

Г.В. ЩАГИНА

**МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ДИСЦИПЛИНЕ
«МАТЕМАТИЧЕСКИЕ И СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ
НАУЧНЫХ ДАННЫХ»**

Направление подготовки: 44.02.04 Психолого-педагогическое образование
Профиль: Психолого-педагогическое образование

Уровень образования – магистратура

Форма обучения – заочная

Методические рекомендации

Челябинск

2021

МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ РФ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Южно-уральский государственный гуманитарно-педагогический университет»
(ФГБОУ ВО «ЮУрГГПУ»)
Профессионально-педагогический институт
Кафедра подготовки педагогов профессионального обучения и предметных методик

**МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ДИСЦИПЛИНЕ
«МАТЕМАТИЧЕСКИЕ И СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ
НАУЧНЫХ ДАННЫХ»**

Направление подготовки: 44.02.04 Психолого-педагогическое образование
Профиль: Психолого-педагогическое образование

Уровень образования – магистратура

Форма обучения – заочная

Методические рекомендации

Челябинск
2021

Шагина Г.В. Методические рекомендации по дисциплине «Математические и статистические методы обработки научных данных». Направление подготовки: 44.02.04 Психолого-педагогическое образование. Профиль: Психолого-педагогическое образование. [Текст] /Сост. Г.В. Шагина. – Челябинск, Изд-во 2021. – 28 с.

Применение методов математической статистики (статистических методов) для обработки результатов эмпирического исследования является одним из требований к выпускной квалификационной работе магистранта, обучающегося по направлению Психолого-педагогическое образование, профиль: Психология и педагогика образования личности. Курс «Математические и статистические методы обработки научных данных» призван подготовить магистрантов к применению математических методов при обработке эмпирических данных, полученных в результате исследовательской работы в области психологии и педагогики образования личности.

Методические рекомендации содержат малую часть математических и статистических методов, применяемых для обработки научных данных.

Для того, чтобы студентам подобрать правильный математический аппарат для обработки полученных научных данных в Приложении имеется таблица, содержащая краткую классификацию задач и методов их статистического решения.

АВТОР-СОСТАВИТЕЛЬ

Г.В. Шагина – кандидат педагогических наук, доцент кафедры подготовки педагогов профессионального обучения Профессионально-педагогического института ФГБОУ ВО «Южно-Уральский государственный гуманитарно-педагогический университет»

©Г.В. Шагина, 2021

Челябинск
2021

СОДЕРЖАНИЕ

1. Общее представление об обработке данных.....	5	
2. Методы статистической обработки данных.....	6	
2.1. Первичная статистическая обработка данных.....	6	
2.2. Вторичная статистическая обработка данных.....	11	
Список литературы.....		16
Приложение.....		18

*Нет хороших и плохих методов,
есть хорошее и плохое их применение.
Б. М. Теплов*

1. Общее представление об обработке данных

Обработка данных психолого-педагогических исследований направлена на решение следующих задач:

- упорядочивание полученного материала;
- обнаружение и ликвидация ошибок, недочетов, пробелов в сведениях;
- выявление скрытых от непосредственного восприятия тенденций, закономерностей и связей;
- обнаружение новых фактов, которые не ожидалось и не были замечены в ходе эмпирического процесса;
- выяснение уровня достоверности, надежности и точности собранных данных и получение на их базе научно обоснованных результатов.

Различают количественную и качественную обработку данных.

Количественная обработка – это работа с измеренными характеристиками изучаемого объекта.

Качественная обработка представляет собой способ проникновения в сущность объекта путем выявления его не измеряемых свойств.

В количественном исследовании доминирует аналитическая составляющая познания, что отражено и в названиях количественных методов обработки эмпирического материала: корреляционный анализ, факторный анализ и т. д. Реализуется количественная обработка с помощью математико-статистических методов.

В качественной обработке преобладают синтетические способы познания. Обобщение проводится на интерпретационном этапе исследовательского процесса. Обычно результатом качественной обработки является интегрированное представление о множестве свойств объекта или множестве объектов в форме классификаций и типологий. Качественная обработка в значительной мере апеллирует к методам логики.

Противопоставление друг другу качественной и количественной обработки довольно условно. Количественный анализ без последующей качественной обработки бессмыслен, так как сам по себе не приводит к приращению знаний, а качественное изучение объекта без базовых количественных данных в научном познании невозможно. Без количественных данных научное познание – чисто умозрительная процедура.

Единство количественной и качественной обработки наглядно представлено во многих методах обработки данных: факторном и таксономическом анализе, шкалировании, классификации и др. Наиболее распространены такие приемы количественной обработки, как классификация, типологизация, систематизация, периодизация, казуистика.

Качественная обработка естественным образом выливается в описание и объяснение изучаемых явлений, что составляет уже следующий уровень их изучения, осуществляемый на стадии интерпретации результатов. Количественная же обработка полностью относится к этапу обработки данных.

2. Методы статистической обработки данных

2.1. Первичная статистическая обработка данных

Методами статистической обработки результатов эксперимента называются математические приемы, формулы, способы количественных расчетов, с помощью которых показатели, получаемые в ходе эксперимента, можно обобщать, приводить в систему, выявляя скрытые в них закономерности.

Данные – это основные элементы, подлежащие классифицированию или разбитые на категории с целью обработки.

Некоторые из методов математико-статистического анализа позволяют вычислять так называемые элементарные математические статистики, характеризующие выборочное распределение данных, например:

- выборочное среднее,
- выборочная дисперсия,
- мода,
- медиана и ряд других.

Иные методы математической статистики позволяют судить о динамике изменения отдельных статистик выборки, например:

- дисперсионный анализ,
- регрессионный анализ.

С помощью третьей группы методов выборочных данных можно достоверно судить о статистических связях, существующих между переменными величинами, которые исследуют в данном эксперименте:

- корреляционного анализа;
- факторного анализа;
- методов сравнения.

Все методы количественной обработки принято подразделять на первичные и вторичные.

Первичная статистическая обработка нацелена на упорядочивание информации об объекте и предмете изучения. На этой стадии «сырые» сведения группируются по тем или иным критериям, заносятся в сводные таблицы.

Первичный статистический анализ всей совокупности полученных в исследовании данных дает возможность охарактеризовать ее в предельно сжатом виде и ответить на два главных вопроса: 1) какое значение наиболее характерно для выборки; 2) велик ли разброс данных относительно этого характерного значения, т. е. какова «размытость» данных.

Для решения первого вопроса вычисляются *меры центральной тенденции*, для решения второго – *меры изменчивости (или разброса)*. Эти статистические показатели используются в отношении количественных данных, представленных в порядковой, интервальной или пропорциональной шкале.

Меры центральной тенденции – это величины, вокруг которых группируются остальные данные. Данные величины являются как бы обобщающими всю выборку показателями, что, во-первых, позволяет судить по ним обо всей выборке, а во-вторых, дает возможность сравнивать разные выборки, разные серии между собой. К мерам центральной тенденции в обработке результатов психолого-педагогических исследований относятся: выборочное среднее, медиана, мода.

Выборочное среднее (M) (или среднеарифметическое) – это результат деления суммы всех значений (X) на их количество (N).

$$M = \frac{\sum X}{N}$$

Пример 1. В исследовании объема вербальной механической памяти, тест «10 слов» в группе из 12 испытуемых ($n = 12$), получены следующие результаты (количество запомненных слов): 5, 4, 5, 6, 7, 3, 6, 2, 8, 6, 9, 7. Найдите выборочное среднее.

Выборочное среднее вычисляется:

$$M = \frac{5+4+5+6+7+3+6+2+8+6+9+7}{12} = \frac{68}{12} = 5,6$$

Для данной выборки выборочное среднее $M = 5,6$.

Медиана (Me) – это значение, выше и ниже которого количество отличающихся значений одинаково, т.е. это центральное значение в последовательном ряду данных. Медиана не обязательно должна совпадать с конкретным значением. Совпадение происходит в случае нечетного числа значений (ответов), несовпадение – при четном их числе. В последнем случае

медиана вычисляется как среднее арифметическое двух центральных значений в упорядоченном ряду.

Пример 2. Найдем медиану выборки: 5, 4, 5, 6, 7, 3, 6, 2, 8, 6, 9, 7.

Упорядочим выборку: 2, 3, 4, 5, 5, 6, / 6, 6, 7, 7, 8, 9. Поскольку здесь имеется четное число элементов, то существует две "середины" - 6 и 6. В этом случае медиана определяется как среднее арифметическое этих значений.

$$\text{Me} = \frac{6+6}{2} = 6$$

Пример 3. Найдем медиану выборки с нечетным количеством значений: 9, 3, 5, 8, 4, 11, 13.

Сначала упорядочим выборку по величинам входящих в нее значений. Получим: 3, 4, 5, 8, 9, 11, 13. Поскольку в выборке семь элементов, четвертый по порядку элемент будет серединой ряда. Таким образом, медианой будет четвертый элемент - 8

Мода (Mo) – это значение, наиболее часто встречающееся в выборке, т. е. значение с наибольшей частотой. Если все значения в группе встречаются одинаково часто, то считается, что моды нет. Если два соседних значения имеют одинаковую частоту и больше частоты любого другого значения, мода есть среднее этих двух значений. Если то же самое относится к двум несмежным значениям, то существует две моды, а группа оценок является бимодальной.

Пример 4. В ряду значений 2, 3, 4, 5, 5, 6, 6, 6, 7, 7, 8, 9 модой является 6, потому что 6 встречается чаще любого другого числа.

Пример 5. В выборке 1, 2, 2, 2, 5, 5, 5, 6 частоты рядом расположенных значений 2 и 5 совпадают и равняются 3. Эта частота больше, чем частота других значений 1 и 6 (у которых она равна 1). Следовательно, модой этого

ряда будет величина $\frac{(2+5)}{2} = 3,5$

Обычно выборочное среднее применяется при стремлении к наибольшей точности в определении центральной тенденции. Медиана вычисляется в том случае, когда в серии есть «нетипичные» данные, резко влияющие на среднее. Мода используется в ситуациях, когда не нужна высокая точность, но важна быстрота определения меры центральной тенденции.

Вычисление всех трех показателей производится также для оценки распределения данных. При нормальном распределении значения выборочного среднего, медианы и моды одинаковы или очень близки.

Меры разброса (изменчивости) – это статистические показатели, характеризующие различия между отдельными значениями выборки. Они позволяют судить о степени однородности полученного множества, его компактности, а косвенно и о надежности полученных данных и вытекающих

из них результатов. Наиболее используемые в психолого-педагогических исследованиях показатели: среднее отклонение, дисперсия, стандартное отклонение.

Размах (Р) – это интервал между максимальным и минимальным значениями признака. Определяется легко и быстро, но чувствителен к случайностям, особенно при малом числе данных.

Среднее отклонение (МД) – это среднеарифметические разницы (по абсолютной величине) между каждым значением в выборке и ее средним.

$$МД = \frac{\sum d}{N},$$

где $d = |X - M|$,

M – среднее выборки,

X – конкретное значение,

N – число значений.

Множество всех конкретных отклонений от среднего характеризует изменчивость данных, но если не взять их по абсолютной величине, то их сумма будет равна нулю и мы не получим информации об их изменчивости. Среднее отклонение показывает степень сгущенности данных вокруг выборочного среднего. Кстати, иногда при определении этой характеристики выборки вместо среднего (M) берут иные меры центральной тенденции – моду или медиану.

Дисперсия (D) характеризует отклонения от средней величины в данной выборке.

$$D = \frac{\sum d^2}{N} \text{ для больших выборок } (N > 30);$$

$$D = \frac{\sum d^2}{(N - 1)} \text{ для малых выборок } (N < 30),$$

Стандартное отклонение (δ). Из-за возведения в квадрат отдельных отклонений d при вычислении дисперсии полученная величина оказывается далекой от первоначальных отклонений и потому не дает о них наглядного представления. Чтобы этого избежать и получить характеристику, сопоставимую со средним отклонением, проделывают обратную математическую операцию – из дисперсии извлекают квадратный корень. Его положительное значение и принимается за меру изменчивости, именуемую среднеквадратическим, или стандартным, отклонением:

$$\delta = \sqrt{D} = \sqrt{\frac{\sum d^2}{N}} \text{ для больших выборок } (N > 30);$$

$$\delta = \sqrt{D} = \sqrt{\frac{\sum d^2}{(N - 1)}} \text{ для малых выборок } (N < 30),$$

Пример 6. Вычислим дисперсию (δ^2) и среднеквадратичное отклонение для следующего ряда: 2, 4, 6, 8, 10. Прежде всего, найдем среднее (M) для данного ряда, оно равно 6.

Из каждого элемента ряда вычтем величину среднего этого ряда. Полученные величины характеризуют то, насколько каждый элемент отклоняется от средней величины в данном ряду. Экспериментальные данные этой задаче, необходимые для расчета дисперсии, представим в виде (таблица 1).

Таблица 1

Первичный результат (x_i)	$x_i - M$	$(x_i - M)^2$
2	- 4	16
4	- 2	4
6	0	0
8	2	4
10	4	16
$M = 6$		$\sum (x_i - M)^2 = 40$

Далее разности возводят в квадрат суммируются. Полученную сумму квадратов разностей делим на объем данной выборки. В нашем примере

получится следующее: $\sigma^2 = \frac{40}{6} = 6,8$

$$\sigma = \sqrt{\frac{40}{6}} = \sqrt{6,8} = 2,58$$

- стандартное отклонение.

О чем же свидетельствует стандартное отклонение равное 2,58? Оно позволяет сказать, что большая часть результатов данного исследования располагается в пределах 2,58 от средней, т. е. между 3,42 (6 - 2,58) и 8,58 (6 + 2,58).

Для того чтобы лучше понять, что подразумевается под «большой частью результатов», необходимо рассмотреть те свойства стандартного отклонения, которые проявляются при нормальной или приблизительно нормальной кривой распределения, т. к. здесь существует прямое соответствие между σ и относительным количеством случаев. На рис. 1 по горизонтальной оси отложены интервалы, соответствующие отклонению в 1σ , 2σ и 3σ вправо и влево от среднего значения (M). Процент случаев, приходящийся на интервал $M + 1\sigma$ в нормальном распределении, равен 34,13. Поскольку кривая симметрична, 34,13 случаев приходится также на интервалы от $M - 1\sigma$, так, что диапазон от $- 1\sigma$ до $+ 1\sigma$ охватывает 68,26 % случаев. Почти все случаи (99,72%), т. е. почти все показатели лежат в пределах от $- 3\sigma$ до $+ 3\sigma$ относительно среднего значения.

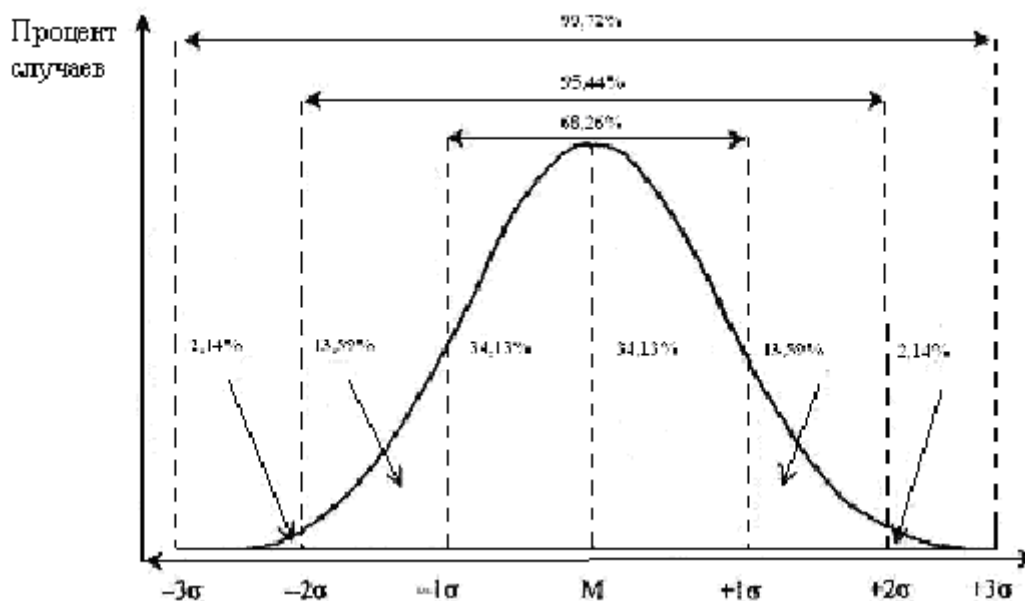


Рис. 1. – Процентное соотношение случаев для кривой нормального распределения

Эта закономерность известна как закон «трех сигм!» и является одной из характеристик нормального распределения.

2.2. Вторичная статистическая обработка данных

С помощью вторичных методов статистической обработки данных проверяются, доказываются или опровергаются гипотезы, связанные с эмпирическим исследованием.

Чаще всего в прикладных психолого-педагогических исследованиях применяются следующие методы вторичной статистической обработки результатов:

Методы сравнения двух или нескольких элементарных статистик (средних, дисперсий и т. д.) относящихся к разным выборкам.

Методы установления статистических связей между переменными (корреляционный анализ).

Методы выявления внутренней статистической структуры эмпирических данных (факторный анализ).

В общем, психологические задачи, решаемые с помощью методов математической статистики, условно можно разделить на несколько групп.

Задачи, требующие установления сходства или различия.

Задачи, требующие группировки и классификации данных.

Задачи, ставящие целью анализ источников вариативности (изменчивости) получаемых психологических признаков.

Задачи, предполагающие возможность прогноза на основе имеющихся данных.

Это далеко не полная сводка типов задач. Рассмотрим решение некоторых из них.

Способы оценки значимости различий. Для сравнения выборочных средних величин, принадлежащих к двум выборкам, которые распределены по нормальному закону, и для решения вопроса о том, отличаются ли средние значения статистически достоверно друг от друга, *используют t-критерий Стьюдента*.

Одним из главных достоинств критерия является широта его применения. Он может быть использован для сопоставления средних у связанных и несвязанных выборок, причем выборки могут быть не равны по величине.

Случай несвязанных выборок.

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

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

$$S_d = \sqrt{S_x^2 + S_y^2}$$

где (1)

Рассмотрим сначала равночисленные выборки. В этом случае $n_1 = n_2 = n$, тогда выражение (1) будет вычисляться следующим образом:

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

В случае неравночисленных выборок $n_1 \neq n_2$, выражение будет вычисляться следующим образом:

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

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.

Понятно, что при численном равенстве выборок $k = 2 \cdot (n - 2)$.

Рассмотрим пример использования t - критерия Стьюдента для несвязанных и неравных по численности выборок.

Пример 7. [9]

Критерий Фишера позволяет сравнивать величины выборочных дисперсий двух рядов наблюдений. Для вычисления $F_{\text{эмп}}$ нужно найти отношение дисперсий двух выборок, причем так, чтобы большая по величине

дисперсия находилась бы в числителе, а меньшая знаменателе. Формула

$$F_{\text{эмп}} = \frac{S_x^2}{S_y^2}$$

вычисления по критерию Фишера F такова:

Где $S_x^2 = \left(\frac{1}{n_1}\right) \times \sum (x_i - \bar{x})^2$

и $S_y^2 = \left(\frac{1}{n_2}\right) \times \sum (y_i - \bar{y})^2$

Поскольку, согласно условию критерия, величина числителя должна быть больше или равна величине знаменателя, то значение $F_{\text{эмп}}$ всегда будет больше

или равно единице, т.е. $F_{\text{эмп}} \geq 1$. Число степеней свободы определяется также

просто: $df_1 = n_1 - 1$ для первой (т.е. для той выборки, величина дисперсии которой больше) и $df_2 = n_2 - 1$ для второй выборки.

Пример 8. [10]

Корреляционный анализ.

Корреляцией называют зависимость между двумя переменными величинами.

Переменная – это любая величина, которая может быть измерена и чье количественное выражение может варьировать.

При изучении корреляций стараются установить, существует ли какая-то связь между двумя показателями в одной выборке (например, между ростом и весом детей или между уровнем IQ и школьной успеваемостью), либо между двумя различными выборками (например, при сравнении пар близнецов), и если эта связь существует, то увеличение одного показателя сопровождается возрастанием (положительная корреляция) или уменьшением (отрицательная корреляция) другого.

Коэффициент корреляции – это величина, которая может варьировать в пределах от +1 до -1. В случае полной положительной корреляции этот коэффициент равен (+1), а при полной отрицательной (-1).

В случаи, если коэффициент корреляции равен 0, обе переменные полностью независимы друг от друга.

В гуманитарных науках корреляция считается сильной, если ее коэффициент выше 0,60; если же он превышает 0,90, то корреляция считается очень сильной.

Можно выделить несколько видов корреляционного анализа: линейный, ранговый, парный и множественный. Мы рассмотрим два вида корреляционного анализа - *линейный и ранговый*.

Коэффициент корреляции Пирсона.

Линейный корреляционный анализ позволяет установить прямые связи между переменными величинами по их абсолютным значениям. Формула расчета коэффициента корреляции построена таким образом, что если связь между признаками имеет линейный характер, коэффициент Пирсона точно устанавливает тесноту этой связи. Поэтому он называется также коэффициентом линейной корреляции Пирсона.

Расчет коэффициента корреляции Пирсона в Excel [11].

Пример задачи на расчёт коэффициента корреляции Пирсона: 20 школьникам были даны тесты на наглядно-образное и вербальное мышление. Измерялось среднее время решения заданий теста в секундах. Психолога интересует вопрос: существует ли взаимосвязь между временем решения этих задач? Переменная X - обозначает среднее время решения наглядно-образных, а переменная Y - среднее время решения вербальных заданий тестов.

Для применения коэффициента корреляции Пирсона необходимо соблюдать следующие условия:

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

Распределения переменных X и Y должны быть близки к нормальному.

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

Коэффициент корреляции рангов Спирмена.

Коэффициент корреляции рангов, предложенный К. Спирменом, относится к непараметрическим показателям связи между переменными, измеренными в ранговой шкале. При расчете этого коэффициента не требуется никаких предположений о характере распределений признаков в генеральной совокупности. Этот коэффициент определяет степень тесноты связи порядковых признаков, которые в этом случае представляют собой ранги сравниваемых величин.

Величина коэффициента корреляции Спирмена также лежит в интервале +1 и -1. Он, как и коэффициент Пирсона, может быть положительным и отрицательным, характеризуя направленность связи между двумя признаками, измеренными в ранговой шкале.

В принципе число ранжируемых признаков (качеств, черт и т. п.) может быть любым, но сам процесс ранжирования большего, чем 20 числа признаков – затруднителен.

Расчет коэффициента корреляции Спирмена в Excel [12].

Пример задачи на расчёт коэффициента корреляции Спирмена: психолог выясняет, как связаны между собой индивидуальные показатели готовности к

школе, полученные до начала обучения в школе у 11 первоклассников и их средняя успеваемость в конце учебного года.

Внедрение в научные исследования автоматизированных средств обработки данных позволяет быстро и точно определять любые количественные характеристики любых массивов данных. Разработаны различные программы для компьютеров, по которым можно проводить соответствующий статистический анализ практически любых выборок. Из массы статистических приемов в психологии наибольшее распространение получили следующие:

- 1) комплексное вычисление статистик;
- 2) корреляционный анализ;
- 3) дисперсионный анализ;
- 4) регрессионный анализ;
- 5) факторный анализ;
- 6) таксономический (кластерный) анализ;
- 7) шкалирование.

Познакомиться с характеристиками этих методов можно в специальной литературе, приведенной в методических рекомендациях.

СПИСОК ЛИТЕРАТУРЫ

1. Афанасьев, В. Н. Статистическая методология в научных исследованиях : учебное пособие для аспирантов / В. Н. Афанасьев, Н. С. Еремеева, Т. В. Лебедева. — Оренбург : Оренбургский государственный университет, ЭБС АСВ, 2017. — 246 с. — ISBN 978-5-7410-1703-6. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/78841.html> (дата обращения: 01.11.2021). — Режим доступа: для авторизир. Пользователей
2. Баврин, И. И. Математическая обработка информации : учебник для студентов всех профилей направления «Педагогическое образование» / И. И. Баврин. — Москва : Прометей, 2016. — 262 с. — ISBN 978-5-9908018-9-9. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/58146.html> (дата обращения: 01.11.2021). — Режим доступа: для авторизир. Пользователей
3. Глотова, М. Ю. ИКТ и математические методы обработки данных : учебное пособие / М. Ю. Глотова, Е. А. Самохвалова. — Москва : Московский педагогический государственный университет, 2019. — 244 с. — ISBN 978-5-4263-0767-4. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/94642.html> (дата обращения: 01.11.2021). — Режим доступа: для авторизир. Пользователей
4. Маглеванный, И. И. Математические основы первичной обработки экспериментальных данных : методические материалы по прикладной статистике / И. И. Маглеванный, Т. И. Карякина. — Волгоград : Волгоградский государственный социально-педагогический университет, «Перемена», 2015. — 42 с. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/40738.html> (дата обращения: 01.11.2021). — Режим доступа: для авторизир. Пользователей
5. Мирзоев, М. С. Основы математической обработки информации : учебное пособие / М. С. Мирзоев. — Москва : Прометей, 2016. — 316 с. — ISBN 978-5-906879-01-1. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/58165.html> (дата обращения: 01.11.2021). — Режим доступа: для авторизир. Пользователей
6. Перевозкин, С. Б. Методы математической статистики в научно-исследовательской работе психолога : учебное пособие / С. Б. Перевозкин, Ю. М. Перевозкина. — Новосибирск : Новосибирский государственный университет экономики и управления «НИНХ», 2017. — 162 с. — ISBN 978-5-7014-0797-6. — Текст : электронный // Электронно-библиотечная система IPR

BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/87132.html> (дата обращения: 01.11.2021). — Режим доступа: для авторизир. пользователей

7. Романко, В. К. Статистический анализ данных в психологии : учебное пособие / В. К. Романко. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 313 с. — ISBN 978-5-00101-802-5. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/89075.html> (дата обращения: 01.11.2021). — Режим доступа: для авторизир. пользователей

Интернет-ресурсы

8. ©МатБюро - Решение задач по математике, экономике, статистике

https://www.matburo.ru/ex_ms.php?p1=mspsy

9. <https://statpsy.ru/t-student/primer-t-test-ind/>

10. <http://cito-web.yspu.org/link1/metod/met125/node33.html>

11. <https://statpsy.ru/pearson/raschet-pearson-excell>

12. <https://statpsy.ru/spearman/spearman-excell>

ПРИЛОЖЕНИЕ

Приложение 1

Краткая классификация задач и методов их статистического решения

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

		коэффициент корреляции Пирсона Линейная и криволинейная регрессии
5. Анализ изменений признака под влиянием контролируемых условий	а) под влиянием одного фактора	S критерий Джонкира L критерий тенденций Пейджа Однофакторный дисперсионный анализ Критерий Линка и Уоллеса Критерий Немени Множественное сравнение независимых выборок
	б) под влиянием двух факторов одновременно	Двухфакторный дисперсионный анализ

Работать с данной таблицей рекомендуется следующим образом:

1. По первому столбцу таблицы выбирается задача, стоящая в исследовании.
2. По второму столбцу таблицы определяются условия решения задачи, например, сколько выборок обследовано или на какое количество групп может быть разбита обследованная выборка.
3. Выбирается соответствующий статистический метод. Можно выбрать несколько методов и сравнить их результаты.

Практические задачи для самостоятельного решения

1. В трех выпускных классах средней школы подсчитывался средний балл успеваемости.

Получены следующие результаты:

Пол	11-а класс		11-б класс		11-в класс		Число учащихся	Балл
	Число учащихся	Балл	Число учащихся	Балл	Число учащихся	Балл		
Девушки	18	3,62	15	3,90	17	3,75		
Юноши	12	3,44	13	3,58	13	3,70		

Вычислите средний балл успеваемости у девушек и юношей выпускных классов.

2. Имеется следующая совокупность экспериментальных данных: 15, 18, 20, 17, 15, 15, 20, 20, 16, 19, 18, 17, 20, 15, 16, 19, 17, 18. Вычислите для данной выборки среднее значение, дисперсию, выборочное отклонение, медиану, моду.

3. Обследованы две группы испытуемых (по 10 человек) на уровень личностной тревожности (УЛТ) по Спилбергу. Получены следующие результаты:

№№	Группа 1		Группа 2		Группа 3		Группа 4		Группа 5	
УЛТ-1 24	42	29	39	26	37	40	33	44	38	
УЛТ-2 34	40	26	47	29	31	38	43	45	42	

Определите средние значения УЛТ, выборочные отклонения для каждой группы испытуемых, сравните их между собой, сделайте выводы.

4. В трех классах начальной школы подсчитывалась среднее количество слов, читаемых в минуту. Получены следующие результаты:

Пол	1-а класс	1-б класс	1-в класс
Девочки	78	75	67
Мальчики	62	53	73

Вычислите среднее количество слов, читаемых в минуту девочками и мальчиками начальных классов.

5. Обследованы две группы испытуемых (по 10 человек) на уровень напряженности психологической защиты (НПЗ). Получены следующие результаты:

№№	1	2	4	5	6	7	8	9	10		
НПЗ –1		34	40	30	50	66	37	42	33	48	30
НПЗ -234		40	26	47	23	30	38	45	45	42	

Определите средние значения НПЗ; выборочные отклонения для каждой группы испытуемых, сравните их между собой; моду; медиану; сделайте выводы.

6. Студенты математического и филологического факультетов были обследованы на уровень социального интеллекта. В исследовании приняли участие 12 филологов и 13 математиков. Получены следующие результаты:

№№	1	2	3	4	5	6	7	8	9	10	11	12
13												
Филологи		139	137	137	135	134	137	140	141	139	137	136
135												
Математики	140	137	135	140	142	138	140	139	140	135	138	
142	133											

Докажите или опровергните мнение о различиях в социальном интеллекте между математиками и филологами, с помощью методов математической статистики. Обоснуйте выбор критерия. Сделайте выводы.

7. Студент факультета психологии, выполняя курсовую работу, исследовал уровень личностной тревожности учеников выпускных классов и подростков. Получены следующие данные:

№№	1	2	3	4	5
Юноши	12	14	10	9	
Подростки	14	12	13	11	15

Определите достоверность различий в уровне тревожности между юношами и подростками. Обоснуйте выбор критерия. Сделайте выводы.

8. В исследовании эмоционального выгорания приняли участие врачи, психологи, продавцы. Исследователи поставили перед собой задачу выявить различия в проявлении эмоционального выгорания у представителей разных профессий. Были обследованы 5 психологов, 5 врачей и 4

№№	1	2	3	4	5
Психологи	57	51	52	46	57
Врачи	61	52	59	60	60
Продавцы	52	49	58	44	

Определите достоверность различий в уровне эмоционального выгорания между психологами, врачами, продавцами. Обоснуйте выбор критерия. Сделайте выводы.

9. При исследовании у студентов 1-5 курсов проявления сотрудничества в конфликте были получены следующие данные:

№№	1	2	3	4	5
1 курс	5	6	8	8	4
2 курс	7	6	7	9	7
3 курс	8	6	8	8	9
4 курс	9	8	10	10	9
5 курс	9	9	12	8	10

Прав ли исследователь, утверждая, что частота использования стиля «сотрудничество» увеличивается со временем обучения в вузе. Обоснуйте выбор критерия. Сделайте выводы.

10. При изучении агрессивного поведения были обследованы студенты исторического и архитектурного факультетов по методике А. Басса-А.Дарки. Получены следующие результаты:

№№	1	2	3	4	5	6	7	8	Средние значения
Историки	6	5	8	13	4	8	5	7	7
Архитекторы	7	5	11	8	5	6	7	5	8

Определите достоверность различий по уровню агрессии у студентов историков и архитекторов.

11. Педагогические коллективы детского сада (д.с.) и школы (ш.) были обследованы на уровень эмоционального выгорания (ЭВ) по методике В.В. Бойко. Получены следующие результаты:

№№	1	2	3	4	5	6	7	8	Среднее значение
ЭВ-Д.с.	46	32	28	33	34	48	52	47	40
ЭВ-Ш.	47	52	44	48	52	36	48	53	47,5

Определите достоверность различий по уровню эмоционального выгорания у педагогов детского сада и школы, используя критерий Стьюдента.

Докажите или опровергните мнение о различиях в проявлении альтруизма юношами и девушками, с помощью методов математической статистики. Обоснуйте выбор критерия. Сделайте выводы.

12. Студент факультета психологии, выполняя курсовую работу, исследовал уровень агрессии у студентов и курсантов. Были обследованы 20 человек: 10 курсантов и 10 студентов. Получены следующие данные:

№№	1	2	3	4	5	6	7	8	9	10	11	12
Студенты	12	14	10	9	10	9	14	15	11	8		
Курсанты	14	12	13	11	15	11	12	12	9	17		

Определите достоверность различий в уровне агрессии между студентами и курсантами. Обоснуйте выбор критерия. Сделайте выводы.

13. В исследовании социального интеллекта приняли участие педагоги, строители, музыканты. Исследователи поставили перед собой задачу выявить различия социального интеллекта у представителей разных профессий. Были обследованы 12 педагогов, 8 строителей и 4 музыканта.

№№	1	2	3	4	5	6	7	8	9	10	11	12
Педагоги	37	31	32	26	37	24	18	46	59	19	41	
Строители	21	32	39	30	40	40	29	26				
Музыканты	32	29	38	24								

Определите достоверность различий в уровне социального интеллекта между строителями, педагогами, музыкантами. Обоснуйте выбор критерия. Сделайте выводы.

14. При исследовании ситуативной тревожности студентов с 1 по 4 курс были получены следующие данные:

№№	1	2	3	4	5
1 курс	10	11	9	8	14
2 курс	11	10	7	9	
3 курс	10	6	8	8	12
4 курс	9	8	5	10	9

Прав ли исследователь, утверждая, что ситуативная тревожность снижается со временем обучения в вузе. Обоснуйте выбор критерия. Сделайте выводы.

15. Психолог проводит тренинг партнерского общения. Его задача – выяснить будет ли эффективен данный конкретный вариант тренинга для изменения показателей готовности к сотрудничеству участников?

№№										
До	4	1	5	4	3	5	6	2	6	5
После	4	1	5	5	3	5	3	6	7	3

16. Психолог изучает длительность подготовки домашнего задания учениками начальных, средних и старших классов (в мин.). Увеличивается ли учебная нагрузка детей со временем школьного обучения?

№	1	2	3	4	5	6	7
Нач.классы	50	70	20	20	60	50	60
Ср.классы	70	200	50	80	70	60	50
Ст.классы	120	90	100	130	120	110	45

17. Пяти школьникам предъявляют анаграммы. Фиксируется время решения каждого задания. Выясняется вопрос – будут ли найдены статистически значимые различия между временем решения трех различных анаграмм?

№№	Время решения анаграммы 1 в сек.			Время решения анаграммы 2 в сек.			Время решения анаграммы 3 в сек.		
1	5	235	7						
2	7	604	20						
3	2	93	5						
4	2	171	8						

18. Выпускникам педагогического университета было предложено оценить уровень психологической готовности к работе учителем по 5-балльной шкале до и после педагогической практики. Изменились ли представления студентов о работе учителя во время прохождения практики?

№№	1	2	3	4	5	6	7	8	9	10
До	4	5	4	4	3	5	4	2	4	5
После	4	4	5	5	4	4	5	4	4	4

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

№№	Уровень тревожности «до» тренинга			Уровень тревожности «после» тренинга		
№№	Уровень тревожности «до» тренинга	Уровень тревожности «до» тренинга	Уровень тревожности «до» тренинга	Уровень тревожности «после» тренинга	Уровень тревожности «после» тренинга	Уровень тревожности «после» тренинга
1	24	22	10	24	21	
2	12	12	11	33	30	
3	40	23	12	38	26	
4	30	31	13	39	38	
5	40	32	14	25	23	
6	35	24	15	28	22	
7	40	40	16	36	22	
8	32	12	17	37	36	
9	40	22	18	32	38	

20. Психолог проводит с младшими школьниками коррекционную работу по формированию навыка внимания, используя для оценки результатов коррекционную пробу. Задача состоит в том, чтобы определить, будет ли уменьшаться количество ошибок внимания у младших школьников после специальных коррекционных упражнений?

№№	Количество ошибок «До» упражнений			Количество ошибок «После» упражнений		
№№	Количество ошибок «До» упражнений	Количество ошибок «До» упражнений	Количество ошибок «До» упражнений	Количество ошибок «После» упражнений	Количество ошибок «После» упражнений	Количество ошибок «После» упражнений
1	24	22	10	22	21	
2	12	12	11	33	34	
3	42	41	12	78	56	
4	30	31	13	79	78	
5	40	32	14	25	23	
6	55	44	15	28	22	
7	50	50	16	16	12	
8	52	32	17	17	16	
9	50	32	18	12	18	

21. Шести школьникам предъявляют тест Равена. Фиксируется время решения каждого задания. Выясняется вопрос – будут ли найдены статистически значимые различия между временем решения первых трех заданий теста?

№№	Время решения первого задания теста в сек.	Время решения второго задания теста в сек.	Время решения третьего задания теста в сек.
8	3	5	
4	15	12	
6	23	15	
3	6	6	
7	12	3	
15	24	12	

22. Анализируя процесс выполнения школьниками заданий теста Равена, психолог предположил, что времени на решение второго и четвертого заданий затрачивается больше по сравнению с первым и третьим заданием. Докажите или опровергните предположение исследователя.

№№	Время решения задания 1 в сек.	Время решения задания 2 в сек.	Время решения задания 3 в сек.	Время решения задания 4 в сек.
1	8	5	12	3
2	4	12	13	15
3	6	15	20	23
4	3	6	12	6
5	7	3	8	12
6	15	12	7	24

23. В эксперименте испытуемый должен произвести выбор левого или правого стола с заданиями. В инструкции психолог подчеркивает, что задания на обоих столах одинаковы. Из 150 испытуемых правый стол выбрали 98 человек, а левый 52. можно ли утверждать, что подобный выбор левого или правого стола равновероятен или он обусловлен какой-либо причиной неизвестной психологу?

24. Одинаков ли уровень подготовленности учащихся в двух школах, если в первой школе из 100 человек поступили в вуз 82 человека и во второй школе из 87 человек поступили в вуз 44?

25. В эксперименте психологу необходимо использовать шестигранный игральный кубик с цифрами на гранях от 1 до 6. Для чистоты эксперимента необходимо получить «идеальный» кубик, т.е. такой, чтобы при достаточно большом числе подбрасываний (120), каждая его грань выпадала бы примерно равное число раз. Задача состоит в выяснении того, будет ли данный кубик близок к идеальному?

Грани	1	2	3	4	5	6
Частоты эмпирические	18	23	15	21	25	18
Частоты теоретические	20	20	20	20	20	20

Учебное издание

Г.В. Шагина

Методические рекомендации по дисциплине
«Математические и статистические методы обработки научных данных»
по направлению подготовки: 44.02.04 Психолого-педагогическое образование,
профиль: Психолого-педагогическое образование

Подписано в печать 20.11.2021

Формат 60x90/16. Усл. печ. л. 8,7

Тираж 100

Из-во ЗАО «Библиотека А. Миллера»

Отпечатано с готового оригинала-макета в типографии ЮУрГГПУ
Пр. Ленина, 69