ПРИМЕНЕНИЕ СТАТИСТИЧЕСКИХ МЕТОДОВ В ПСИХОЛОГО-ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

Учебное пособие

Троицк 2009

Федеральное агентство по образованию
Троицкий филиал государственного образовательного учреждения
высшего профессионального образования
ЧЕЛЯБИНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ПРИМЕНЕНИЕ СТАТИСТИЧЕСКИХ МЕТОДОВ В ПСИХОЛОГО-ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

Учебное пособие

Троицк 2009

Применение статистических методов в психолого – педагогических исследованиях: Учебное пособие / Сост. С.В. Нужнова. - Троицкий филиал ГОУ ВПО «ЧелГУ».- Троицк, 2005. – 120 с.

В учебном пособии в соответствии с ГОС высшего профессионального образования раскрыты основные разделы курса «Математические основы психологии»

Данное учебное пособие предназначено для студентов, обучающихся по специальности (направлению) «Педагогика и психология».

.

Составитель: Нужнова С. В., кандидат педагогических наук, доцент кафедры математики и информатики Троицкого филиала ГОУ ВПО «ЧелГУ».

Рецензенты: Лежнева Н. В., доктор педагогических наук, профессор кафедры педагогики и психологии Троицкого филиала ГОУ ВПО «ЧелГУ»,
Валеева Н. Х., кандидат педагогических наук, преподаватель ФГОУ СПО «ТАТУК ГА».

Рекомендовано к изданию учебно – методической комиссией Троицкого филиала ГОУ ВПО «ЧелГУ»

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

При изучении дисциплины «Математические основы психологии» рассматриваются основные методы математической статистики, которые наиболее часто используются в обработке результатов психологических и педагогических исследований. Современная математическая статистика представляет собой большую и сложную систему знаний. Статистика - это область науки, имеющая дело со сбором, анализом и интерпретацией данных. Массовые социально-экономические явления и процессы измерения однородных объектов, обладающие качественной общностью, обнаруживают определенные закономерности, математическая статистика создает методы выявления этих закономерностей.

Статистические методы применяются при обработке материалов психологических исследований для того, чтобы извлечь из тех количественных данных, которые получены в экспериментах, при опросе и наблюдениях, возможно больше полезной информации. В частности, в обработке данных, получаемых, при испытаниях по психологической диагностике, это будет информация об индивидуально-психологических особенностях испытуемых. Вообще психологические исследования обычно строятся с опорой на количественные данные.

Нельзя рассчитывать на то, что каждый психолог, сделавший диагностику своей специальностью, овладеет этими знаниями. Между тем, статистика нужна психологу постоянно в его повседневной работе. Специалисты статистики разработали целый комплекс простых методов, которые совершенно доступны любому человеку, не забывшему то, что он выучил еще в средней школе.

Уместное грамотное применение этих методов позволит практику и исследователю, во всяком случае, проведя начальную обработку, получить общую картину того, что дают количественные результаты его исследований, оперативно проконтролировать ход исследований.

Следует рассмотреть типы задач, с которыми чаще всего имеет дело психолог. Соответственно существуют и статистические методы, которые приложимы для обработки психологических материалов, направленных на решение этих задач.

Первый тип задач. Психологу нужно дать сжатую и достаточно информативную характеристику психологических особенностей какой-то выборки, например, школьников определенного класса. Чтобы подойти к решению этой задачи, необходимо располагать результатами диагностических испытаний; эти испытания, разумеется, следует заранее спланировать так, чтобы они давали информацию о тех особенностях группы, которые в этом конкретном случае интересуют психолога. Это могут быть особенности умственного развития, психофизиологические особенности, данные об изменении работоспособности и т. д.

Получив все экспериментальные результаты и материалы наблюдений, следует подумать о том, как их подать пользователю в компактном виде, чтобы при этом свести к минимуму потерю информации. В перечне статистических методов, используемых при решении подобных задач, обычно находят свое место и параметрические и непараметрические методы, о возможностях применения тех и других, как было сказано выше, судят по самому полученному материалу.

Второй тип задач. Это, пожалуй, наиболее часто встречающиеся задачи в исследовательской и практической деятельности психолога: сравниваются между собой несколько выборок, чтобы установить, являются ли выборки независимыми, принадлежат ли одной и той же совокупности. Так, проведя эксперименты в восьмых классах двух различных школ, психолог сравнивает эти выборки между собой.

К этому же типу относятся задачи с определением тесноты связи двух рядов показателей, полученных на одной и той же выборке; в такой обработке чаще всего применяют метод корреляций.

Третий тип задач — это задачи, в которых обработке подлежат временные ряды, ряды, в которых расположены показатели, меняющиеся во времени; их называют также динамическими рядами. В предшествующих типах задач фактор времени не принимался во внимание и материал анализировался так, как будто он весь поступил в руки исследователя в одно и то же время. Такое допущение можно оправдать тем, что за тот короткий период времени, который был затрачен на собирание материала, он не претерпел существенных перемен. Но психологу приходится работать и с таким материалом, в котором наибольший интерес представляют как раз его изменения во времени. Допустим, психолог намерен изучить изменение работоспособности школьников в течение учебной четверти. В этом случае информативными будут показатели, по которым можно судить о динамике работоспособности. Берясь за такой материал, психолог должен понимать, что при анализе динамических рядов нет смысла пользоваться средним арифметическим ряда, так как среднее арифметическое замаскирует нужную информацию о динамике. Показатели лонгитюдного исследования — это также динамические ряды, и при их обработке следует пользоваться методами, предназначенными для таких рядов.

Четвертый тип задач — задачи, возникающие перед психологом, занимающимся конструированием диагностических методик, проверкой и обработкой результатов их применения. Психологическая диагностика, в особенности тестология, имеет целый ряд канонических правил, применение которых должно обеспечивать высокое качество информации, получаемой посредством диагностических методик. Так, методика должна быть надежной, гомогенной, валидной. По упрочившимся в тестологии правилам, все эти свойства проверяются статистическими методами.

Статистика как таковая не создает новой научной информации. Эта информация либо содержится, либо не содержится (к сожалению, и так бывает) в полученных исследователем материалах. Назначение статистики состоит в том, чтобы извлечь из этих материалов больше полезной информации. Вместе с тем, статистика показывает, что эта информация не случайна и что добытые данные имеют определенную и значимую вероятность.

Статистические методы раскрывают связи между изучаемыми явлениями. Однако необходимо твердо знать, что как бы ни была высока вероятность таких связей, они не дают права исследователю признать их причинно-следственными отношениями. Статистика, например, утверждает, что существует значимая связь между двигательной скоростью и игрой в теннис. Но отсюда еще не вытекает, будто двигательная скорость и есть причина успешной игры. Нельзя, по крайней мере в некоторых случаях, исключить и того, что сама двигательная скорость явилась следствием успешной игры.

Чтобы подтвердить или отвергнуть существование причинно-следственных отношений, исследователю зачастую приходится продумывать целые серии экспериментов. Если они будут правильно построены и проведены, то статистика поможет извлечь из результатов этих экспериментов информацию, которая необходима исследователю, чтобы либо обосновать и подтвердить свою гипотезу, либо признать ее недоказанной.

1. ОСНОВНЫЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

1.1. Статистические шкалы

Измерение - это присваивание числового значения какому-либо явлению, объекту или событию определенного числа в соответствии с определенными правилами. Например, измерение роста человека. Во-первых, возникает вопрос: что такое рост человека? Рост измеряют в сантиметрах, метрах. Но о чем может нам сказать следующее измерение: рост человека равен 180 см. Можно ли сказать, что это высокий человек? Появляется необходимость сравнения роста этого человека с ростом других людей или ростом группы людей, объединенных по какому-либо признаку (например, по национальности, по возрасту и т.д.). С. Стивенсон предложил 4 типа классификации измерений и соответственно 4 типа измерительных шкал: шкалу наименований (или номинативную), шкалу порядка, шкалу интервалов и шкалу отношений. Зная типические особенности каждой шкалы, нетрудно установить, к какой из шкал следует отнести подлежащий статистической обработке материал.

Шкала наименований. К этой шкале относятся материалы, в которых изучаемые объекты отличаются друг от друга по их качеству. При обработке таких материалов нет никакой нужды в том, чтобы располагать эти объекты в каком-то порядке, исходя из их характеристик. В принципе, объекты можно располагать в любой последовательности. Вот пример: изучается состав международной научной конференции. Среди участников есть французы, англичане, датчане, немцы и русские. Имеет ли значение порядок, в котором будут расположены участники при изучении состава конференции? Можно расположить их по алфавиту, это удобно, но ясно, что никакого принципиального значения в этом расположении нет. При переводе этих материалов на другой язык (а значит, и на другой алфавит) этот порядок будет нарушен. Можно расположить национальные группы по числу участников. Но при сравнении этого материала с материалом другой конференции найдем, что, вряд ли этот порядок окажется таким же. Отнесенные к шкале наименований объекты можно размещать в любой последовательности, в зависимости от цели исследования.

При статистической обработке такого рода материалов нужно считаться с тем, каким числом единиц представлен каждый объект. Имеются весьма эффективные статистические методы, позволяющие по этим числовым данным прийти к научно значимым выводам (например, метод хи-квадрат).

Шкала порядка. Если в шкале наименований порядок следования изучаемых объектов практически не играет никакой роли, то в шкале порядка — это видно из ее названия — именно на эту последовательность переключается все внимание. К этой шкале в статистике относят такие исследовательские материалы, в которых рассмотрению подлежат объекты, принадлежащие к одному или нескольким классам, но отличающиеся при их сравнении одного с другим — «больше—меньше», «выше—ниже» и т. п.

Проще всего показать типические особенности шкалы порядка, если обратиться к публикуемым итогам любых спортивных соревнований. В этих итогах последовательно перечисляются участники, занявшие соответственно первое, второе, третье и прочие по порядку места. Но в этой информации об итогах соревнований нередко отсутствуют или отходят на второй план сведения о фактических достижениях спортсменов, а на первый план ставятся их порядковые места. Допустим, шахматист Д. занял в соревнованиях первое место. Каковы же его достижения? Оказывается, он набрал 12 очков. Шахматист Е. занял второе место. Его достижение— 10 очков. Третье место занял

Ж. с 8-ю очками, четвертое —3. с 6-ю очками и т. д. В сообщениях о соревновании разница в достижениях при размещении шахматистов отходит на второй план, а на первом остаются их порядковые места. В том, что именно порядковому месту отводится главное значение, есть свой смысл. В самом деле, в нашем примере 3. набрал 6. а Д.— 12 очков. Это абсолютные их достижения — выигранные ими партии. Если попытаться истолковать эту разницу в достижениях чисто арифметически, то пришлось бы признать, что 3. играет вдвое хуже, чем Д. Но с этим нельзя согласиться. Обстоятельства соревнований не всегда просты, как не всегда просто и то, как провел их тот или другой участник. Поэтому, воздерживаясь от арифметической абсолютизации, ограничиваются тем, что устанавливают: шахматист 3. отстает от занявшего первое место Д. на три порядковых места.

Заметим, что в других соревнованиях расклад абсолютных достижений может быть иным: занявший первое место может всего на пол-очка опережать ближайших участников. Важно, что он набрал наибольшее количество очков. Только от этого зависит его порядковое место.

Шкала интервалов. К ней относятся такие материалы, в которых дана количественная оценка изучаемого объекта в фиксированных единицах. Вернемся к опытам, которые провел психолог с Саней. В опытах учитывалось, сколько точек может поставить, работая с максимально доступной ему скоростью, сам Саня и каждый из его сверстников. Оценочными единицами в опытах служило число точек. Подсчитав их, исследователь получил то абсолютное число точек, которое оказалось возможным поставить за отведенное время каждому участнику опытов. Главная трудность при отнесении материалов к шкале интервалов состоит в том, что нужно располагать такой единицей, которая была бы при всех повторных измерениях тождественной самой себе, т. е. одинаковой и неизменной. В примере с шахматистами (шкала порядка) такой единицы вообще не существует.

В самом деле, учитывается число партий, выигранных каждым участником соревнований. Но ясно, что партии далеко не одинаковы. Возможно, что участник соревнований, занявший четвертое место — он выиграл шесть партий, — выиграл труднейшую партию у самого лидера! Но в окончательных итогах как бы принимается, что все выигранные партии одинаковы. В действительности же этого нет. Поэтому при работе с подобными материалами уместно их оценивать в соответствии с требованиями шкалы порядка, а не шкалы интервалов. Материалы, соответствующие шкале интервалов, должны иметь единицу измерения.

Шкала отношений. К этой шкале относятся материалы, в которых учитываются не только число фиксированных единиц, как в шкале интервалов, но и отношения полученных суммарных итогов между собой. Чтобы работать с такими отношениями, нужно иметь некую абсолютную точку, от которой и ведется отсчет.

1.2. Генеральная совокупность и выборка

Сравнение это определение, установление сходства и различия, общего и особенного между анализируемыми объектами, явлениями. Для сравнения нужно провести измерение других объектов. Но при этом следует уточнить, кого или что мы измеряем: мужчин, женщин, взрослых, детей; рост, вес, внимание, память и т.п. Если мы измеряем рост мужчины, то его рост необходимо сравнивать с ростом других мужчин. При описании роста всех мужчин на Земле или в государстве они образуют совокупность, которая на языке статистики называется **генеральной совокупностью**. Общее число мужчин совокупности (N) называется объемом генеральной совокупности, а измеряемая величина, в данном случае рост, является случайной величиной и обозначается через буквы латинского алфавита, например через X, рост каждого отдельного мужчины в этом случае будет обозначаться через - $X_1, X_2, X_3, \dots, X_п$.

Насколько реально измерить всех мужчин? Чаще пользуются для сравнения не всем количеством, а какой-либо группой. Группа каких-либо явлений, событий, являющихся частью практически бесконечной генеральной совокупности, называется **выборкой**. Число показателей, составляющих эту группу, обозначается n и называется объемом выборки. Объем выборки, как правило, значительно меньше объёма генеральной совокупности.

Задача исследователя заключается в том, чтобы подобрать такую выборку, которая была бы **репрезентативной**, то есть представляла собой совокупность наиболее полно отражающую генеральную совокупность. Составить такую выборку, в точности повторяющую все разнообразные сочетания признаков, которые имеются в элементах генеральной совокупности, вряд ли возможно. Поэтому некоторые потери в информации оказываются неизбежными. Важно, чтобы были сохранены в выборке существенные, с точки зрения данного исследования, признаки генеральной совокупности. Возможны случаи, и для их обнаружения есть статистические методы, когда задачи исследования требуют создания двух выборок одной совокупности, при этом нужно установить, не взяты ли выборки из разных совокупностей. Эти и другие подобные замечания нужно иметь в виду психологу при обработке результатов выборочных исследований.

Итак, чтобы сказать, что мужчина, у которого рост 180 см, высокий или низкий, нужно измерить рост группы мужчин, проживающих на разных континентах или в странах. Если мы определяем рост мужчины - японца, то в качестве генеральной совокупности будут приняты все мужчины - японцы. Если же мы измеряем рост жителя Европы, то в качестве генеральной совокупности будут приняты все мужчины - европейцы.

1.3. Вариационный ряд

При изучении какого – либо признака выборочной совокупности осуществляют его измерение и получают отдельных значения, которые называются вариантами (от слова вариация – изменение). Для того, что бы рассмотреть и проанализировать исходные данные, их необходимо каким – либо образом представить. Основные формы представления выборки из генеральной совокупности следующие.

1. Представление выборки в несгрупированном виде, путём обычного перечисления вариант - X: $x_1, x_2, x_3, \ldots, x_п$.
2. Представление выборки в упорядоченном виде: расположение вариант либо в порядке возрастания (чаще всего) либо в порядке убывания.
3. Представление выборки в сгруппированном виде, когда вместе с вариантами указываются числа (называемые частотами) равными числу повторений данной варианты в выборке.

В широком смысле все эти формы представления исходных данных носят названия вариационного ряда. В математической статистике даётся более узкое понятие вариационного ряда.

Упорядоченный **вариационный ряд** в сгруппированном виде (или ряд статистического распределения) это расположенные в порядке возрастания варианты с соответствующими значениями частот или относительных частот.

Вариационный ряд (или ряд статистического распределения) представляет собой таблицу, содержащую элементы двух видов: варианты и частоты (или относительные частоты).

X_i	x_1	x_2	…	$x_п$
m_i	m_1	m_2	…	m_3

или

X_i	x_1	x_2	…	$x_п$
n_i	n_1	n_2	…	n_3

Частоты (m_i)- это количественные значения одинаковых вариант или групп вариант, показывающие как часто встречаются те или иные варианты в ряду распределения. Сумма всех частот определяет численность всей совокупности, ее объем.

Относительными частотами ($n_i = m_i/n$) называют частоты, выраженные в долях единицы или в процентах к общему числу. Соответственно сумма всех относительных частот равно 1 или 100%.

Например, по результатам теста, измеряющим способности к абстрактному мышлению, 20 испытуемых получили следующие оценки в баллах. Первичная запись вариант представлена в таблице 1.3.1.

Таблица 1.3.1

№	Ф. И. О	Оценка	№	Ф. И. О	Оценка
1	Андреев Т.	28	11	Максимова З.	35
2	Белоус Ф.	40	12	Напольских С.	29
3	Бережков Д.	44	13	Николаев Т.	36
4	Блинов К.	19	14	Петров И.	39
5	Ежов В.	32	15	Самарин Т.	38
6	Зайцева Т.	39	16	Соловьева Н.	37
7	Иванов К.	26	17	Труничева Е.	33
8	Капралова В.	42	18	Уфимцев Я.	44
9	Кораблев П.	47	19	Царицына И.	39
10	Любимов С.	32	20	Щукина Л.	21

Рассматривая этот вариационный ряд, можно заметить, что есть крайние значения: обозначим их как X_{min}- минимальное, X_{max}- максимальное. Минимальное значение: $X_{min} = 19$, максимальное: $X_{max} = 48$ Разница между максимальным и минимальным значениями называется **размахом** (N) вариационного ряда. В примере размах вариационного ряда равен N = 29. Сравнение этих показателей нескольких вариационных рядов дает нам приближенную информацию о близости (отдаленности) инвариант в группах. Чем меньше значение размаха, тем более близкие значения инвариант в ряду, тем более похожи между собой испытуемые по измеряемому признаку.

Если в вариационном ряду значение максимальной варианты сильно отличается от варианты, наиболее близкой к максимальной варианте, или значение минимальной варианты чрезмерно мало по сравнению с другой вариантой, приближающейся к минимальной, то необходимо проверить принадлежность таких вариант к данной выборке. Такие резко отклоняющиеся варианты возникают из-за грубой ошибки при измерении или подсчете, или относятся к другой популяции. В подобных случаях резко отклоняющиеся варианты, чуждые данной совокупности, необходимо исключить из выборки. Однако это исключение должно быть обосновано посредством количественной оценки отклонения таких вариант.

Для проверки максимальной варианты на «выскакивание» применяется формула:

$$\Theta = \frac{X_n - X_{n-1}}{X_n - X_2}$$

где X_n - максимальная варианта;

X_{n-1} - варианта, следующая по величине за максимальной;

X_2 - варианта, предшествующая по величине минимальной варианте;

Θ - критерий принадлежности максимальной варианты к совокупности.

Например, в вариационном ряду: 82, 77, 74, 73, 66, 64, 63, 62, 54, 44, 34 - следует проверить варианты 82 и 34. Проверяем значение 82:

$$\Theta = \frac{X_n - X_{n-1}}{X_n - X_2} = \frac{82-77}{82-44} = 0,132$$

Полученное значение критерия 0,132 сравниваем с табличным (приложение 4). Для n= 11 табличное Θ = 0,450. Если полученное значение критерия меньше табличного, то проверяемую варианту нельзя исключать из выборки.

Для проверки **минимальной варианты** на «выскакивание» применяется формула

$$\Theta = \frac{X_2 - X_1}{X_{n-1} - X_1}$$

где Θ - критерии принадлежности минимальной варианты: X_1 - минимальная варианта; X_2 - варианта, стоящая рядом с минимальной;

X_{n-1} - вторая по значению варианта после максимальной.

Проверим варианту 34:

$$\Theta = \frac{44-34}{77-34} = 0,203$$

Полученное значение Θ = 0,203 меньше табличного, поэтому отбрасывать данную варианту нельзя. В случае если резко отклоняются по величине не только крайние варианты, но и соседние с ними, принадлежность их к данной выборке оценивается по тем же формулам.

Примерами вариационных рядов могут быть:

- интервальный ряд, который используется для записи результатов измерения вариационного признака являющегося непрерывной случайной величиной, например рост учащихся:

Таблица 1.3.2

Рост (x)	140-143	143-146	146-149	149-152	152-155	155-158	158-161
m	16	53	121	193	229	186	121

- часто в психолого-педагогических исследованиях используется дихотомический вариационный ряд:

Таблица 1.3.3

Уровни распределения измеряемой величины	Группы		
	1	2	3
Низкий (0-6 баллов)	10	12	9
Средний (7-13 баллов)	5	8	11
Высокий (14 – 20 баллов)	5	4	7

Выбор вида вариационного ряда в каждом конкретном эксперименте осуществляется самим исследователем в зависимости от задач исследования, объема выборки и типа измеряемой величины.

1.4. Способы группировки данных

Удобно ли искать минимальное и максимальное значения в таблице 1.3.1? Эти значения находятся в середине таблицы, расположение оценок выбрано в соответствии фамилий, расположенных в алфавитном порядке. Однако, такой способ расположения вариант не удобен для анализа. Поэтому чаще всего используется другой способ - упорядочивание вариант по убывающей или **возрастающей** величине.

Таблица 1.4.1

№	ФИО	Оценка	№	ФИО	Оценка
1	Кораблев П.	47	11	Николаев Т.	36
2	Бережков Д.	44	12	Максимова З.	35
3	Уфимцев Я.	44	13	Труничева Е.	33
4	Капралова В.	42	14	Ежов В.	32
5	Белоус Ф.	40	15	Любимов С.	32
6	Зайцева Т.	39	16	Напольских С.	29
7	Петров И.	39	17	Андреев Т.	28
8	Царицына И.	39	18	Иванов К.	26
9	Самарин Т.	38	19	Щукина Л.	21
10	Соловьева Н.	37	20	Блинов К.	19

В таблице 1.4.1 приведены данные в убывающем порядке. Теперь легко определить, какие значения являются крайними, а также какое место в группе занимает каждый испытуемый: это его порядковый номер.

Ранжирование. Если группа большая, то будут возникать ситуации, когда несколько человек получают одинаковые оценки, результаты. В таком случае нельзя сказать, что один занимает положение выше другого, им необходимо присвоить одинаковый ранг или место. При этом сумма мест должна соответствовать количеству вариант. При присваивании рангов вариантам (это и есть ранжирование) пользуются следующими правилами ранжирования:

1. Равным значениям присваивается одинаковый ранг, являющийся средним арифметическим местом в числе одинаковых значений. Например, в табл. 1.4.1. испытуемые Бережков и Уфимцев имеют по 44 балла, занимая 2 и 3 место по порядку. Их ранг равен 2,5.

2. Последний ранг должен соответствовать общему числу вариант (n).

Пример ранжирования результатов приведен в таблице 1.4.2.

Таблица 1.4.2.

ФИО	Оценка	Место	Ранг	ФИО	Оценка	Место	Ранг
Кораблев П.	47	1	1	Николаев Т.	36	11	11
Бережков Д.	44	2	2,5	Максимова З.	35	12	12
Уфимцев Я.	44	3	2,5	Труничева Е.	33	13	13
Капралова В.	42	4	4	Ежов В.	32	14	14,5
Белоус Ф.	40	5	5	Любимов С.	32	15	14,5
Зайцева Т.	39	6	7	Напольских С.	29	16	16
Петров И.	39	7	7	Андреев Т.	28	17	17
Царицына И.	39	8	7	Иванов К.	26	18	18
Самарин Т.	38	9	9	Щукина Л.	21	19	19
Соловьева Н.	37	10	10	Блинов К.	19	20	20

Табулирование. При большом количестве вариант (n больше 30) возможна группировка значений без повторной записи одинаковых значений. При этом дополнительно записывается число, указывающее количество одинаковых вариант.

Таблица 1.4.3

Оценка (X)	m_i	Оценка (X)	m_i	Сумма
47	1	35	1	n = 20
44	2	33	1	
42	1	32	2	
40	1	29	1	
39	3	28	1	
38	1	26	1	
37	1	21	1	
36	1	19	1	

Это число, как было определено выше, называется частотой, и обозначается m_i. Сумма частот равняется количеству вариант (n). В таблице 1.4.3. приводится пример табулирования значений из таблицы 1.4.2.

Если значение вариант упорядочить в порядке возрастания, то получим статистическое распределение.

Если выборка очень велика (несколько десятков или сотен), то для обобщения результатов необходимо несколько этапов группировки данных. В такой ситуации, как правило, существует настолько широкий диапазон оценок, что целесообразнее сгруппировать их по величинам. Каждая группа вариант, на которые разбивается выборка, называется разрядом оценок (или классом). Для получения достаточно точных результатов группировки число интервалов рекомендуется выбирать по таблице 1.4.5.

Таблица 1.4.5

Число разрядов (классов) в зависимости от объема выборки

Объем	Число	Объем	Число
30-50	2-6	100-200	9-12
60-100	7-10	200 и более	12-17

Пример. В таблице 1.4.6 приведены данные коэффициента интеллекта 70 взрослых людей.

Таблица 1.4.6

141	104	101	130	148_{max}	92	87
115	91	96	100	133	124	92
123	132	118	98	101	107	97
124	118	146	107	110	11	138
121	129	106	135	97	108	108
107	110	101	129	105	105	110
116	113	123	105	110	116	113
123	83_{min}	127	112	114	105	127
114	113	106	139	95	105	95
105	106	109	102	102	102	89

Нужно построить вариационный ряд по сгруппированным частотам. Порядок действий при построении распределения сгруппированных частот будет следующим:

1. Пользуясь таблицей 1.4.5, выбирается число классов, их будет - 7.

2. Определяется величина классового интервала (или, иначе, ширина класса), которая зависит от принятого числа классов и от объема выборки:

$$X_{max} – X_{min} +1 = 148 - 83 + 1 = 66$$

3. Определение интервала разрядов: 66: 7 = 9,43 - округлим до 10. Таким образом, интервал разрядов определяется по формуле

$$C = \frac{X_{\max} - X_{\min}}{k-1}$$

где X_{max} - максимальная варианта выборки;

X_{min} - минимальная варианта выборки;

k - число классов, которое выбирается по таблице 1.4.5.

4. Расчет границ классов. Для этого рекомендуется начинать табулирование с величины, кратной разрядному интервалу. Начальное число должно обязательно включать минимальное значение величины варианты. Произвольное начало отсчета границ классов следует вы-

бирать так, чтобы крайние варианты по возможности оказались примерно в середине своих классов. Пример табулирования значений коэффициентов интеллекта взрослых испытуемых приведен в таблице 1.4.7.

Таблица 1.4.7

№	Граница класса	Середина класса	Подсчет вариант	Частота m_i
1	77-87	82	II	2
2	88-98	93	IIII.IIIII	10
3	99-109	104	IIIII.IIIII.IIIII.IIIII.III	23
4	110-120	115	IIIII.IIIII.IIIII.I	16
5	121-131	126	IIIII.IIII.I	11
6	132-142	137	IIIII.I	6
7	143-153	148	II	2
	C =11	K = 7		n = 70

Число вариант, относящихся к одному классу, будет также называться **частотой:** Так, для класса № 1 данного примера частота m_1 = 2, для класса № 2 частота m_2=10 и т.д.

Как уже говорилось выше частоты можно выразить в относительных величинах ($n_i = m_i/n$): долях единицы или в процентах от общего числа вариант. Такой способ обобщения результатов позволяет проводить сравнение большого числа фактов и разделять эти явления по группам или классам.

Способы графического изображения данных

Графическое изображение данных позволяет более наглядно изучать, анализировать то или иное явление. Распространенными способами изображения данных являются **гистограмма и полигон.**

Гистограмма - это последовательность столбцов, каждый из которых опирается на один разрядный интервал, откладываемый на оси абсцисс, а высота его отражает число (процент) случаев или частоту в этом разряде, которая откладывается по оси ординат. Гистограмма строится на основе табулирования.

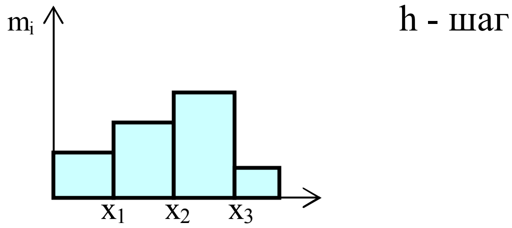

Рисунок 1. Графическое изображение гистограммы

Полигон распределения во многом напоминает гистограмму. При построении полигона распределения на оси абсцисс также откладываются значения признака, но только отмечаются не границы разрядов, а их середины; по оси ординат - частоты. Против середины каждого разряда на уровне, соответствующем частоте разряда, ставятся точки, которые затем соединяются прямыми линиями. Для завершения фигуры слева и справа откладываются две середины разрядов с нулевой частотой. Эти середины разрядов соединяются прямыми линиями с точками соседних разрядов.

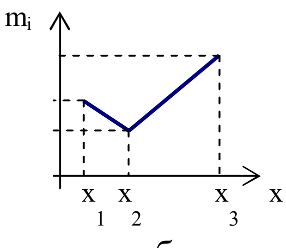

Рисунок 2. Графическое изображение полигона

Так же для представления результатов психолого – педагогических исследований широко используют различные виды диаграмм: точечные, круговые, лепестковые и т.д., которые легко строятся в программе EXCEL при обработке данных на компьютере.

2. ОСНОВНЫЕ ХАРАКТЕРИСТИКИ СТАТИСТИЧЕСКОГО РАСПРЕДЕЛЕНИЯ

2.1. Характеристики, отражающие центральную тенденцию выборки

Центральную тенденцию выборки позволяют оценить такие статистические характеристики, как мода, медиана и среднее арифметическое значение.

Мода (M_0) – это значение варианты имеющей наибольшую частоту появления. Другими словами это значение варьирующего признака, встречающееся наиболее часто. Это наиболее простая мера центральной тенденции. В случае, когда все значения в группе встречаются одинаково часто, принято считать, что группа не имеет моды. Если несколько вариант имеют наибольшие одинаковые значения частоты, то за моду принимают их среднее арифметическое.

Например. При измерении некоторого показателя X, получили следующие значения: 3; 3; 4; 5; 5; 5; 6; 6.

Мода этого ряда будет равна числу 5.

При измерении некоторого показателя X, получили следующие значения: 3; 3; 4; 5; 5; 5; 6; 6; 6.

Мода этого ряда будет равна числу (5 + 6) / 2 = 5, 5.

Медиана (M_e) - это значение варианты, которое делит упорядоченное множество вариант пополам, так, что одна половина значений оказывается больше медианы, а другая - меньше. Это значение делит всю выборку на две равные части.

Если выборка содержит нечетное число различных вариант, например, 11, 14, 18, 20, 21, то $M_e = 18$.

Если выборка содержит четное число вариант, например, 12; 14; 15; 16; 17; 18 то медиана есть точка, лежащая посредине между двумя центральными значениями, когда они упорядочены:

$$M_e = (15 + 16)/2 = 15{,}5$$

Среднее арифметическое значений вариант - это среднее совокупности n значений. Обозначается через $\overline{X}$ (или М) и определяется как

$\overline{X} = \dfrac{X_1 + X_2 + \ldots + X_n}{n} = \dfrac{\sum_{i=1}^{n} X_i}{n}$, для без частотного вариационного ряда и

$\overline{X} = \dfrac{\sum_{i=1}^{n} X_i\ m_i}{n}$, если вариационный ряд записан по сгруппированным частотам.

Замечание: если ко всем значениям варианты прибавить некоторое число, то средняя выборки увеличится на это же число. Если все значения варианты умножить на некоторое число, то ее среднее значение также увеличится на это же число.

2.2. Меры изменчивости

На практике часто интересует вопрос, как сильно каждый результат (значение варианты) отклоняется от среднего значения. Легко можно представить, что две группы результатов измерений имеют одинаковые средние, но различные значения измерений**.**

Например, два ряда: первый - 2, 8, 2 - среднее значение $\overline{X}$ = 4; второй ряд - 5, 2, 5 - среднее значение также $\overline{X}$ = 4. Различия же между этими рядами существенные. Поэтому меры центральной тенденции всегда необходимо дополнять показателями вариации.

Самой простой характеристикой изменчивости ряда является размах варьирования. Однако он улавливает только крайние отклонения, но не отражает отклонений всех результатов.

Среднее отклонение ***(d)*** - это отклонение каждого значения от среднего, вычисляется как X_1 - $\bar{X}$. Совокупность всех n отклонений характеризует изменчивость в исходных данных. Сумма положительных и отрицательных отклонений в группе данных всегда равна нулю.

Если рассматривать отклонения от $\bar{X}$ без учета знака, то сумма этих расстояний будет характеризовать изменчивость данных

$$\bar{d} = \frac{\sum_{i=1}^{n} | X_i - \bar{X} |}{n}$$

Среднее отклонение используется как показатель индивидуальных особенностей в некоторых психофизиологических методиках.

Дисперсия (D или σ^2) - это одна из основных мер варьирования признака в статистической совокупности, величина колебания вариант около их средней арифметической. Если при расчете среднего отклонения используется абсолютная величина отклонения варианты от среднего, то при расчете дисперсии отклонения от средней перед их усреднением возводятся в квадрат, благодаря чему все отклонения становятся положительными, а также получают большую величину, что повышает их значимость и, следовательно, определяет большую чуткость при измерении варьирования признака.

Если число измерений не больше 30, то

$$D = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}$$ для без частотного ряда.

$$D = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2 \, m_i}{n-1}$$ для вариационного ряда с заданной частотой

Если выборка больше 30, то

$$D = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}$$ для без частотного ряда и

$$D = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2 \ m_i}{n}$$ для вариационного ряда с заданной частотой.

В формулах по вычислению дисперсии в зависимости от числа вариант получается много слагаемых, поэтому удобнее пользоваться для её подсчета таблицами. Например, вычисление дисперсии для без частотного ряда представлено в таблице 2.2.1.

Таблица 2.2.1

№	X_i	$(X_i - \overline{X})$	$(X_i - \overline{X})^2$
1	46	-7.16	51.26
2	50	-3.16	9.28
3	59	5.84	34.10
4	60	6.84	46.78
5	55	1.84	3.38
6	49	-4.16	17.30
	$\sum = 319$		$\sum = 164.10$

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{319}{6} = 53{,}16, \text{ а } D = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1} = \frac{164{,}10}{5} = 32{,}82$$

Среднее квадратическое отклонение (σ - сигма) показывает, насколько часто отклоняются индивидуальные значения среднего, и измеряется в тех самых единицах, что и среднее арифметическое. По данному критерию мы судим о плотности выборки: чем больше σ, тем меньше плотность результатов выборки. Чем ближе инварианты к $\overline{X}$, чем плотнее выборка, тем выше ее однородность.

Среднее квадратическое отклонение вычисляется по формуле: $\sigma = \sqrt{D}$

Например: $\sigma = \sqrt{D} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}}$ для без частотного ряда, когда объём выборки меньше 30.

2. 3. Коэффициент вариации

Можно ли сравнивать выборки по однородности, если изучаемые показатели имеют разные единицы **измерения?** Например, в одной группе измерялся рост, а в другой - вес. Соответственно, $\overline{X}_1$ = 184 см σ_1 =6,36 см; $\overline{X}_2$ = 89,9 кг и σ_2 = 4,44 кг. Сравнить эти выборки можно только в относительных единицах измерения. Это позволяет сделать коэффициент вариации (V), который определяется по формуле

$$V = \frac{100 \cdot \sigma}{\overline{X}}$$

В зависимости от величины V принято считать, что:

1. выборка однородна, если V находится в пределах от 0 до 10%
2. средняя степень однородности - V от 11 до 20%;
3. низкая степень однородности - V-от 21% и больше.

Ошибка средней арифметической (m или s) характеризует «колеблемость» средней и вычисляется по формуле:

$$m = \frac{\sigma}{\sqrt{n}}$$

Замечание 1. Чем больше объем выборки, тем меньше будет вариация средних величин. Если, например, мы возьмем две выборки: первая - 10 человек, а вторая - 1000, то средние результаты во втором случае, вероятнее всего, будут ближе друг к другу (и одновременно ближе к средней генеральной совокупности).

Замечание 2. Все показатели, рассмотренные в данном параграфе, называются первичной статистикой и являются обязательными при описании и анализе, сравнении изучаемых явлений.

2.4. Пример полной первичной статистической обработки экспериментальных данных

В эксперименте участвовало 50 испытуемых. Каждый из них выполнил по 25 проб, по 1 минуте каждая. Вычислена средняя каждого испытуемого. Полученный ряд упорядочен и все индивидуальные результаты представлены в последовательности от меньшего к большему: 85 — 93 — 93 — 99 — 101 — 105 — 109 — 110 —111 — 115 — 115 — 116 — 116 — 117 — 117 — 117 — 118 — 119 — 121 — 121 — 122 — 124 — 124 — 124 — 124 — 125 — 125 — 125 — 127 — 127 — 127 — 127 — 127 — 128 — 130 — 131 — 132 — 132 — 133 — 134 — 134 — 135 — 138 — 138 — 140 — 143 — 144 — 146 — 150 — 158

Мода данного вариационного ряда $M_0 = 127$, как наиболее часто встречающееся значение.

Медиана данного вариационного ряда $Me = (124 + 125) / 2 = 124{,}5$

Для дальнейшей обработки удобнее эти первичные данные представить в виде вариационного ряда распределённого по сгруппированным частотам. Благодаря группировке отчетливее выступает присущее данному ряду распределение величин по их частотам. Отчасти упрощается и вычисление основных числовых характеристик вариационного ряда.

Для получения достаточно точных результатов группировки число интервалов выбирается по таблице 1.4.5. Объем данной выборки N = 50. Рекомендуется выбрать 6 интервалов, однако так как $X_{min} = 85$, максимальное: $X_{max} = 158$, то размах варьирования достаточно большой ($N = X_{max} - X_{min} = 73$), нам больше подойдёт если возьмём большее число интервалов, например 9. Для облегчения вычислений составим сводную таблицу 2.4.1.

Таблица 2.4.1

Границы интерва-лов	x_i	Результат раз-носки	m_i	$m_i\ \bar{x}$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$m_i\ (x_i - \bar{x})^2$
1	2	3	4	5	6	7	8
83 - 91	87	I	1	87	36	1296	1296
92 - 100	96	III	3	288	27	729	2187
101 - 109	105	III	3	315	18	324	972
110 - 118	114	III III III I	10	1140	9	81	810
119 - 127	123	III III III III IIII	16	1968	0	0	0
128 - 136	132	III III III I	9	1188	9	81	729
137 - 145	141	IIII II	5	705	18	321	1620
146 - 154	150	III III III I	9	300	27	729	1458
155 - 163	159	I	1	159	36	1296	1296
		$\sum$	50	6150			10368

Таблица 2.4.1 состоит из 8 столбцов:

1-й столбец — группы, полученные после разбиения изучаемого ряда.

2-й столбец — средние значения каждой группы; этот столбец показывает, в каком диапазоне варьируют величины изучаемого ряда, т. е. иксы.

3-й столбец показывает результаты «ручной» разноски величин ряда или иксов: каждая величина занесена в соответствующую ее значению группу в виде черточки.

4-й столбец — это итог подсчета результатов разноски, значение частот.

5-й столбец показывает, сколько раз встречалась каждая величина ряда — это произведение величин второго столбца на величины 4-го столбца по строчкам. Итоги 4-го и 5-го столбцов дают суммы, необходимые для вычисления среднего арифметического

6-й столбец показывает разность среднего арифметического и среднего значения (т. е. значения x) по каждой группе.

7-й столбец — квадрат этих разностей

8-й столбец показывает, сколько раз встречался каждый квадрат разности, суммирование величин этого столбца дает итог, необходимый для вычисления среднеквадратического отклонения

В заголовках 5-го и 8-го столбцов указывается, насколько часто встречается та или другая величина.

Используем формулы:

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

Остается показать, как вычисляются по формулам среднее арифметическое и среднее квадратическое отклонения. Обратимся к величинам, полученным в таблице:

$$\bar{x} = \frac{6150}{50} = 123$$

При составлении таблицы это число было заранее вычислено, без него нельзя было бы получить числовые значения 6, 7, 8 столбцов таблицы.

$$\sigma = \sqrt{\frac{10368}{50}} = \sqrt{207{,}3} = 14{,}4$$

Коэффициент вариации: $V = \frac{100 \cdot \sigma}{\bar{x}}$, $\qquad V = \frac{100 \cdot 14{,}4}{123} = 11{,}7$

Выполнив все эти вычисления, психолог может представить информацию об изучении двигательной скорости с помощью примененной методики в 6-х классах. Согласно результатам изучения в 6-х классах двигательной скорости получены следующие материалы первичной статистической обработки: размах вариационного ряда N = 73, мода M_0 = 127, медиана данного вариационного ряда Me = 124,5, среднее арифметическое — 123; среднее квадратическое отклонение—14,4; коэффициент вариации — 11,7.

3. ПАРАМЕТРИЧЕСКИЕ И НЕПАРАМЕТРИЧЕСКИЕ МЕТОДЫ СТАТИСТИКИ

3.1. закон нормального распределения

Приступая к статистической обработке своих исследований, психолог должен решить, какие методы ему более подходят по особенностям его материала — параметрические или непараметрические. При изучении некоторых варьирующих признаков выборок (имеющих достаточно большой объём) совокупностей, при составлении рядов распределения по сгруппированным данным рекомендуется представить это распределение в виде диаграммы. На диаграмме изображается полигон распределения. Контуры этого полигона помогут решить вопрос о статистических методах обработки. Нередко эти контуры напоминают контуры колокола, с наивысшей точкой в центре полигона и с симметричными ветвями, отходящими в ту и другую сторону. Такой контур соответствует кривой нормального распределения. Это понятие было введено в математическую статистику К.Ф. Гауссом (1777—1855), поэтому кривую именуют также кривой Гаусса. Он же дал математическое описание этой кривой.

Варьируемая величина называется **нормально распределенной**, если ее плотность вероятности ***f(x)*** определяется по формуле:

$$f(x) = \frac{1}{\sqrt{2\pi}} \quad \lambda^{-\frac{(x-a)^2}{2\sigma^2}}$$

$$a = \overline{X} = M \text{ (математическое ожидание) } \sigma = \sqrt{D}$$

Плотностью вероятности распределения случайной величины (принимающей значения варьирующего признака) в математической статистике называют дифференциальную функцию распределения величины по вероятности, которая равна пределу частоты появления значений при объёме выборки (n) стремящемуся к бесконечности.

$$f(x) \rightarrow \frac{m_i}{n}$$

Кривая нормального распределения - особый вид частного распределения, используемого для характеристики распределений большого числа переменных, которые отражают явления, имеющие биологическую, социальную, психологическую природу. На ее основе создаются критерии нормы.

Основные свойства кривой нормального распределения:

1. Плотность распределения f*(x)* зависит от параметров $a = \overline{X}$ (или М) и σ.

2. Кривая плотности распределения симметрична относительно прямой, перпендикулярной оси абсцисс, и проходящей через точку $a = \overline{X}$ (или М) – среднее значение случайной величины равной значениям варьируемого признака.

3. Максимальное значение функции достигает при значении равном среднему значению величины ($\overline{X}$) или математическому ожиданию (М).

4. Кривая плотности распределения имеет две точки перегиба, абсциссами которых являются значения М ± σ.

Если всю площадь под кривой принять за 100%, т.е. все возможные результаты измерений, то теперь можно предсказать, какой процент от всей выборки будут составлять те или иные показатели (рисунок 3.1.1).

Для оценки варьирования результатов измерений используют следующие соотношения: М $\pm\sigma$ - интервал включает 68% всех результатов; М ± 2σ включает 95,45%; интервал М ±3σ включает 97,78 % результатов.

Частота случаев, попадающих в интервалы, измеряемые сигмами, показаны на рисунке 3.1.1.

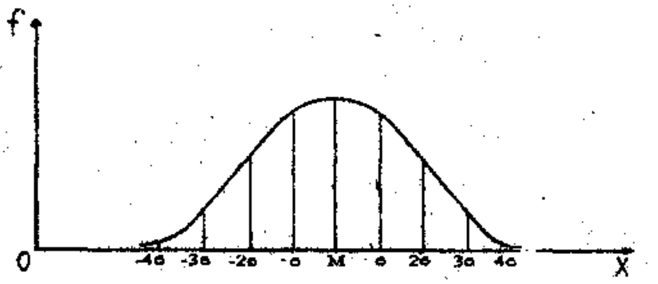

Рисунок 3.1.1.

В самом общем виде интервал на оси распределения вариант, определяемый точками М $\pm\sigma$, принято считать зоной средних значений или статистической нормой. Интервал от (М $+1\sigma$) до 2σ принято считать зоной выше или больше средних значений. И интервал от (М$+2\sigma$) и более (до 3σ) принято считать зоной высоких или больших значений по сравнению с М. Аналогично можно рассмотреть зоны, лежащие левее точки М. Интервал от (М – σ) до - 2 σ принято считать зоной ниже или меньше среднего значений вариационного ряда. И интервал от – 2σ до – 3σ принято считать зоной низких или малых значений по сравнению с М.

Итак, располагая такими показателями, как М и σ, можно разделить весь вариационный ряд на 3 подгруппы и говорить об особенностях того или иного испытуемого данной группы по измеряемому признаку.

При стандартизации тестовых оценок используется нормальное распределение со следующими характеристиками: М = 0, σ = 1. Площадь под такой кривой равна 1. Это распределение называется **стандартным (единичным) нормальным распределением**.

Для построения кривой Гаусса, (или кривой нормального распределения) теоретически требуется бесчисленное количество случаев. Практически же приходится довольствоваться тем фактическим материалом, который накоплен в исследовании. Если данные, которыми располагает исследователь, при их внимательном рассмотрении или после переноса их на диаграмму, лишь в незначительной степени расходятся с кривой нормального распределения, то это дает право исследователю применять в статистической обработке параметрические методы, исходные положения которых основываются на нормальной кривой распределения Гаусса. В математической статистике существуют более точные способы установления нормальности распределении, которые будут изложены ниже.

Нормальное распределение называют параметрическим потому, что для построения и анализа кривой Гаусса достаточно иметь всего два параметра: среднее арифметическое, значение которого должно соответствовать высоте перпендикуляра, восстановленного в центре кривой, и так называемое среднее квадратическое или стандартное отклонение — величины, характеризующей размах колебаний данной кривой.

Параметрические методы обладают для исследователя многими преимуществами, но нельзя забывать о том, что применение их правомерно только тогда, когда обрабатываемые данные показывают распределение, лишь несущественно отличающееся от Гауссова.

При невозможности применить параметрические, надлежит обратиться к непараметрическим методам. Эти методы успешно разрабатывались в последние 3— 4 десятилетия, и их разработка была вызвана, прежде всего, потребностями ряда наук, в частности, психологии. Они показали свою высокую эффективность. Вместе с тем, они не требуют сложной вычислительной работы.

3.2. Проверка нормальности распределения результативности признака по асимметрии и эксцессу

Нормальность распределения признака можно проверить путем расчета показателей асимметрии и эксцесса и сопоставления их с критическими значениями.

Асимметрия – это величина, определяемая по формуле

$$A = \frac{\sum (x_i - \bar{x})^3}{n\,\sigma^3}$$ (для без частотного ряда) **и**

для ряда с частотой $A = \frac{\sum (x_i - \bar{x})^3 m_i}{n\,\sigma^3}$

Эксцесс соответственно: $E = \frac{\sum (x_i - \bar{x})^4}{n\,\sigma^4}$ -3 и

$$E = \frac{\sum (x_i - \bar{x})^4 m_i}{n\,\sigma^4} - 3.$$

По гипотезе М.А. Плохинского, если показатели асимметрии и эксцесса превышают по абсолютной величине свою ошибку репрезентативности в 3 раза и более, то можно говорить о достоверном отличии эмпирических распределений от нормального

$$t_A = \frac{|A|}{m} \geq 3, \quad \text{где} \quad m_A = \sqrt{\frac{6}{n}} \qquad t_E = \frac{|E|}{m} \geq 3, \quad \text{где} \quad m_E = 2\sqrt{\frac{6}{n}}$$

Например, пусть некоторый психологический признак совокупности имеет следующие 16 значений: 11, 13, 12, 9, 10, 11, 8, 10. 15, 14, 8, 7, 10, 10, 5, 8. Нужно проверить имеет ли этот признак нормальное распределение или нет?

Для расчётов ассиметрии и эксцесса заполним вспомогательную таблицу 3.2.1.

Таблица 3.2.1.

№	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$	$(x_i - \bar{x})^4$
1	11	1	1	1	1
2	13	3	9	27	81
3	12	2	4	8	16
4	9	-1	1	-1	1
5	10	0	0	0	0
6	11	1	1	1	1
7	8	-2	4	-8	16
8	10	0	0	0	0
9	15	5	25	125	625
10	14	4	16	64	256
11	8	-2	4	-8	16
12	7	-3	9	-27	81
13	10	0	0	0	0
14	10	0	0	0	0
15	5	-5	25	-125	625
16	8	-2	4	-8	16
$\sum$	161	1	103	49	1735

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

$D = \frac{103}{15} = 6{,}9$; $\sigma = \sqrt{6{,}9} = 2{,}6$; $A = \frac{49}{275{,}3} = 0{,}17$

$E = \frac{1735}{16 * 46} - 3 = \frac{1735}{736} - 3 = 2{,}4 - 3 = -0{,}65$

$t_A = \frac{|0{,}17|}{0{,}61} = 0{,}26 < 3$, $m_A = 0{,}61$ $t_E = \frac{|0{,}65|}{1{,}21} = 0{,}54 < 3$, $m_E = 1{,}21$

Таким образом получили, что показатели асимметрии и эксцесса не превышают по абсолютной величине свою ошибку репрезентативности в 3 раза, следовательно можно говорить о том, что данное распределение сколь угодно близко к нормальному.

Иногда для проверки распределения на нормальность используют критерий Е.И. Пустыльника:

- если выполняются неравенства $A_{эмп.} < A_{критич.}$ и $E_{эмп.} < E_{критич.}$, то распределение исследуемого признака будет близко к нормальному.

Для определения $A_{кр.}$ и $E_{кр.}$ используют формулы Е. И. Пустыльника:

$$A_{кр.} = 3\sqrt{\frac{6(n-1)}{(n+1)(n+3)}} \qquad E_{кр.} = 5\sqrt{\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}}$$

В нашем примере:

$$A_{кр.} = 3\sqrt{\frac{6(16-1)}{(16+1)(16+3)}} = 3\sqrt{\frac{90}{323}} = 3\sqrt{0{,}28} = 3 \cdot 0{,}53 = 1{,}59$$

$$E_{кр.} = 5\sqrt{\frac{384(16-2)(16-3)}{289(16+3)(16+5)}} = 5\sqrt{\frac{69888}{115311}} = 5\sqrt{0{,}61} = 5 \cdot 0{,}78 = 3{,}9$$

Следовательно, $A_{эмп.} < A_{критич.}$ (0,17 < 1,59) и $E_{эмп.} < E_{критич.}$(-0,65< 3,9).

Критерий Е.И. Пустыльника так же подтверждает что данное распределение сколь угодно близко к нормальному.

3.3. Подсчёт теоретических частот нормального распределения

Вариационный ряд, составленный по экспериментальным данным, характеризует эмпирическое распределение изучаемого признака. Если оказывается возможным установить некоторый известный теоретический закон (Пуассона, равномерный и пр.), которому подчинены эмпирические данные, дальнейшее изучение признака значительно упрощается. Наиболее часто в психологических исследованиях встречаются признаки, значения которых подчиняются нормальному закону распределения.

Выбрав закон, которому возможно подчинено рассматриваемое распределение, можно при произвольном значении признака определить теоретическую частоту f_m, с которой признак принимает соответствующее значение.

Если предполагается, что изучаемый признак в исследуемой совокупности распределён по нормальному закону, теоретические частоты определяются по формуле:

$$f_m = \frac{\varphi(t)\, n\, \lambda}{\sigma}$$

Где σ - это среднее квадратическое отклонение ($\sigma = \sqrt{D}$),

n – это объём совокупности (число значений варьирующего признака),

λ – величина интервала (шаг),

φ(t) – значение малой функции Лапласа для параметра t.

Параметр t вычисляется по формуле:

$$t = \frac{x_i - \overline{x}}{\sigma}$$

Значения функции Лапласа табулированы (см. приложение 1).

Например. Получено распределение 1000 студентов некоторого учебного заведения по росту

Рост	134-137	137-140	140-143	143-146	146-149	149-152	152-155	155-158	158-161	161-164	164-167	167-170	170-173
m	1	4	16	53	121	193	229	186	121	53	71	5	1

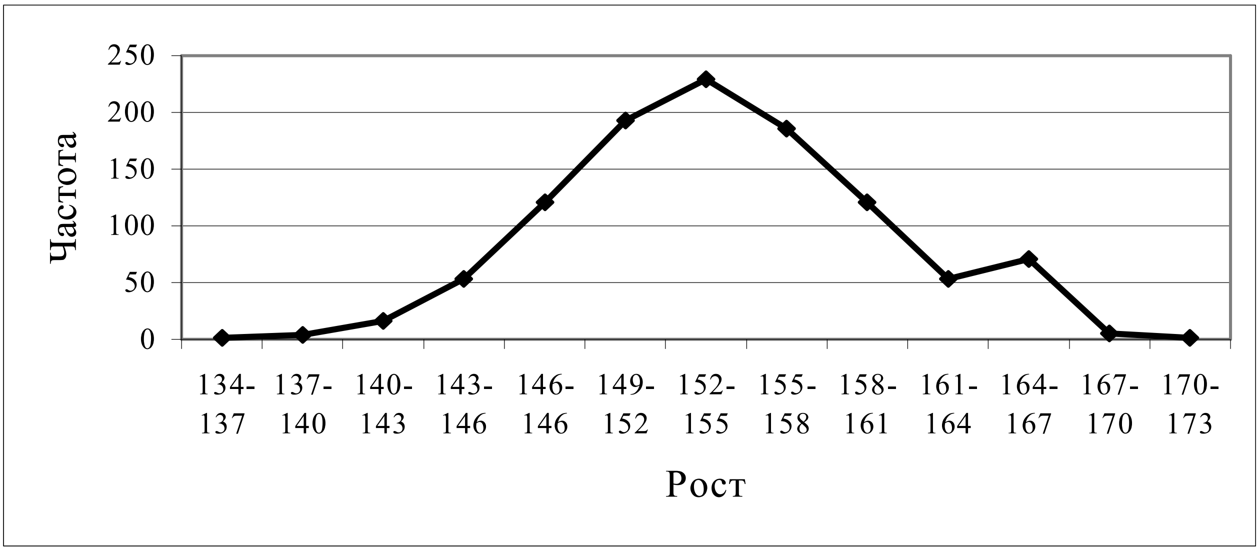

Рисунок 3.3.1.

Требуется: установить теоретический закон распределения, найти его параметры и вычислить теоретический ряд частот.

Таблица 3.3.1.

Рост	Середина интервалов (x_i)	Частота m_i	$x_i\ m_i$	$x_i - \overline{x}$	$(x_i - \overline{x})^2 m_i$
134-137	135,5	1	135,5	-18	324
137-140	138,5	4	554	-15	901
140-143	141,5	16	2264	-12	2306
143-146	144,5	53	7658,5	-9	4299
146-149	147,5	121	17847,5	-6	4365
149-152	150,5	193	29046,5	-3	1737
152-155	153,5	229	35151,5	0	0
155-158	156,5	186	29109	3	1674
158-161	159,5	121	19299,5	6	4356
161-164	162,5	53	8612,5	9	4293
164-167	165,5	17	2813,5	12	2448
167-170	168,5	5	842,5	15	1125
170-173	171,5	1	171,5	18	324
		$\sum$	153506	$\sum$	28152

Вариационный ряд распределения студентов по росту является непрерывным. Известно, что рост довольно большой совокупности людей, как правило, имеет нормальное распределение, что видно из рисунка 3.3.1.

Для нахождения параметров нормального распределения составим вспомогательную таблицу 3.3.1.

Средняя: $\overline{x} = \frac{153506}{1000} = 153{,}506$ см. Дисперсия: $D = \frac{28152}{1000} = 28{,}152$

$\sigma = \sqrt{D} = \sqrt{28{,}152} = 5{,}31$

Вычислим теоретический ряд частот, для этого составим вспомогательную таблицу 3.3.2.

Учитываем, что n = 1000, а $\lambda = 3$ (как размер интервала),

φ (t) - находим по таблице значений малой функции Лапласа (см. приложений 1).

Таблица 3.3.2.

x_i	m_i	$x_i - \overline{x}$	$t = \frac{x_i - \overline{x}}{\sigma}$	φ (t)	$f_m = \frac{\varphi(t) n \lambda}{\sigma}$
1	2	3	4	5	6
135,5	1	-18	-3,39	0,0013	0,7
138,5	4	-15	-2,82	0,0075	4,2
141,5	16	-12	-2,26	0,0310	17,5
144,5	53	-9	-1,69	0,0975	55,1
147,5	121	-6	-1,13	0,2107	119,0
150,5	193	-3	-0,56	0,3410	192,6
153,5	229	0	0	0,3989	225,4
156,5	186	3	0,56	0,3410	192,6
159,5	121	6	1,13	0,2107	119,0
162,5	53	9	1,69	0,0975	55,1
165,5	17	12	2,26	0,0310	17,5
168,5	5	15	2,82	0,0075	4,2
171,5	1	18	3,39	0,0013	0,7

Из сравнения значений эмпирических частот (столбец 2) и теоретических частот (столбец 6) можно сделать предположение, что они различаются незначительно и наше эмпирическое распределение скорей всего распределено по нормальному закону. Однако более точные выводы относительно нормальности распределения можно сделать только после проверки данного предположения при помощи статистического критерия. Наиболее часто для этих целей используется критерий χ^2 – Пирсона (с которым мы познакомимся в последующих пунктах).

3. 4. Особенности статистической обработки непараметрического ряда

Далеко не все материалы, получаемые в психологических исследованиях, подлежат обработке параметрическими методами. Если после ознакомления с изучаемым рядом исследователь убеждается в том, что этот ряд не имеет свойств нормального распределения, ему остается перейти на методы непараметрической статистики. С их помощью могут быть получены и центральная тенденция изучаемого ряда — мода, медиана и величина, позволяющая судить о диапазоне варьирования и о строении изучаемою ряда — квантильное отклонение.

Например, после проведения диагностических испытаний уровня умственного развития учеников 6-го класса все полученные данные были упорядочены, т. е. расположены в последовательности от меньшей величины к большей. Испытания проходили 18 учащихся. Буквами обозначены учащиеся, числами — полученные ими баллы по тесту, столбцы под буквами R— ранги.

Таблица 3.4.1.

Уч-ся	Баллы	Ранги (R)	Уч-ся	Баллы	Ранги (R)
А	25	1	К	68	10
Б	28	2	Л	69	11,5
В	39	4	М	69	11,3
Г	39	4	Н	70	14,5
Д	39	4	О	70	14,5
Е	45	6	П	70	14,5
Ж	50	7	Р	70	14,5
З	52	8,5	С	74	17,5
И	52	8,5	Т	74	17,5

Напомним, что процедура ранжирования состоит в следующем: все числа ряда в их последовательности получают по своим порядковым местам присваиваемые им ранги. Если какие-нибудь числа повторяются, то всем повторяющимся числам присваивается один и тот же ранг — средний из общей суммы занятых этими числами мест. Так числу «28» в изучаемом ряду присвоен ранг «2». Затем следуют трижды повторяющиеся числа «39». На них приходятся занятые ими ранговые места «3», «4», «5». Поэтому этим числам присваивается один и тот же средний ранг, в данном случае — «4». Поскольку места до 5 включительно заняты, то следующее число получает ранг «6» и т. д.

При обработке ряда, не имеющего признаков нормального распределения, иначе — непараметрического ряда — для величины, которая выражала бы его центральную тенденцию, более всего пригодна ранговая медиана, т. е. величина, расположенная в середине ряда. Ее определяют по срединному рангу по формуле

$$Me = \frac{n+1}{2}$$

где Ме — означает ранговую медиану, n — число членов ряда. При нечетном числе членов ряда ранговая медиана — целое число, при четном — число с 0,5. Заметим, что числовое значение медианы может и не быть в составе самого обрабатываемого ряда.

Возьмем к примеру ряд в семь членов: 3—5—6—7—9—10—11.
Проранжировав этот ряд, имеем: 1—2—3—4—5—6—7.
Ранговая медиана в таком ряду равна:

$$Me = \frac{7+1}{2} = 4$$

что соответствует величине 7.

Возьмем ряд в восемь членов: 3—5—6—7—9—10—11—12,
Проранжировав этот ряд, имеем: 1—2—3—4—5—6—7—8,
Ранговая медиана в этом ряду равна:

$$Me = \frac{8+1}{2} = 4{,}5$$

Этому рангу соответствует середина между двумя величинами, имеющими ранг 4 и ранг 5, т. е. между семью и девятью.

Медиана этого ряда равна:

$$Me = \frac{7+9}{2} = 8$$

Следует обратить внимание на то, что величины 8 в составе ряда нет, но таково значение медианы этого ряда.

Вернемся к изучаемому ряду. Он состоит из 18 членов. Его ранговая медиана равна

$$Me = \frac{19+1}{2} = 9{,}5$$

Она расположится между 9-й и 10-й величиной ряда.
9-я величина ряда — 52, 10-я величина ряда — 68. Медиана занимает срединное место между этими величинами, следовательно,

$$Me = \frac{52+68}{2} = 60$$

По обе стороны от этой величины находится по 50% величин ряда.

Характеристику распределения численностей в непараметрическом ряду можно получить из отношения его квантилей. **Квантилью** называется величина, отграничивающая 1/4 всех величин ряда Квантиль первая — ее обозначение Q_1 вычисляется по формуле:

$$Q_1 = \frac{R_1 + R_{n/2(лев)}}{2}$$

это полусумма первою и последнего рангов первой — левой от медианы половины ряда; квантиль третья, обозначаемая Q_3, вычисляется по формуле:

$$Q_3 = \frac{R_{n/2} + R_{n(прав)}}{2}$$

т. е. как полусумма первого и последнего рангов второй, правой от медианы, половины ряда. Берутся порядковый значения рангов по их последовательности в ряду. В обрабатываемом ряду

$$Q_1 = \frac{1+9}{2} = 5 \quad Q_3 = \frac{10+18}{2}$$

Рангу 5 в этом ряду соответствует величина 39, а ранг 14 — величина 70.

Для характеристики распределения в непараметрическом ряду вычисляется среднее квантильное отклонение, обозначаемое Q. Формула для Q такова:

$$Q_1 = \frac{Q_3 - Q_1}{2}$$

Для обрабатываемого ряда $Q_1 = \frac{70-39}{2} = 15{,}5$

Параметрический ряд относится к шкале интервалов, непараметрический — к шкале порядка. Но встречаются также ряды, относящиеся к шкале наименований. Наиболее краткая характеристика такого ряда может быть получена с помощью моды, величины, которая выражает наивысшее числовое значение величин данного ряда, и «n» — числом членов ряда. Следует заметить, что моду можно лишь условно считать выражением центральной тенденции в ряду, относящемся к шкале наименований. Она выражает наиболее типичную величину ряда.

Рассмотрим пример, где речь идет об участниках некой конференции, в их числе 3 англичанина, 2 датчанина, 5 немцев, 1 русский и 2 француза. Мода в данном ряду приходится на участников конференции — немцев. Число членов ряда равно — 13, а мода — Mo = 5.

На предшествующих страницах были рассмотрены статистические методы, применяющиеся для задач первого типа. Для задач, требующих установления сходства или различия, ставящих целью анализ источников вариативности, получаемых психологических признаков и предполагающих возможность прогноза на основе имеющихся данных, применяются методы относящихся к группе, так называемых методов вторичной статистической обработки экспериментальных данных.

4. ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

4.1.Общие принципы проверки статистических гипотез

При осуществлении психолого-педагогического эксперимента экспериментатор должен проверить статистическими методами некоторые положения.

Например, если исследовать уровень развития интеллекта детей в неполной семье и уровень интеллекта в полной семье, можно выдвинуть такое научное предположение, что средний коэффициент интеллекта у детей из неполных семей отличается от среднего коэффициента у детей в полной семье. Такое предположение называют **статистической гипотезой**, которую можно подтвердить или опровергнуть.

Статистическая гипотеза – формальное предположение о том, что сходство или различие некоторых параметрических характеристик генеральной совокупности или выборочных совокупностей случайно или неслучайно.

При проверке статистических гипотез используются два понятия:

1. H_o – нулевая гипотеза, это гипотеза о сходстве.
2. H_1 – альтернативная гипотеза - гипотеза о различии.

Принятие H_o свидетельствует об отсутствии различий.

Например, $\overline{x}_э - \overline{x}_m = 0$ или $a_1 - a_2 = 0$

и $\sigma_э - \sigma_т = 0$ или $D_1 - D_2 = 0$.

Принятие H_1 свидетельствует о наличии различий.

Например, $\sigma_э - \sigma_т \neq 0$

При проверке гипотезы, т.к. все данные у нас имеют вероятностную природу, неизбежны ошибки.

При проверке гипотез возможны ошибки двух родов:

1. в результате статистической проверке нулевая гипотеза (H_o) отклоняется, но на самом деле она верна;
2. когда нулевая гипотеза (H_o) не отклоняется, но на самом деле верна альтернативная гипотеза (H_1).

Ошибки тесно связаны с понятием уровня статистической значимости.

Уровень значимости – это вероятность ошибки первого рода при принятии решения. Для обозначения уровня значимости используют Р = 0,05; Р= 0,01; Р = 0,001. Вероятность ошибки (Р) показывает процент ошибки, допустимый при статистическом исследовании.

Например, в группе из 100 учеников проводилось исследование по уровню утомляемости в начале учебного года и в конце учебного года. Была принята H_o, заключающееся в том, что существенных различий между уровнем утомляемости в начале года и в конце года нет (с уровнем значимости 0,05). Это означает, что различия могут быть только в пяти процентах от общего количества испытуемых, т.е. различия могут наблюдаться у пяти детей из этих ста.

В основном для любой гипотезы сначала вычисляют по соответствующим (выбранному статистическому критерию) методикам (формулам) эмпирический показатель характеристики ($x_э$) исследуемого признака, затем сравнивают его с критическим значением ($x_к$), вычисленным по таблице, соответствующей выбранному теоретическому распределению.

$$x_к = \begin{cases} x_к, & p = 0{,}05 \\ x_к, & p = 0{,}01 \end{cases}$$

Эти числа определяют так называемую числовую ось значимости. Числовая ось значимости разбита на 3 зоны: зону не значимости, зону неопределенности и зону значимости (рисунок 4.1.1).

зона незначимости зона неопред. зона значимости

H_o $x_к\,(p = 0{,}05)$ $x_к\,(p = 0{,}01)$ H_1

Рисунок 4.1.1. Ось значимости

Эмпирический показатель характеристики $X_э$ исследуемого признака будет располагаться в одной из этих зон. Если $x_э$ попал в зону незначимости, то в этом случае принимается H_o (различий нет), если $x_э$ попал в зону значимости, то принимается H_1 (т.е. H_o отвергается). Если $x_э$ попала в зону неопределенности, в этом случае психолог должен интуитивно решить сам, какую гипотезу принять, в зависимости от точности, важности и значимости исследования (или выбрать другой более мощный статистический критерий).

Иногда используют другой подход, когда определяют не вероятность ошибки (α), а вероятность достоверности значения, которое обозначается β причём $\alpha = 1 - \beta$. В этом случае находят по соответствующей таблице критическое значение предельно допустимого значения ($P = 0{,}95$, $P = 0{,}99$).

Этапы принятия статистического решения

1. Формулировка H_o и H_1.
2. Определение объема выборки.
3. Выбор соответствующего уровня значимости ($\approx 0{,}05,\ 0{,}01,\ 0{,}001$).
4. Выбор статистического метода (критерия), который зависит от типа решаемой психологической задачи.
5. Вычисление соответствующего эмпирического значения по экспериментальным данным, согласно выбранному статистическому методу (критерию).
6. Нахождение по таблице для выбранного статистического метода критических значений, соответствующих уровню значимости

P = 0,05 и P=0,01 или уровень допустимого значения P = 0,95, P = 0,99.

7. Построение оси значимости и нанесение на нее табличных критических значений и эмпирических значений измеряемой величины.
8. Формулировка принятия решения (принятие гипотезы).

Обращаясь к таблицам уровней значимости, исследователь обнаруживает во многих из них специальный столбец с указанием степеней свободы, относящихся к полученному параметру или коэффициенту. Уровень значимости прямо зависит от того, каким числом степеней свободы обладает данный коэффициент или параметр Число независимых величин, участвующих в образовании того или другого параметра, называется **числом степеней свободы** этого параметра. Оно равно общему числу величин, по которым вычисляется параметр, минус число условий, связывающих эти величины. Число степеней свободы и способы его определения всегда даются в окончательных формулах, которыми пользуется исследователь при статистической обработке своих материалов.

4.2. Статистические критерии

Одним из направлений в статистическом анализе является сравнение, проверка гипотезы о разности или сходстве двух эмпирических выборок, о принадлежности полученной выборки генеральной совокупности. Такое доказательство необходимо, чтобы, например, установить возрастные нормы, различия в успешности деятельности, влияние корригирующих воздействий и т.д. Для этого необходимо пользоваться определенными статистическими критериями.

Статистический критерий - это правило, обеспечивающее принятие верного статистического решения, то есть принятие или отклонение гипотезы H_o с высокой вероятностью.

Статистические критерии обозначают также метод расчета определенного числа и само это число. По соотношению эмпирического и критического значений критерия можно судить о том, подтверждается ли или опровергается гипотеза. Критерии делятся на параметрические и непараметрические.

Параметрические критерии - это критерии, у которых в формулу расчета включены параметры распределения (М и σ). К ним относятся t критерий Стьюдента, критерий F и др.

Непараметрические критерии - критерии, не включающие в формулу расчета параметры распределения и основанные на оперировании частотами или рангами. К ним относятся критерий Q Розенбаума, критерий знаков и др.

И те, и другие критерии имеют свои преимущества и недостатки. Выбор критерия доказательства различий зависит от следующих особенностей групп:

1) зависимы или независимы между собой группы;

2) подчиняются или не подчиняются они закону нормального распределения;

3) от объема выборок.

Зависимыми называются группы, между которыми есть существенные внутренние связи: выборки, состоящие из родителей и детей, братьев и сестер (генетическая связь); одна и та же группа, исследуемая дважды (например, до и после педагогического воздействия).

Независимые - это группы, между которыми нет существенных связей (например, группы одного возраста разных профессий, группы разного возраста или пола, не являющиеся родственниками).

Мощность критерия — это вероятность принятия при его применении правильного решения для отклонения гипотезы H_o; чем выше эта вероятность, тем больше мощность критерия. Мощность любого критерия увеличивается вместе с увеличением объема сравниваемых выборок, а также со снижением того уровня значимости, на который ориентируется исследователь. Другими словами, если выборки велики, то принятие правильного решения относительно гипотезы H_0 увеличивается. Ориентация на высокий уровень значимости, например, 0,990 или 0,999 предполагает применение достаточно мощного критерия.

Одним из самых распространенных в психолого-педагогических исследованиях статистических критериев является так называемый универсальный критерий χ^2 – Пирсона, применяемый для выявления сходства или различия в распределениях либо двух различных совокупностей, либо сходства или различия распределения исследуемой совокупности с теоретическим распределением (равномерным, нормальным, показательным и т.д.), рассмотрению которого посвящен следующий параграф.

5. КРИТЕРИЙ X^2 - ПИРСОНА

5.1. Общие положения применения критерия χ^2 - Пирсона

Допустим, что сравниваются две выборки, составленные соответственно из выпускников двух школ. Часть выпускников каждой школы сдавала экзамены в вузы. Из первой школы сдавали экзамены 100 человек, из них 82 успешно сдали экзамены, не сдали —18. Таково распределение численности в первой выборке. Из второй школы сдавали экзамены в вузы 87 человек, выдержали 44 человека, не сдали — 43. Таково распределение численностей во второй выборке. Достаточно ли этих данных, чтобы утверждать, что подготовленность к вузовским экзаменам выпускников этих школ неодинакова? На первый взгляд, разница налицо: лучше подготовлены выпускники первой школы. Однако при таком раскладе численностей возможно влияние случайности. Поэтому встает вопрос: можно ли, считаясь с представленными распределениями, прийти к статистически обоснованному выводу о мере подготовленности к экзаменам в вузы той и другой выборки.

Метод, с помощью которого подвергаются статистическому анализу описанные распределения численностей, получил название «хи» - квадрат, его обозначают греческой буквой χ^2 с показателем степени. Он был разработан математиком Пирсоном. Метод χ^2 — весьма универсален, применим во многих исследованиях, он пригоден для статистического анализа распределения численностей разнообразных количественных материалов, относящихся ко всем статистическим шкалам, в том числе и к шкале наименований.

Формула определения критерия χ^2 -квадрат:

$$\chi^2 = \sum \frac{(f_э - f_m)^2}{f_m}$$

где $f_э$ — наблюдаемые (эмпирические) численности (частоты);

f_m — предполагаемые (теоретические) численности (частоты).

Критерий χ^2 – один из наиболее часто использующихся в психологических исследованиях, поскольку он позволяет решать большое число различных задач. Этот критерий используется в двух вариантах:

1. Как расчет согласия эмпирического значения и предполагаемого теоретического. В этом случае проверяется H_0 об отсутствии различий между теоретическим и эмпирическим распределением.
2. Как расчет однородности двух независимых экспериментальных выборок. В этом случае проверяется гипотеза H_0 об отсутствии различий между двумя эмпирическими распределениями.

Ограничения критерия

1. Объем выборки должен быть достаточно большим $N \geq 30$. При $N < 30$ критерий дает весьма приближенные значения. Точность

повышается с ростом N.

2. Теоретическая частота для каждой ячейки таблицы не должна быть меньше пяти. Например, если составлять распределение числа обращений в пожарную охрану за неделю (семь дней), то должно быть 7*5=35 обращений.
3. Выбранные разряды должны вычерпывать все распределения, то есть охватывать весь диапазон вариативности признаков. Группировка на разряды должна быть одинаковой во всех сопоставляемых распределениях.
4. Необходимо вносить поправку на непрерывность при сопоставлении распределений признаков, которые принимают два значения. При внесении поправки значение χ^2 уменьшается.
5. Разряды должны быть неперекрещивающиеся. Если наблюдение отнесено к одному разряду. То оно уже не может быть отнесено не к одному другому разряду.

5.2. Алгоритм расчета критерия χ^2 Пирсона

Техника вычисления критерия χ^2довольно проста. Рассмотрим пример со сдачей экзаменов в вузы выпускниками 1-ой и 2-ой школ, приведенный в начале предыдущего параграфа. В условии сказано, что всего намерены были сдавать экзамены 187 человек, из этого числа на долю 1-ой школы приходится 53,5% (100 человек), а на долю 2-ой школы — 46,5% (87 чел.). Предположим, что выпускники той и другой школы подготовлены одинаково, тогда и доли сдавших и не сдавших будут такие же, как доли их представленности в общем числе сдающих. Всего сдало экзамены 126 выпускников. Согласно высказанному предположению, 53,5% от этого числа должны бы были прийтись на 1-ую школу — это составит 66,9 от 126 —и 46,5% на 2-ую школу, что составит 58,9 от 126. Такое же рассуждение повторяем и относительно не сдавших. Их всего 61 человек. На 1-ую школу, как нам известно, должно, по предположению, прийтись 53,5% от этого числа, т. е. 33,0 от 61, а на долю 2-ой школы — 46,5%, т. е. 28,1 от 61. Нуль-гипотеза, имеющая в данном раскладе тот смысл, что между выпускниками нет различия, при таком соотношении сдавших и не сдавших подтвердилась бы. Однако в условиях этого исследования показано другое распределение. Количество выпускников 1-ой школы, сдавших экзамены, составляет 82, а не 66,9, как можно было бы предположить, исходя из нуль - гипотезы. Соответственно количество выпускников 2-й школы, сдавших экзамены, составляет в действительности всего 44, а не 58,9. Точно также, сравнивая количество не сдавших (по условию с предполагаемым распределением) найдем по 1-ой школе 18, а не 33, а по 2-ой школе — 43, а не 28,1. Расхождение между действительными (наблюденными) распределениями и распределениями, которые могли бы иметь место, если исходить из нуль - гипотезы, налицо. Они-то и учитываются при вычислении χ^2. Все сказанное удобно представить в виде таблицы-графика распределения численностей (частот). Количества, которые были бы получены при принятии нуль - гипотезы, заключены в скобки. В правом углу буквенное обозначение клетки.

Таблица 5.2.1.

Школы	Число сдавших	Число не сдавших	Всего	Долевые отношения
1	82 А (69,9)	18 В (33,0)	100 (100)	53,5%
2	44 С (58,9)	43 Д (28,1)	87 (87)	46,5%
Всего	126	61	187	100%

Получены разности по клеткам (знак разности несущественен).

Клетки: А $f_А$ = 82-66,9=15,1

В $f_В$ = 18-33=15,0

С $f_С$ = 44-58,9=14,9

Д $f_Д$ = 43-28,1=14,9

Формула определения критерия χ^2 -квадрат:

$$\chi^2 = \sum \frac{(f_э - f_m)^2}{f_m}$$

где fэ — наблюдаемые (эмпирические) численности;

f*m* — предполагаемые (теоретические) численности.

В рассмотренном примере эмпирический критерий Пирсона $\chi_э{}^2$ равен

$$\chi_э{}^2 = \frac{15{,}1^2}{66{,}9} + \frac{15^2}{33} + \frac{14{,}9^2}{58{,}9} + \frac{14{,}9^2}{28{,}1} = \frac{228}{66{,}9} + \frac{225}{33} + \frac{222}{58{,}9} + \frac{222}{28{,}1} = 3{,}4 + 6{,}8 + 3{,}8 + 7{,}9 = 21{,}9$$

Для получения числа степеней свободы нужно воспользоваться формулой (только для χ^2): μ = (к—1) (с—1) = (2—1)(2—1) = 1, где к — число столбцов, с — число строк с анализируемым материалом.

Обратимся для нахождения $\chi^2{}_К$ к таблице уровней значимости для одной степени свободы (приложение 2).

$$\chi_К{}^2 = \begin{cases} 3{,}841 & p = 0{,}05 \\ 6{,}635 & p = 0{,}01 \end{cases}$$

Эти числа определяют числовую ось значимости. Числовая ось значимости разбита на 3 зоны: зону не значимости, зону неопределенности и зону значимости (рисунок 5.2.1).

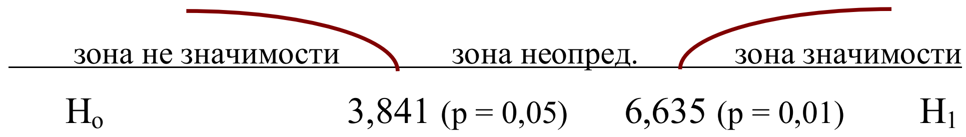

Рисунок 5.2.1. Ось значимости

Значение эмпирического критерия $\chi_э{}^2 = 21{,}9$ расположено на числовой оси значимости в зоне значимости. Следовательно, полученная величина $\chi_э{}^2$ вполне достаточна для отклонения H_0. Есть все основания для содержательного вывода о различной степени подготовленности выпускников 1-ой и 2-ой школ к экзаменам в вузы.

Алгоритм расчета критерия χ^2 Пирсона

1. Занести в таблицу значения вариант (или наименование классов, разрядов) и соответствующие им эмпирические частоты.
2. Рядом с каждой эмпирической частотой записать теоретическую частоту. Теоретическая частота находится по формулам, правилам соответствующего распределения (равномерного, нормального и т.д.).
3. Подсчитать разности между эмпирической и теоретической частотой по каждой строке и записать их в третий столбец.
4. Возвести в квадрат полученные разности, и занести их в четвертый столбец.
5. Разделить полученный квадрат разности на теоретическую частоту и записать результаты в пятый столбец.
6. Просуммировать значения пятого столбца, полученную сумму обозначить как $\chi^2_э$.
7. Определить число степеней свободы.
8. Определить по таблице критическое значение $\chi^2{}_к$ для данного

числа степеней свободы.

9. Построить ось значимости. Определить по оси значимости в какую из трех зон (зону не значимости, зону неопределенности и зону значимости) попадает $\chi^2_э$. Принять статистическое решение: принимается или отвергается гипотеза H_0

Замечание: Если в своей статистической обработке вы пользуетесь однополюсной осью (чаще всего для уровня p = 0,05), то если $\chi^2_э$ меньше критического значения критерия Пирсона, то расхождения между распределениями статистически недостоверны. Если $\chi^2_э$ равно критическому значению или превышает его, то расхождения между распределениями статистически достоверны.

Пример. Имеем распределение 1000 студентов некоторого учебного заведения по росту. Предположим, что это распределение является нормальным. Что бы подтвердить эту гипотезу при помощи критерия Пирсона, сравним эмпирические частоты этого распределения с теоретическими, посчитанными для него ранее в.п.3.3. Согласно алгоритма, составим таблицу 5.2.2.

Таблица 5.2.2.

Рост	$f_э$	$f_т$	$f_э - f_т$	$(f_э - f_т)^2$	$(f_э - f_т)^2 / f_т$
134-137	1	0,7	0,3	0,09	0,1286
137-140	4	4,2	-0,2	0,04	0,0095
140-143	16	17,5	-1,5	2,25	0,1285
143-146	53	55,1	-2,1	4,41	0,0800
146-149	121	119,0	2	4	0,0336
149-152	193	192,6	0,4	0,16	0,0008
152-155	229	225,4	3,6	12,96	0,0575
155-158	186	192,6	6,6	43,56	0,2262
158-161	121	119,0	2	4	0,0336
161-164	53	55,1	-2,1	4,41	0,0800
164-167	17	17,5	-0,5	0,25	0,0143
167-170	5	4,2	0,8	0,64	0,1524
170-173	1	0,7	0,3	0,09	0,1286
				$\sum$	1,074

Следовательно, $\chi_э{}^2 = 1{,}074$

Число степеней свободы:

$$\mu = (к—1)\ (с—1) = (2—1)(13—1) = 12$$

Найдем $\chi^2{}_K$ по таблице уровней значимости для 12 степеней свободы (приложение 2).

$$\chi_K{}^2 = \begin{cases} 21{,}026 & p = 0{,}05 \\ 26{,}217 & p = 0{,}01 \end{cases}$$

Построим числовую ось значимости (рисунок 5.2.2).

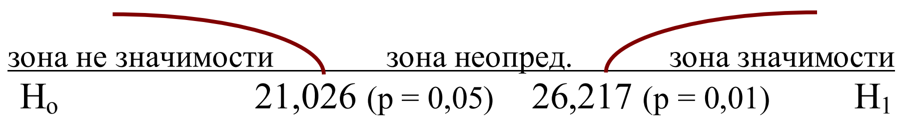

Рисунок 5.2.2. Ось значимости

Значение эмпирического критерия $\chi_э{}^2 = 1{,}074$ расположено на числовой оси значимости в зоне не значимости. Следовательно, принимаем гипотезу H_0 – об отсутствии различий между эмперическим распределением и теоретическим. Следовательно, данное распределение студентов по росту является нормальным, т.е. подчинено нормальному закону распределения.

6. ПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ РАЗЛИЧИЙ

6. 1. Критерий Стьюдента для независимых выборок

Предположим, что оба ряда распределений двух выборок подчинены нормальному закону. Сравнение величин центральных тенденций, например средних арифметических, не даёт ответа на вопрос о том, относятся ли выборки к одной совокупности. Даже если, средние арифметические будут тождественными, этого явно недостаточно для ответа на вопрос – существуют ли различия между этими выборками? Ответ на данный вопрос можно получить, применив один из параметрических критериев различий.

Для выявления различий между двумя выборками по известным средним чаще всего используют t – критерий Стьюдента.

В общем виде, для расчетов t – критерия Стьюдента применяют следующие формулы:

$$t = \frac{\left|\overline{X} - \overline{Y}\right|}{S}, \qquad \text{где } S = \sqrt{\sigma_x^2 + \sigma_y^2}$$

Конкретно S вычисляют для одинаковых по объему (n) выборок по формуле:

$$S = \sqrt{\frac{\sum(X_i - \overline{X})^2 + \sum(Y_i - \overline{Y})^2}{(n-1)n}}$$

Для разных по объему ($n_1 \neq n_2$) выборок по формуле:

$$S = \sqrt{\frac{n_1 + n_2}{n_1 n_2}} \cdot \sqrt{\frac{\sum(X_i - \overline{X})^2 + \sum(Y_i - \overline{Y})^2}{n_1 + n_2 - 2}}$$

В обоих случаях подсчет степеней свободы осуществляется по формуле: $k = n_1 + n_2 - 2$.

Рассмотрим пример использования t – критерия Стьюдента для несвязных выборок одинакового объема.

Двум группам шестиклассников по 6 человек было дано задание бросать мяч в корзину. Группы обучались по разным программам.

Можно ли считать, что разница в программах сказалась на конечной результативности школьников? Для сравнения было взято число попаданий в корзину. Всего было дано по 10 проб (см. Таблица 6.1.1).

Таблица 6.1.1.

№	Группы		Отклонения от среднего		Квадраты отклонения	
	X	$У$	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1	2	5	-1	0	1	0
2	4	4	1	-1	1	1
3	6	2	3	-3	9	9
4	4	8	1	3	1	9
5	1	6	-2	1	4	1
6	1	5	-2	0	4	0
$\sum$	18	30			20	20

Ход вычислений показывает:

1 выборка: $\sum x_i = 18 \qquad \bar{x} = \dfrac{\sum x_i}{n} = \dfrac{18}{6} = 3 \qquad \sum (x_i - \bar{x})^2 = 20$

2 выборка: $\sum y_i = 30 \qquad \bar{y} = \dfrac{\sum y_i}{n} = \dfrac{30}{6} = 5 \qquad \sum (y_i - \bar{y})^2 = 20$

Используем формулу: $S = \sqrt{\dfrac{\sum (X_i - \bar{X})^2 + \sum (Y_i - \bar{Y})^2}{(n-1)n}}$

$$S = \sqrt{\frac{20+20}{5 \cdot 6}} = \sqrt{\frac{40}{30}} = \sqrt{1{,}33} = 1{,}15$$

$$t_э = \frac{|\bar{X} - \bar{Y}|}{S} = \frac{2}{1{,}15} = 1{,}74$$

k (число степеней свободы) = $n_1 + n_2 - 2 = 10$.

По таблице уровней значимости t - критерия Стьюдента (приложение 3) находим при 10 степенях свободы $t_к$:

$$t_к = \begin{cases} 2{,}23 & p = 0{,}05 \\ 3{,}17 & p = 0{,}01 \end{cases}$$

Строим ось значимости:

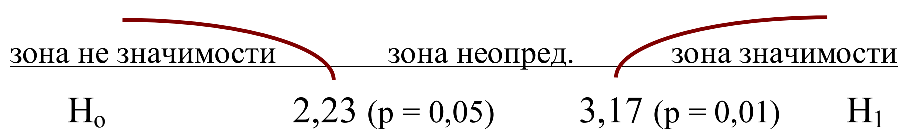

Рисунок 6.1.1. Ось значимости

$t_э$ = 1,74 находится в зоне не значимости, следовательно принимается гипотеза H_o: выборки отличаются одна от другой несущественно.

Для вычисления t- критерия существует несколько формул, различающихся только техникой расчетов.

6.2. Критерий Стьюдента для зависимых выборок

Если выборки зависимые, то доказательство различия средних арифметических можно произвести по следующей формуле:

$$t = \frac{d}{\sigma_d / \sqrt{n}}$$

где d- среднее значение разности всех пар показателей;

σ_d - стандартное отклонение среднего значения разности;

n - число пар наблюдений.

Например: на курсах повышения квалификации слушатели обучались некоторым приемам мнемотехники. Чтобы проверить эффективность обучения, проводилось тестирование перед началом курса и после его окончания: предлагалось запомнить на слух 20 слов, фиксировалось число правильно повторенных слов.

Гипотеза H_o - владение приёмами мнемотехники в группе испытуемых после обучения не изменилось и гипотеза H_1 - изменения после обучения оказались существенными.

Для расчет t-критерия Стьюдента заполним вспомогательную таблицу 6.2.1.

При заполнении таблицы были выполнены следующие действия:

1. Для каждого испытуемого найти разность первого и второго замеров (d_i)

2. Возвести разность в квадрат (d_i^2).

3. Подсчитать суммы d_i (столбик 4) и d_i^2 (столбик 5).

4. Рассчитать среднее значение разности $d = \frac{\sum_{i=1}^{n} d_i}{n} = 49/10 = 4.9$

Таблица 6.2.1.

№	X_{i1} (до)	X_{i2} (после)	$d_i = X_{i2} - X_{i1}$	d_i^2
1	10	18	8	64
2	6	10	4	16
3	7	7	0	0
4	5	8	3	9
5	8	12	4	16
6	4	7	3	9
7	7	15	8	64
8	6	14	8	64
9	5	11	6	36
10	8	13	5	25
		$\sum$	49	303

5. Рассчитать стандартное отклонение для разностей по формуле:

$$\sigma_d = \sqrt{\frac{\sum d_i^2 - (\sum d_i)^2 / n}{n-1}}$$

$$\sigma_d = \sqrt{[303 - 49^2/10]/10-1} = \sqrt{(303-240{,}1)/9} = \sqrt{7} \approx 2{,}7$$

6. Рассчитать величину $t_э$: $t = \frac{\bar{d}}{\sigma_d/\sqrt{n}} = \frac{4.9}{2{,}7/\sqrt{10}} \approx 5{,}77$

7. Определить степень свободы k = n-1 = 9 и табличные значения критерия:

$$t_к = \begin{cases} 2{,}262 & p = 0{,}05 \\ 3{,}250 & p = 0{,}01 \end{cases}$$

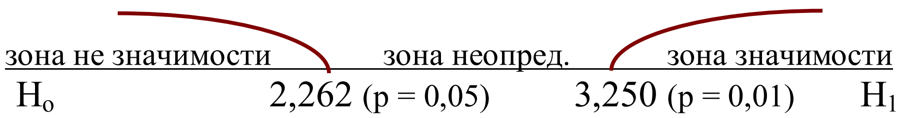

Рисунок 6.2.1. Ось значимости

Таким образом, полученное значение $t_э$ = 5,77 превышает табличные значения для всех уровней значимости. Поэтому принимаем гипотезу H_1, то есть делаем вывод об эффективности обучения.

Критерий Стьюдента предполагает, что наблюдения в группах имеют нормальное распределение. Но не всегда распределение является нормальным.

7. НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ РАЗЛИЧИЙ

7.1. Критерий знаков G для связных выборок

Нередко сравнивая на глазок результаты до и «после» какого либо воздействия (например тренинга) психолог видит тенденции повторного измерения — большинство показателей может увеличиваться или напротив уменьшаться. Наиболее простым путем оценки различий казалось бы является подсчет процентов в изменениях в ту или другую сторону «до» и «после» и сравнение полученных процентов между собой. На основе этого сравнения можно было бы придти к заключению, что если наблюдаются различия в процентах, то имеет место различие и в сравниваемых психологических характеристиках «до» и «после». Подобный подход категорически неприемлем, поскольку для процентов нельзя определить уровень достоверности в их различиях. Делать какие либо выводы из экспериментального материала возможно только на основе статистических процедур специально сконструированных так, что на их основе можно определить уровень достоверности различий. Проценты, взятые сами по себе не дают возможности делать статистически достоверные выводы. Поэтому для того чтобы доказать эффективность какого либо воздействия необходимо выявить статистически значимую тенденцию в смещении (сдвиге) показателей.

Для решения подобных статистических задач психолог может использовать целый ряд критериев различия. Один из наиболее простых критериев различия — критерии знаков **G**. Этот критерий относится к непараметрическим, и применяется только для связанных (зависимых) выборок. Он дает возможность установить, насколько однонаправленно изменяются значения признака при повторном измерении связанной, однородной выборки. Критерий знаков применяется к данным, полученным в ранговой, интервальной и шкале отношений.

Решим с использованием критерия знаков следующую задачу:

Психолог проводит групповой тренинг. Его задача — выяснить будет ли эффективен данный конкретный вариант тренинга для снижения уровня тревожности участников?

Для решения этой задачи психолог с помощью теста Тейлора дважды *выявляет* уровень тревожности у 14 участников до и после проведения тренинга. Результаты измерения приведем в таблице 7.1.1, включив в нее столбец, необходимый для расчета по критерию знаков **G.**

Таблица 7.1.1.

№п/п	Уровень тревожности «до» тренинга	Уровень тревожности «после» тренинга	Сдвиг
1	30	34	+ 4
2	39	39	0
3	35	26	-9
4	34	33	- 1
5	40	34	-6
6	35	40	+ 5
7	22	25	+ 3
8	22	23	+ 1
9	32	33	+ 1
10	23	24	+ 1
11	16	15	- 1
12	34	27	-7
13	33	35	+ 2
14	34	37	+ 3

В столбце, обозначенном словом «Сдвиг» для каждого у частника отдельно определяют насколько изменился его уровень тревожности после проведения тренинга. Сдвиг — это величина разности между уровнями тревожности одного и того же участника «после» и «до» тренинга. Но не наоборот. Величины сдвигов обязательно должны быть даны в соответствующем столбце таблице с учетом знаков.

В критерии знаков по результатам полученным в столбце таблицы обозначенном словом «Сдвиг» подсчитываются суммы нулевых положительных и отрицательных сдвигов. При использовании критерия знаков необходимо учитывать только сумму положительных и отрицательных сдвигов, а сумму нулевых — отбрасывать.

Проведем необходимый подсчет для нашей задачи:

- общее число (сумма) нулевых сдвигов = 1;
- общее число (сумма) положительных сдвигов = 8;
- общее число (сумма) отрицательных сдвигов = 5.

Таким образом отбросив нулевые сдвиги получаем 13 ненулевых сдвигов. При этом подсчет показал, что сдвиги имели место и что большая часть из них положительна.

Напомним, что критерии знаков **G** предназначен для установления того, как изменяются значения признака при повторном измерении связной выборки в сторону увеличения или уменьшения. Поэтому, анализируя соотношение положительных и отрицательных сдвигов в нашей задаче, решаем вопрос можно ли утверждать, что после проведения тренинга наблюдается достоверный сдвиг в сторону уменьшения уровня тревожности участников?

Для решения этого вопроса необходимо ввести два обозначения:

Первое — сумма сдвигов, получившаяся наибольшей носит название *типичного* сдвига и обозначается буквой *n.* Типичный сдвиг используется при работе с таблицей (приложения 5) в которой приводятся критические величины 5% и 1% уровней значимости данного критерия.

Второе — сумма сдвигов получившаяся наименьшей носит название — нетипичного сдвига и обозначается как — $G_{мп}$. Эта величина располагается на оси значимости. В нашем случае $G_{мп} = 5$.

В целом типичный и нетипичный сдвиги рассматриваются как дополнительные друг к другу.

Подчеркнем, что в том случае когда величины типичного и нетипичного сдвигов оказываются равными, критерий знаков неприменим.

Оценка статистической достоверности различии по критерию знаков производится по таблице (приложения 5). В ней в столбце обозначенным буквой *n* приведены величины типичных сдвигов а в столбцах имеющих обозначение соответствующее уровнями значимости $p = 0,05$ и $p = 0,01$ — так называемые критические величины. Условно их также можно считать нетипичными сдвигами. Они обозначаются как $G_к$ и с ними сравнивается полученное значение нетипичного сдвига $G_{мп}$.

Итак, оцениваем уровень достоверности различий нашей задачи. Для этого необходимо воспользоваться таблицей (приложения 5).

При $n = 8$ (это число типичных сдвигов) равно:

$G_к = 1$ для $p = 0,05$ и

$G_к = 0$ для $p = 0,01$.

Ось значимости:

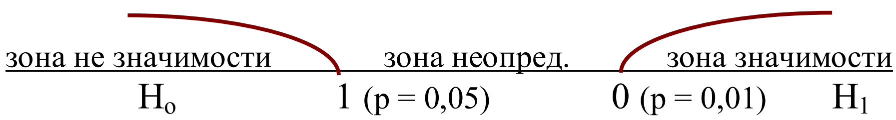

Рисунок 7.1.1. Ось значимости

После построения оси значимости (только нужно учитывать, что значения для данного критерия на оси значимости увеличиваются в противоположном направлении от общепринятого, таким же свойством обладают: критерий Т – Вилкоксона, U-критерий Манна - Уитни) будет видно, что наше значение $G_{мп} = 5$ попало в зону не значимости.

Полученный выше результат может быть переформулирован также в терминах нулевой и альтернативной гипотез, поскольку преобладание типичного положительного направления сдвига в данном конкретном эксперименте является случайным, то должна быть принята гипотеза H_0 об отсутствии различии или о наличии сходства. Возвращаясь к психологической задаче, укажем, что согласно критерию знаков примененный психологом способ тренинга неудовлетворителен, поскольку не дает статистически достоверных изменений в состоянии участников тренинга.

7.2. Критерий Т – Вилкоксона для связных выборок

Из непараметрических методов для по - парного сравнения удобен для пользования критерий Т - Вилкоксона, правда, на весьма небольших выборках этот критерий оказывается недостаточно мощным; его лучше применять на выборках объемом от 12 и более элементов.

Для использования этого критерия (его называют также знаково-ранговым) следует проранжировать, сначала не обращая внимания на знаки, весь перечень разностей между рядами «до» и «после». Если разность у отдельных испытуемых в отдельных случаях нулевая, то эта разность из ранжирования исключается и не входит в суммирование рангов.

Например, психологу было предложено ответить на такой вопрос: влияют ли занятия физкультурой на общее самочувствие занимающихся школьников? Исследование он построил так: школьников просили отмечать на линейной шкале свое самочувствие до занятий физкультурой и после них.

Статистической обработке подлежат по - парные сравнения — показания одного и того же испытуемого до и после воздействия. Ниже приводится таблица с результатами показаний школьников о самочувствии. Нуль-гипотеза формулируется так: сравнение рядов до и после воздействия не дает оснований утверждать, что по измеряемому признаку произошли существенные изменения.

Выборка, подвергнутая изучению, состояла из 12 человек, результаты представлены в таблице 7.2.1.

Таблица 7.2.1.

№	«До» занятий	«После»	Разность рядов «после» и «до»		Ранги R
			d	$\lvert d \rvert$	
1	3,2	3,8	+0,6	0,6	3,5
2	1,6	1,0	–0,6	0,6	3,5
3	5,7	8,4	+2,7	2,7	12
4	2,8	3,6	+0,8	0,8	6
5	5,5	5,0	–0,5	0,5	1,5
6	1,2	3,5	+2,3	2,3	11
7	6,1	7,3	+1,2	1,2	7
8	2,9	4,8	+1,9	1,9	9
9	3,1	3,8	+0,7	0,7	5
10	3,9	4,4	+0,5	0,5	1,5
11	4,2	2,1	–2,1	2,1	10
12	5,0	6,8	+1,8	1,8	8

Для того, что бы проверить правильность вычислений подсчитаем общую сумму рангов:

$$3{,}5 + 3{,}5 + 12 + 6 + 1{,}5 + 11 + 7 + 9 + 5 + 1{,}5 + 10 + 8 = 78$$

и по формуле: $\sum \mathrm{R} = \frac{n\,(n+1)}{2} = \frac{12\,(12+1)}{2} = 78$

Так как результаты совпадают, то мы проранжировали разности правильно.

Далее нужно просуммировать раздельно ранги разности с положительным знаком и ранги разностей с отрицательным знаком:

$\sum R_{+} = 63$

$\sum R_{-} = 15$

Значение критерия $T_{эм}$ равно меньшей по абсолютной величине сумме рангов. В этом примере меньшей по абсолютной величине суммой рангов является сумма рангов разностей отрицательных $T_{эм.} = 15$.

Но прежде чем отыскивать уровень значимости $T_{кр}$, нужно обратить внимание на то, что в данном случае критерий Т - Вилкоксона — это двусторонний критерий. Как это понимать? Различают односторонние и двусторонние критерии. Отвергая нуль-гипотезу, выдвигают альтернативную ей гипотезу. При этом возникает вопрос: в какую сторону направлено отличие альтернативной гипотезы от H_0 — в положительную или отрицательную. Если исследование предполагает равно возможными и ту, и другую направленность, то следует принять двусторонний критерий. Возможна, вместе с тем, такая постановка исследования, когда учитывается лишь одна направленность результатов. Так, сравнивая две выборки учащихся по освоении ими научных химических понятий, исследователь ставит ограниченную задачу — рассмотреть только возможность преобладания в этом освоении одной выборки над другой. В этом исследовании применим односторонний критерий.

При описании статистических методов всегда указывается, какого рода критерий подлежит применению — односторонний или двусторонний. В таблицах уровней значимости обычно значения для одностороннего и для двустороннего критерия даются либо в особых столбцах, либо в таблице указывается, какому значению одностороннего критерия соответствует значение двустороннего, и наоборот.

Возвращаясь к рассматриваемому примеру, следует признать, что для него при обработке с помощью критерия Т - Вилкоксона применим двусторонний критерий: различия между показателями «до» и «после» в одних строках положительные, в других отрицательные, учитываются те и другие.

В таблице (приложение 5) уровней значимости для критерия Т - Вилкоксона, находим:

$$\mathrm{T}_{\text{кр}} = \begin{cases} 17 & p = 0{,}05 \\ 9 & p = 0{,}01 \end{cases}$$

Ось значимости:

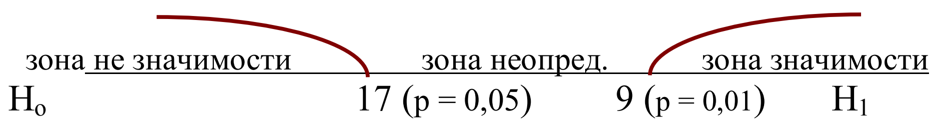

Рисунок 7.2.1. Ось значимости

Наше эмпирическое значение $\mathrm{T}_{\text{эм.}} = 15$ попадает в зону неопределенности, ближе к зоне не значимости. В этом случае, исследователь сам решает – какое решение принять. Скорее всего гипотезу H_0 (сравнение рядов до и после воздействия не дает оснований утверждать, что по измеряемому признаку произошли существенные изменения) не следует отклонять.

7.3. Сравнение независимых выборок по U-критерию Манна - Уитни

Предположим, что психологу нужно решить такую задачу: есть ли различия между выборками школьников одного и того же класса; одна выборка включает школьников, которые после контрольной работы проходили дополнительное обучение по коррекционным программам, другая — из школьников, такого обучения не проходивших. В рассматриваемом примере выборки малы, а при установлении существенной разницы между ними, т. е при отказе от H_o желательно, чтобы уровень значимости был как можно выше, но не ниже 0,05.

Формулы вычисления критерия Манна—Уитни таковы:

$$U = n_1 \cdot n_2 + \frac{n_2(n_2+1)}{2} - R_1$$

или

$$U = n_1 \cdot n_2 + \frac{n_1(n_1+1)}{2} - R_2$$

В примере сравнению подлежат результаты контрольной работы выборки А из 4-х школьников, проходивших обучение по □рите□циионным программам, и выборки Б, состоящей из 7-и школьников, никакого коррекционного обучения не проходивших Последовательность действий, предусматриваемых вычислением всех нужных для решения задачи величин, такова:

1. Выписать в любом порядке число успешно решенных заданий школьниками, сначала выборки А, затем выборки Б.

2. Проранжировать число успешно решенных заданий, объединив обе выборки.

3. Найти сумму рангов выборок А и Б раздельно.

Эти три действия дадут все необходимые для вычисления □ритерия данные. Эти данные занесём в таблицу 7.3.1.

Таблица 7.3.1.

	Выборка А (4 чел)				Выборка Б (7 чел)						
Исполнители	И	К	Л	М	Н	О	П	Р	С	Т	У
Задания	8	6	9	10	4	5	7	5	4	3	3
Ранг	9	7	10	11	3,5	5,5	8	5,5	3,5	1,5	1,5
Сумма рангов	R_A=37				R_B=29						

$n_1=4$; $n_2=7$; $N= n_1 + n_2 = 11$

Для проверки расчетов вычисляется сумма рангов по таблице и по формуле, если они равны, то операция ранжирования выполнена верно:

$$R_A + R_B = \frac{N}{2}(1+N);\ \text{т. Е. } 37+29 = \frac{11}{2}(1+11),\ \text{т. Е. } 66 = 66.$$

Для вычисления $U_{эм}$ используем формулу:

$$U = n_1 \cdot n_2 + \frac{n_2(n_2+1)}{2} - R_1$$

Выбираем большее значение объёма выборки n_2=7 и большую сумму рангов, она: $R_1 = 37$

Получим:

$$U_{эм} = 4 \cdot 7 + \frac{7(7+1)}{2} - 37 = 56 - 37 = 19$$

Критические значения $U_к$ с заданным уровнем значимости находим по таблице (приложение 6):

$$U_к = \begin{cases} 4 & p = 0{,}05 \\ 1 & p = 0{,}01 \end{cases}$$

Ось значимости:

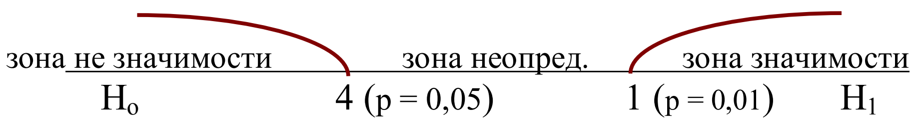

Рисунок 7.3.1. Ось значимости

$U_{эм}$ = 19 расположено в зоне значимости, следовательно принимаем гипотезу H_1 : различия между выборками школьников, которые проходили дополнительное обучение и школьников, такого обучения не проходивших, будут значимыми.

8. КОРРЕЛЯЦИОННАЯ ЗАВИСИМОСТЬ

8.1. Понятие о корреляции

В работе практического психолога часто приходится анализировать зависимость между двумя или несколькими переменными величинами (признаками).

Если две какие-либо характеристики, полученные для одного и того же «объекта», имеют тенденцию изменяться совместно, так что создается возможность предсказать одну из них по значению другой, то говорят, что эти характеристики коррелируют друг с другом. Соответственно в статистике **корреляция** выражает степень взаимосвязи между такими характеристиками. Количественно эта степень взаимосвязи выражается с помощью коэффициента корреляции.

Вообще в основном различают две формы взаимосвязи. Это функциональная зависимость, когда каждому значению независимой переменной (аргумента) соответствует определенное значение зависимой переменной (функции). Например, площадь круга $S = \pi r^2$. Вторая форма зависимости это статистическая зависимость, когда какому-либо одному фактору соответствует не одно, а несколько значений какого-либо другого фактора, причем варьирующих в каких-либо пределах. Если статистическая зависимость устанавливается между средними каких либо факторов, то такая зависимость носит название корреляционной зависимости.

При изучении корреляционных зависимостей между двумя признаками обычно решаются следующие две задачи:

1) установить формы связи между функцией У и аргументом Х (описание закона изменения величины условных средних $У_х$ в связи с изменением Х). Эта задача решается путем нахождения уравнения регрессии (с ним вы можете познакомиться в литературе по статистике);

2) оценка тесноты связи между Y и X.

Решение этой задачи требует ответов на следующие два вопроса:

1) есть ли вообще между Y и X корреляционная зависимость, то есть, наблюдается ли закономерное изменение условных средних $У_х$ в связи с изменением X?

2) если корреляционная зависимость существует, то в какой степени она отличается от функциональной зависимости?

Эту задачу и помогает решить корреляционный анализ. Корреляционная модель предполагает, что обе переменные - случайные величины, совместное распределение которых является нормальным двумерным распределением.

Виды корреляций различают по направлению. **Прямая положительная** корреляция: при увеличении причинного фактора увеличивается и следственный фактор.

Например, при увеличении количества повторений увеличивается объем запоминаемого материала.

Прямая отрицательная корреляция: уменьшение причинного фактора вызывает уменьшение следственного фактора.

Например, уменьшение физической нагрузки приводит к снижению частоты сердечных сокращений.

Обратная положительная корреляция: уменьшение причинного фактора вызывает увеличение следственного фактора (сокращение времени на обдумывание о. зета ведет к увеличению числа ошибок).

Обратная отрицательная корреляция: увеличение причинного фактора вызывает уменьшение следственного фактора (чем более опытный, тем меньше он совершает ошибок).

Коэффициент корреляции (от англ. «*correlation*» – взаимосвязь) является количественной мерой связи двух переменных. Расчет коэффициента корреляции по результатам исследования позволяет проверить гипотезу о наличии связи между интересующими признаками, оценить силу и направление связи.

Впервые идею количественного измерения связи двух признаков с помощью коэффициента корреляции предложил ученик и сотрудник лаборатории Френсиса Гальтона Карл Пирсон. Эта идея оказала чрезвычайно большое влияние на развитие психологии как эмпирической науки.

Количественно коэффициент корреляции может принимать значения от -1 до +1. Связь считается слабой, если не превышает 0,30; средней в диапазоне значений 0,31 - 0,69; сильная связь - это 0,70 - 0,99. При значении +1 или - 1 связь становится фактически функциональной. При r = 0 корреляционная связь отсутствует.

Замечание. Корреляция как метод статистического анализа в психологических исследованиях применяются очень часто. Всем, кто работает с применением корреляционного анализа, т. е. выясняет посредством этого метода тесноту связи двух рядов, следует знать, что коэффициент, как бы высок он ни был, нельзя интерпретировать как показатель наличия причинной связи между коррелируемыми рядами. Если коэффициент и может быть как-то использован в обсуждении вопроса о возможных причинных связях, то только в том случае, когда содержательная логика исследования и выдвигаемые при этом теоретические соображения позволяют опереться как на один из аргументов и на значение коэффициента корреляции. Наконец, полезно помнить, что корреляции по Пирсону (с определенными ограничениями и в определенных сочетаниях) создают ту базу, на которой открываются возможности к переходу к т. н. факторному анализу.

И так, при анализе корреляционной связи вычисляют коэффициент корреляции, для этого существуют различные способы, рассмотрим самые распространенные из них.

8.2. Коэффициент корреляции Пирсона

Коэффициент корреляции Пирсона обладает высокой степенью точности количественной характеристики связи между факторами. Однако он характеризует только наличие линейной связи между исследуемыми признаками. Вычисление данного коэффициента осуществляется только для параметрических (имеющих нормальное распределение) рядов.

В общем виде формула для подсчета коэффициента корреляции Пирсона такова:

$$r = \frac{\sum_{i=1}^{n}(X_i - \overline{X}) \cdot (Y_i - \overline{Y})}{n \cdot \sigma_x \cdot \sigma_y}$$

где X_i ,У_i сравниваемые количественные признаки;

$\overline{X}$, $\overline{Y}$ - выборочные средние арифметические; n - число сравниваемых наблюдений; σ_x, σ_y - стандартные отклонения в сопоставляемых рядах; $(X_i - \overline{X}) \cdot (Y_i - \overline{Y})$ - произведение моментов.

Имеется много различных видов этой формулы, представляющих собой ее преобразования. Исследователь сам выбирает удобную для себя формулу.

Для практических расчетов малых выборок более удобна формула:

$$r = \frac{n\sum X_i Y_i - (\sum X_i) \cdot (\sum Y_i)}{\sqrt{n(\sum X_i)^2 - (\sum X_i)^2} \cdot \sqrt{n(\sum X_i)^2 - (\sum X_i)^2}}$$

Удобство этой формулы в том, что она оперирует непосредственными исходными данными X_i и У_i.

Можно использовать формулу:

$$r = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}}$$

Полученный коэффициент корреляции, сравнивают с критическим, найденным по таблице (приложение 7), причем число степеней свободы находят как: f = n – 2, где n — объем выборки.

Пример. Группе школьников были предложены два теста с выполнением задач теста «Аналогии» и теста «Классификации». Данные тестирования представлены в двух числовых рядах. Исследователю нужно ответить на вопрос, насколько тесно связаны эти два ряда. Данные по тесту «Аналогии» обозначим «Х», а по тесту «Классификации» — «У». Для упрощения расчетов введены некоторые тождества:

$$\sum(x-\bar{x})^2 = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$\sum(y-\bar{y})^2 = \sum y^2 - \frac{(\sum y)^2}{n}$$

$$\sum(x-\bar{x})(y-\bar{y}) = \sum xy - \frac{\sum x \sum y}{n}$$

Используем эти тождества в вычислениях.

Таблица 8.2.1.

Испытуемые	x	y	x^2	y^2	xy
А	1	3	1	9	3
Б	2	4	4	16	8
В	3	5	9	25	15
Г	3	6	9	36	18
Д	4	6	16	36	24
Е	4	7	16	49	28
Ж	4	7	16	49	28
З	5	8	25	64	40
И	5	8	25	64	40
К	6	8	36	64	48
Л	6	8	36	64	48
М	7	9	49	81	63
Н	8	9	64	81	72
О	9	10	81	100	90
П	10	11	100	121	110
n=15	77	109	487	859	635

$$\sum (x-\bar{x})^2 = 487 - \frac{5929}{15} = 487 - 395{,}27 = 91{,}73$$

$$\sum (y-\bar{y})^2 = 895 - \frac{11881}{15} = 859 - 792{,}07 = 66{,}93$$

$$\sum (x-\bar{x})(y-\bar{y}) = 635 - \frac{77 \cdot 109}{15} = 653 - \frac{8393}{15} = 635 - 559{,}53 = 75{,}47$$

$$r = \frac{75{,}47}{\sqrt{91{,}73 \cdot 66{,}93}} = \frac{75{,}47}{\sqrt{6139{,}49}} = \frac{75{,}47}{78{,}72} = 0{,}96$$

Получили: $r_{эм} = 0{,}96$.

Определим уровень значимости данного результата.

Число степеней свободы k = n – 2=15 – 2 = 13.

По таблице (см. приложение 7) уровней значимости находим, что при 13 степенях свободы:

$$r_K = \begin{cases} 0{,}51 & p = 0{,}05 \\ 0{,}64 & p = 0{,}01 \end{cases}$$

Строим ось значимости:

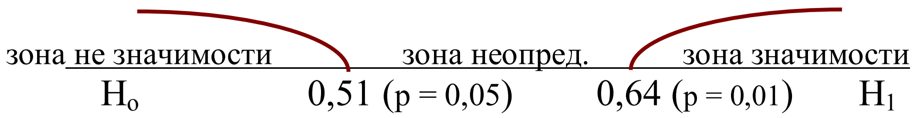

Рисунок 8.2.1. Ось значимости

Полученный коэффициент корреляции $r_{эм} = 0{,}96$ находится в зоне значимости, следовательно между результатами по тесту «Аналогии» и по тесту «Классификации» имеется связь. Высокий уровень значимости свидетельствует о том, что эта связь с высокой вероятностью будет воспроизводиться в таких же экспериментах. Причем, величина коэффициента ($r_{эм} = 0{,}96$) близка к 1, что говорит о тесной корреляционной связи, а то что коэффициент положительный показывает, что это прямая положительная корреляция: чем выше результаты по тесту «Аналогии», тем выше результаты и по тесту «Классификации», и наоборот.

Рассмотрим пример, в котором вычисления осуществляются по корреляционной таблице, составленной в другой форме. Пусть исследователю нужно установить существует ли корреляционная связь между двумя признаками Х и У (таблица 8.2.2).

Таблица 8.2.2

X_i	Y_i	$X_i - \overline{X}$	$Y_i - \overline{Y}$	$(X_i - \overline{X})^2$	$(Y_i - \overline{Y})^2$	$(X_i - \overline{X})(Y_i - \overline{Y})$
9	51	-3	1	9	1	-3
13	40	1	-10	1	100	-10
16	64	4	14	16	196	56
11	54	-1	4	1	16	-4
20	60	8	10	64	100	80
8	51	-4	1	16	1	-4
12	31	0	-19	0	361	0
9	55	-3	5	9	25	-15
7	43	-5	-7	25	49	35
15	51	3	1	9	1	3
				Σ = 150	Σ = 850	Σ = 138

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \qquad r = \frac{138}{\sqrt{150 \cdot 850}} = \frac{138}{357} = 0{,}386$$

По таблице (см. приложение 7) уровней значимости находим, что при 10 (n = 10) степенях свободы:

$$r_K = \begin{cases} 0{,}68 & p = 0{,}05 \\ 0{,}71 & p = 0{,}01 \end{cases}$$

Полученный коэффициент корреляции $r_{эм}$ = 0,386 находится в зоне не значимости. Следовательно корреляционной связи между двумя этими признаками Х и У нет.

8.3. Вычисление коэффициента корреляции по Спирмену (коэффициента ранговой корреляции)

Коэффициент корреляции, рангов предложенный К. Спирменом, относится к непараметрическим показателям связи между переменными, измеренными в ранговой шкале.

Формула рангового коэффициента линейной корреляции Спирмена такова:

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

где d – разность рангов ряда X и ряда У (R_x - R_y).

Это один из простых способов установления меры связи между факторами. Само название метода указывает на то, что связь определяется между рангами (обе переменные измеряются в шкалах порядка). Число пар должно быть больше 4, но меньше 40 ($4<n<40$). Применяется тогда, когда достаточно получить приблизительную информацию.

Этапы вычисления коэффициента ранговой корреляции Спирмена:

1. Найти ранги первого фактора (R_1), найти ранги второго фактора (R_1).
2. Найти разность между R_1 и R_2.
3. Разность возвести в квадрат $(R_1 – R_2)^2$.

4. Далее расчет по формуле $r = 1 - \frac{6\sum_{i=1}^{n}(R_1 - R_2)^2}{n(n^2 - 1)}$

5. По таблице уровней значимости (приложение 8) находим критические значения коэффициента для p=0,05 и p=0,01 по числу испытуемых n.

6. Строим ось значимости (в зависимости от предпочтений ис-

следователя, ось может быть либо однополюсной, либо двух полюсной). Принимаем статистическое решение.

Рассмотрим пример: при исследовании абстрактной памяти и конкретной памяти группы школьников, получены следующие данные (таблица 8.3.1.).

Таблица 8.3.1.

ФИО	Абстрактная память	Конкретная память	R_1	R_2	R_1-R_2	$(R_1 - R_2)^2$
Алябьев Н.	47	30	1	1	0	0
Баранов В.	44	25	2	4	2	4
Володина К.	43	27	3	2	1	1
Гурьянов Г.	42	25	4	4	0	0
Дмитриева А.	40	24	5	6	1	1
Ермаков Е.	39	21	7	7	0	0
Жарова С.	39	20	7	8	1	1
Зиновьев М.	39	14	7	10	3	9
Ильюшенко В.	37	25	9	4	5	25
Короткова Н.	35	19	10	9	1	1
n= 10					$\sum$	42

$$r = 1 - \frac{6 * 42}{10(10^2 - 1)} = 0,74$$

По таблице (приложение 8) находим критические значения коэффициента ранговой корреляции при n = 10:

$$r_K = \begin{cases} 0,64 & p = 0,05 \\ 0,79 & p = 0,01 \end{cases}$$

Строим ось значимости:

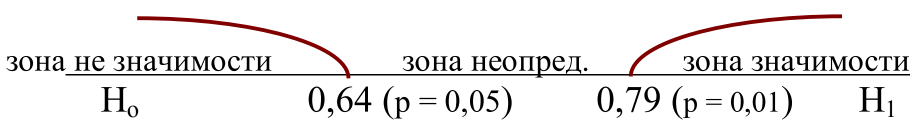

Рисунок 8.3.1. Ось значимости

Эмпирическое значение коэффициента корреляции Спирмена: $r_{эм}$= 0,74 попадает в зону неопределенности.

Таким образом, связь оказывается значимой (но только на уровне 95% значимости), прямой положительной.

Рассмотрим ещё один пример. Группе школьников были предложены два теста, задания которых были построены на материале школьных дисциплин гуманитарного цикла: литературе и истории. Но в первом тесте для выполнения заданий требовалась актуализация умственного действия аналогии, а во втором — умственного действия классификации. Данные тестирования представлены в двух числовых рядах. Исследователю нужно ответить на вопрос, насколько тесно связаны эти два ряда. При строгой постановке эксперимента это исследование должно было пролить свет на то, какую роль играют умственные действия, указанные выше, на усвоение знаний в гуманитарном цикле. Исследовалась выборка школьников объемом в 15 человек. Данные приведены в таблице 8.3.2.

Таблица 8.3.2.

Испытуемые	X	R_x	Y	R_y	R_x - R_y	d^2
А	1	1	3	1	0	0
Б	2	о	4	2	0	0
В	3	35	5	3	0,5	0,25
Г	3	35	6	4,5	1	1
Д	4	6	6	45	1,5	2,25
Е	4	6	7	6,5	0,5	0,25
Ж	4	6	7	6,5	0,5	0,25
З	5	8,5	8	9,5	1	1
И	5	8,5	8	9,5	1	1
К	6	10,5	8	9,5	1	1
Л	6	10,5	8	9,5	1	1
М	7	12	9	12,5	0,5	0,25
Н	8	13	9	12,5	0,5	0,25
О	9	14	10	14	0	0
П	10	15	11	15	0	0

n=15 $\sum = 8,5$

Производится раздельное ранжирование ряда «х» и ряда «у» Вычисляется разность рангов d попарно. Знак разности не существенен, так как по формуле нужно d возвести в квадрат. Далее действия определяются формулой:

$$r = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{51}{3360} = 1 - 0{,}2 = 0{,}98$$

По таблице уровней значимости для n=15:

$$r_K = \begin{cases} 0{,}52 & p = 0{,}05 \\ 0{,}68 & p = 0{,}01 \end{cases}$$

Эмпирическое значение коэффициента корреляции Спирмена: $r_{эм}$= 0,98 попадает в зону значимости (ось значимости постройте самостоятельно). Таким образом, связь оказывается значимой (на уровне 99% значимости), прямой положительной.

9. ПРОГНОЗ НА ОСНОВЕ ВРЕМЕННЫХ РЯДОВ

Предположим, что психологу дано задание дать информацию о состоянии умственной работоспособности школьников 8-х классов, начиная со второй недели учебного года и до 9-й недели включительно. Одной из методик, с помощью которых можно фиксировать состояние умственной работоспособности, считается тест Крепелина — он состоит из большого количества примеров, в каждом из них нужно складывать два двузначных числа; учитывается общее число правильно решенных примеров. Каждые 3 минуты испытуемые по сигналу экспериментатора отмечают черточкой сделанное. Общая длительность эксперимента в зависимости от возраста составит 9, 12 или 15 минут. Этой методикой и воспользовался психолог. Он начал с того, что сформировал из учащихся, средние успехи которых оценивались за предыдущее полугодие баллами 4 и 5, выборку из 10 человек. Все они изъявили желание участвовать в эксперименте. С этими учащимися психолог в течение первой недели учебного года провел по 12 тренировочных занятий; это было необходимо, иначе рост продуктивности вследствие упражняемости замаскировал бы изменение в динамике работоспособности. Затем начался эксперимент: по субботам после уроков учащиеся этой выборки в течение 12 минут работали с тестом Крепелина. Эксперимент, как было сказано, продолжался 8 недель. Были получены следующие данные, средние по всей выборке (Таблица 9.1).

Визуальная оценка полученного динамического ряда свидетельствует о снижении умственной работоспособности, в чем, конечно, нет ничего удивительного. Однако снижение идет не вполне равномерно.

Таблица 9.1.

Недели эксперимента	I	II	III	IV	V	VI	VII	VIII
Средняя продуктивность по тесту Крепелина	92	94	90	92	81	74	78	70
Средние по триадам		92	92	88	82	77	74	

Основная тенденция изменения умственной работоспособности вполне ясна. Наблюдаемые, в общем, незначительные отклонения от этой тенденции могут быть устранены методом сглаживания. В этом случае применим метод скользящей средней. Для сглаживания суммируются три показателя «у» — в данном примере это показатели продуктивности по тесту, — далее, опуская по одному показателю, суммируются одна за другой триады. Средняя каждой триады принимается за показатель сглаженной ломанной, если ориентироваться по графику. Смысл проводимого действия состоит в том, что основная тенденция динамики умственной работоспособности выступает более отчетливо.

Судя по показателям, полученным после сглаживания, в течение первых 3-х экспериментальных недель значительного снижения работоспособности не наблюдается, а далее идет непрерывное и резкое ее снижение. Сглаживание устранило колебания в работоспособности, отмеченные после VII недели. При сглаживании по триадам общее число точек уменьшается на 2.

Какое значение имеет выделение посредством сглаживания основной тенденции? Если условия, благодаря которым возникла основная тенденция, сохранятся, то и эта тенденция с высокой вероятностью сохранится и, таким образом, по основной тенденции может быть построен прогноз, как будут развиваться изучаемые явления в будущем. Но такой прогноз возможен только при стабильности действий определенных условий. Для его построения нужен не только формальный, но и содержательный анализ; он же позволяет раскрыть значение факторов, вызвавших отклонения в ту или другую сторону от основной тенденции.

Техника MCI ода скользящей средней дает возможность выбирать различные способы объединения показателей для сглаживания. Таковыми могут быть не только триады, но при достаточно большом числе показателей (порядка 30—40 и более) для выведения скользящей средней могут быть выбраны пентаны (объединения пяти показателей) и даже септиды (семь показателей).

Нужно иметь в виду, что наглядный и простой метод скользящей средней малопригоден для сглаживания динамики процессов, развитие которых во времени не имеет линейной формы. Сглаживание методом скользящей средней в таких случаях может привести к искажению действительной тенденции развивающегося процесса. Исследователю следует внимательно всмотреться в материал, подлежащий сглаживанию, чтобы решить, имеет ли он право воспользоваться этим методом. Если криволинейная зависимость отражена в достаточно больших отрезках ломаной, то каждый из этих отрезков в отдельности может быть подвергнут сглаживанию. Таково ограничение в использовании метода скользящей средней.

Анализируя основную тенденцию в ее приближении к прямой, можно заметить, что метод не дает меры наклона угла, который образуется между полученной после сглаживания приближающейся к прямой ломаной и осью абсцисс. Между тем, узнав этот угол, исследователь получит информацию о том, с какой скоростью изменяются изучаемые явления во времени: чем круче наклон и, соответственно, чем меньше внешний угол сглаженной кривой с осью абсцисс, тем больший путь проходит за единицу времени изменяющийся процесс.

Точные сведения о мере наклона отрезка прямой, полученного после сглаживания, дает метод наименьших квадратов.

Для нахождения параметров отрезка прямой, который после сглаживания представит основную тенденцию изменяющегося ряда, проделываются вычисления по определенным формулам.

Формула прямой: у = a + bx, где у означает показатели ряда, x — единицы времени, по которым прослеживаются изменения изучаемого ряда. Надлежит узнать величины a и b. Величина a необходима для установления точки, с которой берет свое начало отрезок прямой, b — необходимо для установления степени наклона отрезка прямой по отношению к оси абсцисс (оси иксов). Для вычисления вышеуказанных параметров a и b имеется система двух уравнений с двумя неизвестными:

$$n \cdot a + \sum x \cdot b = \sum y$$

$$\sum x \cdot a + \sum x^2 \cdot b = \sum x \cdot y$$

x и y в этой формуле рассчитываются из фактических данных изучаемого ряда.

Порядок вычислений дан в примере. Шестиклассники Саня и Толя в течение пяти дней упражнялись в бросках мяча в корзину. Показатели Сани приведены в таблице (x — единица времени, y — число попаданий мячом в корзину. В таблице приведены вычисления и других, требуемых формулой, величин. Число членов ряда n = 5).

Таблица 9.2.

x	y	x^2	xy
1	3	1	3
2	4	4	8
3	6	9	18
4	5	16	20
5	8	25	40
$\sum x = 15$	$\sum y = 26$	$\sum x^2 = 55$	$\sum xy = 89$

$$5a + 15b = 26$$

$$15a + 55b = 89$$

Нахождение неизвестных a и b производится обычным способом исключения одного неизвестного. Члены первого уравнения для этого умножаются на 3: $15a + 45b = 78$

Из второго уравнения вычитается первое:

$$10b = 11 \; ; \; b = 1,1$$

Подставив числовое значение b в первое уравнение, можно получить числовое значение a: $5a + 16,5 = 26$ $5a = 9,5;$ $a = 1,9$

Поскольку известны оба параметра отрезка прямой, можно определить все значения параметров по пяти точкам, по формуле прямой $y = a + bx$.

В этом случае $y=1,9+1,1x$.

$$y_1 = 1,9 + 1,1 = 3,0$$

$$y_2 = 1,9 + 2,2 = 4,1$$

$$y_3 = 1,9 + 3,3 = 5,2$$

$$y_4 = 1{,}9 + 4{,}4 = 6{,}3$$

$$y_5 = 1{,}9 + 5{,}5 = 7{,}4.$$

Для y_1 x был равен единице, для y_2 он равен 2 и так далее.

Как было сказано ранее, сверстник Сани Толя упражнялся в том же умении. Так же, как и у Сани, количество дней упражнение было равно 5. Ниже приводятся результаты Толи и показаны все другие величины, которые необходимы для вычисления величин, требуемых формулой.

Таблица 9.3.

x	y	x^2	xy
1	3	1	3
2	6	4	12
3	5	9	15
4	8	16	32
5	10	25	50
$\sum x = 15$	$\sum y = 32$	$\sum x^2 = 55$	$\sum xy = 112$

Обозначения здесь такие же, что и в предыдущем примере. Буквы заменяются их числовыми значениями:

$$5a + 15b = 32 \qquad 15a + 55b = 112$$

Члены первого уравнения умножаются на 3: $15a + 45b = 96$

Из второго уравнения вычитается первое: $10b = 16$; $b = 1{,}6$

Из первого уравнения получаем значение а: $5a + 24 = 32$

$$5a = 8;\ a = 1{,}6.$$

Можно получить сглаженные показатели по дням упражнений у Толи.

$$y_1 = 1{,}6 + 1{,}6 = 3{,}2$$

$$y_2 = 1{,}6 + 3{,}2 = 4{,}8$$

$$y_3 = 1{,}6 + 4{,}8 = 6{,}4$$

$$y_4 = 1{,}6 + 6{,}4 = 8{,}0$$

$$y_5 = 1{,}6 + 8{,}0 = 9{,}6.$$

Если перенести на график результаты сглаживания, то следует обратить внимание на то, как различаются отрезки прямой по их наклону по отношению к оси абсцисс.

10. КОНСТРУИРОВАНИЕ ДИАГНОСТИЧЕСКИХ МЕТОДИК

Американская психологическая ассоциация (АПА) периодически издает «Стандартные требования к педагогическим и психологическим тестам», специальный кодекс требований к диагностическим методикам; это пособие нужно как авторам методик, так и тем, кто методиками пользуется.

Некоторые из этих требований могут считаться дискуссионными, но полезность кодекса в целом несомненна. Его выполнение, с одной стороны, обеспечивает объективность методик и их обоснованность, а с другой — препятствует проникновению в арсенал методик психологической диагностики дилетантских поделок, произвольных наборов всевозможных заданий, заимствованных из популярных журналов или сочиненных самим автором. Самые общие и самые необходимые к исполнению требования можно было бы свести всего к двум: диагностические методики должны быть надежными и валидными. Реализация этих требований осуществляется посредством прочно вошедших в психологическую диагностику статистических методов.

Чтобы получить коэффициент надежности, характеризующий гомогенность методики, ее внутреннюю согласованность, прибегают к приему, называемому расщеплением методики. Эксперимент проводится с выборкой, желательно, порядка 100, по не менее 50 испытуемых.

Полученные от каждого участника выборки ответы на вопросы или решения заданий делятся на четные и нечетные— по их нумерации в методике. По каждой половинке методики выписывается число правильно выполненных каждым испытуемым заданий. Два эти ряда коррелируют между собой.

Допустим, что методика состоит из 24 заданий. Тогда максимальное число выполненных заданий в каждой половинке будет равно 12. В нижеследующей таблице 30 представлены результаты первых 16 испытуемых и показана техника вычисления коэффициента надежности (**гомогенности**) – r.

Таблица 10.1.

Вычисление коэффициента надежности методики А (гомогенность)

Имена испытуемых	Правильно решено		Ранг четных задании	Ранг нечетных заданий	d	d^2
	четных заданий	нечетных заданий				
А	10	И	10,5	13,5	3	9
Б	8	8	8	8,5	0,5	0,25
В	3	7	3	6,5	3,5	12,25
Г	3	3	3	2	1	1
Д	11	12	12,5	15,5	3	9
Е	12	10	15	11	4	16
Ж	12	12	15	15,5	0,5	0,25
З	9	8	9	8,5	0,5	0,25
И	7	7	6,5	6.5	0	0
К	6	6	6	6	0	0
Л	7,	5	6,5	4	2,5	0,25
М	11	10	12,5	11	1,5	2,25
Н	3	4	3	3	1	1
О	2	2	1	1	0	0
П	10	11	10,5	13,5	3	9
Р	12	10	15	11	4	16

$\sum d^2 = 82{,}5$

$$n = 16\ ;\ r = 1 - \frac{6 \cdot \sum d^2}{n(n^2 - 1)} = 1 - \frac{6 \cdot 82{,}5}{16 \cdot 255} = 1 - \frac{495}{4080} = 1 - 012 = 0{,}88$$

Проделана обычная ранговая корреляция. По таблице уровней значимости $r_{0,99} = 0{,}64$; полученный коэффициент превышает эту величину. Принято считать, что коэффициент надежности не должен быть ниже

0,8. Полученный коэффициент удовлетворяет этому требованию.

Есть поправочная формула Спирмена—Брауна (SB) к коэффициенту надежности - гомогенности, получаемому путем расщепления. Поскольку при прочих равных условиях получаемый коэффициент будет тем выше, чем больше заданий содержится в методике, следует принять во внимание, что прием расщепления уменьшает число заданий вдвое — в этом и состоит этот прием. Поправочная формула такова:

$$r_{SB} = \frac{2 \cdot r_{\frac{1}{2}\frac{I}{II}}}{1 + r_{\frac{1}{2}\frac{I}{II}}}$$ в нашем примере $r_{SB} = \frac{2 \cdot 0{,}88}{1 + 0{,}88} = \frac{1{,}76}{1{,}88} = 0{,}94$

r_{SB} – коэффициент с учетом поправки, а $r_{\frac{1}{2}\frac{I}{II}}$ – коэффициент, вычисленный при коррелировании двух половинок методики. Если этот последний равен 0,88, то после поправки Спирмена—Брауна коэффициент будет равен 0,94.

Поправочную формулу Спирмена—Брауна можно применять только в тех случаях, когда методика делится на половинки (расщепление). Если же в методике в процессе обработки не меняют числа заданий, то поправочная формула неприменима.

Величина коэффициента надежности - гомогенности зависит от социально-психологических особенностей той выборки, по результатам испытания которой этот коэффициент устанавливался. Поэтому при опубликовании методики, приводя ее основные характеристики, автор должен указать, на каком контингенте проводилась проверка надежности.

При вычислении коэффициента надежности методики, характеризующего стабильность данных, получаемых с помощью этой методики, первый коррелируемый ряд представляет собой результаты первого, а второй — повторного испытания: ею рекомендуют проводить примерно через шесть недель после первого. При обоснованной необходимости этот срок может изменяться. Эти два ряда коррелируют между собой. Корреляция проводится по обычным правилам. Это прием «тест-ретест».

Для установления надежности методики существуют и некоторые другие приемы. Так, для получения коэффициента надежности практикуется прием параллельных форм. Авторы, конструирующие методику, создают две ее формы; условно назовем их формой А и формой Б. Обе формы должны быть вполне однородны по психологической направленности, по доступности содержания заданий и по их трудности. В одном варианте формы А и Б предъявляются испытуемым одна следом за другой, причем в одной половине выборки испытуемым сначала предлагается форма А, а за ней форма Б, а в другой половине выборки, наоборот, сначала форма Б, а затем А. Результаты, полученные по той и другой форме, коррелируют между собой, и полученный коэффициент трактуется как коэффициент надежности. Нетрудно заметить, что этот прием близок приему расщепления, с той разницей, что методика как бы удвоена и сравниваются не четные и нечетные задания, а две половины этой удвоенной методики. Это дает право скорее трактовать получаемый коэффициент как коэффициент надежности - гомогенности, а не надежности-стабильности. Поскольку проверке подвергается набор заданий в целом, поправочную формулу Спирмена — Брауна применять не следует.

Другой вариант использования приема параллельных форм состоит в том, что одна из форм предлагается испытуемым через какой-то интервал времени после другой, что сближает этот прием с приемом «тест - ретест». При проведении этого приема необходимо убедиться в том, что обе формы высоко коррелируют между собой согласно только что изложенному приему по надежности - гомогенности. Результаты обоих испытаний затем коррелируют между собой. Полученный коэффициент может трактоваться как коэффициент надежности - стабильности. Выше указывалось, что в приеме «тест - ретест» рекомендуется интервал между испытаниями в шесть недель. Для этого варианта приема параллельных форм этот интервал может быть уменьшен, так как испытуемый при выполнении заданий не сможет опираться на память.

Из предшествующего изложения явствует, что в приемах установления надежности главную роль играет статистический метод корреляций. Несколько по-иному обстоят дела при проверке валидности методики.

Если показатели того критерия, который взят для получения коэффициента внешней валидности, имеют примерно ту же меру рассеяния, меру вариативности, что и мера рассеяния показателей самой методики, то применение корреляции правомерно. Допустим, автор методики намерен установить ее валидность, сравнивая успешность выполнения методики с учебной деятельностью. Валидность устанавливается на выборке школьников. В этом случае, как показывает практика, суммарные оценки за одну учебную четверть или за полугодие покажут примерно тот же размах колебаний, что и размах колебаний по методике; методика состоит из 20 заданий, и при ее выполнении показан размах колебаний от 3 до 20. Суммарные оценки успеваемости после того, как они подсчитаны за полгода, имеют размах колебаний от 14 до 36. Такие ряды вполне возможно между собой коррелировать.

Но в некоторых случаях для получения коэффициента валидности приходится сравнивать успешность выполнения диагностической методики, допустим, в тех же пределах колебаний — от 3 до 20, и производственные достижения, которые имеют всего три ступени оценок: ниже средних, средние и выше средних. Корреляцией в этом случае воспользоваться нельзя, если иметь в виду линейную корреляцию, о которой идет речь в этом пособии. Однако могут быть использованы некоторые другие статистические методы, показывающие существование или отсутствие связи между распределением двух рядов численностей. Простейший способ получения коэффициента валидности в описываемом случае и в других подобных случаях — метод «хи» - квадрат Пирсона. Всех испытуемых, прошедших диагностический эксперимент, делят на три равные группы — их и сопоставляют с тремя группами, на которые были поделены испытуемые при оценке их профессиональной успеваемости. В изучаемой выборке — 90 человек. Они делятся по профессиональным достижениям на три группы: первая — в ней 30 испытуемых — лица с профессиональными достижениями ниже среднего уровня; вторая — в ней 40 испытуемых — это лица со средними достижение и третья — в ней 20 испытуемых, их достижения выше среднего уровня. Первая группа составляет 33,3% выборки, вторая — 44,4% и третья — 22,2% выборки. Ниже показана техника вычисления (см. таблица 10.2).

Эксперимент, данные которого представлены в этой таблице-графике, предпринимался, чтобы установить валидность психологической оценки. Нуль-гипотеза формулируется так: психологическая оценка не имеет никакого значения для профессиональных достижений; поэтому она никак не скажется на распределении численностей в таблице-графике χ^2-квадрат Пирсона.

Принятие нуль - гипотезы может произойти в том случае, если в каждой из групп по профессиональной успешности испытуемые будут распределены независимо от их психологической оценки. Тогда испытуемые, получившие психологическую оценку ниже среднего, распределятся по всем трем группам в тех же процентных отношениях, в каких они распределились и по профессиональным достижениям. Напомним эти отношения: 33,3 – 44,4 – 22,2 Психологическую оценку «ниже среднего» получили всего 30 испытуемых. 33,3% этого числа должны были бы попасть в группу с профессиональными достижениями ниже среднего уровня. 33,3% от 30 составляет 10. В группу с достижениями среднего уровня должны попасть 44,4% от 30, что составляет 13,3. Наконец, в группу с достижениями выше среднего уровня должны были бы попасть 22,2%, это составляет 6,7% от 30.

Таблица 10.2.

	Профессиональные достижения			
	Ниже среднего	Средние	Выше среднего	Итого
Психологические оценки	A 20 (10)	B 5 (13,3)	C 5 (6,7)	30
	D 5 (10)	E 15 (13,3)	F 10 (6,7)	30
	G 5 (10)	H 20 (13,3)	J 5 (6,7)	30
Итого	30	40	20	90

Те же рассуждения повторяются и относительно испытуемых, имеющих психологическую оценку «среднюю» и относительно испытуемых, имеющих психологическую оценку «выше среднего». Однако фактически наблюдается не такое распределение. Возникает вопрос: можно ли, учитывая это фактическое распределение, отвергнуть нуль-гипотезу и признать, что психологическая оценка влияет на профессиональные достижения? Это раскроет методика χ^2-квадрат Пирсона.

В клетках таблицы графика представлены как фактически наблюдаемые численности, так и предполагаемые согласно нуль - гипотезе; они заключены в скобки.

Как известно, формула методики χ^2- квадрат Пирсона такова:

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e}$$

где f_0 — фактически наблюденные численности,

f_e — предполагаемые численности.

Для получения значения χ^2- квадрат нужно просуммировать по клеткам:

$$\frac{(f_0 - f_e)^2}{f_e}$$

Клетки

A: $\frac{(20-10)^2}{10} = \frac{10^2}{10} = 10$

B: $\frac{(5-13,3)^2}{13,3} = \frac{8,2^2}{13,3} = \frac{68,89}{13,3} = 5,2$

C: $\frac{(5-6,7)^2}{6,7} = \frac{-1,7^2}{6,7} = \frac{2,89}{6,7} = 0,4$

D: $\frac{(5-10)^2}{10} = \frac{-5^2}{10} = \frac{25}{10} = 2,5$

E: $\frac{(15-13,3)^2}{13,3} = \frac{1,7^2}{13,3} = \frac{2,89}{13,3} = 0,2$

F: $\frac{(10-6,7)^2}{6,7} = \frac{3,3^2}{6,7} = \frac{10,89}{6,7} = 1,6$

G: $\frac{(5-10)^2}{10} = \frac{-5^2}{10} = \frac{25}{10} = 2,5$

H: $\frac{(20-13,3)^2}{13,3} = \frac{6,7^2}{13,3} = \frac{44,89}{13,3} = 3,7$

J: $\frac{(5-6,7)^2}{6,7} = \frac{-1,7^2}{6,7} = \frac{2,89}{6,7} = 0,4$

$$\chi^2 = 10 + 5,2 + 0,4 + 2,5 + 0,2 + 1,6 + 2,5 + 3,7 + 0,4 = 26,5$$

fd – число степеней свободы.

В этом примере fd= (к— 1) (с— 1) = (3—1) (3— 1) = 4

$\chi^2_{0,99}$ при 4-х степенях свободы равен 11,34.

Сравнивая полученную в эксперименте величину χ^2 с величиной $\chi^2_{0,99}$, указанной в таблице значимостей, можно заключить: полученная в эксперименте величина (χ^2 = 26,5) свидетельствует о валидности примененной психологической методики.

Величина χ^2 - квадрат с указанием ее значимости служит в подобных случаях показателем или коэффициентом валидности. Этот же метод применим, если оценка дается не по трем ступеням, как в рассмотренном примере, а по пяти (значительно ниже средней, ниже средней, средняя, выше средней, значительно выше средней и т. д.). Техника вычислений при такой дифференциации оценок аналогична показанной выше.

В современной диагностике применяются не только перечисленные в этом разделе вычислительные методы, но и многие другие. Однако можно полагать, что, ограничив свою цель изложением простейших статистических методов, нет необходимости обращаться к сложным и сложнейшим. Читатели, заинтересовавшиеся проблемами статистических методов в диагностике, могут обратиться к другим пособиям и источникам.

11. ЭЛЕМЕНТЫ ПЛАНИРОВАНИЯ В ПСИХОЛОГИЧЕСКИХ И ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

Нельзя начинать исследование, не уяснив его цель,— это аксиома. Однако наблюдения показывают, что не все ее принимают. Нередко можно обнаружить смещение двух категорий целей: цель исследования и цель исследователя. Но полное доминирование цели исследователя и безразличное отношение к цели исследования не должно иметь места. Планирование должно исходить из цели исследования.

Есть два главных источника, стимулирующих возникновение исследований: либо они отвечают на запросы, выдвигаемые практикой, которую обслуживает данная наука, либо они возникают из нужд самой науки и имеют целью совершенствовать познание тех сфер жизни, которым посвящена данная наука. Стоит отметить, что детальное планирование необходимо и в том, и в другом случае. Мнение, будто практические исследования могут делаться без заранее продуманного плана, безусловно, ошибочно, так как только правильно спланированное исследование может в своих выводах дать ответ на те вопросы, ради решения которых оно и задумывалось

Различают планирование исследований, не нуждающихся в эксперименте, и исследований, включающих эксперимент как необходимую свою часть. Что касается первых, то их планы, в принципе, не отличаются от планов исследований в других науках. В вводной части (она будет примерно такой же и в экспериментальных исследованиях) очерчивается место данного исследования в потоке современной науки, кратко реферируются работы, затрагивающие ту же проблематику, указываются источники и формулируется замысел исследования и его цель. Далее планируется само исследование. Все без исключения исследования вообще могут рассматриваться как система доказательств, обосновывающих выводы, в которых содержится и цель, поставленная автором.

Этот план не должен рассматриваться как обязательный. Особенности работы могут заставить автора в той или иной степени отойти от него, дополнить его или сократить. В исследованиях, включающих эксперимент, во вводной части должно быть показано, зачем оказался нужным эксперимент и каковы принципы его построения.

Планирование эксперимента в психологическом исследовании предполагает предварительное обсуждение следующих пяти пунктов:

А. Каков планируемый объект эксперимента, другими словами, какова та выборка испытуемых, которых намерен привлечь автор? В зависимости от того, каких испытуемых возьмет автор, ему придется обдумать и следующий пункт.

Б. Если необходимо работать со школьниками, то эксперимент должен быть согласован со школьными режимами— годовым, еженедельным и ежедневным, с учетом умственной нагрузки школьников. Необходимо считаться и с периодом подготовки к экзаменам и их сдачей. С первыми двумя пунктами тесно связан третий пункт.

В. Нужны методики, которые, с одной стороны, учитывали бы особенности исследуемого контингента, а с другой, — непосредственно вели бы к цели исследования. Поскольку намечены методики и время их проведения, возникает следующий пункт плана.

Г. Материалы эксперимента нуждаются в адекватной обработке и почти всегда в привлечении статистики. Планируются такие статистические методы, результаты которых непосредственно направлены на достижение цели исследования. Все перечисленные пункты подготавливают планирование последнего пункта.

Д. Сколько и какой квалификации работников нужно для проведения эксперимента, какая понадобится аппаратура и каких средств потребует эксперимент?

Исследователь, обдумывая предстоящий эксперимент, должен иметь в виду, что полученные выводы будут относиться не только к выборке испытуемых, непосредственно участвующих в эксперименте, но и к той совокупности, к которой принадлежит эта выборка. Чтобы этот расчет оправдался, нужно с достаточной определенностью представить, что же это за совокупность. Поэтому важно вести эксперимент не со случайным набором испытуемых, а с испытуемыми, образующими репрезентативную выборку, воспроизводящую все характерные психологические признаки совокупности. С этих же позиций репрезентативности нужно рассмотреть вопрос об объеме выборки. Не всегда целесообразно планировать участие большой выборки в несколько сотен или тысяч испытуемых — в такой выборке почти неизбежно утратится репрезентативность, в ней, возможно, будет представлено несколько совокупностей, каждая из которых так или иначе повлияет на результаты эксперимента: их интерпретация потеряет ясность. Поэтому предпочтительнее работать с малыми и средними выборками, объемом до 30—50—100 испытуемых. Чтобы решить, сколько же конкретно следует взять участников эксперимента, придется провести пилотажный или подготовительный мини-эксперимент. Проведение такого эксперимента поможет выявить два необходимых момента: гомогенность выборки, ее сравнительно малую вариативность по тем признакам, которые, при прочих равных условиях, изучаются в эксперименте, и такой ее объем, который обеспечит получение всех показателей как внутри выборки, так и в ее сопоставлениях на должном уровне статистической значимости. О последнем моменте свидетельствует следующее наблюдение: допустим, что в пилотажном эксперименте на выборке объемом в 10 испытуемых получен коэффициент корреляции между двумя признаками, равный 0,60. Этот коэффициент свидетельствует о том, что коррелируемые ряды связаны между собой. Однако этот коэффициент ниже уровня 0,95 значимости, который принят в психологических исследованиях. При увеличении выборки до 12 человек коэффициент окажется на приемлемом уровне значимости — несколько выше коэффициента общепринятого уровня, а он равен 0,576. Вывод, который придется сделать исследователю: выборка должна состоять не из 10 испытуемых, а минимум из 12—15. Этот объем позволит получить значимый коэффициент. Но определить объем выборки без пилотажного эксперимента не представляется возможным. Если автор претендует на более высокий уровень значимости, то по таблице уровней значимости он установит и объем выборки. Чем выше гомогенность выборки, тем яснее ее отнесенность к той или другой совокупности. Вместе с тем, высокая гомогенность может рассматриваться как предпосылка того, что желательные уровни статистической значимости действительно могут быть достигнуты с увеличением выборки.

При планировании эксперимента исследователю надлежит обратить внимание на то, чтобы в подборе испытуемых для своей выборки он избежал вольных или невольных ошибок, порождаемых стремлением работать с выборкой, обеспечивающей получение желательных результатов. Надежным заслоном против таких ошибок является обращение к таблице случайных чисел. Так, исследователю предстоит отобрать из двух классов одну выборку: число учеников в обоих классах составляет 60 человек, а выборку исследователь намерен составить из 15 человек. Возможно, что ему посоветуют взять лучших, или дисциплинированных, или усердных и т. п. Но те признаки, которыми советуют руководствоваться исследователю, несущественны для его цели; допустим, что он намерен изучить наиболее яркие проявления гуманитарных способностей. Чем руководствоваться исследователю при отборе испытуемых в свою выборку? Ему следует обратиться к таблице случайных чисел.

Чтобы воспользоваться этой таблицей, сначала нужно выписать подряд, одну за другой, в любой последовательности фамилии учеников, из числа которых исследователь намерен образовать нужную ему выборку. Далее, открыв таблицу случайных чисел на любой странице, следует взять, например, два первых двузначных числа из любого из десяти столбцов, напечатанных на этой странице. Идя сверху вниз, нужно последовательно приписывать эти двузначные числа к фамилиям учеников. В выборку попадут ученики, к чьим фамилиям будут приписаны первые пятнадцать чисел, начиная с наименьшего. Исследователь волен взять не первые два числа, а два последних или два средних и идти не сверху вниз, а снизу вверх. Необходимо только сохранять тот порядок, который был избран для работы с таблицей случайных чисел в данном конкретном исследовании.

Ниже приведен фрагмент одной из страниц таблицы случайных чисел.

5489 5583 3156 0835 1988 3912 0938 7460 0869 4420
3522 0935 7877 5665 7020 9555 7379 7124 7878 5544
7555 7579 2550 2487 9477 0864 2349 1012 8250 2633
5759 3564 5080 9074 7001 6249 3294 6368 9102 2672

и т. д.

Допустим, исследователь решил, идя сверху вниз, воспользоваться первыми двумя числами третьего столбца. Тогда идущий первым по порядку ученик получит приписанное к своей фамилии число 31, второй по порядку — число 78, третий — число 25, четвертый — 50 и далее, следуя вниз по столбцу. После того, как числа будут приписаны всем 60 ученикам, будут отобраны те, кто получил первые по порядку 15 чисел. Эта несложная процедура исключит произвольность в отборе испытуемых.

Рекомендации, содержащиеся выше, помогут спланировать пилотажный эксперимент, а затем и исследование в его окончательном варианте.

Такое построение работы поможет сэкономить силы, средства и время и, в конечном счете, прийти к поставленной цели, либо доказав и подтвердив гипотезу автора, либо отказавшись от нее. В том и другом случае прояснится дальнейший путь развития исследований, уточняющих и углубляющих разработку проблемы.

Дело, однако, не только в этом. Неточно спланированное исследование, сколько бы сил в него ни вложили, вряд ли продвинет вперед науку и вряд ли поможет практике. Всегда останется сомнение в действенности его выводов. А это приведет к тому, что возникнет необходимость в новых, тождественных по целям исследованиях, станут вероятными противоречивые выводы.

Поэтому умение планировать экспериментальное исследование составляет важное и необходимое звено в профессиональной подготовке и конечной квалификации психолога.

ЗАКЛЮЧЕНИЕ

Мы рассмотрели основные статистические понятия и методы математической обработки результатов психологических наблюдений. Приведенные статистические критерии не являются единственными при количественном анализе эмпирического материала. Разумеется, что современная наука в области математической статистики создала большой арсенал методов, которыми может пользоваться психолог-исследователь. Однако описание других, более сложных статистических методов, требующих глубокой математической подготовки и умении в использовании компьютерных программ, не входило в нашу задачу.

Надеемся, что данное пособие поможет начинающим психологам избежать некоторых ошибок при применении методов математической статистики и уменьшит их неприязнь к этим методам.

ГЛОССАРИЙ

Генеральная совокупность – это совокупность всех субъектов, обладающих исследуемым признаком.

Выборочная совокупность представляет собой выборку из генеральной совокупности меньшим объемом.

Объем выборки – это число исследуемых в выборке.

Размах вариационного ряда – разность между наибольшим значением варианты и наименьшим значением варианты.

Мода ***(*** m_0 ***)*** – значение варианты, частота или относительная частота которой имеет наибольшее значение.

Медиана ***(*** m_e ***)*** - варианта особого ряда, которая делит вариационный ряд пополам. Если число вариант нечетное, то медиана равна центральной варианте. Если число вариант четное, то медиана равна среднему арифметическому двух центральных.

Ранжирование вариационного ряда - это присваивание порядкового номера значению варианты в порядке возрастания.

Графическое представление результатов. Графическое представление результатов осуществляется при помощи полигона и гистограммы.

Полигон – ломанная кривая, соединяющая точки, соответствующие значению варианты и ее частоте.

Гистограмма **–** ступенчатая фигура из прямоугольников, высота которых равна либо частоте, либо относительной частоте, либо относительной частоте, деленной на шаг: ширина равна шагу варьируемого признака.

Выборочная средняя – это величина равная сумме произведений значений варианты на соответствующие им относительные частоты.

Дисперсия **–** это сумма произведений квадратов отклонений значений варианты на её относительные частоты

Среднее квадратическое отклонение или нормальное отклонение - это отклонение значения варьирующего признака от среднего значения. Оно равно корню квадратному из дисперсии.

Параметрические способы обработки результатов применяются к величинам, имеющим нормальное распределение.

Статистическая гипотеза – формальное предположение о том, что сходство или различие некоторых параметрических характеристик генеральной совокупности или совокупностей случайно или неслучайно. При проверке статистических гипотез используются два понятия: H_o – нулевая гипотеза, это гипотеза о сходстве; H_1 – альтернативная гипотеза - гипотеза о различии.

Уровень значимости – это вероятность ошибки первого рода при принятии решения. Для обозначения уровня значимости используют Р = 0,05, Р= 0,01, Р = 0,001.

Этапы принятия статистического решения

1. формулировка H_o и H_1;
2. определение объема выборки;
3. выбор соответствующего уровня значимости (≈ 0,05, 0,01, 0,001);
4. выбор статистического метода, который зависит от типа решаемой психологической задачи;
5. вычисление соответствующего эмпирического значения по экспериментальным данным, согласно выбранному статистическому методу;
6. нахождение по таблице для выбранного статистического метода критических значений, соответствующих уровню значимости Р = 0,05 и Р=0,01;
7. построение оси значимости и нанесение на нее табличных критических значений и эмпирических значений измеряемой величины;
8. формулировка принятия решения (принятие гипотезы).

Классификация психологических задач, решаемых с помощью статистических методов

В общем их можно разделить на 4 группы:

1. Задачи, требующие установления сходства или различия.
2. Задачи, требующие группировки и классификации.
3. Задачи, ставящие целью анализ источников вариативности, получаемых психологических признаков.
4. Задачи, предполагающие возможность прогноза на основе имеющихся данных.

Критерий χ^2 - Пирсона, один из наиболее часто используемых в психолого-педагогических исследованиях, поскольку он позволяет решать большое число различных задач. Этот критерий используется в двух вариантах:

1. Как расчет согласия эмпирического значения и предполагаемого теоретического. В этом случае проверяется H_o об отсутствии различий между теоретическим и эмпирическим распределением.
2. Как расчет однородности двух независимых экспериментальных выборок. В этом случае проверяется гипотеза H_o об отсутствии различий между двумя эмпирическими распределениями.

Ограничения критерия χ^2- Пирсона:

1. Объем выборки должен быть достаточно большим $N \geq 30$. При $N < 30$ критерий дает весьма приближенные значения. Точность повышается с ростом N.
2. Теоретическая частота для каждой ячейки таблицы не должна быть меньше пяти.
3. Выбранные разряды должны вычерпывать все распределения, то есть охватывать весь диапазон вариативности признаков. Группировка на разряды должна быть одинаковой во всех сопоставляемых распределениях.

4. Необходимо вносить поправку на непрерывность при сопоставлении распределений признаков, которые принимают два значения. При внесении поправки значение χ^2 уменьшается.
5. Разряды должны быть неперекрещивающиеся. Если наблюдение отнесено к одному разряду. То оно уже не может быть отнесено не к одному другому разряду.

Рекомендации к выбору критерия различия

1. Прежде всего, следует определить, является ли выборка связной (зависимой) или несвязной (независимой).
2. Следует определить однородность и неоднородность выборки.
3. Затем следует оценить объем выборки и, зная ограничения каждого критерия по объему, выбрать соответствующий критерий.
4. В первую очередь выбирают наименее трудоемкий критерий.
5. Если используемый критерий не выявил различия, следует применять более мощный, но и более трудоемкий критерий.
6. Если в распоряжении психолога имеется несколько критериев, то следует выбирать те из них, которые наиболее полно используют информацию содержащуюся в экспериментальных данных.
7. При малом объеме выборки следует увеличивать величину уровня значимости не менее 1%, т.к. небольшая выборка и низкий уровень значимости приводит к увеличению вероятности принятия ошибочных решений.

Функциональная зависимость – это когда значению одной совокупности ставится в соответствие единственное значение другой совокупности.

Статистическая зависимость – когда значению одной совокупности ставится в соответствие среднее значение другой совокупности.

Корреляционной называется статистическая зависимость между значениями или средними признаками одной выборки и значениями (средними) другой выборки.

Линейная корреляция описывается уравнениями: у = ρх + b,   у – у = r (х - х). Эти уравнения носят названия *уравнения регрессии (*r – коэффициент регрессии), ρ и b – коэффициенты, которые называются параметрами линейной регрессии:

$$\rho = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} \qquad b = \frac{\sum x^2 \sum y - \sum x \sum xy}{n \sum x^2 - (\sum x)^2}$$

Коэффициент корреляции Пирсона устанавливает тесноту связи линейного характера между признаками:

$$r = \frac{\sum (x - \bar{x}) \sum (y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

Знак коэффициента корреляции очень важен для интерпретации полученной связи. Если знак «+», то большему значению одного признаками соответствует большее значение другого признака. Такая зависимость еще называется ***прямопропорциональной.***

Если полученный знак « - », то наоборот, большей величине одного признака соответствует меньшая величина другого признака. Соответственно получается ***обратнопропорциональная зависимость.***

Приложение 1

Таблица значений малой функции Лапласа

X	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	989	989	988	986	984	982	980	977	973
0,1	0,3970	965	961	956	951	945	939	932	925	918
0,2	0,3910	902	894	885	876	867	857	847	836	825
0,3	0,3814	802	790	778	765	752	739	726	712	697
0,4	0,3683	668	653	637	621	605	589	572	555	538
0,5	0,3521	503	485	467	448	429	410	391	372	352
0,6	0,3332	312	292	271	251	230	209	187	166	144
0,7	0,3123	101	079	056	034	011	2989	2966	2943	2920
0,8	0,2897	874	850	827	803	780	756	732	709	685
0,9	0,2661	637	613	589	565	541	516	492	468	444
1,0	0,2420	396	371	347	323	299	275	251	227	203
1,1	0,2179	155	131	107	083	059	036	012	1989	1965
1,2	0,1942	919	895	872	849	826	804	781	758	736
1,3	0,1714	691	669	647	626	604	582	561	539	518
1,4	0,1497	476	456	435	415	394	374	354	334	315
1,5	0,1295	276	257	238	219	200	182	163	145	127
1,6	0,1109	092	074	057	040	023	006	0989	0973	0957
1,7	0,0941	925	909	893	878	863	848	833	818	804
1,8	0,0790	775	761	748	734	721	707	694	681	669
1,9	0,0655	644	632	620	608	596	584	573	562	551
2,0	0,0541	529	519	508	498	488	478	468	459	449
2,1	0,0440	431	422	413	404	396	387	379	371	363
2,2	0,0355	347	339	332	325	317	310	303	297	290
2,3	0,0283	277	270	264	258	252	246	241	235	229
2,4	0,0224	210	213	208	203	198	194	189	184	180
2,5	0,0175	171	167	163	158	154	151	147	143	139
2,6	0,0136	132	129	126	122	119	116	113	110	107
2,7	0,0104	101	099	096	093	091	088	086	084	081
2,8	0,0079	077	075	073	071	069	067	065	063	061
2,9	0,0060	058	056	055	053	051	050	048	047	046
3,0	0,0044	043	042	040	039	038	037	036	035	034
3,1	0,0033	032	031	030	029	028	027	026	025	025
3,2	0,0024	023	022	022	021	020	020	019	018	018
3,3	0,0017	017	016	016	015	015	014	014	013	013
3,4	0,0012	012	012	011	011	010	010	010	009	009
3,5	0,0009	008	008	008	008	007	007	007	007	006
3,6	0,0006	006	006	006	005	005	005	005	005	004
3,7	0,0004	004	004	004	004	004	003	003	003	003
3,8	0,0003	003	003	003	003	002	002	002	002	002

Приложение 2

Таблица значений критерия χ^2 - Пирсона

Число степеней свободы	Уровень значимости		Число степеней свободы	Уровень значимости	
	0,05	**0,01**		**0,05**	**0,01**
1	3,841	6,635	32	46,194	53,486
2	5,991	9,210	33	47,400	54,776
3	7,815	11,345	34	48,602	56,061
4	9,488	13,277	35	49,802	57,342
5	11,070	15,086	36	50,998	58,619
6	12,592	16,812	37	52,192	59,892
7	14,067	18,475	38	53,384	61,162
8	15,507	20,090	39	54,572	62,428
9	16,919	21,666	40	55,758	63,691
10	18,307	23,209	41	56,942	64,950
11	19,675	24,725	42	58,124	66,206
12	21,026	26,217	43	59,304	67,459
13	22,362	27,688	44	60,481	68,709
14	23,685	29,141	45	61,656	69,957
15	24,996	30,578	46	62,830	71,201
16	26,296	32,000	47	64,001	72,443
17	27,587	33,409	48	65,171	73,683
18	28,869	34,805	49	66,339	74,919
19	30,144	36,191	50	67,505	76,154
20	31,410	37,566	51	68,669	77,386
21	32,671	38,932	52	69,832	78,616
22	33,924	40,289	53	70,993	79,843
23	35,172	41,638	54	72,153	81,069
24	36,415	42,980	55	73,311	82,292
25	37,652	44,314	56	74,468	83,513
26	38,885	45,642	57	75,624	84,733
27	40,113	46,963	58	76,778	85,950
28	41,337	48,278	59	77,931	87,166
29	42,557	49,588	60	79,082	88,379
30	43,773	50,892	61	80,232	89,591
31	44,985	52,191	62	81,381	90,802

Приложение 3

Таблица значимости t-распределения Стьюдента

Число степеней свободы	Уровни значимости			Число степеней свободы	Уровни значимости		
	0,05	**0,01**	**0,001**		**0,05**	**0,01**	**0,001**
1	12,71	63,66	636,6	17	2,110	2,898	3,965
2	4,303	9,925	31,60	18	2,101	2,878	3,922
3	3,182	5,841	12,92	19	2,093	2,861	3,883
4	2,776	4,604	8,610	20	2,086	2,845	3,850
5	2,571	4,032	6,869	21	2,080	2,831	3,819
6	2,447	3,707	5,959	22	2,074	2,819	3,792
7	2,365	3,499	5,408	23	2,069	2,807	3,767
8	2,306	3,355	5,041	24	2,064	2,797	3,745
9	2,262	3,250	4,781	25	2,060	2,787	3,725
10	2,228	3,169	4,587	26	2,056	2,779	3,707
11	2,201	3,106	4,437	27	2,052	2,771	3,690
12	2,179	3,055	4,318	28	2,048	2,763	3,674
13	2,160	3,012	4,221	29	2,045	2,756	3,659
14	2,145	2,977	4,140	30	2,042	2,750	3,646
15	2,131	2,947	4,073	∞	1,960	2,576	3,291
16	2,120	2,921	4,015				

Приложение 4

Таблица оценки сильно отклоняющихся вариант на «выскакивание». Значения критериев Θ_{max} и Θ_{min} на уровнях значимости 0,01 и 0,05

n	**0,01**	**0,05**	**n**	**0,01**	**0,05**	**n**	**0,01**	**0,05**
4	0,991	0,955	13	0,520	0.410	22	0,414	0,320
5	0,916	0,807	14	0,502	0,395	23	0,407	0,314
6	0,805	0,689	15	0,486	0,381	24	0,400	0,309
7	0,740	0,610	16	0,472	0,369	25	0,394	0,304
8	0,683	0,554	17	0,460	0,359	26	0,389	0,299
9	0,635	0,512	18	0,449	0,349	27	0,383	0,295
10	0,597	0,477	19	0,439	0,341	28	0,378	0,291
11	0,566	0,450	20	0,430	0,334	29	0,374	0,287
12	0,541	0.428	21	0,421	0,327	30	0,369	0,283

Приложение 5

Критические значения критерия Т - Вилкоксона для уровней статистической значимости $p < 0,05$ и $p < 0,01$

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

Критические значения критерия U Виякоксона-Манна-Уитни

n_1 / n_2	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
	p = 0,05																		
3	—	0																	
4	—	0	1																
5	0	1	2	4															
6	0	2	3	5	7														
7	0	2	4	6	8	11													
8	1	3	5	8	10	13	15												
9	1	4	6	9	12	15	18	21											
10	1	4	7	11	14	17	20	24	27										
11	1	5	8	12	16	19	23	27	31	34									
12	2	5	9	13	17	21	25	30	34	38	42								
13	2	6	10	15	19	24	28	33	37	42	47	51							
14	3	7	11	16	21	26	31	36	41	46	51	56	61						
5	3	7	12	18	23	28	33	39	44	50	55	61	66	72					
16	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83				
17	3	9	15	20	26	33	39	45	51	57	64	70	77	83	69	96			
18	4	9	18	22	28	35	41	48	55	61	68	75	62	88	95	102	10		
19	4	10	17	23	30	37	44	51	55	65	72	80	87	94	101	109	11	123	
20	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	12	130	138
	p = 0,01																		
5	_	_	0	1															
6	—	—	1	2	3														
7	—	0	1	3	4	6													
8	—	0	2	4	6	7	9												
9	—	1	3	5	7	9	11	14											
10	—	1	3	6	8	11	13	16	19										
11	—	1	4	7	9	12	15	18	22	25									
12	—	2	5	8	11	34	17	21	24	28	31								
13	0	2	5	9	12	16	20	23	27	31	35	39							
14	0	2	6	10	13	17	22	25	30	34	38	43	47						
15	0	3	7	11	15	19	24	28	33	37	42	47	51	55					
16	0	3	7	12	16	21	26	31	36	41	46	51	56	61	66				
17	0	4	8	13	18	23	28	33	38	44	49	55	50	66	71	77			
18	0	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	65		
19	1	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	
20	1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

(продолжение таблицы)

p = 0,05																		
n_1	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**
21	19	26	34	41	49	57	65	73	81	89	97	105	113	121	130	138	146	154
22	20	28	36	44	52	50	69	77	65	94	102	111	119	128	136	145	154	162
23	21	29	37	46	55	63	72	81	90	99	107	116	125	134	143	152	161	170
24	22	31	39	48	57	66	75	65	94	103	113	122	131	141	150	160	169	179
25	23	32	41	50	50	69	79	89	98	108	118	128	137	147	157	167	177	187
26	24	33	43	53	62	72	82	93	103	113	123	133	143	154	154	174	185	195
27	25	35	46	55	65	75	86	96	107	118	128	139	150	160	171	182	193	203
28	26	36	47	57	68	79	89	100	111	122	133	144	156	167	178	189	200	212
29	27	38	48	59	70	82	93	104	116	127	139	150	162	173	185	196	208	220
30	28	39	50	62	73	85	96	108	120	132	144	156	168	180	192	204	216	228
31	29	41	52	54	76]	88	100	112	124	137	149	161	174	166	199	211	224	236
32	30	42	54	66	78	91	103	116	129	141	154	167	180	193	206	219	232	245
33	31	43	55	65	81	94	107	120	133	146	159	173	186	199	213	226	239	253
34	32	45	58	71	84	97	110	124	137	151	154	178	192	206	219	233	247	261
35	33	46	59	73	86	100	114	128	142	155	170	184	198	212	226	241	255	269
36	35	48	61	75	89	103	117	132	146	160	175	189	204	219	233	248	283	278
37	36	49	63	77	92	106	121	135	150	165	180	196	210	225	240	255	271	286
38	37	51	65	79	94	109	124	139	155	170	185	201	216	232	247	263	273	294
39	38	$2	67	82	97	112	128	143	159	175	190	206	222	238	254	270	286	302
40	39	53	69	84	100	115	131	147	153	179	196	212	228	245	261	278	294	311
p = 0,01																		
21	Ю	16	22	29	35	42	49	55	83	70	77	84	91	98	105	113	120	127
22	10	17	23	30	37	45	52	59	66	74	81	89	96	104	111	119	127	134
23	И	18	25	32	39	47	55	62	70	78	66	94	102	109	117	125	133	141
24	12	19	26	34	42	49	57	66	74	82	90	98	107	115	123	132	140	149
25	12	20	27	35	44	52	50	69	77	86	95	103	112	121	130	138	147	156
26	13	21	29	37	46	54	53	72	81	90	99	108	117	1.26	136	145	154	163
27	14	22	30	39	48	57	66	75	85	94	103	113	122	132	142	151	161	171
28	14	23	32	41	50	59	69	78	88	96	108	118	128	138	148	158	168	178
29	15	24	33	42	52	62	72	82	92	102	112	123	133	143	154	164	175	185
30	15	25	34	44	54	64	75	85	95	106	117	127	138	149	160	171	182	192
n_1	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**10**	**20**	**21**
n_2	p= 0,01																	
31	16	26	36	46	66	67	77	88	99	110	121	132	143	155	166	177	188	20
32	17	27	37	47	58	69	90	91	103	114	126	137	149	160	172	184	195	20
33	17	28	38	49	50	72	83	95	106	118	130	142	154	166	178	1У	202	21
34	18	29	40	51	62	74	86	98	110	122	134	147	159	172	184	197	209	22
35	19	30	41	53	64	77	89	101	114	126	139	152	164	177	190	203	216	22
36	19	31	42	54	67	79	92	104	117	130	143	156	170	183	196	210	223	23
37	20	32	44	56	69	81	95	108	121	134	148	161	175	189	202	216	230	24
38	21	33	45	58	71	84	97	111	125	138	152	166	180	194	208	223	237	25
39	21	34	46	59	73	86	100	114	128	142	157	171	185	200	214	229	244	25
40	22	35	48	61	75	89	103	117	132	146	161	176	191	206	221	236	251	26

Приложение 7

Критические значения коэффициента корреляции Пирсона

k = n - 2	P		k = n - 2	P	
	0,05	0,01		0,05	0,01
5	0,75	0,87	27	0,37	0,47
6	0,71	0,63	28	0,36	0,46
7	0,67	0,80	29	0,36	0,46
8	0,63	0,77	30	0,35	0,45
9	0,60	0,74	за	0,33	0,42
10	0,66	0,71	40	0,30	0,39
11	0,65	0,66	45	0,29	0,37
12	0,63	0,66	50	0,27	0,35
13	0,51	0,64	60	0,25	0,33
14	0,50	0,62	70	0,23	0,30
15	0,48	0,61	80	0,22	0,28
16	0,47	0,59	90	0,21	0,27
17	0,46	0,58	100	0,20	0,25
18	0,44	0,66	125	0,17	0,23
19	0,43	0,55	160	0,16	0,21
20	0,42	0,54	200	0,14	0,18
21	0,41	0,53	300	0,11	0,15
22	0,40	0,52	400	0,10	0,13
23	0,40	0,51	500	0,09	0,12
24	0,39	0,50	700	0,07	0,10
25	0,38	0,49	900	0,06	0,09
26	0,37	0,48	1000	0,06	0,09

Приложение 8

Критические значения коэффициента корреляции рангов Спирмена

n	P		n	P		n	P	
	0,05	0,01		0,05	0,01		0 05	0,01
5	0,94	—	17	0,48	0,62	29	0,37	0,48
6	0,85	—	18	0,47	0,60	30	0,36	0,47
7	0,78	0,94	19	0,46	0,58	31	0,36	0,46
8	0,72	0,68	20	0,45	0,57	32	0,36	0,45
9	0,68	0,83	21	0,44	0,56	33	0,34	0,45
10	0,64	0,79	22	0,43	0,64	34	0,34	0,44
11	0,61	0,76	23	0,42	0,53	35	0,33	0,43
12	0,68	0,73	24	0,41	0,52	36	0,33	0,43
13	0,56	0,70	25	0,49	0,51	37	0,33	0,43
14	0,64	0,68	26	0,39	0,50	38	0,32	0,41
15	0,52	0,68	27	0,38	0,49	39	0,32	0,41
16	0,60	0,64	28	0,38	0,48	40	0,31	0,40

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

1. Артемьева, Е.Ю. Сборник задач по теории вероятностей и математической статистике для психологов [текст] / Е.Ю. Артемьева. - М.: Изд-во МГУ, 1969.
2. Бирхгофф, Г. Математика и психология [текст] / Г. Бирхгофф. - М.: Сов. радио, 1977.
3. Ватутин, В.А. Теория вероятностей и математическая статистика в задачах [текст]: учебное пособие для вузов / В.А. Ватутин, Г.И. Ивченко, Ю.И. Медведев и др. – М.: Дрофа, 2003.
4. Герасимов, В.П. Математическое обеспечение психологических исследований [текст]: учебное пособие / В.П. Герасимов. - Бийск: НИЦ БиГПИ, 1997.
5. Гласс, Дж. Статистические методы в педагогике и психологии [текст] / Дж. Гласс, Дж. Стэнли. - М.: Прогресс, 1976.
6. Гмурман, В.Е. Теория вероятностей и математическая статистика. [текст]: учебное пособие для вузов / В.Е. Гмурман. - М.: Высшая школа, 2000.
7. Горелова, Г.В. Теория вероятностей и математическая статистика в примерах и задачах с применением Excel [текст]: учебное пособие для вузов / Г.В. Горелова, И.А. Кацко. - Ростов н/Д.: Феникс, 2002.
8. Грес, П.В. Математика для гуманитариев [текст]: учебное пособие / П.В. Грес. - М.: Юрайт, 2000.
9. Ермолаев, А.Ю. Математическая статистика для психологов [текст]: учебник / А.Ю. Ермолаев. - М.: Московский психолого – социальный институт: Флинта, 2003.
10. Клайн, П. Справочное руководство по конструированию тестов. Введение в психометрическое проектирование [текст] / Клайн П. - Киев: «ПАН Лтд.», 1994.

11. Козлов, В.Н. Математика и информатика [текст] / В.Н. Козлов. - СПб.: Питер, 2004.
12. Основы математической статистики [текст]: учебное пособие для институтов физической культуры) / Под общ. ред. В.С. Иванова. - М.: Физкультура и спорт, 1990.
13. Сидоренко, Е.В. Методы математической обработки в психологии [текст] / Е.В. Сидоренко. - СПб.: ООО «Речь», 2003.
14. Сосновский, Б.А. Лабораторный практикум по общей психологии [текст] / Б.А. Сосновский / Под ред. Гамезо В.М. - М.: Просвещение, 1979.
15. Тюрин, Ю.Н. Статистический анализ данных на компьютере [текст] / Ю.Н. Тюрин, А.А. Макаров / Под ред. В.Э. Фигурнова. - М.: ИНФРА–М, 1998.