниям (по направлению, по форме, по силе).

*По направлению выделяют:*

– положительную корреляцию, если коэффициент корреляции имеет положительный знак:  $0 < r \leq +1$ ;

– отрицательную корреляцию, если коэффициент корреляции имеет отрицательный знак:  $-1 \leq r < 0$ .

Так, коэффициент корреляции обычно изменяется от  $-1$  до  $+1$  и показывает направление отношений (“+” – прямая зависимость, “–” – обратная).

*Корреляционные связи классифицируют также по форме.*

Выделяют два вида связей:

– прямолинейная корреляционная связь, например связь между количеством тренировок на тренажере и количеством правильно решенных задач при контрольном тестировании. В этом случае графически связь можно представить в виде прямой линии:

– криволинейная корреляционная связь, например связь между уровнем мотивации и эффективностью выполнения задачи (закон Йеркса-Додсона). В этом случае связь можно представить графически в виде кривой, которая вначале возрастает (до оптимального значения мотивации), а затем начинает убывать. Это означает, что при повышении мотивации эффективность выполнения задачи сначала возрастает, затем достигается максимальная эффективность выполнения задачи, соответствующая оптимальному значению мотивации, после чего дальнейшему повышению мотивации сопутствует снижение эффективности

Корреляционные связи классифицируют по силе (степени или тесноте). При этом используют две системы классификаций – общую и частную. Общая классификация корреляционных связей по силе:

- 1)  $r < 0,19$  – очень слабая связь;
- 2)  $0,20 < r < 0,29$  – слабая;
- 3)  $0,30 < r < 0,49$  – умеренная;
- 4)  $0,50 < r < 0,69$  – средняя;
- 5)  $r > 0,70$  – сильная (тесная).

Существует еще один подход к **величине коэффициента корреляции** показывает степень зависимости (большие числа показывают большую степень зависимости). Корреляционные связи различаются по величине следующим образом:

- $r=0$  – нет никакой связи;  
 $r=0.01-0.3$  – слабая связь;  
 $r=0.31-0.7$  – умеренная связь;  
 $r=0.71-0.99$  – сильная связь;  
 $r=1$  – совершенная связь.

Сила связи не зависит от ее направленности и определяется по абсолютному значению коэффициента корреляции. Например: а) взаимосвязь между количеством съеденными булочками и вероятностью набрать вес  $r = + 0,94$  – сильная положительная связь, т.е. чем больше человек скушает булочек, тем больше у него шансов стать толстым; б) взаимосвязь между количеством съеденных булочек и оценками у студентов  $r = -0,8204$  – сильная отрицательная связь, то есть чем больше ест булочек студент, тем ниже у него оценки.

Частная классификация корреляционных связей по силе:

- 1) высокая значимая корреляция – при  $r$ , соответствующем уровню статистической значимости  $p \leq 0,01$ ;
- 2) значимая корреляция – при  $r$ , соответствующем уровню статистической значимости  $p \leq 0,05$ ;
- 3) тенденция достоверной связи – при  $r$ , соответствующем уровню статистической значимости  $p \leq 0,10$ ;
- 4) незначимая корреляция – при  $r$ , не достигающем уровня статистической значимости.

В этой классификации  **$p$  – это уровень значимости**, полученный при проверке нулевой гипотезы  $H_0$  о равенстве нулю коэффициента корреляции между интересующими нас признаками в генеральной совокупности. При этом может оказаться, что при малом объеме выборки сильная корреляция будет недостоверной, а при больших объемах выборки слабая корреляция – достоверной. Чем больше объем выборки, тем меньшей величины коэффициента корреляции оказывается достаточно, чтобы корреляция была признана достоверной. В психологических исследованиях ориентируются на вто-

ную классификацию, поскольку она учитывает объем выборки. Однако надо помнить, что сильная корреляция – это корреляция с коэффициентом корреляции  $r > 0,70$ , а не просто корреляция высокого уровня значимости. Не имеет смысла анализировать статистически незначимую связь.

Для коэффициента корреляции подсчитывают уровень статистической значимости  $p$ . Это уровень значимости, полученный при проверке нуль-гипотезы о равенстве нулю коэффициента корреляции между интересующими нас переменными в генеральной совокупности. Если  $p < 0,05$ , то говорят, что коэффициент корреляции оказался значимым, и, следовательно, можно отвергнуть нуль-гипотезу об отсутствии связи и принять гипотезу о наличии зависимости между интересующими нас переменными и в генеральной совокупности. Не имеет смысла интерпретировать незначимые коэффициенты корреляции.

Мера случайности полученного результата, равная вероятности того, что в генеральной совокупности этот результат (различия, связь) отсутствует. Чем меньше эта вероятность (значение  $p$ -уровня), тем выше статистическая значимость результата. Результат считается статистически достоверным (значимым), если  $p$ -уровень не превышает 0,05. Так, интерпретация  $p$ -уровня коэффициента корреляции производится аналогично тому, как это делалось для параметрических и непараметрических критериев:

- если  $p$ -уровень  $\leq 0,05$ , то связь между переменными является статистически значимой;
- если  $p$ -уровень  $> 0,05$ , то связь между переменными является статистически не значимой.

Также при интерпретации  $p$ -уровня коэффициента корреляции важным является не только сам факт значимости, но и ее уровень. Традиционно  $p$ -уровень корреляции дифференцируется на три уровня:

- $p \leq 0,05 > 0,01$  – низкая статистическая значимость,
- $p \leq 0,01 > 0,001$  – средней силы статистическая значимость,
- $p \leq 0,001$  – высокая статистическая значимость.

## **9.2. Характеристики коэффициентов корреляции**

Наиболее распространенными коэффициентами корреляции являются – Пирсона, Спирмена и Кендалла. Их общей особенностью является то, что они отражают взаимосвязь двух признаков, измеренных в количественной шкале – ранговой или метрической.

Приемы для измерения меры корреляции различны для количественных и качественных переменных. Для переменных, измеренных в номинальной или порядковой шкале, используются следующие эмпирические меры тесноты связи:

- а) коэффициент ассоциации, или тетрахорический показатель связи;
- б) коэффициент взаимной сопряженности Пирсона и Чупрова. Коэффициент Пирсона ( $r$ -Pearson). Мера корреляции, подходящая для двух непрерывных (метрических) переменных, измеренных на одной и той же выборке.
- в) коэффициент Фехнера;
- г) коэффициент корреляции рангов Спирмена – мера корреляции, подходящая для двух переменных, измеренных в ранговой шкале. 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.

Первые три метода измерения меры корреляции могут быть заменены методами сопоставления и сравнения, которые обладают большими возможностями, дают основание для достоверных выводов.

## **Метод ранговой корреляции Спирмена.**

Основными характеристиками метода являются:

- универсальность, поскольку он применим к любым ранжированным или количественно измеренным данным;
- простота, т.к. позволяет проводить расчеты вручную;
- широкие возможности – позволяет сопоставлять не индивидуальные показатели, а индивидуальные иерархии (профили), что невозможно никакими другими способами.

*Назначение.* Позволяет определить силу и направление корреляционной связи между двумя признаками или между двумя иерархиями признаков.

*Описание.* Для подсчета ранговой корреляции необходимо иметь два ряда значений, которые могут быть проранжированы.

Такими рядами могут быть:

- два признака, измеренные на одной и той же группе испытуемых; например коэффициент IQ и успеваемость у учащихся класса;
- две индивидуальные иерархии признаков, выявленные у двух испытуемых по одному и тому же набору признаков; например иерархия ценностей по методике Р. Рокича;
- две групповые иерархии признаков, при этом ранжируются среднегрупповые значения, полученные на двух группах испытуемых по одному и тому же набору признаков; например, групповые последовательности предпочтений при выборе из нескольких альтернатив;
- индивидуальная и групповая иерархия признаков, при этом ранжируются индивидуальные значения и среднегрупповые значения по тому же набору признаков, но при исключении этого отдельного испытуемого.

*Ограничения:*

- 1) данные получены в порядковой шкале, но могут быть измерены также в шкале интервалов или равных отношений, а затем проранжированы;
- 2) характер распределения переменных не имеет значения;
- 3) объем выборки удовлетворяет условию  $5 < n \leq 40$ ;
- 4) при большом числе сопряженных рангов дает огрубленное значение.

*Гипотезы:*  $H_0$ : корреляция между переменными X и Y (между иерархиями X и Y) не отличается от нуля.  $H_1$ : корреляция между переменными X и Y (между иерархиями X и Y) достоверно отличается от нуля.

*Алгоритм:*

1. Определить, какие два признака или две иерархии признаков будут участвовать в сопоставлении, как X и Y, внести их в первый и второй столбец таблицы.
2. Проранжировать значения переменной X, начиная ранг 1 наименьшему значению, в соответствии с правилами ранжирования. Занести ранги в третий столбец таблицы по порядку номеров испытуемых или признаков.
3. Проранжировать значения переменной Y по тем же правилам. Занести ранги в четвертый столбец таблицы по порядку номеров испытуемых или признаков.
4. Подсчитать разности  $d$  между рангами X и Y по каждой строке таблицы и занести их в пятый столбец.
5. Возвести каждую разность в квадрат. Эти значения занести в шестой столбец таблицы.
6. Подсчитать сумму квадратов.
7. При наличии одинаковых рангов рассчитать поправки.
8. Рассчитать коэффициент ранговой корреляции  $r_s$  по определенным формулам: при отсутствии одинаковых рангов или при наличии одинаковых рангов.

9. Определить по таблице критических значений для данного  $n$  и соответствующего уровня значимости  $\alpha$  критическое значение коэффициента корреляции  $r_s$  кр. Если эмпирическое значение превышает критическое или равно ему ( $r_s \geq r_{s \text{ кр}}$ ), то корреляция достоверно отличается от нуля при соответствующем уровне значимости  $\alpha$ .

$r$ -Пирсона применяется для изучения взаимосвязи двух метрических переменных, измеренных на одной и той же выборке

**Расчет коэффициента корреляции Пирсона** можно производить с помощью статистических пакетов. Частными случаями коэффициента линейной корреляции Пирсона являются ф-коэффициент сопряженности (или коэффициент ассоциации), рангово бисериальный коэффициент корреляции, точечный бисериальный коэффициент корреляции.

Рассмотрим подробнее ф-коэффициент сопряженности (ассоциации), который используется для изучения связи двух бинарных переменных. Бинарная переменная имеет только две градации, обозначаемые как 0 и 1. Например, пол: мужской – 0, женский – 1; семейное положение: холост – 0, женат – 1.

Для вычисления ф-коэффициента сопряженности строят специальные таблицы, которые называют таблицами сопряженности (таблицами кросstabуляции). Это таблицы совместного распределения частот двух и более номинальных признаков, измеренных на одной группе объектов. Столбцы такой таблицы соответствуют градациям одного номинального признака, а строки – градациям другого

Величина ф-коэффициента лежит в интервале +1 и -1, характеризуя направление и силу связи двух дихотомически измеренных признаков.

Ограничения:

- 1) сравниваемые переменные должны быть измерены в дихотомической шкале;
- 2) число варьирующих признаков в сопоставляемых переменных  $X$  и  $Y$  должно быть одинаково;
- 3) требуется приблизительное равенство количества значений 0 и 1 по каждой переменной.

Гипотезы:  $H_0$ : корреляция между переменными  $X$  и  $Y$  не отличается от нуля.  $H_1$ : корреляция между переменными  $X$  и  $Y$  достоверно отличается от нуля.

Алгоритм:

1. На основе эмпирических данных строят таблицу сопряженности.
2. Вычисляют эмпирическое значение ф-коэффициента сопряженности по формуле.
3. Поскольку для этого коэффициента корреляции нет стандартных таблиц критических значений, то их рассчитывают с помощью  $t$ -критерия Стьюдента по формуле
4. Проверяют величину на уровень значимости по таблице критических значений для критерия Стьюдента.

Анализ взаимосвязи признаков, измеренных в номинальной шкале, может проводиться и в том случае, если каждый из них имеет более чем два уровня. Для таблиц сопряженности размером больше чем  $2 \times 2$  можно только констатировать факт наличия (отсутствия) связи и ничего нельзя сказать о степени взаимосвязи.

### **Коэффициент корреляции Пирсона**

На величину коэффициента корреляции не влияет то, в каких единицах измерения представлены признаки. Следовательно, любые линейные преобразования признаков (умножение на константу, прибавление константы) не меняют значения коэффициента корреляции. Исключением является умножение одного из признаков на отрицательную константу: коэффициент корреляции меняет свой знак на противоположный.

Таблица 5 – Применение корреляции Спирмена и Пирсона

Коэффициенты	Переменные	
	зависимая	независимая
Корреляция Пирсона	метрическая	метрическая
Корреляция Спирмена	ранговая	ранговая
	метрическая	ранговая
	метрическая	метрическая

Так, корреляция Пирсона есть мера линейной связи между двумя переменными. Она позволяет определить, насколько пропорциональна изменчивость двух переменных. Если переменные пропорциональны друг другу, то графически связь между ними можно представить в виде прямой линии с положительным (прямая пропорция) или отрицательным (обратная пропорция) наклоном. На практике связь между двумя переменными, если она есть, является вероятностной и графически выглядит как облако рассеивания эллипсоидной формы. Этот эллипсоид, однако, можно представить (аппроксимировать) в виде прямой линии, или линии регрессии. Линия регрессии – это прямая, построенная методом наименьших квадратов: сумма квадратов расстояний (вычисленных по оси Y) от каждой точки графика рассеивания до прямой является минимальной.

Особое значение для оценки точности предсказания имеет дисперсия оценок зависимой переменной. По сути, дисперсия оценок зависимой переменной Y – это та часть ее полной дисперсии, которая обусловлена влиянием независимой переменной X. Иначе говоря, отношение дисперсии оценок зависимой переменной к ее истинной дисперсии равно квадрату коэффициента корреляции.

Квадрат коэффициента корреляции зависимой и независимой переменных представляет долю дисперсии зависимой переменной, обусловленной влиянием независимой переменной, и называется коэффициентом детерминации. Коэффициент детерминации, таким образом, показывает, в какой степени изменчивость одной переменной обусловлена (детерминирована) влиянием другой переменной.

Коэффициент детерминации обладает важным преимуществом по сравнению с коэффициентом корреляции. Корреляция не является линейной функцией связи между двумя переменными.

Поэтому, среднее арифметическое коэффициентов корреляции для нескольких выборок не совпадает с корреляцией, вычисленной сразу для всех испытуемых из этих выборок (т.е. коэффициент корреляции не аддитивен). Напротив, коэффициент детерминации отражает связь линейно и поэтому является аддитивным: допускается его усреднение для нескольких выборок.

Дополнительную информацию о силе связи дает значение коэффициента корреляции в квадрате – коэффициент детерминации: это часть дисперсии одной переменной, которая может быть объяснена влиянием другой переменной. В отличие от коэффициента корреляции коэффициент детерминации линейно возрастает с увеличением силы связи.

Коэффициенты корреляции Спирмена и  $\tau$ -Кендалла (ранговые корреляции)

Если обе переменные, между которыми изучается связь, представлены в порядковой шкале, или одна из них – в порядковой, а другая – в метрической, то применяются ранговые коэффициенты корреляции: Спирмена или  $\tau$ -Кенделла. И тот, и другой коэффициент требует для своего применения предварительного ранжирования обеих переменных.

Коэффициент ранговой корреляции Спирмена – это непараметрический метод, который используется с целью статистического изучения связи между явлениями. В этом случае определяется фактическая степень параллелизма между двумя количественными рядами.

ми изучаемых признаков и дается оценка тесноты установленной связи с помощью количественно выраженного коэффициента.

Если члены группы численностью были ранжированы сначала по переменной  $x$ , затем – по переменной  $y$ , то корреляцию между переменными  $x$  и  $y$  можно получить, просто вычислив коэффициент Пирсона для двух рядов рангов. При условии отсутствия связей в рангах (т.е. отсутствия повторяющихся рангов) по той и другой переменной, формула для Пирсона может быть существенно упрощена в вычислительном отношении и преобразована в формулу, известную как Спирмена.

Мощность коэффициента ранговой корреляции Спирмена несколько уступает мощности параметрического коэффициента корреляции.

Коэффициент ранговой корреляции целесообразно применять при наличии небольшого количества наблюдений. Данный метод может быть использован не только для количественно выраженных данных, но также и в случаях, когда регистрируемые значения определяются описательными признаками различной интенсивности.

Коэффициент ранговой корреляции Спирмена при большом количестве одинаковых рангов по одной или обеим сопоставляемым переменным дает огрубленные значения. В идеале оба коррелируемых ряда должны представлять собой две последовательности несовпадающих значений

Альтернативу корреляции Спирмена для рангов представляет корреляция  $\tau$ -Кендалла. В основе корреляции, предложенной М. Кендаллом, лежит идея о том, что о направлении связи можно судить, попарно сравнивая между собой испытуемых: если у пары испытуемых изменение по  $x$  совпадает по направлению с изменением по  $y$ , то это свидетельствует о положительной связи, если не совпадает – то об отрицательной связи.

Коэффициенты корреляции были специально разработаны для численного определения силы и направления связи между двумя свойствами, измеренными в числовых шкалах (метрических или ранговых). Как уже упоминалось, максимальной силе связи соответствуют значения корреляции +1 (строгая прямая или прямо пропорциональная связь) и -1 (строгая обратная или обратно пропорциональная связь), отсутствию связи соответствует корреляция, равная нулю.

Дополнительную информацию о силе связи дает значение коэффициента детерминации: это часть дисперсии одной переменной, которая может быть объяснена влиянием другой переменной.

### **9.3. Содержательный анализ корреляционных матриц**

**К**орреляционная матрица.

1. Результат вычисления корреляций одного типа для каждой пары из множества переменных, измеренных в количественной шкале на одной выборке.

2. Таблица, в которой на пересечениях столбцов и строк представлены показатели корреляции каждой переменной с каждой. Размер матрицы (число ячеек на пересечении переменных) пропорционален числу коррелируемых переменных. Корреляционная плеяда (или коррелограмма, или структурограмма)

Корреляционная матрица – это таблица, которая показывает степень взаимосвязи между различными переменными. Коэффициент корреляции может варьироваться от -1 до 1:

- 1 – сильная прямая связь (например, чем больше запросов по одному бренду, тем больше по другому);
- -1 – сильная обратная связь (например, чем больше запросов по одному бренду, тем меньше по другому);
- 0 – отсутствие связи.

Корреляционные матрицы помогают:

- Выявлять взаимосвязи между брендами, категориями или продуктами.
- Анализировать влияние внешних факторов.
- Определять этапы воронки, которые наиболее сильно влияют на конверсию.
- Формировать гипотезы.

На практике корреляционная матрица обычно используется по трем причинам:

**1. Корреляционная матрица удобно суммирует набор данных.**

Матрица корреляции — это простой способ суммировать корреляции между всеми переменными в наборе данных. Например, предположим, что у нас есть следующий набор данных, содержащий следующую информацию для 1000 студентов:

Было бы очень сложно понять взаимосвязь между каждой переменной, просто взглянув на необработанные данные. К счастью, корреляционная матрица может помочь нам быстро понять корреляции между каждой парой переменных.

**2. Корреляционная матрица служит диагностикой регрессии.**

Одним из ключевых предположений множественной линейной регрессии является то, что ни одна независимая переменная в модели не имеет сильной корреляции с какой-либо другой переменной в модели.

Когда две независимые переменные сильно коррелируют, это приводит к проблеме, называемой мультиколлинеарностью, и может затруднить интерпретацию результатов регрессии.

Один из самых простых способов обнаружить потенциальную проблему мультиколлинеарности — посмотреть на матрицу корреляции и визуально проверить, сильно ли коррелируют какие-либо переменные друг с другом.

**3. Корреляционная матрица может использоваться в качестве входных данных для других анализов.**

Корреляционная матрица используется в качестве входных данных для других сложных анализов, таких как исследовательский факторный анализ и модели структурных уравнений.

## МОДУЛЬ 2

# МЕТОДЫ СРАВНЕНИЯ ВЫБОРОК

### Тема 10. Параметрические методы сравнения выборок

10.1. Анализ различий. Т-тест (критерий Стьюдента).

10.2. Сравнение долей признака: угловое преобразование φ-Фишера.

#### 10.1. Анализ различий. Т-тест (критерий Стьюдента)

**T**-тесты предназначены для установления различий между двумя группами респондентов. Основные характеристики переменных, участвующих в t-тестах (табл. 6).

Таблица 6 – Т-тесты для независимых выборок

Т-тесты для независимых выборок			
Зависимые переменные		Независимые переменные	
Количество	Тип	Количество	Тип
Одна	Дихотомическая интервальная	Любое	Интервальная
Т-тесты для зависимых выборок			
Зависимые переменные		Независимые переменные	
Количество	Тип	Количество	Тип
-	-	Две	Интервальная
Т-тесты для одной выборки			
Зависимые переменные		Независимые переменные	
Количество	Тип	Количество	Тип
-	-	Любое	Интервальная

**Т-тест (или критерий Стьюдента)** представляет собой параметрический статистический метод, предназначенный для проверки гипотез относительно средних значений совокупностей. Фундаментальная идея этого метода заключается в сравнении средних значений выборок с учётом их дисперсий. Уильям Госсет, публиковавший свои работы под псевдонимом «Стьюдент», разработал данный метод в начале XX века, работая на пивоваренной фабрике Guinness для контроля качества продукции.

Математическая основа t-теста – t-распределение (распределение Стьюдента), которое учитывает неопределённость, возникающую при работе с малыми выборками. В отличие от нормального распределения, t-распределение имеет более тяжёлые «хвосты», что делает статистические выводы более консервативными и надёжными при ограниченном объёме данных.

Критерий t-Стьюдента для зависимых выборок. Проверяет гипотезу о том, что средние значения двух генеральных совокупностей, из которых извлечены сравниваемые зависимые выборки, отличаются друг от друга.

**Исходные предположения:**

1. Каждому представителю одной выборки поставлен в соответствие представитель другой выборки.

2. Данные двух выборок положительно коррелируют.

3. Распределение в обеих выборках соответствует нормальному закону.

Структура исходных данных: имеется по два значения изучаемого признака(ов).

Принцип работы t-теста основан на формулировании двух конкурирующих гипотез:

**Нулевая гипотеза:** предположение об отсутствии статистически значимых различий между средними значениями групп.

**Альтернативная гипотеза:** предположение о наличии статистически значимых различий между средними значениями групп.

T-статистика рассчитывается как отношение разницы средних значений к стандартной ошибке этой разницы. Полученное значение сравнивается с критическим значением t-распределения при заданном уровне значимости  $\alpha$  (обычно 0,05 или 0,01) и соответствующем числе степеней свободы.

T-тест является мощным инструментом анализа, однако его применение требует соблюдения ряда условий, включая нормальное распределение данных в выборках и наличие независимых наблюдений. При нарушении этих условий более уместными могут оказаться непараметрические методы, такие как критерий Манна – Уитни или критерий Вилкоксона.

Статистический критерий Стьюдента имеет несколько модификаций, каждая из которых адаптирована для решения специфических аналитических задач. Выбор конкретного типа t-теста зависит от характера имеющихся данных и поставленных исследовательских вопросов.

#### **Разновидности t-тестов:**

1. **Одновыборочный t-тест** – применяется для сравнения среднего значения одной выборки с некоторым фиксированным значением  $\mu_0$  (например, с эталоном или стандартом).

В результате t-теста для одной выборки можно выяснить, отличается ли значительно реальное среднее значение какой-либо переменной от стандарта. В исследованиях при помощи данного теста определяют, отличается ли среднее значение какого-либо параметра для определенной целевой группы респондентов от среднего значения по всей выборке.

2. **Двухвыборочный t-тест для независимых выборок** – используется для сравнения средних значений двух независимых групп. Имеет две разновидности:

– Т-тест с равными дисперсиями: применяется, когда дисперсии в сравниваемых группах предположительно равны

$$t = (\bar{x}_1 - \bar{x}_2) / (sp * \sqrt{(1/n_1 + 1/n_2)})$$

$$\text{где } sp = \sqrt{((n_1 - 1) * s_1^2 + (n_2 - 1) * s_2^2) / (n_1 + n_2 - 2)}$$

Число степеней свободы  $df = n_1 + n_2 - 2$ .

– Т-тест Уэлча – модификация для случая неравных дисперсий:

$$t = (\bar{x}_1 - \bar{x}_2) / \sqrt{s_1^2/n_1 + s_2^2/n_2}$$

Число степеней свободы рассчитывается по формуле Саттертуэйта:

$$df = (s_1^2/n_1 + s_2^2/n_2)^2 / [(s_1^2/n_1)^2 / (n_1 - 1) + (s_2^2/n_2)^2 / (n_2 - 1)]$$

Обратите внимание: зависимая переменная есть только для t-тестов независимых выборок. Для других видов t-тестов (зависимых выборок и одной выборки) зависимая переменная отсутствует. Это связано с тем, что в последнем случае анализу подвергается фактически одна и та же выборка респондентов. В качестве тестируемых независимых переменных во всех случаях используются только переменные с интервальнойной

шкалой. Порядковые переменные могут использоваться только после преобразования их к интервальному виду.

В случае t-тестов для независимых выборок под независимыми выборками понимаются бинарные категории (то есть варианты ответа) какой-либо переменной. Например, мужчины и женщины (вопрос Пол респондента), покупатели и не покупатели какого-либо продукта и т. д. То есть когда есть два уровня группирующей (зависимой) переменной и несколько независимых переменных, на основании которых и будет выполняться различие между группами зависимой переменной.

*3. Парный t-тест используется для сравнения средних значений двух связанных выборок* (например, измерений «до» и «после» у одних и тех же объектов).

$$t = \bar{d} / (sd / \sqrt{n})$$

где  $\bar{d}$  – среднее значение разностей парных наблюдений,  $sd$  – стандартное отклонение разностей,  $n$  – количество пар. Число степеней свободы  $df = n - 1$ .

Тип t-теста	Использование	Предпосылки	Робастность
Одновыборочный	Сравнение с известным стандартом	Нормальность распределения	Средняя
Двухвыборочный с равными дисперсиями	Сравнение двух групп с близкими вариациями	Нормальность, гомоскедастичность	Низкая
Двухвыборочный Уэлча	Сравнение групп с разными вариациями	Только нормальность	Высокая
Парный	Анализ зависимых измерений	Нормальность разностей	Средняя

Все t-тесты относятся к классу параметрических методов и основаны на предположении о нормальности распределения данных. При значительных отклонениях от нормальности или наличии выбросов точность результатов может снижаться. В таких случаях рекомендуется применять либо трансформированные данные, либо альтернативные непараметрические методы.

#### **Рекомендации для эффективного применения t-тестов:**

1. Необходимо визуализировать данные перед анализом – используйте гистограммы, графики плотности и квантильные графики для оценки распределения.
2. Необходимо сохранять полный отчет о результатах – помимо p-значения, фиксируйте t-статистику, степени свободы и доверительные интервалы.
3. Необходимо рассчитывать размер эффекта – p-значение показывает только статистическую значимость, но не величину эффекта. Для этой цели используйте коэффициент Коэна ( $d$ ) или другие метрики.
4. Необходимо использовать адекватные методы для множественных сравнений – при необходимости применяйте поправку Бонферрони, процедуру Хольма или метод контроля FDR.

Ключевые компоненты результатов t-теста, требующие интерпретации:

1. Р-значение вероятность получить наблюдаемые или более экстремальные результаты при условии истинности нулевой гипотезы. Меньшее р-значение указывает на более сильные доказательства против нулевой гипотезы.
2. Т-статистика – показывает, насколько сильно средние значения выборок отличаются друг от друга относительно вариабельности данных. Большой модуль t-статистики указывает на более выраженное различие.
3. Доверительные интервалы – диапазоны, в которых с заданной вероятностью находится истинное значение параметра. Широкие интервалы указывают на низкую точность оценки.
4. Степени свободы (df) – параметр, зависящий от размера выборок и влияющий на форму t-распределения.

Частые ошибки при интерпретации результатов t-теста:

- ошибки значимости: интерпретация статистической значимости как практической значимости. Статистически значимый результат может иметь минимальную практическую ценность, особенно при больших выборках;
- дихотомическое мышление: восприятие р-значения как бинарного показателя (значимо/незначимо) без учета континуума доказательств. Р-значение 0.051 интерпретируется радикально иначе, чем 0.049, хотя разница минимальна;
- игнорирование размера эффекта: фокусирование только на р-значении без оценки величины наблюдаемого эффекта. Маленький эффект может быть статистически значимым при больших выборках, но иметь ограниченную практическую ценность;
- некорректные выводы о причинно-следственных связях: интерпретация статистически значимой разницы как доказательства причинно-следственной связи, особенно в наблюдательных исследованиях;
- обобщение за пределы выборки: распространение выводов на популяции, существенно отличающиеся от изученных выборок.

Таким образом, Т-тесты являются фундаментальным инструментом проверки гипотез, без которого невозможно представить современную науку и аналитику данных. Правильное понимание их математической основы, областей применения и потенциальных ограничений позволяет исследователям делать обоснованные выводы на основе эмпирических данных.

## **10.2. Сравнение долей признака: угловое преобразование φ-Фишера**

**Критерий Фишера** предназначен для сопоставления двух выборок (независимых или связанных) по частоте встречаемости интересующего исследователя эффекта. Он позволяет оценить достоверность различия между долями обеих выборок, в которых эффект проявился. Это непараметрическая альтернатива t-критерию для зависимых выборок. Критерий применяется в ситуациях, когда исследователь проводит два измерения (например, при разных условиях) одних и тех же субъектов и желает установить наличие или отсутствие различия. φ-критерий Фишера называется угловым преобразованием Фишера, поскольку в нем исходная доля, нормированная на единицу (P), переводится в величину угла, нормированного на величину  $\pi = 3,14\dots$  по формуле

Критерий φ-Фишера:

- Применяется для проверки гипотезы о равенстве дисперсий двух выборок. Его относят к критериям рассеяния.
- Имеет смысл перед использованием критерия t-Стьюардента предварительно проверить гипотезу о равенстве дисперсий. Если она верна, то для сравнения средних

можно воспользоваться критерием t-Стьюдента (гипотезы о равенстве средних значений в двух выборках).

- Критерий Фишера основан на дополнительных предположениях о независимости и нормальности выборок данных. Перед его применением рекомендуется выполнить проверку нормальности распределения признака.
- В регрессионном анализе критерий Фишера позволяет оценивать значимость линейных регрессионных моделей.
- В частности, он используется в шаговой регрессии для проверки целесообразности включения или исключения независимых переменных (признаков) в регрессионную модель.
- В дисперсионном анализе критерий Фишера позволяет оценивать значимость факторов и их взаимодействия.

*Проверяемые гипотезы:*

$H_0$ : Доля испытуемых, у которых проявился исследуемый эффект, в выборке 1 не больше, чем в выборке 2. При нулевой гипотезе (отсутствие эффекта обработки) число положительных разностей имеет биномиальное распределение со средним, равным половине объема выборки (положительных разностей будет примерно столько же, сколько отрицательных). Основываясь на биномиальном распределении, можно вычислить критические значения.

$H_1$ : Доля испытуемых, у которых проявился исследуемый эффект, в выборке 1 больше (меньше), чем в выборке 2.

## Тема 11. Непараметрические методы сравнения выборок

- 11.1. Сравнение двух независимых совокупностей (критерии Манна – Уитни).
- 11.2. Сравнение трех и более независимых совокупностей. Критерий Крускала – Уоллиса как непараметрический аналог дисперсионного анализа для независимых совокупностей.

### 11.1. Сравнение двух независимых совокупностей (критерии Манна – Уитни)

#### **Н**епараметрические критерии.

1. Методы сравнения двух и более выборок по признаку. Распределение признака в выборке не соответствует нормальному виду.
2. Представляют собой функции, зависящие непосредственно от варианта данной совокупности с их частотами; служат для проверки рабочих гипотез независимо от формы распределения совокупностей, из которых взяты сравниваемые выборки.

**К**ритерий У Манна – Уитни (U; Mann – Whitney U-test) – непараметрический критерий, позволяющий проверить гипотезу о том, что значения двух совокупностей, из которых извлечены сравниваемые независимые выборки, отличаются друг от друга. Критерий Манна – Уитни представляет непараметрическую альтернативу t-критерию для независимых выборок. Опция предполагает, что данные расположены таким же образом, что и в t-критерии для независимых выборок. В частности, данные должны содержать группирующую переменную, имеющую, по крайней мере, два

разных кода для однозначной идентификации принадлежности каждого наблюдения к определенной группе.

Критерий U Манна – Уитни предполагает, что рассматриваемые переменные измерены, по-крайней мере, в порядковой шкале (ранжированы). Заметим, что во всех ранговых методах делаются поправки на совпадающие ранги.

Интерпретация теста, по существу, похожа на интерпретацию результатов t-критерия для независимых выборок за исключением того, что U-критерий вычисляется как сумма индикаторов парного сравнения элементов первой выборки с элементами второй выборки.

U-критерий – наиболее мощная (чувствительная) непараметрическая альтернатива t-критерию для независимых выборок; фактически, в некоторых случаях он имеет даже большую мощность, чем t-критерий.

## ***11.2. Сравнение трех и более независимых совокупностей. Критерий Крускала – Уоллиса как непараметрический аналог дисперсионного анализа для независимых совокупностей***

Критерий Н Крускала – Уоллеса (H; Kruskal – Wallis H Test) позволяет проверить гипотезы о различии более двух независимых выборок по уровню выраженности изучаемого признака.

Критерий Крускала – Уоллиса основан на рангах (а не на исходных наблюдениях) и предполагает, что рассматриваемая переменная непрерывна и измерена как минимум в порядковой шкале. Критерий проверяет гипотезу: имеют ли сравниваемые выборки одно и то же распределение или же распределения с одной и той же медианой. Таким образом, интерпретация критерия схожа с интерпретацией параметрической однофакторной ANOVA за исключением того, что этот критерий основан на рангах, а не на средних значениях.

Для применения критерия Крускала – Уоллиса должны выполняться следующие условия:

- Результаты выборок случайно и независимо друг от друга извлекаются из соответствующих генеральных совокупностей.
- Анализируемая переменная является непрерывной.
- Наблюдения допускают ранжирование как внутри, так и между группами.
- Все с генеральных совокупностей имеют одинаковую изменчивость.
- Все с генеральных совокупностей имеют одинаковый вид.

Процедура Крускала – Уоллиса имеет меньше ограничений, чем F-критерий. Процедура Крускала – Уоллиса предусматривает ранжирование только по всем выборкам в совокупности. Общее распределение должно быть непрерывным, но его вид значения не имеет. Если эти условия не выполняются, критерий Крускала – Уоллиса по-прежнему можно применять для проверки гипотезы о различиях между с генеральными совокупностями. Альтернативная гипотеза утверждает, что среди с генеральными совокупностями существует хотя бы одна, которая отличается от остальных какой-нибудь характеристикой – либо средним значением, либо видом. С другой стороны, для применения F-критерия переменная должна быть числовой, а с выборок должны извлекаться из нормально распределенных генеральных совокупностей, имеющих одинаковую дисперсию.

Ограничения:

- если  $\alpha=0,05$ , то при сопоставлении трех выборок допускается, чтобы в одной выборке было 3 значения, а во-второй и третьей тогда по 2 значения.
- если  $\alpha=0,01$ , то необходимо чтобы в каждой выборке было не менее 3 наблюдений, или чтобы по крайней мере в одной из выборок было 4 наблюдения, а в двух других – по 2.
- при большом количестве выборок и испытуемых в каждой выборке необходимо пользоваться таблицей критических значений критерия  $\chi^2$  при числе степеней свободы  $v = c-1$ .

- при множественном сопоставлении выборок достоверные различия между какой-либо конкретной парой (или парами) их могут оказаться стертыми

Алгоритм Н-теста:

- 1) обе выборки соединяются в единую выборку. При этом запоминается, к какой выборке относится каждый элемент массива
- 2) общая выборка ранжируется по возрастанию
- 3) разбить единую выборку на прежние две выборки
- 4) подсчитать сумму рангов отдельно по каждой выборке. Проверить, совпадает ли общая сумма рангов с расчетной.
- 5) найти значение H

$$H = \frac{12}{N(N+1)} \sum_{i=1}^c \frac{R_i^2}{n_i} - 3(N+1),$$

где  $N = \sum n_i$  – общий объем наблюдений,

$R_i$  – ранговая сумма i-той группы,

$n_i$  – объем i-той группы

- 6) найти эмпирическое значение Нэмп, внеся поправку на одинаковые ранги

$$H_{emp} = \frac{H}{1 - \frac{\sum (t^3 - t)}{N^3 - N}},$$

где  $t$  – объем группы одинаковых рангов.

- 7) Найти критическое значение критерия Нкрит:

- при числе групп  $c=3$  – по таблице критических значений Н-критерия;
- при числе групп  $c>3$  – по таблице критических значений критерия  $\chi^2$ .

- 8) Сравнить найденное эмпирическое значение критерия Нэмп и критическое значение критерия Нкрит

- если  $H_{emp} < H_{crit}$  ( $p>\alpha$ ), то нулевая гипотеза  $H_0$  принимается;
- если уровень значимости  $H_{emp} \geq H_{crit}$  ( $p\leq\alpha$ ), то нулевая гипотеза  $H_0$  отклоняется.

# **СПИСОК ЛИТЕРАТУРЫ**

## **Основная литература**

1. Основы математической статистики в психологии: учеб.-метод. пособие: в 2 ч. Ч. 1 / сост.: Н.П. Радчикова, Н.А. Литвинова; М-во образования Респ. Беларусь, Белорус. гос. пед. ун-т. – 7-е изд. – Минск: БГПУ, 2008. – 88 с.
2. Сидоренко, Е.В. Методы математической обработки в психологии / Е.В. Сидоренко. – СПб.: СПЦ, 1996. – 350 с.
3. Статистика. Практикум: учебное пособие для студентов учреждений высшего образования по экономическим специальностям / Н.В. Агабекова [и др.]; под ред. Н.В. Агабековой. – Минск: РИВШ, 2023. – 367, [1] с.: табл.
4. Белановская, М.Л. Статистические методы в психологии: учебное пособие для студентов учреждений высшего образования по специальностям «Психология», «Практическая психология» / М.Л. Белановская. – Минск: РИВШ, 2022. – 295, [1] с.: ил., табл.

## **Дополнительная литература**

1. Бююль, А. SPSS: искусство обработки информации. Анализ статистических данных / А. Бююль, П. Цефель. – СПб.: ДиаСофтиЮп, 2005. – 608 с.
2. Гласс Дж., Стэнли Дж. Статистические методы в педагогике и психологии. – М.: Прогресс, 1976.
3. Ермолаев О.Ю. Математическая статистика для психологов. – М.: Московский психолого-социальный институт: Флинта, 2003. – 336 с.
4. Калачева И.В. Статистические методы в психологии: учебно-методическое пособие. – Могилев: МГУ имени А.А. Кулешова, 2017.
5. Кремень М.А. Математические методы в научных исследованиях: для педагогов и психологов. – Мн., 1998.
6. Макарова Н.В. Статистический анализ медико-биологических данных с использованием пакетов статистических программ Statistica, SPSS, NCSS, SYSTAT: метод. рекомендации / под ред. С.С. Алексанина; ФГБУ ВЦЭРМ им. А.М. Никифорова МЧС России. – СПб.: Полиграф. центр Санкт-Петербургского ун-та Гос. противопожарной службы МЧС России, 2012. – 178 с.
7. Наследов А.Д. SPSS: Компьютерный анализ данных в психологии и социальных науках. – СПб.: Питер, 2005. – 416 с.
8. Наследов А.Д. Математические методы в психологических исследованиях. Анализ и интерпретация данных. – СПб.: Речь, 2008. – 392 с.
9. Папоян С.С. Математические методы в социальной психологии. – М.: Наука, 1983.
10. Сидоренко Е.В. Методы математической обработки в психологии. – СПб.: «Речь», 2000.
11. Сидоренко Е.В. Методы математической обработки в психологии. – 2-е изд. – СПб., 2000.
12. Тюрин Ю.Н., Макаров Н.Н. Статистический анализ данных на компьютере. – М., 1998.
13. Шелехова Л.В. Математические методы в психологии и педагогике: в схемах и таблицах: учеб. пособие. – М.: Лань, 2015. – 224 с.

Учебное издание

**КУХТОВА Наталья Валентиновна**

**МЕТОДЫ СТАТИСТИЧЕСКОГО АНАЛИЗА ДАННЫХ**

Курс лекций

Технический редактор

*Г.В. Разбоева*

Компьютерный дизайн

*Е.А. Барышева*

Подписано в печать 2025. Формат 60x84<sup>1/16</sup>. Бумага офсетная.

Усл. печ. л. 3,26. Уч.-изд. л. 3,64. Тираж . Заказ .

Издатель и полиграфическое исполнение – учреждение образования  
«Витебский государственный университет имени П.М. Машерова».

Свидетельство о государственной регистрации в качестве издателя,  
изготовителя, распространителя печатных изданий

№ 1/255 от 31.03.2014.

Отпечатано на ризографе учреждения образования  
«Витебский государственный университет имени П.М. Машерова».  
210038, г. Витебск, Московский проспект, 33.