ДАЛЬНЕВОСТОЧНЫЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ТИХООКЕАНСКИЙ ИНСТИТУТ
ДИСТАНЦИОННОГО ОБРАЗОВАНИЯ И ТЕХНОЛОГИЙ

Л. С. Титкова

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ПСИХОЛОГИИ

ВЛАДИВОСТОК
2002 г.

Содержание

Учебно-методическое пособие

Данное учебно-методическое пособие является дополнением к электронному учебнику по курсу «Математические методы в психологии» и предназначено для контроля и проверки теоретических знаний по дисциплине.

Учебное пособие включает в себя список вопросов, необходимых для теоретического изучения данного курса и тестер, для проверки знаний.

Студент должен:

- овладеть системой знаний о применении математических методов в психологии;
- научится интерпретировать, полученные результаты;
- подготовиться к выполнению практических заданий, которые предусмотрены в изучаемом курсе.

I. ВОПРОСЫ
для подготовки к тестированию

1. Основные разделы математической статистики.
2. Измерения в психологии. Номинальная шкала.
3. Порядковые шкалы.
4. Интервальная шкала.
5. Шкалы равных отношений.
6. Распределение признака. Нормальное распределение, его особенности.
7. Нормальное распределение. Закон трех сигм.
8. Асимметрия и эксцесс.
9. Проверка нормальности распределения результативного признака.
10. Меры центральной тенденции.
11. Меры изменчивости. Оценка разброса.
12. Генеральная совокупность, свойства и параметры совокупности, виды совокупностей.
13. Выборка. Классификация выборки. Репрезентативность.
14. Статистические гипотезы. Виды статистических гипотез.
15. Статистический критерий. Виды статистических критериев.
16. Уровень статистической значимости.
17. Ошибка первого рода. Вероятность ошибки первого рода.
18. Мощность статистического критерия.
19. Ошибка второго рода. Вероятность ошибки второго рода.
20. Выявление различий в распределении признака. Обоснование задачи сравнения распределений признака.
21. Критерий χ^2 - Пирсона. Применение, ограничения критерия.
22. Понятие о корреляционной зависимости и корреляционной связи.
23. Характеристики корреляционной зависимости.
24. Формула ранговой корреляции Спирмена.
25. Метод линейной корреляции Браве-Пирсона.
26. Выявление различий в уровне исследуемого признака (Q-критерий Розенбаума).
27. Классификация сдвигов. Типический и нетипический сдвиг.
28. Оценка достоверности сдвига в значениях исследуемого признака (Т – критерий Вилкоксона).
29. Критерий t-Стьдента.
30. Критерий F-Фишера.

II. МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

1. Учебно-методическое пособие курса «Математические методы в психологии»
2. Электронный учебник по дисциплине «Математические методы в психологии»

III. ТЕХНИЧЕСКИЕ СРЕДСТВА ОБЕСПЕЧЕНИЯ ДИСЦИПЛИНЫ

1. Персональный компьютер

Введение

Слово «статистика» часто ассоциируется со словом «математика», и это пугает студентов, связывающих это понятие со сложными формулами, требующими высокого уровня абстрагирования.

Статистика – это, прежде всего, способ мышления, и для ее применения нужно лишь иметь немного здравого смысла и знать основы математики. В нашей повседневной жизни мы, сами о том не догадываясь, постоянно занимаемся статистикой. Хотим ли мы спланировать бюджет, рассчитать потребление бензина автомашиной, оценить усилия, которые потребуются для усвоения какого-то курса, с учетом полученных до сих пор отметок, предусмотреть вероятность хорошей и плохой погоды по метеорологической сводке или вообще оценить, как повлияет то или иное событие на наше личное или совместное будущее, – нам постоянно приходится отбирать, классифицировать и упорядочивать информацию, связывать ее с другими данными так, чтобы можно было сделать выводы, позволяющие принять верное решение.

Все эти виды деятельности мало отличаются от тех операций, которые лежат в основе научного исследования и состоят в синтезе данных, полученных на различных группах объектов в том или ином эксперименте, в их сравнении с целью выяснить черты различия между ними, в их сопоставлении с целью выявить показатели, изменяющиеся в одном направлении, и, наконец, в предсказании определенных фактов на основании тех выводов, к которым приводят полученные результаты. Именно в этом заключается цель статистики в науках вообще, особенно в гуманитарных. В последних нет ничего абсолютно достоверного, и без статистики выводы в большинстве случаев были бы чисто интуитивными и не могли бы составлять солидную основу для интерпретации данных, полученных в других исследованиях.

Статистика содержит три главных раздела, к которым относятся: описательная статистика, индуктивная статистика и корреляционный анализ.

1) ***Описательная статистика****,* позволяет описывать, подытоживать и воспроизводить в виде таблиц или графиков данные того или иного *распределения,* вычислять *среднее* для данного распределения и его *размах* и *дисперсию*.
2) Задача ***индуктивной статистики заключается в проверке*** того, можно ли распространить результаты, полученные на отдельной *выборке,* на всю *популяцию,* из которой взята эта выборка. Иными словами, правила этого раздела статистики позволяют выяснить, до какой степени можно путем индукции обобщить на большее число объектов ту или иную закономерность, обнаруженную при изучении их ограниченной группы в ходе какого-либо наблюдения или эксперимента. Таким образом, при помощи индуктивной статистики делают какие-то выводы и обобщения, исходя из данных, полученных при изучении выборки.
3) ***Корреляционный анализ*** позволяет узнать, насколько связаны между собой две переменные, с тем чтобы можно было предсказывать возможные значения одной из них, если мы знаем другую.

Для получения зачета по данной дисциплине необходимо выполнить четыре лабораторные работы. Тексты и методические рекомендации по выполнению лабораторных работ даны в приложениях разделов со второго по пятый. Лабораторные работы выполняются на компьютере в программе “Excel”. Выполненные лабораторные работы должны отсылаться по электронной почте по адресу mailto:titkova@psy.dvgu.ru

Рабочая учебная программа

дисциплины
Математические методы в психологии
психология
специальность____

факультет_Психологии и социальной работы
кафедра_____психосоциальной работы
курс ______________________________,
лекции__(час.)
практические занятия______(час.)
семинарские занятия_______________(час.)
лабораторные работы_ _____(час.)
всего часов____120___________________
самостоятельная работа__120_____(час.)
реферативные работы не предусмотрены
контрольные работы предусмотрены

экзамен ___________________________
(семестр)
зачет______________________________
(семестр)

Рабочая программа составлена на основании авторских разработок и государственного стандарта

Рабочая программа обсуждена на заседании кафедры психосоциальной работы
«___»________________2002 г.

Заведующий кафедрой Кулебякин Е.В._____

Составитель (ли) Титкова Лидия Степановна
(Ф.И.О., ученое звание)

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Программа курса «Математические методы в психологии» составлена в соответствии с требованиями государственного образовательного стандарта высшего профессионального образования.

Цель курса – ориентация студентов в сущности применения математических методов в психологических науках.

Преподавание курса связано с другими курсами государственного образовательного стандарта: «Основы психодиагностики», «Общий психологический практикум», «Экспериментальная психология».

По завершению обучения по дисциплине студент должен:

– овладеть системой знаний о применении математических методов в психологии;

– владеть умениями применения статистических критериев в психологии и интерпретации, полученных результатов.

I. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

1. Основные понятия, используемые в математической обработке данных

Рассматриваются виды данных в статистике. Определение количественных, качественных и порядковых данных. Характеристика типов шкал, применяемых в психологии. Операции с числами, возможные с каждым типом шкал измерения. Ограничения в использовании различных типов шкал. Перевод данных из одного типа шкал в другой тип измерения.

Понятие генеральной совокупности и выборки. Свойства и параметры совокупности. Репрезентативность. Классификация выборок по способу отбора, объему, схеме испытаний и репрезентативности.

Понятие проблемы и гипотезы. Принципы фальсифицируемости и верифицируемости. Научная и статистическая гипотеза. Нулевая и альтернативная гипотезы.

Определение статистического критерия. Параметрические и непараметрические критерии. Уровни статистической значимости. Ошибка первого рода. Ось значимости. Мощность критериев и ошибка второго рода.

2. Методы описательной статистики

Представление количественных данных. Различные этапы представления данных. Несгруппированные ряды. Упорядоченные ряды. Ранжирование данных. Распределение частот. Числовые характеристики распределения данных. Оценка средних величин. Мода, медиана и средняя арифметическая. Оценка разброса данных. Коэффициенты вариации. Ассиметрия и эксцесс.

3. Нормальный закон распределения случайной величины

Нормальный закон распределения случайной величины. Понятие распределения признака и нормального распределения признака; основные характеристики нормального распределения.

Построение кривой нормального распределения. Формула для нахождения теоретических частот (m'), алгоритм построения кривой нормального распределения.

Проверка нормальности распределения результативного признака. Даются формулы для расчета критических значений А (асимметрия) и Е (эксцесс) Пустыльника Е.И. и Плохинского Н.А.

4. Меры связи между признаками

Понятие корреляционного анализа; корреляционной связи и корреляционной зависимости; методы для расчета коэффициента корреляции: метод ранговой корреляции Спирмена; метод Браве-Пирсона. Интерпретация корреляции.

5. Методы проверки статистических гипотез

Описание и применение статистических критериев: t-критерий Стьюдента, F-критерий Фишера, Q-критерий Розенбаума, Т-критерий Вилкоксона, χ^2-критерий Пирсона.

6. Многомерный анализ данных

Двумерный регрессионный анализ. Двухфакторный дисперсионный анализ. Дисперсионный анализ (ДА).

7. Многомерный анализ данных. Факторный анализ

Применение ФА в психологии как одного из методов многомерного количественного описания (измерения, анализа) наблюдаемых переменных. Разведочный и конфирматорный ФА. Этапы факторного анализа. Статистические показатели для оценки результатов факторного анализа.

8. Многомерное шкалирование

Понятие многомерного шкалирования (метрическое и неметрическое) (МШ). Построение пространственной модели субъективного расстояния в психологическом пространстве. Метод ортогональных проекций. Построение метрической модели.

II. САМОСТОЯТЕЛЬНАЯ РАБОТА СТУДЕНТОВ (120 часов)

1) Знакомство с научной литературой по психологии и математическим методам, применяемым в психологии.
2) Выполнение лабораторных работ.

III. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Основная

1) Гласс Дж. и Стенли Дж., Статистические методы в педагогике и психологии. / Пер. с англ. М., 1976
2) Сидоренко Е.В. Методы математической обработки в психологии. СПб, 1996
3) Суходольский Г.В. Основы математической статистики для психологов. Л., 1972
4) Тарасов С.Г. Основы применения математических методов в психологии. СПб, 1998

Дополнительная

1) Громыко Г.А. Статистика. МГУ, 1981
2) Елисеева И.И., Юзбашев М.М., Общая теория статистики. М., 1995
3) Наследов А.Д. Методы обработки многомерных данных в психологии. СПб., 1999
4) Суходольский Г.В. Математико-психологические модели деятельности. СПб, 1994
5) Толстова Ю.Н. Анализ социологических данных. М., 2000

IV. МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

1) Учебно-тематический план курса «Математические методы в психологии».
2) Электронный учебник по дисциплине «Математическим методам в психологии».

V. ТЕХНИЧЕСКИЕ СРЕДСТВА ОБЕСПЕЧЕНИЯ ДИСЦИПЛИНЫ

1) Персональный компьютер

Раздел 1. Основные понятия, используемые в математической обработке данных

Содержание

Данные и их разновидности. Рассматриваются виды данных в статистике. Определение количественных, качественных и порядковых данных.

Измерительные шкалы. Характеристика типов шкал, применяемых в психологии. Операции с числами, возможные с каждым типом шкал измерения. Ограничения в использовании различных типов шкал. Пример перевода данных из одного типа шкал в другой тип измерения.

Генеральная совокупность и выборка. Понятие генеральной совокупности и выборки. Свойства и параметры совокупности. Репрезентативность. Классификация выборок по способу отбора, объему, схеме испытаний и репрезентативности.

Статистические гипотезы. Понятие проблемы и гипотезы. Принципы фальсифицируемости и верифицируемости. Научная и статистическая гипотеза. Нулевая и альтернативная гипотезы.

Статистические критерии. Определение статистического критерия. Параметрические и непараметрические критерии. Уровни статистической значимости. Ошибка первого рода. Ось значимости. Мощность критериев и ошибка второго рода.

1.1. Данные и их разновидности

Данные в статистике – это основные элементы, подлежащие анализу. Данными могут быть какие-то количественные результаты, свойства, присущие определенным членам популяции, место в той или иной последовательности – любая информация, которая может быть классифицирована или разбита на категории с целью обработки.

Построение ***распределения ряда данных*** – это разделение первичных данных, полученных на выборке, на классы или категории с целью получить обобщенную упорядоченную картину, позволяющую их анализировать. Существуют три типа данных:

1. ***Количественные данные,*** получаемые при измерениях (например, данные о весе, размерах, температуре, времени, результатах тестирования и т.п.). Их можно распределить по шкале с равными интервалами.
2. ***Порядковые данные,*** соответствующие местам этих элементов в последовательности, полученной при их расположении в возрастающем порядке.
3. ***Качественные данные,*** представляющие собой какие-то свойства элементов выборки или популяции. Их нельзя измерить, и единственной их количественной оценкой служит частота встречаемости.

Из всех этих типов данных только количественные данные можно анализировать с помощью методов, в основе которых лежат *параметры* (такие, например, как средняя арифметическая, мода, дисперсия и т.д.). Но даже к количественным данным такие методы можно применить лишь в том случае, если число этих данных достаточно, чтобы проявилось нормальное распределение.

1.2. Измерительные шкалы

Поскольку психология имеет дело с психологическими процессами, то она оперирует по необходимости различными числовыми показателями, выражающими частоты, протяженности и напряженность связи между различными характеристиками. Предпосылка всех операций с количественными выражениями свойств психологических процессов и характеристик – первичное измерение качественных признаков или их квантификация. Проблема первичного измерения лишь частично математическая. Чтобы по определенным правилам приписать числа свойствам объекта психологии, надо уяснить их содержательную структуру, найти соответствие между нею и инструментом измерения. Это задачи качественно-количественного анализа. Измерению подлежат любые свойства психологических объектов: качественные и количественные. С количественными дело обстоит просто, для них уже есть общепринятые эталоны измерения (год, рубль, один человек). Качественные характеристики не имеют установленных эталонов измерения. Их приходится конструировать в соответствии с природой изучаемого объекта.

Введем некоторые определения.

Признаки и переменные – это измеряемые психологические явления. Такими явлениями могут быть время решения задачи, количество ошибок.

Значения признака определяются при помощи специальных шкал наблюдения. Психологические переменные являются случайными величинами, поскольку неизвестно заранее, какое именно значение они примут.

Измерение – это приписывание числовых форм объектам или событиям в соответствии с определенными правилами.

С. Стивенсом предложена классификация из 4 типов шкал измерения:
1) номинативная, или номинальная, или шкала наименований;
2) порядковая, или ординальная, шкала;
3) интервальная, или шкала равных интервалов;
4) шкала равных отношений.

Номинативная шкала – это шкала, классифицирующая по названию. Название же не измеряется количественно, оно лишь позволяет отличить один объект от другого или одного субъекта от другого. Номинативная шкала – это способ классификации объектов или субъектов, распределения их по ячейкам классификации.

Простейший случай номинативной шкалы – дихотомическая шкала, состоящая всего лишь из двух ячеек, например: «имеет братьев и сестер – единственный ребенок в семье»; «иностранец – соотечественник»; «проголосовал «за» – проголосовал «против»» и т.п.

Расклассифицировав все объекты, реакции или испытуемых, можно перейти, от наименований к числам, подсчитав количество наблюдений в каждом классе. Номинальная шкала позволяет подсчитывать частоты встречаемости разных наименований или значений признака и затем работать с этими частотами. Единица измерения, которой мы оперируем – это одно наблюдение.

Операции с числами для номинативной шкалы.

1) Нахождение частот распределения по пунктам шкалы с помощью процентирования или в натуральных единицах. Нетрудно подсчитать численность каждой группы и отношение этой численности к общему ряду распределения (частоты).
2) Поиск средней тенденции по модальной частоте. Модальной (Мо) называют группу с наибольшей численностью. Эти две операции дают представление о распределении психологических характеристик в количественных показателях. Его наглядность повышается отображением в диаграммах.
3) Самым сильным способом количественного анализа является установление взаимосвязи между рядами свойств, расположенных неупорядоченно. С этой целью составляют перекрестные таблицы. Помимо простой процентовки в таблицах перекрестной классификации можно подсчитать критерий сопряженности признаков по Пирсону.

Порядковая шкала – это шкала, классифицирующая по принципу «больше – меньше». Если в шкале наименований было безразлично, в каком порядке расположены классификационные ячейки, то в порядковой шкале они образуют последовательность от ячейки «самое малое значение» к ячейке «самое большое значение» (или наоборот).

Это полностью упорядоченная шкала наименований, она устанавливает отношения равенства между явлениями в каждом классе и отношения последовательности в понятиях больше, меньше между всеми без исключения классами.

Упорядоченные номинальные шкалы общеупотребимы при опросах общественного мнения. С их помощью измеряют интенсивность оценок каких-то психологических свойств, суждений, событий, степени согласия или несогласия с предложенными утверждениями. Весьма часто употребляемая разновидность шкал этого типа – ранговые. Они предполагают полное упорядочение каких-то объектов.

Операции с числами.

Интервалы в этой шкале не равны, поэтому числа обозначают лишь порядок следования признаков. И операции с числами – это операции с рангами, но не с количественным выражением свойств в каждом пункте.

1) Числа поддаются монотонным преобразованиям: их можно заменить другими с сохранением прежнего порядка. Так вместо ранжирования от 1 до 5 можно упорядочить тот же ряд в числах от 2 до10. Отношения между рангами останутся неизменными.
2) Суммарные оценки по ряду упорядоченных номинальных шкал – хороший способ измерять одно и то же свойство по набору различных индикаторов.
3) Для работы с материалом, собранным по упорядоченной шкале, можно использовать, помимо модальных показателей (Mo), поиск средней тенденции с помощью медианы (Me), найти среднее арифметической (M) и сделать оценку разброса данных с помощью дисперсии (D) и стандартного отклонения (σ).
4) Наиболее сильный показатель для таких шкал – корреляция рангов по Спирмену или по Кендаллу. Ранговые корреляции указывают на наличие или отсутствие функциональных связей в двух рядах признаков, измеренных упорядоченными шкалами.

Интервальная шкала – это шкала, классифицирующая по принципу «больше на определенное количество единиц – меньше на определенное количество единиц». Каждое из возможных значений признака отстоит от другого на равном расстоянии.

Шкала интервалов представляет собой полностью упорядоченный ряд с измеренными интервалами между пунктами, причем отсчет начинается с произвольно от выбранной величины (нет абсолютного нуля).

Операции с числами в интервальной метрической шкале богаче. Чем в номинальных шкалах.

1) Точка отсчета на шкале выбирается произвольно.
2) Все методы описательной статистики.
3) Возможности корреляционного и регрессионного анализа. Можно использовать коэффициент парной корреляции Пирсона и коэффициенты множественной корреляции, что может предсказать изменения в одной переменной в зависимости от изменений в другой или в целом ряде переменных.

Шкала равных отношений – это шкала, классифицирующая объекты или субъектов пропорционально степени выраженности измеряемого свойства. В шкалах отношений классы обозначаются числами, которые пропорциональны друг другу: 2 так относится к 4, как 4 к 8. Это предполагает наличие абсолютной нулевой точки отсчета. Считается, что в психологии примерами шкал равных отношений являются шкалы порогов абсолютной чувствительности (Стивенс С., 1960; Гайда В.К., Захаров В.П., 1982). Возможности человеческой психики столь велики, что трудно представить себе абсолютный нуль в какой-либо измеряемой психологической переменной. Абсолютная глупость и абсолютная честность – понятия скорее житейской психологии.

Возможны преобразования из одной шкалы в другую. Результаты, полученные по шкале интервалов, могут быть преобразованы в ранги или переведены в номинативную шкалу. Рассмотрим, например, первичные результаты шести испытуемых по шкале экстраверсии-интроверсии теста Айзенка (табл. 1).

Таблица 1

Испытуемые	Шкала интервалов	Шкала рангов	Номинативная шкала
А.	20	5	Э
Б.	15	4	Э
В.	22	6	Э
Г.	6	3	И
Д.	3	1	И
Е.	4	2	И

Первый столбец – имена испытуемых, второй столбец – балл за выраженность качества (реализована шкала интервалов), третий столбец – в соответствии с исходным баллом испытуемым приписаны ранги (первый ранг получает испытуемый, имеющий наименьший балл, второй ранг – испытуемый, имеющий следующий по величине балл, и т.д.), четвертый столбец – в соответствии с исходными баллами испытуемые распределены на два класса: интроверты (И) – баллы от 0 до 12, экстраверты (Э) – от 13 до 24. Отметим, что каждый раз при переходе от одной шкалы к другой теряется часть информации об испытуемых. Так, при ранжировании оказываются следующими друг за другом испытуемые Д. и Е. имеющие различие первичных оценок в один балл, и испытуемые Б. и Г., имеющие различие первичных оценок в шесть баллов. При распределении испытуемых по классам в один класс попадают сильно различающиеся по первичным оценкам испытуемые.

Мы рассмотрели различные приемы перевода качественных психологических признаков в количественные выражения. Следует отметить, что при описании психологических явлений необходимо всегда отдавать себе отчет в том, какая именно шкала используется, поскольку каждый способ обработки экспериментальных данных рассчитан на определенный тип шкал. Применение математических методов к неадекватным данным приводит к странным, а часто и ложным результатам. Квантификация сложных и далеко не однозначных психологических характеристик накладывает немало ограничений на математические операции с их измерениями. Математик работает с простыми числами, психолог обязан помнить, что в действительности скрывается за величинами, которыми он оперирует.

1) Первое ограничение – *соразмерность количественных показателей*, фиксированных разными шкалами в рамках одного исследования. Более сильная шкала отличается от слабой тем, что допускает более широкий диапазон математических операций с числами. Все, что допустимо для слабой шкалы допустимо и для более сильной, но не наоборот. Поэтому, смешение в анализе мерительных эталонов разного типа приводит к тому, что не используются возможности сильных шкал.
2) Второе ограничение *связано с формой распределения величины фиксированных описанными выше шкалами, которое предполагается нормальным.* Для нормального распределения оценки меры рассеяния совпадают: Mo=Me=M, в скошенном хвосты распределения не влияют на среднюю (M).

Таким образом необходимо внимательно изучать форму распределения с точки зрения его отклонения от нормального.

1.3. Генеральная совокупность и выборка

В математической статистике выделяют два фундаментальных понятия: ***генеральная совокупность*** и ***выборка***.

Совокупностью – называется практически счетное множество некоторых объектов или элементов, интересующих исследователя;

Свойством совокупности называется реальное или воображаемое качество, присущее некоторым всем ее элементам. Свойство может быть случайным или неслучайным.

Параметром совокупности называется свойство, которое можно квантифицировать в виде константы или переменной величины.

Простая совокупность характеризуется:

- отдельным свойством (например: все студенты России);
- отдельным параметром в виде константы или переменной (Все студенты женского пола);
- системой непересекающихся (несовместных) свойств, к примеру: Все учителя и ученики школ г. Владивостока.

Сложная совокупность характеризуется:

- системой, хотя бы частично пересекающихся свойств (Студенты психологического и математических факультетов ДВГУ, окончивших школу с золотой медалью);
- системой параметров независимых и зависимых в совокупности; при комплексном исследовании личности.

Гомогенной или однородной называется совокупность, все характеристики которой присущи каждому ее элементу;

Гетерогенной или неоднородной называется совокупность, характеристики которой сосредоточены в отдельных подмножествах элементов.

Важным параметром является ***объем*** совокупности – количество образующих ее элементов. Величина объема зависит от того, как определена сама совокупность, и какие вопросы нас конкретно интересуют. Допустим нас интересует эмоциональное состояние студента 1-го курса в период сдачи конкретного экзамена в сессию. Тогда генеральная совокупность исчерпывается в течении получаса. Если нас интересует эмоциональное состояние всех студентов 1-го курса, то совокупность будет гораздо больше, и еще больше, если взять эмоциональное состояние всех студентов 1-го курса данного вуза и т.д. Понятно, что совокупности большого объема можно исследовать только выборочным путем.

Выборкой называется некоторая часть генеральной совокупности, то, что непосредственно изучается.

Выборки классифицируются по репрезентативности, объему, способу отбора и схеме испытаний.

Репрезентативная – выборка адекватно отображающая генеральную совокупность в качественном и количественном отношениях. Выборка должна адекватно отображать генеральную совокупность, иначе результаты не совпадут с целями исследования.

Репрезентативность зависит от объема, чем больше объем, тем выборка репрезентативней.

По способу отбора.

Случайная – если элементы отбираются случайным образом. Так как большинство методов математической статистики основывается на понятии случайной выборки, то естественно выборка должна быть случайной.

Неслучайная выборка:

- *механический отбор*, когда вся совокупность делится на столько частей, сколько единиц планируется в выборке и затем из каждой части отбирается один элемент;
- *типический отбор* – совокупность делится на гомогенные части, и из каждой осуществляется случайная выборка;
- *серийный отбор* – совокупность делят на большое число разновеликих серий, затем делают выборку одной какой-либо серии;
- *комбинированный отбор* – сочетаются рассматриваемые виды отбора, на разных этапах.

По схеме испытаний – выборки могут быть *независимые* и *зависимые.*

По объему выборки делят на *малые* и *большие*. К малым относят выборки, в которых число элементов $n \le 30$. Понятие большой выборки не определено, но большой считается выборка в которой число элементов > 200 и средняя выборка удовлетворяет условию $30 \le n \le 200$. Это деление условно.

Малые выборки используются при статистическом контроле известных свойств уже изученных совокупностей.

Большие выборки используются для установки неизвестных свойств и параметров совокупности.

1.4. Статистические гипотезы

Основной задачей статистической проверки гипотез в психологических исследованиях является репрезентативное выборочное описание свойств генеральных совокупностей. Для описания значительных по объему совокупностей психических свойств, состояний, процессов требуется накопление огромного выборочного материала или проведение исследований в национальном масштабе. Поэтому задача репрезентативного описания сводится к задаче проверки однородности выборочных описаний, полученных в разных исследованиях, и к объединению однородных данных.

Для проверки однородности, необходимы:

а) однообразность статистических описаний одних и тех же психических явлений разными авторами;

б) указание на величину объектов выборок, из которых вычислялись статистические оценки параметров и функций.

Начало любого исследования – это постановка проблемы. Самые простые, наивные вопросы являются прототипами проблемы. В неизменных условиях, к которым приспосабливается человек, мир для него беспроблемен. И лишь изменчивость мира и духовная активность людей порождают проблемы.

В отличие от житейской, научная проблема формулируется в терминах определенной научной отрасли. Она должна быть операционализированной.

″Являются ли различия в агрессивности, личностном свойстве людей, генетически детерминированным признаком или зависят от влияний семейного воспитания?″ – это проблема, которая сформулирована в терминах психологии развития и может быть решена определенными средствами.

Постановка проблемы влечет за собой формулировку гипотезы. Гипотеза – это научное предположение, вытекающее из теории, которое еще не подтверждено и не опровергнуто. Научная гипотеза должна удовлетворять:

- принципам фальсифицируемости – быть опровергаемой в эксперименте; принцип фальсифицируемости абсолютен, так как опровержение теории всегда окончательно,
- принципам верифицируемости – быть подтверждаемой в эксперименте, этот принцип относителен, так как всегда есть вероятность опровержения гипотезы в следующем исследовании.

Различают ***научные*** и ***статистические*** гипотезы. Научные гипотезы формулируются как предполагаемое решение проблемы. Статистическая гипотеза – утверждение в отношении неизвестного параметра, сформулированное на языке математической статистики. Любая научная гипотеза требует перевода на язык статистики. После проведения конкретного эксперимента проверяются многочисленные статистические гипотезы, поскольку в каждом психологическом исследовании регистрируется не один, а множество поведенческих параметров. Каждый параметр характеризуется несколькими статистическими мерами: центральной тенденции, изменчивости, распределения. Можно вычислить меры связи параметров и оценить значимость этих связей.

Научные гипотезы. Экспериментальная гипотеза служит для организации эксперимента, а статистическая – для организации процедуры сравнения регистрируемых параметров. Статистическая гипотеза необходима на этапе математической интерпретации данных эмпирических исследований. Большое количество статистических гипотез необходимо для подтверждения или опровержения основной – экспериментальной гипотезы. Экспериментальная гипотеза – первична, статистическая – вторична.

Процесс выдвижения и опровержения гипотез можно считать основным и наиболее творческим этапом деятельности исследователя. Установлено, что количество и качество гипотез определяется общей креативностью (общей творческой способностью) исследователя – «генератора идей».

Гипотеза может отвергаться, но никогда не может быть окончательно принятой. Любая гипотеза открыта для последующей проверки.

Формулирование гипотез систематизирует предположения исследователя и представляет их в четком и лаконичном виде.

Статистические гипотезы. В обычном языке слово «гипотеза» означает предположение. В том же смысле оно употребляется в научном языке, используясь в основном для предположений, вызывающих сомнение. В математической статистике термин «гипотеза» означает предположение, которое не только вызывает сомнения, но и которое мы собираемся в данный момент проверить.

При построении статистической модели приходиться делать много различных допущений и предположений, и далеко не все из них мы собираемся или можем проверить.

Статистическая проверка гипотезы состоит в выяснении того, насколько совместима эта гипотеза с имеющимся результатом случайного выбора.

Определение. *Статистическая гипотеза – это предположение о распределении вероятностей, которое мы хотим проверить по имеющимся данным*. Гипотезы различают простые и сложные:

- *простая гипотеза* полностью задает распределение вероятностей;
- сложная гипотеза указывает не одно распределение, а некоторое множество распределений. Обычно это множество распределений, обладающих определенным свойством.

Статистические гипотезы подразделяются на нулевые и альтернативные.

Нулевая гипотеза - это гипотеза об отсутствии различий. Она обозначается как H_0 и называется нулевой потому, что содержит число 0: $X_1 - X_2 = 0$,
где X_1 и X_2 - сопоставляемые значения признаков.
Нулевая гипотеза - это то, что мы хотим опровергнуть, если перед нами стоит задача доказать значимость различий.
Альтернативная гипотеза - эта гипотеза о значимости различий. Она обозначается H_1. Альтернативная гипотеза - это то, что мы хотим доказать, поэтому ее иногда называют *экспериментальной гипотезой*.

Бываю задачи, когда мы хотим доказать *незначимость различий*, то есть подтвердить нулевую гипотезу. Например, если нам нужно убедиться, что разные испытуемые получают хотя и различные, но уравновешенные по трудности задания, или что экспериментальная и контрольная выборки не различаются между собой по каким-то значимым характеристикам.

Чаще всего требуется доказать *значимость различий*, ибо они более информативны для нас в поиске нового.

Проверка гипотез осуществляется с помощью критериев статистической оценки различий.

1.5. Статистические критерии

"***Статистический критерий*** – это решающее правило, обеспечивающее надежное поведение, то есть принятие истинной и отклонение ложной гипотезы с высокой вероятностью" (Суходольский Г.В.). Статистические критерии обозначают также метод расчета определенного числа и само это число.

В большинстве случаев для того, чтобы мы признали различия значимыми, необходимо, чтобы эмпирическое значение критерия превышало критическое, в некоторых критериях придерживаются противоположного правила. Эти правила оговариваются в описании каждого критерия.

В некоторых случаях расчетная формула критерия включает в себя количество наблюдений в исследуемой выборке, обозначаемое как n. В этом случае эмпирическое значение критерия одновременно является тестом для проверки статистических гипотез. По специальной таблице определяется, какому уровню статистической значимости различий соответствует данная эмпирическая величина.

В большинстве случаев, одно и то же эмпирическое значение критерия может оказаться значимым или незначимым в зависимости от количества наблюдений в выборке (n) или от так называемого количества *степеней свободы,* которое обозначается как ν.

Число степеней свободы. Число степеней свободы равно числу классов вариационного ряда минус число условий, при которых он был сформирован. К числу таких условий относятся: *объем выборки, средние и дисперсии*.

Если мы расклассифицировали наблюдения по классам какой-либо номинативной шкалы и подсчитали количество наблюдений в каждой ячейке классификации, то мы получаем так называемый частотный вариационный ряд. Единственное условие, которое соблюдается при его формировании – объем выборки n.

Допустим у нас три класса: "Умеет работать на ПК – умеет выполнять лишь определенные операции – не умеет работать".

Выборка состоит из 50 человек. Если в первом классе – 20 человек, во втором классе – 20 человек, то в третьем должны оказаться 10 человек. Мы ограничены только одним условием – объемом выборки. Мы не свободны в определении количества испытуемых в третьем классе, "свобода" простирается только на первые два класса

$$\nu = c-1 = 3-1 = 2$$

Аналогичным образом, если бы у нас была классификация из 10 разрядов или классов, то мы были бы свободны только в 9 и т.д.

Зная n и/или число степеней свободы, по специальным таблицам можно определить критические значения критерия и сопоставить с ними полученное эмпирическое значение.

Среди возможных статистических критериев выделяют: *односторонние и двусторонние, параметрические и непараметрические, более и менее мощные.*

Односторонние и двусторонние. Понятие *одностороннего* либо *двустороннего* критерия связано с формулировкой гипотез. Если "нулевая" гипотеза формулируется о равенстве ($X_1 = X_2$), то для проверки используется двусторонний критерий. Если же "нулевая" гипотеза формулируется о неравенстве, то возможны три варианта:

1) если $X_1 \neq X_2$, то используется двусторонний критерий;
2) если $X_1 > X_2$ или $X_1 < X_2$, то односторонний критерий.

Параметрические критерии – это некоторые функции от параметров совокупности, они служат для проверки гипотез об этих параметрах или для их оценивания. *Параметрические* критерии включают в формулу расчета параметры распределения, т.е. средние и дисперсии.

Непараметрические критерии – это некоторые функции от функций распределения или непосредственно от вариационного ряда наблюдавшихся значений изучаемого случайного явления. Они служат только для проверки гипотез о функциях распределения или рядах наблюдавшихся значений.

Непараметрические критерии не включают в формулу расчета параметров распределения и основанные на оперировании частотами или рангами.

И те, и другие критерии имеют свои преимущества и недостатки.

Параметрические критерии могут оказаться несколько более *мощными*, чем непараметрические, но только в том случае, *если признак измерен по интервальной шкале и нормально распределен*. Лишь с некоторой натяжкой мы можем считать данные, представленные в стандартизованных оценках, как интервальные. Кроме того, проверка распределения «на нормальность» требует достаточно сложных расчетов, результат которых заранее не известен.

Может оказаться, что распределение признака отличается от нормального, и нам так или иначе все равно придется обратиться к непараметрическим критериям.

Непараметрические критерии лишены всех этих ограничений и не требуют таких длительных и сложных расчетов. По сравнению с параметрическими критериями они ограничены лишь в одном – с их помощью невозможно оценить взаимодействие двух или более условий или факторов, влияющих на изменение признака.

Уровни статистической значимости. *Уровень значимости* – это вероятность того, что мы сочли различия существенными, а они на самом деле случайны.

Когда мы указываем, что различия достоверны на 5% уровне значимости, или при $p \leq 0{,}05$, то мы имеем ввиду, что вероятность того, что они недостоверны, составляет 0,05.

Если же мы указываем, что различия достоверны на 1% уровне значимости, или при $p \leq 0{,}01$, то имеем ввиду, что вероятность того, что они все-таки недостоверны равна 0,01.

Иначе, уровень значимости – это вероятность отклонения нулевой гипотезы, в то время как она верна.

> Ошибка, состоящая в том, что мы отклонили нулевую гипотезу, в то время как она верна, называется ошибкой 1 рода.

Вероятность такой ошибки обычно обозначается как α. Поэтому правильнее указывать уровень значимости: $\alpha \leq 0{,}05$ или $\alpha \leq 0{,}01$.

Если вероятность ошибки – это α, то вероятность правильного решения равна: 1–α. Чем меньше α, тем больше вероятность правильного решения.

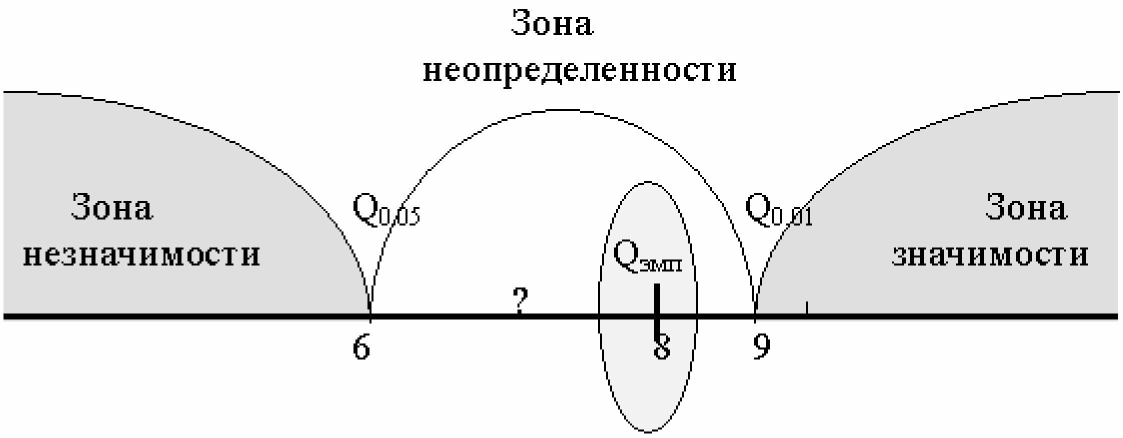
В психологии принять считать низшим уровнем статистической значимости 5%-ный уровень, а достаточным 1%-ный. В таблицах критических значений обычно приводятся значения критериев, соответствующих уровням значимости $p \leq 0{,}05$ и $p \leq 0{,}01$ иногда для $p \leq 0{,}001$. Для некоторых критериев в таблицах указан точный уровень значимости их разных эмпирических значений. Например, для значения критерия Фишера $\varphi=1{,}56$ $p=0{,}06$.

До тех пор пока уровень значимости не достигнет $p=0.05$, мы еще не имеем право отклонить нулевую гипотезу. Будем придерживаться следующего правила отклонения гипотезы об отсутствии различий (H_0) и принятии гипотезы о статистической достоверности различий (H_1).

Правило отклонения H_0 и принятия H_1
Если эмпирическое значение критерия равняется критическому значению, соответствующему $p \leq 0{,}05$ или превышает его, то H_0 отклоняется, но мы еще не можем определенно принять H_1. Если эмпирическое значение критерия равняется критическому значению, соответствующему $p \leq 0{,}01$ или превышает его, то H_0 отклоняется и принимается H_1.

Исключения: критерий знаков G, критерий T Вилкоксона и критерий U Манна-Уитни. Для них устанавливаются обратные соотношения.

Для облегчения принятия решения можно вычерчивать "ось значимости".



Критические значения критерия обозначены как $Q_{0,05}$ и $Q_{0,01}$, эмпирическое значение критерия как $Q_{эмп}$. Оно заключено в эллипс.

Вправо от критического значения $Q_{0,01}$ простирается "зона значимости" – сюда попадают эмпирические значения Q, которые ниже $Q_{0.01}$ и, следовательно, значимые.

Влево от критического значения $Q_{0.05}$ простирается "зона незначимости", – сюда попадают эмпирические значения Q, которые ниже $Q_{0,05}$ и, следовательно, незначимы.

В нашем примере, $Q_{0,05}=6$; $Q_{0,01}=9$; $Q_{эмп}=8$.

Эмпирическое значение критерия попадает в область между $Q_{0,05}$ и $Q_{0,01}$. Это "зона неопределенности": мы уже можем отклонить гипотезу о недостоверности различий (H_0), но еще не можем приять гипотезы об их достоверности (H_1).

Практически, можно считать достоверными уже те различия, которые не попадают в зону незначимости, сказав, что они достоверны при $p \leq 0,05$.

Мощность критерия. Важнейшей характеристикой любого статистического критерия является его *мощность*.

Мощность критерия – это его способность выявлять различия, если они есть. Иначе, это его способность отклонить нулевую гипотезу об отсутствии различий, если она неверна.

Ошибка, состоящая в том, что мы приняли нулевую гипотезу, в то время как она неверна, называется ошибкой II рода.

Вероятность ошибки второго рода статистического критерия обозначим как β, тогда величина 1–β будет *мощностью критерия*. Ясно, что мощность может принимать любые значения от 0 до 1. Чем ближе мощность к единице, тем эффективнее критерий.

Мощность определяется эмпирическим путем. Одни и те же задачи могут быть решены с помощью разных критериев, при этом обнаруживается, что некоторые критерии позволяют выявить различия там, где другие оказываются неспособными это сделать.

Основанием для выбора критерия может быть не только его мощность, но и другие его характеристики, а именно:

а) простота;

б) более широкий диапазон исследования (по отношению к данным, определенным по номинативной шкале, или по отношению к большим n);

в) применимость по отношению к неравным по объему выборкам;

г) большая информативность результатов.

Раздел 2. Методы описательной статистики

Содержание

Представление количественных данных. Различные этапы представления данных. Несгруппированные ряды. Упорядоченные ряды. Ранжирование данных. Распределение частот.

Числовые характеристики распределения данных. Оценка средних величин. Мода, медиана и средняя арифметическая. Оценка разброса данных. Коэффициенты вариации. Асимметрия и Эксцесс.

Изучение раздела заканчивается выполнением лабораторной работы №1 (см. Приложение 2.1. и 2.2.).

2.1. Представление количественных данных

Для анализа и интерпретации количественных данных необходимо их обобщить. Первый этап представления – это упорядочивание данных по величине от максимальной до минимальной. Такое представление называют ***несгруппированным рядом.*** В небольшом классе этого часто вполне достаточно. Рассмотрим пример.

Группа детей шестилетнего возраста была протестирована по методике Керна-Йерасика (тест на школьную зрелость). Результаты тестирования по вербальной шкале занесены в таблицу 2.1.

Таблица 2.1. Результаты тестирования детей

№ исп.	Вербальный интеллект
1	14
2	13
3	14
4	14
5	14
6	13
7	12
8	12
9	15
10	13
11	13
12	13
13	13
14	9
15	13
16	13

Оценки проставлялись в алфавитном порядке так, как записаны дети. Однако в подобной форме показатели интеллекта не слишком удобны, и мы можем лишь с трудом судить, например, о том, будет ли первый по списку ученик с показателем по вербальной шкале равным 14 обладать самым высоким или только средним уровнем интеллекта по сравнению с остальными детьми в группе. Упорядочим ряд данных по убыванию:

15, 14, 14, 14, 14, 13, 13, 13, 13, 13, 13, 13, 13, 12, 12, 9 – это несгруппированный ряд данных.

Можно ***проранжировать*** эти данные, присваивая 1 ранг наибольшему значению. Таким образом, число 15 будет иметь 1 ранг; затем следует число 14, которое повторяется 4 раза, этому числу принадлежит 4 ранга – 2, 3, 4 и 5. Общий ранг вычисляем следующим образом: (2+3+4+5)/4=3,5, т.е. складываем все ранги и делим на число повторений. Таким же образом посчитаем ранг числа 13, он будет равен: (6+7+8+9+10+11+12+13)/8=9,5, ранг числа 12 равен 14,5 и числа 9 равен 15. Запишем это в таблице 2.2.

Таблица 2.2. Ранжирование несгруппированного упорядоченного ряда

№ п/п	Вербальный Интеллект	ранг
9	15	1
1	14	3,5
3	14	3,5
4	14	3,5
5	14	3,5
2	13	9,5
6	13	9,5
10	13	9,5
11	13	9,5
12	13	9,5
13	13	9,5
15	13	9,5
16	13	9,5
7	12	14,5
8	12	14,5
14	9	16

Этот список можно сократить, классифицируя оценки по ***распределению частот***, иногда называемому просто распределением. В таблице 2.3. различные показатели вербального интеллекта размещаются по величине в данном случае от 15 до 9, а справа от каждой оценки указывается число ее повторений. Каждое число справа называется частотой и обозначается f, сумма частот обозначается n.

Таблица 2.3. Распределение частот

Сгруппированные показатели	f, частоты
15	1
14	4
13	8
12	2
9	1
n=	16

Для большого числа оценок, скажем, 100 или более может иметь смысл обобщение данных. Как правило, существует настолько широкий диапазон оценок, что целесообразнее сгруппировать их по величинам, например, в группы, объединяющие все оценки от 9 до 12 включительно, от 13 до 14 и т.д. Каждая такая группа называется *разрядом оценок.* В случае полного размещения по группам обычно говорят о *распределении сгруппированных частот.* Хотя и не существует четкого правила выбора количества разрядов, предпочтительнее образовывать не менее 12 и не более 15 разрядов. Иметь менее 12 разрядов рискованно из-за возможного искажения результатов, в то время как наличие более 15 разрядов затрудняет работу с таблицей.

2.2. Числовые характеристики распределения данных

Мы рассмотрели частотное распределение значений рассматриваемого признака. Каждое распределение может дать представление об изучаемой совокупности. Однако, этим анализ распределения данных признака не ограничивается, т.к. частотное распределение ничего не говорит о статистических закономерностях, которые описывали бы числовые характеристики изучаемой совокупности.

К характеристикам распределения, описывающим количественно его структуру и строение, относятся:

- характеристики положения;
- рассеивания;
- асимметрии и эксцесса.

Оценка центральной тенденции

К характеристикам положения относятся следующие оценки центральной тенденции: мода (Mo), медиана (Me), квантили и среднее арифметическое ($\overline{M}$).

Важное значение имеет такая величина признака, которая встречается чаще всего в изучаемом ряду, в совокупности. Такая величина называется ***модой (Mo).*** В дискретном ряду Mo определяется без вычисления, как значение признака с наибольшей частотой (например, по данным таблицы 2.1. Mo= 13).

При расчете моды может возникнуть несколько ситуаций:

1. Два значения признака, стоящие рядом, встречаются одинаково часто. В этом случае мода равна среднему арифметическому этих двух значений. Например, в следующем ряду данных: 12, 13, 14, 14, 14, 16, 16, 16, 18, 19

Mo= (14+16)/2= 15.

2. Два значения, встречаются также одинаково часто, но не стоят рядом. В этом случае говорят, что ряд данных имеет две моды, т.е. он бимодальный.

3. Если все значения данных встречаются одинаково часто, то говорят, что ряд не имеет моды.

Чаще всего встречаются ряды данных с одним модальным значением признака. Если в ряду данных встречается два или более равных значений признака, то говорят о неоднородности совокупности.

Вторая числовая характеристика ряда данных называется *медианой (Me)* – это такое значение признака, которое делит ряд пополам. Иначе, медиана обладает тем свойством, что половина всех выборочных значений признака меньше её, половина больше. При нечетном числе элементов в ряду данных, медиана равна центральному члену ряда, а при четном среднему арифметическому двух центральных значений ряда. В нашем примере (таблица 2.1.) Me=(13+13)/2=13. Вычисление медианы имеет смысл только для порядкового признака.

Среднее арифметическое **значение признака:**

$$M = \frac{\sum_i x_i}{n},$$

где x_i – значения признака, n – количество данных в рассматриваемом ряду.

Среднее арифметическое значение признака, вычисленное для какой-либо группы, интерпретируется как значение наиболее типичного для этой группы человека. Однако бывают случаи, когда подобная интерпретация несостоятельна (в случае, если существует большая разница между минимальным и максимальным значениями признака).

Квантиль – это такое значение признака, которое делит распределение в заданной пропорции: слева 0,5%, справа 99,5%; слева 2,5%, справа 97,5% и т.п. Обычно выделяют следующие разновидности квантилей:

1) Квартили Q_1, Q_2, Q_3 – они делят распределение на четыре части по 25% в каждой;
2) Квинтили K_1, K_2, K_3, K_4 – они делят распределение на пять частей по 20% в каждой;
3) Децили D_1, ...,D_9, их девять, и они делят распределение на десять частей по 10% в каждой;

4) Процентили P_1, P_2 ...,P_{99}, их девяносто девять, и они делят распределение на сто частей по 1% в каждой части.

Поскольку процентиль – наиболее мелкое деление, то все другие квантили могут быть представлены через процентили. Так, первый квартиль – это двадцать пятый процентиль, первый квинтиль – второй дециль или двадцатый процентиль, и т.п.

Характеристики рассеивания

Используя для описания ряда значений признака, только меру центральной тенденции, можно сильно ошибиться в оценке характера изучаемой совокупности. Это хорошо видно на следующем примере. Допустим, мы изучаем средний возраст в двух группах, состоящих каждая из 6-ти человек. Значения признака распределились следующим образом:

1 группа – 10, 10, 10, 50, 50, 50

2 группа – 30, 30, 30, 30, 30, 30

Подсчитав среднее значение в каждой из групп, получим $\overline{M}_1= 30$ и $\overline{M}_2=30$. Т.е. мы получили одинаковые значения, тогда как совершенно очевидно, что выборки взяты из разных совокупностей. Ошибка произошла из-за разброса значений возраста в этих группах.

Существует несколько способов оценки степени разброса или рассеивания данных. Основными характеристиками рассеивания являются: *размах (R)*, *дисперсия (D), среднеквадратическое* (стандартное) *отклонение* (σ – сигма), *коэффициент вариации(V).*

Простейший из параметров распределения, ***размах*** – это разность между максимальным и минимальным значениями признака: $R = x_{max} – x_{min}$.

Дисперсия показывает разброс значений признака относительно своего среднего арифметического значения, то есть насколько плотно значения признака группируются вокруг $\overline{M}$; чем больше разброс, тем сильнее варьируются результаты испытуемых в данной группе, тем больше индивидуальные различия между испытуемыми:

$$D = \frac{\sum_i (x_i - M)^2}{n-1} .$$

Из формулы видно, что дисперсия имеет "квадратный размер": если величина измерена в баллах, то дисперсия характеризует ее разброс в "баллах в квадрате", и т.п. Большую наглядность в отношении разброса имеет среднеквадратическое отклонение, так как его размерность соответствует размерности измеряемой величины:

$$\sigma = \sqrt{D} .$$

Коэффициент вариации вообще не имеет размерности, что позволяет сравнивать вариативность случайных величин, имеющих различную природу:

$$V = \frac{\sigma}{M} \cdot 100\% .$$

Характеристики ассиметрии и эксцесса

Мера ***асимметрии*** – коэффициент асимметрии *(As),* рассчитываемый по формуле

$$As = \frac{\sum_i (x_i - M)^3}{n\sigma^3} .$$

Асимметрия характеризует степень асимметричности распределения. Коэффициент асимметрии изменяется от минус до плюс бесконечности ($-\infty<A_s<+\infty$), для симметричных распределений As=0.

Мера ***эксцесса*** (островершинности) – коэффициент эксцесса *(E_x)*, рассчитываемый по формуле:

$$Ex = \frac{\sum_i (x_i - M)^4}{n\sigma^4} - 3 .$$

Коэффициент эксцесса также изменяется от минус до плюс бесконечности ($-\infty<E_x<+\infty$), и $E_x=0$ для нормального распределения.

Пример 2.1. У студентов первого и второго курса был исследован уровень депрессивного расстройства по методике Бэка. Сделать сравнительный анализ, используя методы описательной статистики. Результаты тестирования даны в таблице 4.1.

Таблица 4.1.

1 курс	2 курс
30	24
27	17
23	17
22	15
19	15
19	14
18	14
16	13
15	12
14	12
13	11
12	11
12	8
12	8
10	7
10	7
10	4
10	0

1. Оценка центральной тенденции

	1 курс	2 курс
мода	10	17
медиана	15	12
среднее	16	12

Медиана и среднее арифметическое значение на первом курсе выше, чем на втором, из чего можно сделать вывод, что уровень депрессивного расстройства на 1 курсе превышает уровень расстройства на втором курсе. Однако мода на 2-ом курсе значительно выше, чем на первом, т.е. преобладают более высокие значения уровня депрессивного расстройства.

2. Оценка разброса данных

	1 курс	2 курс
дисперсия	37,0	23,9
ст. отклонение	6,1	4,9
к. вариации	37,5	39,4

Дисперсия и стандартное отклонение на первом курсе выше, чем на втором, что говорит о более широком разбросе данных и следовательно, можно сделать вывод о том, что вторая выборка более однородна.

3. Асимметрия и Эксцесс

	1 курс	2 курс
ассиметрия	0,9	0,5
эксцесс	0,06	0,64

Первая выборка(1 курс) имеет небольшую положительную левостороннюю асимметрию (0,9) и незначительный положительный эксцесс (0,06). Вторая выборка (2 курс) также имеет небольшую положительную ассиметрию (0,5) и положительный эксцесс(0,64).

Приложение 2.1

Лабораторная работа № 1

Задание. Выявление центральных тенденций распределения. Оценка разброса данных и отклонения от нормального распределения.

Цель задания. Освоение расчета моды, медианы, среднего арифметического, дисперсии и стандартного отклонения системы упорядоченных событий на ПК. Оценка меры отклонения распределения от нормального на ПК.

Аппаратура. Персональный компьютер.

Математическое обеспечение. Операционная система WINDOWS и EXCEL 7.0.

Теоретическое обеспечение.

1) Система упорядоченных событий. Ранжирование.
2) Меры оценки центральной тенденции.
3) Оценка разброса данных. Дисперсия, стандартное отклонение.
4) Асимметрия и эксцесс.

Этапы обработки данных.

1) Занести данные в таблицу Excel (две выборки).
2) Упорядочить данные (по убыванию) в каждой выборке.
3) Рассчитать моду, медиану и среднее.
4) Сделать сравнительный анализ, полученных результатов.
5) Посчитать дисперсию, стандартное отклонение.
6) Посчитать коэффициент вариации.
7) Рассчитать асимметрию и эксцесс.
8) Сделать интерпретацию результатов.

Приложение 2.2

Задачи к лабораторной работе № 1

Вариант 1

При определении степени выраженности некоторого психического свойства в двух группах, опытной и контрольной, баллы распределились следующим образом:

Опытная группа – 18, 15, 16, 11, 14,15, 16,16, 20, 22, 17, 12, 11, 12, 18, 19, 20

Контрольная – 26, 8, 11, 12, 25, 22, 13, 14, 21, 20, 15, 16, 17, 16, 9, 11, 16

Дать сравнительную характеристику степени выраженности этого свойства в данных группах.

Вариант 2

При определении степени выраженности некоторого психического свойства в двух группах, опытной и контрольной, баллы распределились следующим образом:

Опытная группа – 19, 16, 17, 12, 15,16, 17,17, 21, 23, 18, 13, 12, 13, 19, 20, 21

Контрольная – 27, 9, 12, 13, 26, 23, 14, 15, 22, 21, 16, 16, 18, 17, 10, 12, 17

Дать сравнительную характеристику степени выраженности этого свойства в данных группах.

Вариант 3

При определении степени выраженности некоторого психического свойства в двух группах, опытной и контрольной, баллы распределились следующим образом:

Опытная группа – 16, 13, 14, 9, 10,13, 14,14, 18, 20, 15, 10, 9, 10, 16, 17, 18

Контрольная группа – 24, 6, 9, 10, 23, 20, 11, 12, 19, 18, 13, 14, 12, 14, 7, 9, 14

Дать сравнительную характеристику степени выраженности этого свойства в данных группах.

Вариант 4

При определении степени выраженности некоторого психического свойства в двух группах, опытной и контрольной, баллы распределились следующим образом:

Опытная группа – 15, 12, 13, 8, 11,12, 13,13, 17, 19, 14, 9, 8, 9, 15, 16, 17

Контрольная – 23, 5, 9, 9, 22, 19, 10, 11, 18, 17, 12, 13, 14, 13, 6, 8, 13

Дать сравнительную характеристику степени выраженности этого свойства в данных группах.

Вариант 5

Была исследована группа детей с заболеванием крови до лечения препаратами и после лечения. В таблицу занесены показатели L крови по результатам медицинского обследования. Сделать сравнительный анализ результативности лечения данным препаратом, используя методы описательной статистики.

Таблица. Результаты лабораторного обследования детей

№ респон.	до лечения	после лечения
	L	L
1.	20,5	2,3
2.	12,1	7,5
3.	13,6	3,8
4.	40,5	3,8
5.	9,6	4,8
6.	33	8,8
7.	77,2	13
8.	8,7	4,7
9.	3,5	3,9
10.	13,8	4,8
11.	7,4	5,7
12.	29,4	9
13.	116	13
14.	21,9	0,9

Вариант 6

Была исследована группа детей с заболеванием крови до лечения препаратами и после лечения. В таблицу занесены показатели tr крови по результатам медицинского обследования. Сделать сравнительный анализ результативности лечения данным препаратом, используя методы описательной статистики.

Таблица. Лабораторные данные обследования детей

№ респон.	до лечения	после леч.
	tr	tr
1.	280	86
2.	230	280

3.	100	30
4.	60	30
5.	90	170
6.	80	210
7.	8	230
8.	36	230
9.	50	156
10.	90	102
11.	17	161
12.	42	15
13.	42	60
14.	30	20

Вариант 7

Была исследована группа детей с заболеванием крови до лечения препаратами и после лечения. В таблицу занесены показатели нв крови по результатам медицинского обследования. Сделать сравнительный анализ результативности лечения данным препаратом, используя методы описательной статистики.

Таблица. Лабораторные данные обследования детей

№ респон.	до лечения	после лечения
	нв	нв
1.	112	82
2.	60	78
3.	84	110
4.	60	130
5.	60	130
6.	40	104
7.	76	108
8.	60	129
9.	84	110
10.	40	88
11.	112	105
12.	46	73
13.	64	85
14.	70	80

Вариант 8

Для проверки эффективности новой развивающей программы были созданы две группы детей шестилетнего возраста. На первом этапе дети обеих групп были протестированы по методике Керна-Йерасика (школьная зрелость). Результаты тестирования по невербальной шкале занесены в таблицу. Сделать сравнительный анализ школьной зрелости детей этих групп.

Таблица. Результаты тестирования по невербальной шкале (сырые баллы)

номер испыт.	1 этап	
	эксп.	контр.
1	29	34
2	31	31
3	31	28
4	25	27

5	25	30
6	19	23
7	22	21
8	20	28
9	14	29
10	16	31
11	27	17
12	24	22
13	32	21
14	27	15
15	14	33
16	24	29

Вариант 9

У участников психологического исследования, в число которых входила группа педагогов и группа не педагогов, был исследован уровень конфликтности. Полученные данные занесены в таблицу 1. и таблицу 2. Можно ли утверждать, что уровень конфликтности педагогов выше, чем у не педагогов?

Таблица 1. Педагоги

N респ.	10 конфл
1	32
2	31
3	32
7	29
8	32
9	28
10	32
12	32
13	25
16	39
17	29
18	31
19	35
20	32
22	26
25	31
26	20
30	33
31	22
32	30

Таблица 2. Не педагоги

N респ.	10 конфл
4	24
5	25
6	25
11	24
14	30
15	27
21	28
23	28
24	30
27	30
28	31
29	29
32	30
34	24
35	33
36	32
37	35

Вариант 10

У участников психологического исследования, в число которых входила группа педагогов и группа не педагогов, был исследован уровень конфликтности. Полученные данные занесены в таблицу 1. и таблицу 2. Можно ли утверждать, что уровень конфликтности педагогов выше, чем у не педагогов?

Таблица 1. Педагоги

N респ.	агрессив.

Таблица 2. Не педагоги

N респ.	Агрессив.

1	36	4	35
2	41	5	38
3	41	6	38
7	41	11	33
8	41	14	41
9	40	15	41
10	37	21	44
12	39	23	38
13	35	24	36
16	39	27	42
17	40	28	43
18	45	29	38
19	45	32	40
20	42	34	35
22	42	35	37
25	45	36	48
26	41	37	46
30	36		
31	34		
32	40		
33	39		

Вариант 11

Для проверки эффективности новой развивающей программы были созданы две группы детей шестилетнего возраста. Затем одна группа детей обучалась по обычной программе, а вторая по экспериментальной. В конце учебного года в каждой группе посчитали средний балл по успеваемости каждого ребенка. Сделать сравнительный анализ успеваемости детей этих групп.

Таблица. Средние баллы по успеваемости

номер испыт.	успеваемость уч-ся	
	экпер.	контроль
1	4,67	3,78
2	3,95	4,36
3	3,89	4,37
4	4,87	4,19
5	3,95	4,67
6	3,89	3,95
7	3,51	3,86
8	4,18	3,51
9	3,71	3,64
10	4,19	4,18
11	3,81	4,32
12	4,38	4,65
13	4,31	4,67

Вариант 12

Для проверки эффективности новой развивающей программы были созданы две группы детей шестилетнего возраста. После эксперимента дети обеих групп были протестированы по методике Керна-Йерасика (школьная зрелость). Результаты тестирования по вербальной шкале занесены в таблицу. Сделать сравнительный анализ школьной зрелости детей этих групп.

Таблица. Результаты тестирования по вербальной шкале (сырые баллы)

	эксп.	контр.
1	14	13
2	13	13
3	11	14
4	8	12
5	12	14
6	13	14
7	13	12
8	13	13
9	11	15
10	12	13
11	14	11
12	13	12
13	12	14
14	14	9
15	10	14
16	13	13

Вариант 13

Для проверки эффективности новой развивающей программы были созданы две группы детей шестилетнего возраста. После эксперимента дети обеих групп были протестированы по методике Керна-Йерасика (школьная зрелость). Результаты тестирования по невербальной шкале занесены в таблицу. Сделать сравнительный анализ школьной зрелости детей этих групп с помощью методов описательной статистики.

Таблица. Результаты тестирования по невербальной шкале (сырые баллы)

номер испыт.	3 этап	
	эксп.	контр.
1	33	39
2	33	30
3	37	38
4	33	36
5	34	31
6	33	37
7	31	35
8	29	32
9	29	39
10	35	34
11	31	30
12	29	32
13	31	36
14	34	29
15	26	39
16	26	36

Раздел 3. Нормальный закон распределения случайной величины

Содержание

Нормальный закон распределения случайной величины. В разделе дается понятие распределения признака и нормального распределения признака; основные характеристики нормального распределения.

Построение кривой нормального распределения. Дается формула для нахождения теоретических частот (m'), алгоритм построения кривой нормального распределения и рассматривается пример построения кривой.

Проверка нормальности распределения результативного признака. Даются формулы для расчета критических значений А (асимметрия) и Е (эксцесс) Пустыльника Е.И. и Плохинского Н.А..

Изучение Раздела 3 заканчивается лабораторной работой №2. Задание для выполнения лабораторной работы, тексты и таблица даны в Приложениях 3.1, 3.2 и 3.3 в конце раздела.

3.1. Нормальный закон распределения случайной величины

Распределением признака называется закономерность встречаемости разных его значений (Плохинский Н.А., 1970, с. 12).

В психологических исследованиях чаще всего ссылаются на нормальное распределение.

Нормальное распределение характеризуется тем, что крайние значения признака в нем встречаются достаточно редко, а значения, близкие к средней величине – достаточно часто. Нормальным такое распределение называется потому, что оно очень часто встречалось в естественно-научных исследованиях и казалось "нормой" всякого массового случайного проявления признаков. Это распределение следует закону, открытому тремя учеными в разное время: Муавром в 1733 г. в Англии, Гауссом в 1809 г. в Германии и Лапласом в 1812 г. во Франции (Плохинский Н.А., 1970, с.17). График нормального распределения представляет собой привычную глазу психолога-исследователя так называемую колоколообразную кривую (см. рис.3.1).

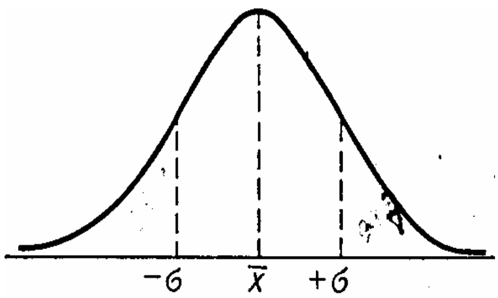


Рисунок 3.1. Кривая нормального распределения

Нормальное распределение выражается следующей формулой:

$$f_{отн} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i - M)^2}{2\sigma^2}}.$$

где $f_{отн.}$ – относительные частоты появления каждого конкретного значения случайной величины x_i. Предполагается, что переменная x_i, может принимать бесконечно большие и бесконечно малые значения, количество измерений бесконечно, а интервал квантования мал.

По этой формуле при различных значениях среднего арифметического *(М)* и стандартного отклонения (σ) получается семейство нормальных кривых.

Нормальное распределение имеет колоколообразную форму, асимптотически приближается к оси X (то есть может принимать сколь угодно малые значения по ординате при стремлении икс-значений к плюс или минус бесконечности), значения моды, медианы и среднего арифметического равны между собой.

Свойством нормальных распределений является наличие определенного количества случайной величины (случаев, испытуемых), приходящегося на интервалы между значениями σ, обычно это количество измеряют в процентах от общего числа случаев, испытуемых. Считается, что нормальное распределение характеризует такие случайные величины, на которые воздействует большое количество разнообразных факторов, причем сила воздействия одного отдельно взятого фактора значительно меньше суммы воздействий остальных факторов. В результате получается, что чаще наблюдаются некоторые средние значения измеряемого параметра, реже крайние, и чем сильнее отличается какое-то значение от среднего, тем реже оно встречается. Многие биологические параметры распределены подобным образом (рост, вес и т.п.). Психологи полагают, что большинство психологических свойств, качеств (интеллект, свойства личности и т.п.) также имеет нормальное распределение, именно из этой посылки исходят при проведении стандартизации тестовых методик.

Параметры распределения – это его числовые характеристики, указывающие, где "в среднем" располагаются значения признака, насколько эти значения изменчивы и наблюдается ли преимущественное появление определенных значений признака. Наиболее практически важными параметрами являются математическое ожидание ($\overline{M}$), дисперсия (D), стандартное отклонение (σ), показатели асимметрии и эксцесса.

В реальных психологических исследованиях мы оперируем не параметрами, а их приближенными значениями, так называемыми оценками параметров. Это объясняется ограниченностью обследованных выборок. Чем больше выборка, тем ближе может быть оценка параметра к его истинному значению. В дальнейшем, говоря о параметрах, мы будем иметь в виду их оценки.

О чем же свидетельствует стандартное отклонение? Оно позволяет сказать, что большая часть исследуемой выборки располагается в пределах σ от средней. Что это значит? Статистики показали, что при нормальном распределении «большая часть» результатов, располагающаяся в пределах одного стандартного отклонения по обе стороны от средней, в процентном отношении всегда одна и та же и не *зависит от величины стандартного отклонения:* она соответствует 68% популяции (т.е. 34% ее элементов располагается слева и 34%-справа от средней):

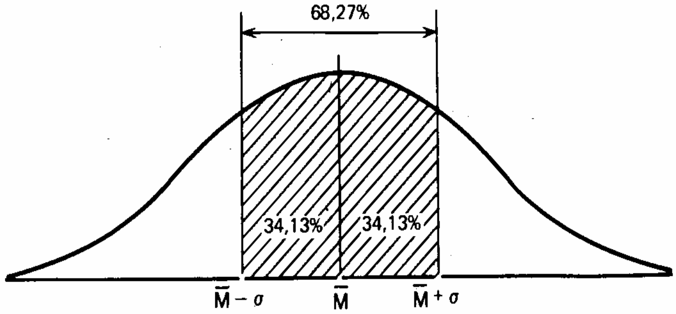


Рисунок 3.2. Кривая нормального распределения

Точно так же рассчитали, что 94,45% элементов популяции при нормальном распределении не выходит за пределы двух стандартных отклонений от средней:

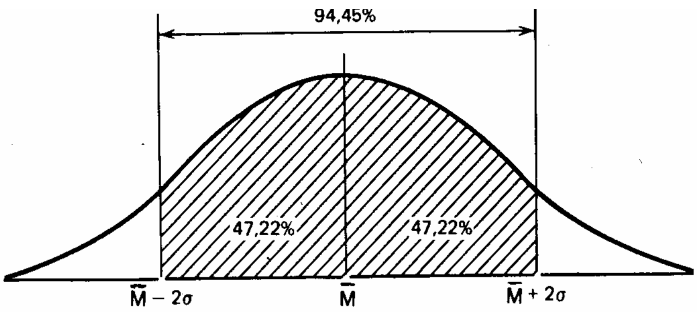


Рисунок 3.3. Кривая нормального распределения

В пределах трех стандартных отклонений умещается почти вся популяция-99,73%.

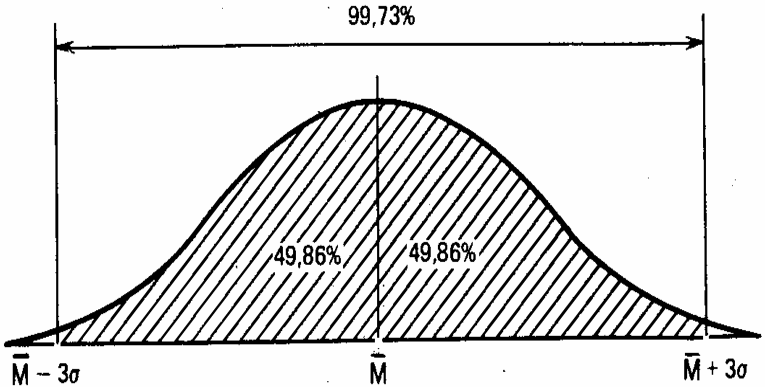


Рисунок 3.4. Кривая нормального распределения

В тех случаях, когда какие-нибудь причины благоприятствуют более частому появлению значений, которые выше или, наоборот, ниже среднего, образуются асимметричные распределения. При левосторонней, или положительной, асимметрии в распределении чаще встречаются более низкие значения признака, а при правосторонней, или отрицательной – более высокие (см. Рис. 3.5). Для симметричных распределений A=0;

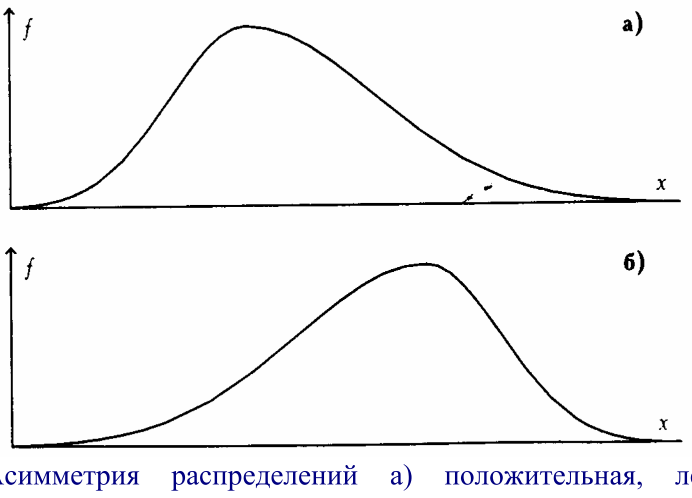


Рисунок 3.5. Асимметрия распределений а) положительная, левосторонняя, б) отрицательная, правосторонняя

В тех случаях, когда какие-либо причины способствуют преимущественному появлению средних или близких к средним значений, образуется распределение с положительным эксцессом. Если же в распределении преобладают крайние значения, причем одновременно и более низкие, и более высокие, то такое распределение характеризуется отрицательным эксцессом и в центре распределения может образоваться впадина, превращающая его в двувершинное (см. Рис. 3.6).

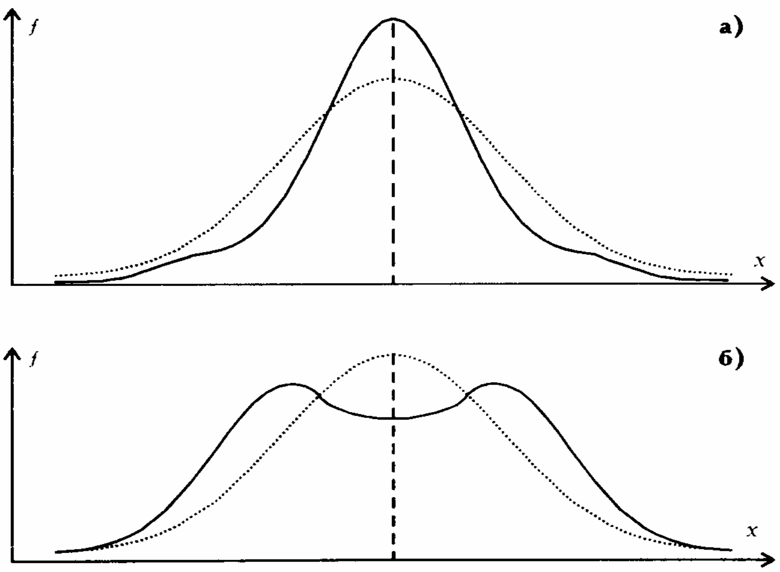


Рисунок 3.6. Эксцесс а) положительный, б) отрицательный

В распределениях с нормальной выпуклостью E=0.

Параметры распределения оказывается возможным определить только по отношению к данным, представленным, по крайней мере, в интервальной шкале. Параметры распределения не учитывают истинной психологической неравномерности секунд, миллиметров и других физических единиц измерения.

На практике психолог-исследователь может рассчитывать параметры любого распределения, если единицы, которые он использовал при измерении, признаются разумными в научном сообществе.

3.2. Построение кривой нормального распределения по эмпирическим данным

Есть несколько способов построения кривой нормального распределения по эмпирическим данным, если есть основания предположить близость данного распределения к нормальному. По одному из этих способов теоретические частоты (m') отыскиваются по формуле:

$$m' = \frac{Nh}{\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}},$$

где $\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} = \varphi(t)$—табулированная величина, отыскиваемая по отклонениям t, а Nh/σ – константа, на которую умножаются значения φ(t) и которая определяет теоретические частоты исходя из общей численности единиц совокупности и числа выделяемых групп.

Алгоритм расчета теоретических частот

1) данные сортируются по убыванию;
2) подсчитывается частота повторения каждого значения, это и есть эмпирическая частота (m);
3) рассчитывается средняя арифметическая ряда ($\bar{M}$);
4) рассчитывается среднее квадратическое отклонение (σ);
5) находится нормированное отклонение каждого варианта от средней арифметической, т. е. $t = \frac{x - \bar{M}}{\sigma}$;
6) для найденных t по таблицам определяется φ(t) (см. Приложение 3.2);
7) рассчитывается константа Nh/σ; N- объем выборки,

$$h = \frac{x_{max} - x_{min}}{1+3.322*\lg N}$$

8) каждое значение φ(t) умножается на константу.
Результаты умножения (после округления до целых чисел) будут искомыми частотами *(m')* теоретической кривой нормального распределения.

Пример. На группе из 30 добровольцев-студентов, был проделан опыт по изучению глазодвигательной координации. Задача заключалась в том, чтобы поражать предъявляемые на дисплее движущиеся мишени. Были предъявлены 10 последовательностей из 25 мишеней. Построить кривую распределения величины, отражающей количество пораженных мишеней. (Т.3.4)

Таблица 3.4. Результаты эксперимента

Исп-емые	к-во пораж. мишеней
1	12
2	21
3	10
4	15
5	15
6	19
7	17
8	14
9	13
10	11
11	20
12	15
13	15
14	14
15	17

Таблица 3.5. Поэтапный расчет теоретических частот

до воз.	m	$x_i - \bar{x}$	t	φ (t)	8,5* (t)	m'
10	1	5,2	1,34	0,1626	1,3824	1,4
11	1	4,2	1,08	0,2227	1,8933	1,9
12	1	3,2	0,82	0,285	2,423	2,4
13	1	2,2	0,57	0,3391	2,8829	2,9
14	2	1,2	0,31	0,3802	3,2323	3,2
15	4	0,2	0,05	0,3984	3,3871	3,4
17	2	-1,8	-0,46	0,3589	3,0513	3,1
19	1	-3,8	-0,98	0,2468	2,0982	2,1
20	1	-4,8	-1,24	0,1849	1,572	1,6
21	1	-5,8	-1,49	0,1315	1,118	1,1

ср.зн= 15,2 15

Дисп= 15,067 n =15 h= (x- x) / n = 2,2

ст.отк.= 3,88 lg15 = 1,18 Const= Nh/σ = 8,5

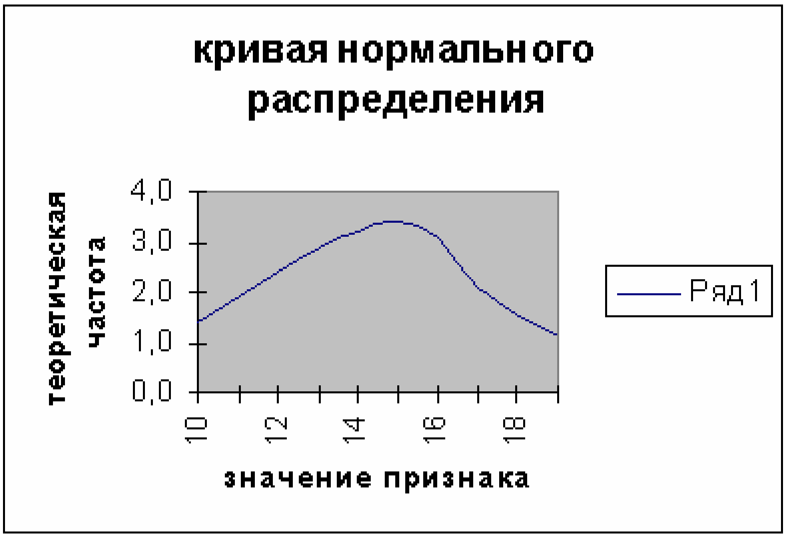


Рисунок 3.7. Кривая распределения признака

3.3. Проверка нормальности распределения результативного признака

Если мы применяем параметрические методы (к примеру, формулу для расчета коэффициента корреляции Браве-Пирсона или дисперсионный анализ) которые следует применять только тогда, когда известно или доказано, что распределение признака является нормальным (Суходольский Г.В., 1972; Шеффе Г., 1980 и др.), то в этом случае нам необходимо убедиться в нормальности распределения результативного признака. Нормальность распределения результативного признака можно проверить путем расчета показателей асимметрии и эксцесса и сопоставления их с критическими значениями (Пустыльник Е.И., 1968, Плохинский Н.А.. 1970 и др.). Рассмотрим применение метода Е.И. Пустыльника на примере.

Действовать будем по следующему алгоритму:

1) рассчитаем критические значения показателей асимметрии и эксцесса по формулам Е.И. Пустыльника и сопоставим с ними эмпирические значения;

2) если эмпирические значения показателей окажутся ниже критических, сделаем вывод о том, что распределение признака не отличается от нормального.

Расчет эмпирических показателей асимметрии и эксцесса будем производить по формулам данным ранее.

Сначала сделаем расчет промежуточных значений, который удобно выполнять поэтапно, занося данные в таблицу (Таблица 3.6.).

Таблица 3.6. Расчет промежуточных значений

№	x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$	$(x_i - \bar{x})^4$
1	11	0,94	0,884	0,831	0,781
2	13	2,94	8,644	25,412	74,712
3	12	1,94	3,764	7,301	14,165
4	9	−1,06	1,124	−1,191	1,262
5	10	−0,06	0,004	−0,000	0,000
6	11	0,94	0,884	0,831	0,781
7	8	−2,06	4,244	−8,742	18,009
8	10	−0,06	0,004	−0,000	0,000
9	15	4,94	24,404	120,554	595,536
10	14	3,94	15,524	61,163	240,982
11	8	−2,06	4,244	−8,742	18,009
12	7	−3,06	9,364	−28,653	87,677
13	10	−0,06	0,004	−0,000	0,000
14	10	−0,06	0,004	−0,000	0,000
15	5	−5,06	25,604	−129,554	655,544
16	8	−2,06	4,244	−8,742	18,009
Суммы	161		102,944	30,468	1725,467

Для расчетов в таблице, необходимо значение среднего арифметического, которое вычисляется по формуле:

$$\bar{x} = \frac{\sum x_i}{n}$$

где x_i - каждое наблюдаемое значение признака;
n - количество наблюдений.

В данном случае:

$$\bar{x} = \frac{161}{16} = 10{,}06$$

Стандартное отклонение (сигма) вычисляется по формуле:

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

где x_i - каждое наблюдаемое значение признака;
$\bar{x}$ - среднее значение (среднее арифметическое);
n - количество наблюдений.

В данном случае:

$$\sigma = \sqrt{\frac{102{,}944}{16-1}} = \sqrt{6{,}893} = 2{,}62$$

Подставляя в формулы для расчета А и Е полученные значения n, σ и соответствующие значения из таблицы, получаем:

$$A = \frac{+30{,}468}{16 \cdot 2{,}62^3} = +0{,}106$$

$$E = \frac{1725{,}467}{16 \cdot 2{,}62^4} - 3 = -0{,}711$$

Теперь рассчитаем критические значения для показателей *А* и Е по формулам Е.И. Пустыльника:

$$A_{кр} = 3 \cdot \sqrt{\frac{6 \cdot (n-1)}{(n+1) \cdot (n+3)}}$$

$$E_{кр} = 5 \cdot \sqrt{\frac{24 \cdot n \cdot (n-2) \cdot (n-3)}{(n+1)^2 \cdot (n+3) \cdot (n+5)}}$$

где n - количество наблюдений.

В данном случае:

$$A_{кр} = 3 \cdot \sqrt{\frac{6 \cdot (16-1)}{(16+1) \cdot (16+3)}} = 3 \cdot \sqrt{\frac{90}{323}} = 1{,}58$$

$$E_{кр} = 5 \cdot \sqrt{\frac{24 \cdot 16 \cdot (16-2) \cdot (16-3)}{(16+1)^2 \cdot (16+3) \cdot (16+5)}} = 5 \cdot \sqrt{\frac{69888}{115311}} = 3{,}89$$

$A_{эмп}=0{,}106$

$A_{эмп} \angle A_{кр}$

$E_{эмп}$=-0,711

$E_{эмп} \angle E_{кр}$

Так как эмпирические значения А и Е меньше критических значений, то можно сделать следующий вывод: распределение результативного признака в данном примере не отличается от нормального распределения.

Приложение 3.1

Лабораторная работа № 2

Задание. Построить график кривой распределения признака. Рассчитать Асимметрию и Эксцесс. Проверить с помощью формул Е.И. Пустыльника отклонение данного распределения от нормального. Сделать заключение.

Цель задания. Освоение построения графика нормального распределения и проверки отклонения данного распределения от нормального.

Аппаратура. Персональный компьютер.

Математическое обеспечение. Операционная система WINDOWS-95 и пакет EXCEL 7.0.

Теоретическое обеспечение.

1) Нормальное распределение.
2) Характеристики нормального распределения (асимметрия и эксцесс).
3) Построение кривой распределения признака (расчет теоретических частот).
4) Формулы Е.И. Пустыльника для расчета критических значений А и Е.

Этапы обработки данных.

1) Занести данные в таблицу 1.
2) Упорядочить данные (по убыванию) в каждой выборке.
3) Посчитать асимметрию (А), эксцесс (Е), среднее арифметическое ($\bar{x}$), стандартное отклонение (σ).
4) Посчитать эмпирическую частоту встречаемости каждого признака (m).
5) Данные занести в таблицу 2. и рассчитать теоретические частоты (m').
6) С помощью Мастера диаграмм (Excel) построить кривую.
7) Рассчитать критические значения А и Е по формулам Е.И. Пустыльника.
8) Дать выводы об отклонении данного распределения от нормального.

Приложение 3.2

Таблица 1

Значения функции $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$

Целые и десятые доли t	Сотые доли t: 0	1	2	3	4	5	6	7	8	9
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3876	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2592	2468	2444
1,0	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1926	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0241	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0303	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0,0043	0,0042	0,0040	0,0039	0,0038	0027	0,0036	0,0035	0,0034
3,1	0033	0032	0031	0030	0029	0028	0,0037	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,5	0009	0008	0008	0008	0011	0010	0010	0010	0009	0009
3,6	0006	0006	0006	0005	0008	0007	0007	0005	0007	0006
3,7	0004	0004	0004	0004	0005	0005	0005	0007	0005	0004
3,8	0003	0003	0003	0003	0004	0004	0003	0003	0003	0003
3,9	0002	0002	0002	0002	0003	0002	0002	0002	0002	0002
4,0	0,0001	0,0001	0,0001	0,0001	0002	0002	0002	0002	0001	0001
4,1	0,0001338	—	—	—	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
4,5	0,0000160	—	—	—	—	—	—	—	—	—
5,0	0,0000015	—	—	—	—	—	—	—	—	—

Приложение 3.3

Лабораторная работа №1

Вариант 1

При определении степени выраженности некоторого психического свойства в опытной группе были получены следующие результаты.

Опытная группа – 18, 15, 16, 11, 14,15, 16,16, 16, 22, 17, 12, 11, 12, 18, 19, 20

Построить кривую распределения признака и дать заключение об отклонении данного распределения от нормального.

Вариант 2

При определении степени выраженности некоторого психического свойства в контрольной группе были получены следующие результаты.

Контрольная – 14, 8, 13, 12, 25, 22, 13, 14, 21, 20, 14, 16, 17, 16, 9, 11, 16

Построить кривую распределения признака и дать заключение об отклонении данного распределения от нормального.

Вариант 3

При определении степени выраженности некоторого психического свойства в опытной группе были получены следующие результаты.

Опытная группа – 19, 16, 17, 12, 15,16, 17,17, 21, 23, 18, 13, 13, 13, 19, 20, 21

Построить кривую распределения признака и дать заключение об отклонении данного распределения от нормального.

Вариант 4

При определении степени выраженности некоторого психического свойства в контрольной группе были получены следующие результаты.

Контрольная – 27, 16, 15, 13, 23, 23, 14, 15, 22, 21, 16, 16, 18, 17, 10, 12, 17

Построить кривую распределения признака и дать заключение об отклонении данного распределения от нормального.

Вариант 5

При определении степени выраженности некоторого психического свойства в опытной группе были получены следующие результаты.

Опытная группа – 16, 13, 14, 9, 10,13, 14,14, 18, 20, 15, 10, 9, 10, 16, 17, 18

Построить кривую распределения признака и дать заключение об отклонении данного распределения от нормального.

Вариант 6

При определении степени выраженности некоторого психического свойства в контрольной группе были получены следующие результаты.

Контрольная группа – 24, 6, 9, 10, 23, 20, 11, 12, 19, 18, 13, 14, 12, 14, 7, 9, 14

Построить кривую распределения признака и дать заключение об отклонении данного распределения от нормального.

Вариант 7

При определении степени выраженности некоторого психического свойства в опытной группе были получены следующие результаты.

Опытная группа – 15, 12, 13, 8, 11,12, 13,13, 17, 19, 14, 9, 8, 9, 15, 16, 17

Построить кривую распределения признака и дать заключение об отклонении данного распределения от нормального.

Вариант 8

При определении степени выраженности некоторого психического свойства в контрольной группе были получены следующие результаты.

Контрольная – 23, 5, 9, 9, 22, 19, 10, 11, 18, 17, 13, 13, 14, 13, 6, 8, 13

Построить кривую распределения признака и дать заключение об отклонении данного распределения от нормального.

Вариант 9

По методике Цунга была исследована группа студентов факультета психологии. Измерялся уровень депрессивного состояния. Построить кривую распределения уровня депрессивного состояния у студентов-психологов. Отличается ли распределение признака от нормального?

Результаты тестирования: 53 51 49 47 46 45 44 44 42 42 42 41 41 41 41 40 40 40 39 39 39 38 38 37 36 36

Вариант 10

По методике Цунга была исследована группа студентов факультета психологии. Измерялся уровень депрессивного состояния. Построить кривую распределения уровня депрессивного состояния у студентов-психологов. Отличается ли распределение признака от нормального?

Результаты тестирования: 39 39 37 37 36 36 36 35 35 35 35 35 34 34 34 33 32 31 30 30 30 29 29 28 25

Вариант 11

По методике Цунга была исследована группа студентов не психологического факультета. Измерялся уровень депрессивного состояния. Построить кривую распределения уровня депрессивного состояния у студентов-психологов. Отличается ли распределение признака от нормального?

Результаты тестирования: 52 51 48 48 47 46 46 46 46 45 45 44 41 40 39 38 38 38 37 37 37 37 37 36 36 36 36 35 35 35 35

Вариант 12

По методике Цунга была исследована группа студентов не психологического факультета. Измерялся уровень депрессивного состояния. Построить кривую распределения уровня депрессивного состояния у студентов-психологов. Отличается ли распределение признака от нормального?

Результаты тестирования: 40 39 38 38 38 37 37 37 37 37 36 36 36 36 35 35 35 35 34 34 34 34 33 33 33 31 31 30 29 28 27 26 26 25 25

Вариант 13

Следующие данные представляют собой оценки взрослых людей в тесте на определение коэффициента интеллектуальности Стенфорда-Бине: 141 92 100 132 97 110 106 107 105 83 127 95 109 108 104 104 87 133 118 124 111 135 110 110 127 114 105 102 92 94 101 115 124 98 118

Отличается ли распределение признака от нормального?

Вариант 14

Следующие данные представляют собой оценки взрослых людей в тесте на определение коэффициента интеллектуальности Стенфорда-Бине: 138 97 101 116 112 113 95 102 131 121 130 91 92 101 146 121 108 129 113 114 106 105 102 86 107 148 96 123 107 129 108 105 123 105 139 106 89 134 103

Отличается ли распределение признака от нормального?

Раздел 4. Меры связи между признаками

Содержание

Корреляционный анализ. В разделе дается понятие корреляционного анализа; корреляционной связи и корреляционной зависимости; методы для расчета коэффициента корреляции: метод ранговой корреляции Спирмена; метод Браве-Пирсона. Интерпретация корреляции.

Изучение раздела 4 заканчивается выполнением лабораторной работы №3. Задание к лабораторной работе №3, варианты лабораторных работ и таблица критических значений коэффициента корреляции даны в Приложениях 5.1, 5.2 и 5.3 в конце раздела.

4.1. Общие положения

Анализ связей между признаками – главный вид задач, встречающийся практически в любом эмпирическом исследовании. Изучение связей между переменными, интересует исследователя не само по себе, а как отражение соответствующих причинно-следственых отношений.

При изучении *корреляций* стараются установить, существует ли какая-то связь между двумя показателями в одной выборке (например, между ростом и весом детей или между уровнем *IQ* и школьной успеваемостью) либо между двумя различными выборками (например, при сравнении пар близнецов), и если эта связь существует, то сопровождается ли увеличение одного показателя возрастанием (положительная корреляция) или уменьшением (отрицательная корреляция) другого.

Иными словами, корреляционный анализ помогает установить, можно ли предсказывать возможные значения одного показателя, зная величину другого.

Первоначальное значение термина "корреляции" – взаимная связь (Oxford Advanced Learner's Dictionary of Current English, 1982). Когда говорят о корреляции, используют термины "корреляционная связь" и "корреляционная зависимость".

Корреляционная связь – <u>это согласованные изменения двух признаков или большего количества признаков (множественная корреляционная связь).</u> Корреляционная связь отражает тот факт, что изменчивость одного признака находится в некотором соответствии с изменчивостью другого "Стохастическая" связь имеется тогда, когда каждому из значений одной случайной величины соответствует специфическое (условное) распределение вероятностей значений другой величины, и наоборот, каждому из значений этой другой величины соответствует специфическое (условное) распределение вероятностей значений первой случайной величины".

Корреляционная зависимость – <u>это изменения, которые вносят значения одного признака в вероятность появления разных значений другого признака</u>.

Стохастическая означает вероятностная. Связи между случайными явлениями называют вероятностными или стохастическими связями. Этот термин подчеркивает их отличие от детерминированных или функциональных связей в физике или математике (связь площади треугольника с его высотой и основанием, связь длины окружности с ее радиусом и т.п.). В функциональных связях каждому значению первого признака всегда соответствует (в идеальных условиях) совершенно определенное значение другого признака. В корреляционных связях каждому значению одного признака может соответствовать определенное распределение значений другого признака, но не определенное его значение.

Оба термина – корреляционная связь и корреляционная зависимость – часто используются как синонимы. Между тем, согласованные изменения признаков и отражающая это корреляционная связь между ними может свидетельствовать не о зависимости этих признаков между собой, а зависимости обоих этих признаков от какого-то третьего признака или сочетания признаков, не рассматриваемых в исследовании.

Зависимость подразумевает влияние, связь – любые согласованные изменения, которые могут объясняться сотнями причин. *Корреляционные связи не могут рассматриваться как свидетельство причинно-следственной связи, они свидетельствуют лишь о том, что изменениям одного признака, как правило, сопутствуют определенные изменения другого, но находится ли причина изменений в одном из признаков или она оказывается за пределами исследуемой пары признаков, нам неизвестно*.

Говорить в строгом смысле о зависимости мы можем только в тех случаях, когда сами оказываем какое-то контролируемое воздействие на испытуемых или так организуем исследование, что оказывается возможным точно определить интенсивность не зависящих от нас воздействий. Воздействия, которые мы можем качественно определить или даже измерить, могут рассматриваться как независимые переменные. Признаки, которые мы измеряем и которые, по нашему предположению, могут изменяться под влиянием независимых переменных, считаются зависимыми переменными. Согласованные изменения независимой и зависимой переменной действительно могут рассматриваться как зависимость.

Если в исследование включены независимые переменные, которые мы можем, по крайней мере, учитывать, например, возраст, то можно считать выявляемые между возрастом и психологическими признаками корреляционные связи корреляционными зависимостями. В большинстве же случаев нам трудно определить, что в рассматриваемой паре признаков является независимой, а что – зависимой переменной.

Корреляционные связи различаются по ***форме, направлению и степени*** (силе).

По форме корреляционная связь может быть прямолинейной или криволинейной. ***Прямолинейной*** может быть, например, связь между количеством тренировок на тренажере и количеством правильно решаемых задач в контрольной сессии. ***Криволинейной*** может быть, например, связь между уровнем мотивации и эффективностью выполнения задачи (см. Рис. 4.1). При повышении мотивации эффективность выполнения задачи сначала возрастает, затем достигается оптимальный уровень мотивации, которому соответствует максимальная эффективность выполнения задачи; дальнейшему повышению мотивации сопутствует уже снижение эффективности.

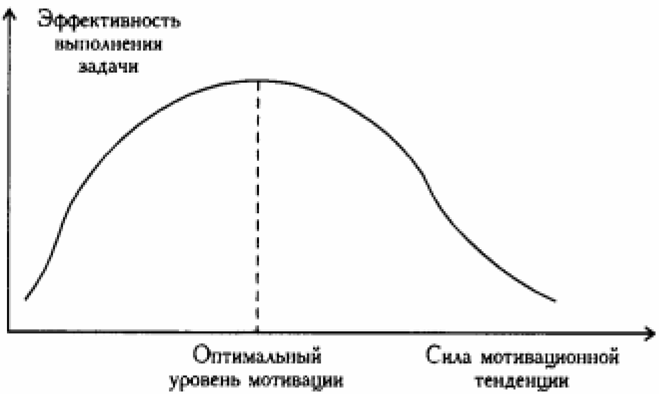


Рисунок 4.1. Связь между эффективностью решения задачи и силой мотивационной тенденции

По направлению корреляционная связь может быть положительной ("прямой") и отрицательной ("обратной"). При ***положительной*** прямолинейной корреляции более высоким значениям одного признака соответствуют более высокие значения другого, а более низким значениям одного признака – низкие значения другого (см. Рис. 4.2). При ***отрицательной*** корреляции соотношения обратные.

При положительной корреляции коэффициент корреляции имеет положительный знак, например r=+0,207, при отрицательной корреляции – отрицательный знак, например r=-0,207.

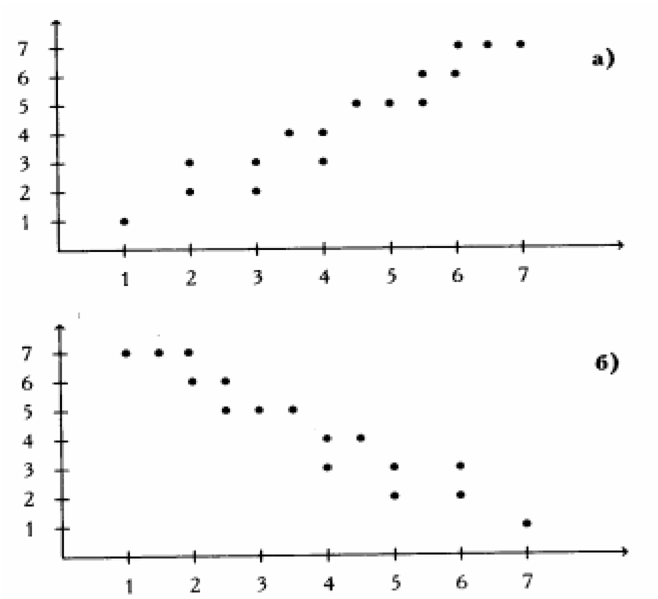
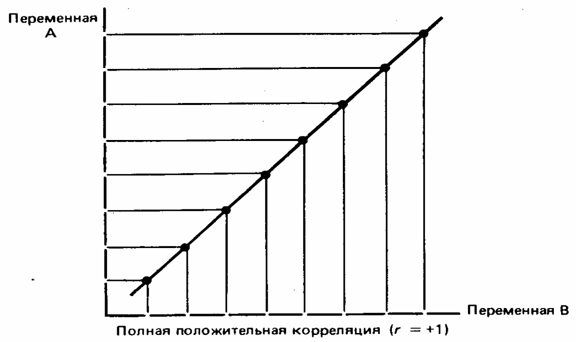


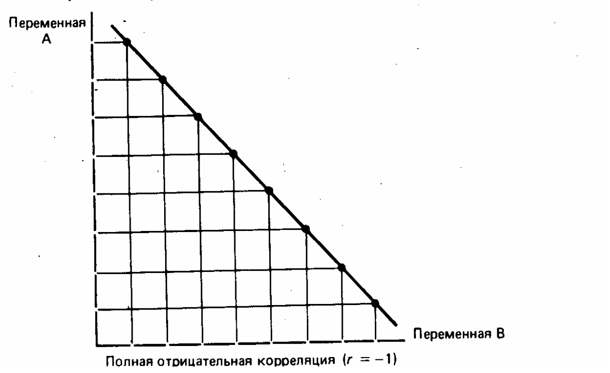
Рисунок 4.2. Схема прямолинейных корреляционных связей: а) положительная (прямая) связь, б) отрицательная (обратная) связь

Степень, сила или теснота корреляционной связи определяется по величине коэффициента корреляции.

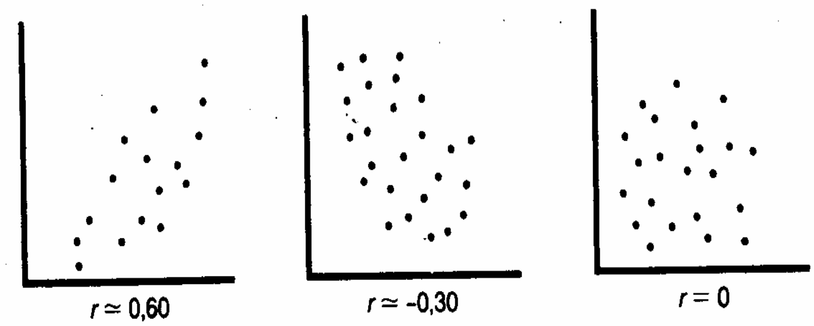
Сила связи не зависит от ее направленности и определяется по абсолютному значению коэффициента корреляции.

Коэффициент корреляции – это величина, которая может варьировать в пределах от +1 до –1. В случае полной положительной корреляции этот коэффициент равен плюс 1, а при полной отрицательной – минус 1. На графике этому соответствует прямая линия, проходящая через точки пересечения значений каждой пары данных:





В случае же если эти точки не выстраиваются по прямой линии, а образуют «облако», коэффициент корреляции по абсолютной величине становится меньше единицы и по мере округления этого облака приближается к нулю:



В случае если коэффициент корреляции равен 0, обе переменные полностью независимы друг от друга.

Используется две системы классификации корреляционных связей по их силе: общая и частная.

Общая классификация корреляционных связей:

1) сильная, или тесная при коэффициенте корреляции $r>0,70$;
2) средняя при $0,50<r<0,69$;
3) умеренная при $0,30<r<0,49$;
4) слабая при $0,20<r<0,29$;
5) очень слабая при $r<0,19$.

Частная классификация корреляционных связей:

1) высокая значимая корреляция при r, соответствующем уровню статистической значимости $\rho\leq0.01$
2) значимая корреляция при r, соответствующем уровню статистической значимости $\rho\leq0,05$;
3) тенденция достоверной связи при r, соответствующем уровню статистической значимости $\rho\leq0,10$;
4) незначимая корреляция при r, не достигающем уровня статистической значимости.

Две эти классификации не совпадают. Первая ориентирована только на величину коэффициента корреляции, а вторая определяет, какого уровня значимости достигает данная величина коэффициента корреляции при данном объеме выборки. Чем больше объем выборки, тем меньшей величины коэффициента корреляции оказывается достаточно, чтобы корреляция была признана достоверной. В результате при малом объеме выборки может оказаться так, что сильная корреляция окажется недостоверной. В то же время при больших объемах выборки даже слабая корреляция может оказаться достоверной.

Обычно принято ориентироваться на вторую классификацию, поскольку она учитывает объем выборки. Вместе с тем, необходимо помнить, что сильная, или высокая, корреляция – это корреляция с коэффициентом $r>0,70$, а не просто корреляция высокого уровня значимости.

В качестве мер корреляции используются:

1) эмпирические меры тесноты связи, многие из которых были получены еще до открытия метода корреляции, а именно:
   а) коэффициент ассоциации, или тетрахорический показатель связи;
   б) коэффициенты взаимной сопряженности Пирсона и Чупрова;
   в) коэффициент Фехнера;
   г) коэффициент корреляции рангов;
2) линейный коэффициент корреляции r;
3) корреляционное отношение η;
4) множественные коэффициенты корреляции и др.

В психологических исследованиях чаще всего применяется коэффициент линейной корреляции r – Пирсона и методы ранговой корреляции Спирмена и Кендала. Однако метод Пирсона является параметрическим и поэтому не лишен недостатков, свойственных параметрическим методам (необходимо, чтобы данные были измерены в интервальных шкалах или распределение не отличалось от нормального). Параметрическими являются также методы определения корреляционного отношения и подсчета множественных коэффициентов корреляции.

Метод ранговой корреляции Спирмена, является непараметрическим методом, он является универсальным и работает с данными измеренными в любых шкалах и прост в применении.

Уникальность метода ранговой корреляции состоит в том, что он позволяет сопоставлять не индивидуальные показатели, а индивидуальные иерархии, или профили, что недоступно ни одному из других статистических методов, включая метод линейной корреляции. Коэффициент ранговой корреляции рекомендуется применять в тех случаях, когда нам необходимо проверить, согласованно ли изменяются разные признаки у одного и того же испытуемого и насколько совпадают индивидуальные ранговые показатели у двух отдельных испытуемых или у испытуемого и группы

4.2. Коэффициент ранговой корреляции r_s Спирмена

Метод ранговой корреляции Спирмена позволяет определить тесноту (силу) и направление корреляционной связи между *двумя признаками* или *двумя профилями (иерархиями)* признаков.

Для подсчета ранговой корреляции необходимо располагать двумя рядами значений, которые могут быть проранжированы. Такими рядами значений могут быть:

1) *два признака,* измеренные в одной и той же *группе* испытуемых;

2) *две индивидуальные иерархии признаков,* выявленные у двух испытуемых по одному и тому же набору признаков;

3) две *групповые иерархии признаков,*

4) *индивидуальная и групповая* иерархии признаков.

Вначале показатели ранжируются отдельно по каждому из признаков.

Как правило, меньшему значению признака начисляется меньший ранг.

В первом случае (два признака) ранжируются индивидуальные значения по первому признаку, полученные разными испытуемыми, а затем индивидуальные значения по второму признаку.

Если два признака связаны положительно, то испытуемые, имеющие низкие ранги по одному из них, будут иметь низкие ранги и по другому, а испытуемые, имеющие высокие ранги по одному из признаков, будут иметь по другому признаку также высокие ранги. Для подсчета r_s необходимо определить разности *(d)* между рангами, полученными данным испытуемым по обоим признакам. Затем эти показатели d определенным образом преобразуются и вычитаются из 1. Чем меньше разности между рангами, тем больше будет r_s, тем ближе он будет к +1.

Если корреляция отсутствует, то все ранги будут перемешаны и между ними не будет никакого соответствия. Формула составлена так, что в этом случае r_s окажется близким к 0.

В случае отрицательной корреляции низким рангам испытуемых по одному признаку будут соответствовать высокие ранги по другому признаку, и наоборот. Чем больше несовпадение между рангами испытуемых по двум переменным, тем ближе r_s к -1.

Во втором случае (два индивидуальных профиля), ранжируются индивидуальные значения, полученные каждым из 2-х испытуемым по определенному (одинаковому для них обоих) набору признаков. Первый ранг получит признак с самым низким значением; второй ранг – признак с более высоким значением и т.д. Очевидно, что все признаки должны быть измерены в одних и тех же единицах, иначе ранжирование невозможно. Например, невозможно проранжировать показатели по личностному опроснику Кеттелла (16PF), если они выражены в "сырых" баллах, поскольку по разным факторам диапазоны значений различны: от 0 до 13, от 0 до 20 и от 0 до 26. Мы не можем сказать, какой из факторов будет занимать первое место по выраженности, пока не приведем все значения к единой шкале (чаще всего это шкала стенов).

Если индивидуальные иерархии двух испытуемых связаны положительно, то признаки, имеющие низкие ранги у одного из них, будут иметь низкие ранги и у другого, и наоборот. Например, если у одного испытуемого фактор E (доминантность) имеет самый низкий ранг, то и у другого испытуемого он должен иметь низкий ранг, если у одного испытуемого фактор C (эмоциональная устойчивость) имеет высший ранг, то и другой испытуемый должен иметь по этому фактору высокий ранг и т.д.

В третьем случае (два групповых профиля), ранжируются среднегрупповые значения, полученные в 2-х группах испытуемых по определенному, одинаковому для двух групп, набору признаков. В дальнейшем линия рассуждений такая же, как и в предыдущих двух случаях.

В случае 4-ом (индивидуальный и групповой профили), ранжируются отдельно индивидуальные значения испытуемого и среднегрупповые значения по тому же набору признаков, которые получены, как правило, при исключении этого отдельного испытуемого – он не участвует в среднегрупповом профиле, с которым будет сопоставляться его индивидуальный профиль. Ранговая корреляция позволит проверить, насколько согласованы индивидуальный и групповой профили.

Во всех четырех случаях значимость полученного коэффициента корреляции определяется по количеству ранжированных значений *N*. В первом случае это количество будет совпадать с объемом выборки n. Во втором случае количеством наблюдений будет количество признаков, составляющих иерархию. В третьем и четвертом случае N – это также количество сопоставляемых признаков, а не количество испытуемых в группах. Подробные пояснения даны в примерах. Если абсолютная величина r_s достигает критического значения или превышает его, корреляция достоверна.

Гипотезы.

Возможны два варианта гипотез. Первый относится к случаю 1, второй – к трем остальным случаям.

Первый вариант гипотез

H_0: Корреляция между переменными А и Б не отличается от нуля.

H_1: Корреляция между переменными А и Б достоверно отличается от нуля.

Второй вариант гипотез

H_0: Корреляция между иерархиями А и Б не отличается от нуля.

H_1: Корреляция между иерархиями А и Б достоверно отличается от нуля.

Ограничения коэффициента ранговой корреляции

1. По каждой переменной должно быть представлено не менее 5 наблюдений. Верхняя граница выборки определяется имеющимися таблицами критических значений.

2. Коэффициент ранговой корреляции Спирмена r_s 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. В случае, если это условие не соблюдается, необходимо вносить поправку на одинаковые ранги.

Коэффициент ранговой корреляции Спирмена подсчитывается по формуле:

$$r_s = 1 - \frac{6 \cdot \sum(d^2)}{N \cdot (N^2 - 1)}$$

где d - разность между рангами по двум переменным для каждого испытуемого;

N - количество ранжируемых значений, в данном случае количество испытуемых.

Если в обоих сопоставляемых ранговых рядах присутствуют группы одинаковых рангов, перед подсчетом коэффициента ранговой корреляции необходимо внести поправки на одинаковые ранги $T_а$ и $T_в$:

$$T_а = \Sigma\,(а^3 - а)/12,$$
$$T_в = \Sigma\,(в^3 - в)/12,$$

где *а* – объем каждой группы одинаковых рангов в ранговом ряду А, в – объем каждой группы одинаковых рангов в ранговом ряду В.

Для подсчета эмпирического значения r_s используют формулу:

$$r_s = 1 - 6 \cdot \frac{\sum d^2 + T_a + T_b}{N \cdot \left(N^2 - 1\right)}$$

Расчет коэффициента ранговой корреляции Спирмена r_s

1. Определить, какие два признака или две иерархии признаков будут участвовать в сопоставлении как переменные А и В.
2. Проранжировать значения переменной А, начисляя ранг 1 наименьшему значению, в соответствии с правилами ранжирования (см. П.2.3). Занести ранги в первый столбец таблицы по порядку номеров испытуемых или признаков.
3. Проранжировать значения переменной В, в соответствии с теми же правилами. Занести ранги во второй столбец таблицы по порядку номеров испытуемых или признаков.
4. Подсчитать разности d между рангами А и В по каждой строке таблицы и занести в третий столбец таблицы.
5. Возвести каждую разность в квадрат: d^2. Эти значения занести в четвертый столбец таблицы.
6. Подсчитать сумму квадратов Σd^2.
7. При наличии одинаковых рангов рассчитать поправки:

$T_а = \Sigma\,(a^3 - a)/12$,

$T_в = \Sigma\,(в^3 - в)/12$,

где *а* – объем каждой группы одинаковых рангов в ранговом ряду А; *в* – объем каждой группы одинаковых рангов в ранговом ряду В.

8. Рассчитать коэффициент ранговой корреляции r_s по формуле:

а) при отсутствии одинаковых рангов

$$r_s = 1 - 6 \cdot \frac{\Sigma d^2}{N(N^2 - 1)},$$

б) при наличии одинаковых рангов

$$r_s = 1 - 6 \cdot \frac{\Sigma d^2 + T_a + T_в}{N(N^2 - 1)},$$

где Σd^2 – сумма квадратов разностей между рангами; $T_а$ и $T_в$ – поправки на одинаковые ранги;

N – количество испытуемых или признаков, участвовавших в ранжировании.

9. Определить по Таблице (см. Приложение 4.3) критические значения r_s для данного N. Если r_s, превышает критическое значение или, по крайней мере, равен ему, корреляция достоверно отличается от 0.

Пример 4.1.При определении степени зависимости реакции употребления алкоголя на глазодвигательную реакцию в испытуемой группе были получены данные до употребления алкоголя и после употребления. Зависит ли реакция испытуемого от состояния опьянения?

Результаты эксперимента:

До:16, 13, 14, 9, 10, 13, 14, 14, 18, 20, 15, 10, 9, 10, 16, 17, 18.

После: 24, 9, 10, 23, 20, 11, 12, 19, 18, 13, 14, 12, 14, 7, 9, 14.

Сформулируем гипотезы:

H_0: корреляция между степенью зависимости реакции до употребления алкоголя и после не отличается от нуля.

H_1: корреляция между степенью зависимости реакции до употребления алкоголя и после достоверно отличается от нуля.

Таблица 4.1. Расчет d^2 для рангового коэффициента корреляции Спирмена r_s при сопоставлении показателей глазодвигательной реакции до эксперимента и после (N=17)

№п/п	**До**		**После**		**d**	**d^2**
	значения	**ранг**	**значения**	**ранг**		
1	16	12,5	24	17	-4,5	20,25

2	13	6,5	6	1	7,5	56,25
3	14	8,5	9	3,5	5	25
4	9	1,5	10	5	-3,5	12,25
5	10	3,5	23	16	-12,5	156,25
6	13	6,5	20	15	-8,5	72,25
7	14	8,5	11	6	2,5	6,25
8	14	8,5	12	7	1,5	2,25
9	18	15,5	19	14	1,5	2,25
10	20	17	18	13	4	16
11	15	11	13	8,5	2,5	6,25
12	10	3,5	14	10,5	-7	49
13	9	1,5	13	8,5	-7	49
14	10	3,5	14	10,5	-7	49
15	16	12,5	7	2	10,5	110,25
16	17	14	9	3,5	10,5	110,25
17	18	15,5	14	10,5	-5	25
						767,75

Так как, мы имеем повторяющиеся ранги, то в данном случае будем применять формулу с поправкой на одинаковые ранги:

$$r_s = 1 - 6 \cdot \frac{\sum d^2 + T_a + T_b}{N \cdot \left(N^2 - 1\right)}$$

$T_a = ((2^3\text{-}2)+(3^3\text{-}3)+(2^3\text{-}2)+(3^3\text{-}3)+(2^3\text{-}2)+(2^3\text{-}2))/12=\mathbf{6}$

$T_b =((2^3\text{-}2)+(2^3\text{-}2)+(3^3\text{-}3))/12=\mathbf{3}$

Найдем эмпирическое значение коэффициента Спирмена:

$r_s = 1\text{-}\ 6*((767{,}75+6+3)/(17*(17^2\text{-}1)))=\mathbf{0{,}05}$

По таблице (приложение 4.3) находим критические значения коэффициента корреляции для N=17:

$$r_{кр} = \begin{cases} 0{,}48\ (p \le 0{,}05) \\ 0{,}62\ (p \le 0{,}01) \end{cases}$$

Получаем

$\mathbf{r_s=0{,}05\angle r_{кр(0,05)}=0{,}48}$

Вывод: H_1гипотеза отвергается и принимается H_0. Т.е. корреляция между степенью зависимости реакции до употребления алкоголя и после не отличается от нуля.

4.3. Коэффициент корреляции Браве-Пирсона

Для вычисления этого коэффициента применяют следующую формулу (у разных авторов она может выглядеть по-разному):

$$r = \frac{(\Sigma x_i y_i) - n \cdot \overline{M}_1 \cdot \overline{M}_2}{(n-1) \cdot \sigma_x \cdot \sigma_x},$$

где:

$\Sigma x_i y_i$ – сумма произведений данных из каждой пары,

n – число пар,

$\overline{M}_1$ – средняя для данных переменной X,

$\overline{M}_2$ – средняя для данных переменной *Y,*

σ_x – стандартное отклонение для распределения *x,*

σ_y – стандартное отклонение для распределения *y.*

Так как, в примере 4.1 из предыдущего раздела распределение обоих признаков не отличается от нормального, то мы можем найти коэффициент линейной корреляции Браве-Пирсона и установить существует ли зависимость между приемом алкоголя и реакцией испытуемого.

Таблица 4.2. Расчет $\Sigma x_i y_i$

№ п/п	До(x_i)	После(y_i)	x_i y_i
1	16	24	384
2	13	6	78
3	14	9	126
4	9	10	90
5	10	23	230
6	13	20	260
7	14	11	154
8	14	12	168
9	18	19	342
10	20	18	360
11	15	13	195
12	10	14	140
13	9	12	108
14	10	14	140
15	16	7	112
16	17	9	153
17	18	14	252
		$\Sigma x_i y_i$=	3292

Значения $\overline{M}_1$=13,9, $\overline{M}_2$= 13,8, σ_x=3,4, σ_y=5,3, тогда числитель равен 3292 – 17*13,9*13,8=**31,06**.

В знаменателе имеем (17-1)*3,4*5,3=**290,6.**

$r_{эмп}$= 31,06/ 290,6 =**0,11**

Критические значения те же, что и в примере 4.1:

$$r_{кр} = \begin{cases} 0{,}48\ (p \le 0{,}05) \\ 0{,}62\ (p \le 0{,}01) \end{cases}$$

Имеем $\mathbf{r_{эмп}=0{,}11 \angle r_{кр(0,05)}=0{,}48}$

Вывод: H_1гипотеза отвергается и принимается H_0. Т.е. корреляционной зависимости между глазодвигательной реакцией до употребления алкоголя и после нет.

4.4. Интерпретация коэффициентов корреляции

Причинность и корреляция. Наличие корреляции двух переменных отнюдь не означает, что между ними существует причинная связь. Несмотря на то, что сосуществование (корреляцию) событий можно использовать для выявления причинных связей наряду с другими методологическими подходами, монопольное применение корреляции к анализу причинности рискованно и может вводить в заблуждение. Во-первых, даже в тех случаях, когда можно предположить существование причинной связи между двумя переменными, которые коррелированы, r сам по себе ничего не говорит о том, вызывает ли x появление y или y вызывает появление x. Во-вторых, часто наблюдаемая связь существует благодаря другим переменным, а не двум рассматриваемым. В-третьих, взаимосвязи переменных в педагогике и общественных науках почти всегда слишком сложны, чтобы их объяснением могла служить единственная причина. Успеваемость в школе – результат многочисленных влияний, да и сама по себе она является сложным понятием, которое нельзя описать адекватно при помощи какого бы то ни было одного измерения. Мы рассмотрим некоторые проблемы, возникающие при попытке выявить причинные связи с помощью корреляции. Вероятно, справедливо, что существует положительная корреляция между средним заработком преподавателей в школах и процентом выпускников, поступивших в колледж. Значит ли это, что высокооплачиваемое школьное преподавание вызывает появление лучше подготовленных абитуриентов колледжа? Увеличится ли процент выпускников, поступивших в колледж, если повысить плату преподавателям? Конечно, утвердительные ответы на эти вопросы не объяснить одной ассоциативной связью. Связь между двумя факторами не проста, кроме того, еще не упоминалась одна существенная переменная, которая характеризует финансовые и экономические условия жизни общества и определяет его возможность нести расходы как по оплате преподавателей, так и по обучению в колледжах. Наряду с этим, экономическая и финансовая обстановка отчасти зависит от интеллектуальных возможностей населения, другой переменной, вносящей вклад и в более высокую оплату педагогов и в повышенную посещаемость колледжей молодежью.

Установлено, что процент «исключенных» из школ отрицательно коррелирует с числом учебников, приходящихся на ученика в библиотеках этих школ. Но здравый смысл подсказывает нам, что нагромождение книг в библиотеке не больше повлияет на число исключенных, чем наем ленивого служащего на магическое увеличение школьной библиотеки. Если бы только здравый смысл всегда служил нам так хорошо!

Многие исследователи не останавливаются на том ложном выводе, что корреляция свидетельствует *на первый взгляд о* причинной зависимости, а выводят также и другое заключение. Они приписывают причинной связи определенное направление. Рассмотрим более внимательно правдоподобный пример. Предположим, что в большой группе учащихся коэффициент корреляции между тревожностью *(X)* и результатом теста IQ (Y) равен -0,60. Означает ли это, что большое волнение привело к тому, что учащиеся плохо выдержали испытание, а более спокойные ученики, не травмированные страхом, оказались в состоянии успешно проявить свои способности? Этот вывод склонны делать некоторые исследователи. Но разве не столь же правдоподобно считать, что сам этот тест есть фактор, вызывающий беспокойство? Не могли ли тупые ученики бояться испытания их интеллекта, а способные найти эксперимент приятным и не вызывающим беспокойства? В данном случае вопрос в том, можно ли сказать, что X вызывает Y или что Y вызывает X? Обычный коэффициент корреляции между X и Y не может дать ответ на этот вопрос. Без экспериментальной проверки связи сами по себе часто трудно интерпретировать. Искусный экспериментальный подход к той же самой задаче предполагал бы формирование группы тревожных учеников и сравнение их оценок с оценками контрольной группы.

Хотя корреляция прямо не указывает на причинную связь, она может служить ключом к разгадке причин. При благоприятных условиях на ее основе можно сформулировать гипотезы, проверяемые экспериментально, когда возможен контроль других влияний, помимо тех немногочисленных, которые подлежат исследованию. Существуют также хорошо разработанные процедуры, в частности в социологии, для вывода причин из связанных данных.

Иногда отсутствие корреляции может иметь более глубокое воздействие на нашу гипотезу о причинной связи, чем наличие сильной корреляции. Нулевая корреляция двух переменных может свидетельствовать о том, что никакого влияния одной переменной на другую не существует, при условии, что мы доверяем результатам измерений и что произведение моментов r Пирсона, измеряющее только частный тип связи, подходит для измерения более общего типа связи, называемой «причинной». Но все это мало помогает: требуются методы обнаружения причинных связей, а не методы иллюстрации беспричинных явлений.

Идентичные группы с различными средними. Существенная корреляция между двумя переменными – это факт, который в разных ситуациях можно объяснить по-разному. Некоторые корреляции – результат измерения причины и ее действия, например, когда X – пища, съеденная за месяц, а Y – вес, приобретенный за то же время. Другие корреляции возникают при измерениях двух переменных с общей причиной или влиянием, например когда X – успеваемость по английскому языку, а У – по общественным наукам. Иногда возникают иные корреляции, когда объединяются две различные группы, в каждой из которых X и У не имеют связи.

Предположим, что девочки проявляют большую тревожность, чем мальчики, при проверке, например, по шкале выраженной тревожности Тейлора. Хорошо известно, что девочки, как правило, имеют более высокие оценки по английскому языку по сравнению с мальчиками, особенно в средних классах. Диаграмма рассеивания тревоги и успеваемости по английскому для 15 мальчиков и 15 девочек могла быть подобна той, которая представлена на рис. 4.4.

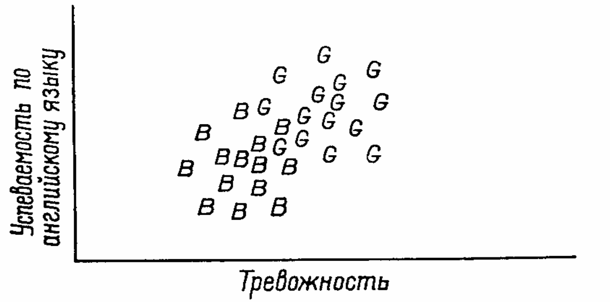


Рисунок 4.4. Диаграмма рассеивания оценок тревожности и успеваемости по английскому языку

На рис. 4.4 видна довольно сильная положительная связь между тревогой и успехами в английском, когда объединяются оценки мальчиков и девочек. Свидетельствует ли это о том, что тревожность (напряжение) заставляет учащегося усерднее трудиться и тем самым стимулирует большие достижения? Вовсе нет. Если бы это было так, то почему никому не удалось установить какую-либо связь между двумя переменными отдельно для мальчиков и девочек?

На рис. 4.4. видно, что ненулевые корреляции могут получиться в тех случаях, когда объединяются отдельные группы, например мальчики и девочки с различными средними. В результате такого объединения могут наблюдаться либо положительные, либо отрицательные связи.

Идентификация подгрупп с различными средними по *X* и У не исключает возможности корреляции *X* и У. Однако она допускает более рациональное объяснение того, почему r существенно отличается от нуля.

Нелинейность и формы маргинальных распределений переменных. Из всех способов, которыми могут быть связаны измерения двух переменных, r оценивает только один. Величина r представляет собой меру степени *линейной связи X* и У. Если *X* и У жестко линейно связаны, то точки диаграммы рассеивания будут расположены на одной прямой, как это показано в табл. 4.4. Если мы разбросаем точки на таком графике над и под прямой случайным образом и приблизительно на одинаковые расстояния, то получим различные степени линейных в своей основе связей между *X* и *Y*. Если точки на диаграмме рассеивания ориентируются – хотя и отклоняются случайным образом – относительно *кривой,* связь *X и Y* может быть существенно *криволинейной.* Из того, что r измеряет только линейную связь между *X* и *Y,* следует, что различные виды нелинейных связей *X* и *Y* могут дать такие значения r, которые подозрительно близки к нулю, если интерпретировать их без учета диаграммы рассеивания.

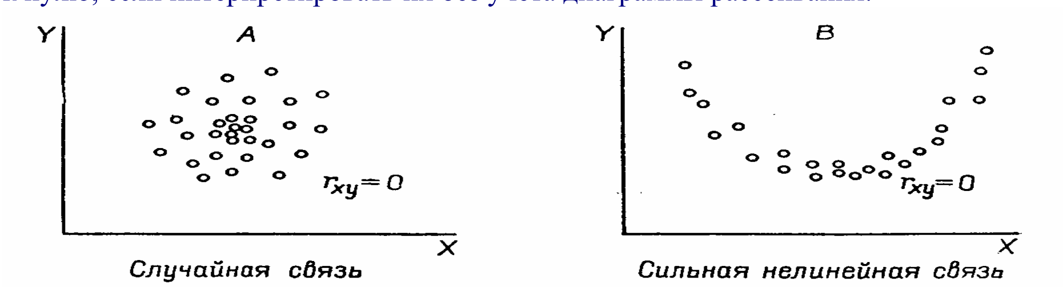


Рисунок 4.5. Два примера близкой к нулю корреляции

Если известно, что *X* и *Y,* в общем, тесно связаны линейно, то смысл r совершенно ясен. Однако если *X* и *Y* имеют некую нелинейную связь, то близкие к нулю значения r могут быть получены даже несмотря на то, что *X* и *Y* сильно связаны. Рис. 4.5. содержит две разные диаграммы рассеивания, каждая из которых имеет близкие к нулям коэффициенты корреляции.

Хотя обе диаграммы рассеивания *А и В* на рис. 4.5. имеют нулевые коэффициенты корреляции, в *В* есть существенная связь между *X* и *Y,* а в *А* нет никакой систематической связи между ними. Одной иллюстрации на рис. 4.5. по-видимому, достаточно для предупреждения против опрометчивого вывода о том, что две переменные не *связаны* только потому, что r=0. Оценки педагогических и психологических тестов часто дают «потолочные» или «подвальные» эффекты у нетипичных групп, то есть испытания могут быть слишком легкими или слишком трудными, ибо многие получают максимальную или минимальную оценку. Диаграмма рассеивания оценок теста А который характеризуется «потолочным эффектом», и теста *В с* «подвальным эффектом» могла бы быть подобна диаграмме рис. 4.6.

Величина r для данных рис. 4.6. невелика; вероятно, она приблизительно равна 0,30. Оказывается, что в области, для которой оба теста эквивалентны по трудности, они связаны более сильно. Считают, что если бы тест *А* был более трудным, а тест *В* – более легким без радикального изменения их содержания, то величина $r_{ав}$ увеличилась бы. Диаграмма рассеивания для подобных измененных тестов, возможно, обладала бы меньшей нелинейностью, чем имеющаяся. (Этот пример показывает другой важный момент: степень связи между любыми двумя переменными – независимо от того, как эта связь выражена, – зависит от характера измерения переменных.

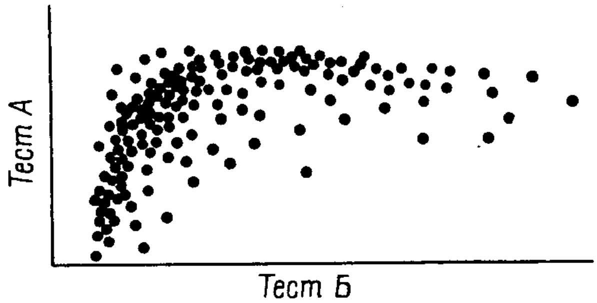


Рисунок 4.6. Диаграмма рассеивания оценок для теста А и теста В

Например, мы обычно считаем, что характеристики *веса* и *роста* довольно сильно связаны между собой у взрослых людей; но нетрудно представить себе весьма плохие способы измерения этих переменных – например, измерение с помощью субъективных суждений четырехлетних детей, оценки веса и роста которых не показали бы почти никакой корреляции).

Дополнительные замечания об интерпретации r. Кэрролл (1961) представил интересный доклад о том, как интерпретация r зависит от формы распределений *X* и У и их совместного распределения. Его статья содержит отличное изложение многих вопросов, затрагиваемых здесь лишь бегло, и отчасти будет понятна учащемуся, чье ознакомление с корреляцией не выходит за рамки этой и двух последующих глав. Он приводит следующее наблюдение одновременно над интерпретацией r и обучением студентов статистике: «Студентам недостаточно точно объясняют, что пределы [от -1 до+1] и выражения [«сильно связанный», «умеренно связанный», «не связанный»] непосредственно относятся к определенным статистическим моделям. Две наиболее часто применяемые модели – нормальная двумерная поверхность и модель линейной регрессии. Для вычисления коэффициента Пирсона не требуется никаких предположений, но интерпретация его смысла определенно зависит от области, в пределах которой данные приводятся в соответствие с подходящей статистической моделью для выполнения этой интерпретации. Поскольку реальные данные отклоняются от модели, под которую их подгоняют (например, двумерной нормальной поверхности), то пределы коэффициента корреляции могут сужаться, а предлагаемая интерпретация терять смысл».

Приложение 4.1

Лабораторная работа № 3

Задание. Расчет корреляционной связи между двумя признаками.
Цель задания. Освоение метода корреляционного анализа с помощью ПК.
Аппаратура. Персональный компьютер.
Математическое обеспечение. Операционная система WINDOWS и EXCEL 7.0.

Теоретическое обеспечение.

1) Корреляционная связь и зависимость.
2) Формулы для расчета критических значений А и Е.
3) Метод Пирсона.
4) Метод ранговой корреляции Спирмена.
5) Интерпретация результатов корреляционного анализа.

Этапы обработки данных.

1) Занести данные в таблицу Excel (две выборки).
2) Рассчитать $A_{эмп.}$, $E_{эмп.}$, $A_{кр}$ и $E_{кр}$. Сделать заключение о распределении признака в каждой выборке и отклонении его от нормального.
3) Сделать выбор метода корреляционного анализа (ранговая корреляция Спирмена либо метод Пирсона)
4) Сделать расчет по выбранной формуле.
5) Сравнить эмпирическое значение коэффициента корреляции с критическим (по таблице в Приложении 4.3.).
6) Дать интерпретацию полученных результатов.

Приложение 4.2

Лабораторная работа №3

Вариант 1

У участников психологического эксперимента был измерен уровень соперничества (по тесту Томаса) и стиль общения (по тесту Журавлева). Полученные данные занесены в таблицу 1. Можно ли утверждать, что люди склонные к соперничеству предпочитают деспотический стиль общения?

Таблица 1.

N респ.	лет	Уровень сопернич.	Деспотич. стиль общ.
1	27	7	15
2	38	7	22
3	34	3	22
7	24	2	15
8	34	3	9
9	22	5	7
10	42	2	0
12	23	5	11
13	33	2	10
16	26	4	43
17	24	4	9
18	36	8	6
19	34	5	37
20	38	4	24
22	45	5	30
25	38	11	60
26	36	4	13
30	34	3	20
31	40	4	10
32	27	1	21
33	49	9	67

Вариант 2

У участников психологического эксперимента был измерен уровень соперничества (по тесту Томаса) и стиль общения (по тесту Журавлева). Полученные данные занесены в таблицу 1. Можно ли утверждать, что люди, склонные к соперничеству предпочитают коллегиальный стиль общения?

Таблица 1.

N респ.	лет	Уровень сопернич.	Десптич стиль общ.	Коллегиал. стиль общ.	Либерал. стиль общ.
1	27	7	15	51	9
2	38	7	22	75	4
3	34	3	22	52	7
7	24	2	15	73	7
8	34	3	9	75	9
9	22	5	7	57	3
10	42	2	0	52	2
12	23	5	11	57	2
13	33	2	10	47	2
16	26	4	43	29	24
17	24	4	9	44	11
18	36	8	6	73	0
19	34	5	37	30	6
20	38	4	24	46	11
22	45	5	30	35	58
25	38	11	60	10	8
26	36	4	13	62	20
30	34	3	20	49	12
31	40	4	10	13	38
32	27	1	21	11	55
33	49	9	67	18	25

Вариант 3

У участников психологического эксперимента был измерен уровень соперничества (по тесту Томаса) и стиль общения (по тесту Журавлева). Полученные данные занесены в таблицу 1. Можно ли утверждать, что люди склонные к соперничеству предпочитают либеральный стиль общения?

Таблица 1.

N респ.	лет	Урвень соперн.	Деспотич. стиль общ.	Коллегиал. стиль общ.	Либерал. стиль общ.
1	27	7	15	51	9
2	38	7	22	75	4
3	34	3	22	52	7
7	24	2	15	73	7
8	34	3	9	75	9
9	22	5	7	57	3
10	42	2	0	52	2
12	23	5	11	57	2
13	33	2	10	47	2
16	26	4	43	29	24
17	24	4	9	44	11
18	36	8	6	73	0
19	34	5	37	30	6
20	38	4	24	46	11

22	45	5	30	35	58
25	38	11	60	10	8
26	36	4	13	62	20
30	34	3	20	49	12
31	40	4	10	13	38
32	27	1	21	11	55
33	49	9	67	18	25

Вариант 4

У участников психологического эксперимента был измерен уровень эмпатии и стиль общения (по тесту Журавлева). Полученные данные занесены в таблицу 1.Можно ли утверждать, что люди с высоким уровнем эмпатии склонны к либерализму?

Таблица 1.

N респ.	лет	Уровень эмпатии	Деспотич. стиль общ.	коллегиал. стиль общ.	либеральн. стиль общ.
1	27	32	15	51	9
2	38	51	22	75	4
3	34	54	22	52	7
7	24	56	15	73	7
8	34	47	9	75	9
9	22	56	7	57	3
10	42	69	0	52	2
12	23	55	11	57	2
13	33	69	10	47	2
16	26	46	43	29	24
17	24	53	9	44	11
18	36	62	6	73	0
19	34	55	37	30	6
20	38	53	24	46	11
22	45	57	30	35	58
25	38	43	60	10	8
26	36	53	13	62	20
30	34	60	20	49	12
31	40	50	10	13	38
32	27	43	21	11	55
33	49	38	67	18	25

Вариант 5

У участников психологического эксперимента был измерен уровень эмпатии и стиль общения (по тесту Журавлева). Полученные данные занесены в таблицу 1. Есть ли зависимость между уровнем эмпатии и деспотическим стилем общения? Если да, то какая?

Таблица 1.

N респ.	лет	Уровень эмпатии	деспотичес. стиль общ.	колегиальн. стиль общ.	либеральн. стиль общ.
1	27	32	15	51	9
2	38	51	22	75	4
3	34	54	22	52	7
7	24	56	15	73	7
8	34	47	9	75	9
9	22	56	7	57	3
10	42	69	0	52	2
12	23	55	11	57	2
13	33	69	10	47	2

16	26	46	43	29	24
17	24	53	9	44	11
18	36	62	6	73	0
19	34	55	37	30	6
20	38	53	24	46	11
22	45	57	30	35	58
25	38	43	60	10	8
26	36	53	13	62	20
30	34	60	20	49	12
31	40	50	10	13	38
32	27	43	21	11	55
33	49	38	67	18	25

Вариант 6

У участников психологического эксперимента был измерен уровень эмпатии и стиль общения (по тесту Журавлева). Полученные данные занесены в таблицу 1.Можно ли утверждать, что люди с высоким уровнем эмпатии выбирают коллегиальный стиль общения?

Таблица 1.

N респ.	лет	Уровень эмпатии	деспотичес. стиль общ.	колегиальн. стиль общ.	либеральн. стиль общ.
1	27	32	15	51	9
2	38	51	22	75	4
3	34	54	22	52	7
7	24	56	15	73	7
8	34	47	9	75	9
9	22	56	7	57	3
10	42	69	0	52	2
12	23	55	11	57	2
13	33	69	10	47	2
16	26	46	43	29	24
17	24	53	9	44	11
18	36	62	6	73	0
19	34	55	37	30	6
20	38	53	24	46	11
22	45	57	30	35	58
25	38	43	60	10	8
26	36	53	13	62	20
30	34	60	20	49	12
31	40	50	10	13	38
32	27	43	21	11	55
33	49	38	67	18	25

Вариант 7

У группы участников психологического эксперимента был измерен уровень конфликтности и уровень агрессивности. Данные занесены в таблицу 1. Можно ли утверждать, что конфликтность зависит от уровня агрессивности?

N респ.	Уровень агрессивности	Уровень конфликтности
1	36	32
2	41	31
3	41	32
4	35	24

5	38	25
6	38	25
7	41	29
8	41	32
9	40	28
10	37	32
11	33	24
12	39	32
13	35	25
14	41	30
15	41	27
16	39	3
17	40	29
18	45	31
19	45	35
20	42	32
21	44	28
22	42	26
23	38	28
24	36	30
25	45	31
26	41	20
27	42	31
28	43	30
29	38	29
30	36	33
31	34	22
32	40	30
33	39	33
34	35	24
35	37	33
36	48	32
37	46	35

Вариант 8

У группы участников психологического эксперимента был измерен уровень агрессивности и определен тип акцентуации личности (по Леонгарду). В таблицу 1 занесены данные участников с акцентуацией по застревающему типу. Можно ли утверждать, что между уровнем агрессивности и акцентуацией (по застревающему типу) существует зависимость?

Таблица 1.

N респ.	Уровень агрессивн.	Степень акцентуации
1	36	20
2	41	20
3	41	14
4	35	16
5	38	14
6	38	22
7	41	8
8	41	6
9	40	14
10	37	16
11	33	12

12	39	18
13	35	14
14	41	18
15	41	16
16	39	18
17	40	14
18	45	12
19	45	22
20	42	16
21	44	20
22	42	8
23	38	16
24	36	14
25	45	18
26	41	4
27	42	16
28	43	14
29	38	16
30	36	22
31	34	12
32	40	18
33	39	18
34	35	12
35	37	14
36	48	16
37	46	16

Вариант 9

У школьников был измерен коэффициент развития вербального интеллекта. В конце года посчитали общий балл по успеваемости. Есть ли зависимость между развитием вербального интеллекта и успеваемостью? Результаты тестирования занесены в таблицу 1.

Таблица 1.

№ респ.	Уровень вербального интеллекта	Общий балл по успеваемости
1	34	4,38
2	31	3,71
3	27	3,51
4	29	4,31
5	31	4,67
6	35	4,03
7	29	4,38
8	29	3,61
9	31	3,66
10	33	4,19
11	34	3,95
12	26	3,95
13	28	3,89
14	33	4,87
15	32	4,18
16	32	3,89

Вариант 10

У школьников был измерен коэффициент развития вербального интеллекта. В конце года посчитали общий балл по успеваемости. Есть ли зависимость между развитием вербального интеллекта и успеваемостью? Результаты тестирования занесены в таблицу 1.

Таблица 1.

№ респ.	Уровень вербального интеллекта	Общий балл по успеваемости
1	35	4,32
2	39	4,65
3	29	3,78
4	36	4,18
5	31	3,95
6	30	3,64
7	34	4,36
8	39	4,67
9	32	4,87
10	35	4,37
11	37	4,89
12	31	4,19
13	24	3,51
14	36	4,32
15	38	4,67
16	30	3,86

Вариант 11

В начале учебного года у школьников был измерен коэффициент развития невербального интеллекта. В конце года посчитали общий балл по успеваемости. Есть ли зависимость между развитием вербального интеллекта и успеваемостью? Результаты тестирования занесены в таблицу 1.

Таблица 1.

№ респ.	Уровень невербального интеллекта	Общий балл по успеваемости
1	14	4,38
2	12	3,71
3	12	3,51
4	13	4,31
5	14	4,67
6	12	4,03
7	11	4,38
8	13	3,61
9	13	3,66
10	13	4,19
11	12	3,95
12	13	3,95
13	12	3,89
14	8	4,87
15	13	4,18
16	11	3,89

Вариант 12

В начале учебного года у школьников был измерен коэффициент развития невербального интеллекта. В конце года посчитали общий балл по успеваемости. Есть ли зависимость между развитием вербального интеллекта и успеваемостью? Результаты тестирования занесены в таблицу 1.

Таблица 1.

№ респ.	Уровень невербального интеллекта	Общий балл по успеваемости
1	13	4,32
2	14	4,65
3	9	3,78
4	14	4,18
5	12	3,95
6	11	3,64
7	13	4,36
8	15	4,67
9	13	4,87
10	12	4,37
11	14	4,89
12	14	4,19
13	14	3,51
14	12	4,32
15	14	4,67
16	13	3,86

Вариант 13

У школьников был измерен коэффициент развития вербального интеллекта. В конце года посчитали коэффициент адаптации ребенка к школе. Есть ли зависимость между развитием вербального интеллекта и адаптацией ребенка к школе? Результаты тестирования занесены в таблицу 1.

Таблица 1.

№ респ.	Уровень вербального интеллекта	Коэффициент адаптации
1	34	1710
2	31	45,8
3	27	45,8
4	29	7,1
5	31	28,5
6	35	15,8
7	29	11,4
8	29	31,4
9	31	28,5
10	33	24,3
11	34	42,8
12	26	60
13	28	40
14	33	2,1
15	32	34,2
16	32	38,2

Вариант 14

У школьников был измерен коэффициент развития вербального интеллекта. В конце года посчитали коэффициент адаптации ребенка к школе. Есть ли зависимость между развитием вербального интеллекта и адаптацией ребенка к школе? Результаты тестирования занесены в таблицу 1.

Таблица 1.

№ респ.	Уровень вербального интеллекта	Коэффициент адаптации
1	35	8,1
2	39	4,2
3	29	41,4
4	36	8,5
5	31	24,3
6	30	34,8
7	34	15,7
8	39	7,1
9	32	20
10	35	11,4
11	37	5,7
12	31	24,3
13	24	50
14	36	8,5
15	38	1
16	30	12,8

Вариант 15

У школьников был измерен коэффициент развития невербального интеллекта. В конце года посчитали коэффициент адаптации ребенка к школе. Есть ли зависимость между развитием вербального интеллекта и адаптацией ребенка к школе? Результаты тестирования занесены в таблицу 1.

Таблица 1.

№ респ.	Уровень невербального интеллекта	Коэффициент адаптации
1	14	1710
2	12	45,8
3	12	45,8
4	13	7,1
5	14	28,5
6	12	15,8
7	11	11,4
8	13	31,4
9	13	28,5
10	13	24,3
11	12	42,8
12	13	60
13	12	40
14	8	2,1
15	13	34,2
16	11	38,2

Вариант 16

У школьников был измерен коэффициент развития невербального интеллекта. В конце года посчитали коэффициент адаптации ребенка к школе. Есть ли зависимость между развитием вербального интеллекта и адаптацией ребенка к школе? Результаты тестирования занесены в таблицу.

Таблица.

№ респ.	Уровень невербального интеллекта	Коэффициент адаптации
1	13	8,1
2	14	4,2
3	9	41,4
4	14	8,5
5	12	24,3
6	11	34,8
7	13	15,7
8	15	7,1
9	13	20
10	12	11,4
11	14	5,7
12	14	24,3
13	14	50
14	12	8,5
15	14	1
16	13	12,8

Приложение 4.3

Таблица 4.1. Критические значения коэффициентов линейной корреляции Пирсона и ранговой корреляции Спирмена

n	α				n	α			
	0,10	0,05	0,01	0,001		0,10	0,05	0,01	0,001
5	0,805	0,878	0,959	0,991	46	0,246	0,291	0,376	0,469
6	0,729	0,811	0,917	0,974	47	0,243	0,288	0,372	0,465
7	0,669	0,754	0,875	0,951	48	0,240	0,285	0,368	0,460
8	0,621	0,707	0,834	0,925	49	0,238	0,282	0,365	0,456
9	0,582	0,666	0,798	0,898	50	0,235	0,279	0,361	0,451
10	0,549	0,632	0,765	0,872	51	0,233	0,276	0,358	0,447
11	0,521	0,602	0,735	0,847	52	0,231	0,273	0,354	0,443
12	0,497	0,576	0,708	0,823	53	0,228	0,271	0,351	0,439
13	0,476	0,553	0,684	0,801	54	0,226	0,268	0,348	0,435
14	0,458	0,532	0,661	0,780	55	0,224	0,266	0,345	0,432
15	0,441	0,514	0,641	0,760	56	0,222	0,263	0,341	0,428
16	0,426	0,497	0,623	0,742	57	0,220	0,261	0,339	0,424
17	0,412	0,482	0,606	0,725	58	0,218	0,259	0,336	0,421
18	0,400	0,468	0,590	0,708	59	0,216	0,256	0,333	0,418
19	0,389	0,456	0,575	0,693	60	0,214	0,254	0,330	0,414
20	0,378	0,444	0,561	0,679	61	0,213	0,252	0,327	0,411
21	0,369	0,433	0,549	0,665	62	0,211	0,250	0,325	0,408
22	0,360	0,423	0,537	0,652	63	0,209	0,248	0,322	0,405
23	0,352	0,413	0,526	0,640	64	0,207	0,246	0,320	0,402
24	0,344	0,404	0,515	0,629	65	0,206	0,244	0,317	0,399
25	0,337	0,396	0,505	0,618	66	0,204	0,242	0,315	0,396
26	0,330	0,388	0,496	0,607	67	0,203	0,240	0,313	0,393
27	0,323	0,381	0,487	0,597	68	0,201	0,239	0,310	0,390
28	0,317	0,374	0,479	0,588	69	0,200	0,237	0,308	0,388
29	0,311	0,367	0,471	0,579	70	0,198	0,235	0,306	0,385
30	0,306	0,361	0,463	0,570	80	0,185	0,220	0,286	0,361
31	0,301	0,355	0,456	0,562	90	0,174	0,207	0,270	0,341
32	0,296	0,349	0,449	0,554	100	0,165	0,197	0,256	0,324
33	0,291	0,344	0,442	0,547	110	0,158	0,187	0,245	0,310
34	0,287	0,339	0,436	0,539	120	0,151	0,179	0,234	0,297
35	0,283	0,334	0,430	0,532	130	0,145	0,172	0,225	0,285
36	0,279	0,329	0,424	0,525	140	0,140	0,166	0,217	0,275
37	0,275	0,325	0,418	0,519	150	0,135	0,160	0,210	0,266
38	0,271	0,320	0,413	0,513	200	0,117	0,139	0,182	0,231
39	0,267	0,316	0,408	0,507	250	0,104	0,124	0,163	0,207
40	0,264	0,312	0,403	0,501	300	0,095	0,113	0,149	0,189
41	0,260	0,308	0,398	0,495	350	0,088	0,105	0,138	0,175
42	0,257	0,304	0,393	0,490	400	0,082	0,098	0,129	0,164
43	0,254	0,301	0,389	0,484	450	0,078	0,092	0,121	0,155
44	0,251	0,297	0,384	0,479	500	0,074	0,088	0,115	0,147
45	0,248	0,294	0,380	0,474	600	0,067	0,080	0,105	0,134

Раздел 5. Методы проверки статистических гипотез

Содержание

В данном разделе содержится описание и применение статистических критериев:

t – критерий Стьюдента, используется для установления сходства-различия средних арифметических значений в двух выборках или в более общем виде, для установления сходства-различия двух эмпирических распределений;

F – критерий Фишера, используется для установления сходства-различия дисперсий в двух независимых выборках;

Q – критерий Розенбаума, используется для оценки *различий между двумя* выборками *по уровню* какого-либо признака, количественно измеренного.

T – критерий Вилкоксона, применяется для сопоставления показателей, измеренных в *двух* разных условиях на *одной* и той же выборке испытуемых. Он позволяет установить *направленность* изменений, *и* их *выраженность*.

χ^2-критерий Пирсона, используется:

1) для сопоставления *эмпирического* распределения признака с *теоретическим* – равномерным, нормальным или каким-то иным;

2) для сопоставления *двух, трех или более эмпирических* распределений одного и того же признака.

Для каждого критерия в разделе рассматривается соответствующий пример на его применение.

Изучение раздела 5 заканчивается выполнением лабораторной работы №4. Задание к лабораторной работе №4 и варианты лабораторных работ даны в Приложениях 5.1. и 5.2. в конце раздела. В Приложении 5.3. даны таблицы для критических значений статистических критериев.

5.1. t-критерий Стьюдента

t-Критерий Стьютдента используется для:

1) установления сходства-различия средних арифметических значений в двух выборках ($\overline{M}_1 \leftrightarrow \overline{M}_2$) или в более общем виде, для установления сходства-различия двух эмпирических распределений;

2) установления отличия от нуля некоторых мер связи: коэффициента линейной корреляции Пирсона, ранговой корреляции Спирмена, точечно-бисериальной и рангово-бисериальной корреляции (r_{xy}, r_s, $r_{pb} \leftrightarrow$"0") и коэффициента линейной регрессии ($R_{xy} \leftrightarrow$ "O"):

3) установления сходства-различия двух дисперсий в двух зависимых выборках.

Ограничения:

1) это параметрический критерий, поэтому необходимо, чтобы распределение признака, по крайней мере, не отличалось от нормального распределения;
2) для независимых и зависимых выборок разные формулы расчета;

Гипотезы

1) независимые выборки:
   H_0: средние значения признака в обоих выборках не различаются,
   H_1: средние значения признака в обоих выборках статистически значимо различаются.
2) зависимые выборки:
   H_0: разности оценок испытуемых в двух состояниях не отличаются от нуля,
   H_1: разности оценок испытуемых в двух состояниях статистически значимо отличаются от нуля.

Рассмотрим случай 1.

Пример 5.1.(независимые выборки). Предположим, имеется две независимые выборки школьников, интеллект которых развивали в течение некоторого времени по двум различным методикам, требуется установить, какая из методик лучше (табл.5.1). Предварительно было выяснено, что начальный уровень интеллекта был одинаковым в обеих выборках. Задача сравнения двух методик может быть переформулирована на язык статистики как задача сравнения средних арифметических значений интеллекта в обеих выборках.

Таблица 5.1.

Числовые характеристики	1-я выборка	2-я выборка
n	30	32
M	103	110
σ	10	12

Гипотезы:

H_0: средние значения уровня интеллекта в обоих выборках не различаются,

H_1: средние значения уровня интеллекта в обоих выборках статистически значимо различаются.

В данном случае для получения эмпирического значения *t*-критерия используется следующая формула:

$$t = \frac{|M_1 - M_2|}{\sqrt{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}},$$

где: n_1, n_2 – количество испытуемых в 1-й и 2-й выборках; $\overline{M}_1, \overline{M}_2$ – средние арифметические значения в 1-й и 2-й выборках; σ_1, σ_2 – стандартные отклонения в 1-й и 2-й выборках.

Количество степеней свободы для нахождения критического значения критерия:

$$Df = n_1+n_2-2.$$

(В рассматриваемых примерах критические значения t-критерия приводятся для ненаправленных гипотез).

Тогда:

$$t = \frac{|103-110|}{\sqrt{29\cdot 10^2 + 31\cdot 12^2}} \cdot \sqrt{\frac{30\cdot 32\cdot (30+32-2)}{30+32}} \approx 2{,}486 .$$

Таким образом, получаем $t_{эмп}$=2,486

Критические значения *t-критерия* находим по таблице 1 (приложение 5.3.) для df=30+32-2=60.

$$t_{кр} = \begin{cases} 2{,}0 \quad \text{для } p \le 0{,}05 \\ 2{,}66 \text{ для } p \le 0{,}01 \end{cases}$$

Полученное эмпирическое значение t-критерия превышает критическое для α=*0,05*, но оказывается меньше критического для α=0.01, т.е.

$2{,}0 < T_{кр} = 2{,}486 < 2{,}66$

Вывод: H_0 гипотеза отклоняется и можно сделать вывод о статистически значимом различии средних арифметических значений в двух выборках для $\rho \le 0.05$ и о преимуществах второй методики по сравнению с первой.

Строгое использование t-критерия предполагает, что обе выборки извлечены из нормальных совокупностей. Однако многие авторы не считают это условие достаточно жестким, указывая на возможность использования t-критерия в ситуациях, когда нет серьезных оснований сомневаться в нормальности распределения признака в генеральной совокупности, даже если это нельзя подтвердить статистически.

При зависимых выборках возникает корреляция результатов, поскольку измерения проводятся на одних и тех же испытуемых в различных условиях (x и y)', чтобы учесть влияние корреляции, применяется другая формула:

$$t = \frac{\left(\sum d_i\right)}{\sqrt{\sum d_i^2 - \left(\sum d_i\right)^2 / n}} \sqrt{\frac{n-1}{n}},$$

где $d_i = x_i - y_i$, то есть разность значений признака для каждого испытуемого. Количество степеней свободы df=n–1. Проверяется статистическая гипотеза о соответствии распределения разностей t-распределению Стьюдента с нулевым средним значением.

Пример 5.2. (зависимые выборки). Допустим, проводится измерение ситуативной тревожности до и после психотерапевтического воздействия с помощью некоторого опросника (табл.5.2). Исследователя интересует вопрос, приводит ли воздействие к изменению уровня тревожности.

Гипотезы:

H_0: разности оценок у испытуемых ситуативной тревожности до и после психотерапевтического воздействия не отличаются от нуля,

H_1: разности оценок у испытуемых ситуативной тревожности до и после психотерапевтического воздействия статистически значимо отличаются от нуля

Таблица 5.2.

Испытуемые	"до" (x_i)	"после"(y_i)	$d_i = x_i - y_i$	$(d_i)^2 = (x_i - y_i)^2$
1	30	20	10	100
2	33	17	16	256
3	41	21	20	400
4	50	43	7	49
5	36	39	–3	9
6	45	11	34	1156
7	31	28	3	9
8	25	20	5	25
n=8			Σd_i=92	Σd_i^2=2004

Подставив в формулу найденные значения Σd_i и Σd_i^2 получим:

$$t = \frac{92}{\sqrt{2004 - 92^2 / 8}} \cdot \sqrt{\frac{8-1}{8}} \approx 2{,}798, \quad df = 8 - 1 = 7$$

Имеем: $t_{эмп}$=2,798

Находим по таблице 1 критические значения (Приложение 5.3.)

$$t_{кр} = \begin{cases} 2{,}365 \text{ для } p \le 0{,}05 \\ 3{,}499 \text{ для } p \le 0{,}01 \end{cases}$$

Отсюда: $2{,}365 < t_{эмп} = 2{,}798 < 3{,}499$

Вывод: Принимается H_1 гипотеза. Различия в уровнях тревожности до и после психотерапевтического воздействия следует признать статистически значимыми (p<0,05), так как эмпирическое значение превышает первое критическое, но меньше второго. Следовательно, психотерапевтическое воздействие действительно снижает тревожность.

Случай 2. При проверке отличия от нуля мер связи (коэффициентов корреляции) эмпирическое значение t-критерия вычисляется по формуле

$$t = r\sqrt{\frac{n-2}{1-r^2}} \quad ,$$

где r – коэффициент корреляции, n – количество испытуемых. Количество степеней свободы df= n–2. Вывод об отличии меры связи от нуля делается при превышении эмпирического значения критерия над критическим для α=0.05 и соответствующего числа степеней свободы, то есть аналогично рассмотренным выше случаям.

Для коэффициентов линейной корреляции Пирсона и ранговой корреляции Спирмена можно непосредственно использовать таблицы критических значений (Приложение 5.3., таблица 2). Эмпирическое значение коэффициента корреляции берется по абсолютной величине.

В некоторых пособиях и учебниках приводятся отдельные таблицы критических значений *для* коэффициента ранговой корреляции Спирмена, по В.Ю. Урбаху. Значения в них отличаются от критических для коэффициентов линейной корреляции Пирсона. Программа Statistica не делает различий между этими типами корреляций.

5.2. F-критерий Фишера (для сравнения дисперсий)

F-критерий Фишера используется для:

1) установления сходства-различия дисперсий в двух независимых выборках ($D_1 \leftrightarrow D_2$);

2) установления отличия от нуля коэффициента детерминации ($\eta^2 \leftrightarrow$"О");

3) установления наличия-отсутствия влияния фактора в дисперсионном анализе.

Случай 1

Эмпирическое значение F-критерия *для* сравнения двух дисперсий в независимых выборках находят по очень простой формуле:

$$F = \frac{D_1}{D_2} = \frac{\sigma_1^2}{\sigma_2^2} \quad ,$$

где D_1 – большая дисперсия, D_2 – меньшая дисперсия [Подстановка в числитель большей дисперсии необходима для использования таблиц критических значений, в которых приводится только правое критическое значение (больше единицы). Статистические программы рассчитывают и левое критическое значение (меньше единицы)].

Количество степеней свободы определяется отдельно для числителя и отдельно для знаменателя:

$df_{числ}= n_{числ}-1$

$df_{знам} =n_{знам} -1$

Пример 5.3.Две группы испытуемых обучались некоторым моторным навыкам по двум разным методикам, фиксировалось количество ошибочных действий, до обучения результаты в обеих группах имели одинаковый разброс. Какая из методик даст наибольшее выравнивание результатов внутри группы после обучения (табл.5.3.).

Таблица 5.3.

Числовые характеристики	1-я группа	2-я группа
n	21	16
σ	4	6
D	16	36

Подставляя в формулу получим:

$F_{эмп}$= 36/16 = 2,25.

$df_{числ}$= 16-1 = 15

$df_{знам}$ =21-1 = 20

Поскольку нам заранее не известно, какая из методик может обладать меньшей дисперсией, мы используем ненаправленную гипотезу и, следовательно, двусторонний критерий. Находим по таблице 3 (Приложение) критическое значение $F_{кр}$ для α = 0,05 ($\alpha/2+\alpha/2$ = 0,05) и $df_{числ}$=15, $df_{знам}$=20, $F_{крит}$=2,573.

Получим: $F_{эмп}=2{,}25 \angle F_{крит}=2{,}573$

Вывод: Так как эмпирическое значение меньше критического, то статистически значимых различий дисперсий в первой и второй группах нет и, следовательно, стабилизация навыка при обучении по обеим методикам одинакова.

Замечание. Для сравнения дисперсий в зависимых выборках более строгим будет применение t-критерия Стьюдента.

Случай 2

В случае определения отличия от нуля коэффициента детерминации эмпирическое значение F-критерия рассчитывается так:

$$F = \frac{\eta^2 (N-r)}{(1-\eta^2)(r-1)} ,$$

где: N – общее число испытуемых, r-число интервалов квантования, исходя из которых рассчитывалось η^2.

При определении критического значения число степеней свободы для числителя:

$df_{числ}$=r–1,

для знаменателя:

$df_{знам}$=N–r.

(Коэффициент детерминации – η^2, определяет общую меру связи – корреляционное отношение. Он определяется по формуле:

$$\eta^2 = 1 - \frac{ss_{внтр}}{ss_{общ.}} ,$$

Здесь:

$SS_{внтр}$ – сумма квадратов отклонений от внутригруппового (условного) среднего;

$SS_{общ}$ – сумма квадратов отклонений от общего для всех измерений среднего (безусловного среднего);

Следует отметить, что в отличие от линейной корреляции коэффициент детерминации устанавливает два типа связей: зависимость *x* от *y* и зависимость *y* от *x* ($\eta^2_{x/y}$, $\eta^2_{y/x}$). То есть сначала одна переменная рассматривается как зависимая, другая – как независимая, затем наоборот).

Пример 5.4. В таблице 5.4. даны результаты тестирования по двум методикам 12 испытуемых. Отличается ли коэффициент детерминации от нуля?

Таблица 5.4.

Испытуемые	1	2	3	4	5	6	7	8	9	10	11	12
1-я методика (X)	10	10	10	18	18	18	26	26	26	34	34	34
2-я методика (Y)	4	5	6	7	8	9	9	8	7	6	5	4

В нашем случае, имеем η^2=0,77, N=12, r=4. Подставляя в формулу,

$$F = \frac{0{,}77 \cdot (12-4)}{(1-0{,}77)\cdot(4-1)} \approx 8.93 \quad .$$

получаем $F_{эмп}$=8,93.

Критические значения F-критерия для α =0,05, $df_{числ}$=4-1=3, $df_{знам}$=12-4=8 находим по таблице:

$$F_{кр} = \begin{cases} 4{,}066 \text{ для } p \le 0{,}05 \\ 7{,}591 \text{ для } p \le 0{,}01 \end{cases}$$

$F_{эмп}=8{,}93>F_{кр}(\rho\le 0{,}01)=7{,}591$

Вывод: Коэффициент детерминации η^2 статистически значимо отличается от нуля ($\rho<0.01$).

5.3. Выявление различий в уровне исследуемого признака. Q-критерий Розенбаума

В психологии часто приходится проводить исследования на выявление различий между двумя, тремя и более выборками испытуемых. Это может быть, например, задача определения психологических особенностей больных детей по сравнению со здоровыми, или различий между работниками государственных предприятий и частных фирм и т.д.

Иногда по выявленным в исследовании статистически достоверным различиям формируется "групповой профиль" или "усредненный портрет" человека той или иной профессии, статуса, соматического заболевания и др.

В последние годы все чаще встает задача выявления психологического портрета специалиста новых профессий: "успешного менеджера", "успешного политика", "успешного торгового представителя", "успешного коммерческого директора" и др. Такого рода исследования не всегда подразумевают участие двух или более выборок. Иногда обследуется одна, но достаточно представительная выборка численностью не менее 60 человек, а затем внутри, этой выборки выделяются группы более и менее успешных специалистов, и их данные по исследованным переменным сопоставляются между собой. В самом простом случае критерием для разделения выборки на "успешных" и "неуспешных" будет средняя величина по показателю успешности. Однако такое деление является довольно грубым: лица, получившие близкие оценки по успешности, могут оказаться в противоположных группах, а лица, заметно различающиеся по оценкам успешности, – в одной и той же группе.

Это может исказить результаты сопоставления групп или, по крайней мере, сделать различия между группами менее заметными.

Чтобы избежать этого, можно попробовать выделить группы "успешных" и "неуспешных" специалистов более строго, включая в первую из них только тех, чьи значения *превышают* среднюю величину не менее чем на 1/4 стандартного отклонения, а во вторую группу – только тех, чьи значения не менее чем на 1/4 стандартного отклонения *ниже* средней величины. При этом все, кто оказывается в зоне средних величин, $M\pm 1/4\ \sigma$, выпадают из дальнейших сопоставлений. Если распределение близко к нормальному, то выпадет примерно 19,8% испытуемых. Если распределение отличается от нормального, то таких испытуемых может быть и больше. Чтобы избежать потерь, можно сопоставлять не две, а три группы испытуемых: с высокой, средней и низкой профессиональной успешностью.

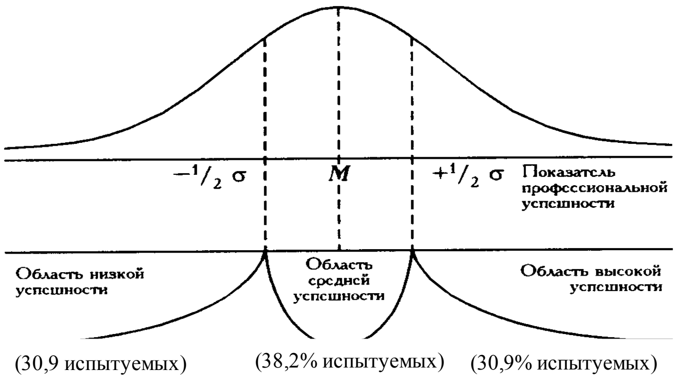


Рис 5.1. Схематическое изображение процесса разделения выборки на группы с низкой, средней и высокой профессиональной успешностью

На Рис. 5.1 представлена схема разделения выборки на группы с низкой, средней и высокой профессиональной успешностью по критерию отклонения значений от средней величины на 1/2 стандартного отклонения. При таком строгом критерии в "среднюю" группу попадают (при нормальном распределении) около 38,2% всех испытуемых, а в крайних группах оказывается по 30,9% испытуемых.

Чем меньше испытуемых оказывается в группах, тем меньше у нас возможностей для выявления достоверных различий, так как критические значения большинства критериев при малых n строже, чем при больших n.

Таким образом, при нестрогом разделении испытуемых на группы мы теряем в точности, а при строгом – в количестве испытуемых.

Сопоставление уровневых показателей в разных выборках может быть необходимой частью комплексных диагностических, учебных, психокоррекционных и иных программ. Оно помогает обратить внимание на те особенности обследованных выборок, которые должны быть учтены и использованы при адаптации программ к данной группе в процессе их конкретного воплощения.

Решение о выборе того или иного критерия принимается на основе того, сколько выборок сопоставляется и каков их объем.

Назначение критерия.

Критерий используется для оценки *различий между двумя* выборками *по уровню* какого-либо признака, количественно измеренного. В каждой из выборок должно быть не менее 11 испытуемых.

Описание критерия.

Это очень простой непараметрический критерий, который позволяет быстро оценить различия между двумя выборками по какому-либо признаку. Однако если критерий Q не выявляет достоверных различий, это еще не означает, что их действительно нет.

В этом случае стоит применить критерий φ* – Фишера. Если же Q-критерий выявляет достоверные различия между выборками с уровнем значимости $\rho \leq 0{,}01$, то можно ограничиться только им и избежать трудностей применения других критериев.

Критерий применяется в тех случаях, когда данные представлены, по крайней мере, в порядковой шкале. Признак должен варьировать в каком-то диапазоне значений, иначе сопоставления с помощью Q-критерия просто невозможны. Например, если у нас только 3 значения признака, 1, 2 и 3, – нам очень трудно будет установить различия. Метод Розенбаума требует достаточно тонко измеренных признаков.

Применение критерия начинается с упорядочивания значений признака в обеих выборках по нарастанию (или убыванию) признака. Для того чтобы не запутаться, в этом и во многих других критериях рекомендуется первым рядом (выборкой, группой) считать тот ряд, где значения выше, а вторым рядом – тот, где значения ниже.

Гипотезы.

H_0: *Уровень признака в выборке 1 не превышает уровня признака в выборке* 2.

H_1: *Уровень признака в выборке 1 превышает уровень признака в выборке 2.*

Графическое представление критерия Q.

На Рис. 5.2. представлены три варианта соотношения рядов значений в двух выборках. В варианте (а) все значения первого ряда выше всех значений второго ряда. Различия, безусловно, достоверны, при соблюдении условия, что $n_1, n_2 \geq 11$.

В варианте (б), напротив, оба ряда находятся на одном и том же уровне: различия недостоверны. В варианте (в) ряды частично перекрещиваются, но все же первый ряд оказывается гораздо выше второго. Достаточно ли велики зоны S_1 и S_2, в сумме составляющие Q, можно определить по Таблице 4 Приложения 1, где приведены критические значения Q для разных *n*. Чем величина Q больше, тем более достоверные различия мы сможем констатировать.

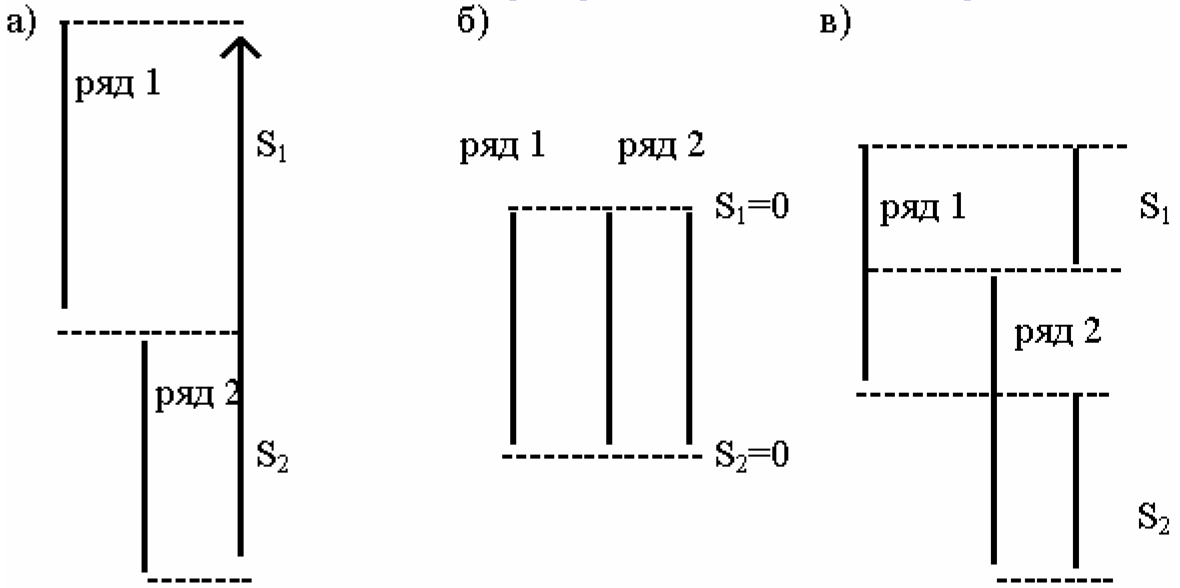


Рис. 5.2. Возможные соотношения рядов значений в двух выборках; S_1 – зона значений первого ряда, которые выше максимального значения 2-го ряда; S_2 – зона значений второго ряда, которые меньше минимального значения 1-го ряда; штриховкой отмечены перекрещивающиеся зоны двух рядов.

Ограничения критерия Q

1) В каждой из сопоставляемых выборок должно быть не менее 11 наблюдений. При этом объемы выборок должны примерно совпадать. Е.В. Гублером указываются следующие правила:
   а) если в обеих выборках меньше 50 наблюдений, то абсолютная величина разности между n_1 и n_2 не должна быть больше 10 наблюдений;
   б) если в каждой из выборок больше 51 наблюдения, но меньше 100, то абсолютная величина разности между n_1 и n_2 не должна быть больше 20 наблюдений;
   в) если в каждой из выборок больше 100 наблюдений, то допускается, чтобы одна из выборок была больше другой не более чем в 1,5-2 раза (Гублер Е.В., 1978, с. 75).
2) Диапазоны разброса значений в двух выборках должны не совпадать между собой, в противном случае применение критерия бессмысленно. Между тем, возможны случаи, когда диапазоны разброса значений совпадают, но, вследствие разносторонней асимметрии двух распределений, различия в средних величинах признаков существенны (Рис. 5.3., 5.4).

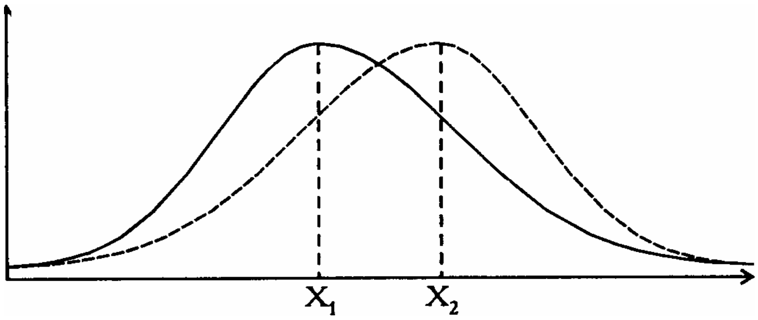


Рис. 5.3. Вариант соотношения распределений признака в двух выборках, при котором критерий Q беспомощен

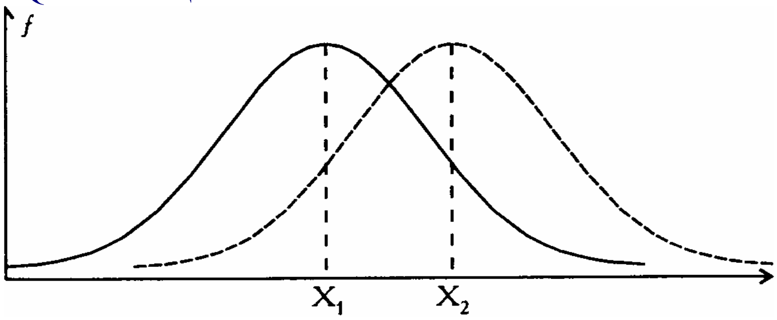


Рис. 5.4. Вариант соотношения распределений признака в двух выборках, при котором критерий Q может быть могущественным.

Пример 5.5. У предполагаемых участников психологического эксперимента, моделирующего деятельность воздушного диспетчера, был измерен уровень вербального и невербального интеллекта с помощью методики Д. Векслера. Было обследовано 26 юношей в возрасте от 18 до 24 лет (средний возраст 20,5 лет). 14 из них были студентами физического факультета, а 12 – студентами психологического факультета Ленинградского университета (Сидоренко Е.В., 1978). Показатели вербального интеллекта представлены в Таблице 5.5. Можно ли утверждать, что одна из групп превосходит другую по уровню вербального интеллекта?

Таблица 5.5. Индивидуальные значения вербального интеллекта в выборках студентов физического (n_1=14) и психологического (n_2=12) факультетов

Студенты-физики		Студенты-психологи	
Код имени испытуемого	Показатель вербального интеллекта	Код имени испытуемого	Показатель вербального интеллекта
1. И.А	132	1. Н.Т.	126
2. К.А.	134	2. О.В.	127
3. К.Е.	124	3. Е.В.	132
4. П.А.	132	4. . Ф.О.	120
5. С.А.	135	5. И.Н.	119
6. Ст.А.	132	6. И.Ч.	126
7. Т.А.	131	7. И.В.	120
8. Ф.А.	132	8. К.О.	123
9. Ч.И.	121	9. Р.Р.	120
10. Ц.А.	127	10. Р.И.	116
11. См.А.	136	11. О.К.	123
12. К.Ан.	129	12. Н.К.	115
13. Б.Л.	136		
14. Ф.В.	136		

Упорядочим значения в обеих выборках, а затем сформулируем гипотезы:

H_0: Студенты-физики не превосходят студентов-психологов по уровню вербального интеллекта.

H_1: Студенты-физики превосходят студентов-психологов по уровню вербального интеллекта.

Таблица 5.6.Упорядоченные по убыванию вербального интеллекта ряды индивидуальных значений в двух студенческих выборках

1 ряд - студенты-физики				2 ряд - студенты-психологи			
1	См.А.	136					
2	Б.Л.	136					
3	Ф.В.	136	S_1				
4	С.А.	135					
5	К.А.	134					
6	И.А.	132			1	Е.В.	132
7	П.А.	132					
8	Ст.А.	132					
9	Ф.А.	132					
10	Т.А.	131					
11	К.Ан.	129					
12	Ц.А.	127			2	О.В.	127
					3	Н.Т.	126
					4	И.Ч.	126
13	К.Е.	124					
					5	К.О.	123
					6	О.К.	123
14	Ч.И.	121					
					7	Ф.О.	120
					8	И.В.	120
				S_2	9	Р.Р.	120
					10	И.Н.	119
					11	Р.И.	116
					12	Н.К.	115

Как видно из Табл. 5.6. мы обозначили ряды: первый, тот, что "выше" – ряд физиков, а второй, тот, что "ниже" – ряд психологов.

По Табл. 5.6. определяем количество значений первого ряда, которые больше максимального значения второго ряда: $S_1=5$.

Теперь определяем количество значений второго ряда, которые меньше минимального значения первого ряда: $S_2=6$.

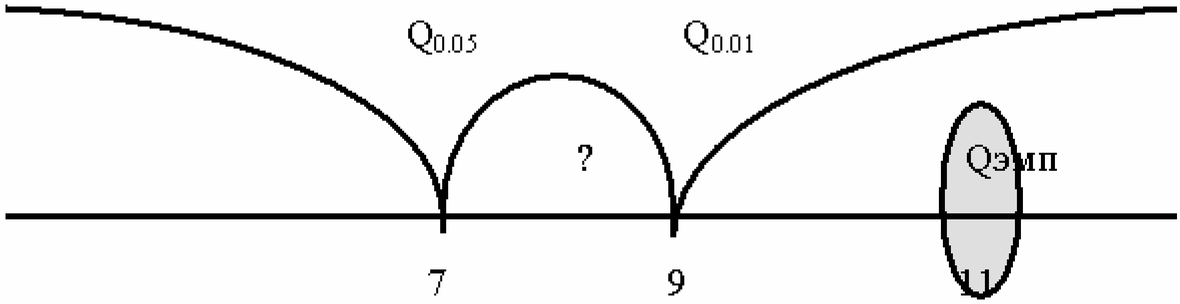
Вычисляем $Q_{эмп}$, по формуле:

$Q_{эмп} = S_1+S_2= 5+6=11$

По Табл. 3 Приложения 5.3. определяем критические значения Q для $n_1=14$, $n_2=12$:

$$Q_{кр} = \begin{bmatrix} 7 \ (p \le 0{,}05) \\ 9 \ (p \le 0{,}01) \end{bmatrix}$$

Чем больше расхождения между выборками, тем больше величина Q. Построим "ось значимости".



Qэмп > Qкр (р ≤0,01)

Ответ: H_0 отклоняется. Принимается H_1. Студенты-физики превосходят студентов-психологов по уровню вербального интеллекта ($p<0,01$).

Замечание. В тех случаях, когда эмпирическая величина критерия оказывается на границе зоны незначимости, мы имеем право утверждать лишь, что различия достоверны при $p \le 0,05$, если же оно оказывается между двумя критическими значениями, то мы можем утверждать, что $\rho<0,05$.

Если эмпирическое значение критерия оказывается на границе зоны значимости, $p \le 0,01$, в зоне значимости – что $p<0,01$.

АЛГОРИТМ
Подсчет критерия Q Розенбаума

1) Проверить, выполняются ли ограничения: n_1 ,$n_2 \ge 11$, $n_1 \approx n_2$.
2) Упорядочить значения отдельно в каждой выборке по степени возрастания признака. Считать выборкой 1 ту выборку, значения в которой предположительно выше, а выборкой 2 – ту, где значения предположительно ниже.
3) Определить самое высокое (максимальное) значение в выборке 2.
4) Подсчитать количество значений в выборке 1, которые выше максимального значения в выборке 2. Обозначить полученную величину как S_1.
5) Определить самое низкое (минимальное) значение в выборке 1.
6) Подсчитать количество значений в выборке 2, которые ниже минимального значения выборки. Обозначить полученную величину как S_2.
7) Подсчитать эмпирическое значение Q по формуле: $Q=S_1+S_2$.
8) По табл. 3 Приложения 5.3 определить критические значения Q для данных n_1 и n_2. Если Qэмп равно $Q_{0.05}$ или превышает его, H_0 отвергается.
9) При n_1, $n_2>26$ сопоставить полученное эмпирическое значение с $Q_{кр}=8$ ($p \le 0,05$) и $Q_{кр}=10$ ($p \le 0,01$). Если $Q_{эмп}$. превышает или по крайней мере равняется Qкр=8, H_0 отвергается.

5.4. Т-критерий Вилкоксона

Назначение критерия.

Критерий применяется для сопоставления показателей, измеренных в *двух* разных условиях на *одной* и той же выборке испытуемых.

Он позволяет установить не только *направленность* изменений, но *и* их *выраженность*. С его помощью мы определяем, является ли сдвиг показателей в каком-то одном направлении более интенсивным, чем в другом.

Описание критерия Т.

Этот критерий применим в тех случаях, когда признаки измерены по крайней мере по шкале порядка, и сдвиги между вторым и первым замерами тоже могут быть упорядочены. Для этого они должны варьировать в достаточно широком диапазоне. В принципе, можно применять критерий Т и в тех случаях, когда сдвиги принимают только три значения: -1, 0 и +1, но тогда критерий Т вряд ли добавит что-нибудь новое к тем выводам, которые можно было бы получить с помощью критерия знаков. Вот если сдвиги изменяются, скажем, от -30 до +45, тогда имеет смысл их ранжировать и потом суммировать ранги.

Суть метода состоит в сопоставлении выраженности сдвигов в том и ином направлениях по абсолютной величине. Для этого сначала ранжируются все абсолютные величины сдвигов, а потом суммируются ранги. Если сдвиги в положительную и в отрицательную сторону происходят случайно, то суммы рангов абсолютных значений их будут примерно равны. Если же интенсивность сдвига в одном из направлений перевешивает, то сумма рангов абсолютных значений сдвигов в противоположную сторону будет значительно ниже, чем это могло бы быть при случайных изменениях.

Первоначально исходят из предположения о том, что типичным сдвигом будет сдвиг в более часто встречающемся направлении, а нетипичным, или редким, сдвигом – сдвиг в более редко встречающемся направлении.

Гипотезы.

H_0: *Интенсивность сдвигов в типичном направлении не превосходит интенсивности сдвигов в нетипичном направлении.*

H_1: *Интенсивность сдвигов в типичном направлении превышает интенсивность сдвигов в нетипичном направлении.*

Ограничения в применении критерия T Вилкоксона

1) Минимальное количество испытуемых, прошедших измерения в двух условиях – 5 человек. Максимальное количество испытуемых – 50 человек, что диктуется верхней границей имеющихся таблиц. Критические значения T приведены в Табл. VI Приложения 1.
2) Нулевые сдвиги из рассмотрения исключаются, и количество наблюдений n уменьшается на количество этих нулевых сдвигов (McCall R., 1970, p. 36). Можно обойти это ограничение, сформулировав гипотезы, включающие отсутствие изменений, например: "Сдвиг в сторону увеличения значений превышает сдвиг в сторону уменьшения значений и тенденцию сохранения их на прежнем уровне".

Пример 5.6. В выборке курсантов военного училища (юноши в возрасте от 18 до 20 лет) измерялась способность к удержанию физического волевого усилия на динамометре. Сначала у испытуемых измерялась максимальная мышечная сила каждой из рук, а на следующий день им предлагалось выдерживать, на динамометре с подвижной стрелкой мышечное усилие, равное 1/2 максимальной мышечной силы данной руки. Почувствовав усталость, испытуемый должен был сообщить об этом экспериментатору, но не прекращать опыт, преодолевая усталость и неприятные ощущения – "бороться, пока воля не иссякнет". Опыт проводился дважды; вначале с обычной инструкцией, а затем, после того, как испытуемый заполнял опросник самооценки волевых качеств по методике А.Ц. Пуни (Пуни А.Ц., 1977), ему предлагалось представить себе, что он уже добился идеала в развитии волевых качеств, и продемонстрировать соответствующее идеалу волевое усилие. Подтвердилась ли гипотеза экспериментатора о том, что обращение к идеалу способствует возрастанию волевого усилия? Данные представлены в табл. 5.7.

Таблица 5.7. Расчет критерия T при сопоставлении замеров физического волевого усилия

Код имени испытуемого		Длительность удержания усилия на динамометре (с)		Разность ($t_{после}$-$t_{до}$)	Абсолютное значение разности	Ранговый номер разности
		До измерения волевых качеств ($t_{до}$)	После измерения волевых качеств ($t_{после}$)			
1.	Г.	64	25	-39	39	11
2.	Кос.	77	50	-27	27	8
3.	Крв.	74	72	+ 3	3	1
4.	Кур.	95	76	-19	19	6
5.	Л.	105	67	-38	38	9,5
6.	М.	83	75	- 8	8	4
7.	Р.	73	77	+4	4	2,5
8.	С.	75	71	- 4	4	2,5
9.	Т.	101	63	-38	38	9,5
10.	Х.	97	122	+25	25	7
11.	Ю.	78	60	-18	18	5
Сумма						66

Для подсчета этого критерия нет необходимости упорядочивать ряды значений по нарастанию признака. Можно использовать алфавитный список испытуемых, как в данном случае.

Первый шаг в подсчете критерия T – вычитание каждого индивидуального значения "до" из значения "после" (можно вычитать значения "после" из значений "до", это никак не повлияет на расчет критерия. Но лучше во всех случаях придерживаться одной системы, чтобы не запутаться самим).

Мы видим из табл. 5.7., что 8 полученных разностей – отрицательные и лишь 3 – положительные.

Это означает, что у 8 испытуемых длительность удержания мышечного усилия во втором замере уменьшилась, а у 3 – увеличилась. Мы столкнулись с тем случаем, когда уже сейчас нельзя сформулировать статистическую гипотезу, соответствующую первоначальному предположению исследователя. Предполагалось, что обращение к идеалу будет увеличивать длительность мышечного усилия, а экспериментальные данные свидетельствуют, что лишь в 3 случаях из 11 этот показатель действительно увеличился. Можно сформулировать лишь гипотезу, предполагающую несущественность сдвига этого показателя в сторону снижения.

Гипотезы.

H_0: *Интенсивность сдвигов в сторону уменьшения длительности мышечного усилия не превышает интенсивности сдвигов в сторону ее увеличения.*

H_1: *Интенсивность сдвигов в сторону уменьшения длительности мышечного усилия превышает интенсивность сдвигов в сторону ее увеличения.*

На следующем шаге все сдвиги, независимо от их знака, должны быть проранжированы по выраженности. В Таблице 5.7. в четвертом слева столбце приведены абсолютные величины сдвигов, а в последнем столбце (справа) – ранги этих абсолютных величин. Меньшему значению соответствует меньший ранг. При этом сумма рангов равна 66, что соответствует расчетной:

$$\sum R_i = \frac{N \cdot (N+1)}{2} = \frac{11 \cdot (11+1)}{2} = 66$$

Теперь отметим те сдвиги, которые являются нетипичными, в данном случае – положительными. В Таблице 5.7. эти сдвиги и соответствующие им ранги выделены цветом. Сумма рангов этих «редких» сдвигов составляет эмпирическое значение критерия Т:

$$T = \sum R_r$$

где R_r - ранговые значения сдвигов с более редким знаком.

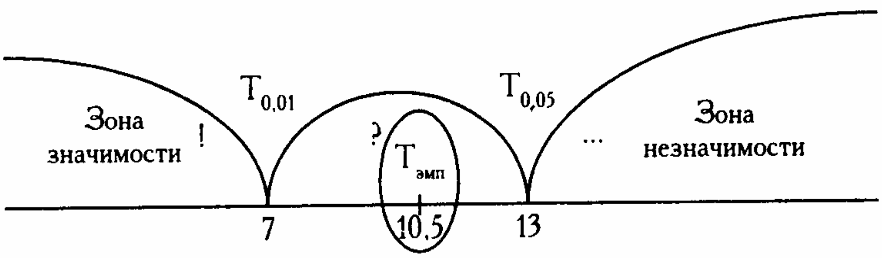
Итак, в данном случае,

$$T_{эмп} = 1 + 2{,}5 + 7 = 10{,}5$$

По таблице 4 Приложения 5.3. находим критические значения для Т-критерия Вилкоксона для n=11:

$$T_{кр.} = \begin{cases} 13\,(p \le 0{,}05) \\ 7\,(p \le 0{,}01) \end{cases}$$

Построим “ось значимости”.



Зона значимости в данном случае простирается влево, действительно, если бы "редких", в данном случае положительных, сдвигов не было совсем, то и сумма их рангов равнялась бы нулю. В данном же случае эмпирическое значение Т попадает в зону неопределенности:

$T_{эмп.} < T_{кр(0,05)}$

Ответ: H_0 отвергается. Интенсивность отрицательного сдвига показателя физического волевого усилия превышает интенсивность положительного сдвига ($p<0{,}05$).

Попытаемся графически отобразить интенсивность отрицательных и положительных сдвигов. На Рис. 5.5. слева сдвиги представлены в секундах, а справа – в своих ранговых значениях. Мы видим, что ранжирование несколько уменьшает площади сопоставляемых облаков, или "фронтов".

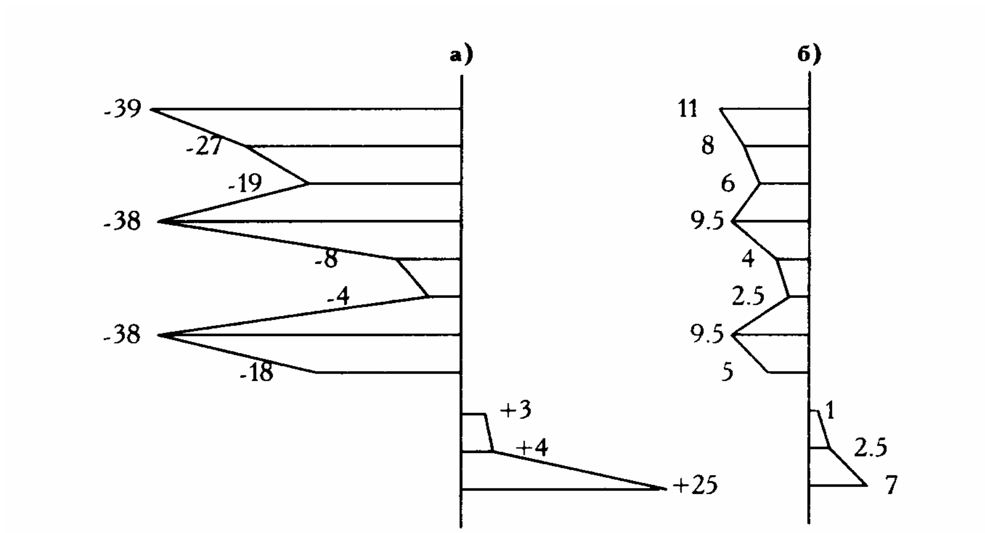


Рисунок 5.5. Графическое представление отрицательных и положительных сдвигов; слева в секундах, справа в ранговых значениях

Таким образом, продолжительность удержания мышечного волевого усилия во втором замере снижается, и этот сдвиг неслучаен. Инструкция, ориентирующая испытуемого на соответствие идеалу в развитии воли, оказалась гораздо менее мощным фактором, чем какая-то иная сила – возможно, мышечное утомление, может быть, разочарование в себе или в возможностях данного психологического эксперимента. А может быть, в момент второго замера просто перестает действовать какой-то мощный фактор, который был активен вначале?

Представим выполненные действия в виде алгоритма:

АЛГОРИТМ
Подсчета критерия Т Вилкоксона

1) Составить список испытуемых в любом порядке, например, алфавитном.
2) Вычислить разность между индивидуальными значениями во втором и первом замерах ("после" – "до"). Определить, что будет считаться "типичным" сдвигом и сформулировать соответствующие гипотезы.
3) Перевести разности в абсолютные величины и записать их отдельным столбцом (иначе трудно отвлечься от знака разности).
4) Проранжировать абсолютные величины разностей, начисляя меньшему значению меньший ранг. Проверить совпадение полученной суммы рангов с расчетной.
5) Отметить кружками или другими знаками ранги, соответствующие сдвигам в "нетипичном" направлении.
6) Подсчитать сумму этих рангов по формуле:
   $T=\sum R_r$, где *Rr* – ранговые значения сдвигов с более редким знаком.
7) Определить критические значения Т для данного n по табл. 4 Приложения 5.3.
   Если Тэмп. меньше или равен Ткр., сдвиг в "типичную" сторону по интенсивности достоверно преобладает.

5.5. Выявление различий в распределении признака. χ^2-критерий Пирсона

Распределения могут различаться по средним, дисперсиям, асимметрии, эксцессу и по сочетаниям этих параметров. Рассмотрим несколько примеров.

На Рис. 5.6 представлены два распределения признака. Распределение 1 характеризуется меньшим диапазоном вариативности и меньшей дисперсией, чем распределение 2. В распределении 1 чаще встречаются значения признака, близкие к средней, а в распределении 2 чаще встречаются более высокие и более низкие, чем средняя, значения признака.

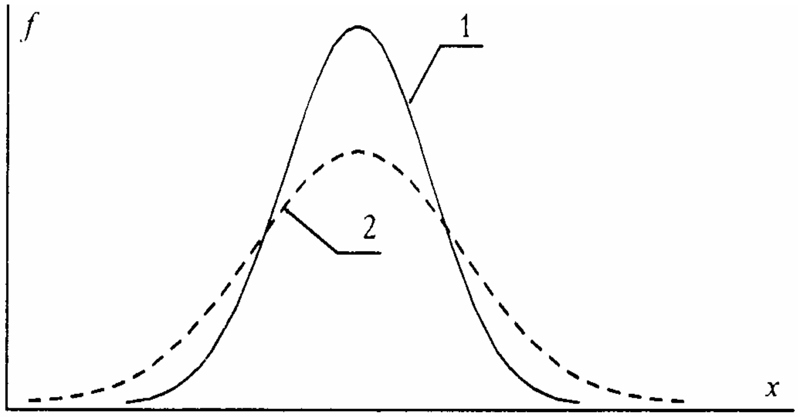


Рисунок 5.6. Кривые распределения признака с меньшим диапазоном вариативности признака(1) и большим диапазоном вариативности признака (2)

Именно такое соотношение может наблюдаться в распределении фенотипических признаков у мужчин (кривая 2) и женщин (кривая 1). Фенотипическая дисперсия мужского пола должна быть больше, чем женского (Геодакян В.А., 1974; 1993). Мужчины – это авангардная часть популяции, ответственная за поиск новых форм приспособления, поэтому у них чаще встречаются редкие крайние значения различных фенотипических признаков. Эти отклонения, по мнению В.А. Геодакяна носят "футуристический" характер, это "пробы", включающие как будущие возможные пути эволюции, так и ошибки (Геодакян В.А., 1974, с. 381). В то же время женская часть популяции ответственна за сохранение уже накопленных изменений, поэтому у них чаще встречаются средние значения фенотипических признаков.

Анализ реально получаемых в исследованиях распределений может позволить нам подтвердить или опровергнуть данные теоретические предположения.

На Рис. 5.7 представлены два распределения, различающиеся по знаку асимметрии: распределение 1 характеризуется положительной асимметрией (левосторонней), а распределение 2 – отрицательной (правосторонней).

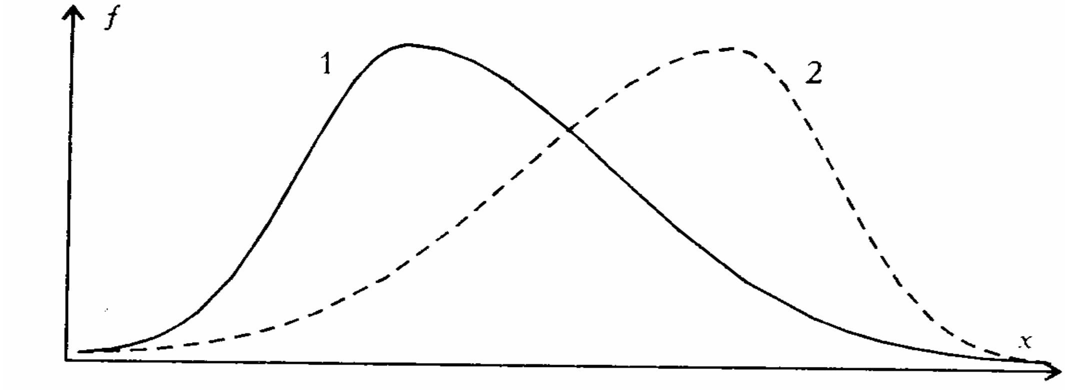


Рисунок 5.7. Кривые распределения признака с положительной асимметрией (1) и отрицательной (2)

Данные кривые могут отражать распределение времени решения простой задачи (кривая 1) и трудной задачи (кривая 2). Простую задачу большинство испытуемых решают быстро, поэтому большая часть значений группируется слева. В то же время сама простота задачи может привести к тому, что некоторые испытуемые будут думать над нею очень, очень долго, дольше даже, чем над сложной.

Трудную задачу большинство испытуемых решают в тенденции дольше, чем простую, но в то же время почти всегда находятся люди, которые решают ее мгновенно. Если мы докажем, что распределения статистически достоверно различаются, это может стать основой для построения классификаций задач и типологий испытуемых.

Часто бывает полезно сопоставить полученное эмпирическое распределение с теоретическим распределением. Например, для того, чтобы доказать, что оно подчиняется или, наоборот, не подчиняется нормальному закону распределения.

В практических целях эмпирические распределения должны проверяться на "нормальность" в тех случаях, когда мы намерены использовать параметрические методы и критерии.

Традиционные для отечественной математической статистики критерии определения расхождения или согласия распределений – это метод χ^2 – К. Пирсона и λ-критерий Колмогорова-Смирнова.

Оба эти метода требуют тщательной группировки данных и довольно сложных вычислений. Кроме того, возможности этих критериев в полной мере проявляются на больших выборках (n>30). Тем не менее, они могут оказаться столь незаменимыми, что исследователю придется пренебречь экономией времени и усилий. Например, они незаменимы в следующих двух случаях:

1) в задачах, требующих доказательства неслучайности предпочтений в выборе из нескольких альтернатив;
2) в задачах, требующих обнаружения точки максимального расхождения между двумя распределениями, которая затем используется для перегруппировки данных с целью применения критерия φ^* (углового преобразования Фишера).

Назначения критерия.

Критерий χ^2 применяется в двух целях:

1) для сопоставления *эмпирического* распределения признака с *теоретическим* – равномерным, нормальным или каким-то иным;
2) для сопоставления *двух трех или более эмпирических* распределений одного и того же признака На самом деле области применения критерия χ^2 многообразны (см., например: Суходольский Г.В., 1972, с. 295), но в данном руководстве мы ограничиваемся только этими двумя, наиболее часто встречающимися на практике, целями).

Описание критерия.

Критерий χ^2 отвечает на вопрос о том, с одинаковой ли частотой встречаются разные значения признака в эмпирическом и теоретическом распределениях или в двух и более эмпирических распределениях.

Преимущество метода состоит в том, что он позволяет сопоставлять распределения признаков, представленных в любой шкале, начиная от шкалы наименований. В самом простом случае альтернативного распределения "да – нет", "допустил брак – не допустил брака", "решил задачу – не решил задачу".

При сопоставлении эмпирического распределения с теоретическим мы определяем степень расхождения между эмпирическими и теоретическими частотами.

При сопоставлении двух эмпирических распределений мы определяем степень расхождения между эмпирическими частотами и теоретическими частотами, которые наблюдались бы в случае совпадения двух этих эмпирических распределений. Формулы расчета теоретических частот будут специально даны для каждого варианта сопоставлений.

Чем больше расхождение между двумя сопоставляемыми распределениями, *тем больше* эмпирическое *значение* χ^2.

Гипотезы.

Возможны несколько вариантов гипотез, в зависимости от задач, которые мы перед собой ставим.

Первый вариант:

H_0: Полученное эмпирическое распределение признака не отличается от теоретического (например, равномерного) распределения.

H_1: Полученное эмпирическое распределение признака отличается от теоретического распределения.

Второй вариант:

H_0: Эмпирическое распределение 1 не отличается от эмпирического распределения 2.

H_1: Эмпирическое распределение 1 отличается от эмпирического распределения 2.

Третий вариант:

H_0: Эмпирические распределения 1, 2, 3,... не различаются между собой.

H_1: Эмпирические распределения 1, 2, 3, ... различаются между собой.

Критерий χ^2 позволяет проверить все три варианта гипотез.

Ограничения критерия.

1) Объем выборки должен быть достаточно большим: n>.30. При n<30 критерий χ^2 дает весьма приближенные значения. Точность критерия повышается при больших *n*.
2) Теоретическая частота для каждой ячейки таблицы не должна быть меньше 5: $f \geq 5$. Это означает, что если число разрядов задано заранее и не может быть изменено, то мы не можем применять метод, χ^2 не накопив определенного минимального числа наблюдений. Если, например, мы хотим проверить наши предположения о том, что частота обращений в телефонную службу Доверия неравномерно распределяются по 7 дням недели, то нам потребуется 5*7=35 обращений. Таким образом, если количество разрядов *(k)* задано заранее, как в данном случае, минимальное число наблюдений (n_{min}) определяется по формуле: n_{min}=k*5.
3) Выбранные разряды должны "вычерпывать" все распределение, то есть охватывать весь диапазон вариативности признаков. При этом группировка на разряды должна быть одинаковой во всех сопоставляемых распределениях.
4) Необходимо вносить "поправку на непрерывность" при сопоставлении распределений признаков, которые принимают всего 2 значения. При внесении поправки значение χ^2 уменьшается (см. Пример с поправкой на непрерывность).
5) Разряды должны быть неперекрещивающимися: если наблюдение отнесено к одному разряду, то оно уже не может быть отнесено ни к какому другому разряду. Сумма наблюдений по разрядам всегда должна быть равна общему количеству наблюдений.

Правомерен вопрос о том, что считать числом наблюдений – количество выборов, реакций, действий или количество испытуемых, которые совершают выбор, проявляют реакции или производят действия. Если испытуемый проявляет несколько реакций, и все они регистрируются, то количество испытуемых не будет совпадать с количеством реакций. Мы можем просуммировать реакции каждого испытуемого, как, например, это делается в методике Хекхаузена для исследования мотивации достижения или в Тесте фрустрационной толерантности С. Розенцвейга, и сравнивать распределения индивидуальных сумм реакций в нескольких выборках.

В этом случае числом наблюдений будет количество испытуемых. Если же мы подсчитываем частоту реакций определенного типа в целом по выборке, то получаем распределение реакций разного типа, и в этом случае количеством наблюдений будет общее количество зарегистрированных реакций, а не количество испытуемых.

С математической точки зрения правило независимости разрядов соблюдается в обоих случаях: одно наблюдение относится к одному и только одному разряду распределения.

Но считать ли наблюдением каждого испытуемого или каждую исследуемую реакцию испытуемого – вопрос, решение которого зависит от целей исследования.

Пример 5.7. (с поправкой на непрерывность).

В исследовании порогов социального атома профессиональных психологов просили определить, с какой частотой встречаются в их записной книжке мужские и женские имена коллег-психологов. Попытаемся определить, отличается ли распределение, полученное по записной книжке женщины-психолога X, от равномерного распределения. Эмпирические частоты представлены в табл. 5.8

Таблица 5.8. Эмпирические частоты встречаемости имен мужчин и женщин в записной книжке психолога X

Мужчин	Женщин	Всего человек
22	45	67

Гипотезы.

H_0: Распределение мужских и женских имен в записной книжке X не отличается от равномерного распределения.

H_1: Распределение мужских и женских имен в записной книжке X отличается от равномерного распределения.

Количество наблюдений n=67; количество значений признака $k=2$. Рассчитаем теоретическую частоту: $f_{теор.}$=n/k=33.5

Число степеней свободы ν=k-1.

Далее все расчеты производим по известному алгоритму, но с одним добавлением: перед возведением в квадрат разности частот мы должны уменьшить абсолютную величину этой разности на 0,5 (см. табл. 5.8, четвертый столбец).

Таблица 5.8. Расчет критерия χ^2 при сопоставлении эмпирического распределения имен с теоретическим равномерным распределением

	Разряды - принадлежность к тому или иному полу	Эмпирическая частота $f_{эj}$	Теоретическая частота $f_т$	$(f_{эj}- f_т)$	$(f_{эj}- f_т-0,5)$	$(f_{эj}- f_т-0,5)^2$	$\frac{(f_{эj}- f_т-0,5)^2}{f_t}$
1	Мужчины	22	33,5	-11,5	11	121	3,61
2	Женщины	45	33,5	+11,5	11	121	3,61
	Суммы	67	67	0			7,22

По таблице 5 Приложения 5.3 находим критические значения критерия для ν=1.

$$\chi^2_{кр} = \begin{cases} 3,841 \ (\rho \le 0,05) \\ 6,635 \ (\rho \le 0,01) \end{cases}$$

$\chi^2_{эмп}$=7,22

$\chi^2_{эмп}>\chi^2_{кр}$

Ответ: H_0 отклоняется, принимается H_1. Распределение мужских и женских имен в записной книжке психолога X отличается от равномерного распределения ($\rho<0,01$).

Пример 5.8 (сравнение двух эмпирических распределений). Определим, различаются ли распределения мужских и женских имен у психолога А и психолога В, тоже женщины. Эмпирические частоты приведены в табл. 5.10.

Таблица 5.10. Эмпирические частоты: встречаемости имен мужчин и женщин в записных книжках психолога X. и психолога С.

	Мужчин		Женщин		Всего человек
Психолог А.	22	А	45	Б	67
Психолог В.	59	В	109	Г	168
Суммы	81		154		235

Гипотезы.

H_0: Распределения мужских и женских имен в двух записных книжках не различаются.

H_1: Распределения мужских и женских имен в двух записных книжках различаются между собой.

Теоретические частоты рассчитываем по уже известной формуле:

$$f_{теор}=\frac{\left(\begin{array}{c}\text{Сумма частот по}\\ \text{соответствующей строке}\end{array}\right)\cdot\left(\begin{array}{c}\text{Сумма частот по}\\ \text{соответствующему столбцу}\end{array}\right)}{(\text{Общее количество наблюдений})}$$

А именно, для разных ячеек таблицы эмпирических частот,

$f_{А\ теор}=67\cdot 81/235=23{,}09$

$f_{Б\ теор}=67\cdot 154/235=43{,}91$

$f_{В\ теор}=168\cdot 81/235=57{,}91$

$f_{Г\ теор}=168\cdot 154/235=110{,}09$

Число степеней свободы $\nu=(k-1)\cdot(c-1)=1$

Все дальнейшие расчеты проводим по алгоритму (табл. 5.11).

Таблица 5.11. Расчет критерия при сопоставлении двух эмпирических распределений мужских и женских имен

Ячейки таблицы эмпирических частот		Эмпирическая частота $f_{эj}$	Теоретическая частота f_t	$(f_{эj}-f_т)$	$f_{эj}-f_т-0{,}5$	$(f_{эj}-f_т-0{,}5)^2$	$\frac{(f_{эj}-f_т-0{,}5)^2}{f_t}$
1	А	22	23,09	-1.09	0.59	0.35	0,015
2	Б	45	43.91	+1,09	0,59	0,35	0,008
3	В	59	57,91	+1.09	0,59	0,35	0,006
4	Г	109	110,09	-1,09	0,59	0.35	0,003
Суммы		235	235,00	0			0,032

Критические значения χ^2 при $\nu=1$ нам известны по предыдущему примеру:

$$\chi^2_{кр.}=\begin{cases}3{,}841\ (\rho\le 0{,}05)\\ 6{,}635\ (\rho\le 0{,}01)\end{cases}$$

$\chi^2_{эмп}=0{,}03$

$\chi^2_{эмп}<\chi^2_{кр.}$

Ответ: H_0 принимается. Распределения мужских и женских имен в записных книжка двух психологов совпадают.

Замечание. *Поправки на непрерывность можно избежать, если подобного рода задачи решать с помощью φ*-критерия Фишера.*

АЛГОРИТМ

Расчет критерия χ^2

1) Занести в таблицу наименования разрядов и соответствующие им эмпирические частоты (первый столбец).
2) Рядом с каждой эмпирической частотой записать теоретическую частоту (второй столбец).
3) Подсчитать разности между эмпирической и теоретической частотой по каждому разряду (строке) и записать их в третий столбец.
4) Определить число степеней свободы по формуле: ν=k-1, где k – количество разрядов признака. Если ν=1, внести поправку на "непрерывность".
5) Возвести в квадрат полученные разности и занести их в четвертый столбец.

6) Разделить полученные квадраты разностей на теоретическую частоту и записать результаты в пятый столбец.
7) Просуммировать значения пятого столбца. Полученную сумму обозначить как $\chi^2_{.эмп.}$
8) Определить по табл. 5 Приложения 5.3 критические значения для данного числа степеней свободы ν.

Если $\chi^2_{.эмп}$ меньше критического значения, расхождения между распределениями статистически недостоверны.

Если $\chi^2_{.эмп}$ равно критическому значению или превышает его, расхождения между распределениями статистически достоверны.

Приложение 5.1

Лабораторная работа № 4

Задание. Применение методов индуктивной статистики для проверки статистических гипотез в психологических исследованиях.

Цель задания. Освоение методов индуктивной статистики.

Аппаратура. Персональный компьютер.

Математическое обеспечение. Операционная система WINDOWS и EXCEL 7.0.

Теоретическое обеспечение.

1) Генеральная совокупность и выборка.
2) Статистические гипотезы. Проверка статистических гипотез. Уровни значимости.
3) Основные статистические критерии, применяемые в психологических исследованиях.

Этапы обработки данных.

1) Занести данные в таблицу Excel (две выборки).
2) Рассчитать отклонения каждого распределения от нормального.
3) Сделать выбор статистического критерия, опираясь на результаты п.2.
4) Сделать расчет по выбранной формуле (эмпирическое значение).
5) Сравнить эмпирическое значение коэффициента с критическим (по таблице).
6) Дать интерпретацию полученных результатов.

Приложение 5.2

Лабораторная работа №4
Вариант 1

Была исследована группа детей с заболеванием крови до лечения препаратами и после лечения. В таблицу занесены показатели er крови по результатам медицинского обследования. Сделать сравнительный анализ результативности лечения данным препаратом, используя статистические критерии.

Таблица. Лабораторные данные (результаты обследования детей с ОЛЛ)

№ респон.	до лечения	после леч.
	er	er
1.	4,2	2,65
2.	2	2,38
3.	2,33	3,5
4.	2,4	3,5
5.	1,8	4,8

6.	0,8	4
7.	2	3,7
8.	2,1	4,12
9.	2,8	4,2
10.	1,46	2,89
11.	3,85	3,7
12.	1,64	2,58
13.	2,6	3,2
14.	1,7	2,1

Вариант 2

Была исследована группа детей с заболеванием крови до лечения препаратами и после лечения. В таблицу занесены показатели L крови по результатам медицинского обследования. Сделать сравнительный анализ результативности лечения данным препаратом, используя статистические критерии.

Таблица. Лабораторные данные (результаты обследования детей с ОЛЛ)

№ респон.	до лечения	после леч.
	L	L
1.	20,5	2,3
2.	12,1	7,5
3.	13,6	3,8
4.	40,5	3,8
5.	9,6	4,8
6.	33	8,8
7.	77,2	13
8.	8,7	4,7
9.	3,5	3,9
10.	13,8	4,8
11.	7,4	5,7
12.	29,4	9
13.	116	13
14.	21,9	0,9

Вариант 3

Для проверки эффективности новой развивающей программы были созданы две группы детей шестилетнего возраста. Одна группа (экспериментальная) занималась по новой программе, вторая (контрольная) – по старой. После эксперимента дети обеих групп были протестированы по методике Керна-Йерасика (школьная зрелость). Результаты тестирования по вербальной шкале занесены в таблицу. Можно ли сделать заключение об эффективности новой программы и ее преимуществе перед старой.

Таблица. Результаты тестирования по вербальной шкале (сырые баллы)

№ исп.	эксп.	контр.
1	12	14
2	13	13
3	10	14
4	8	14
5	13	14
6	12	13
7	15	12
8	12	12

9	10	15
10	11	13
11	13	13
12	11	13
13	11	13
14	14	9
15	11	13
16	13	13

Вариант 4

Для проверки эффективности новой развивающей программы были созданы две группы детей шестилетнего возраста. Одна группа (экспериментальная) занималась по новой программе, вторая (контрольная) – по старой. После эксперимента дети обеих групп были протестированы по методике Керна-Йерасика (школьная зрелость). Результаты тестирования по вербальной шкале занесены в таблицу. Можно ли сделать заключение об эффективности новой программы и ее преимуществе перед старой.

Таблица. Результаты тестирования по вербальной шкале (сырые баллы)

№ исп.	эксп.	контр.
1	14	13
2	13	13
3	11	14
4	8	12
5	12	14
6	13	14
7	13	12
8	13	13
9	11	15
10	12	13
11	14	11
12	13	12
13	12	14
14	14	9
15	10	14

Вариант 5

Для проверки эффективности новой развивающей программы были созданы две группы детей шестилетнего возраста. Одна группа (экспериментальная) занималась по новой программе, вторая (контрольная) – по старой. После эксперимента дети обеих групп были протестированы по методике Керна-Йерасика (школьная зрелость). Результаты тестирования по невербальной шкале занесены в таблицу. Можно ли сделать заключение об эффективности новой программы и ее преимуществе перед старой.

Таблица. Результаты тестирования по вербальной шкале (сырые баллы)

номер исслед.	2 этап	
	эксп.	контр.
1	29	32
2	29	33