Федеральное агентство по образованию РФ
АМУРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ
Зав. кафедрой Педагогики и психологии
____________________А.В. Лейфа
«_____»__________________2007г.

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

УЧЕБНО-МЕТОДИЧЕСКИЙ КОМПЛЕКС

для специальности 050711 «Социальная педагогика»

Составитель: Т.А. Юрьева

Благовещенск 2007 г.

Печатается по решению

редакционно-издательского совета

факультета математики и информатики

Амурского государственного университета

Т.А Юрьева

Учебно-методический комплекс дисциплины «Математические методы в педагогических исследованиях» для специальности 050711. – Благовещенск: АмГУ, 2007. – 130 с.

СОДЕРЖАНИЕ

1. РАБОЧАЯ ПРОГРАММА

1.1 Цели и задачи учебной дисциплины «Математические методы в педагогических исследованиях» и ее место в учебном процессе

Цели преподавания дисциплины «Математические методы в педагогических исследованиях».

Основной целью преподавания дисциплины «Математические методы в педагогических исследованиях» является обучение студентов математическим и статистическим методам обработки и анализа экспериментальных данных в практических и научных в педагогических исследованиях.

Задачи преподавания дисциплины.

- на примерах математических понятий и методов продемонстрировать сущность научного подхода, специфику математики, ее роль в развитии в педагогических исследований;
- научить студентов приемам исследования и решения, математически формализованных задач педагогики.

Перечень учебных дисциплин с указанием разделов, усвоение которых необходимо для изучения осознания учебных тем, вопросов курса «Математические методы в педагогических исследованиях».

При работе над практикумом читатели должны опираться на систему базовых математических знаний, приобретенных при изучении высшей математики, теории вероятностей и математической статистики, и понимать качественный смысл тех количественных преобразований в области педагогики.

После изучения дисциплины студент должен знать и уметь использовать:

- соотношение понятий педагогики и математики
- математический аппарат педагога; язык, модели, методы, линейные модели
- статистические методы для планирования экспериментов и обработки получаемых результатов
- методы измерения, тесты и рейтинг

1.2 Содержание учебной дисциплины «Математические методы в педагогических исследованиях».

В течении периода изучения математических методов в педагогических исследованиях студенты обязаны прослушать теоретический курс в объеме 38 часов и закрепить материал на практических занятиях.

Программа курса составлена в объеме необходимом для изучения общенаучных гуманитарных и специальных дисциплин, развития навыков математического мышления для специальностей гуманитарного профиля, необходимого для обработки информации и использования математических моделей в педагогических исследованиях.

Распределение учебного времени

	Тема	В часах		
		Лекции (заочное)	Практические занятия (заочное)	Сам. раб. (заочное)
1.	**Структура курса «Математические методы в педагогических исследованиях».** Предмет и содержание прикладной статистики, ее связь с педагогическими исследованиями. Характер данных, встречающихся в педагогической практике и педагогических исследованиях. Понятие эксперимента, измерения, измерительные шкалы.	2 ( )	-	
2.	**Начала теории измерений**. Множества. Бинарные отношения и их свойства. Отображения. Гомоморфизм. Алгебраическая структура. Понятие шкалы. Измерение. Номинальная шкала и ее свойства. Отношение порядка. Порядковая шкала и ее свойства. Числовые шкалы и их свойства. Шкалы интервалов, отношений и абсолютная шкала. Связь и преобразование числовых шкал. Преобразование данных из одного типа шкалы в другой и связанные с этим ограничения и опасности.	4 ( )	2	
3.	**Математическая статистика**. **Критическая статистика.** Понятие статистической оценки параметров. Характеристики оценок.	4 ( )		

	Статистические гипотезы и их проверка. Интервал правдоподобия и критическая область. Уровень значимости. Ошибки первого и второго рода.			
4.	**Параметрические критерии.** Критерий Стьюдента, критерий Фишера.	4 ( )	2	
5.	**Основы корреляционного анализа.** Коэффициент корреляции r Пирсона и его свойства как меры связи. Ранговая корреляция, коэффициенты ρ Спирмена и τ Кендалла.	4( )	4( )	
6.	**Основы регрессионного анализа.** Линейная и нелинейная регрессия, и оценка ее качества. Множественная регрессия.	4 ( )	1	
7.	**Непараметрические критерии.** Критерии различий, изменений. Критерии согласия. X^2 и теорема Пирсона. Критерий однородности.	6 ( )	4	
8.	**Факторный анализ**. Область применения и основные этапы факторного анализа. Дополнительные статистические показатели для оценки факторного анализа.	4 ( )	2	
9.	**Дисперсионный анализ.** Понятие об одно-, двух- и многофакторном дисперсионном анализе. Таблица сопряженности для числовых и номинальных признаков. Теорема Пирсона-Фишера.	4 ( )	2	
10.	**Метрическое и неметрическое многомерное шкалирование.** Матрица сходств и различий. Построение пространственной модели стимулов: Метод ортогональных проекций. Понятие метрики. Метрическая модель. Кластерный анализ.	2	2	
11.	**ВСЕГО**	38( )	19( )	

1.3 Формы текущего контроля знаний студентов

Результативность работы обеспечивается системой контроля, которая при очной форме обучения включает опрос студентов на практических занятиях, проверку выполнения домашних заданий, контрольные работы.

Каждое практическое занятие начинается с проверки домашнего задания, опроса по теоретическому материалу.

На лекциях и практических занятиях проводятся мини контрольные работы.

Данная программа предусматривает в течение семестра проведение двух плановых контрольных работы

Студенты заочной формы обучения текущий контроль усвоения материала осуществляют самостоятельно по контрольным вопросам.

Студентам дневной и заочной формы обучения в течении изучения дисциплины выполняется индивидуальные исследовательские задания.

Тематическое планирование практических занятий и формы текущего контроля

1 семестр			
Тема занятия	**Час.**	**Форма контроля**	
		К	**Индивидуальное задание**
1. Начала теории измерений. Шкалы	2		
2. Параметрические критерии	2	+	
3. Корреляционный анализ	4		
4. регрессионный анализ	1		
4. Критерии различий. Критерии изменений	2		
5. Критерии согласия.	2	+	
6. Факторный анализ.	2		
7. Дисперсионный анализ (многофакторный)	2		
8. Понятие метрики. Метрическая модель. Кластерный анализ.	2		

1.4 Итоговый контроль

Студенты дневного отделения допускаются к сдаче зачета при условии выполнения ими на положительную оценку всех форм текущего контроля, предусмотренных программой.

Студенты-заочники допускаются к зачету в установленном порядке, определенном «Положением о курсовых экзаменах и зачетах» АмГУ.

Зачет проводится по билетам, содержащих 10 заданий по вопросам из различных разделов программы. Отметка зачтено ставится при выполнении не менее 7 заданий. Либо по тестам, при условии правильного выполнения не менее 70% заданий.

ВОПРОСЫ К ЗАЧЕТУ

1. Предмет и содержание прикладной статистики.
2. Характер данных, встречающихся в педагогической практике. Понятие эксперимента, измерения.
3. Бинарные отношения и их свойства. Примеры.
4. Гомоморфизм, структура, шкала. Примеры.
5. Классификация шкал.
6. Неметрические шкалы: наименований и порядка. Примеры.
7. Метрические шкалы: интервалов и отношений. Примеры.
8. Построение статистической оценки методом моментов.
9. Статистические гипотезы и их проверка.
10. Односторонний и двусторонний интервал правдоподобия.
11. Назначение уровня значимости и его связь с риском ошибки.
12. Критерии согласия.
13. Критерии однородности.
14. Анализ связи признаков.
15. Факторный анализ.
16. Кластерный анализ.
17. Метрическое и неметрическое многомерное шкалирование.
18. Построение метрической модели.
19. Сущность дисперсионного анализа.
20. Однофакторный дисперсионный анализ.
21. Особенности двух факторного дисперсионного анализа.
22. Особенности дисперсионного анализа данных с повторными измерениями.
23. Специфика и допущение многомерного дисперсионного анализа.

24.Корреляционный анализ.

25.Регрессионный анализ.

1.7. Учебно-методические материалы

Основная литература**:**

1. Ермолов. О.Ю. Математическая статистика для психологов, 2003г.

2. Сидоренко Е.В. Методы математической обработки в психологии. СПб.: ООО «Речь», 2001г.

3. Двоерядкина Н.Н., Киселева А.Н., Юрьева Т.А. Математические методы в психологии и социологии: Практикум. Благовещенск: Амурский гос. ун-т, 2005г.

Дополнительная литература:

1. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике: Учебное пособие для вузов. М.:Высш. шк., 2001г.

2. Гусев А.Н. Дисперсионный анализ в экспериментальной психологии. М.:уч.-мет. Коллектор «Психология», 2000г.

3. Гусев А.Н., Измайлов Н.А., Михайлевская М.Б. Измерения в психологии. М.:-Наука, 1997г.

4. Айвазян С. А., Мхитарян В. С. Прикладная статистика: Учебник для вузов – М.: ЮНИТИ, 1998г.

5. Шеенко П.С. Практикум по математической статистике для нематематических факультетов: Учебное пособие. – Комсомольск-на-Амуре: ГПУ, 2001г.

2. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПРОФЕССОРСКО-ПРЕПОДАВАТЕЛЬСКОМУ СОСТАВУ

2.1 Методические рекомендации по проведению лекций

Лекция - традиционно ведущая форма обучения в вузе. Ее основная дидактическая цель - формирование ориентировочной основы для последующего усвоения студентами учебными материала. Будучи главным звеном дидактического цикла обучения, она выполняет научные, воспитательные и мировоззренческие функции, вводит студента в творческую лабораторию лектора.

Лекция - методологическая и организационная основа для всех форм учебных занятий, в том числе самостоятельных. Методологическая основа - так как вводит студента в науку вообще, придает учебному курсу концептуальность, а организационная - так как все другие формы учебных занятий так или иначе "завязаны" на лекцию, чаще всего логически следуют за ней, опираются на нее содержательно и тематически.

Содержание лекции устанавливается на основе учебной программы данной дисциплины. Каждая лекция требует такого построения, чтобы студенты могли конспектировать ее в виде четко ограниченных, последовательных и взаимосвязанных положений, тезисов с выводами и заключениями.

Все отдельные лекции лекционного курса требуют, поэтому взаимосвязи, последовательности и единства цели. Важной является связь лекционного материала с другими курсами и видами обучения.

Как правило, отдельная лекция состоит из трех основных частей: введения, изложения содержательной части и заключения:

1. Вводная часть. Формирование цели и задачи лекции. Краткая характеристика проблемы. Показ состояния вопроса. Список литературы. Иногда установление связи с предыдущими темами.

2. Изложение. Доказательства. Анализ, освещение событий. Разбор фактов. Демонстрация опыта. Характеристика различных точек зрения. Определение своей позиции. Формулирование частных выводов. Показ связей с практикой. Достоинства и недостатки принципов, методов, объектов рассмотрения. Область применения.

3. Заключение. Формулирование основного вывода. Установка для самостоятельной работы. Методические советы. Ответы на вопросы.

Лектор не может не считаться с общим уровнем подготовки и развитием студентов, но в то же время ему не следует ориентироваться как на слабо подготовленных студентов, так и на особо одаренных студентов. Ориентиром, очевидно, должны быть студенты, успевающие по данному предмету, представляющие основной состав лекционных потоков.

В лекционных курсах необходимо последовательно, от лекции к лекции повышать уровень научного изложения и наблюдать, чтобы лекции были посильны и интересны большинству студентов. Особенно велика развивающая роль лекций как формы научного мышления на первых курсах обучения. Здесь наряду с учебной информацией лекция организует и направляет самостоятельную работу студентов, вызывает потребность дополнительного приобретения знаний путем самообразования. Поэтому на первых курсах необходимо основное педагогическое и психологическое внимание уделять системности лекционного изложения и рекомендациям для самостоятельной работы. Развивающая роль лекционного преподавания на первых курсах нуждается в большей доступности изложения материала, в более четкой форме логического построения, в замедленном функционировании основных положений и выводов. На первых курсах нужны конкретные указания о связи лекций с учебниками, пособиями, заданиями и другой самостоятельной работой. Необходимо на младших курсах приучить студентов вести записи лекций, так как правильное конспектирование не только фиксирует основное содержание лекций, но и активизирует восприятие лекционного материала и организует внимание студентов к предмету.

Применение на лекциях вспомогательных средств, главным образом демонстрационных, повышает интерес к изучаемому материалу, обостряет и направляет внимание, усиливает активность восприятия, способствует прочному запоминанию.

2.2 Методические рекомендации по проведению практических и лабораторных занятий

Практические занятия целесообразно начинать с проверки знания и понимания студентами теоретического материала и умения применять эти знания для решения типовых задач.

Основной блок практического занятия составляет решение задач, поскольку истинное качество усвоения информации достигается именно в процессе формирования умений ее востребовать и применять. Задачи должны различаться по степени обобщенности действий и по виду самостоятельной деятельности, выполняемой студентами.

Организация процесса решения задач на практических занятиях должна обеспечивать самостоятельность студентов, основанная на индивидуальном и дифференцированном подходе, когда каждый студент работает в оптимальном для него темпе. Однако некоторые виды работ целесообразно организовывать, объединяя студентов в малые группы. В конце занятия обязательным является подведение итогов работы.

На лабораторных занятиях по «Математическим методам в психологии» формируется и совершенствуется практический уровень владения компьютерными средствами обработки статистических данных, основывающийся на применении теоретических знаний. Для проведения лабораторных занятий со студентами по дисциплине «Математические методы в психологии» используются компьютерные классы. Занятия в компьютерном классе предполагают индивидуальную или парно-групповую формы организации обучения.

Этапы проведения лабораторной работы следующие:

- Контрольный опрос студентов для проверки готовности к выполнению лабораторной работы (до 10 мин).
- Выдача индивидуального задания и пояснения о порядке выполнения индивидуального задания (до 5 мин).
- Выполнение индивидуального задания (около 1 ч.)
- Оформление результатов работы. Сдача выполненной работы преподавателю (до 10 мин).
- Получение домашнего задания (1-2 мин.)
- Приведение в порядок рабочего места, в том числе закрытие всех рабочий окон и уничтожение созданных на винчестере индивидуальных файлов (3-4 мин).

Индивидуальные задания для лабораторных работ должны быть представлены конкретно-практическими и творческими задачами.

На первой ступени изучения темы выполняются конкретно-практические задачи, при решении которых формируется минимальный набор умений. Преподаватель опосредованно руководит самостоятельной познавательной деятельностью студентов, консультирует студентов при возникновении непосильных затруднений в ходе решения задачи, обращает внимание группы "опасные" места решения. Отработка минимального набора навыков завершается во внеаудиторное время при выполнение домашней работы. Принимая во внимание сложность доступа некоторыми студентами к компьютерной техники во внеаудиторное время, домашние задания по " Математическим методам в психологии " должны носит большей части моделирующий характер.

Вторая ступень изучения темы дифференцируется в зависимости от степени усвоения его обязательного уровня. Студенты, овладев основами теории и усвоив содержание типовых методов и приемов решения задач, приступают к решению творческих задач. Если уровень знаний и умений, демонстрируемых студентом при контрольном обследовании, не соответствует установленным требованиям, студент вновь возвращается к стандартным упражнениям, но под более пристальным наблюдением преподавателя.

Студенты, пропустившие лабораторные занятия, должны их выполнить во внеаудиторное время и отчитаться до начала экзаменационной сессии.

2.3 Методические рекомендации по проведению по организации контроля знаний

В Университете качество освоения образовательных программ оценивается путем осуществления текущего контроля успеваемости, проведения промежуточных аттестаций и итогового контроля по окончании семестра.

На первом занятии до сведения студентов доводятся требования и критерии оценки знаний по дисциплине.

Целью текущего контроля успеваемости является оценка качества освоения студентами образовательных программ в течение всего периода обучения. К главной задаче текущего контроля относится повышение мотивации студентов к регулярной учебной работе, самостоятельной работе, углублению знаний, дифференциации итоговой оценки знаний.

Текущий контроль успеваемости осуществляется систематически и, как правило, преподавателем, ведущим практические занятия. Формами текущего контроля являются письменные опросы, автоматизированное тестирование, аудиторные контрольные работы, домашние задания. В течение семестра преподавателем должно быть проведено не менее 4 контрольных проверок знаний по каждому студенту из учебной группы.

Результаты текущего контроля служат основанием для прохождения студентом промежуточной аттестации.

Итоговый контроль (зачет или экзамен) преследуют цель оценить работу студентов за курс, полученные теоретические знания, их прочность, развитие творческого мышления, навыки самостоятельной работы, умение синтезировать полученные знания и применять их при решении практических задач. Задания итогового контроля состоят из двух частей: письменного теоретического опроса и практических заданий.

Во время проведения итогового контроля (зачета или экзамена) студентам не разрешается пользоваться вспомогательными материалами Их использование, а также попытки общения с другими студентами или иными лицами, в т.ч. с применением электронных средств связи, перемещения без разрешения экзаменатора и т.д., являются основанием для удаления студента из аудитории с последующим выставлением в ведомость неудовлетворительной оценки.

Критериями ОЗ - оценки знаний студента являются:

- ТМ - уровень освоения теоретического материала, предусмотренного программой курса;
- ПЗ - умение использовать теоретические знания при решении практических задач;
- СХ - социальные характеристики: посещаемость занятий; корректное общение с преподавателем; прилежание и трудолюбие; общая эрудиция; активность на занятиях;
- ТК – результаты текущего контроля.

Каждый критерий и итоговая оценка знаний студентов оценивается в баллах («отлично», «хорошо», «удовлетворительно», «неудовлетворительно»).

Итоговая оценка знаний студентов рассчитывается:

ОЗ = 0,25*ТМ+0,25*ПЗ+0,1*СХ+0,4*ТК.

3. ОРГАНИЗАЦИЯ УЧЕБНОЙ ДЕЯТЕЛЬНОСТИ СТУДЕНТОВ

3.1 Конспекты лекций

Структура курса «Математические методы в педагогических исследования». Предмет и содержание прикладной статистики, ее связь с психологией. Характер данных, встречающихся в психологической практике и психологических исследованиях. Понятие эксперимента, измерения, измерительные шкалы. Анализ данных на компьютере. Статистические пакеты; возможности и ограничения конкретных компьютерных методов обработки данных.

Математическая статистика — это наука о случайных явлениях. Под явлением понимается любой подлежащий изучению объект независимо от его конкретного содержания. По степени количественной определенности явления рассматриваются как отдельные события, величины, функции и как системы событий, величин, функций. Если изучаемые объекты можно трактовать как следствия многочисленных разнообразных по действию причин, то интересующие исследователя свойства этих объектов определяются неоднозначно и могут быть предсказаны лишь в результате массовых наблюдений не полностью, а с большей или меньшей вероятностью. Такие объекты и характеризуются как случайные явления (случайные события, величины, функции или *их* системы).

Математическая статистика занимается математическим описанием случайных явлений, т.е. построением вероятностных моделей, а также проверкой их пригодности. Поэтому выделяют два раздела: описательную статистику и статистику «проверяющую» (статистическую проверку гипотез); соответственно разделяется и методический аппарат. Понятия и методы описательной статистики создаются в теории вероятностей, а понятия и методы статистической проверки гипотез создаются в специальных теориях (например, в теории статистических решений) либо в приложениях теории вероятностей к конкретным наукам.

В словесном описании результатов наблюдения и эксперимента трудно избежать элементов субъективизма, которые проявляются чаще всего в преждевременных выводах и необоснованных обобщениях. Известно, что качественное описание является недостаточно точным, поскольку с помощью языковых средств сложно передать дифференцированность изучаемых явлений и особенно своеобразие их динамики. Одно только качественное описание не позволяет определить также и степень ошибки наблюдения или эксперимента. Но это вовсе не означает, что в психологии необходимо отказаться от качественного анализа в пользу оперирования исключительно количественными показателями. Мы хотим подчеркнуть лишь то обстоятельство, что количественный анализ результатов исследования должен не только предшествовать, но и обязательно следовать за качественным анализом.

Это особенно важно для интерпретации результатов исследования. Именно таким образом может быть преодолен субъективизм, так как формулируемые суждения и выводы становятся более независимыми от личности исследователя и обеспечивается возможность их проверки. Знание различных приемов обработки и анализа результатов наблюдений и эксперимента с помощью статистических показателей является обязательным для педагога. Поэтому первый раздел нашего практикума посвящен анализу видов психоло-педагогических измерений и способов статистической обработки получаемых при этом результатов. Однако следует иметь в виду, что знакомство с материалом данного раздела не может заменить студенту систематического изучения математической статистики. Нами будут рассмотрены лишь элементарные статистические методики, без которых нельзя обойтись на практических занятиях по общей и экспериментальной психологии.

Процесс измерения лежит в основе любой эмпирической науки. Беглый взгляд на историю науки показывает, что совершенствование принципов и техники измерения было основным фактором, обеспечивающим ее движение вперед. Самого высокого уровня развития на сегодняшний день достигли те ее области, которым быстрее удалось преодолеть трудности, связанные с разработкой методологических и методических проблем измерения. Прежде чем рассмотреть приемы измерения, используемые в психолого-педагогическом эксперименте, и способы статистической обработки его результатов, познакомимся с основными методами психолого-педагогического исследования.

Исходя из порядка операций с объектами в научном исследовании, Б. Г. Ананьев разработал классификацию методов современной психологии. В основу ее он положил целостный цикл психологического исследования и все методы распределил по четырем группам. В первую группу, которую можно назвать группой организационных методов, Ананьев относит сравнительный, лонгитюдный (т. е. исследование одних и тех же лиц в течение длительного времени) и комплексный методы: «Они действуют на протяжении всего исследования, и их эффективность определяется по конечным результатам исследования...». Вторая группа методов включает известные также по традиционным классификациям эмпирические способы добывания научных данных. В эту группу входят: обсервационные методы (наблюдение и самонаблюдение), экспериментальные методы (лабораторные, полевые, психолого-педагогические), психодиагностические методы (тесты, анкеты, опросники, интервью, беседы), праксиметрические методы (приемы анализа процессов и продуктов деятельности: хронометрии, профессиографическое описание, оценка выполненных работ), моделирование (математическое, кибернетическое) и биографические методы (приемы исследования жизненного пути, изучение документации). Третью группу методов составляют приемы обработки результатов эксперимента и наблюдений. Ананьев в эту группу относит как стандартные приемы статистической обработки данных (количественная обработка), так и приемы качественного анализа, включая дифференциацию материала по классам, разработку типологии, составление психологической казуистики (описание случаев). Четвертая группа методов — интерпретационные методы — представлена в классификации Ананьева вариантами генетического и структурного методов. Генетический метод интерпретирует весь обработанный материал исследования в характеристиках развития, а структурный метод — в характеристиках типов связей между отдельными компонентами структуры изучаемой личности или структуры социальной группы.

Приступая к выбору методики, экспериментатор должен иметь четкое представление о том, что именно он хочет измерить, и удовлетворят ли результаты измерения требованиям адекватного решения исследовательской или практической задачи. В первую очередь ему надлежит доказать валидность, надежность и объективность избранной методики. Под валидностью методики понимается адекватность ее предмету исследования. Количественно валидность определяется путем установления взаимосвязи между результатами, полученными с помощью данной методики, и каким-либо из внешних критериев. Поясним сказанное примером. Очевидно, что успешность обучения в какой-то степени обусловлена уровнем интеллектуального развития обучаемого, и поэтому в качестве внешнего критерия правомерно рассматривать оценку его успеваемости. Допустим, что было проведено тестовое исследование умственного развития группы лиц, например студентов, с помощью избранной методики. Так вот, применяемая методика может считаться валидной лишь в том случае, если между результатами тестирования и оценкой успеваемости в обучении будет обнаружена положительная взаимосвязь. Не менее важным аспектом оценки качества методики является ее надежность. Под надежностью психологической методики понимается точность производимых с ее помощью измерений. Иначе говоря, через надежность определяется пригодность данной методики в качестве измерительного инструмента. Наконец, объективность методики характеризует степень независимости результатов измерения от пользователя данной методики. Объективными результаты будут лишь в том случае, если, во-первых, они независимы от личностных особенностей пользователя и, во-вторых, исключен произвол в их обработке и интерпретации.

Для проверки валидности и надежности методик чаще всего привлекаются количественные (статистические) критерии оценки. Объективность методики можно обосновать исходя из положений общей теории измерений и специфики их в отношении психолого-педагогического исследовании. Конечно, это вовсе не означает, что психологическое исследование исчерпывается измерением. Однако знание разнообразных измерительных процедур вооружает педагога исследовательским инструментом, с помощью которого он способен решать психолого-педагогические задачи.

<u>Начала теории измерений</u>. Множества. Бинарные отношения и их свойства. Отображения. Гомоморфизм. Алгебраическая структура. Понятие шкалы. Измерение. Номинальная шкала и ее свойства. Отношение порядка. Порядковая шкала и ее свойства. Числовые шкалы и их свойства. Шкалы интервалов, отношений и абсолютная шкала. Связь и преобразование числовых шкал. Преобразование данных из одного типа шкалы в другой и связанные с этим ограничения и опасности.

С точки зрения теории измерения, все множество различных измерительных процедур, применяемых в педагогики и психологии, является процедурами построения шкалы эмпирической переменной, иначе говоря, процедурами психологического шкалирования. В понимании большинства психологов шкалирование — это совокупность экспериментальных и математических приемов для измерения особенностей психических процессов и состояний. Вслед за С. С. Стивенсоном в настоящее время понятие «шкалирование» рассматривают в качестве синонима понятия «измерение». Под шкалированием психологических процессов, свойств, объектов или событий понимается процесс приравнивания к этим процессам, свойствам, объектам или событиям чисел по определенным правилам, а именно таким образом, чтобы в отношениях чисел отображались отношения явлений, подлежащих измерению.

Итак, измерение состоит в отображении эмпирических систем с помощью математических систем, а целью такого рода отображения является частичная замена действий, производимых с реальными предметами, формальными действиями с числами. Область чисел выполняет функцию модели определенных свойств предметов и в качестве средства познания дает возможность более глубоко проникать в объективно существующие свойства и взаимосвязи. В этом смысле шкалирование (измерение) служит главной силой, преобразующей психологию из науки описательной, следующей за фактами, в науку, уме куцую предсказывать новые факты.

Понятно, что относительно разных эмпирических систем мы должны использовать разные методики измерения, т. е. применять измерительные шкалы разных типов. Понимание исследователем формальных аспектов измерения является необходимым условием для адекватного выбора им измерительных инструментов и процедур, а также для применения адекватных методов анализа полученных в наблюдении и эксперименте данных. Основываясь на правилах измерения, принято различать несколько типов шкал, с каждым из которых могут быть соотнесены конкретные процедуры шкалирования. При этом каждый тип шкалы может быть охарактеризован соответствующими числовыми свойствами. Рассмотрим более подробно основные свойства разных типов шкал, эмпирические операции, допустимые на уровне этих шкал, а также статистические приемы обработки и анализа исходных или, как их чаще называют, первичных результатов исследования.

Шкалы наименований, или номинативные шкалы. Шкала наименований представляет собой взаимнооднозначное отображение некоторой эмпирической системы в числовой системе. Таким образом, шкала наименований отображает взаимнооднозначное соответствие между классами эквивалентности, т. е. классами эмпирических объектов — обозначений. Само название «шкала наименований» указывает на то, что в этом случае шкальные значения играют роль лишь названий классов эквивалентности.

Шкалы наименований подчиняются законам равенства. То есть объект *A* может быть равен объекту *B* по признаку *X,* так что $X_A=X_B$; но по отношению к третьему объекту С по признаку А он может быть неравным: $X_A \neq X_C$. Любая другая связь между шкальными значениями, за исключением равенства, не имеет отношения к данному случаю, так как для данного типа шкал не существует никакого дополнительного определения. Шкала наименований представляет собой наиболее общую форму шкал. Все типы шкал в каждом отдельном случае являются некоторыми видами шкал наименований, но обладающими при этом теми или иными дополнительными свойствами. При построении шкал наименований должны быть выполнены следующие требования: во-первых, каждый член некоторого множества объектов должен быть отнесен лишь к одному классу объектов (или к собирательному классу «прочие объекты») и, во-вторых, ни один из объектов не может быть отнесен одновременно к двум или большему числу классов. К примеру, если принять, что глаза у людей могут быть только светлым и или темными, то все люди по этому признаку разделяются на две группы. При этом люди с множеством оттенков глаз: голубых, серо-зеленых и серых попадут в класс «люди со светлыми глазами», а те, у которых глаза карие и темно-коричневые, — в класс «люди с темными глазами». Из приведенного примера видно, что отношения эквивалентности по заданному признаку между классифицируемыми объектами, как правило, грубее реальных отношений, существующих между объектами.

С формальной точки зрения установление классов эквивалентности как будто не вызывает никаких затруднений. В действительности, как это было показано предыдущим примером, понятие «равенство» можно трактовать более узко или более широко в зависимости от «тонкости» или «грубости» используемой классификации по заданному признаку. При построении шкал наименований главными являются качественные различия, а количественные не принимаются во внимание. Поэтому числа, используемые в качестве обозначений классов эквивалентности в этих шкалах, не отражают количественных различий выраженности изучаемого признака.

Примерами дихотомической классификации могут быть: «нормальный/анормальный», «женатый/ холостой», «решает задачу/не решает задачу» и т. п. В случае так называемой истинной дихотомии классы могут быть четко разделены по определенному признаку, например: «мужской/ женский пол».

Однако бывают классификации с менее жесткими переходами признака, т, е. с довольно произвольными границами между классами эквивалентности, например: «способен к концентрации внимания/не способен к концентрации внимания». Именно с такого рода классификациями чаще всего и имеет дело психолог. Это так называемые квазидихотомические классификации. Построение и использование шкал с квазидихотомические границами классов вызывает ряд затруднений. Первая трудность, которая при этом возникает, состоит в установлении границы классов. В частности, какой же будет в нашем примере критерий «способности» к концентрации внимании, как определить точку в континууме «концентрация внимания», дифференцирующую людей на «способных» и «неспособных» к концентрации внимания?

Разберем другой пример из области психологии мышления. На первый взгляд альтернатива «решил задачу/не решил задачу» вполне может быть расценена как истинно-дихотомическая классификация. И действительно, в принципе для отнесения любого конкретного решения к классу «решил задачу» достаточно соотнести получаемый в нем результат с результатом, полученным достаточно большой группой людей, аналогичным образом решивших данную задачу. Все остальные решения можно тогда отнести к классу «не решил задачу». Однако возникает вопрос: действительно ли данный человек решил эту задачу? И вот почему: вполне возможно, во-первых, что решение было случайным, т.е. случайно данный результат совпал с результатом решения других людей и, во-вторых, что этот класс задач заранее был известен данному человеку. Но, как правило, такого рода сопровождающие факторы, например в психодиагностических тестах, совершенно не учитываются.

В шкале наименований с числами, которые мы приписываем объектам или классам объектов, нельзя производить никаких арифметических действий. Числа, обозначающие классы, нельзя суммировать, вычитать, умножать и делить. Дело в том, что структура шкалы остается ин вариантной по отношению к перемене обозначений (наименований) и к изменению последовательности, т. е. разного рода перестановкам. Следовательно, операция присвоения чисел классам объектов является совершенно произвольной операцией и ей не соответствуют операции, производимые с реальными объектами. Поэтому классы объектов можно обозначать любыми символами— произвольными числами, буквами или другими знаками при одном условии: каждый символ будет использован исключительно для обозначения одного класса объектов и одновременно ни один класс объектов не будет обозначаться двумя или большим числом символов.

Из вышесказанного уже очевидны те ограничения, которые накладываются на использование статистических приемов обработки результатов, полученных на уровне шкалы наименовании. Поскольку операции арифметического характера не допускаются, то в качестве меры центральной тенденции можно использовать лишь моду. Модальный класс объектов определяют после подсчета абсолютных или относительных частот, т. е. встречаемости того или иного результата в каждом классе. В качестве меры тесноты взаимосвязи между различными массивами измерений можно использовать некоторые коэффициенты корреляции. Для оценки статистической значимости различий между частотами или между модами можно использовать критерий хи-квадрат.

Шкалы порядка, или ординальные шкалы. В порядковых измерениях символы, в частности числа, присваивают классам объектов так, чтобы первые отображали не только равенство или неравенство, эквивалентность или неэквивалентность, но и упорядоченность объектов в отношении измеряемого свойства. В шкалах порядка массы объектов, как и в случае шкал наименовании, являются дискретными. И хотя числа можно сравнивать, всегда надо помнить, что в шкалах порядка их величины имеют лишь относительное, а не абсолютное значение. Например, если какой-то один класс объектов обозначен большим числом, чем другой, то мы понимаем, что по измеряемой характеристике первый превосходит второй, но при этом нам неизвестно, насколько велико это различие. Дело в том, что в самих измерительных операциях, связанных с установлением порядка, не содержится никаких данных о величине различий. Шкала порядка отображает монотонное возрастание или убывание измеряемого признака с помощью монотонно возрастающих или монотонно уменьшающихся чисел. Оценить направление изменения признака можно только в том случае, если шкала порядки содержит не меньше трех классов, которые образуют последовательность. Из-за того, что в шкале порядка устанавливается последовательность классов, любые преобразования, связанные с перестановками элементов этой шкалы, недопустимы.

К числу постулатов, которым подчиняются преобразования шкал порядка, относятся постулаты трихотомии, асимметрии и транзитивности. Рассмотрим явление трихотомии. Если два объекта *A* и *B* обладают признаком *X*, то между ними по данному признаку может существовать одно из трех отношений: $X_A < X_B$, или $X_A = X_B$, или $X_A > X_B$. В соответствии с постулатом асимметрии справедливым будет следующее утверждение: если между объектами А и В по признаку X обнаружено неравенство $X_A > X_B$, то никогда не может быть $X_A < X_B$, или $X_A = X_B$,. Наконец, в соответствии с постулатом транзитивности можно утверждать, что если три объекта А, В и С обладают признаком *X* и между ними по признаку X существуют отношения $X_A < X_B$ и $X_B < X_c$. то из этого следует, что $X_A < X_c$. Следовательно, для порядковых шкал допустимы любые преобразования типа x'=f(x), где f(x) представляет собой любое монотонное преобразование, не изменяющее последовательности элементов. Это означает, что для преобразования шкал порядка можно пользоваться возведением в степень, извлечением корня, логарифмированием.

Упорядочивание объектов может быть униполярным или биполярным. При униполярном установлении порядка объекты или классы объектов соотносят, используя в качестве индикатора степень выраженности одного единственного свойства. Например, шкала порядка для оценки умственной отсталости может содержать следующие классы: «нет отклонения от нормы/отклонение слабое/отклонение среднее/отклонение сильное».

При биполярном упорядочивании исходят, как правило, из популярных проявлений какого-то свойства, которые фиксируются в виде двух «точек отсчета» на шкале. Примером биполярной шкалы в психологическом исследовании является методика семантического дифференциала. В этом случае для построения шкалы первоначально производят отбор некоторого множества понятий, которые могут характеризовать, по мнению исследователя, изучаемые психические свойства испытуемого. Затем каждому понятию находят антоним (например: «общительный — замкнутый», «сильный — слабый», «уравновешенный — неуравновешенный»). Очевидно, что между каждыми двумя такими понятиями располагается несколько промежуточных оценочных категорий. Словесное определение промежуточных категорий очень часто вызывает у исследователей значительные трудности, поскольку и языке, как правила, мы легче находим понятия для обозначения экстремальных степеней выраженности какого-то свойства и труднее — для промежуточных.

Примерами использования в психологии порядковых шкал могут служить первичные результаты тестовых испытаний группы лиц, первичные результаты при использовании некоторых личностных опросников, работы со шкалами самооценки и т. п. Можно сказать, что результаты большинства психологических исследований представляют собой ординальные величины, т. е. выражающиеся порядковыми числами. Об этом необходимо помнить, поскольку характер первичных результатов накладывает ряд ограничений на возможность использования тех или других статистических приемов их обработки и анализа. Поскольку в порядковых шкалах не определена единая точка отсчета величин, то и для их элементов, как и для элементов шкал наименований, непригодны способы расчета, требующие арифметических действий, — в частности сложение и вычитание. В качестве меры положения классов объектов для преобразовании шкал порядка кроме моды *(Mo)* могут быть использованы еще и медиана *(Me),* полуквартильные отклонения (Q_1 и Q_2), а в качестве меры тесноты взаимосвязи классов — коэффициент ранговой корреляции Ч. Спирмена.

Шкалы интервалов. Когда шкала обладает всеми свойствам к порядковой шкалы и дополнительно к этому определены еще расстояния между ее единицами, то такую шкалу называют шкалой интервалов. Иначе говоря, классы объектов шкал интервалов всегда дискретны и упорядочены по степени

возрастания (или убывания) измеряемого свойства. Кроме того, в эти шкалах одинаковым разностям степени выраженности измеряемого свойства соответствуют равные разности между приписываемыми им числам. Шкалы интервалов имеют равные единицы измерения, однако способ их определения является произвольным, следовательно, и сами единицы произвольны. При этом неизвестна абсолютная величина отдельных значений по шкале, поскольку шкала интервалов не имеет естественной нулевой точки отсчета. Последняя может быть произвольно смещена. Шкалам интервалов присущи все те отношения, которые характерны для номинативных и порядковых шкал. Кроме того, для них возможно использование арифметических действий. Основными операциями с элементами интервальных шкал являются операции установления равенства, разностей, сопоставление больше — меньше в отношении измеряемых свойств, а также утверждение равенства интервалов и равенства разностей между значениями одной шкалы. Наряду со всеми ранее указанными свойствами номинативных и порядковых шкал шкалы интервалов подчиняются еще и следующим постулатам сложения:
a+Ь=Ь+a и (a+Ь)+c=a+(Ь+c), если a=p и Ь>0, то a+Ь>p, если a=p и Ь=q, то a+Ь=p+q.

С интервальными шкалами допускаются, следовательно, любые линейные преобразования типа *x'=ax+Ь* для *a>0,* при которых сохраняется только последовательность градаций измеряемого свойства объектов, но и величина относительных расстояний между классами объектов. Возможность смещения точки отсчета отражена в константе *Ь,* а величина единиц шкалы связана с константой *a.*

Хотя психологические измерения дают нам преимущественно ординальные величины, их обработка часто осуществляется с помощью приемов, допустимых на уровне интервальных шкал. То есть большинство исследователей исходят из равенства интервалов между полученными при измерении величинами. Такой подход основывается чаще всего на следующих предпосылках: во-первых, что измеряемая переменная (то или иное свойство объектов) в генеральной совокупности имеет нормальное распределение, и, вторых, что различные показатели одной и той же переменной обнаруживают линейную корреляцию. Действительно, на основании этого можно допустить, что интервалы в шкале равны, так как чем более линейна зависимость, тем более равными должны быть интервалы в шкале. Итак, при конструировании шкалы интервалов используют три произвольные операции: установление величин единиц измерения, определение нулевой точки и определение направления, в котором ведут отсчет по отношению к нулевой точке.

Благодаря равенству единиц на уровне шкал интервалов возможна характеристика формы распределения эмпирических величин с помощью стандартных статистических показателей: средней арифметической величины, среднего квадратичного отклонения, показателей асимметрии и эксцесса. Использование линейных преобразований приводит к изменению лишь средней арифметической и/или среднего квадратичного отклонения, не меняя показателей симметрии и эксцесса. Наиболее частыми линейными преобразованиями, которые находят применение, как в области психометрии, так и в области психофизики, являются центрирование и нормирование результатов измерении. Под центрированием понимается такое линейное преобразование, при котором средняя арифметическая величина становится равной нулю, в то время как направление шкалы и величина ее единиц остаются неизменными. Под нормированием понимают такое линейное преобразование результатов измерения, при котором их средняя арифметическая величина становится равной нулю, а среднее квадратичное отклонение равным ± 1. Из сказанного, очевидно, что для обработки и анализа эмпирических данных, полученных на уровне шкал интервалов, допустимы любые приемы статистической обработки, а именно расчет основных характеристик распределения, а также меры взаимосвязи количественных переменных (коэффициентов корреляции). В случае наличия нормальных распределений первичных результатов для их сравнения можно применить также все известные критерии оценки значимости различий, как между значениями их средних величин, так и дисперсии, т. е. размаха распределения

Примером интервальных шкал, используемых в психологии, являются стандартизованные тестовые шкалы психодиагностики: шкалы Векслера, шкалы Тёрстена и шкала Т. Гилфорда.

Шкалы отношений. Конструирование шкал отношений предполагает наряду с наличием свойств предыдущих шкал существование постоянной естественной нулевой точки отсчета, в которой измеряемый признак полностью отсутствует. Следовательно, шкалы отношений характеризуются тем, что в них, во-первых, классы объектов разделены и упорядочены согласно измеряемому свойству, во-вторых, равным разностям между классами объектов соответствуют равные разности между приписываемыми им числами, в-третьих, числа, приравниваемые классам объектов, пропорциональны степени выраженности измеряемого свойства. Последнее не было свойственно рассмотренным выше шкалам.

Основными операциями, допустимыми на уровне шкал отношений, являются все те операции, которым подчиняются шкалы всех перечисленных выше типов, и дополнительно — операции установления равенства отношений между отдельными значениями шкалы. Это возможно благодаря существованию на шкале естественного, абсолютного, нуля. Поэтому лишь для данной шкалы числа, являющиеся точками (значениями) на шкале, соответствуют реальному количеству измеряемого свойства, что позволяет производить с ними любые арифметические действия — оперирование суммами, произведениями и частными. Для шкал отношений допустимы любые мультипликативные преобразования. Однако недопустимы (об этом часто забывают!) никакие операции прибавления или вычитания константных величин, что приводит, как было показано на примере шкал интервалов, к сдвигу точки отсчета. Дополнительно к указанным для описанных выше шкал измерения приемам статистической обработки данных для величин шкалы отношений можно рассчитывать, например, геометрические и гармонические средние, а также коэффициенты изменчивости измеряемого признака.

Считалось, что шкалы отношений не встречаются в психологических измерениях. Однако Стивенс, исходя из постулата о допустимости непосредственного измерения психических процессов, показал возможность построения шкал отношений в психофизике. Для этой шкалы он разработал ряд измерительных процедур, предусматривающих прямое шкалирование. Среди них наиболее известными стали методики фракционирования и мультипликации предъявляемых стимулов. К этой же группе методик можно отнести и методики оценки величин стимулов и непосредственной оценки их отношений. Общим для всех перечисленных методик прямого шкалирования является то, что в качестве измерительного инструмента выступает сам испытуемый, который оценивает количественные отношения между раздражителями.

Математическая статистика.

Критическая статистика. Понятие статистической оценки параметров. Характеристики оценок. Статистические гипотезы и их проверка. Интервал правдоподобия и критическая область. Уровень значимости. Ошибки первого и второго рода.

В процессе статистического анализа иногда бывает необходимо сформулировать и проверить предположения относительно величины независимых параметров или закона распределения изучаемой генеральной совокупности (совокупностей). Такие предположения называются *статистическими гипотезами.*

Статистические гипотезы подразделяются на нулевые и альтернативные.

Выдвинутая гипотеза называется нулевой (основной). Ее принято обозначать H_0. Обычно нулевая гипотеза – это гипотеза об отсутствии различий.

По отношению к высказанной нулевой гипотезе всегда можно сформулировать *альтернативную (конкурирующую),* противоречащую ей. Альтернативную гипотезу принято обозначать H_1 .

Цель статистической проверки гипотез состоит в том, чтобы на основании выборочных данных принять решение о справедливости нулевой гипотезы H_0.

Так как проверка статистических гипотез осуществляется на основании выборочных данных, то такое решение неизбежно сопровождается некоторой, хотя возможно и очень малой, ошибкой.

Ошибка, состоящая в том, что мы отклонили нулевую гипотезу, в то время как она верна, называется *ошибкой I рода*, а ее вероятность – *уровнем значимости* α.

Ошибка, состоящая в том, что мы приняли нулевую гипотезу, в то время как она неверна, называется *ошибкой II рода,* а ее вероятность обозначают β. Величину равную 1 –β называют *мощностью критерия.*

Мощность критерия определяется эмпирическим путем, а уровень значимости задается исследователем. В психологических и социологических исследованиях низшим уровнем значимости принято считать α.=0,05 а достаточным α= 0,01.

Статистический критерий — это правило (формула), по которому определяется мера расхождения результатов выборочного наблюдения с высказанной гипотезой H_0.

Значение критерия, рассчитываемое по специальным правилам на основании выборочных данных, называется *наблюдаемым* значением критерия.

Значения критерия, определяемые на заданном уровне значимости α по таблицам распределения случайной величины, выбранной в качестве критерия, называются *критическими точками*.

В психолого-педагогических и социологических исследованиях принято определять значения критерия при α=0,01 и α=0,05. Полученные критические точки делят совокупность значений критерия на область допустимых значений или зону незначимости (область принятия нулевой гипотезы), критическую область или зону значимости (область принятия альтернативной гипотезы) и зону неопределенности.

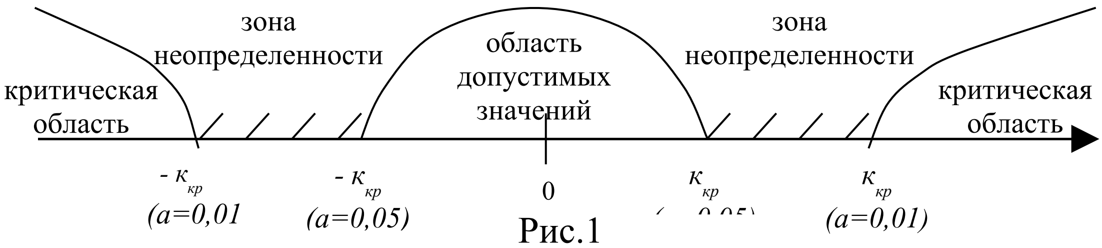

Рис.1

Чаще всего критерии, используемые при психологических и социологических исследованиях, имеют положительные значения. Поэтому для простоты при решении прикладных задач изображают только неотрицательную часть оси значимости (рис.2).

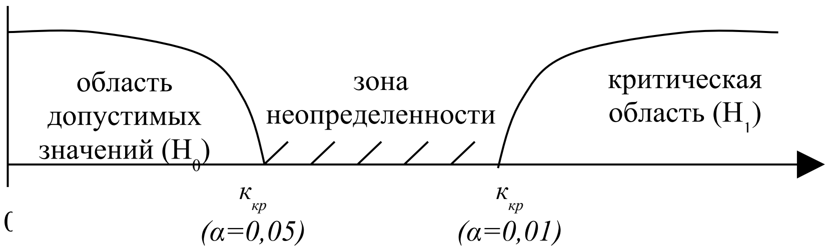

Рис.2

Основной принцип проверки статистических гипотез состоит в следующем:

- если наблюдаемое значение критерия принадлежит критической области, то нулевая гипотеза H_0 отклоняется и принимается конкурирующая H_1;

- если наблюдаемое значение критерия принадлежит области допустимых значений, то нулевую гипотезу H_0 нельзя отклонить;

-если наблюдаемое значение критерия принадлежит зоне неопределенности, то мы уже можем отклонить нулевую гипотезу H_0, но еще не можем принять конкурирующую H_1.

Критерии делятся на параметрические – включающие в формулу расчета параметры распределения (средние и дисперсии) и непараметрические – основанные на оперировании частотами или рангами.

Параметрические критерии. Критерий Стьюдента, критерий Фишера.

1. t-критерий Стьюдента используется

a) для сравнения выборочной средней $\overline{x}$ с некоторым известным числовым значением a_0.

Возможны гипотезы:

H_0: $\overline{x} = a_0$ выборочная средняя генеральной совокупности равна заданному числу a_0.

H_1: $\overline{x} \neq a_0$ ($\overline{x} < a_0$, $\overline{x} > a_0$) выборочная средняя генеральной совокупности не равна (меньше, больше) заданному числу a_0.

Наблюдаемое значение t-критерия рассчитывается по формуле:

– если дисперсия генеральной совокупности неизвестна

$$t_{набл} = \frac{\overline{x} - a_0}{S}\sqrt{n} \quad (1)$$

– если дисперсия генеральной совокупности известна

$$t_{набл} = \frac{\overline{x} - a_0}{\sigma_{ген}}\sqrt{n} \quad (2)$$

где $\overline{x}$ - выборочная средняя;

a_0 — числовое значение генеральной средней;

S — исправленное среднее квадратическое отклонение;

$\sigma^2_{ген}$ - известная дисперсия генеральной совокупности;

n — объем выборки.

Критическое значение $t_{кр}$ следует находить с помощью таблиц распределения Стьюдента по уровню значимости α и числу степеней свободы $k = n - 1$.

b) для обнаружения различия между средними значениями $\overline{x}$, $\overline{y}$ двух выборок.

Возможны гипотезы:

H_0: $\overline{x} = \overline{y}$ средние значения двух выборок равны,

H_1: $\overline{x} \neq \overline{y}$ средние значения двух выборок не равны.

Наблюдаемое значение t-критерия рассчитывается по формуле:

– для независимых выборок

$$t_{набл} = \frac{\bar{x} - \bar{y}}{\sqrt{(n_1 - 1)S_x^2 + (n_2 - 1)S_y^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \quad (3)$$

– для зависимых выборок

$$t_{эмп} = \frac{\sum d}{\sqrt{\frac{n\sum d^2 - (\sum d)^2}{n - 1}}} \quad (4)$$

где $S_x^2 = \frac{1}{n_1 - 1}\sum (x - \bar{x})^2$ - выборочная дисперсия 1 выборки;

$S_y^2 = \frac{1}{n_2 - 1}\sum (y - \bar{y})^2$ - выборочная дисперсия 2 выборки;

$\bar{x}$ - среднее значение признака для 1 выборки;

$\bar{y}$ - среднее значение признака для 2 выборки;

n_1 – объем 1 выборки;

n_2 – объем 2 выборки;

d — разность между результатами в каждой паре («*после»минус «до»*);

n —число пар данных в зависимых выборках

Критическое значение $t_{кр}$ следует находить с помощью таблиц распределения Стьюдента по уровню значимости α и числу степеней свободы $k = n_1+n_2-2$.

t-критерий для независимых выборок можно использовать для сравнения средних показателей экспериментальной группы с контрольной группой.

t-критерий для зависимых выборок очень полезен в тех ситуациях, когда две сравниваемые группы основываются на одной и той же совокупности наблюдений (субъектов), которые тестировались *дважды* (например, *до* и *после* эксперимента).

2. *F* — критерий Фишера-Снедекора используют

а) для сравнения разброса значений двух выборок, т.е. для проверки гипотезы о равенстве дисперсий.

Возможны гипотезы:

H_0 : $S_x^2 = S_y^2$ - разброс значений признака относительно среднего одинаковый в обеих выборках.

H_1: $S_x^2 \neq S_y^2$ - разброс значений признака не совпадает.

Наблюдаемое значение F — критерия рассчитывается по формуле:

$$F_{набл} = \frac{S_x^2}{S_y^2}, \qquad (5)$$

где S_x^2 — большая (по величине) выборочная дисперсия;

S_y^2 — меньшая (по величине) выборочная дисперсия.

Критическое значение $F_{крит}$ следует находить с помощью таблицы распределения Фишера-Снедекора по уровню значимости α и числу степеней свободы $k_1 = n_1 – 1$ и $k_2 = n_2 - 1$,

где k_1 — число степеней свободы большей (по величине) дисперсии;

k_2 — число степеней свободы меньшей (по величине) дисперсии;

n_1 — объем выборки большей (по величине) дисперсии;

n_2 — объем выборки меньшей (по величине) дисперсии.

b) для выявлений тенденций изменения признака в трех и более выборках при переходе от условия к условию (однофакторный дисперсионный анализ).

Возможны гипотезы:

для независимых выборок: (влияние разных условий на разных испытуемых)

H_0 : разные условия не влияют на изменение значений признака;

H_1: условия влияют на изменение значений признака.

для зависимых выборок (одни и те же испытуемые, но в разных условиях) возможно два варианта гипотез:

а) H_0: условия не влияют на изменение признака;

H_1: условия влияют на изменение значений признака.

б) H_0: индивидуальные различия испытуемых не влияют на изменение значений признака;

H_1: индивидуальные различия между испытуемыми влияют на изменение значений признака.

Наблюдаемое значение критерия рассчитывается по формулам:

для независимых выборок $F_{набл} = \frac{S^2_{факт}}{S^2_{сл}}$, (6)

для зависимых выборок а) $F_{набл} = \frac{S^2_{факт}}{S^2_{сл}}$, (7)

б) $F_{набл} = \frac{S^2_{испыт}}{S^2_{сл}}$, (8)

где $S^2_{факт} = \frac{n \cdot \sum \left(\bar{x}_{гр} - \bar{x}\right)^2}{к_{факт}}$ - характеризует изменение признака, обусловленное действием фактора (условия).

$S^2_{общ} = \frac{\sum \left(x_{ij} - \bar{x}\right)^2}{к_{общ}}$ - характеризует общее изменение признака по всем выборкам и наблюдениям.

$S^2_{испыт} = \frac{m \cdot \sum \left(\bar{x}_{испыт} - \bar{x}\right)^2}{к_{испыт}}$ - характеризует изменение признака, обусловленное индивидуальными особенностями испытуемых.

$S^2_{сл} = \frac{\sum \left(x_{ij} - \bar{x}\right)^2 - n \cdot \sum \left(\bar{x}_{гр} - \bar{x}\right)^2}{к_{общ} - к_{факт}}$ - для независимых выборок

$S^2_{сл} = \frac{\sum \left(x_{ij} - \bar{x}\right)^2 - n \cdot \sum \left(\bar{x}_{гр} - \bar{x}\right)^2 - m \cdot \sum \left(\bar{x}_{испыт} - \bar{x}\right)^2}{к_{общ} - к_{факт} - к_{испыт}}$ - для зависимых выборок.

n – количество наблюдений,

m – количество выборок (групп),

$\bar{x}_{гр}$ - среднее значение признака в каждой группе,

$\bar{x}$ - среднее значение признака по всей совокупности,

$\bar{x}_{испыт}$ - среднее значение признака каждого испытуемого,

$к$ – число степеней свободы:

$к_{общ} = n \cdot m - 1$

$к_{факт} = m - 1$

$к_{испыт} = n - 1$

Критические значения $F_{крит}$ следует находить с помощью таблицы распределения Фишера-Снедекора по уровню значимости α и числу степеней свободы и $k_2 = k_{сл}$, и в зависимости от проверяемых гипотез $k_1 = k_{факт}$ или $k_1 = k_{испыт}$

Основы корреляционного анализа. Коэффициент корреляции r Пирсона и его свойства как меры связи. Ранговая корреляция, коэффициенты ρ Спирмена и τ Кендалла.

Исследователя нередко интересует, как связаны между собой две или более переменных в одной или нескольких изучаемых группах.

Чаще всего признак определяется воздействием многих факторов: средовых, генетических, социальных, экологических и т.д. Поэтому связи между психологическими или социологическими признаками имеют не функциональный, а статистический характер, когда в среднем определенному значению одного признака соответствует не одно какое-либо значение, а целый спектр, распределяющихся в вариационный ряд числовых значений. Такого рода зависимость между случайными величинами называется корреляционной, или корреляцией.

Задача корреляционного анализа сводится к установлению направления и формы между варьирующими признаками, измерению тесноты, и, наконец, к проверке значимости коэффициентов корреляции.

Переменные *x* и *y* могут быть измерены в разных шкалах. Это обстоятельство определяет выбор соответствующего коэффициента линейной корреляции.

Тип шкалы		Мера связи
Переменная *x*	Переменная *y*	
Интервальная или отношений	Интервальная или отношений	Коэффициент Пирсона r_{xy}
Ранговая, интервальная или отношений	Ранговая, интервальная или отношений	Коэффициент Спирмена g_{xy}
Ранговая	Ранговая	Коэффициент τ Кендалла
Дихотомическая	Дихотомическая	1 Коэффициент ассоциации φ
Дихотомическая	Ранговая	Рангово-бисериальный $R_{гв}$
Дихотомическая	Интервальная или отношений	Бисериальный $R_{бис}$

Все коэффициенты по абсолютной величине не могут превосходить 1.

а) Коэффициент корреляции Пирсона вычисляется по формуле:

$$r_{xy} = \frac{\sum_i (x_i - \bar{x}) * (y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 * \sum_i (y_i - \bar{y})^2}} = \frac{n * \sum_i x_i y_i - \sum_i x_i * \sum_i y_i}{\sqrt{(n * \sum_i x_i^2 - (\sum_i x_i)^2) * (n * \sum_i y_i^2 - (\sum_i y_i)^2)}} \quad (9)$$

где x_i – значения, переменных принимаемые переменной x,

y_i - значения, переменных принимаемые переменной y,

$\bar{x}$ – средняя по x,

$\bar{y}$ - средняя по y.

Оценка значимости осуществляется при числе степеней свободы $k=n-2$.

б) Коэффициент корреляции рангов Спирмена вычисляется по формуле:

$$\rho = 1 - \frac{6 \cdot \sum_i (d_i)^2}{n \cdot (n^2 - 1)}, \quad (10)$$

где n – количество ранжируемых признаков

d_i – разность между рангами по двум переменным для каждого испытуемого.

При наличии одинаковых рангов в числитель добавляются поправки на одинаковые ранги:

$$D_1 = \frac{n^3 - n}{12}, \quad (11)$$

$$D_2 = \frac{k^3 - k}{12}, \quad (12)$$

где n – число одинаковых рангов в первом столбце,

k – число одинаковых рангов во втором столбце.

По каждой группе одинаковых рангов вводится своя поправка.

Критически значения определяются при уровне значимости равном числу значений признака, по таблице критических значений ρ Спримена.

в) Коэффициент ассоциации φ вычисляется по формуле:

$$\varphi = \frac{p_{xy} - p_x \cdot p_y}{\sqrt{p_x \cdot (1 - p_x) \cdot p_y \cdot (1 - p_y)}}, \quad (13)$$

где p_x – частота или доля признака, имеющего 1 по х, $(1\text{-}p_x)$ – частота или доля признака, имеющего 0 по х, p_y – частота или доля признака, имеющего 1 по у, $(1\text{-}p_y)$ – частота или доля признака, имеющего 0 по у, p_{xy} – частота или доля признака, имеющих 1 и по х и по у.

г) Коэффициент корреляции τ Кендалла вычисляется по формуле:

$$\tau = 1 - \frac{4 \cdot Q}{N \cdot (N - 1)}, \tag{14}$$

где Q – число инверсий (подсчет инверсий осуществляется суммированием числа рангов второго признака меньше каждого из рангов второго признака, при условии, что ранги первого признака упорядочены по возрастанию);

N – число ранжируемых признаков.

д) Бисериальный коэффициент корреляции вычисляется по формуле:

$$R_{бис\ эмп} = \frac{\bar{x}_1 - \bar{x}_0}{\sigma_y} * \sqrt{\frac{n_1 * n_0}{N * (N - 1)}}, \tag{15}$$

где $\bar{x}_1$ - среднее по тем элементам переменной y, которым соответствует признак 1 в переменной x,

$\bar{x}_0$ - среднее по тем элементам переменной y, которым соответствует признак 0 в переменной x,

n_1 – число единиц в переменной x,

n_0 – число нулей в переменной x,

$N = n_1 + n_0$,

σ_y - среднее квадратическое отклонение переменной y.

е) Рангово-бисернальный коэффициент корреляции вычисляется по формуле:

$$R_{rb\ эмп} = \frac{(\bar{x}_1 - \bar{x}_0)^2}{N}, \tag{16}$$

где $\bar{x}_1$ - средний ранг по тем элементам переменной y, которым соответствует признак 1 в переменной x;

$\bar{x}_0$ - средний ранг по тем элементам переменной y, которым соответствует признак 0 в переменной x;

N – количество элементов в переменной x.

Для измерения нелинейной корреляционной связи между признаками используется корреляционное отношение.

Показатели корреляционного отношения вычисляется по формулам:

$$h_{yx} = \sqrt{\frac{\sum f_x (\bar{y}_x - \bar{y})^2}{\sum f_y (y_i - \bar{y})^2}}, \qquad (17)$$

$$h_{xy} = \sqrt{\frac{\sum f_y (\bar{x}_y - \bar{x})^2}{\sum f_x (x_i - \bar{x})^2}}, \qquad (18)$$

где $\bar{x}$ и $\bar{y}$ общие, а $\bar{x}_y$ и $\bar{y}_x$ - групповые средние арифметические, f_y и f_x частоты рядов *X* и *У*.

Для вычисления корреляционного отношения h_{xy} (*У* по *X*) или h_{xy} (*X* по *У*) необходимо:

1) расположить по порядку исходные данные по *X* от меньшей к большей, при этом сохранив значения соответствующих величин *У* по отношению к *X*;
2) определить частоты переменной $X(f_x)$;
3) подсчитать арифметическое средние по переменной *У* для соответствующей частоты $(\bar{y}_x)$;
4) расположить по порядку исходные данные по *У* от меньшей величине к большей, при этом сохранив значения соответствующих величин *X* по отношению к *У*;
5) определить частоты переменной $У(f_y)$;
6) подсчитать арифметические средние по переменной *X* для соответствующей частоты $(\bar{x}_y)$;
7) определить $\bar{x}$ и $\bar{y}$;
8) произвести расчет по формулам (17) и (18);
9) определить значимость полученных показателей.

Для расчета уровней значимости коэффициентов корреляции проверяются гипотезы:

H_0: коэффициент корреляции между признаками статистически значимо не отличается от нуля;

H_1: коэффициент корреляции между признаками статистически значимо отличается от нуля.

Все коэффициенты корреляции, не имеющие стандартных таблиц для нахождения критических значений, оценивают с помощью t – критерия Стьюдента по формуле:

$$t_{эмп} = |r_{эмп}| * \sqrt{\frac{n-2}{1-r^2_{эмп}}}, \qquad (19)$$

где $r_{эмп}$ – соответствующий коэффициент корреляции (корреляционное отношение),

n – число коррелируемых значений.

Критические значения $t_{кр}$ определяется по таблице значений для t – критерия Стьюдента. Число степеней свободы будет равно $k=n-2$.

Основы регрессионного анализа. Линейная и нелинейная регрессия, и оценка ее качества. Множественная регрессия.

Взаимосвязь между переменными величинами может быть описана разными способами. Например, эту связь можно описать с помощью различных коэффициентов корреляции (линейных, частных, корреляционного отношения и т. п.). В то же время эту связь можно выразить по-другому: как зависимость между аргументом (величиной *X*) и функцией *Y*. В этом случае задача будет состоять в нахождении зависимости вида $X=F(Y)$ или, напротив, в нахождении зависимости вида $X=F(Y)$. При этом изменение функции в зависимости от изменений одного или нескольких аргументов называется регрессией.

Графическое выражение регрессионного уравнения называют линией регрессии. Линия регрессии выражает наилучшее предсказание зависимой переменной (*Y*) по независимым переменным (*X*). Эти независимые переменные, а их может быть много, носят название *предикторов*.

Регрессию выражают с помощью двух уравнений регрессии, которые в самом простом случае выглядят, как уравнения прямой, а именно так:

$$Y = a0 + a1 \cdot X \qquad (1)$$

$$X = b0 + b1 \cdot Y \qquad (2)$$

В уравнении 1 Y – зависимая переменная, а X – независимая переменная, $a0$ свободный член, а $a1$ – коэффициент регрессии, или угловой коэффициент, определяющий наклон линии регрессии по отношению к осям координат.

В уравнении 2 X – зависимая переменная, а Y – независимая переменная, $b0$ свободный член, а $b1$ – коэффициент регрессии, или угловой коэффициент, определяющий наклон линии регрессии по отношению к осям координат.

Линии регрессии пересекаются в точке O ($\bar{x}$, $\bar{y}$), с координатами, соответствующими средним арифметическим значениям корреляционно связанных между собой переменных X и Y. Линия AB, проходящая через точку O, соответствует линейной функциональной зависимости между переменными величинами X и Y равен $r_{x\,y}$= 1. При этом наблюдается такая закономерность: чем сильнее связь между X и Y , тем ближе обе линии регрессии к прямой AB. При отсутствии связи между X и Y линии регрессии оказываются под прямым углом по отношению друг к другу и в этом случае r_{xy}=0.

Количественное представление связи (зависимости) между X и Y (между Y и X) называется регрессионным анализом. Главная задача регрессионного анализа заключается , собственно говоря, в нахождении коэффициентов $a0$, $b0$, $a1$ и $b1$ и определении уровня значимости полученных аналитических выражений (1) и (2), связывающих между собой переменные X и Y .

При этом коэффициенты регрессии $a1$ и $b1$ показывают, насколько в среднем величина одной переменной изменяется при изменении на единицу меры другой. Коэффициент регрессии $a1$ в уравнении (1) можно подсчитать по формуле:

$$a1 = r_{xy} \cdot \frac{S_y}{S_x} \qquad (3)$$

а коэффициент $b1$ в уравнении (2) по формуле (4)

$$b1 = r_{xy} \cdot \frac{S_x}{S_y} \qquad (4)$$

где r_{yx} – коэффициент корреляции между переменными X и Y;

S_x – среднеквадратическое отклонение, подсчитанное для переменной X;

S_y – среднеквадратическое отклонение, подсчитанное для переменной Y.

Коэффициенты регрессии можно вычислить также без подсчёта среднеквадратических отклонений по следующим формулам:

$$a1 = r_{xy} \cdot \sqrt{\frac{\sum (y_i - \bar{y})^2}{\sum (x_i - \bar{x})^2}} \qquad (5)$$

$$b1 = r_{xy} \cdot \sqrt{\frac{\sum (x_i - \bar{x})^2}{\sum (y_i - \bar{y})^2}} \qquad (6)$$

В том случае, если неизвестен коэффициент корреляции, коэффициенты регрессии можно вычислить по следующим формулам:

$$a1 = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \qquad (7)$$

$$b1 = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum (y_i - \bar{y})^2} \qquad (8)$$

Величины $a1$, $b1$ и r_{xy} взаимосвязаны. Более того, зная две из них – всегда можно получит третью. Например, зная величины $a1$ и $b1$ можно легко получить r_{xy}:

$$r_{xy} = \sqrt{a1 \cdot b1} \qquad (9)$$

Формула (9) достаточно очевидна, поскольку, умножив $a1$, вычисленный по формуле (3) на $b1$, вычисленный по формуле (24), получим:

$$r_{xy} \cdot \frac{S_y}{S_x} \cdot r_{xy} \cdot \frac{S_x}{S_y} = r_{xy} \cdot r_{xy}$$

Формула (9) очень важна, поскольку она позволяет по известным значениям коэффициентов регрессии $a1$ и $b1$ определить коэффициент корреляции, и можно проверить правильность расчёта коэффициента корреляции. Как и коэффициент корреляции, коэффициенты регрессии характеризуют только линейную связь и при положительной связи имеют знак плюс, при отрицательной – знак минус.

В свою очередь свободные члены $a0$ и $b0$ в уравнениях регрессии придётся вычислять по следующим формулам. Для подсчёта свободного члена $a0$ уравнения регрессии (1) используется формула:

$$a0 = \frac{\sum y_i \cdot \sum x_i^2 - \sum x_i \cdot \sum x_i y_i}{\sum x_i^2 - \sum (x_i)^2} \qquad (10)$$

Для подсчёта свободного члена $b0$ уравнения регрессии (2) используется формула:

$$b0 = \frac{\sum x_i \cdot \sum y_i^2 - \sum y_i \cdot \sum x_i y_i}{\sum y_i^2 - \sum (y_i)^2} \qquad (11)$$

Вычисления по формулам (7), (8), (10) и (11) достаточно сложны, поэтому при расчётах коэффициентов регрессии используют, как правило, более простой метод. Он заключается в решении двух систем уравнений. При решении одной системы находятся величины $a0$ и $a1$, и при решении другой – $b0$ и $b1$.

Общий вид системы уравнений для нахождения величин $a0$ и $a1$ таков:

$$\begin{cases} a0 \cdot N + a1 \cdot \sum x_i = \sum y_i \\ a0 \cdot \sum x_i + a1 \cdot \sum (x_i \cdot x_i) = \sum y_i \cdot x_i \end{cases} \qquad (12)$$

Общий вид системы уравнений для нахождения величин $b0$ и $b1$ таков:

$$\begin{cases} a0 \cdot N + a1 \cdot \sum x_i = \sum y_i \\ a0 \cdot \sum x_i + a1 \cdot \sum (x_i \cdot x_i) = \sum y_i \cdot x_i \end{cases} \qquad (13)$$

В системах уравнений (12) и (13) используются следующие обозначения:

N – число элементов в переменной *X* или переменной *Y*,

$\sum x_i$ – сумма всех элементов переменной *X*,

$\sum y_i$ – сумма всех элементов переменной *Y*,

$\sum (y_i \cdot y_i)$ – произведение всех элементов переменной *Y* друг на друга,

$\sum (x_i \cdot x_i)$ – произведение всех элементов переменной *X* друг на друга,

$\sum (y_i \cdot x_i)$ – попарное произведение всех элементов переменной *X* на соответствующие элементы переменной *Y.*

Для применения метода линейного регрессионного анализа необходимо соблюдать следующие условия:

1. Сравниваемые переменные *X* и *Y* должны быть измерены в шкале интервалов или отношений.

2. Предполагается, что переменные X и Y имеют нормальный закон распределения.
3. Число варьирующих признаков в сравниваемых переменных должно быть одинаковым.

Множественная линейная регрессия

Предположим, что психолог при анализе успешности обучения подростков в дополнение к независимой переменной IQ рассматривает другие независимые переменные, влияющие, по его мнению, на успеваемость, например такие, как мотивация, личностные особенности и т.п. В этом случае можно построить линейное уравнение множественной регрессии, в которое будут входить все вышеназванные переменные. В общем случае, зависимость между несколькими переменными величинами выражают уравнением множественной регрессии, которая может быть как линейной, так и не линейной. В простейшем случае множественная линейная регрессия выражается уравнением с двумя независимыми переменными величинами X и Y и имеет вид (20). Y в данном случае является зависимой переменной.

$$Y = a + b \cdot X + c \cdot Z \qquad (20)$$

Где a – свободный член, b и c – параметры уравнения (20).

Уравнение (20) может решаться относительно зависимой переменной Z, тогда X и Y являются независимыми переменными и уравнение множественной регрессии имеет следующий вид:

$$Z = a + b \cdot X + c \cdot Y \qquad (21)$$

Можно решить уравнение (20) и относительно X, тогда Z и Y будут независимыми переменными, а уравнение будет иметь вид:

$$X = a + b \cdot Y + c \cdot Z \qquad (22)$$

При проведении конкретных расчётов выбор зависимых и независимых переменных определяется планом эксперимента.

Решение уравнений (20), (21) и (22) состоит в том, что находятся величины a, b и c на основе решения системы из трёх уравнений:

Для решения уравнения (20) система имеет следующий вид:

$$\begin{cases} a\cdot N + b\cdot\sum x_i + c\cdot\sum z_i = \sum y_i \\ a\cdot\sum x_i + b\cdot\sum (x_i\cdot x_i) + c\cdot\sum x_i\cdot z_i = \sum x_i\cdot y_i \\ a\cdot\sum z_i + b\cdot\sum (x_i\cdot z_i) + c\cdot\sum (z_i\cdot z_i) = \sum y_i\cdot z_i \end{cases} \quad (23)$$

Для решения уравнения (21) система будет выглядеть следующим образом:

$$\begin{cases} a\cdot N + b\cdot\sum x_i + c\cdot\sum y_i = \sum z_i \\ a\cdot\sum x_i + b\cdot\sum (x_i\cdot x_i) + c\cdot\sum y_i\cdot x_i = \sum x_i\cdot z_i \\ a\cdot\sum y_i + b\cdot\sum (x_i\cdot y_i) + c\cdot\sum (y_i\cdot y_i) = \sum y_i\cdot z_i \end{cases} \quad (24)$$

Для решения уравнения (22) система будет иметь следующий вид:

$$\begin{cases} a\cdot N + b\cdot\sum y_i + c\cdot\sum z_i = \sum x_i \\ a\cdot\sum y_i + b\cdot\sum (y_i\cdot y_i) + c\cdot\sum y_i\cdot z_i = \sum x_i\cdot y_i \\ a\cdot\sum z_i + b\cdot\sum (x_i\cdot z_i) + c\cdot\sum (z_i\cdot z_i) = \sum x_i\cdot z_i \end{cases} \quad (25)$$

В общем случае уравнение регрессии представляет собой сложный полином, описывающий зависимость сразу между несколькими переменными. Такое уравнение множественной регрессии имеет вид:

$$Y = b_0 + b_1\cdot X_1 + b_2\cdot X_2 + \ldots + b_p\cdot X_p \quad (26)$$

Где X_1, X_2, X_3 и т.п. – интересующие психолога независимые переменные, а Y – зависимая переменная.

Для применения метода множественной линейной регрессии необходимо соблюдать следующие условия:

1. Сравниваемые переменные должны быть измерены в шкале интервалов или отношений.
2. Предполагается, что все переменные имеют нормальный закон распределения.
3. Число варьирующих признаков в сравниваемых переменных должно быть одинаковым.

Непараметрические критерии. Критерии различий, изменений. Критерии согласия. X^2 и теорема Пирсона. Критерий однородности.

Непараметрические критерии не включают в формулу расчета параметров распределения, основаны на оперировании частотами или рангами.

1. Критерии различий

а) Q-критерий Розенбаума

Критерий используется для оценки различий между двумя выборками по уровню какого-либо признака, количественно измеренного.

Возможны гипотезы:

H_0: Уровень признака в выборке 1 не превышает уровня признака в выборке 2.

H_1: Уровень признака в выборке 1 превышает уровень признака в выборке 2.

Ограничения критерия Q

1) В каждой из сопоставляемых выборок должно быть не менее 11 наблюдений. При этом объемы выборок должны примерно совпадать. Если в обеих выборках меньше 50 наблюдений, то абсолютная величина разности между объемами выборок n_1 и n_2, соответственно, не должна быть больше 10 наблюдений. Если в каждой из выборок больше 51 наблюдения, но меньше 100, то абсолютная величина разности между объемами выборок n_1 и n_2, соответственно, не должна быть больше 20 наблюдений. Если в каждой из выборок больше 100 наблюдений, то допускается, чтобы одна из выборок была больше другой не более чем в 1,5-2 раза.
2) Диапазоны разброса значений в двух выборках должны не совпадать между собой, в противном случае применение критерия бессмысленно.
3) Измерение может быть проведено по шкале порядка, интервалов или отношений.
4) Выборки должны быть независимыми.

Эмпирическое значение критерия подсчитывается по формуле:

$$Q_{эмп} = S_1 + S_2, \tag{20}$$

где S_1 – количество наблюдений в выборке 1, которые выше максимального значения в выборке 2;

S_2 - количество наблюдений в выборке 2, которые ниже минимального значения выборки 1.

Значения в выборках должны быть упорядочены по возрастанию признака.

Критические значения Q-критерия определяются по таблице для данных n_1 и n_2 и для выбранного уровня значимости. Если $Q_{эмп}$ не меньше $Q_{кр}$, то H_o отвергается.

б) U- критерий Манна-Уитни

Критерий предназначен для оценки различий между двумя выборками по уровню какого-либо признака, количественно измеренного.

Возможны гипотезы:

H_0: Уровень признака в группе 2 не ниже уровня признака в группе 1.

H_1: Уровень признака в группе 2 ниже уровня признака в группе 1.

Ограничения критерия U

1) В каждой выборке должно быть не менее 3 наблюдений. Допускается, чтобы в одной выборке было 2 наблюдения, но тогда во второй их должно быть не менее 5.
2) В каждой выборке должно быть не более 60 наблюдений.
3) Измерение может быть проведено по шкале интервалов или отношений.
4) Выборки должны быть несвязными.

Наблюдения обеих выборок необходимо объединить и проранжировать по степени нарастания признака.

Эмпирическое значение критерия рассчитывается по формуле:

$$U_{эмп} = (n_1 \cdot n_2) + \frac{n_x \cdot (n_x + 1)}{2} - T_x, \qquad (21)$$

где n_1, n_2 - количество испытуемых в выборках 1 и 2 соответственно,

T_x – большая из ранговых сумм,

n_x – количество испытуемых в группе с большей суммой рангов.

Критическое значение критерия определяется по таблице для данных n_1 и n_2 и выбранного уровня значимости. Если $U_{эмп}$ больше $U_{кр}$, то принимается H_o.

в) H - критерий Крускала – Уоллиса

Критерий предназначен для оценки различий между тремя и более выборками по уровню какого-либо признака. Он позволяет установить, что уровень признака изменяется при переходе от группы к группе, но не указывает направление этих изменений.

Возможны гипотезы:

H_o: Между выборками 1, 2, 3 и т. д. существуют лишь случайные различия по уровню исследуемого признака.

H_1: Между выборками 1, 2, 3 и т. д. существуют неслучайные различия по уровню исследуемого признака.

Ограничения критерия H

1) Измерение может быть проведено по шкале интервалов или отношений.
2) Выборки должны быть независимыми.
3) При сопоставлении 3-х выборок допускается, чтобы в одной из них $n=3$, а в двух других $n=2$. Но при таких численных составах установить различия можно лишь на низшем уровне значимости.
4) При большем 5 количестве выборок и испытуемых в каждой выборке необходимо пользоваться таблицей критических значений критерия χ^2. Число степеней свободы при этом определяется как ν=c-1, где c – количество сопоставляемых выборок.

Наблюдения всех выборок необходимо объединить и проранжировать по степени нарастания признака.

Эмпирическое значение критерия H подсчитывается по формуле:

$$H_{эмп}=\frac{12}{N(N+1)}\cdot\sum_{i=1}^{c}\frac{T_i^2}{n_i}-3\cdot(N+1), \qquad (22)$$

где N – общее количество испытуемых в объединенной выборке,

n_i – количество испытуемых в каждой выборке,

T_i^2 – квадраты сумм рангов по каждой i–ой выборке.

Если эмпирическое значение критерия меньше критического значения, то H_o принимается.

г) S – критерий тенденций Джонкира

Критерий S предназначен для выявления тенденций изменения признака при переходе от выборки к выборке при сопоставлении трех и более выборок.

Гипотезы:

H_o: Тенденция возрастания значений признака при переходе от выборки к выборке является случайной.

H_1: Тенденция возрастания значений признака при переходе от выборки к выборке не является случайной.

Ограничения критерия

1) Измерение может быть проведено по шкале интервалов или отношений.
2) Выборки должны быть независимыми.
3) Количество наблюдений в каждой выборке должно быть одинаковым.
4) Нижняя граница применимости критерия: не менее трех выборок и не менее двух элементов в каждом наблюдении. Верхняя граница определяется таблицей приложения: не более 6 выборок и не более 10 наблюдений в каждой выборке.

Выборки необходимо располагать по возрастанию суммы значений признака слева на право.

Эмпирическое значение критерия рассчитывается по формуле:

$$S_{эмп}=2A-B, \qquad (23)$$

где A – общая сумма инверсий,

$B = \frac{c \cdot (c-1)}{2} \cdot n^2$ – максимально возможное значение величины А.

Под числом инверсий понимается число значений признака, больших каждого конкретного значения рассматриваемой выборки и расположенных правее от нее.

Критические значения критерия определяются по таблице, в соответствии с выбранным уровнем значимости, количеством выборок (c) и числом наблюдений (n) в каждой выборке.

Если эмпирическое значение критерия меньше критического значения, то принимается гипотеза H_o.

2. Критерии изменений

В психологических, педагогических, социологических исследованиях часто бывает важно доказать, что в результате действия каких-либо факторов произошли достоверные изменения (сдвиги) в измеряемых показателях. В зависимости от воздействующих факторов выделяют временные, ситуационные, умозрительные, измерительные сдвиги, сдвиги под влиянием экспериментальных воздействий и структурные сдвиги. Для оценки достоверности сдвигов применяют критерии изменений.

а) G-критерий знаков

Критерий знаков предназначен для установления общего направления сдвига исследуемого признака. Он позволяет установить, в какую сторону в выборке в целом изменяются значения признака при переходе от первого измерения ко второму: изменяются ли показатели в сторону улучшения, повышения или усиления или, наоборот, в сторону ухудшения, понижения или ослабления.

Возможны гипотезы:

H_0: Преобладание типичного направление сдвига является случайным.

H_1: Преобладание типичного направление сдвига не является случайным.

Ограничения критерия G

1) Измерение может быть проведено по шкале порядка, интервалов или отношений.
2) Выборка должна быть однородной и связной.
3) Объем выборки должен быть равным от 5 до 300.
4) При равенстве типичных и нетипичных сдвигов критерий знаков неприменим.

Эмпирическое значение критерия $G_{эмп}$ принимают равным числу нетипичных сдвигов, т. е. не преобладающих сдвигов в сторону увеличения или уменьшения показателя.

Критическое значение критерия $G_{кр}$ определяют по таблице в соответствии с выбранным уровнем значимости и объемом выборки без учета нулевых сдвигов. Если $G_{эмп}$ не превосходит $G_{кр}$, то гипотеза H_0 отвергается.

б) Парный критерий Т – Вилкоксона

Критерий применяется для сопоставления показателей, измеренных в двух разных условиях на одной и той же выборке испытуемых. Он позволяет установить не только направленность изменений, но и их выраженность. С его помощью мы определяем, является ли сдвиг показателей в каком-то одном направлении более интенсивным, чем в другом.

Возможны гипотезы:

H_0: Интенсивность сдвигов в типичном направлении не превосходит интенсивности сдвигов в нетипичном направлении.

H_1: Интенсивность сдвигов в типичном направлении превышает интенсивности сдвигов в нетипичном направлении.

Ограничения критерия Т

1) Измерение может быть проведено по любой шкале, кроме номинальной.
2) Выборка должна быть связной.
3) Объем выборки должен быть равным от 5 до 50.

Эмпирическое значение критерия подсчитывают по формуле:

$$T_{эмп} = \sum R_r \ , \qquad (24)$$

где R_r – ранговые значения сдвигов с более редким знаком.

Критическое значение критерия $T_{кр}$ определяется для данного объема выборки и выбранного уровня значимости по таблице. Если $T_{эмп}$ не превосходит $T_{кр}$, то гипотеза H_0 отвергается.

в) Критерий χ_r^2 Фридмана

Критерий применяется для сопоставления показателей, измеренных в трех или более условиях на одной и той же выборке испытуемых. Он позволяет установить, что величины показателей от условия к условию изменяются, но при этом не указывает на направление изменений.

Возможны гипотезы:

H_0: Между показателями, полученными (измеренными) в разных условиях, существуют лишь случайные различия.

H_1: Между показателями, полученными (измеренными) в разных условиях, существуют неслучайные различия.

Ограничения критерия χ_r^2

1) Измерение может быть проведено по шкале интервалов или отношений.
2) Выборка должна быть связной.
3) В выборке должно быть не менее двух испытуемых, каждый из которых имеет не менее трех показателей. Количество измерений не может превышать 100.

Эмпирическое значение критерия вычисляется по формуле:

$$\chi^2_{r\ эмп} = \left[\frac{12}{n\cdot c\cdot(c+1)}\cdot\sum_{i=1}^{c} T_i^2\right] - 3\cdot n\cdot(c+1)\ , \qquad (25)$$

где c – количество условий,

n – количество испытуемых,

T_i – суммы рангов по каждому из условий.

Критическое значение критерия $\chi^2_{r\ кр}$ определяем при выбранном уровне значимости и данном объеме выборки по правилам:

1) При $c=3$ и $n\leq 9$, критические значения определяются по таблице 8-А приложения.
2) При $c=4$ и $n\leq 4$, критические значения определяются по таблице 8-Б приложения.
3) При большем числе измерений и испытуемых критические определяются по таблице для критерия χ^2. В этом случае число степеней свободы определяется по формуле $\nu = c - 1$.

Если $\chi^2_{r\ эмп}$ не меньше $\chi^2_{r\ кр}$ то гипотеза H_0 отклоняется.

г) L – критерий тенденций Пейджа.

Критерий L Пейджа применяется для сопоставления показателей, измеренных в трех и более условиях на одной и той же выборке испытуемых. Критерий позволяет выявить тенденции в изменении величин признака при переходе от условия к условию, а также указывает на направление этих изменений.

Возможны гипотезы:

H_0: Увеличение индивидуальных показателей при переходе от первого условия ко второму, а затем к третьему и далее, случайно.

H_1: Увеличение индивидуальных показателей при переходе от первого условия ко второму, а затем к третьему и далее, неслучайно.

Ограничения критерия L

1) Измерение может быть проведено по ранговой шкале, шкале интервалов или отношений.
2) Выборка должна быть связной.
3) В выборке должно быть не менее двух и не больше 12 испытуемых, каждый из которых имеет не менее трех показателей. Максимальное число условий – 6.

Эмпирическое значение критерия определяется по формуле:

$$L_{\text{эмп}} = \sum_{i=1}^{c} (T_i \cdot i), \qquad (26)$$

где c – количество условий,

T_i – суммы рангов по каждому из условий,

i –порядковый номер, приписанный каждому условию, после упорядочения по возрастанию сумм рангов.

Критическое значение критерия $L_{кр}$ определяем при выбранном уровне значимости, данном объеме выборки и данном количестве условий по таблице. Если $L_{эмп}$ не меньше $L_{кр}$, то гипотеза H_0 отклоняется.

3. Критерии согласия

1. χ^2 - критерий Пирсона.

Основная расчетная формула эмпирического значения χ^2 :

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(f_э - f_m\right)^2}{f_m}, \tag{27}$$

где k - количество разрядов признака,

f_m - теоретическая частота,

$f_э$ - эмпирическая частота.

Если количество разрядов равно двум (принимает минимально возможное значение), то в расчетную формулу вносится поправка на непрерывность:

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(|f_э - f_m| - 0{,}5\right)^2}{f_m}. \tag{28}$$

Критическое значение $\chi^2_{кр}$ определяется по таблице 10 приложения в соответствии с определенным числом степеней свободы и уровнем значимости.

Используется в двух вариантах:

а) для сопоставления эмпирического распределения с теоретическим; в этом случае проверяется нулевая гипотеза H_0 об отсутствии различий между теоретическим и эмпирическим распределением.

В качестве теоретических распределений могут выступать, например, равномерное или нормальное распределения.

В случае равномерного распределения теоретические частоты подсчитываются по формуле:

$$f_m = \frac{n}{k}, \tag{29}$$

где n – количество наблюдений.

Число степеней свободы $v=k-1$.

В случае нормального распределения теоретическая частота подсчитывается по формуле:

$$f_m = \frac{nh}{\sigma} \cdot \varphi(u_i), \tag{30}$$

где h – шаг (разность между двумя соседними значениями признака),

$u_i = \frac{x_i - \bar{x}}{\sigma}$,

$\varphi(u)$- нормированная дифференциальная функция Лапласа.

Число степеней свободы рассчитывают как $k-3$.

б) как расчет однородности двух и более независимых экспериментальных выборок; в этом случае проверяется гипотеза H_0 об отсутствии различий между эмпирическими (экспериментальными) распределениями.

в) для сравнения показателей внутри одной выборки; в этом случае проверяется гипотеза H_0: сравниваемые признаки не влияют друг на друга.

В этом случае исходные данные удобно представлять в виде таблицы сопряженности:

Разряды	Эмпирические частоты					
	первое распределение	второе распределение	…	j-ое распределение	…	с-ое распределение
1	n_{11}	n_{12}	…	n_{1j}	…	n_{1c}
2	n_{21}	n_{22}	…	n_{2j}	…	n_{2c}
…	…	…	…	…	…	…
i	n_{i1}	n_{i2}	…	n_{ij}	…	n_{ic}
…	…	…	…	…	…	…
k	n_{k1}	n_{r2}	…	n_{kj}	…	n_{kc}

Где n_{ij} – эмпирическая частота, соответствующая *i*-ому разряду *j*-ого распределения.

Для каждой ячейки таблицы, соответствующая теоретическая частота рассчитывается по формуле:

$$f_{mij} = \frac{\left(\sum\limits_{j=1}^{c} n_{ij}\right)\cdot\left(\sum\limits_{i=1}^{k} n_{ij}\right)}{\sum\limits_{i=1}^{k}\sum\limits_{j=1}^{c} n_{ij}} \qquad (31)$$

или $$f_{mij} = \frac{(\text{сумма частот по строке})(\text{сумма частот по столбцу})}{\text{общее количество наблюдений}}. \qquad (32)$$

Число степеней свободы определяется по формуле*: ν=(k-1)(c-1).*

В случае, когда число переменных в двух сравниваемых выборках велико, можно использовать для вычисления $\chi^2_{эмп}$ следующую формулу:

$$\chi^2_{эмп} = 4\sum_{i=1}^{k}\frac{(f_{k1})^2}{f_{k1}+f_{k2}} - 2n, \qquad (33)$$

где f_{k1}- частоты первого распределения,

f_{k2}- частоты второго распределения,

n- число элементов в каждой выборке.

Если число значений в выборках различно, то используют формулу:

$$\chi^2_{эмп} = \frac{(n_1 + n_2)^2}{n_1 n_2}\left(\sum_{i=1}^{k} \frac{(f_{k1})^2}{f_{k1} + f_{k2}} - \frac{n_1^2}{n_1 + n_2}\right). \qquad (34)$$

Для применения критерия хи-квадрат необходимо соблюдать следующие условия:

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

2. λ- критерий Колмогорова-Смирнова

Критерий λ предназначен для сопоставления двух распределений:

а) эмпирического распределения с теоретическим;

б) одного эмпирического распределения с другим эмпирическим распределением.

Для применения λ- критерия необходимо соблюдать следующие условия:

1) измерение может быть проведено по шкале интервалов или отношений;
2) выборки должны быть случайными и независимыми;
3) желательно, чтобы суммарный объем двух выборок был большим или равным 50. С увеличением объема выборки точность критерия повышается;
4) эмпирические данные должны допускать возможность упорядочения по возрастанию или убыванию какого-либо признака и обязательно отражать какое-то его однонаправленное изменение.

Возможны гипотезы:

H_0: Различия между двумя распределениями не достоверны.

H_1: Различия между двумя распределениями достоверны.

Эмпирическое значение критерия для сопоставления эмпирического распределения с теоретическим определяется по формуле:

$$d_{эмп} = \max \frac{\left| f_э^* - f_т^* \right|}{n}, \tag{35}$$

где $f_э^*$ - накопленные эмпирические частоты,

$f_т^*$ - накопленные теоретические частоты.

Критическое значение критерия определяется по таблице, при определенном уровне значимости и данном объеме выборки. Если число элементов выборки больше 100, то величина критических значений вычисляется по формуле:

$$d_{кр} = \begin{cases} 1{,}36 \big/ \sqrt{n} \\ 1{,}63 \big/ \sqrt{n} \end{cases}. \tag{36}$$

Эмпирическое значение критерия для сопоставления эмпирического распределения с другим эмпирическим распределением определяется по формуле:

$$\lambda_{эмп} = d_{\max} \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}, \tag{37}$$

где n_1-количество наблюдений в первой выборке,

n_2- количество наблюдений во второй выборке,

d_{max}- наибольшая абсолютная величина разности между накопленными частостями по каждому разряду.

Уровень значимости соответствующий полученному значению λ определяется по таблице.

3. Критерий Фишера – φ

Критерий Фишера предназначен для сопоставления двух рядов выборочных значений по частоте встречаемости какого-либо признака. Этот критерий можно применять для оценки различий в любых двух выборках зависимых или независимых. С его помощью можно сравнивать показатели одной и той же выборки, измеренные в разных условиях.

Эмпирическое значение критерия подсчитываются по формуле:

$$\varphi_{эмп} = (\varphi_1 - \varphi_2) \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}, \tag{38}$$

где φ_1 - величина, определяемая по таблице, соответствующая большей процентной доле,

φ_2 - величина, определяемая по таблице, соответствующая меньшей процентной доле.

Уровень значимости соответствующий полученному значению φ определяется по таблице.

Для применения критерия Фишера φ необходимо соблюдать следующие условия:

1) измерение может быть проведено по любой шкале;
2) характеристики выборок могут быть любыми;
3) нижняя граница – в одной из выборок может быть только два наблюдения, при этом во второй должно быть не менее 30 наблюдений. Верхняя граница не определена;
4) нижние границы двух выборок должны содержать не меньше 5 элементов (наблюдений) в каждой.

Факторный анализ. Область применения и основные этапы факторного анализа. Дополнительные статистические показатели для оценки факторного анализа.

Основные понятия факторного анализа

Факторный анализ – статистический метод, который используется при обработке больших массивов экспериментальных данных. Задачами факторного анализа являются: сокращение числа переменных (редукция данных) и определение структуры взаимосвязей между переменными, т.е. классификация переменных, поэтому факторный анализ используется как метод сокращения данных или как метод структурной классификации.

Важное отличие факторного анализа от всех описанных выше методов заключается в том, что его нельзя применять для обработки первичных, или, как говорят, «сырых», экспериментальных данных, т.е. полученных непосредственно при обследовании испытуемых. Материалом для факторного анализа служат корреляционные связи, а точнее – коэффициенты корреляции Пирсона, которые вычисляются между переменными (т.е. психологическими признаками), включенными в обследование. Иными словами, факторному анализу подвергают корреляционные матрицы, или, как их иначе называют, матрицы интеркорреляций. Наименования столбцов и строк в этих матрицах одинаковы, так как они представляют собой перечень переменных, включенных в анализ. По этой причине матрицы интеркорреляций всегда квадратные, т.е. число строк в них равно числу столбцов, симметричные, т.е. на симметричных местах относительно главной диагонали стоят одни и те же коэффициенты корреляции.

Необходимо подчеркнуть, что исходная таблица данных, из которых получается корреляционная матрица, не обязательно должна быть квадратной. Например, психолог измерил три показателя интеллекта (вербальный, невербальный и общий) и школьные отметки по трем учебным предметам (литература, математика, физика) у 100 испытуемых – учащихся девятых классов. Исходная матрица данных будет иметь размер 100×6, а матрица интеркорреляций размер 6×6, поскольку в ней имеется только 6 переменных. При таком количестве переменных матрица интеркорреляций будет включать 15 коэффициентов и проанализировать ее не составит труда.

Однако представим, что произойдет, если психолог получит не 6, а 100 показателей от каждого испытуемого. В этом случае он должен будет анализировать 4950 коэффициентов корреляции. Число коэффициентов в матрице вычисляется по формуле $\frac{n \cdot (n+1)}{2}$ и в нашем случае равно соответственно $\frac{(100 \times 99)}{2} = 4950$. очевидно, что провести визуальный анализ такой матрицы – задача труднореализуемая. Вместо этого психолог может выполнить математическую процедуру факторного анализа корреляционной матрицы размером 100×100 (100 испытуемых и 100 переменных) и таким путем получить более простой материал для интерпретации экспериментальных результатов.

Главное понятие факторного анализа – фактор. Это искусственный статистический показатель, возникающий в результате специальных преобразований таблицы коэффициентов корреляции между изучаемыми психологическими признаками, или матрицы интеркорреляций. Процедура извлечения факторов из матрицы интеркорреляций называется факторизацией матрицы. В результате факторизации из корреляционной матрицы может быть извлечено разное количество факторов вплоть до числа, равного количеству исходных переменных. Однако факторы, выделяемые в результате факторизации, как правило, неравноценны по своему значению.

Элементы факторной матрицы называются *«факторными нагрузками, или весами»*; и они представляют собой коэффициенты корреляции данного фактора со всеми показателями, использованными в исследовании. Факторная матрица очень важна, поскольку она показывает, как изучаемые показатели связаны с каждым выделенным фактором. При этом факторный вес демонстрирует меру, или тесноту, этой связи.

Поскольку каждый столбец факторной матрицы (фактор) является своего рода переменной величиной, то сами факторы также могут коррелировать между собой. Здесь возможны два случая: корреляция между факторами равна нулю, в таком случае факторы являются независимыми (ортогональными). Если корреляция между факторами больше нуля, то в таком случае факторы считаются зависимыми (облическими). Подчеркнем, сто ортогональные факторы в отличие от облических дают более простые варианты взаимодействий внутри факторной матрицы.

В качестве иллюстрации ортогональных факторов часто приводят задачу Л. Терстоуна, который, взяв ряд коробок разных размеров и формы, измерил в каждой из них больше 20 различных показателей и вычислил корреляции между ними. Профакторизовав полученную матрицу интеркорреляций, он получил три фактора, корреляция между которыми была равна нулю. Этими факторами были «длина», «ширина» и «высота».

Для того чтобы лучше уловить сущность факторного анализа, разберем более подробно следующий пример.

Предположим, что психолог у случайной выборки студентов получает следующие данные:

- V_1 – вес тела (в кг);
- V_2 – количество посещений лекций и семинарских занятий по предмету;
- V_3 – длина ноги (в см);
- V_4 – количество прочитанных книг по предмету;
- V_5 – длина руки (в см);
- V_6 – экзаменационная оценка по предмету

(V – от английского слова variable – переменная).

При анализе этих признаков не лишено оснований предположение о том, что переменные V_1, V_3 и V_5 – будут связаны между собой, поскольку чем больше человек, тем больше он весит и тем длиннее его конечности. Сказанное означает, что между этими переменными должны получиться статистически значимые коэффициенты корреляции, поскольку эти три переменные измеряют некоторое фундаментальное свойство индивидуумов в выборке, а именно: их размеры. Точно так же вероятно, что при вычислении корреляций между V_2, V_4 и V_6 тоже будут получены достаточно высокие коэффициенты корреляции, поскольку посещение лекций и самостоятельные занятия будут способствовать получению более высоких оценок по изучаемому предмету.

Таким образом, из всего возможного массива коэффициентов, который получается путем перебора пар коррелируемых признаков V_1 и V_2, V_1 и V_3 и т.д., предположительно выделяется два блока статистически значимых корреляций. Остальная часть корреляций – между признаками, входящими в разные блоки, вряд ли будет иметь статистически значимые коэффициенты, поскольку связи между такими признаками, как размер конечности и успеваемость по предмету, имеют скорее всего случайный характер. Итак, содержательный анализ 6 наших переменных показывает, что они по сути дела, измеряют только две обобщенные характеристики, а именно: размеры тела и степень подготовленности к предмету.

К полученной матрице интеркорреляций, т.е. вычисленным попарно коэффициентам корреляций между всеми шестью переменными V_1 – V_6, допустимо применить факторный анализ. Его можно проводить и вручную, с помощью калькулятора, однако процедура подобной статистической обработки очень трудоемка. По этой причине в настоящее время факторный анализ проводится на компьютерах, как правило, с помощью стандартных статистических пакетов. Во всех современных статистических пакетах есть программы для корреляционного и факторного анализов. Компьютерная программа по факторному анализу по существу пытается «объяснить» корреляции между переменными в терминах небольшого числа факторов (в нашем примере двух).

Предположим, что, используя компьютерную программу, мы получили матрицу интеркорреляций всех шести переменных и подвергли ее факторному анализу. В результате факторного анализа получилась таблица, которую называют «факторной матрицей», или «факторной структурной матрицей».

Переменная	Фактор 1	Фактор 2
V_1	0,91	0,01
V_2	0,20	0,96
V_3	0,94	-0,15
V_4	0,11	0,85
V_5	0,89	0,07
V_6	-0,13	0,93

По традиции факторы представляются в таблице в виде столбцов, а переменные в виде строк. Заголовки столбцов таблицы 1 соответствуют номерам выделенных факторов, но более точно было бы их называть «факторные нагрузки», или «веса», по фактору 1, то же самое по фактору 2. как указывалось выше, факторные нагрузки, или веса, представляют собой корреляции между соответствующей переменной и данным фактором. Например, первое число 0,91 в первом факторе означает, что корреляция между первым фактором и переменной V_1 равна 0,91. чем выше факторная нагрузка по абсолютной величине, тем больше ее связь с фактором.

Из таблицы видно, что переменные V_1, V_3 и V_5 имеют большие корреляции с фактором 1 (фактически переменная 3 имеет корреляцию близкую к 1 с фактором 1). В то же время переменные V_1, V_3 и V_5 имеют корреляции близкие к 0 с фактором 2. подобно этому фактор 2 высоко коррелирует с переменными V_2, V_4 и V_6 и фактически не коррелирует с переменными V_1, V_3 и V_5.

В данном примере очевидно, что существует две структуры корреляций, и, следовательно, вся информация таблицы 1 определяется двумя факторами. Теперь начинается заключительный этап работы – интерпретация полученных данных. Анализируя факторную матрицу, очень важно учитывать знаки факторных нагрузок в каждом факторе. Если в одном и том же факторе встречаются нагрузки с противоположными знаками, это означает, что между переменными, имеющими противоположные знаки, существует обратно пропорциональная зависимость. Отметим, что при интерпретации фактора для удобства можно изменить знаки всех нагрузок по данному фактору на противоположные.

Факторная матрица показывает также, какие переменные образуют каждый фактор. Это связано, прежде всего, с уровнем значимости факторного веса. По традиции минимальный уровень значимости коэффициентов корреляции в факторном анализе берется равным 0,4 или даже 0,3 (по абсолютной величине), поскольку нет специальных таблиц, по которым можно было бы определить критические значения для уровня значимости в факторной матрице. Следовательно, самый простой способ увидеть какие переменные «принадлежат» фактору это значит отметить те из них, которые имеют нагрузки выше чем 0,4 (или меньше чем – 0,4). Укажем, что в компьютерных пакетах иногда уровень значимости факторного веса определяется самой программой и устанавливается на более высоком уровне, например 0,7.

Так, из таблицы 1, следует вывод, что фактор 1 – это сочетание переменных V_1, V_3 и V_5 (но не V_2, V_4 и V_6, поскольку их факторные нагрузки по модулю меньше чем 0,4). Подобно этому фактор 2 представляет собой сочетание переменных V_2, V_4 и V_6.

Выделенный в результате факторизации фактор представляет собой совокупность тех переменных из числа включенных в анализ, которые имеют значимые нагрузки. Нередко случается, однако, что в фактор входит только одна переменная со значимым факторным весом, а остальные имеют незначимую факторную нагрузку. В этом случае фактор будет определяться по названию единственной значимой переменной.

В сущности, фактор можно рассматривать как искусственную «единицу» группировки переменных (признаков) на основе имеющихся между ними связей. Эта единица является условной, потому что, изменив определенные условия процедуры факторизации матрицы интрекорреляций, можно получить иную факторную матрицу (структуру). В новой матрице может оказаться иным распределение переменных по факторам и их факторные нагрузки.

В связи с этим в факторном анализе существует понятие «простая структура». Простой называют структуру факторной матрицы, в которой каждая переменная имеет значимые нагрузки только по одному из факторов. А сами факторы ортогональны, т.е. не зависят друг от друга. В нашем примере два общих фактора независимы. Факторная матрица с простой структурой позволяет провести интерпретацию полученного результата и дать наименование каждому фактору. В нашем случае фактор первый – «размеры тела», фактор второй – «уровень подготовленности».

Сказанное выше не исчерпывает содержательных возможностей факторной матрицы. Из нее можно извлечь дополнительные характеристики, позволяющие более детально исследовать связи переменных и факторов. Эти характеристики называются «общность» и «собственное значение2 фактора.

Однако, прежде чем представить их описание, укажем на одно принципиально важное свойство коэффициента корреляции, благодаря которому получают эти характеристики. Коэффициент корреляции, возведенный в квадрат (т.е. помноженный сам на себя), показывает, какая часть дисперсии (вариативности) признака является общей для двух переменных, или, говоря проще, насколько сильно эти переменные перекрываются. Так, например, две переменные с корреляцией 0,9 перекрываются со степенью $0{,}9 \times 0{,}9 = 0{,}81$. это означает, что 81% дисперсии той и другой переменной являются общими, т.е. совпадают. Напомним, что факторные нагрузки в факторной матрице – это коэффициенты корреляции между факторами и переменными, поэтому, возведенная в квадрат факторная нагрузка характеризует степень общности (или перекрытия) дисперсий данной переменной и данного фактором.

Если полученные факторы не зависят друг от друга («ортогональное» решение), по весам факторной матрицы можно определить, какая часть дисперсии является общей для переменой и фактора. Вычислить, какая часть вариативности каждой переменной совпадает с вариативностью факторов, можно простым суммированием квадратов факторных нагрузок по всем факторам. Из таблицы 13.1, например, следует, что $0{,}91 \times 0{,}91 + 0{,}01 \times 0{,}01 = 0{,}8282$, т.е. около 82% вариативности первой переменной «объясняется» двумя первыми факторами. Полученная величина называется *общностью* переменной, в данном случае переменной V_1.

Переменные могут иметь разную степень общности с факторами. Переменная с большей общностью имеет значительную степень перекрытия (большую долю дисперсии) с одним или несколькими факторами. Низкая общность подразумевает, что все корреляции между переменными и факторами невелики. Это означает, что ни один из факторов не имеет совпадающей доли вариативности с данной переменной. Низкая общность может свидетельствовать о том, что переменная измеряет нечто качественно отличающееся от других переменных, включенных в анализ. Например, одна переменная, связанная с оценкой мотивации среди заданий, оценивающих способности, будет иметь общность с факторами способностей близкую к нулю.

Малая общность может также означать, что определенное задание испытывает на себе сильное влияние ошибки измерения или крайне сложно для испытуемого. Возможно, напротив, также, что задание настолько просто, что каждый испытуемый дает на него правильный ответ, или задание настолько нечетко по содержанию, что испытуемый не понимает суть вопроса. Таким образом, низкая общность подразумевает, что данная переменная не совмещается с факторами по одной из причин: либо переменная измеряет другое понятие, либо переменная имеет большую ошибку измерения, либо существуют искажающие дисперсию признака различия между испытуемыми в вариантах ответа на это задание.

Наконец, с помощью такой характеристики, как собственное значение фактора, можно определить относительную значимость каждого из выделенных факторов. Для этого надо вычислить, какую часть дисперсии (вариативности) объясняет каждый фактор. Тот фактор, который объясняет 45% дисперсии (перекрытия) между переменными в исходной корреляционной матрице, очевидно является более значимым, чем другой, который объясняет только 25% дисперсии. Эти рассуждения, однако, допустимы, если факторы ортогональны, иначе говоря, не зависят друг от друга.

Для того чтобы вычислить собственное значение фактора, нужно возвести в квадрат факторные нагрузки и сложить их по столбцу. Используя данные таблицы 13.1 можно убедиться что собственное значение фактора 1 составляет $(0,91 \times 0,91 + 0,20 \times 0,20 + 0,94 \times 0,94 + 0,11 \times 0,11 + 0,84 \times 0,84 + (-0,13) \times (-0,13)) = 2,4863$. Если собственное значение фактора разделить на число переменных (6 в нашем примере), то полученное число покажет, какая доля дисперсии объясняется данным фактором. В нашем случае получится

$\frac{2,4863}{6} \cdot 100\% = 41,4\%$. Иными словами, фактор 1 объясняет около 41% информации (дисперсии) в исходной корреляционной матрице. Аналогичный подсчет для второго фактора даст 41,5%. В сумме это будет составлять 82,9%.

Таким образом, два общих фактора, будучи объединены, объясняют только 82,9% дисперсии показателей исходной корреляционной матрицы. Что случилось с «оставшимися» 17,1%? Дело в том, что рассматривая корреляции между 6 переменными, мы отмечали, что корреляции распадаются на два отдельных блока, и поэтому решили, что логично анализировать материал в понятиях двух факторов, а не 6, как и количество исходных переменных. Другими словами, число конструктов, необходимых, чтобы описать данные, уменьшилось с 6 (число переменных) до 2 (число общих факторов). В результате факторизации часть информации в исходной корреляционной матрице была принесена в жертву построению двухфакторной модели. Единственным условием, при котором информация не утрачивается, было бы рассмотрение шестифакторной модели.

Условия применения факторного анализа

Факторный анализ может быть уместен, если выполняются следующие критерии.

1. Нельзя факторизировать качественные данные, полученные по шкале наименований, например, такие, как цвет волос (черный / каштановый / рыжий) и т.п.
2. Все переменные должны быть независимыми, а их распределение должно приближаться к нормальному.
3. Связи между переменными должны быть приблизительно линейны, по крайней мере, не иметь явно криволинейного характера.
4. В исходной корреляционной матрице должно быть несколько корреляций по модулю выше 0,3. В противном случае достаточно трудно извлечь из матрицы какие-либо факторы.

5. Выборка испытуемых должна быть достаточно большой. Рекомендации экспертов варьируют. Наиболее жесткая точка зрения рекомендует не применять факторный анализ, если число испытуемых меньше 100, поскольку стандартные ошибки корреляции в этом случае окажутся слишком велики.

Однако если факторы хорошо определены (например, с нагрузками 0,7, а не 0,3), экспериментатору нужна меньшая выборка, чтобы выделить их. Кроме того, если известно, что полученные данные отличаются высокой надежностью (например, используются валидные тесты), то можно анализировать данные и по меньшему числу испытуемых.

Приемы для определения числа факторов

Разработано несколько приемов для выбора «правильного» числа факторов из корреляционной матрицы. Определение числа выделяемых факторов, вероятно, наиболее важное решение, которое необходимо принять при проведении факторного анализа. Неверное решение может привести к бессмысленным результатам при обработке самого четкого набора данных. Нет ничего страшного в том, чтобы попытаться выполнить несколько вариантов анализа, базирующегося на разном числе факторов, и использовать несколько различных приемов, определяющих выбор факторов.

Первые руководящие принципы – это теория, здравый смысл, а также прошлый опыт. При этом психолог должен установить:

- не способствует ли увеличение числа факторов уменьшению доли нагрузок в диапазоне от –0,4 до +0,4? Если это так, то это увеличение скорее всего не имеет смысла;
- не появляются ли какие либо большие корреляции между факторами при осуществлении облических вращений. Последнее может указывать, что было извлечено слишком много факторов, и два фактора проходят через один и тот же кластер переменных. Корреляции между факторами больше, чем приблизительно 0,5 могут косвенно свидетельствовать об этом;

- не разделились ли какие-либо хорошо известные факторы на две или большее количество частей. Например, если во множестве предшествующих исследований было показано, что набор заданий формирует только один фактор (например экстраверсия), а вам кажется, что в вашем анализе, они все же формируют два фактора, вероятно, что было извлечено слишком много факторов.

Существует ряд способов определения числа факторов, с которыми связаны исследуемые переменные величины. Наиболее надежны из них – определение числа вкладов ряда первых *m* факторов в общую дисперсию. Обычно, если сумма вкладов первых *m* факторов составляет 90 или 95%, этой величиной ограничивают число анализируемых факторов.

Иллюстрирует это приведенный ниже пример в таблице:

	Собственные значения 10 факторов Метод главных компонент			
Факторы	Собственные значения факторов	% общей дисперсии	Кумулят. соб. знач.	Кумулят. %
1	6,118369	61,18369	6,11837	61,1837
2	1,800682	18,00682	7,91905	79,1905
3	,472888	4,72888	8,39194	83,9194
4	,407996	4,07996	8,79993	87,9993
5	,317222	3,17222	9,11716	91,1716
6	,293300	2,93300	9,41046	94,1046
7	,195808	1,95808	9,60626	96,0626
8	,170431	1,70431	9,77670	97,7670
9	,137970	1,37970	9,91467	99,1467
10	,085334	,85334	10,00000	100,0000

Как можно видеть из таблицы, первый фактор (значение 1) объясняет 61% процент общей дисперсии, фактор 2 (значение 2) – 18% процентов, и т.д. Четвертый столбец содержит накопленную или кумулятивную дисперсию. Напомним, что дисперсии, выделяемые факторами, называются *собственными значениями*.

Таким образом, из 10 факторов первые 5 объясняют 91% всей дисперсии, их анализом можно ограничиться. Фактически, однако, только первые два фактора несут на себе основную нагрузку, и реально исследователи в такой ситуации нередко пренебрегают оставшимися тремя, которые все вместе объясняют не более 12%.

В заключение отметим, что проблема определения имеет ряд дискуссионных вопросов. Существуют несколько методов определения количества факторов, но они достаточно сложны и их реализация возможна только на ЭВМ.

Вращение факторов

Вращение факторов изменяет положение факторов по отношению к переменным таким образом, что получаемое решение легко интерпретировать. Как упоминалось выше, факторы идентифицируют, наблюдая, какие переменные имеют большие и/или нулевые нагрузки по ним. Решения, которые не подчиняются интерпретации, – это те решения, в которых большое число переменных имеет нагрузки «среднего уровня» по фактору, т.е. нагрузки порядка 0,3. они слишком малы, чтобы рассматриваться как «выступающие» и использоваться для идентификации фактора, и все же слишком велики, чтобы их можно было игнорировать безо всякого риска.

Вращение (ротация факторов) перемещает факторы относительно переменных таким образом, что каждый фактор начинает обладать несколькими существенными нагрузками и несколькими нагрузками близкими к нулю. Иными словами, цель вращения – преобразовать факторную матрицу таким образом, чтобы получилась простая структура, в которой каждый фактор имеет некоторое количество больших нагрузок и некоторое маленьких, и подобно этому каждая переменная имеет существенные нагрузки только по некоторым факторам.

Приведем пример факторной матрицы «до» и «после» вращения.

	До вращения	До вращения	После вращения (Варимакс)	После вращения (Варимакс)
	Фактор 1	Фактор 2	Фактор1	Фактор2
Экстраверсия	0.37	0.29	0.60	0.00
Тревожность	0.42	0.52	0.74	0.00
Нейротизм	0.43	–0.43	0.13	0.75
Агрессивность	0.51	–0.32	0.06	0.89

Эта таблица демонстрирует, насколько проще интерпретировать факторы, полученные после вращения, по сравнению с факторами, имевшимися до вращения. Факторное решение до вращения (левая половина таблицы 3) трудно интерпретировать, поскольку все переменные имеют почти равные нагрузки как по первому, так и по второму фактору. После вращения (правая половина таблицы 3) получается простая структура, провести интерпретацию которой становится значительно проще. Распределение нагрузок по факторам дает основание утверждать, что первый фактор измеряет экстраверсию и тревожность, второй – нейротизм и агрессивность.

В практике факторного анализа используются разные варианты вращения факторов, при этом выделяются два основных метода вращения – *ортогональное* и *косоугольное (обликическое)*.

Сущность ортогонального вращения заключается в том, что при вращении остается верным предположение о независимости факторов.

Ортогональное вращение бывает четырех видов: *варимакс*, *квартимакс*, *эквимакс* и *биквартимакс*.

При использовании метода *варимакс* минимизируется количество переменных, имеющих высокие нагрузки на данный фактор, при этом максимально увеличивается дисперсия фактора. Это способствует упрощению описания фактора за счет группировки вокруг него только тех переменных, которые в большей степени связаны с ним, чем остальные.

Квартимакс, напротив, минимизирует количество факторов, необходимых для объяснения данной переменной. Этот метод усиливает возможности интерпретации переменных. Он позволяет выделить один фактор с достаточно высокими нагрузками на большинство переменных.

Последующие два метода являются комбинациями *варимакса* и *квартимакса*. Однако, как показывает практика, психологи предпочитают использовать метод *варимакс*.

Что касается методов косоугольного вращения, то они также позволяют упростить описание факторного решения за счет введения предположения о коррелированности факторов. В статистических программах на ЭВМ большое распространение получил метод *облимин*. Этот метод эквивалентен методу *эквимакс* для ортогонального вращения.

Использование факторного анализа в психологии

Факторный анализ широко используется в психологии в разных направлениях, связанных с решением как теоретических, так и практических проблем.

В теоретическом плане использование факторного анализа связано с разработкой так называемого факторно-аналитического подхода к изучению структуры личности, темперамента и способностей. Использование факторного анализа в этих сферах основано на широко принятом допущении, согласно которому наблюдаемые и доступные для прямого измерения показатели являются лишь косвенными и/или частными внешними проявлениями более общих характеристик. Эти характеристики, в отличие от первых, являются скрытыми, так называемыми латентными переменными, поскольку они представляют собой понятия или конструкты, которые не доступны для прямого измерения. Однако они могут быть установлены путем факторизации корреляционных связей между наблюдаемыми чертами и выделением факторов, которые (при условии хорошей структуры) можно интерпретировать как статистическое выражение искомой латентной переменной.

Хотя факторы имеют чисто математический характер, предполагается, что они репрезентируют скрытые переменные (теоретически постулируемые конструкты или понятия), поэтому названия факторов нередко отражают сущность изучаемого гипотетического конструкта. Так, факторный анализ, который был разработан в начале XX века Ч. Спирменом для исследования структуры способностей, позволил ввести в психологию понятие общего фактора способностей – фактора *g*. Впоследствии Л. Терстоун выдвинул и экспериментально апробировал модель, которая включала 12 факторов способностей. Факторно-аналитичесие исследования темперамента и личности в зарубежной психологии охватывают целый ряд теорий прошлого и настоящего, включая теории Г. Олпорта, Р. Кэттела, Г. Айзенка и других.

В отечественной психологии факторный анализ наиболее широко использовался в дифференциальной психологии и психофизиологии при изучении свойств нервной системы человека в работах Б.М. Теплова и его школы. Теплов придавал большое значение этому виду статистической обработки данных, подчеркивая, что факторный анализ – ценное орудие в любой области, где можно хотя бы в виде предварительной гипотезы предположить наличие некоторых основных параметров, функций, свойств, образующих «структуру» данной области явлений.

В настоящее время факторный анализ широко используется в дифференциальной психологии и психодиагностике. С его помощью можно разрабатывать тесты, устанавливать структуру связей между отдельными психологическими характеристиками, измеряемыми набором тестов или заданиями теста.

Еще один аспект использования факторного анализа заключается в так называемой «редукции» данных или «концептуальной чистке» большого количества тестов, разработанных с различных теоретических позиций для измерения личностных особенностей. В результате факторизации матрицы корреляций, полученной на большой выборке испытуемых при использовании различных личностных тестов, можно более точно выявить структуру личностных особенностей, определяемых используемыми тестами.

Факторный анализ используется также для стандартизации тестовых методик, которая проводится на репрезентативной выборке испытуемых.

Дисперсионный анализ. Понятие об одно-, двух- и многофакторном дисперсионном анализе. Таблица сопряженности для числовых и номинальных признаков. Теорема Пирсона-Фишера.

Допустим, имеется несколько переменных факторов, которые либо классифицированы, либо упорядочены, либо, наконец, измерены. Требуется установить, влияют ли эти факторы на изучаемую переменную (случайную величину или случайную функцию). Исследование влияния переменных факторов на изучаемую переменную по дисперсиям называется дисперсионным анализом (ДА).

Пусть изучаемая случайная величина X, о которой известно, что при определенном комплексе условий X имеет генеральную дисперсию D_0 [X]. Требуется проверить, влияет ли на X некоторый фактор (условие) A, до сих пор не принимавшийся во внимание. Проводиться серия наблюдений переменной X при условии A (при действии фактора A) в дополнение к предыдущему комплексу условий. В результате получаем выборку из n значений X ($x_1, x_2, x_3, \ldots, x_n$), имеющую дисперсию D [X]. Очевидно, что если фактор A не влияет на X, то D [X]= D_0 [X] – в идеальном случае. Практически это не так, но различие не должно превосходить случайного (с заданной степенью надежности), что легко определяется по F – критерию Фишера. Иначе говоря, если D [X] ≈ D_0 [X] – в пределах случайных флюктуаций – можно считать, что либо фактор A не влияет X, либо это влияние несущественно при данном объеме (n) выборки.

Если D [X]> D_0 [X], и это не случайно, то приходиться признать, что фактор A влияет на переменную X. Тогда считая действие фактора A независимым от других условий, можем написать:

$$D[X]= D_0[X]+D_A[X],$$

где D_0 [X] – дисперсия, обусловленная случайными влияниями неконтролируемых условий; D_A[X] – дисперсия, характеризующая влияние фактора A на изучаемую переменную X.

Влияние фактора A на переменную X может быть различным. Здесь выделяются три случая. Первый случай: фактор A влияет только на среднюю величину X, тогда D_A[X] есть характеристика рассеивания средних значений переменной X под влиянием фактора A. Второй случай: фактор A влияет только на рассеивание значений X, т.е. на D_0 [X], тогда D_A[X] есть некоторая «добавка» к D_0 [X], выражающая степень влияния фактора A. Третий случай: фактор A влияет и на M[X], и на D_0[X], тогда D_A[X] суммирует эти влияния.

Если имеет место второй или третий случай, то влияние фактора A можно учесть, применяя параметрические и непараметрические критерии, меры корреляции, уравнения регрессии. Если же имеет место первый случай, то наряду с рассмотренными методами применяют дисперсионный анализ. Иначе говоря, ДА используется лишь в первом случае. Его мы и будем рассматривать в этой главе.

Пусть теперь в дополнение к фактору A требуется испытать влияние фактора B. Здесь возможны два варианта:

1)*Факторы A и B независимы*. Проводится новая серия наблюдений над переменной X при условии, что кроме A на M[X] действует фактор B. В результате получаем выборку из n значений переменной X, имеющую дисперсию D*[X]. Очевидно, что

$$D^*[X] = D[X] + D_B[X],$$

где D_B[X] – доля дисперсии, обусловленная действием фактора B. Если влияние фактора B значимое (согласно F – критерию и принятому доверительному уровню), то член D_B[X] существенно отличен от нуля.

$$D^*[X] = D_0[X] + D_A[X] + D_B[X],$$

где в лево части – общая дисперсия выборки, а в правой части члены D_A[X] и D_B[X] выражают долю рассеивания средних значений M[X] переменной X под влиянием факторов A и B соответственно; член D_0 [X] выражает случайные флюктуации переменной под влиянием всех других (не A и не B) факторов (условий) опыта и называется *остаточной дисперсией.*

Заметим, что член D [X] выражал влияние (в общем случайное) всех других, кроме B, факторов (в том числе и A); это тоже остаточная дисперсия. Вообще *под остаточной дисперсией понимается часть общей дисперсии выборки, которая не входит в долю дисперсии по данному фактору или группе факторов*.

2)*факторы А и В зависимы*. Тогда помимо трех слагаемых в правой части появляется еще одно слагаемое:

$$D^*[X] = D_0[X] + D_A[X] + D_B[X] + D_{AB}[X].$$

Слагаемое $D_{AB}[X]$ выражает долю общей дисперсии $D^*[X]$, обусловленную совместным влиянием факторов А и В на математическое ожидание изучаемое переменной X.

Обобщим сказанное о двух факторах на произвольное количество факторов. Пусть на случайную перемену. Тогда если эти факторы независимы попарно и в совокупности, то

$$D[X] = D_0[X] + D_{A_1}[X] + D_{A_2}[X] + \ldots D_{A_m}[X],$$

где D [X] – общая дисперсия выборки, полученная при воздействии указанных факторов $A_1, A_2, \ldots, A_m$; D_0 [X] – генеральная дисперсия, которую мы здесь принимаем за остаточную дисперсию. Члены $D_{A_i}[X]$ (при i=1,2,…,m) выражают парциальное действие соответствующих факторов на математическое ожидание переменной X. Если факторы попарно независимы, то добавляются справа слагаемые, выражающие совместное действие всех возможных пар факторов; число таких слагаемых определяется числом сочетаний из m по 2. Если факторы к тому же зависимы по три, по четыре и т. д. – по m, то добавляются еще члены, выражающие еще долю в общей дисперсии сочетаний факторов по три, по четыре и т.д.

На основе сказанного можно видеть, что общая дисперсия выборки, полученной при влиянии m факторов, определяется как дисперсия суммы m+1 случайных величин, где дополнительная случайная величина отображает влияние неучитываемых условий (ей соответствует остаточная дисперсия). При этом для независимых факторов общая дисперсия есть линейная сумма факторных и остаточной (всего m+1) дисперсий, а для зависимых (в общем случае) сюда добавляется сумма дисперсии всех факторных взаимодействий по два, по три и т.д. – по m.

Таким образом, сущность ДА состоит в том, чтобы представить общую дисперсию в виде суммы дисперсий, обусловленных влиянием контролируемых и неконтролируемых условий опыта и , оценивая дисперсионные отношения, определить меру влияния контролируемых условий (факторов) на середине значения изучаемой переменной.

Как мы видели, одной из предпосылок ДА является использование F-критерия Фишера для проверки значимости влияния факторов. Поскольку F – Фишера основывается на предположении о нормальном распределении генеральных совокупностей, из которых взяты выборки, то это обстоятельство весьма существенно для возможности применения рассматриваемого метода в целом. Дисперсионный анализ следует применять, когда известно (или доказано), что выборки нормально распределены. Это первое важное условие грамотного применения ДА. В противном случае истинность выводов ничем не обоснована и не гарантируется.

Второй предпосылкой ДА является, как показано, выделение факторных и остаточной дисперсий из общей дисперсии. Чтобы это выделение было возможным, необходимо, чтобы остаточная дисперсия не изменялась от опыта к опыту под влиянием контролируемых и неконтролируемых факторов. В противном случае изменение остаточной дисперсии не позволят однозначно решить, что вносят в рассеивание средних данных контролируемые факторы. Поэтому прежде чем применять дисперсионный анализ, необходимо убедиться в том, что дисперсия выборочных серий не меняется однонаправлено от серии к серии. Если такое изменение происходит, необходимо стабилизировать дисперсию и лишь потом изменять ДА.

Основная задача ДА состоит в том, чтобы из произвольного числа факторов, которые могут (как предполагается априори) влиять на изучаемую переменную, выделить сравнительно небольшое количество факторов, влияющих наиболее существенно. Эта основная задача ДА, смотря по обстоятельствам конкретизируется по-разному. В частности, иногда выделяют два этапа ДА. Первый связан с оценкой общего, недифференцированного влияния одной или нескольких факторов на среднее значение изучаемой переменной. Второй этап состоит в исследовании специфического, парциального действия факторов.

Оценка общего влияния факторов позволяет сравнительно быстро минимизировать первоначально обычно большое количество факторов. В результате первого этапа ДА многие из априори выбранных факторов отбрасываются как несущественные. Оставшиеся небольшое количество факторов затем исследуется подробно, для того чтобы определить «вес» каждого из факторов и факторных комбинаций.

Таким образом, конкретизация основной задачи ДА может идти в трех направлениях: во-первых, оценка парциального влияния отдельных факторов и, в-третьих, оценка парциального влияния различных комбинаций факторов. Отметим, что последнее позволяет обоснованно осуществить выбор между линейной или нелинейной аппроксимациями регрессии.

До сих пор мы предлагали генеральную дисперсию $D_0[X]$ известной и оценивали действие нового фактора, сравнивая две выборки: выборку с новым фактором и выборку без него. Но обычно имеется всего одна выборка, которая состоит из наблюдений, полученных при различных сочетаниях нескольких факторов, о влиянии которых отсутствует априорная информация. Требуется по этой выборке определить генеральную дисперсию и оценить парциальное и совместное действия всех исследуемых факторов. Тогда практически задача состоит в «расщеплении» общей дисперсии выборки на слагаемые, выражающие влияние факторов и остаточную дисперсию. После этого требуется проверить, значимо ли влияние факторов по отдельности и в комбинациях, и на основе результатов проверки отобрать факторные комбинации, существенно влияющие на исследуемую переменную, для их дальнейшего детального изучения.

Планирование и проведение эксперимента, а также схема обсчета данных существенно зависят от числа исследуемых факторов, от количества градаций (уровней) каждого из них, от количества повторных (параллельных) испытаний, от того все или только некоторые сочетания факторов на всех уровнях исследуются. В соответствии с этим выделяют следующие виды ДА.

В зависимости от количества факторов (К): однофакторный (К=1), двухфакторный (К=2), трехфакторный (К=3) и т.д. – многофакторный (мультифакторный, при произвольном К) ДА.

В зависимости от количества градаций (m) каждого из факторов выделяют ДА на двух-, трех-, четырех-, и т.д. уровнях. В этой связи обычно говорят об уровнях организации (планирования) К - факторного эксперимента, т.е. о ДА Km , где К – число факторов, а m – число их градаций. Обычно стремятся, чтобы m было для всех К одинаково, это значительно упрощает ДА.

В зависимости от того, есть повторные испытания при каждом сочетании факторов на каждом уровне или они отсутствуют, выделяют ДА без повторных (параллельных) испытаний и ДА с повторными испытаниями. В последнем случае также стремятся, чтобы число повторных испытаний (n) для всех K_i m_j было одним и тем же. Но в общем случае оно может быть переменным, и это несколько усложняет расчеты.

Наконец, в зависимости от того, все ли сочетания факторов на всех уровнях (K_i m_j) используются в исследовании или же часть таких сочетаний пропущена, выделяют полный факторный и дробный факторный ДА.

Рассмотрим принципиальную схему такого ДА, когда имеется всего один фактор А, исследуемый на произвольном количестве (m) уровней. При этом, в сущности, не важно, какие значения принимает фактор А на каждом из рассматриваемых уровней, т.е. не важно, как эти уровни количественно определенны, лишь бы это было не двусмысленно. Одним из уровней может быть и отсутствие фактора А. Сначала рассмотрим ДА при одинаковом количестве параллельных испытаний на каждом из уровней.

При одинаковом количестве параллельных испытаний мы имеем схему эксперимента $m \times n$, где mчисло уровней фактора A_j (j= 1,2,3,…,m) и n – количество параллельных испытаний на каждом j – уровне. Результаты этой схемы табулируются, как показано в верхней части табл. 1.

Прежде всего, необходимо убедиться в стабильности дисперсий по уровням, применяя критерии Бартлетта или Кохрана. Если дисперсии нестабильны, их нужно стабилизировать. Дальнейшая процедура состоит в разложении общей дисперсии совокупности на части, соответствующие рассеиванию внутри уровней и между уровнями. Для этого необходимо вычислить и проанализировать соотношение определенных сумм квадратов, показанных в нижней части табл.1. Общая дисперсия выборки, представленной в табл. 1, определяется как: $D[X] = \frac{1}{N-1}\sum_{j=1}^{m}\sum_{i=1}^{n}(x_{ij} - \bar{\bar{x}})^2$

Разложим двойную сумму в формуле на составляющие ее части, одна из которых выражает рассеивание внутри уровней (столбцов табл. 1.), а другая – рассеивание между уровнями.

Таблица 1- Исходная таблица однофакторного ДА при одинаковом количестве повторных испытаний.

Номера повторных испытаний	Уровни (j) фактора А (j=1,2,3,…,m)					
	A1	A2	A3	...	Am	
1	X11	X12	X13	…		
2	X21	X22	X23	…	X2m	
3	X31	X32	X33	…	X3m	
…	…	…	…	…	…	N=mn
n	Xn1	Xn2	Xn3	…	xnm	
$\sum_{i=1}^{n} x_{ij} = S_j$	S1	S2	S3	…	Sm	$\sum_{j=1}^{m} S_j = \sum_{j=1}^{m}\sum_{i=1}^{n} x_{ij}$
M[xj]= $\bar{x}_j$	$\bar{x}$	$\bar{x}$	$\bar{x}$	…	$\bar{x}_m$	$\bar{x} = \frac{1}{m}\sum_{j=1}^{m}\bar{x}_j = M[x_{ij}]$
$\sum_{i=1}^{n} x_{ij}^{2}$	$\sum_{i=1}^{n} x_{i1}^{2}$	$\sum_{i=1}^{n} x$	$\sum_{i=1}^{n} x^{2}_{i}$	…	$\sum_{i=1}^{n} x_{im}^{2}$	$\sum_{j=1}^{m}\sum_{i=1}^{n} x_{ij}^{2}$
$(\sum_{i=1}^{n} x_{ij})^2 = S_j^2$	S_1^2	S_2^2	S_3^2	…	S_m^2	$\sum_{j=1}^{m} S_j^2 = \sum_{j=1}^{m}(\sum_{i=1}^{n} x_{ij})^2$
$\frac{S_j^2}{n}$	$\frac{S_1^2}{n}$	$\frac{S_2^2}{n}$	$\frac{S_3^2}{n}$	…	$\frac{S_{mn}^2}{n}$	$\frac{1}{n}\sum_{j=1}^{m} S_j^2$

Это разложение представлено следующим уравнением:

$$\sum_{j=1}^{m}\sum_{i=1}^{n}(x_{ij} - \bar{\bar{x}})^2 = \sum_{j=1}^{m}\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2 + n\sum_{j=1}^{m}(\bar{x}_j - \bar{\bar{x}})^2$$

Преобразуя члены этого уравнения и вводя обозначения, получаем рабочие формулы для вычисления сумм квадратов отклонений:

$$Q = \sum_{j=1}^{m} \sum_{i=1}^{n} (x_{ij} - x)^2 = \sum_{j=1}^{m} \sum_{i=1}^{n} x_{ij}^2 - \frac{1}{N}(\sum_{j=1}^{m} \sum_{i=1}^{n} x_{ij})^2,$$

$$Q_0 = \sum_{j=1}^{m} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2 = \sum_{j=1}^{m} \sum_{i=1}^{n} x_{ij}^2 - \frac{1}{n}\sum_{j=1}^{m} (\sum_{i=1}^{n} x_{ij})^2,$$

$$Q_A = n\sum_{j=1}^{m} (\bar{x}_j - \bar{\bar{x}})^2 = \frac{1}{n}\sum_{j=1}^{m} (\sum_{i=1}^{n} x_{ij})^2 - \frac{1}{N}(\sum_{j=1}^{m} \sum_{i=1}^{n} x_{ij})^2$$

$Q=Q_0+Q_A$

Теперь, чтобы получить соответствующие дисперсии, нужно последнее равенство разделить на число степеней свободы.

Создатель дисперсионного анализа Р. Фишер предложил вычислять дисперсии с учетом их числа степеней свободы, а именно:

$$D[X] = \frac{Q}{mn - 1},$$

$$D_0[X] = \frac{Q_0}{m(n-1)},$$

$$D_A[X] = \frac{1}{n}[\frac{Q_A}{m-1} - \frac{Q_0}{m(n-1)}]$$

При таком расчете дисперсий общая дисперсия есть средняя арифметическая случайной и факторной дисперсий, взвешенных по числу степеней свободы:

$$D[X] \overset{?}{=} \frac{mn - m}{mn - 1} D_0[X] + \frac{m - 1}{mn - 1} D_A[X].$$

При $m \geq 2$ и $n \geq 2$ дробные множители перед $D_0[X]$ и $D_A[X]$ в уравнении всегда меньше единицы, т.е. всегда

$$D[X] < D_0[X] + D_A[X]$$

И основное уравнение дисперсионного анализа оказывается нарушенным. Заметим при этом, что общая дисперсия $D[X] = \frac{Q}{mn - 1}$ вычисляется правильно. Следовательно, по методике Фишера всегда получаются завышенные значения $D_0[X]$ и $D_A[X]$. К сожалению, методика Фишера излагается почти во всех руководствах по математической статистике.

Учитывая это, мы сначала будем рассматривать для примера метод Фишера, а затем излагать ДА так, чтобы основное уравнение не нарушалось.

Результаты ДА всегда представляют в виде таблицы, где указаны вариативность значения сумм квадратов отклонений количество степеней свободы, оцениваемые дисперсии, а также часто граничное значение F – критерия и доверительная вероятность. Результаты ДА по методике Фишера представлены в таблице 2. В табл. 3 показаны результаты ДА по исправленной методике вычисления дисперсий. Можно видеть, что различия между этими методиками, как и отмечалось, затрагивают только способ вычисления дисперсий и порядок проверки дисперсионного отношения, но не затрагивают основной процедуры расчета сумм квадратов отклонений.

Таблица2

Табулированные результаты однофакторного ДА по методу Фишера при одинаковом количестве повторных испытаний на всех уровнях.

вариативность	Суммы Квадратов отклонений	Кол-во Степеней свободы	Среднее Квадратов отклонение	Оцениваемые Компоненты дсперсии	Значения F – критерия
Между уровнями (факторная)	Q_A	m-1	$\frac{Q_A}{m-1}$	$nD_A[X]+D_0[X$	$F_q =$
Внутри уровней (остаточная)	Q_0	mn-m=m(n--1)	$\frac{Q_0}{m(n-1)}$	$D_0[X]$	$= \frac{nD_A[X]+D_0[X]}{D_0[X]}$
Общая	Q	mn-1	$\frac{Q}{mn-1}$		

Общее количество степеней свободы для вычисления всех дисперсий в табл. 3 одинаковое: mn-1. Тогда искомые дисперсии определяются следующим уравнениями:

$$D_A[X] = \frac{Q_A}{mn-1}$$

$$D_0[X] = \frac{Q_0}{mn-1}$$

$$D[X] = \frac{Q}{mn-1}$$

Таблица .3

Табулированные результаты однофакторного ДА по исправленному методу при одинаковом количестве повторных испытаний на всех уровнях.

Вариативность	Суммы Квадратов отклонений	Кол-во Степеней свободы	Оцениваемая дисперсия	Значение F - критерия	Вероятность ошибки второго рода
Между Уровнями (факторная)	Q_A	mn-1	DA[X]	$F_q = \frac{D_0[X] + nD_A[X]}{D_0[X]}$	$\beta \le 0{,}01$
Внутри Уровней (случайная)	Q0	mn-1	D0[X]		Или
Общая	Q	mn-1	D[X]		$\beta \le 0.05$

По Фишеру сначала определяют дисперсионное отношение оцениваемых компонент дисперсии, а затем вычисляют сами дисперсии (см. табл. 2). По исправленной методике ДА сначала непосредственно из сумм квадратов отклонений вычисляют выборочные оценки дисперсий и лишь потом рассчитывают дисперсионное отношение непосредственно через $D_0[X]$ и $D_A[X]$. При этом, естественно, величину F – критического следует определять по таблицам F – распределения для того же числа степеней свободы, что и по Фишеру: для суммы дисперсий в числителе m – 1, а для знаменателя mn – 1. Если отношение $\frac{1}{D_0[X]}(D_0[X] + nD_A[X]) < F$ - критического, то влияние фактора А принимается как несущественное. Тогда в качестве оценки генеральной дисперсии вместо величины $D_0[X]$ можем принять общую дисперсию D[X], которая обусловлена только случайными флюктуациями внутри и между уровнями. Если значение $\frac{1}{D_0[X]}(D_0[X] + nD_A[X]) > F$ - критического, то следует отбросить гипотезу о несущественном влиянии фактора А и принять альтернативную гипотезу о существенном его влиянии.

При неодинаковом количестве повторных испытаний схема расчета несколько усложняется. Имеем план эксперимента: $\sum_{j=1}^{m} n_j$, где n_j – объем выборки при j-м уровне фактора А.

Результаты эксперимента табулируются, так же как и в предыдущем случае (табл. 1), но добавляется еще одна строка для значений n_j. Несколько видоизменяются формулы для расчета сумм квадратов отклонений. Они теперь имеют вид:

$$Q = \sum_{j=1}^{m} \sum_{i=1}^{n_j} (x_{ij} - \bar{\bar{x}})^2 = \sum_{j=1}^{m} \sum_{i=1}^{n_j} x_{ij}^2 - \frac{1}{N}(\sum_{j=1}^{m} \sum_{i=1}^{n_j} \ )^2 ,$$

$$Q_0 = \sum_{j=1}^{m} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2 = \sum_{j=1}^{m} \sum_{i=1}^{n_j} x_{ij}^2 - \sum_{j=1}^{m} \frac{1}{n_j}(\sum_{i=1}^{n_j} x_{ij})^2$$

$$Q_A = n\sum_{j=1}^{m} (\bar{x}_j - \bar{\bar{x}})^2 = \sum_{j=1}^{m} \frac{1}{n_j}(\sum_{i=1}^{n_j} x_{ij})^2 - \frac{1}{N}(\sum_{j=1}^{m} \sum_{i=1}^{n_j} x_{ij})^2 ,$$

где $$N = \sum_{j=1}^{m} n_j , \qquad n = \frac{1}{m}\sum_{j=1}^{m} n_j .$$

Если пользоваться методом Фишера, то в качестве оцениваемых компонентов дисперсии имеем

$$\frac{Q_A}{m-1} = (\frac{1}{m}\sum_{j=1}^{m} n_j)D_A[X] + D_0[X],$$

$$\frac{Q_0}{m(n-1)} = D_0[X]$$

$$F = \frac{1}{D_0[X]}\left\{(\frac{1}{m}\sum_{j=1}^{m} n_j)D_A[X] + D_0[X] - \right\}$$

Ниже мы будем для расчета дисперсий пользоваться исправленным методом, а дисперсионное отношение вычислять по Фишеру.

Основное уравнение ДА при двух факторах A и B, влияющих, как предполагается априори, на среднее значение изучаемой переменной X, в общем случае выглядит так:

$$D[X] = D_A[X] + D_B[X] + D_{AB}[X] + D_0[X],$$

где $D_A[X]$ – компонент общей дисперсии (D[X]), обусловленный парциальным влиянием фактора A; $D_B[X]$ – компонент общей дисперсии, обусловленный парциальным влиянием фактора B; $D_{AB}[X]$ – компонент, обусловленный совместным влиянием (взаимодействием) факторов A и B; $D_0[X]$ – остаточная дисперсия. Член $D_{AB}[X]$ удается определить лишь при наличии повторных испытаний. Но в ряде случаев практически важно по минимальному количеству наблюдений выявить, существенно ли общее и парциальное влияние двух факторов без учета их взаимодействия. Поэтому сначала рассмотрим схему двухфакторного ДА, когда повторные наблюдения отсутствуют, затем изложим схему с повторными наблюдениями.

При отсутствии повторных испытаний и произвольном количестве уровней каждого из двух факторов имеем схему испытаний t × q, где t – количество уровней фактора A и q – количество уровней фактора B. Соответственно имеются уровни $A_j\,(j = 1,2,3\ldots,t)$ и уровни $B_i\,(j = 1,2,3\ldots,q)$ - всего tq испытаний.

Если влияние факторов существенно, то оно должно сказываться на вариативности между столбцами, обусловленной фактором A, и на вариативности между строками, обусловленной фактором B. Поэтому для расчета факторных дисперсий необходимо определить суммы квадратов отклонений по столбцам и по строкам, кроме того, нужны соответствующие суммы для общей и остаточной дисперсии. Эти четыре суммы квадратов отклонений вычисляются по следующим формулам:

$$Q = \sum_{i=1}^{q}\sum_{j=1}^{t} x_{ij}^2 - \frac{1}{qt}\left(\sum_{qt}^{t} S_j\right)^2,$$

$$Q_A = \frac{1}{q}\sum_{j=1}^{t}\left(S_j\right)^2 - \frac{1}{qt}\left(\sum_{j=1}^{t} S_j\right)^2$$

$$Q_B = \frac{1}{t}\sum_{i=1}^{q}\left(S_i'\right)^2 - \frac{1}{qt}\left(\sum_{j=1}^{t} S_j\right)^2$$

$$Q_0 = \sum_{i=1}^{q}\sum_{j=1}^{t} x_{ij}^2 - \frac{1}{t}\sum_{i=1}^{q}\left(S_i'\right)^2 - \frac{1}{q}\sum_{j=1}^{t}\left(S_j\right)^2 + \frac{1}{qt}\left(\sum_{j=1}^{t} S_j\right)^2$$

где Q – общая сумма квадратов отклонений; Q_A и Q_B - суммы квадратов отклонений по факторам A и B; Q_0 - остаточная сумма. Легко видеть, что выполняется равенство

$$Q = Q_A + Q_B + Q_0$$

Вычисление отдельных слагаемых, образующих суммы, не представляет затруднений. Отметим лишь, что на основе равенства суммы по столбцам и суммы по строкам член $\frac{1}{qt}\left(\sum_{j=1}^{t} S_j\right)^2$ может быть заменен равными ему членом $\frac{1}{qt}\left(\sum_{j=1}^{q} S'_i\right)^2$.

Таблица 4-Первичные данные для двухфакторного ДА без повторных наблюдений

	A_1	A_2	…	A_j	…	A_t	Суммы По строкам	Квадраты Сумм по строкам
B_1	X_{11}	X_{12}	…	X_{1j}	…	X_{1t}	S_1'	$(S_1')^2$
B_2	X_{21}	X_{22}	…	X_{2j}	…	X_{2t}	S_2'	$(S_2')^2$
…	…	…	…	…	…	…	…	…
B_i	X_{i1}	X_{i2}	…	X_{ij}	…	X_{it}	S_i'	$(S_i')^2$
…	…	…	…	…	…	…	…	…
Bq	X_{q1}	X_{q2}	…	X_{qj}	…	X_{qt}	S_q'	$(S_q')^2$
Суммы По столбцам	S_1	S_2	…	S_j	…	S_t	$qt = \sum_{q=1}^{t} S_j = \sum_{i=1}^{q} S_i'$	$\sum_{i=1}^{q} (S_1')^2$
Квадраты Сумм по столбцам	$S_1^{\ 2}$	$S_2^{\ 2}$	…	$S_j^{\ 2}$	…	$S_t^{\ 2}$	$\sum_{i=1}^{t} (S_j)^2$	

По Фишеру, в рассматриваемом случае учитываются следующие степени свободы. Для факторных дисперсий – по числу уровней фактора минус одна: t – 1 и q – 1. Для остаточной дисперсии – общее количество степеней свободы минус суммарное число уровней без единицы: tg – (q+t-1)=(t – 1) (q – 1). Для общей суммарное число степеней свободы без одной: tg – 1. Легко видеть что

tg – 1 = (t – 1)(q – 1)+ (t – 1) + (q – 1).

по методу Фишера оцениваемые компоненты дисперсии вычисляются следующим образом:

$$\frac{Q_A}{t-1} = qD_A[X] + D_0[X],$$
$$\frac{Q_B}{q-1} = tD_B[X] + D_0[X],$$
$$\frac{Q_0}{(t-1)(q-1)} = D_0[X]$$

На основе правых частей равенств и соответствующего количества степеней свободы вычисляются и проверяются дисперсионные отношения:

$$F^{(A)}_{P;q-1/(t-1)(q-1)} = \frac{qD_A[X] + D_0[X]}{D_0[X]}$$
$$F^{(B)}_{P;q-1/(t-1)(q-1)} = \frac{tD_B[X] + D_0[X]}{D_0[X]}$$

Схема трехфакторного дисперсионного анализа похожа на схему двухфакторного ДА с одинаковым количеством повторных наблюдений, только вместо них берется n уровней третьего фактора $C_к$, где к = 1,2,…,n.

Если при трех факторах отсутствуют повторные наблюдения, то нельзя определить взаимодействие ABC. поэтому начнем рассмотрение сразу со схемой трехфакторного эксперимента, в котором при каждой комбинации факторов имеются повторные наблюдения. Итак, имеем три фактора: A_j (j=1,2,…,t), B_i (i=1,2,…,g) и C_k (1,2,…,n). Кроме того, для каждой комбинации уровней A_j B_i C_k имеется l повторных наблюдений значений X_{ijkm} переменной X(m =1,2,…,l) - всего tgnl наблюдений.

Таблица 6 - Первичные суммы для вычисления сумм квадратов центральных отклонений.

	A1		A2		
	C1	C2	C1	C2	
B_1	$\sum A_1B_1C_1$	$\sum A_1B_1C_1$	$\sum A_2B_1C_1$	$\sum A_2B_1C_2$	$\sum B_1$
B_2	$\sum A_1B_2C_1$	$\sum A_1B_2C_2$	$\sum A_2B_2C_1$	$\sum A_2B_2C_2$	$\sum B_2$
	$\sum A_1C_1$	$\sum A_1C_1$	$\sum A_2C_1$		Общая сумма всех наблюдений
	$\sum A_1$		$\sum A_2$		
	$\sum C_1$		$\sum C_2$		
	$\sum A_1B_1$	$\sum A_1B_2$	$\sum A_2B_1$	$\sum A_2B_2$	
	$\sum B_1C_1$	$\sum B_1C_2$	$\sum B_2C_1$	$\sum B_2C_2$	

Примечание:
$\sum A_j = \sum A_jC_1 + \sum A_jC_2$; $\sum \mathrm{C_k} = \sum A_1C_k + \sum A_2C_k$; $\sum \mathrm{A}_jB_iC_1 + \sum A_jB_iC_2$; $\sum \mathrm{B}_iC_k = \sum A_1B_iC_k$ + $\sum A_2B_iC_k$.

Пусть наблюдения сгруппированы побочно (в блоке l повторных наблюдений) в черновых таблицах. Суммируя значения X_{ijkm} в каждом из блоков, определим суммы $\sum_{m=1}^{l} A_j B_i C_k$, которые запишем в табл. 6. Затем вычислим остальные суммы для комбинаций факторов по два и для парциального действия факторов. Эти суммы для простоты обозначены и рассчитываются так, как показано в табл. 6.

Суммы квадратов центральных отклонений вычисляются по следующим формулам:

$$Q_A = \frac{1}{ngl}\sum_j \left(\sum A_j\right)^2 - \frac{1}{tgnl}\sum {}^2 общ$$

$$Q_B = \frac{1}{tnl}\sum_i \left(\sum B_i\right)^2 - \frac{1}{tngl}\sum {}^2 общ$$

$$Q_C = \frac{1}{tgl}\sum_k \left(\sum C_k\right)^2 - \frac{1}{tngnl}\sum {}^2 общ$$

$$Q_{AB} = \frac{1}{nl}\sum_j \sum_i \left(\sum A_j B_i\right)^2 - Q_A - Q_B - \frac{1}{tgnl}\sum {}^2 общ$$

$$Q_{AC} = \frac{1}{gl}\sum_j \sum_k \left(\sum A_j B_k\right)^2 - Q_A - Q_C - \frac{1}{tgnl}\sum {}^2 общ$$

$$Q_{BC} = \frac{1}{tl}\sum_i \sum_k \left(\sum B_i C_k\right)^2 - Q_B - Q_C - \frac{1}{tgnl} общ$$

$$Q_{ABC} = \frac{1}{l}\sum_j \sum_i \sum_k \left(\sum A_j B_i C_k\right)^2 - Q_A - Q_B - Q_C - Q_{AB} - Q_{AC} - Q_{BC} - \frac{1}{tgnl}\sum {}^2 общ$$

$$Q_0 = \sum_J \sum_I \sum_k \left(x_{ijklm}\right)^2 - Q_A - Q_B - Q_C - Q_{AB} - Q_{AC} - Q_{BC} - Q_{ABC} - \frac{1}{tgnl}\sum {}^2 общ$$

$$Q = \sum_j \sum_i \sum_k \sum_m \left(x_{ijkm}\right)^2 - \frac{1}{tngl}\sum {}^2 общ$$

Нетрудно видеть, что основное уравнение дисперсионного анализа в системе уравнений выполняется:

$$Q = Q_A + Q_B + Q_C + Q_{AB} + Q_{AC} + Q_{BC} + Q_{ABC} + Q_0$$

Определяя соответствующие дисперсии, следует разделить равенство почленно на общее число степеней свободы (tgnl – 1). источники вариативности, количество степеней свободы и дисперсионные отношения в общем виде записаны в табл. 5. Процедуры расчета и проверки значимости компонент общей дисперсии аналогичны рассмотренным выше.

Таблица 5 - Табулированные результаты трехфакторного дисперсионного анализа с повторными наблюдениями.

Источники вариативности	Количество степеней свободы	Дисперсионные отношения
A	**t – 1**	$F^{(A)} = \frac{1}{D_0} \cdot (gnD_A + nD_{AB} + gD_{AC} + lD_{ABC} + D_0)$
B	**g-1**	$F^{(B)} = \frac{1}{D_0} \cdot (tnD_B + nD_{AB} + tD_{BC} + lD_{ABC} + D_0)$
C	**n-1**	$F^{(C)} = \frac{1}{D_0} \cdot (tgD_C + gD_{AC} + tD_{BC} + lD_{ABC} + D_0)$
AB	**(t-1)(g-1)**	$F^{(AB)} = \frac{1}{D_0}(nD_{AB} + lD_{ABC} + D_0)$
AC	**(t-1)(n-1)**	$F^{(AC)} = \frac{1}{D_0}(gD_{AC} + lD_{ABC} + D_0)$
BC	**(g-1)(n-1)**	$F^{(BC)} = \frac{1}{D_0}(tD_{BC} + lD_{ABC} + D_0)$
ABC	**(t-1)(g-1)(n-1)**	$F^{(ABC)} = \frac{1}{D_0}(lD_{ABC} + D_0)$
Случайные	**Tgn(l-1)**	
итого	**Tgnl-1**	

Примечание: обозначение дисперсий сокращенны: D_A - вместо $D_A[X]$, D_{AB} – вместо $D_{AB}[X]$ и т.п.

1. Изложенная схема трехфакторного анализа с повторными наблюдениями без труда упрощается до схемы без повторных наблюдений: в табл. 4. все $\sum A_j B_i C_k$ - заменяются одним числом; а Q_{ABC} выпадает, так как ее нельзя выделить из Q_0; из табл. 5 выпадают значения lD_{ABC} и $F^{(ABC)}$.

2.Для упрощения расчетов дисперсионного комплекса полезно применять линейные преобразования исходной переменной $x_{ijklm} = ay_{ijkm} + b$

т.е. перейти к новой переменной $y_{ijkm} = \frac{x_{ijkm} - b}{a}$ с переносом начала отсчета b и изменением масштаба. Следует помнить, что при этом значение Q_i оказывается измененными в a^2 раз (уменьшенными, если a>1, и увеличенными, если a<1), и это необходимо учитывать при окончательном расчете дисперсий D_i..

Метрическое и неметрическое многомерное шкалирование. Матрица сходств и различий. Построение пространственной модели стимулов: Метод ортогональных проекций. Понятие метрики. Метрическая модель. Кластерный анализ.

Главная задача многомерного шкалирования – найти минимальное число субъективных признаков, определяющих различение стимулов человеком, и вычислить значение признаков, которыми характеризуются данные стимулы. Решение задачи многомерного шкалирования основано на использовании понятия психологического пространства, точки которого представляют исходные стимулы. Аналогично геометрическим представлениям вводится система координат, число которых определяется числом простых субъективных признаков. Это число задает размерность психологического пространства. Оси координат представляют собой шкалы соответствующих субъективных признаков, и положение точек-стимулов в пространстве задано шкальными значениями признаков. Число субъективны шкал и шкальные значения стимулов характеризуют пространственную модель многомерного шкалирования. Метрическая модель строится, когда определяется субъективное расстояние на основе сходства и различия между стимулами в психологическом пространстве.

Формально общая задача многомерного шкалирования выражается следующим образом. По заданной симметричной матрице различий между стимулами $D = \begin{pmatrix} D_{11} & \dots & D_{1n} \\ \dots & \dots & \dots \\ D_{n1} & \dots & D_{nn} \end{pmatrix}$ нужно построить метрическую и пространственную модели стимулов, т.е. определить размерность пространства и координаты точек-стимулов в этом пространстве $X = \begin{pmatrix} X_{11} & \dots & X_{1n} \\ \dots & \dots & \dots \\ X_{n1} & \dots & X_{nn} \end{pmatrix}$ таким образом, чтобы матрица расстояний, вычисленных между точками на основании метрической модели расстояния $d = \begin{pmatrix} d_{11} & \dots & d_{1n} \\ \dots & \dots & \dots \\ d_{n1} & \dots & d_{nn} \end{pmatrix}$ была бы в смысле некоторого критерия возможно более близка к исходной матрице различий.

Модели различий в многомерном шкалирование достаточно широко применяются в психологии, главным образом для изучения индивидуальной специфики оценок сложных стимулов различными людьми. В результате кластерного анализа при помощи предварительно заданных переменных формируются группы наблюдений. Под наблюдениями здесь понимаются отдельные личности (респонденты) или любые другие объекты. Члены одной группы (одного кластера) должны обладать схожими проявлениями переменных, а члены разных групп различными.

Самой распространенной мерой для определения расстояния между двумя точками на плоскости, образованной координатными осями x и y, является евклидова мера:

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

При кластерном анализе с тремя переменными можно ввести ещё одну ось — ось z и рассматривать размещение наблюдений, а также проводить расчёт расстояния по формуле евклидовой меры в трёхмерном пространстве.

При наличии более трёх переменных определение расстояния между двумя точками x и y в любом n-мерном пространстве для математиков не представляет особого труда. Формула Евклида в таких случаях приобретает следующий вид:

$$\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

Кластерный анализ объединяет различные процедуры, используемые для проведения классификации. В результате применения этих процедур исходная совокупность объектов разделяется на кластеры или группы (классы) схожих между собой объектов. Под кластером обычно понимают группу объектов, обладающую свойством плотности (плотность объектов внутри кластера выше, чем вне его), дисперсией, отделимостью от других кластеров, формой (например, кластер может иметь очертания гиперсферы или эллипсоида), размером. Конечно, данное определение не является строгим (его вообще не существует). Если вы представите созвездия на небе или горы на географической карте – это и будут кластеры.

Наиболее часто методы кластерного анализа используют в социологии, биологии, медицине, археологии, маркетинговых исследованиях, экономике.

Сложность задач кластерного анализа состоит в том, что реальные объекты являются многомерными, то есть описываются не одним, а несколькими параметрами, и объединение объектов в группы проводится в пространстве многих измерений, что весьма непросто. Кроме того, данные могут носить нечисловой характер. В целом методы кластеризации делятся на агломеративные (агломерат – скопление) и итеративные дивизивные (division – деление, разделение).

В агломеративных, или объединительных, методах происходит последовательное объединение наиболее близких объектов в один кластер. Процесс такого последовательного объединения можно показать на графике в виде дендограммы, или дерева объединения.

Исходными данными для анализа могут быть собственно объекты и их параметры, например, марки машин и их параметры: цена, время для разгона до скорости 60 миль/час и другие.

Данные для анализа могут быть представлены матрицей расстояний между объектами, в которой на пересечении строки с номером i и столбца с номером j записано расстояние между объектами. Переход от объектов к расстоянию между объектами – важный момент.

Расстояние между объектами – одна из мер сходства. Интуитивно понятно, что чем меньше расстояние между объектами, тем они более схожи. Но как выбрать естественную метрику, то есть как естественно для данной задачи измерить расстояние между объектами?

Часто используют евклидову метрику, например, если объект описывается двумя параметрами, то он может быть изображён точкой на плоскости, а расстояние между объектами – это расстояние между точками, вычисленное по теореме Пифагора.

Если вы не будете возводить в квадраты координатные расстояния, а просто возьмёте их абсолютные значения и просуммируете, то получите так называемое манхэттенское расстояние, или «расстояние городских кварталов». Такое расстояние связано с перемещением человека по улицам города, а не с движением по ровной местности (перемещаться можно только по линиям, параллельным осям координат в декартовой системе координат).

Применение только евклидовой метрики или манхэттенской метрики, и даже какой угодно другой метрики – метрика Минковского, не решает всех проблем. Например, при проведении массового опроса вы имеете ответы типа «хуже – лучше», или «да – нет», или «да – нет – не знаю» и т.д. И как быть с такого рода данными – неясно, они существенно отличаются от чисел, с которыми мы привыкли иметь дело в математике. Такого рода данные не всегда естественно представлять точками в евклидовом пространстве.

В реальных задачах евклидова метрика может оказаться вовсе не подходящей. В этих задачах используется понятие «меры сходства» объектов (расстояние – одна из мер сходства). Часто эта мера сходства является эмпирической, сходство измеряется непосредственно. Например, проводится массовый опрос: сходно ли ваше отношение к двум явлениям политической жизни? или к двум товарам? к двум цветам? И констатация фактов приводит к мере сходства.

Важной мерой сходства, которая традиционно используется в социальных науках, являются статистические коэффициенты корреляции, например, коэффициент корреляции Пирсона.

Для бинарных данных часто просто вычисляют количество параметров, которые совпадают у объектов. Далее это число делят на общее число параметров и получают меру сходства. Меры, построенные таким образом, называют коэффициентами ассоциативности.

При проведении кластерного анализа отдельные кластеры могут формироваться при помощи пошагового слияния, для которого существует ряд различных. Важную роль играют иерархические и партиционные методы, причём последние применяются в подавляющем большинстве случаев. В иерархических методах каждое наблюдение образовывает сначала свой отдельный кластер. На первом шаге два соседних кластера объединяются в один; этот процесс может продолжаться до тех пор, пока не останутся только два кластера. расстояние между кластерами является средним значением всех расстояний между всеми возможными парами точек из обоих кластеров.

Дистанционные меры и меры подобия зависят от вида переменных, участвующих в анализе, то есть выбор меры зависит от типа переменной и шкалы, к которой она относится: интервальная переменная, частоты или бинарные (дихотомические) данные.

3.2 Методические указания для подготовки к практическим занятиям

При подготовке к практическим занятиям целесообразно пользоваться планом, представленным в первом разделе.

Тщательно проработать лекционный материал и соответствующие учебные пособия по теме каждого практического занятия. Прорешать типовые задачи домашнего задания и при необходимости обратиться к преподавателю за консультацией. При подготовке к практическим занятиям рекомендуется опираться на контрольные вопросы:

Занятие № 1

1. Предмет и содержание прикладной статистики.
2. Характер данных, встречающихся в педагогической практике. 3. 3. Понятие эксперимента, измерения.
4. Бинарные отношения и их свойства. Примеры.
5. Гомоморфизм, структура, шкала. Примеры.
6. Классификация шкал.
7. Неметрические шкалы: наименований и порядка. Примеры.
8. Метрические шкалы: интервалов и отношений. Примеры.

Литература

1. Ермолов. О.Ю. Математическая статистика для психологов, 2003г.

2. Сидоренко Е.В. Методы математической обработки в психологии. СПб.: ООО «Речь», 2001г.

Занятие № 2

1. Понятие случайной величины.
2. Основные числовые характеристики случайных величин.
3. Статистическая оценка.
4. Понятие доверительного интервала.
5. Сгруппированный ряд.
6. Гистограмма.

Литература

1. Двоерядкина Н.Н., Киселева А.Н., Юрьева Т.А. Математические методы в психологии и социологии: Практикум. Благовещенск: Амурский гос. ун-т, 2005г.

2. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике: Учебное пособие для вузов. М.:Высш. шк., 2001г.

Занятие № 3

1. F-критерий Фишера. Понятие и назначение.
2. t-критерий Стьюдента. Понятие и назначение.

Литература

1. Двоерядкина Н.Н., Киселева А.Н., Юрьева Т.А. Математические методы в психологии и социологии: Практикум. Благовещенск: Амурский гос. ун-т, 2005г.

2. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике: Учебное пособие для вузов. М.:Высш. шк., 2001г.

Занятие № 4

1. Понятие функциональной, статистической и корреляционной связи.

2. Меры связи случайных величин.
3. Классификация корреляционных связей.
4. Оценка достоверности корреляционной связи.
5. Основные задачи регрессионного анализа.
6. Уравнение регрессии. Интерпретация коэффициентов уравнения парной линейной регрессии.

Литература

1. Ермолов. О.Ю. Математическая статистика для психологов, 2003г.

2. Двоерядкина Н.Н., Киселева А.Н., Юрьева Т.А. Математические методы в психологии и социологии: Практикум. Благовещенск: Амурский гос. ун-т, 2005г.

3. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике: Учебное пособие для вузов. М.: Высш. шк., 2001г.

Занятие № 5

1. Понятие множественной регрессии.
2. Линейная и нелинейная модели множественной регрессии.
3. Оценка достоверности коэффициентов линейной и нелинейной множественной регрессии.

Литература

1. Ермолов. О.Ю. Математическая статистика для психологов, 2003г.

2. Сидоренко Е.В. Методы математической обработки в психологии. СПб.: ООО «Речь», 2001г.

3. Двоерядкина Н.Н., Киселева А.Н., Юрьева Т.А. Математические методы в психологии и социологии: Практикум. Благовещенск: Амурский гос. ун-т, 2005г.

Занятие № 6

1. Назначение и ограничение в применение критериев различий.
2. Назначение и ограничение в применение критериев изменений.
3. Алгоритм проверки статистических гипотез.

Литература

1. Ермолов. О.Ю. Математическая статистика для психологов, 2003г.

2. Сидоренко Е.В. Методы математической обработки в психологии. СПб.: ООО «Речь», 2001г.

3. Двоерядкина Н.Н., Киселева А.Н., Юрьева Т.А. Математические методы в психологии и социологии: Практикум. Благовещенск: Амурский гос. ун-т, 2005г.

Занятие № 7

3. Основные задачи дисперсионного анализа.
4. Ограничения в применении дисперсионного анализа.

Литература

1. Ермолов. О.Ю. Математическая статистика для психологов, 2003г.

2. Гусев А.Н. Дисперсионный анализ в экспериментальной психологии. М.:уч.-мет. Коллектор «Психология», 2000г.

3. Айвазян С. А., Мхитарян В. С. Прикладная статистика: Учебник для вузов – М.: ЮНИТИ, 1998г.

4. Шеенко П.С. Практикум по математической статистике для нематематических факультетов: Учебное пособие. – Комсомольск-на-Амуре: ГПУ, 2001г.

Занятие № 8

1. Понятие и назначение факторного анализа.
2. Понятие фактора, факторной нагрузки.
3. Методы факторизации корреляционной матрицы.
4. Собственные значения факторов.

Литература

1. Ермолов. О.Ю. Математическая статистика для психологов, 2003г.

2. Айвазян С. А., Мхитарян В. С. Прикладная статистика: Учебник для вузов – М.: ЮНИТИ, 1998г.

3. Шеенко П.С. Практикум по математической статистике для нематематических факультетов: Учебное пособие. – Комсомольск-на-Амуре: ГПУ, 2001г.

Занятие № 9

1. Задача кластерного анализа.
2. Понятие дендрограммы, кластера.
3. Эвклидово расстояние.
4. Пространство коэффициентов корреляции.
5. Метод К средних.

Литература

1. Двоерядкина Н.Н., Киселева А.Н., Юрьева Т.А. Лабораторный практикум по математическим методам в психологии. Благовещенск: Амурский гос. ун-т, 2006г

2. Шеенко П.С. Практикум по математической статистике для нематематических факультетов: Учебное пособие. – Комсомольск-на-Амуре: ГПУ, 2001г.

3.3 Методические указания по оформлению индивидуального задания

1. Отчет пишется на стандартных листах писчей бумаги. Все листы заполняются только с одной стороны. Номера листов проставляются в верхнем правом углу. Текст на листе ограничивается рамкой: сверху и снизу — 2 — 2,5 см., слева и справа — 2 — 2.5 см.

Каждый отчет начинается с *титульного листа,* который служит обложкой работы. Сверху на нем указывается принадлежность студента к учебному заведению, факультету, специализации или кафедре. В середине листа укалывается название изучаемой темы или раздела и название учебного задания. Ниже и справа указывается фамилия и инициалы студента, номер академической группы, фамилия и инициалы преподавателя. Внизу титульного листа отмечают год выполнения работы.

Эта страница служит также для отметок преподавателя о выполнении учебного задания и замечаний по поводу подготовленного студентом отчета.

2. Структура отчета о выполнении учебного задания в практикуме

Теоретическое введение и постановка проблемы (не более 3-х листов). В данном разделе отчета дается общая характеристика изучаемого метода, его характерных особенностей, даются определения необходимых терминов.

Формулировка цели и конкретных задач работы в соответствии с общей проблемой, рассмотренной в теоретическом введении (не более 0.5 листа).

Описание методики. В этот раздел входят следующие пункты: 1) сведения об испытуемых, дата и время проведения опыта; 2) описание использованной аппаратуры и программного обеспечения; 3) описание параметров стимуляции; 4) подробное описание процедуры опыта: какие стимулы предъявлялись, в каком порядке, какие ответы и в какой форме давал испытуемый; приводится инструкция испытуемому: если опыт состоял из нескольких серий, указывается их порядок. В том случае, если процедура опыта была нарушена, указывается причина.

Если опыт проводился на компьютере, следует указать имя файла результатов.

Результаты. В этой части необходимо описать полученные данные, методы их обработки и привести основные результаты. Если использовались нестандартные способы обработки результатов, то их описанию стоит уделить особое внимание. Если использовались методы статистического анализа, то необходимо привести соответствующие формулы.

Итоговые результаты проведенных измерений сводятся в одну таблицу. Каждая таблица должна быть пронумерована и иметь соответствующее название, где нужно четко выразить основное содержание данной таблицы.

Рисунки, иллюстрирующие основное содержание работы, должны быть также пронумерованы и начерчены на координатной бумаге. В том случае, если по рисунку производится вычисление каких-либо результатов (как, например, в случае с психометрической функцией), то следует обратить внимание на выбор подходящего масштаба. Если рисунок делается с помощью компьютерной программы, то стоит позаботиться о введении координатной сетки. На графиках должны быть указаны все параметры, необходимые для однозначного понимания графика. Подрисуночные подписи и обозначения на графиках должны давать полную информацию, чтобы не возникала необходимость для понимания графика обращаться к тексту отчета.

Рисунки и таблицы рекомендуется выполнять на отдельных листах.

Обсуждение результатов и выводы должны соответствовать целям и задачам работы. В том случае, если получен нестандартный и неожидаемый результат, то, безусловно, следует уделить особое внимание его интерпретации и попытаться объяснить причины его появления. Если работа выполнялась в рамках какой-либо модели, то следует сделать четкое заключение о соответствии полученных результатов ее предположениям.

Выводы должны быть короткими и конкретными. *Литературные ссылки* оформляются в соответствии с требованиями, предъявляемыми ГОСТом к научным статьям.

4. ОРГАНИЗАЦИЯ КОНТРОЛЯ ЗНАНИЙ

4.1 Комплект заданий для контрольных работ

Контрольная работа №1

Вариант 0

Задача 1. На основании наблюдений за работой 25 кандидатов на должность секретаря-референта установлено, что в среднем они тратили 7 минут на набор одной страницы сложного текста на компьютере при выборочном стандартном отклонении S - 2 минуты. При предположении, что время (X) набора текста имеет нормальный закон распределения:

а) Оцените количество претендентов на работу, которые набрали текст быстрее, чем за 5 минут.

б) Предполагалось, что среднее время набора страницы текста должно составить 5,5 минуты. Не противоречат ли полученные данные этой гипотезе?

Задача 2. Считается, что у студентов, обучающихся на специальности «Психология» средний балл аттестата равен 4,5. Приемная комиссия отметила снижение оценок в аттестатах у некоторых абитуриентов. Для проверки этого факта вычислен средний балл аттестата у 20 случайно выбранных абитуриентов, поступающих на специальность «Психология». Результаты представлены в таблице 2.4. Можно ли утверждать, что средний балл аттестата соответствует норме?

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
ср.балл	4,2	4	5	4,5	4,3	4,6	4,8	4,1	4	4,5	4,7	4,2	4,1	4,3	4,2	4	4,6	4,5	4,2	5

Контрольная работа №2

Вариант 0

Задача 1. Одинаков ли уровень профессиональной ориентации на экономическом и социальном факультетах, если из закончивших экономический факультет 98 человек, стали работать по специальности 64 человека, а из закончивших факультет социальных наук 63 человек – 39 стали работать по специальности.

Задача 2. Психолог сравнивает два эмпирических распределения, в каждом из которых было обследовано 70 человек по тесту интеллекта. Различаются ли между собой эти два распределения?

Уровни интеллекта	60	70	80	90	100	110	120	130	140
1-е распределение	1	3	8	17	26	7	5	2	1
2-е распределение	0	1	2	19	20	13	10	4	1

Задача 3. По данным наблюдения определить, имеется ли у людей сопряженность между цветом волос и цветом глаз. Частоты совместного распределения приведены в таблице.

	светлые	русые	черные	рыжие
голубые	177	71	17	14
серые	95	119	75	25
карие	12	44	23	8

4.2 Темы индивидуальных заданий

1. Исследование факторов, определяющих специфику национального характера.

2. Исследование направленности и мотивов выбора профессии старшеклассников.

3. Исследование различий самооценки у девушек и юношей с помощью корреляционного анализа.

4. Исследование связи между вербальной агрессией и личностной тревожностью.

5. Влияния стиля руководства на климат в коллективе.

6. Мотивация «На оценку» и следования ее уровня на различных стадиях младшего школьного возраста.

7. Выявление факторов, определяющих положение человека в семье.

8. Исследование факторов влияющих на выбор страны для зарубежной поездки.

4.4 Тесты для самоконтроля

Вариант 0.

1. Бинарное отношение является отношением эквивалентности, если оно

А) симметрично, транзитивно и рефлексивно

Б) рефлексивно, симметрично

В) транзитивно, арефлексивно и полно

2. Признак, характеризующийся некоторым свойством или состоянием объекта наблюдения является

А) качественным Б) количественным

3. Любое взаимно-однозначное преобразование допустимо для

А) номинальной шкалы Б) порядковой шкалы

В) интервальной шкалы Г) шкалы отношений

4. На отношениях эквивалентности и порядка как на единственных строится шкала

А) отношений Б) порядка

В) интервальная Г) номинальная

5. Постоянная, естественная нулевая точка отсчета существует для

А) номинальной шкалы Б) порядковой шкалы

В) интервальной шкалы Г) шкалы отношений

6. Какие из указанных статистических приемов обработки допустимы для величин измеренных по шкале наименований

А) определение медианы

Б) определение средней арифметической

В) определение моды

Г) определение средней геометрической

7. Средняя арифметическая является мерой

А) положения Б) рассеяния В) связи

8. Коэффициент корреляции Пирсона является мерой

А) положения Б) рассеяния В) связи

9.Корреляционные связи по направленности бывают

А) криволинейные Б) обратные В) линейные

Г) прямые

10.t – критерий Стьюдента является

А) параметрическим Б) непараметрическим

11. Выявление различий в распределений признака осуществляется посредством метода

А) Т – критерий Вилкоксона Б) Q – критерий Розенбаума

В) χ^2 - критерий Пирсона Г) F – критерий Фишера

12. Оценка сдвига значений исследуемого признака осуществляется посредством метода

А) Т – критерий Вилкоксона Б) Q – критерий Розенбаума

В) χ^2 - критерий Пирсона Г) F – критерий Фишера

13. При выборке в 42 человека М(Х)=3,34, Д(Х)=91. Какого объема следует взять выборку, чтобы оценить среднее арифметическое значение величины Х с ошибкой не большей α =0,05. (t=1,96)

А) 16458 Б) 12485 В) 102

14. Более общим является понятие

А) корреляционная связь Б) корреляционная зависимость

15. Последовательное объединение наиболее близких объектов в один кластер происходит в:

А) агломеративных методах кластеризации

Б) итеративных дивизивных методах кластеризации

16. Методы измерения порогов чувствительности относятся к

А) нольмерному шкалированию

Б) одномерному шкалированию

В) многомерному шкалированию

17. Является ли нормальным распределение признака если A_s=0,106, E_x=-0,711, количество наблюдений n=16

А) да Б) нет

Начала теории измерений

1. Шкалой, классифицирующей по принципу «больше - меньше», является:
   А) номинальная;
   Б) порядковая;
   В) интервальная.
2. Интервальная шкала измерения классифицирует по принципу:
   А) больше на определенное количество единиц – меньше на определенное количество единиц;
   Б) больше - меньше;
   В) объекты и субъекты пропорциональны степени выраженности измеряемого свойства.
3. Единицей измерения в шкале порядка является:
   А) 1 выбор;
   Б) 1 ранг;
   В) 1 наблюдение.
4. При построении номинальной шкалы главным являются:
   А) качественные различия объектов;
   Б) количественные различия объектов;
   В) и те, и другие.
5. Примером интервальной шкалы является:
   А) шкала Кельвина;
   Б) шкала Цельсия;
   В) шкала порогов чувствительности.
6. Какие из указанных статистических приемов обработки допустимы для величин измеренных по шкале порядка:
   А) коэффициент корреляции Пирсона;
   Б) дисперсия;
   В) мода.
7. Допустимые преобразования для шкалы отношений:
   А) строго возрастающие преобразования;
   Б) линейные преобразования;
   В) преобразования подобия.
8. Шкала стенов, предложенная Р.Б. Кеттеллом относится к шкале:
   А) номинальная;
   Б) порядковая;
   В) интервальная.
9. Классификацию четырех типов шкал предложил:
   А) Кеттелл;
   Б) Стьюдент;
   В) Стивенс.
10. Данные образуют генеральную совокупность при наблюдении:
    А) внешнем;

Б) сплошном;
В) включенном.

11. Строго возрастающие преобразования допустимы для:
А) номинальной шкалы;
Б) порядковой шкалы;
В) интервальной шкалы;
Г) шкалы отношений.

12. На отношении эквивалентности как на единственном строится шкала:
А) отношений;
Б) порядка;
В) интервальная;
Г) номинальная.

13. Какие из указанных статистических приемов обработки допустимы для величин измеренных по шкале интервалов
А) определение медианы;
Б) определение средней арифметической;
В) определение моды;
Г) определение средней геометрической;

14. Бинарное отношение является отношением нестрогого порядка, если оно:
А) асимметрично, транзитивно, рефлексивно;
Б) рефлексивно, симметрично, транзитивно;
В) транзитивно, арефлексивно, симметрично.

15. Линейные преобразования допустимы для:
А) номинальной шкалы;
Б) порядковой шкалы;
В) интервальной шкалы;
Г) шкалы отношений.

16. Бинарное отношение является отношением толерантности, если оно:
А) симметрично, транзитивно и рефлексивно;
Б) рефлексивно, симметрично;
В) транзитивно, арефлексивно и полно;

17. Бинарное отношение является отношением строгого порядка, если оно:
А) асимметрично, транзитивно;
Б) рефлексивно, симметрично;
В) транзитивно, арефлексивно.

Проверка статистических гипотез. Распределение признака

1. Статистическими гипотезами называют:
А) предположения относительно величины независимых параметров распределения изучаемой генеральной совокупности (совокупностей);
Б) предположения относительно закона распределения изучаемой генеральной совокупности (совокупностей);

В) и те, и другие.

2. Цель статистической проверки гипотез состоит в том, чтобы

а) на основании выборочных данных принять решение о справедливости нулевой гипотезы H_0;

б) на основании выборочных данных принять решение о справедливости альтернативной гипотезы H_1;

в) на основании выборочных данных принять решение о несправедливости альтернативной гипотезы H_1;

3. К параметрам распределения признака не относятся:

а) математическое ожидание;

б) асимметрия;

в) мода.

4. Ошибка первого рода предполагает:

а) отклонение нулевой гипотезы, в то время как она верна;

б) принятие нулевой гипотезы, в то время как она неверна;

в) принятие альтернативной гипотезы.

5. Вероятность ошибки второго рода обозначается:

а) α;

б) β;

в) γ.

6. Закономерность встречаемости разных значений признака называют:

а) нормальным распределением;

б) распределением признака;

в) эксцессом.

7. Проверка статистических гипотез осуществляется с помощью:

а) параметрических критериев;

б) непараметрических критериев;

в) и тех, и других.

8. Основной принцип проверки статистических гипотез состоит в следующем:

а) если наблюдаемое значение критерия принадлежит критической области, то нулевая гипотеза H_0 отклоняется и принимается конкурирующая H_1;

б) если наблюдаемое значение критерия принадлежит области допустимых значений, то нулевую гипотезу H_0 нельзя отклонить;

в) если наблюдаемое значение критерия принадлежит зоне неопределенности, то мы уже можем отклонить нулевую гипотезу H_0, но еще не можем принять конкурирующую H_1;

г) все утверждения справедливы.

9. Среднее квадратическое отклонение является мерой:

А) положения;

Б) рассеяния;

В) связи.

10. Коэффициент линейной корреляции Пирсона является мерой

А) положения;
Б) рассеяния;
В) связи.

Кластерный анализ

1. Выберите верные утверждения:
   a) Кластерный анализ - форма визуализации данных;
   b) Кластерный анализ является доказательством статистической гипотезы;
   c) Кластерный анализ генерирует статистическую гипотезу.

2. Древовидная фигура, где более близкие ветви соответствуют более близким объектам или группам объектов называют:
   a) дендрограмма;
   b) гистограмма;
   c) кластер.

3. Метрика имеет вид: ρ(x,y)=1, если x≠y и ρ(x,y)=0, если x=y. Подходит ли она для кластеризации классов?
   a) да;
   b) нет.

4. Метод k-средних относится к
   a) агломеративным методам кластеризации;
   b) итеративным дивизивным методам кластеризации.

5. Кластерный анализ относится к классу методов:
   a) нольмерной статистики;
   b) одномерной статистики;
   c) многомерной статистики.

6. Последовательное объединение наиболее близких объектов в один кластер происходит:
   А) в агломеративных методах кластеризации;
   Б) в итеративных дивизивных методах кластеризации.

7. Состав групп по методу k-средних определяется таким образом, чтобы:
   а) дисперсия расстояний между группами была максимальной, а внутри групп – минимальной;
   б) дисперсия расстояний между группами была минимальной, а внутри групп – минимальной;
   в) дисперсия расстояний между группами была максимальной, а внутри групп – максимальной;
   г) дисперсия расстояний между группами была минимальной, а внутри групп – максимальной.

8. Матрица попарных расстояний может рассчитываться с помощью:

   а) Эвклидова расстояния;
   б) метрики коэффициентов корреляции;
   в) метрики для категоризированных данных;
   г) всех перечисленных.

9. К методам кластеризации относятся:

а) метод ближнего соседа;
б) метод дальнего соседа;
в) метод соединения;
г) метод k-средних.

10. По методу Уорда (Варда) группировки определяются так, чтобы:

а) дисперсия расстояний внутри групп была минимальной;
б) расстояние между группами принималось как расстояние между средними;
в) расстояние между группами принималось как среднее расстояние между всеми парами объектов, относящихся к разным группам.

5. КАРТА ОБЕСПЕЧЕННОСТИ ДИСЦИПЛИНЫ КАДРАМИ ПРОФЕССОРСКО-ПРЕПОДАВАТЕЛЬСКОГО СОСТАВА