ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ КАЗЕННОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ УНИВЕРСИТЕТ МИНИСТЕРСТВА ВНУТРЕННИХ ДЕЛ
РОССИЙСКОЙ ФЕДЕРАЦИИ ИМЕНИ В.Я. КИКОТЯ»

Е. А. Слесарева

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ПСИХОЛОГИИ

Учебное пособие

Москва
2019

ББК 88.6
С47

Рецензенты:
старший инспектор по особым поручениям УОПК ДГСК МВД России ***М. А. Базанина****; доцент кафедры математики и моделирования систем Воронежского института МВД России кандидат физико-математических наук, доцент полковник полиции* ***В. Н. Думачев***

Слесарева, Е. А.
С47 ***Математические методы в психологии*** : учебное пособие / Е. А. Слесарева. – М. : Московский университет МВД России имени В.Я. Кикотя, 2019. – 109 с.
ISBN 978-5-9694-0741-1

Учебное пособие предназначено для курсантов и слушателей, обучающихся по специальности 37.05.02 «психология служебной деятельности» (специализация: психологическое обеспечение служебной деятельности сотрудников правоохранительных органов).

В пособии рассматриваются теоретические вопросы курса, после каждой темы предлагается перечень вопросов для самоконтроля.

ББК 88.6

ISBN 978-5-9694-0741-1

ОГЛАВЛЕНИЕ

Введение

Данное учебное пособие охватывает теоретический материал по темам: «Применение математических и статистических методов в психологии», «Основные понятия выборочного метода», «Статистическая обработка результатов исследований», «Сравнения выборок. Понятие статистических критериев и их виды», «Характеристики взаимосвязи признаков», «Выявление различий между признаками», «Методы многомерного анализа данных», а также содержит показательные примеры задач.

Представленные в пособии математические методы необходимы для освоения курсов психодиагностики и экспериментальной психологии, а также для выполнения курсовых и дипломных работ.

Обучающиеся должны:

Знать измерения в психологии, первичную обработку и представления эмпирических психологических связей, математико-статистические методы и процедуры, используемые для анализа и обработки результатов психологических исследований.

Уметь строить статистические предсказания; применять математические методы, адекватные методологическому аппарату исследования; получать, обрабатывать и интерпретировать данные исследований с помощью математико-статистического аппарата.

Владеть методами первичной и вторичной обработки полученных данных, способами определения эффективности применения математического метода; базовыми методами и процедурами проведения психологических исследований и экспериментов, обработки и описания эмпирических данных, анализа и интерпретации полученных результатов.

ТЕМА 1. Применение математических и статистических методов в психологии

1.1. Предмет и методы математической статистики

Математическая статистика – научная дисциплина, изучающая разработку методов регистрации, описания и анализа статистических экспериментальных данных, полученных в результате наблюдений массовых случайных явлений.

Исходя из определения, можно выделить основные *задачи математической статистики*:

- определение закона распределения случайной величины или системы случайных величин;
- проверка значимости гипотез;
- определение неизвестных параметров распределения.

Любое статистическое исследование начинается с эксперимента. Числовые показатели, составляющие эксперимент, называются *статистическими* и представляют собой первоначальный материал исследования. Для придания им научной и (или) практической ценности используются методы математической статистики.

Предмет статистического исследования представляет собой статистические совокупности – множество однокачественных варьирующих предметов.

К методам математической статистики можно отнести:

1) сбор данных (статистическое наблюдение, публикация);
2) обобщение данных (сводка, группировка);
3) представление данных (таблицы и графики);
4) анализ и интерпретация числовых данных (расчет средних величин, вариационного анализа, КРА, ряды динамики, индексы).

В основе всех методов математической статистики лежит теория вероятностей. Однако не стоит забывать, что в этой научной дисциплине подбирается подходящая теоретико-вероятностная модель, исходя из статистических данных, тогда как в теории вероятностей считается заданной модель явления и производится расчет возможного реального течения этого явления.

1.2. Типы шкал и их сравнительная характеристика

Понятие статистического измерения

Измерение представляет собой процедуру, позволяющую сравнить измеряемый объект с некоторым эталоном и выразить его в определенном масштабе или шкале.

Для работы с измерением индивидуально-психологических особенностей (креативность, нейротизм, импульсивность, свойства нервной системы и т. п.) в психодиагностике разрабатываются специальные измерительные процедуры. Так как психологи зачастую используют экспериментальные методы и модели исследования психических явлений в ходе эксперимента, изучаемые характеристики могут преобразовываться в количественные показатели, применяемые для статистической обработки.

Каждый вид измерения предусматривает наличие единиц измерения. В психологических науках не представляется возможным использовать стандартные единицы измерения (градус, метр, ампер и т. д.), поэтому чаще всего при определении значения психологического признака используются специальные измерительные шкалы.

Существует *четыре типа измерительных шкал* (или способов измерения):

1) номинативная, номинальная или шкала наименований;
2) порядковая, ординарная или ранговая шкала;
3) интервальная или шкала равных интервалов;
4) шкала равных отношений или шкала отношений.

Измерения, реализуемые с помощью номинативной и порядковой шкал, считаются качественными (неметрические шкалы), а с помощью интервальной шкалы – количественными (метрические шкалы).

Результат измерения, осуществленного в определенной шкале, представляется в числовых кодах, позволяющих психологу-исследователю использовать соответствующие статистические операции к полученным экспериментальным данным.

Номинативная шкала

Сущность измерения по номинативной шкале (номинальной или шкале наименований) состоит в присваивании тому или иному явлению определенного обозначения или символа (численного, буквенного и т. п.). Процедура измерения сводится к классификации свойств, группировке объектов, объединению их в классы, группы при условии, что объекты, принадлежащие одному классу, идентичны (или

аналогичны) друг другу в отношении какого-либо признака или свойства, тогда как объекты, различающиеся по этому признаку, попадают в разные классы. Примером измерения по данной шкале в психологии является классификация людей по типам темперамента: сангвиник, холерик, флегматик и меланхолик.

Номинальная шкала подразумевает качественное отличие различных свойств и признаков, а не количественные операции над нами. Пункты номинативной шкалы невозможно ранжировать, а также определять, являются ли одни из них более или менее значимыми.

Наиболее элементарной номинативной шкалой является *дихотомическая.* При измерениях по ней исследуемые признаки можно кодировать двумя символами или цифрами, например, 0 и 1 или 2 и 6, или буквами А и Б, а также любыми двумя отличающимися друг от друга символами. Признак, измеренный по дихотомической шкале, называется *альтернативным.*

В дихотомической шкале все объекты, признаки или изучаемые свойства разбиваются на два независимых класса, и ставится вопрос о том, проявился ли интересующий признак у испытуемого или нет. Например, в исследовании признак «аттестованный сотрудник» проявился у 52 пациентов из 80, т. е. 52 пациентам можно поставить, цифру 1, соответствующую признаку «аттестованный сотрудник», остальным – цифру 0 – «неаттестованный сотрудник». Таким образом, образуются два непересекающихся множества, применительно к которым можно только подсчитать количество индивидов, обладающих тем или иным признаком.

Номинальная шкала позволяет подсчитать частоту встречаемости признака, число испытуемых, явлений и т. д., оказавшихся в конкретной группе и располагающих данным свойством. Например, нам необходимо определить количество аттестованных и неаттестованных сотрудников, вставших на учет в поликлинике. Для этого кодируем аттестованных сотрудников цифрой 1, а неаттестованных – цифрой 0. Затем подсчитываем общее количество цифр (кодов) 1 и 0. Это и есть подсчет частоты признака.

Единицей измерения, используемой в номинальной шкале, является количество наблюдений (испытуемых, свойств, реакций и т. п.). Общее число наблюдений берется за 100 %, и в таком случае у нас появляется возможность рассчитать процентное соотношение сотрудников из вышеуказанного примера. В том случае, если количество групп разбиения больше, чем две, то можно определить про-

центный состав испытуемых (респондентов) в каждой группе. Также можно найти группу, состоящую из наибольшего числа респондентов, т. е. группу с наибольшей частотой измеренного признака. Эта группа носит название *моды*.

Однако результаты измерений, полученные в номинальной шкале, можно обработать небольшим числом статистических методов.

Еще одним примером измерения по номинативной шкале является разбиение людей на типы поведения в конфликтной ситуации (соперничество, сотрудничество, компромисс, избегание, приспособление).

Порядковая шкала

Измерение по порядковой шкале разделяет всю совокупность измеренных признаков на множества, связанные между собой отношениями: «больше – меньше», «выше – ниже», «сильнее – слабее» и т. д. В данной шкале все признаки располагаются по рангу – от самого большого (высокого, сильного, умного и т. п.) до самого маленького (низкого, слабого, глупого и т. п.) или наоборот.

Примером порядковой шкалы может являться балл по одной из учебных дисциплин (от 10 до 1 балла). Порядковая (ранговая) шкала должна состоять как минимум из трех классов (групп): например, «да», «не знаю», «нет» или «низкий», «средний», «высокий» и т. д. с тем расчетом, чтобы можно было расставить измеренные признаки по порядку. Именно поэтому эта шкала и называется порядковой или ранговой шкалой.

От классов просто перейти к числам, если считать, что низший класс получает ранг (код или цифру) 1, средний – 2, высший – 3 (или наоборот). Большее количество классов разбиений всей экспериментальной совокупности позволяет шире использовать возможности статистической обработки полученных данных и проверки статистических гипотез.

Ранжирование выполняется следующим образом: максимальный ранг приписывается наиболее значимому признаку, а минимальный – наименее значимому. Остальным признакам, в соответствии со степенью их значимости, приписываются промежуточные цифры (ранги). Процедура ранжирования является формальной, поэтому в зависимости от предпочтения можно проставлять величины рангов и в противоположном порядке, т. е. наиболее значимому признаку приписать ранг 1, а наименее значимому – максимальный ранг.

Необходимо отметить, что ранжированию подлежат не только качественные, но и количественные признаки каких-либо измеренных психологических свойств.

Для проверки правильности ранжирования используется следующая формула:

$$S_n = n \cdot \frac{n+1}{2},$$

где n – количество ранжируемых признаков.

В том случае, если определены одинаковые ранги исследуемых признаков, им приписывается величина среднего арифметического от количества условных рангов, проставляемых по порядку их величин.

Рассмотрим основные правила ранжирования:

1. Наименьшему числовому значению присваивается ранг 1.

2. Наибольшему числовому значению присваивается ранг, равный количеству ранжируемых величин.

3. В случае, если несколько исходных числовых значений оказались равными, то им присваивается ранг, равный средней величине тех рангов, которые эти величины получили бы, если бы они стояли по порядку друг за другом и не были бы равны. Напомним, что к этому случаю можно отнести как первые, так и последние величины исходного ряда для ранжирования.

4. Общая сумма реальных рангов должна совпадать с расчетной, определяемой по формуле: $S_n = n \cdot \frac{n+1}{2}$.

5. Не рекомендуется ранжировать более чем 20 величин (признаков, качеств, свойств и т. п.), поскольку в этом случае ранжирование в целом оказывается малоустойчивым.

6. При необходимости ранжирования достаточно большого числа объектов их следует объединять по какому-либо признаку в достаточно однородные классы (группы), а затем уже ранжировать полученные классы (группы).

Пример. Проранжировать следующий ряд значений: 40, 35, 36, 42, 39, 39, 35, 42, 39.

Решение. Сначала упорядочим данные в порядке возрастания: 35, 35, 36, 39, 39, 39, 40, 42, 42. Затем заполним следующую таблицу:

Таблица 1

Ранжирование статистических рядов

Ряд значений	Начальный ранг	Окончательный ранг
35	1	1,5
35	2	1,5
36	3	3
39	4	5
39	5	5
39	6	5
40	7	7
42	8	8,5
42	9	8,5
Сумма	**45**	**45**

В 3-м столбце окончательные ранги вычисляются, как средние арифметические рангов из 2-го столбца, если соответствующие значения в первом столбце равны.

Например, в первых двух строках в 1-м столбце стоят значения 35, во 2-м – 1 и 2, следовательно, в 3-м должно стоять (1 + 2) / 2 = 1,5 и т. д. Общая сумма рангов должна совпадать со значением, полученным по соответствующей формуле. Для нашего примера n = 9, следовательно, общая сумма рангов равна 45 ($S_9 = 9 \cdot \frac{9+1}{2} = 45$).

Проранжировать ряд значений можно с помощью пакета MS Excel (рис. 1). Для этого используем функцию РАНГ.СР. В графе Число указываем первое значение первой выборки. В графе Ссылка определяем список чисел (объединенные эмпирические данные 1-й и 2-й выборок). В графе Порядок указываем 1 (ИСТИНА, если сортировка производилась по возрастанию) или 0 (ЛОЖЬ, если данные сортировались по убыванию).

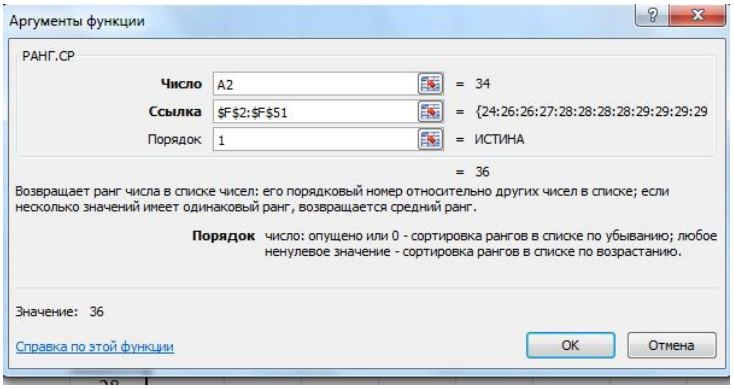

Рис. 1. Применение функции РАНГ.СР

Не забываем определить, что ячейки списка абсолютные. Используя функцию автозаполнения, заполним все ячейки 1-й и 2-й выборок.

Интервальная шкала

Интервальная шкала используется в том случае, если каждое из возможных значений измеренных величин стоит от ближайшего на равном расстоянии.

Интервал представляет собой долю или часть измеряемого свойства между двумя соседними позициями на шкале. Размер интервала – величина, фиксированная и постоянная на всех участках шкалы. Для измерения с помощью интервальной шкалы устанавливают специальные единицы измерения – в психологии это стены. При работе с этой шкалой измеряемому свойству или предмету присваивается число, равное количеству единиц измерения, эквивалентное количеству имеющегося свойства. Важной особенностью интервальной шкалы является то, что у нее нет естественной точки отсчета (ноль условен и не указывает на отсутствие измеряемого свойства).

Только измерение по строго стандартизированной тестовой методике, при условии того, что распределение значений в репрезентативной выборке достаточно близко к нормальному, может считаться измерением в интервальной шкале (рис. 2).

Рис. 2. Шкала интервалов

Не стоит забывать, что к экспериментальным данным, полученным в интервальной шкале, применимо достаточно большое число статистических методов.

Пример. Построить шкалу стенов для распределения с параметрами: $M = 11{,}8$ и $\sigma = 1{,}6$, выбрав количество интервалов равное 5.

Решение. На числовой оси влево и вправо от среднего значения откладываются интервалы равные 3σ. То есть интересующий нас интервал – (7; 16,6). Длина найденного интервала равна 9,6. Разделим ее на количество интервалов, которое по условию задачи равно 5, и найдем длину шага разбиения:

$$h = \frac{9{,}6}{5} = 1{,}92.$$

Также влево и вправо от среднего значения отложим интервалы, равные *h*/2. Получаем интервал (10,84; 12,76). Данный интервал – это нулевой стен. Теперь к правому краю интервала прибавляем *h* и получаем правую границу 1-го стена − 14,68. Таким образом, 1-й стен – это интервал (12,76; 14,68). Все, что на числовой оси находится правее 1-го стена, – 2-й стен. Другими словами, 2-й стен не имеет правой границы и равен интервалу (14,68; +∞). Аналогично 1-й стен – это интервал (8,92; 10,84), а 2-й – (−∞; 8,92) (рис. 3).

Любому «сырому» значению, попадающему, например, в интервал (8,92; 10,84), присваивается значение −1 по шкале стенов. Скажем, если полученное эмпирическое значение какого-либо свойства равно 10,04, то ему присваивается значение −1; если 8,06, то −2 и т. д.

Пример. Построение шкалы в единицах стандартного отклонения (шкалы стенов) – «стандартную десятку». За точку отсчета принимается среднее значение *M* = 10,2 в «сырых» баллах. Влево и вправо отмеряются интервалы длиной 0,5σ (σ = 2,4). Построенная шкала стенов показана на рис. 3.

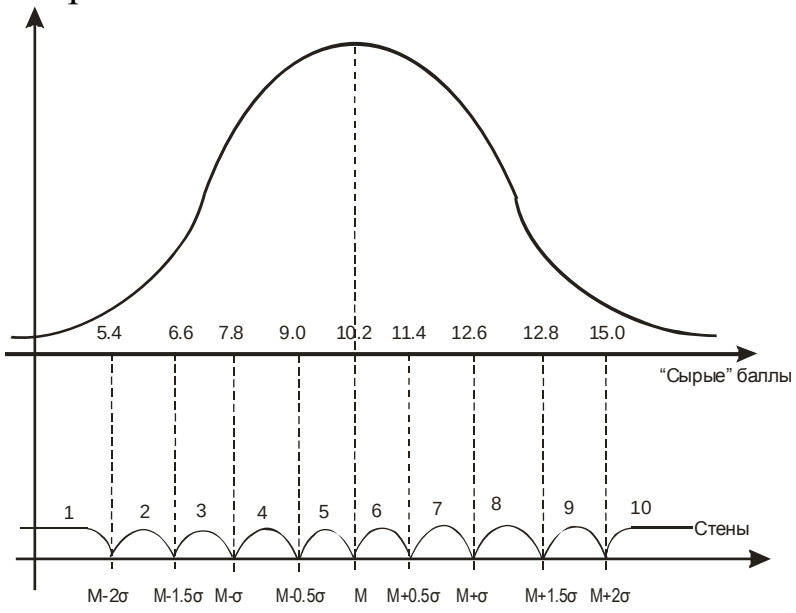

Рис. 3. Шкала стенов

Шкала отношений

Особенностью шкалы отношений (шкалы равных отношений) является наличие твердо фиксированного нуля, означающего полное отсутствие какого-либо свойства или признака. Например, может отсутствовать агрессивность личности – тогда уровень агрессии данного индивида равен нулю. Данная шкала является наиболее информативной шкалой, допускающей любые математические операции и использование разнообразных статистических методов.

Шкала отношений по сути очень близка интервальной, поскольку если строго фиксировать начало отсчета, то любая интервальная шкала превращается в шкалу отношений.

Именно в шкале отношений производятся точные и сверхточные измерения в таких науках, как физика, химия, микробиология и др. Измерения по шкале отношений производятся и в близких к психологии науках, таких как психофизика, психофизиология, психогенетика.

1.3. Основные этапы статистической обработки

1. *Предварительный этап.* После проведения исследования и обработки протокола (бланка ответов) составляется исходная матрица первичных данных, например, следующего вида (табл. 2).

Таблица 2

Коэффициент развития вербального интеллекта на психологическом факультете

№ п/п	Возраст	Пол	Коэффициент развития вербального интеллекта на первом курсе	Коэффициент развития вербального интеллекта на третьем курсе
1			34	38
2			31	39
3			27	29
4			29	36
5			31	36
6			35	37
7			29	34
8			29	39
9			31	32
…..	….	…..	…..	….
24			28	36
25			26	32

Заполнение таблицы исходных данных является началом работы в любом пакете статистического анализа на компьютере.

2. *Первый этап.* На данном этапе описываются результаты, полученные в данной выборке, соответственно каждому признаку. Этот этап является обязательным в любом исследовании – с него начинается анализ результатов. Необходимо представить среднестатистического испытуемого в данной выборке.

3. *Второй этап.* Установление взаимосвязей между признаками, т. е. проведение корреляционного анализа.

4. *Третий этап.* Сравнение выборок с целью выявления различий между ними или установления их сходства.

5. *Четвертый этап.* Использование многомерных методов анализа в зависимости от решаемых исследовательских задач.

Вопросы для самоконтроля

1. Предмет, методы и задачи математической статистики.
2. Понятие статистического измерения.
3. Сравнительная характеристика шкалы наименований.
4. Сравнительная характеристика порядковой шкалы.
5. Сравнительная характеристика шкалы отношений.
6. Сравнительная характеристика интервальной шкалы.
7. Основные правила ранжирования.
8. Основные этапы статистической обработки.

ТЕМА 2. Основные понятия выборочного метода

2.1. Понятие статистической совокупности

Существует множество определений статистики. Приведем наиболее точные, на наш взгляд.

Статистика: 1) это наука, изучающая массовые явления, происходящие в обществе с точки зрения качества и количества; 2) это наука, разрабатывающая методы статистики для исследований, проводящихся в различных областях.

Самые существенные и общие признаки, свойства, связи явлений и предметов, существующих объективно, отражают определенные понятия или категории. С их помощью статистика изучает собственный предмет. Одной из таких категорий является статистическая совокупность.

Статистическая совокупность представляет собой множество меняющихся явлений, существующих в пространстве и во времени. Все они относительно однородны и имеют признаки, сближающие или различающие их.

Свойства статистической совокупности

1. *Неразложимость.* При появлении или исчезновении каких-либо элементов статистической совокупности ее качественная основа не разбивается. Например, характеристика общности сотрудников органов внутренних дел не меняется в зависимости от того, что каждый год сюда поступают на работу новобранцы, а пенсионеры выходят на пенсию.

2. *Однородность.* Статистическая совокупность всегда имеет минимум один общий признак для всех ее элементов. И в то же время это свойство не является одинаковым для них – этот признак может иметь разные значения для разных единиц. То, насколько статистическая совокупность однородна, устанавливается во время исследования. И зависит это свойство, прежде всего, от тех целей и задач, с которыми оно проводится. Например, проводя исследование психологических особенностей юношей подросткового возраста, в выборку мы не можем включить ни девушек подросткового возраста, ни детей, ни мужчин.

3. *Вариация.* Во время перехода от одного элемента статистической совокупности к другому происходит изменение количественного значения признака. Оно может быть одинаковым для всех компонентов. В таком случае, чтобы получить представление о том, что та-

кое статистическая совокупность, можно не изучать ее полностью, а рассмотреть только один составляющий элемент. На возникновение вариации влияет целый комплекс причин и условий. Их нахождением и выяснением занимаются экономические дисциплины. А статистика лишь оценивает с количественной точки зрения, как каждая причина воздействует на вариацию определенного признака. Информация об этом помогает принимать правильные управленческие решения.

Группы статистических совокупностей

1. Созданные естественным путем. Они уже образуют единство, независимо от того, подвергаются исследованию или нет. Например, совокупности работников образования или здравоохранения. Данные совокупности подлежат измерению.

2. Сформированные искусственным образом, специально для проведения исследования. Например, статистическая совокупность курсантов разного типа поведения в конфликтных ситуациях.

3. Гипотетические (предполагаемые) множества. Это совокупности стохастические (например, небесных тел, существующих во всей Галактике).

Любая статистическая совокупность имеет характерные признаки.

Классификация признаков статистической совокупности

В зависимости от характера выражения признаки бывают:

• Описательные (атрибутивные). Они выражаются при помощи слов (группа здоровья, национальность, пол, принадлежность к тому или иному роду деятельности). Эти признаки позволяют подытожить, какое количество единиц обладает тем или иным значением.

• Количественные. Измеряются в числах (например, возраст, стаж работы, величина дохода или расхода и др.). Эти признаки позволяют подытожить не только количество единиц, обладающих определенным значением, но и их среднее или суммарное значение отдельно по совокупностям.

2.2. Генеральная и выборочная совокупности и их характеристики

Статистическое исследование состоит из множества данных, полученных в результате измерения одного или нескольких признаков. Основные параметры генеральной и выборочной совокупностей представлены в табл. 3.

Таблица 3

Основные параметры генеральной и выборочной совокупностей

Характеристики параметров распределения	Совокупность	
	Генеральная	Выборочная
Объем выборки	N	n
Альтернативный признак		
Численность единиц совокупности, обладающих признаком x	N_x	n_x
Доля единиц, обладающих изучаемым признаком x	$p = \frac{N_x}{N}$	$w = \frac{n_x}{n}$
Дисперсия	$\sigma^2 = p \cdot (1-p)$	$\sigma^2 = w \cdot (1-w)$
Среднее квадратическое отклонение	$\sigma = \sqrt{p \cdot (1-p)}$	$\sigma = \sqrt{w \cdot (1-w)}$
Количественный признак		
Среднее значение признака	$\mu = \frac{\sum x_i}{N}$	$\bar{x} = \frac{\sum x_i}{n}$
Дисперсия	$\sigma = \frac{\sum (x_i - \mu)^2}{N-1}$	$s^2 = \frac{\sum x_i - x)^2}{n-1}$
Среднее квадратическое отклонение	$\sigma^2 = \sqrt{\frac{\sum (x_i - \mu)^2}{N-1}}$	$s^2 = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$

Выборка – это реально наблюдаемая совокупность объектов, статистически представленная рядом наблюдений $x_1, x_2, \ldots, x_n$ случайной величины X. А гипотетически существующая (домысливаемая) совокупность объектов представляет собой генеральную совокупность. *Генеральная совокупность* может быть конечной (число наблюдений N = const) или бесконечной ($N = \infty$), а выборка из генеральной совокупности всегда ограничена количеством наблюдений. Количество наблюдений n, образующих выборку, называется объемом выборки. Если объем выборки n достаточно велик ($n \to \infty$), выборка считается большой, в противном случае она называется выборкой ограниченного объема. Выборка считается малой, если при измерении одномерной случайной величины X объем выборки не превышает 30 ($n <= 30$), а при измерении одновременно нескольких (k) признаков в многомерном пространстве отношение n к k не превышает 10 ($n/k < 10$).

Выборка образует вариационный ряд, если ее элементы являются порядковыми статистиками, т. е. выборочные значения случайной ве-

личины X упорядочены по возрастанию (ранжированы), значения же признака называются вариантами.

Пример. Практически одна и та же случайно отобранная совокупность объектов курсантов Московского университета МВД России имени В.Я. Кикотя может рассматриваться как выборка из генеральной совокупности всех вузов Москвы, как выборка из генеральной совокупности всех подразделений МВД, а также из всех вузов России.

Основные характеристики параметров генеральной и выборочной совокупностей

В основе статистических выводов проведенного исследования лежит распределение случайной величины X, наблюдаемые же значения ($x_1, x_2, \ldots, x_n$) называются реализациями случайной величины X (n – объем выборки). Распределение случайной величины X в генеральной совокупности носит теоретический, идеальный характер, а ее выборочный аналог является эмпирическим распределением. Некоторые теоретические распределения заданы аналитически, т. е. их параметры определяют значение функции распределения $F(x)$ в каждой точке пространства возможных значений случайной величины X. Для выборки же функцию распределения определить трудно, а иногда невозможно, поэтому параметры оценивают по эмпирическим данным, а затем их подставляют в аналитическое выражение, описывающее теоретическое распределение. При этом предположение (или гипотеза) о виде распределения может быть как статистически верным, так и ошибочным. Но в любом случае восстановленное по выборке эмпирическое распределение лишь грубо характеризует истинное. Важнейшими параметрами распределений являются математическое ожидание μ и дисперсия σ^2.

По своей природе распределения бывают непрерывными и дискретными. Наиболее известным непрерывным распределением является нормальное. Выборочными аналогами параметров μ и σ^2 для него являются: среднее значение ($\overline{X}$) и эмпирическая дисперсия (S^2). Среди дискретных в социально-экономических исследованиях наиболее часто применяется альтернативное (дихотомическое**)** распределение. Параметр математического ожидания μ этого распределения выражает относительную величину (или долю) единиц совокупности, которые обладают изучаемым признаком χ (она обозначена буквой ρ); доля совокупности, не обладающая этим признаком, обо-

значается буквой q ($q = 1 - p$). Дисперсия же σ^2 альтернативного распределения также имеет эмпирический аналог S^2.

Долей выборки k_n называется отношение числа единиц выборочной совокупности к числу единиц генеральной совокупности:

$$k_n = \frac{n}{N}.$$

Выборочная доля w – это отношение единиц, обладающих изучаемым признаком x к объему выборки n:

$$w = \frac{n_n}{n}.$$

Пример. Среди курсантов психологического факультета Московского университета МВД России имени В.Я. Кикотя (300 чел.) при 5 % выборке (участвуют в самодеятельности) доля выборки k_n в абсолютной величине составляет 15 чел. Если же в этой выборке 2 чел. танцуют сальсу, то выборочная совокупность танцующих сальсу w составит 0,13 ($w = 2/15 = 0,13$ или 13 %).

2.3. Сплошное и выборочное исследования

Выборочное исследование представляет собой способ систематического сбора данных о поведении, характере, темпераменте и т. д. людей посредством опроса специально подобранной группы респондентов, дающих информацию о себе и своем мнении.

Выборочное исследование осуществляется с помощью отбора из общей совокупности единиц исследования (генеральной совокупности) небольшой части (выборочной совокупности), максимально точно отражающей основные параметры целого. Процедура построения выборки основана на методах математической статистики и опирается на принципы теории вероятности. Выборочное исследование более экономично и не менее надежно, чем сплошное, хотя требует более искусной методики и техники.

Вопросы, на которые отвечают случайно попавшие в выборочную совокупность респонденты, могут быть как письменными, так и устными. В первом случае выборочное исследование называется анкетированием, во втором – интервьюированием. Помимо традиционного анкетирования и традиционного интервью используются такие методики, как интернет-опрос, телефонное интервью и т. д.

Сплошное исследование представляет собой исследование, охватывающее все единицы генеральной совокупности.

Примером сплошного исследования, охватывающим все единицы генеральной совокупности, может служить перепись населения.

2.4. Объем выборки и виды выборок

Объем выборки – число случаев, включенных в выборочную совокупность.

Выборки можно условно разделить на большие и малые, так как в математической статистике используются различные подходы в зависимости от объема выборки. Как правило, выборки объемом более 30 можно отнести к большим.

Для того чтобы определить объем выборки, необходимо сформулировать задачи исследования. Психолог может изучать единичные случаи, если те по каким-либо причинам представляют особый интерес для науки (одаренные дети, имеющие свои неповторимые особенности, сросшиеся сиамские близнецы Маша и Даша). Когда психолог ставит целью изучение характеристик, присущих многим представителям генеральной совокупности, возникает вопрос о наиболее приемлемом объеме выборки. В этих случаях очевидно, что больший объем выборки позволяет получить более надежные результаты. Объем выборки зависит также от степени однородности изучаемого явления. Как правило, чем более однородно изучаемое явление, тем меньше может быть объем выборки. Например, психолог изучает выраженность уровня маскулинности – феминности у мастеров спорта по хоккею. Поскольку подобная группа спортсменов представляет собой достаточно однородную выборку, то ее объем может быть весьма небольшим, например, в пределах одной команды – 12–20 человек.

Для психологических исследований рекомендуется использовать экспериментальную и контрольную группы таким образом, чтобы численность обеих сравниваемых групп была не менее 30–35 испытуемых в каждой.

Виды выборок

При сравнении двух (и более) выборок важным параметром является их зависимость. Если можно установить гомоморфную пару (когда одному случаю из выборки *X* соответствует один и только один случай из выборки *Y* и наоборот) для каждого случая в двух выборках (и это основание взаимосвязи является важным для измеряемого на выборках признака), такие выборки называются *зависимыми*, например:

- пары близнецов;

- два измерения какого-либо признака до и после экспериментального воздействия;
- супруги и т. д.

В случае, если такая взаимосвязь между выборками отсутствует, то эти выборки считаются *независимыми*, например:

- курсанты и слушатели;
- психологи и математики.

Соответственно, зависимые выборки всегда имеют одинаковый объем, а объем независимых выборок может отличаться.

Существуют два варианта составления выборок. После того, как объект отобран и над ним произведено наблюдение, он может быть возвращен либо не возвращен в генеральную совокупность. В связи с этим выборки подразделяют на повторные и бесповторные.

Повторной называют выборку, при которой отобранный объект (перед отбором следующего) возвращается в генеральную совокупность. *Бесповторной* называют выборку, при которой отобранный объект в генеральную совокупность не возвращается. На практике обычно пользуются бесповторным случайным отбором.

Требования к выборке

1. Однородность выборки.

К выборке предъявляется ряд обязательных требований, определенных прежде всего целями и задачами исследования. Планирование эксперимента обязательно включает в себя учет как объема выборки, так и ряда ее особенностей. Одно из важнейших требований в психологических исследованиях – однородность выборки, означающая, что психолог, изучая, например, курсантов психологического факультета, не может включать в эту же выборку курсантов следственного факультета. Основаниями для формирования однородной выборки могут служить разные характеристики, такие как уровень интеллекта, национальность, отсутствие определенных заболеваний и т. д., в зависимости от целей исследования.

2. Репрезентативность выборки.

Суть репрезентативности выборки заключается в следующем: для того, чтобы по данным выборки можно было достаточно уверенно судить об интересующем нас признаке генеральной совокупности, необходимо, чтобы объекты выборки правильно его представляли. В силу закона больших чисел можно утверждать, что выборка будет репрезентативной, если ее осуществить случайно: каждый объект вы-

борки отобран случайно из генеральной совокупности, если все объекты имеют одинаковую вероятность попасть в выборку. Итак, *репрезентативная выборка* – это такая выборка, в которой все основные признаки генеральной совокупности представлены приблизительно в той же пропорции и с той же частотой, с которой данный признак выступает в данной генеральной совокупности. Иными словами, репрезентативная выборка представляет собой меньшую по размеру, но точную модель той генеральной совокупности, которую она должна отражать. В той степени, в какой выборка является репрезентативной, выводы, основанные на изучении этой выборки, можно с большой долей уверенности считать применимыми ко всей генеральной совокупности. Это распространение результатов называется генерализуемостью.

2.5. Формирование репрезентативной выборки

Для репрезентативности выборки используют различные способы отбора. Принципиально эти способы можно подразделить на две группы:

1) если генеральная совокупность не разбивается на части, получается простой случайный повторный или бесповторный отбор;
2) если генеральная совокупность разбивается на части, получается отбор одного из трех видов: типический, механический или серийный.

Первый метод – формирование простой случайной выборки. В этом случае выборка состоит из элементов, отобранных из генеральной совокупности таким образом, чтобы каждый элемент этой совокупности имел бы равные возможности (равную вероятность) попасть в выборку. Полученная таким образом выборка называется простой случайной выборкой.

Осуществить простой случайный отбор можно различными способами. Например, для извлечения n объектов из генеральной совокупности объема N поступают так: выписывают номера от 1 до N на карточках, которые тщательно перемешивают, и наугад вынимают одну карточку; объект, имеющий одинаковый номер с извлеченной карточкой, подвергают обследованию; затем карточка возвращается в пачку, и процесс повторяется, т. е. карточки перемешиваются, наугад вынимают одну из них и т. д. Так поступают n раз; в итоге получают простую случайную повторную выборку объема n. Если извлеченные карточки не возвращать в пачку, то выборка будет простой, случайной, бесповторной.

Получить простую случайную выборку можно путем обычной жеребьевки (по аналогии с лотереей) или с помощью специальных таблиц случайных чисел, в которых числа расположены в случайном порядке. Для того, чтобы отобрать, например, 30 объектов из пронумерованной генеральной совокупности, открывают любую страницу таблицы случайных чисел и выписывают подряд 30 чисел; в выборку попадают те объекты, номера которых совпадают с выписанными случайными числами. Если бы оказалось, что случайное число таблицы превышает число *N*, то такое случайное число пропускают. При осуществлении бесповторной выборки случайные числа таблицы, уже встречавшиеся ранее, следует также пропустить.

Второй метод основывается на понятии стратифицированной случайной выборки. Для этого необходимо разбить элементы генеральной совокупности на страты (группы) в соответствии с некоторыми характеристиками. Например, при исследовании поведения личности в конфликтных ситуациях, генеральную совокупность желательно разбить на группы, различающиеся по должности, возрасту, стажу работы, образованию, специальности. В результате случайная выборка производится отдельно из каждой группы (страты), следовательно, является стратифицированной случайной выборкой.

Рассмотрим способы расклада генеральной совокупности на страты.

1. *Типический.* Объекты отбираются не из всей генеральной совокупности, а из каждой ее типической части. Например, если в Университете МВД России готовят специалистов различных специальностей, то отбор производят не из общего числа курсантов, а из определенного исследуемого факультета.

2. *Механический.* Генеральная совокупность механически делится на столько групп, сколько объектов должно войти в выборку, и из каждой группы отбирается один объект.

Например, если нужно отобрать 10 % испытуемых генеральной совокупности, то отбирается каждый десятый; если нужно отобрать 20 % испытуемых генеральной совокупности, то отбирается каждый пятый и т. д.

3. *Серийный.* Объекты отбирают из генеральной совокупности не по одному, а сериями, которые подвергают сплошному обследованию. Например, если исследуется психологический факультет, то подвергают сплошному обследованию одну группу. Серийный отбор используют тогда, когда исследуемый признак колеблется в различных сериях незначительно.

На практике часто применяется *комбинированный отбор*, при котором сочетаются указанные выше способы. Например, иногда разбивают генеральную совокупность на группы одинакового объема, затем простым случайным отбором выбирают несколько групп, и, наконец, из каждой группы простым случайным отбором извлекают отдельные объекты.

Вопросы для самоконтроля

1. Определение статистической совокупности.
2. Свойства статистической совокупности.
3. Группы статистической совокупности.
4. Генеральная и выборочная совокупности. Их характеристики.
5. Определение сплошного и выборочного исследований.
6. Виды выборок и их объем.
7. Требования, предъявляемые к выборке.
8. Формирование репрезентативной выборки.
9. Способы расклада генеральной совокупности на страты.

ТЕМА 3. Статистическая обработка результатов исследований

3.1. Параметры теоретических распределений и их эмпирическая оценка

Пусть из генеральной совокупности извлечена выборка, причем x_1 наблюдалось n_1 раз, $x_2 - n_2$ раз, …, $x_k - n_k$ раз и $\sum_{i=1}^{k} n_i = n -$ объем выборки. Наблюдаемые значения x_i называют *вариантами*, а последовательность вариант, записанных в возрастающем порядке, – *вариационным рядом*. Числа наблюдений называют частотами, а их отношения к объему выборки $\frac{n_i}{n} = W_i$ – *относительными частотами*.

Пример. Задано распределение частот выборки объема $n = 26$:

x_i	4	8	14
n_i	6	12	8

Написать распределение относительных частот.

Решение. Найдем относительные частоты, для чего разделим частоты на объем выборки:

$$W_1 = \frac{6}{26} = 0{,}23;\ W_2 = \frac{12}{26} = 0{,}46;\ W_1 = \frac{8}{26} = 0{,}31.$$

Напишем распределение относительных частот:

x_i	4	8	14
W_i	0,23	0,46	0,31

Контроль: 0,23 + 0,46 + 0,31 = 1.

Эмпирическая функция распределения

Пусть известно статистическое распределение частот количественного признака X. Число наблюдений, при которых наблюдалось значение признака, меньшее $x - n_x$. Объем выборки – n. Относительная частота события $X < x$ равна $\frac{n_x}{n}$. Если x изменяется, то изменяется и относительная частота, т. е. относительная частота $\frac{n_x}{n}$ представляет собой функцию от x.

Данная функция называется эмпирической в связи с тем, что она определяется эмпирическим путем.

Эмпирическая функция распределения – функция $F^*(x)$, которая определяет для каждого значения x относительную частоту события $X < x$. Рассчитывается по формуле:

$$F^*(x) = \frac{n_x}{n},$$

где n_x – число вариант, меньших x;

n – объем выборки.

Теоретическая функция распределения представляет собой функцию распределения $F(x)$ генеральной совокупности.

Теоретическая функция $F(x)$ определяет вероятность события $X < x$, а эмпирическая функция $F^*(x)$ определяет относительную частоту этого же события. Из теоремы Бернулли следует, что относительная частота события $X < x$, т. е. эмпирическая функция распределения, стремится по вероятности к вероятности $F(x)$ этого события. То есть при больших объемах теоретическая и эмпирическая функции распределения незначительно отличаются друг от друга.

Свойства эмпирической функции распределения

1. Значения эмпирической функции принадлежат отрезку [0; 1].
2. $F^*(x)$ неубывающая функция.
3. Если x_1 – наименьшая варианта, то $F^*(x) = 0$ при $x \le x_1$;

   $x_к$ – наибольшая варианта, $F^*(x) = 1$ при $x > x_к$.

Таким образом, эмпирическая функция распределения $F^*(x)$ оценивает теоретическую функцию распределения генеральной совокупности $F(x)$.

Пример. Построить эмпирическую функцию распределения данного вариационного ряда.

x_i	4	8	12
n_i	16	23	28

Решение. Найдем объем выборки: 16 + 23 + 28 = 67.

Наименьшая варианта равна 4, следовательно, $F^*(x) = 0$ при $x \le 4$.

Значение $x < 8$ ($x_1 = 4$) наблюдалось 16 раз, следовательно, $F^*(x) = \frac{16}{67} = 0{,}24$ при $4 < x \le 4$.

Значение $x < 12$ ($x_1 = 4$, $x_2 = 8$) наблюдалось 16 + 23 = 39 раз, следовательно, $F^*(x) = \frac{39}{67} = 0{,}58$ при $8 < x \le 12$.

$x_к = 10$ – наибольшая варианта, а значит $F^*(x) = 1$ при $x > 12$.

$$F^*(x) = \begin{cases} 0 \text{ при } x \le 4 \\ 0{,}24 \text{ при } 4 < x \le 8 \\ 0{,}58 \text{ при } 8 < x \le 12 \\ 1 \text{ при } x > 12. \end{cases}$$

Построим график этой функции (рис. 4):

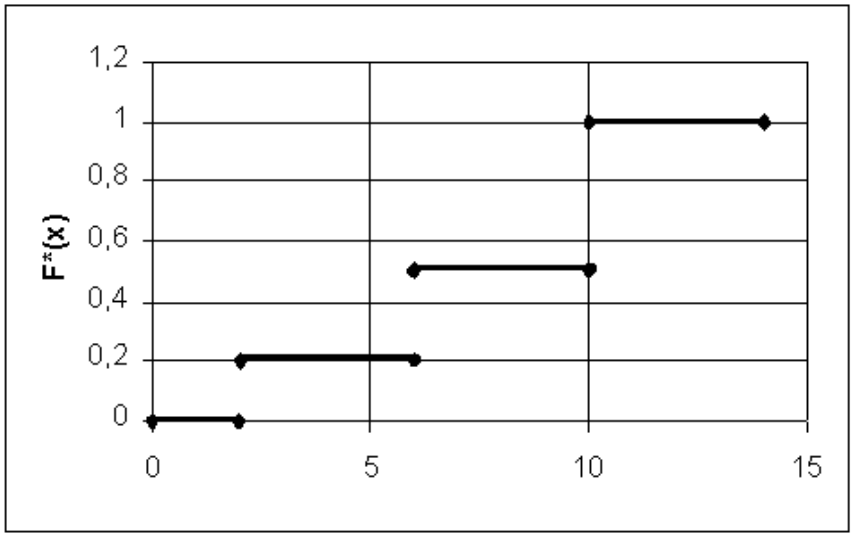

Рис. 4. Эмпирическая функция распределения выборки

3.2. Статистическое распределение выборки

Статистическое распределение выборки представляет собой перечень вариант и соответствующих им частот (относительных частот). Статистическое распределение можно задать также в виде последовательности интервалов и соответствующих им частот (в качестве частоты, соответствующей интервалу, принимают сумму частот, попавших в этот интервал).

Вариационный ряд – ряд, в котором сопоставлены (по степени возрастания или убывания) варианты и соответствующие им частоты.

Варианты – отдельные количественные выражения признака. Обозначаются x_i. Предполагается, что вариантой является уникальное значение признака без учета количества повторов.

Например, в вариационном ряду показателей веса у женщин: 45, 50, 50, 60, 60, 60, 65, 65, 70, 75, 80 – вариантами являются 7 значений:

Частота – число повтора вариант. Обозначается n_i. Сумма всех частот равна объему выборки (n).

В рассмотренном выше примере частоты принимают следующие значения:

x_i	45	50	60	65	70	75	80
n_i	1	2	3	2	1	1	1

Виды вариационных рядов

1. *Простой* – это ряд, в котором каждая варианта встречается только по одному разу (все частоты при этом равны 1).

2. *Взвешенный* – ряд, в котором одна или несколько вариант встречаются неоднократно.

Вариационный ряд служит для описания больших массивов чисел. Именно в этой форме изначально представляются собранные данные большинства психологических исследований. Для характеристики вариационных рядов рассчитываются специальные показатели, такие как: средние величины, показатели вариабельности (дисперсии), показатели репрезентативности выборочных данных.

Показатели вариационного ряда

Для того чтобы определить, какое значение наиболее характерно для выборки, используют меры центральной тенденции (среднее арифметическое, мода, медиана).

Меры центральной тенденции – величины, вокруг которых группируются остальные данные.

Среднее арифметическое значение – это обобщающий показатель, характеризующий размер изучаемого признака. Среднее арифметическое обозначается $\bar{x}$ и рассчитывается по формуле:

$$\bar{x} = \frac{(x_1 + x_2 + \ldots + x_n)}{n} = \frac{1}{n} \cdot \left(\sum_{i=1}^{n} x_i \right).$$

Среднее арифметическое заменяет индивидуальные варьирующие значения случайной величины на некоторую уравненную величину, сохраняющую основные свойства всех остальных значений. Но лишь в тех случаях, когда распределение случайной величины является равномерным или значения, близкие к среднему, встречаются часто, а удаленные от среднего – редко. Нахождение среднего арифметического не всегда отражает основные свойства совокупности данных. Например, если наблюдается большой разброс значений случайной величины или в совокупности присутствуют значения, резко отличающиеся от остальных.

Например, в фирме ООО «Рассвет» 30 сотрудников, из которых 27 получают 30 000 руб., двое – 5 000 руб., генеральный директор – 300 000 руб. Рассчитав среднюю заработную плату, получим: 37 333 руб. Однако фактически один сотрудник получает в десять раз больше, двое в шесть раз меньше основной части сотрудников.

В таком случае предлагается отбрасывать крайние минимальное и максимальное значения либо применить другую меру центральной тенденции (медиану).

Медиана – значение варианты, делящей вариационный ряд пополам: по обе стороны от нее находится равное число вариант. Обозначается как *Md.*

Для нахождения медианы необходимо, прежде всего, отсортировать ряд в порядке возрастания. В приведенном выше примере медиана равна 30 000.

Например, в вариационном ряду: 55, 45, 50, 65, 50, 55, 60 – медианой является значение 55. В том случае, если в вариационном ряду четное количество значений, медиану определяют как среднее арифметическое двух центральных значений.

Если отдельные значения выборки повторяются, $\bar{x}$ называют взвешенной средней и рассчитывают следующим образом:

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{k} x_i \cdot f_i ,$$

где f_i – частоты повторяющихся значений.

Мода – средняя величина вариационного ряда, соответствующая наиболее часто повторяющейся варианте. То есть это варианта, которой соответствует наибольшая частота. Обозначается как M_o. Мода рассчитывается только для взвешенных рядов, так как в простых рядах ни одна из вариант не повторяется, и все частоты равны единице.

Например, в вариационном ряду значений частоты роста: 165, 170, 170, 175, 175, 175, 175, 180, 185 – значение моды составляет 175, так как данная варианта встречается 4 раза, следовательно, ее частота наибольшая.

Правила нахождения моды

1. Если значения в выборке встречаются одинаковое количество раз, вариационный ряд не имеет моды. Например: 160, 160, 170, 170, 180, 180.

2. Если два (и более) смежных значения имеют одинаковую частоту, большую по сравнению с другими частотами, мода рассчитывается как среднее арифметическое этих значений. Например: 45, 50, 50, 55, 55, 60, 65, 70. В данном вариационном ряду частоты рядом расположенных значений 50 и 55 совпадают и равняются 2. Эта ча-

стота больше частоты других значений: 45, 60, 65, 70 (равна 1). Модой представленного ряда будет величина: $\frac{(50+55)}{2}=52{,}5$.

3. Когда два и более не смежных значения в выборке имеют равные частоты и они больше других частот, то определяют две (или более) моды. Например, в вариационном ряду: 160, 160, 160, 165, 165, 170, 175, 175, 175, 180, 180, 185 – модами являются значения 160 и 175. В данном случае выборка называется бимодальной. Существуют мультимодальные выборки, имеющие больше двух выборок.

Особенности мер центральной тенденции

1. Иногда среднее арифметическое принимает значение, вообще не встречающееся в распределении, либо представляет собой дробное значение для дискретной величины (например, среднее количество проданных автомобилей в автосалоне будет выражаться дробным числом, хотя ни в одном автосалоне оно не встречается).

2. Медиана представляет собой реальное значение, только если количество элементов в выборке нечетное. Использование медианы особенно удобно, если наблюдается большой размах выборки. Однако медиана неинформативна, если какое-либо значение в выборке встречается очень часто и оно расположено либо в начале вариационного ряда, либо в конце.

3. Для показателей, измеренных в шкале наименований, мода является единственной мерой центральной тенденции. Например, необходимо описать основные жизненные ориентации курсантов психологического факультета. Данную переменную можно измерить в шкале наименований. Наиболее часто встречаемая жизненная ориентация будет являться модой, единственно возможной и понятной всем мерой центральной тенденции. Если в распределении непрерывной случайной величины моды нет или каждое значение встречается не более одного раза, желательно применять среднее арифметическое или медиану.

Меры изменчивости (рассеивания, разброса)

Характеризуют различия между отдельными значениями выборки. Позволяют судить о степени однородности полученного множества.

Разброс (размах) выборки – разность между максимальной и минимальной величинами конкретного вариационного ряда. Рассчитывается по формуле:

$$R = X_{max} - X_{min}.$$

Дисперсия – наиболее часто используемая мера рассеяния случайной величины. Представляет собой среднее арифметическое квадратов отклонений значений переменной от ее среднего значения. Рассчитывается по формуле:

$$D = \frac{1}{n} \cdot \sum_{i=1}^{n} (X_i - \overline{X})^2,$$

где n – объем выборки;

i – индекс суммирования;

$\overline{X}$ – среднее арифметическое значение.

Алгоритм вычисления дисперсии

1. Рассчитать среднее по выборке.
2. Для каждого элемента выборки рассчитать его отклонение от средней, т. е. образовать множество T.
3. Каждый элемент полученного множества (T) возвести в квадрат.
4. Определить сумму этих квадратов.
5. Полученную сумму разделить на объем выборки (n). В том случае, если величина выборки мала, деление осуществляется на $n - 1$.
6. Полученная величина называется дисперсией.

Пример. Рассчитаем дисперсию для вариационного ряда: 4, 5, 6, 7, 8, 12.

Решение. Найдем среднее ряда $\overline{X} = 7$.

Рассмотрим величины: $(X_i - \overline{X})$ для каждого элемента.

Из каждого элемента представленного ряда вычтем величину среднего этого ряда. Мы получили величины, характеризующие, насколько каждый элемент отклоняется от средней величины в данном ряду. Запишем полученную совокупность разностей как множество T. В результате получим:

$T = (4 - 7 = -3; 5 - 7 = -2; 6 - 7 = -1; 7 - 7 = 0; 8 - 7 = 1; 12 - 7 = 5)$.

Образовался новый ряд чисел. При сложении всех элементов данного ряда получается ноль: $(-3) + (-2) + (-1) + 0 + 1 + 5 = 0$.

Необходимо отметить, что сумма такого ряда $\sum(X_i - \overline{X})$ всегда равна нулю. Для того чтобы избавиться от нуля, каждое значение разности $(X_i - \overline{X})$ возводят в квадрат, суммируют их, а затем делят на число элементов. В итоге получаем следующее:

$$\sum_{i=1}^{n}(X_i - \overline{X})^2 = (-3)\cdot(-3) + (-2)\cdot(-2) + (-1)\cdot(-1) + 0\cdot 0 + 1\cdot 1 + 5\cdot 5 = 40;$$

$$D = \frac{40}{7} = 5{,}7.$$

Пример. Из группы необходимо отправить на олимпиаду по математике трех курсантов. Оценки за первый семестр каждого из курсантов представлены в табл. 4.

Решение. Рассчитаем средние арифметические значения для каждого вариационного ряда.

Таблица 4

Успеваемость курсантов за 1 семестр

ФИО	Оценки за 1 семестр								Средние значения
	Сентябрь		Октябрь		Ноябрь		Декабрь		
	к/р	тест	к/р	тест	к/р	тест	к/р	тест	
Бабина Н. А.	3	4	3	4	4	3	3	4	3,50
Белкина Р. О.	4	5	5	3	5	5	3	4	4,25
Волкова Н. Е.	4	4	4	4	4	4	5	5	4,25
Воробьев К. И.	5	4	3	2	5	5	5	5	4,25
Воронцова А. И.	4	5	3	2	4	3	3	4	3,50
Зябликов В. А.	3	5	3	5	4	3	3	5	3,88
Иванов Н. Г.	4	3	5	3	4	3	5	3	3,75
Калинин В. В.	4	5	4	5	3	3	5	5	4,25
Перепелкин П. Р.	4	4	3	4	3	4	3	4	3,63
Орлова Г. О.	3	4	5	3	5	3	4	2	3,63

Как видно из табл. 3.1, опираясь только на средние оценки, лидерами являются: Белкина Р. О., Волкова Н. Е., Воробьев К. И., Калинин В. В. Рассчитаем дисперсию для представленных вариационных рядов (табл. 5).

Таблица 5

Значения дисперсии

ФИО	Средние значения	Дисперсия
Бабина Н. А.	3,5	0,25
Белкина Р. О.	4,25	0,6875
Волкова Н. Е.	4,25	0,1875
Воробьев К. И.	4,25	1,1875
Воронцова А. И.	3,5	0,75
Зябликов В. А.	3,88	0,8594

Продолжение табл. 5

Иванов Н. Г.	3,75	0,6875
Калинин В. В.	4,25	0,6875
Перепелкин П. Р.	3,63	0,2344
Орлова Г. О.	3,63	0,9844

Представленная выше табл. 5 свидетельствует о том, что наименьшее значение дисперсии наблюдается у Волковой Н. Е. (0,1875). Это означает, что в оценках данной курсантки наблюдается определенная стабильность.

Дисперсия не всегда удобна для интерпретации. Например, в эксперименте измерялся рост в сантиметрах, тогда размерность дисперсии будет являться характеристикой площади, а не линейного размаха (так как при определении дисперсии сантиметр возводят в квадрат).

Стандартное отклонение (среднее квадратическое отклонение) – мера вариабельности вариационного ряда. Представляет собой интегральный показатель, объединяющий все случаи отклонения вариант от средней. Определяет, насколько далеко и как часто варианты распространяются от средней арифметической. Обозначается как σ и рассчитывается как квадратный корень, извлеченный из дисперсии выборки:

$$\sigma = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (X_i - \overline{X})^2} .$$

Рассчитаем стандартное отклонение для предыдущего примера: $\sigma = \sqrt{\frac{40}{7}} = 2{,}39$.

Необходимо отметить, что размерность стандартного отклонения и размерность исходного ряда совпадают.

3.3. Формы представления эмпирических распределений

Для наглядности строят различные графики статистического распределения и, в частности, полигон и гистограмму.

Полигон частот представляет собой ломаную, отрезки которой соединяют точки (x_1; n_1), (x_2; n_2), ..., (x_k; n_k). Для построения полигона частот на оси абсцисс откладывают варианты x_i, а на оси ординат – соответствующие им частоты n_i. Точки (x_i; n_i) соединяют отрезками прямых и получают полигон частот.

Полигоном относительных частот называют ломаную, отрезки которой соединяют точки (x_1; W_1), (x_2; W_2), ..., (x_k; W_k). Для построения

полигона относительных частот на оси абсцисс откладывают варианты x_i, а на оси ординат – соответствующие им относительные частоты *Wi*. Точки (x_i; W_i) соединяют отрезками прямых и получают полигон относительных частот.

Пример. Построим полигон относительных частот следующего распределения (рис. 5):

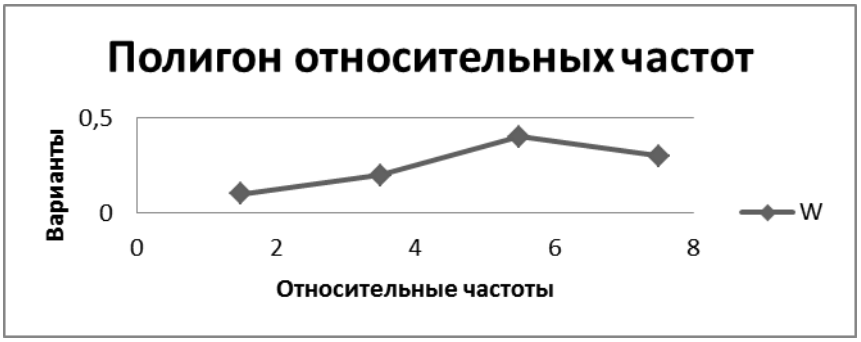

Рис. 5. Полигон относительных частот

X_i	1,5	3,5	5,5	7,5
W_i	0,1	0,2	0,4	0,3

Если наблюдается непрерывный признак, целесообразно строить гистограмму, для построения которой необходимо разбить на несколько частичных интервалов (длиной *h*) интервал всех наблюдаемых значений признака, а затем для каждого из них найти сумму частот вариант, попавших в *i*-й интервал.

Гистограммой частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиною *h*, а высоты равны отношению $\frac{n_i}{h}$ (плотность частоты).

Для построения гистограммы частот на оси абсцисс откладывают частичные интервалы, а над ними проводят отрезки, параллельные оси абсцисс на расстоянии $\frac{n_i}{h}$. Площадь *i*-го частичного прямоугольника равна $h\frac{n_i}{h} = n_i$ – сумме частот вариант *i*-го интервала; следовательно, площадь гистограммы частот равна сумме всех частот, т. е. объему выборки (рис. 6).

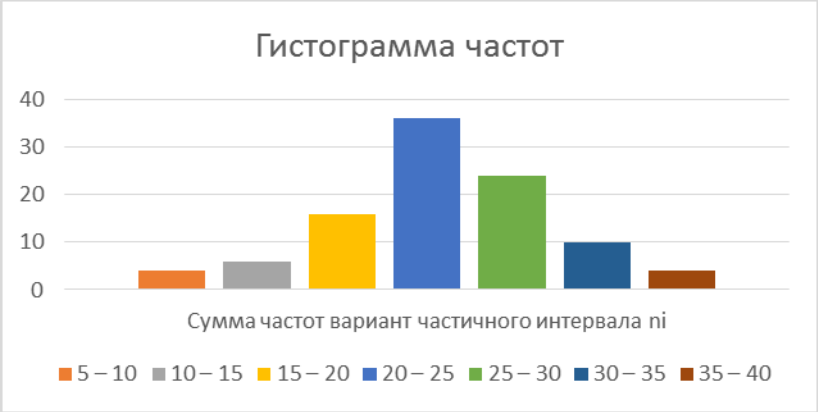

Рис. 6. Гистограмма частот распределения

На рисунке изображена гистограмма частот распределения объема $n = 100$, приведенного в следующей таблице.

Таблица 6

Гистограмма частот распределения

Частичный интервал длиной $h = 5$	Сумма частот вариант частичного интервала n_i	Плотность частоты $\frac{n_i}{h}$
5 – 10	4	0,8
10 – 15	6	1,2
15 – 20	16	3,2
20 – 25	36	7,2
25 – 30	24	4,8
30 – 35	10	2,0
35 – 40	4	0,8

Гистограммой относительных частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиной h, а высоты равны отношению $\frac{W_i}{h}$ (плотность относительной частоты). Для построения гистограммы относительных частот на оси абсцисс откладывают частичные интервалы, а над ними проводят отрезки, параллельные оси абсцисс, на расстоянии $\frac{W_i}{h}$. Площадь i-го частичного прямоугольника равна $h\frac{W_i}{h} = W_i$ – относительной частоте вариант, попавших в i-й интервал. Следовательно, площадь гистограммы относительных частот равна сумме всех относительных частот, т. е. единице.

3.4. Нормальное распределение

Распределение признака представляет собой закономерность встречаемости разных его значений.

Как правило, в психологических исследованиях ссылаются на нормальное распределение. Оно характеризуется тем, что крайние значения признака в нем встречаются довольно редко, а значения, близкие к средней величине, – достаточно часто. График нормального распределения представлен на рис. 7.

Параметрами распределения являются его числовые характеристики, определяющие, где в среднем располагаются значения признака, а также насколько эти значения изменчивы и наблюдается ли преимущественное появление определенных значений признака.

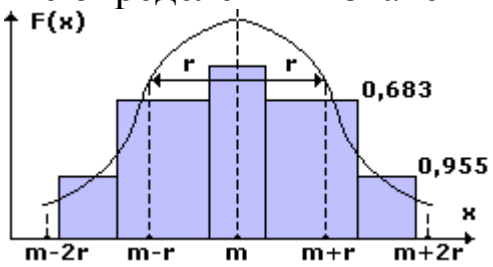

Рис. 7. График нормального распределения

Характерной чертой графика является то, что его форма и положение определены двумя параметрами: средней арифметической (μ) и стандартным отклонением (σ). В том случае, если стандартное отклонение постоянно, а величина средней меняется, форма графика не трансформируется, но он смещается вправо (при увеличении μ) или влево (при уменьшении μ) по оси абсцисс. Если же средняя остается постоянной, а стандартное отклонение меняется, наблюдается изменение ширины графика (при уменьшении σ график становится более узким и поднимается; при уменьшении σ он расширяется и опускается (рис. 8). Однако в данных случаях колоколообразный график остается строго симметричным относительно средней.

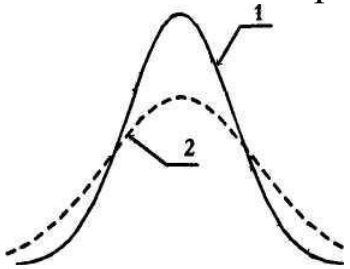

Рис. 8. Кривые распределения признака с меньшим диапазоном вариативности признака (1) и большим (2)

Если большинство значений близки к средним, 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 вырисовывается впадина, превращающая график в двувершинный. Рис. 9 иллюстрирует два вида распределения: 1) распределение с положительной асимметрией (левосторонней); 2) распределение с отрицательной асимметрией (правосторонней).

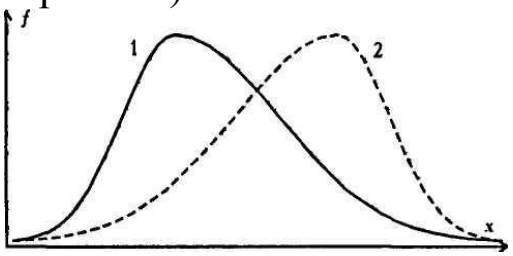

Рис. 9. Кривые распределения признака

Если существуют причины, благоприятствующие более частому появлению более высоких или низких значений (относительно среднего), появляются асимметричные распределения. При левосторонней (положительной) асимметрии в распределении чаще наблюдаются более низкие значения признака, а при правосторонней (отрицательной) – более высокие.

Эмпирические распределения исследуются на «нормальность», если необходимо воспользоваться параметрическими методами и критериями.

Алгоритм определения нормальности распределения признака

1. Рассчитать среднее арифметическое значение (M).

2. Определить стандартное отклонение (σ).

3. Вычислить показатели асимметрии и эксцесса. Показатель асимметрии (A) вычисляется по формуле:

$$A_{эмп} = \frac{\sum (x_i - M)^3}{N \cdot \sigma^3}.$$

Для симметричных распределений $A = 0$.

Показатель эксцесса (E) определяется по формуле:

$$E_{эмп} = \frac{\sum (x_i - M^4)}{N \cdot \sigma^4}.$$

В распределениях с нормальной выпуклостью $E = 0$.

1. Рассчитать критические значения асимметрии и эксцесса по формулам:

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

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

где n – количество наблюдений.

2. Сравнить эмпирические и критические значения асимметрии и эксцесса. Если $A_{эмп} < A_{кр}$; $E_{эмп} < E_{кр}$, то можно сделать вывод о том, что распределение результативного признака не отличается от нормального распределения.

Еще одной особенностью нормального распределения является совпадение величин средней арифметической, моды и медианы. Увеличение отклонения величины признака от среднего значения свидетельствует об уменьшении частоты встречаемости (вероятности) данного признака в распределении.

В психологических исследованиях нормальное распределение встречается нечасто. Однако данный факт не освобождает исследователей от необходимости проведения оценки характера распределения, так как от него зависит правильность выбора метода статистической обработки.

Пример. На психологическом факультете было проведено исследование развития вербального интеллекта на первом курсе (в начале обучения) и на третьем курсе. Данные представлены в табл. 7.

Таблица 7

Коэффициент развития вербального интеллекта на психологическом факультете

№ п/п	Коэффициент развития вербального интеллекта на первом курсе	Коэффициент развития вербального интеллекта на третьем курсе
1	34	38
2	31	39
3	27	29
4	29	36
5	31	36
6	35	37
7	29	34
8	29	39
9	31	32
10	33	35

Продолжение табл. 7

11	34	37
12	26	31
13	28	34
14	33	36
15	32	38
16	32	37
17	33	39
18	35	37
19	28	37
20	31	35
21	30	33
22	33	37
23	32	35
24	28	36
25	26	32

Психологу необходимо сделать вывод о статистической значимости полученных различий и эффективности обучения. Представленные выборки являются связанными и характеризуют коэффициент вербального интеллекта курсантов до обучения в университете и спустя три года после поступления.

Решение. Необходимо определить, имеют ли сравниваемые данные нормальный закон распределения. Для этого воспользуемся инструментами MS Excel:

– отсортируем исходные данные по возрастанию;

– найдем средние значения выборок посредством функции СРЗНАЧ;

– определим стандартное отклонение, используя функцию СТАНДОТКЛОН;

– возвратим нормальное значение данным, применив функцию НОРМАЛИЗАЦИЯ;

– возвратим стандартное нормальное интегральное распределение, используя функцию НОРМ.СТ.РАСП. В графе Z укажем значение, выданное функцией НОРМАЛИЗАЦИЯ. Так как нам необходимо построить функцию плотности вероятности, в графе Интегральная укажем ЛОЖЬ.

На основании полученных данных (столбцы НОРМАЛИЗАЦИЯ и НОРМ.СТ.РАСП.) построим график распределения для первой выборки (рис. 10) и для второй выборки (рис. 11).

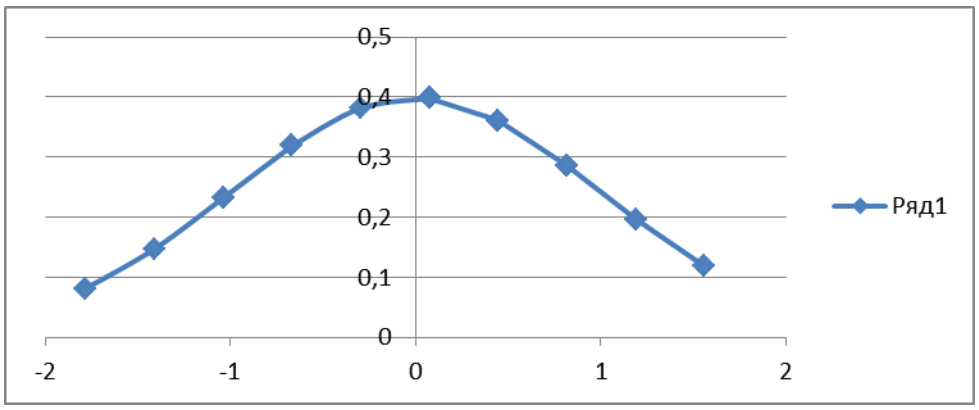

Рис. 10. График распределения для первой выборки

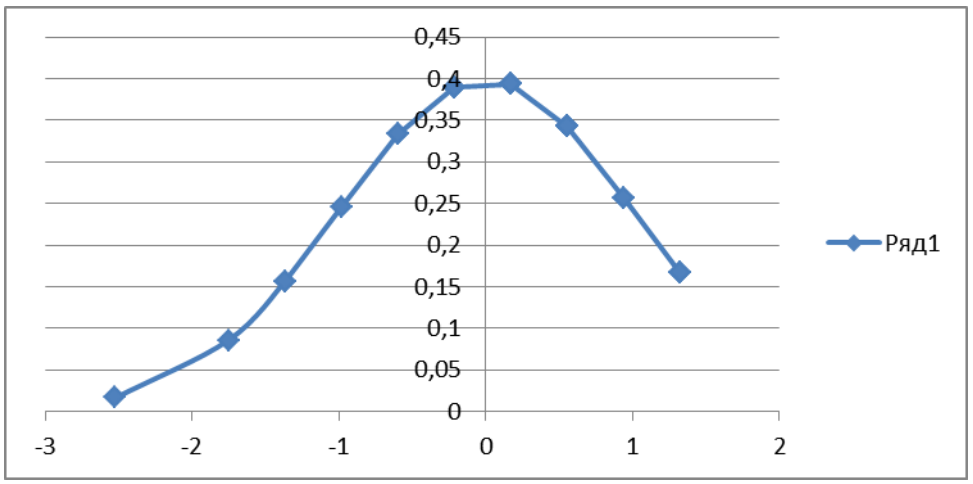

Рис. 11. График распределения для второй выборки

Представленные выше графики свидетельствуют о том, что обе выборки соответствуют нормальному закону распределения.

Вопросы для самоконтроля

1. Понятие вариационного ряда, вариант, относительных частот.
2. Понятие эмпирической функции распределения. Ее свойства.
3. Статистическое распределение выборки.
4. Виды вариационных рядов.
5. Показатели вариационного ряда. Меры центральной тенденции.
6. Особенности мер центральной тенденции.
7. Формы представления эмпирических распределений. Полигон частот, полигон относительных частот.
8. Формы представления эмпирических распределений. Гистограмма частот, гистограмма относительных частот.
9. Проверка нормальности распределения.

ТЕМА 4. Сравнения выборок. Понятие статистических критериев и их виды

4.1. Статистические критерии согласия

При проведении психологического исследования зачастую возникает необходимость сравнения результатов обследования того или иного психологического явления в разных условиях (до и после определенного воздействия, обследование контрольной и экспериментальной групп). Для решения данной задачи применяют различные статистические критерии.

Статистический критерий представляет собой определенное правило, позволяющее принять истинную и отклонить ложную гипотезы с высокой вероятностью.

Статистические критерии позволяют определить степень статистической достоверности различий между показателями, измеренными в соответствии с задачами статистического исследования.

Рассчитав эмпирическое значение того или иного статистического критерия, психолог сравнивает данный показатель с критическим значением. В том случае, если эмпирическое значение превышает критическое, можно говорить о значимости различий (за исключением U-критерия Манна-Уитни, критерия знаков).

Как правило, значимость эмпирического критерия зависит от объема выборки (n) и количества степеней свободы (df или ν).

Число степеней свободы ν равно числу классов вариационного ряда минус число условий, при которых он был сформирован. Такими условиями являются: объем выборки (n), средние значения, дисперсии.

При классификации наблюдений по группам какой-нибудь шкалы наименований и при подсчете количества наблюдений формируется частотный вариационный ряд. При его образовании должно выполняться условие соблюдения объема выборки (n). Например, при выборе профессии сформировалось три группы:

– нравится профессия;

– безразлична профессия;

– не нравится профессия.

Объем выборки – 65 испытуемых. Нам известно, что в первой группе 15 человек; во второй – 35; следовательно, в третьей – 15. В том случае, если нам неизвестно точное количество испытуемых третьей группы, но мы знаем количество испытуемых в первых двух

группах, нам не представится сложным определить число испытуемых, которым не нравится предлагаемая профессия.

Мы несвободны в подсчете числа испытуемых в третьей группе – «свобода» распространяется лишь на первые две ячейки классификации:

$$v = c - 1 = 3 - 1 = 2.$$

Таким образом, *степени свободы* – это количество значений, способных свободно варьироваться, при условии, что известна информация вроде выборочного среднего.

Аналогичным образом, если бы у нас была классификация из 10 разрядов, то мы были бы свободны только в 9 из них; если бы у нас было 100 классов – то в 99 из них и т. д.

Выделяют *два вида статистических критериев*: параметрические и непараметрические.

Для параметрических критериев характерен определенный тип распределения генеральной совокупности, а также применение параметров данной совокупности (средние, дисперсии и т. д.). Примерами параметрических критериев являются t-критерий Стьюдента, критерий согласия распределения выборок (критерий Фишера).

Непараметрические критерии основываются на оперировании частотами или рангами и не включают в формулу расчета параметры распределения. К непараметрическим критериям относятся: Q-критерий Розенбаума, Т-критерий Вилкоксона и т. д.

Более подробно рассмотрим особенности применения параметрических и непараметрических критериев.

Параметрические критерии применяют в том случае, если необходимо:

1. Оценить различия в средних значениях двух (и более) выборок (t-критерий Стьюдента).

2. Оценить различия в дисперсиях двух и более выборок (F-критерий Фишера).

3. Определить тенденции изменения признака при переходе от условия к условию, при наличии нормального распределения (однофакторный дисперсионный анализ).

4. Оценить взаимодействие двух и более факторов в их влиянии на изменения признака (двухфакторный дисперсионный анализ).

Необходимо помнить об определенных условиях применения параметрических критериев:

– значения признака должны быть измерены по интервальной шкале;

– наличие нормального распределения признака;

– необходимо соблюдение равенства дисперсий в ячейках комплекса (в дисперсионных анализах).

В случае соблюдения данных условий параметрические критерии оказываются более мощными по сравнению с непараметрическими. Однако в том случае, если эти условия не выполнимы, исследователь для решения своей задачи имеет право использовать непараметрические критерии. В таком случае непараметрические критерии оказываются более мощными, чем параметрические.

Непараметрические критерии позволяют:

1. Оценить средние тенденции. Например, определить, как часто в выборке № 1 встречаются более высокие, а в выборке № 2 – более низкие значения признака (Q-критерий Розенбаума, U-критерий манна-Уитни и т. д.).

2. Оценить различия в диапазонах вариативности признака (F-критерий Фишера).

3. Выявить тенденции изменения признака при переходе от условия к условию при любом распределении признака (L-критерий тенденций Пейджа и S-критерий тенденций Джонкира).

Необходимо отметить, что существенным недостатком непараметрических критериев является невозможность оценить взаимодействие двух или более условий, влияющих на изменение признака (эту задачу решают с помощью двухфакторного дисперсионного анализа).

Приведем методы сравнительного анализа при определении статистической значимости (табл. 8).

Таблица 8

Методы сравнительного анализа

Непараметрические методы				**Параметрические методы**			
Зависимые выборки		Независимые выборки		Зависимые выборки		Независимые выборки	
2 выборки	3 и более выборок	2 выборки	3 и более выборок	2 выборки	3 и более выборок	2 выборки	3 и более выборок
G-критерий знаков, T-критерий Вилкоксона	L-критерий тенденций Пейджа, χ^2 Фридмана	U-критерий Манна-Уитни	H-критерий Крускала-Уоллиса	t-критерий Стьюдента для зависимых выборок	Критерий Фишера (дисперсионный анализ ANOVA)	t-критерий Стьюдента для независимых выборок	Критерий Фишера (дисперсионный анализ ANOVA)

4.2. Понятие статистической гипотезы

Статистические гипотезы – предположения о свойствах и параметрах генеральной совокупности.

Полученные в экспериментах выборочные данные всегда ограничены и носят в значительной мере случайный характер. Именно поэтому для их обработки применяются методы математической статистики, позволяющие обобщить закономерности, полученные на выборке, и распространить их на всю генеральную совокупность.

Данные, полученные в результате эксперимента на какой-либо выборке, позволяют оценить генеральную совокупность. Однако в силу действия случайных вероятностных причин подобная оценка всегда будет сопровождаться погрешностью. Поэтому подобного рода оценки должны рассматриваться как предположительные, а не как окончательные утверждения.

По мнению Г. В. Суходольского, под *статистической гипотезой* обычно принимают формальное предположение о том, что сходство или различие некоторых параметрических или функциональных характеристик случайно или, наоборот, неслучайно.

Таким образом, гипотеза – это предположение о параметре генеральной совокупности.

Сущность проверки статистической гипотезы заключается в следующем:

1. Установить согласованность экспериментальных данных и выдвинутой гипотезы.

2. Определить допустимость причисления разницы между гипотезой и результатом статистического анализа экспериментальных данных к случайным причинам.

Каждая проверка гипотез предполагает наличие основной (нулевой) и альтернативной гипотез.

Принято считать, что нулевая гипотеза H_0 – это гипотеза о сходстве, а альтернативная H_1 – гипотеза о различии. Таким образом, принятие нулевой гипотезы H_0 свидетельствует об отсутствии различий, а гипотезы H_1 – о наличии различий. *Альтернативная (экспериментальная) гипотеза* – это то, что необходимо доказать в ходе эксперимента.

Пусть выборки извлечены из нормально распределенных генеральных совокупностей, причем одна выборка имеет параметры μ_x и σ_x, а другая μ_y и σ_y. Нулевая гипотеза основывается на том, что

$\mu_x = \mu_y$ и $\sigma_x = \sigma_y$, т. е. разность двух средних $\mu_x - \mu_y = 0$ и разность двух стандартных отклонений $\sigma_x - \sigma_y = 0$. Принятие альтернативной (экспериментальной) гипотезы H_1 свидетельствует о наличии различий и исходит из предположения, что $\mu_x - \mu_y \neq 0$ и $\sigma_x - \sigma_y \neq 0$.

Например, психолог изучил ценностные ориентации курсантов (России) и слушателей (дружественных стран) психологического факультета. В результате эксперимента выявлены различия в ценностных ориентациях курсантов (России) и слушателей (дружественных стран). Можно ли на основании проведенного эксперимента сделать вывод о том, что на выбор ценностных ориентаций курсантов и слушателей влияют особенности этноидентичности?

Принимаемый в таких случаях вывод представляет собой *статистическое решение*. Отметим, что такое решение вероятно.

Если при проверке гипотезы экспериментальные данные могут противоречить гипотезе H_0, эта гипотеза *отклоняется.* В том случае, если экспериментальные данные согласуются с гипотезой H_0, она *не отклоняется*. Как правило, в таких случаях говорят, что гипотеза H_0 *принимается*. Это свидетельствует о том, что статистическая проверка гипотез, основанная на экспериментальных данных, неизбежно связана с риском (вероятностью) принять ложное решение. При этом возможны ошибки двух родов (табл. 9). Ошибка первого рода произойдет, если будет принято решение отклонить гипотезу H_0, хотя в действительности она будет верной. Ошибка второго рода произойдет, если будет принято решение не отклонять гипотезу H_0, хотя в действительности она будет неверной. Вышесказанное представим в табл. 9.

Таблица 9

Ошибки при проверке статистических гипотез

Результаты проверки гипотезы H_0	Возможные состояния проверяемой гипотезы	
	Верна гипотеза H_0	Верна гипотеза H_1
Гипотеза H_0 отклоняется	Ошибка первого рода	Правильное решение
Гипотеза H_0 не отклоняется	Правильное решение	Ошибка второго рода

Всегда есть вероятность, что исследователь может ошибиться в своем статистическом решении. Такие ошибки могут быть только двух родов. В связи с тем, что исключить ошибки при принятии статистических гипотез невозможно, необходимо минимизировать возможные последствия (принятие неверной статистической гипотезы) посредством увеличения объема выборки.

Пример 1. На заводе, выпускающем автомобили, о качестве продукции судят по результатам выборочного контроля. Если выборочная доля брака не превышает 5 %, то партия принимается (H_1). Если при проверке выдвинутой гипотезы (H_1) допущена ошибка первого рода, то приемщик забракует годную продукцию. В результате ошибки второго рода потребителю отправят бракованный автомобиль.

Пример 2. Рассматривается уголовное дело о краже ценностей. Сформулируем гипотезы: H_0 – обвиняемый невиновен в совершении преступления; H_1 – подсудимый виновен. При вынесении приговора судья может допустить ошибки первого или второго рода. Ошибка первого рода означает, что суд приговорил невиновного, в то время как он не совершал преступления. Ошибкой второго рода будет вынесение оправдательного приговора, когда на самом деле подсудимый виновен.

4.3. Уровни статистической значимости

Уровень значимости – это вероятность того, что мы сочли различия существенными, а они на самом деле случайны, т. е. вероятность отклонения нулевой гипотезы (H_0) в то время как она верна.

Уровень p-значимости – минимальный уровень значимости, при котором будет отвергнута основная гипотеза, притом что она является истиной.

Когда исследователь говорит о том, что различия достоверны на 5 %-ом уровне значимости или при $p \le 0{,}05$, вероятность недостоверности этих различий составляет 0,05 (обозначают $\alpha = 0{,}05$, вероятность правильного решения $1 - \alpha$).

Если исследователь утверждает, что различия достоверны на 1 %-м уровне значимости или при $p \le 0{,}01$, значит, вероятность того, что они недостоверны, составляет 0,01 ($\alpha = 0{,}01$).

В математической статистике принято считать:

– низшим уровнем статистической значимости 5 % ($p \le 0{,}05$);

– достаточным уровнем статистической значимости 1 % ($p \le 0{,}01$);

– высшим уровнем статистической значимости 0,1 % ($p \le 0{,}001$).

В статистических таблицах критических значений, как правило, приводятся значения критериев, соответствующих уровням статистической значимости $p \le 0{,}05$, $p \le 0{,}01$, $p \le 0{,}001$.

Пока уровень статистической значимости не достигнет $p = 0{,}05$, исследователь не может отклонить нулевую гипотезу. Если эмпирическое значение критерия равняется критическому значению, соот-

ветствующему $p \le 0{,}05$ или больше него, H_0 отклоняется, но определенно принять гипотезу H_1 исследователь не может.

H_0 отклоняется и принимается H_1 однозначно в том случае, если эмпирическое значение критерия равно критическому значению, соответствующему $p \le 0{,}01$ или больше его.

Однако есть исключения, например, G-критерий знаков, T-критерий Вилкоксона, U-критерий Манна-Уитни.

Пример. Рассмотрим построение «оси значимости» для t-критерия Стьюдента. Исследователь рассчитал эмпирическое значение t-критерия Стьюдента ($t_{эмп} = 4{,}1$) и сравнил его с критическим. Затем отобразил полученные результаты на «оси значимости» (рис. 12):

$$t_{кр} = \begin{cases} 2{,}13 \ для \ p \le 0{,}05 \\ 2{,}95 \ для \ p \le 0{,}01 \\ 4{,}07 \ для \ p \le 0{,}001. \end{cases}$$

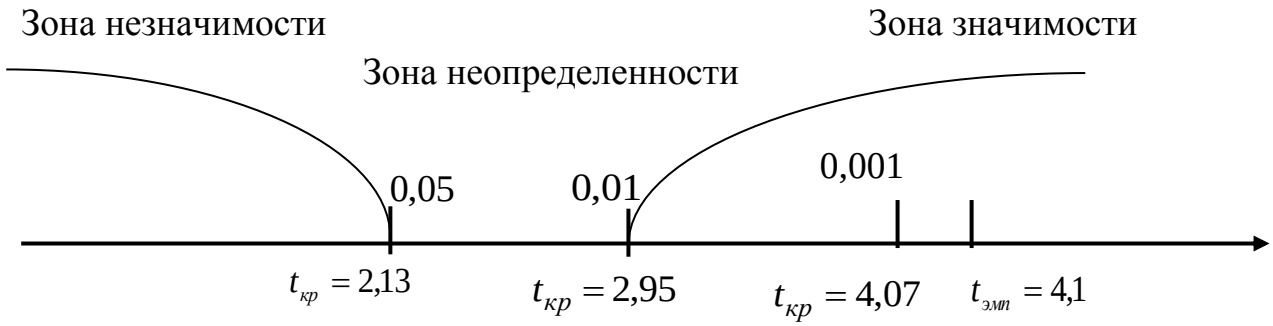

Рис. 12. Пример «оси значимости» для t-критерия Стьюдента

Эмпирическое значение критерия попадает в зону значимости, а значит можно определенно отклонить гипотезу H_0 и принять альтернативную гипотезу H_1.

В том случае, если $t_{эмп} = 2{,}41$, «ось значимости» будет выглядеть следующим образом (рис. 13):

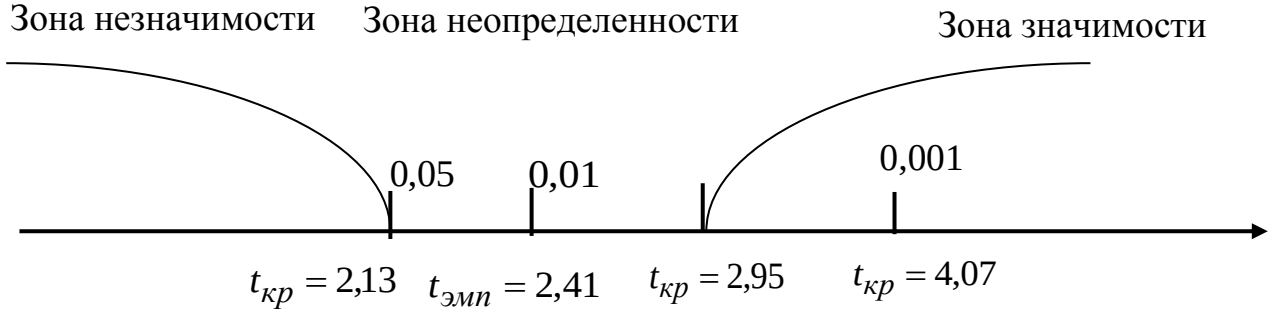

Рис. 13. Пример «оси значимости» для t-критерия Стьюдента

Как видно из рис. 13, эмпирическое значение t-критерия Стьюдента попадает в зону между 0,05 и 0,01, называемую зоной неопреде-

ленности. Это означает, что исследователь уже может отклонить гипотезу H_0, но однозначно принять гипотезу H_1 не может. Однако уже можно считать достоверными те различия, которые не попали в зону незначимости, опираясь на то, что они достоверны при $p \leq 0{,}05$ или указав точный уровень значимости полученного эмпирического значения, например, $p \leq 0{,}03$ (критерии Крускела-Уоллиса, Пейджа, Фридмана и т. д.). Необходимо отметить, что это именно тот случай, когда исследователь может допустить ошибки первого или второго. Для их устранения рекомендуется увеличить объем выборки.

Величины 0,01, 0,05 и 0,001 – это так называемые стандартные уровни статистической значимости. При статистическом анализе психолог должен, в зависимости от задач и гипотез исследования, выбрать необходимый уровень значимости.

Правила отклонения нулевой гипотезы H_0 и принятия альтернативной гипотезы H_1

1. Эмпирическое значение критерия меньше критического значения, соответствующего $p \leq 0{,}05$, значит, гипотеза H_0 принимается.

2. Эмпирическое значение критерия равняется критическому значению, соответствующему $p \leq 0{,}05$ или превышает его, следовательно, H_0 отклоняется, но однозначно принять альтернативную гипотезу H_1 исследователь не может.

3. Эмпирическое значение критерия равняется критическому значению, соответствующему $p \leq 0{,}01$, или превышает его. Следовательно, H_0 отклоняется, и принимается альтернативная гипотеза H_1.

4.4. Мощность критериев

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

Ошибка принятия нулевой гипотезы, в то время как она неверна, называется ошибкой второго рода. Вероятность совершения такой ошибки обозначается как β. Мощность критерия – это способность не допустить ошибку второго рода. Данное утверждение можно записать:

Мощность критерия $= 1 - \beta$.

Определить мощность критерия можно эмпирическим путем. Применение различных критериев обусловлено тем, что некоторые критерии позволяют выявить различия там, где другим это недо-

ступно, или позволяют выявить более высокий уровень статистической значимости.

При выборе статистического критерия необходимо руководствоваться следующими характеристиками:

– простота;

– более широкий диапазон использования (например, по отношению к данным, определенным по шкале наименований или по отношению к большим объемам выборки);

– использование критерия в отношении неравных по объему выборок;

– большая информативность результатов.

Принятие статистического решения разбивается на следующие этапы:

1. Формулировка нулевой гипотезы H_0 и альтернативной (экспериментальной) гипотезы H_1.
2. Определение объема выборки.
3. Выбор соответствующего выбора значимости (5 %, 1 %, 0,1 %).
4. Выбор статистического метода.
5. Расчет соответствующего эмпирического значения по одному из выбранных статистических методов.
6. Нахождение критических значений по таблицам.
7. Построение оси значимости и нанесение на нее критических и эмпирических значений выбранного статистического метода.
8. Формулировка принятия решения (выбор соответствующей гипотезы).

Вопросы для самоконтроля

1. Понятие статистического критерия.
2. Понятие степеней «свободы».
3. Особенности применения параметрических критериев.
4. Особенности применения непараметрических критериев.
5. Сущность проверки статистической гипотезы.
6. Понятие альтернативной гипотезы.
7. Понятие нулевой гипотезы.
8. Сущность проверки статистической гипотезы.
9. Ошибки при проверке статистических гипотез.
10. Понятие уровня статистической значимости.
11. Правила отклонения нулевой гипотезы H_0 и принятия альтернативной гипотезы H_1.
12. Мощность критериев.
13. Этапы принятия статистического решения.

ТЕМА 5. Параметрические критерии различий

Ключевым моментом обработки результатов исследования является выбор подходящего статистического критерия. Для реализации данной задачи приведем классификацию методов решения психологических задач (табл. 9).

Таблица 9

Классификация психологических задач и методов их решения

Задачи	Условия	Методы
Выявление различий в уровне исследуемого признака	2 выборки испытуемых	Q-критерий Розенбаума
		U-критерий Манна-Уитни
		t-критерий Стьюдента
	3 и более выборок испытуемых	S-критерий тенденций Джонкира
		H-критерий Крускала-Уоллиса
Оценка сдвига значений исследуемого признака	2 замера на одной и той же выборке испытуемых	T-критерий Вилкоксона
		G-критерий знаков
		F-критерий Фишера
	3 и более замеров на одной и той же выборке испытуемых	Критерий Фридмана
		L-критерий тенденций Пейджа
Выявление различий в распределении признака	При сопоставлении эмпирического распределения с теоретическим	Биноминальный критерий
		Критерий Пирсона
		Критерий Колмогорова-Смирнова
	При сопоставлении двух эмпирических признаков	F-критерий Фишера
Выявление степени согласованности изменений	Двух признаков	Коэффициент ранговой корреляции Спирмена или Кендалла
	Двух иерархий или профилей	Критерий линейной корреляции Пирсона
Анализ изменения признака под влиянием контролируемых условий	Под влиянием одного фактора	L-критерий тенденций Пэйджа
		Однофакторный дисперсионный анализ
		Критерий Барлетта
		G-критерий Кохрена
	Под влиянием нескольких факторов одновременно	Факторный дисперсионный анализ

5.1. t-критерий Стьюдента

Уильям Госсет разработал t-критерий для оценки качества пива в компании Гиннесс. В связи с обязательствами перед компанией по неразглашению коммерческой тайны, статья У. Госсета вышла в 1908 г. в журнале «Биометрика» под псевдонимом «Student» (Студент).

t-критерий Стьюдента оценивает различия средних величин $\overline{X}$, $\overline{Y}$ двух выборок *X*, *Y*, распределенных по нормальному закону. Данные в выборке должны быть измерены по интервальной шкале или шкале отношений. Данный критерий позволяет сопоставить средние значения у зависимых и независимых выборок. При сравнении двух (и более) выборок важным параметром является их зависимость. Соответственно, зависимые выборки всегда имеют одинаковый объем, а у независимых выборок он может отличаться.

5.1.1. Двухвыборочный t-критерий Стьюдента для независимых выборок

Расчет t-критерия Стьюдента для двух независимых выборок возможен двумя способами:

– когда дисперсии известны;

– когда дисперсии неизвестны, но равны друг другу.

Алгоритм расчета

1. Проверить нормальность закона распределения по одному из критериев согласия.

2. Определить средние арифметические значения $\overline{X}$ и $\overline{Y}$ для каждой выборки по формуле:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i ,$$

где x_i – значение *i*-го результата наблюдения.

3. Определить $t_{эмп}$ – эмпирическое значение t-критерия Стьюдента:

$$t = \frac{\left|\overline{X} - \overline{Y}\right|}{S_d},$$

где $S_d = \sqrt{S_x^2 + S_y^2}$ – среднеквадратическое отклонение;

$S_x{}^2$ и $S_y{}^2$ – оценки дисперсий.

Для равночисленных выборок ($n_1 = n_2 = n$) среднеквадратическое отклонение рассчитывается по формуле:

$$S_d = \sqrt{S_x^2 + S_y^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2 + (y_i - \bar{y})^2}{(n-1)n}}.$$

Для неравночисленных выборок ($n_1 \neq n_2$) среднеквадратическое отклонение определяется:

$$S_d = \sqrt{S_x^2 + S_y^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2 + (y_i - \bar{y})^2}{(n_1 + n_2 - 2)} \frac{n_1 + n_2}{n_1 n_2}}.$$

Расчет числа степеней свободы для неравночисленных выборок:

$$\upsilon = df = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2.$$

Расчет степеней свободы для равночисленных выборок: $\upsilon = 2n - 2$.

4. Эмпирическое значение $t_{эмп}$ критерия Стьюдента сравнивается с критическим значением $t_{кр}$ (по таблице критических значений) для рассчитанного числа степеней свободы.

Нулевая гипотеза H_0 при заданном уровне значимости α принимается, если эмпирическое значение $t_{эмп.} < t_{кр}$.

Пример. Психолог определил индивидуальный уровень агрессивности личности (по опроснику определения индивидуального уровня агрессивности личности А. Басса и А. Дарки) у курсантов психологического и экономического факультетов. Исследователь выдвинул гипотезу о том, что курсанты психологического факультета менее агрессивны, чем курсанты экономического факультета.

Таблица 10

Расчет параметров для определения двухвыборочного t-критерия Стьюдента (независимые выборки)

№ п/п	Уровень агрессивности курсантов психологического факультета	Уровень агрессивности курсантов экономического факультета	Отклонения от среднего значения		Квадраты отклонений	
			$(x_i - \bar{x})$	$(y_i - y)$	$(x_i - \bar{x})^2$	$(y_i - y)^2$
1	36	42	−2,00	−1,78	4,00	3,16
2	37	42	−1,00	−1,78	1,00	3,16
3	37	43	−1,00	−0,78	1,00	0,60
4	38	43	0,00	−0,78	0,00	0,60
5	38	43	0,00	−0,78	0,00	0,60

Продолжение табл. 10

6	38	44	0,00	0,222	0,00	0,05
7	39	45	1,00	1,222	1,00	1,49
8	39	46	1,00	2,222	1,00	4,94
9	39	46	1,00	2,222	1,00	4,94
10	39	47	1,00		1,00	
Сумма	**380**	**441**	0,00	0,00	**10,00**	**28,9**
Средние значения	**38**	**44,1**				

Рассчитаем эмпирическое значение двухвыборочного t-критерия Стьюдента по формуле:

$$t = \frac{\left|\overline{X} - \overline{Y}\right|}{S_d};$$

$$\left|\overline{X} - \overline{Y}\right| = \left|38 - 44{,}1\right| = 6{,}1;$$

$$S_d = \sqrt{S_x^2 + S_y^2} = \sqrt{\frac{\sum(x_i - \overline{x})^2}{(n-1)} + \frac{\sum(y_i - \overline{y})^2}{(n-1)}} = \sqrt{\frac{10}{10-1} + \frac{28{,}9}{10-1}} = \sqrt{4{,}32} = 2{,}08;$$

$$t_{эмп} = \frac{|6{,}1|}{2{,}08} = 2{,}93.$$

Рассчитаем степени свободы:

$$\upsilon = df = (n_1 - 1) + (n_2 - 1) = (10-1) + (10-1) = 18.$$

По таблице критических значений определяем $t_{кр}$ для 17 степеней свободы:

$$t_{кр} = \begin{cases} 2{,}10 \ \textit{для}\ p \le 0{,}05 \\ 2{,}88\ \textit{для}\ p \le 0{,}01 \\ 3{,}92\ \textit{для}\ p \le 0{,}001. \end{cases}$$

Критические значения t-критерия Стьюдента можно определить с помощью MS Excel. Для этого воспользуемся функцией СТЬЮДРАПСОБР (рис. 14).

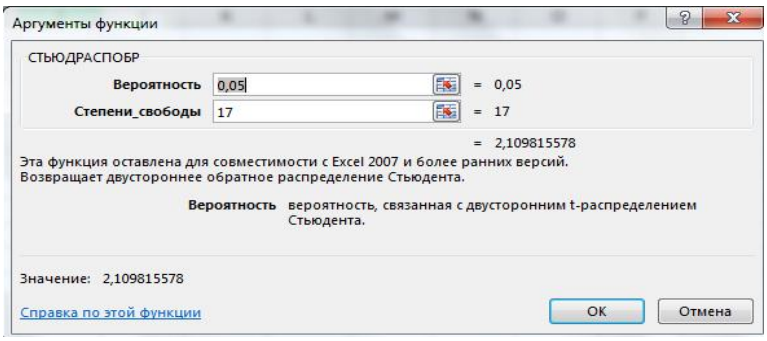

Рис. 14. Применение функции СТЬЮДРАПСОБР

В графе «вероятность» укажем интересующий нас уровень значимости (0,05, 0,01 или 0,01), а в графе «степени свободы» укажем число степеней свободы.

Строим ось значимости (рис. 15):

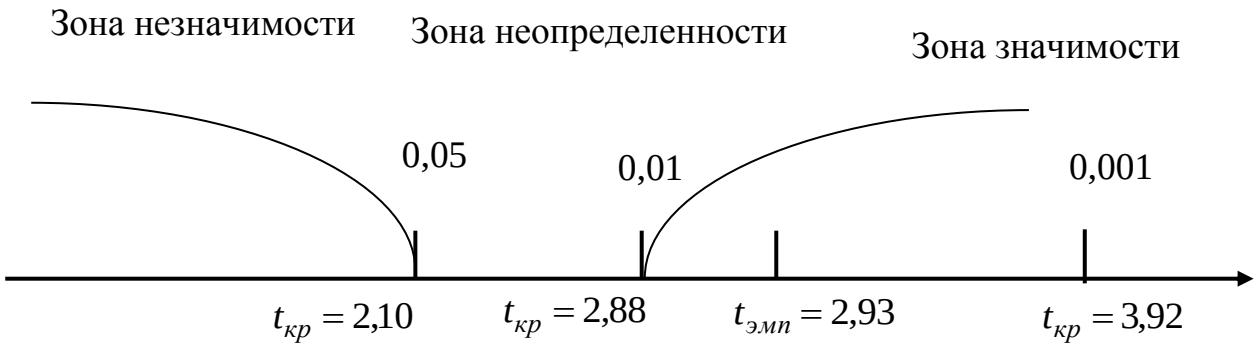

Рис. 15. Пример «оси значимости» для t-критерия Стьюдента

Таким образом, на 1 %-м уровне значимости подтвердилась гипотеза H_1 (курсанты психологического факультета менее агрессивны, чем курсанты экономического факультета). Гипотеза H_0 (о сходстве) отклоняется.

5.1.2. Двухвыборочный t-критерий Стьюдента для зависимых выборок

Данный критерий можно использовать только для двух зависимых выборок, при условии нормального распределения данных этих выборок.

Гипотезы:

H_0 : среднее значение в выборке не отличается от нуля.

H_1 : среднее значение в выборке отличается от нуля.

Алгоритм расчета

1. Проверить нормальность закона распределения по одному из критериев согласия.

2. Рассчитать $\sigma_i = x_i - y_i$ $(i = 1\ldots n)$ – попарные разности вариант x_i и y_i результаты измерений для i-го объекта до и после воздействия некоторого фактора. Величину σ_i будем считать независимой для разных объектов и нормально распределенной.

3. Определить сумму попарных разностей $\sum_{i=1}^{n}\sigma_i$ и вспомогательных параметров $\sum_{i=1}^{n}\sigma_i^{\,2}$ и $\left(\sum_{i=1}^{n}\sigma_i\right)^2$.

4. Рассчитать $t_{эмп}$ – эмпирическое значение критерия $\upsilon = (n-1)$ степенями свободы по формуле:

$$t = \frac{\sum_{i=1}^{n}\sigma_i}{\sqrt{\frac{n\sum_{i=1}^{n}\sigma_i^{\,2} - \left(\sum_{i=1}^{n}\sigma_i\right)^2}{n-1}}},$$

где n – объем выборки.

5. Рассчитанное эмпирическое значение $t_{эмп}$ t-критерия Стьюдента сравнивается с критическим значением $t_{кр}$ по таблице критических значений[1] для данного числа степеней свободы.

Нулевая гипотеза H_0 при заданном уровне значимости α принимается, если эмпирическое значение $t_{эмп} < t_{кр}$.

Пример. Исследователь-психолог разрабатывает методику снижения уровня агрессивности. Для этого он определил индивидуальный уровень агрессивности личности (по опроснику определения индивидуального уровня агрессивности личности А. Басса и А. Дарки) у курсантов психологического факультета до и после применения данной методики. Результаты исследования представлены в табл. 11.

[1] См. Приложение 1.

Таблица 11

Расчет параметров для определения t-критерия Стьюдента (зависимые выборки)

№ п/п	Уровень агрессивности курсантов психологического факультета до воздействия (x)	Уровень агрессивности курсантов психологического факультета после воздействия (y)	Попарные разности вариант $\sigma_i = x_i - y_i$	Квадрат попарных разностей вариант $\sigma_i^2 = (x_i - y_i)^2$
1	39	35	4	16
2	40	36	4	16
3	41	37	4	16
4	42	37	5	25
5	42	37	5	25
6	43	38	5	25
7	43	38	5	25
8	44	39	5	25
9	44	40	4	16
10	45	40	5	25
Сумма $\sum$	**423**	**377**	**46**	**214**

$$t_{эмп} = \frac{\sum_{i=1}^{n} \sigma_i}{\sqrt{\frac{n\sum_{i=1}^{n} \sigma_i{}^2 - (\sum_{i=1}^{n} \sigma_i)^2}{n-1}}} = \frac{46}{\sqrt{\frac{10 \cdot 214 - (46)^2}{10-1}}} = \frac{46}{\sqrt{2{,}661}} = \frac{46}{1{,}63} = 28{,}22.$$

Число степеней свободы: $\upsilon = n - 1 = 10 - 1 = 9$.

Определяем критические значения для t-критерия Стьюдента для 18 степеней свободы:

$$t_{кр} = \begin{cases} 2{,}26 \ \textit{для}\ p \le 0{,}05 \\ 3{,}25 \ \textit{для}\ p \le 0{,}01 \\ 4{,}78 \ \textit{для}\ p \le 0{,}001. \end{cases}$$

Строим ось значимости (рис. 16):

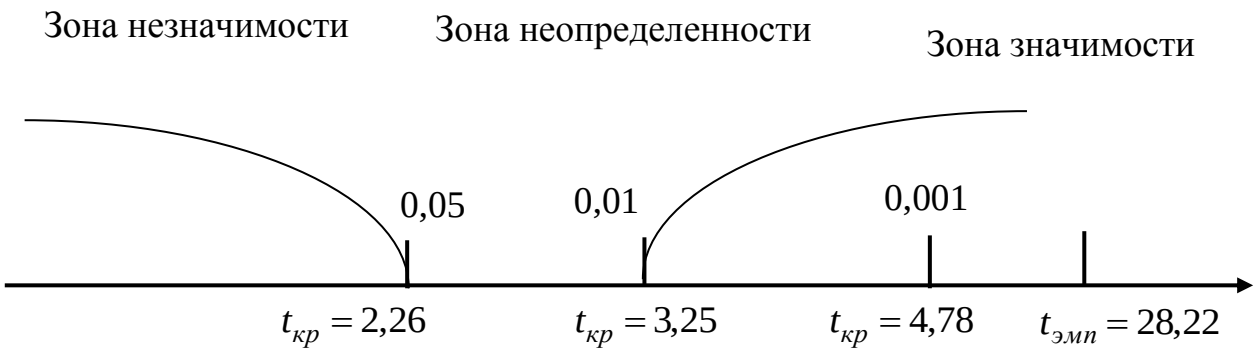

Рис. 16. Пример «оси значимости» для t-критерия Стьюдента

Таким образом, на 0,1 %-м уровне значимости подтвердилась экспериментальная гипотеза о том, что методика снижения уровня агрессивности дает положительные результаты: снижается уровень агрессивности.

Так как наши выборки зависимые, воспользуемся инструментом анализа данных MS Excel – парный двухвыборочный t-тест для средних. В результате получим следующие данные (табл. 12).

Таблица 12

Результаты анализа парного двухвыборочного t-теста для средних

	Выборка 1	**Выборка 2**
Среднее	42,3	37,7
Дисперсия	3,566666667	2,677777778
Наблюдения	10	10
Корреляция Пирсона	0,967144224	
Гипотетическая разность средних	0	
df	9	
t-статистика	28,16913204	
P (T <= t) одностороннее	2,17691E − 10	
t-критическое одностороннее	1,833112933	
P (T <= t) двустороннее	4,35382E − 10	
t-критическое двустороннее	2,262157163	

Как показывает рис. 16, эмпирическое значение t-критерия равно 28,17 и выше критического значения (2,26). Таким образом, на 0,1 %-м уровне значимости первоначальное предположение подтвердилось: коэффициент вербального интеллекта у курсантов психологического факультета существенно увеличился.

5.2. F-критерий Фишера

Критерий Фишера позволяет сравнить дисперсии двух выборок. Данные этих выборок должны быть измерены по шкале интервалов или по шкале отношений, а также распределены по нормальному закону.

Для определения $F_{эмп}$ необходимо найти отношение дисперсий двух выборок, причем большая по величине дисперсия записывается в числителе, а меньшая – в знаменателе. Поэтому значение F-критерия Фишера больше или равно 1.

Гипотезы:

H_0: дисперсия выборки 1 не отличается от дисперсии в выборке 2.

H_1: дисперсия выборки 1 отличается от дисперсии в выборке 2.

Алгоритм расчета

1. Проверить данные на нормальность распределения.

2. Рассчитать средние арифметические значения $\bar{x}_1$ и $\bar{x}_2$ для каждой выборки по формуле:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i ,$$

где x_i – значение *i*-го результата наблюдения.

3. Рассчитать значение S_{x1}^2 и S_{x2}^2 – дисперсии для каждой выборки по формуле:

$$S_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 .$$

4. Определить число степеней свободы по выборкам:

$df_1 = (n_1 - 1)$ – по первой выборке и $df_2 = (n_2 - 1)$ – по второй выборке.

5. Рассчитать $F_{эмп}$ – эмпирическое значение критерия по одной из формул:

$F_{эмп} = \frac{S_{x1}^2}{S_{x2}^2}$ или $F_{эмп} = \frac{S_{x2}^2}{S_{x1}^2}$ с учетом того, что дисперсия в числителе должна быть больше дисперсии в знаменателе.

6. Найденное эмпирическое значение критерия Фишера $F_{эмп}$ сравнить с критическим значением $F_{кр}$ по таблице критических значений[1] для данного числа степеней свободы υ.

[1] См. Приложение 2.

Если $F_{эмп} < F_{кр}$, то нулевая гипотеза H_0 о равенстве дисперсий в выборках при заданном уровне значимости α принимается.

Пример. В двух группах измерялся уровень стрессоустойчивости по методике Т. Холмса и Р. Раге. Полученные значения средних достоверно не различались. Перед исследователем встала задача определить, есть ли различия в степени однородности показателей стрессоустойчивости между группами.

Сравним дисперсии уровня стрессоустойчивости в обеих группах (табл. 13).

Таблица 13

Уровень стрессоустойчивости в исследуемых группах

№ п/п	Уровень стрессоустойчивости в 1-й группе	Уровень стрессоустойчивости во 2-й группе
1	150	144
2	150	150
3	160	152
4	160	152
5	169	156
6	179	160
7	179	168
8	180	170
9	180	180
10	185	190
11	200	200
12	200	210
13	210	220
14	210	260
Сред. знач.	179,43	179,43
Дисперсия	411,03	1106,11

Представленная таблица свидетельствует о том, что средние значения в исследуемых группах совпадают (179,43). Соответственно, значение t-критерия Стьюдента незначимо. Рассчитаем оценки дисперсий для исследуемых групп: $S_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 411{,}03$ и $S_y^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 1106{,}11$.

Определим значение F-критерия Фишера:

$$F_{эмп} = \frac{S_y^2}{S_x^2} = \frac{1106{,}11}{411{,}03} = 2{,}7\,.$$

Рассчитаем степени свободы:

$df_1 = (n_1 - 1) = 14 - 1 = 13$ для 1-й выборки;

$df_1 = (n_1 - 1) = 14 - 1 = 13$ для 2-й выборки.

$$F_{кр} = \begin{cases} 2{,}55 \ для\ p \le 0{,}05 \\ 3{,}85 \ для\ p \le 0{,}01 \end{cases}$$

Построим ось значимости (рис. 17):

Рис. 17. Пример «оси значимости» для F-критерия Фишера

Таким образом, эмпирическое значение F-критерия Фишера попало в зону неопределенности. Это означает, что гипотеза H_0 может быть отвергнута, а гипотеза H_1 может быть принята на 5 %-м уровне значимости. Психолог может сделать вывод о том, что между двумя выборками, в которых исследовался уровень стрессоустойчивости, имеется различие.

Вопросы для самоконтроля

1. Классификация психологических задач и методов их решения.
2. Условия применения t-критерия Стьюдента для независимых выборок.
3. Условия применения t-критерия Стьюдента для зависимых выборок.
4. Алгоритм расчета t-критерия Стьюдента для независимых выборок.
5. Алгоритм расчета t-критерия Стьюдента для зависимых выборок.
6. Условия применения F-критерия Фишера.
7. Алгоритм расчета F-критерия Фишера.

ТЕМА 6. Статистические критерии различий

6.1. Непараметрические критерии для связанных выборок

Иногда, визуально сравнивая полученные результаты исследования, например, до и после воздействия, психолог замечает увеличение или уменьшение данных и для оценки различий подсчитывает проценты изменений, а также сравнивает полученные проценты между собой. Затем на основе подобного сравнения делает вывод о том, что если имеются различия в процентах, то различия есть и в сравниваемых показателях до и после воздействия. Однако данное утверждение категорически неверно, так как для данных в процентах нельзя определить уровень достоверности в их различиях. На основе процентов невозможно сделать статистически достоверные выводы. Доказать эффективность какого-либо воздействия или методики возможно только с помощью определения статистически значимой тенденции в смещении (сдвиге) данных. Для решения статистических задач такого рода применяют критерии различий. Рассмотрим подробнее некоторые их них.

Для помощи в принятии решения о выборе критерия для сопоставлений приведем следующий алгоритм (рис. 18).

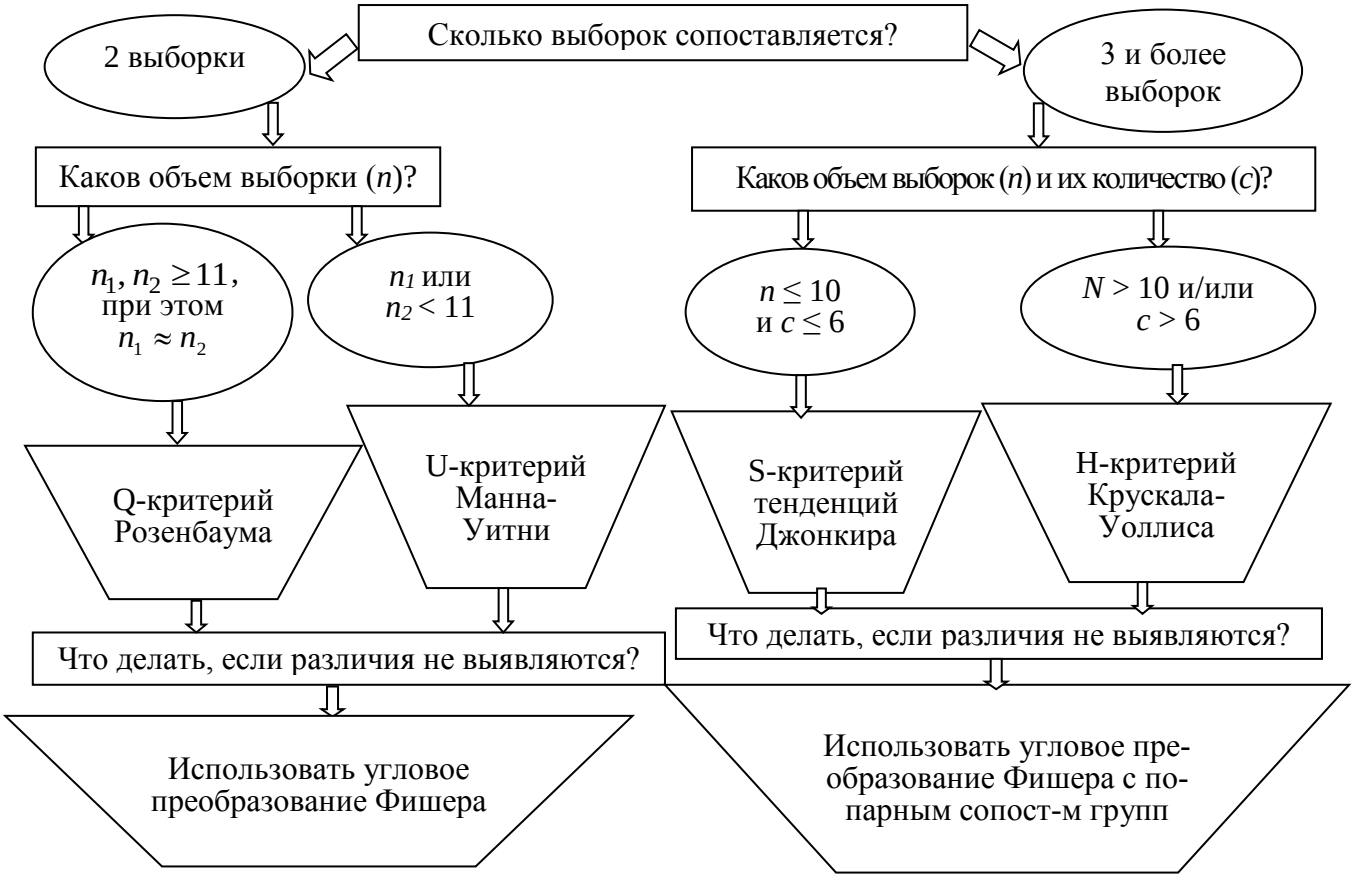

Рис. 18. Алгоритм принятия решения о выборе критерия для сопоставлений

6.1.1. G-критерий знаков

Данный критерий применим для зависимых однородных выборок, а также для данных, полученных в порядковой, интервальной и шкале отношений. Объем выборок должен совпадать. При большом числе сравниваемых парных показателей критерий знаков достаточно эффективен. Он позволяет определить, насколько однонаправленно изменяются значения признака при повторном измерении зависимой, однородной выборки.

Алгоритм расчета для зависимых выборок

1. Определить сдвиг, т. е. изменение показателей выборки до и после воздействия.

2. Подсчитать количество нулевых, отрицательных и положительных сдвигов.

3. Нулевые сдвиги отбросить. Наибольшая сумма сдвигов называется *типичным сдвигом* и обозначается *n*. Наименьшая сумма сдвигов носит название *нетипичный сдвиг* и обозначается как $G_{эмп}$.

4. По таблице критических значений[1] определим $G_{кр}$ для соответствующих уровней значимости.

5. Построить ось значимости, указав на ней $G_{кр}$ и $G_{эмп}$.

Необходимо отметить следующее:

– если типичный и нетипичный сдвиги равны, критерий знаков неприменим;

– в критерии знаков ось значимости располагается справа налево. То есть чем больше количество нетипичных сдвигов, тем меньше вероятность того, что суммарный сдвиг окажется статистически достоверным;

– G-критерий знаков является простым для вычисления, однако он наименее мощный из всех критериев (если данный критерий показал значимые различия на 1 %-м уровне значимости, то более мощные критерии подтвердят эти различия).

Пример. Психолог разработал методику снижения уровня конфликтности. Ему необходимо определить статистическую достоверность данной методики.

В группе, состоящей из 25 чел., был измерен уровень конфликтности до и после применения методики. Полученные результаты, а также их изменения приведены в табл. 14.

---

[1] См. Приложение 3.

Таблица 14

Уровень конфликтности до и после психологического воздействия

№ п/п	Уровень конфликтности до воздействия	Уровень конфликтности после воздействия	Сдвиг
1	32	31	1
2	28	27	1
3	30	31	−1
4	32	29	3
5	18	17	1
6	29	27	2
7	35	34	1
8	26	25	1
9	34	33	1
10	19	18	1
11	20	19	1
12	22	21	1
13	34	33	1
14	21	19	2
15	24	22	2
16	24	21	3
17	26	25	1
18	28	27	1
19	30	28	2
20	31	32	−1

Подсчитаем количество сдвигов:

Нулевые – 0.

Отрицательные – 2.

Положительные – 18.

Таким образом, типичным является положительный сдвиг, а нетипичным – отрицательный (n = 18, $G_{эмп}$ = 2). Оценим уровень достоверности различий. Для этого в таблице критических значений найдем $G_{кр}$ для n = 18.

$$G_{кр} = \begin{cases} 5 \ для \ p = 0{,}05 \\ 3 \ для \ p = 0{,}01. \end{cases}$$

Формула показывает, что количество нетипичных сдвигов при 5 %-м уровне достоверности не должно превышать 5, а для 1 %-го уровня – 3.

Построим ось значимости (рис. 19):

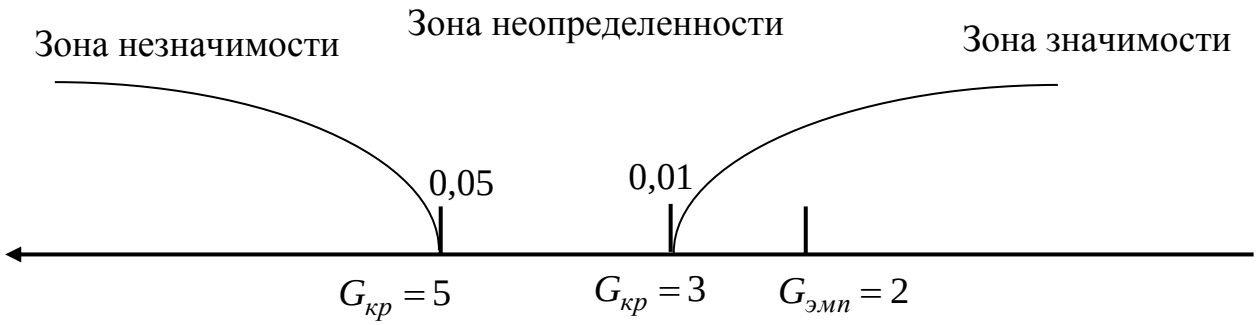

Рис. 19. Пример «оси значимости» для G-критерия знаков

Ось значимости показывает нам, что $G_{эмп}$ попало в зону значимости. Следовательно, исследователь может сделать вывод о том, что полученный в результате эксперимента сдвиг уровня конфликтности значим на 1 %-м уровне. То есть в результате применения методики конфликтность испытуемых снизилась статистически достоверно (гипотеза H_1 о различиях принята, а гипотеза H_0 о наличии сходства отклонена на 1 %-м уровне).

6.1.2. Парный Т-критерий Вилкоксона

Данный критерий применяется для сравнения двух связанных выборок по уровню какого-либо количественного признака, измеренного в порядковой, интервальной или шкале отношений.

Парный Т-критерий Вилкоксона основан на сопоставлении абсолютных величин выраженности сдвигов в том или ином направлении. Сначала все абсолютные величины сдвигов ранжируются, а затем суммируются ранги. Если сдвиги в ту или иную сторону происходят случайно, то и суммы их рангов окажутся примерно равны. Если же интенсивность сдвигов в одну сторону больше, то сумма рангов абсолютных значений сдвигов в противоположную сторону будет значительно ниже, чем это могло бы быть при случайных изменениях.

Парный Т-критерий Вилкоксона применяется для оценки различий двух выборок, в разных условиях или в разное время. Данный тест способен выявить направленность и выраженность изменений: т. е. являются ли показатели больше сдвинутыми в одном направлении, чем в другом.

Классическим примером ситуации, в которой может применяться Т-критерий Вилкоксона для зависимых выборок, является исследование до и после определенного воздействия.

Условия применения

1. Показатели выборок должны быть измерены по порядковой, интервальной или шкале отношений.
2. Объем выборки должен быть не менее 5 и не более 50.
3. Применяется только для связанных (зависимых) выборок.
4. Объем выборок должен быть равным.

Алгоритм расчета

1. Определить величины сдвигов с учетов знаков.
2. Определить абсолютную величину сдвигов.
3. Проранжировать абсолютные величины сдвигов.
4. Рассчитать сумму рангов.
5. Сравнить эмпирическую сумму рангов с теоретической, рассчитанной по формуле:

$$R = \frac{N \cdot (N+1)}{2}.$$

1. Отметить нетипичные сдвиги.
2. Определить $T_{эмп}$ (сумма рангов нетипичных сдвигов).
3. Определить $T_{кр}$ по таблице критических значений[1] для общего числа испытуемых.
4. Сравнить $T_{эмп}$ и $T_{кр}$ для избранного уровня статистической значимости (p = 0,05 или p = 0,01).

Если $T_{эмп} \leq T_{кр}$, то принимается гипотеза H_1 (альтернативная гипотеза). Чем меньше значение $T_{эмп}$, тем выше достоверность различий.

Если $T_{эмп} > T_{кр}$, то принимается гипотеза H_0 (статистическая значимость изменений показателя отсутствует).

Пример. Психолог проводит исследование депрессии у обучающихся вузов посредством теста-опросника А. Т. Бека. После первого замера уровня депрессии он провел психологический тренинг. Затем сделал второй замер. Ему необходимо определить статистическую значимость полученных результатов. Данные обоих замеров приведены в табл. 15.

[1] См. Приложение 4.

Таблица 15

Результаты исследования депрессии у студентов (2 замера)

№ п/п	1-й замер депрес-сии	2-й замер депрес-сии	Сдвиг	Абсо-лютная величина сдвига	Ранги	Нетипич-ные сдви-ги
1	20	20	0	0	2	
2	21	19	−2	2	14,5	
3	18	16	−2	2	14,5	
4	17	16	−1	1	8	
5	20	18	−2	2	14,5	
6	25	21	−4	4	18	
7	27	25	−2	2	14,5	
8	21	20	−1	1	8	
9	17	17	0	0	2	
10	18	17	−1	1	8	
11	23	22	−1	1	8	
12	16	15	−1	1	8	
13	26	23	−3	3	17	
14	28	28	0	0	2	
15	20	21	1	1	8	8
16	22	21	−1	1	8	
17	23	24	1	1	8	8
18	24	23	−1	1	8	
Сумма					**171**	**16**

Искомую таблицу мы заполнили в MS Excel:

1. Для столбца Сдвиг применили формулу: значение 2-го замера − значение 1-го замера. Для заполнения всего столбца использовали функцию Автозаполнение.

2. Абсолютную величину ранга рассчитали с помощью функции ABS (возвращает модуль числа). Для заполнения всего столбца использовали функцию Автозаполнение.

3. Ранги определили посредством функции РАНГ.СР. Числом является абсолютная величина сдвига; в графе Ссылка указали массив абсолютных величин сдвига, зафиксировали его $; в графе Порядок указали ненулевое значение, если необходима сортировка по возрастанию, и 0, если сортируем по убыванию. Для заполнения всего столбца использовали функцию Автозаполнение.

4. Нетипичные сдвиги рассчитали с помощью функции ЕСЛИ (в логическом выражении указали, что если сдвиг ≤ 0, то истинным значением будет являться «…» (пустая строка), иначе отобразится

значение ранга. Для заполнения всего столбца использовали функцию Автозаполнение.

5. Суммы рангов и нетипичных сдвигов подсчитали посредством функции СУММ.

Как видно из табл. 15, $T_{эмп}$ = 16 (сумма нетипичных сдвигов). По таблице критических значений определим $T_{кр}$ для n = 18.

	p	
n	0,05	0,01
18	47	32

Так как в нашем примере типичным является отрицательный, а нетипичным положительный сдвиги, то на 5 %-м уровне значимости сумма рангов нетипичных сдвигов не должна превышать 47, а на 1 %-м уровне значимости – 32:

$$T_{\text{кр}} = \begin{cases} 47 \ для \ p \le 0{,}05 \\ 32 \ для \ p \le 0{,}01. \end{cases}$$

Построим ось значимости (рис. 20):

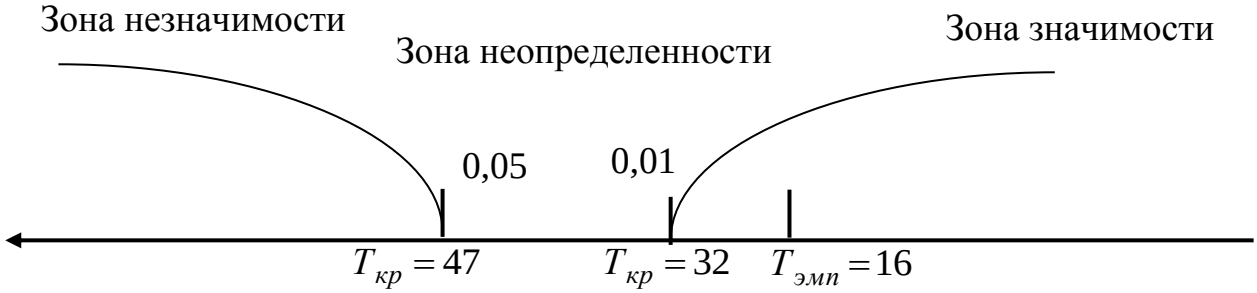

Рис. 20. Пример «оси значимости» для Т-критерия Вилкоксона

Ось значимости показывает, что $T_{эмп}$ попало в зону значимости, и, следовательно, изменения в показателях депрессии неслучайны и значимы на 1 %-м уровне. Таким образом, гипотеза H_1 о наличии различий принимается, а гипотеза H_0 о сходстве отклоняется.

6.1.3. Критерий Фридмана

Данный критерий применим для трех и более связанных выборок испытуемых, однако не позволяет выявить направление изменений. Особенностью критерия Фридмана является то, что в зависимости от числа измерений применяются разные таблицы критических значений.

Алгоритм расчета

1. Проранжировать экспериментальные данные по строкам.
2. Определить сумму полученных рангов по столбцам.
3. Определить сумму теоретических рангов по формуле:

$$R = \frac{n \cdot c \cdot (c+1)}{2},$$

где c – число столбцов;

n – число строк или испытуемых.

4. Сравнить суммы рангов (теоретическую и эмпирическую).

5. Рассчитать эмпирическое значение критерия Фридмана по формуле:

$$\chi^2_{фр_{эмп}} = \left[\frac{12}{n \cdot c \cdot (c+1)} \cdot \sum_{i=1}^{c} (R_i^2 \right] - 3 \cdot n \cdot (c+1),$$

где: n – количество испытуемых;

c – количество столбцов;

R_i – сумма рангов i-го столбца.

6. По таблице критических значений определить $\chi^2_{фр_{кр}}$ для определенного n.

Как упоминалось выше, при определении $\chi^2_{фр_{кр}}$ используются разные таблицы критических значений (в зависимости от числа испытуемых).

Правила применения таблиц критических значений для определения $\chi^2_{фр_{кр}}$

1. Критические значения $\chi^2_{фр_{кр}}$ для количества условий c = 3 и числе испытуемых от 2 до 9 $(2 \le n \le 9)$ $\chi^2_{фр_{кр}}$ определяются по соответствующей таблице.

2. Критические значения $\chi^2_{фр_{кр}}$ для количества условий c = 4 и числе испытуемых от 2 до 4 $(2 \le n \le 4)$ $\chi^2_{фр_{кр}}$ определяются по таблице[1].

3. При большом количестве измерений и испытуемых критические значения $\chi^2_{фр_{кр}}$ определяются по таблице критических значений для критерия χ^2. Число степеней свободы определяется по формуле:

$$\nu = c - 1,$$

где c – количество условий измерений.

Пример. Психолог проводит исследование депрессии у обучающихся посредством теста-опросника А. Т. Бека. После первого замера

[1] См. Приложение 5.

уровня депрессии исследователь провел психологический тренинг. Затем сделал второй замер, применил авторскую методику снижения уровня депрессии и в третий раз провел замер. Ему необходимо определить статистическую значимость полученных результатов. Данные всех замеров у обучающихся приведены в табл. 16.

Таблица 16

Результаты исследования депрессии у студентов (3 замера)

№ п/п	1-й замер депрессии	2-й замер депрессии	3-й замер депрессии
1	20	20	18
2	21	19	18
3	18	16	16
4	17	16	14
5	20	18	17
6	25	21	19

Решим задачу в MS Excel:

1. Добавим в исходную таблицу три столбца: Ранги 1-го замера, Ранги 2-го замера, Ранги 3-го замера.

2. Заполним столбец Ранги 1-го замера. Для этого проранжируем экспериментальные данные по строкам, применив функцию РАНГ.СР. В графе Число укажем показатель депрессии у 1-го испытуемого в 1-м замере; в графе Ссылка укажем диапазон значений 1-го испытуемого во всех замерах; в графе Порядок поставим 1 для сортировки по возрастанию. Используя функцию Автозаполнение, заполним столбцы Ранги 2-го замера и Ранги 3-го замера для 1-го испытуемого.

3. Аналогично рассчитаем ранги для каждого испытуемого по каждому замеру.

4. Определим сумму полученных рангов по столбцам (замерам), используя функцию СУММ.

5. Определим общую сумму рангов. Полученные данные представлены в табл. 17.

Таблица 17

Результаты исследований по 3 замерам с подсчетов рангов

№ п/п	1-й замер депрессии	2-й замер депрессии	3-й замер депрессии	Ранги 1-го замера	Ранги 2-го замера	Ранги 3-го замера
1	20	20	18	2,5	2,5	1
2	21	19	18	3	2	1
3	18	16	16	3	1,5	1,5
4	17	16	14	3	2	1

Продолжение табл. 17

5	20	18	17	3	2	1
6	25	21	19	3	2	1
Сумма				17,5	12	
Общая сумма						**36**

1. Сравним полученную сумму рангов (36) с теоретической, рассчитанной по формуле: $R = \frac{n \cdot c \cdot (c+1)}{2} = \frac{6 \cdot 3 \cdot (3+1)}{2} = 36$. Так как суммы рангов совпали, делаем вывод о правильности ранжирования.

2. Подставим необходимые значения из табл. 17 в формулу расчета $\chi^2_{фр\ эмп}$:

$$\chi^2_{фр_{эмп}} = \left[\frac{12}{n \cdot c \cdot (c+1)} \cdot \sum_{i=1}^{c} (R_i^2)\right] - 3 \cdot n \cdot (c+1) = \left[\frac{12}{6 \cdot 3 \cdot (3+1)} \cdot (17{,}5^2 + 12^2 + 6{,}5^2)\right] - 3 \cdot 6 \cdot (3+1) = 10{,}08.$$

3. По таблице критических значений определим $\chi^2_{фр_{кр}}$ для количества испытуемых равному 6. Выбираем p наиболее близкие к 0,05 и 0,01. Для приведенного примера эти значения равны 0,052 и 0,012.

$$\chi^2_{фр_{кр}} = \begin{cases} 6{,}33 \text{ } для\ p \le 0{,}05 \\ 8{,}33 \text{ } для\ p \le 0{,}01. \end{cases}$$

Построим ось значимости (рис. 21):

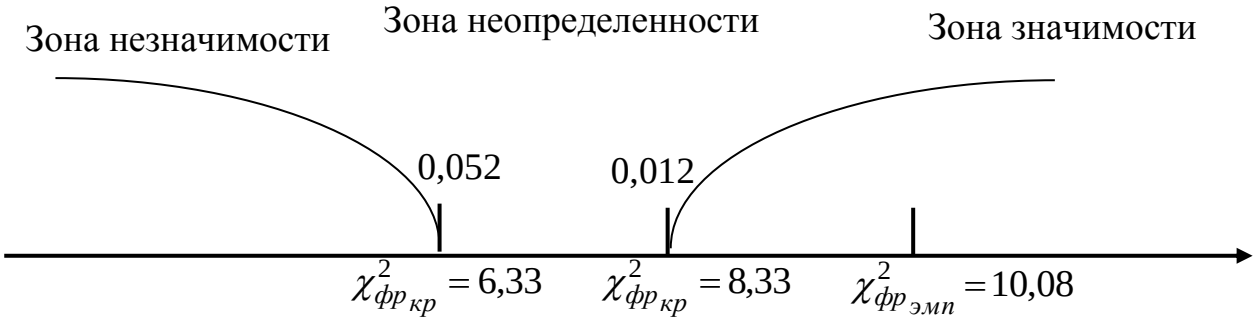

Рис. 21. Пример «оси значимости» для критерия Фридмана

Таким образом, $\chi^2_{фр_{эмп}}$ попало в зону значимости, следовательно, различия в показателях депрессии для 3 замеров значимы. То есть принимается гипотеза H_1 о различиях, а гипотеза H_0 о сходствах отклоняется.

6.1.4. L-критерий тенденций Пейджа

Данный критерий позволяет выявить различия, а также указывает на направление в изменении величин признака.

Условия применения

1. Объем выборки не должен быть больше 12.
2. Число измерений не должно превышать 6.
3. Измерение должно быть произведено в порядковой, интервальной или шкале отношений.

Алгоритм расчета

1. Проранжировать экспериментальные данные по строкам.
2. Определить сумму полученных рангов по столбцам.
3. Определить сумму теоретических рангов по формуле:

$$R = \frac{n \cdot c \cdot (c+1)}{2},$$

где c – число столбцов (измерений);

n – число строк (испытуемых).

4. Сравнить суммы рангов (теоретическую и эмпирическую).
5. Проранжировать сумму рангов замеров.
6. Рассчитать эмпирическое значение L-критерия тенденций Пейджа по формуле:

$$L_{эмп} = \sum_{i=1}^{c} (R_i \cdot i),$$

где R_i – сумма рангов i-го столбца в упорядоченной таблице;

i – порядковый номер столбца, образовавшегося в новой таблице;

c – число измерений (столбцов).

7. По таблице критических значений определить $L_{кр}$ для определенного количества измерений (c) и числа испытуемых (n). Сравнить с Lэмп. Построить ось значимости.

Пример. Воспользуемся примером из предыдущего параграфа. Как видно из табл. 18, ранги 1, 2, 3 замеров равны 17,5; 12; 6,5 соответственно (так как данная таблица взята из предыдущего примера, пункты 1–4 алгоритма расчета L-критерия тенденций Пейджа выполнены, перейдем к п. 5).

Таблица 18

Результаты расчетов критерия Фридмана

№ п/п	1-й замер депрессии	Ранги 1-го замера	2-й замер депрессии	Ранги 2-го замера	3-й замер депрессии	Ранги 3-го замера
1	20	2,5	20	2,5	18	1
2	21	3	19	2	18	1
3	18	3	16	1,5	16	1,5
4	17	3	16	2	14	1
5	20	3	18	2	17	1
6	25	3	21	2	19	1
Сумма		17,5		12		6,5
Общая сумма						**36**

Проранжируем сумму рангов замеров. Каждой величине нового ряда поставим в соответствие его ранг (в формуле $L_{эмп} = \sum_{i=1}^{c}(R_i \cdot i)$ обозначен как R_i). Итак, $6{,}5 \rightarrow i = 1$; $12 \rightarrow i = 2$; $17{,}5 \rightarrow i = 3$.

Рассчитаем $L_{эмп} = (6{,}5 \cdot 1) + (12 \cdot 2) + (17{,}5 \cdot 3) = 83$.

Определим критические значения L-критерия Пейджа для 3 измерений и 6 испытуемых:

$$L_{кр} = \begin{cases} 83 \ для\ p \le 0{,}001 \\ 81 \ для\ p \le 0{,}01 \\ 79 \ для\ p \le 0{,}05. \end{cases}$$

Построим ось значимости (рис. 22):

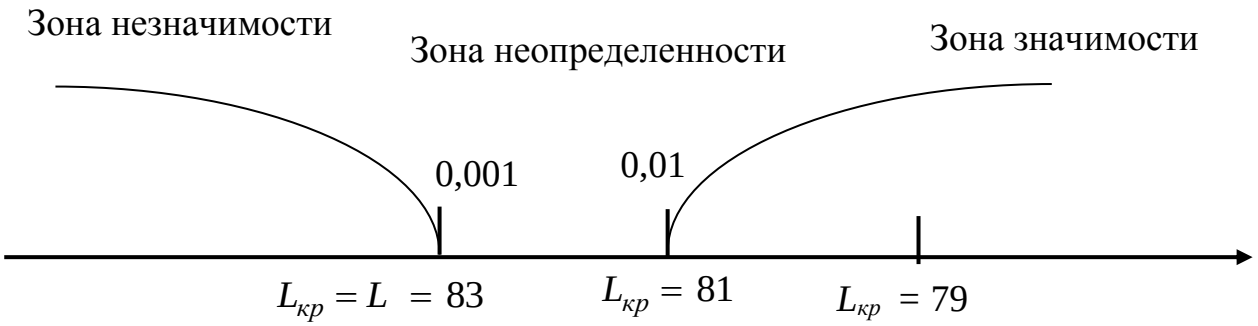

Рис. 22. Пример «оси значимости» для L-критерия тенденций Пейджа

Ось значимости показывает, что $L_{эмп}$ совпало с $L_{кр}$ ($L_{эмп}$ = 83, $L_{кр}$ = 83), следовательно, можно говорить о том, что различия в показателях депрессии для 3 замеров значимы на 0,1 %-м уровне значимости. Анализируя данные, полученные с помощью критерия Фридмана и L-критерия тенденций Пейджа, можно прийти к выводу о том, что данные критерии не согласуются. Но это не так. Критерий Фридмана

и L-критерий тенденций Пейджа обращаются к разным сторонам анализируемых данных, тем самым характеризуя различные аспекты обрабатываемых данных. Напомним о том, что критерий Фридмана выявляет наличие различий в измеренных признаках (показателях), а L-критерий тенденций Пейджа выявляет тенденцию в изменениях величин измеряемых признаков (показателей).

6.2. Непараметрические критерии для несвязанных выборок

6.2.1. U-критерий Манна-Уитни

Напомним, что данный критерий предназначен для оценки различий между двумя выборками по уровню какого-либо признака, количественно измеренного. Он позволяет выявлять различия между малыми выборками, когда $n_1 \cdot n_2 \geq 3$ или $n_1 = 2$, $n_2 \geq 5$.

Данный метод определяет, достаточно ли мала зона перекрещивающихся значений между двумя рядами. Чем меньше область перекрещивающихся значений, тем более вероятно, что различия достоверны.

Эмпирическое значение U-критерия Манна-Уитни отражает то, насколько велика зона совпадения между рядами. Поэтому чем меньше $U_{эмп}$, тем более вероятно, что различия достоверны.

Гипотезы:

H_0: уровень признака в группе 2 не ниже уровня признака в группе 1;

H_1: уровень признака в группе 2 ниже уровня признака в группе 1.

Существует несколько ограничений критерия Манна-Уитни:

1. В каждой выборке должно быть не менее 3 наблюдений: n_1, $n_2 \geq 3$.
2. В каждой выборке должно быть не более 60 наблюдений; n_1, $n_2 \leq 60$.

Алгоритм нахождения

1. Объединим эмпирические данные из 1-й и из 2-й выборок.
2. Произведем сортировку по возрастанию.
3. Проранжируем эмпирические данные. Для этого используем функцию РАНГ.СР. В графе Число указываем первое значение первой выборки. В графе Ссылка определяем список чисел (объединенные эмпирические данные 1-й и 2-й выборок). В графе Порядок указываем 1 (ИСТИНА, если сортировка производилась по возрастанию) или 0 (ЛОЖЬ, если данные сортировались по убыванию).

Не забываем определить, что ячейки списка абсолютные. Используя функцию автозаполнения, заполним все ячейки 1-й и 2-й выборок.

4. Найдем сумму рангов 1-й и 2-й выборок (534,5 и 740,5).

5. Для проверки правильности ранжирования определим общую сумму рангов (1275) и сравним ее с расчетной суммой по формуле:

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

где N – общее количество ранжируемых наблюдений (значений).

Несовпадение реальной и расчетной сумм рангов будет свидетельствовать об ошибке, допущенной при начислении рангов или их суммировании.

6. Определить большую из двух ранговых сумм.

7. Определить значение U по формуле:

$$U = (n_1 \cdot n_2) + \frac{n_x \cdot (n_x + 1)}{2} - T_x,$$

где n_1 – количество испытуемых в выборке 1;

n_2 – количество испытуемых в выборке 2;

T_x – большая из двух ранговых сумм;

n_x – количество испытуемых в группе с большей суммой рангов.

8. Определить критические значения U по таблице критических значений[1]. Если $U_{эмп} > U_{кр}$, H_0 принимается. Если $U_{эмп} \leq U_{кр}$, H_0 отвергается. Чем меньше значения U, тем достоверность различий выше.

Таким образом, мы показали как, избегая сложных расчетов, применить наиболее часто используемые критерии статистической значимости, а также определить соответствие эмпирических данных нормальному закону распределения.

6.2.2. Q-критерий Розенбаума

Данный критерий применим для оценки различий между двумя выборками по уровню какого-либо признака, количественно измеренного. Объем выборок не менее 11 испытуемых.

Если же Q-критерий выявляет достоверные различия между выборками с уровнем значимости $p < 0{,}01$, можно ограничиться только им и избежать трудностей применения других критериев.

---

[1] См. Приложение 6.

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 будет очень трудно установить различия.

Условия применения

1. Объем каждой выборки не менее 11.

2. Если в обеих выборках меньше 50 наблюдений, то абсолютная величина разности между n_1 и n_2 не должна быть больше 10 наблюдений.

3. 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.

4. 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.

Алгоритм расчета

1. Проверить, выполняются ли ограничения: $n_1, n_2 \geq 11$, $n_1, n_2 \approx n_2$.

2. Отсортировать значения в каждой выборке по возрастанию. Пронумеровать выборки, причем 1-й будет являться та выборка, значения в которой предположительно выше, а 2-й – где значения предположительно ниже.

3. Определить максимальное значение во 2-й выборке.

4. Определить количество значений в 1-й выборке, которые выше максимального значения в выборке 2. Обозначить полученную величину как S_1.

5. Определить минимальное значение в 1-й выборке.

6. Определить количество значений во 2-й выборке, которые ниже минимального значения выборки 1. Обозначить полученную величину как S_2.

7. Рассчитать $Q_{эмп}$ по формуле:

$$Q_{эмп} = S_1 + S_2.$$

8. Определить $Q_{кр}$ по таблице критических значений[1].

Пример. У курсантов был измерен уровень вербального интеллекта с помощью методики Д. Векслера. Было обследовано 26 кур-

[1] См. Приложение 7.

сантов в возрасте от 18 до 24 лет (средний возраст 20,5 лет): 12 из них были курсантами факультета информационной безопасности, а 14 – курсантами психологического факультета Московского университета МВД России. Полученные данные представлены в табл. 19. Психолог должен определить, есть ли различия в показателях вербального интеллекта у курсантов разных факультетов.

Таблица 19

Показатели вербального интеллекта у курсантов факультета информационной безопасности и психологического факультетов

Психологический факультет		Факультет информационной безопасности	
№ п/п	Показатель вербального интеллекта	№ п/п	Показатель вербального интеллекта
1	132	1	126
2	134	2	127
3	124	3	132
4	132	4	120
5	135	5	119
6	132	6	126
7	131	7	120
8	132	8	123
9	121	9	120
10	127	10	116
11	136	11	123
12	129	12	115
13	136		
14	136		

Как видно из представленных данных, объем первой и второй выборок равен 14 и 12 соответственно. Обе выборки примерно равны.

Отсортировав данные, видим, что максимальное значение (136) располагается в выборке психологического факультета, значит, она и будет являться выборкой № 1. Максимальное значение в выборке № 2 (факультет информационной безопасности) – 132. Определяем количество значений в выборке № 1, которые выше максимального значения в выборке № 2 (5). Обозначим полученную величину как S_1. Минимальным значением в выборке № 1 является 121. Определим количество значений в выборке № 2, которые ниже минимального в выборке № 1 (6). Обозначим полученную величину как S_2. Теперь рассчитаем $Q_{эмп}$ по формуле:

$$Q_{эмп} = S_1 + S_2 = 5 + 6 = 11.$$

Определим $Q_{кр}$ по таблице критических значений для $n_1 = 14$, $n_2 = 12$:

$$Q_{кр} = \begin{cases} 7 \ для \ p \le 0{,}05 \\ 9 \ для \ p \le 0{,}01. \end{cases}$$

Построим ось значимости (рис. 23):

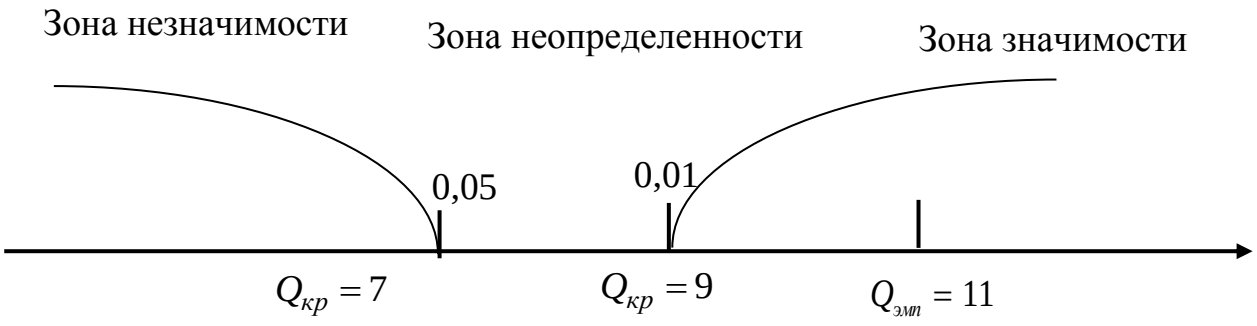

Рис. 23. Пример «оси значимости» для Q-критерия Розенбаума

Как показывает ось значимости, $Q_{эмп} > Q_{кр}$, а значит, принимается гипотеза H_1. Курсанты психологического факультета превосходят курсантов факультета информационной безопасности по уровню вербального интеллекта ($p < 0{,}01$).

Вопросы для самоконтроля

1. Условия применения G-критерия знаков для зависимых выборок.
2. Алгоритм расчета G-критерия знаков для зависимых выборок.
3. Условия применения парного Т-критерия Вилкоксона.
4. Алгоритм расчета парного Т-критерия Вилкоксона.
5. Условия применения критерия Фридмана.
6. Алгоритм расчета критерия Фридмана.
7. Условия применения L-критерия тенденций Пейджа.
8. Алгоритм расчета L-критерия тенденций Пейджа.
9. Условия применения U-критерия Манна-Уитни.
10. Алгоритм нахождения U-критерия Манна-Уитни.
11. Условия применения Q-критерия Розенбаума.
12. Алгоритм расчета Q-критерия Розенбаума.

ТЕМА 7. Корреляционный и регрессионный анализ

7.1. Корреляционный анализ

Корреляционный анализ широко применяется в психологических и педагогических исследованиях. Он применяется тогда, когда две переменные представлены в числовой шкале. Сутью корреляционного анализа является проверка статистической связи между несколькими переменными.

Психолога-исследователя зачастую волнует вопрос: есть ли связь и какая именно: между двумя или несколькими переменными в исследуемых группах? Например, могут ли сотрудники с высоким уровнем конфликтности оперативно решать задачи в особых условиях? Или: как влияет склонность к депрессиям на успеваемость студентов? Ответить на подобные вопросы помогает корреляционный анализ.

Планируя исследование, необходимо выделить изучаемые свойства, операционализовать их, определить шкалы, с помощью которых они будут измерены, и на основе эмпирических результатов выбрать метод, позволяющий проверить поставленные гипотезы.

Существует два вида взаимосвязей между явлениями и их признаками: функциональная (строго детерминированная) и статистическая (стохастически детерминированная).

Функциональная связь представляет собой вид причинной зависимости, при котором определенному значению факторного признака соответствует одно или несколько точно заданных значений результативного признака.

Стохастическая связь представляет собой вид причинной зависимости, проявляющейся не в каждом отдельном случае, а в общем, в среднем, при большом числе наблюдений.

Частным случаем стохастической связи является корреляционная связь.

Корреляционная связь – зависимость среднего значения результативного признака от изменения факторного признака.

Виды корреляционных связей

1) парная корреляция – связь между двумя признаками (результативным и факторным);

2) частная корреляция – зависимость между результативным и одним факторным признаками при фиксированном значении других факторных признаков;

3) множественная корреляция – зависимость результативного и двух или более факторных признаков, включенных в исследование.

Этапы изучения корреляционных зависимостей

1. Предварительный анализ свойств совокупности.
2. Определение факта наличия связи, установление ее направления и формы.
3. Измерение степени тесноты связи между признаками.
4. Построение регрессионной модели, т. е. нахождение аналитического выражения связи.
5. Оценка адекватности модели, ее экономическая интерпретация и практическое использование.

Коэффициент корреляции – количественная мера взаимосвязи нескольких переменных, показывающая согласованность их изменений.

Показатели коэффициента корреляции

Корреляционную связь можно охарактеризовать несколькими показателями:

1. *Сила связи* или *теснота*. Она передает, насколько значения переменных разбросаны относительно средних значений:

1) сильная или тесная при коэффициенте корреляции $r \geq 0{,}70$;

2) средняя при $0{,}50 \leq r \leq 0{,}69$;

3) умеренная при $0{,}30 \leq r \leq 0{,}49$;

4) слабая при $0{,}20 \leq r \leq 0{,}29$;

5) очень слабая при $r \leq 0{,}19$.

2. *Направление связи.* Значения переменных варьируются относительно друг друга в двух направлениях:

– положительная или прямая связь, т. е. высоким значениям одного признака характерны высокие значения второго признака.

– отрицательная или обратная корреляция, т. е. высоким значениям одной переменной соответствуют низкие значения другой переменной.

Значение коэффициента корреляции не всегда отражает наличие связи.

3. *Надежность.* Для установления наличия взаимосвязи применяют уровень значимости, относительно которого принимаются гипотезы о неслучайности выявленной взаимосвязи. Данный показатель характеризует связь между переменными. В связи с тем, что распределение переменных является вероятностным, существуют различные

причины, определяющие полученное на конкретном объекте соотношение признаков. Для этого проверяют гипотезы о значимости статистической связи. Основная гипотеза, исследуемая в корреляционном анализе H_0, устанавливает значимость отличий от 0. Альтернативная гипотеза H_1 выставляется как противоположная главной, т. е. связь значимо отличается от нуля. Таким образом, изменчивость двух признаков согласована.

Чем больше объем выборки и чем больше абсолютная величина коэффициента корреляции, тем выше его статистическая значимость. Стоит отметить, что при большой численности выборки слабые корреляционные связи также могут достигать высокого уровня статистической значимости.

Интерпретация коэффициента корреляции

Выделим несколько причин, по которым может быть не выявлена корреляционная связь:

1. Наличие точек выброса, что может существенно менять коэффициенты корреляции, в том числе изменять знак корреляции.

2. Выраженная асимметрия распределения одного или нескольких признаков (характерна для больших объемов выборок). Для устранения возможно провести стандартизацию переменных и на основе новых значений найти коэффициент корреляции.

3. Неоднородность выборки. В выборочной совокупности могут быть представлены на самом деле несколько групп из различных генеральных совокупностей. Неоднородность выборки можно обнаружить с помощью нахождения моды. Если выделяется две и более моды, то необходимо проверить выборку на неоднородность.

4. Нелинейность связи. Чаще всего описание взаимосвязи осуществляется в виде линейной функции, хотя корреляционная связь может быть описана и различными кривыми.

Корреляционный анализ позволяет определить взаимосвязь, но с его помощью нельзя выделить причинно-следственные связи. Различие в способах нахождения соотношения факторов чаще всего связано с типами шкал измерения (номинативная, порядковая, интервальная и шкала отношений).

Если перед исследователем стоит задача определить связь между двумя факторами под влиянием третьего, то можно применить критерий частной корреляции.

В связи с тем, что переменные X и Y могут быть измерены в разных шкалах, необходимо правильно выбрать соответствующий коэф-

фициент корреляции. Представленная ниже таблица позволяет определить соотношения между типами измерительных шкал, в которых могут быть измерены переменные *X* и *Y* соответствующими мерами связи.

Таблица 20

Зависимость меры связи и шкал измерений

Типы шкал		**Мера связи**
Переменная *X*	**Переменная *У***	
Интервальная или шкала отношений	Интервальная или шкала отношений	Коэффициент Пирсона r_{xy}
Порядковая, интервальная или шкала отношений	Порядковая, интервальная или шкала отношений	Коэффициент Спирмена r_s
Ранговая	Ранговая	Коэффициент Кендалла τ
Дихотомическая	Дихотомическая	Коэффициент φ
Дихотомическая	Ранговая	Рангово-бисериальный коэффициент R_{rb}
Дихотомическая	Интервальная или шкала отношений	Бисериальный коэффициент $R_{бис}$

Основным критерием нахождения взаимосвязи переменных является *r-критерий Пирсона,* применяемый тогда, когда два фактора являются количественными, нормально распределены, измерены в интервальной шкале и идентичны по форме распределения.

7.1.1. Коэффициент корреляции Пирсона r_{xy}

Коэффициент корреляции Пирсона r_{xy} позволяет установить силу и направление корреляционной связи между двумя признаками.

Расчет коэффициента корреляции Пирсона предполагает наличие двух рядов значений, таких как:

– два признака, измеренные в одной и той же группе испытуемых;

– два ряда значений по одному и тому же признаку, полученные в двух группах испытуемых.

Вычисляют коэффициент корреляции Пирсона по формуле:

$$r_{xy} = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \cdot (y_i - \bar{y})^2}},$$

где x_i – значения, принимаемые переменной *x*;

y_i – значения, принимаемые переменной *y*;

$\bar{x}$ – среднее значение по *X*;

$\bar{y}$ – среднее значение по У.

Гипотезы:

H_0: корреляция между переменными X и Y не отличается от нуля;

H_1: корреляция между переменными X и Y достоверно отличается от нуля.

Условия применения

1. Сравниваемые переменные должны быть получены в интервальной шкале или шкале отношений.

2. Распределения переменных X и Y должны быть близки к нормальному.

3. Число варьирующих признаков в сравниваемых переменных X и Y должно быть одинаковым.

4. По каждой переменной должно быть представлено не менее 5 наблюдений. Верхняя граница выборки определяется имеющимися таблицами критических значений, а именно $n \leq 5000$.

Алгоритм расчета

1. Обозначить признаки как переменные X и Y.

2. Вычислить произведение переменных X и Y.

3. Рассчитать значение каждой переменной (отдельно X и Y) в квадрате.

4. Определить суммы: каждой переменной (X и Y), квадратов каждой переменной (X^2 и Y^2), последовательных произведений переменных друг на друга в соответствующих ячейках.

5. По формуле рассчитать коэффициент корреляции Пирсона:

6. По таблице критических значений определить $r_{xy(кр)}$, сравнить его с $r_{xy(эмп)}$. Если $r_{xy(эмп)} \geq r_{xy(кр)}$, то корреляция достоверно отличается от нуля (H_0 отвергается).

В связи с тем, что вычисления данной статистики довольно трудоемки, рассмотрим способы ее расчета с помощью статистического пакета в MS Excel.

Первый способ. Выделяем ячейку, в которой хотим получить вычисляемое значение. В полном алфавитном перечне выберем функцию PEARSON. В появившемся окне в поле *Массив 1* вводим диапазон ячеек первой переменной, а в поле *Массив 2* – диапазон ячеек второй переменной. Нажимаем кнопку ОК. Оба аргумента должны содержать одинаковое количество значений. В ячейке, которую мы выделили для записи функции, появится значение коэффициента корреляции Пирсона.

Второй способ. На вкладке Данные выберем раздел Анализ данных. В появившемся окне найдем функцию Корреляция. В поле Входной интервал укажем диапазон значений обеих переменных (рис. 24). Определим входной интервал, установив курсор на необходимой ячейке.

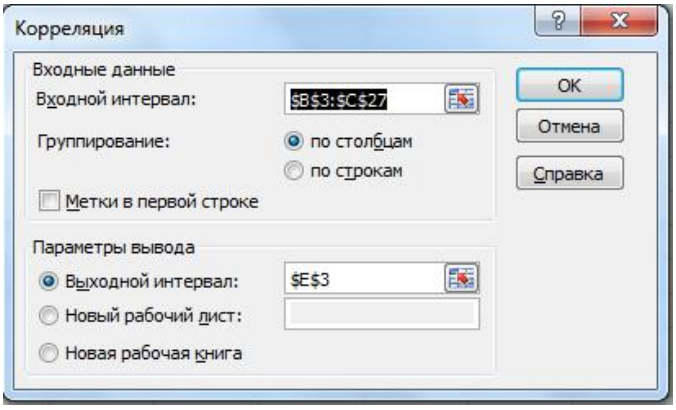

Рис. 24. Расчет корреляционной связи

В результате получим корреляционную матрицу (рис. 25).

	Столбец 1	*Столбец 2*
Столбец 1	1	
Столбец 2	0,321848924	1

Рис. 25. Корреляционная матрица

Затем сопоставим полученное значение с критическим и, в зависимости от результата, примем или отклоним основную гипотезу как наиболее вероятную.

Пример. Изучив влияние психологических детерминант на уровень конфликтности среди сотрудников органов внутренних дел, обозначились следующие корреляционные связи.

Просматривается корреляционная связь конфликтности с фактором «моральная нормативность» (рис. 26). Это объясняется следующим: высокие показатели этого фактора характерны людям, которые требовательны к себе, настойчивы, ответственны, с выраженными волевыми чертами. Однако отрицательную связь можно объяснить тем, что при низкой нормативности поведения отсутствует установка на общепринятые правила и стандарты поведения. Пропадает зависимость от них. Освобождение от влияния группы делает деятельность более плодотворной.

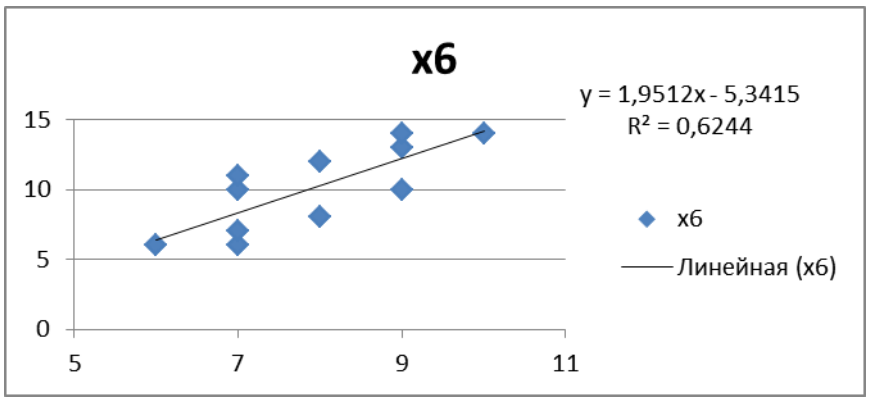

Рис. 26. Положительная корреляционная связь моральной нормативности с уровнем конфликтности

Также наблюдается отрицательная связь фактора «эмоциональная чувствительность» с конфликтностью, что логично, так как для стратегии «соперничество» свойственна самоуверенность, суровость, некоторая жестокость и черствость по отношению к окружающим (рис. 27). Отрицательный показатель данного фактора характеризует практичность, независимость, скептическое отношение к некоторым аспектам жизни, порой самодовольство.

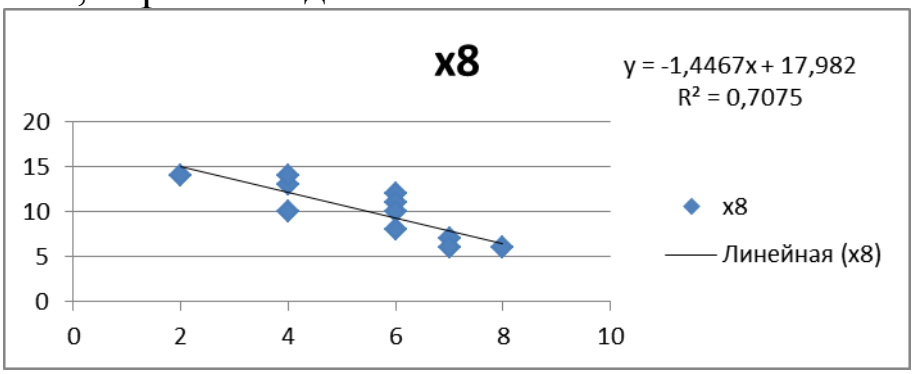

Рис. 27. Связь эмоциональной чувствительности и уровня конфликтности

7.1.2. Коэффициент корреляции Спирмена r_s

В том случае, если данные выборки не имеют нормального распределения, применяют коэффициент Спирмена r_s, который позволяет оценить меру взаимосвязи без предварительного анализа выборочных данных. Данный коэффициент используют и для количественных, и для порядковых значений.

Для подсчета ранговой корреляции необходимо располагать двумя рядами значений, которые могут быть проранжированы. Примерами таких рядов являются:

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 (например, личностные профили по 16-факторному опроснику Р. Б. Кеттелла, последовательности предпочтений в выборе из нескольких альтернатив и др.);

3) две групповые иерархии признаков;

4) индивидуальная и групповая иерархии признаков.

Коэффициент корреляции Спирмена r_s рассчитывается по формуле:

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

где n – объем выборки;

d – разность между рангами по двум переменным для каждого испытуемого;

$\Sigma(d^2)$ – сумма квадратов разностей между рангами.

Гипотезы:

При подсчете ранговой корреляции возможно выдвинуть два вида гипотез.

1. Если два признака измерены в одной и той же группе испытуемых:

H_0: корреляция между переменными X и Y не отличается от нуля;

H_1: корреляция между переменными X и Y достоверно отличается от нуля.

2. Если исследуются: две индивидуальные иерархии признаков, выявленные у двух испытуемых по одному и тому же набору признаков; две групповые иерархии признаков; индивидуальная и групповая иерархии признаков:

H_0: корреляция между иерархиями X и Y не отличается от нуля;

H_1: корреляция между иерархиями X и Y достоверно отличается от нуля.

Условия применения

1. Наличие не менее 5 наблюдений. Верхняя граница выборки определяется имеющимися таблицами критических значений, а именно $n \leq 40$. В случае использования большего, чем 40, числа ранжируемых признаков, уровень значимости коэффициента корреляции следует определять по таблице критических значений для коэффициента корреляции Пирсона.

2. Коэффициент корреляции Спирмена r_s, при большом количестве одинаковых рангов по одной или обеим сопоставляемым переменным дает приблизительные значения. Предполагается, что оба коррелируемых ряда должны представлять собой две последовательности несовпадающих значений. Если это условие не соблюдается, необходимо вносить поправку на одинаковые ранги.

Алгоритм расчета

1. Обозначить признаки (иерархии) как переменные *X* и *Y*.
2. Проранжировать значения переменной *X*.
3. Проранжировать значения переменной *Y*.
4. Вычислить разности *d* между рангами *X* и *Y* по каждому испытуемому.
5. Определить d^2 (квадрат разностей рангов).
6. Рассчитать сумму квадратов $\Sigma(d^2)$.
7. При наличии одинаковых рангов рассчитать поправки:

$$T = \frac{\sum(x^3 - x)}{12};\ T = \frac{\sum(y^3 - y)}{12},$$

где *x* – объем каждой группы одинаковых рангов в ранговом ряду *X*;

y – объем каждой группы одинаковых рангов в ранговом ряду *Y*.

8. Вычислить коэффициент ранговой корреляции r_s.
9. Формула расчета коэффициента ранговой корреляции Спирмена rs при наличии одинаковых рангов:

$$r_s = 1 - 6 \cdot \frac{\sum(d^2) + T_x + T_y}{n \cdot (n^2 - 1)},$$

где $\Sigma(d^2)$ – сумма квадратов разностей между рангами;

T_x и T_y – поправки на одинаковые ранги;

n – объем выборки.

10. По таблице критических значений определить $r_{s(кр)}$, сравнить его с $r_{s(эмп)}$. Если $r_{s(эмп)} \geq r_{s(кр)}$, то корреляция достоверно отличается от нуля (H_0 отвергается).

7.1.3. Коэффициент корреляции τ Кендалла

Коэффициент корреляции τ Кендалла является непараметрическим и применяется для работы с данными, измеренными в порядковой шкале. С помощью данного коэффициента возможно установить

силу и направление корреляционной связи между двумя признаками или двумя иерархиями признаков.

Условия применения

1. Исследуемые данные должны быть измерены в ранговой шкале.
2. Необходимо равенство количества варьирующих признаков в сравниваемых переменных *X* и *Y*.
3. При вычислении данного коэффициента нельзя использовать одинаковые ранги.
4. Для определения уровня достоверности коэффициента τ применяется формула расчета и таблица критических значений для t-критерия Стьюдента при $df = n - 2$.

Алгоритм расчета

1. Обозначить признаки (иерархии), принимающие участие в сопоставлении как переменные *X* и *Y*.
2. Проранжировать значения переменной *X* и *Y* (наименьшему значению присваивается 1-й ранг).
3. Отсортировать значения рангов переменной *X* по возрастанию, соответственно поменяются местами и соответствующие ранги переменной *Y*. Записать ранги *X* в первый столбец таблицы по порядку номеров испытуемых или признаков.
4. Записать ранги переменной *Y* в соответствующие ячейки второго столбца таблицы.
5. Определить количество совпадений (*P*) и инверсий (*Q*).
6. Определим правильность подсчета числа совпадений по формулам:

$$P + Q = \frac{(n-1)\cdot n}{2};$$

$$\tau = 1 - \frac{4\cdot Q}{(n-1)\cdot n};$$

$$\tau = \frac{4\cdot P}{(n-1)\cdot n} - 1.$$

7. Установим уровень значимости коэффициента корреляции с помощью t-критерия Стьюдента по формуле:

$$t = \tau \cdot \sqrt{\frac{n-2}{1-\tau^2}},$$

где τ – полученное эмпирическое значение коэффициента корреляции.

Критические значения определяются по таблице критических значений для t-критерия Стьюдента.

Если выполняется неравенство $t_{эмп} \geq t_{кр}$, то H_0 отклоняется.

7.2. Регрессионный анализ

Взаимосвязь между переменными можно описать не только с помощью различных коэффициентов корреляции, но и как зависимость между аргументом *X* и функцией *Y*. В таком случае необходимо определить зависимость типа $Y = F(X)$ или $X = F(Y)$.

Регрессия – это изменение функции в зависимости от изменений одного или нескольких аргументов.

Линия регрессии – это графическое выражение регрессионного уравнения.

Предикторы – независимые переменные (*X*), предсказанные зависимой переменной (*Y*).

Зависимость переменных можно выразить следующим образом:

$$y_i = a + b \cdot x_i + \varepsilon_i,$$

где y_i – зависимая переменная;

x_i – независимая переменная или причина, фактор;

a – свободный член – прогнозируемое значение зависимой переменной при условии, что независимая переменная равна 0;

b – коэффициент регрессии, задающий угол наклона линии, показывает влияние независимой переменной;

ε_i – случайная ошибка, которая может быть связана с ошибками измерения, неучтенные факторы и т. п.

Главной задачей регрессионного анализа является оценка неизвестных параметров a и b, задающих линейную зависимость x и y. Эту задачу решает метод наименьших квадратов, который рассчитывается как расстояние, вычисленное по оси *Y*:

$$\sum (y_i - \hat{y}_i)^2 = \sum e_i^2 = \min,$$

где y_i – истинное *i*-значение *Y*;

$\hat{y}_i$ – оценка *i*-значения *Y* при помощи линии регрессии;

$e_i = y_i - \hat{y}_i$ – ошибка оценки.

Нахождение наименьшего расстояния можно продемонстрировать на графике (рис. 28).

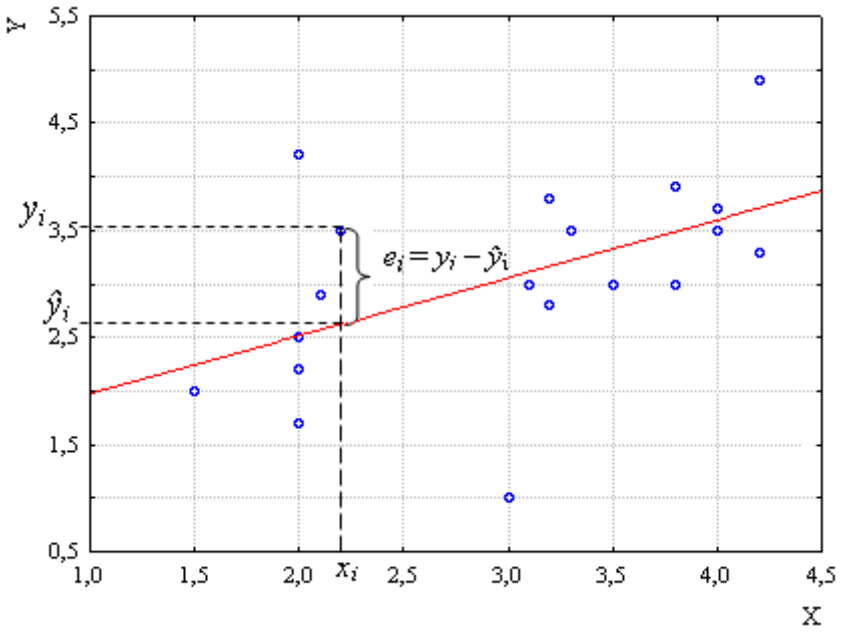

Рис. 28. График рассеяния и линия регрессии (e_i – ошибка оценки для одного из объектов)

Следовательно, если на некоторой выборке измерены две переменные, коррелирующие друг с другом, то, вычислив коэффициенты регрессии, можно на основе известных значений одной переменной предсказать неизвестные значения другой переменной. Данная модель позволяет спрогнозировать изменения *y* при определенных значениях *x*. Коэффициент регрессии *b* показывает среднее изменение зависимой переменной с изменением *x* на одну единицу его измерения. *b* часто представляют через коэффициент корреляции, но это не одно и то же. Они связаны следующим образом:

$$b = r_{xy}\frac{\sigma_y}{\sigma_x},\ a = M_y - b \cdot M_x.$$

Если $|r_{xy}| = 1$, то предсказание будет наиболее точным: каждому значению *x* будет соответствовать одно значение *y*. В случае $|r_{xy}| = 0$, предсказательная ценность регрессии ничтожна, так как при любом *x* значение *y* будет равно среднему значению.

Особое значение для оценки точности предсказания имеет дисперсия оценок зависимой переменной. При коэффициенте корреляции дисперсия оценок равна 0, а при коэффициенте равном единице она достигает истинного значения (не искаженного ошибками измерения и другими факторами, которые вносят изменчивость в связи между переменными), достигая своего максимума. Дисперсия оценок

зависимой переменной Y – это та часть ее полной дисперсии, которая обусловлена влиянием независимой переменной X. Отношение же дисперсий оценок к истинной дисперсии, как можно увидеть на основе определенных преобразований, равно квадрату коэффициента корреляции. В связи с этим коэффициент корреляции, возведенный в квадрат, показывает, в какой степени изменчивость одной переменной обусловлена влиянием другой переменной. Данный коэффициент получил название *коэффициент детерминации* r_{xy}^2.

Коэффициент детерминации обладает рядом преимуществ по сравнению с коэффициентом корреляции. Корреляция, полученная на разных выборках при объединении этих выборок в одну, может измениться. А коэффициент детерминации, в свою очередь, допускает усреднение для нескольких выборок, в связи с чем возможно перенесение полученных результатов на другие случаи (предсказывание подобных случаев). Коэффициент детерминации позволяет оценить степень влияния независимой переменной на зависимую (коэффициент корреляции не имеет такой возможности).

В связи с тем, что многие явления могут быть поняты при условии влияния нескольких переменных, стало необходимым устанавливать связи между множеством переменных. Данную задачу решает множественный регрессионный анализ.

Множественный регрессионный анализ позволяет проанализировать связь между несколькими независимыми переменными и одной зависимой, а также определить степень влияния той или иной переменной на зависимую, при условии, что все другие переменные остаются неизменными.

Связь между зависимой и независимыми переменными можно выразить в виде линейного уравнения:

$$y_i = b + b_1 \cdot x_1 + b_2 \cdot x_2 + \ldots + b_{p-1} \cdot x_{p-1} + b_p \cdot x_p + \varepsilon_i,$$

где y_i – зависимая переменная;

$x_1, x_2, \ldots, x_{p-1}, x_p$ – независимые переменные;

b – свободный член;

$b_1, b_2, \ldots, b_{p-1}, b_p$ – параметры модели;

ε_i – ошибка предсказания.

Например, исследователю необходимо определить факторы, оказывающие влияние на текущую успеваемость курсантов: интеллектуальные способности, различные черты характера, коммуникативные способности, особенности темперамента, наличие определенных цен-

ностей и др. Данную задачу можно решить с помощью множественного регрессионного анализа.

Условия выбора переменных для применения множественного регрессионного анализа

1. Необходимо обозначить переменную, которая будет рассматриваться как зависимая (должна быть представлена одним значением).

2. Зависимая переменная должна быть измерена в количественной шкале (если зависимая переменная представлена качественно, то используется логистическая регрессия).

3. Независимые переменные могут быть представлены и качественно, и количественно, однако необходимо помнить о том, что они не должны коррелировать друг с другом (так как в регрессионном анализе оценивается влияние переменной как изолированной характеристики на изучаемое свойство). Для выполнения данного условия необходимо провести корреляционный анализ.

Чтобы построить качественную предсказательную модель, необходимо определить возможные причины, способные изменяться самостоятельно или исследователем. Такая модель демонстрирует наиболее и наименее существенные показатели. Коэффициенты b дают возможность оценить степень влияния и соотношение между зависимой и независимой переменными.

С помощью уравнения регрессии возможно провести оценку предполагаемых значений зависимой переменной и реальных результатов. Сравнение таких значений позволяет установить позиции, недооцененные в реальной ситуации (располагаются ниже линии регрессии) и переоцененные (располагаются выше линии регрессии).

После вычисления регрессионных коэффициентов по значениям независимых переменных могут быть вычислены оценки зависимой переменной Y:

$$\hat{Y} = b + b_1 \cdot x_1 + b_2 \cdot x_2 + \ldots + b_p \cdot x_p.$$

Для оценки возможных погрешностей предсказания используют анализ остатков и ошибок, представляющий собой сравнение значений зависимой переменной Y_i с их оценками $\hat{Y}_i$ по выборке, для которых значения по Y_i известны.

Значения $\hat{Y}_i$ могут быть определениями для испытуемых с неизвестными значениями зависимой переменной. Чтобы оценить ошибки и качество полученной модели, необходимо найти разность между

исходными и оцененными значениями, а также коэффициент множественной корреляции.

Коэффициент множественной корреляции – мера линейной связи одной переменной с множеством других переменных. Он принимает положительные значения от 0 (отсутствие связи) до 1 (строгая прямая связь).

При наличии нескольких переменных β-коэффициент регрессии зависит не только от корреляции данной переменной с зависимой, но и от вида корреляционной связи независимой переменной с другими независимыми переменными. Знак β-коэффициента соответствует знаку коэффициента корреляции. Абсолютная величина β-коэффициента является максимальной, т. е. равна коэффициенту корреляции с зависимой переменной, если данная независимая переменная не коррелирует ни с одной из других независимых переменных. Чем сильнее данная независимая переменная связана с другими независимыми переменными, тем меньше β-коэффициент.

Произведение коэффициента β_i на коэффициент корреляции r_{iY} этой переменной с зависимой – это вклад данной переменной в дисперсию зависимой переменной. Если корреляция с зависимой переменной выше, то и он выше. В связи с этим ценность независимой переменной в множественном регрессионном анализе определяется не только ее корреляционной связью с зависимой переменной, но и ее уникальностью – слабой корреляцией с другими переменными.

Коэффициент множественной детерминации – сумма вкладов всех переменных в зависимую переменную.

Коэффициент множественной детерминации равен коэффициенту множественной корреляции в квадрате:

$$R^2 = \sum_{i=1}^{p} \beta_i \cdot r_{iY} .$$

Данный коэффициент демонстрирует, насколько полученная модель позволяет предсказать изменения зависимой переменной, оценивая это через долю дисперсии, объясняемой данным набором переменных. Например, R^2 = 0,64, значит, 64 % дисперсии зависимой переменной объясняется исходными переменными. Построение качественной модели связано с увеличением коэффициента множественной детерминации.

Изменчивость зависимой переменной связана не только с влиянием независимых переменных, но и с влиянием неучтенных факторов.

Чтобы найти долю влияния этих факторов, необходимо вычесть множественный коэффициент детерминации из общей дисперсии зависимой переменной:

$$D_e = 1 - R^2,$$

где D_e – часть дисперсии, обусловленная влиянием неучтенных факторов, или дисперсии ошибки оценки.

Поэтому основным показателем множественного регрессионного анализа является коэффициент множественной детерминации, который является мерой линейной взаимосвязи одной переменной с совокупностью других переменных. R^2 принимает только положительные значения, изменяясь в пределах от 0 до 1.

Интерпретация уравнения множественной регрессии

Проанализировать полученную модель возможно с точки зрения ее предсказательных возможностей. Прежде всего, необходимо оценить коэффициент множественной детерминации. Если он приближен к 1 – модель качественная, если к 0 – необходимо проанализировать признаки, выступающие как независимые переменные, а также их связь с зависимой, верность определения причин, выбор той или иной модели анализа данных.

Рассматривается коэффициент множественной корреляции, который оценивается с помощью *F*-критерия Фишера (сопоставление изменчивости переменной с изменчивостью при условии наличия факторов, переменных). Если различия статистически значимы, то можно говорить о достоверности коэффициента множественной регрессии, что дает основания для перехода к содержательной интерпретации.

Результаты представлены в виде таблицы, где дано значение β-коэффициента для каждой переменной и *p*-уровень, на основании которого происходит включение переменных в модель. Если *p*-уровень выше заданного значения (0,05 или 0,01), то переменная исключается из модели и вновь происходит нахождение β-коэффициентов. Для описания структуры связей между переменными необходимо проанализировать β-коэффициенты. Они не только показывают, какой вклад каждая переменная внесла в общую модель, но и интерпретируются как коэффициенты корреляции: по тесноте и направлению.

Свободный коэффициент *B* содержательно не рассматривается, так как он показывает значение, которое будет у зависимой перемен-

ной при условии, что все независимые переменные равны 0. Однако он является важным при построении предсказательной модели.

Пример. Психолог-исследователь (на основе методики Р. Б. Кеттелла) с помощью регрессионного анализа получил профессиональные портреты для некоторых специальностей:

Психотерапевт = 0,72 A + 0,29 B + 0,29 H + 0,29 N,

Психодиагност = 0,31 A + 0,78 B + 0,47 N.

Буквами обозначены шкалы из опросника Р. Б. Кеттелла, в котором дана следующая характеристика обозначенных шкал:

Шкала A – сердечность, B – интеллектуальные способности, H – смелость/осторожность, N – проницательность/бестактность.

Приведенная выше модель показывает, что для психотерапевта больший вес имеют: сердечность, доброта, общительность, эмоциональность, также важными являются высокие умственные способности, отзывчивость, дружелюбие и точный ум, эмоциональная сдержанность, проницательность. В то время как для диагноста наибольшее влияние оказывает фактор B – высокие умственные способности, умение быстро соображать, а также проницательность, четкий ум, общительность, открытость.

Сравнивая полученные результаты, можно сказать, что в деятельности терапевта успешными будут корректные и отзывчивые люди, умеющие расположить к себе другого, поддержать позитивное взаимодействие. Профессия психодиагноста характерна для проницательных людей с высокими интеллектуальными способностями.

Полученную модель можно использовать для предсказания успешности в той или иной деятельности отдельного человека. Например, в результате проведения опроса конкретный человек набрал по факторам A – 5 стенов, B – 7 стенов, H – 4 стена, N – 5 стенов. Подставив эти значения в уравнение, получим значение 8,24 (при разбросе возможных значений зависимой переменной от 1,59 до 15,9). Вышеизложенное позволяет сделать прогноз, что, скорее всего, человек в психотерапии будет не очень успешным. Если его же значения поставим в модель психодиагноста, то получим значение 9,36 (при разбросе значений от 1,56 до 10,56). В данном случае можно прогнозировать, что человек будет иметь более высокие показатели в психодиагностической практике.

Вопросы для самоконтроля

1. Виды взаимосвязи между явлениями и их признаками.
2. Виды корреляционных связей.
3. Этапы изучения корреляционных зависимостей.
4. Показатели коэффициента корреляции.
5. Интерпретация коэффициентов корреляции.
6. Коэффициент корреляции Пирсона r_{xy} (условия применения и алгоритм расчета коэффициента корреляции Пирсона).
7. Коэффициент корреляции Спирмена r_s (условия применения и алгоритм расчета коэффициента ранговой корреляции Спирмена r_s).
8. Коэффициент корреляции τ Кендалла (условия применения и алгоритм расчета коэффициента корреляции τ Кендалла).
9. Понятие регрессионного анализа.
10. Условия выбора переменных для применения множественного регрессионного анализа.
11. Интерпретация уравнения множественной регрессии.

Библиографический список

Аникеева Н. В., Усачева И. В. Диагностика волевой регуляции курсантов образовательных организаций системы МВД // Международный журнал психологии и педагогики служебной деятельности. – 2017. – № 2. – С. 16–18.

Бояринцева Т. И., Мастихина А. А. Теория графов : методические указания. – М. : Московский государственный технический университет им. Н.Э. Баумана, 2014. – 37 с.

Волков Б. С. Методология и методы психологического исследования : учеб. пособие / Б. С. Волков, Н. В. Волкова, А. В. Губанов. – М. : Академический Проект, 2010. – 382 с.

Веретенников В. Н. Высшая математика. Математический анализ функций одной переменной. – СПб. : Российский государственный гидрометеорологический университет, 2013. – 254 с.

Дьячук А. А. Математические методы в психологических и педагогических исследованиях : учебное пособие. – Красноярск : Красноярский гос. пед. ун-т им. В.П. Астафьева, 2013. – 347 с.

Ермолаев-Томин О. Ю. Математические методы в психологии : учеб. для бакалавров. – 4-е изд., перераб. и доп. – М. : Юрайт, 2013. – 511 с.

Зыкова Н. Ю., Лапкова О. С., Хлоповских Ю. Г. Методы математической обработки данных психолого-педагогического исследования : учебное пособие. – Воронеж : Изд-во ВГУ, 2008. – 84 с.

Кинякин В. Н., Слесарева Е. А. Системы регрессионных (одновременных) уравнений // Вестник экономический безопасности. – 2016. – № 4. – С. 265–270.

Лавров И. А. Задачи по теории множеств, математической логике и теории алгоритмов : учебное пособие. – 5-е изд., испр. – М. : Физматлит, 2004. – 256 с.

Малахов А. Н. Математика. Высшая математика : учебное пособие. – М. : Евразийский открытый институт, 2009. – 64 с.

Ермолаев О. Ю. Математическая статистика для психологов : учебник. – 2-е изд., испр. – М. : Московский психолого-социальный институт ; Флинта, 2003. – 336 с.

Первитская А. М. Математические методы в психологии : учеб. пособие. – Курган : Изд-во Курганского гос. ун-та, 2013. – 70 с.

Слесарева Е. А., Смирнов Д. Е. Информационные технологии как средство решения практических задач в деятельности психолога // Государственная служба и кадры. – 2016. – № 4.

Слесарева Е. А. Корреляционный анализ как один из методов обработки результатов психологического исследования // Психология и педагогика служебной деятельности. – 2016. – № 4.

Слесарева Е. А., Задохина Н. В. Критерии статистической значимости как неотъемлемая часть экспериментальной психологии // Вестник Московского университета МВД России. – 2017. – № 4.

Слесарева Е. А. Математические методы в психологии как средство обработки экспериментальных данных // Сборник материалов Всероссийской научно-технической конференции «Математические методы и информационные технологии управления в науке, образовании и правоохранительной сфере». – М., 2017. – С. 325–329.

Щербакова Ю. В. Теория вероятностей и математическая статистика : учебное пособие. – Саратов : Научная книга, 2012.

Приложения

Приложение 1

Таблица критических значений t-критерия Стьюдента

(при уровне значимости 0,10, 0,05 и 0,01)

v	**Уровень значимости**		
Степени свободы	**0,1**	**0,05**	**0,01**
1	6,3138	12,706	63,657
2	2,92	4,3027	9,9248
3	2,3534	3,1825	5,8409
4	2,1318	2,7764	4,6041
5	2,015	2,5706	4,0321
6	1,9432	2,4469	3,7074
7	1,8946	2,3646	3,4995
8	1,8595	2,306	3,3554
9	1,8331	2,2622	3,2498
10	1,8125	2,2281	3,1693
11	1,7959	2,201	3,1058
12	1,7823	2,1788	3,0545
13	1,7709	2,1604	3,0123
14	1,7613	2,1448	2,9768
15	1,753	2,1315	2,9467
16	1,7459	2,1199	2,9208
17	1,7396	2,1098	2,8982
18	1,7341	2,1009	2,8784
19	1,7291	2,093	2,8609
20	1,7247	2,086	2,8453
21	1,7207	2,0796	2,8314
22	1,7171	2,0739	2,8188
23	1,7139	2,0687	2,8073
24	1,7109	2,0639	2,7969
25	1,7081	2,0595	2,7874
26	1,7056	2,0555	2,7757
27	1,7033	2,0518	2,7707
28	1,7011	2,0484	2,7633
29	1,6991	2,0452	2,7564
30	1,6973	2,0423	2,75
40	1,6839	2,0211	2,7045
60	1,6707	2,0003	2,6603
120	1,6577	1,9799	2,6174
200	1,6525	1,9719	2,6006

Приложение 2

Таблица критических значений F-критерия Фишера

Стандартные значения критерия Фишера, используемые для оценки достоверности различий между двумя выборками

Степени свободы	**Уровень значимости**			**Степени свободы**	**Уровень значимости**		
v	**0,05**	**0,1**	**0,01**	***v***	**0,05**	**0,1**	**0,01**
3	10,13	34,12	167,5	28	4,2	7,64	13,5
4	7,71	21,2	74,1	29	4,18	7,6	13,4
5	6,61	16,26	47	30	4,17	7,56	13,3
6	5,99	13,74	35,5	32	4,15	7,5	13,2
7	5,59	12,25	29,2	34	4,13	7,44	13,1
8	5,32	11,26	25,4	36	4,11	7,39	13
9	5,12	10,56	22,9	38	4,1	7,35	12,9
10	4,96	10,04	21	40	4,08	7,31	12,8
11	4,84	9,65	19,7	42	4,07	7,27	12,7
12	4,75	9,33	18,6	44	4,06	7,24	12,5
13	4,67	9,07	17,8	46	4,05	7,21	12,4
14	4,6	8,86	17,1	48	4,04	7,19	12,3
15	4,54	8,68	16,6	50	4,03	7,17	12,2
16	4,49	8,53	16,1	55	4,02	7,12	12,1
17	4,45	8,4	15,7	60	4	7,08	12
18	4,41	8,28	15,4	65	3,99	7,04	11,9
19	4,38	8,18	15,1	70	3,98	7,01	11,6
20	4,35	8,1	14,8	80	3,96	6,96	11,5
21	4,32	8,02	14,6	100	3,94	6,9	11,4
22	4,3	7,94	14,4	125	3,92	6,84	11,3
23	4,28	7,88	14,2	150	3,91	6,81	11,2
24	4,26	7,82	14	200	3,89	6,76	11
25	4,24	7,77	13,9	400	3,86	6,7	10,9
26	4,22	7,72	13,7	1000	3,85	6,66	10,8
27	4,21	7,68	13,6				

Приложение 3

Таблица критических значений G-критерия знаков

(для уровней статистической значимости $\rho \leq 0{,}05$ и $\rho \leq 0{,}01$, по Д. Б. Оуэну)

Преобладание «типичного» сдвига является достоверным, если $G'_{эмп}$ ниже или равно $G'_{0,05}$, и тем более достоверным, если $G'_{эмп}$ ниже или равен $G'_{0,01}$.

n	Уровни стат. значимости (ρ)		n	Уровни стат. значимости (ρ)		n	Уровни стат. значимости (ρ)	
	0,05	**0,01**		**0,05**	**0,01**		**0,05**	**0,01**
5	0	–	27	8	7	49	18	15
6	0	–	28	8	7	50	18	16
7	0	0	29	9	7	52	19	17
8	1	0	30	10	8	54	20	18
9	1	0	31	10	8	56	21	18
10	1	0	32	10	8	58	22	19
11	2	1	33	11	9	60	23	20
12	2	1	34	11	9	62	24	21
13	3	1	35	12	10	64	24	22
14	3	2	36	12	10	66	25	23
15	3	2	37	13	10	68	26	23
16	4	2	38	13	11	70	27	24
17	4	3	39	13	11	72	28	25
18	5	3	40	14	12	74	29	26
19	5	4	41	14	12	76	30	27
20	5	4	42	15	13	78	31	28
21	6	4	43	15	13	80	32	29
22	6	5	44	16	13	82	33	30
23	7	5	45	16	14	84	33	30
24	7	5	46	16	14	86	34	31
25	7	6	47	17	15	88	35	32
26	8	6	48	17	15	90	36	33

Приложение 4

Таблица критических значений Т-критерия Вилкоксона

($\rho \leq 0{,}05$ и $\rho \leq 0{,}01$ по Wilcoxon F.)

«Типичный» сдвиг является достоверно преобладающим по интенсивности, если $T'_{эмп}$ ниже или равно $T'_{0{,}05}$, и тем более достоверно преобладающим, если $T'_{эмп}$ ниже или равен $T'_{0{,}01}$.

n	Уровни стат. значимости (ρ)		n	Уровни стат. значимости (ρ)	
	0,05	**0,01**		**0,05**	**0,01**
5	0	–	28	130	101
6	2	–	29	140	110
7	3	0	30	151	120
8	5	1	31	163	130
9	8	3	32	175	140
10	10	5	33	187	151
11	13	7	34	200	162
12	17	9	35	213	173
13	21	12	36	227	185
14	25	15	37	241	198
15	30	19	38	256	211
16	35	23	39	271	224
17	41	27	40	286	238
18	47	32	41	302	252
19	53	37	42	319	266
20	60	43	43	336	281
21	67	49	44	353	296
22	75	55	45	371	312
23	83	62	46	389	328
24	91	69	47	407	345
25	100	76	48	426	362
26	110	84	49	446	379
27	119	92	50	466	397

Приложение 5

Таблица критических значений критерия x^2_r Фридмана

(для количества условий c = 4, 2 < n < 4)

Различия между условиями можно считать достоверными на указанном в таблице уровне значимости, если $x^2_{r(эмп)}$ достигает соответствующего критического значения или превышает его (по Greene J., D'Ohvera M.).

	n = 2		n = 3		n = 4		
x^2_r	P	x^2_r	P	x^2_r	P	x^2_r	P
0	1	0	1	0	1	5,7	0,141
0,6	0,958	0,6	0,958	0,3	0,992	6	0,105
1,2	0,834	1	0,91	0,6	0,928	6,3	0,094
1,8	0,792	1,8	0,727	0,9	0,9	6,6	0,077
2,4	0,625	2,2	0,608	1,2	0,8	6,9	0,068
3	0,542	2,6	0,524	1,5	0,754	7,2	0,054
3,6	0,458	3,4	0,446	1,8	0,677	7,5	0,052
4,2	0,375	3,8	0,342	2,1	0,649	7,8	0,036
4,8	0,208	4,2	0,3	2,4	0,524	8,1	0,033
5,4	0,167	5	0,207	2,7	0,508	8,4	0,019
6	0,042	5,4	0,175	3	0,432	8,7	0,014
		5,8	0,148	3,3	0,389	9,3	0,012
		6,6	0,075	3,6	0,355	9,6	0,0069
		7	0,054	3,9	0,324	9,9	0,0062
		7,4	0,033	4,5	0,242	10,2	0,0027
		8,2	0,017	4,8	0,2	10,8	0,0016
		9	0,0017	5,1	0,19	11,1	0,00094
				5,4	0,158	12	0,000072

Таблица критических значений U-критерия Манна-Уитни

(для уровней статистической значимости $\rho \leq 0{,}05$ и $\rho \leq 0{,}01$)

Различия можно считать достоверными, если $U_{эмп}$ ниже или равен $U_{0,05}$ и тем более достоверными, если $U_{эмп}$ ниже или равен критическому значению $U_{0,01}$.

n_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
n_2	$\rho = 0{,}05$																		
3	–	0																	
4	–	0	1																
5	0	1	2	4															
6	0	2	3	5	7														
7	0	2	4	6	8	11													
8	1	3	5	8	10	13	15												
9	1	4	6	9	12	15	18	21											
10	1	4	7	11	14	17	20	24	27										
11	1	5	8	12	16	19	23	27	31	34									
12	2	5	9	13	17	21	26	30	34	38	42								
13	2	6	10	15	19	24	28	33	37	42	47	51							
14	3	7	11	16	21	26	31	36	41	46	51	56	61						
15	3	7	12	18	23	28	33	39	44	50	55	61	66	72					
16	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83				
17	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96			
18	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109		
19	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	
20	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138

ρ = 0,01																			
5	–	–	0	1															
6	–	–	1	2	3														
7	–	0	1	3	4	6													
8	–	0	2	4	6	7	9												
9	–	1	3	5	7	9	11	14											
10	–	1	3	6	8	11	13	16	19										
11	–	1	4	7	9	12	15	18	22	25									
12	–	2	5	8	11	14	17	21	24	28	31								
13	0	2	5	9	12	16	20	23	27	31	35	39							
14	0	2	6	10	13	17	22	26	30	34	38	43	47						
15	0	3	7	11	15	19	24	28	33	37	42	47	51	56					
16	0	3	7	12	16	21	26	31	36	41	46	51	56	61	66				
17	0	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77			
18	0	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88		
19	1	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	
20	1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

Окончание прил. 6

n_1	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
n_2	$\rho = 0{,}05$																	
21	19	26	34	41	49	57	65	73	81	89	97	105	113	121	130	138	146	154
22	20	28	36	44	52	60	69	77	85	94	102	111	119	128	136	145	154	162
23	21	29	37	46	55	63	72	81	90	99	107	116	125	134	143	152	161	170
24	22	31	39	48	57	66	75	85	94	103	113	122	131	141	150	160	169	179
25	23	32	41	50	60	69	79	89	98	108	118	128	137	147	157	167	177	187
26	24	33	43	53	62	72	82	93	103	113	123	133	143	154	164	174	185	195
27	25	35	45	55	65	75	86	96	107	118	128	139	150	160	171	182	193	203
28	26	36	47	57	68	79	89	100	111	122	133	144	156	167	178	189	200	212
29	27	38	48	59	70	82	93	104	116	127	139	150	162	173	185	196	208	220
30	28	39	50	62	73	85	96	108	120	132	144	156	168	180	192	204	216	228
31	29	41	52	64	76	88	100	112	124	137	149	161	174	186	199	211	224	236
32	30	42	54	66	78	91	103	116	129	141	154	167	180	193	206	219	232	245
33	31	43	56	68	81	94	107	120	133	146	159	173	186	199	213	226	239	253
34	32	45	58	71	84	97	110	124	137	151	164	178	192	206	219	233	247	261
35	33	46	59	73	86	100	114	128	142	156	170	184	198	212	226	241	255	269
36	35	48	61	75	89	103	117	132	146	160	175	189	204	219	233	248	263	278
37	36	49	63	77	92	106	121	135	150	165	180	195	210	225	240	255	271	286
38	37	51	65	79	94	109	124	139	155	170	185	201	216	232	247	263	278	294
39	38	52	67	82	97	112	128	143	159	175	190	206	222	238	254	270	286	302
40	39	53	69	84	100	115	131	147	163	179	196	212	228	245	261	278	294	311
$\rho = 0{,}01$																		

Таблица критических значений Q-критерия Розенбаума

(для уровней статистической значимости $\rho \leq 0{,}05$ и $\rho \leq 0{,}01$ по Е. В. Гублеру, А. А. Генкину)

Различия можно считать достоверными, если $Q_{эмп} \geq$ критического значения $Q_{0,05}$ и тем более достоверными, если $Q_{эмп} \geq$ критического значения $Q_{0,01}$.

n	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
	ρ=0,05															
11	6															
12	6	6														
13	6	6	6													
14	7	7	6	6												
15	7	7	6	6	6											
16	8	7	7	7	6	6										
17	7	7	7	7	7	7	7									
18	7	7	7	7	7	7	7	7								
19	7	7	7	7	7	7	7	7	7							
20	7	7	7	7	7	7	7	7	7	7						
21	8	7	7	7	7	7	7	7	7	7	7					
22	8	7	7	7	7	7	7	7	7	7	7	7				
23	8	8	7	7	7	7	7	7	7	7	7	7	7			
24	8	8	8	8	8	8	8	8	8	8	7	7	7	7		
25	8	8	8	8	8	8	8	8	8	8	8	7	7	7	7	
26	8	8	8	8	8	8	8	8	8	8	8	7	7	7	7	7

	ρ=0,01																
11	9																
12	9	9															
13	9	9	9														
14	9	9	9		9												
15	9	9	9		9	9											
16	9	9	9		9	9	9										
17	10	9	9		9	9	9	9									
18	10	10	9		9	9	9	9	9								
19	10	10	10		9	9	9	9	9	9							
20	10	10	10		10	9	9	9	9	9	9						
21	11	10	10		10	9	9	9	9	9	9	9					
22	11	11	10		10	10	9	9	9	9	9	9	9				
23	11	11	10		10	10	10	9	9	9	9	9	9	9			
24	12	11	11		10	10	10	10	9	9	9	9	9	9	9		
25	12	11	11		10	10	10	10	10	9	9	9	9	9	9	9	
26	12	12	11		11	10	10	10	10	10	9	9	9	9	9	9	9

Окончание прил. 7

Учебное пособие

Слесарева Екатерина Александровна,
кандидат психологических наук

Математические методы в психологии

Редактор *Лосева О. С.*
Корректор *Степанова А. А.*
Компьютерная верстка *Лосевой О. С.*

Московский университет МВД России имени В.Я. Кикотя
117437, г. Москва, ул. Академика Волгина, д. 12

Подписано в печать 17.04.2019 г.	Формат 60×84 1/16	Тираж 56 экз.
Заказ № 1759	Цена договорная	Объем 3,82 уч.-изд. л. 6,33 усл. печ. л.