**2 курс, мат. статистика. Помощь в подготовке**

**ГЛОССАРИЙ**

**Асимметрия (А)** – один из параметров распределения, показывающий вероятность большего отклонения множества измеренных эмпирически значений от среднего математического ожидания. При нормальном распределении А = 0. Если большее число полученных данных имеет значения ниже среднеарифметического, то такая асимметрия является положительной (или левосторонней, т.к. длинный хвост кривой находится слева). При преобладании в обследованной выборке значений выше среднеарифметического асимметрия характеризуется как отрицательная (или правосторонняя).

**Биноминальный критерий (m)** – многофункциональный критерий, позволяющий сравнить эмпирическую частоту появления эффекта с теоретической (заданной, ожидаемой).

**Бисеральный коэффициент корреляции** **(Rбис**) – коэффициент корреляции, используемый для оценки связи между переменными, измеренными в разных шкалах (одна – в дихотомической, другая – в шкале интервалов или отношений).

**Выборки независимые** – выборки, состоящие из разных испытуемых.

**Выборки зависимые** – выборки, образованные из одних и тех же людей, которые неоднократно проходят процедуру измерения. Связанными или сопряженными называются также выборки, в которых каждому участнику одной группы соответствует участник из другой группы (муж – жена, психолог – клиент и др.).

**Выборочная совокупность (выборка)** – часть генеральной совокупности, которая изучается эмпирически. Важнейшее требование к выборке – ее репрезентативность, или адекватность отражения генеральной совокупности по количественным и качественным параметрам. Поэтому заключения, полученные на репрезентативной выборке, возможно переносить (генерализовать) на всю генеральную совокупность.

**Генеральная совокупность** **(популяция)** – это конечное или бесконечное число объектов, выступающих предметом изучения.

**Гипотеза статистическая (Н)** – предположение, проверяемое с помощью критериев статистической оценки различий. Статистические гипотезы представляют собой два основных вида: нулевая гипотеза (Н0) – предположение об отсутствии различий – и альтернативная (Н1) – предположение о значимости различий.

**Дендрограмма** (или дерево расстояний) – метод графического представления результатов кластерного анализа.

**Дискриминантный анализ** – разновидность регрессионного анализа. Объединяет в себе группу статистических процедур, направленных на разделение (дискриминацию) многомерных совокупностей на группы. Позволяет также предсказывать индивидуальные значения зависимой переменной.

**Дисперсионный анализ (F-тест)** –один из распространенных в психологии параметрических методов многомерного анализа, позволяющий оценить изменчивость признака под воздействием одного и более факторов. Дисперсионный анализ позволяет также выявить взаимодействие между факторами, влияющими на измеряемую переменную.

**Дисперсия** –мера изменчивости для шкал интервалов и отношений. Чем больше значение дисперсии, тем больше отклонение индивидуальных значений измеренного признака от среднегруппового.

**Измерение** – процедура установления взаимно однозначного соответствия определенных параметров объектов и символов, основными из которых выступают числа. Образно говоря, это преобразование реальности в числа.

**Кластерный анализ** – метод многомерного анализа, предназначенный для классификации качественных данных. В результате процедур образуются классы (или кластеры), состоящие из наиболее однородных объектов.

**Корреляционный анализ** – раздел математической статистики, объединяющий в себе методы оценки связи между двумя измеренными переменными.

**Коэффициент ассоциации Пирсона (φ)** – коэффициент, используемый для оценки связи между двумя переменными, измеренными в дихотомической шкале. Называется также коэффициентом сопряженности.

**Коэффициент корреляции** – показатель степени связи между двумя переменными. Как правило, числовое значение располагается в диапазоне от –1 до +1: «–1» – прямая линейная отрицательная связь; «+1» – прямая линейная положительная связь. Коэффициент корреляции, приближающийся к нулю, трактуется как отсутствие зависимости между переменными.

**Коэффициент корреляции Кендалла (τ)** – коэффициент, используемый для проведения корреляционного анализа двух переменных, выраженных в порядковой шкале.

**Коэффициент корреляции Пирсона (r)** – коэффициент, применяемый для оценки связи между двумя переменными, измеренными в метрических шкалах (интервалов и отношений).

**Коэффициент корреляции Спирмена (rs)** – коэффициент, используемый для оценки связи между двумя переменными, измеренными в порядковой шкале. Называется также коэффициентом ранговой корреляции.

**Критерий Вилкоксона (Т)** – непараметрический критерий, оценивающий сдвиг значений порядковой переменной в двух условиях на одной выборке.

**Критерий знаков (G)** – непараметрический критерий для зависимых выборок, оценивающий сдвиг значений номинативной или порядковой переменной в двух условиях.

**Критерий Колмогорова-Смиронова (λ)** – непараметрический критерий для порядковых шкал. Применяется для решения различных задач (сопоставление двух и более эмпирических распределений одного признака, сопоставление эмпирического и теоретического распределений и др.).

**Критерий Крускала-Уоллиса (H)**– непараметрический критерий, оценивающий различия в уровне количественно измеренного признака между тремя и более независимыми выборками. Позволяет доказать, что переменная существенно изменяет свои значения в разных группах, но не дает возможности определить направление этих изменений.

**Критерий Макнамары (М)** – непараметрический критерий для зависимых выборок, оценивающий сдвиг значений признака, измеренного в дихотомической шкале.

**Критерий Манна-Уитни (U)**– непараметрический критерий, позволяющий оценить различия между двумя независимыми выборками по уровню количественно измеренного признака. Более удобен в расчетах для малых, чем для средних по объему выборках.

**Критерий Пирсона (*χ²)*** – непараметрический критерий для переменных, измеренных в номинативной шкале. Используется для решения широкого класса исследовательских задач (сопоставление двух и более эмпирических распределений одного признака, сравнение эмпирического и теоретического распределений и др.).

**Критерий Розембаума *(Q)*** – непараметрический критерий, оценивающий уровень различий порядковой переменной в двух независимых выборках.Применим в выборках разного объема.

**Критерий статистический** – 1) решающее правило, обеспечивающее высокую вероятность правильности принятия решения об истинности/ложности проверяемой статистической гипотезы; 2) метод расчета определенного числа; 3) само число, полученное в итоге проведенных расчетов, или эмпирическое значение критерия.

**Критерий тенденций Джонкира (S)** – непараметрический критерий, оценивающий уровень различий и направление изменений количественно измеренной переменной в трех и более независимых выборках.

**Критерий углового преобразования Фишера (φ)** – многофункциональный критерий, применяемый для решения самого широкого класса задач в зависимых и независимых выборках любого объема.

**Критерий Фридмана (χ²r)** –непараметрический критерий, позволяющий оценивать достоверность сдвига значений количественно измеренной переменной при условии наличия не менее трех измерений в одной выборке.

**Медиана (Ме)** –мера центральной тенденции для количественных шкал, представляющая собой значение измеренного признака, делящее упорядоченный ряд значений строго пополам.

**Меры рассеяния (вариации, изменчивости)** – методы дескриптивного анализа, дающие информацию об индивидуальных значениях переменной, измеренной в групповой совокупности.

**Меры среднего уровня (центральной тенденции)** – методы дескриптивного анализа, обеспечивающие усредненную характеристику совокупности объектов по измеренному признаку.

**Метрика** – одинаковая единица расстояния между соседними точками на шкале измерений. В зависимости от отсутствия/наличия метрики измерительные шкалы представляют два вида: неметрические (наименований и порядка) и метрические (интервалов и отношений).

**Многомерное шкалирование** – методмногомерного анализа,по целям аналогичный факторному анализу (сокращение числа исходных переменных). Может использоваться для любых количественно измеренных переменных.

**Мода (Мо)** – мера центральной тенденции, представляющая собой значение измеренной переменной, наиболее часто встречающееся в исследованной совокупности. Является единственной мерой для номинативной шкалы.

**Мощность критерия** – способность критерия различать статистически значимый результат (или не допустить ошибку второго рода о сохранении ложной нулевой гипотезы).

**Наблюдение (наблюдаемое значение)** – обозначение, принятое в психологии для фиксации значений измеренного признака у отдельных испытуемых. Используются также синонимы «индивидуальный показатель/значение», «вариата», «дата».

**Непараметрические критерии** – критерии, используемые в индуктивной статистике для оценки различий данных, выраженных в неметрических шкалах.

**Объем выборки** – число элементов, образующих совокупность. В статистике приняты следующие обозначения объема: N – число элементов в генеральной совокупности – и n – число элементов в выборочной совокупности.

**Параметрические критерии** – критерии, используемые виндуктивной статистике для оценки различий данных, измеренных в метрических шкалах. Эти критерии основываются на расчете параметров нормального распределения.

**Переменная (признак)** – любой изменяющийся параметр реальности, в том числе и психической. Об уровне или показателе переменной говорят только относительно количественно измеренных признаков, значения которых можно расположить в континууме низкий – высокий: низкий уровень тревожности, средний уровень вербальной агрессии, высокий показатель социального интеллекта и др.

**Рангово-бисеральный коэффициент корреляции (Rrb)** – коэффициент корреляции, используемый для оценки связи между переменными, измеренными в разных шкалах (одна – в дихотомической, другая – в порядковой).

**Рандомизация (случайный отбор из популяции)** – наилучшая стратегия для организации репрезентативной выборки. Результаты исследования групп, составленных случайным образом, обладают высокой внутренней и внешней валидностью.

**Распределение** – закономерностьвстречаемости разных значений измеренного признака в выборочной совокупности.

**Регрессионный анализ** –метод определения причинной связи между переменными, позволяющий по изменениям значений независимой переменной/переменных (или предикторов) предсказывать изменения зависимой переменной.

**Репрезентативная выборка** – выборка, адекватно отражающая качественные и количественные параметры генеральной совокупности. Репрезентативность выборки определяется ее объемом и способом отбора испытуемых.

**Статистика индуктивная (теория статистического вывода)** – раздел математической статистики, основывающийся на теории вероятностей. Включает в себя способы обработки данных, позволяющие распространять результаты, полученные эмпирическим путем на ограниченной выборке, на другие группы, не принимавшие участие в исследовании (или на всю генеральную совокупность).

**Статистика описательная (дескриптивный анализ)** – базовый раздел математической статистики. Включает в себя способы обработки данных, позволяющие получить концентрированную информацию об эмпирически исследованной совокупности. Наиболее часто используются проценты, меры среднего уровня и меры рассеяния.

**t-критерий Стьюдента (t-тест)** – один из самых распространенных в психологии параметрических критериев, позволяющий оценивать различия между двумя выборками на основе среднего арифметического.

**Уровень значимости (достоверности)** – вероятность ошибки при принятии статистического решения. В психологии в качестве минимально допустимого принят уровень достоверности не менее 5%, при котором вероятность ошибки не превышает порог случайности. В противном случае решения не являются статистически достоверными. Более ценятся решения, в которых вероятность ошибки ниже 1% или даже 0,1%. В текстовой форме уровень значимости обозначается следующим образом: р ≤ 0,05 (5% вероятности ошибки), р ≤ 0,01 (1% вероятности ошибки), р ≤ 0,001 (0,1% вероятности ошибки).

**Факторный анализ** – методмногомерного анализа,являющийся продолжением корреляционного анализа. Направлен на сокращение числа переменных и выявление скрытых, латентных переменных.

**Шкала** – приписывание чисел объектам. В психологии выделяют четыре типа шкал: наименований, порядка, отношений, интервалов (С. Стивенс). Для каждого типа шкалы существуют конкретные процедуры математико-статистической обработки и представления полученных данных.

**Эксцесс (Е)** – параметр распределения, характеризующий порядок появления определенных значений измеренного признака. При нормальном распределении Е = 0. При положительном эксцессе появляется острый пик значений около математического ожидания, т.е. в выборке преобладают средние значения. При отрицательном эксцессе пик распределения, напротив, становится гладким, т.е. в выборке преимущественно наблюдаются крайние значения измеренной переменной.

**ТЕОРЕТИЧЕСКИЙ МАТЕРИАЛ**

**ВЫБОРОЧНАЯ СОВОКУПНОСТЬ И ТРЕБОВАНИЯ**

**К ЕЕ РЕПРЕЗЕНТАТИВНОСТИ**

1. Генеральная и выборочная совокупности.

2. Репрезентативность выборки.

3. Стратегии контроля состава групп.

**1.** **Генеральная и выборочная совокупности**

Проверка исследовательских гипотез осуществляется в специально организованных эмпирических исследованиях. Одним из первых и главных шагов при их планировании является ответ на вопросы: «Кто будет принимать участие в исследовании?» и «Сколько мне надо обследовать людей?».

***Генеральная совокупность*** *–* это весь объем объектов (конечный или бесконечный), относительно которых сформулирована гипотеза. В зарубежной психологии для обозначения этого множества предпочитают использовать термин «популяция».

Генеральная совокупность, как правило, некоторая ограниченная величина (обозначаемая как N). Но и при своей конечности она все равно включает в себя очень большое число объектов (родители учащихся, пользователи мобильных телефонов, подчиненные, дети с нарушениями поведения и др.). Естественно, что исследователь просто физически не в состоянии изучить всех представителей генеральной совокупности. Поэтому из популяции специально выбирается некоторая группа, которая называется ***выборочной совокупностью или выборкой***. Именно в выборке и проводится сбор эмпирических данных. Объем выборки обозначается как n.

Выборочные совокупности классифицируются по разным основаниям.

***По объему выборки*** представляют три вида:

– малые (n ≤ 30);

– средние (30 ≤ n ≤ 200);

– большие (n > 200).

***По степени связи*** образующих их элементов выборки разделяются на независимые и зависимые.

*Независимые выборки* – выборки, построенные на отсутствии при отборе испытуемых в группы какой-либо возможной связи между ними. К примеру, люди разного пола, возраста, типа профессии и др. Такие выборки еще называются несвязанными.

*Зависимые выборки* (связанные, сопряженные) – выборки, в которых отбор каждого испытуемого в одну группу сопряжен с отбором по определенным критериям конкретного испытуемого в другую группу. Максимальная сила зависимости выборок присутствует при изучении одной группы «до» и «после», т.е. когда в исследовании принимают участие одни и те же испытуемые. Например взрослые «до» и «после» тренинга ассертивности или дети «до» и «после» определенной программы обучения и т.п. Меньшую силу зависимости иллюстрируют, к примеру, исследования детско-родительских отношений (в одну выборку входит ребенок, в другую родитель) или супружеских взаимоотношений (одна выборка образована женами, другая – мужьями и т.п.).

***По структуре представленности*** ***психологического признака*** выделяются гомогенные и гетерогенные выборки.

*Гомогенная (однородная) выборка –* выборка, составленная на основе одной степени выраженности интересующего исследователя параметра генеральной совокупности.

*Гетерогенная (разнородная выборка)* – выборка, в которой интересующий психолога признак имеет разные градации (может быть качественно или количественно разным).

Например, выборка, образованная только людьми с сильной нервной системой, является гомогенной по отношению к силе нервной системы испытуемых; выборка, в которую входят люди со слабым и сильным типами нервной системы, уже гетерогенна. Выборка белорусов однородна по сравнению с выборкой, представленной людьми разных национальностей (если интересующий нас признак – это национальная принадлежность). Чем выборка однороднее, тем меньше может быть ее объем (это одновременно предполагает большую конкретизацию гипотезы, а следовательно, и сужение области распространения полученных выводов).

Таким образом, выборочную совокупность можно рассматривать как уменьшенную модель генеральной совокупности. Только если эта модель будет правильной (соответствующей тем параметрам генеральной совокупности, которые являются предметом исследования), можно получить на основании ее изучения правильные выводы и о генеральной совокупности в целом.

**2. Репрезентативность выборки**

Для практических целей иногда очень важно знать параметры конкретной группы. Например, уровень агрессивности у учащихся отдельного класса, частотой драк в котором обеспокоены многие учителя. Или степень эмоционального выгорания медицинских работников отдельной больницы, пациенты которой регулярно пишут в различные инстанции жалобы на плохое обслуживание и т.п. Но для научного исследования такого знания параметров конкретной группы недостаточно, ведь оно стремится к тому, чтобы на основании результатов изучения отдельной выборки делать заключения обо всей генеральной совокупности. Поэтому важнейшим условием правильности, достоверности сделанных выводов является адекватность составления выборки.

***Репрезентативность*** – свойство выборки точно отражать интересующие исследователя параметры генеральной совокупности. Иначе говоря, репрезентативность – это представительность.

Выборка является ***нерепрезентативной*** по двум основным причинам.

*Недостаточный объем*.

На правильный по сути вопрос, который студентами обычно формулируется очень прагматично: «Сколько людей “брать” для исследования?» можно ответить только учитывая следующие принципиальные аспекты проводимого исследования.

1) Тип исследования (качественное или количественное). В качественных исследованиях, представляющих собой анализ единичных случаев, статистика вообще не обязательна. Относительно количественных исследований существует правило: статистически приемлемой для обобщений группой считать выборку 30 и более респондентов (т.е. средние и большие по объему группы).

2) Масштаб исследовательской проблемы, поскольку именно для ее решения и создается выборка (проблема отдельного учреждения, города, страны и др.).

3) Цель и задачи исследования. Минимальный объем выборки допустим для установления взаимосвязи между психологическими переменны-ми: 30–35 человек. Если требуется сравнение двух групп, то каждая из них должна включать не менее 50 человек. Самый большой объем выборки предписан нормами разработки диагностических, стандартизированных методик: не менее 200 человек, а чаще всего порядка нескольких тысяч.

4) Способы дальнейшей математико-статистической обработки полученных эмпирических данных. Для большинства критериев численный состав выборки ограничен минимальным и максимальным значениями. Есть критерии, которые «работают» и на очень малом количестве наблюдений (от 2–3 до 5–6), есть и такие, которые применимы для очень большого числа.

*Несовпадение свойств выборки и генеральной совокупности.*

Здесь проблема заключается как в квалификации исследователя (все ли параметры, которые входят в проверяемую гипотезу, он смог участь), так и в ограниченности проведенного теоретического анализа или специфике используемого теоретического подхода. В этом плане показательна критика результатов экспериментов, полученных в американском бихевиоризме, который часто называют «наукой о белых крысах и студентах-психологах». Действительно, справедливо возникает вопрос, насколько результаты, полученные на молодых людях, заинтересованных в психологии, правомерно переносить на людей другого возраста, другого уровня образования, имеющих иную систему жизненных ценностей и т.п.

При анализе адекватности представленности в выборке качественных и количественных параметров генеральной совокупности возникает еще одна парадоксальная с точки зрения логики проблема. Чтобы выборка точно отражала генеральную совокупность, исследователь должен знать те самые параметры генеральной совокупности. Но если он их уже знает, зачем тогда вообще проводить исследование в отдельной группе? На самом деле, параметры генеральной совокупности известны далеко не все и не всегда. И обозначенный парадокс решается как раз посредством применения математико-статистического аппарата, который позволяет проверить принадлежность выборки генеральной совокупности и оценить возможность перенесения полученных эмпирических данных на всю генеральную совокупность.

**3.** **Стратегии контроля состава групп**

В психологических исследованиях существуют определенные ограничения в подборе состава выборки. Ограничения эти связаны с тем, что исследователь по каким-либо причинам не имеет возможности выбора из всей генеральной совокупности, а имеет доступ к каким-то реально существующим выборкам.

***Случайный отбор испытуемых из генеральной совокупности*** – наиболее оптимальная стратегия организации выборочной совокупности. Эта стратегия называется также рандомизацией и является строгим критерием экспериментирования. Этой стратегии полезно придерживаться и в других типах эмпирических исследований.

Считается, что составленная посредством случайного отбора испытуемых выборка хорошо репрезентирует популяцию, поскольку каждый ее член имел одинаковую вероятность стать участником исследования. Результаты исследований, полученные в случайных группах, обладают не только высокой внешней валидностью (испытуемые правильно отражают состав популяции, поэтому полученные данные можно смело распространять на всю генеральную совокупность), но и внутренней валидностью (поскольку индивидуальные различия людей предположительно случайно распределятся в эмпирических выборках, то группы будут эквивалентными).

Вариантом стратегии случайного отбора выступает отбор из популяции по заданному критерию. Этот критерий задает определенные границы и тем самым делает отбор более целенаправленным (например, пол, возраст, уровень образования, принадлежность к субкультуре и др.).

Основным методом создания случайной выборки является таблица случайных чисел, сконструированная таким образом, что вероятность появления каких-либо соседних чисел не подчиняется никаким статистически закономерностям. Например, генеральная совокупность клиентов некоторой фирмы составляет 3 тысячи человек, а нам нужна выборка в 40 респондентов. В соответствии с имеющейся клиентской базой каждому из списка присваивается свой порядковый 4-значный номер (поскольку именно четыре разряда имеет число генеральной совокупности): 0001, 0002, 0003 и т.п. Затем следует обратиться к таблице случайных чисел, на которой выбирается любая точка, от которой начинают считываться 40 чисел (либо по горизонтали, либо по вертикали). 40 клиентов, записанных под этими случайными числами, и составляют искомую вороятностную выборку. Сегодня сгенерировать последовательность случайных чисел возможно и с помощью специальных компьютерных программ.

Как следует из описания, обращение к таблице случайных чисел довольно трудоемкая процедура. Случайную последовательность можно получить более простым образом: подбрасывая монетку или кубик.

***Стратегия попарного уравнивания*** целесообразна в тех случаях, когда испытуемые могут быть подвергнуты предварительному тестированию, измерению и если требуются не очень большие выборки. Реализуется эта стратегия пошагово:

– первоначально измеряется уровень интересующей психолога переменной в выборке потенциальных испытуемых (например, уровень тревожности среди беременных женщин);

– затем сырые оценки упорядочиваются или ранжируются (либо по степени увеличения, либо по степени уменьшения тревоги);

– из испытуемых с наиболее близкими оценками составляются пары;

– испытуемых из каждой пары случайным образом (например, орел/решка) распределяют по двум группам.

Таким образом, в нашем примере две выборки беременных женщин оказываются эквивалентными по значимому для исследователя признаку – уровню тревожности.

***Стратегия случайного распределения слоев*** представляет собой смешанную по отношению к двум описанным выше технологию составления выборки. Слоем называется группа, обладающая одинаковым уровнем выраженности измеренной переменной (в примере с беременными женщинами можно выделить три слоя: группа с низким уровнем тревожности, группа со средним уровнем данной переменной и группа с высоким уровнем). Затем испытуемых из каждого слоя случайным образом распределяют по выборкам для проведения дальнейшего исследования.

Данная стратегия имеет преимущество перед рандомизацией при составлении гетерогенной выборки (например, если исследователь намерен распространить полученные результаты на всех взрослых, то в зависимости от частной гипотезы требуется выделение слоев по полу, по возрасту, по типу профессии и др.).

***Случайный отбор групп*** применяется в так называемых полевых исследованиях, т.е. при проведении сбора данных в реальных условиях жизнедеятельности. Например, как бы ни была сама по себе значима проверка новой системы мотивации сотрудников, она, тем не менее, не позволяет расформировывать уже действующие трудовые коллективы и создавать новые на основании тех критериев, которые значимы для исследователя. В подобных случаях из уже существующих групп (школьный класс, воинское подразделение, спортивная команда и др.) случайным образом отбирают выборки для проведения исследования.

Таким образом, определение стратегии составления выборки обусловлено многими факторами: исследовательской гипотезой; параметрами генеральной совокупности, которые обязательно надо проконтролировать; объемом генеральной совокупности; трудоемкостью используемых исследователем процедур получения данных; реально существующими возможностями проведения исследования и др. Однако, несмотря на все ограничения, исследователь должен стремиться максимально приблизить свою выборку к случайной, чтобы обеспечить ее репрезентативность.

**Проблема измерения в психологии**

1. Сущность процедуры измерения.

2. Типы измерительных шкал.

3. Особенности обработки данных, представленных в разных измерительных шкалах.

**1. Сущность процедуры измерения**

С момента появления научной психологии поиск объективных методов изучения психической реальности и ее объяснения выступал одной из самых значимых проблем, которые сохраняют свою актуальность и сегодня. Причиной этого является своеобразие онтологии психических явлений, прежде всего их субъективная данность и внешняя представленность только в опосредованной, превращенной форме. Сегодня большинство представителей мирового психологического сообщества разделяют точку зрения на реальность существования психических явлений и возможность их изучения объективными методами.

Такими объективными общенаучными методами, посредством которых ученые получают эмпирические данные, выступают три: наблюдение, измерение и эксперимент. Как и другие научные методы, измерение представляет собой совокупность правил, принятых ученым сообществом в качестве нормы исследования, направленного на получение объективного и достоверного знания о реальности.

Измерение основывается на допущении, что возможно точное отображение эмпирической структуры в символическую. Иначе говоря, измерение – это операция установления взаимно однозначного соответствия неких объектов и неких символов (в основном чисел). Практически измерение осуществляется как фиксация состояния объекта (предмета изучения) на основе фиксации состояния другого объекта (прибора). Для более полного понимания сущности измерения необходимо обратить внимание на следующее:

– наличие взаимодействия объекта изучения и прибора (например, женщина стала на напольные весы);

– обусловленность изменений показателей прибора изменениями состояния объекта («Ого, 72 килограмма, плюс два за неделю»);

– интерпретация показателей прибора как характеристик реального объекта («Я потолстела!!!»).

С измерительными приборами каждый из нас достаточно хорошо знаком по собственному жизненному опыту (сантиметр, транспортир, термометр и др.). Вопрос состоит в том, что выступает прибором в психологических исследованиях? Такие психологические «приборы» могут быть разными: во-первых, это сам субъект (или эксперт); во-вторых, специальные измерительные инструменты (психометрические и психодиагностические методики различного типа). В зависимости от позиции, в которой осуществляется измерение, различается и его содержание.

Таблица 1 – Содержание измерения в зависимости от позиции

|  |  |  |  |
| --- | --- | --- | --- |
| Кто проводит  измерение | Что измеряется | Цель измерения | Итог  измерения |
| Испытуемый  (позиция субъекта) | Стимулы | Оценка испытуемым своего субъективного опыта или некоторого объекта | Отнесение испытуемого к определенному классу, определение расположения измеренного стимула в пространстве выраженности признаков |

Продолжение таблицы 1

|  |  |  |  |
| --- | --- | --- | --- |
| Исследователь  (позиция внешнего наблюдателя) | Психологические переменные испытуемых (например, поведение, мышление, коммуникативные навыки, личностные качества и др.) | Оценка испытуемого (поведения, мышления, коммуникативных навыков, личностных качеств и др.) | Определение сходства или различия между людьми, установление индивидуальных особенностей |

Измерение в психологии одновременно выступает и как самостоятельный метод и как компонент экспериментального метода. Измерение как самостоятельный метод служит для выявления индивидуальных различий субъектов и широко используется как в исследовательских, так и в практических целях. Основной инструментальной формой измерения как самостоятельного метода являются тесты, которые по своей сути представляют собой сжатую во времени и упрощенную процедуру измерения. Измерение как компонент эксперимента – это уже методика регистрации состояния зависимой переменной и изменений этого состояния в ответ на экспериментальное воздействие.

**2. Типы измерительных шкал**

Измерение предполагает наличие неких единиц, в которых осуществляется данная процедура. В естественных науках существуют эталоны, или точные образцы, установленной единицы измерения (килограмм, метр и др.). Параметры психической реальности таких стандартных единиц измерения не имеют. Поэтому для установления значения психологического признака (переменной) требуются особые измерительные шкалы.

В психологии принята классификация шкал, предложенная в 40-е гг. ХХ в. американским психофизиологом С. Стивенсоном. Шкалы отличаются друг от друга правилами приписывания чисел измеряемым психологическим параметрам и единицами измерения (в терминологии С. Стивенсона «измерительными палочками»), которые задают допустимые преобразования для данной шкалы. Названный ученый ввел различение четырех типов шкал: наименований, порядка, интервалов и отношений.

***Шкала наименований (номинальная или номинативная)*** – это качественное разделение объектов на непересекающиеся множества. Процедура основана на сравнительном анализе и представляет собой классификацию, группировку объектов. Главные требования: 1) один класс должны образовать только эквивалентные друг другу объекты; 2) один и тот же объект не может входить в разные классы. После разделения на классы каждому из них приписывается свой условный символ – либо буквенный, либо числовой. Нельзя приписывать одни и те же числа разным классам, и нельзя один и тот же класс обозначать разными цифрами.

А (или 1) Б (или 2) В (или 3) и т.д.

Наиболее простым вариантом шкалы наименований является дихотомическая шкала, в которой проявления изучаемой переменной разделяются на два непересекающихся класса. Признак, измеренный в данной шкале, называется альтернативным. Например: по критерию места проживания людей можно разделить на две группы – городские и сельские жители; по наличию определенных умений – умеющие танцевать танго и не умеющие танцевать этот танец; по критерию пола – женщины и мужчины и др. Выделенные классы можно обозначить цифрами 0 и 1, 7 и 4 или любыми другими символами. Иными словами, для номинативной шкалы не существенно, в каком порядке представлены измеряемые признаки.

Легкость использования данной шкалы только кажущаяся, так как обращение к ней сопровождается решением определенных содержательных проблем. Одной из основных таких проблем выступает определение критерия, в соответствии с которым устанавливаются отношения равномерности (идентичности) внутри класса, которые позволяют отнести каждый элемент только в один класс. В психологии встречаются описания результатов, построенные таким образом, что каждый класс определяется по новому критерию (например, первый – по особенностям мотивации испытуемых, второй – по особенностям их мышления и т.д.). В таком случае присутствует грубая ошибка – элементы перемещаются из класса в класс с каждой новой сменой критерия. Второй серьезной проблемой является относительность многих психологических терминов и культурная обусловленность их содержания. Относительное понятие определяется не путем указания образующих его признаков, а путем противопоставления другим терминам (интеллигентный, эгоистичный, мужественный и др.), что сильно затрудняет определение однозначного критерия для деления на классы.

***Шкала порядка (порядковая,******или ранговая)*** приписывает числа объектам в зависимости от выраженности измеряемого свойства. Она устанавливает порядок следования (меньше – больше, слабее – сильнее, ре-же – чаще и др.) и представляет собой линейную упорядоченность объектов на некоторой оси признака. На этой шкале есть деления, но не установлена метрика между ними, а значит и расстояние между разными соседними делениями может быть разным.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

меньше 1 2 3 4 5 6 7 больше

слабее сильнее

реже чаще

Иначе говоря, в этой шкале мы можем установить только последовательность или ранг выраженности признака. В порядковой шкале должно быть не менее трех классов (например, выше среднего – средний – ниже среднего; всегда – часто – редко и др.). При числовом кодировании данных главным требованием является соблюдение последовательности или порядка, т.е. каждая последующая цифра должна быть больше/меньше предыдущей. Само направление ранжирования является произвольным, т.е. исследователь может самостоятельно выстраивать шкалу либо по степени возрастания, либо по степени убывания измеренного признака.

Порядковая шкала считается самой распространенной в психологии. Она является непараметрической, поэтому для статистической обработки данных применяются непараметрические критерии.

***Шкала интервалов (интервальная)*** определяет величину различий между объектами в проявлении свойства. Размер интервала представляет собой одинаковую единицу измерения на всем протяжении шкалы и позволяет измерить долю или часть психологического признака.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1 2 3 4 5 6 7 8 9 10

В отличие от порядковой шкалы здесь присутствует равномерность делений, что позволяет ответить на вопрос: «На сколько больше/меньше выражен измеренный признак?». Отличительной особенностью данной шкалы является отсутствие естественной точки отсчета (или нуля). Поскольку эта шкала метрическая, т.е. установлена одинаковая единица расстояния между соседними точками, то для нее используется параметрическая статистика. Измерение в интервальной шкале осуществляется посредством специальных единиц (стены и стенайлы).

Вопрос о том, насколько эта шкала представлена в психологии, остается дискуссионным на протяжении нескольких десятилетий. Сторонники тестирования интеллекта полагают, что измерение IQ осуществляется посредством интервальной шкалы. Другие специалисты с ними не соглашаются, в частности, указывая на то, что под вопросом остается само содержание теоретического конструкта «интеллект». Тем не менее в качестве примеров использования в психологии интервальной шкалы можно ссылаться на тесты интеллекта или на 16-факторный личностный опросник Р. Кеттелла, в котором сырые баллы при обработке результатов переводятся в стены.

***Шкала отношений*** –отличается от шкалы интервалов тем, что у нее появляется нулевая отметка. Она позволяет ответить на вопрос: во сколько раз тот или иной субъект имеет более высокие или более низкие значения, чем другой.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

0 1 2 3 4 5 6 7 8 9

В психологии шкала отношений используется крайне редко, в основном для измерения времени (выполнения заданий, решения задач, реакции и др.), а также в психофизиологии для измерения абсолютных порогов чувствительности. Это объясняется тем, что психологические переменные в своем большинстве не имеют естественного нуля (например, как оценить «0 памяти», «0 темперамента» и др.?!).

Таким образом, используемые в психологии измерительные шкалы отличаются друг от друга правилами приписывания чисел, а значит, и предполагают разные способы математико-статистической обработки данных, полученных с их помощью.

**3. Особенности обработки данных, представленных в разных измерительных шкалах**

Существующие измерительные шкалы имеют разную мощность, или дифференцирующую способность, т.е. дают разный объем информации об измеренном параметре психической реальности. По возрастанию мощности шкалы расположены следующим образом: наименований, порядка, интервалов, отношений. Более мощная шкала открывает и большие возможности статистической обработки. В обобщенном виде возможности обработки эмпирических данных для названных шкал измерений представлены в таблице 2.

Как следует из таблицы 2, номинативная шкала как самая слабая по мощности, имеет и самый ограниченный арсенал средств обработки данных. Преобразования шкал могут быть двух видов: понижение мощности шкалы и повышение мощности шкалы. Понижение мощности шкалы означает, что все способы обработки и графического представления результатов, используемые для более слабых шкал, подходят и для более сильной (именно поэтому для шкалы отношений подходит абсолютно любая статистика). Понижение мощности шкалы означает обязательную потерю части эмпирической информации. Другой вид преобразований – повышение мощности шкалы – очень сложная процедура, которую необходимо специально обосновывать.

Таблица 2 – Обработка данных, представленных в различных шкалах

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Шкала | | Результат | Статистика | Графическое представление |
| Неметрическая | Наименований | Типологизация,  классификация | Непараметрическая | Диаграмма |
| Порядка | Установление порядка выраженности свойства, ранжирование | Непараметрическая | График |
| Метрическая | Интервалов | Сравнение: на сколько больше/меньше выражено? | Параметрическая | Гистограмма  (полигон распределений) |
| Отношений | Сравнение: во сколько раз больше/меньше? | Любые статистики | Любая форма |

**ОПИСАТЕЛЬНАЯ СТАТИСТИКА**

1. Описательная статистика и область ее применения.

2. Меры среднего уровня.

3. Меры вариации.

4. Распределение и его параметры.

1. **Описательная статистика и область ее применения**

Описательная статистика выступает в качестве базового раздела математической статистики, который называется также дескриптивной статистикой или дескриптивным анализом. Описательная статистика представляет собой способы обработки данных, позволяющие получить компактную информацию о конкретной, эмпирически исследованной выборке. Иначе говоря, в результате дескриптивного анализа появляются такие числа, которые характеризуют некоторую интересующую исследователя ситуацию (среднее количество времени, затрачиваемое на разговоры по телефону; число несчастных случаев в различные времена года; результаты прохождения централизованного тестирования; уровень тревожности сотрудников организации после смены руководства и др.). Назначение описательной статистики – сжать множество значений измеренной психологической переменной в группе респондентов до одного числа. Такое «сжатие», концентрация дает возможность для интерпретации данных и для их дальнейшей статистической обработки.

К базовым методам дескриптивной статистики относятся:

– процентные показатели;

– меры среднего уровня;

– меры рассеяния (или вариации);

– парные коэффициенты связи.

В психологических исследованиях наиболее популярны подсчет частоты встречаемости измеренного признака в процентах, а также меры среднего уровня и вариации, которые различаются в зависимости от типа измерительной шкалы.

**2. Меры среднего уровня**

Это меры, обеспечивающие усредненную характеристику совокупности объектов по измеренному признаку. Для разных шкал измерений используются разные меры среднего уровня или меры центральной тенденции: мода (для номинативной шкалы), медиана, среднее арифметическое (для порядковой, интервальной и шкал отношений).

Меры центральной тенденции отражают уровень выраженности измеренного признака.

***Мода (Мо)*** *–* значение, которое встречается во множестве наблюдений наиболее часто (в буквальном смысле мода – это популярность, типичность).

Например, 10 респондентов-женщин в качестве наиболее предпочитаемых домашних питомцев назвали следующих животных: хомяк, рыбки, кошка, собака, хомяк, черепаха, кошка, хомяк, собака, морская свинка. В данном примере наиболее частым ответом будет «хомяк» (это и есть мода или модальное значение). В другой выборке из 10 респондентов-мужчин был получен несколько иной ряд животных: собака, кошка, рыбки, попугай, собака, собака, кошка, черепаха, кошка, хомяк. В этом случае одинаково часто среди опрошенных встречаются два выбора – собака и кошка. Это означает, что здесь наблюдается две моды, и такая совокупность называется бимодальной. Если фиксируется три и более мод, то говорится о полимодальной совокупности, и такие выборки еще можно характеризовать по наименьшей и наибольшей моде. Если ни одна мода не прослеживается, тогда считается, что данное распределение моды не имеет.

Мода может рассчитываться и для более мощных, чем номинативная, шкал измерений. К примеру, время решения задач-анаграмм в группе из 9 учащихся составило в минутах: 1, 1, 3, 4, 1, 2, 1, 2, 7 (моду составляет 1 минута как наиболее часто встречающееся число).

***Медиана (Ме)*** *–* значение, делящее упорядоченное (по возрастанию или убыванию) множество значений измеренного признака пополам. Одна половина значений оказывается меньше медианы, а другая – больше.

Исследованная выборка может иметь нечетное или четное количество измерений. Для нечетной медиана выступает центральным значением. Например, для времени решения задач-анаграмм упорядоченный по возрастанию ряд наблюдений (9 измерений) выглядит следующим образом: 1, 1, 1, 1, 2, 2, 3, 4, 7. В этой выборке Ме = 2. Для выборки с четным количеством измерений медиана рассчитывается как среднее между двумя центральными значениями. Допустим, время решения задач для 10 испытуемых выглядит следующим образом: 1, 1, 1, 2, 3, 3, 4, 7. В этом примере центральными значениями в упорядоченном ряду выступают 2 и 3. Соответственно медиана Ме = (2 + 3)/2 = 2,5.

***Среднее арифметическое (Х)*** *–* это суммавсех значений измеренного признака, деленная на количество суммированных значений. Среднее арифметическое обозначается также и другими символами: М – для выборочной совокупности, µ – для генеральной совокупности.

|  |  |
| --- | --- |
| M = | Σxi |
| n |

Сумма обозначается буквой Σ, отдельные значения признака как х1, х2 …хi, n – общее количество измеренных значений. Например, в случае времени решения задач-анаграмм М = 2,44 (Σ = 1 + 1 + 3 + 4 + 1 + 2 + 1 + 2 + + 7 = 22, n = 9). Расчеты показывают, что в данном случае среднее арифметическое близко по значению с медианой (Ме = 2). Это говорит о том, что распределение измеренных признаков в данной выборке приближается к симметричному. Однако оно не совпадает с модой (Мо = 1). Такое несовпадение объясняется тем, что среднее арифметическое очень чувствительно к экстремально большим или малым значениям переменной (их еще называют «выбросами», обусловленными индивидуальными вариациями испытуемых). В обсуждаемом случае таким «выбросом» выступает очень длительное время у одного испытуемого, равное 7 минутам. Поэтому при необходимости сравнения нескольких выборок по среднему арифметическому измеренного признака важно, чтобы в этих группах подобные «выбросы» отсутствовали. Если они присутствуют, то в качестве мер центральных тенденций лучше ограничиться модой или медианой, которые не чувствительны к индивидуальным вариациям.

Таким образом, интерпретация мер центральной тенденции заключается в следующем: мода – наиболее частое значение, медиана – срединное в упорядоченном ряду данных, среднее арифметическое – наиболее ожидаемое значение.

**3. Меры вариации**

Данные меры называются также мерами изменчивости или мерами рассеяния. Они дают исследователю информацию об индивидуальных различиях измеренного признака. Для количественно измеренных переменных нижняя граница мер изменчивости равна 0 (т.е. объекты не отличаются друг от друга по определенному признаку). Верхняя граница – это всегда открытая величина, поскольку она определяется особенностями изучаемой переменной (в качестве иллюстрации открытости верхней границы уместно вспомнить исследования объема памяти, описанные А.Р. Лурией в «Маленькой книжке о большой памяти» и не обнаружившие этот предел у испытуемого Ш.).

К наиболее часто используемым в психологии мерам вариации относятся: размах, дисперсия, стандартное отклонение.

***Размах (R)*** *–* самая простая мера вариации, указывающая на диапазон изменчивости. Это разница между максимальным и минимальным значением измеренного признака. Например, если в студенческой группе минимальный балл, полученный на экзамене, составил 4, а максимальный – 9, то R = 5.

***Дисперсия (S²)*** *–* мера изменчивости для данных, полученных в метрических шкалах (интервалов и отношений). Это характеристика отклонения от среднего. Она вычисляется по формуле:

|  |  |  |
| --- | --- | --- |
| S2 = | Σ(xi – М)2 | , |
| n – 1 |

где хi – каждое наблюдаемое значение измеренного признака;

М – среднее арифметическое значение признака;

n – количество наблюдений.

Дисперсия рассчитывается по следующему алгоритму:

– находится среднее арифметическое по выборке;

– для каждого элемента выборки вычисляется его отклонение от среднего арифметического;

– каждый элемент полученного множества значений возводится в квадрат;

– высчитывается сумма этих квадратов;

– сумма делится на число наблюдений, уменьшенное на 1 (n – 1).

Например, две группы взрослых студентов-заочников (n = 10 в каждой) выполняли тесты достижений в разное учебное время. Первая группа работала днем во вторник, вторая – в пятницу вечером. При оценке теста анализировалось количество допущенных ошибок. Рассчитаем дисперсию для этих двух групп.

Таблица 3 – Расчет значений для вычисления дисперсии

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| № исп. | 1 группа | | | 2 группа | | |
| хi | хi – M | (xi – M)² | xi | xi – M | (xi – M)² |
| 1 | 4 | 4 – 2,8 = 1,2 | 1,44 | 3 | 3 – 4,5 = – 1,5 | 2,25 |
| 2 | 4 | 4 – 2,8 = 1,2 | 1,44 | 7 | 7 – 4,5 = 2,5 | 6,25 |
| 3 | 3 | 3 – 2,8 = 0,2 | 0,04 | 10 | 10 – 4, 5 = 5,5 | 30,25 |
| 4 | 4 | 4 – 2,8 = 1,2 | 1,44 | 10 | 10 – 4,5 = 5,5 | 30,25 |
| 5 | 4 | 4 – 2,8 = 1,2 | 1,44 | 5 | 5 – 4,5 = 0,5 | 0,25 |
| 6 | 4 | 4 – 2,8 = 1,2 | 1,44 | 4 | 4 – 4,5 = – 0,5 | 0,25 |
| 7 | 0 | 0 – 2,8 = – 2,8 | 7,84 | 2 | 2 – 4,5 = – 2,5 | 6,25 |
| 8 | 3 | 3 – 2,8 = 0,2 | 0,04 | 1 | 1 – 4,5 = – 3,5 | 12,25 |
| 9 | 1 | 1 – 2,8 = – 1,8 | 3,24 | 1 | 1 – 4,5 = – 3,5 | 12,25 |
| 10 | 1 | 1 – 2,8 = – 1,8 | 3,24 | 2 | 2 – 4,5 = – 2,5 | 6,25 |
|  | M = 2,8 | Σ+ = 6,4  Σ– = 6,4 | Σ = 21,6 | М = 4,5 | Σ+ = 14  Σ– = 14 | Σ = 106,5 |

Следует обратить внимание, что при определении «xi – M» суммы вычисленных отклонений от средней со знаком «+» и со знаком «–» должны быть равны. Их совпадение – это показатель правильности сделанных математических расчетов.

Подставим значения в формулу дисперсии. Дисперсия для первой группы: S² = 21,6/9 = 2,4; дисперсия для второй группы: S² = 106,5/9 = 11,8. Очевидно, что у второй группы дисперсия выше, а значит, показатели отдельных испытуемых значительно больше отклоняются от среднего группового показателя числа допущенных ошибок, чем у участников первой группы.

Вообще дисперсия может принимать значения от 0 до бесконечности. Если S² = 0, то это означает отсутствие изменчивости или постоянство значений измеренного признака в данной группе. Однако для интерпретации психологических данных дисперсия оказывается не всегда удобной величиной, поскольку ее размерность может не совпадать с размерностью измеренного признака (даже в обсуждаемом примере бросается в глаза, что дисперсия во второй группе студентов значительно выше максимального количества ошибок у отдельного испытуемого: S² = 11,8, а xmax = 10). Для того чтобы приблизить размерность данных статистической обработки к тем единицам, в которых измерена переменная, из дисперсии извлекают квадратный корень. Полученную величину называют стандартным отклонением.

***Стандартное (или среднеквадратичное) отклонение***– положительное значение квадратного корня из дисперсии. Обозначается эта величина буквами S – стандартное отклонение в исследованной выборке – или σ (сигма) – стандартное отклонение в генеральной совокупности. Последнее обозначение считается допустимым использовать и для обозначения данного расчетного показателя в конкретной выборке.

Найдем стандартное отклонение для первой и второй групп по формуле: σ = √S². Для первой группы: σ = √2,4 = 1,54; для второй группы: σ = √11,8 = 3,43. Из примера видно, что переходе от дисперсии к стандартному отклонению их показатели сохраняются, т.е. по-прежнему индивидуальный разброс числа допущенных в тесте ошибок выше во второй группе студентов, чем в первой. Однако теперь изменчивость выражается в таких конкретных числах, которые оказываются более сопоставимы с исходной измерительной шкалой, а потому более удобны исследователю для последующего их объяснения.

В качественной интерпретации числовых значений дисперсии и стандартного отклонения следует руководствоваться правилом: чем выше значения S² или σ, тем больше разбросаны значения переменной относительно среднего, и наоборот.

**4. Распределение и его параметры**

***Распределение признака*** *–* это закономерность встречаемости его значений в исследованной выборке. Появление в результате измерения значений (как некоторой случайной, конкретной величины) подчиняется различным законам.

Предельным законом, описывающим распределение, является закон нормального распределения. Нормальным это распределение названо потому, что оно наиболее часто обнаруживалось в эмпирических исследованиях живой и неживой природы. Поэтому, как наиболее встречаемый вариант, он и был назван нормальным, т.е. присущим всем массовидным феноменам. Предельным этот закон считается потому, что все другие законы распределения в определенной мере к нему приближаются. Хотя указанный закон был открыт независимо друг от друга тремя учеными в разное время, он также называется именем наиболее известного из них немецкого математика Ф. Гаусса. В психологии Ф. Гальтон доказал, что нормальному распределению подчиняются и психологические характеристики, в частности уровень развития способностей. В современной науке принят постулат, что все свойства в генеральной совокупности имеют нормальное распределение.

Нормальное распределение характеризуется тем, что в нем редко встречаются крайние значения признака (как низкие, так и высокие), а наиболее часто наблюдаются значения, близкие к средней величине. Графически нормальное распределение представляет собой симметричную колоколообразную (или холмообразную) кривую (рисунок 1).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *μ* ± *3 σ* | | | | | | | | |
|  | 99,7%  *μ* ± *2 σ* | | | | | | |  |
|  | 95,4%  *μ* ±  *σ* | | | |  | |
|  |  | |
|  | 68,3% | |  |
|  |  |
|  | |  |
|  |

*μ – 3 σ μ – σ* Мо *μ + σ μ + 3 σ*

*μ – 2 σ* Ме *μ + 2 σ*

М

Рисунок 1 – Кривая нормального распределения Гаусса

*Основные характеристики нормального распределения*

1.Совпадение моды, медианы и среднего арифметического.

2. Правило «трех сигм», описывающее зависимость между средним арифметическим, дисперсией и данными измерения: 68,3% значений измеренного признака располагаются в пределах М ± 1σ (именно поэтому кривая на рисунке 1 имеет характерный изгиб в точках, расположенных на расстоянии в одну σ от М); 95,5% значений измерений находятся в границах М ± 2σ; 99,7% – в зоне М ± 3σ. Иначе говоря, чем сильнее конкретное значение признака отклоняется от среднего, тем ниже вероятность его появления.

Однако далеко не все эмпирические распределения подчиняются нормальному закону, что объясняется разными причинами, основными из которых выступают ограниченность обследованной выборки и специфика самих измеряемых психологических переменных. Для того чтобы оценить, каково распределение значений в конкретной исследованной выборке, используются специальные параметры.

Параметры распределения – это те числовые характеристики, по которым можно судить о среднем значении измеренного признака и о его изменчивости. Наиболее важные параметры – это математическое ожидание (среднее арифметическое), дисперсия, показатели ассиметрии и эксцесса (точнее, в конкретной исследованной выборке – это оценки параметров, но допустимо использовать и термин параметры).

***Асимметрия* (А)** – характеристика формы распределения, показывающая вероятность большего отклонения от средней в какую-либо сторону (греч. asymmetria – несоразмерность). Иначе говоря, асимметрия – это скошенность кривой. Ассиметрия рассчитывается по формуле:

|  |  |  |
| --- | --- | --- |
| A = | Σ(xi – М)2  n . σ3 | . |

При нормальном распределении А = 0. На рисунке 2 отражен пример ассиметричного распределения.

y

|  |
| --- |
|  |

0 Мо Ме М x

Рисунок 2 – Кривая левосторонней асимметрии

При положительной асимметрии (кривая скошена влево, рисунок 2) более длинная часть кривой плотности распределения значений (хвост кривой) лежит правее среднего арифметического. Это означает, что большая часть эмпирических данных имеет более низкие значения, чем это можно было бы ожидать математически (т.к. мода и медиана ниже по значению, чем средняя арифметическая М > Ме > Мо). При отрицательной, правосторонней асимметрии (кривая скошена вправо), наоборот, и, соответственно, большая часть измерений имеет более высокие, чем ожидаемые математически значения (М < Ме < Мо).

***Эксцесс***– параметр распределения, характеризующий порядок преимущественного появления определенных значений измеренного признака (от лат. excessus – выход, отступление, уклонение). Эксцесс рассчитывается по формуле:

|  |  |  |
| --- | --- | --- |
| E = | Σ(xi – М)4  n . σ4 | – 3. |

Данная формула или коэффициент эксцесса измеряет пикообразность распределения. При нормальном распределении Е = 0.

На рисунке 3 изображены разные варианты эксцесса. При положительном эксцессе пик распределения около математического ожидания острый (островершинное распределение, Ех > 0), т.е. в обследованной выборке преобладают средние или близкие к ним значения. При отрицательном эксцессе пик распределения гладкий (плосковершинное распределение, он может быть также и вогнутым, Ех < 0), что свидетельствует о преимущественном появлении крайних значений измеренного признака.

f Ex = 0 Ex > 0

E x < 0

|  |
| --- |
|  |

x

Рисунок 3 – Примеры эксцесса

Расчет параметров нормальности распределения необходим исследователю в следующих случаях принятия решений:

– о выборе методов дальнейшей статистической обработки (в случае использования параметрической статистики в формулы расчета входят параметры распределения, чаще всего М и σ);

– о степени обобщения полученных результатов (в частности, при оценке возможности перенесения выводов на другие, чем исследованная, выборки необходимо оценить распределение).

**ИНДУКТИВНАЯ СТАТИСТИКА**

1. Индуктивная статистика и область ее применения.

2. Статистические гипотезы.

3. Непараметрические и параметрические критерии.

**1. Индуктивная статистика и область ее применения**

Этот раздел статистики называется также теорией статистического вывода. Он объединяет в себе такие методы обработки данных, с помощью которых можно делать заключения о больших группах (или о генеральной совокупности) на основе измерений, проведенных в небольших группах. Иначе говоря, индуктивная статистика позволяет переносить обобщения, полученные на ограниченных выборках, на те группы, которые эмпирически не исследовались.

Ценность научного исследования определяется двумя основными факторами: экономичность его осуществления (по затраченному времени, человеческим ресурсам и др.) при возможности получения таких результатов, которые можно было бы экстраполировать как можно шире (на другие выборки, на другие ситуации и др.). Например, исследовав группу пенсионеров в 100 человек на предмет одиночества, делать заключения об особенностях этого переживания всеми пожилыми людьми; изучив личностные особенности детей из семей с разными типами воспитания, утверждать, что эти особенности, обнаруженные в группах по 30 детей, присущи и другим детям, воспитывающимся в аналогичных условиях и т.п. Индуктивная статистика как раз и предназначена для того, чтобы, оберегая исследователя от лишних затрат, позволить ему сделать обоснованные широкие заключения.

Таким образом, индуктивная статистика позволяет делать выводы о значимости различий (в двух и более исследованных выборках; в одной выборке при разных замерах; в конкретной исследованной выборке и генеральной совокупности и др.). Все статистические выводы основываются на математической теории вероятностей.

1. **Статистические гипотезы**

Индуктивная статистика строится на принципе проверки статистических гипотез. Их формулировки необходимы для того, чтобы исследовательские идеи представить в более лаконичном и структурированном виде. Существует несколько типов гипотез.

***Нулевая гипотеза (Но)*** *–* предположение об отсутствии различий.

***Альтернативная гипотеза (Н1)*** – предположение о значимости различий. Эта гипотеза называется также рабочей, т.е. проверяемой эмпирическим путем.

Гипотезы также могут быть направленными и ненаправленными.

***Направленная гипотеза*** – отражает направление отношений между переменными (оптимисты добиваются в жизни больших успехов, чем пессимисты; девочки более конформны, чем мальчики…).

***Ненаправленная гипотеза*** – предположение, не содержащие указаний на возможное направление отношений между переменными (оптимисты и пессимисты отличаются по уровню достижения успеха; степень конформизма мальчиков и девочек различна…).

При принятии решения об отвержении нулевой гипотезы в пользу альтернативной (или решения о значимости различий) обязательно учитывается уровень достоверности. В психологии принято полагать, что такие решения можно принимать, если уровень достоверности или случайности возникновения найденного различия не превышает 5%. Еще более достоверным заключение будет при меньшей вероятности случая. Уровень достоверности в текстовой форме принято обозначать следующим образом: ρ ≤ 0,05 (5%-я возможность появления случайного результата, или 5 из 100); ρ ≤ 0,01 (1%-я вероятность, или 1 из 100); ρ ≤ 0,01 (0,1%-я вероятность, или 1 из 1 000).

Если уровень достоверности превышает допустимый 5%-й порог появления случайных различий, то считается, что найденная разница вполне может оказаться случайной. Это не позволяет принять надежное решение и отбросить нулевую гипотезу в пользу альтернативной.

При расчете большинства математических критериев уровни достоверности (или 5% и 1%-й уровни значимости) уже известны и представлены в специальных таблицах. Эти уровни еще называют критическими (или табличными) значениями. Полученное путем расчетов эмпирическое значение критерия обязательно сравнивается с его критическим значением.

Еще одной величиной, которая оказывается значимой для принятия решения о достоверности проверяемой статистической гипотезы, является число степеней свободы (обозначается как v или df). Это число представляет собой разницу между числом классов вариационного ряда и числом условий, при которых он был сформирован. К основным таким условиям относятся объем выборки (n), средние и дисперсии. Чаще всего число степеней свободы определяется объемом выборки (это в основном те критерии, для которых n не входит в формулу расчета). Критические значения для разных степеней свободы также отражены в специальных таблицах наряду с уровнем достоверности. В тексте это принято обозначать следующим образом: «при n = 56 критические значения критерия составляют…» или «при v = 3 критические значения критерия равны…».

При осуществлении статистических выводов сохраняется возможность совершения ошибок. Типы этих ошибок и их последствия обобщенно отражены в таблице 4.

Таблица 4 – Ошибки статистического вывода

|  |  |  |
| --- | --- | --- |
| Тип ошибки | Суть ошибки | Последствия |
| Ошибка первого рода  (α или ρ) | Отвержение истинной нулевой гипотезы | Неверное подтверждение рабочей гипотезы |
| Ошибка второго рода (β) | Сохранение ложной нулевой гипотезы | Неспособность определить статистически значимый результат |

Главное заключается в том, что вероятностью возникновения ошибок можно управлять. Как следует из таблицы 4, способность критерия не допустить ошибки первого рода – это и есть проанализированный выше уровень его значимости, т.е. чем выше задаваемый порог (оптимально ρ ≤ 0,001), тем меньше вероятность принять неверное решение о правильности рабочей гипотезы. Способность критерия не допустить ошибку второго рода – это его мощность, которая различается у разных критериев. Поэтому при выборе конкретного критерия для собственных расчетов следует помнить, что чем он мощнее, тем надежнее полученные результаты.

1. **Непараметрические и параметрические критерии**

Все методы, используемые в индуктивной статистике для доказательства значимости различий, делятся на две группы: непараметрические и параметрические.

***Непараметрические критерии*** – это критерии, не учитывающие параметров распределения. В первую очередь они предназначены для обработки данных, полученных в неметрических шкалах (номинативной и порядковой), поэтому оперируют частотами или рангами.

Преимуществами непараметрических критериев выступают:

– широта применения, т.к. они могут быть использованы для абсолютно любых шкал измерений;

– достаточная простота расчетов (большинство из которых не сложно производить «вручную»).

***Параметрические критерии*** – это критерии, включающие в формулу расчета параметры распределения (среднее арифметическое и дисперсия). Их использование ограничено следующими условиями:

– представленность эмпирических данных в метрических шкалах (интервалов и отношений);

– нормальное (или приближенное к нему) распределение измеренного признака.

При невозможности соблюдения указанных условий исследователь вынужден обращаться к непараметрическим критериям.

Самими известными и эффективными из параметрических критериев выступают критерий Стьюдента (t) и дисперсионный анализ. (Стьюдент – это псевдоним, под которым вынужден был скрываться от своего работатодателя, беспокоящегося о сохранении конфиденциальных данных о компании, английский ученый-статистик У. Госсет).

*t критерий Стьюдента (или t-тест)* используется при сравнении двух выборок. Сущность этого метода заключается в сравнении средних двух распределений. Рассчитывается t-критерий по-разному для разных выборок: для зависимых (парный критерий Стьюдента) и для независимых (непарный критерий Стьюдента).

*Дисперсионный анализ (тест F Снедекора)* – процедура, к которой обращаются при необходимости сравнения трех и более выборок с нормальным распределением. Сущность данного метода состоит в оценке разброса средних (в качестве показателя разброса используется дисперсия) как внутри каждой выборки, так и в целом для всей исследованной совокупности. Если все выборки принадлежат одной и той же популяции, то разброс значений между разными выборками должен совпадать с разбросом данных внутри каждой из этих выборок. Дисперсионный анализ позволяет решить также очень важную исследовательскую задачу оценки взаимодействия нескольких факторов, влияющих на изменение переменной.

Таким образом, параметрические критерии являются более мощными, но только для анализа признаков с нормальным распределением. Однако подобное распределение в реальных психологических исследованиях встречается нечасто. Поэтому более распространенными в психологии являются непараметрические статистики. Еще одна причина их популярности состоит в том, что наиболее часто встречаемой измерительной шкалой выступает порядковая.

**НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ В РЕШЕНИИ ЗАДАЧ**

**ПСИХОЛОГИЧЕСКОГО ИССЛЕДОВАНИЯ**

1. Выявление различий в уровне выраженности измеренного признака для независимых выборок.

2. Оценка сдвига значения измеренного признака для зависимых выборок.

3. Определение различий в распределении признака.

4. Многофункциональные критерии.

**1. Выявление различий в уровне выраженности измеренного признака для независимых выборок**

Представленные далее критерии разработаны для неметрических шкал. Однако их можно считать универсальными, поскольку применение данных способов оценки различий допустимо и для более мощных метрических шкал измерений (в результате процедуры понижения мощности).

***Номинативная шкала измерений***

***χ² – критерий Пирсона***(читается «хи квадрат») – критерий, позволяющий оценивать вероятность появления разных значений признака. Этот критерий отличается широтой использования, поскольку он позволяет определить достоверность различий в следующих основных случаях: 1) между двумя и более эмпирическими распределениями одной психологической переменной; 2) между эмпирическим и теоретическим распределениями (например, нормальным).

*Условия применения критерия χ²:*

– средний объем выборки (т.е. n ≥ 30: чем больше объем, тем точнее значения критерия);

– количество значений для одного разряда переменной должно быть не менее 5 (Например, проверяется гипотеза о том, что представители молодежных субкультур чаще переживали безответную любовь, чем молодые люди, придерживающиеся традиционной культуры. «Молодежная субкультура» – это номинативная переменная, которая в случае конкретного исследования может быть представлена следующими четырьмя разрядами: «готы», «панки», «ролевики», «эмо». Т.е. количество несчастных влюбленных – в соответствии с гипотезой – в каждом разряде не должно быть менее 5.);

– достоверность различий оценивается по специальной таблице критических значений с учетом степеней свободы: v = k – 1, где k – количество разрядов признака (таблица ограничена 100 степенями свободы);

– решение о достоверности различий принимается, если χ²эмп ≥ χ²кркак минимум при ρ ≤ 0,05.

***Порядковая шкала измерений***

***Q – критерий Розембаума*** – довольно простой в подсчетах (но одновременно и не очень мощный) критерий, позволяющий оценить различия между двумя выборками по количественно измеренному признаку.

*Условия применения критерия Розембаума (Q):*

– в выборке должно быть не менее 11 респондентов;

– количество респондентов в двух выборках может не совпадать, однако разница не должна быть очень значимой, желательно, чтобы она не превышала 20 человек (для средних по объему выборок);

– количественно измеренный признак должен варьировать (три значения маловато для установления различий);

– наличие «хвостов», т.е. несовпадение упорядоченных по величине признака рядов наблюдений в двух выборках. Для лучшего понимания «хвост» можно представить наглядно следующим образом:

|  |  |  |
| --- | --- | --- |
| хвост  1 1 2 2 2 | 3 4 4 5 5 6 6 6  3 4 4 5 6 6 6 | 7 7 9 9 9  хвост |

– достоверность различий оценивается по специальной таблице критических значений с учетом числа респондентов в каждой выборке; таблица ограничена для n = 26, если выборки больше, то считается, что Qкр = 8 для ρ ≤ 0,05 и Qкр = 10 ρ ≤ 0,01;

– решение о значимости различий принимается, если Qэмп ≥ Qкркак минимум при ρ ≤ 0,05.

***U – критерий Манна-Уитни*** – критерий, позволяющий оценить различия между двумя выборками по уровню количественно измеренного признака. Существуют различные способы расчета критерия и таблиц критических значений.

*Условия применения критерия Манна-Уитни (U):*

– критерий хорошо применим для малых выборок – от 3 человек в каждой (допустима также одна выборка объемом в два респондента, но тогда в другой должно быть не менее пяти человек);

– верхний порог выборок для данного критерия n = 60, однако уже при n = 20 расчеты становятся крайне трудоемкими, т.к. метод предполагает ранжирование значений;

– количество респондентов в двух выборках может не совпадать, но разница не должна быть очень существенной (тем более что речь идет и так о малых по объему группах);

– достоверность различий оценивается по специальной таблице критических значений с учетом числа респондентов в каждой выборке;

– решение о значимости различий принимается, если Uэмп ≤ Uкркак минимум при ρ ≤ 0,05.

***H – критерий Крускала-Уоллиса*** – критерий, используемый для оценки различий в уровне количественно измеренного признака между тремя и более выборками. Позволяет доказать, что переменная существенно изменяет свои значения в разных группах, но не дает возможности определить направление этих изменений. Это своеобразное продолжение критерия U – Манна-Уитни для более чем двух выборок.

*Условия применения критерия Крускала-Уоллиса (H):*

– критерий хорошо применим для малых выборок – от 3 человек в каждой, возможна также и пропорция 4:2:2 (допустима также пропорция 3:2:2, но тогда уровень значимости будет минимальным – 5%);

– достоверность различий оценивается по специальной таблице критических значений с учетом числа респондентов в каждой выборке (в таблице отражены значения для 3 выборок с максимальным объемом 5 человек);

– решение о значимости различий принимается, если Hэмп ≥ Hкр;

– при большем количестве выборок или увеличении их объема необходимо обратится к таблице критических значений χ², с учетом степеней свободы по следующей формуле: v = c – 1, где c – количество сопоставляемых выборок.

**S – *критерий тенденций Джонкира***

Данный критерий также применяется для оценки различий в уровне количественно измеренной переменной между тремя и более выборками. Он позволяет не только зафиксировать наличие различий, но и определить направления изменения признака при переходе от выборки к выборке. Критерий Джонкира может рассматриваться и как мера связи между двумя признаками, измеренными как минимум в порядковых шкалах (т.е. если сами выборки образованы на основании некоторого количественного признака – стаж деятельности, уровень интеллекта и т.п. – и если интересующая психолога переменная также выражена количественно – уровень стрессоустойчивости, конфликтности и др.).

*Условия применения критерия тенденций Джонкира (S):*

– сравниваемые выборки должны быть одинаковы по объему;

– критерий хорошо применим для малых выборок – от 2 до 10 человек в каждой;

– достоверность различий оценивается по специальной таблице критических значений с учетом числа респондентов в каждой выборке (в таблице отражены значения для 6 выборок);

– решение о значимости различий принимается, если Sэмп ≥ Sкр;

– при большем количестве выборок или увеличении их объема необходимо обратится к таблице критических значений критерия Крускала-Уоллиса.

**2. Оценка сдвига значения измеренного признака для зависимых выборок**

Представленные далее статистические методы позволяют определить существование различий в уровне исследованного признака, если измерение проводилось на связанных выборках.

***Номинативная шкала измерений***

***М – критерий Макнамары***

Данный критерий используется для дихотомической шкалы измерений (замеры «до» и «после»).

*Условия применения критерия Макнамары (М):*

– для проведения расчетов значим не объем выборки (он может быть любым), а число наблюдений в каждом замере, которое фиксируется в таблице 2 на 2, особенно число так называемых «сдвигов» или изменений. Например, в группе подростков проводился анализ их отношения к татуировкам «до» и «после» беседы о последствиях такой модификации тела. В таблицу под соответствующими буквами подставляются следующие числовые значения:

А – количество подростков, которые «до» и «после» беседы ответили, что им татуировки нравятся;

В – количество подростков, которые изменили свое мнение, а именно: до беседы они утверждали, что им татуировки нравятся, а после нее, что не нравятся (сдвиг значений в одну сторону);

С – количество учащихся, которые также изменили свое мнение, а именно: до информирования им татуировки не были симпатичны, а после беседы, напротив, они стали положительно относится к такой модификации (сдвиг значений в другую сторону);

D – количество подростков, которые и «до» и «после» остались убеждены в том, что татуировки им не нравятся.

|  |  |  |
| --- | --- | --- |
| После беседы  До беседы | Нравятся | Не нравятся |
| Нравятся | А | В |
| Не нравятся | С | D |

Критерий Макнамары неприменим:

– если В = С (в этом случае следует обратиться к критерию χ² Пирсона);

– если В + С ≤ 20, используется один способ расчета критерия, значения которого находятся в специальной таблице значений для Mэмп; критические значения фиксированы: для ρ ≤ 0,05 Мкр = 0,025, для ρ ≤ 0,01 Мкр = 0,005;

– если В + С > 20, то используется другой способа расчета критерия, для которого также существуют свои фиксированные критические значения, а именно: для ρ ≤ 0,05 Мкр = 3,841, для ρ ≤ 0,01 Мкр = 6,635;

– статистическая гипотеза о значимости различий принимается, если эмпирические значения критерия Макнамары превышают существующие критические значения.

***G – критерий знаков***

Используется для определения общего направления сдвига измеренной переменной (которая может быть как номинативной, так и порядковой) в исследуемой выборке. Он позволяет установить случайность/неслучайность «типичного» сдвига (т.е. такого изменения, которое «на глаз» встречается в выборочной совокупности чаще).

*Условия применения критерия знаков (G):*

– nmin = 5, nmax = 300;

– нулевые реакции (т.е. те, в которых не произошло изменений между «до» и «после») из расчетов исключаются, и, соответственно, выборка уменьшается на это число;

– достоверность различий оценивается по специальной таблице критических значений для n без нулевых реакций (а не для всей исследованной совокупности);

– решение о значимости различий принимается, если Gэмп ≤ Gкр.

**Q** – **критерий Кохрена (или Кочрена)**

Данный критерий применяется, если возникает необходимость сравнить переменную, которая измерялась на одной выборке более трех раз.

*Условия применения критерия Кохрена (Q):*

– качественно определяются условия измерений, а сама интересующая переменная должна быть представлена количественно, поскольку этот тест основан на расчете однородности дисперсий;

– достоверность различий оценивается по специальной таблице значений, согласно которой степени свободы ограничены количеством выборок (20) и условиями измерения (9);

– Н1 принимается, если Gэмп превышает критические табличные значения.

***Порядковая шкала измерений***

Вышеописанный ***G – критерий знаков*** может успешно применяться и для данной шкалы измерений (особенно если значения переменной не сильно варьируют).

***T – критерий Вилкоксона*** предназначен для оценки различий между двумя количественно измеренными показателями на одной выборке респондентов. В отличие от критерия знаков он может оценивать не только направление сдвигов, но и их интенсивность или степень выраженности.

*Условия применения критерия Вилкоксона (Т):*

– имеющиеся наблюдения должны варьировать в достаточно широком диапазоне, поскольку выраженность сдвига оценивается по абсолютной разнице значений «до» и «после», которые потом ранжируются (если их всего три, то проще ограничиться критерием знаков);

– nmin = 5, nmax = 50;

– аналогично критерию знаков нулевые сдвиги из расчетов исключаются и соответственно, выборка уменьшается на это число;

– достоверность различий оценивается по специальной таблице критических значений для n, из которой уже исключены нулевые сдвиги;

– решение о значимости различий принимается, если Тэмп ≤ Ткр.

***χ²r*** – ***критерий Фридмана***

Является продолжением критерия Вилкоксона для трех и более замеров количественной переменной. Позволяет фиксировать только наличие изменений, но не их направление.

*Условия применения критерия Фридмана (χ²r):*

– необходимо как минимум 2 испытуемых, участвовавших в трех замерах;

– для малой выборки (n ≤ 9) и для небольшого числа замеров (с ≤ 4) используются специальные таблицы критических значений;

– для большей выборки или при увеличении числа условий требуется обратиться к таблице критических значений χ² Пирсона, с учетом степеней свободы по формуле: v = с – 1.

***L – критерий тенденций Пейджа***

Выступает продолжением критерия Фридмана и позволяет не только зафиксировать различия, но и установить их направление.

*Условия применения критерия тенденций Пейджа (L);*

– ограниченное число испытуемых: nmin = 2, nmax = 12;

– ограниченное число условий: сmin = 3, сmax = 6;

– при невыполнении данных условий следует остановиться на χ²r – критерии Фридмана и констатировать только неслучайность изменений.

**3. Определение различий в распределении признака**

К этой категории методов непараметрической статистики психолог обращается для решения следующих основных задач:

во-первых, если недостаточно уже рассмотренных методов обработки эмпирических данных и он вынужден использовать более мощные параметрические критерии. Такой переход возможен только если распределение интересующей его переменной в выборочной совокупности соответствует нормальному;

во-вторых, если для обоснования сделанных заключений ему необходимо сравнить параметры распределения исследованной выборочной совокупности с генеральной (в которой априори все признаки подчиняются закону нормального распределения).

***Номинативная шкала измерений***

Если психологические данные получены в номинативной шкале, то вышеуказанные задачи решаются посредством уже описанного выше ***χ² – критерия Пирсона.***

***Порядковая шкала измерений***

Те же задачи, что и критерий Пирсона качественных данных, для данных, измеренных количественно, решает критерий Колмогорова – Смирнова.

***λ – критерий Колмогорова – Смирнова***

Данный критерий позволяет оценивать различия между одним и более эмпирическими распределениями, а также различия между эмпирическим и теоретическим распределениями.

*Условия применения критерия Колмогорова – Смирнова (λ):*

– средний и большой объем сравниваемых выборок (nmin = 50);

– расчет основан на определении рангов, поэтому значения переменной обязательно должны быть упорядочены либо по убыванию, либо по возрастанию;

– оценка достоверности различий производится по специальным таблицам критических значений;

– решение о достоверности различий принимается, если эмпирическое значение критерия равно или превышает указанные в таблицах критические значения.

1. **Многофункциональные критерии**

К этой группе относятся критерии, которые применимы для решения различных задач и подходят для разных шкал измерений. Наиболее распространенными в обработке психологических данных являются критерий биноминального распределения и критерий Фишера.

**m – биноминальный критерий**

Данный критерий позволяет оценить случайность интересующего исследователя эффекта в одной выборке испытуемых. Расчет критерия основывается на сопоставлении долей единицы в эмпирическом и предполагаемом распределениях (теоретически заданном или ожидаемом).

*Условия применения биноминального критерия (m):*

– выборка может иметь малый объем n ≤ 30;

– минимальное количество наблюдений должно быть равно 5 (возможно и меньше, но только для определенных случаев), максимальное число наблюдений 300 (может быть и 50, что обусловлено разными таблицами критических значений);

–  оценка достоверности различий производится по специальным таблицам критических значений;

– решение о достоверности различий принимается, если эмпирическое значение m-критерия равно или превышает указанные в таблицах критические значения.

**φ – критерий углового преобразования Фишера**

Один из самых универсальных критериев. Его применяют для решения самых разнообразных задача (оценка различий в одной или нескольких выборках, сравнения показателей распределения). К нему как к «палочке-выручалочке» обращаются, если другие критерии неприменимы или их использование оказывается очень трудоемким. Этот критерий построен на оценке долей, выраженных в процентах.

*Условия применения углового преобразовании Фишера (φ):*

– выборка может иметь любой объем: 5 ≤ n ≤ ∞;

– выборки могут быть зависимыми и независимыми;

– количество измерений также может быть различным (два, три и больше);

– решение о достоверности различий принимается с помощью специальной таблицы критических значений, если φэмп ≤  φкр.

Помимо указанных многофункциональных критериев к этой группе можно смело отнести уже знакомый по предыдущим описаниям критерий χ² Пирсона, поскольку он также предназначен для решения различных задач и может быть использован не только для номинативной, но и для более мощных измерительных шкал.

**КОРРЕЛЯЦИОННЫЙ АНАЛИЗ**

1. Корреляционный анализ и область его применения.

2. Выбор коэффициента корреляции в зависимости от шкалы измерения.

3. Коэффициенты корреляции для данных, полученных в разных шкалах измерений.

**1. Корреляционный анализ и область его применения**

Корреляционный анализ является одним из разделов математической статистики. Это статистический метод оценки связи между двумя переменными (обозначаемые буквами X и Y).

***Корреляция*** (лат. correlation – соотношение) – это согласованность изменения двух признаков, т.е. соотношение изменчивости одного с изменчивостью другого.

Необходимость в корреляционном анализе возникает, если перед исследователем стоит задача получить информацию об одной переменной с помощью другой или задача предсказания значений одной переменной по значениям другой. Этот метод может применяться и в одной выборочной совокупности, и в разных.

В корреляционном анализе возможная связь между переменными оценивается по следующим параметрам:

– форма связи (линейная или нелинейная);

– знак связи (положительная или отрицательная);

– теснота связи (высокая или низкая).

***Коэффициент корреляции*** – это показатель степени связи между двумя переменными. Его значения (для большинства критериев) находятся в диапазоне от –1 до +1.

Коэффициент корреляции, приближающийся к +1, означает наличие линейной, положительной и высокой тесноты связи между переменными. Положительный коэффициент корреляции трактуется как наличие прямо пропорциональной зависимости между переменными: чем больше возрастают значения одной из них, тем больше увеличиваются и значения другой. Диаграмма рассеяния значений двух переменных при такой связи отражена на рисунке 4.

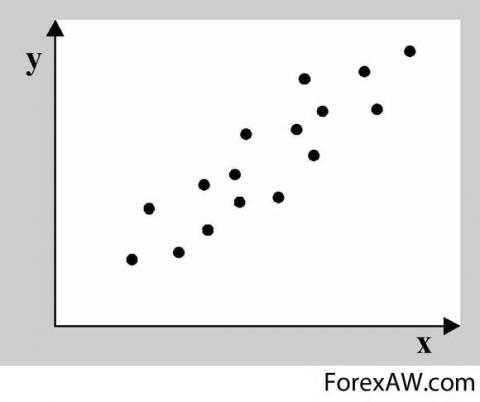


Рисунок 4 – Графическое изображение прямой корреляционной связи

Коэффициент корреляции, близкий к –1, свидетельствует о наличии линейной, отрицательной и высокой зависимости между измеренными признаками. Отрицательный коэффициент корреляции означает существование обратно пропорциональной зависимости между переменными: чем больше возрастают значения одной из них, тем больше убывают значения другой (рисунок 5).

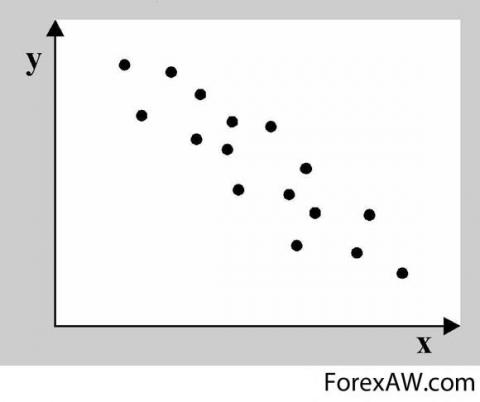


Рисунок 5 – Графическое изображение

обратной корреляционной связи

Коэффициент корреляции, приближающийся к 0 (рисунок 6), говорит об отсутствии связи между переменными (нелинейная форма и отсутствие связи).

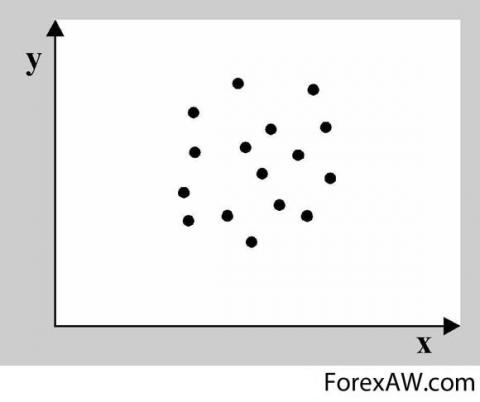


Рисунок 6 – Графическое изображение отсутствия связи

между значениями двух переменных

Нулевая гипотеза при проведении корреляционного анализа формулируется как отсутствие связи, альтернативная – как наличие связи (положительной или отрицательной).

Как и другие статистические критерии, коэффициент корреляции имеет свои уровни значимости. Обнаруженную связь можно считать достоверной, если эмпирическое значение коэффициента превышает его критическое значение на допустимом уровне (т.е. ρ не должно превышать 5%-й порог случайности) при определенных степенях свободы.

В процессе качественной интерпретации коэффициента связи исключительно важно оставаться в пределах линейной связи между переменными и не пытаться объяснять даже сильную зависимость иной, причинной связью. Задачу установления причинно-следственных отношений между переменными можно решить лишь экспериментальным путем. Корреляционный анализ устанавливает только связь между переменными, но он не может дать данных о том, как одна переменная влияет на изменения другой.

При корреляционном анализе возможно также появление ложных корреляций, т.е. установление значимой корреляционной зависимости между заведомо не связанными переменными (например, изменение скорости реакции в зависимости от времени года; взаимосвязь стиля общения и темперамента руководителя и т.п.). Причина описанной ошибки заключается не в неправильности математических расчетов, а в недостаточности теоретического анализа, в ходе которого какой-то важный фактор не был замечен психологом.

**2. Выбор коэффициента корреляции в зависимости от шкалы измерения**

Математической статистикой разработаны коэффициенты, позволяющие оценивать связи между переменными, измеренными в различных шкалах.

***Шкала наименований***

Для переменных (X и Y), измеренных в дихотомической шкале, используется ***коэффициент ассоциации Пирсона (φ)***. К. Пирсон – выдающийся английский философ, математик и биолог. Он считается основателем математической статистики и биометрии (совместно с Ф. Гальтоном). Созданный К. Пирсоном коэффициент может рассчитываться разными способами (один из которых выступает в подготовке таблицы 2 на 2 поля, т.е. две переменные, каждая из которых принимает только два состояния), поэтому для его обозначения используется и другой термин *«коэффициент сопряженности»*.

*Условия применения коэффициента ассоциации φ:*

– переменные X и Y представлены в дихотомической шкале (0 и 1);

– переменные X и Y имеют одинаковое число наблюдений;

– для определения уровней достоверности эмпирического значения φ отсутствуют специальные таблицы, поэтому он рассчитывается по формуле:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Тф = | ׀rэмп׀ | . | √ | n – 2  1 – rэмп2 | ; |

– при оценке уровня достоверности коэффициента ассоциации φ используются таблицы критических значения для t-критерия Стьюдента (для числа степеней свободы: n – 2).

Следует отметить, что относительно понимания дихотомической шкалы между специалистами существуют разногласия. Некоторые (к примеру, В.Н. Дружинин) считает ее редуцированной интервальной шкалой, другие (Дж. Глас, Дж. Стенли, Е.Е. Сидоренко и др.) относят к номинативным шкалам. Однако при наличии разногласий в понимании сущности данной шкалы для проведения корреляционного анализа подобных признаков обращаются именно к коэффициенту ассоциации.

В случае, если две номинативные переменные принимают более чем два значения, то используется модифицированная версия коэффициента ассоциации – ***коэффициент корреляции Крамера (v)****.* Числовые значения данного коэффициента имеют диапазон от 0 (отсутствие связи) до 1 (полная связь). Поэтому интерпретация данного коэффициента имеет свою специфику, а именно: он показывает только тесноту связи, но не указывает ее направление.

***Шкала порядка***

Для переменных, измеренных в ранговой шкале, используется два коэффициента корреляции: коэффициент корреляции Спирмена (rs) и коэффициент корреляции Кендалла (τ, читается «тау»). Ч. Спирмен – выдающийся английский психолог, разработавший множество методов математической статистики (в том числе и факторный анализ), М. Кендалл – известный английский статистик.

***Коэффициент корреляции Спирмена rs*** *и его условия:*

– переменные X и Y представлены в ранговой шкале (иногда этот коэффициент так и называется – коэффициент ранговой корреляции);

– параметры нормальности распределения не учитываются;

– переменные X и Y имеют одинаковое число наблюдений;

– объем выборки ограничен от n = 5 до n = 40;

– при оценке уровня достоверности коэффициента корреляции rs используются специальные таблицы критических значений для rs Спирмена (со степенями свободы n – 2).

Если измеренные значения переменных превышают 40 наблюдений, то можно обратиться к таблице критических значений коэффициента корреляции Пирсона.

Более простым в расчете считается коэффициент корреляции Кендалла. Этот анализ основан на попарном сравнении между собой наблюдений. Поэтому эта простота относится только к маленьким выборкам, т.к. чем большее число значений переменной, тем большее количество пар надо перебрать.

***Коэффициент корреляции Кендалла τ*** *и условия его применения:*

– переменные X и Y представлены в порядковой шкале;

– параметры нормальности распределения не учитываются;

– переменные X и Y имеют одинаковое число наблюдений;

– использование одинаковых рангов не допускается;

– для оценки уровня достоверности коэффициента корреляции Кендалла специальные таблицы отсутствуют. Поэтому критические значения рассчитываются по вышеописанной формуле Тф, а для их определения используются таблицы критических значения t-критерия Стьюдента (со степенями свободы n – 1).

По сравнению с коэффициентом Спирмена коэффициент Кендалла более точен к определению ρ-уровня.

***Шкалы интервалов и отношений***

Для установления связи между метрическими переменными применяется ***коэффициент корреляции Пирсона (r)***.

*Условия применения коэффициента корреляции Пирсона r:*

– переменные X и Y представлены либо в интервальной шкале, либо в шкале отношений;

– распределение значений обеих переменных приближается к нормальному;

– переменные X и Y имеют одинаковое число наблюдений;

– объем выборки ограничен от n = 5 до n = 100;

– при оценке уровня достоверности коэффициента корреляции r используются специальные таблицы критических значений для r Пирсона (со степенями свободы n – 2).

Кроме вышеперечисленных особенностей применения коэффициента корреляции Пирсона, он, как параметрический критерий, оказывается довольно чувствительным к «выбросам» крайних значений переменных. Поэтому при наличии экстремальных значений измеренного признака даже при нормальном распределении лучше понизить мощность шкалы и использовать коэффициент ранговой корреляции Спирмена.

**3. Коэффициенты корреляции для данных, полученных в разных шкалах измерений**

В психологических исследованиях довольно часто разные переменные измеряются в разных шкалах. Поэтому для этих случаев в корреляционном анализе также есть свои специальные критерии, а именно рангово-бисеральный и бисеральный коэффициенты.

***Рангово-бисеральный коэффициент корреляции (Rrb)*** *и условия его применения:*

– одна переменная представлена в дихотомической шкале, другая – в порядковой;

– параметры нормальности распределения не учитываются;

– переменные X и Y имеют одинаковое число наблюдений;

– для определения уровня достоверности применяется формула для Тф, а для его оценки также следует обратиться к таблицам критических значений t-критерия Стьюдента (со степенями свободы n – 2).

***Бисеральный коэффициент корреляции (Rбис)*** *и условия его использования:*

– переменная Х представлена в дихотомической шкале, переменная Y – в метрической (интервалов или отношений);

– значения переменной Y должны описываться законом нормального распределения;

– переменные X и Y имеют одинаковое число наблюдений;

– для определения уровня достоверности применяется формула для Тф, а для его оценки также следует обратиться к таблицам критических значений t-критерия Стьюдента (со степенями свободы n – 2).

Показатель этих коэффициентов корреляции также имеет значения от –1 до +1, однако при интерпретации полученных с их помощью результатов этот знак не учитывается.

В заключение следует отметить, что возможности корреляционного анализа не исчерпываются вышеописанными. Существуют и более сложные варианты связей между переменными. Например, множественная корреляция (между тремя переменными X, Y и Z), частная корреляция (связь двух зависимых переменной с третьей), корреляционное отношение (выраженное двумя показателями, поскольку связь устанавливается с двух сторон, как с позиции X, так и позиции Y) и др.

**МЕТОДЫ МНОГОМЕРНОГО АНАЛИЗА**

1. Дисперсионный анализ.

2. Факторный анализ.

3. Многомерное шкалирование.

4. Кластерный анализ.

5.Регрессионный анализ.

6. Дискриминантный анализ.

**1. Дисперсионный** **анализ**

Дисперсионный анализ – метод, предложенный Р. Фишером (F-тест). Называется также ANOVA (Аnalysis of Variance – анализ вариативности). Представляет собой анализ изменчивости признака под влиянием каких-либо контролируемых факторов. Чем выше эмпирическое значение F, тем больше признак варьирует под воздействием интересующего исследователя фактора (или их взаимодействия). Дисперсионный анализ позволяет устанавливать только изменение признака, но не направление этих изменений.

Прежде чем приступать к дисперсионному анализу, психолог должен содержательно обосновать, почему какие-либо переменные выступают причиной (это и есть факторы), а другие – следствиями (они называются результативными признаками).

*Условия применения дисперсионного анализа:*

– фактор может быть выражен в любой шкале (в том числе в номинативной: например, по признаку пола создаются две группы; по уровням мотивации – низкому, среднему, высокому – три группы и др.);

– результативные признаки должны быть выражены в метрических шкалах, а их значения в выборке соответствовать нормальному распределению (при определенных условиях это требование можно и обойти);

– выборки могут быть независимыми и зависимыми;

– сравнивать возможно две и более выборок;

– выборки могут быть любого объема (от 2 до ∞);

– если выборки разного объема, то значимым выступает требование соблюдения равенства дисперсий.

Дисперсионный анализ имеет различные виды, конкретный выбор которого зависит:

– от числа контролируемых факторов, в соответствии с которыми выделяются однофакторный, двухфакторный, мультифакторный дисперсионный анализ;

– количества уровней каждого из контролируемых факторов (двух-, трехуровневый и т.п. дисперсионный анализ);

– числа исследуемых сочетаний факторов: если анализируются сочетания всех факторов на всех уровнях, то необходим полный дисперсионный анализ; если часть сочетаний опускается, тогда проводится дробный дисперсионный анализ.

Аналогами однофакторного дисперсионного анализа выступают непараметрические критерии – Н критерий Крускала-Уоллиса (для независимых выборок) и критерий χ²r Фридмана (для зависимых выборок), позволяющие оценить различия измеренного признака в трех и более условиях. Однако названные непараметрические критерии предназначены только для выборок малого объема. Поэтому при увеличении выборки следует обратиться к однофакторному дисперсионному анализу.

Дисперсионный анализ является единственным и незаменимым методом при решении задачи выявления одновременного воздействия двух и более факторов на интересующий признак (речь идет о двухфакторном и мультифакторном анализах). Иначе говоря, только благодаря дисперсионному анализу можно выявить взаимодействие нескольких факторов в их влиянии на один и тот же результативный признак.

**2. Факторный анализ**

Данный метод был специально разработан для решения психологических задач, а позднее оказался очень востребованным и другими науками. Идея его создания принадлежит Ф. Гальтону, высказавшему идею о наличии некоторого общего, скрытого фактора способностей, благодаря которому можно объяснить их индивидуальную изменчивость. В 30-е гг. ХХ в. эта идея была воплощена Ч. Спирменом в однофакторном анализе, итогом которого стала модель общего интеллекта. Позднее были разработаны многофакторные модели интеллекта (Ч. Терстоун) и личности (Р. Кеттелл).

Психическая реальность является многомерной. Например, индивидуально-психологические особенности личности измеряются очень большим числом переменных, что значительно затрудняет их анализ и интерпретацию. Кроме того, переменные, как правило, каким-то образом взаимосвязаны друг с другом, за ними может стоять некоторый общий фактор/факторы. Поэтому возникает необходимость это исходное множество объединить, укрупнить для перехода на более высокий уровень обобщения и понимания изучаемых феноменов. Именно эту задачу позволяет решить факторный анализ.

Факторный анализ – это метод выявления корреляционных отношений между переменными, позволяющий определить структуру взаимосвязи между ними и сократить их число. Иначе говоря, этот метод выступает еще как метод структурной классификации и метод сокращения (редукции) данных.

*Условия применения факторного анализа:*

– переменные могут быть измерены только в метрических шкалах;

– переменные должны быть независимыми;

– распределение измеренных признаков должно приближаться к нормальному;

– оптимально, чтобы связи между переменными имели линейный характер и хотя бы несколько переменных коррелировали друг с другом на высоком уровне значимости;

– рекомендуемый объем выборки равен 100 (при меньшем объеме появляются слишком большие стандартные ошибки корреляций).

Спецификой факторного анализа выступает то, что обрабатываются не сами «сырые» данные, а корреляционные матрицы (или матрицы интеркорреляций). Существуют различные виды факторизации: метод главных компонент, центроидный, максимального правдоподобия и др. Вне зависимости от конкретной процедуры фактор выступает как некоторый искусственный математический показатель, появляющийся в результате преобразования матрицы корреляций. В один фактор объединяются переменные, максимально друг с другом взаимосвязанные, другой фактор включает другие переменные, также значимо коррелирующие друг с другом и т.п. Т.е., образно говоря, фактор представляет собой пучок взаимосвязанных друг с другом переменных. Количество выделяемых факторов определяется исследовательскими задачами и тем, что для психолога выступает неслучайным в общем контексте всей научной работы (поэтому существует несколько правил отбора значимых факторов: критерий Кайзера, критерий «каменистой осыпи», процент объясненной дисперсии и др.).

Результаты факторного анализа чаще всего представляются как перечисление образующих фактор шкал с указанием их факторной нагрузки. К примеру, в исследовании модели успешного современного человека в группе 50 взрослых людей (Е.И. Медведская, 2013) был выделен следующий фактор:

понимать самого себя 0,985

отсеивать ненужную информацию 0,822

понимать язык тела 0,777

слушать другого человека 0,687

систематизировать знания 0,551

разбираться в других людях 0,490

хорошо одеваться 0,418

Как следует из представленного примера, факторный анализ дает взаимосвязь наиболее коррелирующих шкал (причем эта корреляция не всегда логична с точки зрения теории), а перед психологом стоит довольно сложная задача обобщения обнаруженных разнородных взаимосвязей и их интерпретации как непротиворечивого целого.

**3. Многомерное шкалирование**

Данный метод представляет собойопределенную альтернативу факторному анализу. Многомерное шкалирование также предназначено для редукции данных, выявления латентных факторов с целью их последующего объяснения. Если в факторном анализе такое сокращение достигается посредством обработки корреляционных матриц, то при многомерном шкалировании анализируется матрица сходства расстояний.

В итоге процедур каждому измеренному объекту приписываются определенные числовые значения на осях координат. Эти координаты образуют некоторое пространство (как правило, геометрическое, Эвклидово) со значительно меньшей размерностью, чем исходный массив данных. Координаты нового смоделированного пространства представляют собой те самые незаметные глазу, латентные переменные, обнаруженные в итоге математических процедур.

Если сохранить размерность пространства равной числу измеренных переменных, то можно получить точное воспроизведение исходной матрицы расстояний между объектами. Однако целью многомерного шкалирование выступает выявление нескольких важнейших факторов (которые также называются латентными переменными или вспомогательными шкалами), поэтому созданная пространственная модель, безусловно, несколько отличается от исходной. Но и в ней те объекты, которым в исходной матрице соответствуют большие меры различий, находятся далеко друг от друга, а те, которые более подобны друг другу, располагаются, соответственно, ближе.

Для оценки качества подгонки результативной модели исходной матрице сходств расстояний используется специальная мера, которая называется стресс. Чем меньше его значение, тем лучше математически смоделированная матрица расстояний репрезентирует исходную, эмпирическую структуру взаимосвязей между переменными.

*Условия применения многомерного шкалирования:*

– переменные могут быть измерены в порядковой шкале и не подчиняться закону нормального распределения (неметрический метод многомерного шкалирования);

– переменные могут быть представлены в шкалах интервалов или отношений (метрический метод многомерного шкалирования);

– число переменных должно быть достаточно большим.

Результаты многомерного шкалирования в основном представляются графически. К примеру, в исследовании зрительного восприятия 18 букв русского алфавита 50 взрослым респондентами при попарном представлении букв предлагалось оценить их по шкале «похожи – не похожи» (Р.М. Фрумкина, П.Ф. Андрукович, А.Ю. Терехина, 1976). Было построено 50 матриц сходств, при обработке которых методом многомерного шкалирования обнаружилось два латентных фактора, представленных на рисунке 7.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Прямые элементы – остроугольные элементы |  |  |  |  |  |  |  |  |  |  |  | **О** |  |  |  |  |  |  | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
|  |  |  |  |  |  |  |  |  |  |  | **С** |  |  |  |  |  |  | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **З** | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **В** |  | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  | **Р** |  |  |  | |
|  |  |  | **К** |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **Б** |  |  | |
|  | **У** |  |  |  |  |  |  |  |  |  |  |  | **Е** |  |  |  |  | |
|  |  |  |  | **М** |  |  |  | **Н** |  |  |  |  |  |  |  |  |  | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
|  |  |  |  | **Н** |  |  |  |  |  | **Г** |  |  |  |  |  |  |  | |
|  |  | **Л** |  |  |  |  |  | **П** |  |  |  |  |  |  |  |  |  | |
|  |  |  | **А** |  |  |  |  |  |  | **Т** |  |  |  |  |  |  |  | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | |
|  |  |  |  |  |  |  |  | **Д** |  |  |  |  |  |  |  |  |  | |
|  | | | | | | | | | | | | | | | | | |

Остроугольные элементы – прямоугольные элементы

Рисунок 7 – Пространство букв русского алфавита,

полученное в результате многомерного шкалирования

В отличие от факторного анализа, методы многомерного шкалирования не так распространены в отечественной психологии, что обусловлено целым рядом причин, в том числе и отсутствием четкого теоретико-методического обоснования соответствия формальной модели реально существующей, содержательной области психической реальности.

**4. Кластерный анализ**

Данный вид анализа применяется для обработки качественных данных. Кластерный анализ объединяет различные процедуры, позволяющие сгруппировать множество объектов в классы (кластер или группа, гроздь) таким образом, что сходные объекты попадают в один класс. В итоге процедур кластеризации обнаруживаются группы подобных между собой объектов.

Суть кластерного анализа заключается в нахождении расстояния между всей совокупностью изученного множества объектов. Результаты представляются графически в виде дендрограммы. В качестве примера на рисунке 8 представлена часть дендрограммы, полученной в итоге кластер-анализа матрицы сходства, полученной на 20 респондентах в оценке 50 рисунков, посвященных охране природы (В.Ф. Петренко, А.Г. Шмелев, А.А. Нистратов, 1978). Цифры сверху обозначают шаги объединения, цифры справа – номера анализируемых рисунков.

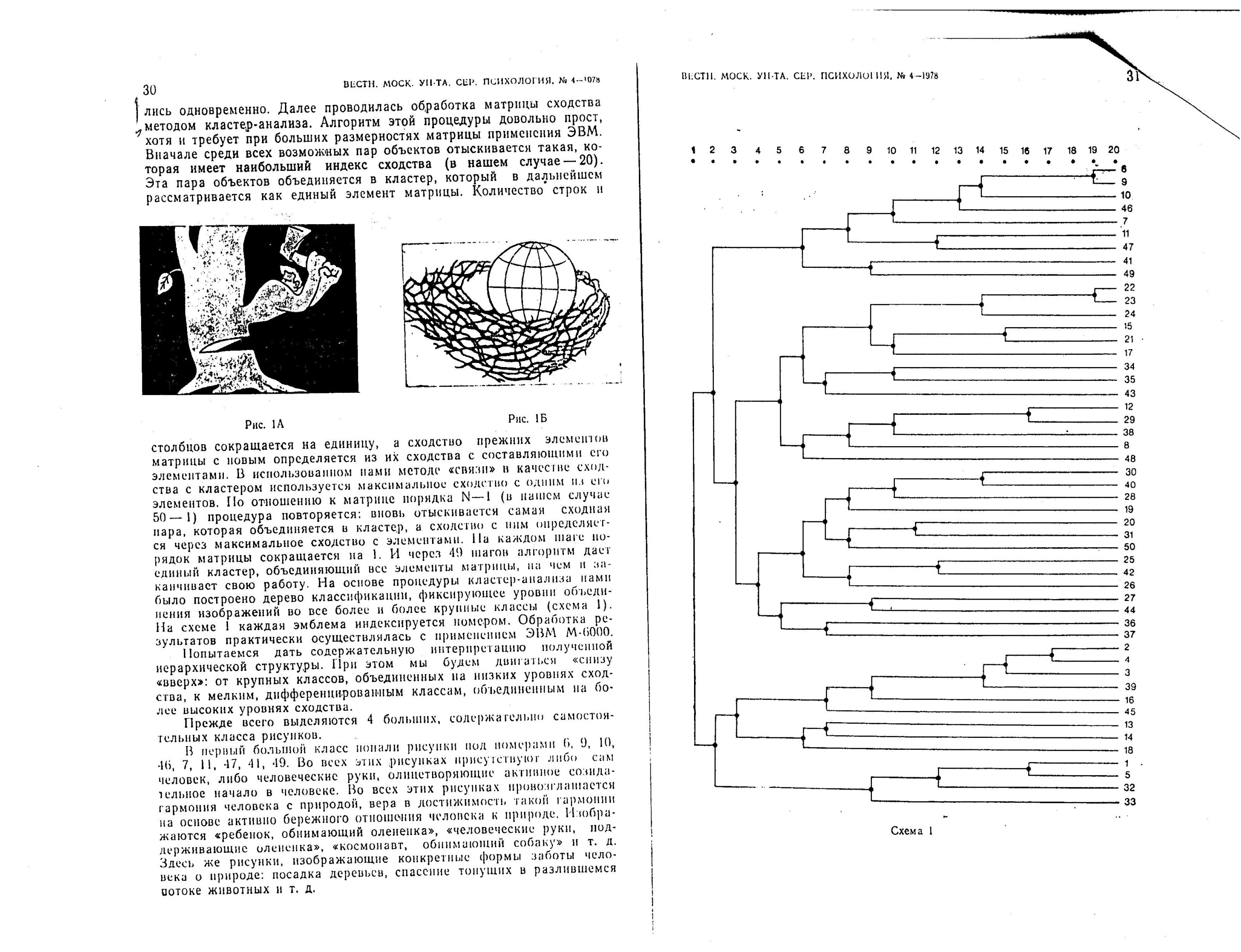


Рисунок 8 – Пример дендрограммы

Как следует из представленной денродграммы, в первый класс вошли рисунки под номерами 6, 9, 10, 46, 7, 11, 47, 41, 49; всего в данном исследовании было выявлено четыре кластера.

Кластеры обладают следующими свойствами: плотность (скопление точек в кластере по сравнению с другими); дисперсия (степень рассеивания точек в пространстве относительно центра кластера), форма (расположение точек в пространстве, круглая или удлиненная), отделимость (степень перекрытия кластеров).

Существуют различные методы кластеризации. Поскольку они довольно эвристичны, то требуется тщательно подходить к их выбору, поскольку при разных методах, применимых к одному массиву данных, можно получить совершенно различные результаты классификации. Среди методов наиболее распространены два: иерархический агломеративный (объединительный) и итеративный метод k-средних.

Объединительный метод заключается в последовательном объединении наиболее близких объектов в один кластер. Работает он с матрицей сходства объектов. Иерархическое объединение в классы может осуществляться разными способами:

– метод одиночной связи (на первом шаге объединяются два объекта, наиболее близких между собой; на втором шаге к ним присоединяется еще один объект, наиболее сходный с одним из объектов кластера и т.п.);

– метод полной связи (объединяющий в один кластер все объекты, имеющие меру сходства вышезаданного порогового значения);

– метод «средней связи» (основан на оценке среднего арифметического мер сходства объектов кластера);

– центроидный метод (расстояние между кластерами определяется как евклидово расстояние между центрами этих кластеров, на каждом шаге объединяются два кластера, имеющие минимальное расстояние) и др.

Итеративный метод k-средних предназначен для обработки самих объектов. Он позволяет получить заранее заданное исследователем число кластеров. Благодаря этому методу все множество объектов разбивается на определенное количество кластеров, максимально различающихся между собой. Конкретный объект войдет в тот класс, расстояние до которого минимально.

Кластерный анализ по сравнению с факторным предоставляет больше возможностей для сопоставления данных с различными теоретическими гипотезами. Однако в отличие от факторного анализа он не так формализован, т.е. не существует четких правил: какой именно метод кластеризации использовать для данных определенного типа, какую меру выбирать для иерархизации, на основе какой метрики вычислять меру сходства между объектами – и других однозначных предписаний. Поэтому его использование предполагает достаточно высокую квалификацию психолога.

**5. Регрессионный анализ**

Регрессионный анализ – это количественное выражение причинной связи между зависимой и независимыми переменными. Его назначение состоит в том, чтобы по значениям следствий (независимых переменных, которые называются также предикторами) спрогнозировать причину (зависимую переменную). Иначе говоря, регрессионный анализ позволяет построить модель предсказания: к примеру, зная результаты успеваемости в школе, предсказать успеваемость в университете; на основе известных показателей вербального и невербального интеллекта спрогнозировать степень успешности в определенной деятельности и т.п.

Регрессией называется изменение функции в зависимости от изменений одного или нескольких аргументов. В самом общем виде регрессионную зависимость можно представить в виде функции: Y = f (X1, X2 … Xk), где Y – зависимая переменная, а X – предикторы. Если функция f является линейной (прямая зависимость между переменными), то речь идет о линейной регрессии, если нет – о нелинейной модели (логарифмической, экпоненциальной и др.). В зависимости от вида функции выделяются и различные виды регрессионного анализа.

В психологии чаще используются линейные виды регрессионного анализа. Регрессия рассчитывается по специальным уравнениям. В простейшем случае уравнение линейной регрессии включает две переменные y = bx + a. Если в уравнении используется несколько независимых переменных, регрессия называется множественной. Уравнение множественной линейной регрессии имеет вид y = ∑ bixi + a. В уравнениях особое значение имеют коэффициенты регрессий, показывающие, насколько в среднем величина одной переменной изменяется при изменении на единицу меры другой. При линейной связи коэффициенты регрессии, аналогично коэффициентам корреляции, при положительной связи имеют знак «+», а при отрицательной, соответственно, знак «–». Поэтому, зная значения одной переменной, можно довольно однозначно предсказать значения другой, что и является ценным для психолога.

*Условия использования метода линейного регрессионного анализа:*

*–*переменные должны быть измерены в метрических шкалах (при определенных условиях возможен анализ и дихотомических данных);

*–*значения переменных должны подчиняться закону нормального распределения;

*–* в переменных должно присутствовать одинаковое число варьирующих признаков.

**6. Дискриминантный анализ**

Данный метод, разработанный Р. Фишером, является разновидностью регрессионного анализа. Он объединяет в себе группу статистических процедур, направленных на разделение (дискриминацию) многомерных совокупностей на группы. При этом достигается максимальная однородность внутри групп и минимальная между группами. Таким образом, дискриминантный анализ позволяет решать две задачи: классификации наблюдений по группам и интерпретации межгрупповых различий.

Число групп в дискриминантном анализе задается заранее и определяется целями исследования. Если надо создать две группы, то используются линейные дискриминантные функции, если три и более – квадратичные. Суть дискриминантного анализа (на примере линейных функций) заключается в следующем. Есть две выборки, у которых измерено множество переменных. И есть третья выборка многомерных объектов. Дискриминантный анализ позволяет найти правило, благодаря которому объекты из третьей выборки можно отнести либо к первой, либо ко второй (т.е. классифицировать).

*Условия применения дискриминантного анализа:*

– зависимая переменная измеряется в шкале наименований,

– независимые переменные должны быть выражены в шкалах интервалов или отношений, а их значения подчиняться закону нормального распределения;

– число переменных не ограничено, но при этом число объектов должно их превышать минимум на два;

– должно присутствовать два и более класса;

– в каждый класс должно входить минимум два объекта;

– недопустимы переменные с линейной корреляцией, равной единице (т.к. они полностью повторяют друг друга).

В психологии дискриминантный анализ широко используется в психодиагностике при разработке тестов, а также для обработки экспериментальных данных. Как разновидность регрессионного анализа дискриминантный анализ позволяет решать еще одну важную психологическую задачу, а именно: на основе независимых переменных не просто предсказать изменение зависимой переменной в целом по группе, а узнать дискретные, индивидуальные проявления этой зависимой переменной.

В завершение следует отметить, что в психологии, кроме рассмотренных выше, используются и другие виды многомерного анализа, или анализа с латентными переменными (ковариационный, структурного моделирования, медиаторный и др.). Сегодня они все осуществляются с помощью компьютерных программ, поэтому собственно математические вычисления уже не выступают проблемой. Главной задачей при применении совокупности сложных математических процедур для современных исследователей выступает собственно психологическая интерпретация полученных данных, которая требует тщательного теоретического обоснования результатов и творческого подхода к их анализу.

**ПРЕДСТАВЛЕНИЕ РЕЗУЛЬТАТОВ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ В ТЕКСТОВОЙ ФОРМЕ**

1. «Сырые» данные и их оформление.

2. Первичная обработка и способы представления ее итогов.

3. Вторичная обработка и описание ее результатов в научном тексте.

**1. «Сырые» данные и их оформление**

Под «сырыми» данными принято понимать результаты обработки индивидуальных протоколов испытуемых, которые организуются в общий массив данных по выборочной совокупности. Как правило, для компактности и удобства дальнейшей обработки эти результаты представляются в табличной форме. Такие таблицы называются сводными. В них по горизонтали принято обозначать респондентов, по вертикали – полученные каждым из них результаты по конкретным методикам. Примеры оформления таких таблиц представлены далее.

Таблица 5 – Итоги интерпретации теста «Дерево»

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Исп. | Диагностические критерии | | | | |
| Размер  рисунка | форма ствола | тип дерева | крона | листья |
| 1 | две трети  листа | морковка  (острие вниз) | лиственное | замкнутая | опавшие |
| 2 | треть листа | ровная | лиственное | открытая | есть |
| 3… | треть листа | морковка  (острие вверх) | хвойное | замкнутая | тщательно  прорисованы |
| 40 | менее трети листа | морковка  (острие вверх) | лиственное | открытая | хаотичны |

Таблица 6 – Результаты изучения эмпатии в группе программистов

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Исп. | Канал эмпатии | | | | | | Общ.  балл |
| рацио-нальный | эмоцио-  нальный | интуи-тивный | установки | проник.  способ. | иденти-фикация |
| А.В. | 4 | 2 | 4 | 3 | 3 | 2 | 18 |
| С.И. | 5 | 3 | 3 | 2 | 3 | 3 | 16 |
| К.Х. |  |  |  |  |  |  |  |
| Д.В. | 5 | 4 | 5 | 3 | 4 | 4 | 25 |

При составлении подобных сводных таблиц следует обратить внимание на следующее:

– для сохранения конфиденциальности психологической информации респонденты обязательно зашифровываются (номерами, инициалами или другим образом);

– содержанием таблиц может выступать информация качественного и количественного характера;

– возможно одновременное размещение в таблице итогов обработки несколько методик, если это не делает ее слишком громоздкой.

«Сырые» данные, даже если они структурированы в сводную таблицу, являются только вспомогательным материалом, поэтому они всегда помещаются в приложения.

**2. Первичная обработка и способы представления ее итогов**

Под первичной обработкой данных понимается упорядочивание и обобщение «сырых» данных (группировка, например, распределение по уровням выраженности, по наличию/отсутствию качественных признаков; расчет среднегрупповых показателей и т.д.). Первичная обработка является обязательным элементом исследования. В итоге ее обнаруживаются те результаты, которые можно получить в ходе довольно простых подсчетов. Говоря сленговым языком – это то, что «видно невооруженным взглядом», т.е. без применения специальных статистических мер.

Отражение в текстовой форме итогов первичной обработки включает в себя три обязательных компонента:

1) представление;

2) обсуждение;

3) интерпретация.

***Представление данных обработки***

В тексте работы результаты первичной обработки представляются в виде таблиц или рисунков (диаграмм, графиков, гистограмм и др.). Обязательным требованием является только однократное их упоминание, т.е. нельзя одни и те же данные отражать и в форме таблицы, и в форме рисунка. Выбор формы определяется самим исследователем.

На все таблицы и рисунки должны быть ссылки в работе. Эти ссылки осуществляются с помощью следующих оборотов:

– итоги классификации испытуемых по наличию этнических стереотипов отражены на диаграмме (рисунке);

– итоги распределения респондентов по уровням развития креативности содержатся в таблице;

– для большей наглядности среднегрупповые данные по шкалам семантического дифференциала отражены в виде графика (рисунка).

***Обсуждение результатов обработки***

Обсуждение предполагает своеобразный мысленный диалог с другим человеком, которому автор работы описывает полученные им статистические результаты. В ходе этого описания имеющиеся числовые данные (представленные в виде таблиц, рисунков и т.п.) не просто пересказываются словами. Очень важно эти данные прокомментировать и расставить акценты, подчеркнув именно те результаты, которые важны для решения поставленных задач. Таким образом, числовые данные должны быть проанализированы качественно, т.е. за каждым числом должно стоять именно психологическое объяснение: что это означает, о чем свидетельствует, о каких тенденциях говорит и т.п.

Стилистически ошибочным считаются обороты «Из таблицы видно…», «Из графика видно…» и т.п. Более грамотным будет использование следующих выражений:

– Данные, содержащие в таблице 1, демонстрируют…

– Результаты измерения, представленные в таблице 2, показывают…

– Анализ данных таблицы 3 по горизонтали позволят говорить о…

– На диаграмме отражено преобладание …

– Графически линия группы А расположена выше линии группы Б, что свидетельствует о ….

***Интерпретация полученных результатов***

Интерпретация полученных результатов обязательно должна включать следующие моменты.

1)Соотнесение эмпирических данных

– с поставленной гипотезой (соответствуют / не соответствуют);

– с результатами работ других авторов (совпадают / не совпадают, не имеют аналогов и т.п.).

2)Объяснение результатов исследования: почему были полученные такие результаты? Как это можно объяснить? Можно выдвинуть дополнительные гипотезы, свои предложения по усовершенствованию методики и т.п.

Обсуждение и интерпретация полученных данных являются наиболее творческими этапами исследования, требующими внимательного и вдумчивого отношения.

**3. Вторичная обработка и описание ее результатов в научном тексте**

Вторичная обработка данных состоит в выявлении скрытых от непосредственного анализа отношений. Для этого используются индуктивная статистика, корреляционный анализ, методы многомерного анализа.

Обращение к статистическим критериям требуется психологу для подтверждения обоснованности сделанных им заключений. Какой бы сложности и объема расчеты не проводились, в общем контексте исследования они выступают только вспомогательным инструментом. Поэтому существует следующее правило:

1) в основном тексте приводятся только конечные результаты расчетов, а именно полученные эмпирические значения определенных критериев и их критические значения для выбранных уровней значимости;

2) в приложениях размещаются промежуточные данные: формулировки статистических гипотез, формулы расчета критериев, проведенные по ним математические расчеты, построение оси значимости.

Результаты расчета статистических критериев могут быть представлены в текстовой или табличной форме.

В текстовой форме используются общепринятые обозначения критериев и следующие примерные выражения:

– Оценка сдвигов в двух замерах осуществлялась с помощью Т критерия Вилкоксона (Тэмп = 13 приТкр = 10 для ρ ≤ 0,05). Таким образом, значимые различия в исследованной группе отсутствуют…

– Установление различий выраженности тревожности в пяти различных ситуациях проводилось с помощью критерия тенденций Пейджа: Lэмп = 485 (при Lкр = 477 ρ ≤ 0,05 Lкр = 487 ρ ≤ 0,01). Таким образом, выдвинутая гипотеза подтверждается статистически при допустимой 5%-й вероятности ошибки…

– Для определения достоверности различий был использован t-критерий Стьюдента (расчеты представлены в приложении В). Расчеты показали, что t = 2,8 при критическом t = 2,4 для p ≤ 0,01. Это говорит о наличии статистически значимых различий в …

Если результатов статистических расчетов достаточно много, то более удобной и компактной выступает табличная форма их репрезентации. При этом в таблицах принято одновременно размещать исходные для проведения расчетов данные (количественные показатели, проценты, среднее арифметическое и др.) и полученные эмпирические значения критериев. Для отражения уровней значимости в таблице отводится отдельная колонка либо используются условные обозначения в виде «звездочек», которые ставятся как индекс рядом с полученным эмпирическим значением использованного статистического критерия: \* – р ≤ 0,05, \*\* – р ≤ 0,01, \*\*\* – р ≤ 0,001. Отсутствие звездочки означает низкий уровень достоверности или р ≥ 0,05. Пример содержится в таблице 7.

Таблица 7 – Мотивация использования Интернета

подростками разного пола (в %)

|  |  |  |  |
| --- | --- | --- | --- |
| Варианты ответов | Мальчики | Девочки | Критерий  Фишера |
| Ищу информацию по школьным предметам | 60 | 100 | 6,85\*\*\* |
| Читаю книги, нахожу интересные факты | 50 | 70 | 2,05\* |
| Играю в компьютерные игры | 70 | 36 | 3,47\*\*\* |
| Сижу в социальных сетях | 64 | 96 | 4,42\*\*\* |

Аналогичным образом (в текстовой или табличной форме) оформляются и результаты корреляционного анализа. Как правило, сегодня он осуществляется с помощью разнообразных статистических пакетов, которые также принято указывать. Например:

– В итоге корреляционного анализа (проводимого с помощью программы Excel) значимых связей выявлено не было: r = 0,24 (при критическом r = 0,4 для p ≤ 0,05);

– Корреляционный анализ осуществлялся с помощью статистического пакета SPSS (v. 13). В его итоге были установлены достоверные на уровне ρ ≤ 0,01 положительные взаимосвязи эмоционального истощения с двумя стилями детско-родительских отношений: отвержение (rs = 0,48) и инвалидизация (rs = 0,51).

Аналогично данным первичной обработки, представление в тексте полученных математико-статистических расчетов обязательно должно сопровождаться их обсуждением и интерпретацией.

На результаты математико-статистической обработки должны быть также ссылки в заключении работы. В соответствии со сложившейся традицией в заключении конкретные цифры не приводятся, а заменяются вербальными выражениями, отражающими содержательную сущность использованных методов обработки.

Таблица 8 – Формулировки результатов статистических расчетов

в заключении работы

|  |  |
| --- | --- |
| Примеры неправильной  формулировки | Образцы правильной  формулировки |
| 56% опрошенных | Фактически половина опрошенных |
| 80% испытуемых | Четверо из пяти испытуемых |
| 30% | Каждый третий респондент |
| t = 3,4 (ρ ≥ 0,01) | Имеются статистически значимые различия |
| χ² = 122, ρ ≥ 0,05 | Выявленные различия не являются существенными |
| G = 2 (ρ ≥ 0,05) | Обнаружен достоверный сдвиг значений |
| r = 0,82 | Выявлена значимая положительная корреляция между Х и У |