МВД России
Санкт-Петербургский университет

М. Г. Баринова

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ПСИХОЛОГИИ

Учебное пособие

Санкт-Петербург
2022

План-график 2022, поз. 19

УДК 159.9; 519.2
ББК 88.4; 22.17
Б 24

Баринова М. Г.

Б 24 Математически методы в психологии: учебное пособие. — Санкт-Петербург: Изд-во СПб ун-та МВД России, 2022. — 144 с.

ISBN 978-5-91837-563-1

Учебное пособие «Математические методы в психологии» соответствует ФГОС ВО по специальности 37.05.02 Психология служебной деятельности. В учебном пособии изложены основы математических методов для психологических исследований. Рассмотрены основные статистические критерии, использующиеся в научных исследованиях в области психологии. Описаны основные виды анализа данных психологических исследований.

Предназначено для курсантов и слушателей, обучающихся по специальности 37.05.02 Психология служебной деятельности, а также для научно-педагогических работников образовательных организаций системы МВД России.

УДК 159.9; 519.2
ББК 88.4; 22.17

Рецензенты:
Борисова С. Е., кандидат психологических наук, доцент
(Орловский институт МВД России им. В. В. Лукьянова);
Евсеева И. Г., кандидат психологических наук, доцент
(Московский университет МВД России им. В. Я. Кикотя)

ISBN 978-5-91837-563-1

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ 4
Глава 1. ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКИХ И СТАТИСТИЧЕСКИХ МЕТОДОВ В ПСИХОЛОГИИ 5
1.1. Роль и место математико-статистических методов в психологии 5
1.2. Основные направления применения математики в психологии 7
1.3. Границы применения математических методов в психологии 9
1.4. Основные понятия математической статистики 11
Глава 2. ОСНОВНЫЕ ПОНЯТИЯ ВЫБОРОЧНОГО МЕТОДА 23
2.1. Вероятность как математическая основа 24
2.2. Генеральная совокупность и выборка 28
2.3. Случайность и случайный выбор 31
2.4. Способы визуализации данных 44
Глава 3. СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ 49
3.1. Статистическая значимость 49
3.2. Статистические гипотезы 52
3.3. Зависимые и независимые выборки 62
3.4. Степени свободы 63
Глава 4. СРАВНЕНИЯ ВЫБОРОК. ВИДЫ СТАТИСТИЧЕСКИХ КРИТЕРИЕВ 66
4.1. Шкалы измерения 66
4.2. Параметрические и непараметрические критерии 71
4.3.Сравнительный анализ и его виды 74
Глава 5. ХАРАКТЕРИСТИКИ ВЗАИМОСВЯЗИ ПРИЗНАКОВ 86
5.1. Корреляционный анализ 86
5.2. Регрессионный анализ 93
5.3. Дисперсионный анализ 101
Глава 6. ВЫЯВЛЕНИЕ РАЗЛИЧИЙ МЕЖДУ ПРИЗНАКАМИ 108
6.1. Методы классификации 108
6.2. Кластерный анализ, его виды 111
6.3. Иерархические и интегративные методы кластерного анализа 114
Глава 7. МЕТОДЫ МНОГОМЕРНОГО АНАЛИЗА ДАННЫХ 118
7.1. Множественный регрессионный анализ, его виды 119
7.2. Множественный дискриминантный анализ 121
7.3. Факторный анализ 124
Глава 8. МАТЕМАТИЧЕСКИЕ ПАКЕТЫ СТАТИСТИЧЕСКОГО АНАЛИЗА ДАННЫХ 132
8.1. Компьютерные системы для анализа данных 132
8.2. Виды статистических пакетов 134
8.3. Общая информация о пакете SPSS и его структуре 136
ЗАКЛЮЧЕНИЕ 141
СПИСОК РЕКОМЕНДУЕМЫХ НОРМАТИВНЫХ ПРАВОВЫХ АКТОВ И ЛИТЕРАТУРЫ 142

ВВЕДЕНИЕ

Освоение учебной дисциплины «Математические методы в психологии» является важным этапом становления психолога органов внутренних дел. В образовательных организациях МВД России данная дисциплина обеспечивает знаниями в области проведения научных исследований и обработки полученных эмпирических данных.

Психолог должен владеть методами сбора и обработки эмпирического материала, данные методы позволяют планировать эксперимент и прогнозировать его результаты. Одним из методов психологического исследования является наблюдение, но это — субъективные метод. Математико-статистические методы позволяют стандартизировать экспериментальные данные, что позволяет более-менее однозначно их трактовать.

Применение математических методов в психологии позволяет сравнивать и описывать количественные и качественные характеристики психологических особенностей, моделировать психические и психологические процессы и состояния, делать аргументированные выводы и прогнозы.

Дисциплина «Математические методы в психологии» позволяет систематизировать знания об обработке эмпирического материала, аргументированно доказывать результаты и выводы, полученные при проведении психологического исследования. Кроме того, она имеет свой методологический и понятийный аппарат. В процессе освоения дисциплины, реализующей доказательный подход к исследованию, рассматриваются основные способы сравнения, анализа и обработки данных, в том числе с помощью современных статистических программ.

Данное пособие ставит целью закрепление и углубление академических знаний, приобретенных на лекциях, а также развитие профессиональных компетенций, связанных с обработкой и анализом данных эмпирического психологического исследования.

Глава 1. ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКИХ И СТАТИСТИЧЕСКИХ МЕТОДОВ В ПСИХОЛОГИИ

Учебные вопросы:

1.1. Роль и место математико-статистических методов в психологии.
1.2. Основные направления применения математики в психологии.
1.3. Границы применения математических методов в психологии.
1.4. Основные понятия математической статистики.

1.1. Роль и место математико-статистических методов в психологии

Первоначально статистикой (*statistics*) называлось изучение государственных дел. Слово статистика происходит от латинского *status*, что означает «состояние дел». В итальянском языке слово *stato* означает «государство». Тех, кто обладал знаниями о состоянии дел и об устройстве государства, сейчас мы называем их политиками, раньше называли *statista* (статиста). Первоначально с помощью статистики рассматривали состояние дел в государстве, то есть можно сказать, что она возникла на основе практических потребностей общественной жизни. Для ведения войн, сбора дани, налогов, необходима информация о количестве жителей, их возраст и пол, размерах земель, количестве скота и животных и т. д., также известно, что регулярно проводилась перепись населения.

Для расширения торговли, международной торговли необходимы знания о составе населения, городах, географическом положении, промышленности, политическом устройстве и т. д. Такие сведения известны с XIII века, вначале в Венецианской республике, затем Италии и Голландии, где издавались справочники, содержащие данную информацию.

Статистические методы формировались на основе регулярных данных о состоянии дел в государстве и на основе честных азартных игроков, пытающихся выиграть, основываясь на математических закономерностях. Таким образом, развивались измерения, упорядочение, табулирование, совершенствовалась перепись населения и возникла описательная статистика. С другой стороны, возникла теория статистического вывода, основывающаяся на теории вероятностей.

Термин «статистика» ввел в научный оборот в 1746 году, немецкий философ Готфрид Ахенваль (1719–1772), написавший книгу о государствоведении, он же стал основателем новой науки.

В XVII веке во Франции небольшие исследования для азартных игроков, которые хотели понять природу удачи, проводили такие математики, как Блез Паскаль (1623–1662) и Пьер Ферма (1601–1665), в дальнейшем на основе этих исследований возникла теория вероятностей. Теория вероятностей исследует закономерности возникновения случайных событий или явлений (например, какова вероятность при подбрасывании монеты выпадения орла).

Иоганн Карл Фридрих Гаусс (1777–1855) — признанный современниками «король математики», в первой половине XIX века занимался математикой, механикой, физикой, астрономией и геодезией. В своей книге по астрономии «Теория движения небесных тел» он предположил возможность определения орбиты планеты в зависимости от трех полных наблюдений, основанных на созданной им теории ошибок измерения. Для минимизации влияния этих ошибок использовал метод наименьших квадратов, который сам же и предложил. Гаусс очень подробно исследовал нормальный закон распределения, который в истории остался гауссианой. Нормальное распределение сегодня также называют распределением Гаусса или Гаусса — Лапласа.

18 апреля 1822 года Иоганн Фридрих Гербарт (1776–1841), немецкий философ, психолог и педагог, в Королевском немецком научном обществе представляет доклад «О возможности и необходимости применять в психологии математику», в котором обосновывает включение математики в психологические исследования для подтверждения их научности. В 1824 году Гербарт выпустил книжное издание «Психология как наука, заново основанная на опыте, метафизике и математике».

Адольф Кетле (1796–1874) считается основателем современной статистики, разработавшим проблемы практического применения теории вероятностей в общественных науках. Он основал статистические общества и статистические учреждения, создал теорию «среднего человека» — человека, обладающего всеми средними, типичными качествами (физическими, нравственными, интеллектуальными), характерными для конкретного общества.

Уильям Сили Госсет (1876–1937) учёный из Великобритании, начинал свои научные исследования в области пивоварения, так как он работал в пивоваренной компании Гиннесса (англ. *Guinness*). Устав запрещал работникам компании публиковать результаты исследований, проведенных в данной компании. Тем не менее компания разрешила опубликовать теоретические проблемы, связанные со статистикой, но без имён служащих компании, чтобы конкуренты не узнали о пользе статистики для пивоварения. В итоге, с 1908 года в журнале «Биометрика» под псевдонимом Student стали публиковаться результаты статистических исследований У. Госсета.

В XX веке математико-статистические исследования становятся необходимой частью научной психологии, формируются психологическая теория измерений, в психологии поведения используется математическое моделирование, дисперсионный анализ применяется в генетической психологии и т. д. Современная психологическая наука доказывает свои теории с помощью математико-статистических методов.

В настоящее время термин «статистика» имеет множество значений:

1. Общественная *наука*, собирающая, упорядочивающая, анализирующая и составляющая факты в виде чисел, количественных характеристик общественных явлений и национальных хозяйств, то есть является орудием познания для установления специфических закономерностей массовых явлений.

2. Форма *практической деятельности* для сбора и обобщения различных явлений и процессов общественной жизни.

3. *Числовые данные*, дающие информацию о различных сторонах общественной жизни, явлений, процессов.

Термин «статистика» часто употребляется как эквивалент «статистических методов». Статистические методы — это методы, используемые для сбора, анализа и интерпретации данных.

Сегодня большинство психологических концепций считаются научно доказанными, если они подтверждены статистическими методами. В психологии нет собственных единиц измерения, миллиметры и секунды — это единицы физических измерений.

Математическая статистика в психологии позволяет обосновать достоверность выводов — она инструмент, являющийся важной частью исследования, вторым шагом, после планирования.

1.2. Основные направления применения математики в психологии

Математические методы в психологии применяются при обработке полученных в исследовании данных, для установления психологических закономерностей. Любое психологическое исследование требует математической обработки полученных данных с помощью специальных программ — MS Excel или специализированных статистических пакетов.

Математико-статистические методы помогают психологу:

1) доказывать правильность и обоснованность используемых методических приемов и методов;
2) обосновывать экспериментальные планы;
3) обобщать данные эксперимента;
4) доказывать наличие зависимостей между экспериментальными данными;
5) выявлять скрытые закономерности;
6) выявлять значимые различия между группами;
7) строить статистические предсказания;
8) избегать содержательных ошибок и др.

Тем не менее, необходимо всегда помнить, что математико-статистические методы — это инструмент, помогающий психологу разобраться в многочисленном и сложном экспериментальном материале. Любое психологическое исследование начинается с постановки задачи, выдвижения гипотез, тщательного планирования, поэтому математика должна способствовать объективному мышлению психолога.

Математическая статистика — это область математики, основанная на теории вероятностей, которая помогает систематизировать и анализировать собранные эмпирические данные. Она включает в себя несколько разделов.

Описательная статистика, или дескриптивная статистика, (англ. *descriptive statistics*) включает в себя статистические измерения, табулирование (создание таблиц), представление и описание полученных эмпирических данных. Данные могут быть представлены как количественно (возраст, вес), так и качественно (пол, национальность). Большое количество данных необходимо уменьшить, при этом не потеряв важную информацию. Человеческий разум не в состоянии одновременно воспринимать и анализировать большое количество полученных в исследовании данных. Описательная статистика позволяет представить массив данных в воспринимаемом человеком виде с помощью описания, обобщения или представления в желаемом виде.

Теория статистического вывода или индуктивная статистика (англ. *inferential statistics, inductive statistics*) — формализированная система, основанная на предсказании свойств генеральной совокупности на основе свойств выборки. Данные выводы делаются на основе методов описательной статистики.

Вариационная статистика занимается изучением изменений случайных величин и вероятностных событий.

Планирование и анализ экспериментов (англ. experimental design techniques) позволяет найти и проверить наличие причинных связей между явлениями. В этом направлении статистических методов большое значение имеет правильно составленный план эксперимента.

Необходимо помнить, что математические методы в психологии являются вспомогательным инструментом и находятся в тесной связи с психодиагностикой, психологической коррекцией и консультацией, а также другими исследовательскими видами деятельности психолога.

В психологии, как науке, использующей математико-статистические методы, применяется:

— математическое моделирование (создание идеальной модели, проведение экспериментов с ней и экстраполирование полученных данных на предполагаемый объект модели);

— статистика (сбор, измерение, мониторинг и анализ количественных данных);

— психометрия (психологическое измерение, связанное с созданием и валидацией психологических измерительных инструментов, разработка и проверка психологических тестов);

— оценивание количественных данных (можно оценить и сравнить, например, скорость реакции, объем памяти, коэффициент интеллекта и т. д.).

Практический психолог проводит психологическое исследование, без которого сложно представить его профессиональную деятельность. При проведении исследования необходимо решать теоретические задачи выбора подхода, психологических технологий, организации сбора и интерпретации данных, обработки данных, применяя стандартизированные пакеты программного обеспечения. Развитие математической компетентности способствует развитию научного мышления.

1.3. Границы применения математических методов в психологии

В математических методах в психологии не может быть четкой инструкции к применению при обработке полученных эмпирических данных, их скорее можно назвать набором идей, которые знающий человек может использовать с пользой для себя, для подтверждения своих гипотез. Из неправильно собранного, малоинформативного, неправильно представленного материала нельзя получить ценную информацию, а иногда даже можно прийти к ложным представлениям и предрассудкам.

Математическая статистика является наукой только для того, кто ей владеет, для думающего исследователя. Известно, что при изменении схемы использования математической технологии, на основании одних и тех же статистических данных, можно получить совершенно разные результаты.

Сегодня любое научное психологическое исследование не обходится без использования математико-статистического аппарата. Исследователь должен определить, какие факторы он будет учитывать, а какие — считать случайными, таким образом, похожие эксперименты, при учете разных факторов могут привести к разным выводам. Поэтому в каждом эксперименте исследователи детально описывают наибольшее количество учитываемых факторов.

Психологические явления сложно описать с помощью точных, математических значений, поэтому при проведении исследования стараются достичь более-менее стабильных внешних условий, но необходимо помнить, что внутренние условия невозможно полностью формализовать.

Психологические исследования строятся на основе количественных данных, полученных на основе психодиагностических данных и формализации внешних условий.

Следует понимать, что универсальных математических методов нет, не может быть единственного четкого алгоритма обработки экспериментальных данных. Необходимо точно знать, какие математические приемы и способы можно применять в каждом конкретном эксперименте, точно знать границы их использования и ограниченность выводов. Непонимание данных ограничений может привести к трудностям как в применении математических методов, так и в интерпретации полученных с помощью их экспериментальных данных.

Предположение «тренинг *A* более эффективен, чем тренинг *B*» подразумевает, что «большая часть клиентов, участвовавших в тренинге *A*, по сравнению с клиентами, участвовавшими в тренинге *B*, показали лучшие результаты по определенным показателям». В научных исследованиях мы не используем неопределенную терминологию «потому что», которую используют маленькие дети, если их спрашивают: «Почему ты это сделал?». Данный ответ неопределенный, и имеет множество оттенков и различных значений.

Применение математических методов в психологических исследованиях имеет своей целью познание количественных и структурных характеристик изучаемых качественных отношений. Основная задача заключается в предсказании различных психологических характеристик и явлений.

Математико-статистические методы позволяют описывать, систематизировать большие массивы данных, исследовать причинные закономерности. Они могут помочь найти ответы на вопросы типа: Каков средний возраст слушателей иностранного факультета? Какой процент из них женат? Сколько из них имеют детей? Составляют ли те иностранные слушатели, кто знает английский язык, большинство, по сравнению с курсантами других факультетов? Зависит ли успеваемость от пола? Подвержены ли слушатели стран Африки влиянию группы в большей степени, чем слушатели стран СНГ?

Овладение математико-статистическими методами требует определенной подготовки. Статистка — это направление прикладной математики, которая также предполагает использование логического мышления и исследовательской интуиции.

1.4. Основные понятия математической статистики

Большое значение в первоначальной математической статистике имеет описательная статистика.

Описательная (дескриптивная) статистика способствует обобщению первичных результатов, собранных в процессе исследования. В основном она занимается группировкой данных, построением графиков распределения частот, выявлением тенденций распределения и оценкой разброса данных.

Для описательного анализа количественных данных используют меры центральной тенденции и меры изменчивости. К мерам центральной тенденции относятся: мода, медиана, среднее значение. К мерам изменчивости — рассеяние, асимметрия, эксцесс, стандартное отклонение, дисперсия.

Разберём все эти понятия более подробно. Начнем с основных, лежащих в основе изучение математической статистики.

Генеральная совокупность — конечная или бесконечная большая совокупность предметов, в отношении которых мы будем распространять полученные выводы, определенная часть изучаемых предметов называется *выборкой*. В статистическом выводе, при превышении объема генеральной совокупности по сравнению с выборкой больше, чем в сто раз, в принципе, все равно, конечная или бесконечная совокупность, разные методы исследования дают в сущности одинаковые результаты. Например, при исследовании агрессивности школьников старших классов генеральной совокупностью будут все старшеклассники (их в России больше миллиона), а выборкой будут непосредственно школьники, у которых изучалась эта агрессивность (30 — 60 — 100 — 500 человек). Выборка из совокупности выделяется специально, особым образом, для исследования свойств совокупности. Таким образом, выборка всегда значительно меньше генеральной совокупности.

При изучении выборки обычно используют методы, ориентированные на бесконечные совокупности, и говорят, что генеральная совокупность фактически бесконечна (если мы сегодня исследуем агрессивность школьников старших классов, то выводы мы будем распространять не только на сегодняшних школьников, но и на тех, которые когда-либо будут учиться в старших классах школы).

Выборка, или статистическая совокупность — это множество объектов, объединенных единой закономерностью и изменяющихся на основе выявленных статистических закономерностей.

Таким образом, генеральной совокупностью можно считать неограниченно большую совокупность измерений, индивидов или явлений, о свойствах которых мы хотим узнать при проведении исследования с помощью методов математической статистики.

Вычисленные для генеральной совокупности значения описательного анализа называются параметрами, а для выборки — статистиками. Некоторые авторы выделяют различия в терминологии для генеральной совокупности и выборки. Статистики обозначают латинскими буквами, а параметры — греческими. Например, символом X обозначают выборочное среднее, а символом M — генеральное среднее; символ D (или d) обозначает выборочную дисперсию, а σ^2 — генеральную, символом μ — генеральное среднее арифметическое, а $\bar{x}$ — выборочную среднюю арифметическую.

Вычисленную по выборке статистику можно выделить как оценку параметра совокупности. Оценка показывает некоторую информацию о параметре, например, выборочное среднее X показывает среднее значение совокупности.

Символом N принято обозначать объем совокупности или общее количество вариант в выборке, количество измерений.

Вариантой называют переменную, полученную в ходе измерений, которая изменяется, то есть варьируется.

Мода (M_o) — определенное значение в выборке, встречающееся наиболее часто, мера положения.

Мода — это значение случайной величины, которое имеет наибольшую вероятность повторения.

Пример: рост слушателей в определенной группе из 9 человек

160, 162, 164, 166, 170, 170, 170, 176, 182.

$M_o = 170$.

Существуют определенные правила выделения моды:

1) если все значения в выборке встречаются одинаково часто, то такая выборка не имеет моды.

Пример: 160, 160, 164, 164, 170, 170, 176, 176

$M_o = 0$, моды нет;

2) когда в выстроенном ряду два соседних значения имеют одинаковую частоту и они больше частоты других значений, то мода определяется как среднее этих двух значений.

Пример: 160, 160, 164, 164, 164, 170, 170, 170, 176.

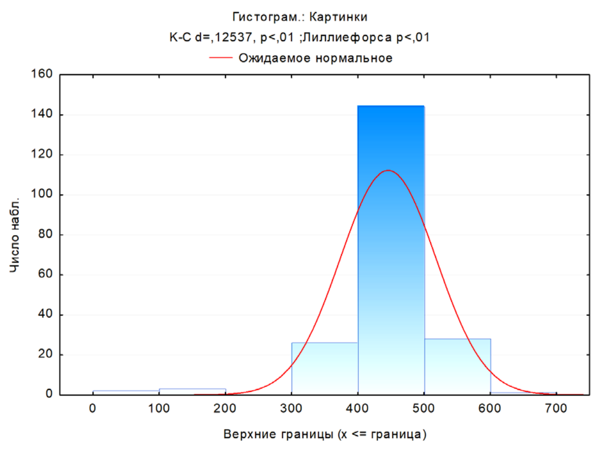
$M_o = 167$;

3) если два не соседних значения встречаются одинаково часто, при этом чаще любого другого значения, то мы можем определить две моды, тогда будем говорить о бимодальной группе значений.

Пример: 160, 164, 164, 164, 166, 168, 170, 170, 170, 176, 176.

$M_o = 164$ и 170.

Моду можно представить на гистограмме для непрерывных случайных величин (о них мы поговорим в следующей теме), среди распределений можно выделить унимодальные (рис. 1) и полимодальные, которые больше одной моды (рис. 2).



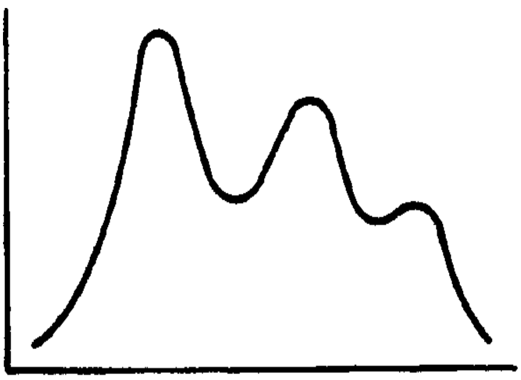


Рис. 1. Унимодальная гистограмма Рис. 2. Полимодальная гистограмма

Медиана (M_e) — мера положения, которое делит упорядоченную выборку пополам, то есть меньше — ровно половина упорядоченной выборки и больше — тоже ровно половина выборки. Медиана делит выборку пополам. Существует формула для вычисления медианы:

для нечетной выборки для четной выборки

$$M_e = \frac{X_{N+1}}{2},\ M_e = \frac{X_{\frac{N}{2}} + X_{\frac{N}{2}+1}}{2}.$$

Если выборка состоит из нечетного количества значений, то M_e будет соответствовать среднему значению упорядоченной выборки.

Упорядоченная выборка — значения числового ряда, выстроенные в порядке возрастания, то есть справа значение всегда больше левого.

Если выборка состоит из четного количества значений, то M_e можно отметить как точку, расположенную между двумя центральными значениями, среднее арифметическое между ними.

Пример: уровень интеллекта в группе
105, 107, 107, 110, 110, 115, 120, 120, 122
$M_e = 110$
105, 107, 107, 110, 115, 120, 120, 122
$M_e = 112{,}5$

Среднее значение является одной из важных и часто употребляемых статистик, мера положения, вычисляемая из набора количественных данных, мера расположения центра данных.

В математической статистике используются разные виды средних значений — среднеарифметическое, среднеквадратическое, среднегармоническое и др. Все средние значения, по определению, больше минимального и меньше максимального значения.

Среднее значение случайной величины в теории вероятностей называется «математическое ожидание». Теория вероятностей является разделом математики, изучает случайные события и величины, их свойства и операции над ними. Вероятность — количественная оценка возможности наступления события (например, завтра будет дождь с вероятностью 80 %, означает, что скорее всего завтра будет дождь). В математической статистике значение вероятности принимает от 0 до 1, где 1 — событие будет, оно достоверное, а 0 — невозможное событие, то есть его не будет.

В математической статистике среднее значение обозначает меру положения, с одной стороны, и математическое ожидание — с другой. С третьей стороны, существует понятие математического ожидания генеральной совокупности. Математическое ожидание означает среднее значение случайной величины и обычно оценивается как среднеарифметическое всех значений. Таким образом, под средним или средним значением мы будем иметь в виду среднее арифметическое полученных значений.

Среднее арифметическое ($X_{ср}$) — это одно из разновидностей среднего значения, также является мерой положения и определяется как величина, характеризующая среднее значение признака при учете общего количества значений, то есть оно равно сумме всех значений, деленной на их количество.

Мы будем обозначать как $\overline{X}$ и вычислять по формулам:

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{N} \quad \text{или} \quad \overline{X} = \frac{1}{N}\sum_{i=1}^{n} X_i,$$

где X_1, X_2….. 1, 2 и т. д. значения,

X_n — последнее значение,

N — количество всех значений,

$\sum_{i=1}^{n}$ — сумма всех значений, начиная с 1 (i=1) и заканчивая последним (n).

Пример: уровень интеллекта в группе
105, 107, 107, 110, 110, 115, 120, 120, 122

$\overline{X}$ = 1/9*(105+107+107+110+110+115+120+120+122) = 1/9*1016 = 112,9

Среднее арифметическое (М) — среднее значение для генеральной совокупности. Любое значение из генеральной совокупности является случайной величиной (X), тогда математическое ожидание равно выборочной средней, М (X) = $\overline{X}$

$$M = \sum_{i=1}^{n} P_i X_i,$$

где P_i — вероятности, X_i — значения.

У среднего можно выделить следующие свойства:

1. Нулевое — при вычитании из X_i среднего, получается величина отклонения, сумма всех отклонений должна быть равна 0:

$$\sum_{i=1}^{n}\left(X_i-\overline{X}\right)=0$$

2. При увеличении или уменьшении каждого значения на постоянное число С среднее увеличится или уменьшится на это число.

3. При умножении или делении каждого значения на постоянное число С, среднее также увеличится или уменьшится в С раз.

Средняя геометрическая G — мера центральной тенденции, мера положения, которое можно определить, как центральное значение в геометрической прогрессии. Средняя геометрическая всегда меньше средней арифметической. Чаще всего средняя геометрическая используется для оценки среднего эффекта тренировки, вычисляется только для положительных значений по формуле:

$$\overline{X} = \sqrt[n]{x_1 * x_2 * x_3 * \ldots * x_n} = \sqrt[n]{\Pi x_i},$$

где П — произведение всех значений.

Пример: требуется рассчитать среднюю интенсивность тренировок при выполнении упражнений на пресс

День тренировки	1	2	3	4	5	6
Кол-во упражнений	20	22	22	26	25	28

$$\overline{X} = \sqrt[6]{20*22*22*26*25*28} = \sqrt[6]{176176000} = 23{,}68$$

Ответ: интенсивность тренировок при выполнении упражнений на пресс составляет 23,68 раз, то есть в среднем испытуемый выполнял 23,68 раз упражнений на пресс в течение 6 дней тренировки.

Средняя гармоническая Н — мера центральной тенденции, мера положения, величина, обратная средней арифметической, если средняя арифметическая это x_i, то средняя гармоническая это $1/x_i$. Показатель средней гармонической Н всегда меньше показателя средней геометрической G. Чаще всего используется как мера положения распределения частоты колебаний или скорости, вычисляется по формуле:

$$H=\frac{N}{\sum\frac{f_i}{x_i}},$$

где N — количество переменных;

f_i — частота i-ой переменной;

x_i — переменная.

Для первоначального математико-статистического анализа данных можно упорядочить значения по мере возрастания, от меньшего к большему, если мы знаем общее количество измерений, то мы можем рассчитать квантили распределения.

Выброс. Среднее значение показывает нам типичного представителя исследуемой выборки, но в выборке, в нашем примере с ростом слушателей (160, 164, 164, 164, 166, 168, 170, 170, 170, 176, 176), могут оказаться очень высокие (196 *см*) или очень низкие (150 *см*). Если в одной группе присутствует очень высокий слушатель, то его показатель значительно изменит среднее значение, и оно уже не будет показывать типичного представителя исследуемой выборки. Показатель, сильно отличающийся от остальных данных, называется выбросом. Определить выбросы можно с помощью гистограмм, диаграмм, а также 3σ (что такое σ описано ниже). Данные, выходящие за пределы ±3σ, считаются выбросами и их можно не учитывать в дальнейших вычислениях.

Квантиль — значение переменной, делящее все распределение на равные части. В математической статистике обычно применяют такие квантили, как квартили, децили и центили. Квантиль — обобщающее понятие. *Квартили (Q_1, Q_2, Q_3)* делят распределение на четыре части, поэтому их три, *децили (D1,…,D9)* на десять, *центили (P1,…,P99)* на сто.

Чаще всего используют квартили:

0,25 — квартиль называется первым или нижним;

0,5 — квартиль называют медианой, или вторым квартилем;

0,75 — квартиль это третий или верхний квартиль.

Среднее квартильное отклонение — мера рассеивания при совпадении среднего значения и медианы, его можно вычислить по формуле:

$$Q = \frac{Q_B - Q_H}{2},$$

где Q_B — верхний квартиль,

Q_H — нижний квартиль.

Также можно определить *полуинтерквартильное* (вероятное, срединное) отклонение, которое определяется как половина интервала рассеивания, в которую входит медиана и половина объема выборки:

$$E = 0{,}5\ (Q_3 — Q_1),$$

где E — полуинтерквартильное отклонение,

Q_3 и Q_1 — третий и первый квартили.

Меры изменчивости (рассеивания, разброса). Мерами изменчивости называют статистические данные, показывающие различия между отдельными значениями. Они дают возможность узнать о степени однородности полученных данных и могут косвенно говорить о надежности выводов.

Размах — одна из простых мер изменчивости, показывает на шкале расстояние, в пределах которого изменяются показатели. Выделяют два типа размаха — включающий и исключающий. Исключающий размах по-другому называют вариационным. Размах чувствителен к выбросам. В некоторых случаях, для получения достоверных результатов, отсекают 25 % самых высоких показателей и 25 % самых низких — это называется межквартильным размахом.

Вариационных размах показателя — количественная мера, показывающая разность между максимальным и минимальным значениями. Например, исключающий размах значений 105, 107, 107, 110, 110, 115, 120, 120, 122 равен 122 – 105 = 17. Значения –4, –2, 0, 2, 4, 6 имеют исключающих размах 6 – (–4) = 10.

Включающий размах можно узнать, вычислив разницу между естественной верхней границей интервала с максимальным значением, и естественной нижней границей интервала с минимальным значением. Например, мы измеряем рост слушателей с точностью до сантиметра: 160, 164, 164, 164, 166, 168, 170, 170, 170, 176, 176 *см*. В действительности минимальный рост может быть в пределах от 159,5 до 160,4, тогда нижняя граница будет 159,5, верхняя же — 176,5 *см*., следовательно, включающий размах равен 176,5 – 159,5 = 17, что на единицу больше, чем исключающий размах (176 – 160 = 16).

Размах распределения можно узнать с помощью формулы:

$$d = X_{max} — X_{min},$$

где X_{max} — наибольшее значение показателя,
X_{min} — наименьшее значение показателя.

Асимметрия (A_s)

Коэффициент асимметрии является мерой асимметрии распределения случайной величины. Асимметрия — количественная мера «скошенности» распределения, высчитывается по формуле:

$$A_s = \frac{\sum d^3}{\sigma^3 \cdot n}$$

где $d = x_i$

σ — стандартное отклонение случайной величины,

n — количество исследуемых показателей.

При значении $A_s = 0$, график считается симметричным, при

$A_s > 0$ — «скошенным» влево, отрицательная асимметрия, при

$A_s < 0$ — «скошенным» вправо, положительная асимметрия (рис. 3).

Эксцесс (E_x) — количественная мера симметричного распределения, показывающая его «горбатость».

Коэффициент эксцесса (E_x) определяется как

$$E_x = \frac{\sum d^4}{\sigma^4 \cdot n} - 3,$$

$$-3 < E_x < \infty.$$

При нормальном распределении $E_x = 0$, $A_s = 0$.

При значении $E_x > 0$ наблюдается положительный эксцесс, при $E_x < 0$ — отрицательный (рис. 3).

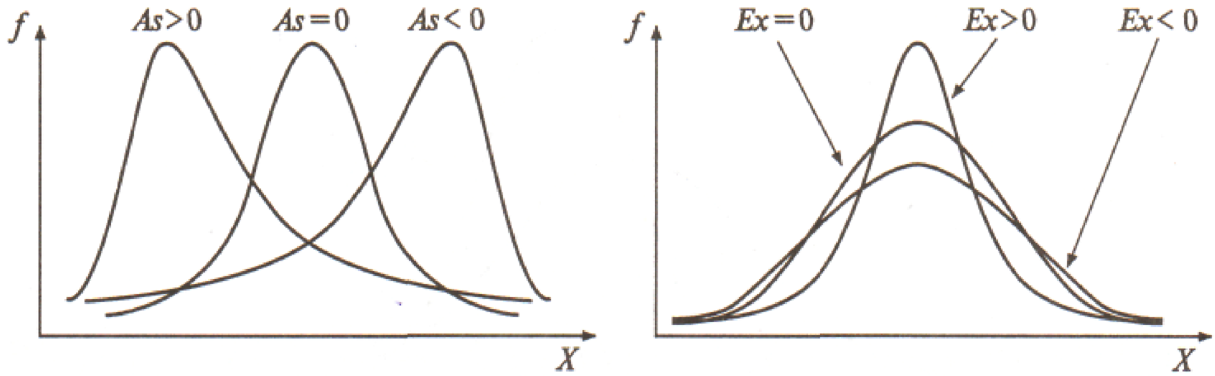


Рис. 3. Графики распределений с разными значениями асимметрии и эксцесса

Показатели асимметрии и эксцесса можно использовать как показатель для проверки соответствия полученного распределения нормальному.

Вариация — различие значений измеряемого показателя. Вариационный размах, как описано выше, показывает разницу между максимальным и минимальным значением показателя. Вариантой называют переменную, результаты измерений которой варьируются, изменяются, она обозначается как x_i и является единицей выборки, результатом одного измерения.

По степени вариации можно говорить об однородности выборки, её типичности.

Дисперсия (D). Дисперсию иногда называют флуктуацией или девиантой, сам термин можно перевести как «рассеяние». Величина дисперсии показывает разброс данных. Дисперсии сложно дать определение, легче её вычислить по формуле:

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

Дисперсию можно определить как показатель, равный среднему значению квадрата отклонений отдельных показателей от средней арифметической, или как разность между квадратами среднего квадратического и среднего арифметического.

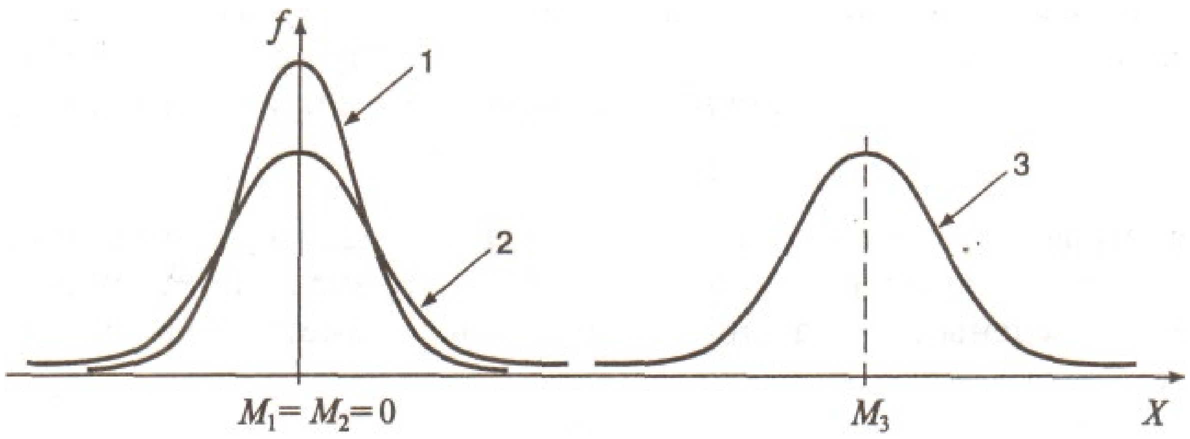


Рис. 4. Графики распределений с разными средними и разными дисперсиями

На рисунке представлены графики распределения частот: с разной дисперсией ($D_1 < D_2$), одинаковой дисперсией ($D_2 = D_3$) и разными средними арифметическими ($M_2 < M_3$).

Пример: рост слушателей: 160, 164, 164, 164, 166, 168, 170, 170, 170, 176, 176 *см*.

1. Найдем среднее значение

$\bar{X}$ = (160+164+164+164+166+168+170+170+170+176+176)*1/11=

$\bar{X}$ = 1848/11–1=168

2. Найдем дисперсию

$D = ((160 - 168)^2 + (164 - 168)^2 + (164 - 168)^2 + (164 - 168)^2 + (166 - 168)^2 + (168 - 168)^2 + (170 - 168)^2 + (170 - 168)^2 + (170 - 168)^2 + (176 - 168)^2 + (176 - 168)^2))/10 = (8^2 + 4^2 + 4^2 + 4^2 + 2^2 + 0 + 2^2 + 2^2 + 2^2 + 8^2 + 8^2)/10 = (64 + 16 + 16 + 16 + 4 + 0 + 4 + 4 + 4 + 64 + 64)/10 = 256/10 = 25{,}6$

$D = 25{,}6$

Стандартное отклонение. Наиболее употребительным термином психологов при использовании математико-статистических методов является стандартное или среднеквадратичное отклонение, мера рассеяния.

Среднеквадратичное отклонение показывает, насколько в среднем каждое значение отклоняется от средней арифметической.

Стандартное отклонение вычисляется как корень из дисперсии:

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

Среднее значение и среднеквадратичное отклонение часто используют для описания выборки, при нормальном распределении большая часть выборки, ~68 %, расположена в пределах одного среднеквадратичного отклонения (± σ).

Коэффициент вариации V. Карл Пирсон предложил коэффициент вариации или вариативности, показатель степени изменчивости данных, который выражается в отношении стандартного отклонения к среднему значению, выраженному в процентах.

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

В данной формуле видно, что коэффициент вариации зависит от среднего (*M*), следовательно, для сравнения выборок необходимо брать выборки с одинаковым началом и средним арифметическим, при нормальном распределении, иначе сравнение приведет к ложным выводам.

Таким образом, к мерам центральной тенденции, или мерам положения, относятся мода, медиана и среднее значение, они учитываются в определенных условиях.

Независимо от условий распределения вероятностей упорядоченных данных можно определить их меры изменчивости. Данные отличаются графически, то есть они имеют разное положение на числовой оси; рассеиванием, то есть размахом и эксцессом; асимметрией рассеивания значений переменной. Также к мерам изменчивости относят дисперсию и стандартное отклонение, которые неустойчивы к выбросам, большая часть выборки при нормальном распределении (~68 %) входит в значения ±σ. Таким образом, зная дисперсию и стандартное отклонение, значения которых можно встретить в научных публикациях, можно представить типичный показатель, а также типичный диапазон изменений.

Выраженность мер изменчивости зависит как от особенностей упорядочивания данных, так и от условий получения данных. Используя количественные характеристики выявленных психологических показателей, психологи получают возможность сравнения этих показателей у разных групп людей или в разных ситуациях. Другими словами, если распределения вероятностей показывают психологические характеристики, то при сравнении распределений мы опосредованно сравниваем эти психологические характеристики.

Вопросы и задания для самоконтроля

1. На основе чего возникла теория вероятностей?
2. Дайте определение распределению Гаусса.
3. Сопоставьте понятия описательной статистики.
4. Дайте определение психометрией.
5. Приведите примеры, где необходимо применение математико-статистических методов.
6. Сравните понятия генеральная совокупность и выборка.
7. Перечислите меры центральной тенденции.
8. Перечислите меры изменчивости.
9. Приведите пример необходимости вычисления среднего арифметического и как оно может обозначатся.
10. Сравните понятия дисперсия и стандартное отклонение.

Глава 2. ОСНОВНЫЕ ПОНЯТИЯ ВЫБОРОЧНОГО МЕТОДА

Учебные вопросы:

2.1. Вероятность как математическая основа.
2.2. Генеральная совокупность и выборка.
2.3. Случайность и случайный выбор.
2.4. Способы визуализации данных.

Выборочный метод математической статистики — это исследование свойств генеральной совокупности на основе исследования свойств выборки. Теория выборочного метода основана на теории выбора из конечной совокупности и теории выбора из бесконечной совокупности.

Конечная и бесконечная совокупность для выборочного метода отличается тем, что для конечной совокупности объекты выборки чаще всего неслучайны (например, те, кто учится на отлично), а для бесконечной обычно изучаются свойства случайных объектов.

В соответствии с теорией вероятностей выборка тогда будет адекватно отражать свойства генеральной совокупности, когда выбор объектов исследования происходит случайно, то есть любой из объектов генеральной совокупности имеет шанс попасть в выборку.

Выборки также делятся на повторные и бесповторные. На практике чаще встречаются бесповторные, когда объект, попавший в выборку, исключается из дальнейшего исследования и повторно не может попасть в данную выборку. Выбор с возвращением, или повторная выборка, обычно бывает в теоретических исследованиях — в практических, особенно психологических, это очень осложнит исследование.

При достаточно большом объеме генеральной совокупности (N) и относительно небольшой выборке (n) различия между повторной и бесповторной выборками $n \ll N$ практически отсутствуют, и даже исчезают при бесконечной генеральной совокупности и конечном объеме выборки.

С помощью выборочного метода можно изучать как количественные свойства генеральной совокупности, так и качественные.

Изучая качественные свойства, необходимо выбрать количество объектов (A), обладающих исследуемым качественным свойством (например, среди всех курсантов мы можем исследовать только девушек, пол является качественным свойством, так как его нельзя определить количественно, то есть мы не можем сказать, насколько выражено у девушки свойство девушки, насколько она ближе к девушке, чем к юноше). Если N мы обозначим всех курсантов, то $A = \mu N/n$, где μ — число девушек в выборке n.

Изучая количественные свойства, необходимо определить среднее значение $M = (X_1+X_2+....+X_n)/N$, которое редко удается определить, поскольку не часто известны данные всей генеральной совокупности, тогда оно заменяется выборочным средним значением $X_{ср} = (X_1+X_2+....+X_n)/n$.

В математической статистике результаты психологических наблюдений принято называть выборкой. Предполагается, что можно выбрать любое количество наблюдений.

Для лучшего понимания основ выборочного метода необходимо рассмотреть понятие вероятности.

2.1. Вероятность как математическая основа

Вероятность — оценка возможности появления определенного события. Существует три варианта развития событий:

— достоверное — обязательно происходит в процессе исследования;

— невозможное — не может произойти в процессе исследования;

— случайное — может произойти, а может не произойти в процессе исследования.

Вероятностью события А называют отношение количества положительных исходов (m) к общему числу событий (n):

$$P(A) = m/n$$

$$0 \le P(A) \le 1,$$

при этом P = 0 — вероятность невозможного события равна 0,

P = 1 — вероятность достоверного события равна 1,

вероятность случайного события находится в интервале от 0 до 1.

Описано несколько определений вероятности.

1. Классическое — вероятность случайного события А — это отношение числа равновероятных событий А к числу всех возможных событий N.

2. Геометрическое — существует геометрическая область G, вероятность попадания в любую точку подмножества G зависит от размера области и не зависит от формы и расположения.

3. Частотное (статистическое) — вероятность — это предел частоты встречаемости события А при неограниченном увеличении количества наблюдений, их однородности и независимости.

Классическое определение вероятности
(ключевое слово — равновероятных)

Исходя из собственного опыта нам известно, что одни события встречаются чаще других. Например, в Санкт-Петербурге часто идет дождь, а солнце зимой появляется редко; простудные заболевания встречаются часто, а переломы костей редко. В данных примерах слова «часто» и «редко» демонстрируют ожидаемую вероятность появления того или иного случайного события. Но на основе известных данных (прошлых событий) мы можем попытаться предсказать вероятность их появления, для чего определим понятие частота.

Частота f_i — количество событий, демонстрирующее, сколько раз встречается показатель x_i , при этом сумма всех частот должна быть равна всей выборке: $\sum f_i = n$.

При проведении исследования определенное событие может произойти, а может и не произойти. С увеличением количества исследований отношение определенного события к общему числу всех событий вероятно будет меняться. Например, слушатель может оказаться экстравертом, а может и нет. Частота данного события будет уменьшаться с увеличением количества обследованных и будет зависеть от общего числа испытуемых (при исследовании 10 слушателей 4 могут оказаться экстравертами, при исследовании 20 — 8 экстраверты, при исследовании 100 человек — 40 экстраверты; экстравертов становится меньше, а испытуемых больше). В то же время видно, что чем больше наблюдений, тем больше мы выявляем экстравертов (40 человек больше, чем 4).

Частоту появления случайного события можно использовать для сравнения в выборках только при условии равного количества испытуемых в данных выборках. При условии неравнозначности выборок мы пользуемся понятием частость — мера возможности появления случайных событий.

Частость, или относительная частота, выраженная в долях или процентах по отношению к общему количеству случаев. Частость — доля каждой частоты f_i по отношению ко всем случаям n. Сумма всех частостей равна 1 или 100 %. Частость экстраверсии гораздо меньше зависит от общего объема выборки. Если взять пример выше, то частость экстраверсии равна 0,4 или 40 %, независимо от того, столько человек было обследовано 10, 20 или 100, частость при этом не изменялась.

Если кидать игральный кубик и смотреть, как часто будет выпадать единица, то можно нарисовать график, показывающий зависимость выпадения единицы от общего количества выбросов кубика, где по оси ординат — величина рассеивания, по оси абсцисс — количество исходов. Если мы попробуем определить частость выпадения единицы для 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, то увидим, что частость будет меняться, точнее, будет уменьшаться при увеличении количества выбросов кубика. Если из 10 исходов в 3 будут выпадать единицы, то из 20 исходов в 5 будут выпадать единицы, из 30 — в 7 и т. д.

$f_1 = 3/10 = 0{,}3$

$f_2 = 5/20 = 0{,}25$

$f_3 = 7/30 = 0{,}23$

$f_4 = 9/40 = 0{,}225$

$f_5 = 10/50 = 0{,}2$

$f_6 = 12/60 = 0{,}2$

$f_7 = 13/70 = 0{,}186$

$f_8 = 14/80 = 0{,}175$

$f_9 = 16/90 = 0{,}178$

$f_{10} = 19/100 = 0{,}19$

Таким образом, частость появления какого-либо события зависит от количества испытаний. Чем чаще испытуемый будет выполнять тест на запоминание 20 слов, тем больше слов он каждый раз будет запоминать, так как каждый раз при запоминании он тренирует свою память. Эти два примера показывают, что сущность проведения исследования может влиять на результат: от нас никак не зависит выпадение единицы на кубике, но множество событий может влиять на память испытуемого.

При изучении частости появления события большое значение имеет количество испытаний, чем их больше, тем больше вероятность того, что частость показывает объективную вероятность исследуемого события и не зависит от количества испытаний. Однако в некоторых случаях такая зависимость существует. Например, когда мы бросаем кубик 1000 раз и из них 200 раз выпадает единица, значит, когда бросаем 100 раз, единица выпадает 20 раз, а когда 10, то 2 раза и уже не зависит от количества испытаний. Количество испытаний имеет значение при установлении этой закономерности.

Вероятность — количественная оценка возможности появления определенного события; значение, к которому стремится частость события при неограниченном числе испытаний. Практически невозможно бесконечно проводить испытания, поэтому эмпирически узнать вероятность события невозможно. Теоретически мы можем определить, например, вероятность выпадения решки при бросании монеты ½, так как монета имеет всего две грани; вероятность выпадения 1 на шестигранном кубике равна 1/6, но все же правильнее будет определение вероятности экспериментальным путем.

Как частость, так и вероятность могут принимать любые значения от 0 до 1. Вероятность определяет среднюю частость.

Геометрическое определение вероятности

Если предположить, что попадание в любую точку области G (рис. 5) равновозможно, то попадания случайной точки в область g можно вычислить как отношение площадей:

$$P(g) = S(g) / S(G)$$

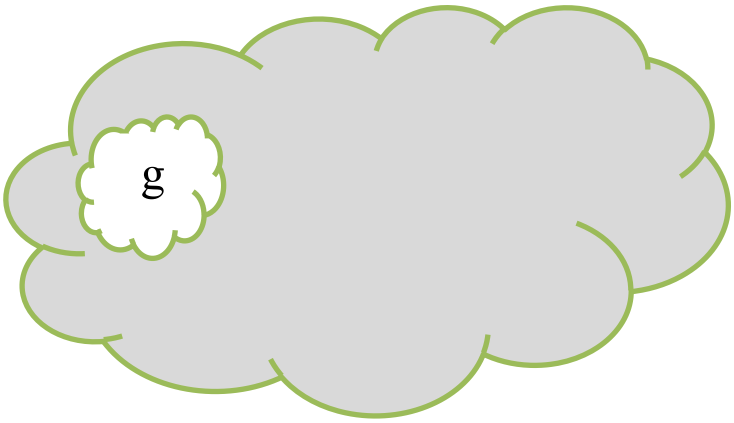


Рис. 5. Вероятность попадания в область g

Геометрическое определение вероятности наглядно демонстрирует вероятностное соотношение между событиями на основе длин, площадей, объемов, и в некоторых случаях позволяет наглядно представить какую-либо задачу.

Вероятность — вероятностная мера попадания точки (события) в определенных условиях (G) в определенную область (g), которая может произойти, а может и не произойти.

Вероятностная мера — мера вероятности случайного события, ее количественная оценка.

Статистическое определение вероятности

Статистическое определение основано на экспериментальных данных, на данных наблюдений. По данным эксперимента вычисляют частность события.

Вероятностью события А называют число, к которому стремится относительная частота события А при увеличении количества испытаний:

$$P(A) = \lim_{n\to\infty} W(A) = \lim_{n\to\infty} m_A/n,$$

где m — число появлений события А,

n — общее число событий

Статистической вероятностью называется число, в районе которого наиболее часто встречается относительная частота события при неограниченно большом количестве испытаний.

2.2. Генеральная совокупность и выборка

Совокупность похожих объектов исследуют относительно определенного признака, количественно или качественно. Например, у группы курсантов изучают акцентуации характера с помощью опросника Шмишека. В результате мы получим количественные оценки каждого курсанта по каждой шкале опросника, также можно качественно проанализировать полученные результаты и создать психологический портрет данной группы курсантов. Сплошное исследование всех курсантов (из разных городов, курсов) практически не проводится.

Если генеральная совокупность состоит из большого количества объектов, то проводят не сплошное исследование, а выборочное. Выборочную совокупность называют выборкой.

Генеральная совокупность (от лат. *generis* – общий, родовой) — совокупность всех возможных объектов, в отношении которых планируется исследование и относительно которых будут сделаны выводы. Например, курсанты — генеральная совокупность, то есть это все обучающиеся в военных вузах и вузах МВД России, на всех курсах, во всех учебных заведениях, всех городов России. Выборкой будет называться определенная группа курсантов, участвующая в исследовании.

Генеральная совокупность теоретически состоит из конечного числа объектов, но если она очень большая, то для упрощения математико-статистических расчетов считается, что она бесконечна и используются формулы для бесконечного множества.

Выборочная совокупность, или выборка — часть элементов генеральной совокупности, отобранная определенным образом. Выборка обладает определенными характеристиками:

— качественная (как и что мы выбираем);

— количественная (её объем).

Объем выборки определяется несколькими факторами:

— число анализируемых групп;

— точность результатов и ценность полученной информации;

— стоимость выборки, затраты исследования (время, место, исследователи) должны быть сопоставимы с результатами, при этом необходимо помнить, что чем больше выборка, тем более достоверны результаты;

— разброс значений (если все одинаковые, то нет смысла в большой выборке).

Выборка может быть повторной, когда исследуемый объект после исследования возвращается обратно в генеральную совокупность и имеет шанс быть выбранным повторно, и бесповторной, когда исследуемый объект исследуется только один раз.

Для того чтобы выводы, полученные при исследовании выборки, можно было распространить на генеральную совокупность, выборка должна быть *репрезентативной*, то есть обладать всеми значимыми свойствами генеральной совокупности. Например, при исследовании курсантов для получения репрезентативной выборки надо взять курсантов военных вузов, курсантов вузов МВД, желательно взять из 3–4 разных городов, разных видов войск, обязательно взять курсантов с разных курсов и разного пола, желательно также взять курсантов, обучающихся на «отлично», «хорошо» и «удовлетворительно», курсантов, проживающих дома, в общежитии, в казарме.

Существуют различные варианты отбора выборки:

1. Выбор производится из всей генеральной совокупности:

— простой случайный (объекты выбираются путем жеребьевки) повторный отбор;

— простой случайный бесповторный отбор.

2. Генеральная совокупность делится на части и из каждой части выбираются объекты для исследования:

— *простой случайный отбор* — элементы исследования отбираются случайным образом, часто с помощью жеребьевки;

— *типический (статифицированный)* — объекты отбираются из каждой типической части генеральной совокупности;

— *механический* — всю генеральную совокупность механически разделяют на группы (в зависимости от количества предполагаемых объектов в выборке групп должно быть столько, сколько элементов должна содержать выборка);

— *серийный* — объекты отбираются из генеральной совокупности не единично, а сериями, которые попадают в выборку;

— *комбинированный* — комбинирование способов отбора и видов выборки;

— *многоступенчатый*.

Выборки бывают зависимые и независимые (рис. 6).

Зависимая выборка — чаще всего это повторные исследования одной и той же группы (лонгитюдные исследования), когда одного человека в одной выборке можно сравнить с человеком в другой выборке. Зависимыми также называют сравнение двух и более переменных в одной группе, а также выборки близнецов, сиблингов, мужей и их жен (при изучении, например, семейных ценностей) и другие. В зависимых выборках один случай в одной выборке соответствует только одному случаю в другой выборке — гомоморфная пара.

При независимых выборках сравнивают одну группу с другой, используя, например, средние значения.

Зависимые выборки также называют «связанные», а независимые — «несвязанные».

В независимых выборках можно менять респондентов. Например, если кто-то из группы заболел, можно взять несколько человек из другой группы. В зависимых выборках такое невозможно — если кто-то заболел в одной группе, то данный респондент со своей парой просто выбывают из исследования, и надо искать другую пару респондентов.

Зависимая выборка

	до	после	
Маша	10	7	Маша
Катя	7	5	Катя
Наташа	9	6	Наташа
Вика	10	8	Вика
Света	8	6	Света
Лиза	7	5	Лиза
Алина	10	7	Алина
Полина	9	7	Полина

Независимая выборка

	до	после	
Маша	10	7	Маша
Катя	7	5	Катя
Наташа	9	6	Наташа
Вика	10	8	Вика
Света	8	6	Света
Лиза	7	5	Лиза
Алина	10	7	Алина
Полина	9	7	Полина
Ср.знач	8,75	6,375	

Рис. 6. Сравнение зависимой и независимой выборок

Психологи в своих экспериментах обычно используют следующие стратегии отбора в экспериментальные группы (выборки):

1. Рандомизация или случайный отбор.

Для создания простых случайных выборок, самый простой выбор — по типу жребия. В этом случае каждый элемент генеральной совокупности имеет шанс попасть в выборку.

2. Попарный отбор.

Выборки должны быть эквивалентными по значимым характеристикам, чем больше будет совпадений, тем меньше влияния неисследуемых параметров. Чаще всего используется при создании экспериментальной и контрольной групп, оптимальный вариант — близнецы.

3. Стратометрический[1] отбор.

Генеральная совокупность делится на группы с определёнными характеристиками, например, пол, возраст, звание, курс, факультет, страна, национальность, средний бал и т. д., затем из каждой группы выбираются испытуемые.

4. Приближённое моделирование.

Формирование модели, описывающей характеристики определенной группы. Выводы, сделанные на ограниченной выборке, распространяются на широкую популяцию. Например, при изучении психологических характеристик курсантов второго курса экономического факультета Санкт-Петербургского университета МВД России полученные выводы могут быть распространены на всех курсантов Санкт-Петербургского университета МВД России. Данную стратегию можно применять очень ограниченно, поскольку такая выборка может очень сильно изменяться в зависимости от набора курсантов и при отсутствии уверенности в том, что каждый набор будет полностью совпадать по психологическим характеристикам курсантов.

2.3. Случайность и случайный выбор

Случайными обычно называют события или характеристики, причин появления или изменения которых мы не знаем.

Выводы, сделанные на основе выборки, дают нам некое представление о характеристиках совокупности. Как уже указывалось выше, существует несколько стратегий отбора выборок. Чаще всего используется простой случайный выбор, поскольку остальные стратегии требуют больше времени и усилий. В элементарной статистике обычно говорят

[1] Страты — слои.

о простой случайной выборке, поэтому в дальнейшем под выборкой мы будем подразумевать именно простую случайную выборку.

Выборка должна быть репрезентативной, то есть ее характеристики должны соответствовать характеристикам генеральной совокупности. Надо понимать, что все характеристики генеральной совокупности невозможно учесть в выборке, поэтому необходимо выбрать те, которые могут повлиять на процесс или результат исследования. Самыми простыми характеристиками, которые необходимо учитывать при формировании выборки, являются пол и возраст, остальные больше зависят от темы исследования и возможного влияния их на результат. При случайной выборке невозможно предсказать ее репрезентативность, но чем больше объем выборки, тем более вероятна ее репрезентативность.

Например, мы можем отобрать в выборку слушателей первого курса иностранного факультета для выявления отношения к семейным ценностям. Есть вероятность, что все слушатели будут представлены одной страной, следовательно, полученные выводы мы можем распространить только на эту страну, но мы хотели исследовать всех представителей иностранного факультета. Случайная выборка случайно оказалась не репрезентативной.

В случайной выборке можно предположить, по каким характеристикам она будет нерепрезентативной, и если есть такие характеристики, то необходимо увеличить объем выборки. Например, исследуя курсантов военных вузов, мы, скорее всего, будем изучать курсантов мужского пола, но сейчас достаточно много девушек обучается в военных вузах. В связи с этим для репрезентативности выборки необходимо в наше исследование включить несколько курсантов женского пола.

Статистический вывод строится на оценке характеристик в выборке и последующем размышлении о том, насколько данная выборка в действительности оценивает исследуемые характеристики, насколько она соответствует параметрам генеральной совокупности. Убрать несоответствие можно относительно простым способом, увеличив объем выборки.

Случайные события

Случайными называют события, удовлетворяющие трем условиям:

1) такие события можно многократно повторять (кидать кубик);
2) при повторениях условия не меняются (центр тяжести кубика не меняется);
3) исходы таких событий неоднозначны.

Понятие вероятности подразумевает как минимум два исхода (при случайном выборе курсанта в журнале мы можем выбрать юношу, а можем девушку).

Опыт — понятие, подразумевающее наблюдение или исследование появления какого-либо события.

Событие — то, что происходит, какой-либо факт, интересующий психолога.

Условия опыта — существующие события или состояния испытуемого, оказывающие воздействие на опыт. Условий опыта обычно множество, и их все невозможно учесть, поэтому обычно выделяют наиболее значимые, наиболее влияющие на проведение опыта. Иногда учитывают наиболее желательные условия.

События, которые могут произойти при данных условиях опыта, называются исходами.

Испытанием называют сочетание условий опыта и исходов.

В теории вероятностей можно выделить три типа событий:

— детерминированные;

— случайные;

— неопределенные.

Детерминированное или определенное событие должно обязательно произойти, а если уже произошло, то оно имеет причину. Нам известно, что 1-е января наступит — это детерминированное событие. Если событие в результате испытания наступает всегда, его называют достоверным; если никогда — невозможным, это детерминированные события. Такие события можно предсказать с высокой вероятностью.

Неопределенным называют событие, исход которого непредсказуем, чаще всего это уникальные события.

Случайным, как мы уже знаем, называют событие, которое может произойти, а может и не произойти. События также могут быть равновероятными и разновероятными (неравновероятными), но при этом сумма этих событий должна равняться единице.

В теории вероятностей есть закон больших чисел — среднее значение результатов опыта, проведенного большое количество раз, соответствует математическому ожиданию этого распределения — вероятность события А расположена между невозможным событием (0 на шкале) и достоверным событием (1 на шкале). Другими словами, если при бросании кубика 100 раз 17 раз выпадала единица, то при бросании кубика 1000 раз велика вероятность 167 раз выпадения единицы и т. д., но при сохранении тех же условий опыта.

Условия опыта в психологических исследованиях практически невозможно полностью повторить (в одну и ту же реку нельзя войти дважды). Изучая автомобильные аварии в целях их предупреждения, необходимо учитывать психофизиологическое состояние водителя, состояние покрытия, техническое состояние автомобиля, погодные условия и т. д. Учесть все условия практически невозможно, поэтому аварии считаются случайными событиями. Неизвестно, когда и как произойдет авария, считается, что это зависит от случая, то есть является случайным.

Случайное событие происходит тогда, когда один из параметров будет иметь определенную величину, определенное значение, такие величины, значения которых зависят от случая, называются случайными величинами.

Случайные величины делятся на две группы:

— дискретные, принимают конкретные значения, эти значения конечны, а если бесконечны, то их можно подсчитать;

— непрерывные, принимают всевозможные значения из конечного или бесконечного интервала.

Если случайные события возникают последовательно, то можно определить их статистическую закономерность — причинную связь (правда, только для больших выборок), при которой конкретное состояние определяет последующее состояние с определенной вероятностью. Например, если у родителей карие и голубые глаза, то с вероятностью 50 % у ребенка будут карие глаза. Если известны начальные условия (цвет глаз родителей), то можно предсказать дальнейшее развитие системы с определенной долей вероятности.

События, для которых исход неоднозначен, не всегда являются случайными. Будет ли курсант или слушатель учиться без троек, болеть, попадет ли в тире в мишень и т. д. — исход этих событий неопределенный, но не случайный. Данные события нельзя многократно повторять (если будет один и тот же слушатель, то при многократном повторении он выучит все, а если будут разные слушатели, то изменяются первоначальные условия), их исход неоднозначен, его нельзя предсказать, они относятся к неопределенным, но никак не случайным.

Системой событий называется несколько событий, изучаемых вместе. Как и любая система, система событий обладает другими свойствами по сравнению с каждым отдельным событием. События в системе могут происходить одновременно или нет, зависеть друг от друга или нет, по-разному упорядочены.

Вероятностное пространство[1] связано со случайными событиями — это перечень всевозможных исходов событий, определенных заранее, и при этом неизвестно, какой именно из исходов наступит. Таким образом, случайные ситуации становятся более определенными.

Распределения вероятностей событий

Распределением вероятностей событий называют описание вероятности значения конкретного события, его можно показать в виде графика, таблицы или формулы.

Вероятность события А в системе событий можно представить в виде формулы:

$$0 \le P \le 1 \qquad P_1 + P_2 + ... + P_K = \sum_{i=1}^{K} P_i = 1.$$

где P (probability) — вероятность, значение которой может изменятся от 0 до 1, P_1, P_2….P_k — вероятность возникновения 1, 2…k события.

Распределения бывают:

— условными и безусловными. Условное распределение случайной величины (A, B) происходит при выполнении определенного условия, что другая случайная величина будет иметь определенное значение: P (A|B) = P (AB) / P (A);

— изолированными и совмещёнными;

— общими и частными,

— распределения, сочетающие различные из перечисленных признаки.

Все случайные события можно разделить на группы в зависимости от того, могут они появляться вместе в одном опыте или не могут. В одну группу будут входить события, которые не могут одновременно появиться в одном опыте, такие события называются *несовместными* или несовместимыми. Другую группу, соответственно, будут составлять события, которые могут появиться в одном опыте, а могут и не появиться, такие события называются *совместными* или совместимыми.

Также одно событие может оказать влияние на появление другого события в этой же системе или в другой. *Зависимые* события — события, влияющие на появление другого события; *независимые* события не влияют на возможность появления другого события.

---

[1] Понятие введено А. Н. Колмогоровым (1903–1987) в начале XX века.

Квантификация или количественное определение события возникает, когда мы события заменяем числами, превращаем качественные данные в количественные. Количественные данные можно сравнивать, прогнозировать.

Величиной будем называть какое-либо явление, как физическое, физиологическое, так и психическое, которое может быть представлено в виде чисел. Если условия изменяются, а числа нет, то такую величину называют *константой*. Если при изменении условий происходит изменение чисел по определенной закономерности, то такую величину называют *переменной*. Если при постоянных условиях величина принимает только одно значение и заранее неизвестно — какое, то это *случайная* величина.

Случайная величина — это функция (f) среди множества событий, когда каждому событию соответствует определенное число. Для случайной величины характерна статистическая устойчивость.

Погрешности измерений считаются случайными величинами.

Случайные, так же как и неслучайные, величины могут быть дискретными и непрерывными, об этом было сказано выше. К дискретным случайным величинам можно отнести, например, количество испытуемых, количество случаев и т. д. К непрерывным случайным величинам мы отнесем время реакции, индивидуальную минуту и т. д. Непрерывная случайная величина определяется конечным числом заданных интервалов.

Ряд распределения дискретной случайной величины представляет собой функцию, выражающую значения случайной величины, зависимые от вероятности (частоты, частости) ее появления. Данный ряд можно представить в виде таблицы или в виде диаграммы (рис. 7).

X	x_1	x_2	x_3	x…	x_n
P	p_1	p_2	p_3	p….	p_n

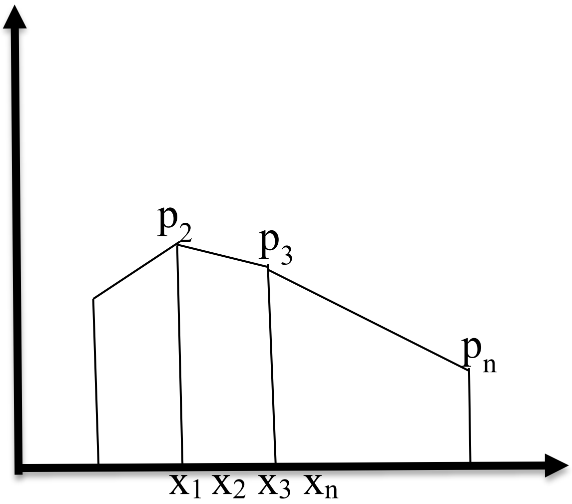


Рис. 7. Ряд распределения дискретной случайной величины X

Также распределение вероятностей значений дискретной случайной величины можно представить в виде кумулятивного ряда распределений, то есть совокупности вероятностей. Совокупность вероятностей образует полигон распределения (рис. 8) или гистограмму распределения вероятностей (рис. 9).

Рис. 8. Полигон распределения вероятностей

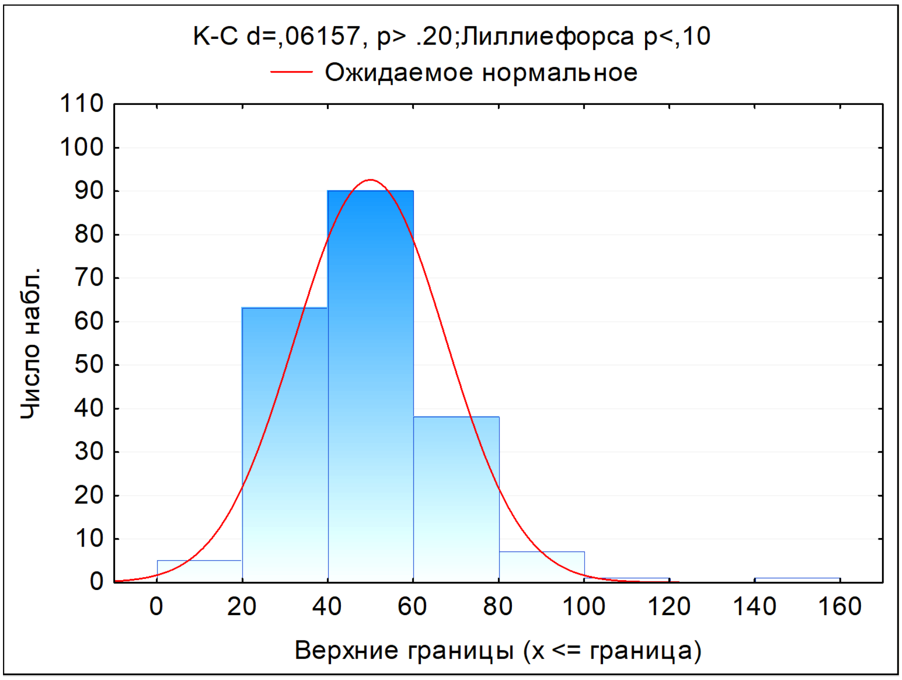


Рис. 9. Гистограмма распределения вероятностей

Законом распределения случайной величины называется математическое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями. Графическое выражение данного распределения называется *кривой распределения*.

Кривая распределения демонстрирует особенности генеральной совокупности, полученные на основе выборки, и позволяет наглядно представить выявленные закономерности.

Эмпирические кривые распределения делятся на две большие группы (рис. 10, 11):

1) одновершинные:

— симметричные,

— умеренно асимметричные или скошенные,

— крайне асимметричные,

— U-образные;

2) многовершинные.

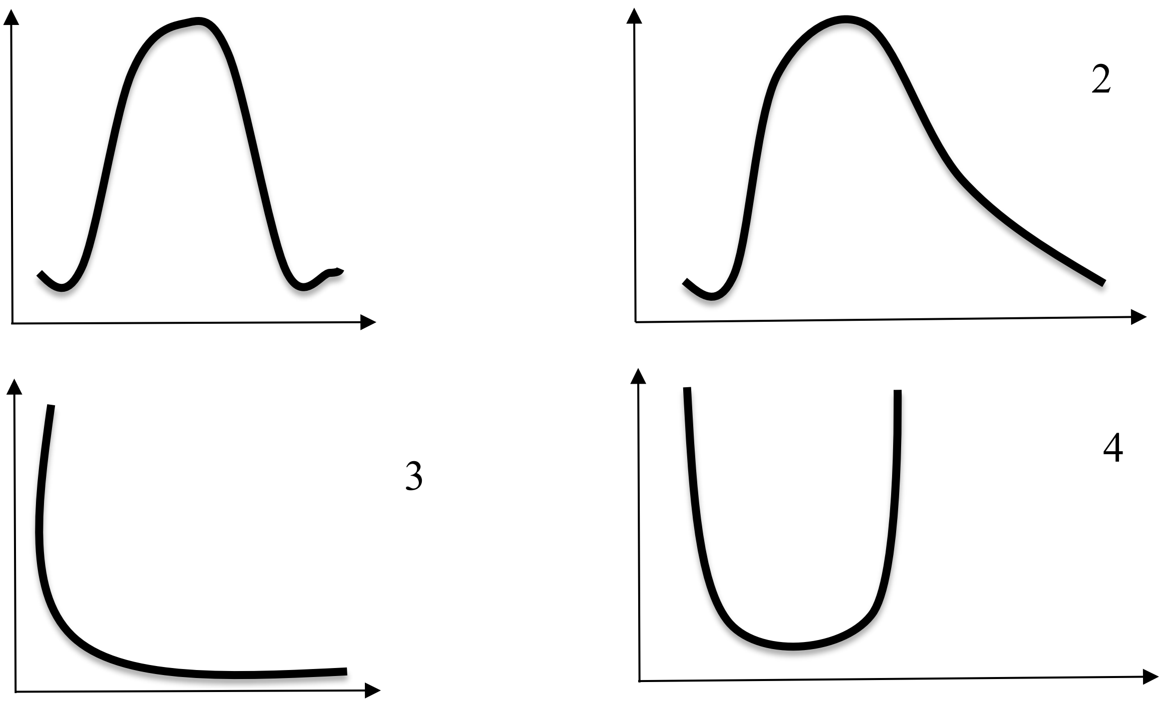


Рис. 10. Одновершинные виды распределений
1– симметричное, 2 – умеренно асимметричное или скошенное,
3 — крайне ассиметричное, 4 –U-образное

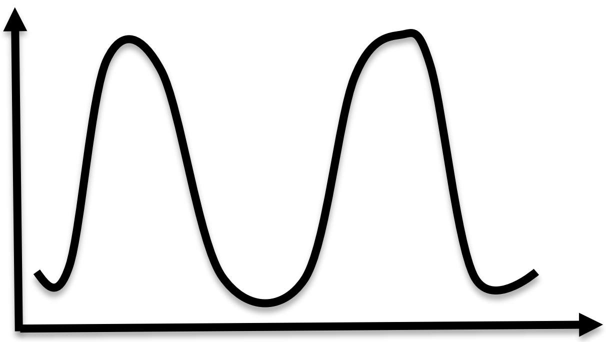


Рис. 11. Пример многовершинного распределения

Распределение вероятностей случайных величин как дискретных, так и непрерывных, которые могут отличаться расположением на числовой оси (рис. 12).

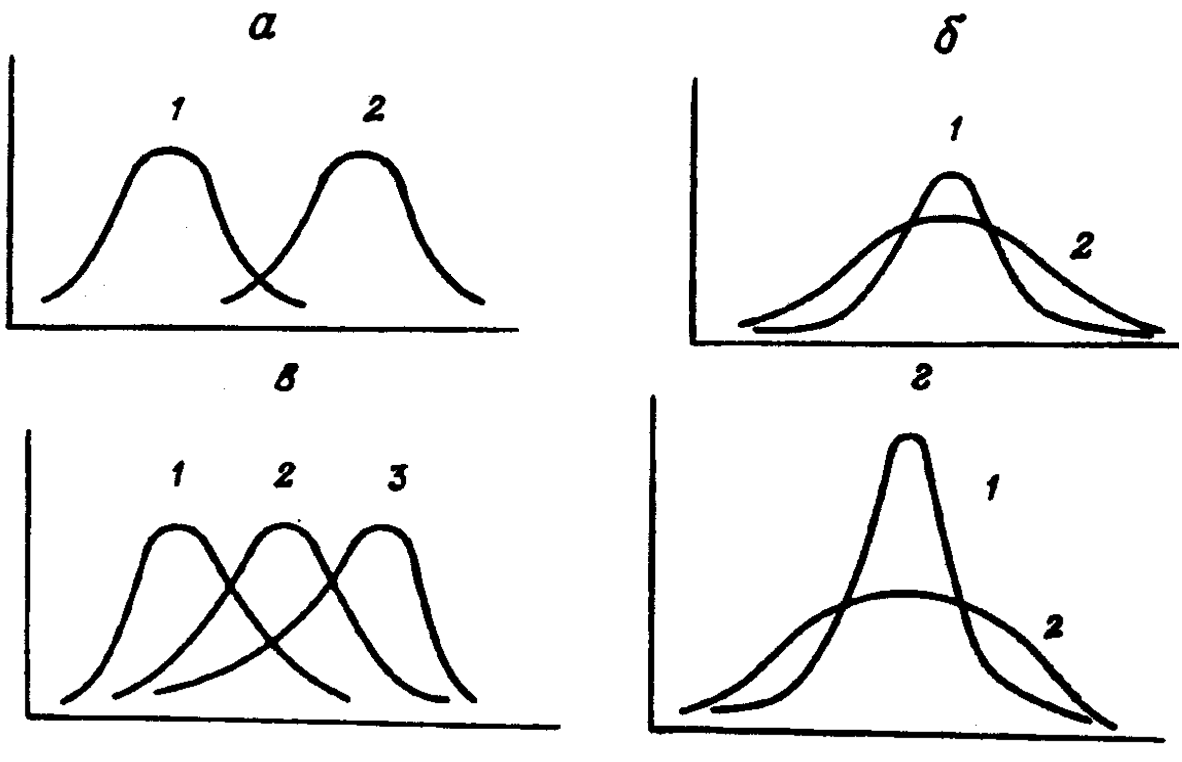


Рис. 12. Основные свойства распределения вероятностей случайной величины

По оси абсцисс — значения X, по оси ординат — вероятности f(x); *а* — кривые различаются положением на числовой оси; *б* — положение кривых одинаково, отличаются рассеиванием, у 1 — меньше; *в* — кривые отличаются положением и асимметрией, 2 — симметричная, 1 — отрицательная симметрия, 3 — положительная; *г* — отличия кривых в дисперсии и эксцессе, эксцесс больше у 1.

Представленные свойства распределения вероятностей случайной величины зависят как от самой величины, так и от условий, в которых она была выявлена. Данные свойства выражены количественно и, в основном, именно они чаще всего встречаются в исследованиях.

Следовательно, количественные меры, представленные визуально, позволяют сравнивать случайные величины в зависимости от выраженности отдельных мер, таких как положение, рассеивание, асимметрия, эксцесс, а также можно увидеть основные функции эмпирического распределения. Данные количественные меры называются числовыми характеристиками или параметрами.

В математической статистике существует несколько законов распределения случайной величины для анализа психологических исследований большое значение имеет закон нормального распределения.

Рассмотрим основные законы распределения:

I. Дискретной случайной величины:

1. Биномиальный закон распределения.
2. Геометрическое распределение.
3. Распределение Пуассона (Закон редких явлений).

II. Непрерывных случайных величин:

1. Равномерное распределение.
2. Закон нормального распределения.

Биномиальный закон распределения подходит для независимых событий, чтобы узнать, сколько раз происходит некоторое событие в серии из определенного числа независимых наблюдений, происходящий в одинаковых условиях.

$$P(m) = C_n^m * p^m * q^{n-m},$$

где $P(m)$ — вероятность наступления события А,

m — количество опытов, в которых наступает событие А,

n — общее количество опытов,

$q = 1-p$ — вероятность того, что событие А не произойдет,

C — количество комбинаций

Например, узнать вероятность рождения мальчика, если в семье планируется двое детей:

$n = 2$, p (вероятность рождения мальчика) = 0,5; q (вероятность рождения девочки) = 0,5;

$p(0)$ вероятность рождения 0 мальчиков, $m = 0$

$p(0) = C_2^0 * p^0 * q^2 = 1*1*0{,}5^2 = 1*1*0{,}25 = 0{,}25$

$p(1)$ вероятность рождения 1 мальчика, $m = 1$

$p(1) = C_2^1 * p^1 * q^1 = 0,5$

p (2) вероятность рождения 2 мальчиков, m = 2

$p(2) = C_2^2 * p^2 * q^0 = 0,25$

Все вероятности в сумме дают единицу, 0,25 + 0,5 + 0,25 = 1

Составляем таблицу вероятностей:

x	0	1	2
p	0,25	0,5	0,25

С помощью *геометрического распределения* оценивают вероятность проведения некоторого числа испытаний для достижения успеха. Например, вероятность попадания в мишень 0,7, X — число выстрелов, определить вероятность попадания в мишень после 2, 3, 4 и 5 выстрелов, математическое ожидание и дисперсию.

$P(x=k) = pq^{k-1}$, k = 1,2,3…. $M(x) = 1/p$ $D(x) = q/p^2$

$P(x=2) = pq = 0,7*0,3 = 0,21$

$P(x=3) = pq^2 = 0,7*0,3^2 = 0,063$

$P(x=4) = pq^3 = 0,7*0,3^3 = 0,0189$

$P(x=5) = pq^4 = 0,7*0,3^4 = 0,006$

$M(x) = 1/p = 1/0,7$

$D(x) = q/p^2 = 0,3/0,7^2 = 0,61$

Закон распределения Пуассона представляет собой число событий, произошедших за ограниченное время, при условии, что события независимы и происходят с четкой средней интенсивностью.

Распределение вероятностей называется равномерным, если на интервале (a, b) дифференциальная функция распределения имеет постоянное значение, плотность вероятности является постоянной величиной:

$$f(x) = C$$

Закон нормального распределения, или закон Гаусса, — закон распределения вероятности непрерывных случайных величин. Величина подчиняется нормальному распределению, когда на нее влияет большое количество случайных помех. Нормальным оно названо потому, что часто встречается в распределениях.

Закон нормального распределения определяет вероятность какого-либо значения из всего спектра значений. Он имеет большое значение в психологических исследованиях, лежит в основе измерений, проверки гипотез, создания тестов. Его особенностью является то, что вид графика определяется двумя параметрами — средним арифметическим (X) и стандартным отклонением (σ), внешний вид похож на колокол, поэтому такой график называют колоколообразной кривой или кривой Гаусса (рис. 13). Площадь под кривой нормального распределения равна 1. По оси абсцисс располагается величина изучаемого признака, по оси ординат — частота встречаемости.

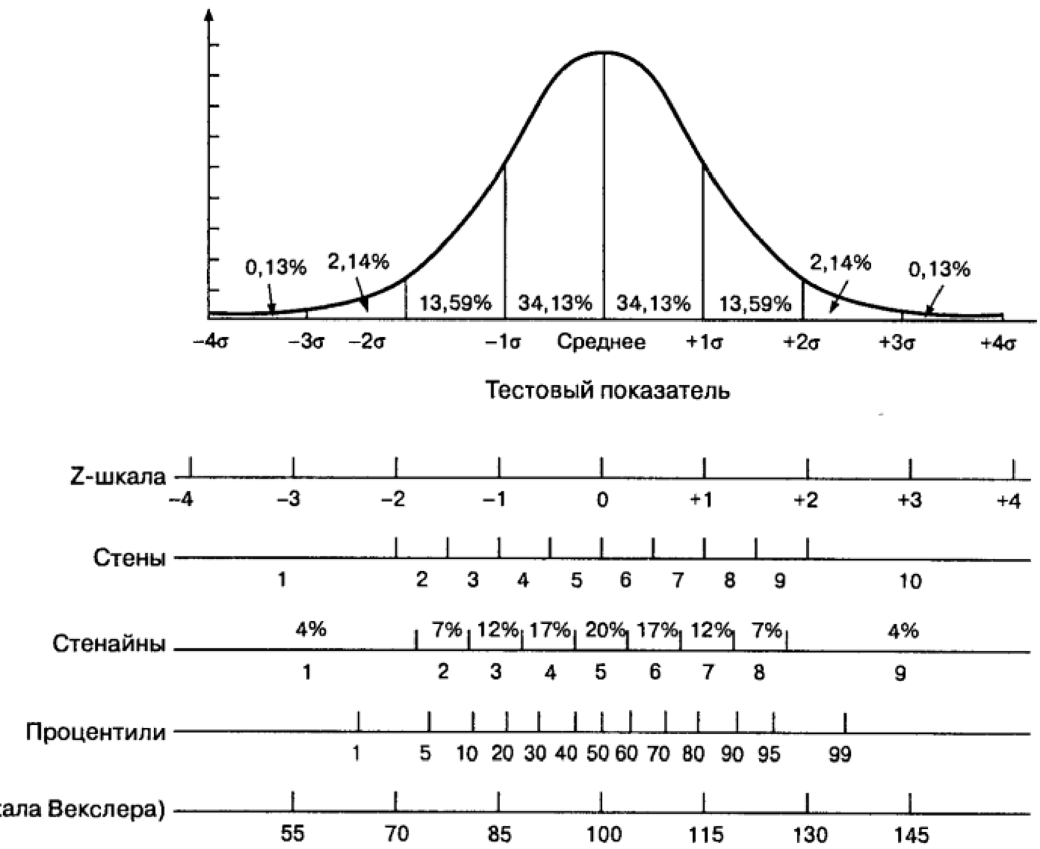


Рис. 13. Кривая нормального распределения и тестовые шкалы, по А. Д. Наследову[1]

Для нормального распределения характерны совпадение средних значений с модой и медианой, на рисунке 13 это значение 0 или среднее:

$$X_{ср} = M_o = M_e$$

Таким образом, чем больше значение величины отклоняется от среднего значения, тем меньше вероятность его встретить. Наиболее часто встречаются средние значения, в интервал:

[1] Наследов А. Д. Математические методы психологического исследования. Анализ и интерпретация данных. СПб: Речь, 2012. С. 55.

34,13 % + 34,13 % = 68,26 %, M±1σ, попадает наибольшее количество всех значений;

13,59 % + 13,59 % = 27,18 %, M±2σ – M±1σ;

27,18 % + 68,26 % = 95,44 %, M±2σ;

M±3σ дают 99,73 % из всех значений нормального распределения.

Таким образом, если наша выборка показала нормальное распределение, то, зная среднее значение и значение стандартного отклонения, мы можем узнать с вероятностью 99,73 % максимальное и минимальное предполагаемые значения.

На рисунке 13 представлена Z шкала, которая помогает привести любые значения, полученные в исследовании, к стандартным, для того чтобы можно было сравнивать разные показатели, представленные в разных шкалах. Например, можно сравнить показатель интеллекта, IQ и скорость реакции, если провести Z преобразование по следующей формуле:

$$Z = (x_i - M) / \sigma_x$$

Z-оценки показывают относительное расположение конкретного данного по сравнению со всеми результатами выборки, независимо от среднего значения и стандартного отклонения. Недостатком Z-оценки являются отрицательные значения и дробные величины, поэтому Z-оценки обычно преобразуют в другие стандартные шкалы (шкала стенайнов, шкала стенов, шкала процентилей), для примера представлена шкала IQ методики Векслера (M = 100, σ = 15) (рис. 13).

Для нормального распределения используются специфические критерии, основанные на показателях средних значений.

Сопоставить полученную в эмпирическом исследовании кривую с кривой нормального распределения можно с помощью критерия «хи-квадрат» Пирсона $\chi^2_{эмп}$, критерия Колмогорова-Смирнова (используется, если известны среднее значение и стандартное отклонение), критерия Лиллиефорса (если среднее значение и стандартное отклонение неизвестны, но вычисляются по выборке), критерия Шапиро-Уилка (является наиболее мощным, универсальным, применяется если неизвестны возможные отклонения от нормальности, $8 < n < 50$) и другие.

2.4. Способы визуализации данных

Математическая статистика помогает собирать и обрабатывать данные. Для лучшего понимания представленных данных существуют различные способы визуализации. Известно, что информация воспринимается лучше, если она наглядно представлена.

Одним из наиболее простых способов визуализации статистической информации являются таблицы. Статистическая таблица — способ оптимального представления статистических данных при помощи чисел, расположенных по определенному правилу. Пример визуализации данных в виде таблицы представлен ниже.

Таблица 1. Характеристика выборки сотрудников правоохранительных органов (исходя из половой принадлежности)[1]

Выборка	*N*	$\bar{X}\pm\sigma$	*min*	*max*
Мужчины	155 (75 %)	0,75±0,03	1	
Женщины	51 (25 %)		0	
Возраст	206	35,25±3,44	22	52
Стаж службы	206	13,80±2,45	1	28
Звание	206	2,90±0,91	младший начальствующий состав	старший начальствующий состав
Подразделение (экстремальный характер деятельности)	206	0,81±0,2	0	2

Таблица частот (табл. 2) наглядно показывает частоту встречаемости исследуемого или значимого для исследования элемента. Чтобы узнать относительную частоту встречаемости какого-либо признака, надо абсолютную частоту разделить на объем всей выборки. Относительная частота показывает вероятность появления события в данной выборке. Накопленная частота, или кумулята, показывает сумму накопленных частот.

---

[1] Баринова М. Г., Зуева Е. Г., Психофизиологическая готовность сотрудников правоохранительных органов к деятельности в экстремальных ситуациях: монография. СПб.: Изд-во СПб ун-та МВД России, 2019. С. 46–48

Таблица 2. Распределение частот

Рост слушателей иностранного факультета	f_a абсолютная частота	f_o относительная частота	f_{cum} накопленная частота
160	1	0,0625	0,0625
164	1	0,0625	0,125
166	3	0,1875	0,3125
168	4	0,25	0,5625
170	3	0,1875	0,75
176	2	0,125	0,875
178	1	0,0625	0,9375
182	1	0,0625	1
∑ (сумма)	16	1,000	

Эти же показатели можно представить в виде гистограммы (рис. 14), где по оси абсцисс будет рост слушателей, а по оси ординат — частота встречаемости

Рис. 14. Гистограмма распределения

Также мы можем построить полигон распределения (рис. 15) и круговую диаграмму, показывающую в процентном соотношении рост слушателей (рис. 16).

Рис. 15. Полигон распределения

Рис. 16. Круговая диаграмма

Есть еще средства визуализации, позволяющие показать сразу две изучаемые характеристики, например, рост и вес слушателей иностранного факультета (рис. 17).

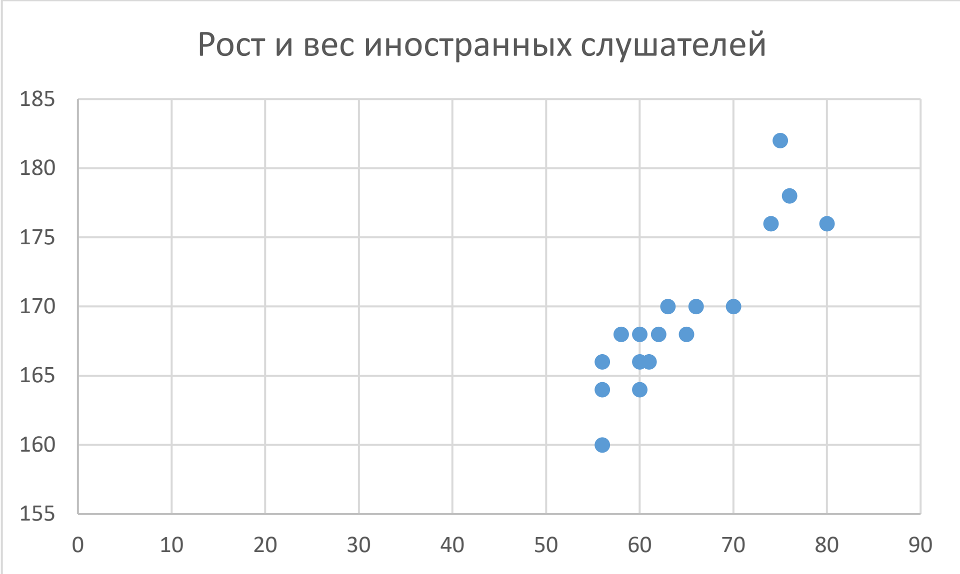


Рис. 17. Точечная диаграмма или диаграмма рассеяния

Пузырьковая диаграмма показывает три характеристики исследуемой выборки, на рисунке 18 изображена пузырьковая диаграмма, где по основным осям расположены рост и вес слушателей, а размер «пузыря» или точки указывает на третье свойство, в нашем случае — средний балл по физкультуре, полученный за последнюю сессию.

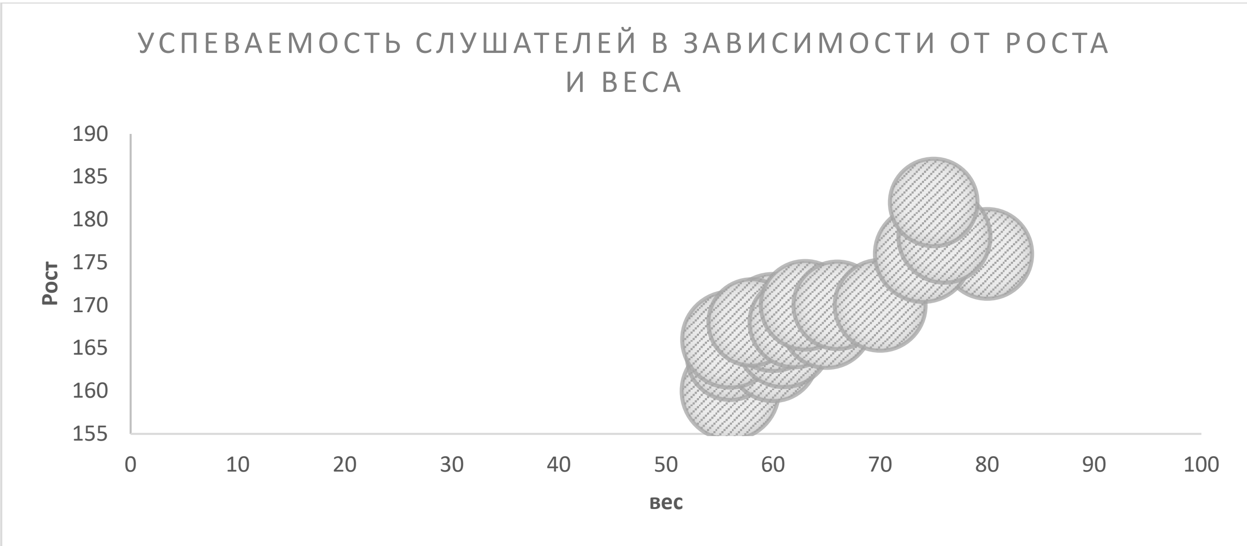


Рис. 18. Пузырьковая диаграмма

Еще одна диаграмма позволяет визуально отразить медиану, общий и межквартильный размах, по ней можно предположить, является ли представленное распределение нормальным, также на ней хорошо видны выбросы (рис. 19).

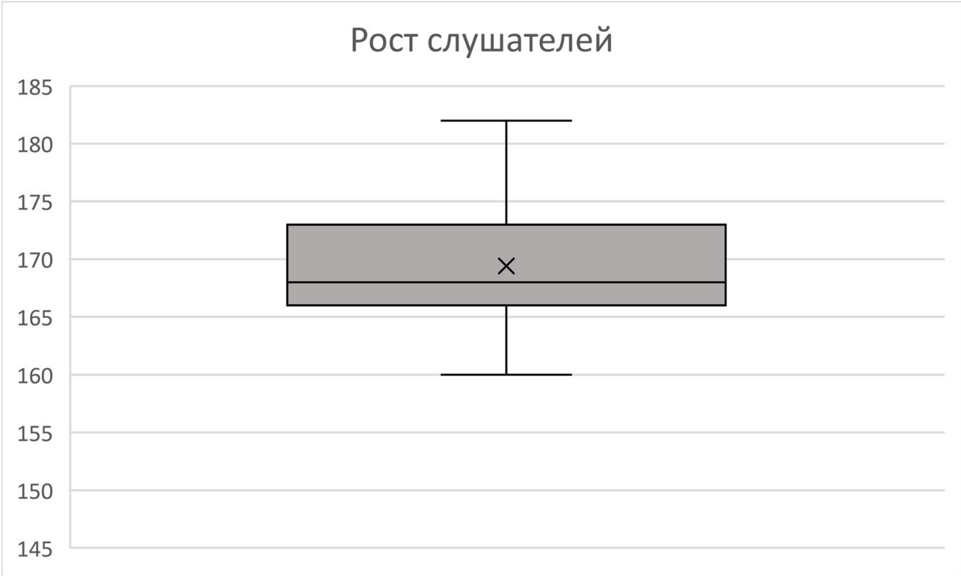


Рис. 19. «Ящик с усами» или боксплот

Диаграмма должна быть понятна, на ней не должно быть ничего лишнего, цвета рекомендуется использовать не яркие и не больше трех для презентации и черно-белые для печати (курсовые, выпускные работы, статьи).

Вопросы и задания для самоконтроля

1. Сравните понятия конечная и бесконечная совокупность.
2. Дайте определение вероятности.
3. Определите понятие «репрезентативная выборка».
4. Сравните понятия «зависимая» и «независимая» выборка.
5. Приведите пример необходимости рандомизации.
6. Дайте определение понятию дискретные случайные величины.
7. Сравните понятия «зависимые» и «независимые» события.
8. Нарисуйте полигон распределения.
9. Постройте гистограмму своих оценок.
10. Постройте диаграмму вашей группы в зависимости от пола.

Глава 3. СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ

Учебные вопросы:

3.1. Статистическая значимость.
3.2. Статистические гипотезы.
3.3. Зависимые и независимые выборки.
3.4. Степени свободы.

Статистическая обработка результатов психологического исследования необходима для извлечения полезной информации из количественных данных, для аргументированного доказательства сформулированных гипотез. Статистические методы получили распространение в научных исследованиях во второй половине XX века.

3.1. Статистическая значимость

Статистическая значимость показывает степень репрезентативности выборки, то есть степень достоверности результатов выборки для генеральной совокупности. Степень, или уровень достоверности обозначается *p*. Чем выше уровень достоверности, тем менее надежны результаты исследования.

Статистическая значимость зависит от трех параметров: истинность гипотезы; нормальное распределение; *p*-значение.

Истинность/ложность гипотезы мы рассмотрим в следующем вопросе. Нормальное распределение, как мы уже рассмотрели в предыдущей теме, связано с правилом трех сигм — в одну сигму входит 68 % всей выборки, в две сигмы — 95 %, в три сигмы — 99,7 %.

Р-уровень показывает вероятность ошибки. В психологических исследованиях принят уровень $p \leq 0{,}05$, который считается приемлемой границей статистической значимости. Например, $p = 0{,}1$ означает, что есть 10 % вероятность того, что полученные в эксперименте результаты являются случайными и характеризуют именно эту выборку. Для психологических исследований такое значение p не считается доказательством истинности гипотезы. $P = 0{,}01$ показывает, что существует 1 % вероятность, что выявленная связь является случайной, скорее всего, она характерна для всей генеральной совокупности.

Вероятность ошибки (р — уровень)

←

Уровень значимости (надежность) результатов

→

Е. В. Сидоренко предлагает «ось значимости» (рис. 20), которая показывает, что значения, лежащие правее 0,05, являются незначимыми, а правее 0,01 — значимыми[1].

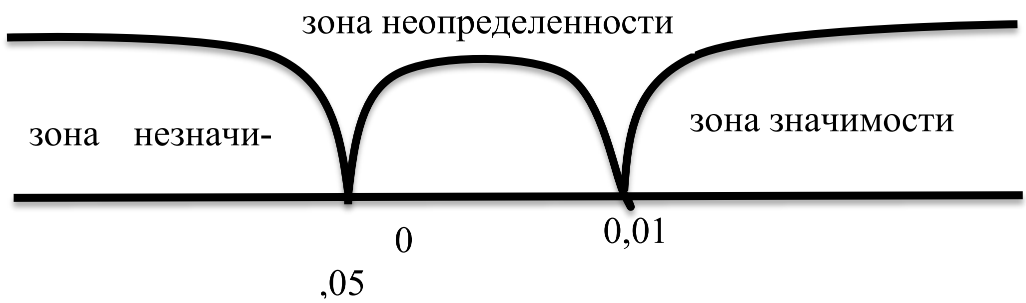


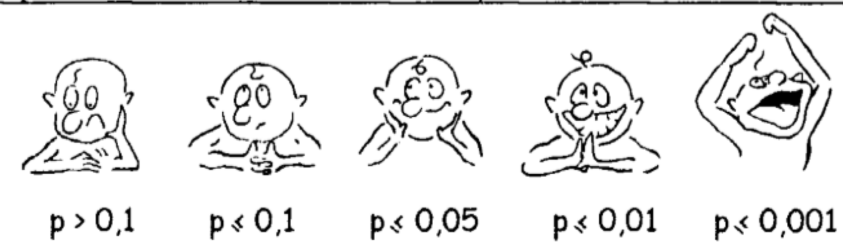
Рис. 20. Ось значимости (по Е.В. Сидоренко)

А. Д. Наследов представил соотношение р-уровня и степени значимости α = 0,05 (α — ошибка первого рода, описана ниже) (табл. 3).

Таблица 3. Уровни значимости по А.Д. Наследову[2]

Традиционная интерпретация уровней значимости при α = 0,05

Уровень значимости	Решение	Возможный статистический вывод
$p > 0{,}1$	Принимается H_0	«Статистически достоверные различия не обнаружены»
$p \leqslant 0{,}1$	сомнения в истинности H_0, неопределенность	«Различия обнаружены на уровне статистической тенденции»
$p \leqslant 0{,}05$	значимость, отклонение H_0	«Обнаружены статистически достоверные (значимые) различия»
$p \leqslant 0{,}01$	высокая значимость, отклонение H_0	«Различия обнаружены на высоком уровне статистической значимости»



[1] Сидоренко Е. В. Методы математической обработки в психологии. СПб.: Речь, 2010. С. 31.

[2] Наследов А. Д. Указ. соч. С. 105.

Существуют общие рекомендации объема выборки, для повышения достоверности статистических выводов:

— от 200 до 2 500 испытуемых необходимо при разработке или адаптации новой методики;

— от 25 испытуемых в каждой выборке необходимо для сравнения двух и более групп;

— от 30–35 испытуемых необходимо для выявления взаимосвязи между показателями в одной выборке.

Объем выборки также зависит от изменчивости исследуемого признака — чем больше изменчивость, тем больше объем. Уменьшая однородность выборки (например, исследуем только мужчин, или только в возрасте от 17 до 22 лет, и т. д.) уменьшается изменчивость и, соответственно, генеральная совокупность.

Статистически значимые выводы исследования доказывают выдвинутую статистическую гипотезу, но при статистически незначимых выводах можно рассмотреть причины и продумать варианты повышения статистической значимости. К таким вариантам относятся:

— увеличение выборки исследования (например, при выборке в 30 испытуемых и в 100 испытуемых можно получить выводы разной статистической значимости);

— рассмотрение выборки на предмет выскакивающих значений (исключить слишком большие и очень маленькие значения, например, в соответствии с правилом 3σ);

— выбор других параметров измерений (заменить психологические методики).

Выбор уровня значимости может быть произвольным, но надо понимать, что он зависит от величины выборки — обычно указывают минимальный уровень значимости, позволяющий отвергнуть гипотезу, соответственно, и экспериментальный эффект будет также минимальным. Кроме того, надо понимать, что мы всегда можем отвергнуть H_1, но никогда не сможем ее доказать; когда мы отвергаем H_1, то автоматически принимаем H_0, например, на уровне $p = 0{,}05$, это две противоположные гипотезы.

3.2. Статистические гипотезы

Начнем с того, что гипотеза — это предположение, требующее проверки. Любое научное исследование начинается с выдвижения гипотез, а целью исследования является их проверка.

Гипотезы делятся на научные и статистические.

Научные гипотезы — это предположения, требующие для своего выдвижения определенный уровень теоретических научных знаний, а для проверки этих предположений — практические навыки научной работы. Не все суждения могут быть гипотезами, для каждой из них можно сформулировать противоположное суждение, и оно должно быть реальным.

Часто формулируется не одна, а несколько гипотез, взаимосвязанных логически, и часть из которых можно проверить в эксперименте. Например, можно предположить, что самоизоляция повышает тревожность, или что агрессивность зависит от возраста. В первом случае надо исследовать тревожность в период самоизоляции и в обычное время, во втором — агрессивность в разные возрастные периоды.

Статистические гипотезы — это предположения о характеристиках случайной величины, которые подтверждаются или опровергаются с помощью статистических методов.

При выдвижении гипотезы выделяют несколько переменных, каждую из которых необходимо отнести к зависимой или независимой. *Зависимой* называют переменную (ЗП), изменения которой зависят от изменений независимой переменной. *Независимая* переменная (НП) контролируется экспериментатором и изменяется в зависимости от него, она должна принимать минимум два значения.

Например, мы хотим выяснить, зависит ли уровень интеллекта от объема памяти. Объем памяти будет независимой переменной, для этого мы должны исследуемую выборку разделить на группы, в зависимости от объема памяти (минимум 2 группы (большой и маленький объем), но лучше больше). Следующим этапом определяем уровень интеллекта в нашей выборке — это будет зависимая переменная, так как в нашей гипотезе уровень интеллекта зависит от объема памяти. Сравниваем выделенные группы (группа с большим объемом памяти и группа с маленьким объемом памяти) по уровню интеллекта. Если с помощью статистических критериев выявляются значимые различия между группами ($p \leq 0{,}05$), то можно сделать вывод, что объем памяти влияет на уровень интеллекта. Если же значимых различий не выявляется ($p > 0,05$), то можно сделать вывод о том, что интеллект не зависит от объема памяти.

В психологии выделяют два типа независимых переменных:

— исследователь сам изменяет переменную (например, ограничивает время опыта);

— исследователь отбирает выборки в соответствии с определенной чертой или характеристикой (зависимость интеллекта от объема памяти).

Кроме зависимых и независимых переменных на результат исследования могут также воздействовать внешние переменные (погрешность), латентные, скрытые от непосредственного наблюдения, и дополнительные — те внешние переменные, которые мы можем учесть в исследовании.

Виды статистических гипотез изображены на рисунке.

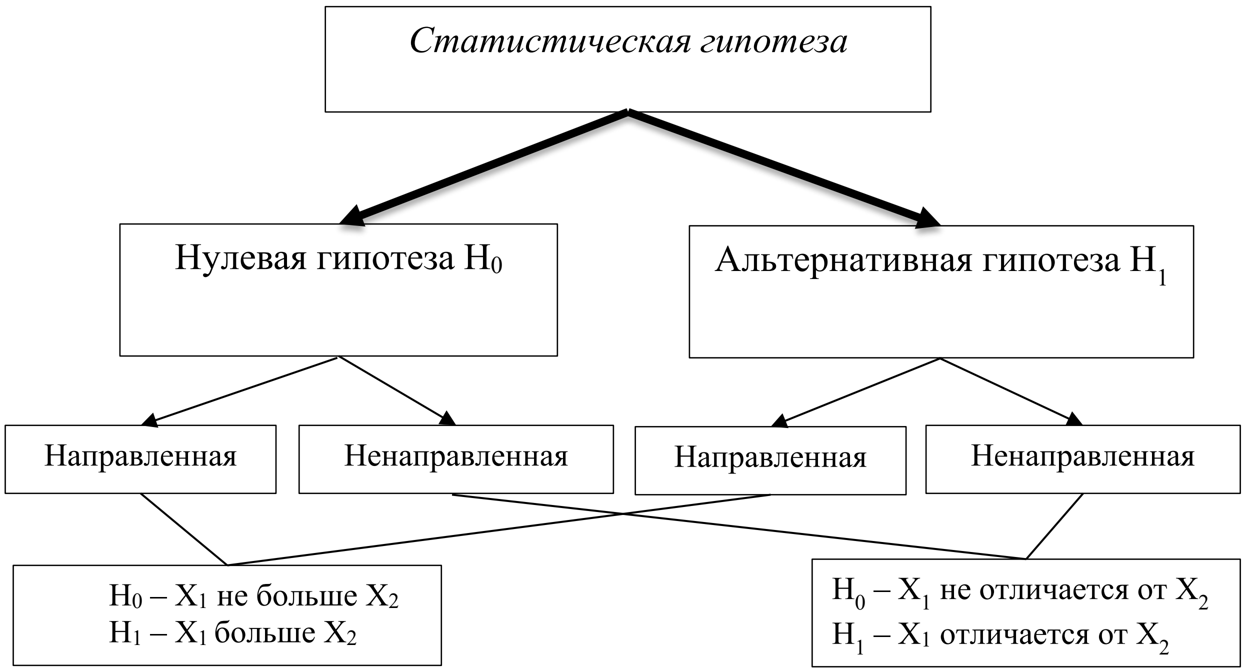


Рис. 21. Виды статистических гипотез

1. При формулировании статистических гипотез выделяют *нулевую* или основную гипотезу *(H_0)*, утверждающую, что между исследуемыми явлениями нет связи ($X_1 - X_2 = 0$), и *альтернативную,* или конкурирующую *(H_1)*, отклоняющую нулевую гипотезу, доказывающую, что связь имеется, также эту гипотезу еще называют *экспериментальной*. Если принимается нулевая гипотеза, то обязательно отвергается альтернативная, и наоборот.

Например, мы хотим определить, одинаковый ли объем памяти у слушателей иностранного факультета и курсантов 4 факультета. Сформулируем гипотезы:

— H_0 — объем памяти слушателей иностранного факультета и курсантов 4 факультета не отличается;

— H_1 — объем памяти слушателей иностранного факультета и курсантов 4 факультета различен.

2. Выделяют *направленные и ненаправленные* статистические гипотезы (приведенный выше пример содержит ненаправленные гипотезы).

Направленные гипотезы показывают направление различий:

— H_0 — объем памяти слушателей иностранного факультета не превосходит объем памяти курсантов 4 факультета;

— H_1 — объем памяти слушателей иностранного факультета больше, чем объем памяти курсантов 4 факультета.

Статистический критерий — математическое правило (формула), с помощью которой принимается или отвергается статистическая гипотеза в соответствии с уровнем значимости.

Мощность статистического критерия — способность критерия выявлять истинные различия; показывает вероятность правильного отклонения нулевой гипотезы, если она неверна.

Таблица 4. Мощность статистического критерия

	Действительное положение (пока неизвестное)	
	верна H_0	верна H_1 (не верна H_0)
принятие H_0	+	ошибка II рода = β
отвержение H_0	ошибка I рода = α	+ $1-\beta$ = мощность критерия

Когда мы отвергаем H_0, но на самом деле она верна, различий в действительности не существует, вероятность α показывает процент случайного нахождения различий, в то время как их нет. $\alpha=0,05$ говорит о том, что в 5 % возможно неверное нахождение различий, это ошибка первого (I) рода. Вероятность ошибки первого рода называется *уровнем значимости критерия*.

Ошибка, связанная с неправильным принятием нулевой гипотезы, когда различия существуют, но они не найдены, называется ошибкой второго (II) рода, и обозначается β. Вероятность ошибки II рода называется *мощностью критерия*. Мощность показывает возможность не сделать ошибку II рода, поэтому она равна 1–β.

Мощность критерия зависит от объема выборки (n), ее можно узнать только эмпирически, когда мы решаем одинаковые задачи с помощью разных критериев. При этом одни критерии выявляют различия в одних и тех же выборках, а другие — нет, или же они показывают разный уровень значимости (p).

Критерий состоит:

— из формул расчета эмпирических статистик;

— определения степеней свободы;

— теоретического распределения в зависимости от степени свободы;

— условий принятия H_0, насколько соответствует эмпирическое распределение теоретическому.

Алгоритм проверки статистической гипотезы:

1. Формулируем нулевую (H_0) и альтернативную (H_1) гипотезы.
2. Задаем желаемый уровень значимости (в психологии это обычно 0,05 или 0,01, но для пилотажного исследования может быть немного больше).
3. Сопоставляем распределение с нормальным.
4. В зависимости от предыдущего пункта и выборки выбираем статистический критерий.
5. Определяем уровень значимости (p).
6. Подтверждаем нулевую или альтернативную гипотезу.

Анализ статистических критериев является обязательным элементом проверки эмпирических данных, принятым в психологии, особенно, если полученные результаты не являются аксиомой. Наличие или отсутствие утверждений, сформулированных в гипотезах, требует доказательства и оценки их достоверности.

Уровень статистических гипотез является обязательным элементом проверки достоверности выводов и вероятности ошибок принятия или отвержения гипотез.

Статистические гипотезы не могут утверждать наличие причинно-следственных связей между зависимой и независимой переменной. Они показывают выборочные значения показателей зависимых переменных.

Статистическая гипотеза опровергает или подтверждает случайный характер выводов исследования на основе выборки и оценивает вероятность ошибки.

Распространенной в психологии статистической гипотезой является гипотеза о принятии или отвержении нулевой гипотезы, являющейся основанием для подтверждения экспериментального эффекта. Психолог хочет узнать, действительно ли имеет место воздействие независимой переменной на зависимую, но статистическую гипотезу так сформулировать нельзя. Статистические гипотезы содержат конкретные, выявляемые в сопоставлении выборок, факты. Необходимо учитывать ограничения, связанные с распространением полученных в эксперименте результатов, на проверяемые психологические гипотезы. Статистические и психологические (научные) гипотезы имеют отличия.

Когда мы отвергаем нулевую гипотезу (H_0), то обязательно должны указать уровень значимости, например, $p = 0{,}05$, и, соответственно, принять альтернативную гипотезу (H_1) о наличии различий между выборками. Это означает, что 95 % — *доверительный интервал*.

Рассмотрим понятие доверительного интервала. Доверительный интервал показывает, насколько параметры из выборочной совокупности отличаются от параметров генеральной совокупности — это допустимое расхождение выборочных значений от значений генеральной совокупности.

Пример: требуется узнать, какой рост у слушателей 8 факультета. Всех слушателей измерить сложно, поэтому возьмем выборку из 17 человек и зададим интервал, например, 95 % ($p=0{,}05$), и будем уверены, что 95 слушателей из 100 при нормальном распределении попадут в этот интервал. Это называется доверительный интервал.

Для построения доверительного интервала необходимо нормальное распределение, также необходимо знать среднее значение ($\bar{x}$), стандартное отклонение генеральной совокупности (S) и задать уровень значимости ($p=0{,}05$). Если стандартное отклонение генеральной совокупности неизвестно, то правильнее использовать доверительный интервал распределения Стьюдента, но в приведенном примере используется стандартное отклонение выборки (σ). Рассмотрим, как найти стандартное отклонение и доверительный интервал в Excel, (рис. 22, 23).

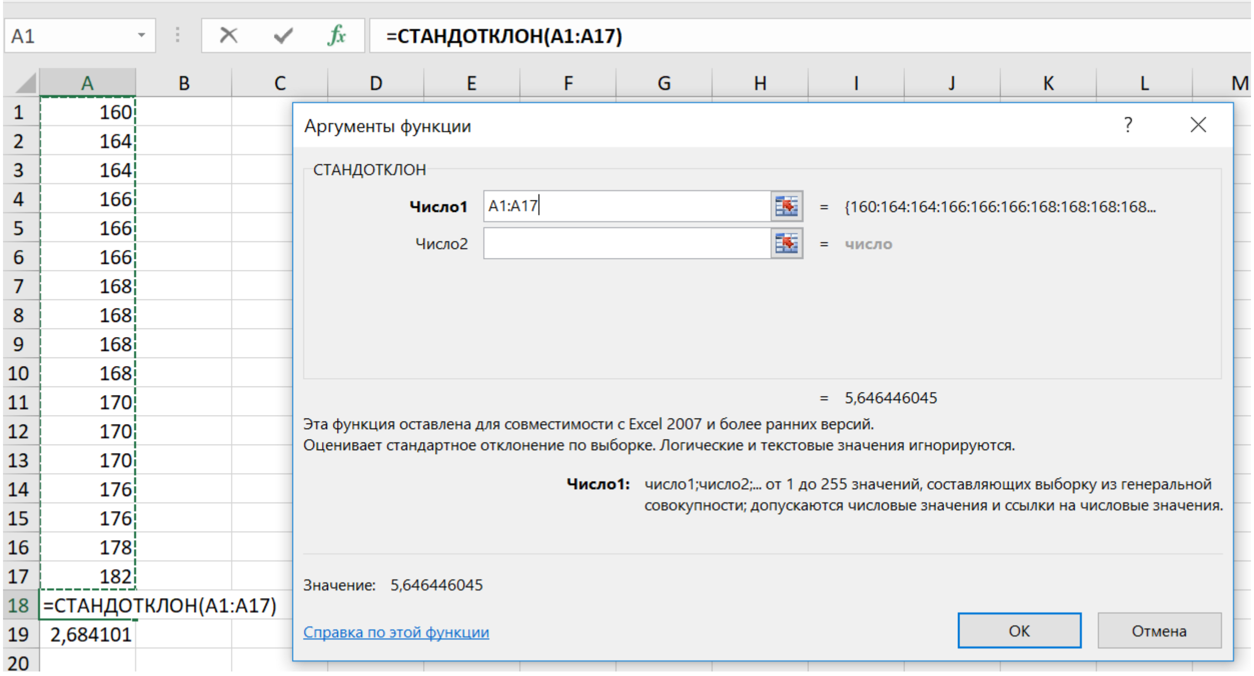


Рис. 22. Нахождение стандартного отклонения (σ) в Excel

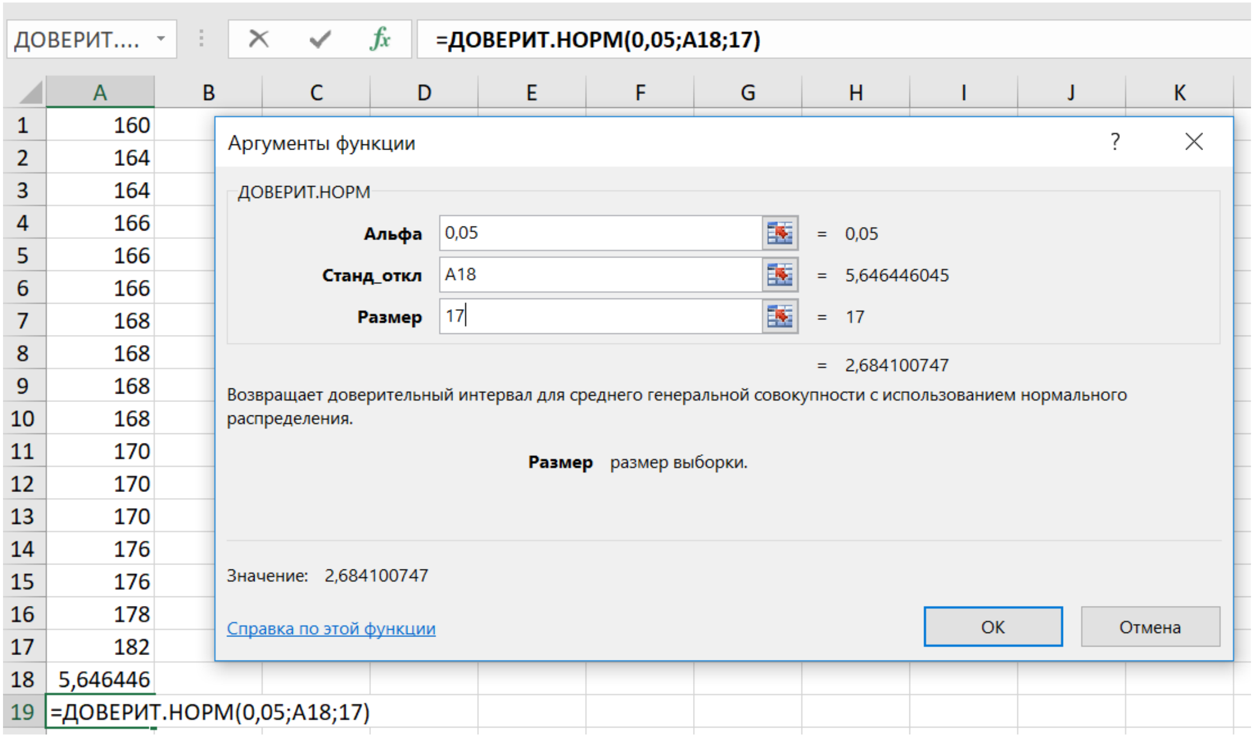


Рис. 23. Нахождение доверительного интервала в Excel

Доверительный интервал вычисляется относительно среднего значения (рис. 24) и равен: среднее значение ±доверительный интервал. Таким образом, в нашем случае он равен 169,41 ± 2,68 при $p = 0{,}05$, то есть 95 % слушателей иностранного факультета имеют рост от 166,73 до 172,09 *см*.

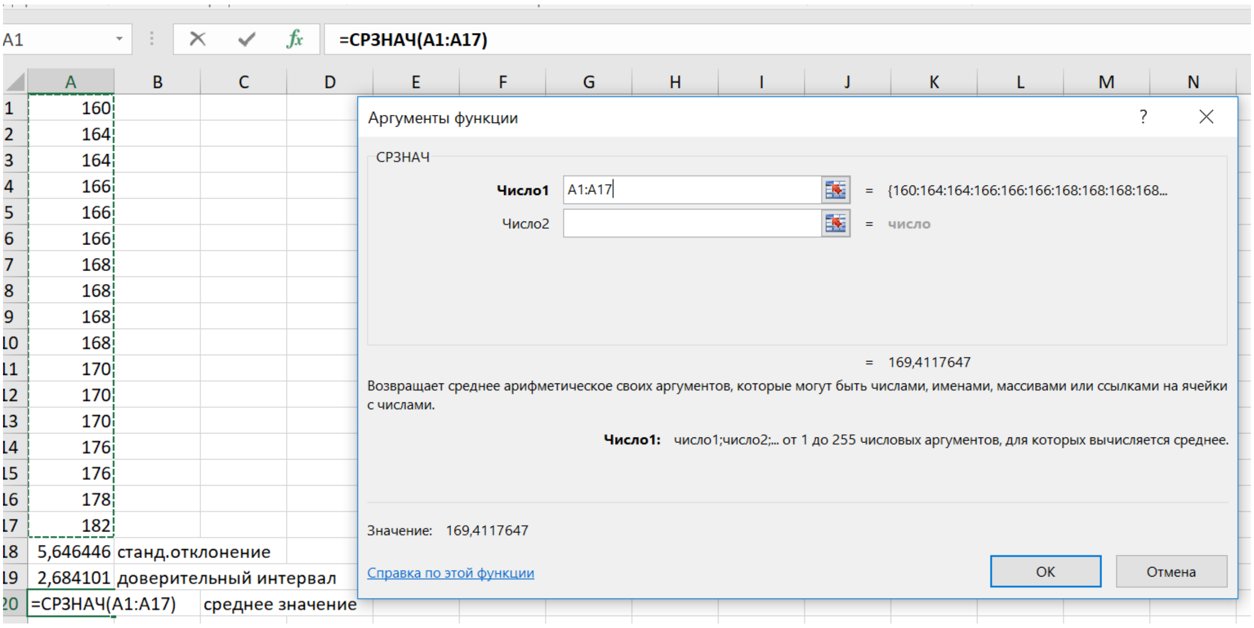


Рис. 24. Нахождение среднего значения в Excel

С помощью программы Excel можно определить доверительный интервал при нормальном распределении, используя распределение Стьюдента, которое не зависит от стандартного отклонения генеральной совокупности (S), а использует стандартное отклонение выборки (σ) (рис. 25). Получается, что 95 % слушателей иностранного факультета имеют рост от 166,51 до 172,31 *см*.

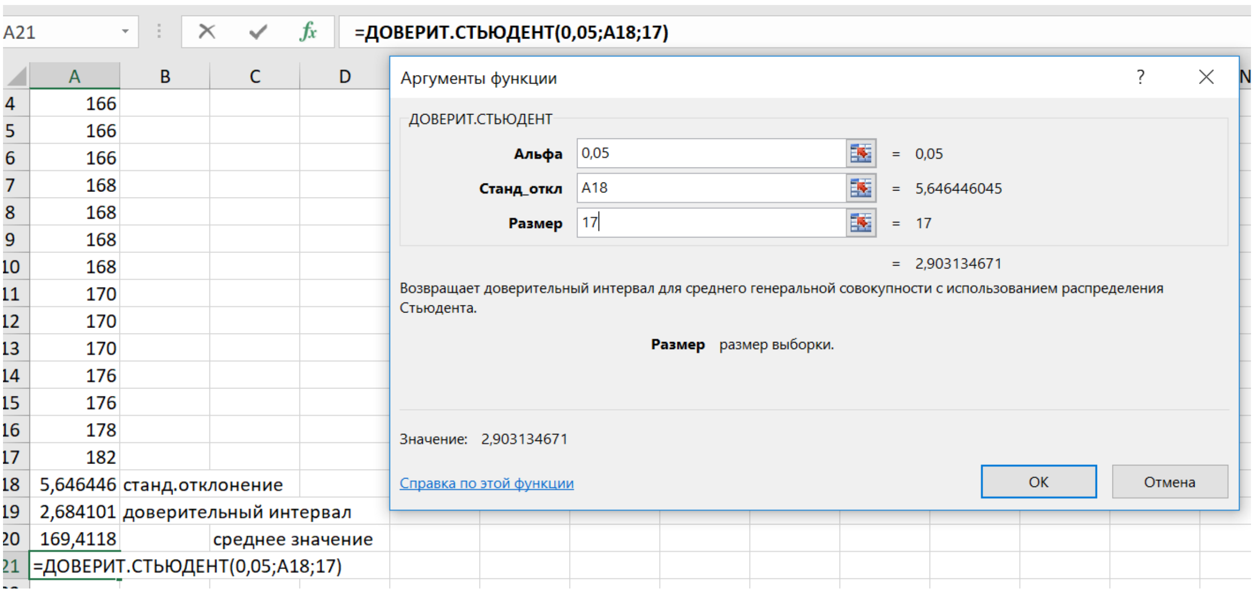


Рис. 25. Нахождение доверительного интервала с помощью распределения Стьюдента

Как мы видим, доверительный интервал в первой и второй формулах немного отличаются, потому что небольшая выборка ($n = 17$), при $n \geq 30$ распределение Стьюдента (t) приближается к нормальному, и разница доверительного интервала также практически исчезает.

Ошибка репрезентативности (m) — средняя ошибка средней величины помогает оценивать достоверность результатов исследования. Она характерна для выборочного исследования, величина этой ошибки показывает, насколько выводы, полученные с помощью выборки, различаются с выводами, которые могли бы быть получены при сплошном исследовании генеральной совокупности.

Ошибка репрезентативности (m) уменьшается при увеличении выборки (n). Средняя ошибка средней арифметической вычисляется по формуле:

$$m_M = \sigma/\sqrt{n},$$

в исследованиях она обычно обозначается как $M \pm m$.

Средняя ошибка относительной величины (P) вычисляется:

$$m_p = \sqrt{\frac{Pq}{n}}$$

где $q = 100$, если величина выражена в процентах, ответ также будет выражен в процентах,

n — количество наблюдений, если $n < 30$, то рекомендуется взять $n-1$.

Пример: возьмем предыдущий пример, с ростом слушателей и рассчитаем ошибку репрезентативности, или, как ее еще называют, стандартную ошибку в Excel.

Для того чтобы найти стандартную ошибку в Excel необходимо активировать «Пакет анализа» (рис. 26); выбираем «Описательная статистика» (рис. 27); выбираем интервал данных для подсчета стандартной ошибки (рис. 28); находим стандартную ошибку (рис. 29).

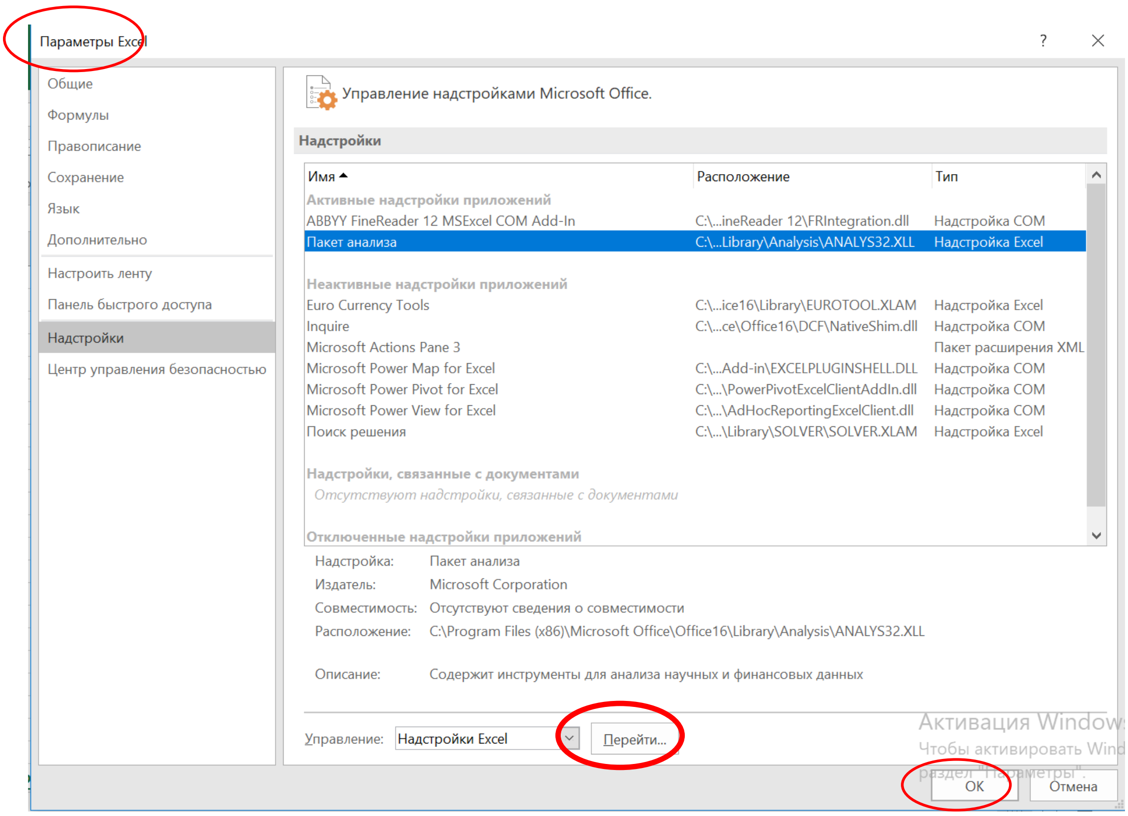


Рис. 26. Активация «Пакета анализа» в Excel

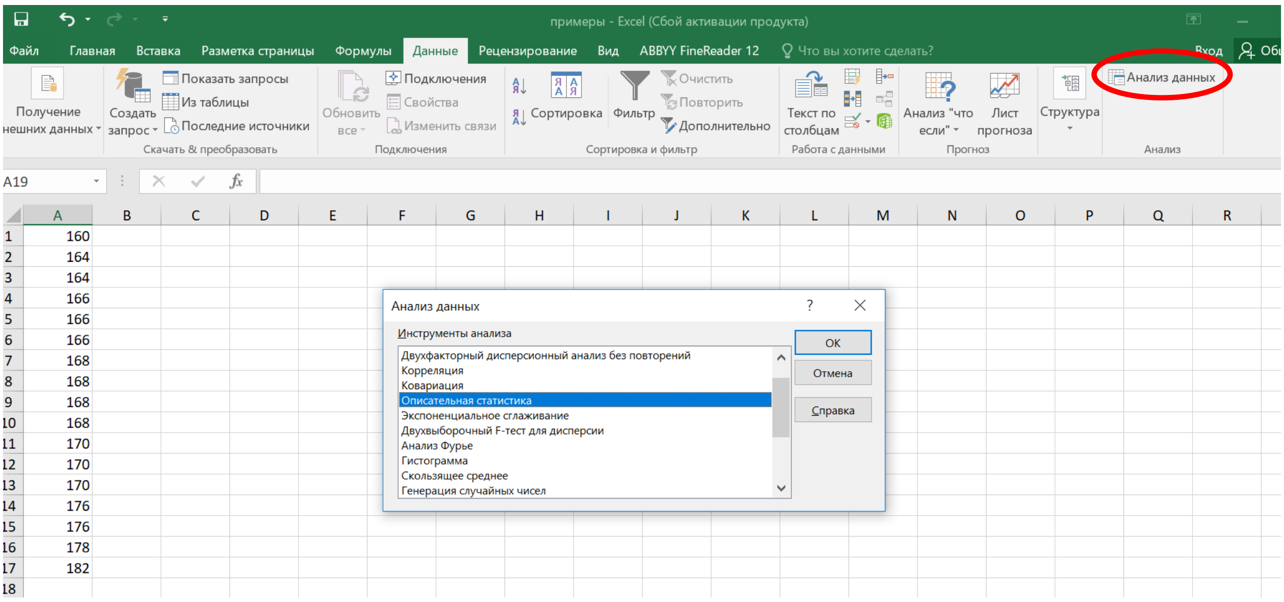


Рис. 27. Выбор «Описательная статистика» в Excel

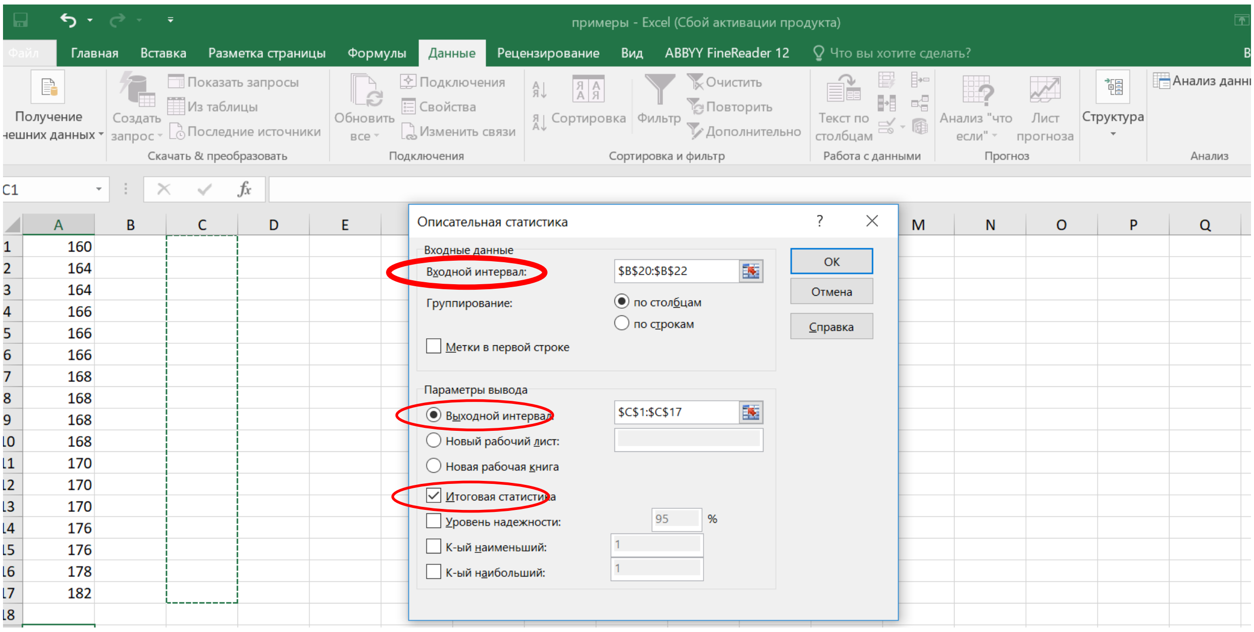


Рис. 28. Выбор интервала данных для подсчета ошибки репрезентативности

	A	B	C	D	E
1	160		*Столбец1*		
2	164				
3	164		Среднее	169,4118	
4	166		Стандартная ошибка	1,369464	
5	166		Медиана	168	
6	166		Мода	168	
7	168		Стандартное отклонение	5,646446	
8	168		Дисперсия выборки	31,88235	
9	168		Эксцесс	0,312231	
10	168		Асимметричность	0,755333	
11	170		Интервал	22	
12	170		Минимум	160	
13	170		Максимум	182	
14	176		Сумма	2880	
15	176		Счет	17	
16	178				
17	182				

Рис. 29. Стандартная ошибка (ошибка репрезентативности), вычисленная в Excel

Статистические выводы в психологии основываются на вероятностных заключениях, но надо понимать, что вероятностно оценивается не сформулированная гипотеза в отношении зависимых и независимых переменных, влияние дополнительных переменных, а достоверность установления этой зависимости. Дополнительные переменные ограничивают распространённость выводов, но влияние их статистически не просчитывается.

3.3. Зависимые и независимые выборки

Часто экспериментатор для доказательства своей гипотезы сравнивает две или более выборки. Эти выборки могут отличаться в зависимости от соотношения участников.

Независимые выборки — это разные группы людей, вероятность попадания в одну из групп никак не зависит от другой группы, независимый отбор.

Зависимая выборка, напротив, характеризуется тем, что каждому участнику в одной выборке соответствует участник другой выборки, попарный отбор.

Лонгитюдное исследование — типичная зависимая выборка, где каждый участник первого исследования соответствует самому себе, но через определенное время. Сравнение результатов до и после проведения тренинга также является зависимой выборкой. При сравнении человека с самим собой мы практически исключаем влияние неизвестных нам дополнительных переменных, таких как состояние здоровья, семейное положение, количество детей и др. В меньшей степени зависимости обладают выборки родители – дети, сиблинги, мужья – жены (при изучении предпочтений). Количество участников выборок всегда одинаковое ($n_1 = n_2$) (см. рис. 6).

В независимых выборках мы сравниваем все характеристики одной группы, при нормальном распределении — средние значения, при ненормальном — ранговые значения, с характеристиками другой группы, часто в психологии такие группы называются экспериментальная и контрольная. Объем выборки в таких группах может немного различаться.

Необходимо помнить, что частично зависимые или частично независимые выборки нельзя применять в психологическом исследовании, так как они непредсказуемо нарушают репрезентативность.

В таблице 5 представлено классическое исследование воздействия тренинга, вместо выражений «до проведения тренинга» и «после тренинга» должны стоять показатели, на которые, по мнению экспериментатора, воздействует тренинг.

Таблица 5. Пример зависимых и независимых выборок

Экспериментальная группа (ЭГ)	*Контрольная группа (КГ)*
1. ЭГ до проведения тренинга	2. КГ до проведения тренинга
Тренинг	
3. ЭГ после тренинга	4. КГ после тренинга

Для начала следует сравнить 1 и 2 группы, чтобы убедиться в том, что выборки не отличаются по дополнительным переменным (пол, возраст, опыт, условия и др.). При сравнении 1 и 2 групп мы должны получить $p > 0{,}05$, то есть наши группы не отличаются по сравниваемым показателям. Это будет сравнение независимых выборок.

Затем следует убедиться в эффективности тренинга, для этого сравниваем 1 и 3 группы, это — зависимые выборки, потому что необходимо сравнить каждого человека с самим собой до и после тренинга. Такое исследование называется *оценка достоверности сдвига*.

Следующим этапом необходимо доказать, что сдвиг произошел именно в результате тренинга, а не в результате воздействия дополнительных, неконтролируемых переменных (например, время). Для этого необходимо сравнить группы 3 и 4, и если получим $p \leq 0{,}05$, то это будет одним из доказательств эффективности тренинга.

Также можно сравнить группы 2 и 4 для доказательства отсутствия воздействия дополнительных переменных, при $p > 0{,}05$, группы будут одинаковыми по изучаемым психологическим характеристикам, до и после тренинга, хотя с ними тренинг не проводился, то есть неконтролируемых изменений не произошло.

Оценка достоверности различий и оценка достоверности сдвигов проводится с помощью особых статистических критериев, о которых будет говориться в следующей теме.

Выделяют две методологии психологического исследования:

— R-методология, которая подразумевает исследование изменчивости какого-либо качества или свойства, связанного с воздействием фактора или другого свойства, выборка состоит из испытуемых (например, факторный анализ);

— Q-методология, которая изучает изменения, происходящие внутри или с испытуемым, в зависимости от стимульного воздействия (условия, интенсивность и др.), выборка состоит из стимулов.

3.4. Степени свободы

В первом учебном вопросе данной темы было рассмотрено понятие статистической значимости, или р-уровень. При подсчете эмпирических значений с помощью различных критериев необходимо знать критические значения этих критериев, часть из которых также зависит и от размера выборки (n) и показателей степени свободы (*df* или ν).

Степень свободы — это показатель, зависящий от объема выборки, число, показывающее количество значений, которые могут быть случайными.

КРИТИЧЕСКИЕ ЗНАЧЕНИЯ КРИТЕРИЯ *t*-СТЬЮДЕНТА (для проверки ненаправленных альтернатив — двусторонний критерий)

df	*p*				*df*	*p*			
	0,10	0,05	0,01	0,001		0,10	0,05	0,01	0,001
1	6,314	12,70	63,65	636,61	46	1,679	2,013	2,687	3,515
2	2,920	4,303	9,925	31,602	47	1,678	2,012	2,685	3,510
3	2,353	3,182	5,841	12,923	48	1,677	2,011	2,682	3,505
4	2,132	2,776	4,604	8,610	49	1,677	2,010	2,680	3,500
5	2,015	2,571	4,032	6,869	50	1,676	2,009	2,678	3,496
6	1,943	2,447	3,707	5,959	51	1,675	2,008	2,676	3,492
7	1,895	2,365	3,499	5,408	52	1,675	2,007	2,674	3,488
8	1,860	2,306	3,355	5,041	53	1,674	2,006	2,672	3,484
9	1,833	2,262	3,250	4,781	54	1,674	2,005	2,670	3,480
10	1,812	2,228	3,169	4,587	55	1,673	2,004	2,668	3,476
11	1,796	2,201	3,106	4,437	56	1,673	2,003	2,667	3,473
12	1,782	2,179	3,055	4,318	57	1,672	2,002	2,665	3,470
13	1,771	2,160	3,012	4,221	58	1,672	2,002	2,663	3,466
14	1,761	2,145	2,977	4,140	59	1,671	2,001	2,662	3,463
15	1,753	2,131	2,947	4,073	60	1,671	2,000	2,660	3,460
16	1,746	2,120	2,921	4,015	61	1,670	2,000	2,659	3,457
17	1,740	2,110	2,898	3,965	62	1,670	1,999	2,657	3,454
18	1,734	2,101	2,878	3,922	63	1,669	1,998	2,656	3,452
19	1,729	2,093	2,861	3,883	64	1,669	1,998	2,655	3,449
20	1,725	2,086	2,845	3,850	65	1,669	1,997	2,654	3,447
21	1,721	2,080	2,831	3,819	66	1,668	1,997	2,652	3,444
22	1,717	2,074	2,819	3,792	67	1,668	1,996	2,651	3,442
df	*p*				*df*	*p*			
	0,10	0,05	0,01	0,001		0,10	0,05	0,01	0,001
29	1,699	2,045	2,756	3,659	74	1,666	1,993	2,644	3,427
30	1,697	2,042	2,750	3,646	75	1,665	1,992	2,643	3,425
31	1,696	2,040	2,744	3,633	76	1,665	1,992	2,642	3,423
32	1,694	2,037	2,738	3,622	78	1,665	1,991	2,640	3,420
33	1,692	2,035	2,733	3,611	79	1,664	1,990	2,639	3,418
34	1,691	2,032	2,728	3,601	80	1,664	1,990	2,639	3,416
35	1,690	2,030	2,724	3,591	90	1,662	1,987	2,632	3,402
36	1,688	2,028	2,719	3,582	100	1,660	1,984	2,626	3,390
37	1,687	2,026	2,715	3,574	110	1,659	1,982	2,621	3,381
38	1,686	2,024	2,712	3,566	120	1,658	1,980	2,617	3,373
39	1,685	2,023	2,708	3,558	130	1,657	1,978	2,614	3,367
40	1,684	2,021	2,704	3,551	140	1,656	1,977	2,611	3,361

Рис. 30. Пример таблицы критических значений по А. Д. Наследову[1]

[1] Наследов А. Д. Указ. соч. С. 355.

На рисунке 30 показан пример таблицы критических значений для критерия t-Стьюдента двух независимых выборок, в первом столбике необходимо найти степень свободы для конкретной выборки, а затем в строчке искать эмпирическое значение для выявления уровня статистической значимости.

Если в исследовании две независимые выборки, то число степеней свободы можно рассчитать по формуле:

для первой выборки $df_1 = n_1 - 1$;

для второй выборки $df_2 = n_2 - 1$;

для двух исследуемых выборок $df = (n_1 - 1) + (n_2 - 1) = (n_1 + n_2) - 2$.

Например, мы хотим сравнить группу слушателей и курсантов мужского пола, чтобы узнать, различается ли рост российских курсантов и иностранных слушателей. Для иностранных слушателей возьмем пример, рассмотренный раньше, где $n = 17$, $\sigma = 5{,}65$, $\bar{x} = 169{,}41$; у курсантов $n = 20$, $\sigma = 5{,}65$, $\bar{x} = 177{,}2$.

Формула для расчета критерия t-Стьюдента:

$$M_x \quad -$$
$$M_y \quad \overline{\sqrt{\_\_ + \_\_}}$$
$$t = \qquad \sigma_x^2$$

$t = (169{,}41 - 177{,}2) / \sqrt{(5{,}65^2/17) + (5{,}65^2/20)} =$

$= (-0{,}79)/\sqrt{(31{,}92/17) + 1{,}596} = -0{,}79/\sqrt{1{,}88+1{,}596} = -0{,}79/1{,}864 = -0{,}42$

$df = (17+20) - 2 = 35$

$t_{кр} > t_{эмп} \Longrightarrow H_0$ $(p > 0{,}05)$, различия статистически не значимы.

Рост курсантов и слушателей не отличается.

Математические методы систематизации, обработки и применения статистических данных и их выводы для доказательства научности проводимого исследования, называются *математическим моделированием.*

Вопросы и задания для самоконтроля

1. Перечислите показатели, от которых зависит статистическая значимость.
2. Дайте определение р-уровню.
3. Сравните объем выборки при разработке или адаптации новой методики и изучения стрессоустойчивости.
4. Сравните научную и статистическую гипотезы.
5. Назовите отличия зависимой и независимой переменной.
6. Определите понятие альтернативная гипотеза.
7. Сравните направленную и ненаправленную гипотезы, приведите примеры.
8. Сравните мощность статистического критерия и уровень значимости.
9. Дайте определение доверительному интервалу.
10. Определите понятие степень свободы.

Глава 4. СРАВНЕНИЯ ВЫБОРОК. ВИДЫ СТАТИСТИЧЕСКИХ КРИТЕРИЕВ

Учебные вопросы:

4.1. Шкалы измерения.

4.2. Параметрические и непараметрические критерии.

4.3. Сравнительный анализ и его виды.

Для того чтобы адекватно описать данные, полученные в эмпирическом исследовании, необходимо понимать тип данных и их распределение. Рассмотрим наиболее популярные методы сравнения выборок.

В любом исследовании первоначально необходимо определить, какой тип данных будет важен для доказательства экспериментальной гипотезы.

Выделяют два типа данных — количественные и качественные. Количественные шкалы можно определить по тому, что они имеют реальное цифровое выражение (например, рост, вес, средний балл успеваемости, количество пропущенных по болезни дней и др.). Качественные шкалы дают характеристику свойств объекта, (например, пол, национальность, семейное положение и др.).

Для интерпретации количественных данных используют более мощные статистические критерии, то есть таким данным ученые доверяют больше. С другой стороны, в любом исследовании невозможно обойтись без качественных данных, они являются одним из доказательств репрезентативности выборки.

4.1. Шкалы измерения

Любое измерение предполагает наличие единиц измерения. Впервые об измерении в психологии заговорил Густав Теодор Фехнер (1801–1887) в «Элементах психофизики». С тех пор, конечно, многое изменилось. Совершенствуются измерительные шкалы, психологические измерения становятся более надежными и валидными.

Способ представления данных и способ их группировки называется шкалой измерения. Шкалирование — упорядочивание изменений. Шкала измерения определяется двумя основными параметрами — способом сбора и объединения данных, и применяется для описания переменных. От нее зависит метод, который будет использоваться для анализа данных. Шкалы измерений (они также называются шкалами С. Стивенса) делятся на четыре основных: номинативная, порядковая (ранговая), относительная (интервальная), абсолютная (шкала отношений).

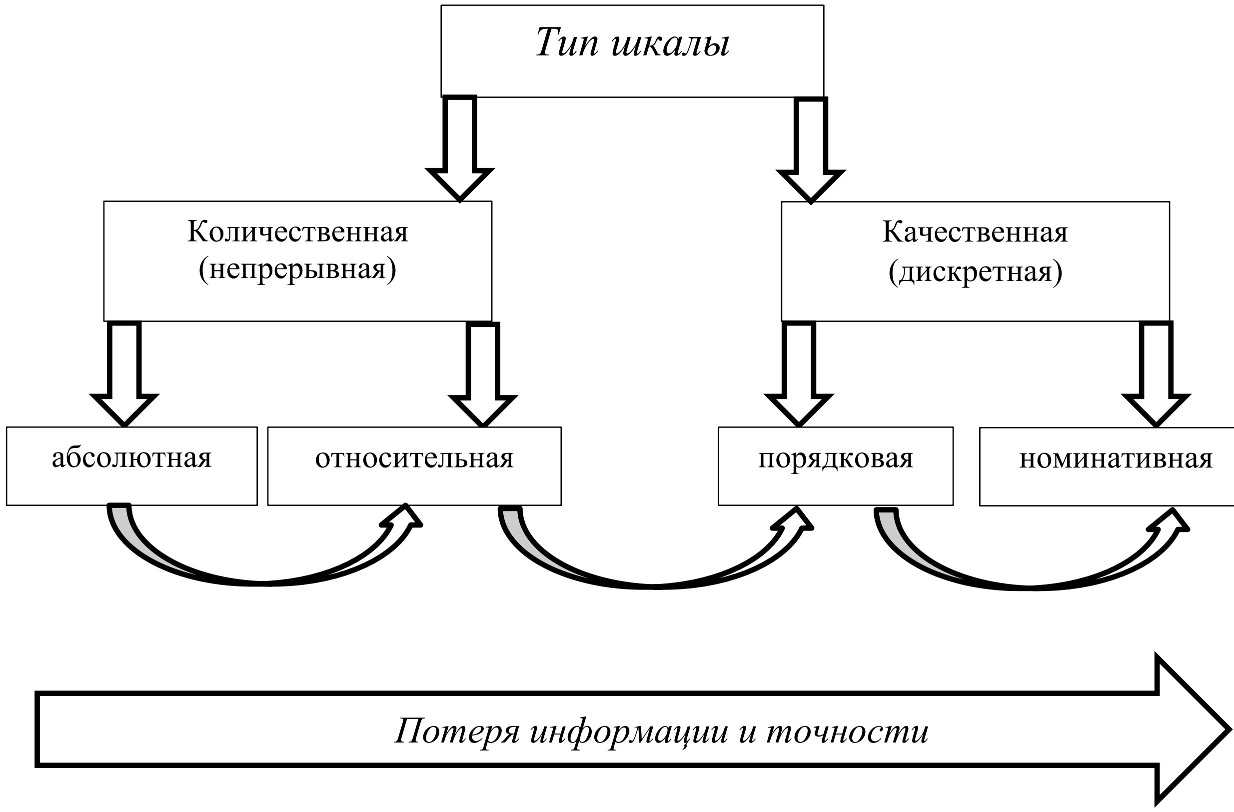


Рис. 31. Типы измерительных шкал

Шкалы измерения можно разделить в зависимости от наличия в них свойств, причем наличие последних означает обязательное наличие предыдущих (табл. 6):

1. Описание или идентификация — позволяют описать полученные в результате исследования данные и присвоить им цифровые значения. (Пример: при делении выборки в зависимости от пола, мужчинам присваивается цифра 0, женщинам — 1, это присвоение дает возможность включить в статистические измерения (внести в Excel) данный параметр.)

2. Порядок или ранг — упорядочивание значений от наименьшего к наибольшему, или, наоборот, в зависимости от типа данных. (Пример: можно упорядочить выборку в зависимости от коэффициента IQ, в данном случае, скорее всего, получится выстроить значения от наибольшего к наименьшему (132, 130, 123, 120, 118, 117, 116, 115).

3. Равенство интервалов или расстояние, показывает, что между соседними значениями обязательно одинаковый интервал. Ранг, или место на соревнованиях, не может гарантировать, что занявший второе место отстал ровно на 10 секунд от первого, а занявший третье место отстал ровно на 10 секунд от второго и т. д. Равенство интервалов гарантирует именно такую точность (например, можно разделить

возраст обучающихся иностранных слушателей на равные интервалы 15–20, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49 лет). Сравнение показателей по этой шкале возможно с формулировкой «первый элемент отличается от второго на пять единиц».

4. Абсолютный ноль или наличие начальной точки — такие шкалы редко используются в психологических исследованиях. Абсолютный ноль показывает начало отсчета признака от точки, в которой этот признак отсутствует. Сравнение показателей по этой шкале возможно с формулировкой «первый элемент отличается от второго в три раза».

Таблица 6. Характеристика измерительных шкал

Типы шкал	*Характеристики шкал*			
	описание	*порядок*	*равенство интервалов*	*абсолютный ноль*
Номинативная	пол: 1— мужчина 2— женщина	–	–	–
Порядковая	мужчины 1 — 5 побед женщины 2 — 3 победы	место на соревнованиях	–	–
Относительная	рост: высокие — 180–170 низкие — 169–160	рост: 175–180 — 1 170–174 — 2 165–169 — 3 160–164 — 4	рост: 160–164 165–169 170–174 175–180	–
Абсолютная	рост: высокие — 180–170 низкие — 169–160	рост: 175–180 — 1 170–174 — 2 165–169 — 3 160–164 — 4	рост: 175–180 170–174 165–169 160–164	рост: 160, 164, 166, 168, 170, 176

Абсолютная шкала или *шкала отношений* — непрерывная метрическая шкала, есть точка отсчета, нулевая точка (например, 0 сек. времени; 0 *кг* веса; 0 *см* роста и т. д., можно сказать, что один человек выше другого в 1,2 раза, или один тяжелее другого в 1,2 раза, один испытуемый решил задачу в 2 раза быстрее, чем другой испытуемый). В психологических исследованиях абсолютная шкала редко используется, мы не может сказать, что у человека IQ равен 0, при дебильности IQ = 80. С абсолютной шкалой возможны все описанные выше статистические действия (показатели центральной тенденции, меры изменчивости и др.). Значения абсолютной шкалы можно перевести в любую другую.

Относительная или *интервальная шкала* — дискретная, метрическая, в ней значения упорядочены, между значениями выделены равные интервалы, отражает, насколько одно значение отличается от другого (например, интеллект обучающихся в вузах на 20 баллов выше тех, кто имеет среднее образование, или один испытуемый решил задачу на 6 секунд быстрее). Относительная шкала также позволяет вычислять меры центральной тенденции и меры изменчивости. Немного изменится, по сравнению с абсолютной шкалой, формулировка выводов. Значения относительной шкалы можно перевести в любую шкалу, кроме абсолютной.

Порядковая или *ранговая шкала* основана на упорядочивании значений переменной в порядке убывания или возрастания. Значения ранговой шкалы можно перевести только в номинативную шкалу.

Рассмотрим способы перевода значений в ранговую шкалу. Рост слушателей: 160, 164, 164, 164, 166, 168, 170, 170, 170, 176, 176 *см*. Выстроим ряд в порядке убывания и присвоим каждому значению порядковый номер. При построении ранговой шкалы необходимо, чтобы одинаковые значения имели одинаковый ранг (на соревнованиях при одинаковом времени дают одно место). Чтобы выполнить это условие, необходимо у одинаковых значений найти среднее значение их порядковых номеров. Например, рост 170 имеют три слушателя, при этом они имеют 3, 4 и 5 порядковые номера. У них будет одинаковый ранг и, чтобы его найти, надо сложить их порядковые номера и разделить на 3, так как у нас 3 значения: (3+4+5)/3 = 4. Следовательно, их ранг будет 4. Для проверки правильности присвоения рангов сумма порядковых номеров и сумма рангов должны совпасть $\sum_1 = \sum_2$ (табл. 7).

Перевести значения в ранговые можно также в программе Excel (рис. 32). Для этого необходимо выстроить указанные значения, если необходимо, то упорядочить, затем выбрать функцию РАНГ, в ячейке «число» задать значение, которому необходимо присвоить ранг, в ячейке «ссылка» задать весь ряд значений, в ячейке «порядок» указать значение 0 или 1, 0 — по убыванию, 1 — по возрастанию.

Таблица 7. Перевод значений в ранговые

Рост	*Порядковый номер*	*Ранг*
176	1	1,5
176	2	1,5
170	3	4
170	4	4
170	5	4
168	6	6
166	7	7
164	8	9
164	9	9
164	10	9
160	11	11
	$\sum_1 = 66$	$\sum_2 = 66$

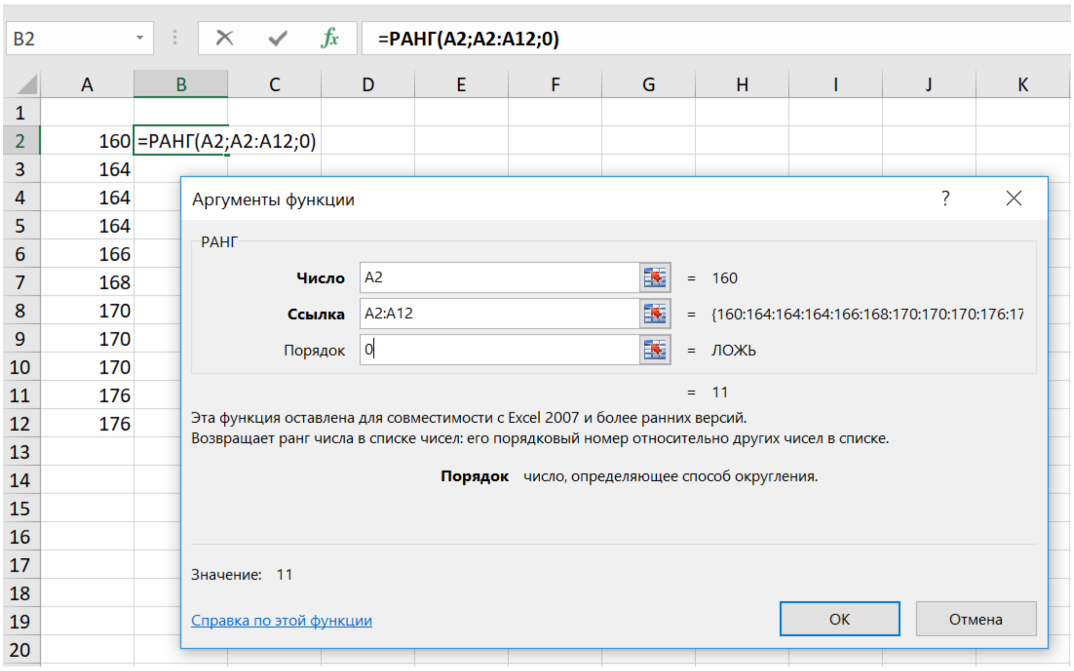


Рис. 32. Присвоение рангов в программе Excel

Номинативная или *шкала наименований*, неметрическая, состоит из однородных групп, сформированных по определенному признаку. Группы между собой нельзя сравнить, можно только их назвать. Например, мужчины/женщины — мы не можем сравнить, сколько в мужчинах женского и сколько в женщинах мужского, можно их только назвать. Или национальность, цвет волос, цвет глаз, какие-то предпочтения — одним нравится кофе, другим чай и т. д. В этом случае учитывается только один показатель. Одним из свойств этой шкалы является произвольное наименование показателей, а также возможность их переименования. Номинативную шкалу нельзя перевести ни в какую другую, это самое «грубое» деление выборки.

На рисунке 31 большая стрелка внизу показывает направление снижения мощности шкал — самой мощной является абсолютная шкала, а минимальной мощностью обладает номинативная шкала. Наиболее мощные шкалы отражают максимальную информацию при сравнении выборок. При прочих равных возможностях следует выбирать максимально мощную для данной выборки шкалу.

4.2. Параметрические и непараметрические критерии

Мы уже разобрали понятие дискретного (точечного) и непрерывного распределения. К дискретным распределениям относится, в частности, распределение Бернулли. Например, если мы знаем, что человек говорит 3 раза ложь и 7 раз правду, какова вероятность, что при разговоре будет ложь? Эту вероятность можно посчитать: 3/10, потому что всего 10 вариантов.

Среди непрерывных распределений нас больше всего интересует нормальное распределение, или кривая Гаусса. При таком распределении параметры выборки по своим значениям приближаются к параметрам генеральной совокупности. Также нормальное распределение обладает определенными свойствами, описанными выше (меры центральной тенденции, меры изменчивости).

После выбора шкалы необходимо определить, является ли представленное распределение вероятностей нормальным. Таким оно может быть только в относительной или абсолютной шкале. При нормальном распределении следует применять параметрические критерии, при отсутствии доказательств нормального распределения оно считается ненормально распределенным, и тогда применяются непараметрические критерии. Параметрические критерии основываются на метрических данных сравнения средних значений, квадратного отклонения и других параметров нормального распределения. Непараметрические критерии включают в свои формулы частоты и ранги. Непараметрическими критериями пользуются для сравнения данных, представленных в номинативной или ранговой шкалах, а также в метрических шкалах при распределении, отличном от нормального.

Различия параметрических и непараметрических критериев также касаются задач, которые они решают.

На рисунке представлен алгоритм выбора статистического критерия для анализа количественных данных, в основе лежит показатель нормальности распределения.

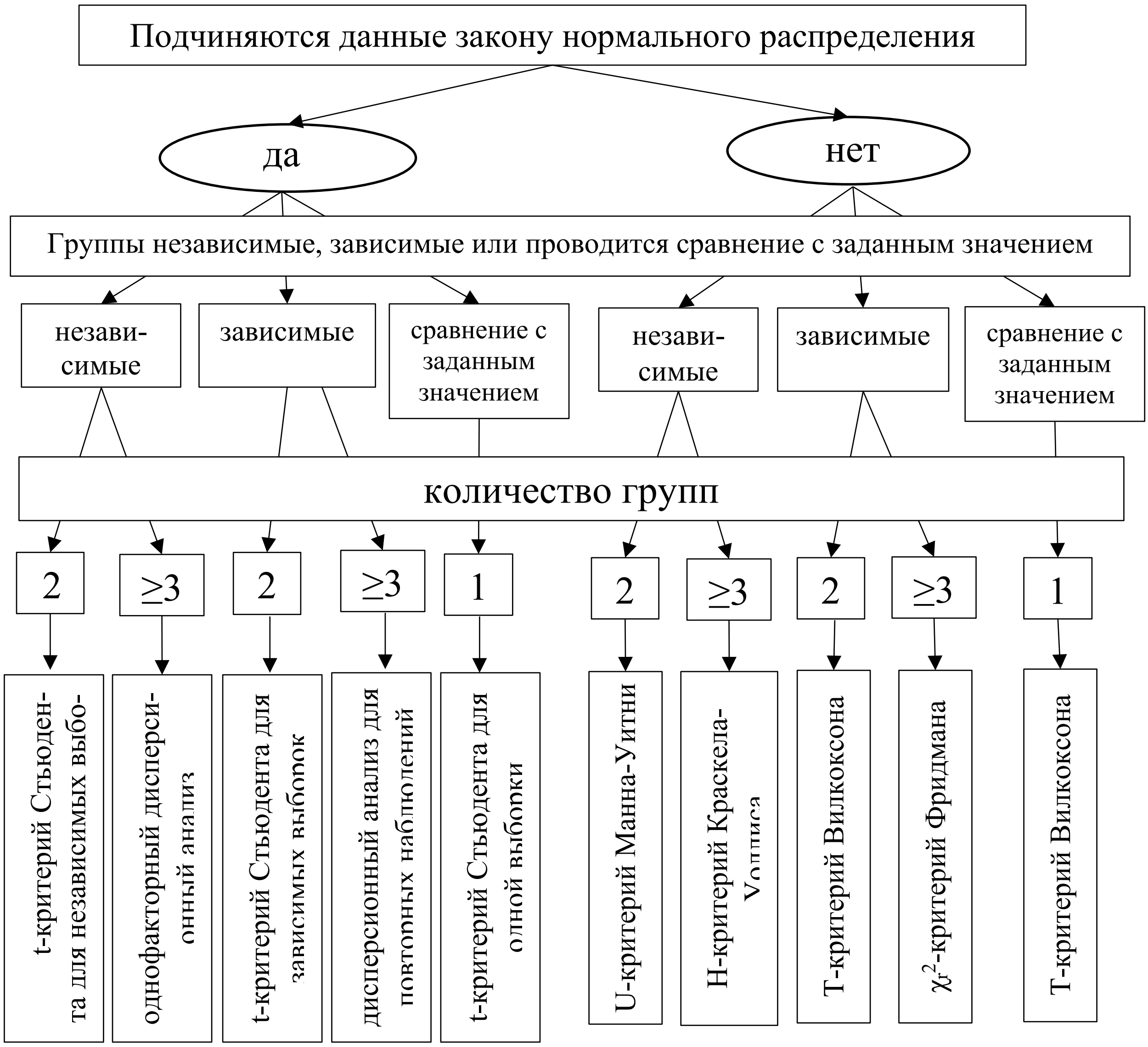


Рис. 33. Алгоритм выбора статистического критерия для анализа количественных данных, по А.М. Гржибовскому[1]

Для проверки соответствия распределения нормальному используют критерии Колмогорова-Смирнова (Kolmogorov–Smirnov test, K–S test, KS test) для больших выборок и критерий Шапиро-Уилка (W) для $8 < n \leq 50$. За нулевую принимается гипотеза, что эмпирическое распределение не отличается от нормального, то есть при $p > 0{,}05$ принимается H_0, при $p \leq 0{,}05$ эмпирическое распределение отличается

[1] Унгуряну Т. Н., Гржибовский А. М. Краткие рекомендации по описанию, статистическому анализу и представлению данных в научных публикациях // Экология человека. — 2011. — № 5. — С. 55–60

от нормального, принимается H_1. Тип распределения проверяется для каждой выборки отдельно.

На рисунке 34 представлен алгоритм выбора статистического критерия, характеризующего свойства объекта, то есть качественные признаки, закодированные цифрами.

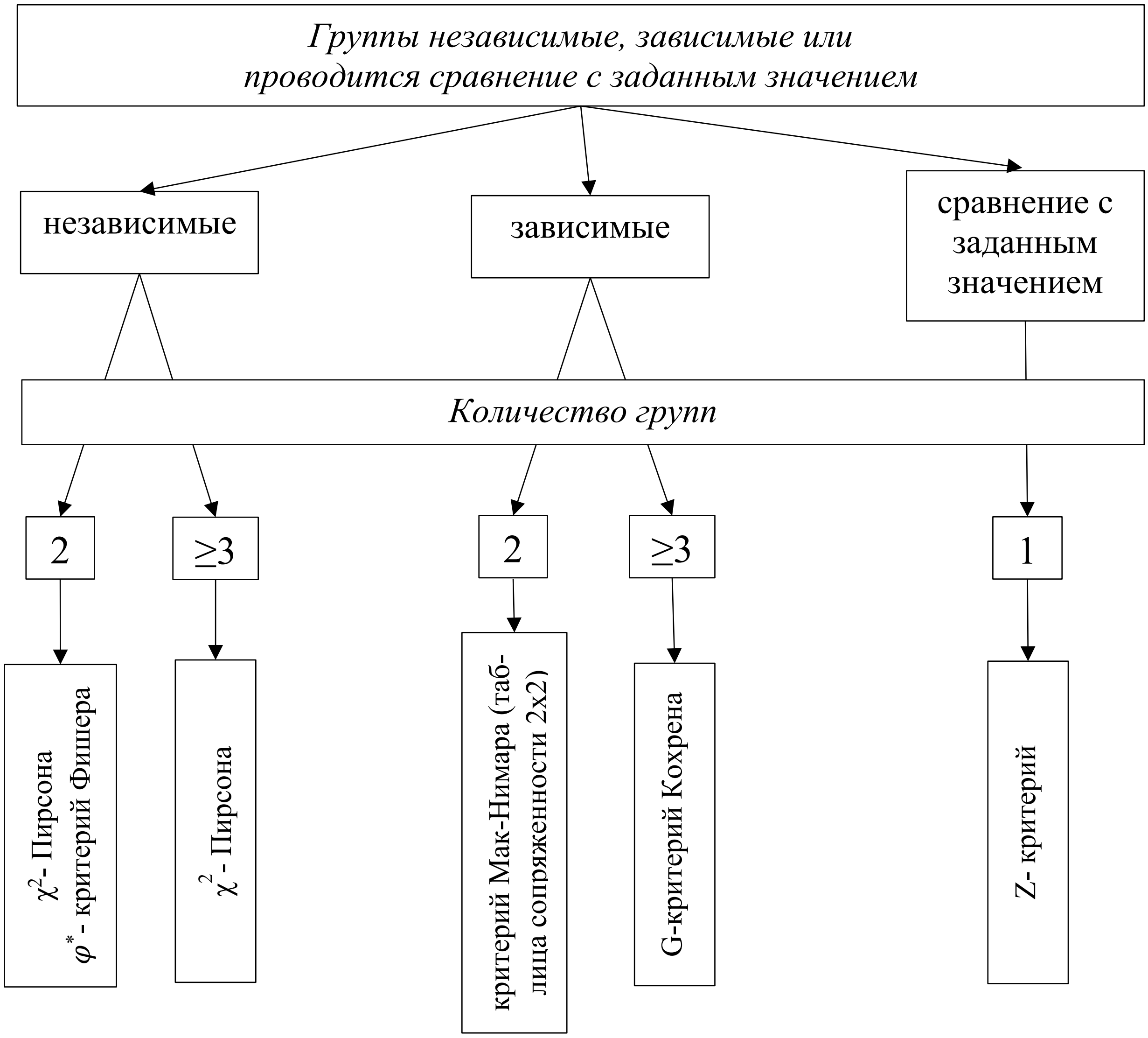


Рис. 34. Алгоритм выбора статистического критерия для анализа качественных данных, по А. М. Гржибовскому

При расчете любых данных желательно высчитывать 95 % доверительный интервал.

При анализе качественных данных рассчитываются процентное соотношение долей и частоты встречаемости качественных параметров в выборке.

При выборе статистического критерия следующим этапом, после определения типа распределения, необходимо определить количество и тип выборок — зависимые или независимые.

Теоретическим называется всем известное распределение генеральной совокупности (например, нормальное, при котором, при известных нескольких показателях, можно рассчитать остальные неизвестные).

Наиболее часто психолог сталкивается с задачей сравнения групп либо в разных условиях, либо до и после воздействия, либо экспериментальной и контрольной группы. Иногда важно выявить динамику изменений или характер изменений психологического параметра в разных группах или в разное время. При решении такого рода задач используются критерии различия, которые позволяют выявить статистически достоверные различия при заданном уровне достоверности.

Существует множество критериев различий, каждый имеет свои особенности и отличия от других. Также эти критерии отличаются по мощности. Множество критериев различий помогает выбирать правильный из них:

— адекватный типу шкалы экспериментальных данных;
— в зависимости от типа выборки связанная/несвязанная;
— позволяющий выявить отличия в разных по объему выборках;
— в зависимости от его мощности.

При выборе критерия имеет значение однородность выборки, которая позволяет распространить выводы на ту или иную генеральную совокупность. Например, при исследовании курсантов, полученные выводы мы не можем распространить на всех студентов, на всех взрослых людей, на всех молодых людей и т. д.

Также выбор критерия зависит от объема выборки, при этом непараметрические критерии позволяют выявлять различия в небольших по объему выборках.

Если для выбранного исследования подходят несколько критериев, то стоит выбрать наиболее мощный.

4.3.Сравнительный анализ и его виды

Попробуем рассмотреть критерии, представленные на рисунках 33 и 34. Начнем с самого простого и самого распространённого в исследованиях двух независимых групп.

t-критерий Стьюдента для двух независимых выборок при нормальном распределении позволяет оценить достоверность различий средних значений. Необходимо также знать среднеквадратичное отклонение и объем выборок, который может быть разным. Формула t-критерия Стьюдента имеет следующий вид:

$$t = \frac{|M_x - M_y|}{\sqrt{\frac{\sigma_x^{2}}{n_x} + \frac{\sigma_y^{2}}{n_y}}}$$

Разберем пример. Допустим, две группы слушателей изучали один и тот же предмет у разных преподавателей. Мы хотим узнать, насколько различаются оценки на экзамене в этих группах. Начальный уровень подготовки был одинаковый, то есть этот предмет слушатели еще не изучали. С помощью таблицы Excel были посчитаны следующие значения (рис. 35).

	группа 1	*группа 2*
n	20	24
M	4,05	4,29
σ	0,69	0,62

	A	B	C	D	E	F	G	H	I	J
1	4		4			*Столбец1*			*Столбец2*	
2	3		4							
3	3		5			Среднее	4,05		Среднее	4,29167
4	4		5			Стандарт	0,15347		Стандарт	0,12739
5	4		3			Медиана	4		Медиана	4
6	5		5			Мода	4		Мода	4
7	5		5			Стандарт	0,68633		Стандарт	0,62409
8	5		4			Дисперси	0,47105		Дисперси	0,38949
9	4		4			Эксцесс	-0,63043		Эксцесс	-0,48464
10	4		5			Асимметр	-0,0624		Асимметр	-0,27981
11	4		5			Интервал	2		Интервал	2
12	3		5			Минимум	3		Минимум	3
13	3		4			Максиму	5		Максиму	5
14	5		4			Сумма	81		Сумма	103
15	5		3			Счет	20		Счет	24
16	4		4							
17	4		4							
18	4		4							
19	4		4							
20	4		5							
21			5							
22			4							
23			4							
24			4							

Рис. 35. Сравнение оценок, полученных на экзамене у двух независимых выборок

$$t = (4{,}05 — 4{,}29) / \sqrt{(0{,}69^2/20) + (0{,}62^2/24)} \approx 0{,}23$$

Формулируем гипотезы:

H_0 — достоверность различий оценок на экзамене у групп разных преподавателей значимо не отличается от нуля;

H_1 — достоверность различий оценок на экзамене у групп разных преподавателей значимо отличается от нуля.

Вычисляем степени свободы $df = n1 + n2 – 2$, $df = 20 + 24 – 2 = 42$. По таблице критических значений (рис. 36) видим, что полученное значение 0,23 < 1,682, то есть p > 0,10.

$$t_{эмп} < t_{кр}\ (p>0{,}05) \Longrightarrow H_0$$

оценки на экзамене у групп разных преподавателей не отличаются.

Значение t-критерия Стьюдента также можно рассчитать с помощью программы Excel (рис. 37). Массив 1 и 2 — это значения двух групп. Хвосты показывают, есть ли направления в различиях, (в нашем случае это не важно, нам надо узнать отличаются ли группы между собой), двусторонняя гипотеза (2), а если нашей задачей будет доказать, что оценки второй группы значимо выше первой, то это будет односторонняя гипотеза, и в хвосте выбираем 1. Тип 1 — зависимые выборки, 2 — независимые выборки. Затем полученное значение все равно следует сравнивать с критическим по таблице.

df	*p*				*df*	*p*			
	0,10	0,05	0,01	0,001		0,10	0,05	0,01	0,001
29	1,699	2,045	2,756	3,659	74	1,666	1,993	2,644	3,427
30	1,697	2,042	2,750	3,646	75	1,665	1,992	2,643	3,425
31	1,696	2,040	2,744	3,633	76	1,665	1,992	2,642	3,423
32	1,694	2,037	2,738	3,622	78	1,665	1,991	2,640	3,420
33	1,692	2,035	2,733	3,611	79	1,664	1,990	2,639	3,418
34	1,691	2,032	2,728	3,601	80	1,664	1,990	2,639	3,416
35	1,690	2,030	2,724	3,591	90	1,662	1,987	2,632	3,402
36	1,688	2,028	2,719	3,582	100	1,660	1,984	2,626	3,390
37	1,687	2,026	2,715	3,574	110	1,659	1,982	2,621	3,381
38	1,686	2,024	2,712	3,566	120	1,658	1,980	2,617	3,373
39	1,685	2,023	2,708	3,558	130	1,657	1,978	2,614	3,367
40	1,684	2,021	2,704	3,551	140	1,656	1,977	2,611	3,361
41	1,683	2,020	2,701	3,544	150	1,655	1,976	2,609	3,357
42	1,682	2,018	2,698	3,538	200	1,653	1,972	2,601	3,340
43	1,681	2,017	2,695	3,532	250	1,651	1,969	2,596	3,330
44	1,680	2,015	2,692	3,526	300	1,650	1,968	2,592	3,323
45	1,679	2,014	2,690	3,520	350	1,649	1,967	2,590	3,319

Рис. 36. Таблица критических значений t-критерия Стьюдента[1]

[1] Наследов А. Д. Указ. соч. С. 356.

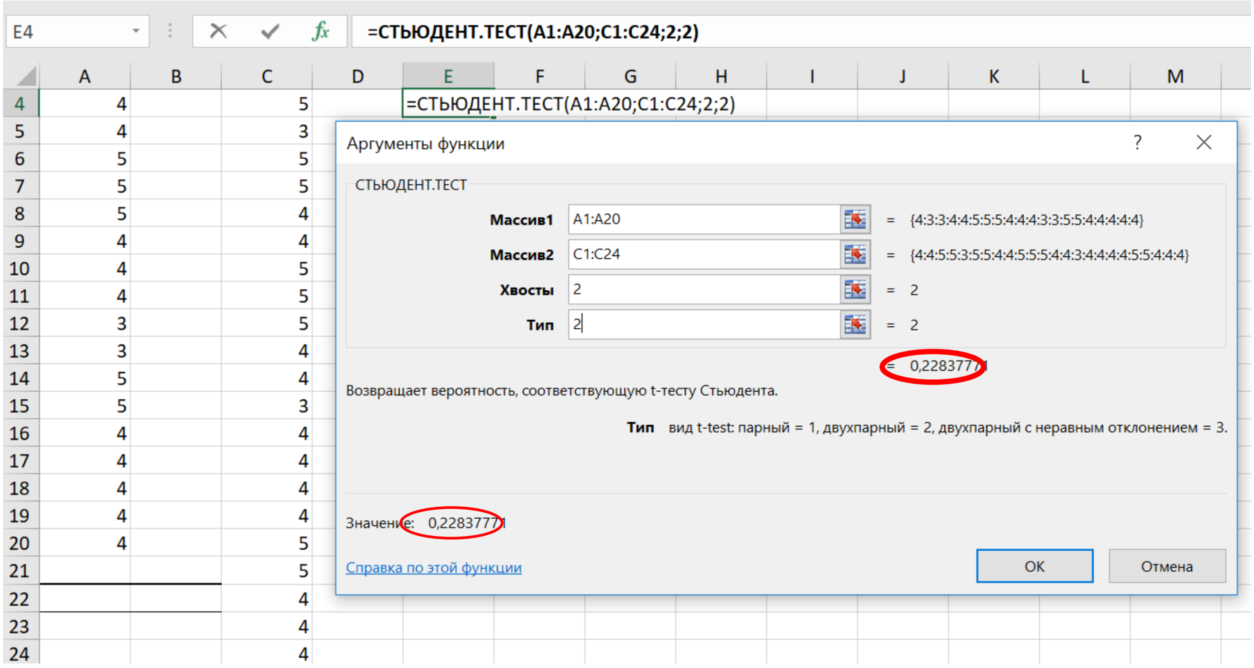


Рис. 37. Нахождение значений t-критерия Стьюдента в Excel

t-критерий Стьюдента для двух зависимых выборок при нормальном распределении позволяет оценить достоверность сдвига значений, имеет вид:

$$t = \left| \frac{M_d}{\sigma_d / \sqrt{n}} \right|$$

где M_d — среднее разности значений,

σ_d — стандартное отклонение разности значений,

$df = n-1$.

Разберем пример. Допустим, мы хотим узнать, изменилась ли успеваемость слушателей в зимней и летней сессии (рис. 38).

H_0 — успеваемость слушателей по результатам зимней и летней сессий не отличается от нуля.

H_1 — успеваемость слушателей по результатам зимней и летней сессий значимо отличается от нуля.

Чтобы воспользоваться формулой t-критерия Стьюдента для зависимых выборок необходимо вычислить разницу значений по каждому слушателю в зимнюю и летнюю сессии, d=y-x, среднее значение разницы — это среднее арифметическое $M_d = \sum d/n$. Затем из каждого значения вычитаем M_d, возводим в квадрат и ищем сумму квадратов.

	A	B	C	D	E
1	зимняя, x	летняя, y	d=y-x	d-Md	
2	4	3	-1	-1,25	1,56
3	4	5	1	0,75	0,56
4	4	5	1	0,75	0,56
5	4	4	0	-0,25	0,06
6	4	4	0	-0,25	0,06
7	5	5	0	-0,25	0,06
8	5	5	0	-0,25	0,06
9	5	5	0	-0,25	0,06
10	4	4	0	-0,25	0,06
11	4	4	0	-0,25	0,06
12	4	3	-1	-1,25	1,56
13	3	4	1	0,75	0,56
14	3	4	1	0,75	0,56
15	4	4	0	-0,25	0,06
16	4	4	0	-0,25	0,06
17	4	5	1	0,75	0,56
18	4	5	1	0,75	0,56
19	4	4	0	-0,25	0,06
20	4	5	1	0,75	0,56
21	4	4	0	-0,25	0,06
22	*n=20*		5		7,7

Рис. 38. Нахождение значений t-критерия Стьюдента для зависимых выборок

$\sigma_d = \sqrt{(d\text{-}M_d)^2/n\text{-}1} = \sqrt{7{,}7/19} = 0{,}6364$

$t = 0{,}25 \ / \ 0{,}6364/\sqrt{19} = 1{,}71$

На рисунке 30 находим критические значения для *df* = 19, полученное значение 1,71 < 1,729, то есть p > 0,10.

$t_{эмп} < t_{кр}\ (p > 0{,}05) \Longrightarrow H_0$

Успеваемость на летней и зимней сессиях у слушателей не различается.

t-критерий Стьюдента для одной выборки позволяет сопоставить среднее значение переменных с неким известным значением (A). H_0: $M_x = A$; H_1: $M_x \neq A$.

Формула для вычисления *df* = n-1:

$$t = \frac{|M - A|}{\sigma/\sqrt{n}}$$

Например, мы хотим узнать, отличается ли успеваемость на 8 факультете в 2021 году, от среднего значения по университету (4,2) α = 0,05. Вычисления в данном случае приводить не будем.

Дисперсионный анализ разберем в следующей теме.

Рассмотрим несколько непараметрических критериев, часто встречающихся в психологических исследованиях.

U-критерий Манна-Уитни считается непараметрическим аналогом t-критерия Стьюдента, показывает достоверность различий двух независимых выборок, у него есть ограничения по объему выборки $3 \leq n \leq 60$.

При подсчете вручную обе выборки объединяются в одну группу, которая упорядочивается и ранжируется, затем выборки разъединяются и производятся действия с полученными рангами. Нулевая гипотеза показывает, что значения одной выборки распределяются среди значений другой выборки. Альтернативная — показывает, что в одной из выборок преобладают минимальные или максимальные значения.

Например, мы хотим сравнить рост курсантов мужского и женского пола.

Таблица 8. Показатели роста курсантов мужского и женского пола

мужчины	168	176	178	182	176	180	184	170	179
женщины	164	170	176	172	168	165	169	171	

рост	164	165	168	168	169	170	170	171	172	176	176	176	178	179	180	182	184
пол	ж	ж	ж	м	ж	ж	м	ж	ж	ж	м	м	м	м	м	м	м
ранг	17	16	14,5	14,5	13	11,5	11,5	10	9	7	7	7	5	4	3	2	1
м				14,5			11,5				7	7	5	4	3	2	1
ж	17	16	14,5		13	11,5		10	9	7							

Подсчитываем сумму рангов в зависимости от пола:

$$\sum_{м} = 55$$

$$\sum_{ж} = 98$$

Формула подсчета U-критерий Манна-Уитни:

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

где n_1 и n_2 — объем 1 и 2 выборки,

n_x — объем выборки с большей суммой рангов (в нашем примере $n_ж$),

T_x ранговая сумма, которая больше (98):

$$U = (9*8) + \frac{8(8+1)}{2} - 98 = 10,$$

По таблице критических значений[1], рисунок 39, находим объем выборок 9 и 8, при $p = 0,05$ $U = 15$; при $p = 0,01$ $U = 9$:

$$U_{эмп} < U_{кр}\ (p \leq 0,05) \Rightarrow H_1$$

Существуют достоверные ($p \leq 0,05$) различия в росте между курсантами мужского и женского пола.

КРИТИЧЕСКИЕ ЗНАЧЕНИЯ КРИТЕРИЯ *U*-МАННА-УИТНИ
(для проверки ненаправленных альтернатив)

$P = 0,05$

N_2	N_1													
	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	12	15	17	20	23	26	28	31	34	37	39	42	45	48
10	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14	22	26	31	36	40	45	50	55	59	64	67	74	78	83
15	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17	28	34	39	45	51	57	63	67	75	81	87	93	99	105
18	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	34	41	48	55	62	69	76	83	90	98	105	112	119	127

$P = 0,01$

N_2	N_1													
	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3			0	0	0	1	1	1	2	2	2	2	3	3
4	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	1	2	3	4	4	6	7	7	8	9	10	11	12	13
6	3	4	5	6	6	9	10	11	12	13	15	16	17	18
7	4	6	7	9	9	12	13	15	16	18	19	21	22	24
8	6	7	9	11	11	15	17	18	20	22	24	26	28	30
9	7	9	11	13	13	18	20	22	24	27	29	31	33	36
10	9	11	13	16	16	21	24	26	29	31	34	37	39	42
11	10	13	16	18	18	24	27	30	33	36	39	42	45	48
12	12	15	18	21	21	27	31	34	37	41	44	47	51	54
13	13	17	20	24	24	31	34	38	42	45	49	53	56	60
14	15	18	22	26	26	34	38	42	46	50	54	58	63	67
15	16	20	24	29	29	37	42	46	51	55	60	64	69	73
16	18	22	27	31	31	41	45	50	55	60	65	70	74	79
17	19	24	29	34	34	44	49	54	60	65	70	75	81	86
18	21	26	31	37	37	47	53	58	64	70	75	81	87	92
19	22	28	33	39	39	51	56	63	69	74	81	87	93	99
20	24	30	36	42	42	54	60	67	73	79	86	92	99	105

Рис. 39. Критические значения U-Манна-Уитни

При сравнении трех и более выборок для оценки различий в зависимости от уровня одного параметра используют непараметрический *Н-критерий Крускала-Уоллиса*, иногда его называют непараметрическим дисперсионным анализом. Он показывает, происходят ли изменения от выборки к выборке, не показывая при этом их направления.

У этого критерия есть ограничения: при трёх выборках есть ограничения по минимальному объему $n_1 = 3$, $n_2 = 2$, $n_3 = 2$ ($p \leq 0,05$). Для $p \leq 0,01$ минимальные объемы выборок $n_1 = 3$, $n_2 = 3$, $n_3 = 3$ или $n_1 = 4$, $n_2 = 2$, $n_3 = 2$.

Критические значения именно для этого критерия предполагают только три выборки, при большем количестве выборок рекомендуется пользоваться таблицей критических значений χ^2-критерия, при этом $df = c-1$, c — число выборок.

[1] Наследов А. Д. Указ. соч. С. 364–365.

Для применения этого критерия необходимо, как и с U-Манна-Уитни, расположить все значения в порядке возрастания или убывания и ранжировать.

H-критерий Краскала-Уоллиса вычисляется по формуле:

$$H = \left(\frac{12}{N(N+1)} * \sum \frac{T}{n}\right) - 3(N+1)$$

где N — общее число испытуемых во всех выборках,

n — число испытуемых в каждой выборке,

T — сумма рангов в каждой выборке.

Например, у нас есть четыре группы слушателей, которые готовятся к семинару в разное время.

Таблица 9. Показатели длительности подготовки к семинару в минутах (N = 17)

N	*1 группа (n=5)*	*2 группа (n=5)*	*3 группа (n=3)*	*4 группа (n=4)*
1.	50	50	45	40
2.	65	55	50	40
3.	70	60	55	45
4.	85	65		50
5.	90	65		
Суммы	360	295	655	175
Средние	72	59	50	43,75

H_0: группы слушателей не различаются по времени подготовки к семинару,

H_1: группы слушателей различаются по времени подготовки к семинару.

В таблице 10 показан подсчет ранговых сумм.

Сумма всех рангов эмпирическая 67 + 53,5 + 19,5 + 13 = 153, должна совпадать с теоретической.

Сумма рангов теоретическая зависит от объема выборки $\sum R_i = \frac{17*(17+1)}{2} = 153$.

$df = 4 - 1 = 3$

$$H_{эмп} = \left(\frac{12}{17(17+1)} * \left(\frac{67^2}{5} + \frac{53{,}5^2}{5} + \frac{19{,}5^2}{3} + \frac{13^2}{4}\right)\right) - 3 * (17+1) = 10{,}62$$

Таблица 10. Ранговые суммы по каждой группе

1 группа (n = 5)		*2 группа (n = 5)*		*3 группа (n = 3)*		*4 группа (n = 4)*	
время, мин	*ранг*	*время, мин*	*ранг*	*время, мин*	*ранг*	*время, мин*	*ранг*
						40	1,5
						40	1,5
				45	3,5	45	3,5
50	6,5	50	6,5	50	6,5	50	6,5
		55	9,5	55	9,5		
		60	11				
65	12,5	65	12,5				
		65	14				
70	15						
85	16						
90	17						
Суммы	67		53,5		19,5		13
Сред-ние	13,4		10,7		6,5		3,25

Таблица 11. Подсчет Т-критерия Вилкоксона на примере подготовки к занятиям слушателей, в минутах

Слушатели	Время подготовки (мин)		Разность ($t_{до} - t_{после}$)	Абсолют-ная разность	Ранг
1	64	25	-39	39	11
2	77	50	-27	27	8
3	74	77	+3	3	1
4	95	76	-19	19	6
5	105	67	-38	38	9,5
6	83	75	-8	8	4
7	73	77	+4	4	2,5
8	75	71	-4	4	2,5
9	101	63	-38	38	9,5
10	97	122	+25	25	7
11	78	60	-18	18	5
					66

В таблице 11 находим критические значения Н-критерия Краскала-Уолисса, используя критические значения χ^2 , так как четыре выборки.

$H_{кр} = 7{,}815$ ($p = 0{,}05$); $H_{кр} = 11{,}345$ ($p = 0{,}01$),

$H_{эмп} > H_{кр} \Rightarrow H_1$ ($p \leq 0{,}05$)

Группы слушателей достоверно ($p \leq 0{,}05$) различаются по времени подготовки к семинару.

Критические значения критерия χ^2 для уровней статистической значимости $\rho \leq 0{,}05$ и $\rho \leq 0{,}01$ при разном числе степеней свободы ν

Различия между двумя распределениями могут считаться достоверными, если $\chi^2_{эмп}$ достигает или превышает $\chi^2_{0,05}$, и тем более достоверными, если $\chi^2_{эмп}$ достигает или превышает $\chi^2_{0,01}$ (по Большеву Л.Н., Смирнову Н.В., 1983).

ν	ρ 0,05	ρ 0,01	ν	ρ 0,05	ρ 0,01	ν	ρ 0,05	ρ 0,01
1	3,841	6,635	35	49,802	57,342	69	89,391	99,227
2	5,991	9,210	36	50,998	58,619	70	90,631	100,425
3	7,815	11,345	37	52,192	59,892	71	91,670	101,621
4	9,488	13,277	38	53,384	61,162	72	92,808	102,816
5	11,070	15,086	39	54,572	62,428	73	93,945	104,010
6	12,592	16,812	40	55,758	63,691	74	95,081	105,202
7	14,067	18,475	41	56,942	64,950	75	96,217	106,393
8	15,507	20,090	42	58,124	66,206	76	97,351	107,582
9	16,919	21,666	43	59,304	67,459	77	98,484	108,771
10	18,307	23,209	44	60,481	68,709	78	99,617	109,958
11	19,675	24,725	45	61,656	69,957	79	100,749	111,144
12	21,026	26,217	46	62,830	71,201	80	101,879	112,329
13	22,362	27,688	47	64,001	72,443	81	103,010	113,512
14	23,685	29,141	48	65,171	73,683	82	104,139	114,695
15	24,996	30,578	49	66,339	74,919	83	105,267	115,876
16	26,296	32,000	50	67,505	76,154	84	106,395	117,057
17	27,587	33,409	51	68,669	77,386	85	107,522	118,236
18	28,869	34,805	52	69,832	78,616	86	108,648	119,414
19	30,144	36,191	53	70,993	79,843	87	109,773	120,591
20	31,410	37,566	54	72,153	81,069	88	110,898	121,767
21	32,671	38,932	55	73,311	82,292	89	112,022	122,942
22	33,924	40,289	56	74,468	83,513	90	113,145	124,116
23	35,172	41,638	57	75,624	84,733	91	114,268	125,289
24	36,415	42,980	58	76,778	85,950	92	115,390	126,462
25	37,652	44,314	59	77,931	87,166	93	116,511	127,633
26	38,885	45,642	60	79,082	88,379	94	117,632	128,803
27	40,113	46,963	61	80,232	89,591	95	118,752	129,973
28	41,337	48,278	62	81,381	90,802	96	119,871	131,141
29	42,557	49,588	63	82,529	92,010	97	120,990	132,309
30	43,773	50,892	64	83,675	93,217	98	122,108	133,476
31	44,985	52,191	65	84,821	94,422	99	123,225	134,642
32	46,194	53,486	66	85,965	95,626	100	124,342	135,807
33	47,400	54,776	67	87,108	96,828			
34	48,602	56,061	68	88,250	98,028			

Рис. 40. Критические значения χ^2-критерия[1]

[1] Сидоренко Е. В. Методы математической обработки в психологии. СПб.: Речь, 2000. С. 328.

Т-критерий Вилкоксона (непараметрический для повторных измерений) используется для зависимых выборок, чаще это повторные исследования одной выборки, показывает направленность и выраженность изменений.

Подсчет его достаточно простой:

$T = \Sigma Rr$, где Rr — нетипичный ранг.

Данный критерий является показателем сдвига значений в других условиях. Например, слушатели готовились к занятиям как обычно, а в другой раз их просили готовиться по определенной методике. Гипотеза исследования заключается в большей эффективности новой методики, чем классический вариант подготовки (табл. 11).

H_0 время подготовки слушателей к занятиям значимо не отличается по новой и старой методике.

H_1 время подготовки слушателей к занятиям значимо меньше при использовании новой методики.

Для проверки подсчитаем теоретическую сумму рангов:

$$\sum R_i = \frac{11*(11+1)}{2} = 66.$$

Эмпирическая сумма рангов нетипичного сдвига (в нашем случае положительные показатели разности времени подготовки)

$$R_1 = 1 + 2{,}5 + 7 = 10{,}5.$$

По таблице критических значений (рис. 41) определяем критические значения для n=11, при $p \leq 0{,}05$ и $p \leq 0{,}01$.

$$H_{кр} = 13\ (p = 0{,}05);\ H_{кр} = 7\ (p = 0{,}01)$$

$$13 > 10{,}5 > 7$$

$$H_{\text{эмп}} < H_{кр} \Longrightarrow H_1\ (p \leq 0{,}05)$$

Время подготовки к занятиям слушателей по новой методике достоверно меньше, по сравнению с обычным способом.

Критические значения критерия T Вилкоксона для уровней статистической значимости $\rho \leq 0{,}05$ и $\rho \leq 0{,}01$

"Типичный" сдвиг является достоверно преобладающим по интенсивности, если $T_{эмп}$ ниже или равен $T_{0,05}$, и тем более достоверно преобладающим, если $T_{эмп}$ ниже или равен $T_{0,01}$ (по Wilcoxon F. et al., 1963).

n	p: 0,05	p: 0,01	n	p: 0,05	p: 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

Рис. 41. Критические значения Т-Вилкоксона[1]

В этой теме мы разобрали отличия параметрических и непараметрических критериев, постарались дать алгоритм выбора критериев и описать наиболее распространенные критерии сравнения выборок.

Вопросы и задания для самоконтроля

1. Сравните количественную и качественную шкалы.
2. Приведите пример качественных шкал.
3. Перечислите параметры нормального распределения.
4. Перечислите критерии использующиеся для проверки соответствия распределения нормальному.
5. Перечислите параметрические критерии.
6. Перечислите непараметрические критерии, приведите пример использования.
7. Приведите пример использования t-критерия Стьюдента.
8. Назовите непараметрический аналог t-критерия Стьюдента.
9. Перечислите критерии, которые применяются для трех и более групп.
10. Переведите эмпирические значения из шкалы отношений в ранговую шкалу (например, рост слушателей в вашей группе).

[1] Сидоренко Е.В. Указ. соч. С. 324.

Глава 5. ХАРАКТЕРИСТИКИ ВЗАИМОСВЯЗИ ПРИЗНАКОВ

Учебные вопросы:

5.1. Корреляционный анализ.
5.2. Регрессионный анализ.
5.3. Дисперсионный анализ.

Количественные данные можно рассматривать не только с помощью сравнительных методов, но и анализируя их взаимосвязь. Характеристики взаимосвязи признаков показывают зависимость параметров, это означает, что при появлении одного из параметров обязательно что-то происходит с другим, зависимым параметром. При этом можно точно сказать, что есть связь, но не можем назвать направленность этой связи.

5.1. Корреляционный анализ

Еще Гиппократ (около 460 до н. э. — около 370 до н. э.) искал зависимость между типом темперамента и соотношением жидкостей в организме человека (гуморальная теория). А. Э. Кречмер (1888–1964) считал, что существует взаимосвязь между типом телосложения и определенными психическими особенностями.

Французский натуралист, зоолог, анатом Жорж Кювье (Жан Леопольд Николя Фредерик, барон Кювье 1769–1832), «отец» палеонтологии, в XVIII веке сравнивал живых животных с окаменелостями, изучал анатомию животных и в начале XIX века вывел закон сосуществования органов, который назвал *закон корреляции*, помогающий восстанавливать целые скелеты из части останков. Также метод корреляции позволил Кювье вывести закон о соотношении органов: при изменении структуры или строения одного органа, неизменно происходят изменения и в других органах.

Понятие корреляции в статистике одним из первых стал использовать Френсис Гальтон (1822–1911), двоюродный брат Чарльза Дарвина. Гальтон известен тем, что ввел в психологию и биологию математические методы исследования. В 1892 году предложил способ вычисления длины ноги по длине руки, который в дальнейшем стал известен как коэффициент корреляции.

Однако формула подсчета коэффициента корреляции была выведена учеником Ф. Гальтона — Карлом Пирсоном (1857–1936), английским математиком.

Корреляцией (corelation) называют согласованность изменений определенных параметров. Например, корреляционными можно назвать изменения, происходящие во время занятий спортом, изменяются частота сердечных сокращений (ЧСС), дыхание, давление и другие физиологические показатели. Эти показатели изменяются согласованно (коррелируют друг с другом), но при этом нельзя сказать, какой из них начинает изменения, поскольку они происходят одновременно. В то же время известно, что при сознательном изменении частоты дыхания происходят согласованные изменения и других физиологических процессов (ЧСС, давления и др.).

Корреляции не показывают причинно-следственную связь, они скорее констатируют факт взаимной связи.

Корреляционные связи разделяются по форме, направлению, силе и значимости:

по форме:

— прямолинейная,

— криволинейная;

по направлению:

— положительная,

— отрицательная;

по силе:

— сильная,

— средняя,

— умеренная,

— слабая,

— очень слабая;

по значимости:

— высокозначимая,

— значимая,

— незначимая.

Простые корреляции называют *прямолинейными* или *линейными*, когда при увеличении (уменьшении) одного признака увеличивается или уменьшается другой, взаимозависимый признак.

Прямолинейными они называются потому, что график их изменений представляет прямую линию (рис. 42).

Криволинейная или нелинейная корреляция, например, может быть представлена параболической зависимостью уровня мотивации и эффективностью решения задач (кривая мотивации Йеркса-Додсона), зависимостью забывания от времени (кривая забывания Эббингауза) и другими.

Нелинейная корреляция показывает, что есть зависимость между изменением одной переменной и изменением другой, но коэффициент этой зависимости является переменной величиной.

Положительная или *прямая корреляция* представлена на рисунке 42, она показывает, что при увеличении одного значения увеличивается и зависимое от него другое значение, а при уменьшении — уменьшается.

Отрицательная или *обратная корреляция* показывает, что при увеличении одного значения, уменьшается другое, а при уменьшении — увеличивается, такая корреляция имеет отрицательное значение r (рис. 43).

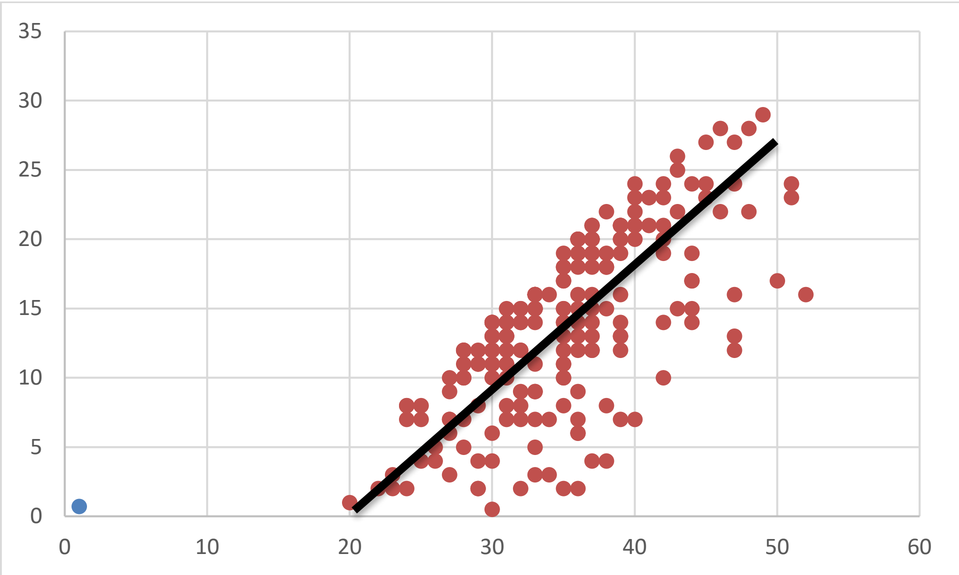


Рис. 42. Линейная положительная корреляция (r = 0,7)

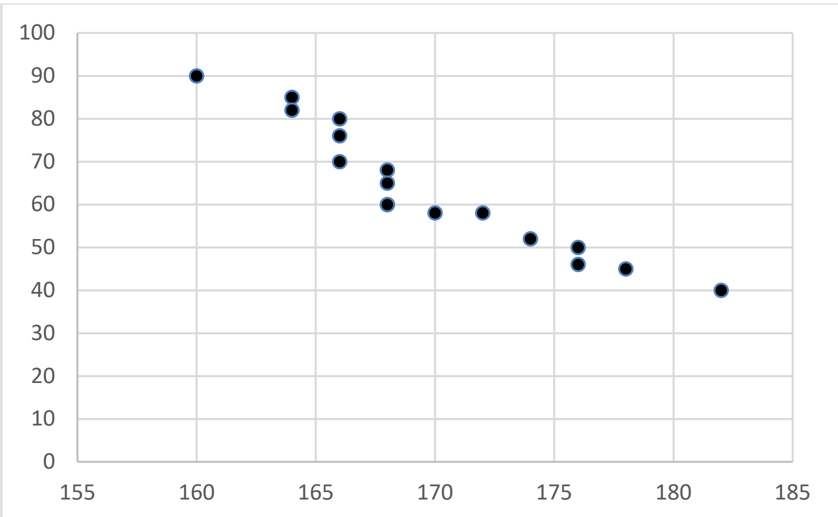


Рисунок 43. Линейная отрицательная корреляция (r = -0,9)

Сила корреляционной связи зависит от величины коэффициента корреляции:

Сила корреляции	*Значения коэффициента*
сильная	$\lvert r \rvert > 0{,}70$
средняя	$0{,}50 < \lvert r \rvert < 0{,}69$
умеренная	$0{,}30 < \lvert r \rvert < 0{,}49$
слабая	$0{,}20 < \lvert r \rvert < 0{,}29$
очень слабая	$\lvert r \rvert < 0{,}19$

Значимость коэффициента корреляции в статистических программах обычно обозначается * при $p \leq 0{,}05$, и ** при $p \leq 0{,}01$; в таблицах критических значений: $p = 0{,}05$, $p = 0{,}01$, $p = 0{,}001$.

Значимость связи коэффициента корреляции	*Уровень достоверности (p)*
высокозначимая	$p \leq 0{,}01$
значимая	$p \leq 0{,}05$
незначимая	$p > 0{,}05$

Корреляционным анализом называют проверку гипотезы о зависимостях между величинами с помощью коэффициента корреляции.

Тип шкалы, в которой представлены данные, определяет вид и коэффициент корреляции:

— линейная или метрическая;

— ранговая;

между номинативными показателями.

В таблице 12 представлен способ вычисления корреляции в зависимости от типа шкал, к которым относятся выбранные значения. Для использования самого мощного критерия взаимосвязи между показателями, коэффициента R — Пирсона, необходимо нормальное распределение данных, а также они должны быть представлены, как минимум, в интервальной, относительной шкале. Если же одна из выборок не имеет нормального распределения или представлена в ранговой шкале, то следует использовать коэффициент R_S-Спирмена. Если же наша выборка составляет данные дихотомической шкалы, то рекомендуется использовать таблицы сопряженности и коэффициент Пирсона для дихотомических шкал.

Таблица 12. Выбор коэффициента корреляции
в зависимости от типа шкал

Тип шкалы		*Коэффициент корреляции*
выборка X	*выборка Y*	
абсолютная или относительная	абсолютная или относительная	R_{xy} коэффициент Пирсона
ранговая	ранговая	R_S коэффициент Спирмена
дихотомическая	дихотомическая	φ коэффициент Пирсона для дихотомических шкал

Коэффициент корреляции Пирсона показывает наличие или отсутствие линейной взаимосвязи между двумя параметрами, представленными в шкале отношений, при нормальном распределении, которые обычно обозначаются как X и Y, а коэффициент корреляции R_{xy}. R_{xy} не может быть меньше –1 и больше 1, чем ближе значение стремится к 1, тем больше точки на графике стремятся к ровной линии. Формула, для подсчета коэффициента корреляции Пирсона[1]:

$$R_{xy} = \frac{\sum(x_i - M_x)(y_i - M_y)}{(n-1)\sigma^x \sigma^y}$$

где x_i y_i — два показателя одного испытуемого,

M_x — среднее арифметическое значений x,

σ^x — стандартное отклонение значений x.

Коэффициент корреляции Пирсона можно вычислить в Excel (рис. 44).

Для решения выдвинутой проблемы зависимости одного показателя от другого необходимо выдвинуть статистические гипотезы:

— H_0 — значение коэффициента корреляции значимо не отличается от нуля, то есть взаимосвязь является случайной;

— H_1 — значение коэффициента корреляции значимо отличается от нуля, то есть взаимосвязь не является случайной.

Для сравнения эмпирического (полученного) значения с критическим необходимо воспользоваться таблицами критических значений и/или понять, значима ли связь. Взаимосвязь считается значимой, если $R_{эмп} < R_{кр} \Longrightarrow H_0$; $R_{эмп} \geq R_{кр} \Longrightarrow H_1$.

[1] Наследов А. Д. Указ. соч. С.70

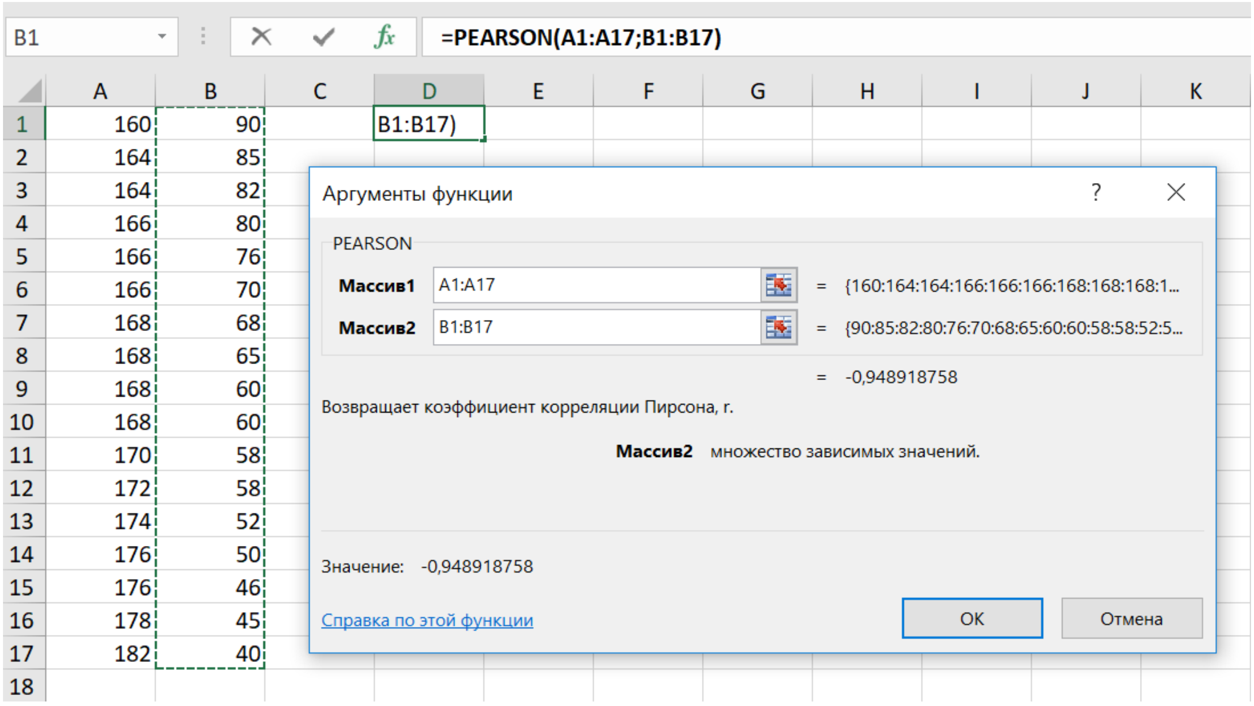


Рис. 44. Вычисление коэффициента корреляции Пирсона в программе MC Excel

Если требуется узнать, есть ли взаимосвязь между переменными, выраженными в дихотомической шкале, которая содержит всего две категории (например, мужчина-женщина, белый-черный, курсант-слушатель и др.), то рекомендуется применять коэффициент Пирсона для дихотомических шкал (критерий согласия Пирсона) на основе таблицы сопряженности. Таблица сопряженности показывает число совместных появлений (табл. 13) исследуемых показателей.

Таблица 13. Общая таблица сопряженности

Признак	*Курсант*	*Слушатель*	*Итого*
Мужчина	A	B	A + B
Женщина	C	D	C + D
Итого	A + C	B + D	N

Формула для вычисления критерия согласия Пирсона:

$$\varphi = \frac{(BC - AD)}{\sqrt{(A + C)) * (B + D) * (A + B) * (C + D)}}$$

Пример: требуется узнать, есть ли зависимость между полом и факультетом, то есть действительно ли на факультете подготовки иностранных слушателей учится больше мужчин, чем на других факультетах университета. Строим таблицу сопряженности:

Признак	*Курсант*	*Слушатель*	*Итого*
Мужчина	200	60	260
Женщина	150	10	160
Итого	350	70	420

Подставляем полученные эмпирическим путем данные в формулу для вычисления коэффициент Пирсона для дихотомических шкал:

$$\varphi=(150*60 — 200*10)/\sqrt{(350*70*260*160)} = 7000/31923{,}91 = 0{,}219.$$

Полученное значение 0,219 говорит о том, что существует слабая зависимость между полом и выбором факультета.

Коэффициент корреляции Спирмена показывает наличие или отсутствие линейной взаимосвязи между двумя параметрами, представленными в ранговой шкале, которые обычно обозначаются как X и Y, а коэффициент корреляции R_s, который вычисляется по формуле[1]:

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

где d — разность рангов для каждого испытуемого. При этом сумма разности рангов всех испытуемых должна быть равна единице.

Корреляционная плеяда — группа признаков, коррелирующих друг с другом.

_ _ _ _ _ _ отрицательная корреляция при $p \le 0{,}05$;

——— положительная корреляции при $p \le 0{,}05$.

Рис. 45. Корреляционная плеяда

[1] Наследов А. Д. Указ. соч. С. 78.

На рисунке корреляционная плеяда, на которой показаны отрицательные взаимосвязи между стажем и тревожностью (чем больше стаж, тем меньше тревожность), стажем и эмоциональностью (чем больше стаж, тем меньше эмоциональность), и положительная взаимосвязь между стажем и возрастом (чем больше стаж, тем больше возраст). Корреляционные плеяды используются для наглядного представления результатов корреляционного анализа, причем в построении плеяд можно использовать не все выявленные корреляции, а только наиболее значимые для данного исследования. Также желательно стоить не одну корреляционную плеяду с множествами связей, а несколько плеяд, подчеркивающих значимость того или иного параметра.

Таким образом, корреляционная связь — согласованные изменения двух признаков, корреляционная зависимость — изменения одного признака, влекущие за собой изменения другого признака. Необходимо также понимать, что согласованные изменения признаков могут появляться вследствие влияния третьего, неизвестного, но влиятельного признака или определенного сочетания признаков, которые не были рассмотрены в данном исследовании.

5.2. Регрессионный анализ

Термин «регрессионный анализ» ввел также Ф. Гальтон, который одним из первых стал использовать корреляции. Он установил, что существует определенное соотношение между ростом взрослых детей и их родителей: у очень низких родителей дети немного выше, а у очень высоких родителей — немного ниже. Эту закономерность он назвал *регрессией*.

Регрессионный анализ исследует изменения показателя среднего значения от изменения показателей других значений, а также способствует изучению взаимосвязи между показателями, представленными в интервальной шкале. Данный анализ выражается в уравнении регрессии, которое показывает количественную связь зависимых (критериальных) переменных от независимых (предикторов, или регрессоров). Если у нас есть регрессионное уравнение и значения независимых переменных, то мы можем вычислить значения зависимых переменных.

Данный анализ выявляет связь между переменными, которую можно представить в виде математической функции на основе теории вероятности. Он также способствует нахождению наиболее значимых факторов, влияющих на зависимую переменную, позволяет их ранжировать, вычислить зависимую переменную.

Однофакторный линейный регрессионный анализ[1] позволяет прогнозировать, как измениться один параметр в связи с изменением другого параметра, то есть при линейной регрессии используются всего два параметра. Предположим, у нас есть значения зависимого и независимого параметров у 10 испытуемых (рис. 46). Данный скаттерплот (scatterplot — диаграмма рассеяния) может быть описан уравнением.

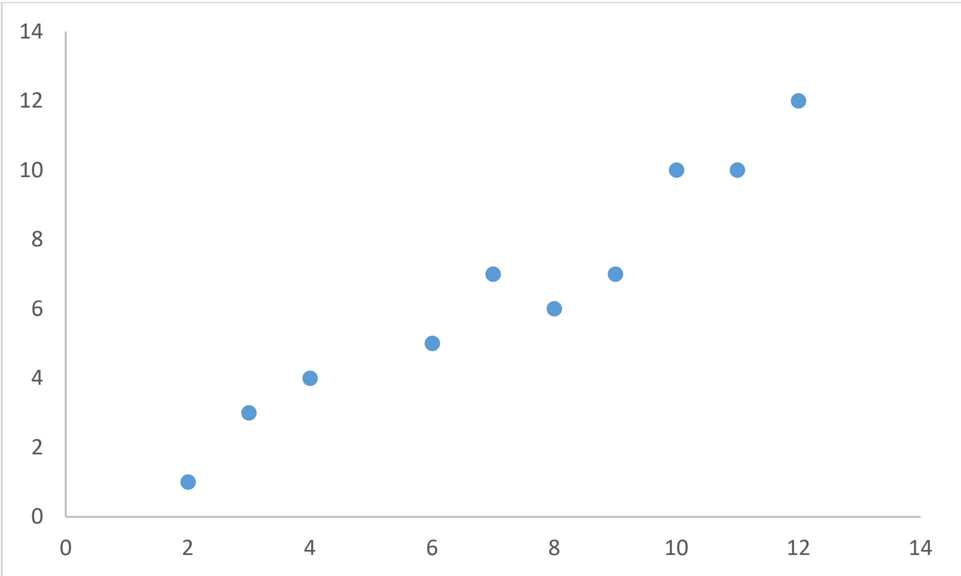


Рис. 46. Скаттерплот зависимости двух параметров у 10 испытуемых

Линия регрессии в статистических программах строится таким образом, чтобы разброс точек был минимальным. Уравнение линейной регрессии имеет вид:

$$Y = b_0 + b_1 * X,$$

где b_0 — точка 0 (свободный член), от которой начинается линия,

b_1 — тангенс угла между осью X и линией регрессии (коэффициент регрессии) (рис. 47).

Метод наименьших квадратов (МНК) позволяет находить оптимальные параметры линейной регрессии, чтобы сумма квадратов ошибок отклонений от линии была минимальной.

[1] Гржибовский А. М., Иванов С. В., Горбатова М. А. Однофакторный линейный регрессионный анализ с использованием программного обеспечения Statistica и SPSS // Наука и здравоохранение. — 2017. — №2. — С. 5–33.

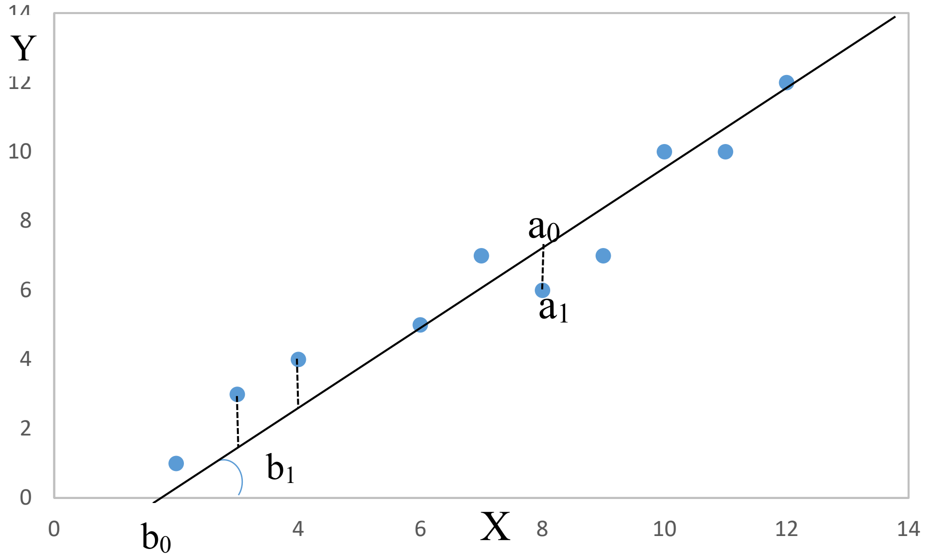


Рис. 47. Линия регрессии для параметров X и Y

Наилучшая линия будет проходить через M_x; M_y.

Регрессионным остатком называют сумму разниц между каждым значением и линией регрессии (a_1 — a_0) (рис.) 47. Для исключения влияния знаков принято складывать квадраты остатков $(a_1 — a_0)^2$.

Для нахождения углового коэффициента наклона линии регрессии необходимо вычислить коэффициент корреляции, а затем подсчитать по формуле:

$$b_1=R_{xy}* \frac{\sigma y}{\sigma x}.$$

Точку, от которой начинается линия регрессии, также можно вычислить по формуле:

$$b_0= M_y - b_y * M_x,$$

где M — среднее значение параметра.

Таким образом, зная значение параметра X, можно вычислить значение параметра Y и построить однофакторную регрессионную модель.

При коэффициенте корреляции равном 0, то есть при отсутствии взаимосвязи между параметрами X и Y, $b_0= M_y$, линия регрессии равна среднему значению Y и будет параллельна оси X.

Однофакторный линейный регрессионный анализ имеет ограничения применения, связанные[1]:

— с репрезентативностью выборки;

[1] Гржибовский А. М., Иванов С. В., Горбатова М. А. Указ. соч.

— шкалы изменения должны быть количественными (интервальными);

— должны быть линейная зависимость между переменными;

— переменные должны быть независимыми;

— остатки имеют одинаковый разброс относительно линии регрессии.

Многомерный регрессионный анализ позволяет прогнозировать, как изменяется зависимая переменная в зависимости от изменения нескольких независимых переменных.

Одним из условий построения данной модели является отсутствие сильных корреляционных связей (R>0,9) между независимыми переменными, так как они могут исключить друг друга, предпочтительнее выбрать одну из переменных. Уравнение многомерной линейной регрессии будит иметь вид:

$$Y = b_0 + b_1 * X_1 + b_2 * X_2 + b_3 * X_3 + \ldots + b_k * X_k$$

Коэффициент многомерной регрессионной модели получают методом наименьших квадратов (МНК), значимость каждого коэффициента оценивается по t-критерию Стьюдента. В психологии значимыми считаются коэффициенты с доверительной вероятностью $\geq 0{,}95$, или $p \leq 0{,}05$.

Коэффициент детерминации или степень соответствия обозначается как R^2 и показывает долю дисперсии зависимой переменной (Y), объясняемую регрессионной моделью.

$$R^2 = 1 - \frac{SSe}{SSt},$$

где SSe — сумма квадратов остатков (error) $SSe=\sum_{j=1}^{n}(y_j - \widehat{y_j})^2$

SSt — сумма квадратов (total)) $SSt=\sum_{j=1}^{n}(y_j - \overline{y_j})^2$,

$\widehat{y_j}$ — модельные значения (значения на линии регрессии),

$\overline{y_j}$ — средние значения.

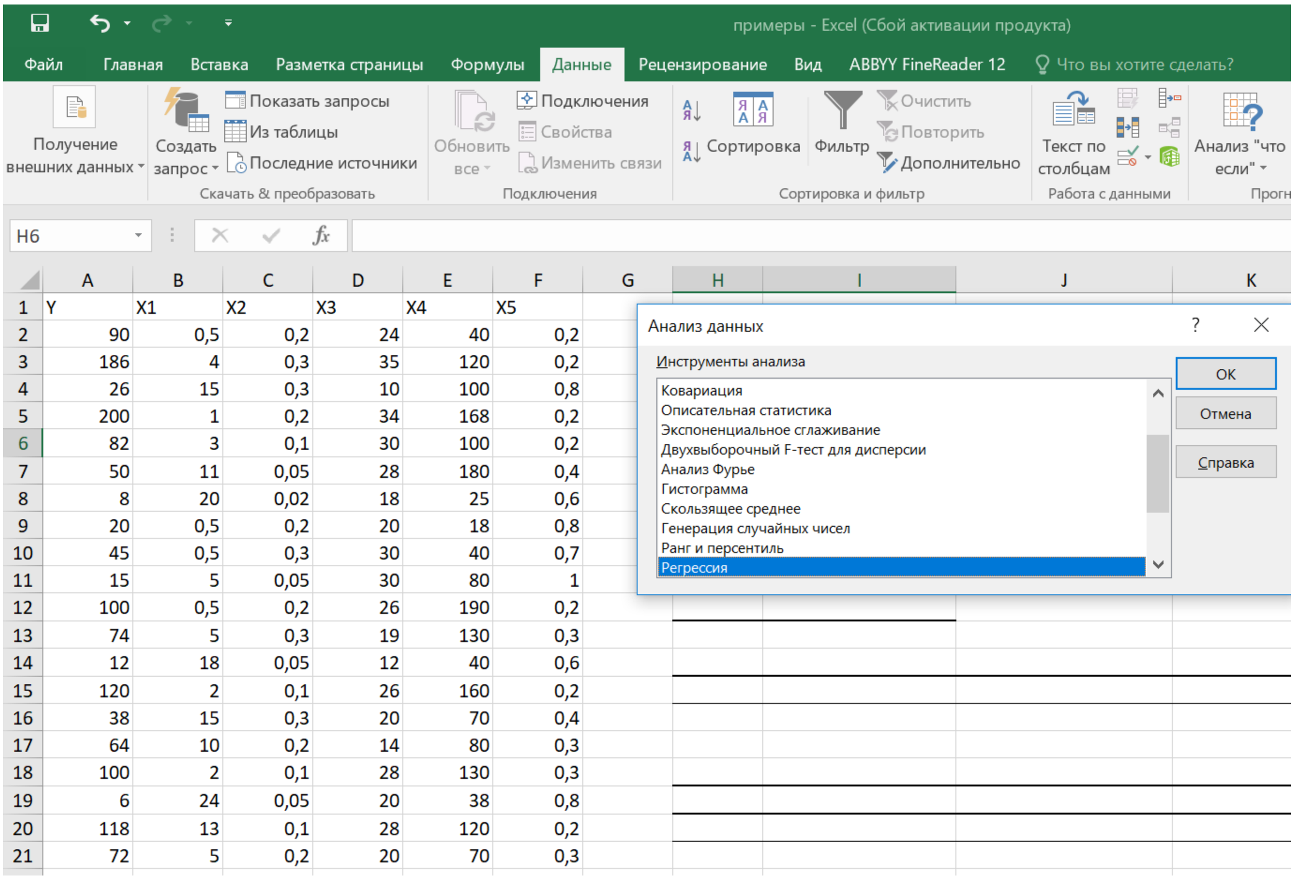
Величина R^2 определяет процентное взаимное влияние переменных. Например, если $R^2 = 0{,}412$, то это значит, что лишь 41,2 % общей дисперсии объясняются взаимным влиянием переменных X и Y, остальные 58,8 % обусловлены влиянием неучтенных в уравнении факторов.

Коэффициент детерминации определяет точность прогноза, его можно вычислить в MS Excel, обозначается он как R^2. Значения располагаются между 0 и 1, при этом значения

$R^2 > 0{,}8$ относятся к моделям хорошего качества, 0,79 < R^2 < 0,5 модели достаточного качества,

R^2 < 0,5 модели низкого качества.

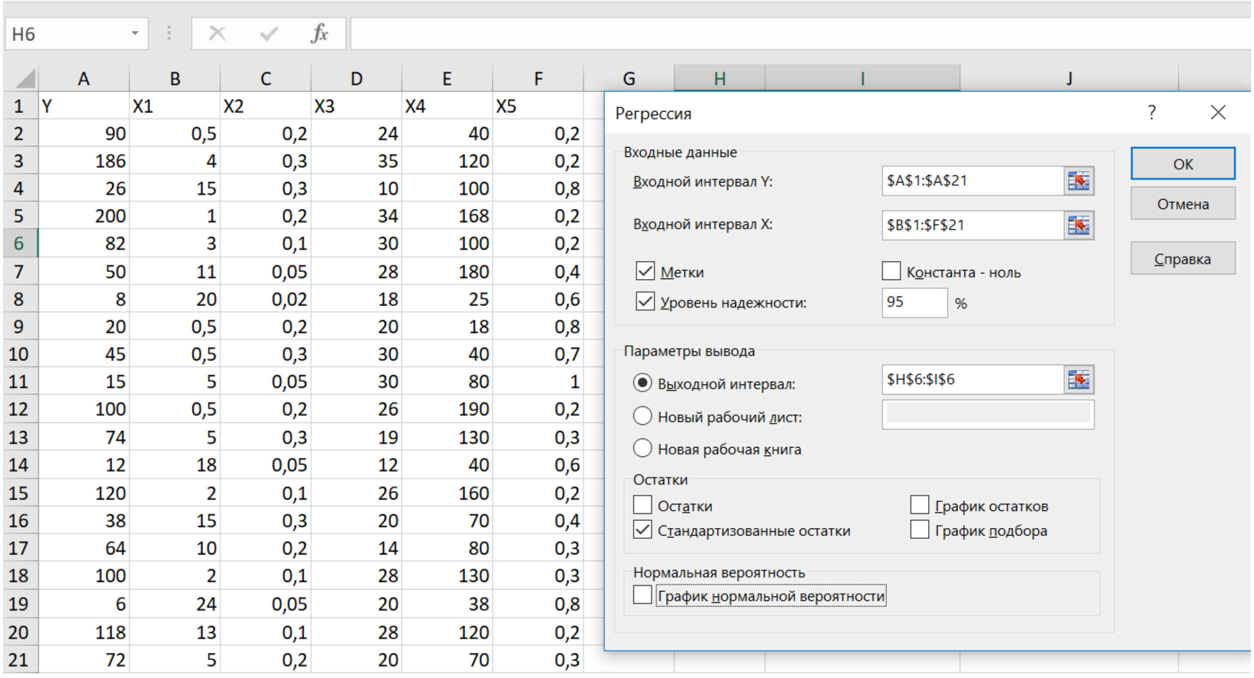
В программе с помощью функций выбираем КВПИРСОН и получаем единственный показатель коэффициента детерминации для двух параметров. Также в программе MS Excel коэффициент детерминации можно выбрать с помощью инструмента «регрессия» (рис. 48.1, 48.2).



	A	B	C	D	E	F
1	Y	X1	X2	X3	X4	X5
2	90	0,5	0,2	24	40	0,2
3	186	4	0,3	35	120	0,2
4	26	15	0,3	10	100	0,8
5	200	1	0,2	34	168	0,2
6	82	3	0,1	30	100	0,2
7	50	11	0,05	28	180	0,4
8	8	20	0,02	18	25	0,6
9	20	0,5	0,2	20	18	0,8
10	45	0,5	0,3	30	40	0,7
11	15	5	0,05	30	80	1
12	100	0,5	0,2	26	190	0,2
13	74	5	0,3	19	130	0,3
14	12	18	0,05	12	40	0,6
15	120	2	0,1	26	160	0,2
16	38	15	0,3	20	70	0,4
17	64	10	0,2	14	80	0,3
18	100	2	0,1	28	130	0,3
19	6	24	0,05	20	38	0,8
20	118	13	0,1	28	120	0,2
21	72	5	0,2	20	70	0,3

Рис. 48.1 Использование инструмента «регрессия» в МС Excel

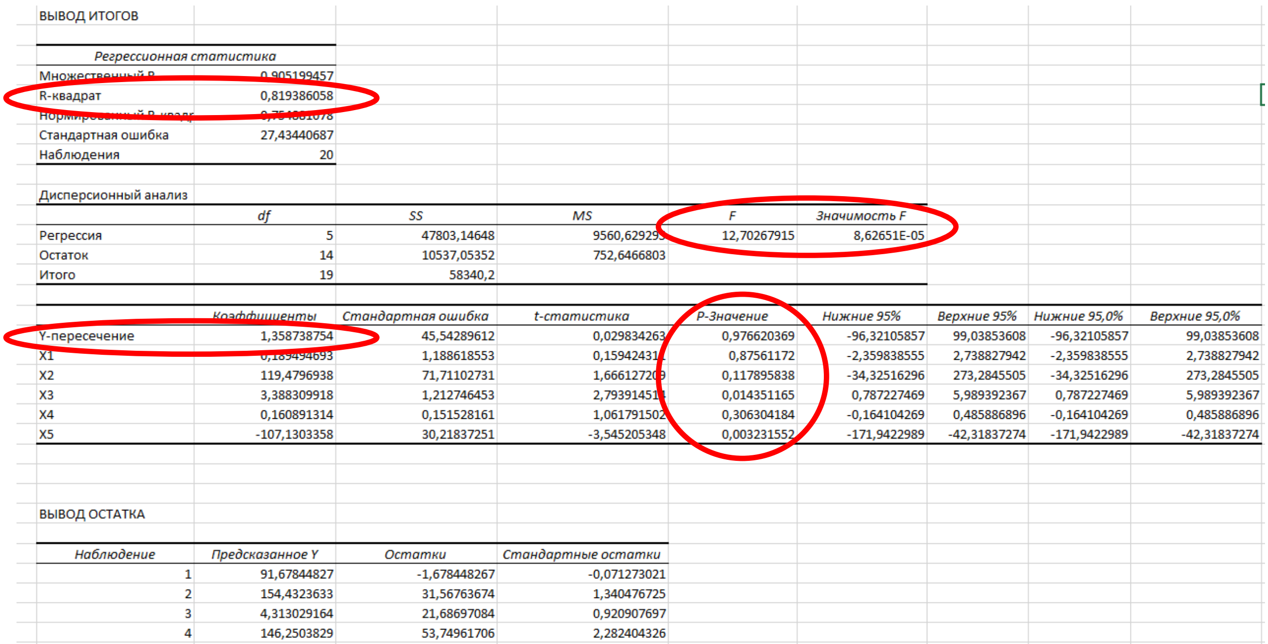
На рисунке 48.2 представлено применение инструмента «регрессия» в МС Excel для нахождения коэффициента детерминации и построения уравнения регрессии.



	A	B	C	D	E	F
1	Y	X1	X2	X3	X4	X5
2	90	0,5	0,2	24	40	0,2
3	186	4	0,3	35	120	0,2
4	26	15	0,3	10	100	0,8
5	200	1	0,2	34	168	0,2
6	82	3	0,1	30	100	0,2
7	50	11	0,05	28	180	0,4
8	8	20	0,02	18	25	0,6
9	20	0,5	0,2	20	18	0,8
10	45	0,5	0,3	30	40	0,7
11	15	5	0,05	30	80	1
12	100	0,5	0,2	26	190	0,2
13	74	5	0,3	19	130	0,3
14	12	18	0,05	12	40	0,6
15	120	2	0,1	26	160	0,2
16	38	15	0,3	20	70	0,4
17	64	10	0,2	14	80	0,3
18	100	2	0,1	28	130	0,3
19	6	24	0,05	20	38	0,8
20	118	13	0,1	28	120	0,2
21	72	5	0,2	20	70	0,3

Рис. 48.2. Использование инструмента «регрессия» в MC Excel

На рисунке 49 представлен результат применения инструмента «регрессия».



ВЫВОД ИТОГОВ

Регрессионная статистика	
Множественный R	0,905199457
R-квадрат	0,819386058
Нормированный R-квадр	0,754881078
Стандартная ошибка	27,43440687
Наблюдения	20

Дисперсионный анализ

	df	*SS*	*MS*	*F*	*Значимость F*
Регрессия	5	47803,14648	9560,62929	12,70267915	8,62651E-05
Остаток	14	10537,05352	752,6466803		
Итого	19	58340,2			

	Коэффициенты	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*	*Верхние 95%*	*Нижние 95,0%*	*Верхние 95,0%*
Y-пересечение	1,358738754	45,54289612	0,029834263	0,976620369	-96,32105857	99,03853608	-96,32105857	99,03853608
X1	0,189494693	1,188618553	0,15942431	0,87561172	-2,359838555	2,738827942	-2,359838555	2,738827942
X2	119,4796938	71,71102731	1,66612720	0,117895838	-34,32516296	273,2845505	-34,32516296	273,2845505
X3	3,388309918	1,212746453	2,79391451	0,014351165	0,787227469	5,989392367	0,787227469	5,989392367
X4	0,160891314	0,151528161	1,061791502	0,306304184	-0,164104269	0,485886896	-0,164104269	0,485886896
X5	-107,1303358	30,21837251	-3,545205348	0,003231552	-171,9422989	-42,31837274	-171,9422989	-42,31837274

ВЫВОД ОСТАТКА

Наблюдение	*Предсказанное Y*	*Остатки*	*Стандартные остатки*
1	91,67844827	-1,678448267	-0,071273021
2	154,4323633	31,56763674	1,340476725
3	4,313029164	21,68697084	0,920907697
4	146,2503829	53,74961706	2,282404326

Рис. 49. Результат применения инструмента «регрессия» MC Excel

На рисунке 49 определим обозначения:

R-квадрат — это коэффициент детерминации;

Y-пересечение показывает значение свободного члена уравнения (b_0), коэффициент показывает, какой будет Y, если все X будут равны 0, то есть зависимость от неописанных в модели показателей.

Достоверность по уровню значимости критерия Фишера ($F = 12,70$; $p = 0,00008$) значительно меньше 0,05, следовательно, модель значима. На основании представленных на рисунке 49 результатов можно построить уравнение регрессии, предварительно выделив значимость каждого параметра X и выбрав только значимые ($p \leq 0,05$):

$$Y = 1,36 + 3,39*X_3 – 107,13*X_5.$$

Построенная модель объясняет 81,9 % зависимости независимой переменной Y от рассматриваемых зависимых переменных X.

Рассмотрим несколько вариантов нелинейной регрессии, которую можно описать с помощью квадратного или полиномиального уравнения, которое можно создать с помощью метода полиномиальной регрессии. Для выбора регрессионного уравнения рекомендуется предварительно построить точечные диаграммы, или скаттерплоты, которые покажут виды зависимостей.

Логистическая регрессия позволяет на основании одной или нескольких переменных предсказать вероятность принадлежности показателя к одной из двух категорий, а также прогнозировать, к какому из двух вариантов относится конкретный показатель и какая вероятность попадания этого показателя к данному варианту. Е. Е. Шарашова, К. К. Холматова, М. А. Горбатова, А. М. Гржибовский[1] описывают условия для логистической регрессии:

— дихотомическая зависимая переменная;

— предпочтительный вариант зависимой переменной должен быть обозначен 1;

— переменные должны быть независимыми;

— отсутствие взаимных корреляций между независимыми переменными;

— линейная зависимость между каждой независимой переменной и логарифмом отношения шансов;

— независимость остатков;

---

[1] Шарашова Е. Е. и др. Применение множественного логистического регрессионного анализа в здравоохранении с использованием пакета статистических программ spss // Наука и здравоохранение. — 2017. — №4. — С. 5–26.

— достаточно большой объем выборки (от 30 на одну независимую переменную, то есть, если три независимые переменные, то будет минимум 90 наблюдений).

Вероятность попадания зависимой переменной к предпочтительному варианту рассчитывается с помощью логарифма шанса. Шанс равен вероятности попадания в группу 1, деленную на вероятность попадания в группу 0:

$$\frac{P}{1-P}, \; P = \frac{1}{1+e^{-y}},$$

где P — вероятность того, что произойдет событие;

e — основание натуральных логарифмов 2,71…,

y — стандартное уравнение регрессии.

Логистическая регрессия вычисляется с помощью статистических программ.

Регрессионная модель позволяет выделить предикторы, или предсказатели, зависимой переменной, но не позволяет говорить о причинно-следственных связях.

Дискриминантный анализ является одним из вариантов множественного регрессионного анализа при условии, что зависимая переменная представлена в номинативной шкале. Дискриминантный анализ считается методом классификации, каждое наблюдение можно отнести только к одному классу или они могут остаться «неизвестными». Классы зависимой переменной называются дискриминантными переменными, или предикторами. Например, классы зависимой переменной могут быть выделены в зависимости от успешности обучения, а дискриминантными переменными — факультет, средний балл аттестата, показатели тревожности, стрессоустойчивости.

Центроид — место типичных наблюдений для данного класса, средние значения дискриминантных переменных данного класса. Обозначается как $\bar{z}_1$ — центроид 1 класса, $\bar{z}_2$ — центроид 2 класса и т. д. Геометрической интерпретацией дискриминантного анализа является определение расстояния каждой точки до центроидов и выбор наименьшего. Наименьшее расстояние позволяет отнести данный объект к соответствующему классу.

Условия применения дискриминантного анализа:

1) объем выборки должен min в 10 раз превышать количество независимых переменных;

2) объем каждого класса должен быть не меньше, чем в 5 раз больше количества независимых переменных;

3) количество независимых переменных не больше 10;

4) необходимо исследование выбросов;

5) необходимо учитывать сильную корреляцию независимых переменных.

Дискриминантная функция имеет вид:

$$z = b_0 + b_1 * x_1 + b_2 * x_2 + \cdots + b_n * x_n,$$

классификация происходит с помощью соотношения показателя с величиной z, которая зависит от количества классификационных групп, в примере рассмотрены 2 группы:

$$z = \frac{\bar{z}_1 - \bar{z}_2}{2}.$$

Дискриминантный анализ помогает вычислить априорные (до проведения классификации) и апостериорные (после проведения или действительные) вероятности классификации. Также можно определить статистическую значимость различий между группами на основе критерия лямбда Уилкса (Wilks' Lambda) λ: если $p < 0{,}05$, то средние двух групп значимо различаются, следовательно, эта переменная может быть предиктором для анализа, если $p > 0{,}05$, то данную переменную можно убрать из анализа.

Коэффициенты корреляции между независимыми переменными не должны быть больше 0,5. Каноническая корреляция анализирует корреляционные зависимости показывает силу взаимосвязи между независимыми переменными и группами. Чем больше расстояние между центроидами, тем лучше разделены группы.

Дискриминантный анализ чаще проводят с помощью статистических программ для анализа данных.

5.3. Дисперсионный анализ

Дисперсионный анализ показывает изменчивость признака в зависимости от изменения другого/других признаков; выявление показателей, влияющих на результаты эксперимента. В англоязычной литературе обозначается как ANOVA — Analysis of Variance, анализ вариативности (дисперсии). Автором данного метода является Рональд Фишер (Fisher R.A.) и, соответственно, используется критерий F Фишера.

Дисперсионный анализ выявляет различия между выборочными средними для трех и более выборок, его задачей является анализ взаимосвязи количественных и качественных переменных, качественными являются зависимые переменные, а количественными — независимые. При воздействии на испытуемых нескольких независимых переменных общую дисперсию можно представить как сумму дисперсий каждой переменной и неизвестными воздействиями (остаточная дисперсия). При сравнении дисперсии каждой независимой переменной с остаточной дисперсией определяется достоверность воздействия данной переменной.

При дисперсионном анализе одни переменные рассматриваются как причины, другие — как следствия. Причины называются независимыми переменными, а следствия — зависимыми переменными или результативными признаками. Дисперсионный анализ — это метод, изучающий изменчивость признака в зависимости от воздействия контролируемых переменных.

Суммарную или общую дисперсию необходимо разложить на отдельные компоненты, зависящие от воздействия той или иной независимой переменной, и проверить гипотезы о значимости воздействия каждой независимой переменной на зависимую. При сравнении компонентов дисперсии друг с другом с помощью критерия F Фишера можно определить долю общей вариативности результативного признака, обусловленную действием независимых переменных.

Для проведения дисперсионного анализа необходимо не менее трех выборок, которые могут быть разными по количеству и быть как связанными, так и несвязанными.

Выделяют несколько видов дисперсионного анализа:

1) в зависимости от количества независимых переменных — однофакторный и многофакторный;

2) в зависимости от количества зависимых переменных — одномерный и многомерный;

3) дисперсионный анализ с повторными измерениями (зависимые выборки);

4) дисперсионный анализ с постоянными факторами, случайными факторами, и смешанные модели с факторами обоих типов.

Виды дисперсионного анализа представлены в таблице 14.

Таблица 14. Виды дисперсионного анализа

Кол-во независимых переменных (k)	*Кол-во зависимых переменных (количественные) (x_i)*	*Вид дисперсионного анализа*
одна	одна	однофакторный
две и больше	одна	многофакторный
одна или больше	две и больше	многомерный

Однофакторный дисперсионный анализ для несвязанных (независимых) выборок. Исследуется влияние одной независимой переменной на зависимую, как изменяется зависимая переменная под влиянием различных градаций фактора. Градаций независимой переменной должно быть не менее трех. В дисперсионном анализе сравнивается вариативность признака, связанная с воздействием независимой переменной, с общей вариативностью, на основе расчета суммы квадратов (SS — sum of squares). Альтернативная гипотеза должна быть сформулирована следующим образом: различия между градациями фактора (зависимыми переменными) более выражены, чем случайные различия внутри каждой группы.

Межгрупповая сумма квадратов отклонений (*between-group*):

$$SS_b = \sum n_i(\bar{x}_i - M)^2,$$

где n_i — количество исследований в i-группе,

$\bar{x}_i$ — среднее значение x в i-группе,

M — общее среднее по всем группам.

Внутригрупповая сумма квадратов отклонений (*within-group*):

$$SS_w = \sum(x - \overline{x_i})^2.$$

Общая сумма квадратов отклонений (*total*):

$$SS_t = \sum(x - M)^2.$$

Межгрупповая дисперсия (факторная) объясняет изменения, связанные с независимой переменной (*mean square)*:

$$MS_b = \frac{SS_b}{k-1},$$

где k — количество градаций фактора.

Внутригрупповая (остаточная) дисперсия объясняет изменения, связанные с неизвестными факторами:

$$MS_w = \frac{SS_w}{N-k},$$

где N — общее количество всех исследований.

Критерий F Фишера: чем больше F, тем больше отличаются средние значения, тем больше значимость различий:

$$F = \frac{MS_b}{MS_w}.$$

Степени свободы числителя: df = k-1, знаменателя: df = N-k.

Коэффициент детерминации вычисляется как отношение межгрупповой к общей сумме квадратов и показывает долю общей дисперсии зависимой переменной, связанной с влиянием фактора. Чем больше коэффициент детерминации, тем больше воздействие исследуемого фактора на дисперсию зависимой переменной. Если этот показатель умножить на сто, то можно узнать процент учтенной дисперсии:

$$R^2 = \frac{SS_b}{SS_t}.$$

В дисперсионном анализе возможно сравнение выборок с разными количествами испытуемых, тогда требуется проверить дисперсии на гомогенность, в статистических пакетах используется критерий Ливена, который определяет уровень значимости различий дисперсии групп. При $p > 0{,}05$ различие дисперсий незначимо, то есть группы являются гомогенными и можно вычислять с помощью ANOVA однофакторный дисперсионный анализ.

Пример: зависит ли агрессивное поведение от длительности стажа службы?

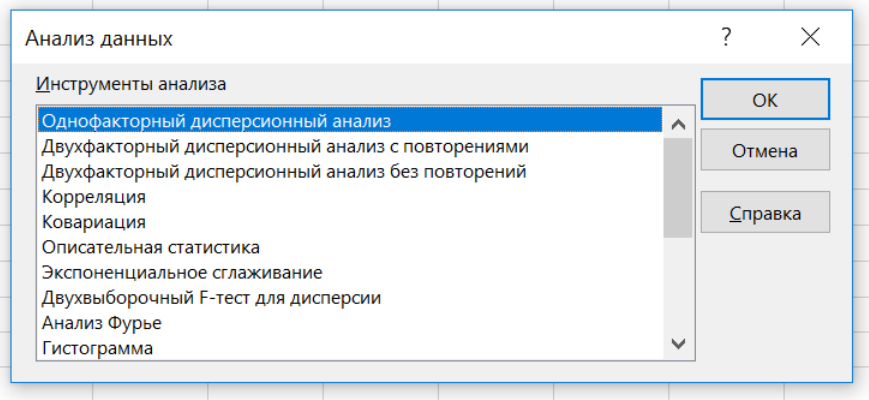
Были отобраны 3 группы полицейских мужского пола, в зависимости от стажа службы: 1 — стаж меньше трёх лет, 2 — стаж больше трёх лет, но меньше десяти, 3 — стаж больше десяти лет. Исследование проводилось с помощью теста руки Вагнера, использовался один количественный показатель агрессивного поведения — общее количество ответов (данные приведены в таблице).

Таблица 15. Показатель открытого агрессивного поведения в зависимости от стажа службы

Меньше 3 лет	*От 3 до 10 лет*	*Больше 10 лет*
16	25	20
20	30	25
23	34	27
18	37	30
25	40	32
21	44	35
24	33	29

В качестве зависимой переменной в данном примере выступает показатель агрессивного поведения, выраженный в количественном показателе, в качестве независимой переменной выступает стаж. Предполагается, что стаж влияет на агрессивное поведение. Расчет произведен с помощью программы MC Excel (рис. 50).

меньше 3 лет	от 3 до 10 лет	больше 10 лет
16	25	20
20	30	25
23	34	27
18	37	30
25	40	32
21	44	35
24	33	29



Однофакторный дисперсионный анализ

ИТОГИ

Группы	*Счет*	*Сумма*	*Среднее*	*Дисперсия*
Столбец 1	7	147	21	10,66667
Столбец 2	7	243	34,71429	39,90476
Столбец 3	7	198	28,28571	23,90476

Дисперсионный анализ

Источник вариации	*SS*	*df*	*MS*	*F*	*P-Значение*	*F критическое*
Между группами	659,142857	2	329,5714	13,27558	0,00028689	3,554557146
Внутри групп	446,857143	18	24,8254			
Итого	1106	20				

Рис. 50. Однофакторный дисперсионный анализ с помощью MC Excel

На рисунке видно, что в данном примере видно, что стаж влияет на агрессивное поведение, принимается H_1, $p = 0{,}000287$, также можно вычислить коэффициент детерминации (R^2).

$$R^2 = 659{,}14 / 1106 = 0{,}596,$$

который показывает, что около 60 % изменений проявления агрессивного поведения связаны со стажем.

Многофакторный дисперсионный анализ исследует влияние нескольких факторов на зависимую переменную. На проявления агрессивного поведения может влиять не только стаж, но и пол. В таблице 16 представлены данные для двухфакторного дисперсионного анализа, где факторами являются стаж и пол.

Таблица 16. Показатель открытого агрессивного поведения в зависимости от стажа службы и пола

	Меньше 3 лет	От 3 до 10 лет	Больше 10 лет
Мужчины	16	25	20
	20	30	25
	23	34	27
	18	37	30
	25	40	32
	21	44	35
	24	33	29
Женщины	20	30	40
	25	40	37
	28	35	32
	30	33	34
	34	38	41
	31	34	40
	29	32	39

Гипотезы можно сформулировать как «на показатели агрессивного поведения влияет стаж службы и/или пол». В нашем примере представлены независимые группы, поэтому в программе МС Excel выберем инструмент «двухфакторный дисперсионный анализ без повторений», результаты которого представлены на рисунке 51. Значимые различия выявлены между строками, следовательно, мужчины и женщины значимо отличаются друг от друга в демонстрации агрессивного поведения.

Таким образом, дисперсионный анализ позволяет выявить зависимости данных экспериментального исследования на основании значимости различий средних значений трех и более выборок. В англоязычной литературе обозначается как ANOVA (от англ. ANalysis Of VAriance).

Характеристики взаимосвязи признаков отражают причинно-следственные отношения, корреляционный анализ позволяет выявить, есть ли согласованные изменения показателей одной переменной в зависимости от изменения другой. Регрессионный анализ позволяет в виде математического выражения представить выявленную корреляционную зависимость. Дисперсионный анализ позволяет выявить воздействие одного или нескольких факторов на зависимые переменные.

Двухфакторный дисперсионный анализ без повторений

ИТОГИ	Счет	Сумма	Среднее	Дисперсия
Строка 1	3	61	20,3333333	20,3333333
Строка 2	3	75	25	25
Строка 3	3	84	28	31
Строка 4	3	85	28,3333333	92,3333333
Строка 5	3	97	32,3333333	56,3333333
Строка 6	3	100	33,3333333	134,333333
Строка 7	3	86	28,6666667	20,3333333
Строка 8	3	90	30	100
Строка 9	3	102	34	63
Строка 10	3	95	31,6666667	12,3333333
Строка 11	3	97	32,3333333	4,33333333
Строка 12	3	113	37,6666667	12,3333333
Строка 13	3	105	35	21
Строка 14	3	100	33,3333333	26,3333333
Столбец 1	14	344	24,5714286	28,1098901
Столбец 2	14	485	34,6428571	23,9395604
Столбец 3	14	461	32,9285714	39,6098901

Дисперсионный анализ

Источник вариации	SS	df	MS	F	P-Значение	F критическое
Строки	766,5714286	13	58,967033	3,60739496	0,002601313	2,11916569
Столбцы	813	2	406,5	24,8682353	9,19788E-07	3,369016359
Погрешность	425	26	16,3461538			
Итого	2004,571429	41				

Рис. 51. Двухфакторный дисперсионный анализ с помощью MC Excel

Вопросы и задания для самоконтроля

1. Кто одним из первых использовал понятие корреляции в статистике?
2. Дайте определение корреляции.
3. Нарисуйте на графике отрицательную корреляцию.
4. Сравните линейную и нелинейную корреляции, нарисуйте их.
5. Дайте определение значимости коэффициента корреляции.
6. Сравните силу коэффициента корреляции и значимость.
7. Нарисуйте корреляционную матрицу и плеяду.
8. Приведите пример использования регрессионного анализа.
9. Сравните дискриминантный анализ и регрессионный.
10. Приведите пример использования дисперсионного анализа.

Глава 6. ВЫЯВЛЕНИЕ РАЗЛИЧИЙ МЕЖДУ ПРИЗНАКАМИ

Учебные вопросы:

6.1. Методы классификации.

6.2. Кластерный анализ, его виды.

6.3. Иерархические и интегративные методы кластерного анализа.

Выявление различий между признаками — часто встречающаяся задача в психологическом исследовании. Для выявления различий в уровне изучаемого параметра при сравнении двух и более групп используются параметрические и непараметрические критерии, описанные в теме: «Сравнения выборок. Виды статистических критериев». В данной теме уделено внимание способам, позволяющим выявить различия, рассмотрены методы классификации и кластерный анализ.

6.1. Методы классификации

Классификацией называется логическое разделение явлений и понятий, характеризующих их, на определенные классы, которые способствуют лучшему представлению специфики явлений. Чем более разнообразны явления, тем больше различий будет между ними и тем больше будет понятна специфика каждого явления

Классификация — деление объектов по классам в зависимости от выраженности признака.

Объект классификации — элементы множества.

Ступень классификации — распределение объектов одной группы.

Признак классификации — качество, лежащее в основе разделения объектов.

Деление обычно происходит от верхней ступени к нижней. Количество ступеней зависит от задач, целей, объектов классификации. Верхними ступенями обычно бывают класс и раздел, средними — группа и вид, нижними — сами объекты.

Множество приемов и способов деления называют методом классификации, а выбор метода зависит от целей классификации.

В психологии используются два метода классификации:

— иерархический;

— фасетный.

Иерархический метод классификации демонстрирует последовательное деление множества на взаимосвязанные подмножества, в котором все объекты имеют определенные степени подчинения (рис. 52).

Для построения иерархической классификации необходимо знать цель и свойства подмножеств (например, оглавление в научной работе).

Существуют следующие правила для построения иерархической классификации:

1) деление следует начинать с наиболее общих признаков;

2) в каждой градации участвует только один признак;

3) классификация должна быть последовательна от общего к частному, от большего к меньшему.

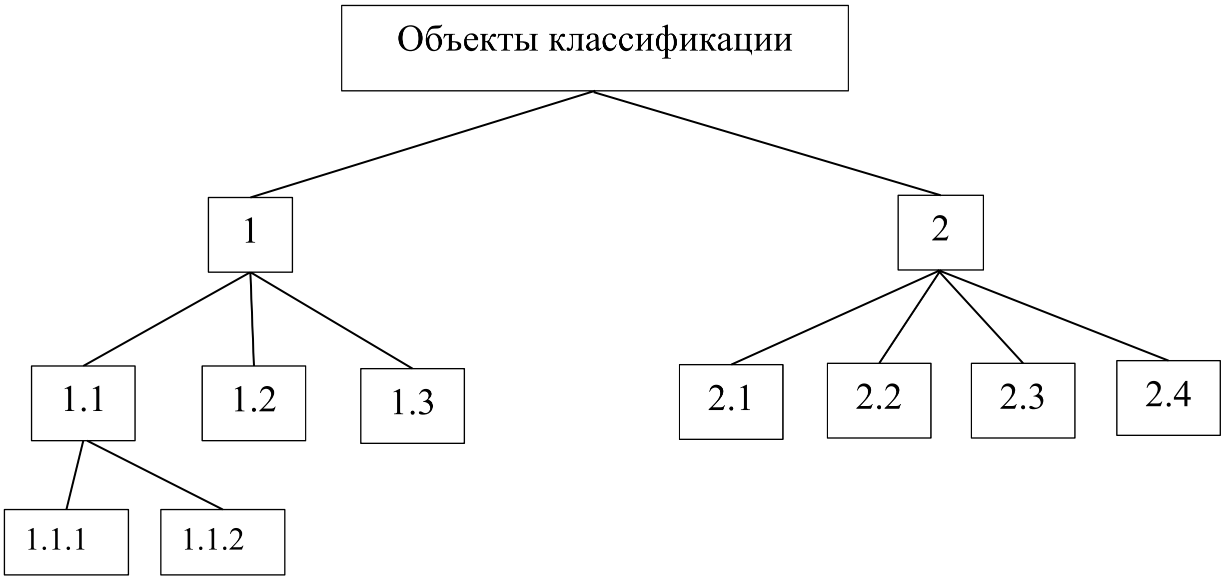


Рис. 52. Иерархический метод классификации

Иерархический метод классификации отличается близкой связью на уровне одной ступени. От выбора признаков, лежащих в основе классификации, зависит исход классификации. Число признаков и ступеней показывают глубину классификации (на рис. 52 глубина равна 3).

Фасетный метод включает в себя несколько независимых, параллельных классификаций по одному признаку, которые разделены на различные, независимые друг от друга группы или фасеты (фасет, от франц. *facette* — скошенная грань). Отдельные группы-фасеты не зависят и не входят друг в друга, как в иерархической классификации. В этой классификации используется большое количество признаков, при этом сама классификация не изменяет их группировку, считается сложнее для построения, чем иерархическая.

<table>
<tr><td colspan="2" rowspan="3"></td><td colspan="4">фасеты</td></tr>
<tr><td colspan="4">факультет</td></tr>
<tr><td>1</td><td>4</td><td>6</td><td>8</td></tr>
<tr><td rowspan="4">значения фасетов</td><td>пол</td><td>м- 20
ж-60</td><td>м-30
ж-30</td><td>м- 50
ж-50</td><td>м-60
ж-20</td></tr>
<tr><td>успеваемост ь</td><td>4,3</td><td>4,1</td><td>4,0</td><td>4,2</td></tr>
<tr><td>населенный пункт</td><td>Санкт-Петербург, Омск, Кострома…</td><td>Санкт-Петербург, Вологда, Севастопол ь</td><td>Санкт-Петербур г</td><td>Ташкент, Астана, Ашхабад, Габороне …</td></tr>
<tr><td>место проживания</td><td>общежити е, дом</td><td>общежити е, дом</td><td>общежи тие, дом</td><td>общежит ие</td></tr>
</table>

Рис. 53. Фасетный метод классификации

Для фасетного метода характерны:

— одинаковая значимость каждого признака классификации;

— независимость признаков;

— возможность увеличения признаков.

Фасетный метод обладает большей гибкостью и, в некоторых случаях, большим удобством, так как позволяет выделять и изучать только несколько фасетов, интересных для исследователя.

К недостаткам фасетного метода можно отнести независимость фасетов, сложность в определении сходства и различия между фасетами.

К недостаткам иерархического метода можно отнести объемность.

Существуют определенные *принципы классификации*, придерживаясь которых можно добиться успеха при выполнении исследовательской работы:

1) необходимо знать цель классификации;

2)выбрать метод;

3) определить количество признаков, участвующих в классификации;

4) принцип единства критерия — позволяет выделить группы одного порядка; нельзя изменять критерий в одной классификации (например, нельзя говорить, что есть курсанты умные, а есть высокие);

5) принцип соразмерности деления — классификация должна быть исчерпывающей (например, при классификации типов темперамента, несмотря на то, что в полиции практически не бывает меланхоликов, необходимо учитывать все четыре типа);

6) принцип альтернативности или взаимоисключения — выделенные признаки не могут быть включены одновременно в разные группы;

7) принцип многоступенчатости — показывает этапность ступеней классификации, что помогает выделять основные и дополнительные свойства и признаки;

8) принцип полноты ступеней классификации (например, нельзя делить курсантов на городских жителей и мужчин до 20 лет и мужчин старше 20 лет).

Все классификации можно разделить на содержательные и искусственные. В содержательных классификациях критерии показывают сущностные признаки (например, статистические критерии бывают параметрические и непараметрические). Искусственные классификации построены на несущественных признаках, но могут быть очень удобны в использовании (например, в программе excel, выбор функции построен в алфавитном порядке).

Классификации могут быть однокритериальными и многокритериальными, в последних классы могут частично пересекаться.

Особым способом классификации считается типология — объединение признаков в соответствии с определенным образцом, типом, которого в реальности, скорее всего, не существует. Необходимо определить, к какому типу в большей или меньшей степени относятся конкретные данные.

Любая классификация всегда относительна и зависит от целей, знаний автора, задач исследования.

6.2. Кластерный анализ, его виды

Кластерный анализ (cluster analysis) применяется для классификации объектов на группы так, чтобы объекты в одной группе были максимально «схожими», а в разных группах максимально «отличными» друг от друга. Такие однородные или гомогенные группы называют кластерами. Кластерный анализ — это описательная процедура, позволяющая увидеть структуру изучаемого объекта.

*Кластер*ом называют группу объектов с определенными характеристиками, высокой плотностью и дисперсией. Результаты кластерного анализа представляются в виде *дендрограммы*, графика в форме дерева (рис. 54).

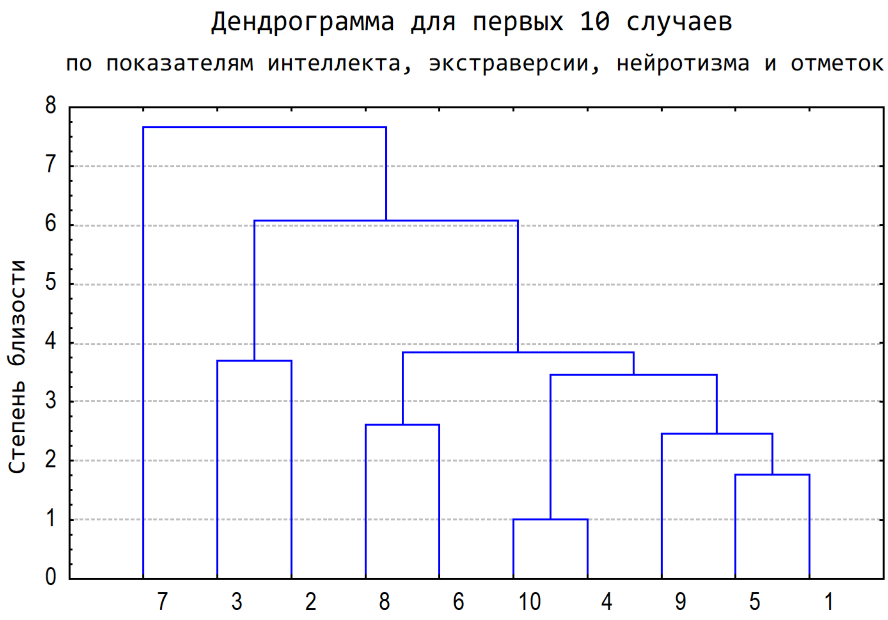


Рис. 54. Дендрограмма по четырем показателям 10 испытуемых

Кластерный анализ упорядочивает объекты по однотипным группам, основываясь на попарном сравнении с предварительно выделенными критериями, объекты каждой пары распределяются в одну или другую группу.

Первоначально все объекты исследования считаются отдельными кластерами, которые в процессе кластеризации объединяются. Объединение кластеров продолжается до того момента, когда останется один, но в любой момент можно остановиться на нужном количестве кластеров.

Задачей кластерного анализа является выявление «точек скопления» и разбиение большого множества на меньшие, однородные подмножества, его сегментация. Критерии оценки разделения на кластеры могут быть различными и зависят от задачи исследования, количество кластеров заранее невозможно предсказать.

Для того, чтобы выполнить кластерный анализ, необходимо:

1) отсутствие корреляций;

2) большое множество значений;

3) нормальное распределение;

4) отсутствие влияния на значения случайных факторов и, в связи с этим, изменение значений;

5) однородность выборки.

Когда в психологических исследованиях требуется классифицировать множество показателей, имеющих множество переменных, применяют кластерный анализ. Сегодня он выполняется в основным с помощью специальных компьютерных программ.

В психологии выделяют два вида кластерного анализа, выбор вида зависит от количества наблюдений: при небольшом объеме исследования — иерархический, при большом — анализ к-средних.

Кластерный анализ состоит из нескольких этапов:

1. Формулировка критерия разделения на кластеры;
2. Выбор способа определения расстояния между двумя точками;
3. Определения метода кластеризации;
4. Определение количества кластеров;
5. Интерпретация полученных кластеров;
6. Оценка достоверности анализа.

Кластерный анализ, или кластеризация, разделяет все данные на кластеры (группы) определенным образом, так, чтобы объекты одного кластера были похожи, а объекты разных кластеров отличны друг от друга. Число кластеров зависит от задач и может меняться в процессе исследования. Кластерный анализ позволяет обработать большой объем исследований и наглядно представить результат.

Кластерный анализ позволяет объединить объекты одновременно по нескольким параметрам.

В кластерах можно выделить некоторые свойства:

— плотность (показывает скопление или рассеивание точек в пространстве);

— дисперсия (показывает разброс точек кластера относительно его центра);

— размер (показывает «радиус» разброса относительно центра);

— форма (может быть круглой, овальной и т. д.);

— отделимость (показывает, насколько кластеры перекрывают друг друга или насколько далеко расположены друг от друга).

Кластерный анализ получил распространение в психологии в связи с широким использованием статистических программ. Сегодня данный вид анализа применяется:

— в психодиагностике — для выявления «похожих» испытуемых;

— при создании тестов — для объединения «похожих» вопросов или заданий;

— в социальной психологии можно увидеть основы формирования групп и др.

6.3. Иерархические и интегративные методы кластерного анализа

Иерархический агломеративный метод (объединяющий) (Agglomerative Nesting, AGNES) анализа данных наиболее часто используется в психологических исследованиях. Данный метод начинается с одного наблюдения, которое образовывает свой кластер. Затем в один кластер объединяют два соседних наблюдения, две точки, наиболее близко расположенные друг к другу, и так продолжается объединение соседних наблюдений до тех пор, пока не останется два кластера, которые объединяются в один.

Иерархический дивизимный метод (разделяющий) (Divisive Analysis, DIANA) диаметрально противоположен агломеративному, начинается с одного кластера, который постепенно делится на меньшие (см. рис. 52).

Интерпретация кластеров зависит от проверки *центроидов* — средних значений по каждой переменной, позволяющих описывать кластеры.

Для проведения кластерного анализа необходимо иметь показатели ряда признаков (например, у ряда испытуемых). Полученные данные можно представить в виде таблицы, а можно в виде графика, точнее, в виде точек в многомерном пространстве. Затем необходимо определить цели и задачи кластеризации, критерии деления на кластеры. Следующим этапом вычисляется расстояние между каждыми двумя точками. Точки с минимальным расстоянием объединяются в один кластер. Затем вычисляется расстояние между кластерами, при минимальных расстояниях кластеры объединяются, и так до тех пор, пока не останется один кластер. После необходимо интерпретировать полученные результаты.

Наиболее распространенными способами определения расстояния между двумя точками, «степень похожести», являются евклидово расстояние (Euclidian Distance) и квадрат евклидова расстояния (Squared Euclidian distance). Расстояние имеет смысл вычислять при условии, что признаки представлены в одинаковых единицах, иначе необходимо их нормализовать.

Евклидово расстояние — это геометрическое расстояние между двумя точками, определяемое по соответствующей формуле или с помощью статистической программы. Квадрат евклидова расстояния применяется, когда следует большее значение придать наиболее отдаленным друг от друга объектам.

Для вычисления расстояния между двумя точками также используются: взвешенное евклидово расстояние; манхэттенское расстояние (хеммингово, сити-блок расстояние или расстояние городских кварталов) высчитывается как разность по координатам; расстояние Чебышева; процент несогласия.

Когда уже выделены определенные кластеры, дальше используют методы кластеризации. *Методом кластеризации* называют способ вычисление расстояния между кластерами. Выделяют следующие методы:

— ближайший сосед (Nearest neighbor), или одиночной связи (Single Linkage). Метод начинается с нахождения наиболее близких «соседей», которые объединяются в кластер, следующий объект также объединяется с наиболее близким кластером-«соседом»;

— самый дальний сосед (Furthest neighbor), или полной связи (Complete Linkage). Этот метод противоположен методу ближайшего соседа и заключается в том, что самый дальний объект из кластера находится ближе всего к новому объекту, с которым будет объединяться, по сравнению с самыми дальними объектами других кластеров;

— межгрупповая связь (Between-groups linkage), рассчитывается как среднее значение всех расстояний между двумя точками из разных кластеров;

— внутригрупповая связь (Within-groups linkage), или средней связи (Average Linkage), вычисляется как среднее арифметическое расстояние между каждыми объектами одного и другого кластера, объединяются кластеры и меньшим расстоянием;

— центроидная кластеризация (Centroid clustering) определяет расстояние между кластерами на основе их центра тяжести;

— медианная кластеризация (Median clustering), также определяет расстояние между кластерами на основе их центра тяжести, но при этом учитывается вес каждого кластера, то есть число объектов кластера;

— метод Варда (Уорда)(Ward's method) использует дисперсионный анализ для оценки расстояния между кластерами.

Для принятия всех точек равноценными необходима нормализация значений. Для этого производится процедура взвешивания, придающая каждому значению определенный вес, показывающий его значимость в кластеризации, более значимой переменной придается больший вес.

Стандартизация данных может быть произведена с помощью z-преобразования:

$$Z = \frac{x - \bar{x}}{\sigma},$$

Z-оценки показывают положение исследуемого показателя относительно результатов всей группы, они не зависят от среднего значения и стандартного отклонения.

При выборе количества кластеров необходимо помнить, что размеры кластеров должны быть значимыми. Количество кластеров также является как важным, так и сложным вопросом — нет определенных критериев, которые помогают выделить правильное количество кластеров. В итоге, это определяет исследователь, основываясь на своем опыте и знаниях.

К *интегративным методам* группировки относят *метод k-средних (K-Means).* Целью данного метода является разделение всех данных на k кластеров, в зависимости от расположения. Для определения расположения или степени близости используют евклидово расстояние.

Необходимо определить центроиды кластеров, а затем сгруппировать объекты вокруг кластера в пределах пороговых значений. В данном методе можно задать необходимое число кластеров. На первоначальном этапе центроиды можно выбрать в зависимости от смысла или логики исследования. Каждая точка относится только к одному кластеру, даже если расстояния примерно одинаковые.

Можно описать примерный алгоритм вычисления k-средних:

1) первоначально выбираем центроиды кластеров и количество кластеров;

2) каждую точку относим к ближайшему кластеру, после заново вычисляем центр кластера;

3) кластеризация заканчивается, когда выбраны все точки и выделено заданное количество кластеров.

Данному методу присущи недостатки: при изначально неправильно выбранном количестве кластеров или центра кластеров результат будет неверным, также этот метод очень чувствителен к выбросам.

Метод k-средних рекомендуется применять для предварительного разделения на группы или для примерного вычисления оптимального количества кластеров и выявления неучтенных данных либо связей в исследовании.

Классификация предполагает упорядочивание результатов исследования в относительно однородные группы, основываясь на попарном сравнении. В результате применения кластерного анализа не происходит потери информации. Несмотря на то, что кластерный анализ известен достаточно давно, распространение в психологических исследованиях он получил лишь в конце XX века в связи с применением математико-статистических пакетов анализа данных. Кластерный анализ можно применять даже при небольшой выборке и при отсутствии нормального распределения данных.

Многомерный статистический метод позволяет провести классификацию, которая способствует более качественному описанию структуры исследуемых показателей и отношения между ними.

Вопросы и задания для самоконтроля

1. Дайте определение классификации.
2. Опишите механизмы, от которых зависит количество ступеней классификации.
3. Приведите пример иерархического метода классификации.
4. Сравните фасетный и иерархический методы классификации.
5. Приведите пример содержательной искусственной классификации.
6. Нарисуйте дендрограмму.
7. Дайте определение понятию кластера.
8. Приведите пример использования кластерного анализа.
9. Приведите пример использования иерархических методов кластерного анализа.
10. Перечислите достоинства метода k-средних.

Глава 7. МЕТОДЫ МНОГОМЕРНОГО АНАЛИЗА ДАННЫХ

Учебные вопросы:

1. Множественный регрессионный анализ, его виды.
2. Множественный дискриминантный анализ.
3. Факторный анализ.

Методы многомерного анализа данных используются для выявления скрытой структуры, латентных связей между явлениями при большом объеме информации, когда для выводов недостаточно знания одномерных статистик — среднего, дисперсии, стандартного отклонения и др. Данные методы позволяют предсказать будущие события.

Предсказание с давних времен интересовало человека, всегда привлекает возможность знать будущее. История знает множество примеров предсказаний, сбывшихся и несбывшихся, однако в прошлом они были основаны на субъективном размышлении одного человека — предсказателя. Такие прогнозы не имели и не могли иметь научного обоснования.

Сегодня прогнозы не могут быть научно не обоснованы, изменилась методология обоснования, что отразилось и на их качестве. Методы, применяемые для прогнозирования, делятся на две группы:

— эвристические, основанные на интуиции;

— математические, основаны на статистических.

Эвристические методы нельзя представить отдельно от лица, делающего прогноз, поскольку от его знаний, опыта, интуиции зависит качество и точность прогноза. К таким методам относят социологические и метод экспертных оценок. В психологических исследованиях используют метод экспертных оценок, при этом эксперты делают выводы на основе собственной интуиции и опыта. Эвристические методы используются, когда практически невозможно проанализировать множество факторов, влияющих на объект прогнозирования. Тогда применяются оценки экспертов, которые могут быть как индивидуальными, так и коллективными.

К математическим методам прогнозирования относятся: множественный регрессионный и дискриминантный анализ, факторный анализ.

7.1. Множественный регрессионный анализ, его виды

Предсказание основано на информации, на каких-то данных. Человека всегда притягивало будущее, знания рассматривались как основа предсказаний. В прошлом предсказания в основном основывались на субъективных представлениях. Современные прогнозы имеют определенную методологию обоснования.

Прогнозирование и планирование в современных условиях зависит от знания законов развития той или иной области, а также информации о настоящем положении дел.

Понятие множественной регрессии впервые использовал английский математик, статистик, биолог и философ, основатель математической статистики, один из основоположников биометрики Пирсон (Pearson) в 1908 году. Данный вид подразумевает анализ связи между независимыми переменными, которые также называются регрессорами или предикторами, и зависимой переменной. В психологических исследованиях в основном используют линейную регрессию, но она может быть и параболической, и гиперболической, и синусоидой.

Множественный регрессионный анализ выявляет и оценивает случайную связь зависимой переменной и нескольких независимых. Он оценивает коэффициент регрессии для каждой независимой переменной, при этом каждая независимая переменная воздействует на зависимую.

Для построения регрессионного уравнения необходимы данные, представленные в метрической (интервальной или абсолютной) шкале и нормальное распределение.

Пример линейной регрессии: гипотеза — курсанты первого курса более тревожны и поэтому больше времени тратят на подготовку к занятиям. Для парного регрессионного анализа необходимо найти зависимость между тревожностью (Y) и временем подготовки к занятиям (X). Линейное уравнение регрессии будет:

$$Y = b_y * X + a,$$

тревожность = b_y * время подготовки + a,

b_y — угловой или регрессионный коэффициент,

a — константа или свободный член.

Зависимость между тревожностью и временем подготовки к занятиям можно вычислить с помощью коэффициента корреляции Пирсона. Угловой коэффициент и константу можно вычислить по формулам:

$$b_y = R_{xy} * \frac{\sigma_y}{\sigma_x}$$

$$a = My — by * Mx$$

Квадрат коэффициента корреляции называется коэффициентом детерминации, который определяет взаимное влияние переменных:

$$R = \sqrt{b_x + b_y}\ ;\ R^2 = b_x * b_y$$

После решения уравнения регрессии можно построить график ожидаемой тревожности для ее прогнозирования у первокурсников.

При одномерной линейной регрессии возможно построение графика в двухмерном пространстве, при двух независимых переменных пространство будет уже трехмерным, и т. д.

Для множественной регрессии необходимо несколько независимых переменных, кроме времени подготовки к занятиям можно взять возраст, средний балл аттестата, интеллект, объем памяти и др. И тогда, после проведения исследования на достаточном объеме выборки, можно будет сказать, что является предиктором тревожности.

Регрессионное уравнение предсказывает значение зависимой переменной от одной или нескольких независимых и долю вклада каждой независимой переменной. Коэффициент детерминации показывает долю дисперсии зависимой переменной, то есть, какой процент зависимой переменной объясняется воздействием независимых переменных. Линия регрессии предсказывает показатели зависимой переменной (Y) по независимым (X). Данная линия строится относительно множества точек, которые имеют определенных разброс. Расстояние каждой точки до линии регрессии называется *остатком*.

Остаточная изменчивость может колебаться в интервале от 0 до 1. Чем меньше разброс остатков вокруг линии регрессии, тем более точным окажется прогноз. Если вычесть из единицы значение остаточной изменчивости, то получится *коэффициент детерминации,* или R-квадрат. Коэффициент детерминации равный 0,4 показывает, что 40 % изменчивости переменной Y вдоль линии регрессии объясняется действием независимых переменных, а 60 % остаточной изменчивости невозможно объяснить действием представленных независимых переменных. Значение коэффициента детерминации показывает, насколько представленные независимые переменные способны объяснить изменчивость зависимой переменной, насколько выявлены все факторы, оказывающие влияние на изменения зависимой переменной. Идеальным значением является 1, которая показывает, что представлены все факторы, объясняющие изменения зависимой переменной.

Существует несколько видов регрессионного анализа:

1. Линейная регрессия;
2. Множественная линейная регрессия;
3. Полиномиальная регрессия, или нелинейная регрессия, моделирует нелинейные взаимосвязи, отличается от линейной регрессии наличием степеней у некоторых независимых переменных ($Y = 1{,}36 + 3{,}39^2{*}X3 - 107{,}13{*}X5$);
4. Логистическая регрессия определяет зависимую переменную как 0 или 1, то есть произошло событие или нет;
5. Гребневая, или ридж-регрессия позволяет уменьшать количество данных при корреляции независимых переменных и неустойчивости оценок коэффициентов множественной линейной регрессии;
6. Лассо-регрессия (LASSO — Least Absolute Shrinkage and Selection Operator) похожа на гребневую регрессию, но использует другие способы уменьшения количества данных, и другие виды регрессии.

Множественный регрессионный анализ также можно провести с помощью специализированных статистических пакетов (например, программы SPSS (Statistical Package for the Social Sciences), статистическая программа для социальных наук (данную программу мы подробнее рассмотрим в следующей теме)).

7.2. Множественный дискриминантный анализ

Дискриминантный анализ является альтернативой множественному регрессионному анализу при зависимой переменной, представленной в ранговой или номинативной шкале и нескольких независимых переменных.

Задачей дискриминантного анализа, так же как и регрессионного, является предсказание зависимой переменной на основе нескольких независимых. Дискриминантный анализ для двух групп — зависимая переменная имеет два значения, множественный дискриминантный анализ — зависимая переменная имеет три и более значений.

Впервые дискриминантный анализ применили в экономике, в кредитном анализе, а разработал его английский статистик Рональд Фишер (1890–1962). Исследователи попытались определить, чем отличаются те, кто оплатил кредит в срок, и те, кто не сделал этого. Дискриминантный анализ позволяет построить модель, предсказывающую по определенным параметрам (предикторам), кто оплатит кредит в срок, а кто не сможет. Предикторы должны быть нормально распределенными и коррелировать с зависимой переменной, то есть с выплатой кредита.

Основной проблемой дискриминантного анализа является выбор предикторов, позволяющих разделить исследуемую группу, а затем по ним отнести конкретный показатель к одной или другой группе.

Дискриминантная функция (переменная или сочетание переменных) способствует выявлению отличий между группами и позволяет прогнозировать, к какой группе можно отнести тот или иной объект исследования. Дискриминантная функция должна отражать различия между группами — чем плотнее расположены показатели около центроида одной группы и чем дальше центроид одной группы расположен от центроида другой, тем больше будет дискриминация рассматриваемых групп.

Дискриминантный анализ основывается на уравнении регрессии при номинативной зависимой переменной. Дискриминантная функция имеет вид:

$$D = b_0 + b_1x_1 + b_2x_2 +..+ b_kx_k,$$

где x — значения независимых переменных,

b — коэффициенты дискриминантного анализа.

Данная функция основана на информации об объектах, составляющих ту или иную группу до проведения анализа, и позволяет отнести новый объект, на основе известных предикторов, к той или иной группе.

Дискриминантные переменные, или предикторы — признаки, позволяющие отнести объект к одной из групп. Количество дискриминантных функций минимум на единицу должно быть меньше количества дискриминантных переменных, но при этом не больше количества групп.

Для проведения дискриминантного анализа достаточно важно:

— нормальное распределение дискриминантных переменных;

— однородность дисперсий в группах;

— наличие корреляций в группах.

Дискриминантный анализ способствует решению следующих задач:

1. Выявление и интерпретация различий между группами; какие переменные наиболее точно выявляют различия;

2. Классификация, то есть отнесение параметра к той или иной группе.

Дискриминантная функция, или прогностическое уравнение, позволяет классифицировать экспериментальные данные.

Для дискриминантного анализа рекомендуется:

1. Количество предикторов в 10–20 раз меньше, чем выборка;

2. Количество предикторов не менее, чем в 5 раз меньше, чем объектов в каждой группе;

3. Количество предикторов не более 10;

4. Исследование выбросов, влияющих на результат;

5. Учитывать корреляцию переменных.

Дискриминантный анализ — один из видов многомерного статистического анализа, который классифицирует объекты в зависимости от сходства или различия с той или иной группой, на основе выделенных предикторов.

В итоге формулируется правило, которое позволяет относить каждый новый объект к одной из групп на основе величины дискриминантной функции.

Выделяют четыре этапа дискриминантного анализа:

1. Разделение выборки на части:

выделяются две выборки, одна — известная, на основе которой составлены группы, вторая — экспериментальная, которую необходимо разделить на группы.

2. Отбор предикторов:

происходит на основе корреляционной матрицы и включает в себя средние коэффициенты корреляции для каждой формируемой группы, вычисляются ковариационные матрицы для всей выборки и для каждой группы, также можно применить однофакторный дисперсионный анализ для различения групп.

3. Вычисление коэффициентов дискриминантного анализа:

выполняется с помощью статистических пакетов анализа данных, основан на минимизации коэффициента Уилкса (λ-лямбда Уилкса (Вилкса)) после анализа с помощью регрессионного уравнения нового предиктора; коэффициент λ показывает отношение внутригрупповой суммы квадратов к общей сумме квадратов, характеризует долю влияния предиктора на дисперсию критерия; со значением λ связаны величины F-критерия Фишера и p-уровня значимости, показывающие, входит ли данный предиктор в уравнение регрессии, если:

— $F > 3{,}84$, то необходимо включать предиктор в уравнение регрессии;

— $F < 2{,}71$, то предиктор не включается в уравнение.

4. Интерпретация результатов:

основываясь на составленном уравнении регрессии, в основе которого лежат показатели предикторов и критерия, вычисленные в известной выборке, зная значения предикторов, можно вычислить критерий для неизвестной выборки, позволяющий каждое значение неизвестной выборки отнести к той или иной группе.

Множественный дискриминантный анализ позволяет прогнозировать значение зависимой переменной, выраженной в номинативной или ранговой шкале. Также дискриминантный анализ позволяет классифицировать данные на группы в соответствии с градацией зависимой переменной. Допускается, что часть данных невозможно отнести ни к одной из групп, такие данные остаются «неизвестными».

7.3. Факторный анализ

Считается, что факторный анализ начинается с разработанного Карлом Пирсоном метода главных компонент, основанного на трудах английского исследователя Френсиса Гальтона (1822–1911), посвященных основам проведения психологических исследований, обработки и интерпретации полученных результатов. В 1889 году Пирсон ввел в научный оборот понятие корреляции.

Чарльза Эдуарда Спирмена (1863–1945), профессора Лондонского и Честерфилдского университетов, разработчика многочисленных методик математической статистики, называют основателем современного факторного анализа. Именно он в 1904 году в психологическом журнале опубликовал статью «Общий интеллект, объективно определенный и измеренный» («General Intelligence Objectively Determined and Measured»), где утверждалось, что интеллект можно объяснить общим (General) и специфическим (Specific) факторами, при этом общий фактор объясняет начало или основу интеллектуальных действий, а специфическийовывается? Факторный — отдельные действия.

Факторный анализ способствует сокращению количества переменных, основываясь на определении структуры связей между ними. В итоге все переменные объединяются в несколько групп, которые называются факторами.

Факторный анализ возможен на основе корреляционной матрицы. Его цель — выявление латентных, внутренних процессов, являющихся причиной взаимовлияний.

В задачи факторного анализа входит либо сокращение количества переменных, либо классификация переменных на основе выявления их структуры взаимосвязей.

Факторы, выделенные с помощью факторного анализа, в сжатом виде содержат информацию о взаимосвязях между переменными. Обратно из факторов извлечь переменные уже невозможно.

В один фактор входят переменные, имеющие сильные корреляционные взаимосвязи, которые фактор способен объяснить. Фактор — это статистический показатель, выделенный искусственно в результате преобразований корреляционной матрицы. Такие преобразования называются *факторизацией*.

Факторный анализ состоит из нескольких этапов:

1. Составление корреляционной матрицы;
2. Извлечение факторов;
3. Вращение факторов;
4. Интерпретация факторов.

Также необходимо учитывать условия, необходимые для проведения факторного анализа:

— количественные данные;

— нормальное распределение показателей;

— независимые значения показателей;

— линейные корреляции;

— корреляции по модулю выше 0,3;

— выборка не менее 100 испытуемых, или не менее чем в 2 раза больше количества переменных.

Качественным критерием правильности и адекватности выполнения факторного анализа считается процент дисперсии, описывающий переменные, входящие в выделенные факторы. Факторы, выделенные в процессе факторизации корреляционной матрицы, как правило, неравнозначны, то есть имеют разный вес.

Факторный вес определяется в результате проведения факторного анализа и обозначает степень проявления отдельного признака в структуре каждого фактора. Факторные веса позволяют ранжировать показатели в каждом факторе.

Факторная нагрузка — величина коэффициентов корреляции первоначального показателя с каждым выделенным фактором — чем больше факторная нагрузка, тем больший вклад вносит показатель в данный фактор. Факторная нагрузка может иметь положительный и отрицательный знак (как и корреляции) и отражает силу связи показателя с фактором. Квадрат факторной нагрузки показывает процент дисперсии, объясняемой данной переменной, а сумма дисперсий всех факторных нагрузок описывает всю дисперсию, объясняемую фактором.

Выделяют несколько методов факторного анализа, которые на основе математической модели описывают взаимосвязи между первоначальными показателями и выделенными факторами.

В психологии наиболее популярным методом является *метод главных компонент* (principal components analysis, PCA), который предполагает независимость выделенных факторов. В геометрическом выражении он сводится к построению новых осей (факторов) на основе геометрической близости точек, связан с дисперсией и математическим ожиданием (средним значением случайной величины). Данный метод заменяет коррелированные показатели некоррелированными факторами, для интерпретации остаются только главные компоненты.

В зависимости от задач, стоящих перед исследователем, выделяют два вида факторного анализа — эксплораторный и конфирматорный. В психологических исследованиях наиболее распространен эксплораторный.

Эксплораторный, или разведывательный, направлен на выделение общих факторов, то есть поиск общего показателя, связанного с другими, коррелирующими между собой. Таким образом, можно уменьшить количество изучаемых переменных. Исследователь заранее не знает ни количество факторов, ни процентное соотношение общей дисперсии в каждый фактор.

Конфирматорный, или подтверждающий, факторный анализ применяется для проверки эффективности модели при известном количестве факторов. Данный анализ предполагает наличие гипотезы о количестве факторов и показателей, входящих в каждый фактор, пытается доказать, что выделенные показатели действительно зависят от латентной переменной — выделенного фактора.

Конфиматорный анализ используется, например, в психодиагностике для проверки пунктов теста, разработанных для определенной шкалы, для определения валидности, то есть действительно ли изменения данных пунктов связаны с изменением шкалы.

Перед тем, как проводить факторный анализ, необходимо понять, можно ли использовать матрицу корреляций для факторизации. Для этого рассчитывается значенияе критеря адекватности выборки Кайзера-Мейера-Олкина, который показывает:

> 0,9 — безусловная адекватность;

> 0,8 — высокая адекватность;

> 0,7 — приемлемая адекватность; > более 0,6 — удовлетворительная адекватность;

> 0,5 — низкая адекватность;

< 0,5 — факторный анализ неприменим к выборке. Критерий сферичности Бартлетта показывает нормальность распределения переменных и проверяет, отличаются ли корреляции от нуля. Значение $p < 0{,}05$ означает, что с данными можно проводить факторный анализ.

В психологических исследованиях чаще применяют эксплораторный факторный анализ. Разберем его на примере[1].

Проводилось исследование 206 сотрудников правоохранительных органов по методикам: опросник темперамента и характера Р. Клонингера; методика «Семантический дифференциал времени», методика определения сложной сенсомоторной реакции (ССМР); методика определения силы нервной системы (теппинг-тест); методика определения воспроизведения временных интервалов «Индивидуальная минута»; субшкала теста Векслера, Гарвардский степ-тест, тест руки Вагнера, САН. Всего было выделено 95 показателей, посредствам которых был проведен корреляционный анализ. На рисунке 55 представлена корреляционная матрица части исследуемых показателей (представлены субшкалы опросника Клонингера, субшкала теста исследования интеллекта Векслера, методика определения воспроизведения временных интервалов «Индивидуальная минута», часть показателей Гарвардского степ-теста, стабилометрическая платформа.

Затем с помощью статистической программы IBM SPSS Statistics был проведен факторный анализ методом главных компонент. С целью упрощения описания столбцов факторной матрицы применялось вращение по методу варимакс. Этим удалось достичь относительно простой структуры группировки показателей, которые были сгруппированы в шесть факторов.

[1] Баринова М. Г., Зуева Е. Г. Психофизиологическая готовность сотрудников правоохранительных органов к деятельности в экстремальных ситуациях: монография. СПб.: Изд-во СПб ун-та МВД России, 2019. 164 с.

Корреляці	ОСЛТ Кло	ПН2	ПН3	ИО1	ИО2	ИО3	ИО4	ЗП1	ЗП2	К1	К4	С1
	0,077	0,14	0,572	0,082	0,054	0,298	0,04	0,24	0,009	0,094	0,661	0,26
	165	165	165	165	165	165	165	165	165	165	165	16
ОСЛТ Кло	1	0,064	,163*	-,197*	-,290**	-0,147	-0,152	0,085	0,051	0,025	0,14	,185*
		0,412	0,037	0,011	0	0,059	0,051	0,277	0,513	0,753	0,074	0,01
	165	165	165	165	165	165	165	165	165	165	165	16
ПН2	0,064	1	0,094	0,066	0,048	-0,064	0,147	-0,002	0,132	-0,103	-0,068	-0,10
	0,412		0,229	0,399	0,539	0,417	0,06	0,982	0,09	0,189	0,387	0,1
	165	165	165	165	165	165	165	165	165	165	165	16
ПН3	,163*	0,094	1	-0,086	-0,103	-0,028	0,032	-0,079	0,109	0,091	-0,009	0,1
	0,037	0,229		0,273	0,187	0,718	0,687	0,315	0,165	0,246	0,912	0,07
	165	165	165	165	165	165	165	165	165	165	165	16
ИО1	-,197*	0,066	-0,086	1	,223**	,226**	,226**	-0,044	-,159*	-0,092	-,185*	-0,14
	0,011	0,399	0,273		0,004	0,004	0,004	0,571	0,041	0,238	0,017	0,0
	165	165	165	165	165	165	165	165	165	165	165	16
ИО2	-,290**	0,048	-0,103	,223**	1	,160*	,305**	,188*	0,085	-0,054	-0,005	-,205**
	0	0,539	0,187	0,004		0,04	0	0,016	0,28	0,487	0,95	0,00
	165	165	165	165	165	165	165	165	165	165	165	16
ИО3	-0,147	-0,064	-0,028	,226**	,160*	1	,244**	0,009	-,188*	-0,12	-0,093	-,191*
	0,059	0,417	0,718	0,004	0,04		0,002	0,913	0,016	0,125	0,234	0,01
	165	165	165	165	165	165	165	165	165	165	165	16
ИО4	-0,152	0,147	0,032	,226**	,305**	,244**	1	0,078	-0,022	0,036	-0,053	-,277**
	0,051	0,06	0,687	0,004	0	0,002		0,317	0,777	0,642	0,496	
	165	165	165	165	165	165	165	165	165	165	165	16
ЗП1	0,085	-0,002	-0,079	-0,044	,188*	0,009	0,078	1	0,15	0,079	0,118	-0,13
	0,277	0,982	0,315	0,571	0,016	0,913	0,317		0,055	0,312	0,132	0,09
	165	165	165	165	165	165	165	165	165	165	165	16
ЗП2	0,051	0,132	0,109	-,159*	0,085	-,188*	-0,022	0,15	1	,161*	,181*	0,05
	0,513	0,09	0,165	0,041	0,28	0,016	0,777	0,055		0,039	0,02	0,48
	165	165	165	165	165	165	165	165	165	165	165	16
К1	0,025	-0,103	0,091	-0,092	-0,054	-0,12	0,036	0,079	,161*	1	,288**	0,03
	0,753	0,189	0,246	0,238	0,487	0,125	0,642	0,312	0,039		0	0,63
	165	165	165	165	165	165	165	165	165	165	165	16
К4	0,14	-0,068	-0,009	-,185*	-0,005	-0,093	-0,053	0,118	,181*	,288**	1	,162*
	0,074	0,387	0,912	0,017	0,95	0,234	0,496	0,132	0,02	0		0,03
	165	165	165	165	165	165	165	165	165	165	165	16
С1	,185*	-0,107	0,14	-0,147	-,205**	-,191*	-,277**	-0,131	0,055	0,037	,162*	

Рис. 55. Корреляционная матрица исследования

Таблица 17. Факторы, отражающие психофизиологическую готовность сотрудников правоохранительных органов в экстремальных ситуациях

№ п/п	*Факторы*	*% общей дисперсии*
1	Актуальное функциональное состояние	14,77
2	Сотрудничество	13,067
3	Сила нервной системы	11,671
4	Позитивное восприятие прошлого	9,961
5	Подвижность нервной системы	9,854
6	Стрессоустойчивость	8,384
Процент общей дисперсии		67,71 %

Для лучшего распределения дисперсии вдоль выделенного фактора производят вращение факторов. Вращение бывает *ортогональным*, когда сохраняется прямоугольная система координат. В психологии наиболее популярным является варимакс-вращение. Метод варимакс — ортогональное вращение факторов, уменьшающее количество переменных с высокими факторными нагрузками, и, соответственно, способствующее более понятной интерпретации фактора. Вращение также может быть *косоугольным* (неортогональным), когда не сохраняется прямоугольная система координат, и факторы коррелируют между собой. Метод прямой облимин и промакс представляют косоугольное вращение. Существуют также другие типы вращений.

Каждый фактор отражает процент общей дисперсии, таким образом выделенные шесть факторов описывают почти 68 % общей дисперсии. Если выделенные факторы объясняют около 70 % общей дисперсии, а следующий (в нашем примере седьмой) фактор менее 5 %, то такая факторная структура считается хорошей.

Таким образом, из 95 исследуемых показателей первоначально были отброшены показатели, у которых отсутствовали корреляционные связи, затем был проведен эксплораторный факторный анализ, в ходе которого было выделено шесть факторов, куда вошли 65 показателей.

В таблице 18 представлена факторная нагрузка по каждому показателю, входящему в данный фактор. Наибольшую факторную нагрузку имеют показатели сложной сенсомоторной реакции: визуальный стимул слова — ложные нажатия — 0,763; визуальный стимул слова — пропуски — 0,740. Следовательно, показатели сложной сенсомоторной реакции наиболее сильно коррелируют с данным фактором, эти показатели отражают актуальное функциональное состояние, поэтому выделенный фактор был так назван.

Также в таблице 18 видно, что все входящие в фактор показатели обладают сильной или средней силы корреляцией с данным фактором. Один из показателей обладает отрицательной корреляцией с данным фактором, что говорит о том, что испытуемые не склонны избегать опасности, скорее наоборот, готовы к встрече с ней.

Таблица 18. Фактор 1. Актуальное функциональное состояние (дисперсия — 14,78 %)

Показатели	*Факторная нагрузка*
Семантический дифференциал времени	
Эмоциональная окраска времени в прошлом	0,557
Величина времени в будущем	0,689
Эмоциональная окраска времени в будущем	0,538
Эмоциональная окраска времени в настоящем	0,622
Величина времени в прошлом	0,602
Величина времени в будущем	0,689
Структура времени в настоящем	0,644
САН	
Самочувствие	0,596
Настроение	0,607
Опросник темперамента и характера Р. Клонингера (TCI)	
Избегание опасности	–0,618
Методика определения параметров сложной сенсомоторной реакции	
Визуальный стимул слова — время реакции	0,713
Визуальный стимул слова — ошибки	0,513
Визуальный стимул слова — ложные нажатия	0,763
Визуальный стимул слова — пропуски	0,740

Факторный анализ является методом, который используют для уменьшения размерности или классификации данных. В результате применения получаем два показателя: 1) процент объясняемой дисперсии и 2) факторные нагрузки. В психологии факторный анализ может применяться:

— для изучения структуры психологических компонентов (личности, темперамента, способностей и т. д.), то есть того, что нельзя непосредственно измерить;

— при конструировании тестов;

— проверке психометрических свойств опросников для его адаптации, также для выявления значения понятий (семантический дифференциал).

Методы многомерного анализа данных способствуют выявлению скрытых закономерностей или значимых взаимосвязей между переменными с помощью различных видов анализа.

В данной теме мы рассмотрели множественный регрессионный анализ, множественный дискриминантный анализ и факторный анализ. Различие используемых методов можно объяснить многообразием психологических явлений. Представленные методы пытаются выявить и объяснить скрытые причинно-следственные связи.

Вопросы и задания для самоконтроля

1. Приведите пример использования дискриминантного анализа.
2. Дайте определение понятию «предиктор».
3. Дайте определение регрессионному анализу.
4. Сравните эвристические и математические методы прогнозирования.
5. Сопоставьте линейную и множественную регрессию.
6. Перечислите цели факторного анализа.
7. Дайте определение понятию «факторные нагрузки».
8. Перечислите этапы факторного анализа.
9. Предложите задачи, в решении которых помогает метод главных компонент.
10. Приведите пример использования факторного анализа.

Глава 8. МАТЕМАТИЧЕСКИЕ ПАКЕТЫ СТАТИСТИЧЕСКОГО АНАЛИЗА ДАННЫХ

Учебные вопросы:

1. Компьютерные системы для анализа данных.

2. Виды статистических пакетов.

3. Общая информация о пакете SPSS и его структуре .

Статистические данные в психологических исследованиях всегда собирались и анализировались. Первоначально сбор и анализ проводился на бумаге, что ограничивало скорость, время, объем данных. Работу психологов упрощают математические пакеты статистического анализа данных.

Сегодня практически у каждого школьника есть компьютер, который позволяет быстро обрабатывать большой объем информации, а психологам — использовать современные математико-статистические программы. В настоящее время используется большое разнообразие как специализированных статистических, так и универсальных пакетов обработки данных.

Использование специализированных статистических пакетов помогает обрабатывать и интерпретировать экспериментальные данные. Каждый статистический пакет обладает своей спецификой, и для правильной интерпретации полученных данных необходимо знать теоретические основы математико-статистических методов.

8.1. Компьютерные системы для анализа данных

Математические пакеты статистического анализа данных широко распространены, выбор определенного пакета зависит от характера решаемых задач, объема экспериментального материала, опытности пользователя. Наиболее популярными в психологических исследованиях являются:

— электронные таблицы MicroSoft Excel (MS Excel);

— программный пакет Statistica;

— программа SPSS;

— программа JASP, бесплатная, созданная для студентов.

Самым простым вариантом обработки большого массива данных являются таблицы Microsoft Excel, которые совместимы с другими статистическими программами. Однако данные таблицы имеют большие ограничения и не позволяют обработать, например, данные, полученные в ранговых шкалах. Достоинства Microsoft Excel заключается в простоте и относительно быстрой обработке небольшого объема данных, а также графическом представлении данных. Одним из недостатков является невозможность использования непараметрических критериев и неметрических шкал. Таблицы Microsoft Excel позволяют подсчитать, например, среднее значение, стандартное отклонение, t-критерий Стьюдента для зависимых и независимых выборок, провести дисперсионный и корреляционный анализ.

Однако расчеты, сделанные при помощи таблицы MS Excel, не признаются авторитетными журналами, в них также сложно построить научные графики.

Для обработки очень большого объема данных или с помощью специфического метода анализа пользуются профессиональными статистическими пакетами, MS Excel для таких расчетов не используется.

Инструменты профессиональной математической статистики можно разделить на две группы:

— профессиональные;

— непрофессиональные статистические расчеты (MS Excel).

Среди профессиональных в настоящее время можно выделить универсальные и узкоспециализированные пакеты анализа данных.

Универсальные статистические пакеты (общего назначения или профессиональные) включают в себя различные статистические методы анализа и способность обрабатывать большой объем данных. В универсальных пакетах отсутствует ориентация на специфическую область и предлагается различные неспецифические виды анализа данных.

Рассмотрим наиболее популярные универсальные статистические пакеты. Уместно будет начать с языка программирования R, позволяющего создавать статистические программы и специализированные пакеты. Этот язык был создан в 1993 году, в университете Окленда, Новая Зеландия, он представляет набор R-функций, справочной информации и примеров, собранных вместе, к тому же это бесплатный продукт.

На основе языка программирования R создано множество статистических пакетов, наиболее популярными являются SPSS, SAS и MatLab, R наиболее используемый программный продукт в научных публикациях с помощью статистического анализа.

Все профессиональные статистические программы можно разделить на три вида:

1) с удобным графическим интерфейсом («нажми сюда и получишь вот это»);

2) с разными языками программирования (нужны навыки программирования);

3) смешанные, в которых есть и графический интерфейс, и возможность создания программ (SAS, STATA, Rcmdr).

8.2. Виды статистических пакетов

Рассмотрим статистические пакеты с графическим интерфейсом, которые делятся на платные и бесплатные. К платным относятся SPSS, STATISTICA, SAS и др., к бесплатным — Past, PSPP, и другие.

Наиболее распространены в психологических исследованиях SPSS, Statistica, Statgraphics, Stata, S-plus, SAS, Deductor, Prognoz Platform и др.

Существуют и профессиональные специализированные пакеты, использующиеся в конкретной предметной области, которые включают в себя небольшое количество статистических методов. Наиболее известные: BioStat, EQS, ЭВРИСТА, GWR4, GeoDA, Arrow Model и др.[1] Каждый пакет ориентирован на определенную область, и с ними могут работать специалисты, хорошо знакомые с таким анализом данных.

Рассмотрим некоторые статистические пакеты более подробно.

SAS создавался в 1960-х годах для сельского хозяйства, сегодня актуальной является 10-я версия, которая применяется, в том числе, и в психологических исследованиях. Она состоит из отдельных модулей, каждый из которых решает определённую задачу. В итоге, пользователь собирает нужные ему модули и решает поставленные статистические задачи.

Stata сегодня имеет 17-ю версию, первая была создана в 1985 году как инструмент анализа научных исследований. Также включает различные модули, содержащие различные наборы средств статистического анализа, совместима с операционными системами Windows, Macintosh и Unix (включая Linux). Может экспортировать результаты в MS Office и SAS. Одним из недостатков является отсутствие совместимости с электронными таблицами, в частности с MS Excel.

Для статистического анализа с 1984 года компания StatSoft разрабатывает программный пакет *Statistica*, предлагающий функции анализа данных, управления данными, добычи данных, визуализации данных с привлечением статистических методов. Компания создала

[1] Цыпин А. П., Сорокин А. С. Статистические пакеты программ в социально-экономических исследованиях // АНИ: экономика и управление. —– 2016. — № 4 (17). — С. 379–384.

различные варианты пакета в зависимости от целей и задач исследователя. Первоначально данная программа была в виде модуля электронной таблицы Lotus. Самостоятельной статистической программой Statistica стала в 1991 году и с тех пор пользуется популярностью. Кроме универсального статистического и графического интерфейса компания создала специализированные модули.

Основные преимущества программы Statistica:

— большие возможности в выборе математических методов;

— графическая интерпретация результатов;

— совместимость с другими программами (например, MS Excel);

— возможность расширение пакета;

— наличие встроенного языка программирования.

Последней версией программы Statistica на 2022 год считается 13-я. Эта программа работает только в Windows и позволяет добавить собственную панель инструментов.

Пакет прикладных программ *SPSS — Statistical Package for Social Science*, работает в операционной системе MS Windows, создан компанией SPSS Inc.

IBM SPSS Statistics на 2022 год имеет 23 русскоязычную версию, и 25 англоязычную, первая англоязычная версия была разработана в 1968 году, версия под Windows вышла в 1992 году, сегодня также есть версии для Mac OS X, Linux и Android.

IBM SPSS Statistics на сегодняшний день является одним из лидеров среди универсальных статистических пакетов, считается гибким и мощным.

Основные преимущества пакета SPSS Statistics:

— большие возможности в выборе математических методов,

— позволяет объединять несколько баз данных,

— выполняет отчеты, позволяющие рассмотреть результаты с разных точек зрения,

— есть встроенный язык программирования,

— поддерживает связь с большинством форматов данных и обмен данными с другими приложениями MS Windows.

Past (Paleontological Statistics Software for Education and Data Analysis) — бесплатное программное обеспечение, предназначенное для анализа статистических данных и их графического представления, имеет достаточно простой интерфейс на английском языке, совместима с электронными таблицами. Имеет достаточно ограниченные методы анализа данных, которые увеличиваются с выпуском новой версии.

PSPP — данный пакет первоначально назывался Fiasco, был создан в конце 1990-х годов как бесплатная альтернатива SPSS, но, к сожалению, его функционал более ограничен. PSPP статистический пакет совместим с MS Excel, имеет англоязычную версию.

STADIA — советская, а позже российская статистическая система, с 1991 года поддерживается НПО «Информатика и компьютеры» при активном участи ведущих специалистов МГУ имени М.В. Ломоносова. В 1977 году был создан первый прототип данной статистической системы, включающий в себя описательную статистику, критерии Пирсона, Стьюдента, Фишера, Спирмена, Вилкоксона, χ^2, гистограмму, одно- и двухфакторный дисперсионный анализ, простую и множественную линейную регрессию, авто- и кросс-корреляционные функции и кластерный анализ. Последние версии подходят для Windows, на 2022 год это 8-я версия, является платной. Данный пакет можно использовать в различных научных областях. Представлен в трех вариантах study, base и prof, отличающихся набором функций и ценой.

Пакет STADIA обладает следующими возможностями:

— представлены различные виды анализа — описательная статистика, дисперсионный, корреляционный, регрессионный, кластерный, факторный анализ и другие виды;

— различные виды графики — функции, диаграммы, гистограммы и др.;

— импорт/экспорт данных и результатов.

8.3. Общая информация о пакете SPSS и его структуре

Пакет SPSS (Statistical Package for Social Science — статистический пакет для социальных наук) сегодня принадлежит компании IBM, поэтому правильным полным названием будет IBM SPSS Statistics.

Данный пакет был разработан в Чикагском университете в 1968 году и сегодня представлен мощной системой статистического анализа и графического представления данных, облегчающем обработку данных научного исследования. Пакет структурно состоит из модулей. Базовый модуль содержит в себе наиболее распространенные виды статистического анализа. Его интерфейс представлен на рисунке 56.

На рисунке представлена вкладка данных, куда можно вставить показатели, или через файл → открыть → данные загрузить, например, готовую электронную таблицу MS Excel. На рисунке 57 представлена вторая вкладка, которая называется «переменные», где можно вставить/изменить названия переменных, вставляя их без пробела.

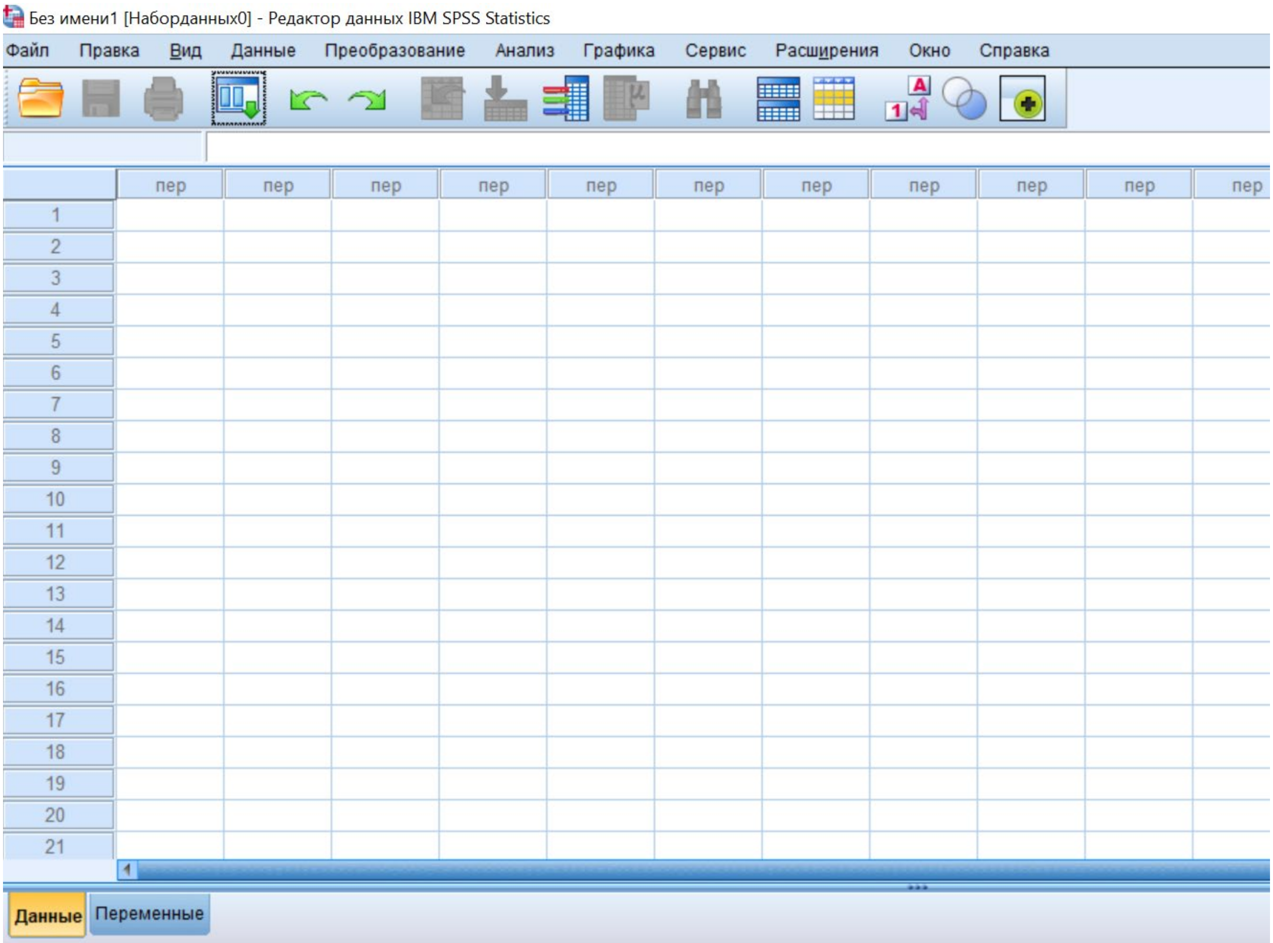
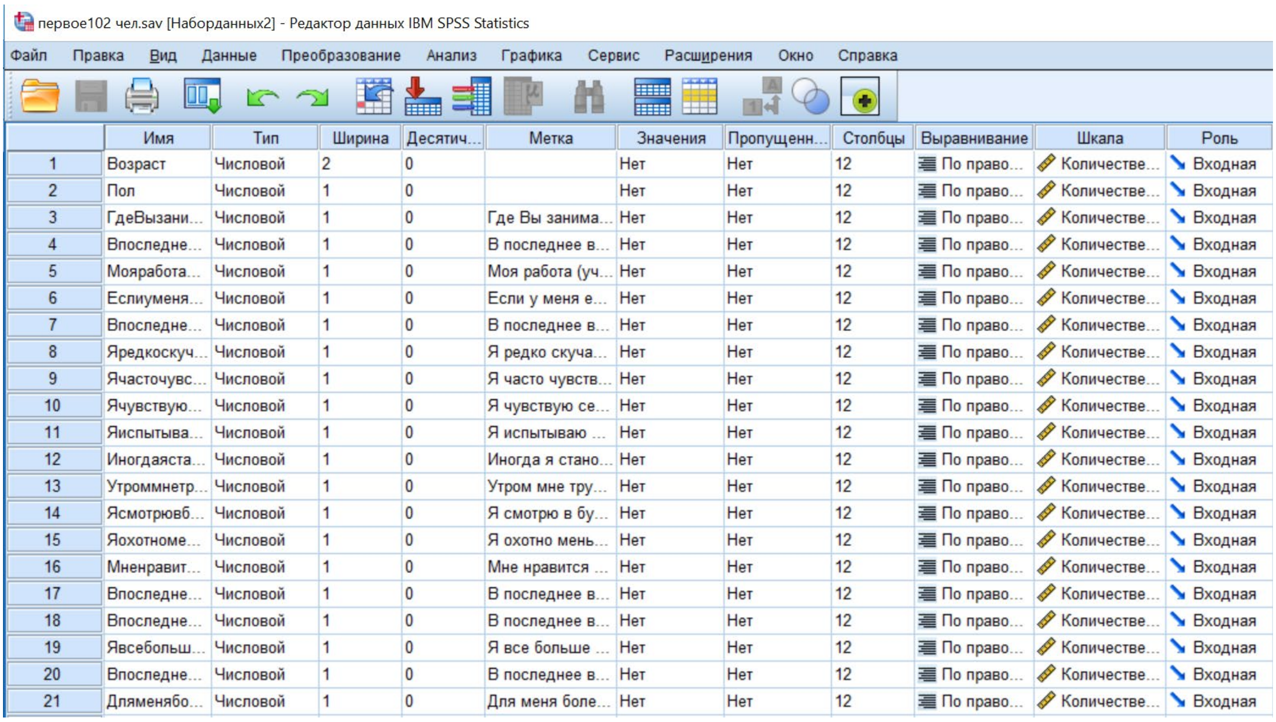


Рис. 56. Пакет IBM SPSS Statistics на русском языке перед вводом данных



	Имя	Тип	Ширина	Десятич...	Метка	Значения	Пропущенн...	Столбцы	Выравнивание	Шкала	Роль
1	Возраст	Числовой	2	0		Нет	Нет	12	По право...	Количестве...	Входная
2	Пол	Числовой	1	0		Нет	Нет	12	По право...	Количестве...	Входная
3	ГдеВызани...	Числовой	1	0	Где Вы занима...	Нет	Нет	12	По право...	Количестве...	Входная
4	Впоследне...	Числовой	1	0	В последнее в...	Нет	Нет	12	По право...	Количестве...	Входная
5	Моляработа...	Числовой	1	0	Моя работа (уч...	Нет	Нет	12	По право...	Количестве...	Входная
6	Еслиуменя...	Числовой	1	0	Если у меня е...	Нет	Нет	12	По право...	Количестве...	Входная
7	Впоследне...	Числовой	1	0	В последнее в...	Нет	Нет	12	По право...	Количестве...	Входная
8	Яредкоскуч...	Числовой	1	0	Я редко скуча...	Нет	Нет	12	По право...	Количестве...	Входная
9	Ячасточувс...	Числовой	1	0	Я часто чувств...	Нет	Нет	12	По право...	Количестве...	Входная
10	Ячувствую...	Числовой	1	0	Я чувствую се...	Нет	Нет	12	По право...	Количестве...	Входная
11	Яиспытыва...	Числовой	1	0	Я испытываю ...	Нет	Нет	12	По право...	Количестве...	Входная
12	Иногдаяста...	Числовой	1	0	Иногда я стано...	Нет	Нет	12	По право...	Количестве...	Входная
13	Утроммнетр...	Числовой	1	0	Утром мне тру...	Нет	Нет	12	По право...	Количестве...	Входная
14	Ясмотрювб...	Числовой	1	0	Я смотрю в бу...	Нет	Нет	12	По право...	Количестве...	Входная
15	Яохотноме...	Числовой	1	0	Я охотно мень...	Нет	Нет	12	По право...	Количестве...	Входная
16	Мненравит...	Числовой	1	0	Мне нравится ...	Нет	Нет	12	По право...	Количестве...	Входная
17	Впоследне...	Числовой	1	0	В последнее в...	Нет	Нет	12	По право...	Количестве...	Входная
18	Впоследне...	Числовой	1	0	В последнее в...	Нет	Нет	12	По право...	Количестве...	Входная
19	Явсебольш...	Числовой	1	0	Я все больше ...	Нет	Нет	12	По право...	Количестве...	Входная
20	Впоследне...	Числовой	1	0	В последнее в...	Нет	Нет	12	По право...	Количестве...	Входная
21	Дляменябо...	Числовой	1	0	Для меня боле...	Нет	Нет	12	По право...	Количестве...	Входная

Рис. 57. Пакет IBM SPSS Statistics на русском языке, вкладка переменные

Во вкладке переменные для статистических вычислений обязательно необходимо проверить столбик с названием «тип», он должен быть числовой; десятичные показывают количество знаков после запятой, в представленном примере (рис. 57) в столбце «десятич…» стоит 0, то есть в данных представлены только целые числа; «шкала» представлена тремя вариантами: количественная, порядковая и номинальная, вычисления параметрических данных осуществляются в количественной шкале, иначе программа напишет, что не сможет вычислить; остальные показатели не влияют на вычисления, а скорее влияют на внешний вид.

На рисунке 58 видно, как с помощью инструмента «Анализ» можно вычислить, например, t-критерий Стьюдента как для зависимых, так и для независимых переменных, одновыборочный (сравнение реальных данных с предполагаемыми), а также однофакторный дисперсионный анализ. Также на этом рисунке видно, что в данный пакет входит множество других видов анализа данных.

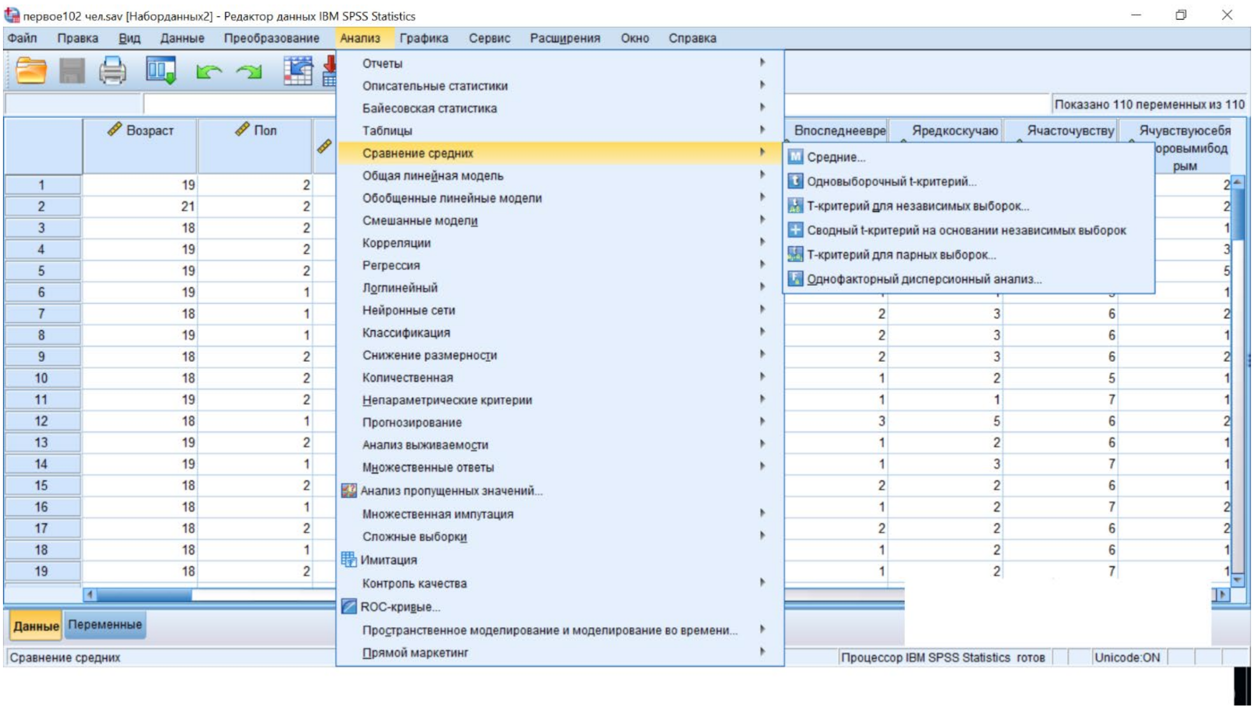
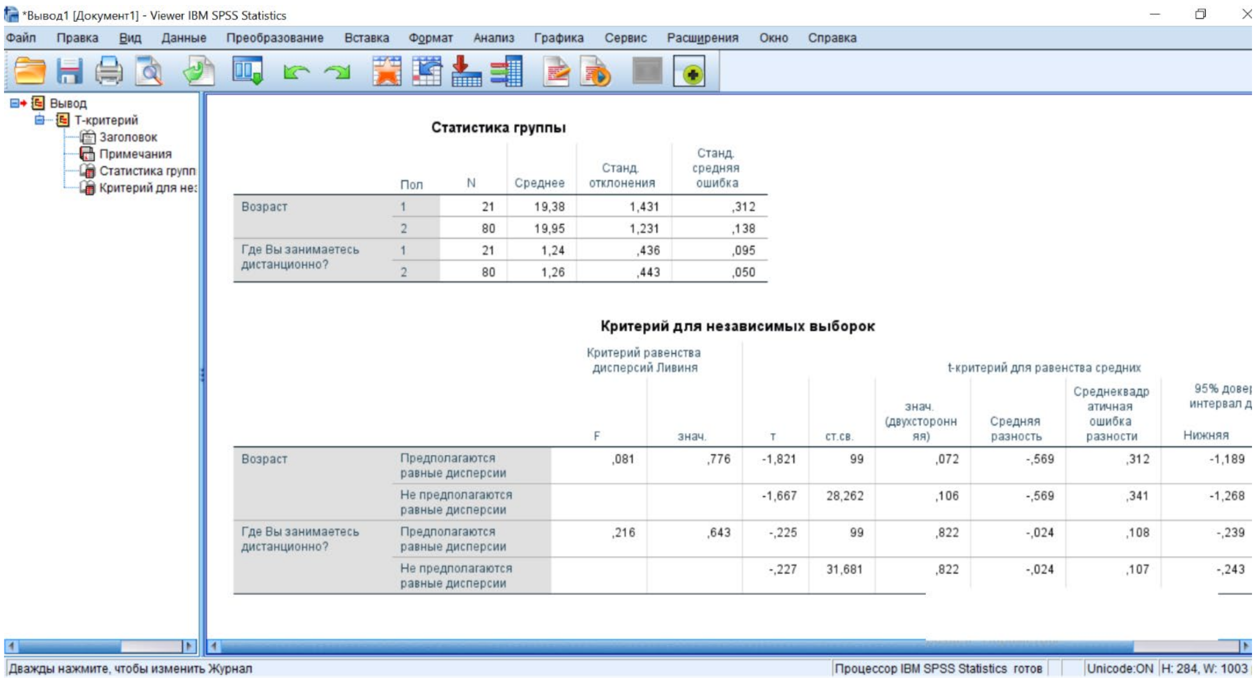


Рис. 58. Пакет IBM SPSS Statistics, инструмент «Анализ», сравнение средних

На рисунке 59 представлены результаты статистического анализа, вкладка называется «Выводы».



Статистика группы

	Пол	N	Среднее	Станд. отклонения	Станд. средняя ошибка
Возраст	1	21	19,38	1,431	,312
	2	80	19,95	1,231	,138
Где Вы занимаетесь дистанционно?	1	21	1,24	,436	,095
	2	80	1,26	,443	,050

Критерий для независимых выборок

		Критерий равенства дисперсий Ливиня		t-критерий для равенства средних					
		F	знач.	т	ст.св.	знач. (двухсторонняя)	Средняя разность	Среднеквадратичная ошибка разности	95% довер интервал д Нижняя
Возраст	Предполагаются равные дисперсии	,081	,776	-1,821	99	,072	-,569	,312	-1,189
	Не предполагаются равные дисперсии			-1,667	28,262	,106	-,569	,341	-1,268
Где Вы занимаетесь дистанционно?	Предполагаются равные дисперсии	,216	,643	-,225	99	,822	-,024	,108	-,239
	Не предполагаются равные дисперсии			-,227	31,681	,822	-,024	,107	-,243

Рис. 59. Результаты статистического анализа пакета IBM SPSS Statistics

На рисунке показано вычисление t-критерия Стьюдента для независимых выборок по признаку «пол». Статистика группы показывает общее количество и средний возраст респондентов, для мужчин это 21 человек и $19{,}38 \pm 1{,}431$, для женщин это 80 человек и $19{,}95 \pm 1{,}231$. В данном пакете t-критерий Стьюдента для независимых выборок обозначается как «Критерий для независимых выборок», с помощь критерия равенства дисперсий Ливиня, который по параметру «возраст» равен 0,081, при $p = 0{,}776$, что показывает равенство дисперсий, следовательно, $t = -1{,}821$, при $p = 0{,}72$, что свидетельствует о том, что выборки по возрасту являются равными.

Одним из достоинств пакета IBM SPSS Statistics, несомненно, является его простота. Для его использования необязательно владеть языком программирования, но владея им, данный пакет дает дополнительные возможности. Еще одним достоинством является удобный и понятный интерфейс. Данный пакет является популярным среди научных исследований в социальных науках, и в частности среди психологов. Достаточно значимым параметром является высокая скорость вычислений, а также, для русскоязычной аудитории, русский язык интерфейса. По этому пакету представлено множество литературы и видеоуроков.

К недостаткам IBM SPSS Statistics можно отнести высокие требования к памяти компьютера и достаточно высокую стоимость пакета.

Нельзя забывать, что любой статистический пакет является всего лишь инструментом в руках исследователя. При недостатке знаний или компетенций любой, даже самый лучший и современный статистический пакет анализа данных, не поможет сделать достоверные выводы.

Математические пакеты статистического анализа данных сегодня являются необходимым инструментом для принятия обоснованных решений, повышающих качественную интерпретацию экспериментальных данных. Одним из наиболее распространённых и давно известных среди психологов является пакет IBM SPSS Statistics, его русскоязычная версия. По данному пакету имеется множество учебников и книг на русском языке. Он позволяет проводить всевозможные виды анализа данных и представлять их графически. Одним из недостатков данного статистического пакета можно отметить его стоимость. Бесплатным вариантом данного пакета, с меньшими возможностями и отсутствием интерфейса на русском языке, является пакет PSPP.

Вопросы и задания для самоконтроля

1. Сравните возможности электронной таблицы MicroSoft Excel и Microsoft Word.
2. Перечислите функции электронных таблиц MS Excel.
3. Сравните профессиональные и непрофессиональные статистические пакеты анализа данных.
4. Дайте определение R-языка.
5. Перечислите платные статистические пакеты.
6. Перечислите бесплатные статистические пакеты.
7. Перечислите возможности применения пакета PSPP.
8. Определите структуру пакета SPSS.
9. Назовите статистические пакеты, которыми вы пользуетесь, сопоставьте их достоинства и недостатки.
10. Перечислите статистические пакеты, с помощью которых можно провести факторный анализ.

ЗАКЛЮЧЕНИЕ

Курсантам и слушателям, обучающимся в образовательных организациях системы МВД России по специальности 37.05.02 «Психология служебной деятельности», крайне важно уметь проводить научные экспериментальные исследования, а также владеть практическими навыками интерпретации полученных результатов.

Данное учебное пособие рассматривает темы, определяющие суть экспериментального психологического исследования и характеризующие основы его интерпретации. В ходе изучения материала обучающийся получает комплексное представление о различных способах обработки экспериментальных данных, методах и интерпретации результатов.

Методы математической статистики помогают психологу в изучении и прогнозировании психологических особенностей и поведения людей. Необходимо понимать, что они являются только необходимым средством и не должны заменять собой цель исследования.

Содержание учебного пособия знакомит с простыми и сложными способами интерпретации данных, заставляет задуматься о методологии и методах познания психических явлений. Раскрытие этих вопросов следует проводить с опорой на серьезные, фундаментальные работы видных отечественных и зарубежных ученых-психологов.

Полезным для освоения методов математической статистики будет их непосредственное применение в научно-исследовательской деятельности.

СПИСОК РЕКОМЕНДУЕМЫХ НОРМАТИВНЫХ ПРАВОВЫХ АКТОВ И ЛИТЕРАТУРЫ

Нормативные правовые акты[*]

1. Федеральный закон от 07.02.2011 № 3-ФЗ «О полиции».
2. Федеральный закон от 30.11.2011 №342-ФЗ «О службе в органах внутренних дел Российской Федерации и внесении изменений в отдельные законодательные акты Российской Федерации».
3. Постановление Правительства РФ от 06.12.2012 № 1259 «Об утверждении правил профессионального психологического отбора на службу в органы внутренних дел Российской Федерации».
4. Приказ МВД России от 02.09.2013 № 660 «Об утверждении Положения об основах организации психологической работы в органах внутренних дел Российской Федерации».
5. Приказ МВД России от 25.12.2020 № 900 «Вопросы организации морально-психологического обеспечения деятельности органов внутренних дел Российской Федерации».

Литература

Основная:

1. Балдин, К. В. Теория вероятностей и математическая статистика: учебник / К. В. Балдин, В. Н. Башлыков, А. В. Рукосуев. — 3-е изд., стер. — Москва: Дашков и К°, 2020. — 472 с.
2. Богаевский, В. А. Математико-статистические методы обработки данных психологических исследований: учебное пособие / В. А. Богаевский, И. А. Паршутин, А. Н. Сударик / ДГСК МВД Рос. — 2-е изд., перераб. и доп. — Москва: [б. и.], 2019. — 72 с.
3. Дорофеев, В. А. Основы регрессионного моделирования для психологов: учебное пособие по дисциплине «Математическая статистика и математические методы в психологии» / В. А. Дорофеев, Ю. А. Мочалова. — Ростов-на-Дону; Таганрог: Изд-во Южн. Фед. ун-та, 2018. — 130 с.
4. Дятлов, А. В. Методы математической статистики в социальных науках (описательная статистика): учебник / А. В. Дятлов, П. Н. Лукичев. — Ростов-на-Дону; Таганрог: Южный федеральный университет, 2018. — 183 с.

Дополнительная:

1. Баринова, М. Г. Психофизиологическая готовность сотрудников правоохранительных органов к деятельности в экстремальных ситуациях: моногра-

[*] Все нормативные правовые акты приводятся по данным официального интернет-портала правовой информации pravo.gov.ru (дата бращения:01.04.2022).

фия / М. Г. Баринова, Е. Г. Зуева. — Санкт-Петербург: Изд-во СПб ун-та МВД России, 2019 — 164 с.

2. Комиссаров, В. В. Практикум по математическим методам в психологии: учебное пособие. — Новосибирск: Изд-во НГТУ, 2012. — 87 с.

3. Лупандин, В. И. Математические методы в психодиагностике: учебное пособие. — Екатеринбург: Изд-во УрГУ, 2012. — 88 с.

4. Остапенко, Р. И. Основы структурного моделирования в психологии и педагогике: учебное пособие. — Москва: Директ-Медиа, 2013. — 123 с.

5. Патронова, Н. Н. Статистические методы в психолого-педагогических исследованиях: учебное пособие / Н. Н. Патронова, М. В. Шабанова. — Архангельск: ИПЦ САФУ, 2013. — 203 с.

6. Туганбаев, А. А. Задачи и упражнения по высшей математике для психологов: учебное пособие. — Москва: Флинта, 2017. — 323 с.

7. Наследов, А. Д. Математические методы психологического исследования. Анализ и интерпретация данных. — Санкт-Петербург: Речь, 2012. — 389 с.

8. Сидоренко, Е. В. Методы математической обработки в психологии. — Санкт-Петербург: Речь, 2010. — 350 с.

9. Унгуряну, Т. Н. Краткие рекомендации по описанию, статистическому анализу и представлению данных в научных публикациях / Т. Н. Унгуряну, А. М. Гржибовский // Экология человека. — 2011. — № 5. — С. 55–60.

10. Цыпин, А. П. Статистические пакеты программ в социально-экономических исследованиях / А. П. Цыпин, А. С. Сорокин // АНИ: экономика и управление. — 2016. — №4 (17). — С. 379–384.

Учебное издание

Баринова Марина Геннадьевна,
кандидат психологических наук

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ПСИХОЛОГИИ

Учебное пособие

Редактор *Свикша Н. О.*
Компьютерная верстка *Фролова А. В.*

ISBN 978-5-91837-563-1

Подписано в печать 28.06.2022. Формат 60×84 $^{1}/_{16}$
Печать цифровая 9 п. л. Тираж 30 экз. Заказ № 85/22

Отпечатано в Санкт-Петербургском университете МВД России
198206, Санкт-Петербург, ул. Летчика Пилютова, д. 1