Министерство образования и науки Российской Федерации
Тольяттинский государственный университет
Гуманитарно-педагогический институт
Кафедра «Педагогика и методики преподавания»

Г.В. Ахметжанова, И.В. Антонова

ПРИМЕНЕНИЕ МЕТОДОВ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ В ПСИХОЛОГО-ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

Электронное учебное пособие

ISBN 978-5-8259-1134-2

УДК 519.22(075.8)
ББК 22.172я73

Рецензенты:
канд. пед. наук, директор муниципального бюджетного учреждения «Лицей № 76 имени В.Н. Полякова» г.о. Тольятти *Ю.С. Коняхина*;
д-р пед. наук, профессор кафедры «Педагогика и методики преподавания» Тольяттинского государственного университета *И.В. Руденко*.

Ахметжанова, Г.В. Применение методов математической статистики в психолого-педагогических исследованиях : электронное учебное пособие / Г.В. Ахметжанова, И.В. Антонова. – Тольятти : Изд-во ТГУ, 2016. – 1 оптический диск.

Пособие содержит сведения об измерениях, статистических гипотезах и основах анализа в психолого-педагогических исследованиях; в пособии имеются задания для самостоятельной работы, статистические таблицы и глоссарий.

Предназначено для студентов магистратуры, обучающихся по направлениям подготовки 44.04.02 «Психолого-педагогическое образование», 44.04.01 «Педагогическое образование».

Текстовое электронное издание.

Рекомендовано к изданию научно-методическим советом Тольяттинского государственного университета.

Минимальные системные требования: IBM PC-совместимый компьютер: Windows XP/Vista/7/8; PIII 500 МГц или эквивалент; 128 Мб ОЗУ; SVGA; CD-ROM; Adobe Acrobat Reader.

Редактор *О.И. Елисеева*
Технический редактор *Н.П. Крюкова*
Компьютерная верстка: *Л.В. Сызганцева*
Художественное оформление,
компьютерное проектирование: *Г.В. Карасева, И.В. Карасев*

Дата подписания к использованию 20.12.2016.
Объем издания 2,3 Мб.
Комплектация издания: компакт-диск,
первичная упаковка.
Заказ № 1-69-16.

Издательство Тольяттинского государственного университета
445020, г. Тольятти, ул. Белорусская, 14,
тел. 8 (8482) 53-91-47, www.tltsu.ru

Содержание

ВВЕДЕНИЕ

В системе профессиональной подготовки специалистов, осваивающих программы магистратуры по педагогическим и психолого-педагогическим направлениям, ведущее значение имеет их умение проводить соответствующие эксперименты в своей будущей профессиональной деятельности. Каждое диссертационное исследование, обязательной составляющей которого является экспериментальная часть, требует владения методами статистической обработки результатов эксперимента.

Вместе с этим в ФГОС ВО по направлению подготовки 44.04.02 «Психолого-педагогическое образование» указано, что у выпускника магистратуры необходимо сформировать следующие компетенции:

– способность к абстрактному мышлению, анализу, синтезу (ОК-1);
– способность использовать научно обоснованные методы и технологии в психолого-педагогической деятельности, владеть современными технологиями организации сбора, обработки данных и их интерпретации (ОПК-2);
– способность проектировать и осуществлять диагностическую работу, необходимую в профессиональной деятельности (ОПК-5).

По направлению подготовки 44.04.01 «Педагогическое образование» выпускник должен обладать способностью:

– к самостоятельному освоению и использованию новых методов исследования, к освоению новых сфер профессиональной деятельности (ОК-3);
– осуществлять профессиональное и личностное самообразование, проектировать дальнейшие образовательные маршруты и профессиональную карьеру (ОПК-4);
– применять современные методики и технологии организации образовательной деятельности, диагностики и оценивания качества образовательного процесса по различным образовательным программам (ПК-1);
– анализировать результаты научных исследований, применять их при решении конкретных научно-исследовательских задач в сфере науки и образования, самостоятельно осуществлять научное исследование (ПК-5).

В данном пособии рассмотрены основные понятия математической статистики, описаны группы статистических методов, с помощью которых осуществляется обработка результатов педагогического эксперимента и формируются перечисленные компетенции. Представлены выводы и рекомендации по выбору определенных статистических критериев при обработке результатов эксперимента. Рассмотрение основных понятий математической статистики и каждого статистического критерия заканчивается соответствующими контрольными вопросами. Приведены компьютерные программы, позволяющие проводить статистический анализ данных в соответствии с характером решаемых задач, объемом обрабатываемых данных, требованиями, предъявляемыми к квалификации пользователя, а также с имеющимся в наличии компьютерным оборудованием. Представлены некоторые интернет-ресурсы, которые окажут практическую помощь магистрантам при изучении основных понятий математической статистики.

Одним из достоинств данного пособия являются примеры диссертационных исследований, использующих статистические методы обработки экспериментальных данных. Рассмотрение каждого статистического критерия заканчивается соответствующим заданием для работы на учебных занятиях. Следует отметить разнообразие заданий:

- выявить ошибки в вычислениях или в приведенных логических рассуждениях по определенному статистическому методу;
- рассмотреть различные случаи, связанные с применением рассматриваемого критерия и соответствующие способы вычисления для каждого из них и т. д.

В конце пособия приведены задания для самостоятельной работы, включающие как теоретические, так и практические вопросы.

Кроме того, интерес представляет приведенный глоссарий, который помогает исследователю работать с основными понятиями математической статистики и методами обработки результатов психолого-педагогического эксперимента.

Следует отметить наличие критериев оценки письменного и устного ответов студентов. В приложении содержатся статистические таблицы.

§ 1. ОСНОВНЫЕ ТИПЫ ИЗМЕРЕНИЙ В ПСИХОЛОГО-ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

1.1. Понятие измерения. Типы измерительных шкал

Измерение – это процедура, с помощью которой измеряемый объект сравнивается с некоторым эталоном и получает численное выражение в определенном масштабе или по особой шкале [13].

Психолого-педагогические переменные за единичными исключениями не имеют собственных измерительных единиц. Поэтому в большинстве случаев значение психолого-педагогического признака определяется при помощи специальных измерительных шкал [5].

Существует четыре типа измерительных шкал:

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

Все находящиеся в одной строчке наименования являются синонимами и поэтому в дальнейшем изложении будут использоваться на равных основаниях.

В процессе проведения эксперимента исследователь может воспользоваться указанными способами, при этом необходимо отметить наличие у каждой измерительной шкалы собственной, отличной от других формы числового представления. Проводя эксперимент, исследователь фиксирует закодированные признаки явления в определенной числовой системе, которая определяется особенностями используемой шкалы. Измерения делятся на качественные и количественные. К качественным относят те измерения, которые осуществляются с помощью двух первых шкал, к количественным – производимые с помощью двух последних. Шкалы, которые приводят к качественным измерениям, называют дискретными; шкалы, которые приводят к количественным измерениям, – непрерывными [11].

Рассмотрим подробно перечисленные выше шкалы.

1.2. Номинативная шкала

Измерение в номинативной шкале выражается в том, что определенному свойству или признаку присваивается соответствующее обозначение или символ (численный, буквенный и т. п.). В процессе измерения по данной шкале происходит процесс классификации или распределение объектов на непересекающиеся классы или группы.

Номинативная шкала указывает на качественное отличие друг от друга определенных свойств или признаков, а какие-либо количественные их изменения не предполагаются. Таким образом, относительно признаков, измеренных по данной шкале, нельзя определить, какой из них больше, а какой меньше. Можно лишь утверждать, что признаки, попавшие в разные группы, различны, что характеризует данную шкалу как качественную.

Необходимо обратить внимание на тот факт, что символы, которые присваиваются объектам в номинативной шкале, являются условными и не обладают определяющей информацией, при этом операции, проводимые с ними, не имеют смысла.

Дихотомической называют самую простую номинативную шкалу. В процессе измерения по данной шкале исследуемые признаки можно кодировать двумя цифрами или символами, например 1 и 0, или 6 и 2, или буквами С и Д. Можно использовать любые символы, которые отличаются друг от друга. Признак, измеренный по дихотомической шкале, называют *альтернативным*.

Отметим, что в дихотомической шкале все объекты, признаки или изучаемые свойства должны быть разбиты на два непересекающихся класса, при этом исследователю необходимо найти ответ на вопрос, проявился или нет интересующий его признак. Использование данной шкалы позволяет экспериментатору выявить частоту встречаемости исследуемого признака.

При использовании номинативной шкалы единицей измерения является количество наблюдений. Общее число наблюдений принимается за 100 %, и вычисляется процентное соотношение. При количестве групп больше двух можно подсчитать процентный состав испытуемых в каждой группе [13].

Примером измерения по шкале наименований может являться классификация студентов по состоянию двух признаков: пол и успешность выполнения контрольного задания. В состоянии каждого признака выделяется в данном случае по две градации: мужчина/женщина и верный ответ/неверный ответ. В результате по состоянию двух признаков коллектив разделяется на четыре класса: женщины, верно выполнившие задание; женщины, неверно выполнившие задание; мужчины, верно выполнившие задание; мужчины неверно выполнившие задание. Далее, всем объектам первого класса припишем, например, число 1, объектам второго класса – число 2, третьего класса – число 3, четвертого – число 4. Классификация некоторой совокупности объектов по нескольким признакам позволяет не только выявить структуру совокупности, но и получить данные, необходимые для выяснения связи двух и более признаков. В приведенном примере такими признаками являются пол студентов и выполнение контрольного задания.

Необходимо отметить, что только небольшое количество статистических методов (критерий хи-квадрат, критерий Макнамары и др.) допускают представление результатов в виде номинативной шкалы. Например, измерения по данной шкале используются с целью проверки определенных статистических гипотез, а также для вычисления показателей корреляции качественных признаков [11].

1.3. Ранговая шкала. Правило ранжирования

Измерение по ранговой шкале расчленяет полную совокупность измеренных признаков на множества, связанные между собой отношениями типа «меньше – больше», «ниже – выше», «слабее – сильнее» и т. п. В ранговой шкале все признаки расположены по рангу – от самого большого до меньшего, от самого сильного до самого слабого и т. п.

При использовании ранговой шкалы должно быть не меньше трех классов для расстановки измеренных признаков по порядку. Поэтому второе название данной шкалы «порядковая шкала».

В процессе кодирования порядковым переменным приписываются любые цифры или коды, но необходимо обратить внимание на обязательное присутствие определенного порядка в кодах [13].

Процедура ранжирования по сути является формальной, поэтому по желанию экспериментатор может проставить величины рангов и в противоположном порядке.

Например, в процессе измерения результатов обучения студентов шкала порядка может быть представлена критериями, которые позволяют расположить результаты по степени увеличения или уменьшения исследуемого признака. Таким образом, нет возможности определить, на какое количество равных единиц по состоянию признака один объект больше или меньше другого.

Шкала порядка имеет ряд определенных особенностей. Так, на выводы, которые были получены экспериментатором с помощью арифметических действий с измерениями, выполненными по данной шкале, влияет выбор балловых оценок. То есть выводы в некоторых случаях могут отражать не объективные свойства рассматриваемых явлений, а случайные свойства, являющиеся следствием характерных особенностей принятой системы балловых оценок. Используя систему измерений по шкале порядка, необходимо помнить об ограничении: с числами (баллами, рангами), приписанными объектам, нельзя производить арифметических действий – вычислять суммы, находить средние значения, дисперсии и т. п. [11].

Процедура ранжирования достаточно проста. В наиболее общем случае для проверки правильности ранжирования столбца или строчки признаков применяется формула суммы всех полученных рангов:

$$\textit{Сумма рангов} = 1 + 2 + 3 + \ldots + = \frac{N \cdot (N+1)}{2}, \tag{1}$$

где N – количество ранжируемых признаков.

Совпадение итогов подсчета рангов по формуле и по реальным результатам ранжирования экспериментальных данных в этом случае будет являться подтверждением правильности ранжирования.

В случае когда экспериментатор работает с таблицей, имеющей большое количество строк и столбцов, при подсчете рангов возможно использование модификации формулы (1):

$$\textit{Сумма рангов} = \frac{\left((k \cdot c + 1) \cdot k \cdot c\right)}{2}, \tag{2}$$

где k – число строк; c – число столбцов.

Расчетная формула обшей суммы рангов для такого способа ранжирования выглядит следующим образом (3):

$$Сумма\ рангов = \frac{n \cdot c \cdot (c+1)}{2}, \tag{3}$$

где n – количество испытуемых в столбце; c – количество столбцов групп испытуемых.

При выполнении измерений могут возникнуть ситуации, когда двум или большему числу качеств приписывают одинаковые ранги.

Одинаковые ранги могут быть присвоены любому числу ранжируемых величин. При этом им также приписывается величина среднего арифметического от количества условных рангов, которые были проставлены по порядку их величин. При ранжировании количественных характеристик можно использовать специальный *алгоритм*.

1. Наименьшему числовому значению приписывается ранг 1.

2. Наибольшему числовому значению приписывается ранг, который равен количеству ранжируемых величин.

3. При равенстве нескольких исходных числовых значений им приписывается ранг, соответствующий средней величине тех рангов, которые эти величины получили бы, если бы они стояли по порядку друг за другом и не были бы равны.

Необходимо отметить, что под этот случай могут попасть и первые, и последние величины исходного ранжируемого ряда.

4. Общая сумма реальных рангов должна совпадать с расчетной формулой (1) [13].

Следует отметить, что не рекомендуется ранжировать более 20 величин (признаков, качеств, свойств и т. п.), так как в этом случае ранжирование становится малоустойчивым. Если есть необходимость ранжирования большого числа объектов, то их нужно объединить по выбранному признаку в однородные группы (классы), а потом ранжировать сформированные группы (классы).

В табл. 1 условные ранги указаны для удобства в скобках, реальные ранги расположены в этом же столбце.

Таблица 1

Ранжирование результатов контрольной работы группы испытуемых

№ испытуемого	Показатель количества баллов за правильно выполненные задания контрольной работы	Ранги	
1	12	(10)	9
2	8	(5)	5
3	7	(4)	4
4	11	(7)	6,5
5	12	(9)	9
6	5	(2)	2
7	3	(1)	1
8	11	(6)	6,5
9	6	(3)	3
10	12	(8)	9
Сумма		55	

Проверим правильность ранжирования по формуле (1), получим:

$$\textit{Сумма рангов} = \frac{10 \cdot (10+1)}{2} = 5 \cdot 11 = 55.$$

Суммируя реальные ранги, получим также сумму рангов, равную 55, значит, ранжирование выполнено верно.

К экспериментальным данным, полученным в этой шкале, применяется множество разнообразных статистических методов.

1.4. Интервальная шкала

В шкале интервалов, или интервальной шкале, каждое из возможных значений измеренных величин отстоит от ближайшего на равном расстоянии.

Интервал, являющийся основным понятием данной шкалы, определяется как часть или доля измеряемого свойства между двумя соседними позициями на шкале. Размером интервала называют фиксированную и постоянную на всех участках шкалы величину.

Используя шкалу интервалов, в процессе измерения применяют специальные единицы измерения. При работе с данной шкалой измеряемому свойству или предмету экспериментатор присваивает число, которое соответствует количеству единиц измерения и является эквивалентом количеству имеющегося свойства.

Характерной особенностью шкалы интервалов является отсутствие естественной точки отсчета (нуль условен и не указывает на отсутствие измеряемого свойства) [13].

Интервальная шкала, или шкала равных единиц, получается при использовании определенного критерия, с помощью которого удается измерить интервал между объектами в состоянии изучаемого свойства, другими словами, можно выяснить, на сколько единиц один объект больше или меньше другого.

В зависимости от характера рассматриваемой данной педагогической наукой проблемы следует принимать тот или иной условный нулевой уровень. Например, когда мы оцениваем выполнение студентами какого-либо задания по числу верных или неверных ответов на вопросы данного задания, то нуль верных ответов не означает полного отсутствии знаний, проверяемых заданием, у этого студента. В педагогике, за исключением шкалы времени, которая, хотя и используется достаточно широко, но играет в теории измерений особую роль, в настоящее время нет измерительных шкал интервального типа [3].

Необходимо отметить, что большинство статистических методов применяются для получения экспериментальных данных именно в интервальной шкале.

1.5. Шкала отношений

Шкалу отношений иногда называют шкалой равных отношений. Характерной особенностью данной шкалы является наличие фиксированного нуля, обозначающего полное отсутствие какого-либо свойства или признака. Шкала отношений – это самая информативная шкала, допускающая использование любых математических операций и разнообразных статистических методов.

Отметим относительную близость шкалы отношений к интервальной шкале, так как при строгой фиксации начала отсчета любая интервальная шкала трансформируется в шкалу отношений [13].

Процесс перехода от одной шкалы измерения к другой (начиная со шкалы наименований) сопровождается расширением класса допустимых для полученных измерений математических операций и методов. В этом смысле наилучшей является шкала отношений [11].

Контрольные вопросы

1. Дайте определение измерения и назовите единицы измерения, существующие в психолого-педагогических исследованиях.
2. Перечислите типы измерительных шкал, выделяемые С. Стивенсом, и дайте их краткую характеристику.
3. С помощью какой операции экспериментальным данным придается форма числового сообщения?
4. Сформулируйте определение качественных и количественных измерений и перечислите их отличия.
5. Дайте определение номинативной шкалы и объясните особенность измерения по этой шкале, укажите единицу измерения в этой шкале.
6. Отметьте особенность дихотомической шкалы и определите ее место среди основных типов шкал.
7. В чем заключается суть измерения в ранговой шкале? Назовите наименьшее количество групп, которое должно быть в ранговой шкале.
8. Поясните смысл процедуры ранжирования данных эксперимента. Сформулируйте алгоритмы ранжирования данных в случае одинаковых и неодинаковых рангов.
9. Какова особенность экспериментальных данных, представленных в шкале интервалов? Назовите единицу измерения в шкале интервалов.
10. Приведите примеры измерений по всем четырем видам шкал в психолого-педагогических исследованиях.

§ 2. ФОРМЫ УЧЕТА РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ

2.1. Таблицы и графики распределения частот

Для наглядного представления экспериментальных данных используются различные формы, облегчающие визуальный анализ полученной в эксперименте информации: таблицы, ряды распределений, графики, гистограммы.

Обработка экспериментальных данных начинается с упорядочения и систематизации собранных данных.

Под группировкой понимают процесс приведения в систему результатов эксперимента и объединение их в относительно однородные группы по определенному признаку.

При проведении эксперимента группировка – это не стандартный технический прием, который позволяет представить исходные данные в ином виде, а операция, позволяющая осмыслить связи между изучаемыми явлениями.

Статистические таблицы представляют наиболее распространенную форму группировки экспериментальных данных. Они, в свою очередь, подразделяются на простые и сложные.

К простым относятся таблицы, которые находят свое применение при альтернативной группировке, в случае противопоставления одной группы испытуемых другой.

Исследователи рекомендуют использовать данный вид таблиц для фиксирования результатов эксперимента, в котором измерение изучаемых признаков производится в номинативной или ранговой шкале.

К сложным таблицам относят многопольные таблицы, используемые экспериментаторами при выяснении причинно-следственных отношений между варьирующими признаками [13].

2.2. Понятие статистического ряда

Особую форму группировки данных представляют собой *статистические ряды,* или числовые значения признака, расположенные в определенном порядке. В зависимости от изучаемого признака статистические ряды можно разделить на атрибутивные, вариационные, ряды динамики, регрессии, ряды ранжированных значений признаков и ряды накопленных частот.

Вариационный ряд распределения – это двойной ряд чисел, который показывает, как числовые значения признака связаны с их повторяемостью в определенной выборке.

Используя в эксперименте вариационный ряд распределения, оперируют понятиями частот или весов варианта, дающими представление о том, сколько раз отдельные варианты могут встретиться в данной совокупности. Частоты обозначаются строчной буквой латинского алфавита f_i и имеют индекс i, который соответствует номеру переменной в вариационном ряду.

Общая сумма частот вариационного ряда равна объему выборки, т. е. $n = \sum_{i=1} f_i$.

Частоты также могут выражаться и в процентах. В данном случае общая сумма частот или объем выборки принимается за 100 %.

Процент каждой отдельной частоты или веса подсчитывается по формуле:

$$n_i\% = \frac{f_i}{n} \cdot 100\%.$$

Процентное представление частот полезно в тех случаях, когда приходится сравнивать вариационные ряды, сильно различающиеся по объемам [13]. По вариационному ряду (табл. 2) видно, что объем выборки равен 21.

Таблица 2

Вариационный ряд

Варианты	x_i	**5**	**6**	**7**	**8**	**9**	**10**	
Частоты	f_i	1	1	2	4	7	6	***n=21***
Кумуляты частот	σ_i	1	2	4	8	15	***21***	

Кумуляты частот показывают места измерений (вариантов) в ранжированном вариационном ряду. Так, на 1-м месте идет «5», на 2-м – «6», на 3-м и 4-м местах – «7», на 5–8-м местах – «8», на 9–15-м местах – «9», на 16–21-м местах – «10».

2.3. Понятие распределения. Основные числовые характеристики распределений. Гистограммы распределения частот

Используя в статистике термин «ряд распределения», имеют в виду распределение частот по вариантам.

Измеренные величины признака в выборке варьируют в пределах от минимального до максимального значения. Этот предел разбивают на классовые интервалы, которые, в зависимости от конкретных данных, могут быть как равными, так и неравными по величине.

Когда по оси абсцисс OX откладываются величины классовых интервалов, а по оси ординат OY – величины частот, которые попадают в определенный классовый интервал, экспериментатор получает так называемую *гистограмму распределения частот*. Гистограмма является графическим изображением данного частотного распределения. Различают несколько видов гистограмм (рис. 1).

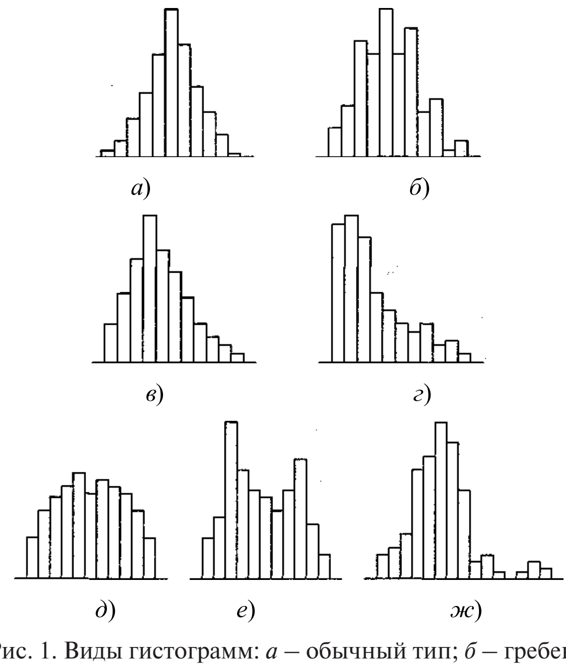

Рис. 1. Виды гистограмм: *а* – обычный тип; *б* – гребенка; *в* – положительно скошенное распределение; *г* – распределение с обрывом слева; *д* – плато; *е* – двухпиковый тип; *ж* – распределение с изолированным пиком

Для экспериментальных данных, полученных по выборке, можно вычислить ряд числовых характеристик [13].

Основными числовыми характеристиками распределений являются:

1) мода (числовое значение, которое встречается в выборке наиболее часто; $\hat{X}$). По данным табл. 2 значение моды равно 9;

2) медиана (числовое значение, которое делит упорядоченное множество данных пополам; $\tilde{X}$ или M_d). Медиану удобно находить с помощью числового значения, показывающего номер медианы или место измерения в ранжированном вариационном ряду. Так, возвращаясь к табл. 2, отметим, что номер медианы ряда равен $(n + 1) : 2$, т. е. $(21 + 1) : 2 = 11$. Так как на 11-м месте в ряду стоит «9», следовательно, значение медианы этого ряда данных равно 9;

3) среднее арифметическое ($\overline{X}$). По табл. 2 его значение равно

$8\frac{4}{7}$ ($\overline{X} = 5 \cdot 1 + 6 \cdot 1 + 7 \cdot 2 + 8 \cdot 4 + 9 \cdot 7 + 10 \cdot 6) : 21 = 180 : 21$);

4) разброс выборки, или размах (разность между максимальной и минимальной величинами вариационного ряда; R). В нашем случае значение разброса выборки равно $10 - 5 = 5$;

5) дисперсия (среднее арифметическое квадратов отклонений значений переменной от ее среднего значения; D). Так,

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

Из табл. 2 имеем значение дисперсии, равное:

$$\frac{1}{21} \cdot \left(\left(5 - 8\frac{4}{7}\right)^2 + \left(6 - 8\frac{4}{7}\right)^2 + 2 \cdot \left(7 - 8\frac{4}{7}\right)^2 + 4 \cdot \left(8 - 8\frac{4}{7}\right)^2 + 7 \cdot \left(9 - 8\frac{4}{7}\right)^2 + 6 \cdot \left(10 - 8\frac{4}{7}\right)^2 \right) =$$

$$= \frac{1}{21} \cdot \left(\left(-3\frac{4}{7}\right)^2 + \left(-2\frac{4}{7}\right)^2 + 2 \cdot \left(-1\frac{4}{7}\right)^2 + 4 \cdot \left(-\frac{4}{7}\right)^2 + 7 \cdot \left(\frac{3}{7}\right)^2 + 6 \cdot \left(1\frac{3}{7}\right)^2 \right) = \frac{274 \cdot 1}{7 \cdot 21} = \frac{274}{147}.$$

Итак, значение дисперсии данного ряда приближенно равно 1,9;

6) степень свободы (число свободно варьирующихся единиц в составе выборки). В общем случае для таблицы экспериментальных данных число степеней свободы определяется по формуле $v = (c - 1) \cdot (n - 1)$, где c — число столбцов, n — число строк (число испытуемых). Для ряда статистических методов расчет числа степеней свободы имеет свою специфику;

7) понятие нормального распределения. Форма и положение графика нормального распределения, имеющего вид колоколообразной кривой (рис. 2), определяется средней (μ) и стандартным отклонением (σ или S), равным квадратному корню из дисперсии; величины моды, медианы и среднего арифметического совпадают, то есть $\hat{X}=\tilde{X}=\bar{X}$). В нашем случае стандартное отклонение равно $S_x = \sqrt{\dfrac{274}{147}} \approx 1{,}4$; $\hat{X} = \tilde{X} = 9$, $\bar{X} \approx 9$, имеем распределение, близкое к нормальному.

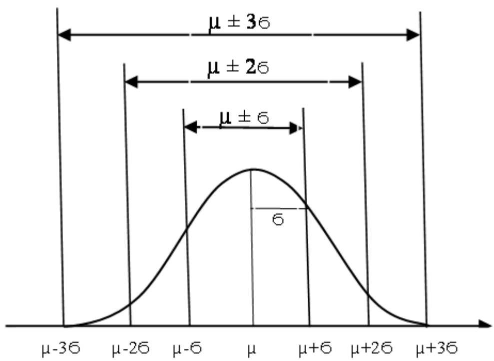

Рис. 2. Параметры μ и σ для нормального распределения

Отметим, что при изображении данных, полученных в ходе психолого-педагогических исследований, чаще всего используются различные виды статистических графиков (линейные графики, столбчатые, круговые и векторные диаграммы).

Так, в соответствии с полученным в примере вариационным рядом можно построить гистограмму распределения частот данных, представленных в табл. 2, где по оси OX отложены числовые значения измерения (признака), по оси OY – их частоты (рис. 3).

По данным, представленным на рис. 3, видно, что полученное распределение действительно является близким к нормальному.

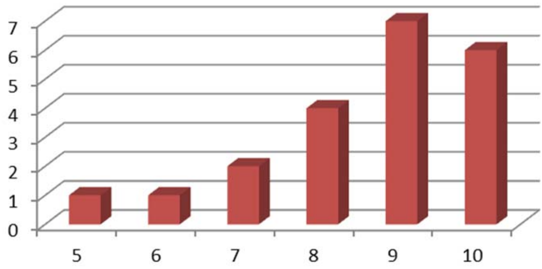

Рис. 3. Гистограмма распределения частот данных табл. 2

Контрольные вопросы

1. Перечислите формы для наглядного представления экспериментальных данных. Укажите цель их использования.
2. Назовите виды статистических таблиц и укажите их отличие друг от друга.
3. Дайте определение статистическому ряду и перечислите виды деления данных рядов.
4. Сформулируйте определение вариационного ряда распределения и укажите, что является его частотой.
5. Что понимают под кумулятами частот? Как они считаются в определенном вариационном ряду?
6. Составьте произвольный вариационный ряд. Найдите по нему кумуляты частот.
7. С какой целью используются в обработке эксперимента гистограммы распределения частот?
8. Перечислите основные числовые характеристики распределений.
9. Дайте определение моды и укажите правила ее нахождения.
10. Сформулируйте определение медианы и укажите способы ее нахождения.
11. Дайте определение дисперсии и приведите примеры нахождения дисперсии для определенного ряда данных.
12. Какую величину называют стандартным отклонением?
13. Какое распределение называют нормальным? Определите, что является характерным для нормального распределения.

§ 3. СТАТИСТИЧЕСКИЕ ГИПОТЕЗЫ В ПСИХОЛОГО-ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

3.1. Понятие статистической гипотезы и сущность ее проверки

Под статистической гипотезой понимается любое предположение о свойствах случайных величин или событий.

В процессе проверки статистической гипотезы экспериментатор устанавливает, насколько согласуются экспериментальные данные и выдвинутая гипотеза и возможно ли отнести расхождение между гипотезой и результатами статистического анализа эксперимента на счет случая. Руководствуясь вышесказанным, делаем вывод о том, что статистическая гипотеза является научной гипотезой, допускающей статистическую проверку.

3.2. Типы статистических гипотез в педагогике

Выделяют четыре основных типа гипотез, применяемых в психолого-педагогических исследованиях [11]:

1) гипотезы о типах вероятностных законов распределения случайных величин, которые характеризуют изучаемое свойство явления или процесса. Обобщенно смысл данной гипотезы можно сформулировать следующим образом: «Некоторое свойство педагогического явления имеет определенный закон распределения (нормальный, экспоненциальный, Пуассона и т. д.)».

При проверке гипотез первого типа используются методы математической статистики, получившие название «критерии согласия». Применение данных критериев возможно только на основе количественных измерений изучаемого свойства педагогического явления в выборке;

2) гипотезы о свойствах тех или других числовых параметров (средних значений, медиан, дисперсий и др.), характеризующих изучаемые случайные величины. Обобщая, гипотезам этого типа можно присвоить следующую формулировку: «Значение параметра, характеризующего некоторое свойство изучаемого педагогическо-

го явления, не меньше (не больше) некоторого заданного значения или заключается в заданных пределах».

Гипотезы второго типа проверяются в основном методами, получившими название параметрических (критерий Стьюдента, критерий Снедекора – Фишера и др.); следует отметить, что данные методы опираются исключительно на количественные измерения;

3) гипотезы о стохастической (вероятностной) зависимости двух и более признаков (факторов), которые характеризуют различные стороны исследуемого процесса или явления. Смысл данных гипотез применительно к педагогике можно сформулировать следующим образом: «Два или более свойств рассматриваемого педагогического явления стохастически зависимы, и зависимость эта подчиняется определенному закону (например, линейному); некоторый фактор или факторы оказывают влияние на изучаемое свойство педагогического явления, и эта стохастическая зависимость подчиняется определенному закону».

Установить факт наличия или отсутствия связи двух и более признаков (проверка гипотез третьего типа) можно с помощью качественных измерений изучаемых признаков, определяя значения коэффициентов Юла, Кендалла, Спирмена и др.;

4) гипотезы о равенстве или различии законов распределения случайных величин, которые характеризуют исследуемое свойство в двух и более совокупностях рассматриваемых явлений. Содержание гипотез четвертого типа применительно к психолого-педагогическим исследованиям можно сформулировать следующим образом: «Состояние одного и того же свойства имеет одинаковое или различное распределение в каждой из двух или более совокупностей учащихся, отличающихся содержанием, методом или организацией обучения или социальной средой».

Проверка гипотез четвертого типа проводится при помощи так называемых критериев значимости. Некоторую их часть можно применять только при количественных измерениях исследуемого свойства, остальные используются и при качественных измерениях.

Использование непараметрических методов математической статистики возможно при проверке гипотез двух последних типов.

Контрольные вопросы

1. Дайте определение понятия выборки.
2. Приведите примеры независимых и зависимых выборок.
3. Сформулируйте отличие полного исследования от выборочного.
4. Дайте определение понятия генеральной совокупности.
5. Перечислите основные требования, предъявляемые к выборкам в соответствии с целями и задачами исследования.
6. Охарактеризуйте понятие репрезентативной выборки.
7. Опишите два метода, обеспечивающие репрезентативность выборки.
8. Перечислите основные типы статистических гипотез, выделяемых в психолого-педагогических исследованиях.
9. Приведите примеры статистических гипотез.
10. Дайте характеристику статистическим методам, с помощью которых осуществляется проверка статистических гипотез.

§ 4. ОБЩИЕ ПРИНЦИПЫ ПРОВЕРКИ СТАТИСТИЧЕСКИХ ГИПОТЕЗ

4.1. Понятие нулевой и альтернативной гипотез

В процессе проверке статистических гипотез используют два понятия: нулевая гипотеза (обозначение H_0) и альтернативная гипотеза (обозначение H_1).

Нулевая гипотеза H_0 – это гипотеза о сходстве, а альтернативная H_1 – гипотеза о различии. Таким образом, принятие нулевой гипотезы H_0 свидетельствует об отсутствии различий, а гипотезы H_1 – о наличии различий [13].

В частности, в качестве нулевой гипотезы H_0 можно рассматривать гипотезу, которая утверждает, что изучаемые выборки взяты из генеральных совокупностей с одинаковым законом распределения, а различие в результатах выборок объясняется чисто случайными причинами. Примером нулевой гипотезы такого типа в психолого-педагогических исследованиях является утверждение о том, что различие в результатах выполнения двумя группами студентов одной и той же контрольной работы вызвано случайными причинами, а на самом деле уровень выполнения этой работы для той и другой группы студентов одинаков; пример альтернативной гипотезы – это утверждение о том, что уровни выполнения работы в двух группах студентов различаются и данные различия являются результатом воздействия неслучайных факторов, например, различных методов обучения.

Нулевая гипотеза H_0 обычно проверяется в процессе сравнения ее с альтернативной гипотезой H_1. В общей теории проверки гипотез обычно представляется множество альтернатив, которые принадлежат определенному классу статистических гипотез [11].

При проведении эксперимента, основываясь на определенных правилах, необходимо сделать выбор: либо отвергнуть H_0 и тем самым принять H_1, либо принять H_0 и тем самым отвергнуть H_1.

4.2. Понятие уровня значимости

В процессе обработки результатов психолого-педагогического эксперимента при обосновании статистического вывода необходимо сделать выбор между принятием или отклонением нулевой гипотезы. Данную границу невозможно четко обозначить, так как в эксперименте всегда присутствуют случайные влияния. Она основывается на понятии уровня значимости. *Уровнем значимости* называется вероятность ошибочного отклонения нулевой гипотезы или, другими словами, вероятность ошибки первого рода при принятии решения. При обозначении данной вероятности чаще всего употребляют букву P.

Стандартными уровнями статистической значимости являются величины 0,05, 0,01 и 0,001.

Принято считать, что низшим уровнем статистической значимости является уровень $P = 0,05$, достаточным – уровень $P = 0,01$ и высшим – уровень $P = 0,001$. Поэтому в статистических таблицах, которые приводятся в приложениях к учебникам по статистике, обычно даются табличные значения для уровней $P = 0,05$, $P = 0,01$ и $P = 0,001$, иногда – для уровней $P = 0,025$ и $P = 0,005$.

В современных статистических пакетах на ЭВМ используются уровни значимости, которые подсчитываются при работе с применяемым статистическим методом. Эти уровни могут иметь различное числовое выражение в интервале от 0 до 1, например, $P = 0,7$, $P = 0,23$ или $P = 0,012$. В первых двух случаях полученные уровни значимости достаточно велики и нельзя говорить о значимости результата; в последнем случае можно говорить о достоверном уровне, потому что показаны результаты на уровне тысячных долей [13].

Сформулируем правило принятия статистического вывода:

I. Подсчет эмпирической статистики или эмпирического значения ($Ч_{эмп}$) по выбранному статистическому методу.

II. Сравнение эмпирического значения ($Ч_{эмп}$) с двумя критическими величинами, соответствующими уровням значимости в 5 % ($Ч_{кр1}$) и в 1 % ($Ч_{кр2}$) для выбранного статистического метода.

Отметим, что величины $Ч_{кр}$ для данного статистического метода можно найти по таблицам, находящимся в приложениях к любому учебнику по статистике.

Найденные по таблицам величины критических значений $Ч_{кр1}$ и $Ч_{кр2}$ представляют в следующей стандартной форме записи:

$$Ч_{кр} = \begin{cases} Ч_{кр1}, \text{найденное по таблицам приложения для } P \leq 0{,}05 \\ Ч_{кр2}, \text{найденное по таблицам приложения для } P \leq 0{,}01. \end{cases}$$

Во всех статистических методах используются свои символические обозначения данных величин: и подсчитанных по соответствующему статистическому методу эмпирической величины, и найденных в таблицах критических величин.

Процесс сравнения эмпирического значения с двумя критическими значениями, которые были найдены в таблицах, можно оптимизировать, расположив все три числа ($Ч_{эмп}$, $Ч_{кр1}$, $Ч_{кр2}$) на оси значимости.

Ось значимости – это прямая, на левом конце которой располагается 0, хотя он, как правило, не отмечается на самой этой прямой, и слева направо идет увеличение числового ряда. На этой прямой выделено три зоны: левая зона – зона незначимости, правая – зона значимости, промежуточная зона – зона неопределенности. Границами всех трех зон являются $Ч_{кр1}$ для $P = 0{,}05$ и $Ч_{кр2}$ для $P = 0{,}01$, как это показано на рис. 4.

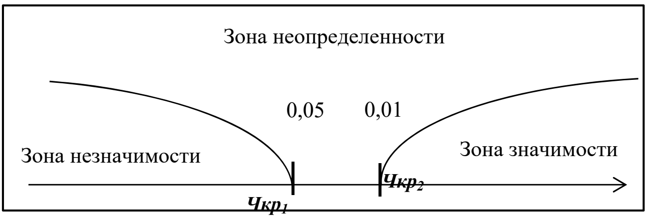

Рис. 4. Ось значимости

III. Выбор статистической гипотезы в зависимости от расположения $Ч_{эмп}$ на оси значимости.

Отметим, что $Ч_{эмп}$, подсчитанное по определенному статистическому методу, должно обязательно попасть в одну из трех зон. Выделяют четыре случая расположения $Ч_{эмп}$ на оси значимости:

1) $Ч_{эмп}$ попало в зону незначимости (рис. 5). В этом случае принимается гипотеза H_0 об отсутствии различий;

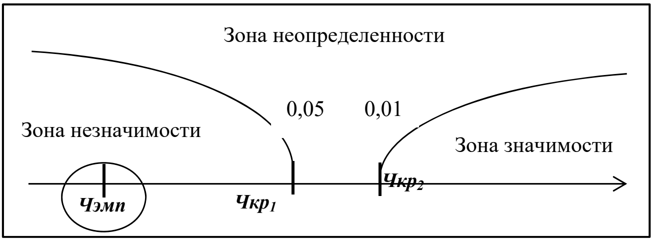

Рис. 5. Ось значимости

2) $Ч_{эмп}$ попало в зону значимости (рис. 6). В этом случае принимается гипотеза H_1 о наличии различий, гипотеза H_0 отклоняется;

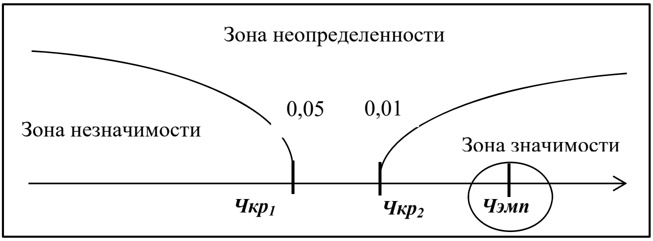

Рис. 6. Ось значимости

3) $Ч_{эмп}$ попало в зону неопределенности (рис. 7). В этом случае перед исследователем стоит дилемма. Так, в зависимости от важности решаемой задачи он может считать полученную статистическую оценку:

а) достоверной на уровне 5 % и принять тем самым гипотезу H_1;

б) недостоверной на уровне 1 % и принять тем самым гипотезу H_0.

Подчеркнем, что это именно тот случай, когда исследователь может допустить ошибки первого или второго рода, о которых пойдет речь в п. 4.3. В этих обстоятельствах рекомендуется увеличить объем выборки;

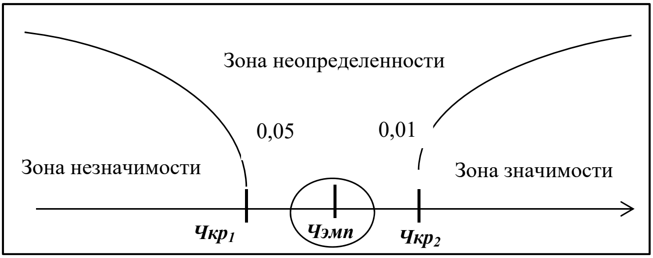

Рис. 7. Ось значимости

4) величина $Ч_{эмп}$ точно совпадает либо с $Ч_{кр1}$, либо с $Ч_{кр2}$.

Так, если $Ч_{эмп}$ точно совпадает:

– с $Ч_{кр1}$, то можно считать, что оценка достоверна точно на уровне в 5 % и принять гипотезу H_1 или, напротив, гипотезу H_0;
– $Ч_{кр2}$, то, как правило, принимается альтернативная гипотеза H_1 о наличии различий, а гипотеза H_0 отклоняется.

4.3. Ошибки первого и второго рода

Проверяя гипотезу при анализе данных психолого-педагогического эксперимента исследователь может столкнуться с ситуацией противоречия экспериментальных данных гипотезе H_0. В таком случае эта гипотеза должна быть отклонена. Если же экспериментальные данные согласуются с гипотезой H_0, она принимается как верная.

Так как принятие того или иного решения при проверке статистических гипотез основано на вероятностных суждениях, то решение об отклонении или принятии нулевой гипотезы может оказаться ошибочным. Логически возможны следующие случаи:

1) H_0 верна, но мы отклоняем H_0, т. е. принимаем H_1;
2) H_0 верна, и мы принимаем H_0, т. е. отклоняем H_1;
3) H_0 неверна, но мы принимаем H_0, т. е. отклоняем H_1;
4) H_0 неверна, и мы отклоняем H_0, т. е. принимаем H_1.

Случаи 2 и 4 отвечают верному решению, а случаи 1 и 3 – ошибочному. Принято считать, что в случае 1 мы допускаем ошибку первого рода, а в случае 3 – ошибку второго рода. Определено, что вероятность ошибки первого рода всегда не превосходит принятого уровня значимости. Поэтому роль уровня значимости по существу заключается в том, что он обеспечивает заранее малую вероятность ошибочных решений при отклонении нулевой гипотезы [6; 11].

Чтобы минимизировать число ошибок при проведении эксперимента, специалисты советуют увеличить объем выборки.

4.4. Этапы принятия статистического решения

В практике процесс принятия статистического решения принято делить на восемь этапов [13].

1. Формулирование нулевой и альтернативной гипотез.
2. Определение объема выборки **N**.
3. Выбор соответствующего уровня значимости или вероятности отклонения нулевой гипотезы. Это может быть величина, меньшая или равная 0,05 (5 % уровень значимости). В зависимости от важности исследования можно выбрать уровень значимости в 0,1 или даже в 0,001.
4. Выбор статистического метода, непосредственно зависящий от того, какая психолого-педагогическая проблема решается при проведении исследования.
5. Вычисление необходимого эмпирического значения по экспериментальным данным в соответствии с выбранным статистическим методом.
6. Нахождение в таблице приложения критических значений, подходящих выбранному статистическому методу и соответствующих уровню значимости для $P = 0,05$ и для $P = 0,01$.
7. Построение оси значимости и нанесение на нее табличных критических значений и эмпирического значения $\mathbf{Ч}_{эмп}$.
8. Формулирование принятого статистического решения (выбор соответствующей гипотезы H_1 или H_0).

4.5. Параметрические и непараметрические критерии

В целом критерии различий можно условно разделить на две группы: параметрические и непараметрические.

Критерий различия, основанный на определенном типе распределения генеральной совокупности (как правило, нормальном) или использующий параметры данной совокупности (средние, дисперсии и т. д.), называют *параметрическим*.

Так, всесторонне используемые на практике критерии, основанные на t-распределении Стьюдента и F-распределении Фишера – Снедекора, являются примерами параметрических критериев. Необходимо отметить, что данные критерии исходят из предположения о нормальном законе распределения случайных величин в генеральных совокупностях [8; 11].

Непараметрическим называют критерий различия, не основанный на предположении о типе распределения генеральной совокупности и не использующий параметры данной совокупности. В практике непараметрические критерии также называют критерием, свободным от распределения.

В сравнении с непараметрическими методами параметрические считаются более сильными.

Возможности практического применения параметрических и непараметрических критериев существенно зависят от характера конкретных измерений, осуществляемых исследователем. При использовании параметрических критериев необходимо применять количественные измерения, т. е. измерения по шкале интервалов или шкале отношений. Большинство непараметрических критериев может быть применено и в том случае, когда результат выборок измерен при помощи шкал наименования и порядка.

Контрольные вопросы

1. Дайте определение нулевой и альтернативной гипотез. Приведите примеры этих гипотез в психолого-педагогических исследованиях.
2. Назовите уровни статистической значимости, наиболее часто используемые в психолого-педагогических исследованиях.
3. Объясните, что означает нижняя граница уровня статистической значимости, равная 0,05.
4. Укажите, какие уровни значимости, подсчитываемые непосредственно в процессе работы с соответствующим статистическим методом, используются в современных статистических пакетах на ЭВМ.

5. Сформулируйте правило принятия статистического вывода, которое использует исследователь при обработке экспериментальных данных.
6. Перечислите все случаи расположения эмпирического значения на оси значимости и поясните, как оно влияет на выбор статистической гипотезы.
7. Охарактеризуйте все логически возможные случаи принятия того или другого решения при проверке статистических гипотез, которое может оказаться ошибочным. В каких из них, возможно допустить ошибку первого рода или ошибку второго рода?
8. Каким образом можно минимизировать ошибки при отклонении нулевой гипотезы в ходе принятия статистического вывода?
9. Сформулируйте этапы принятия статистического решения при обработке результатов психолого-педагогического исследования.
10. Назовите группы критериев, на которые делятся все критерии различий. В чем заключается их отличие? Поясните, какую группу критериев различий чаще всего используют при проведении психолого-педагогического исследования и почему.

§ 5. СРАВНЕНИЕ РЕЗУЛЬТАТОВ ДВУХ ЗАВИСИМЫХ ВЫБОРОК

Отметим, что при подготовке эксперимента исследователю необходимо составить характеристики выборок, используя следующий алгоритм [13]:

1) указать вид выборки: связная (зависимая) или несвязная (независимая);
2) описать однородность или неоднородность выборки;
3) определить объем выборки и выбрать соответствующий критерий с учетом известных ограничений;
4) в случае если выбранный критерий не показал никаких различий, подобрать другой, более трудоемкий критерий;
5) среди нескольких имеющихся критериев выбрать те, которые наиболее оптимально используют информацию, содержащуюся в экспериментальных данных.

Малый объем выборки предполагает увеличение величины уровня значимости (не менее 1 %), так как низкий уровень значимости может сказаться на достоверности полученных результатов.

5.1. Критерий знаков *G*

Существует целый ряд критериев различия, которые могут использоваться исследователем при проведении психолого-педагогического эксперимента. Так, чаще всего на практике используется критерий знаков G для сравнения состояния некоторого свойства у членов двух зависимых выборок на основе измерений, которые были сделаны по шкале не ниже порядковой. Данный критерий относят к непараметрическим и применяют только для связных выборок. Критерий знаков G дает возможность установить, насколько однонаправленно изменяются значения признака при повторном измерении связной однородной выборки. Критерий знаков наиболее применим к данным, которые были получены в ранговой, интервальной и шкале отношений.

При работе с данным критерием используется понятие сдвига. Сдвигом является величина разности между значениями признака одного и того же участника эксперимента «после» и «до».

Необходимо обратить внимание на то, что критерий знаков G предназначен для установления ситуации изменения значения признака при повторном измерении связной выборки: в сторону увеличения или уменьшения.

Введем два необходимых обозначения:

1) сумма сдвигов, получившаяся наибольшей, называемая типичным сдвигом, обозначается буквой n. Типичный сдвиг используется при работе с соответствующей таблицей для данного критерия, в которой приводятся критические величины 5 и 1 % уровней значимости данного критерия;

2) сумма сдвигов, получившаяся наименьшей, которая носит название нетипичного сдвига и обозначается $G_{эмп}$. Величина ($G_{эмп}$) располагается на оси значимости.

В психолого-педагогических исследованиях типичный и нетипичный сдвиги могут рассматриваться как дополняющие друг друга.

Оценка статистической достоверности различий по критерию знаков производится по табл. 1 приложения. В данной таблице, в столбце, обозначенном 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. Вышеуказанные величины условно можно считать нетипичными сдвигами. Они обозначаются как G, и с ними сравнивают полученное значение нетипичного сдвига $G_{эмп}$.

Чтобы наиболее наглядно представить результаты эксперимента, необходимо построить ось значимости, на которой располагаются как величины критических сдвигов, так и величина $G_{эмп}$, т. е. величина нетипичного сдвига.

Изобразим на рис. 8 ось значимости, которая была получена в результате обработки экспериментальных данных проведенного психолого-педагогического исследования.

Так, по рисунку отчетливо видно, что $G_{эмп}$ попало в зону незначимости, т. е. полученный в эксперименте общий положительный сдвиг статистически недостоверен. Другими словами, знания испытуемых не изменились качественно при использовании данного способа воздействия (методика изучения и т. д.).

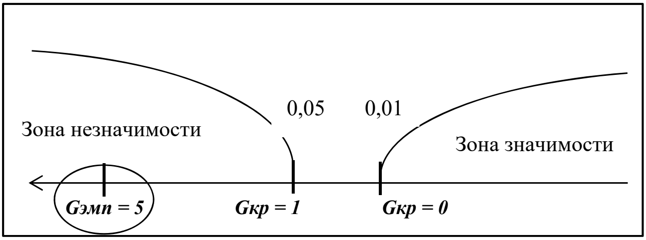

Рис. 8. Ось значимости

Отметим, что в данном случае ось значимости представлена в перевернутом виде. Нулевая отметка находится справа, а числовой ряд увеличивается в противоположную сторону. Эта особенность обоснована тем, что наличие нетипичных сдвигов уменьшает вероятность статистической достоверности суммарного сдвига. Такой вид расположения оси значимости справедлив для *T*-критерия Вилкоксона, критерия Макнамары и *U*-критерия Манна – Уитни.

Рассмотрим примеры использования знакового критерия в педагогике [11].

Пример 1. Студенты выполняли контрольную работу, которая была направлена на диагностику усвоения определенного понятия. Пятнадцати студентам, семь из которых получили оценку «2» и восемь – оценку «3», затем предложили воспользоваться программированным пособием. Оно было составлено с целью сформировать требуемое понятие у студентов с низким уровнем обучаемости. После изучения пособия студенты должны были снова выполнить аналогичную контрольную работу, оцениваемую по пятибалльной системе.

Цель данного эксперимента заключалась в апробации программированного пособия как средства повышения знаний студентов путем самообразования. Результаты двукратного выполнения работы студентами представляют измерения по шкале порядка (пятибалльная шкала) такого качества, как усвоение некоторого понятия.

Пример 2. Психолого-педагогический эксперимент был проведен, чтобы выяснить целесообразность системы упражнений некоторого раздела учебника. С этой целью выяснялось мнение преподавателей относительно числа данных упражнений в начале и в конце учебного года. В связи с тем, что мнения отдельных преподавателей резко расходились в отношении числа упражнений (одни требовали значительного увеличения числа упражнений, другие – уменьшения), было решено результаты опроса каждого преподавателя распределять только по двум категориям: увеличить число таких упражнений и уменьшить число таких упражнений. Всего было дважды опрошено 200 одних и тех же преподавателей. Были получены результаты двукратного измерения мнения каждого преподавателя. По шкале порядка (двухбалльная шкала) измерялось отношение учителя к системе упражнений учебника. Чтобы выявить тенденции изменения мнения участников эксперимента относительно необходимого числа упражнений, в данных условиях целесообразно было применить знаковый критерий.

5.2. Парный Т-критерий Вилкоксона

Помимо критерия знаков G в процессе решения задач, использующих сравнение двух рядов чисел, экспериментатор может применять парный T-критерий Вилкоксона. Данный критерий позволяет выявить не только направленность изменений, но и их выраженность, т. е. он позволяет установить, насколько сдвиг показателей в определенном направлении является более интенсивным, чем в другом.

T-критерий Вилкоксона имеет в основе принцип ранжирования абсолютных величин разности между двумя рядами выборочных значений в первом и втором эксперименте (до и после определенного воздействия). Ранжирование абсолютных величин означает, что знаки разностей не учитываются, однако в дальнейшем наряду с общей суммой рангов находится отдельно сумма рангов как для положительных, так и для отрицательных сдвигов. В случае когда интенсивность сдвига в одном из направлений оказывается большей, то и соответствующая сумма рангов также оказывается больше. Этот сдвиг, как и в случае критерия знаков, называется типичным, а противоположный, меньший по сумме рангов сдвиг – нетипичным. Как и для критерия знаков, эти два сдвига дополняют друг к друга. T-критерий Вилкоксона основывается на величине нетипичного сдвига, который обозначается $T_{эмп}$.

Отметим, что в данном критерии $T_{\text{эмп}}$ равно сумме рангов нетипичных сдвигов, а значения критических величин ($T_{\text{кр}}$) определяются с учетом общего числа испытуемых по соответствующей таблице для данного критерия (табл. 2 прил.).

Применение *T*-критерия Вилкоксона предполагает соблюдение следующих условий:

1) измерение может быть проведено во всех шкалах, кроме номинальной;
2) выборка должна быть связной;
3) число элементов в сравниваемых выборках должно быть равным;
4) *T*-критерий Вилкоксона может применяться при численности выборки от 5 до 50 (на бóльшую величину таблица достоверности не рассчитана).

Отметим, что *T*-критерий Вилкоксона, применяется в тех случаях, когда измерения выполнены хотя бы по интервальной шкале [13].

Рассмотрим примеры использования знакового критерия в педагогике [11].

Пример 1. Было проведено психолого-педагогическое исследование, которое должно было определить влияние просмотра некоего диафильма на уровень усвоения исследуемого понятия. Проверка усвоения понятия проводилась с помощью 10 контрольных заданий, рассчитанных на среднего студента. Работа выполнялась дважды одними и теми же 12 студентами – до просмотра и после просмотра диафильма. Выполнение работы каждым студентом оценивалось числом верных ответов, т. е. в данном эксперименте отметки студентов могли иметь значение от 0 до 10.

Пример 2. Группа студентов выполняла тестовую работу из 19 заданий. Выполнение работы каждым студентом оценивалось числом верных ответов, поэтому оценки студентов могли иметь значение от 0 до 19. Для выявления медианы оценок студентов данной группы было отобрано методом случайного отбора 16 студентов. Оценки этих студентов и были использованы для определения значения медианы, т. е. для проверки гипотезы вида H_0: медиана $M_d = c$ (c – заданное число).

5.3. Критерий Макнамары

Критерий Макнамары относится к числу непараметрических критериев и предназначен для работы с данными, полученными в самой простой из номинальных – в дихотомической шкале.

В психолого-педагогических исследованиях пары (x_i, y_i) могут быть результатами измерения состояния одного и того же свойства у одного и того же ученика до и после применения некоторого педагогического средства, причем x_i – состояние свойства до применения этого средства, а y_i – после его применения; x_i, y_i – измерения по шкале наименований, имеющей две категории, обозначенные «0» и «1». В связи с этим пары (x_i, y_i) могут быть только четырех видов (0, 0), (0, 1), (1, 0), (1, 1).

Для использования этого критерия данные суммируются в виде четырехпольной таблицы (табл. 3), которая называется «таблица 2×2». Поля в этих таблицах обозначаются латинскими буквами *A*, *B*, *C*, *D* или *a*, *b*, *c*, *d* [11].

Для применения критерия Макнамары необходимо соблюдать следующие условия [13]:

1) измерение должно быть проведено в дихотомической шкале;

2) выборка должна быть связной;

3) при количестве измерений $n \leq 20$ для определения величины $M_{эмп}$ используется табл. 3 приложения.

Отметим, что $n = B + C$; $m = \min(B, C)$, т. е. m – это наименьшая из величин B и C. По табл. 3 приложения на пересечении строк m и n находится величина $M_{\text{эмп}}$.

Величины $M_{\text{кр}}$ постоянны и равны 0,025 для 5 % уровня значимости и 0,005 для 1 % уровня значимости;

4) при количестве измерений $n > 20$ $M_{\text{эмп}}$ вычисляется по формуле

$$M_{\text{эмп}} = \frac{(B-C)^2}{(B+C)},$$

а величины $M_{кр}$ постоянны и равны 3,841 для 5 % уровня значимости и 6,635 для 1 % уровня значимости.

Таблица 3

Четырехпольная таблица данных эксперимента

		Классификация y_i		
		$y_i = 0$	$y_i = 1$	
Классификация x_i	$x_i = 0$	A (*число пар, у которых* $x_i = 0, y_i = 0$)	B (*число пар, у которых* $x_i = 0, y_i = 1$)	A + B
	$x_i = 1$	C (*число пар, у которых* $x_i = 1, y_i = 0$)	D (*число пар,* у которых $x_i = 1, y_i = 1$)	C + D
		A + C	B + D	

Рассмотрим примеры использования знакового критерия в педагогике [11].

Пример 1. Был проведен эксперимент, ставящий целью выявить эффективные формы профориентационной работы, проводимой в выпускных классах школы. Был разработан комплекс циклов лекций, бесед, различающихся по содержанию и методике, организованы экскурсии учащихся на предприятия. Степень эффективности каждого цикла определялась процессом формирования у учащихся положительного мнения об определенной профессии. Например, первый цикл был направлен на формирование положительного отношения учащихся к профессии юриста. Учащимся был задан вопрос «Нравится ли вам профессия юриста?» до проведения профориентационного цикла и после него. Предполагались ответы «да» или «нет».

Пример 2. В ходе проведения эксперимента исследователи поставили цель – выявить, насколько влияет форма контроля знаний по определенному разделу программы на результаты контрольного опроса. Для выявления уровня сформированных знаний студентам были предложены письменная работа, состоящая из трех заданий, и тест из 20 вопросов, основанные на идентичном содержательном материале. Основываясь на результатах, экспериментаторы распределили студентов на две категории: усвоившие и неусвоившие. По результатам письменной работы в первую группу были отнесены студенты, получившие оценки «3», «4», «5». По результатам выполнения теста в первую группу были отнесены студенты, верно ответившие на 13 и более вопросов, ко второй группе были отнесены остальные студенты.

Контрольные вопросы

1. Назовите объекты изучения в психолого-педагогических исследованиях.
2. Сформулируйте условия неприменимости критерия знаков.
3. Поясните, почему парный *T*-критерий Вилкоксона является более мощным, чем критерий знаков.
4. Укажите отличия *T*-критерия Вилкоксона и критерия знаков при нахождении эмпирического и критических значений.
5. Уточните, какой должна быть выборка при применении *T*-критерия Вилкоксона.
6. Сформулируйте вывод в случае, когда эмпирическое значение по критерию знаков попало в зону незначимости.
7. Уточните, какие навыки необходимы в процессе применения критерия Макнамары при работе с таблицами критических величин.
8. Можно ли осуществлять обработку результатов анкетирования с помощью критерия Макнамары? Ответ обоснуйте.
9. Перечислите способы подсчета эмпирического значения по критерию Макнамары.
10. Поясните, почему при работе с критерием Макнамары используется только дихотомическая шкала.

§ 6. СРАВНЕНИЕ РЕЗУЛЬТАТОВ ДВУХ НЕЗАВИСИМЫХ ВЫБОРОК

6.1. *U*-критерий Вилкоксона – Манна – Уитни

Данный критерий необходимо применять в том случае, когда исследователь ставит цель выявить различия в распределениях изучаемого свойства у объектов двух совокупностей, основываясь на сравнении результатов изучения данного свойства у членов независимых выборок, которые получены из этих совокупностей. В качестве показателей центральной тенденции рассматриваются медиана и среднее значение.

С помощью данного критерия проверяется одновременно предположение о различии медиан, а также средних значений в рассматриваемых совокупностях. Следует иметь в виду, что различие медиан или средних значений свидетельствует о наличии тенденции членов первой совокупности превосходить по состоянию изучаемого свойства члены второй совокупности либо уступать им. Процесс упорядочения объектов возможен лишь в случае, когда шкала измерений изучаемого свойства не ниже порядковой.

Критерий может найти применение в психолого-педагогических исследованиях, так как позволяет проверить одновременно несколько предположений о характере различия распределений некоторого свойства в двух совокупностях (различие медиан, средних значений распределений обеих совокупностей, а также наличие тенденции объектов первой совокупности быть в среднем больше или, наоборот, меньше членов другой совокупности) на основе порядковых измерений состояния этого свойства.

Таким образом, *U*-критерий более тонко улавливает различие распределений некоторого свойства в двух совокупностях по сравнению с медианным критерием [11].

Отметим, что при обработке результатов эксперимента с помощью *U*-критерия необходимо полученные данные объединить, т. е. представить как один ряд и упорядочить его по возрастанию входящих в него величин. Для *U*-критерия Вилкоксона – Манна – Уитни не так важны численные значения данных, как порядок расположения этих данных [13].

Если предварительно обозначить каждый элемент группы 1 (3; 12; 83; 83) символом x, а группы 2 (2; 9; 43; 75; 75; 90) – y, то общий упорядоченный ряд можно представить в следующем виде:

y	x	y	x	y	y	y	x	x	y
2	3	9	12	43	75	75	83	83	90

Если бы упорядоченный ряд, составленный по данным двух выборок, принял бы вид $x\,x\,x\,x\;\; y\,y\,y\,y\,y\,y$, то такие выборки значительно отличались бы между собой, а такое расположение данных называлось бы идеальным. Критерий U основан на подсчете нарушений в расположении чисел в упорядоченном экспериментальном ряду по сравнению с идеальным рядом.

Инверсия – это всякое нарушение порядка идеального ряда. Одним нарушением (одной инверсией) называют такое расположение чисел, когда перед некоторым числом первого ряда стоит только одно число второго ряда. Если перед некоторым числом первого ряда стоят два числа второго ряда, то возникают две инверсии и т. д. Удобнее подсчитать число инверсий, когда исходные данные представлены в виде таблицы (табл. 4).

Таблица 4

Упорядоченное объединение экспериментальных данных двух групп испытуемых в порядке их возрастания

1	2	3	4
Группа 1	Группа 2	Инверсии X/Y	Инверсии Y/X
–	2	–	0
3	–	1	–
–	9	–	1
12	–	2	–
–	43	–	2
–	75	–	2
–	75	–	2
83	–	5	–
83	–	5	–
–	90	–	4
Суммы инверсий		13	11

Пропуск в 1-м столбце означает, что во 2-м столбце имеется число, занимающее промежуточное значение по отношению к числам 1-го столбца, ограничивающим пропуск. То же самое верно и для 2-го столбца.

Инверсии X/Y – это инверсии 1-го столбца по отношению ко 2-му столбцу; инверсии Y/X – это инверсии 2-го столбца по отношению к 1-му столбцу.

В приведенном выше примере одинаковые величины встречаются только внутри каждой группы испытуемых.

В этом случае $U_{эмп}$ есть минимальная величина из сумм инверсий. Согласно данным примера $U_{эмп} = \min(U(X/Y), U(Y/X)) = 11$. $U_{кр}$ находятся по табл. 4 приложения. Данный случай относится к первому способу расчета по критерию U.

В случае когда две или большее количество одинаковых величин входят в обе сравниваемые группы испытуемых, в столбцах 3 и 4 табл. 4 проставляются ранги так, как будто оба этих столбца образуют собой один упорядоченный ряд. При этом ранги для чисел 1-го столбца помещаются в столбец 3, ранги 2-го столбца – в 4-й столбец.

$U_{эмп}$ во втором случае вычисляется по формуле

$$U_{эмп} = n_1 \cdot n_2 + n_x \cdot \frac{n_x+1}{2} - R_{\max},$$

где n_1 – численное значение 1-й выборки; n_2 – численное значение 2-й выборки; R_{max} – наибольшая по величине сумма рангов (из столбцов 3 или 4); n_x – количество испытуемых в группе с большей суммой рангов.

$U_{кр}$ также находятся по табл. 4 приложения. Рассмотренный случай относится ко второму способу расчета по критерию U.

Для применения критерия Вилкоксона – Манна – Уитни необходимо соблюдать ряд условий [13].

1. Измерение должно быть проведено в шкале интервалов и отношений.
2. Выборки должны быть несвязными.
3. Нижняя граница применимости критерия: $n_1 \geq 3$ и $n_2 \geq 3$ или $n_1 = 2$, а $n_2 \geq 5$.
4. Верхняя граница применимости критерия: $n_1 \leq 60$ и $n_2 \leq 60$.

6.2. Медианный критерий

Медианный критерий необходим, когда требуется выявить в различных центральных тенденциях состояния некоторого свойства в двух совокупностях, основываясь на изучении членов двух независимых выборок из предложенных совокупностей. В данном случае показателем центральной тенденции является медиана измерений изучаемого свойства в каждой из выборок.

Пусть имеются две серии наблюдений, полученных при рассмотрении двух независимых выборок объема n_1 и n_2; x_i – результаты измерений некоторого свойства у объектов первой выборки, составленной из членов одной совокупности; y_i – результаты измерений того же свойства у обектов второй выборки, составленной из членов другой совокупности.

Обе серии наблюдений объединяют в одну выборку, объем которой равен $(n_1 + n_2)$ и обозначается N. Определяется медиана этой выборки m, после чего члены каждой выборки объемами n_1 и n_2 распределяются на две категории: больше общей медианы ($>m$) и меньше или равны общей медиане ($\leq m$).

На основе полученных результатов составляется четырехпольная таблица (табл. 5).

Таблица 5

Четырехпольная таблица данных эксперимента

	Выборка № 1	Выборка № 2
$>m$	A число x_i, больших m	B число y_i, больших m
$\leq m$	C число x_i, меньше или равных m	D число y_i, меньше или равных m

Для проверки гипотез с помощью медианного критерия на основе наблюдений, записанных в форме таблицы 2×2, подсчитывается величина $T_{эмп}$ по формуле

$$T = \frac{N \cdot \left(|A \cdot D - B \cdot C| - \frac{N}{2}\right)^2}{(A + B) \cdot (C + D) \cdot (A + C) \cdot (B + D)}.$$

Критические значения для разных уровней значимости по данному критерию находятся по табл. 5 приложения со степенью свободы $\nu = 1$.

Не рекомендуется использовать медианный критерий для проверки гипотез, если:

а) $20 < n_1 + n_2 < 40$ и хотя бы одно из значений A, B, C, D, записанных в таблице 2×2, меньше 5;

б) около середины ряда, составленного из упорядоченных по возрастанию значений обеих выборок, будет расположено значительное число одинаковых значений, принадлежащих обеим выборкам. Если они группируются около середины ряда, но составляют небольшую часть суммарного объема обеих выборок, то медиана подсчитывается по традиционному правилу и применение критерия возможно.

Иногда в случае группировки нескольких одинаковых значений, принадлежащих обеим выборкам, около середины ряда можно подсчитывать медиану (m) данного ряда по следующей формуле:

$$m = L + \frac{\frac{N}{2} - \sum f}{f} \cdot h,$$

где N – сумма объемов двух выборок; L – нижняя граница интервала значений ряда, к которому принадлежит значение ряда, стоящее на $\frac{N}{2}$ месте (при N нечетном берется значение ряда, стоящее на $\frac{N+1}{2}$ месте); Σf – число значений ряда, меньших L; f – число значений ряда, принадлежащих интервалу, включающему значение ряда, стоящее на $\frac{N}{2}$ месте; h – величина интервала, выбранная при группировке значений ряда.

Использование вышеуказанной формулы целесообразно, если значения обеих выборок записаны в форме интервального ряда. При этом необходимо иметь в виду, что использование формулы возможно только в тех случаях, когда изучаемое свойство педагогического явления измерено по шкале не ниже интервальной. Следует отметить, что применение формулы в подобных случаях имеет смысл, когда результаты измерений выражены дробными числами.

Однако при выполнении этих требований возможно использование критериев Вилкоксона – Манна – Уитни или Колмогорова – Смирнова, позволяющих получить более достоверные выводы по сравнению с медианным критерием. Использование медианного критерия целесообразно лишь в случае измерений изучаемого свойства по шкале порядка с достаточно большим числом градаций.

Чтобы применить медианный критерий, нужно соблюдать следующие условия:

1) обе используемые выборки должны быть случайными выборками из некоторых совокупностей;
2) выборки независимы, и члены каждой выборки независимы между собой;
3) шкала измерений не ниже порядковой, так как при измерениях по шкале наименований невозможно нахождение медианы;
4) число членов в обеих выборках должно быть в сумме больше 20, т. е. $n_1 + n_2 > 20$ [11].

6.3. Критерий хи-квадрат

При применении критерия согласия (хи-квадрат) исследователь может решить большое число разных задач, также исходные данные для этого критерия могут быть получены в любой шкале, начиная со шкалы наименований.

Критерий хи-квадрат используется в двух вариантах:

– как расчет согласия эмпирического распределения и предполагаемого теоретического; в этом случае проверяется гипотеза H_0, указывающая на отсутствие различий между теоретическим и эмпирическим распределениями;

– как расчет однородности двух независимых экспериментальных выборок; в этом случае проверяется гипотеза H_0, указывающая на отсутствие различий между двумя эмпирическими (экспериментальными) распределениями.

Критерий построен таким образом, что при полном совпадении экспериментального и теоретического (или двух экспериментальных) распределений величина $\chi^2_{эмп}$ (хи-квадрат эмпирическое) равна нулю, и чем больше разница между сравниваемыми распределениями, тем больше величина эмпирического значения хи-квадрат.

Основная расчетная формула критерия хи-квадрат

$$\chi^2_{\text{эмп}} = \sum_{i=1}^{k} \frac{(f_{\text{э}} - f_m)^2}{f_m},$$

где $f_{\text{э}}$ – эмпирическая частота; f_m – теоретическая частота; k – количество разрядов признака.

Расчетная формула критерия хи-квадрат для сравнения двух эмпирических распределений в зависимости от вида представленных данных может иметь следующий вид:

$$\chi^2_{\text{эмп}} = \frac{1}{N \cdot M} \sum_{i=1}^{k} \frac{(N \cdot x_i - M \cdot y_i)^2}{x_i + y_i},$$

где N и M – соответственно число элементов в первой и во второй выборке. Эти числа могут совпадать, а могут быть и различными.

Для критерия хи-квадрат оценка уровней значимости определяется по числу степеней свободы, которое обозначается греческой буквой ν (ню) и в большинстве случаев вычисляется по формуле: $\nu = k - 1$, где каждый раз определяется по выборочным данным и представляет собой число элементов в выборке. Если при расчете критерия используется таблица экспериментальных данных, то величина ν рассчитывается следующим образом: $\nu = (k - 1) \cdot (c - 1)$, где k – число строк, а c – число столбцов.

Отметим, что в случае когда сравниваются две выборки, в которых значений переменных так много, что предыдущий способ сравнения оказывается трудновыполнимым, расчет $\chi^2_{\text{эмп}}$ может производиться по другим формулам.

Для применения критерия хи-квадрат необходимо, чтобы были соблюдены следующие условия:

1) измерение может быть проведено в любой шкале;
2) выборки должны быть случайными и независимыми;
3) желательно, чтобы объем выборки был больше или равен 20, так как при увеличении объема выборки повышается точность критерия;
4) теоретическая частота для каждого выборочного интервала не должна быть меньше 5;
5) сумма наблюдений по всем интервалам должна быть равна общему количеству наблюдений;

6) данные, приведенные в таблице критических значений критерия хи-квадрат (табл. 5 приложения), рассчитаны для числа степеней свободы ν, которое каждый раз вычисляется по определенным правилам. Число степеней свободы определяется по формуле $\nu = c - 1$, где c – число альтернатив (признаков, значений, элементов) в сравниваемых переменных. Для таблиц число степеней свободы выводится по формуле $\nu = (k - 1) \cdot (c - 1)$, где k – число столбцов, c – число строк [13].

Контрольные вопросы

1. Поясните назначение *U*-критерия Вилкоксона – Манна – Уитни.
2. Дайте определение понятия инверсии. При работе с каким критерием оно применяется?
3. С чем связано существование различных случаев применения *U*-критерия Вилкоксона – Манна – Уитни? Отличаются ли в них способы нахождения эмпирического значения?
4. Объясните причины применения критерия Вилкоксона – Манна – Уитни для связных выборок.
5. Какой из критериев: Вилкоксона – Манна – Уитни или медианный критерий – более тонко улавливает различие распределений некоторого свойства в двух совокупностях?
6. Поясните, каким образом используется четырехпольная таблица данных эксперимента в медианном критерии.
7. Сформулируйте случаи, в которых не рекомендуется использовать медианный критерий для проверки гипотез.
8. Назовите, в каком из критериев, рассмотренных в этом параграфе, исходные данные могут быть получены в любой шкале.
9. Опишите варианты использования критерия хи-квадрат в психолого-педагогических исследованиях. Приведите основные расчетные формулы для каждого из них.
10. Поясните, с чем связаны различные варианты подсчета числа степеней свободы, применяемого в таблице критических значений по критерию хи-квадрат.

§ 7. ПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ РАЗЛИЧИЙ. КРИТЕРИЙ СТЬЮДЕНТА

Данные критерии носят название «параметрические», потому что в формулу их расчета включаются такие параметры выборки, как среднее, дисперсия и др. В психолого-педагогических исследованиях чаще всего применяют *t-критерий Стьюдента*, оценивающий различия средних для двух выборок, и *F-критерий Фишера*, который оценивает различия между двумя дисперсиями.

Рассмотрим t-критерий Стьюдента.

Критерий Стьюдента направлен на оценку различий величин средних $\bar{X}$ и $\bar{Y}$ двух выборок X и Y, которые распределены по нормальному закону. Важная особенность данного критерия – это возможность его применения в широком спектре исследований. Данный критерий может быть использован для сопоставления средних у связных и несвязных выборок, при этом не обязательно использовать равные по величине выборки.

Случай несвязных выборок

В общем случае формула для расчета по t-критерию Стьюдента такова:

$$t_{\text{эмп}} = \frac{|\bar{X} - \bar{Y}|}{S_d},$$

где $S_d = \sqrt{S_x^2 + S_y^2}$.

Отметим, что для равночисленных выборок (при $n_1 = n_2 = n$) S_d будет вычисляться по формуле

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

для неравночисленных выборок (при $n_1 \neq n_2$) – по следующей формуле:

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

В обоих случаях подсчет числа степеней свободы осуществляется по формуле

$$k = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2,$$

где n_1 и n_2 соответственно величины первой и второй выборки. При численном равенстве выборок $k = 2 \cdot n = 2$.

Случай связных выборок

В случае связных выборок с равным числом измерений в каждой можно использовать более простую формулу t-критерия Стьюдента. Вычисление значения $t_{эмп}$ осуществляется по формуле

$$t_{эмп} = \frac{\bar{d}}{S_d},$$

где $\bar{d} = \frac{\sum d_i}{S_d} = \frac{\sum(x_i - y_i)}{n}$, причем $d_i = x_i - y_i$ – разности между соответствующими значениями переменной X и переменной Y, а $\bar{d}$ – среднее этих разностей; S_d вычисляется, в свою очередь, по формуле

$$S_d = \sqrt{\frac{\sum {d_i}^2 - \frac{\left(\sum d_i\right)^2}{n}}{n \cdot (n-1)}}.$$

Число степеней свободы k находим по формуле $k = n - 1$.

При применении t-критерия Стьюдента должны выполняться следующие условия:

1) измерение может быть проведено в шкале интервалов и отношений;
2) сравниваемые выборки должны быть распределены по нормальному закону [13];

Критические значения для данного критерия находим с учетом определенного числа степеней свободы по табл. 6 приложения.

Контрольные вопросы

1. Назовите параметры выборки, которые могут быть включены в формулу расчета по параметрическим критериям.
2. Перечислите параметрические критерии, которые чаще всего применяют в психолого-педагогических исследованиях.
3. Опишите параметры для двух выборок, с помощью которых оцениваются различия в t-критерии Стьюдента и F-критерии Фишера.
4. Каким образом должны быть распределены параметры двух выборок в t-критерии Стьюдента?

5. Приведите формулы нахождения эмпирического значения для связных и несвязных выборок по данному критерию.
6. Объясните, с чем связаны различия при подсчете числа степеней свободы для связных и несвязных выборок по *t*-критерию Стьюдента.
7. Поясните, в чем состоит отличие в нахождении стандартного отклонения по данному критерию в зависимости от числа элементов в двух несвязных выборках.
8. Охарактеризуйте, как вычисляется число степеней свободы для случая двух связных выборок по *t*-критерию Стьюдента.
9. Обоснуйте применение процедуры измерения по данному критерию с помощью шкалы интервалов или отношений.
10. Поясните, как находятся критические значения для *t*-критерия Стьюдента.

§ 8. ПРИМЕНЕНИЕ КОРРЕЛЯЦИОННОГО И РЕГРЕССИОННОГО АНАЛИЗА В ПСИХОЛОГО-ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЯХ. КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ ПИРСОНА

В процессе проведения психолого-педагогического исследования экспериментатор зачастую задается вопросом, насколько взаимосвязаны две или большее количество переменных в одной или нескольких изучаемых группах. Чтобы охарактеризовать связи в данных исследованиях между переменными величинами, используют понятие функции F, которая ставит в соответствие каждому конкретному значению независимой переменной X определенное значение зависимой переменной Y. Полученная зависимость обозначается как $Y = F(X)$, где X – аргумент, Y – соответствующее ему значение функции $F(X)$. Подобного вида однозначные зависимости между переменными величинами X и Y называются функциональными. Подобные однозначные, или функциональные, связи между переменными величинами встречаются достаточно редко.

Связи между исследуемыми признаками могут иметь не функциональный, а статистический характер. Имеется в виду ситуация, когда в среднем определенному значению одного признака соответствует целый спектр значений, а не одно какое-либо значение. При этом многообразие значений распределяется в вариационный ряд числовых значений, рассматриваемых в качестве зависимой переменной или функции. Подобного рода зависимость между переменными величинами называется корреляционной, или корреляцией. Корреляционная связь – это согласованное изменение двух признаков, которое указывает на то, что изменчивость одного признака находится в соответствии с изменчивостью другого.

Функциональные связи можно обнаружить и измерить на единичных и групповых объектах, но это невозможно проделать с корреляционными связями, применяемыми только на представительных выборках. Изменения признаков, имеющие определенные согласования, и корреляционная связь между ними, отражающая данный процесс, могут свидетельствовать не о зависимости этих признаков между собой, а о существовании зависимости обоих этих признаков от какого-то третьего признака или сочетания признаков, которые не рассматриваются в исследовании. Под зависимостью подразумевается влияние, а связь – это любые согласованные изменения, объясняемые большим количеством разнообразных причин.

Корреляционные связи не являются свидетельством причинно-следственной зависимости, они только указывают на то, что изменения одного признака сопровождают определенные изменения другого. Однако специалист, проводящий исследование, не может сделать вывод о том, что причина изменений заключается в одном из изучаемых признаков, ибо она может оказаться за пределами исследуемой пары признаков. Выделяют следующие виды корреляций: линейная, нелинейная, положительная и отрицательная (рис. 9).

Корреляция считается линейной, если при уменьшении или увеличении одной переменной X вторая переменная Y в среднем либо убывает, либо растет. Корреляция называется нелинейной в случае, когда при увеличении одной величины характер изменения второй не линеен, а подчиняется другим законам.

Корреляция считается положительной в том случае, когда при увеличении переменной X переменная Y в среднем также увеличивается. В том случае когда при увеличении X переменная Y уменьшается, данный процесс называют отрицательной корреляцией.

Иногда нет возможности установить какую-либо зависимость между переменными. Подобную ситуацию характеризуют как отсутствие корреляционной связи. Есть задачи, в которых встречается высокозначимая криволинейная связь, например, полиномиальная или гиперболическая.

Корреляционный анализ имеет следующие задачи:

1) установление направления (положительное или отрицательное) и формы (линейная, нелинейная) связи между варьирующими признаками;
2) измерение тесноты связи;
3) проверка уровня значимости полученных коэффициентов корреляции.

Зависимость между коррелирующими переменными X и Y, как и в математике, выражается при помощи формул и уравнений либо графически.

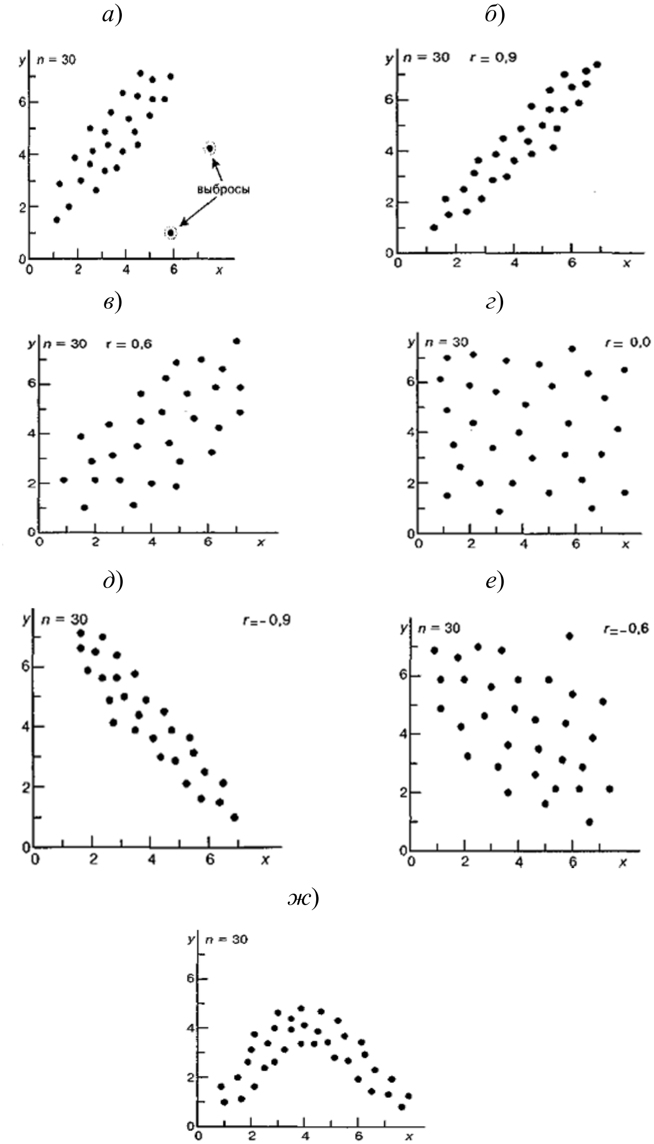

Рис. 9. Типы корреляционных связей: *а* – выбросы; *б* – положительная корреляция; *в* – может быть положительная корреляция; *г* – нет корреляции; *д* – отрицательная корреляция; *е* – может быть отрицательная корреляция; *ж* – нет линейной, но возможна криволинейная корреляция

Графики корреляционных зависимостей строят по уравнениям функций

$$\bar{Y}_x = F(X) \text{ или } \bar{X}_y = F(Y).$$

Приведенные функции называют уравнениями регрессии, в которых $\bar{Y}_x$ и $\bar{X}_y$ — так называемые условные средние арифметические переменных *X* и *Y*.

Переменные *X* и *Y* можно измерить в разных шкалах, этот факт и определяет выбор требуемого коэффициента корреляции (табл. 6).

Величина коэффициента линейной корреляции Пирсона не может превышать (+1) и быть меньше чем (−1). Эти два числа (+1) и (−1) представляют собой границы для коэффициента корреляции. Если при расчете экспериментатор получает величину большую (+1) или меньшую (−1), данный факт указывает на ошибку в вычислениях.

Таблица 6

Соотношение между типами шкал и соответствующими мерами связи

Тип шкалы		Мера связи
Переменная X	Переменная Y	
Интервальная или отношений	Интервальная или отношений	Коэффициент Пирсона
Ранговая, интервальная или отношений	Ранговая, интервальная или отношений	Коэффициент Спирмена
Ранговая	Ранговая	Коэффициент τ Кендалла
Дихотомическая	Дихотомическая	Коэффициент φ
Дихотомическая	Ранговая	Рангово-бисериальный R_{rb}
Дихотомическая	Интервальная или отношений	Бисериальный $R_{\text{бис}}$
Интервальная	Ранговая	Не разработан

В случае когда коэффициент корреляции по модулю получается близким к 1, это соответствует высокому уровню связи между переменными. Так, при корреляции переменной величины с самой собой величина коэффициента корреляции будет равна (+1). Данная связь характеризует прямо пропорциональную зависимость. Если же значения переменной *X* будут распложены в порядке возрастания, а те же значения (обозначенные теперь уже как переменная *Y*) будут располагаться в порядке убывания, то в этом случае корреляция между переменными X и Y будет равна точно (−1). Подобная величина коэффициента корреляции указывает на обратно пропорциональную зависимость.

При интерпретации получаемой связи имеет большое значение знак коэффициента корреляции. В случае когда коэффициент линейной корреляции имеет знак «плюс», связь между коррелирующими признаками имеет следующую особенность: большей величине одного признака (переменной) соответствует большая величина другого признака (другой переменной). Другими словами, если один показатель (переменная) увеличивается, то соответственно увеличивается и другой показатель (переменная). Вышеописанная зависимость называется прямо пропорциональной зависимостью.

Если же коэффициент линейной корреляции имеет знак «минус», значит, большей величине одного признака соответствует меньшая величина другого. То есть при наличии знака «минус» увеличению одной переменной (признака, значения) соответствует уменьшение другой переменной. Подобная зависимость называется обратно пропорциональной. В данном случае выбор переменной, у которой наблюдается тенденция к возрастанию, имеет произвольный характер. Это может быть как переменная X, так и переменная Y. В случае когда экспериментатор считает, что увеличивается переменная X, переменная Y будет соответственно уменьшаться, и наоборот. Приведенные выше положения необходимо четко усвоить, чтобы правильно интерпретировать полученную корреляционную зависимость.

Представляем общий вид формулы для подсчета коэффициента корреляции:

$$r_{xy} = \frac{\sum(x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \cdot \sum(y_i - \bar{y})^2}}, \tag{4}$$

где x_i – значения, принимаемые переменной X; y_i – значения, принимаемые переменной Y; $\bar{x}$ – средняя по X; $\bar{y}$ – средняя по Y.

Расчет коэффициента корреляции Пирсона предполагает, что переменные X и Y распределены нормально.

Формула (4) предполагает, что из каждого значения x_i переменной X должно вычитаться ее среднее значение $\bar{x}$, что неудобно. Поэтому для расчета коэффициента корреляции используют аналог формулы (4), который был выведен из данной формулы:

$$r_{xy} = \frac{\sum(x_i \cdot y_i) - \frac{\sum x_i \cdot \sum y_i}{n}}{\sqrt{S_x \cdot S_y}}, \qquad (5)$$

где $S_x = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$ и $S_y = \sum y_i^2 - \frac{(\sum y_i)^2}{n}$, или модификацию этой формулы

$$r_{xy} = \frac{n \cdot \sum(x_i \cdot y_i) - (\sum x_i \cdot \sum y_i)}{\sqrt{[n \cdot \sum x_i^2 - (\sum x_i)^2] \cdot [n \cdot \sum y_i^2 - (\sum y_i)^2]}}. \qquad (6)$$

Согласно формулам (5) и (6) необходимо подсчитать сумму каждой переменной, сумму квадратов каждой переменной и сумму последовательных произведений переменных друг на друга.

Отметим, что в формуле (4) встречается величина

$$\sum(x_i - \bar{x}) \cdot (y_i - \bar{y}). \qquad (7)$$

При делении на n (число значений переменной X или Y) данная величина называется ковариацией.

Отметим, что выражение (7) может быть подсчитано только в случае, когда число значений переменной X равно числу значений переменной Y и равно n. Формула (7) указывает на то, что в процессе расчета коэффициентов корреляции нельзя произвольно переставлять элементы в коррелируемых столбцах.

В случаях применения коэффициента корреляции Пирсона необходимо соблюдать следующие условия:

1) сравниваемые переменные должны быть получены в интервальной шкале или шкале отношений;
2) распределения переменных X и Y должны быть близки к нормальному;
3) число варьирующих признаков в сравниваемых переменных X и Y должно быть одинаковым;
4) таблицы уровней значимости для коэффициента корреляции Пирсона (табл. 7 приложения) рассчитаны для значений n от 5 до 1000. Оценка уровня значимости по таблицам осуществляется при числе степеней свободы $k = n - 2$ [13].

На практике при наличии связи между двумя переменными подразумевается, что она является вероятностной и графически выглядит, как облако рассеивания эллипсоидной формы. Этот эллипсоид также можно представить в виде прямой линии, или линии регрессии. Линия регрессии является прямой, построенной методом наименьших квадратов: сумма квадратов расстояний (вычисленных по оси *Y*) от каждой точки графика рассеивания до прямой является минимальной:

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

где y_i – истинное *i*-значение *Y*; $\hat{y}_i$ – оценка *i*-значения *Y* при помощи линии (уравнения) регрессии; $e_i = y_i - \hat{y}_i$, – ошибка оценки (рис. 10).

Уравнение регрессии имеет вид

$$\hat{y}_i = b \cdot x_i + a,$$

где *b* – коэффициент регрессии, задающий угол наклона прямой; *a* – свободный член, определяющий точку пересечения прямой оси *Y*.

Если известны средние (M_y; M_x), стандартные отклонения (σ_y; σ_x) и корреляция r_{xy}, то сумма квадратов ошибок минимальна, когда

$$b = r_{xy} \cdot \frac{\sigma_y}{\sigma_x}, a = M_y - b \cdot M_x \text{ [14].}$$

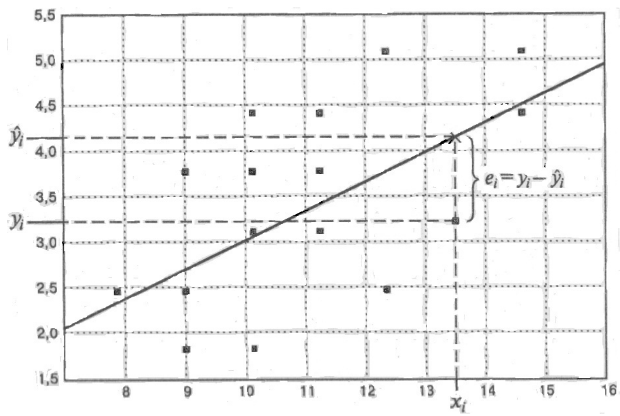

Рис. 10. Диаграмма рассеивания и линия регрессии
(e_i – ошибка оценки для одного из объектов)

Вместе с этим, О.Ю. Ермолаев при расчетах коэффициентов регрессии предлагает использовать более простой метод, заключающийся в решении системы уравнений

$$\begin{cases} a_0 \cdot N + a_1 \cdot \sum x_i = \sum y_i\,, \\ a_0 \cdot \sum x_i + a_1 \cdot \sum (x_i \cdot x_i) = \sum (y_i \cdot x_i). \end{cases}$$

В этой системе уравнений используются следующие обозначения:

1) $Y = a_0 + a_1 \cdot X$ – уравнение регрессии, в котором Y – зависимая переменная; X – независимая переменная; a_0 – свободный член; a_1 – коэффициент регрессии, или коэффициент, определяющий наклон линии регрессии по отношению к осям координат (рис. 11). Чем сильнее связь между X и Y, тем ближе линия регрессии к прямой AB, и наоборот, чем слабее связь между этими величинами, тем больше линия регрессии отклоняется от прямой AB. При отсутствии связи между X и Y линия регрессии Y по X (уравнение регрессии $Y = a_0 + a_1 \cdot X$) и линия регрессии X по Y (уравнение регрессии $X = b_0 + b_1 \cdot Y$) оказываются под прямым углом, $r_{xy} = 0$;
2) N – число элементов в переменной X или переменной Y;
3) Σx_i – сумма всех элементов переменной X;
4) Σy_i – сумма всех элементов переменной Y;
5) $\Sigma(x_i \cdot x_i)$ – произведение всех элементов переменной X друг на друга;
6) $\Sigma(y_i \cdot x_i)$ – попарное произведение всех элементов переменной X на соответствующие элементы переменной Y [13].

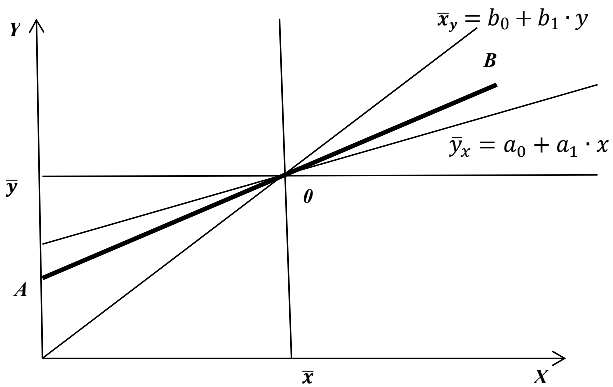

Рис. 11. Линии регрессии Y по X и X по Y в системе прямоугольных координат

В ситуации, когда на некоторой выборке измерены две переменные, коррелирующие друг с другом, можно, вычислив коэффициенты регрессии, получить определенную возможность предсказания неизвестных значений одной переменной (*Y* – *зависимая переменная*) по известным значениям другой переменной (*X* – *независимая переменная*).

Рассмотрим еще один случай применения коэффициента корреляции Пирсона.

Так, при изученияи связи между двумя переменными, представленными в дихотомической шкале, строят обычную четырехклеточную таблицу сопряженности (табл. 7).

В этом случае допустимо применение коэффициента корреляции Пирсона (формула (4)) непосредственно к исходным данным – двум переменным, представленным в дихотомической шкале и принимающим значение 0 или 1, измеренное для каждого члена выборки численностью *N*.

Таблица 7

Четырехклеточная таблица сопряженности

<table>
<tr><td colspan="2" rowspan="2">0</td><td colspan="2">Признак X</td><td>Итог</td></tr>
<tr><td>1</td><td></td><td></td></tr>
<tr><td rowspan="2">Признак Y</td><td>0</td><td>a</td><td>b</td><td>a+ b</td></tr>
<tr><td>1</td><td>c</td><td>d</td><td>c+ d</td></tr>
<tr><td>Итог</td><td></td><td>a + c</td><td>b + d</td><td>N</td></tr>
</table>

Результат применения коэффициента корреляции Пирсона к двум переменным, представленным в дихотомической шкале, называется φ-*к*оэффициентом сопряженности, или коэффициентом четырехклеточной сопряженности Пирсона, или коэффициентом корреляции φ.

Если данные представлены в четырехклеточной таблице сопряженности, то применяется формула (8), существенно упрощающая расчеты, но дающая аналогичный результат [14]:

$$\varphi = \frac{a \cdot d - b \cdot c}{\sqrt{(a+b) \cdot (c+d) \cdot (a+c) \cdot (b+d)}}, \quad (8)$$

где *a*, *b*, *c*, *d* соответствуют обозначениям в четырехпольной табл. 7.

Коэффициент корреляции φ не имеет стандартных таблиц для нахождения критических значений. При необходимости найти критические значения можно с помощью t-критерия Стьюдента по формуле

$$T_{\varphi} = |r_{\text{эмп}}| \cdot \sqrt{\frac{n-2}{1-r_{\text{эмп}}{}^{2}}}, \tag{9}$$

где $r_{\text{эмп}}$ – коэффициент корреляции; n – число коррелируемых признаков.

Величина T_{φ} проверяется на уровень значимости по табл. 6 приложения для t-критерия Стьюдента. Число степеней свободы в этом случае будет равно k, где $k = n - 2$.

В случаях применения коэффициента корреляции φ исследователь должен соблюдать следующие условия:

1) сравниваемые признаки должны быть получены в дихотомической шкале;
2) число варьирующих признаков в сравниваемых переменных X и Y должно быть одинаковым;
3) для оценки уровня достоверности коэффициента корреляции φ следует пользоваться формулой (9) и таблицей критических значений для t-критерия Стьюдента (табл. 6 приложения) при числе степеней свободы $k = (n - 2)$ [13].

Заметим, что в различных исследованиях можно встретить различные варианты степени точности предсказания, а при оценке силы корреляционной связи используются различные шкалы.

Контрольные вопросы

1. Приведите определение понятия корреляционной связи. Чем она отличается от функциональной связи?
2. Охарактеризуйте виды существующих корреляционных связей. Приведите примеры основных типов корреляционных связей.
3. Поясните, в чем состоят задачи корреляционного и регрессионного анализа в психолого-педагогических исследованиях.
4. Обоснуйте существование соотношения между типами шкал, в которых могут быть измерены переменные X и Y, и соответствующими мерами связи.

5. Раскройте понятие коэффициента корреляции; поясните, почему он также называется коэффициентом линейной корреляции Пирсона.
6. Объясните, какие значения может принимать величина коэффициента корреляции и как это можно интерпретировать.
7. Каким образом осуществляется подсчет эмпирического и критических значений коэффициента корреляции, а также оценка уровня значимости для определенного психолого-педагогического исследования?
8. Опишите, что понимают под коэффициентом регрессии и какие методы его расчета существуют. Поясните, какой из них вам кажется наиболее оптимальным.
9. Каким преимуществом обладает коэффициент детерминации по сравнению с коэффициентом корреляции.
10. Приведите различные шкалы оценки величины корреляции по силе связи.
11. Поясните, что общего и различного между коэффициентом корреляции φ и коэффициентом корреляции Пирсона.

§ 9. РЕКОМЕНДАЦИИ ПО ВЫБОРУ И ОБРАБОТКЕ СТАТИСТИЧЕСКИХ МЕТОДОВ В ПСИХОЛОГО-ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

При сравнении результатов двух зависимых выборок используют следующие критерии:

– *критерий Макнамары*. Допускает использование данных, которые были получены с помощью измерений по шкале наименований. Этот критерий можно применять, если результаты эксперимента измерены по шкале наименований, которая имеет только две категории (верно – неверно, согласен – не согласен, да – нет и т. п.). При этом сведение результатов эксперимента в таблицу 2 × 2, необходимое для подсчета значения статистики критерия, ограничивает возможности использования данного критерия в процессе обработки информации, которая содержалась в экспериментальных данных;

– *критерий знаков* рекомендуется применять для сравнения данных, измеренных по шкале порядка. Считается, что этот критерий прост в употреблении, но использует далеко не всю информацию, заключенную в результатах выборок. В некоторых случаях критерий знаков не учитывает возможность отклонения нулевой гипотезы, хотя с достаточно большой вероятностью будет верна альтернативная гипотеза;

– *критерий Вилкоксона*. Его применение возможно при соблюдении следующих требований:

1) исходные данные должны быть измерены по интервальной шкале;
2) плотности распределения сравниваемых случайных величин должны быть симметричны, причем критерий состоятелен для вариантов, которые утверждают, что данные плотности распределения отличаются друг от друга только сдвигом;
3) при проверке таких гипотез критерий Вилкоксона более чувствителен, чем критерий знаков: в некоторых случаях при одном и том же уровне значимости он позволяет отклонить нулевую гипотезу, когда критерий знаков этого сделать не позволяет;
4) подобно критерию знаков, критерий Вилкоксона несостоятелен при проверке нулевой гипотезы против альтернатив более общего вида.

При анализе независимых выборок используют:

- медианный критерий и критерий Вилкоксона – Манна – Уитни, которые предназначены для выявления различий в центральных тенденциях распределений рассматриваемых случайных величин, таких как различия медиан и средних значений. Поэтому указанные критерии, подобно критерию знаков и критерию Вилкоксона для зависимых выборок, оптимально подходят для проверки нулевой гипотезы при альтернативах, которые показывают, что распределения случайных величин отличаются сдвигом. Критерии допускают использование данных, измеренных по шкале не ниже порядковой, но ни один из них (в отличие от критерия Вилкоксона для зависимых выборок) не требует, чтобы измерения были выполнены по шкалам более высокого порядка (интервальная шкала и шкала отношений). Использование этих критериев не опирается на предположение о симметричности распределений каждой из сравниваемых случайных величин;
- критерий χ^2 (хи-квадрат) и критерий Колмогорова – Смирнова. Их использование возможно для проверки нулевой гипотезы при самых общих альтернативных гипотезах. Наибольшую область применения имеет критерий χ^2, так как он допускает использование данных, полученных даже с помощью шкалы наименований и с любым числом категорий (в сравнении с критерием Макнамары, который применим только при двух категориях шкалы наименований).

Отметим, что критерий χ^2 следует применять в тех случаях, когда результаты психолого-педагогического эксперимента не могут быть измерены по интервальной шкале. Если интервальные измерения возможны и данные можно сгруппировать не менее чем в восьми интервалах, то лучше применять критерий Колмогорова – Смирнова; в противном случае следует применять критерий χ^2 [11].

Приведем краткую классификацию задач и методов их статистического решения в различных областях науки (табл. 8).

При работе с таблицей рекомендуется:

– по первому столбцу таблицы выбрать задачу, стоящую в исследовании;

– по второму столбцу определить условия решения задачи, например, сколько выборок обследовано или на какое количество групп может быть разбита исследуемая выборка;
– выбрать соответствующий статистический метод или несколько методов и сравнить полученные результаты [13; 16].

Таблица 8

Краткая классификация задач и статистических методов их решения

Задачи	Условия	Методы
1. Выявление различий в уровне исследуемого признака	2 выборки испытуемых	Критерий Макнамары *Q*-критерий Розенбаума *U*-критерий Манна – Уитни φ-критерий (угловое преобразование Фишера)
	3 и более выборок испытуемых	*S*-критерий Джонкира *H*-критерий Крускала – Уоллиса
2. Оценка сдвига значений исследуемого признака	2 замера на одной и той же выборке испытуемых	*T*-критерий Вилкоксона *G*-критерий знаков φ-критерий (угловое преобразование Фишера) *t*-критерий Стьюдента
	3 и более замеров на одной и той же выборке испытуемых	Критерий $\chi^2_{фр}$ Фридмана *L*-критерий тенденций Пейджа *t*-критерий Стьюдента
3. Выявление различий в распределении признака	При сопоставлении эмпирического распределения с теоретическим	χ^2 (критерий согласия Пирсона) λ-критерий Колмогорова – Смирнова *t*-критерий Стьюдента
	При сопоставлении двух эмпирических распределений	χ^2 (критерий согласия Пирсона) λ-критерий Колмогорова – Смирнова φ-критерий (угловое преобразование Фишера)
4. Выявление степени согласованности изменений	Двух признаков	φ – коэффициент корреляции Пирсона τ – коэффициент корреляции Кендалла *R* – бисериальный коэффициент корреляции η – корреляционное отношение Пирсона

Задачи	Условия	Методы
	Трех и большего числа признаков	ρ — коэффициент ранговой корреляции Спирмена *r* — коэффициент корреляции Пирсона Множественная и частная корреляции Линейная, криволинейная и множественная регрессия Факторный и кластерный анализы
5. Анализ изменений признака под влиянием контролируемых условий	Под влиянием одного фактора	*S*-критерий Джонкира *L*-критерий тенденций Пейджа Однофакторный дисперсионный анализ Критерий Линка и Уоллеса Критерий Кемени Множественное сравнение независимых выборок
	Под влиянием двух факторов одновременно	Двухфакторный дисперсионный анализ

§ 10. СТАТИСТИЧЕСКАЯ ОБРАБОТКА ДАННЫХ С ПОМОЩЬЮ КОМПЬЮТЕРНЫХ ПРОГРАММ

В настоящее время достаточно актуально применение средств статистического анализа данных во многих областях деятельности.

За последние 20 лет было разработано значительное количество компьютерных программ, предназначенных для проведения статистического анализа большого объема данных с целью выявления закономерностей, сравнения вероятных альтернатив выбора, построения прогнозов развития событий, обнаружения связи между явлениями и процессами.

На рынке компьютерных программ для статистического анализа данных существует достаточно высокая конкуренция, зачастую можно наблюдать примеры консолидации и поглощений компаний-разработчиков. Например, один из самых активных игроков на рынке – компания SPSS Inc. – в 1994 г. поглотила компанию Systat Software Inc., а в 1996 г. – BMDP Statistical Software Inc. Данные приобретения привели компанию к усовершенствованию собственных программных продуктов. Так, процесс поглощения BMDP Software несомненно улучшил графические инструменты представления данных в SPSS, а процесс поглощения Systat позволил усовершенствовать технологии обработки и анализа данных, получаемых учеными при проведении биологических и медицинских исследований. В 2009 году компания IBM Inc. поглотила компанию SPSS Inc.

При проведении психолого-педагогического исследования для экспериментаторов актуальным становится вопрос выбора оптимального статистического пакета, чтобы были найдены точные ответы на поставленные в исследовании вопросы. Для выбора наиболее оптимального пакета рекомендуется учитывать следующие параметры:

– соответствие пакета характеру решаемых задач;
– объем обрабатываемых данных;
– требования, предъявляемые к квалификации пользователя (уровень знаний в области статистики);
– имеющееся в наличии компьютерное оборудование.

Статистические пакеты по признаку функциональности делят на три основные группы (табл. 9).

Таблица 9

Классификация статистических пакетов
по признаку функциональности

Универсальные статистические пакеты или пакеты общего назначения	Профессиональные статистические пакеты	Специализированные статистические пакеты
SPSS, Stata, Statistica, Stadia, S-PLUS, Minitab, Systat, Statgraphics	SAS, BMDP	BioStat, MESOSAUR, Datascope

Рассмотрим кратко назначение данных групп статистических пакетов.

Первая группа – универсальные пакеты, или пакеты общего назначения. Они не ориентированы на узкую спецификацию и применимы для анализа данных из различных областей деятельности.

Часто они предоставляют широкий диапазон статистических методов и имеют относительно простой интерфейс.

Многопрофильность универсального пакета дает возможность проводить пробный анализ различных типов данных, при этом используется широкий диапазон статистических методов.

Большая часть применяемых универсальных пакетов имеет много пересечений по составу встроенных статистических процедур.

Статистический пакет считается универсальным, если он соответствует следующим требованиям:

– включает широкий набор стандартных статистических методов;
– достаточно прост для быстрого освоения непрофессиональным пользователем;
– работает с достаточно большими базами данных и отвечает высоким требованиям к вводу, преобразованию и организации хранения данных;
– осуществляет обмен данными с широко распространенными пакетами и базами данных;
– имеет обширный набор средств графического представления данных и результатов их анализа;
– имеет подробное документационное сопровождение и справочную систему, которая позволяет начинающему пользователю лег-

ко находить ответы на вопросы, связанные с работой программы и возможностями применения средств анализа данных.

Вторая группа – профессиональные пакеты. Они отличаются от универсальных тем, что дают возможность работать со сверхбольшими объемами данных, применять узкоспециализированные методы анализа, создавать собственную систему обработки данных.

Профессиональные пакеты являются более дорогостоящими, чем универсальные. Например, стоимость покупки SAS Analytics Pro на один год для индивидуального пользования составляет 5 360 EUR. Эти факторы делают современные профессиональные статистические пакеты слишком тяжеловесными для массового применения в различных областях деятельности.

Третья группа – специализированные пакеты. В определенных областях исследователям приходится сталкиваться с узконаправленными, специфичными данными, требующими тщательного анализа, поэтому работа с такими данными требует применения особых методов статистического анализа, часто не представленных в универсальных пакетах.

Специализированные пакеты дают возможность провести анализ, используя ограниченное число специализированных статистических методов. Их также возможно применять при решении вопросов, которые относятся к отдельно взятой предметной области.

Например, статистический пакет BioStat был создан специально для анализа данных в области биологии и медицины. Российские статистические пакеты MESOSAUR применяются при анализе одномерных и многомерных временных рядов и построении регрессионных моделей; пакет Datascope необходим при проведении анализа многомерных данных.

Наиболее целесообразно применять соответствующие специализированные пакеты в том случае, когда необходимо решать задачи из узкой области или применить ограниченный круг сложных статистических процедур, чтобы проанализировать данные из нескольких областей человеческой деятельности.

Большинство статистических пакетов, представленных на рынке, имеют гибкую модульную структуру, пополняемую и расширяемую посредством пользовательских модулей. Данные модули при необходимости дополнительно закупаются или находятся в свободном доступе в Интернете, что дает возможность адаптировать бóльшую часть пакетов к потребностям каждого пользователя.

Профессионалы считают, что статистический пакет должен иметь следующий обязательный набор характеристик:

- модульность;
- ассистирование при выборе способа обработки данных;
- использование простого проблемно-ориентированного языка для формулировки задания пользователя;
- автоматическую организацию процесса обработки данных;
- ведение банка данных пользователя и составление отчета о результатах проделанного анализа;
- диалоговый режим работы пользователя с пакетом;
- совместимость с другим программным обеспечением.

Как правило, представленные на рынке статистические пакеты регулярно обновляются. Самые распространенные пакеты имеют русскоязычную версию.

На сайте НОУ «ИНТУИТ» представлен обзор наиболее известных программных продуктов, разработанных для статистической обработки данных на компьютере и относящихся к указанным основным группам программ, перечислены их основные характеристики и возможности статистического анализа данных (http://www.intuit.ru/studies/courses/3632/874/lecture/14309).

§ 11. ИНТЕРНЕТ-РЕСУРСЫ

- http://www.protein.bio.msu.su/~akula/index.htm – демо-версия пакета STADIA;
- http://www.statgraphics.com/ – версия 5.1 и демо-версия пакета Statgraphics Plus. Содержит более 250 статистических функций, ранние версии этой программы были популярны у отечественных исследователей;
- http://www.xlstat.com/ – макрос-дополнение XLSTAT-Pro для MS Excel, который включает более 50 статистических функций. Пробная версия макроса – на сайте производителя;
- http://www.learnspss.ru – материалы для освоения технологии работы в SPSS;
- http://flogiston.ru/blog/spss13demo – демо-версии последних версий системы SPSS;
- http://www.minitab.com/ – полнофункциональный пробный вариант программы MINITAB 14, которая работает 30 дней. Это достаточно удобный в работе программный пакет, имеющий хороший интерфейс пользователя, хорошие возможности по визуализации результатов работы. Имеет подробную справку;
- http://www.ncss.com/ – полнофункциональная пробная версия пакета NCSS, работающая 30 дней. Программа развивается с 1981 года и рассчитана на непрофессионалов в области статистической обработки;
- http://www.stata.com/ – официальный сайт пакета STATA;
- http://www.spss.ru/ – русскоязычное представительство компании SPSS, которое предлагает полностью русифицированную версию SPSS 12.0.2 для Windows;
- http://www.ruf.rice.edu/~lane/rvls.html – Rice Virtual Lab in Statistics, база данных ресурсов по статистике, включающая монографию по статистике с гиперссылками к другим сетевым статистическим ресурсам; примеры анализа реальных данных и их интерпретаций (раздел Case Studies) и Analysis Lab – базовый метод статистического анализа;
- http://www.statsoft.ru/home/portal – статистический портал, поддерживаемый компанией Statsoft, с внушительной коллекцией раз-

нообразнейших информационных ресурсов по статистическому анализу, структурированных по темам, методам и областям применения;

- http://www.mathstatica.com – описание работы с пакетом программ mathStatica, интегрированным в систему Mathematica;
- http://www.exponenta.ru – образовательный математический сайт Exponenta.ru, который содержит методические указания к лабораторным работам по математической статистике. Целью работ является обучение студентов решению задач математической статистики с использованием математических пакетов Maple, Mathematica, Matlab и Mathcad и др.;
- http://www.statsoft.com/textbook/stathome.html – электронный учебник по статистике Electronic Statistical Textbook, который предлагает материалы для изучения способов анализа данных по некоторым областям науки: биологии и медицины, а также социологии и ряда бизнес-категорий. Имеется русскоязычная версия этого учебника (http://www.statsoft.ru/home/textbook/default.htm);
- http://newasp.omskreg.ru/probability – учебник по теории вероятностей, написанный группой ученых Института математики Сибирского отделения Российской академии наук, доступен в виде Java-скрипта на русском и английском языках;
- http://distance.ru/4stud/umk/stat/stat.html – курс лекций по статистике, авторами которого являются преподаватели Тюменского государственного университета;
- http://teorver-online.narod.ru/tvms-i.html – интернет-учебник А.Д. Маниты, доцента кафедры теории вероятностей механико-математического факультета МГУ. Учебник по теории вероятностей и математической статистике «ТеорВер-Онлайн» ориентирован на студентов естественно-научных факультетов;
- http://www.edu.ru/modules.php?op=modload&name=Web_Links&file=index&l_op=viewlink&cid=2674&fids[]=2285 – учебники и лекции по математической статистике;
- http://www.statsoft.ru/%20home/download/textbook/default.htm – бесплатный электронный учебник по статистике, предоставленный российским представительством StatSoft Inc., который призван помочь разобраться с основными понятиями статистики и более полно представить диапазон применения статистических методов;

- http://www.tula.net/tgpu/new/New/informatic/g10.htm – сайт, посвященный теме «Элементы математической статистики» учебного пособия Р.Р. Яфаевой и Н.Ю. Игнатовой «Информатика и математика» для студентов заочного отделения гуманитарных специальностей педагогических вузов.

§ 12. ПРИМЕРЫ ДИССЕРТАЦИОННЫХ ИССЛЕДОВАНИЙ, ИСПОЛЬЗУЮЩИХ СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Для проверки гипотез, оценки эффективности разработанных методик в различных методических исследованиях, доказательства результатов наблюдения, построения прогнозов используются различные методы. Наиболее распространенными из статистических методов в исследованиях по методике обучения математике являются критерий хи-квадрат, медианный критерий, критерий ранговой корреляции, критерий Вилкоксона, критерий Колмогорова – Смирнова.

В данном параграфе приведем некоторые результаты диссертационных исследований по специальностям 13.00.01 «Общая педагогика, история педагогики и образования» и 13.00.02 «Теория и методика обучения и воспитания», а также по направлениям подготовки магистров 44.04.02 «Психолого-педагогическое образование» (магистерская программа «Начальное образование»), 44.04.01 «Педагогическое образование» (магистерская программа «Математическое образование»), использующих рассмотренные выше статистические методы обработки экспериментальных данных.

Отметим, что все описанные примеры исследований предназначены для рассмотрения их студентами на практических занятиях и в процессе самостоятельной работы; представлены в авторских вариантах, в отдельных из них при обосновании заключений по проведенному педагогическому исследованию имеются неточности в вычислениях либо ошибки в рассуждениях.

12.1. Критерий знаков *G*

Ермолаев Е.А. Элективные курсы по геометрии в условиях профильного обучения математике в старших классах: автореферат диссертации на соискание ученой степени кандидата педагогических наук: 13.00.02. – Саранск, 2010. – 22 с.

Цель исследования заключалась в выявлении теоретических основ проектирования элективных курсов по геометрии и разработке методики их реализации в условиях профильного обучения математике учащихся старших классов общеобразовательной школы.

Обучающий этап экспериментального исследования был проведен в 2007–2010 гг. в 10-х классах профильных центров г.о. Тольятти и в рамках спецкурса со студентами 4-го курса ТГУ – будущими учителями математики. Его основной целью было определение доступности содержания элективного курса «Площадь. Равновеликие и равносоставленные многоугольники», его влияния на качество знаний и умений учащихся по геометрии в целом, достижения основных целей и задач профильного обучения математике.

Схема обучающего этапа эксперимента имела вид: презентация курса $\rightarrow$ входной контроль $\rightarrow$ проведение курса $\rightarrow$ выходной контроль.

На входном контроле выявлялся исходный уровень обученности геометрии учащихся с помощью диагностической работы базового уровня по теме «Площадь многоугольников». Обработка результатов проводилась с помощью метода хи-квадрат (п. 6.3), который показал отсутствие статистически значимых различий между исходными уровнями обученности у учащихся экспериментальных и контрольных групп (ЭГ и КГ).

В качестве выходного контроля учащимся тех же групп была предложена контрольная работа. Ее содержание построено на основе материала базового курса геометрии, изучаемого в обоих классах по традиционной методике у одного и того же учителя одной из городских школ. Полученные статистические результаты свидетельствовали о наличии существенных различий в уровнях обученности геометрии учащихся указанных групп.

Также проверялось влияние изучения элективного курса «Площадь. Равновеликие и равносоставленные многоугольники» на уровень успеваемости учащихся по геометрии. Для этого сравнивались итоговые оценки по геометрии учащихся, результаты контрольных работ до изучения элективного курса (9 класс) и после него (10 класс). Обработка результатов проводилась с помощью рассматриваемого критерия знаков.

Отметим, что в ЭГ входили учащиеся, которые выбрали данный элективный курс для изучения, в КГ – все остальные. Автором были сформулированы гипотезы для ЭГ: H_0 – изучение элективного курса не способствует повышению уровня успеваемости учащихся по геометрии; H_1 – изучение элективного курса способствует повышению уровня успеваемости учащихся по геометрии. Итоговые оценки учащихся представлены в табл. 10.

Таблица 10

Результаты эксперимента до изучения учащимися элективного курса (9 класс) и после него (10 класс)

Экспериментальные группы						
№ п/п	2007/08 уч. г.			2008/09 уч. г.		
	До эксперимента – 9 класс	После эксперимента – 10 класс	Сдвиг	До эксперимента – 9 класс	После эксперимента – 10 класс	Сдвиг
1	4	5	+1	5	5	0
2	3	4	+1	4	5	+1
3	3	4	+1	3	4	+1
4	4	3	+1	3	4	+1
5	3	4	+1	4	4	0
6	4	5	+1	5	5	0
7	4	4	0	5	5	0
8	4	3	−1	4	4	0
9	5	5	0	3	4	+1
10	4	4	0	4	5	+1
11	4	5	+1	4	4	0
12	4	4	0	3	4	+1
13	3	4	+1	3	4	+1
14	4	3	−1	4	4	0
15	3	4	+1	3	3	0
16	4	3	−1	4	5	+1
17	5	5	0	4	3	−1
18	3	4	+1	3	4	+1
19	3	4	+1	5	5	0
20	4	4	0	5	4	−1
21	3	4	+1	4	5	+1
22	4	4	0	5	5	0
23	4	5	+1	3	4	+1
24	5	5	0	3	4	+1
25	4	5	+1	4	5	+1
26	5	5	0	3	4	+1
27	3	4	+1	3	4	+1

	Экспериментальные группы					
№ п/п	2007/08 уч. г.			2008/09 уч. г.		
	До эксперимента – 9 класс	После эксперимента – 10 класс	Сдвиг	До эксперимента – 9 класс	После эксперимента – 10 класс	Сдвиг
28	3	4	+1	3	4	+1
29	3	4	+1	3	4	+1
30	3	4	+1	4	4	0
31	4	5	+1	4	3	−1
32	4	4	0	3	4	+1
33	4	4	0	4	4	0
34	3	4	+1	3	4	+1
35	–	–	–	5	5	0
36	–	–	–	4	5	+1
37	–	–	–	4	5	+1
38	–	–	–	4	4	0

В 2007/08 учебном году было подсчитано количество сдвигов: нулевых – 11, положительных – 20, отрицательных – 3.

Получаем, что большинство сдвигов положительные. Значит, они будут являться «типичными», а $n = 20$. Нетипичных сдвигов – отрицательных – 3, значит, $G_{\text{эмп}} = 3$.

По табл. 1 приложения найдем критические значения по величине типичного сдвига, в нашем случае для $n = 20$.

Имеем $G_{\text{кр}} = \begin{cases} 5 \text{ для } p \leq 0{,}05, \\ 4 \text{ для } p \leq 0{,}01 \end{cases}$, т. е. при уровне значимости в 5 % сумма нетипичных сдвигов не должна превышать 5, а при уровне значимости 1 % – 3.

В данном случае $G_{\text{эмп}} = 3$, т. е. $G_{\text{эмп}} \leq G_{\text{кр}}$. Поэтому гипотеза H_0 отклоняется и принимается гипотеза H_1 ($p \leq 0{,}01$).

В 2008/09 учебном году также было подсчитано количество сдвигов: нулевых – 14, положительных – 21, отрицательных – 3.

Большинство сдвигов положительные, т. е. они типичные, а $n = 21$. Нетипичных сдвигов – отрицательных – 3, значит, $G_{\text{эмп}} = 3$.

По табл. 1 приложения найдем критические значения по величине типичного сдвига, в нашем случае для $n = 21$.

Имеем $G_{\text{кр}} = \begin{cases} 6 \text{ для } p \le 0{,}05, \\ 4 \text{ для } p \le 0{,}01 \end{cases}$, т. е. при уровне значимости в 5 % сумма нетипичных сдвигов не должна превышать 6, а при уровне значимости 1 % – 4.

В этом случае $G_{\text{эмп}} = 3$, т. е. $G_{\text{эмп}} \le G_{\text{кр}}$. Аналогично первому случаю гипотеза H_0 отклоняется и принимается гипотеза H_1 ($p \le 0{,}01$).

Таким образом, положительный сдвиг в успеваемости учащихся по геометрии не является случайным. Данный факт подтверждает предположение о влиянии рассматриваемого элективного курса на развитие интереса учащихся к изучению геометрии, на уровень знаний учащихся по геометрии.

Полученные результаты послужили основанием для вывода о влиянии содержания и методики реализации элективного курса «Площадь. Равновеликие и равносоставленные многоугольники» на повышение качества геометрического образования в целом.

Задание 1. Докажите обоснованность приведенных автором заключений по проведенному педагогическому исследованию. Составьте алгоритм обработки результатов эксперимента с помощью критерия знаков *G*.

12.2. Парный *T*-критерий Вилкоксона

Лифанова Н.В. Развитие творческого потенциала младших школьников в учебной деятельности: магистерская диссертация по направлению подготовки 44.04.02 «Психолого-педагогическое образование», направленность (профиль) «Начальное образование». – Тольятти, 2016. – 194 с.

Цель исследования: теоретически обосновать, разработать и практически реализовать комплекс психолого-педагогических условий развития творческого потенциала младших школьников в учебной деятельности.

Контрольный этап экспериментального исследования проводился на базе ГБОУ СОШ № 10 г. Жигулёвска в 2015/16 учебном году. Общее количество испытуемых – 31 человек (ученики 2 «Г» класса). Возрастной диапазон обследуемых – 8–9 лет. Эксперимент проводился в естественных условиях образовательного процесса.

На начальном этапе экспериментального исследования автором была выдвинута гипотеза о возможности повышения уровня творческого потенциала младших школьников в результате проведения комплекса мероприятий.

В ходе контрольного этапа эксперимента для диагностики уровня развития творческого потенциала учащихся применялся пакет диагностических методик, разработанный на констатирующем этапе эксперимента. Проведён анализ результатов развития компонентов творческого потенциала младших школьников: мотивационного, познавательно-деятельностного, креативного, интеллектуального. Сделан расчёт интегрального показателя уровня развития творческого потенциала личности.

Сравнительный анализ значений интегрального показателя уровня развития творческого потенциала учащихся на констатирующем и контрольном этапах эксперимента обнаружил, что интегральный показатель творческого потенциала у учащихся 2 «Г» класса изменился неравномерно; у 27 учащихся (87,1 % класса) данный показатель повысился, а у четырех учащихся (12,9 % класса) – понизился. Поэтому для подтверждения либо отклонения выдвинутой гипотезы экспериментального исследования автором были применены методы математической статистики.

Так, с помощью методов математической статистики, позволяющих обобщить закономерности, полученные в результате контрольного этапа эксперимента, осуществлён сравнительный анализ уровня развития отдельных компонентов творческого потенциала младших школьников, а также интегрального показателя творческого потенциала личности с использованием парного T-критерия Вилкоксона.

В табл. 11 приведены соответствующие экспериментальные данные (значения интегрального показателя творческого потенциала личности, полученные в ходе проведения констатирующего и контрольного этапов эксперимента), а также дополнительные столбцы, необходимые для работы по парному T-критерию Вилкоксона.

Сумма рангов S подсчитывается по формуле

$$S = \frac{N \cdot (N-1)}{2},$$

где N – количество испытуемых.

Таблица 11

Результаты эксперимента для интегрального показателя творческого потенциала (ТП) личности учащихся

№ п/п	ТП, констатир. этап эксп.	ТП, контр. этап эксп.	Сдвиг	Абсолютные величины разностей	Ранги	Символ нетипичного сдвига
1	165,6	182,3	16,7	16,7	11	
2	179,0	205,6	26,6	26,6	21	
3	94,8	106,7	11,9	11,9	6	
4	89,2	104,3	15,1	15,1	9	
5	166,1	193,9	27,8	27,8	22	
6	126,9	147,7	20,8	20,8	15	
7	195,1	224,0	28,9	28,9	24	
8	142,1	155,1	13,0	13,0	7	
9	130,8	161,3	30,5	30,5	25	
10	191,2	224,2	33,0	33,0	27	
11	143,8	142,5	−1,3	1,3	2	*
12	121,3	143,6	22,3	22,3	17	
13	106,0	146,2	40,2	40,2	31	
14	131,5	154,9	23,4	23,4	19	
15	171,0	208,3	37,3	37,3	28	
16	145,1	162,5	17,4	17,4	13	
17	149,0	166,0	17,0	17,0	12	
18	138,6	176,4	37,8	37,8	29	
19	125,6	148,2	22,6	22,6	18	
20	136,6	126,4	−10,2	10,2	5	*
21	113,6	107,1	−6,5	6,5	4	*
22	177,2	189,8	12,6	12,6	8	
23	157,7	181,2	23,5	23,5	20	
24	187,2	215,8	28,6	28,6	23	
25	184,5	189,9	5,4	5,4	3	
26	139,0	154,9	15,9	15,9	10	
27	93,1	92,2	−0,9	0,9	1	*
28	116,9	148,8	31,9	31,9	26	
29	117,2	135,1	17,9	17,9	14	
30	177,2	216,9	39,7	39,7	30	
31	101,9	123,3	21,4	21,4	16	
Сумма					496	12

Сумма рангов, полученная путём простого суммирования данных таблицы, и сумма, вычисленная по формуле, совпадают и равны 496. Значит, ранжирование проведено правильно.

В нашем случае типичный сдвиг положительный, нетипичный отрицательный. Нетипичные сдвиги отмечены в таблице специальным символом, причем мы имеем четыре нетипичных сдвига. Суммируем ранги нетипичных сдвигов, это и будет искомая величина $T_{\text{эмп}}$. В нашем случае эта сумма равна 12: $T_{\text{эмп}} = 2 + 5 + 4 + 1 = 12$. По табл. 2 приложения определяем значение $T_{\text{кр}}$ для $n = 31$, т. е. поиск критических величин ведется по общему числу испытуемых. Имеем:

$T_{\text{кр}} = \begin{cases} 163 \text{ для } p \le 0{,}05, \\ 130 \text{ для } p \le 0{,}01 \end{cases}$. Строим ось значимости (рис. 12).

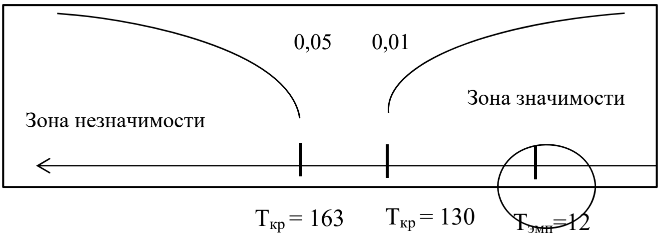

Рис. 12. Ось значимости для уровня развития интегрального показателя творческого потенциала личности учащихся

Анализ оси значимости показывает, что полученная величина $T_{\text{эмп}}$ попадает в зону значимости. Следовательно, можно утверждать, что зафиксированные в эксперименте изменения неслучайны и значимы на уровне 1 %. Поскольку преобладание типичного положительного направления сдвига в данном эксперименте не является случайным, то должна быть принята гипотеза H_1 о наличии различий, а гипотеза H_0 отклонена.

Таким образом, комплекс мероприятий, проведённых в ходе формирующего эксперимента, способствовал повышению уровня творческого потенциала личности исследуемых младших школьников.

Для того чтобы провести анализ развития отдельных компонентов творческого потенциала младших школьников (мотивационный, познавательно-деятельностный, креативный, интеллектуальный), также применим парный T-критерий Вилкоксона.

Рассмотрим анализ развития мотивационного компонента творческого потенциала младших школьников, приведенный автором. Значения мотивационного компонента творческого потенциала личности, полученные в ходе проведения констатирующего и контрольного экспериментов, а также дополнительные столбцы, необходимые для работы по парному *T*-критерию Вилкоксона, представлены в табл. 12.

Таблица 12

Результаты эксперимента для мотивационного компонента творческого потенциала личности учащихся

№ п/п	ТП, констатир. этап эксп.	ТП, контр. этап эксп.	Сдвиг	Абсолютные величины разностей	Ранги	Символ нетипичного сдвига
1	15,0	18,0	3,0	3,0	16,0	
2	19,5	27,0	7,5	7,5	29,0	
3	24,0	22,0	−2,0	2,0	11,0	*
4	15,0	15,5	0,5	0,5	4,5	
5	16,0	23,0	7,0	7,0	26,5	
6	20,5	23,0	2,5	2,5	14,0	
7	28,0	29,0	1,0	1,0	8,0	
8	25,0	24,0	−1,0	1,0	8,0	*
9	16,0	23,0	7,0	7,0	26,5	
10	24,0	29,0	5,0	5,0	21,5	
11	16,5	17,0	0,5	0,5	4,5	
12	11,0	13,0	2,0	2,0	11,0	
13	20,5	21,0	0,5	0,5	4,5	
14	24,0	26,0	2,0	2,0	11,0	
15	21,0	28,0	7,0	7,0	26,5	
16	24,5	22,0	−2,5	2,5	14,0	*
17	21,5	24,0	2,5	2,5	14,0	
18	25,0	25,0	0,0	0,0	1,5	
19	26,5	27,0	0,5	0,5	4,5	
20	28,5	21,5	−7,0	7,0	26,5	*
21	26,5	20,0	−6,5	6,5	24,0	*
22	14,5	27,0	12,5	12,5	31,0	
23	20,0	25,0	5,0	5,0	21,5	
24	21,0	29,0	8,0	8,0	30,0	
25	17,0	18,0	1,0	1,0	8,0	

№ п/п	ТП, кон- статир. этап эксп.	ТП, контр. этап эксп.	Сдвиг	Абсолютные величины разностей	Ранги	Символ нетипичного сдвига
26	20,5	24,0	3,5	3,5	17,0	
27	20,0	16,0	−4,0	4,0	18,5	*
28	22,5	28,0	5,5	5,5	23,0	
29	23,0	23,0	0,0	0,0	1,5	
30	25,0	29,0	4,0	4,0	18,5	
31	7,5	12,0	4,5	4,5	20,0	
Сум- ма					496	102

Строим ось значимости (рис. 13).

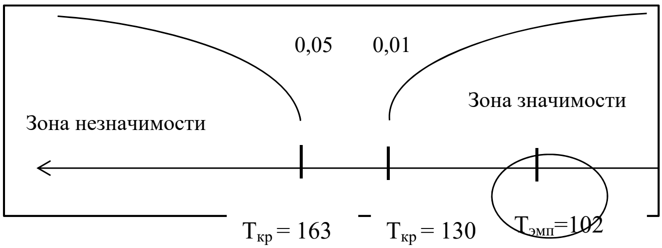

Рис. 13. Ось значимости для уровня развития мотивационного компонента творческого потенциала личности учащихся

Анализ оси значимости показывает, что полученная величина $T_{эмп}$ попадает в зону значимости. Следовательно, можно утверждать, что зафиксированные в эксперименте изменения неслучайны и значимы на уровне 1 %. Поскольку преобладание типичного положительного направления сдвига в данном эксперименте не является случайным, то должна быть принята гипотеза H_1 о наличии различий, а гипотеза H_0 отклонена.

Таким образом, комплекс мероприятий, проведённых в ходе формирующего эксперимента, также способствовал повышению уровня развития мотивационного компонента творческого потенциала личности.

На основании вышесказанного Н.В. Лифановой сделан вывод о том, что признание значимости и необходимости развития творческого потенциала личности младших школьников, качественные изменения в организации, содержании и технологиях учебно-воспитательного процесса способствуют повышению уровня развития творческого потенциала личности и отдельных его компонентов.

Задание 2. Докажите обоснованность приведенных автором заключений по проведенному педагогическому исследованию. Составьте алгоритм обработки результатов эксперимента с помощью парного T-критерия Вилкоксона. Выявите сходство и различия данного критерия с критерием знаков G.

12.3. Критерий Макнамары

Васильева М.А. Профессионально-прикладная направленность обучения математике как средство формирования математической компетентности (на примере аграрного вуза): диссертация на соискание ученой степени кандидата педагогических наук: 13.00.02: Рязань, 2014. – 190 с.

Цель исследования состояла в разработке методики обучения математике студентов аграрных вузов, построенной на профессионально-прикладной направленности обучения, способствующей формированию математической компетентности специалиста.

На входном контроле выявлялся исходный уровень знаний испытуемых с помощью контрольной работы. Обработка результатов проводилась с помощью двустороннего метода хи-квадрат (п. 6.3.), который показал отсутствие статистически значимых различий между исходными уровнями знаний у испытуемых экспериментальной и контрольной групп.

Для выявления различия в качестве знаний студентов в экспериментальной и контрольной группах применили критерий Макнамары.

Нулевая гипотеза H_0: предлагаемая модель профессионально-прикладной направленности обучения студентов математике не оказывает влияния на качество знаний.

Альтернативная гипотеза H_1: предлагаемая модель профессионально-прикладной направленности обучения студентов существенно влияет на качество знаний.

Результаты диагностики студентов можно представить в виде таблицы, где

– a – число студентов, которые на начало и конец эксперимента имели высокий уровень знаний по математике;
– b – число студентов, которые поменяли высокий уровень математических знаний на низкий;
– c – число студентов, которые перешли с низкого уровня математических знаний на более высокий;
– d – число студентов, которые не поменяли уровня математических знаний, оставшись на низком уровне.

Результаты диагностики студентов представлены в табл. 13.

Таблица 13

Четырехпольная таблица данных эксперимента

	Высокий	Низкий
Высокий	$a = 4$	$b = 0$
Низкий	$c = 14$	$d = 7$

В данном случае $n = b + c = 14 + 0 = 14 \leq 20$, поэтому для определения величины $M_{эмп}$ используется таблица биноминального распределения (табл. 3 приложения; $m = 0$ – как наименьшая из величин b и c; $n = b + c = 14$, $M_{эмп} = 0$), величины $M_{кр}$ постоянны и равны 0,025 для уровня значимости 5 % и 0,005 – для уровня значимости 1 % (рис. 14).

Так как $M_{эмп}$ попало в зону значимости, то нулевая гипотеза отклоняется. Иначе говоря, есть достаточные основания считать, что предлагаемая модель профессионально-прикладной направленности обучения студентов существенно влияет на качество знаний. Экспериментальная проверка гипотезы исследования показала, что профессионально-прикладная направленность обучения способствует повышению качества математических знаний.

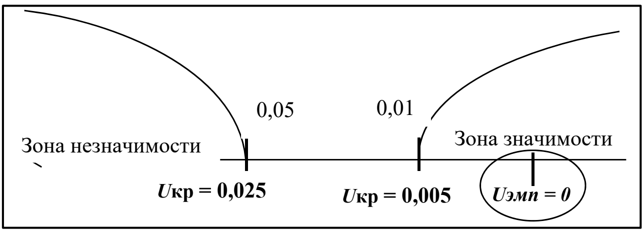

Рис. 14. Ось значимости

Задание 3. Докажите обоснованность приведенных автором заключений по проведенному педагогическому исследованию. Составьте алгоритмы обработки результатов эксперимента с помощью критерия Макнамары для двух различных случаев.

12.4. Критерий Вилкоксона – Манна – Уитни

Фирсова Е.В. Обучение дискретной математике студентов вуза с использованием дистанционных образовательных технологий (на примере специальности/профиля «Прикладная информатика (в экономике)»): диссертация на соискание ученой степени кандидата педагогических наук: 13.00.02. – Коломна, 2014. – 214 с.

Цель исследования – теоретическое обоснование и методическое обеспечение повышения эффективности обучения дискретной математике студентов вуза с использованием дистанционных образовательных технологий.

В эксперименте были выделены три уровня знаний: низкий, средний и высокий (табл. 14).

Контрольной и экспериментальной группам была предложена входная контрольная работа по проверке уровня знаний по математике в системе дистанционного обучения «Прометей». Показателем успешности служило количество правильно выполненных заданий.

Автор поставил вопрос: уровень знаний у студентов контрольной и экспериментальной групп до эксперимента совпадает или различается? Для ответа на вопрос применил *U*-критерий Вилкоксона – Манна – Уитни.

Таблица 14

Шкала распределения уровней знаний студентов в соответствии с числом верно решенных ими задач контрольной работы (переход от шкалы отношений к порядковой шкале)

Уровень знаний	Число правильно решенных задач в контрольной работе
Низкий	Меньше 6
Средний	7–9
Высокий	10–12

Была введена гипотеза H_0 о сходстве уровня знаний у студентов двух групп по количеству правильно решенных заданий входной контрольной работы, а гипотеза H_1 – о наличии различий. Все полученные данные были внесены в таблицу для расчета по критерию U (табл. 15).

Таблица 15

Упорядоченное объединение экспериментальных данных двух групп испытуемых в порядке их возрастания до эксперимента

ЭГ	КГ	Инверсии ЭГ/КГ	Инверсии КГ/ЭГ
–	5	–	0
6	6	1	0
–	7	–	1
–	7	–	1
8	–	4	–
9	9	4	2
–	9	–	3
Суммы инверсий		9	7

Из табл. 15 видно, что в случае инверсии ЭГ/КГ сумма инверсий на 2 больше, чем в случае инверсии КГ/ЭГ. Принято считать, что $U_{эмп}$ = min ((КГ/ЭГ), U(ЭГ/КГ)) = 7. По табл. 4 приложения для $n_1 = 3$ и $n_2 = 6$ находим следующее: $U_{кр1} = 2$ для уровня статистической значимости $p = 0{,}05$; $U_{кр2}$ не существует для уровня статистической значимости $p = 0{,}01$. Построим ось значимости (рис. 15).

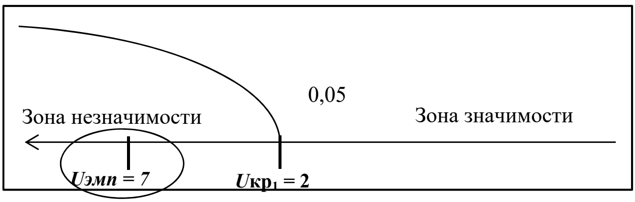

Рис. 15. Ось значимости до проведения эксперимента

Полученное значение попало в зону незначимости, следовательно, применяется гипотеза H_0 о сходстве, а гипотеза H_1 отклоняется. Таким образом, уровень знаний по математике у контрольной группы и экспериментальной группы совпадают.

Подчеркнем, что ось значимости в этом критерии имеет направление справа налево. При этом числовые значения по оси абсцисс по мере увеличения уровней значимости убывают. Последнее закономерно, так как, чем меньше взаимопересечений (инверсий) в двух рядах, тем больше достоверность их различий.

Контрольной и экспериментальной группам для проверки уровня знаний по математике была предложена итоговая контрольная работа. Показателем успешности служило количество правильно выполненных заданий.

Автор поставил вопрос, совпадает или различается уровень знаний у студентов контрольной и экспериментальной групп после эксперимента. Для ответа на вопрос применили рассматриваемый критерий U. Ввели гипотезу H_0 о сходстве уровня знаний у студентов двух групп по количеству правильно решенных заданий итоговой контрольной работы, а гипотезу H_1 – о наличии различий. Внесли все данные в таблицу для расчета по критерию U (табл. 16).

Из табл. 16 видно, что в случае инверсии ЭГ/КГ сумма инверсий на 14 больше, чем в случае инверсии КГ/ЭГ, причем $U_{эмп}$ = min ((КГ/ЭГ), U(ЭГ/КГ)) = 1. По табл. 4 приложения для $n_1 = 3$ и $n_2 = 6$ находим следующее: $U_{кр1} = 2$ для уровня статистической значимости $p = 0{,}05$; $U_{кр2}$ не существует для уровня статистической значимости $p = 0{,}01$. На рис. 16 представлена соответствующая ось значимости.

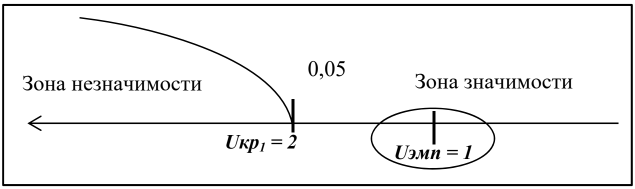

Рис. 16. Ось значимости после проведения эксперимента

Различия между двумя выборками можно считать значимыми ($p \le 0{,}05$), если $U_{эмп}$ ниже или равно $U_{кр}$. $U_{эмп}$ отражает то, насколько велика зона совпадения между рядами значений. Поэтому чем меньше $U_{эмп}$, тем более вероятно, что различия достоверны. Полученное значение попало в зону значимости, следовательно, была применена гипотеза H_1 о наличии различий, а гипотеза H_0 о сходстве отклонена.

Таблица 16

Упорядоченное объединение экспериментальных данных двух групп испытуемых в порядке их возрастания после эксперимента

ЭГ	КГ	Инверсии ЭГ/КГ	Инверсии КГ/ЭГ
–	6	–	0
–	7	–	0
–	8	–	0
–	8	–	0
9	9	4	0
11	11	5	1
12	–	6	–
Суммы инверсий		15	1

Таким образом, получили, что уровень знаний по математике у экспериментальной группы выше, чем у контрольной группы.

Рассмотрим еще одно исследование, в котором применялся *U*-критерий Вилкоксона – Манна – Уитни.

Афанасьева Е.В. Методика обучения учащихся составлению задач как средство гуманитаризации математического образования: магистерская диссертация по направлению подготовки 44.04.01 «Педагогическое образование», магистерская программа «Математическое образование». – Тольятти, 2015. – 90 с.

Цель исследования состоит в обосновании и разработке методики обучения учащихся старших классов общеобразовательной школы самостоятельному составлению задач как средства гуманитаризации математического образования.

Экспериментальная проверка эффективности разработанных методик проводилась на базе МБУ лицей № 57 г. Тольятти в 2014–2015 годах. Весь эксперимент был разбит на три этапа: констатирующий, обучающий и контролирующий.

На констатирующем этапе применялись такие методы исследования, как наблюдение за проведением уроков математики в старших классах и их анализ, беседы с учащимися и учителями, анкетирование учителей (МБУ лицей № 57, МБУ СОШ № 72, 59, 44), проведение диагностической контрольной работы, обработка результатов контрольной работы с помощью непараметрического критерия для несвязных выборок.

Диагностическая контрольная работа проводилась с целью изучения уровня сформированности умений по составлению задач учащимися 10-х классов в условиях профильного обучения.

Результаты контрольной работы оценивались по следующим критериям: за выполнение каждого задания можно получить от 0 до 2 баллов; 2 балла – задание выполнено полностью; 1 балл – задание выполнено не полностью или есть ошибка; 0 баллов – задание не выполнено. Максимальное количество баллов – 12. Баллы, набранные учениками контрольного и экспериментального классов (КК и ЭК) ранжированы и представлены в табл. 17.

Средние баллы выполнения диагностической контрольной работы в КК и ЭК практически одинаковы (7,39 и 7,41 соответственно). Наблюдаемое значение критерия считаем по формуле

$$U_{\text{эмп}} = n_1 \cdot n_2 + n_x \cdot \frac{n_x+1}{2} - R_{\max},$$

где $n_1 = 23$, $n_2 = 24$ – число учащихся КК и ЭК соответственно; $R_{\max} = 578$ – наибольшая по величине сумма рангов; $n_x = 24$ – количество испытуемых в группе с наибольшей суммой рангов. Получим $U_{\text{эмп}} = 274$.

Величины критических значений находим по табл. 4 приложения: $U_{\text{кр}} = \begin{cases} 198, p \leq 0{,}05 \\ 186, p \leq 0{,}01 \end{cases}$. На рис. 17 представлена соответствующая ось значимости. Наблюдаемое значение критерия попало в зону незначимости, следовательно принимается гипотеза H_0 о сходстве, что позволяет сделать вывод об одинаковом уровне сформированности умения составлять математические задачи у учеников КК и ЭК.

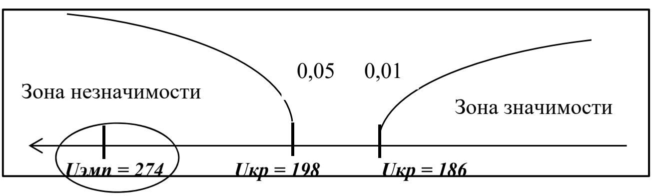

Рис. 17. Ось значимости

Таблица 17

Результаты констатирующего эксперимента

№	КК X	ЭК Y	Ранг $R(X)$	Ранг $R(Y)$	№	КК X	ЭК Y	Ранг $R(X)$	Ранг $R(Y)$
1	10	–	(1) 3		25	7	–	(25) 28	
2	10	–	(2) 3		26	7	–	(26) 28	
3	–	10		(3) 3	27	7	–	(27) 28	
4	–	10		(4) 3	28	7	–	(28) 28	
5	–	10		(5) 3	29	7	–	(29) 28	
6	9	–	(6) 10,5		30	–	7		(30) 28
7	9	–	(7) 10,5		31	–	7		(31) 28
8	9	–	(8) 10,5		32	–	7		(32) 28
9	9	–	(9) 10,5		33	–	7		(33) 28
10	9	–	(10) 10,5		34	6	–	(34) 37	
11	9	–	(11) 10,5		35	6	–	(35) 37	
12	–	9		(12) 10,5	36	6	–	(36) 37	
13	–	9		(13) 10,5	37	–	6		(37) 37
14	–	9		(14) 10,5	38	–	6		(38) 37
15	–	9		(15) 10,5	39	–	6		(39) 37
16	8	–	(16) 19		40	–	6		(40) 37
17	8	–	(17) 19		41	5	–	(41) 43	
18	–	8		(18) 19	42	–	5		(42) 43
19	–	8		(19) 19	43	–	5		(43) 43
20	–	8		(20) 19	44	–	5		(44) 43
21	–	8		(21) 19	45	–	5		(45) 43
22	–	8		(22) 19	46	4	–	(46) 46,5	
23	7	–	(23) 28		47	4	–	(46) 46,5	
24	7	–	(24) 28			Сумма рангов		550	578

Обучающий и контролирующий этапы эксперимента проводились в течение 2014/15 учебного года на базе МБУ лицей № 57.

Цель – проверить эффективность разработанных методик по обучению школьников составлению задач.

Выборка. КК: 23 ученика 11 «Б» класса (11 человек – физико-математический профиль, 12 человек – профиль «обществознание + английский язык»). ЭК: 24 ученика 11 «Д» класса (12 человек – физико-математический профиль, 12 человек – профиль «обществознание + английский язык»).

Ученики ЭК обучались по программе элективного курса «Составление и решение геометрических задач», в рамках урока алгебры составляли задачи по теме «Функция».

Задача обучающего и контролирующего эксперимента – сравнить уровень сформированности умений по составлению задач в КК и ЭК.

В обоих классах была проведена контрольная работа, аналогичная первой диагностической работе. Результаты представлены в табл. 18.

Средний балл выполнения контрольной работы, аналогичной диагностической работе, в ЭК (10,25) выше, чем в КК (8,61). Наблюдаемое значение критерия считаем по формуле

$$U_{\text{эмп}} = n_1 \cdot n_2 + n_x \cdot \frac{n_x+1}{2} - R_{\max},$$

где $n_1 = 23$, $n_2 = 24$ – число учащихся КК и ЭК соответственно; $R_{\max} = 725$ – наибольшая по величине сумма рангов; $n_x = 23$ – количество испытуемых в группе с наибольшей суммой рангов.

Получим $U_{\text{эмп}} = 103$. Величины критических значений находим по табл. 4 приложения: $U_{kp} = \begin{cases} 198, \mathrm{p} \le 0{,}05 \\ 186, \mathrm{p} \le 0{,}01 \end{cases}$.

Наблюдаемое значение критерия попало в зону значимости (рис. 18), следовательно, гипотеза H_0 о сходстве отвергается, что позволяет сделать вывод о существенном различии уровней сформированности умения составлять математические задачи у учеников КК и ЭК. На уровне значимости 5 % можно сделать вывод об эффективности разработанных и примененных методик обучения школьников составлению математических задач.

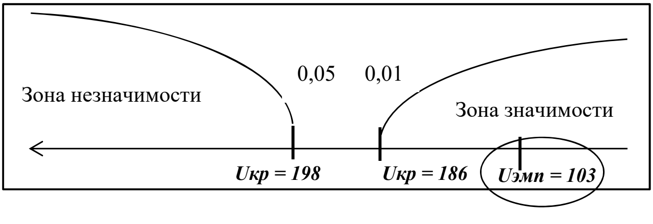

Рис. 18. Ось значимости

Таблица 18

Результаты обучающего эксперимента

№	КК X	ЭК Y	Ранг R(X)	Ранг R(Y)	№	КК X	ЭК Y	Ранг R(X)	Ранг R(Y)
1	–	12		(1) 2,5	25	9	–	(25) 30	
2	–	12		(2) 2,5	26	9	–	(26) 30	
3	–	12		(3) 2,5	27	9	–	(27) 30	
4	–	12		(4) 2,5	28	9	–	(28) 30	
5	11	–	(5) 8		29	9	–	(29) 30	
6	–	11		(6) 8	30	–	9		(30) 30
7	–	11		(7) 8	31	–	9		(31) 30
8	–	11		(8) 8	32	–	9		(32) 30
9	–	11		(9) 8	33	–	9		(33) 30
10	–	11		(10) 8	34	–	9		(34) 30
11	–	11		(11) 8	35	–	9		(35) 30
12	10	–	(12) 18		36	8	–	(36) 39	
13	10	–	(13) 18		37	8	–	(37) 39	
14	10	–	(14) 18		38	8	–	(38) 39	
15	10	–	(15) 18		39	8	–	(39) 39	
16	10	–	(16) 18		40	8	–	(40) 39	
17	10	–	(17) 18		41	8	–	(41) 39	
18	–	10		(18) 18	42	–	8		(42) 39
19	–	10		(19) 18	43	7	–	(43) 44,5	
20	–	10		(20) 18	44	7	–	(44) 44,5	
21	–	10		(21) 18	45	7	–	(45) 44,5	
22	–	10		(22) 18	46	7	–	(46) 44,5	
23	–	10		(23) 18	47	6	–	(47) 47	
24	–	10		(24) 18		Сумма рангов		725	403

Задание 4. Докажите обоснованность приведенных Е.В. Фирсовой и Е.В. Афанасьевой заключений по проведенным педагогическим исследованиям. Какой из двух способов расчета по критерию *U* применили авторы, насколько это правомерно? Найдите ошибки при обработке результатов эксперимента, если таковые имеются.

12.5. Медианный критерий

Горчаков А.С. Развитие математической речи школьников в контексте деятельностного подхода: диссертация на соискание ученой степени кандидата педагогических наук: 13.00.02. – Нижний Новгород, 2014. – 228 с.

Цель исследования заключалась в выявлении и обосновании теоретико-методической концепции и разработке адекватной методики развития математической речи школьников.

По результатам выполнения отдельных заданий автором было установлено, что в контрольном классе речевые умения сформированы недостаточно. Предполагается, что это объясняется отсутствием целенаправленной работы по развитию математической речи школьников. В целом, уровень сформированности речевых умений в экспериментальном классе значительно выше, чем в контрольном классе.

Целью проведения проверочной работы ставилось сравнение качества знаний учащихся контрольного и экспериментального классов и уровня владения математической речью при формулировании математических предложений и ответах на вопросы.

При верных и полных ответах во всех трёх частях работы учащиеся получали 20 баллов. В итоге все результаты учащихся, выраженные в баллах, распределились от 0 до 20.

Для оценки различий результатов, полученных при выполнении проверочной работы в ЭК и КК, воспользуемся статистической обработкой данных. Медианный критерий применяется для выявления различий в состоянии некоторого свойства в двух совокупностях на основе изучения членов двух выборок из этих совокупностей. В этом случае показателем тенденции изменения некоторого свойства служит медиана изменения изучаемого свойства в каждой из выборок. В данном случае исследуемое свойство – развитие математической речи школьников.

В условиях проведённого эксперимента выполняются основные допущения медианного критерия, который был применен для проверки данного предположения:

1) контрольные классы выбраны случайным образом;
2) контрольные и экспериментальные классы являются независимыми;

3) для анализа результатов использовалась интервальная шкала и шкала отношений, т. е. шкалы выше порядковой;
4) число членов в обеих выборках больше 20.

Ряды распределения учащихся ЭК и КК по количеству баллов, полученных за выполнение работы, использовались в качестве показателей развития математической речи школьников и усвоения знаний учащихся. Представим данные учащихся обеих выборок в форме таблицы для более удобного вычисления медианы (табл. 19).

Таблица 19

Результаты выполнения диагностической работы учащимися ЭК и КК по количеству набранных баллов

Число баллов	Количество учащихся, получивших это количество баллов			Накопленная частота
	ЭК	КК	Всего	
20	2	1	3	150
19	1	0	1	147
18	3	1	4	146
17	7	2	9	142
16	8	4	12	133
15	7	7	14	121
14	12	9	21	107
13	15	14	29	86
12	4	12	16	57
11	4	10	14	41
10	6	11	17	27
9	4	3	7	10
8	0	2	2	3
7	0	1	1	1

$n_1 = 73$ $n_2 = 77$

На гистограмме (рис. 19) показано количество учеников по каждому количеству баллов (горизонтальная ось показывает количество набранных баллов, вертикальная – количество учеников).

Имеем две совокупности: ЭК и КК. Пусть X характеризует состояние исследуемого свойства в ЭК, а Y – состояние этого свойства в КК. Обозначим $X[i]$ и $Y[i]$ результаты измерения свойства у учащихся ЭК и КК соответственно, т. е. $X[i]$ и $Y[i]$ – это число баллов, полученных одним учащимся.

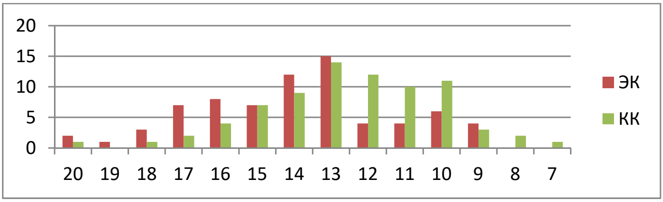

Рис. 19. Гистограмма сравнения учащихся ЭГ и КГ
по набранному количеству баллов

Примем за $N = n_1 + n_2$, то есть $N = 150$ человек – объём объединённой выборки.

Найдём медиану рассматриваемой выборки, то есть такое значение, больше которого 50 % измерений $X[i]$ и $Y[i]$ и меньше которого также 50 %. Иными словами, если все $X[i]$ и $Y[i]$ расположить по возрастающей, то медиана m будет равна

$$m = \left[\begin{array}{l} \dfrac{z(N+1)}{2}, \text{если } N - \text{нечетное;} \\ \dfrac{z\left(\frac{N}{2}\right) + Z\left(\frac{N}{2}+1\right)}{2}, \text{если } N - \text{четное,} \end{array}\right.$$

причем $Z[i] = X[i]$ или $Z[i] = Y[i]$.

Так как число учащихся N в обеих выборках равно 150, т. е. чётное, то медиана численно равна значению, стоящему на 75 месте. В упорядоченном ряду измерений, составленном по результатам двух выборок, это значение равно 13.

Для дальнейших вычислений распределим значения обеих выборок на две категории (табл. 20): больше медианы (>13) и меньше или равное медиане (≤13).

Таблица 20

Распределение результатов по отношению к медиане

	Выборка 1, $X[i]$	Выборка 2, $Y[i]$	
Больше 13	40 (A)	24 (B)	64
Меньше или равно 13	33 (C)	53 (D)	86
	73	77	

Сформулируем проверяемую гипотезу. Нулевая гипотеза H_0: медианы распределения учащихся по числу баллов, полученных за выполнение работы, одинаковы в совокупности учащихся ЭК и КК (то есть $m_1 = m_2$). Альтернативная гипотеза H_1: медианы различны.

Значение статистики медианного критерия находится по формуле

$$T = \frac{N \cdot \left(|A \cdot D - B \cdot C| - \frac{N}{2} \right)^2}{(A + B) \cdot (C + D) \cdot (A + C) \cdot (B + D)}.$$

Вычисляя, получили $T = 8{,}55$. Для уровня значимости $p = 0{,}01$ и одной степени свободы по соответствующей таблице нашли $T_{кр} = 6{,}64$. Итак, получили, что верно неравенство $\boldsymbol{T} > T_{кр}$ (8,55 > 6,64), значит, в соответствии с правилом принятия решения при использовании медианного критерия нулевая гипотеза отклоняется и с вероятностью 99 % принимается альтернативная гипотеза: медианы распределения учащихся по числу баллов за выполнение проверочной работы различны в ЭК и КК. При этом результаты учащихся ЭК имели тенденцию быть выше результатов учащихся КК, т. е. уровень развития математической речи существенно выше у ЭК.

Итак, более высокие результаты диагностики свидетельствуют о более высоких показателях качества знаний и уровня развития математической речи у учащихся экспериментального класса. Статистическая обработка результатов показала значимость наблюдаемых различий.

Таким образом, высказанная гипотеза о влиянии методики обучения на уровень развития математической речи учащихся подтверждена экспериментально.

Задание 5. Докажите обоснованность приведенных А.С. Горчаковым заключений по проведенному педагогическому исследованию. В чем состоит отличие медианного критерия от *U*-критерия Вилкоксона – Манна – Уитни при сравнении результатов двух независимых выборок?

12.6. Критерий хи-квадрат

В педагогических исследованиях значительно чаще встречаются задачи, в которых необходимо сравнить два и более эмпирических распределения.

Приведем некоторые результаты диссертационного исследования, использующего данный критерий.

Сорокина Т.Н. Формирование правовой культуры младших школьников в организациях для детей-сирот и детей, оставшихся без попечения родителей: диссертация на соискание ученой степени кандидата педагогических наук: 13.00.01. – Саранск, 2015. – 161 с.

Цель исследования – теоретически обосновать модель формирования правовой культуры младших школьников в организациях для детей-сирот и детей, оставшихся без попечения родителей, выявить и апробировать педагогические условия эффективного ее функционирования.

В результате проведения эксперимента в организациях для детей-сирот и детей, оставшихся без попечения родителей, автором было установлено, что формирование правовой культуры у младших школьников экспериментальной и контрольной групп более успешно идет у детей из экспериментальной группы, чем у детей из КГ.

Полученные данные контрольного этапа свидетельствовали о том, что к концу педагогического эксперимента в обеих группах младших школьников произошел определенный рост уровня сформированности правовой культуры по всем выделенным критериям, при этом наиболее явные изменения прослеживались в ЭГ, в то время как в КГ они были не столь значительны.

Так, низкий уровень в КГ составляет 61 %, а в ЭГ – 4,5 %; средний уровень в КГ – 37 %, а в ЭГ – 48,5 %; высокий уровень в КГ составляет 2 %, а в ЭГ – 47 %.

Вместе с этим установлено, что за период проведения эксперимента в ЭГ низкий уровень правовой культуры младших школьников уменьшился на 68,5 %, средний уровень увеличился на 21,5 %, а высокий уровень поднялся на 47 %. За этот же период в КГ были выявлены следующие результаты: низкий уровень снизился на 6,5 %, средний уровень вырос на 4,5 %, а высокий – на 2 %, то есть результаты в КГ после эксперимента существенно не изменились.

Для определения совпадений и различий результатов в контрольной и экспериментальной группах автор провел количественный анализ полученных данных с использованием статистического критерия Пирсона (χ^2). Результаты вычислений и сравнение значений критерия Пирсона до и после эксперимента в контрольной и экспериментальной группах позволили распределить учеников по трем категориям: низкого, среднего и высокого уровней (табл. 21).

Таблица 21

Результаты эксперимента до и после применения педагогических условий по сформированности правовой культуры

КОГНИТИВНЫЙ КРИТЕРИЙ ПК					
Группа	Выборки		Категория 1 (низкий уровень)	Категория 2 (средний уровень)	Категория 3 (высокий уровень)
ЭГ	n_1=36	O_{1i}	23	13	0
	n_2=32	O_{2i}	9	14	9
КГ	n_1=37	O_{1i}	24	13	0
	n_2=39	O_{2i}	18	19	2
МОТИВАЦИОННО-ЦЕННОСТНЫЙ КРИТЕРИЙ ПК					
ЭГ	n_1=36	O_{1i}	9	26	1
	n_2=32	O_{2i}	9	14	9
КГ	n_1=37	O_{1i}	9	27	2
	n_2=39	O_{2i}	4	28	7
МИРОВОЗЗРЕНЧЕСКИЙ КРИТЕРИЙ ПК					
ЭГ	n_1=36	O_{1i}	6	27	3
	n_2=32	O_{2i}	9	14	9
КГ	n_1=37	O_{1i}	7	28	3
	n_2=39	O_{2i}	3	27	9
ПОВЕДЕНЧЕСКИЙ КРИТЕРИЙ ПК					
ЭГ	n_1=36	O_{1i}	6	27	3
	n_2=32	O_{2i}	9	14	9
КГ	n_1=37	O_{1i}	7	27	3
	n_2=39	O_{2i}	5	29	4

Результаты анкетирования учащихся в ЭГ и КГ используем для проверки гипотезы о том, что реализация педагогических условий формирования правовой культуры младших школьников в организациях для детей-сирот и детей, оставшихся без попечения родителей, способствует формированию высокого уровня показателей правовой культуры учащихся.

Выборки учащихся случайные и независимые. Анкетирование позволило выделить три категории шкалы измерений, и значительная часть экспериментальных данных представляет цепочки одинаковых значений, поэтому можно воспользоваться двусторонним критерием χ^2 (хи-квадрат), адаптированным для ситуаций, когда экспериментальные данные записаны в форме таблицы $2 \times C$, где C – размерность шкалы измерений (в нашем случае 2×3, так как $C = 3$).

В табл. 21 O_{1i} обозначает число учащихся до реализации педагогических условий формирования их правовой культуры, получивших оценку i ($i = 1, 2, 3$); O_{2i} – число младших школьников после внедрения педагогических условий, получивших оценку i ($i = 1, 2, 3$).

Обозначим p_{1i} ($i = 1, 2, 3$) вероятность выполнения работы учащимися до внедрения модели на оценку i, а p_{2i} ($i = 1, 2, 3$) – вероятность выполнения работы учащимися на оценку i после внедрения. На основе имеющихся данных можно проверить нулевую гипотезу H_0: $p_{1i} = p_{2i}$ для всех $C = 3$ категорий – при альтернативе H_1: $p_{1i} \neq p_{2i}$ хотя бы для одной из $C = 3$ категорий.

Для проверки данной гипотезы подсчет значения статистики критерия χ^2 был произведен по формуле

$$T = \frac{1}{n_1 \cdot n_2} \sum_{i=1}^{3} \frac{(n_1 O_{2i} - n_2 O_{1i})^2}{(O_{1i} + O_{2i})} \qquad (11)$$

при числе категорий $C = 3$.

По табл. 5 приложения находим критическое значение статистики критерия χ^2, приняв $p = 0{,}025$, т. е. примем верное решение с вероятностью $1 - p = 0{,}975$ с числом степеней свободы $v = C - 1 = 3 - 1 = 2$: $T_{кр} = x_{1-p} = 7{,}378$.

Таким образом, если $T_{эмп} > T_{кр}$, то в соответствии с правилом принятия решения полученные результаты дают достаточные основания для отклонения нулевой гипотезы. Иначе говоря, полученные результаты позволяют принять предположение о том, что реализация педагогических условий формирования правовой культуры младших школьников в организациях для детей-сирот и детей, оставшихся без попечения родителей, способствовала формированию высокого уровня показателей правовой культуры у учащихся ЭГ. В противном случае, если $T_{эмп} < T_{кр}$, нулевая гипотеза принимается.

Вернемся к диссертационному исследованию Е.А. Ермолаева на тему «Элективные курсы по геометрии в условиях профильного обучения математике в старших классах».

Напомним, что автором проводился обучающий этап экспериментального исследования в 2007–2010 гг. в 10-х классах профильных центров г.о. Тольятти (лицей № 19, СШ № 41, гимназия № 39) и в рамках спецкурса со студентами 4 курса ТГУ – будущими учителями математики. Основной целью этого этапа было определение:

– доступности содержания элективного курса «Площадь. Равновеликие и равносоставленные многоугольники»;
– влияния курса на качество знаний и умений учащихся по геометрии в целом;
– достижения основных целей и задач профильного обучения математике.

Схема этого этапа эксперимента имела вид: презентация курса $\to$ входной контроль $\to$ проведение курса $\to$ выходной контроль.

На презентации элективного курса выявлялись две группы учащихся: в ЭГ вошли учащиеся, которые выбрали данный курс для изучения, в КГ – все остальные.

Для выявления исходного уровня обученности геометрии до изучения элективного курса учащимся 10-х классов контрольной и экспериментальной групп на входном контроле была предложена диагностическая контрольная работа базового уровня геометрии по теме «Площадь многоугольников».

Для установления совпадения или различия у учащихся обеих групп исходных уровней обученности геометрии использовался критерий χ^2 (табл. 22–23).

Расчет зависимостей данных распределений производится по формуле

$$\chi^2_{\text{эмп}} = \frac{N^2}{n_1 \cdot n_2} \cdot \left(\sum_{i=1}^{k} \frac{f_1^2}{f_1 + f_2} - \frac{n_1^2}{N} \right),$$

где f_1 – частоты первого распределения; f_2 – частоты второго распределения; $N = n_1 + n_2$ – сумма числа элементов в первой и второй выборках.

Автором были сформулированы гипотезы H_0 (уровни обученности по геометрии обеих групп совпадают) и H_1 (уровни обученности по геометрии КГ и ЭГ статистически различны).

Для 2007/08 учебного года получили $N = n_1 + n_2 = 34 + 41 = 75$. Значит,

$$\chi^2_{\text{эмп}} = \frac{75^2}{34 \cdot 41} \cdot \left(16{,}43 - \frac{34^2}{75}\right) = 4{,}10.$$

Для 2008/09 учебного года получили $N = n_1 + n_2 = 38 + 39 = 77$. Значит,

$$\chi^2_{\text{эмп}} = \frac{77^2}{38 \cdot 39} \cdot \left(19{,}54 - \frac{38^2}{77}\right) = 3{,}14.$$

В первом случае число степеней свободы $v = (k - 1) \cdot (c - 1) = (7 - 1) \cdot (2 - 1) = 6$, где k – число интервалов разбиения; c – число столбцов. По табл. 5 приложения критических значений критерия χ^2 для степеней свободы $v = 6$ получим

$$\chi^2_{\text{кр}} = \begin{cases} 12{,}592 \text{ для } p \le 0{,}05, \\ 16{,}812 \text{ для } p \le 0{,}01 \end{cases}.$$

Во втором случае $v = 8$, значит,

$$\chi^2_{\text{кр}} = \begin{cases} 15{,}507 \text{ для } p \le 0{,}05, \\ 20{,}090 \text{ для } p \le 0{,}01 \end{cases}.$$

Таблица 22

Достоверность и объективность выводов входного контроля учащихся 10-х классов в 2007/08 учебном году

Кол-во выполненных заданий	Группы		f_1^2	$f_1 + f_2$	$\frac{f_1^2}{f_1 + f_2}$
	ЭГ, f_1	КГ, f_2			
0	1	4	1	5	0,20
1	7	8	49	15	3,27
2	4	4	16	8	2,00
3	5	9	25	14	1,79
4	9	6	81	15	5,40
5	3	6	9	9	1,00
6	5	4	25	9	2,78
Сумма	34	41			16,43

Таблица 23

Достоверность и объективность выводов входного контроля учащихся 10-х классов в 2008/09 учебном году

Кол-во выполненных заданий	Группы		f_1^2	f_1+f_2	$\frac{f_1^2}{f_1+f_2}$
	ЭГ, f_1	КГ, f_2			
0	1	0	1	1	1,00
1	3	2	9	5	1,80
2	3	5	9	8	1,13
3	7	5	49	12	4,08
4	8	6	64	14	4,57
5	6	8	36	14	2,57
6	4	6	16	10	1,60
7	3	3	9	6	1,50
8	3	4	9	7	1,29
Сумма	38	39			19,54

Результаты в первом и втором случаях $\chi^2_{эмп} = 4,10$ и $\chi^2_{эмп} = 3,14$ попали в зону незначимости, что свидетельствует об отсутствии статистически значимых различий между исходными уровнями обученности у учащихся ЭГ и КГ по геометрии.

В качестве выходного контроля учащимся тех же групп была предложена контрольная работа. Ее содержание построено на основе материала базового курса геометрии, изучаемого ими по традиционной методике у одного и того же учителя. При выполнении работы на каждый вопрос учащиеся должны были дать ответ «да» или «нет». Полученные результаты свидетельствовали о наличии существенных различий в уровнях обученности геометрии учащихся указанных групп. Для сравнения уровней обученности у учащихся обеих групп также использовался критерий χ^2 (табл. 24–25).

Автором были сформулированы две гипотезы: H_0 – элективный курс по геометрии «Площадь. Равновеликие и равносоставленные многоугольники» не влияет на уровень обученности учащихся – и H_1 – элективный курс по геометрии «Площадь. Равновеликие и равносоставленные многоугольники» влияет на уровень обученности учащихся.

Для 2007/08 учебного года

$$\chi^2_{\text{эмп}} = \frac{75^2}{34 \cdot 41} \cdot \left(21{,}44 - \frac{34^2}{75} \right) = 24{,}31.$$

В данном случае число степеней свободы

$$v = (k - 1) \cdot (c - 1) = (11 - 1) \cdot (2 - 1) = 10.$$

По таблице критических значений критерия χ^2 (табл. 5 приложения) для числа степеней свободы $v = 10$ получили

$$\chi^2_{\text{кр}} = \begin{cases} 18{,}307 \text{ для } p \le 0{,}05, \\ 23{,}209 \text{ для } p \le 0{,}01. \end{cases}$$

Таким образом, полученная величина $\chi^2_{\text{эмп}} = 24{,}31$ попала в зону значимости.

Для 2008/09 учебного года

$$\chi^2_{\text{эмп}} = \frac{77^2}{38 \cdot 39} \cdot \left(25{,}33 - \frac{38^2}{77} \right) = 26{,}33.$$

В данном случае число степеней свободы

$$v = (k - 1) \cdot (c - 1) = (13 - 1) \cdot (2 - 1) = 12.$$

По табл. 5 приложения для числа степеней свободы $v = 12$ получили

$$\chi^2_{\text{кр}} = \begin{cases} 21{,}026 \text{ для } p \le 0{,}05, \\ 26{,}217 \text{ для } p \le 0{,}01. \end{cases}$$

Таблица 24

Достоверность и объективность выводов
входного контроля учащихся 10-х классов в 2007/08 учебном году

Кол-во баллов	Группы		f_1^2	$f_1 + f_2$	$\frac{f_1^2}{f_1 + f_2}$
	ЭГ, f_1	КГ, f_2			
6	0	4	0	4	0,00
7	0	2	0	2	0,00
8	2	7	4	9	0,44
9	2	3	4	5	0,80
10	0	5	0	5	0,00
11	1	4	1	5	0,20
12	5	7	25	12	2,08
13	8	4	64	12	5,33
14	5	1	25	6	4,17
15	4	3	16	7	2,29
16	7	1	49	8	6,13
Сумма	34	41			21,44

Таблица 25

Достоверность и объективность выводов входного контроля учащихся 10-х классов в 2008/09 учебном году

Кол-во баллов	Группы		f_1^2	f_1+f_2	$\frac{f_1^2}{f_1+f_2}$
	ЭГ, f_1	КГ, f_2			
4	0	1	0	1	0,00
5	0	1	0	1	0,00
6	0	3	0	3	0,00
7	0	2	0	2	0,00
8	1	6	1	7	0,14
9	1	3	1	4	0,25
10	2	5	4	7	0,57
11	2	4	4	6	0,67
12	4	5	16	9	1,78
13	3	3	9	6	1,50
14	4	1	16	5	3,20
15	11	1	121	12	10,08
16	10	4	100	14	7,14
сумма	38	39			25,33

Таким образом, полученная величина $\chi^2_{эмп} = 26{,}33$ попала в зону значимости. Значит, в обоих случаях гипотеза H_0 отклоняется, и следует принять гипотезу H_1 о том, что в контрольной и экспериментальной группах различия в уровнях обученности геометрии после изучения элективного курса «Площадь. Равновеликие и равносоставленные многоугольники» являются существенными.

Полученные результаты, а также результаты самостоятельных и контрольных работ, устные ответы учащихся, наблюдения за процессом решения задач, анализ выполнения домашних заданий позволили Е.А. Ермолаеву установить, что учащиеся сознательно усваивают предлагаемый материал. Это свидетельствует о том, что содержание предлагаемого элективного курса является доступным для учащихся 10-х классов математического профиля и повышает уровень их обученности.

В качестве подтверждения последнего в рамках эксперимента также проверялось влияние изучения элективного курса «Площадь. Равновеликие и равносоставленные многоугольники» на уровень успеваемости учащихся по геометрии. Для этого сравнивались итоговые оценки по геометрии учащихся, результаты контрольных работ до изучения элективного курса (9 класс) и после него (10 класс).

Для статистической обработки данных в этом случае использовался критерий знаков *G* (п. 5.1).

Задание 6. Докажите обоснованность приведенных Т.Н. Сорокиной и Е.А. Ермолаевым заключений по проведенным педагогическим исследованиям. Выявите целесообразность применения данного критерия с учетом различных случаев сравнения двух экспериментальных распределений.

12.7. Критерий Стьюдента (*t*-критерий)

Приведем некоторые результаты диссертационного исследования, использующего критерий Стьюдента.

Виноградова А.В. Устная работа по геометрии как средство организации коммуникативной деятельности старшеклассников: диссертация на соискание ученой степени кандидата педагогических наук: 13.00.02. – М., 2016. – 248 с.

Цель исследования состоит в теоретическом обосновании и разработке дидактической модели организации коммуникативной деятельности (КД) старшеклассников и методики реализации этой модели в процессе проведения устной работы при обучении геометрии.

На обучающем этапе эксперимента применялся метод диагностирующих проверочных работ. Их основной целью была проверка в КК и ЭК того, как формируются необходимые коммуникативные качества старшеклассников в процессе обучения геометрии при проведении устной работы: а) на начало эксперимента; б) через два месяца после начала; в) через полгода; г) на конец эксперимента.

Результаты проверочных работ на начало и конец эксперимента представлены соответственно в табл. 26–27. Уровни (высокий, хороший, средний и низкий) определялись по четким критериям.

Таблица 26

Уровни сформированности предметной КД старшеклассников на начало эксперимента

Классы	Число учащихся	Высокий	Хороший	Средний	Низкий
КК	75	9	27	30	9
	100 %	12 %	36 %	40 %	12 %
ЭК	75	10	28	29	8
	100 %	13 %	37 %	39 %	11 %

Таблица 27

Уровни сформированности предметной коммуникативной деятельности старшеклассников на конец эксперимента

Классы	Число учащихся	Высокий	Хороший	Средний	Низкий
КК	75	10	29	29	7
	100 %	13 %	39 %	39 %	9 %
ЭК	75	13	32	25	5
	100 %	17 %	43 %	33 %	7 %

Далее автор приводит обобщенные результаты определения предметных уровней сформированности коммуникативной деятельности старшеклассников в контрольном и экспериментальном классах на начало и конец эксперимента (табл. 28–29).

Таблица 28

Сравнительный анализ предметных уровней в экспериментальных и контрольных классах

	Высокий	Хороший	Средний	Низкий
КК в начале эксперимента	9	27	30	9
КК в конце эксперимента	10	29	29	7
ЭК в начале эксперимента	9	27	30	9
ЭК в конце эксперимента	13	32	25	5

Таблица 29

Результаты эксперимента по определению сформированности предметной коммуникативной деятельности старшеклассников

Контрольные классы			Экспериментальные классы		
№ п/п	Кол-во баллов	$(x_i - x_{ср})^2$	№ п/п	Кол-во баллов	$(y_i - y_{ср})^2$
1	12	2,38	1	14	8,23
2	14	12,54	2	13	3,50
3	10	0,21	3	12	0,76
4	11	0,29	4	10	1,28
5	7	11,96	5	10	1,28
6	8	6,04	6	11	0,02
7	14	12,54	7	15	14,97
8	9	2,13	8	8	9,80
9	10	0,21	9	13	3,50
10	13	6,46	10	9	4,54
11	8	6,04	11	12	0,76
12	13	6,46	12	11	0,02
13	9	2,13	13	10	1,28
14	11	0,29	14	9	4,54
15	10	0,21	15	13	3,50
16	14	12,54	16	10	1,28
17	12	2,38	17	14	8,23
18	10	0,21	18	6	26,32
19	12	2,38	19	11	0,02
20	8	6,04	20	7	17,06
21	11	0,29	21	11	0,02
22	8	6,04	22	12	0,76
23	11	0,29	23	15	14,97
24	6	19,88			
Среднеарифметическое КК	10,46		Среднеарифметическое ЭК	11,13	
Сумма		119,96	Сумма		126,61
Стандартная ошибка разности		0,68			
t		9,74			

Составленная на основе полученных данных таблица отражает предметные коммуникативные знания и умения в контрольном и экспериментальном классах, определенные по представленным критериям результатов. Полученные результаты экспериментальной проверки полностью не следует относить к эффекту проведенного эксперимента, поскольку на них оказывает влияние не только экспериментальный фактор, но и вся учебная и внеурочная деятельность обучающегося. Поэтому значимость полученных результатов была проверена с использованием критерия Стьюдента.

Так как полученное в эксперименте эмпирическое значение $t_{эмп} = 9{,}74$ превышает табличное $t_{крит} = 2{,}01$ (при $v = 24 + 23 - 2 = 45$), то автор делает вывод о преимуществе экспериментального обучения.

Это означает, что во всех аналогичных случаях вероятность воспроизведения элементов знаний и умений обучающимися по разработанной методике организации КД старшеклассников в процессе проведения устной работы при обучении геометрии будет больше, чем вероятность воспроизведения знаний и умений учащимися, в обучении которых эта методика не применялась. Достоверность этого вывода около 95 %.

Отмечается, что оценка сформированности КД на предметном уровне не может являться объективной без личностных и социальных составляющих, так как происходит одновременное формирование коммуникативного опыта на метапредметном уровне. В соответствии с разработанной тестовой картой, которая включает вышеназванные уровни сформированности КД старшеклассников, было проведено дополнительное тестирование учащихся (табл. 30, 31).

Таблица 30

Уровни сформированности коммуникативной деятельности старшеклассников на начало эксперимента

Классы	Число учащихся	Высокий	Хороший	Средний	Низкий
КК	75	9	32	27	7
	100 %	12 %	43 %	36 %	9 %
ЭК	75	10	31	28	6
	100 %	13 %	41 %	38 %	8 %

Таблица 31

Уровни сформированности коммуникативной деятельности старшеклассников на конец эксперимента

Классы	Число учащихся	Высокий	Хороший	Средний	Низкий
КК	75	10	35	25	5
	100 %	13 %	47 %	33 %	7 %
ЭК	75	13	36	23	3
	100 %	17 %	48 %	31 %	4 %

Теперь сопоставим результаты определения предметных уровней сформированности КД старшеклассников в контрольных и экспериментальных классах на начало и конец эксперимента (табл. 32).

Таблица 32

Сравнительный анализ результатов в экспериментальном и контрольном классах

	Высокий	Хороший	Средний	Низкий
КК в начале эксперимента	9	32	27	7
КК в конце эксперимента	10	35	25	5
ЭК в начале эксперимента	10	31	28	6
ЭК в конце эксперимента	13	36	23	3

Обобщим результаты в табл. 33.

Таблица 33

Результаты эксперимента по определению сформированности коммуникативной деятельности старшеклассников

Контрольный класс		24	Экспериментальный класс		23
№ п/п	Кол-во баллов	$(x_i - x_{ср})^2$	№ п/п	Кол-во баллов	$(y_i - y_{ср})^2$
1	24	20,25	1	25	0,55
2	22	6,25	2	34	94,85
3	12	56,25	3	31	45,42
4	21	2,25	4	29	22,46
5	11	72,25	5	17	52,72
6	20	0,25	6	30	32,94
7	15	20,25	7	26	3,02
8	33	182,25	8	16	68,24

Контрольный класс		24	Экспериментальный класс		23
№ п/п	Кол-во баллов	$(x_i - x_{ср})^2$	№ п/п	Кол-во баллов	$(y_i - y_{ср})^2$
9	8	132,25	9	33	76,37
10	21	2,25	10	15	85,76
11	15	20,25	11	16	68,24
12	23	12,25	12	27	7,50
13	14	30,25	13	25	0,55
14	24	20,25	14	18	39,20
15	16	12,25	15	23	1,59
16	32	156,25	16	30	32,94
17	7	156,25	17	17	52,72
18	30	110,25	18	19	27,68
19	13	42,25	19	28	13,98
20	22	6,25	20	29	22,46
21	12	56,25	21	9	232,89
22	32	156,25	22	26	3,02
23	21	2,25	23	35	115,33
24	20	0,25			
Среднеарифметическое КК	19,50		Среднеарифметическое ЭК	24,26	
Сумма		1276,00			1100,43
Стандартная ошибка разности		2,12			
t эмпирич.		2,25			
t критич.		2,01			

Так как полученное в эксперименте эмпирическое значение $t_{эмп} = 2,25$ превышает табличное $t_{крит} = 2,01$ (при $v = 24 + 23 - 2 = 45$), то автор делает вывод о преимуществе экспериментального обучения.

Это означает, что во всех аналогичных случаях вероятность воспроизведения проверяемых элементов у обучающихся по разработанной методике будет больше, чем вероятность воспроизведения знаний и умений учащимися, в обучении которых эта методика не применялась. Достоверность этого вывода около 95 %.

Анализ полученных результатов педагогического эксперимента свидетельствует о повышении уровней коммуникативной деятельности учащихся: в экспериментальных классах они выше, чем у учащихся контрольных классов. В целом, результаты эксперимента показывают, что использование предложенной методики организации коммуникативной деятельности старшеклассников при проведении устной работы по геометрии способствует достижению современных образовательных результатов.

Задание 7. Выявите неточности и ошибки автора, если таковые имеются, при обосновании приведенных заключений по проведенному педагогическому исследованию. Какой из двух случаев применения *t*-критерия Стьюдента использовал автор, насколько это оправданно?

12.8. Коэффициент корреляции Пирсона

Приведем некоторые результаты диссертационного исследования, использующего данный критерий.

Трояновская Н.И. Технология формирования действий контроля и оценки учащихся 5–6 классов в обучении математике: диссертация на соискание ученой степени кандидата педагогических наук: 13.00.02. – Нижний Новгород, 2015. – 228 с.

Цель исследования заключается в разработке технологии формирования у учащихся действий контроля и оценки, которая обеспечит развитие контрольно-оценочной самостоятельности учащихся и позволит повысить уровень их математической подготовки.

Формирующий этап эксперимента проводился с целью внедрения технологии формирования действий контроля и оценки учащихся и проверки ее эффективности на разных этапах обучения математике.

В ходе эксперимента важным представлялось подтвердить или опровергнуть гипотезу, сформулированную в ходе теоретического исследования: если в процессе обучения учащихся математике использовать разработанную технологию формирования действий контроля и оценки, то это позволит целенаправленно формировать у учащихся действия контроля и оценки, обеспечивающие развитие контрольно-оценочной самостоятельности учащихся и повышение уровня их математической подготовки.

Эксперимент проводился на базе МАОУ СОШ № 186 «Авторская академическая школа» г. Нижнего Новгорода с 2005 по 2010 г. В нем принимали участие учащиеся 5-х классов.

Для подтверждения эффективности разработанной технологии формирования действий контроля и оценки у учащихся 5-х классов по данным выборок (результаты выполнения заданий итоговой работы ЭК и КК) было построено эмпирическое уравнение регрессии:

$$y = 1{,}2013\,x - 0{,}362,$$

в котором эмпирические коэффициенты регрессии $b = 1{,}2013$, $a = -0{,}362$).

Для оценки параметров α и β регрессионной модели использовался метод наименьших квадратов:

$$S = \sum(y_i - y^*_i)^2 \to \min.$$

Таблица 34

Промежуточные вычисления для расчета параметров регрессии

N	x	y	x^2	y^2	$x \cdot y$
Сумма	18,7	14,5	16,69	11,21	13,28

На основании данных табл. 34 подсчитаны выборочные средние, выборочные дисперсии и среднеквадратичное отклонение:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{18{,}7}{22} \approx 0{,}85;\ \bar{y} = \frac{\sum y_i}{n} = \frac{14{,}5}{22} \approx 0{,}66;$$

$$\overline{x \cdot y} = \frac{\sum x_i y_i}{n} = \frac{13{,}28}{22} \approx 0{,}6;$$

$$S^2(x) = \frac{\sum {x_i}^2}{n} - \bar{x}^2 = \frac{16{,}69}{22} - 0{,}85^2 \approx 0{,}0361;$$

$$S^2(y) = \frac{\sum {y_i}^2}{n} - \bar{y}^2 = \frac{11{,}21}{22} - 0{,}66^2 \approx 0{,}0751;$$

$$S(x) = \sqrt{S^2(x)} = \sqrt{0{,}0361} \approx 0{,}19;$$

$$S(y) = \sqrt{S^2(y)} = \sqrt{0{,}0751} \approx 0{,}274.$$

Далее был определен выборочный линейный коэффициент корреляции по формуле

$$r_{xy} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{S(x) \cdot S(y)} = \frac{0{,}6 - 0{,}85 \cdot 0{,}66}{0{,}19 \cdot 0{,}274} \approx 0{,}833.$$

По шкале Чеддока была определена степень связи – высокая и прямая, так как $0{,}7 < r_{xy} < 0{,}9$.

Для оценки качества параметров регрессии была построена табл. 35.

Таблица 35

Промежуточные вычисления
для оценки качества параметров регрессии

N	x (ЭК)	y (КК)	$y(x)$	$(y_i - \bar{y})^2$	$(y - y(x))^2$	$(x_i - \bar{x})^2$
Сумма	18,7	14,5	14,5	1,65	0,51	0,8

Оценка значимости коэффициента корреляции проводилась с использованием t-критерия Стьюдента.

Были сформулированы следующие гипотезы: H_0 – уровни сформированности действий контроля и оценки учащихся КК и ЭК незначимо отличаются; H_1 – уровни сформированности действий контроля и оценки учащихся КК и ЭК существенно различаются.

При полученных значениях t-критерия Стьюдента в ходе обработки результатов было получено $t_{\text{эмп}} = 6{,}73$.

По таблице Стьюдента (табл. 6 прил.) по уровню значимости $p = 0{,}05$ и числу степеней свободы $k = 20$ была найдена критическая точка $t_{\text{крит}} = 2{,}086$. Так как $|t_{\text{эмп}}| > t_{\text{крит}}$, нулевая гипотеза опровергается на высоком уровне значимости.

Следовательно, уровни сформированности действий контроля и оценки учащихся КК и ЭК существенно различаются, что является результатом внедрения в ЭК разработанной технологии.

На основе результатов выполнения учащимися КК и ЭК заданий итоговой работы был подсчитан коэффициент корреляции Пирсона, который позволяет установить зависимость формирования действий контроля и оценки от использования технологии формирования этих действий.

На основе полученных данных (табл. 36) была найдена зависимость сформированности действий контроля и оценки от применяемой технологии.

Коэффициент Пирсона вычислялся по следующей формуле:

$$\varphi = \frac{a \cdot d - b \cdot c}{\sqrt{(a+b)\cdot(c+d)\cdot(a+c)\cdot(b+d)}},$$

где *a* – количество учащихся ЭК, выполнивших задание; *b* – количество учащихся КК, выполнивших задание; *c* – количество учащихся ЭК, не выполнивших задание; *d* – количество учащихся КК, не выполнивших задание.

Таблица 36

Экспериментальные данные

	Номер задания														
	1			2			3			4			5		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
a	22	18	10	21	20	21	22	12	7	17	18	22	22	17	7
c	0	4	12	1	2	1	0	10	15	5	4	0	0	5	15
b	16	6	4	17	12	19	17	1	2	16	6	20	20	2	0
d	5	15	17	4	9	2	4	20	19	5	15	1	1	19	21
φ	0,37472	0,544904	0,261989	0,150671	0,334367	0,014389	0,231349	0,555062	0,258936	0,024677	0,544904	0,159595	0,159595	0,689515	0,415618

Из табл. 36 видно, что коэффициент корреляции по заданиям № 1.1–2.2, 3.1–5.1, 5.3 лежит в пределах от 0 до ±0,70, что означает ясно выраженную корреляцию. Значит, существует ярко выраженная зависимость результатов от используемой технологии обучения (традиционной в КК и экспериментальной технологии формирования действий контроля и оценки в ЭК).

В ходе эксперимента проводилась проверка эффективности основных положений, на которых базируется разрабатываемая технология, отслеживалась динамика в учении, в развитии детей, включенных в формирующий эксперимент, изучались возможности учителей по использованию новой технологии на практике.

Задание 8. Выявите неточности и ошибки, если таковые имеются, при обосновании приведенных Н.И. Трояновской заключений по проведенному педагогическому исследованию.

Задания для самостоятельной работы

1. По статистическому распределению выборки установите её объем.

x_i	1	2	3	4
n_i	3	2	1	7

2. Дан следующий вариационный ряд:

1 2 3 4 5 6 7 8 9 10

1 1 2 2 4 4 4 5 5 5

Требуется:

1) построить полигон распределения;
2) вычислить выборочную среднюю, дисперсию, моду, медиану;
3) построить выборочную функцию распределения.

3. Дана выборка: 10, 20, 20, 5, 15, 20, 5, 10, 20, 5. Требуется:

а) построить статистический ряд распределения частот;
б) вариационный ряд;
в) найти выборочные моду и медиану;
г) построить гистограмму распределения частот.

4. Используя критерий Пирсона, при уровне значимости 0,05 проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности X по результатам выборки.

X	0,3	0,5	0,7	0,9	1,1	1,3	1,5	1,7	1,9	2,1	2,3
N	7	9	28	27	30	26	21	25	22	9	5

5. Найдите размах значений из ряда данных: 37; 42; 35; 58; 33; 38; 51.

6. Найдите медиану значений из ряда данных: 23; 24; 15; 15; 33; 28; 14; 14.

7. Найдите среднее арифметическое значений из ряда данных: 13, 42, 35, 35, 13, 42, 42.

8. Найдите разность между средним арифметическим и медианой значений из ряда данных: 37, 37, 35, 58, 37, 38, 57.

9. Найдите моду значений из ряда данных: 14; 42; 42; 15; 58; 33; 38; 52, 52.

10. Дан ряд данных: 2,8; 2,9; 3,5; 3,7; 3,8; 4,1; 4,4. Найдите, на сколько отличается медиана от среднего арифметического этого набора чисел.

11. Найти *m* и *n*, если их среднее арифметическое равно 36 и $\frac{1}{5}$ их разности равна 0,8.

12. Стоимость (в рублях) пакета молока «Вкуснотеево» в магазинах микрорайона образует следующий ряд данных: 34, 35, 34, 37, 38, 37, 37. Найдите разность между средним арифметическим и медианой этого ряда данных.

13. Зарплата руководителя отдела компании составляет 70 000 руб., зарплата трёх его заместителей – по 50 000 руб., а зарплата 20 рядовых сотрудников отдела – по 25 000 руб. в месяц. Найти среднее арифметическое зарплат всех сотрудников данного отдела компании.

14. На легкоатлетическом соревновании спортсмены пересекали финишную прямую со следующими интервалами: 37 с; 2 мин 58 с; 1 мин 26 с; 43 с; 2 мин 16 с; 1 мин 42 с. Найти среднее значение этого ряда данных.

15. Оценки за контрольную работу по физике в 9-х классах школы распределились следующим образом: 9 «А» класс – две «двойки», пять «троек», восемь «четвёрок», три «пятёрки»; 9 «Б» класс – две «двойки», двенадцать «троек», две «четвёрки», одна «пятёрка». Найти медиану оценок за контрольную всех учащихся.

16. Рост Вити равен 162 см, а медиана ряда чисел, являющихся показателями роста (в см) всех мальчиков его класса, равна 161. Какие из приведённых ниже утверждений являются неверными?
A. В классе обязательно есть мальчик ростом 161 см.
B. В классе обязательно есть мальчик ростом менее 161 см.
C. В классе обязательно есть мальчик ростом больше Вити.
D. В классе обязательно есть мальчик ростом меньше Вити.

17. Рост Тани равен 154 см, а средний рост всех девочек её класса составляет 153 см. Какие из приведённых ниже утверждений являются верными?
A. В классе обязательно есть девочка ростом 153 см.
B. В классе обязательно есть девочка ростом 152 см.

C. В классе обязательно есть девочка ростом меньше 153 см.
D. В классе обязательно есть девочка, рост которой меньше 154 см, но больше 153 см.

18. В чем заключается отличие Т-критерия Вилкоксона и критерия Фридмана?

19. Объясните, почему критерий Пейджа можно рассматривать как эквивалент критерия Фридмана.

20. Сформулируйте условия для применения критерия Колмогорова – Смирнова.

21. В чем отличие и сходство при обработке результатов эксперимента с помощью корреляций по Пирсону и по Спирмену?

22. Установите, имеется ли достаточно выраженная связь между признаками (измерения по номинативной шкале).

Наличие 1 признака	+	+	–	–	+	–	–	+	+	+	+	–
Наличие 2 признака	+	–	–	+	–	+	–	+	+	–	–	+

23. Установите, имеется ли достаточно выраженная связь между признаками (измерения по ранговой шкале).

1 признак	3	4	4	3	5	2	1	3	3	4	4	4
2 признак	4	4	4	5	3	2	3	3	4	3	2	3

24. Составьте анкету для учителей и анкету для учащихся общеобразовательной школы по определенному педагогическому исследованию. Поясните, каким образом можно обработать их результаты с помощью статистических методов.

25. В одном и том же классе проводился опрос: нравится ли тебе урок «Окружающий мир»? Между срезами была проведена целенаправленная работа, включающая экскурсии на природу для повышения познавательного интереса к данному предмету. Оцените по приведенным данным результативность примененного средства, сформулируйте статистические гипотезы (ответы: нравится – «+», не нравится – «–»):

№ уч-ся	1	2	3	4	5	6	7	8	9	10	11	12
Срез 1	+	+	–	–	+	–	–	+	+	+	+	+
Срез 2	+	–	–	+	+	+	–	+	+	+	–	+

26. Проведены два опроса учащихся младших классов: нравится ли тебе ходить в школу? Между опросами была проведена работа по формированию положительной мотивации. Сформулируйте статистические гипотезы и проверьте результативность этой работы на основе следующих данных (ответы: нравится – «+», не нравится – «–»):

№ уч-ся	1	2	3	4	5	6	7	8	9	10	11	12
Срез 1	+	–	+	–	+	–	+	+	–	+	–	+
Срез 2	+	+	+	+	–	+	+	+	+	–	+	–

27. Для развития логического мышления была разработана специальная методика. Оцените ее результативность на основе двух серий срезов в одном классе (контрольные работы, стандартное оценивание – оценки «2», «3», «4», «5»), сформулируйте статистические гипотезы.

№ уч-ся	Срез 1	Срез 2	№ уч-ся	Срез 1	Срез 2
1	2	3	13	2	2
2	3	3	14	5	4
3	4	4	15	5	5
4	2	4	16	4	4
5	5	3	17	4	3
6	5	4	18	4	5
7	2	3	19	3	4
8	3	4	20	3	4
9	4	4	21	3	3
10	4	5	22	5	4
11	4	3	23	3	2
12	3	3	24	3	3

28. Развитие коммуникабельности учащихся осуществлялось на базе использования коллективных форм работы. Проверьте результативность проделанной работы на основе двух серий срезов в экспериментальном классе, где уровень коммуникабельности замерялся с помощью метода рангового оценивания.

№ уч-ся	1	2	3	4	5	6	7	8	9	10
1-е выполнение	3	5	3	6	3	5	3	3	2	2
2-е выполнение	4	5	3	6	3	4	4	4	3	4

29. Результаты двукратного выполнения работы (в баллах) 10 учащихся записаны в виде таблицы:

№ уч-ся	1	2	3	4	5	6	7	8	9	10
1-е выполнение	2	2	2	3	3	3	3	3	3	2
2-е выполнение	2	3	3	2	3	2	4	4	3	4

Сформулируйте статистические гипотезы и выясните, какая подтвердилась.

30. Для сопоставления уровня визуальной памяти в двух классах были проведены замеры по шкале рангов. Результаты представлены в таблице:

КК 1	3	4	5	4	6	6	4	5	7	4	5	5
ЭК 1	6	4	5	4	3	5	5	6	7	7	7	5
КК 2	4	2	7	6	5	5	4	6	3	4	4	4
ЭК 2	7	3	5	7	5	6	4	7	7	4	3	5

Являются ли различия в уровне развития визуальной памяти у учащихся двух этих классов статистически значимыми?

Требования к содержанию устного ответа студента

Логика рецензирования устного ответа:
- последовательность изложения материала;
- полнота ответа;
- доказательность и самостоятельность суждения;
- умение обобщать и делать выводы;
- грамотность и выразительность речи;
- эмоциональность выступления;
- владение элементами ораторского искусства.

Критерии оценки устного ответа

Оценка «отлично» ставится студенту, который:
- владеет глубокими базисными знаниями и умениями;
- четко воспроизводит учебную информацию;

- осуществляет активный перенос знаний в смежные дисциплины и темы;
- использует научную терминологию;
- аргументирует свою точку зрения и свое отношение к данному вопросу;
- использует ссылки на авторов основных теоретических положений, данные первоисточников;
- материал излагает грамотно, логично, уверенно;
- владеет элементами ораторского искусства;
- допустил 1—2 неточности при освещении второстепенных вопросов, которые легко исправил в процессе выступления.

Оценку «хорошо» получает студент, который:

- показал хорошее знание теоретических основ вопроса;
- выделяет главные идеи и понимает сущность изученного материала;
- осуществляет перенос знаний в смежные дисциплины, темы;
- верно использует научную терминологию;
- аргументирует свою точку зрения и свое отношение к данному вопросу;
- использует данные первоисточников, ссылки на авторов основных теоретических положений при обосновании содержания ответа;
- допускает незначительные по содержанию неточности и устраняет их при ответе на дополнительные вопросы;
- владеет элементами ораторского искусства.

Оценку «удовлетворительно» ставят студенту, если он:

- имеет базисные знания и умения;
- воспроизводя учебную информацию, выделяет главные идеи;
- ориентируется в основных понятиях;
- частично осуществляет перенос знаний в смежные дисциплины, темы;
- испытывает затруднения в осуществлении анализа, сравнения, обобщения;
- недостаточно аргументирует свою точку зрения;
- недостаточно использует данные первоисточников, ссылки на авторов основных теоретических положений при обосновании содержания ответа;
- недостаточно владеет элементами ораторского искусства.

Оценка «неудовлетворительно» ставится студенту, который:

- не владеет основными понятиями;
- неточно воспроизводит материал;
- не выделяет главные идеи;
- не понимает сущности изученного материала;
- не осуществляет перенос знаний в смежные дисциплины, темы;
- продемонстрировал отсутствие логики в ответе;
- не владеет элементами ораторского искусства.

Требования к содержанию письменного ответа студента

1. Содержательное соответствие предложенному тексту.
2. Ссылки на максимальное количество научных источников и электронные базы.
3. Методологическая обоснованность – построение текста в соответствии с уровнями методологии научного знания.
4. Умение анализировать различные теории и результаты исследований.
5. Критический научный анализ излагаемых положений.
6. Способность предложить собственные обоснованные варианты решения теоретических и практических задач.
7. Уважительное отношение к авторам различных теоретических концепций.
8. Четкое выделение понятий, существенных элементов теории и концепций.
9. Структурирование изложения материала.
10. Понятийно-терминологическая обоснованность.

Критерии оценки письменного ответа студента

За письменный ответ выставляется оценка:

– «отлично» – при полном соответствии всем десяти требованиям, отсутствии ошибок, неточностей; при проявлении умения самостоятельно и творчески излагать мысли письменно.

Допускается не более двух неточностей в письменном ответе;

– «хорошо» – при полном соответствии не менее восьми требованиям; при проявлении умения самостоятельно и творчески излагать мысли письменно.

Допускается не более двух неточностей и (или) двух ошибок в письменном ответе;

– «удовлетворительно» – при полном соответствии не менее шести требованиям; при проявлении умения самостоятельно и творчески излагать мысли письменно.

Допускается не более трех неточностей и (или) трех ошибок в письменном ответе;

– «неудовлетворительно» – при полном несоответствии более пяти требованиям; при полном отсутствии текста, имеющего отношение к вопросу.

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Основная литература

1. Березина, Н.А. Учебное пособие по высшей математике : курс лекций [Электронный ресурс] / Н.А. Березина. – Саратов : Научная книга, 2012. – 159 с.
2. Буре, В.М. Теория вероятностей и математическая статистика : учебник [Электронный ресурс] / В.М. Буре, Е.М. Парилина. – СПб. : Лань, 2013. – 416 с.
3. Буре, В.М. Методы прикладной статистики в R и Excel : учеб. пособие [Электронный ресурс] / В.М. Буре, Е.М. Парилина, А.А. Седаков. – СПб. : Лань, 2016. – 152 с.
4. Горлач, Б.А. Теория вероятностей и математическая статистика : учебное пособие [Электронный ресурс] / Б.А. Горлач. – СПб. : Лань, 2013. – 320 с.
5. Кацман, Ю.Я. Теория вероятностей, математическая статистика и случайные процессы : учебник [Электронный ресурс] / Ю.Я. Кацман; Томский политехнический университет. – Томск : ТПУ, 2013. – 130 с.
6. Крупин, В.Г. Высшая математика. Теория вероятностей, математическая статистика, случайные процессы: сборник задач с решениями : учебное пособие для вузов [Электронный ресурс] / В.Г. Крупин, А.Л. Павлов, Л.Г. Попов. – М. : МЭИ, 2013. – 408 с.
7. Свешников, А.А. Сборник задач по теории вероятностей, математической статистике и теории случайных функций : учеб. пособие [Электронный ресурс] / А.А. Свешников ; под общ. ред. А.А. Свешникова. –5-е изд., стер. – СПб. : Лань, 2013. – 446 с.
8. Симчера, В.М. Методы многомерного анализа статистических данных : учеб. пособие [Электронный ресурс] / В.М. Симчера. – М. : Финансы и статистика, 2014. – 400 с.

Дополнительная литература

9. Антонова, И.В. О применении некоторых статистических методов в педагогических исследованиях при обучении математике учащихся общеобразовательной школы / И.В. Антонова // Практика использования естественно-научных методов в прикладных социально-гуманитарных исследованиях: сборник материалов

научно-методического семинара. В 2 ч. Ч. 2 / под научн. ред. И.В. Цветковой. – Тольятти : Изд-во ТГУ, 2014. – С. 185–197.
10. Герасевич, В.А. Современное программное обеспечение для статистической обработки биомедицинских исследований / В.А. Герасевич, А.Р. Аветисов // Белорусский медицинский журнал. – 2005. – № 1. – URL: http://miklebig.narod.ru/docum/statprog.htm (дата обращения: 12.05.2016).
11. Грабарь, М.И. Применение математической статистики в педагогических исследованиях. Непараметрические методы / М.И. Грабарь, К.А. Краснянская. – М. : Педагогика, 1977. – 138 с.
12. Елизарова, Н.Н. Использование программных средств статистической обработки данных при формировании информационного обеспечения управления / Н.Н. Елизарова // Вестник ИГЭУ. – 2009. – Вып. 3. – URL: http://vestnik.ispu.ru/sites/vestnik.ispu.ru/files/publications/76-80_1.pdf (дата обращения: 30.04.2016).
13. Ермолаев, Е.Ю. Математическая статистика для психологов : учебник / Е.Ю. Ермолаев. – 2-е изд., испр. – М. : Московский психолого-социальный институт: Флинта, 2003. – 336 с.
14. Наследов, А.Д. Математические методы психологического исследования. Анализ и интерпретация данных : учебное пособие / А.Д. Наследов. – 3-е изд., стереотип. – СПб. : Речь, 2007. – 392 с.
15. Саранцев, Г.И. Методология обучения математике / Г.И. Саранцев. – Саранск : Красный Октябрь, 2001. – 144 с.
16. Сидоренко, Е.В. Методы математической обработки в психологии / Е.В. Сидоренко. – СПб. : Социально-психологический центр, 1996. – С. 34.

ГЛОССАРИЙ

Измерение – процедура, с помощью которой измеряемый объект сравнивается с некоторым эталоном и получает численное выражение в определенном масштабе или шкале.

Типы измерительных шкал – номинативная, номинальная или шкала наименований; порядковая, ординарная или ранговая шкала; интервальная или шкала равных интервалов; шкала равных отношений, или шкала отношений.

Качественные измерения – измерения, осуществляемые с помощью номинативной и порядковой шкал.

Количественные измерения – измерения, осуществляемые с помощью интервальной шкалы и шкалы отношений.

Дискретные шкалы – шкалы, приводящие к качественным измерениям.

Непрерывные шкалы – шкалы, приводящие к количественным измерениям.

Метрические шкалы предполагают возможность применения единицы измерения (метрики). К ним относятся интервальная шкала и шкала отношений.

Неметрические шкалы не предполагают применения единицы измерения (метрики) (номинативная и порядковая шкалы).

Кодирование – операция, с помощью которой экспериментальным данным придается форма числового сообщения (кода).

Номинативная шкала подразумевает присваивание какому-либо свойству или признаку определенного обозначения или символа; по определенному правилу осуществляется классификация объектов на непересекающиеся классы. Самая простая шкала этого типа – *дихотомическая,* в которой измеряемые признаки можно кодировать двумя символами или цифрами.

Порядковая шкала предполагает приписывание объектам чисел в зависимости от степени выраженности измеряемого свойства; расчленяет всю совокупность измеренных признаков на такие множества, которые связаны между собой отношениями типа «больше – меньше», «выше – ниже», «сильнее – слабее» и т. п. Все признаки располагаются по рангу – от самого большого до самого маленького или наоборот.

Интервальная шкала – измерение, при котором числа отражают не только различия между объектами в уровне выраженности свойства (характеристика порядковой шкалы), но и то, на сколько больше или меньше выражено свойство. При работе с этой шкалой измеряемому свойству или предмету присваивается число, равное количеству единиц измерения, эквивалентному количеству имеющегося свойства. Особенность данной шкалы – отсутствие у нее естественной точки отсчета (нуль условен и не указывает на отсутствие измеряемого свойства).

Шкала отношений – измерение, отличающееся от интервального наличием твердо фиксированного нуля, который означает полное отсутствие какого-либо свойства или признака. Данная шкала является наиболее информативной, допускающей любые математические операции и использование разнообразных статистических методов.

Статистические таблицы – наиболее распространенная форма группировки экспериментальных данных.

Простые таблицы – таблицы, применяемые при альтернативной группировке, когда одна группа испытуемых противопоставляется другой. Их рекомендуется использовать при измерении изучаемых признаков в номинативной или ранговой шкале.

Сложные таблицы – многопольные таблицы, которые могут использоваться при выяснении причинно-следственных отношений между варьирующими признаками. Такие таблицы, как правило, имеют сложное строение, позволяющее одновременно осуществлять разные варианты группировки данных.

Статистические ряды – числовые значения признака, расположенные в определенном порядке.

Вариационный ряд распределения – двойной ряд чисел, показывающий, каким образом числовые значения признака связаны с их повторяемостью в данной выборке.

Частота вариационного ряда – число, показывающее, сколько раз отдельные варианты встречаются в данной совокупности.

Объем выборки – общая сумма частот вариационного ряда.

Кумуляты частот вариационного ряда – числа, которые показывают места измерений (вариантов) в ранжированном вариационном ряду.

Ряд распределения – распределение частот по вариантам. Измеренные величины признака в выборке варьируют в пределах от минимального до максимального значения. Этот предел разбивают на классовые интервалы, которые, в зависимости от конкретных данных, могут быть как равными по величине, так и неравными.

Гистограмма распределения частот – графическое изображение данного частотного распределения, т. е. столбчатая диаграмма, в которой по оси абсцисс OX откладывают величины классовых интервалов, а по оси ординат OY – величины частот, попадающих в данный классовый интервал; при этом над каждым классовым интервалом строится колонка или прямоугольник, площадь которого оказывается пропорциональной соответствующей частоте.

Числовые характеристики распределений – характеристики (меры) используемые для экспериментальных данных, полученных по выборке. Основные: мода; медиана; среднее арифметическое; разброс выборки; дисперсия; степень свободы; понятие нормального распределения.

Мода – числовое значение, которое встречается в выборке наиболее часто; обозначение – $\hat{X}$.

Медиана – числовое значение, которое делит упорядоченное множество данных пополам; обозначение – $\tilde{X}$ или M_d.

Среднее арифметическое – сумма всех значений измеренного признака, деленная на количество суммированных значений; обозначение – $\bar{X}$ или M_x.

Разброс выборки – разность между максимальной и минимальной величинами вариационного ряда; обозначение – R.

Дисперсия – среднее арифметическое квадратов отклонений значений переменной от ее среднего значения; обозначение – D.

Степень свободы – число свободно варьирующих единиц в составе выборки.

Нормальное распределение – распределение вероятностей, у которого форма и положение графика имеют вид колоколообразной кривой;определяется средней (μ) и стандартным отклонением (σ или S), равным квадратному корню из дисперсии; величины моды, медианы и среднего арифметического совпадают, то есть $\hat{X} = \tilde{X} = \bar{X}$.

Стандартное отклонение – величина, равная квадратному корню из дисперсии.

Гипотеза исследования – некоторое предположение, требующее проверки с привлечением фактов; формулируется в отношении связи явлений или свойств в некоторой совокупности объектов.

Статистическая гипотеза – любое предположение о свойствах случайных величин или событий; научная гипотеза, допускающая статистическую проверку.

Нулевая гипотеза – обозначение – H_0; гипотеза о сходстве, ее принятие свидетельствует об отсутствии различий.

Альтернативная гипотеза – обозначение – H_1, гипотеза о различии, ее принятие свидетельствует о наличии различий.

Сущность проверки статистической гипотезы – установление, согласуются ли экспериментальные данные и выдвинутая гипотеза; допустимо ли отмести расхождение между гипотезой и результатом статистического анализа экспериментальных данных за счет случайных причин.

Генеральная совокупность – все множество объектов, в отношении которого формулируется исследовательская гипотеза.

Выборка – ограниченная по численности группа объектов (испытуемых), специально отбираемая из генеральной совокупности для изучения ее свойств.

Критерии обоснованности выводов исследования – репрезентативность выборки и статическая достоверность эмпирических результатов исследования.

Репрезентативность выборки – способность выборки представлять изучаемые явления достаточно полно – с точки зрения их изменчивости в генеральной совокупности.

Статическая достоверность эмпирических результатов исследования определяется с помощью методов статистического вывода, которые предъявляют определенные требования к объему выборки.

Полное исследование – исследование, в котором измерению, наблюдению подвергаются все представители изучаемой генеральной совокупности.

Выборочное исследование – исследование, в котором производится выбор ограниченного числа элементов из изучаемой генеральной совокупности.

Независимая (несвязная) выборка – выборка, в которой процедура эксперимента и полученные результаты измерения некоторого свойства у испытуемых одной выборки не оказывают влияния на особенности протекания этого же эксперимента и результаты измерения этого же свойства у испытуемых другой выборки, т. е. вероятность отбора любого испытуемого одной выборки не зависит от отбора любого из испытуемых другой выборки.

Зависимая (связная) выборка – выборка, в которой процедура эксперимента и полученные результаты измерения некоторого свойства у испытуемых одной выборки оказывают влияние на другую выборку, т. е. каждому испытуемому одной выборки поставлен в соответствие по определенному критерию испытуемый из другой выборки.

Уровень статистической значимости – вероятность ошибочного отклонения нулевой гипотезы или, иными словами, вероятность ошибки первого рода при принятии решения. Для обозначения этой вероятности чаще всего употребляют букву *P*. Стандартными уровнями статистической значимости являются величины 0,05 (низший уровень, допускается пять ошибок в выборке из ста элементов), 0,01 (достаточный уровень) и 0,001 (высший уровень).

Эмпирическое значение по выбранному статистическому методу – эмпирическая статистика или числовое значение (обозначение в общем виде – $Ч_{эмп}$), которое подсчитывается исследователем по выбранному им экспериментальному методу на основании полученных экспериментальных данных. Затем оно сравнивается с двумя критическими величинами, которые соответствуют уровням значимости в 5 % ($Ч_{кр1}$) и в 1 % ($Ч_{кр2}$) для выбранного статистического метода. Критические значения находятся по статистическим таблицам.

Критические величины, соответствующие определенному уровню значимости для выбранного статистического метода – числовые значения, которые находятся по статистическим таблицам в соответствии с выбранным исследователем экспериментальным методом на основании полученных экспериментальных данных. Они соответствуют уровням значимости в 5 % (обозначение в общем виде – $Ч_{кр1}$) и в 1 % ($Ч_{кр2}$) для выбранного статистического метода.

Ось значимости – прямая (ось абсцисс *OX* декартовой системы координат), имеющая три выделенных участка, «зоны»: левая зона – зона незначимости, правая – зона значимости, промежуточная зона – зона неопределенности. Границами всех трех зон являются $Ч_{кр1}$ для $P = 0{,}05$ и $Ч_{кр2}$ для $P = 0{,}01$.

Выбор статистической гипотезы в зависимости от расположения эмпирического значения по выбранному статистическому методу на оси значимости. Выделяют четыре случая расположения $Ч_{эмп}$ на оси значимости:

1) $Ч_{эмп}$ попало в зону незначимости. В этом случае принимается гипотеза H_0 об отсутствии различий;

2) $Ч_{эмп}$ попало в зону значимости. В этом случае принимается гипотеза H_1 о наличии различий, гипотеза H_0 отклоняется;

3) $Ч_{эмп}$ попало в зону неопределенности. В этом случае исследователь в зависимости от важности решаемой задачи либо считает полученную статистическую оценку:
а) достоверной на уровне 5 % и принимает тем самым гипотезу H_1;
б) недостоверной на уровне 1 %, принимает тем самым гипотезу H_0;

4) величина $Ч_{эмп}$ точно совпадает либо с $Ч_{кр1}$, либо с $Ч_{кр2}$. Так, если $Ч_{эмп}$ точно совпадает:
– с $Ч_{кр1}$, то можно считать, что оценка достоверна точно на уровне в 5 % и принять гипотезу H_1 или, напротив, – гипотезу H_0;
– $Ч_{кр2}$, то, как правило, принимается альтернативная гипотеза H_1 о наличии различий, а гипотеза H_0 отклоняется.

Ошибки первого и второго рода при принятии решения об отклонении или принятии нулевой гипотезы. Возможны следующие четыре случая:
1) H_0 верна, но мы отклоняем H_0, т. е. принимаем H_1;
2) H_0 верна, и мы принимаем H_0, т. е. отклоняем H_1;
3) H_0 неверна, но мы принимаем H_0, т. е. отклоняем H_1;
4) H_0 неверна, и мы отклоняем H_0, т. е. принимаем H_1;

Случаи 2 и 4 отвечают верному решению, а случаи 1 и 3 – ошибочному. Говорят, что в случае 1 мы допускаем ошибку первого рода, а в случае 3 – ошибку второго рода. Определено, что вероятность ошибки первого рода всегда не превосходит принятого уровня значимости и обеспечивает заранее малую вероятность ошибочных решений при отклонении нулевой гипотезы.

Параметрический критерий – критерий различия, основанный на конкретном типе распределения генеральной совокупности (как правило, нормальном) или использующий параметры этой совокупности (средние, дисперсии и т. д.). На практике широко используются критерии, основанные на *t*-распределении Стьюдента и *F*-распределении Фишера – Снедекора.

Непараметрический критерий – критерий различия, не базирующийся на предположении о типе распределения генеральной совокупности и не использующий параметры этой совокупности. Для него используется термин «критерий, свободный от распределения». Он выявляет значимые различия и в том случае, если распределение близко к нормальному; при вычислениях вручную он чаще всего является значительно менее трудоемкими, чем параметрический.

Корреляционный анализ – количественный метод определения тесноты и направления взаимосвязи между выборочными переменными величинами. Целью его является оценка тесноты связи между признаками. Теснота связи количественно выражается величиной коэффициентов корреляции.

Парная корреляция – связь между двумя признаками (результативным и факторным или двумя факторными).

Частная корреляция – зависимость между результативным и одним факторным признаками при фиксированном значении других факторных признаков.

Множественная корреляция – зависимость результативного и двух или более факторных признаков, включенных в исследование.

Корреляционная связь – согласованное изменение двух признаков, отражающее тот факт, что изменчивость одного признака находится в соответствии с изменчивостью другого.

Виды корреляционных связей между измеренными признаками. Корреляция бывает линейной и нелинейной, положительной и отрицательной. Кроме того, по форме корреляционная связь может быть прямолинейной или криволинейной; по направлению – положительной («прямой») и отрицательной («обратной»); по силе корреляционная связь чаще всего определяется шкалой Чеддока.

Линейная корреляция: изменение одной переменной на одну единицу всегда приводит к изменению другой переменной на одну и ту же величину, т. е. с увеличением или уменьшением одной переменной X вторая переменная Y в среднем либо также растет, либо убывает.

Нелинейная корреляция: при увеличении одной величины характер изменения второй не линеен, а описывается другими законами.

Положительная корреляция: с увеличением переменной X переменная Y в среднем также увеличивается.

Отрицательная корреляция: с увеличением X переменная Y имеет в среднем тенденцию к уменьшению.

Коэффициент корреляции – количественная мера силы и направления вероятностной взамосвязи двух переменных; принимает значение в диапазоне от –1 до +1.

Графики корреляционных зависимостей строят по уравнениям функций: $\bar{Y}_x = F(X)$ или $\bar{X}_y = F(Y)$, которые называются уравнениями регрессии, где $\bar{Y}_x$ и $\bar{X}_y$ – так называемые условные средние арифметические переменных X и Y.

Регрессионный анализ – количественный метод определения вида математической функции в причинно-следственной зависимости между переменными величинами. Целью его является установление формы зависимости.

Регрессия – зависимость между определёнными переменными, с помощью которой можно спрогнозировать будущее поведение данных переменных. Причём под переменными подразумеваются всевозможные периодические явления вплоть до человеческого поведения.

Линия регрессии – прямая, построенная методом наименьших квадратов.

Уравнение регрессии – уравнение вида $Y = a_0 + a_1 \cdot X$, в котором Y – зависимая переменная; X – независимая переменная; a_0 – свободный член; a_1 – коэффициент регрессии или коэффициент, определяющий наклон линии регрессии по отношению к осям координат.

Расчеты коэффициентов регрессии – метод, заключающийся в решении системы уравнений

$$\begin{cases} a_0 \cdot N + a_1 \cdot \sum x_i = \sum y_i\,, \\ a_0 \cdot \sum x_i + a_1 \cdot \sum (x_i \cdot x_i) = \sum (y_i \cdot x_i). \end{cases}$$

В этой системе уравнений используются следующие обозначения: $Y = a_0 + a_1 \cdot X$ – уравнение регрессии; N – число элементов в переменной X или переменной Y; Σx_i – сумма всех элементов переменной X; Σy_i – сумма всех элементов переменной Y; $\Sigma(x_i \cdot x_i)$ – произведение всех элементов переменной X друг на друга; $\Sigma(y_i \cdot x_i)$ – попарное произведение всех элементов переменной X на соответствующие элементы переменной Y.

Корреляция Пирсона – мера линейной связи между двумя переменными, которая позволяет определить, насколько пропорциональна изменчивость двух переменных. Если переменные пропорциональны друг другу, то графически связь между ними можно представить в виде прямой линии с положительным (прямая пропорция) или отрицательным (обратная пропорция) наклоном.

Формула коэффициента корреляции при линейной зависимости

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

где x_i – значения, принимаемые переменной X; $\boldsymbol{y_i}$ – значения, принимаемые переменной Y; $\bar{x}$ – средняя по X; $\bar{y}$ – средняя по Y.

Формула коэффициента четырехклеточной сопряженности Пирсона или коэффициента корреляции φ

$$\varphi = \frac{a \cdot d - b \cdot c}{\sqrt{(a+b) \cdot (c+d) \cdot (a+c) \cdot (b+d)}},$$

где *a*, *b, c, d* соответствуют обозначениям в четырехпольной таблице:

0		Признак *X*		Итог
		1		
Признак *Y*	0	*a*	*b*	*a* + *b*
	1	*c*	*d*	*c* + *d*
Итог		*a* + *c*	*b* + *d*	*N*

Поиск критических значений для коэффициента корреляции φ осуществляется с помощью *t*-критерия Стьюдента по формуле

$$T_{\varphi} = |r_{\text{эмп}}| \cdot \sqrt{\frac{n-2}{1 - {r_{\text{эмп}}}^2}},$$

где $r_{\text{эмп}}$ – коэффициент корреляции; n – число коррелируемых признаков; величина T_{φ} проверяется на уровень значимости по таблице для *t*-критерия Стьюдента. Число степеней свободы в этом случае будет равно k, где $k = n - 2$.

Коэффициент детерминации – квадрат коэффициента корреляции зависимой и независимой переменных, представляющий долю дисперсии зависимой переменной, обусловленной влиянием независимой переменной; обозначение – ${r_{xy}}^2$. Показывает, в какой степени изменчивость одной переменной обусловлена (детерминирована) влиянием другой переменной; отражает связь между двумя переменными линейно.

Градации величин корреляции по силе связи – по А.Д. Наследову: 1) $r \leq 0{,}3$ – слабая связь; 2) $0{,}3 < r \leq 0{,}7$ – умеренная связь; 3) $r > 0{,}7$ – сильная связь. По шкале Чеддока: 0,1–0,3 – слабая; 0,3–0,5 – умеренная; 0,5–0,7 – заметная; 0,7–0,9 – высокая; 0,9–1 – весьма высокая.

Приложение

СТАТИСТИЧЕСКИЕ ТАБЛИЦЫ

Таблица 1

Критические значения критерия знаков *G* для уровней статистической значимости $p \leq 0{,}05$ и $p \leq 0{,}01$

n	*p*		*n*	*p*		*n*	*p*		*n*	*p*	
	0,05	0,01		0,05	0,01		0,05	0,01		0,05	0,01
5	0	-	**27**	8	7	**49**	18	15	**92**	37	34
6	0	0	**28**	8	7	**50**	18	16	**94**	38	35
7	0	0	**29**	9	7	**52**	19	17	**96**	39	36
8	1	0	**30**	10	8	**54**	20	18	**98**	40	37
9	1	0	**31**	10	8	**56**	21	18	**100**	41	37
10	1	0	**32**	10	8	**58**	22	19	**110**	45	42
11	2	1	**33**	11	9	**60**	23	20	**120**	50	46
12	2	1	**34**	11	9	**62**	24	21	**130**	55	51
13	3	1	**35**	12	10	**64**	24	22	**140**	59	55
14	3	2	**36**	12	10	**66**	25	23	**150**	64	60
15	3	2	**37**	13	16	**68**	26	23	**160**	69	64
16	4	2	**38**	13	11	**70**	27	24	**170**	73	69
17	4	3	**39**	13	1 1	**72**	28	25	**180**	78	73
18	5	3	**40**	14	12	**74**	29	26	**190**	83	78
19	5	4	**41**	14	12	**76**	30	27	**200**	87	83
20	5	4	**42**	15	13	**78**	31	28	**220**	97	92
21	6	4	**43**	15	13	**80**	32	29	**240**	106	101
22	6	5	**44**	16	13	**82**	33	30	**260**	116	110
23	7	5	**45**	16	14	**84**	33	30	**280**	125	120
24	7	5	**46**	16	14	**86**	34	31	**300**	135	129
25	7	6	**47**	17	15	**88**	35	32			
26	8	6	**48**	17	15	**90**	36	33			

Таблица 2

Критические значения ***T***-критерия Вилкоксона
для уровней статистической значимости $\boldsymbol{p \leq 0{,}05}$ и $p \leq \mathbf{0{,}01}$

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
18	47	32	**41**	302	252
19	53	37	**42**	319	266
20	60	43	**43**	336	281
21	67	49	**44**	353	296
22	75	55	**45**	371	312
23	83	62	**46**	389	328
24	92	69	**47**	407	345
25	100	76	**48**	426	362
26	110	84	**49**	446	379
27	119	92	**50**	466	397

Таблица 3

Таблица вероятностей *P* биномиального распределения при ***p* = *q* = 0,5***

m / *n*	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
5	031	188	500	812	969	+ **										
6	016	109	344	656	891	984	+									
7	008	062	226	500	773	938	992	+								
8	004	035	145	363	637	855	965	996	+							
9	002	020	090	254	500	746	910	980	998	+						
10	001	011	055	172	377	623	828	945	989	998	+					
11		006	033	113	274	500	726	887	967	994	+	+				
12		003	019	073	194	387	613	806	927	981	997	+	+			
13		002	011	046	133	291	500	709	867	954	989	998	+	+		
14		001	006	029	090	212	395	605	788	910	971	994	999	+	+	+
15			004	018	059	151	304	500	696	849	941	982	996	+	+	+
16			002	011	038	105	227	402	598	773	896	962	989	998	+	+
17			001	006	025	072	166	315	500	685	834	928	975	994	999	+
18			001	004	015	048	119	240	407	593	760	881	952	985	996	999
19				002	010	032	084	180	324	500	676	820	916	968	990	998
20				001	006	021	058	132	252	412	588	748	868	942	979	994
21				001	004	013	039	095	192	332	500	668	808	905	961	987
22					002	008	026	067	143	262	416	584	738	857	933	974
23					001	005	017	047	105	202	339	500	661	798	895	953
24					001	003	011	032	076	154	271	419	581	729	846	924
25						002	007	022	054	115	212	345	500	655	788	885

* Знаком + в таблице обозначены значения близкие к 1.
** В таблице все величины даны без начального нуля и последующей запятой, так что, если в таблице дано число, например 013, то это число следует читать как 0,013.

Таблица 4

Критические значения ***U***-критерия Вилкоксона-Манна-Уитни для уровней статистической значимости $p \le 0{,}05$ и $p \le 0{,}01$

n_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
n_2	$p = 0{,}05$																		
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	26	30	34	38	42								
13	2	6	10	15	19	24	28	33	37	42	47	51							
14	3	7	11	16	21	26	31	36	41	46	51	56	61						
15	3	7	12	18	23	28	33	39	44	50	55	61	66	72					
16	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83				
17	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96			
18	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109		
19	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	
20	4	11	8	25	32	39	47	54	62	69	77	84	92	100	107	115	123	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	14	17	21	24	28	31								
13	0	2	5	9	12	16	20	23	27	31	35	39							
14	0	2	6	10	13	17	22	26	30	34	38	43	47						
15	0	3	7	11	15	19	24	28	33	37	42	47	51	56					
16	0	3	7	12	16	21	26	31	36	41	46	51	56	61	66				
17	0	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77			
18	0	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88		
19	1	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	
20	1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

n_1	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
$p = 0{,}05$																		
21	19	26	34	41	49	57	65	73	81	89	97	105	113	121	130	138	146	154
22	20	28	36	44	52	60	69	77	85	94	102	111	119	128	136	145	154	162
23	21	29	37	46	55	63	72	81	90	99	107	116	125	134	143	152	161	170
24	22	31	39	48	57	66	75	85	94	103	113	122	131	141	150	160	169	179
25	23	32	41	50	60	69	79	89	98	108	118	128	137	147	157	167	177	187
26	24	33	43	53	62	72	82	93	103	113	123	133	143	154	164	174	185	195
27	25	35	45	55	65	75	86	96	107	118	128	139	150	160	171	182	193	203
28	26	36	47	57	68	79	89	100	111	122	133	144	156	167	178	189	200	212
29	27	38	48	59	70	82	93	104	116	127	139	150	162	173	185	196	208	220
30	28	39	50	62	73	85	96	108	120	132	144	156	168	180	192	204	216	228
31	29	41	52	64	76	88	100	112	124	137	149	161	174	186	199	211	224	236
32	30	42	54	66	78	91	103	116	129	141	154	167	180	193	206	219	232	245
33	31	43	56	68	81	94	107	120	133	146	159	173	186	199	213	226	239	253
34	32	45	58	71	84	97	110	124	137	151	164	178	192	206	219	233	247	261
35	33	46	59	73	86	100	114	128	142	156	170	184	198	212	226	241	255	269
36	35	48	61	75	89	103	117	132	146	160	175	189	204	219	233	248	263	278
37	36	49	63	77	92	106	121	135	150	165	180	195	210	225	240	255	271	286
38	37	51	65	79	94	109	124	139	155	170	185	201	216	232	247	263	278	294
39	38	52	67	82	97	112	128	143	159	175	190	206	222	238	254	270	286	302
40	39	53	69	84	100	115	131	147	163	179	196	212	228	245	261	278	294	311
$p = 0{,}01$																		
21	10	16	22	29	35	42	49	56	63	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	11	18	25	32	39	47	55	62	70	78	86	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	60	69	77	86	95	103	112	121	130	138	147	156
26	13	21	29	37	46	54	63	72	81	90	99	108	117	126	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	98	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
31	16	26	36	46	56	67	77	88	99	110	121	132	143	155	166	177	188	200
32	17	27	37	47	58	69	80	91	103	114	126	137	149	160	172	184	195	207
33	17	28	38	49	60	72	83	95	106	118	130	142	154	166	178	190	202	214
34	18	29	40	51	62	74	86	98	110	122	134	147	159	172	184	197	209	222
35	19	30	41	53	64	77	89	101	114	126	139	152	164	177	190	203	216	229
36	19	31	42	54	67	79	92	104	117	130	143	156	170	183	196	210	223	236
37	20	32	44	56	69	81	95	108	121	134	148	161	175	189	202	216	230	244
38	21	33	45	58	71	84	97	111	125	138	152	166	180	194	208	223	237	251
39	21	34	46	59	73	86	100	114	128	142	157	171	185	200	214	229	244	258
40	22	35	48	61	75	89	103	117	132	146	161	176	191	206	221	236	251	266

n_1	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
n_2	$p = 0,05$																		
21																			
22	171																		
23	180	189																	
24	188	198	207																
25	197	207	217	227															
26	206	216	226	237	247														
27	214	225	236	247	258	268													
28	223	234	245	257	268	279	291												
29	232	243	255	267	278	290	302	314											
30	240	252	265	277	289	301	313	326	338										
31	249	261	274	287	299	312	325	337	350	363									
32	258	271	284	297	310	323	336	349	362	375	389								
33	266	280	293	307	320	334	347	361	374	388	402	415							
34	275	289	303	317	331	345	359	373	387	401	415	429	113						
35	284	298	312	327	341	356	370	385	399	413	428	442	457	471					
36	292	307	322	337	352	367	381	396	411	426	441	456	471	486	501				
37	301	316	332	347	362	378	393	408	424	439	454	470	485	501	516	531			
38	310	325	341	357	373	388	404	420	436	452	467	483	499	515	531	547	563		
39	318	335	351	367	383	399	416	432	448	464	481	497	513	530	546	562	579	595	
40	327	344	360	377	394	410	427	444	460	477	494	511	527	544	561	578	594	611	628
$p = 0,01$																			
21																			
22	142																		
23	150	158																	
24	154	166	174																
25	165	174	183	192															
26	173	182	191	201	210														
27	180	190	200	209	219	229													
28	188	198	208	218	229	239	249												
29	196	206	217	227	238	249	259	270											
30	203	214	225	236	247	258	270	281	292										
31	211	223	234	245	257	268	280	291	303	314									
32	219	231	242	254	266	278	290	302	314	326	338								
33	227	239	251	263	276	288	300	313	325	337	350	362							
34	234	247	260	272	285	298	311	323	336	349	362	375	387						
35	242	255	268	281	294	308	321	334	347	360	374	387	400	413					
36	250	263	277	290	304	318	331	345	358	372	386	399	413	427	440				
37	258	271	285	299	313	327	341	355	370	384	398	412	426	440	454	468			
38	265	280	294	308	323	337	352	366	381	395	410	424	439	453	468	482	497		
39	273	288	303	317	332	347	362	377	392	407	422	437	452	467	482	497	512	527	
40	281	296	311	326	342	357	372	388	403	418	434	449	465	480	495	511	526	542	557

n_1	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
n_2	$p = 0{,}05$																	
41	40	55	70	86	102	118	135	151	168	184	201	218	234	251	268	285	302	319
42	41	56	72	88	105	121	138	155	172	189	206	223	240	258	275	292	310	327
43	42	58	74	91	107	124	142	159	176	194	211	229	247	264	282	300	318	335
44	43	59	76	93	110	128	145	163	181	199	216	235	253	271	289	307	325	344
45	44	61	78	95	113	131	149	167	185	203	222	240	259	277	296	315	333	352
46	45	62	80	97	115	134	152	171	189	208	227	246	265	284	303	322	341	360
47	46	64	81	100	118	137	156	175	194	213	232	251	271	290	310	329	349	369
48	47	65	83	102	121	140	159	178	198	218	237	257	277	297	317	337	357	377
49	48	66	85	104	123	143	163	182	202	222	243	263	283	303	324	344	365	385
50	49	68	87	106	126	146	166	186	207	227	248	268	289	310	331	352	372	393
51	50	69	89	109	129	149	170	190	211	232	253	274	295	316	338	359	380	402
52	51	71	91	111	131	152	173	194	215	237	258	280	301	323	345	366	388	410
53	52	72	92	113	134	155	177	198	220	241	263	285	307	329	352	374	396	418
54	53	74	94	115	137	158	180	202	224	246	269	291	313	336	359	381	404	427
55	54	75	96	118	139	161	184	206	228	251	274	297	319	342	365	389	412	435
56	55	76	98	120	142	164	187	210	233	256	279	302	326	349	372	396	420	443
57	57	78	100	122	145	167	191	214	237	261	284	308	332	355	379	403	427	451
58	58	79	102	124	147	171	194	218	241	265	289	314	338	362	386	411	435	460
59	59	81	103	127	150	174	198	222	246	270	295	319	344	369	393	418	443	468
60	60	82	105	129	153	177	201	225	250	275	300	325	350	375	400	426	451	476
	$p = 0{,}01$																	
41	23	36	49	63	77	91	106	121	136	151	166	181	196	211	227	242	258	273
42	23	37	50	65	79	94	109	124	139	155	170	186	201	217	233	249	265	280
43	24	38	52	66	81	96	112	127	143	159	175	190	207	223	239	255	271	288
44	25	39	53	68	83	99	115	130	146	163	179	195	212	228	245	262	278	295
45	25	40	54	70	85	101	117	134	150	167	183	200	217	234	251	268	285	303
46	26	41	56	71	87	104	120	137	154	171	188	205	222	240	257	275	292	310
47	27	42	57	73	90	106	123	140	157	175	192	210	228	245	263	281	299	317
48	27	43	58	75	92	109	126	143	161	179	197	215	233	251	269	288	306	325
49	28	44	60	77	94	111	129	147	165	183	201	220	238	257	276	294	313	332
50	29	45	61	78	96	114	132	150	168	187	206	225	244	263	282	301	320	339
51	29	46	63	80	98	116	135	153	172	191	210	229	249	268	288	307	327	347
52	30	47	64	82	100	119	137	157	176	195	215	234	254	274	294	314	334	354
53	31	48	65	83	102	121	140	160	179	199	219	239	259	280	300	320	341	361
54	31	49	67	85	104	114	143	163	183	203	224	244	265	285	306	327	348	369
55	32	50	68	87	106	126	146	166	187	207	228	249	270	291	312	333	355	376
56	33	51	69	89	108	129	149	177	190	211	233	254	275	297	318	340	362	384
57	33	52	71	90	111	131	152	173	194	215	237	259	281	302	324	347	369	391
58	34	53	72	92	113	133	155	176	198	220	242	264	286	308	331	353	376	398
59	34	54	73	94	115	136	158	179	201	224	246	268	291	314	337	360	383	406
60	35	55	75	96	117	138	160	183	205	228	250	273	296	320	343	366	390	413

n_1	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
n_2	p = 0,05																		
41	336	353	370	387	404	421	438	456	473	490	507	524	541	559	576	593	610	628	645
42	345	362	380	397	415	432	450	467	485	503	520	538	556	573	591	609	626	644	662
43	353	371	389	407	425	443	461	479	497	515	533	552	570	588	606	624	642	660	679
44	362	380	399	417	436	454	473	491	510	528	547	565	584	602	621	640	658	677	695
45	371	390	408	427	446	465	484	503	522	541	560	579	598	617	636	655	674	693	712
46	380	399	418	437	457	476	495	515	534	554	573	593	612	631	651	670	690	709	729
47	388	408	428	447	467	487	507	527	547	566	586	606	626	646	666	686	706	726	746
48	397	417	437	458	478	498	518	539	559	579	600	620	640	661	681	701	722	742	763
49	406	426	447	468	488	509	530	550	571	592	613	634	654	675	696	717	738	759	780
50	414	435	457	478	499	520	541	562	583	605	626	647	669	690	711	732	754	775	796
51	423	445	466	488	509	531	553	574	596	618	639	661	683	704	726	748	770	791	813
52	432	454	476	408	520	542	564	586	608	630	652	675	697	719	741	763	786	808	830
53	441	463	485	508	530	553	575	598	620	643	666	688	711	734	756	779	802	824	847
54	449	472	495	518	541	564	587	610	633	656	679	702	725	748	771	794	818	841	864
55	458	481	505	528	551	575	598	622	645	669	692	716	739	763	786	810	834	857	881
56	467	491	514	538	562	586	610	634	657	681	705	729	753	777	801	825	850	874	898
57	476	500	524	548	572	597	621	645	670	694	719	743	768	792	816	841	865	890	915
58	484	509	534	558	583	608	633	657	682	707	732	757	782	807	832	856	881	906	931
59	493	518	543	568	594	619	644	669	694	720	745	770	796	821	847	872	897	923	948
60	502	527	553	578	604	630	655	681	707	733	758	784	810	836	862	888	913	939	965
n_2	p = 0,01																		
41	289	304	320	336	351	367	383	398	414	430	446	462	477	493	509	525	541	557	573
42	296	312	328	345	361	377	393	409	425	442	458	474	490	507	523	539	556	572	588
43	304	321	337	354	370	387	403	420	437	453	470	487	503	520	537	553	570	587	604
44	312	329	346	363	380	397	414	431	448	465	482	499	516	533	550	568	585	602	619
45	320	337	354	372	389	407	424	441	459	476	494	511	529	547	564	582	599	617	635
46	328	345	363	381	399	416	434	452	470	488	506	524	542	560	578	596	614	632	650
47	335	353	372	390	408	426	445	463	481	500	518	536	555	573	592	610	629	647	666
48	343	362	380	399	418	436	455	474	492	511	530	549	568	587	606	625	643	662	681
49	351	370	389	408	427	446	465	484	504	523	542	561	581	600	619	639	658	678	697
50	359	378	398	417	437	456	476	495	515	535	554	574	594	613	633	653	673	693	713
51	366	386	406	426	446	466	486	506	526	546	566	587	607	627	647	667	688	708	728
52	374	395	415	435	456	476	496	517	537	558	578	599	620	640	661	682	702	723	744
53	382	403	423	444	465	486	507	528	549	570	591	612	633	654	675	696	717	738	759
54	390	411	432	453	475	496	517	538	560	581	603	624	646	667	689	710	732	753	775
55	398	419	441	462	484	506	527	549	571	593	615	637	659	680	702	724	746	768	790
56	405	427	449	471	494	516	538	560	582	605	627	649	671	694	716	738	761	784	806
57	413	436	458	581	503	526	548	571	593	616	639	662	684	707	730	753	776	799	822
58	421	444	467	490	513	536	559	582	605	628	651	674	697	721	744	767	790	814	837
59	429	452	475	499	522	545	569	592	616	640	663	687	710	734	758	781	805	829	853
60	437	460	484	508	532	555	579	603	627	651	675	699	723	747	772	796	820	844	868

n_1	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
n_2				p = 0,05																
41	662																			
42	679	697																		
43	697	715	732																	
44	714	733	751	770																
45	731	750	762	789	808															
46	749	768	788	807	827	846														
47	766	786	806	826	846	866	886													
48	783	804	824	845	865	886	906	927												
49	800	821	842	863	884	905	926	947	968											
50	818	839	861	882	903	925	946	968	989	1010										
51	835	857	879	901	922	944	966	988	1010	1032	1054									
52	852	875	897	919	942	964	986	1009	1031	1053	1076	1098								
53	870	893	916	938	961	934	1006	1029	1052	1075	1098	1120	1143							
54	887	910	934	957	980	1003	1026	1050	1073	1096	1119	1143	1166	1189						
55	904	928	952	975	999	1023	1046	1070	1094	1113	1141	1165	1189	1213	1236					
56	922	946	970	994	1018	1042	1067	1091	1115	1139	1163	1187	1212	1236	1260	1284				
57	939	964	988	1013	1037	1062	1087	1111	1136	1161	1185	1210	1235	1259	1284	1309	1333			
58	956	981	1007	1032	1057	1082	1107	1132	1157	1182	1207	1232	1257	1282	1308	1333	1358	1383		
59	974	999	1025	1050	1076	1101	1127	1152	1178	1204	1229	1255	1280	1306	1331	1357	1383	1408	1434	
60	991	1017	1043	1069	1095	1121	1147	1173	1199	1225	1251	1277	1303	1329	1355	1381	1407	1433	1460	1486
n_2		p = 0,01																		
41	589																			
42	605	621																		
43	621	637	654																	
44	636	654	671	688																
45	652	670	688	706	723															
46	668	687	705	723	741	759														
47	684	703	722	740	759	777	796													
48	700	719	738	757	776	795	814	834												
49	716	736	755	775	794	814	835	853	872											
50	732	752	772	792	812	832	852	872	892	912										
51	748	769	789	809	830	850	870	891	911	932	952									
52	764	785	806	827	847	868	889	910	931	951	972	993								
53	780	802	823	844	865	886	908	929	950	971	993	1014	1035							
54	796	818	840	861	883	905	926	948	970	991	1013	1035	1057	1078						
55	812	834	857	879	901	923	945	967	989	1011	1034	1056	1078	1100	1122					
56	828	851	873	896	919	941	964	986	1009	1031	1054	1077	1099	1122	1145	1167				
57	844	867	890	913	936	959	982	1005	1028	1051	1074	1098	1121	1141	1167	1191	1213			
58	861	884	907	931	954	978	1001	1024	1048	1071	1095	1118	1142	1165	1189	1213	1236	1260		
59	877	900	924	948	972	996	1020	1044	1068	1091	1115	1139	1163	1187	1211	1235	1254	1283	1307	
60	893	917	941	965	990	1014	1038	1063	1087	1111	1136	1160	1185	1209	1234	1258	1282	1307	1331	1356

Таблица 5

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

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

Таблица 6

Критические значения *t*-критерия Стьюдента
при различных уровнях значимости ***p***

Число сте-пеней свободы ν	*P*			Число сте-пеней свободы ν	*P*		
	0,05	0,01	0,001		0,05	0,01	0,001
1	12,71	63,66	64,60	18	2,10	2,88	3,92
2	4,30	9,92	31,60	19	2,09	2,86	3,88
3	3,18	5,84	12,92	20	2,09	2,85	3,85
4	2,78	4,60	8,61	21	2,08	2,83	3,82
5	2,57	4,03	6,87	22	2,07	2,82	3,79
6	2,45	3,71	5,96	23	2,07	2,81	3,77
7	2,37	3,50	5,41	24	2,06	2,80	3,75
8	2,31	3,36	5,04	25	2,06	2,79	3,73
9	2,26	3,25	4,78	26	2,06	2,78	3,71
10	2,23	3,17	4,59	27	2,05	2,77	3,69
11	2,20	3,11	4,44	28	2,05	2,76	3,67
12	2,18	3.05	4,32	29	2,05	2,76	3,66
13	2,16	3,01	4,22	30	2,04	2,75	3,65
14	2,14	2,98	4,14	40	2,02	2,70	3,55
15	2,13	2,95	4,07	60	2,00	2,66	3,46
16	2,12	2,92	4,02	120	1,98	2,62	3,37
17	2,11	2,90	3,97	∞	1,96	2,58	3,29
P	0,05	0,01	0,001	-	0,05	0,01	0,001

Таблица 7

Критические коэффициента корреляции r_{xy} Пирсона

$\nu = n - 2$	P		$\nu = n - 2$	P	
	0,05	0,01		0,05	0,01
5	0,75	0,87	27	0,37	0,47
6	0,71	0,83	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	35	0,33	0,42
10	0,58	0,71	40	0,30	0,39
11	0,55	0,68	45	0,29	0,37
12	0,53	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	0,56	125	0,17	0,23
19	0,43	0,55	150	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

УСЛОВНЫЕ ОБОЗНАЧЕНИЯ

x_i – числовые значения переменных, полученные в эксперименте, или измерения (варианты)

f_i – частоты вариационного ряда

$n = \sum_{i=1} f_i$ – общая сумма частот вариационного ряда или объем выборки

$n_i\% = \frac{f_i}{n} \cdot 100\ \%$ – процент каждой отдельной частоты вариационного ряда

σ_i – кумуляты частот вариационного ряда

$\hat{X}$ – мода

$\tilde{X}$ или M_d – медиана

$\bar{X}$ или M_x – среднее арифметическое

ν – число степеней свободы

R – разброс выборки или размах

D – дисперсия

S_x – стандартное отклонение выборки

μ – средняя для нормального распределения

σ – стандартное отклонение для нормального распределения

H_0 – нулевая гипотеза или гипотеза о сходстве

H_1 – альтернативная гипотеза или гипотеза о различии

P – уровень статистической значимости

$Ч_{эмп}$ – эмпирическое значение по выбранному статистическому методу

$Ч_{кр1}$ – критическая величина, соответствующая уровню значимости в 5 % для выбранного статистического метода

$Ч_{кр2}$ – критическая величина, соответствующая уровню значимости в 1 % для выбранного статистического метода

ЭК (ЭГ) – экспериментальные классы (группы)

КК (КГ) – контрольные классы (группы)