

Учреждение образования
«Республиканский институт профессионального образования»
Факультет повышения квалификации и переподготовки

Кафедра психологии профессионального образования

СОГЛАСОВАНО
Заведующий кафедрой
Т.Н. Карпович
26.10.2020

СОГЛАСОВАНО
Декан факультета
О.А. Беляева
26.10.2020

**УЧЕБНО-МЕТОДИЧЕСКИЙ КОМПЛЕКС ПО УЧЕБНОЙ
ДИСЦИПЛИНЕ
«ОСНОВЫ ПЕДАГОГИЧЕСКИХ ИЗМЕРЕНИЙ»**

Составитель: Пашкевич Ольга Ивановна, старший преподаватель кафедры психологии профессионального образования

для специальности: 1-08 01 71 «Педагогическая деятельность специалистов»

Рассмотрено и утверждено на заседании Совета УО РИПО 03.11.2020,
протокол №10

Минск 2020

РИПО

ОГЛАВЛЕНИЕ

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА	3
1. ТЕОРЕТИЧЕСКИЙ РАЗДЕЛ	4
1.1 Сущность педагогических измерений. Основные понятия математической статистики в педагогике	4
1.2 Оценивание и измерение в педагогическом процессе	26
1.3 Качественный анализ результатов педагогического исследования	38
1.4 Определение различий между эмпирической и теоретической таблицами распределения	50
1.5 Описательные статистики	58
1.6 Параметрические методы	67
1.7 Непараметрические методы	74
2. ПРАКТИЧЕСКИЙ РАЗДЕЛ	81
2.1 Практическая работа «Качественный анализ результатов педагогического исследования»	81
2.2 Практическая работа «Определение различий между эмпирической и теоретической таблицами распределений»	81
2.3 Практическая работа «Описательные статистики»	82
3. РАЗДЕЛ КОНТРОЛЯ ЗНАНИЙ	83
3.1 Вопросы к зачету	83
3.2 Требования к содержанию самостоятельной работы слушателей	85
4. ВСПОМОГАТЕЛЬНЫЙ РАЗДЕЛ	88
4.1 Фрагмент учебно-тематического плана по специальности переподготовки 1-08 01 71 «Педагогическая деятельность специалистов»	88
4.2 Содержание учебной программы дисциплины «Основы педагогических измерений»	89
4.3 Перечень учебных изданий и информационно-аналитических материалов, рекомендуемых для изучения учебной дисциплины	91
ПРИЛОЖЕНИЯ	94

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Учебно-методический комплекс (далее – УМК) по дисциплине «Основы педагогических измерений» предназначен для обеспечения качества образовательного процесса при осуществлении переподготовки по специальности 1-08 01 71 «Педагогическая деятельность специалистов» с присвоением квалификации «Преподаватель в соответствии с квалификацией по основному образованию».

Структура УМК включает: теоретический раздел, содержащий материалы для теоретического изучения каждой из тем учебной программы; практический раздел, предусматривающий задания для проведения практических работ; раздел контроля знаний, в котором представлены материалы текущей аттестации по дисциплине в виде перечня вопросов к зачету, и средства организации самостоятельной учебной деятельности слушателей. В ЭУМК также входят содержание учебной программы дисциплины «Основы педагогических измерений», фрагмент учебно-тематического плана по специальности переподготовки 1-08 01 71 «Педагогическая деятельность специалистов» и перечень учебных изданий и информационно-аналитических материалов, которые являются источниками дополнительной информации для углубленного изучения каждой из тем. Теоретический и практический разделы ЭУМК представлены в соответствии со структурой и содержанием учебной программы дисциплины. Практический раздел включает задания для проведения трех практических работ в объеме 6 часов. В приложениях содержатся материалы, представляющие собой практическое руководство к выполнению каждой из указанных работ.

В качестве рекомендаций по организации работы с ЭУМК следует указать на первоочередную необходимость усвоения слушателями основных теоретических понятий, определений, методологических и концептуальных основ дисциплины, что, в свою очередь, способствует формированию их познавательного интереса и потребности в приобретении знаний по учебной дисциплине «Основы педагогических измерений». Дальнейшее ознакомление слушателей с современными программными средствами и отработка практических навыков по работе с ними будут способствовать эффективному использованию полученных знаний в практике педагогической деятельности специалиста, а также внедрению результатов педагогических исследований в образовательный процесс.

Основные структурные компоненты ЭУМК прошли успешную апробацию в рамках преподавания учебной дисциплины «Основы педагогических измерений». Представленное их содержание структурных компонентов ЭУМК является необходимым и достаточным для организации эффективного изучения дисциплины на теоретическом и практическом уровнях при осуществлении переподготовки по специальности 1-08 01 71 «Педагогическая деятельность специалистов» с присвоением квалификации «Преподаватель в соответствии с квалификацией по основному образованию».

1. ТЕОРЕТИЧЕСКИЙ РАЗДЕЛ

1.1 Сущность педагогических измерений. Основные понятия математической статистики в педагогике

- *История становления педагогических измерений*
- *Современные направления в изучении и оценке педагогических явлений*
- *Предмет и задачи математической статистики в педагогике*
- *Зависимые и независимые переменные*
- *Уровень статистической значимости (p-уровень). Статистические гипотезы*
- *Классификация задач и методы их решения*
- *Методы и технологии педагогических измерений*

- *История становления педагогических измерений*

Исторически сложилось так, что теории педагогических измерений предшествовала разработка вопросов теории оценивания (Smallwood, M.L.) Оценивание можно определить, как экспертный процесс выставления каких-либо баллов по заранее предписанным правилам.

История возникновения общей теории измерений в общественных науках обстоятельно исследована Б.Райтом (Wright, B.D.). Он был уверен, что создаваемая им теория пригодна для всех общественных наук, включая педагогику. Так оно и случилось, и для того времени это был правильный ход мысли. Но рано или поздно, в науках возникают тенденции расширения и разделения, выделения собственного предмета. А потому естественным образом возникла проблема отделения педагогической теории измерений от общей теории измерений в психологии, социологии и в других общественных науках.

Теория педагогических измерений, как научная сторона процесса тестирования, стала развиваться с небольшим опозданием от теории тестов. Начало теории педагогических измерений положил Эдвард Торндайк (Thorndike E.L.). В своих трудах он затрагивал широкий круг философских, психологических и педагогических вопросов измерений, а также профессионального отбора. Новаторство Э.Л.Торндайка проявилось в том, что он ранее других понял необходимость развития не только теории обучения, но также и двух других теорий – учения (самостоятельного изучения) и теории педагогических измерений. С появлением электронного обучения значимость этих двух теорий была осознана глубже, чем это было раньше.

Со времён Э. Торндайка в педагогическом сознании западных педагогов укоренялась мысль о том, что основная сфера применения тестовых форм – это не столько педагогический контроль, сколько самоконтроль в процессе собственного учения.

Исторически сложилось так, что первыми, больше ста лет назад, появились труды по психологической теории измерений. В 1936 году

теоретики этого направления создали журнал высокого научного уровня, под названием «Psychometrika». Сейчас он издаётся и распространяется по всему миру. Затем стало развиваться общее направление «Психологическое и педагогическое измерение» (<http://www.sagepublications.com>). Язык и методы этого научного направления считались, видимо, по инерции, общими, равно применимыми в психологии и педагогике.

Позже последовало издание журнала по педагогическим измерениям (Journal of Educational Measurement. Published by The National Council on Measurement in Education (NCME)), где были сделаны попытки выделить собственный педагогический предмет исследования, и создать собственные педагогические методы. Успеха, однако, эти попытки не имели из-за отсутствия общей методологии педагогических измерений.

Три теории измерений. Разработка педагогической теории измерений проходила на основе имеющихся теорий. Сейчас наиболее известны три теории измерения.

1. Статистическая или, иначе, классическая теория тестов (classical test theory, СТТ). Так она называлась в начале XX-го века. Сейчас слова «классическая теория» – можно рассматривать как метафору. В подтверждение такого мнения можно привести два аргумента. Во-первых, эта теория имеет точное, а не нынешнее метафорическое название. На английском языке она называется true-score model of measurement. Поэтому её лучше называть не теорией, а моделью измерений. Во-вторых, эта модель – перенос идей теории физических измерений на измерения в психологии и педагогике самого начала XX века. Отсюда упрощённая, мало реалистичная для наших дней, аксиоматика. Её точнее называть моделью истинного и ошибочного компонентов измерения. Разумеется, эта модель уже обросла тысячами исследований подлинно теоретического характера. Поэтому название «теория» реально существует, оно приемлемо, за исключением случаев сравнения этой модели с другими моделями педагогического измерения.

2. Вторую теорию можно очень условно перевести на русский язык как математическая теория измерений (Item Response Theory, IRT) (Аванесов В.С.).

3. Третья теория – больше чем теория. Это одновременно ещё методология и метрическая система. Потому она получила название метрической теоретической системы Г.Раша (Rasch Measurement, RM) (Аванесов В.С.). Когда эту модель измерения называют теорией, то и другие модели можно называть теориями, как это случилось, например, с IRT. Хотя другие, позже появившиеся модели измерения так и остались моделями. Да и более точное название IRT – это Latent Trait Model of Measurement. Если за этими двумя названиями моделей измерения сохранять статус теорий, то отмеченное лексическое несоответствие лучше преодолевать созданием теории педагогических измерений, что и было сделано в статье журнала «Педагогические Измерения (ПИ)» №1, 2004 г.

Все перечисленные – относительно общие теории, применимые не только в педагогике, но и в психологии, социологии, в медицинских и иных

исследованиях. Все теории неоднократно рассматривались в журнале ПИ. Но общие теории не давали, и не могли дать, адекватные ответы на специфические педагогические проблемы. Нужна была специфическая педагогическая теория. Неоднократно приходилось читать о делении теорий измерения на современные и не современные. С подобным подходом трудно согласиться, хотя бы потому, что устаревших (ненужных) теорий в педагогических измерениях нет. Каждая теория и сопутствующие ей методы дают свою, уникальную информацию о качестве заданий. Например, в RM используется проверка заданий на соответствие требованиям модели Раша. Но получаемая при этом статистика оценки адекватности задания (Fit statistics) заметно уступает, по точности, значениям коэффициентов корреляции, используемых в статистической (классической) модели измерения для обоснования пригодности каждого задания при включении его в тест. Не случайно во всех профессионально сделанных вычислительных пакетах значения коэффициентов корреляции приводятся. Сторонники RM иногда не принимают во внимание эту статистику, что надо признать, как недостаток.

Поскольку перечисленные теории одинаково применимы к психологическим, социологическим, медицинским и другим измерениям, их можно отнести к общим междисциплинарным метрическим теориям. Следовательно, педагогическую теорию измерений можно тогда отнести к специальной педагогической теории.

- ***Современные направления в изучении и оценке педагогических явлений***

В педагогической теории измерений выделены три предмета исследования: педагогические понятия теории, принципы отбора педагогического (предметного) содержания заданий и логические принципы формулирования педагогических тестовых заданий.

Существенные признаки педагогических измерений – это педагогически обоснованный отбор содержания теста, точный подбор формы заданий для того или иного содержания, система основных педагогических понятий теории (Аванесов В.С.). Удивительно, что измерения педагогических латентных свойств, в том числе относящихся к педагогической науке, до сего дня некоторые зарубежные авторы считают задачами психометрики, с чем согласиться невозможно.

В работах последних лет был сформулирован понятийный аппарат педагогической теории, даны определения основным понятиям. В качестве таковых выделены: задание в тестовой форме, тестовые задания, педагогический тест, испытуемые, тестовый процесс и содержание теста. Предметом специального анализа стали также два других важных понятия педагогической теории измерений. Первое понятие – форма тестовых заданий – это способ связи, упорядочения элементов задания. Тысячи примеров заданий четырех основных форм, по многим учебным дисциплинам можно найти в книгах автора (Аванесов В.С.).

Показано, что логическое преимущество задания в тестовой форме, по сравнению с привычными вопросами экзаменов и госэкзамена, заключается в возможности превращения заданий в тестовой форме, после ответа испытуемых, в форму высказываний, истинных или ложных. Что создаёт надёжную основу для безошибочного дифференцирования испытуемых по уровню их подготовленности. Задание в тестовой форме выгодно отличается от обычных вопросов свойствами краткости, лучшей понимаемости смысла заданий, быстроты ответа учащихся, технологичности и эффективности.

В современном тестировании часто применяется дифференцирование испытуемых, по их ответам на каждое задание, не только на два, но и на большее число уровней. Это означает, что оценки даются за ответы правильные, неправильные, а также за частично правильные, с разными уровнями градации правильности. Такое оценивание ответов испытуемых называется по-английски Partial Credit Model. Оно имеет преимущества перед традиционным дихотомическим оцениванием посредством единицы и нуля. Главное преимущество – повышение надёжности тестовых результатов.

В качестве главной проблемы теории педагогических измерений названа разработка показателей качества и эффективности образовательного процесса.

Предмет педагогических измерений – уровень подготовленности испытуемых, который представляет собой единство знаний, умений, навыков, представлений. К понятию «подготовленность» можно отнести и часто используемое сейчас понятие «компетентность», представляющую собой прагматическую, или утилитарную, проекцию общего содержания образования.

При педагогических измерениях одно и то же свойство личности проявляется у разных испытуемых в разной степени. Обычно большему проявлению интересующего свойства ставится в соответствие и большее число. Реже поступают наоборот. Это бывает в случаях измерения общественно неодобряемых психологических свойств личности.

Кроме упомянутых выше результатов, в рамках развития педагогической теории измерений исследованы два традиционных критерия качества – надёжности и валидности тестовых результатов, принятых в статистической теории измерений, и два относительно малоизвестных критерия – объективности и эффективности (Аванесов В.С) тестовых результатов, используемых преимущественно в RM и в адаптивном тестировании.

Сформулированы также принципы отбора содержания тестовых заданий, концепция пяти этапов тестового педагогического процесса, дана обновлённая классификация видов проверяемых знаний школьников и студентов. Самое большое внимание в рамках разработки педагогической теории измерений было уделено вопросам формы и принципам композиции тестовых заданий (Аванесов В.С). Три главных свойства заданий подлежат обязательному шкалированию и доскональному теоретическому и эмпирическому педагогическому исследованию – это уровень понимаемости

смысла содержания заданий, мера их трудности и дифференцирующей способности.

Цели и задачи педагогических измерений. Сейчас цели тестирования и педагогических измерений лучше разделять.

Цель тестирования – получение тестовых результатов уровня подготовленности испытуемых, цель же педагогических измерений, в основном – получение шкалированных значений уровня подготовленности испытуемых и уровня трудности заданий.

Цели педагогических измерений можно разделить на метрические, педагогические, психологические и социально-политические.

Пример метрической цели – построить шкалу уровня подготовленности испытуемых или шкалу уровня трудности заданий. В рамках метрической цели нередко используется сравнительно узкое, фокусированное определение цели педагогического измерения – как позиционирование испытуемого на прямой линии, представляющей латентную переменную величину (Wright B.D., Stone M.H.).

Для позиционирования испытуемых на шкале их надо тестировать посредством заданий, определяющих переменную величину, а затем посредством шкалирования убедиться в том, что полученные тестовые результаты помогают определить место испытуемого на латентной переменной величине.

Пример педагогической цели – провести классификацию испытуемых по уровню их подготовленности, для комплектования, например, уровневых классов учащихся школы.

Психологи используют тестовые результаты в целях выявления психологических затруднений в учебном процессе.

И, наконец, пример социально-политической цели – применение тестов для обеспечения равного доступа учащихся к качественному образованию.

Основную задачу педагогических измерений Ю.Н. Каргин видит в поиске оптимальных оценок показателей подготовленности испытуемых и трудности заданий по наблюдаемым результатам тестирования.

Прогресс, достигнутый в 80-е гг. XX в. в методологии педагогических измерений, и, завершившийся созданием теории IRT (Современная теория тестирования (англ. Item Response Theory) – набор методов, позволяющий оценить вероятность правильного ответа испытуемых на задания различной трудности. Она используется для того чтобы избавиться от плохих (неинформативных) вопросов в опросе и предоставить вопросы с адекватными баллами на основе их сложности. На русский язык название Item Response Theory переводится различным образом. Ю.Нейман и В.Хлебников предлагают называть ее «Теория моделирования и параметризации педагогических тестов» (ТМППТ). В.Аванесов – «Математико-статистическая теория оценки латентных параметров заданий теста и уровня подготовленности испытуемых», привел к новым возможностям шкалирования результатов тестирования. В отличие от классической теории

тестов, в рамках которой можно повысить сопоставимость сырых баллов путем их преобразования без изменения природы порядковой шкалы, IRT позволяет перевести наблюдаемые баллы обучаемых из порядковой в количественную шкалу интервального характера. Такое преобразование осуществляется путем введения специальной единицы измерения – логита – и получения оценок подготовленности студентов, не зависящих от трудности заданий теста.

Таким образом, для перехода на количественный уровень измерений необходимо использовать стандартизованные тесты с выбором ответов, автоматизированные формы проверки и обработки данных, а разработку тестов и шкалирование результатов их выполнения вести только на основе теории IRT. Несмотря на снижение объективности, возникающее за счет привлечения экспертов в тех случаях, когда в стандартизованном тесте есть задания со свободно конструируемым ответом, шкалу тестовых баллов, построенную на основе теории IRT по таким заданиям, также относят к количественным измерениям, что не вполне корректно. Для качественных измерений разрабатывают портфолио и анкеты, проводят интервью, собеседования, устные опросы и экзамены в традиционной форме. Иногда при аккредитации учебных заведений привлекается анализ документов, отражающих ход и результаты учебного процесса.

В целом можно отметить, что количественный и качественный уровни измерений имеют существенные отличия по проявлению совокупности свойств оцениваемых характеристик объектов, способам измерения, формам представления результатов, группам допустимых преобразований и операций с различными математическими и статистическими величинами, характеризующими измеряемые признаки. Количественные оценки обычно представляют в виде чисел на шкале, чтобы дифференцировать различные величины переменной. По результатам качественного измерения для оценочных характеристик выбирают дескрипторы в виде слов или символов, хотя нередко в образовании роль символов для удобства размещения на шкале отводят условно выбираемым числам.

В связи с введением компетентного подхода роль качественных оценок, несомненно, будет расти, поскольку оценивание компетенции может проводиться на качественном уровне измерений. На деле приоритет качественного уровня вовсе не означает отказа от тестов при оценивании профессиональных компетенций выпускников вузов. В рамках компетентного подхода количественные оценки должны дополняться многоаспектными качественными оценками и получать развернутую качественную интерпретацию по отношению к задачам профессиональной деятельности выпускников. Для повышения объективности оценок при построении качественных шкал необходимо использовать оценочные средства, разработка которых осуществлялась с опорой на теорию педагогических измерений. Применение таких измерителей должно обеспечивать высокую надежность и валидность оценок.

- ***Предмет и задачи математической статистики в педагогике***

Непрерывное и быстрое расширение областей исследования, в которых удается эффективно использовать математические и статистические методы, составляет одну из характерных черт развития современной науки. Раздвигая традиционные рамки «точных наук», этот процесс вовлекает сегодня в свою сферу биологию и социологию, языкознание и психологию, юриспруденцию, педагогику и историю. Применение методов математической статистики в самых разных областях знания открывает более широкие возможности для глубокого проникновения в сущность и закономерности изучаемых явлений, более точного предсказания их развития в различных условиях, а значит и более эффективного управления ими, практического их использования.

Понятие математической статистики.

Математическая статистика – раздел математики, посвященный математическим методам систематизации, обработки и использования статистических данных для научных и практических выводов. Определение, сформулированное видными отечественными математиками А.Н. Колмогоровым и Ю.В. Прохоровым.

Статистические данные – это данные, полученные в результате обследования большого числа объектов или явлений.

Статистические методы — методы анализа статистических данных. Математическая статистика не раскрывает сущности исследуемых явлений. Она может фиксировать статистически достоверные различия между двумя исследуемыми явлениями, но не может объяснить причины этих различий.

Математическая статистика подразделяется на две основные области: описательную и аналитическую статистику. ***Описательная статистика*** раскрывает методы описания статистических данных, представляет их в форме таблиц, графиков и распределений. ***Аналитическая статистика*** (теория статистических выводов) ориентирована на обработку данных, полученных в ходе эксперимента, с целью формулировки выводов, имеющих прикладное значение для самых различных областей человеческой деятельности.

Математическая статистика исходит из предположения, что наблюдаемая изменчивость окружающего мира имеет два источника:

- действие известных причин и факторов. Они порождают изменчивость, закономерно объяснимую;

- действие случайных причин и факторов. Большинство природных и общественных явлений обнаруживают изменчивость, которая не может быть целиком объяснена закономерными причинами. В таком случае прибегают к ***концепции случайной изменчивости***. Выражение «случайный» в данном контексте означает «подчиняющийся законам теории вероятности».

Проверка психолого-педагогических гипотез и моделей является тоже случайным событием, так как результаты педагогического исследования определяются очень большим количеством заранее непредсказуемых факторов. Определенные закономерности можно выявить только в случае

массовых наблюдений вследствие закона больших чисел. *Закон больших чисел* – это объективный математический закон, согласно которому совместное действие большого числа случайных факторов приводит к результату, почти не зависящему от случая.

Отсюда совершенно очевидным является факт, что педагогические и психологические измерения однозначно связаны со статистическими измерениями, которые, в свою очередь, и являются *предметом математической статистики*.

Статистический подход – это выявление закономерной изменчивости на фоне случайных факторов и причин. Методы математической статистики позволяют оценить параметры имеющихся закономерностей, проверить те или иные гипотезы об этих закономерностях.

Аппарат математической статистики является замечательным по мощности и гибкости инструментом для отсеивания закономерностей от случайностей. Педагогу-исследователю обязательно необходимо накапливать информацию об окружающем мире, пытаясь выделить закономерности из случайностей.

Математическая статистика возникла (XVII в.) и развивалась параллельно с теорией вероятностей. Дальнейшее развитие математической статистики (вторая половина XIX— начало XX в.) обязано, в первую очередь, П. Л. Чебышеву, А. А. Маркову, А. М. Ляпунову, а также К. Гауссу, А. Кетле, Ф. Гальтону, К. Пирсону и др.

В XX в. наиболее существенный вклад в математическую статистику был сделан советскими математиками (В. И. Романовский, Е. Е. Слуцкий, А. Н. Колмогоров, Н. В. Смирнов), а также английскими (Стьюдент, Р. Фишер, Э. Пирсон) и американскими (Ю. Нейман, А. Вальд) учеными.

Задачи математической статистики:

1. Определить способы сбора и группировки статистических данных, сведений, полученных в результате наблюдений или экспериментов.
2. Разработать методы анализа статистических данных в зависимости от цели исследования для получения научных и практических выводов.

Современная математическая статистика разрабатывает способы определения числа необходимых испытаний до начала исследования (планирование эксперимента), в ходе исследования (последовательный анализ) и решает задачи анализа и представления результатов исследования.

Методы математической статистики имеет смысл применять, если необходимо обработать и представить результаты:

1. серии случайных экспериментов, свойства которых частично или полностью неизвестны;
2. эксперимента, проводимого в одних и тех же условиях некоторое число раз. Примером такой серии экспериментов может служить социологический опрос, набор экономических показателей, психолого-педагогический эксперимент.

Математическая статистика является разделом математики, но ее методы нашли широкое применение в различных областях науки. Многие статистические процедуры и методы достаточно понятны, просты и легко применимы на практике.

Степень математизации научных дисциплин служит объективной характеристикой глубины знаний об изучаемом предмете. Так, многие явления физики, химии, техники описываются математическими методами достаточно полно. В результате эти науки достигли высокой степени теоретических обобщений. В педагогике и психологии статистические и математические методы играют подчиненную роль из-за сложности объектов, процессов и явлений, варибельности их характеристик, наличия индивидуальных особенностей испытуемых.

Более основательное знакомство педагогов-исследователей с современными статистическими методами поможет им повысить качество обработки, анализа и интерпретации данных педагогического или психологического эксперимента при проведении научно-исследовательской работы.

Математическая статистика в руках педагогов, психологов, экономистов и социологов должна быть не только мощным инструментом, позволяющим успешно обрабатывать, анализировать и представлять полученные экспериментальные данные, но и средством объективного познания окружающей действительности.

- ***Зависимые и независимые переменные***

Переменная (английский термин *variable*) – это то, что можно измерять, контролировать или, другими словами, то, что можно изменять в исследованиях. Переменные отличаются многими аспектами, особенно ролью, которую они играют в исследованиях (например, шкалой измерения, количеством изучаемых переменных, статусными характеристиками и др.).

Конечная цель всякого исследования или научного анализа состоит в нахождении связей (зависимостей) между переменными. Философия науки учит, что не существует иного способа представления знания, кроме как в терминах зависимостей между количествами или качествами, выраженными какими-либо переменными. Таким образом, развитие науки всегда заключается в нахождении новых связей между переменными. Назначение статистики состоит в том, чтобы помочь объективно оценить зависимости между переменными.

В изучении корреляций (зависимостей, связей, взаимосвязей и т. д.) исследователь не влияет (или, по крайней мере, пытается не влиять) на переменные, а только измеряет их и пытается найти зависимости между некоторыми измеренными переменными. В экспериментальных исследованиях, напротив, варьируются некоторые переменные и измеряются воздействия этих изменений на другие переменные. Например, исследователь

может искусственно увеличивать показатели одной переменной, а затем на определенном уровне измерить показатели другой.

Анализ данных в научном исследовании также заключается в вычислении «корреляций» (зависимостей) между переменными, а именно, между теми, на которые воздействуют, и теми, на которые влияет это воздействие. Тем не менее, эмпирические данные потенциально содержат в себе более качественную информацию. Только экспериментально можно убедительно доказать причинную связь между переменными. Например, если обнаружено, что всякий раз, когда изменяется переменная А, изменяется и переменная В, то можно сделать вывод – переменная А оказывает влияние на переменную В, т. е. между ними имеется причинная зависимость. Результаты корреляционного исследования могут быть проинтерпретированы в каузальных (причинных) терминах на основе некоторой теории, но сами по себе не могут отчетливо доказать причинность.

Переменные дифференцируются на независимые и зависимые. **Независимыми называются переменные**, которые варьируются исследователем. **Зависимые** – это переменные, которые измеряются или регистрируются. На первый взгляд может показаться, что проведение такой дифференциации создает путаницу в терминологии, поскольку можно говорить, что «все переменные зависят от чего-нибудь». Однако четкое проведение такой дифференциации вызвано необходимостью.

Термины зависимая и независимая переменная применяются в основном в экспериментальном исследовании, где экспериментатор манипулирует некоторыми переменными, и в этом смысле они «независимы» от реакций, свойств, характеристик и т. д., присущих объектам исследования. Зависимость проявляется в ответной реакции исследуемого объекта на посланное на него воздействие. Отчасти в противоречии с данным разграничением понятий находится использование их в исследованиях, где не варьируются независимые переменные, а объекты только приписываются к «экспериментальным группам», основываясь на некоторых их априорных свойствах. Например, если в эксперименте юноши сравниваются с девушками относительно такой характеристики как профессиональная направленность, то пол можно назвать независимой переменной, а профессиональную направленность – зависимой. Это экспериментальное сравнение, в свою очередь, может быть описано как поиск связи между переменными: пол и профессиональная направленность.

Переменные и измерительные шкалы. Изучаемые переменные различаются также тем, «насколько хорошо» они могут быть измерены или, другими словами, как много измеряемой информации обеспечивает шкала их измерений. Очевидно, в каждом измерении присутствует некоторая ошибка, определяющая границы «количества информации», которое можно получить в данном измерении. Другим фактором, определяющим количество информации, содержащейся в переменной, является тип шкалы, в которой проведено измерение. Различают следующие типы шкал: номинальная,

порядковая (ординальная), интервальная, относительная (шкала отношения). Соответственно, имеем и четыре типа переменных: номинальные переменные, порядковые переменные, интервальные переменные и относительные переменные.

Наиболее общая классификация, предложенная С. Стивенсом, включает четыре уровня измерений и фиксирует присущие им свойства. Согласно такой классификации, различают шкалы качественные – шкала наименований или классификаций (номинальная шкала), порядковая шкала, и количественные – интервальная шкала и шкала отношений.

Качественные шкалы иногда называют неметрическими (концептуальными), количественные получили название метрических или материальных. Для каждого уровня измерений существуют группы допустимых преобразований и операций с различными математическими и статистическими величинами, характеризующими измеряемые признаки.

На качественном уровне отнесение эмпирических объектов измерения к различным классам проводится по признаку эквивалентности (шкала наименований) или по признаку упорядочения внутри эквивалентных объектов одного класса (порядковая шкала). Для отнесения или упорядочения применяются экспертные методы, при которых оценки на шкале считаются достоверными, если они признаны большинством экспертов.

В этой особенности построения качественных шкал в образовании есть свои плюсы и минусы. Положительным можно считать относительную легкость построения таких шкал, поскольку не привлекаются специальные оценочные средства, а участвуют только эксперты. Негативные характеристики качественных шкал – ограниченная сфера применения и низкая точность измерения. Числа или символы, приписываемые объектам путем экспертного оценивания, субъективны и носят исключительно условный характер. Поэтому эти числа нельзя суммировать или проводить с ними другие математические операции.

Количественные шкалы не предусматривают привлечения экспертов, поэтому представленные в них оценки измеряемых характеристик объектов, обладают более высокой объективностью и поддаются определенным математическим операциям.

К простейшему типу качественных шкал относится шкала наименований (другое название – номинальная шкала), используемая для представления результатов классификации эмпирических объектов измерения, свойства которых проявляются только на уровне оценивания их эквивалентности. При построении шкалы наименований привлекаются эксперты, которые приписывают качественным свойствам объектов некоторые числа или другие символы.

Номинальные шкалы достаточно широко применяются в образовании в тех случаях, когда педагогическое измерение связано с объединением учащихся в группы по какому-либо признаку без установления порядка следования групп. Примером номинальной шкалы могут служить результаты

зачетной сессии, когда все студенты делятся на две группы: получившие и не получившие зачет. Другой пример номинальной шкалы получается при дихотомическом оценивании результатов по отдельным заданиям теста. Если тестируемый студент за правильное выполнение задания теста получает 1, а за неправильное выполнение или пропуск задания теста – 0, то результаты тестирования будут представлены в номинальной шкале.

Так как числа в шкале наименований не несут в себе никаких количественных признаков, а характеризуют только отношение эквивалентности, то их применяют для определения вероятности или частоты появления в ряду наблюдений данного объекта. Поэтому для обработки количественных данных следует использовать не сами числа, а удельный вес количества объектов данного класса.

Если отнесение объектов к классам основано не только на отношении эквивалентности, но и учитывает возрастание или убывание степени проявления измеряемого свойства, то получаемая шкала носит название шкалы порядка. В порядковой шкале производится ранжирование объектов или классов объектов, связанных соотношением больше-меньше.

Порядковые шкалы используются в образовании в тех случаях, когда педагогический контроль осуществляется традиционными способами без применения теории измерений и тестов. Однако результаты тестирования также приводят к порядковой шкале, если их обработка проводится без алгоритмов теории ИРТ. Классический пример порядковой шкалы – привычная четырехбальная шкала, которую иногда неоправданно называют пятибалльной. Каждой группе студентов, проявляющей согласно мнению преподавателя сходные знания, присваивается одинаковый (один из четырех) номер места от двух до пяти.

Если ранжируются не отдельные объекты, а целые классы объектов, то внутри каждой группы порядок не устанавливается. Поэтому часто говорят о том, что порядковая шкала обладает слабым дифференцирующим эффектом, особенно в тех случаях, когда оценивается подготовленность довольно большого числа студентов. Каждой группе объектов присваивается определенный порядковый номер, который позволяет отличить ее представителей от представителей другого класса. Например, по оценкам в порядковой шкале можно ранжировать студентов от слабых к сильным или наоборот, но сделать вывод о том, на сколько один подготовлен лучше другого, нельзя в силу отсутствия единицы измерения.

При измерении признака в порядковой шкале возможны только монотонные преобразования, допускающие умножение на постоянный множитель, возведение в степень и извлечение корня, а также некоторые статистические операции. В частности, в порядковой шкале:

- в качестве средней оценки используют медиану;
- в качестве меры рассеяния применяют квантили;
- в качестве меры связи двух признаков используют ранговый коэффициент корреляции.

Интервальная шкала – количественная. Она используется для упорядочения объектов, свойства которых удовлетворяют отношениям эквивалентности, порядка и аддитивности. В ней определено расстояние между объектами и предусмотрена общая для всех объектов единица измерения, а началом отсчета является условно выбранная нулевая точка. Благодаря существованию единицы измерения, в интервальной шкале возможны все арифметические действия над числами, кроме операции деления, в силу отсутствия абсолютного нуля. В ней также допустимы почти все статистические операции, кроме тех, которые предполагают знание «истинно» нулевой точки.

В интервальной шкале достигается корректная сравнимость результатов педагогических измерений. Хотя она лишена недостатков номинальной и порядковой шкал, в ней остается неизвестной естественная точка отсчета – абсолютный нуль. Поэтому при интерпретации результатов тестирования с помощью интервальной шкалы нуль верных ответов не означает полного отсутствия подготовленности обучаемых.

Для построения интервальной шкалы в педагогических измерениях необходимо обращение к современной теории тестов ИРТ. Ее возможности, связанные с инвариантностью оценок параметров подготовленности обучаемых, и трудности заданий теста в сочетании со специальными преобразованиями эмпирических результатов тестирования позволяют ввести единицу измерения и перевести баллы испытуемых из порядковой шкалы в интервальную.

Шкала отношений отображает свойства объектов, удовлетворяющие отношениям эквивалентности, порядка, аддитивности и пропорциональности. По сравнению со всеми ранее рассмотренными шкалами последняя (шкала отношений) обеспечивает самый высокий уровень измерений. В ней можно выполнять все арифметические и статистические операции, в том числе интерпретировать отношение чисел, приписываемых различным объектам. Использование шкалы отношений в образовании невозможно в силу отсутствия абсолютного нуля, поэтому на сегодняшний день ее удалось реализовать только в рамках физических измерений. Общая типология уровней измерения приведена на рисунке 1.1.1.



Рисунок 1.1.1 - Типология уровней измерения и соответствующих шкал

Помимо четырех уровней измерений, принятых согласно классификации Стивенса, в отечественной литературе по проблемам измерений иногда выделяют пятый уровень и вводят так называемые абсолютные шкалы. Под ними понимают шкалы отношений, обладающие не только естественно выбранным абсолютным нулем, но и естественно определенными единицами измерения, не зависящими от условных соглашений. Такие шкалы не встречаются в образовании, они характерны для оценивания относительных величин в физических измерениях.

Поскольку количественные шкалы предпочтительнее качественных в тех случаях, когда результаты педагогического измерения используются для принятия административно-управленческих решений, то возникает вопрос о необходимых условиях для получения данных на количественном уровне измерений. Такие условия можно разбить на две части, одна из которых относится к самому измерителю, а другая – к методам обработки данных при их шкалировании.

Не нуждается в доказательстве тот факт, что традиционные средства и методы контроля, включающие экзамены, контрольные работы, устные опросы и другое, позволяют получать баллы испытуемых только на качественном уровне в номинальной либо порядковой шкале. Точка зрения на возможности тестов для построения количественных шкал менялась по мере развития теории педагогических измерений и совершенствования технологии разработки тестов. Вначале, когда тесты только появились в образовании, педагогам казалось, что (не понятно, кто они: педагоги или тесты, нельзя удалять) они обеспечивают оценки в количественной шкале. Такое впечатление создавалось в силу того, что при тестировании получались баллы учащихся, подсчитанные на основе количества правильно выполненных заданий теста. Впоследствии возникли серьезные аргументы, подвергающие сомнению количественный характер результатов выполнения тестов.

Действительно, для тестовых баллов нельзя ввести единицу измерения в связи с зависимостью баллов от различной трудности заданий теста. Исключением мог бы стать тот случай, когда в тесте используются задания одинаковой трудности, но это нереально, поскольку не согласуется ни с основными положениями теории педагогических измерений, ни с потребностями учебного процесса. В любой оценке трудности задания всегда присутствуют ошибочные компоненты, поэтому с точки зрения теории измерений невозможно выбрать одинаково трудные задания в тесте. Да они и не нужны, поскольку все студенты отличаются по своей подготовленности, и для оценивания их учебных достижений необходимы различные по трудности задания, адекватные возможностям обучаемых. К тому же специфика содержания отдельных элементов учебного материала никогда не позволит сделать одинаковые по трудности контролирующие задания в любой форме, в том числе и в форме теста.

Таким образом, по мере развития теории и практики тестирования пришло осознание того, что результаты тестирования обучаемых позволяют

установить отношение равенства или отношение «больше-меньше», а разность наблюдаемых (сырых) тестовых баллов при интерпретации отличий в подготовленности учащихся не имеет никакого смысла.

- **Уровень статистической значимости (p-уровень). Статистические гипотезы**

Понятие статистической значимости результата (p-уровень). Статистическая значимость результата (p-уровень) представляет собой оцененную меру уверенности в его «истинности» (в смысле «репрезентативности выборки»). Выражаясь более технически, p-уровень (термин был впервые использован в работе Brownlee, 1960) – это показатель, находящийся в убывающей зависимости от надежности результата. Более высокий p-уровень соответствует более низкому уровню доверия к найденной в выборке зависимости между переменными.

Статистическая значимость представляет собой вероятность ошибки, связанной с распространением наблюдаемого результата на всю совокупность. Например, p-уровень, равный 0,05 (т. е. 1/20), показывает, что имеется 5-процентная вероятность того, что найденная в выборке связь между переменными является лишь случайной особенностью данной выборки. Иными словами, если данная зависимость в совокупности отсутствует при многократном проведении подобных экспериментов, то примерно в одном из двадцати повторений эксперимента можно ожидать такой же или более сильной зависимости между переменными.

Обычно во многих областях науки результат $p = 0,05$ является приемлемой границей статистической значимости. Однако этот уровень все еще включает довольно большую вероятность ошибки (5%), и полученные результаты могут приниматься как достоверные, т.е. будет иметь место тенденция к статистической значимости. Результаты на уровне $p \leq 0,01$ рассматриваются как статистически значимые, а результаты с уровнем $p \leq 0,005$ или $p \leq 0,001$ – как высоко значимые. Следует понимать, что данная классификация уровней значимости достаточно произвольна и является всего лишь неформальным соглашением, принятым на основе практического опыта в той или иной области исследования.

Статистическая и педагогическая гипотезы. Статистическая проверка научной гипотезы. Доказательство статистической достоверности экспериментального влияния существенно отличается от доказательства в математике и формальной логике, где выводы носят более универсальный характер: статистические доказательства не являются столь строгими и окончательными – в них всегда допускается риск ошибиться в выводах и потому статистическими методами не доказываются окончательно правомерность того или иного вывода, а показывается мера правдоподобности принятия той или иной гипотезы.

Педагогическая гипотеза (научное предположение о преимуществе того или иного метода и т. п.) в процессе статистического анализа переводится на

язык статистической науки и заново формулируется, по меньшей мере, в виде двух статистических гипотез. Первая (основная) называется нулевой гипотезой (H_0), в которой исследователь говорит о своей исходной позиции. Он (априори) как бы декларирует, что новый (предполагаемый им, его коллегами или оппонентами) метод не обладает какими-либо преимуществами, и потому с самого начала исследователь психологически готов занять честную научную позицию: различия между новым и старым методами объявляются равными нулю. В другой, альтернативной гипотезе (H_1) делается предположение о преимуществе нового метода. Иногда выдвигается несколько альтернативных гипотез с соответствующими обозначениями.

Например, гипотеза о преимуществе старого метода (H_2). Альтернативные гипотезы принимаются тогда и только тогда, когда опровергается нулевая гипотеза. Это бывает в случаях, когда различия, скажем, в средних арифметических экспериментальной и контрольной групп настолько значимы (статистически достоверны), что риск ошибки отвергнуть нулевую гипотезу и принять альтернативную не превышает одного из трех принятых уровней значимости статистического вывода:

- первый уровень – $p \leq 0,05$ (5%), где допускается риск ошибки в выводе в пяти случаях из ста теоретически возможных таких же экспериментов при строго случайном отборе испытуемых для каждого эксперимента;

- второй уровень – $p \leq 0,01$ (1%), т. е. соответственно допускается риск ошибиться только в одном случае из ста (при тех же требованиях);

- третий уровень – $p \leq 0,001$ (0,1%), т. е. допускается риск ошибиться только в одном случае из тысячи. Данный уровень значимости предъявляет очень высокие требования к обоснованию достоверности результатов эксперимента.

- ***Классификация задач и методы их решения***

Задачи педагогического исследования. Множество задач педагогического исследования предполагает те или иные сопоставления. Например, мы сопоставляем группы испытуемых по какому-либо признаку, чтобы выявить различия между ними по этому признаку. Мы сопоставляем то, что было «до» с тем, что стало «после» наших экспериментальных или любых иных воздействий, чтобы определить эффективность этих воздействий. Мы сопоставляем эмпирическое распределение значений признака с каким-либо теоретическим законом распределения или два эмпирических распределения между собой, с тем, чтобы доказать неслучайность выбора альтернатив или различия в форме распределений. Мы также можем сопоставлять два признака, измеренные на одной и той же выборке испытуемых, для того, чтобы установить степень согласованности их изменений, их сопряженность, корреляцию между ними. Наконец, мы можем сопоставлять индивидуальные значения, полученные при разных комбинациях каких-либо существенных

условий, с тем чтобы выявить характер взаимодействия этих условий в их влиянии на индивидуальные значения признака.

Методы решения задач педагогического исследования. В таблице 1.1.1 приведена краткая классификация задач и методов:

Таблица 1.1.1 - Классификация задач и методы их решения

Задачи	Условия	Методы
1. Анализ различий	а) 2 выборки испытуемых	Q - критерий Розенбаума; U - критерий Манна-Уитни; φ^* - критерий (угловое преобразование Фишера)
	б) 3 и более выборок испытуемых	S- критерий тенденций Джонкира H - критерий Крускала-Уоллиса.
2. Оценка сдвига значений исследуемого признака	а) 2 замера на одной и той же выборке испытуемых	T - критерий Вилкоксона; G - критерий знаков; φ^* - критерий (угловое преобразование Фишера).
	б) 3 и более замеров на одной и той же выборке испытуемых	χ^2 - критерий Фридмана; L- критерий тенденций Пейджа.
3. Выявление различий в распределении признака	а) при сопоставлении эмпирического признака распределения с теоретическим	χ^2 - критерий Пирсона; λ - критерий Колмогорова-Смирнова; m - биномиальный критерий.
	б) при сопоставлении двух эмпирических распределений	χ^2 - критерий Пирсона; λ - критерий Колмогорова-Смирнова; φ^* - критерий (угловое преобразование Фишера).
4. Выявление степени согласованности изменений	а) двух признаков	r_s - коэффициент ранговой корреляции Спирмена.
	б) двух иерархий или профилей	r_s - коэффициент ранговой корреляции Спирмена.
5. Анализ изменений признака под влиянием контролируемых условий	а) под влиянием одного фактора	S - критерий тенденций Джонкира; L - критерий тенденций Пейджа; однофакторный дисперсионный анализ Фишера.
	б) под влиянием двух факторов одновременно	Двухфакторный дисперсионный анализ Фишера.
6. Обобщение эмпирических данных (редукция данных)	Большое количество измерений, большая выборка (не менее 100 испытуемых)	Факторный анализ и его разновидности
7. Построение классификаций	Большое количество измерений, большая выборка	Кластерный анализ и его разновидности

Выбор метода статистического анализа для дальнейшей обработки данных педагогического исследования осуществляется с учетом цели, задач, объема выборки, измерительных шкал.

- **Методы и технологии педагогических измерений**

Общая характеристика измерительных методов и технологий в образовании.

Метод (от гр. methodos – путь вслед за чем-нибудь) – путь достижения какой-либо цели или решения конкретной задачи.

Методы педагогических измерений – это формализованные способы сбора, изучения, обработки и интерпретации данных о явлениях образовательной деятельности.

Основные методы педагогических измерений:

- тестирование,
- анкетирование,
- накопление академических кредитов,
- другие.

Технология (от греч. techne– искусство, мастерство и logos – слово, учение) – это 1) отлаженная система (научное описание) правил, способов, приёмов выполнения действий и операций в каком-нибудь деле; 2) процесс осуществления последовательных действий и операций; 3) научная дисциплина, изучающая технологический процесс.

Методы	Тестирование	Анкетирование	Учет	Количественное оценивание
Технологии	Тестовые	Анкетные	Табельные	
			Кредитные	Рейтинговые

Рисунок 1.1.2 - Основные методы и технологии (формализованные) педагогических измерений

Тестовая технология состоит из: 1) Спецификации; 2) Руководства по использованию технологии; 3) Стимульного материала; 4) Ключа к заданиям.

Педагогическое тестирование. Основные формы тестовых заданий.

Педагогическое тестирование – это измерение (формализованное оценивание) в образовательных целях компетентности обучающихся с помощью педагогических тестов (тестовых технологий), завершающееся количественной оценкой на основе шкал и норм, обоснованных статистически.

Таблица 1.1.2 - Виды педагогического тестирования

По назначению	По способу предъявления и сбора информации
Диагностическое Контрольное Аттестационное Репетиционное Конкурсное	Бланковое Компьютерное: - линейное, - разветвленное, - адаптивное

Педагогический тест – это: 1) Система строго выверенных типовых заданий различной трудности с известными однозначными ответами для проведения методом тестирования проверочной работы обучающихся с анализом и оценкой её результатов; 2) Комплект материалов для тестирования, включающий набор таких заданий, бланки ответов, руководство по проведению тестирования, ключ к заданию.

Таблица 1.1.3 - Виды педагогических тестов

По предмету исследования	По характеру заданий
Предметных знаний Общей подготовленности	Трудности Продуктивности

Основные формы тестовых заданий и их разновидности представлены в таблице 1.1.4:

Таблица 1.1.4 - Основные формы тестовых заданий и их разновидности

Форма проверочных заданий	Разновидности
➤ Закрытая – задания, имеющие готовые ответы и предполагающие выбор респондентом одного или нескольких вариантов из предложенного выбора	➤ Задания для альтернативного выбора готового ответа ➤ Задания для множественного выбора одного правильного ответа ➤ Задания для множественного выбора нескольких правильных ответов
➤ Открытая – задания со свободными ответами без предлагаемых готовых вариантов ответа для выбора	➤ Задания для краткого свободного ответа ➤ Задания для развернутого свободного ответа
➤ Полуоткрытая – задания, нацеливающие исполнителя на распределение готовых ответов в нужном соотношении и требующие комбинаций с предлагаемым материалом для образования определенных пар или последовательных рядов	➤ Задания для подбора готовых ответов в пары ➤ Задания для подбора готовых ответов в последовательный ряд
➤ Комбинированная (смешанная)	➤ Задания, в которых совмещены особенности двух или трех названных выше форм

Существуют общие правила построения тестовых заданий:

1. В основе каждого задания должна лежать одна задача.
2. Задания должны быть реальными, соответствовать взглядам, интересам, возрастным характеристикам.
3. Задания не должны дублировать себя ни по содержанию, ни по назначению.
4. Все тестовые задания имеют образец правильного ответа.

5. Каждое тестовое задание имеет порядковый номер.
6. Формулировка задания должна быть оправданно краткой.
7. В формулировках следует избегать повторов, многозначных, диалектных слов, двусмысленности.
8. Формулировка должна полностью соотноситься с ответом.
9. Оптимальное число пунктов выбора:
 - для выбора одного ответа – 5-7 вариантов ответа;
 - для 2-5 правильных ответов – 6-10 вариантов ответа;
 - для подбора ответов в пары – число вариантов ответов по отношению к числу ожидаемых пар может быть большим в два раза.

При построении предметной части заданий используются принципы:

1. **Заменяемость элементов** (фасетность).

Пример: Буква {б,в,г,д,ж,к,л,м,н,п,р,с,т,ф,ц,ч,ш,щ} в словах {...} обозначает звук:

- ✓ Мягкий
- ✓ Твердый
- ✓ Мягкий и твердый.

2. **Импликация** – построение по схеме «Если..., то...».

Пример: «Если концентрация исходных веществ повышается, то химическое равновесие смещается

- ✓ Вправо
- ✓ Влево.

3. **Обратимость** – возможность менять местами инвариантный и вариативный компоненты.

Пример: «Операция нахождения производной функции называется _____».

И обратное утверждение — «Дифференцированием называется операция нахождения _____».

Педагогическое анкетирование – метод получения необходимой для образовательных целей эмпирической информации при помощи педагогических анкет с последующей их статистической обработкой и количественной (качественной) оценкой данных.

Анкетирование может быть:

- очным;
- заочным;
- анонимным;
- персонифицированным.

Достоинства анкетирования:

- экономичность;
- возможность охвата больших групп людей;
- применимость к разным сторонам жизни людей;
- хорошая формализуемость результатов;
- минимум влияния обследователя на опрашиваемого.

Недостатки анкетирования:

- невозможность контролировать процесс заполнения анкеты, что ведет к несамостоятельности ответов;
- снижение достоверности анкет из-за различия в понимании вопросов.

Педагогическая анкета – это: 1) лист специально разработанных вопросов утверждений и бланк ответов (нередко совмещенных в одном листе) для письменного опроса учащихся, а затем для анализа и оценки полученных данных методом анкетирования; 2) формализованная технология педагогического обследования (письменного опроса), представленная в материалах для анкетирования.

Виды педагогических анкет:

- Анкета интересов.
- Анкета ценностей.
- Анкета фактов жизни.
- Анкета индивидуальных оценок.
- Анкета межличностных отношений.
- Смешанные анкеты.

Вопросы в анкете могут быть:

- ✓ Закрытые;
- ✓ Открытые;
- ✓ Комбинированные;
- ✓ Прямые;
- ✓ Косвенные.

Структура анкеты: а) вступление; б) инструкция по заполнению анкеты; в) серия тематических взаимосвязанных вопросов и, нередко, возможные варианты ответов; г) реквизитная часть (может отсутствовать).

Объем анкеты варьируется от 2-3 до нескольких десятков и сотен вопросов. Большие анкеты нередко имеют разветвленный характер. Профессионально подготовленная анкета обычно стандартизирована. Научная и практическая значимость анкеты зависит от теоретической базы исследователя.

Другие измерительные методы в системе образования.*Учет данных посредством кредитных технологий:*

- ориентированные на накопление кредитов (США),
- ориентированные на перевод кредитов с целью обеспечения академической мобильности (страны азиатско-тихоокеанского региона),
- позволяющие накапливать и переводит кредиты, которые получены внутри или за пределами учебного заведения-провайдера (Великобритания).

Портфолио — это оформленное нужным образом собрание фотографий, проектов, заказов, выполненных специалистом, артистом, школьником. Это своеобразный аналог резюме, с наглядными образцами работ и опыта, раскрывающий творческий потенциал человека по определенному направлению.

Количественное оценивание посредством рейтинга. *Рейтинг* — метод и технологии оценки качества образования через определение и обозначение в баллах позиции субъекта образовательной деятельности среди ему подобных в соответствии с классификацией его свойств по заданной порядковой шкале.

Литература:

1. Ермолаев, О. Ю. Математическая статистика для психологов : Учебник // 2-е изд., испр. / О. Ю. Ермолаев. – Москва : Московский психолого-социальный институт : Флинта. – 336 с. – [Электронный ресурс]. – Режим доступа: <https://e-catalog.nlb.by / Author/Home?author.> – Дата доступа: 15.10.2020
2. Загвязинский, В. И. Методология и методы психолого-педагогического исследования : учеб. пособие для студ. высш. пед. учеб. заведений / В. И. Загвязинский, Р. Атаханов. – 2-е изд., стер. – Москва : Издательский центр «Академия». – 208 с. – [Электронный ресурс]. – Режим доступа: https://prepod.nspu.ru/pluginfile.php/45344/mod_resource/content/0/Zagvjazinskii.pdf. – Дата доступа: 15.10.2020
3. Звонников, В. И. Современные средства оценивания результатов обучения : учебник для студ. учреждений высш. проф. образования / В. И. Звонников, М. Б. Чельшкова. – 5-е изд., перераб. – Москва : Издательский центр «Академия». – 304 с. – (Сер. Бакалавриат). – [Электронный ресурс]. – Режим доступа: https://www.studmed.ru/zvonnikov-vi-chelyshkova-mb-sovremennye-sredstva-ocenivaniya-rezultatov-obucheniya_of8ff339642.html. – Дата доступа: 15.10.2020
4. Математические методы в педагогических исследованиях : учебное пособие / С.И. Осипова, С.М. Бутакова, Т.Г. Дулинец, Т.Б. Шаипова. – Красноярск : Сибирский федеральный университет. – 264 с. – [Электронный ресурс]. – Режим доступа: <http://www.studentlibrary.ru/book/ISBN9785763825060.html>. – Дата доступа: 15.10.2020
5. Пашкевич, О. И. Статистическая обработка эмпирических данных в системе STATISTICA : учебно-методическое пособие / О. И. Пашкевич – Электрон. текстовые данные. – Минск : РИПО. – 148 с. – [Электронный ресурс]. – Режим доступа: <http://www.iprbookshop.ru/67607.html>. – Дата доступа: 15.10.2020
6. Степанова, Е. А. Основы обработки результатов измерений [Электронный ресурс] : учебное пособие / Е. А. Степанова, Н. А. Скулкина, А. С. Волегов. – Электрон. текстовые данные. – Екатеринбург: Уральский федеральный университет, ЭБС АСВ. – 96 с. – Режим доступа: <http://www.iprbookshop.ru/68268.html>. – Дата доступа: 15.10.2020
7. Яковлев, Е. В. Педагогическое исследование: содержание и представление результатов / Е. В. Яковлев, Н. О. Яковлева. – Челябинск : Изд-во РБИУ. – 317 с. – [Электронный ресурс]. – Режим доступа: <http://elib.cspu.ru/xmlui/bitstream/handle/123456789/278/Пед.%20исследование%202010.pdf?sequence=1&isAllowed=y>. – Дата доступа: 15.10.2020

1.2 Оценивание и измерение в педагогическом процессе

- *Оценивание и измерение в педагогическом процессе*
- *Культура измерительной деятельности в образовании*
- *Таблицы первичных эмпирических данных. Приведенные таблицы*

- *Оценивание и измерение в педагогическом процессе*

По наиболее распространенному определению, введенному американским психологом С. Стивенсом в 1946 г., *измерение* – это процедура приписывания чисел некоторым характеристикам объектов в соответствии с определенными правилами. Предложенное Стивенсом определение появилось в результате формального обобщения опыта количественных измерений, широко распространенных в физике и других естественных науках, но его на протяжении многих лет брали за основу и при трактовке представлений об измерениях в образовании.

По мере развития педагогики, психологии, социологии и других эмпирических наук возникла потребность во введении не только количественных, но и качественных оценок, поскольку, помимо величин, встречающихся в естественных науках, существуют другие показатели, отличающиеся по степени проявления того или иного свойства и поддающиеся измерению. Пожалуй, наиболее важную роль здесь сыграли социологические обследования и психологические тесты, по результатам которых выявлялись межкультурные различия либо социальные общности, строящиеся на основе факторного анализа результатов выполнения тестов. По данным факторизации проводилось многомерное шкалирование, результаты которого позволяли объединять сходные группы индивидов. Им приписывались некоторые символы или просто числа, в последнем случае создавалась видимость получения количественных оценок. На деле эти оценки имели качественный характер в силу того, что они получались путем привлечения экспертов.

Связь между количественными и качественными оценками устанавливается на основе традиционно сложившегося разделения их на две группы: классифицирующие или сравнительные оценки относят к категории качественных, метрические – считаются количественными. Качественные оценки всегда бывают менее точными по сравнению с количественными в силу способов и инструментов, применяемых для их получения. Например, на устном экзамене классифицирующие понятия (знающий, аттестованный, подготовленный или «отлично», «хорошо», «удовлетворительно» и др.) определяются субъективно с помощью суждений экзаменатора и содержат в себе, как правило, значительный ошибочный компонент. Поскольку эти суждения принимаются за оцениваемые характеристики объектов, то их надо уметь распознавать, отождествлять, сравнивать, абстрагировать и обобщать. Для удобства оценочные суждения выражают в виде некоторых баллов по традиционной пятибалльной шкале или по любой другой, которая выбирается на основе экспертных соглашений. Приписываемые числа могут быть не

вполне определены в тех случаях, когда для построения качественных шкал не используются измерители. Например, при традиционном контроле у каждого преподавателя есть свои представления о том, за что нужно ставить 5, 4, 3 и т.д.

Измерения можно подразделить на физические, характерные для естественных наук, и нефизические, встречающиеся в науках эмпирических. По времени появления физические измерения предшествовали нефизическим, потребность в последних сформировалась и стала развиваться в связи с запросами педагогики, психологии, экономики и социологии, нуждающихся в отражении на шкале оценочных суждений для их фиксации и сопоставления.

Принципиальное отличие одного класса измерений от другого лежит в области объектов, подвергающихся оценкам. В физических измерениях в большинстве своем присутствуют вполне реальные объекты. В эмпирических – все оцениваемые характеристики носят в основном латентный (скрытый) характер, препятствующий непосредственному измерению. В силу латентности оцениванию подвергаются не сами характеристики, а их эмпирические референты – наблюдаемые признаки характеристик. Выбор таких референтов происходит интуитивно, поэтому их соответствие латентным характеристикам нуждается в тщательном доказательстве, проводимом на основе экспертного и статистического анализа эмпирических результатов измерения.

При оценивании латентные характеристики (переменная А) являются первопричиной, порождающей наблюдаемое множество результатов измерения (переменная В). Однако на практике совершается обратный шаг, поскольку, именно отгалкиваясь от наблюдаемых результатов, можно получить оценки латентных характеристик с определенной ошибкой измерения (рисунок 1.2.1).

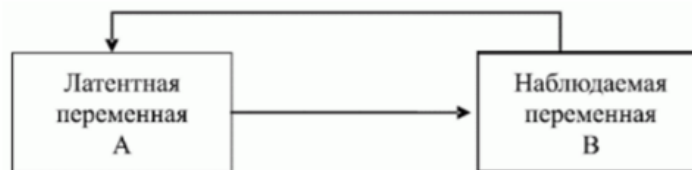


Рисунок 1.2.1 - Связь между латентной и наблюдаемой переменными

Неизбежная латентность переменных измерения, в роли которых в образовании выступают уровни освоения знаний, умений, навыков или компетенций, приводит к необходимости проверки адекватности полученных оценок концептуальной переменной (переменным) измерения и точности оценок, что обеспечивается благодаря коррекции свойств измерителя.

С течением времени в эмпирических науках определение измерений по Стивенсу было признано неудовлетворительным по двум причинам. Во-первых, оно не учитывало того, что результатами измерения могут быть не только числа. Во-вторых, оно никак не регламентировало правила измерения, хотя вся суть процедуры кроется именно в тех самых правилах, в соответствии

с которыми осуществляется приписывание чисел. В связи с этим в 80-х гг. XX в. появилась современная теория измерений, строящаяся на более строгой аксиоматической основе и значительно расширяющая традиционные представления об измерениях в психологии и образовании.

В соответствии с современными воззрениями, педагогическое измерение можно трактовать как конструирование любой функции, осуществляющей изоморфное отображение некоторой эмпирической структуры в соответствующим образом подобранную числовую структуру, представляющую собой шкалу оценок качества учебных достижений или других конструкторов, являющихся целью измерения. Изоморфизм – важное понятие математики, которое определяет ряд условий для взаимно однозначного отображения двух множеств с сохранением их свойств процессе отображения. Хотя само по себе это понятие возникло в высшей алгебре, в наше время оно трактуется довольно широко и используется не вполне строго в понятийном аппарате ряда наук, как, например, в теории педагогических измерений.

Отображающая функция понимается достаточно обобщенно, поскольку позволяет устанавливать различные, не только количественные отношения между объектами, но и проводить любые, в том числе качественные оценки. Под процедурой измерения предлагается понимать совокупность эмпирических операций, позволяющих установить оценки измеряемых характеристик и представить их в количественной или качественной шкале. Поскольку эмпирическая структура и строящаяся по результатам оценивания числовая структура (шкала) изоморфны, то, имея шкалу, можно, не обращаясь непосредственно к измеряемым объектам, восстановить все их свойства, характерные для эмпирической структуры.

Процесс педагогических измерений включает:

- выбор предмета измерения, под которым понимается одна или несколько латентных характеристик объектов (выбор переменных измерения и их числа);
- выбор эмпирических референтов (наблюдаемых характеристик объектов), замещающих латентные переменные;
- выбор измерительных процедур;
- конструирование и использование измерительных инструментов;
- выбор шкалы (если измеряемая переменная одна) или шкал (если измеряют более одной переменной при многомерных измерениях);
- построение отображения результатов измерения на шкалу (шкалы в случае многомерных измерений) по определенным процедурам и правилам;
- анализ и интерпретацию результатов измерения.

Таким образом, процесс педагогических измерений состоит из совокупности операций, позволяющих получить количественные или качественные оценки измеряемых характеристик. Правильная организация процесса педагогического измерения очень важна, поскольку любые нарушения стандартизованных условий проведения, обработки, анализа и

интерпретации результатов измерения снижают обоснованность и точность получаемых оценок. Приведенные компоненты процесса измерения имеют свои аналоги в традиционном контроле, где эти процедуры носят интуитивный характер и основываются на эмпирическом опыте педагога. При тестировании каждая компонента проходит процесс анализа и научного обоснования качества, базирующегося на теории педагогических измерений. Особенно важно такое обоснование, если речь идет об итоговом контроле, результаты которого связаны с принятием важных административных решений при аттестации выпускников или оценке эффективности деятельности учебных заведений. Тогда объектом измерения становятся знания, умения, навыки и компетенции, нередко называемые обобщающим термином «учебные достижения».

Отклонения получаемых результатов измерения от истинных значений измеряемых латентных характеристик представляют большую опасность в эмпирических науках, так как в силу латентного характера измеряемых переменных их не всегда можно своевременно выявить и скорректировать. Величину такого отклонения принимают за ошибку измерения. Возможность оценивания ошибки измерения является тем ключевым критерием, который позволяет провести четкую грань между оценками (количественными или качественными), которые можно считать результатами измерения, и результатами традиционного контроля. Эта же возможность выступает в качестве основного аргумента в защиту тестов. В отличие от традиционных средств контроля, не обеспечивающих никаких данных о точности оценок достижений студентов, тесты позволяют установить величину ошибки измерения и оценить тем самым надежность полученных оценок.

Переменные и компоненты педагогических измерений. Надежность и валидность результатов измерения. В силу существования ошибок измерения оцениваемые характеристики объектов могут принимать более или менее точные значения, поэтому такие характеристики обычно называют переменными измерения. По этой же причине принято говорить о надежных или ненадежных педагогических измерениях, где надежность характеризует степень устойчивости (повторяемости) и точности результатов измерения. Другая характеристика качества результатов педагогического измерения обычно называется валидностью. В самом общем случае она трактуется как адекватность эмпирических результатов поставленной цели измерения.

Цели педагогического измерения могут быть самыми разными, поэтому анализ валидности должен быть многоаспектным. Важнейшее место в нем занимает оценка конструктивной валидности, строящаяся на доказательстве адекватности эмпирических референтов концептуально выделенной переменной (переменных) измерения. Таким образом, в практической работе по созданию теста анализ конструктивной валидности находит свое отражение в работе по проведению доказательства того, что выбранные тестовые задания по своему содержанию и статистическим характеристикам позволяют оценить

именно те аспекты подготовленности обучаемого, которые создатель теста намеревался измерить.

Проблема оценивания конструктивной валидности приобретает особую актуальность в связи с введением компетентностного подхода к трактовке качества результатов образования. Для выявления конструктивной валидности приходится привлекать факторный и корреляционный анализ, оценивать число измеряемых переменных, сопоставлять содержательную трактовку результатов измерения с результатами, полученными по давно существующим и зарекомендовавшим себя тестам.

Измерительный инструмент включает два компонента. Один из них – само измеряющее устройство, роль которого при педагогических измерениях в большинстве случаев отводится тесту. *В самом обобщенном виде под тестом* можно понимать совокупность контрольных заданий в стандартизированной форме, обладающих необходимыми системообразующими статистическими характеристиками и обеспечивающих надежные и валидные оценки концептуально выделенной переменной (переменных) измерения. Таким образом, в самом определении теста заложены требования к его качеству, отсутствующие в традиционных оценочных средствах.

Многие преподаватели, никогда не принимавшие участия в процессе создания теста профессионалами и не имеющие специальной подготовки по теории педагогических измерений, путают тесты с наборами контрольных заданий, необоснованно полагая, что достаточно последние представить в форме задания с выбором ответа, как сразу получится педагогический тест. Этому поверхностному и неправильному восприятию педагогических тестов способствует латентная (скрытая от возможностей непосредственного наблюдения) природа характеристик тестовых заданий, которые проявляются только в процессе применения теста на выборке учащихся и выявляются при обработке и анализе эмпирических данных тестирования. Иными словами, узнать тест это или набор обычных контрольных заданий можно только после анализа результатов апробации заданий на представительной выборке учащихся.

Анализ предваряет специальная обработка данных апробации, основанная на аппарате теории педагогических измерений, позволяющая оценить статистические характеристики заданий, соотнести их с требованиями теории педагогических измерений и наметить пути коррекции характеристик для получения из набора контрольных заданий педагогического теста. Сама по себе форма заданий не имеет значения, поскольку с первого взгляда тестовые задания могут ничем не отличаться от традиционных контрольных и не иметь готовых ответов. Внешним необходимым, но не достаточным признаком теста может служить только число заданий, которых не должно быть меньше 25.

Таким образом, радикальное отличие теста от обычных заданий состоит не в форме представления вопросов и ответов, а в опоре на теорию педагогических измерений в процессе его создания и применения, что позволяет получить ряд важных преимуществ, отсутствующих в тех случаях,

когда не используются тесты. В частности, благодаря аппарату теории педагогических измерений можно:

- повысить полноту охвата содержания дисциплины в контрольных материалах за счет большого числа заданий в тесте (не менее 25);
- проанализировать степень достоверности (уровень объективности) оценок подготовленности учащихся, выявив величину ошибки измерения (надежность результатов измерений);
- проанализировать степень обоснованности оценок подготовленности учащихся, выявив уровень соответствия полученных результатов контроля поставленной цели его проведения путем оценивания величины валидности результатов измерений;
- получить статистические оценки трудности заданий, отличающиеся от экспертных оценок высокой достоверностью;
- повысить эффективность контроля путем оптимизации трудности, валидности и системообразующих свойств (оценок корреляции) тестовых заданий при создании теста;
- обеспечить высокую сопоставимость оценок учащихся за счет представления результатов тестирования в стандартных шкалах, введения норм и выравнивания при сравнительном анализе результатов выполнения тестов.

Второй компонент измерительного инструмента – заранее подготовленная шкала, которая служит для фиксации оценок (количественных или качественных) измеряемой переменной. Согласно наиболее распространенному определению, шкала – это средство упорядочения результатов измерения определенных свойств эмпирических объектов. В процессе такого упорядочения каждому элементу совокупности наблюдаемых эмпирических данных ставится в соответствие определенный балл, устанавливающий положение наблюдаемого элемента на шкале. В ситуации педагогического измерения на шкале откладываются не сырые, а производные баллы, которые подвергаются преобразованию для повышения сопоставимости и удобства интерпретации результатов выполнения педагогического теста.

Шкала с отложенными оценками переменной является целью измерения. При измерениях с высокой надежностью и валидностью шкала адекватно отображает, фиксирует оцениваемые характеристики, представляя их без существенных искажений. Таким образом, суть процедуры измерения состоит в переходе от самих объектов измерения к шкале, на которой выстроены оценки, замещающие исследуемые характеристики этих объектов. Подобная процедура позволяет уйти от объектов измерения и сосредоточить внимание на измеряемых характеристиках. В зависимости от количества измеряемых переменных можно говорить об одномерных (одна переменная) или многомерных (более одной переменной) измерениях. Соответственно по результатам измерения строится одна шкала или несколько, число которых в последнем случае равно числу переменных измерения.

При создании аттестационных тестов в рамках компетентностного подхода по циклам специальных дисциплин обычно используют многомерные измерения, результаты которых откладываются не на одной, а на нескольких шкалах. В других случаях, когда речь идет об итоговой аттестации учащихся на младших курсах, стараются ограничиться одномерными тестами. Вне зависимости от числа переменных построение шкалы возможно только по надежным и валидным данным, обладающим высокой объективностью и адекватностью цели измерения. Поэтому размерность, надежность и валидность являются взаимосвязанными свойствами, характеризующими различные аспекты качества педагогических измерений. Выявление размерности – это необходимый предварительный этап работ по оцениванию надежности и валидности результатов измерений. Таким образом, качество данных, полученных в педагогических измерениях, проходит процесс анализа, а характеристики самих измерителей являются предметом научного обоснования, основанного на специальной теории. В традиционном контроле такое обоснование отсутствует, поскольку не формируются никакие представления об оцениваемых переменных, их соответствии планируемым целям контроля и их точности.

Последний компонент процесса педагогических измерений – обработка, анализ и интерпретация данных – имеет двойное назначение. В первом случае, при котором использование измерительных инструментов предназначается для их апробации, по результатам обработки выявляют качество результатов измерения для проведения коррекции и стандартизации инструментария. Во втором случае, когда уже есть апробированные стандартизованные измерители, полученные данные представляются в удобной для пользователей форме для интерпретации согласно поставленным задачам контроля и управления качеством образования. Среди трех компонентов – обработка, анализ и интерпретация – этап обработки носит первичный характер. Он завершается количественным анализом результатов измерения, данные которого используются либо для научного обоснования качества тестов и преобразования сырых баллов испытуемых в производные баллы (процесс шкалирования), либо для повышения удобства и сопоставимости при интерпретации результатов в учебном процессе. Таким образом, если измеритель разработан, и результаты измерений удовлетворяют требованиям высокой надежности, валидности и сопоставимости, то можно выстраивать интерпретацию результатов обучаемых, проводить мониторинг и принимать обоснованные решения в управлении качеством образования.

- ***Культура измерительной деятельности в образовании***

Культу́ра (от лат. cultura — возделывание, позднее — воспитание, образование, развитие, почитание) — понятие, имеющее огромное количество значений в различных областях человеческой жизнедеятельности. Культура является предметом изучения философии, культурологии, истории,

искусствоведения, лингвистики (этнолингвистики), политологии, этнологии, психологии, экономики, педагогики и др.

В основном, под культурой понимают человеческую деятельность в её самых разных проявлениях, включая все формы и способы человеческого самовыражения и самопознания, накопление человеком и социумом в целом навыков и умений. Культура предстает также проявлением человеческой субъективности и объективности (характера, компетентностей, навыков, умений и знаний).

Культура представляет собой совокупность устойчивых форм человеческой деятельности, без которых она не может воспроизводиться, а значит — существовать.

Культура — это набор правил, которые предписывают человеку определённое поведение с присущими ему переживаниями и мыслями, оказывая на него, тем самым, управленческое воздействие.

Источником происхождения культуры мыслится человеческая деятельность, познание и творчество.

Рассматривая объективность оценочной деятельности педагога как культуру измерительной деятельности следует остановиться на следующих моментах. Хотя объективность является наиболее понятной всем пользователям тестов и часто используется при обсуждении достоинств тестирования в средствах массовой информации, ограничиваться ею при оценивании точности измерений нельзя, поскольку при таком подходе не выдвигается никаких требований к качеству теста. Может создаться впечатление, что для получения объективных данных о подготовленности испытуемых достаточно перейти от традиционных экзаменов к любым наборам заданий в тестовой форме, устранив влияние педагога на оценку тестирования и поручив подсчет баллов испытуемых компьютерной программе.

Типичный пример подобного упрощенного толкования объективности встречается сейчас среди части педагогов, учащихся и их родителей в связи с проведением ЕГЭ. Широкий круг пользователей системы ЕГЭ не видит той работы профессионалов, которая стоит за созданием контрольно-измерительных материалов (КИМ), но в абсолютном большинстве своем полагает, что устранение педагогов из контрольно-оценочного процесса, сведение их участия к минимальному на этапе организации процедуры тестирования и выставления баллов позволяет говорить об объективности результатов ЕГЭ. На деле такое упрощенное толкование наносит вред развитию тестирования в целом, поскольку допускает применение любых, в том числе некачественных, тестов. Оно не опирается на теорию педагогических измерений и лишь частично охватывает научные представления об объективности результатов выполнения теста.

На ранних этапах истории развития педагогических измерений объективность считали неотъемлемой характеристикой результатов тестирования, поскольку стандартизованная проверка результатов выполнения

тестов исключала педагога из оценочного процесса. По мере развития теории и практики применения тестов пришло понимание того, что абсолютная объективность – это недостижимая характеристика результатов любых, в том числе и педагогических, измерений в силу существования неизбежных ошибочных компонентов. На результаты измерений воздействуют различные факторы случайного и систематического характера, которые неизбежно влияют на величину ошибочных компонентов, изменяя их как в сторону увеличения, так и уменьшения. Поэтому в оценочных процессах можно лишь говорить о высокой или низкой объективности, оценка которой основана на выявлении надежности результатов педагогических измерений.

Наиболее полно трактовка термина «объективность измерений» представлена в исследованиях Вебстера. Он был первым ученым, осуществившим методологические исследования в этой области и предложившим восемь толкований обсуждаемого термина, три из которых непосредственно относятся к педагогическим измерениям: процедурная объективность, классическая (традиционная) объективность и инвариантная (специфическая) объективность.

Под процедурной объективностью понимается независимость результатов измерения от условий его проведения и субъективных оценочных суждений педагога, использующего тест. Эта независимость обеспечивается благодаря равенству условий тестирования для всех учащихся, отсутствию списывания или других нарушений дисциплины, правильному инструктированию учащихся, использованию для одной группы тестируемых параллельных (совпадающих по трудности и другим характеристикам) вариантов теста, стандартизации процедуры проверки результатов и максимальной ее автоматизации. Чем выше уровень применения компьютерных технологий в тестировании, исключаящих любое субъективное влияние педагога или экспертов, тем больше величина процедурной объективности результатов тестирования, получаемой при измерении.

Второе углубленное понимание объективности измерений основано на некоторых понятиях и аксиомах традиционной или, как ее называют иначе, классической теории тестирования. К числу таких понятий относится сырой (наблюдаемый, индивидуальный) балл, получаемый простым суммированием результатов испытуемого по отдельным заданиям теста. При дихотомической оценке результатов по заданиям (1 или 0) индивидуальный балл просто равен количеству правильно выполненных заданий теста. Другое понятие классической теории – истинный балл (параметр испытуемого) – отождествляется с абсолютно объективной оценкой испытуемого, свободной от влияния любых ошибок измерения. В отличие от наблюдаемого балла, который меняется в зависимости от теста и способа подсчета результата испытуемого, истинный балл трактуется как не зависящая от средств измерения константа учащегося, характеризующая оцениваемое свойство

испытуемого в момент измерения с нулевой ошибкой измерения и меняющаяся в процессе обучения.

Согласно основной аксиоме классической теории тестов, предполагается, что любой наблюдаемый балл отличается от истинного на величину ошибки измерения (p -уровень). В тех случаях, когда ошибка измерения не превышает выбранные пределы точности измерений, говорят о высокой объективности результатов тестирования и принимают индивидуальные баллы испытуемых за их истинные баллы. Таким образом, углубленное понимание объективности измерений требует оценивания величины ошибки измерения, на размер которой влияют не только условия проведения тестирования, но и характеристики самого теста. Если эти характеристики не удовлетворяют определенным требованиям к качеству измерителей, то ошибка измерения, как правило, становится слишком большой, превышающей допустимую погрешность измерения и не позволяющей говорить о высокой объективности полученных результатов выполнения теста.

Третья трактовка объективности лежит в области современной теории конструирования тестов Item Response Theory (IRT). Преимущества IRT, позволяющие оценить подготовленность обучаемых независимо от трудности заданий теста, обеспечивают достижение так называемой инвариантной объективности измерений, которая превышает объективность, обеспечиваемую при использовании классической теории тестов.

Согласно основным положениям теории IRT, инвариантная объективность характеризуется тем, что на оценку каждого испытуемого выборки не влияют оценки других испытуемых и трудность заданий теста. Благодаря этому оценки подготовленности студентов, обладающие высокой инвариантной объективностью, будут более точными, чем те, которые получаются при использовании традиционной теории при конструировании и применении тестов.

• ***Таблицы первичных эмпирических данных. Приведенные таблицы.***

В математической статистике основным способом упорядоченного представления данных, полученных в результате педагогического исследования, является таблица. Любая таблица состоит из подлежащего и сказуемого. Подлежащее содержит в себе порядковый номер, имя и фамилию испытуемого (или же его код), названия методик (признаков, характеристик, факторов, шкал и др.) по которым он обследуется, и т.д. Само по себе подлежащее не имеет числового измерения и выполняет функцию обозначения той или иной переменной.

Подлежащее может рассматриваться как простое, если его деление носит одноразовый характер, по одному критерию, и как комбинаторное, если деление осуществляется по нескольким критериям. Левая и верхняя шапки таблицы выполняют функцию подлежащего.

Сказуемое является числовой характеристикой и относится к подлежащему, определяет его. Числовой материал, находящийся в таблице, выполняет функцию сказуемого.

Различают следующие виды таблиц – вычислительные и аналитические.

Вычислительные таблицы – таблицы, с помощью которых производят расчеты, обрабатывают данные и – в итоге – на их основе делают количественные выводы.

Аналитические таблицы – таблицы, данные в которых представляются таким образом, чтобы на их основе можно было делать качественные выводы.

Таблицы первичных эмпирических данных (ПЭД). Задача таблиц ПЭД состоит в том, чтобы все данные, полученные в ходе педагогического исследования, представить унифицировано, в одной таблице. На основе таблиц ПЭД строятся так называемые групповые таблицы или ряды распределений, в которых упорядочение осуществляется по одному признаку, и комбинаторные, или таблицы распределения.

На основе таблиц ПЭД строятся также и **приведенные таблицы**, содержащие числовую информацию, подлежащую компьютерной обработке.

Таблицы распределения – n -мерная таблица, где n – количество признаков, по которым проводится анализ.

Эмпирическая таблица распределения – такая таблица распределения, которая строится на основе таблицы первичных эмпирических данных.

Теоретическая таблица распределения – таблица распределения, построенная при условии, что распределение испытуемых носит случайный характер.

Работа в системе STATISTICA начинается с ввода данных. Исходные данные в системе STATISTICA организованы в виде таблиц. Если у пользователя имеется опыт работы с электронными таблицами (типа MS Excel), то он быстро привыкнет и к таблицам системы STATISTICA. Следует отметить, что табличная структура данных STATISTICA позволяет естественно отобразить большинство реальных данных.

Электронная таблица состоит из строк и столбцов. Столбцы таблицы STATISTICA называются Переменные (Variables), а строки – Наблюдения (Cases).

Например, в психолого-педагогических исследованиях наблюдения – это испытуемые, а в качестве переменных могут выступать пол, возраст, название учебного заведения, специальность, успеваемость и т. д. Существует два варианта создания таблицы с данными: непосредственно в системе STATISTICA или в одной из офисных программ MS Word либо MS Excel.

Поскольку система STATISTICA является обычным Windows-приложением, можно легко и быстро импортировать данные, полученные в системе STATISTICA, в другое Windows-приложение, например, в MS Word. Для этого необходимо нажать одновременно кнопки ALT и F3 (импорт в виде рисунка). На экране вместо курсора мыши появится значок «прицел». Используя мышь, надо поместить «прицел» в верхний левый

угол таблицы. Затем нажать левую кнопку мыши, зафиксировав «прицел», и, удерживая кнопку мыши, переместить его в правый нижний угол таблицы. Выделенная часть таблицы будет отмечена прямоугольной рамкой. После того как будет отпущена кнопка мыши, отмеченная часть таблицы будет помещена в буфер обмена. Если теперь открыть нужный документ MS Word и нажать на клавиатуре комбинацию кнопок CTRL и V, то выбранный сегмент таблицы будет скопирован в документ. Импорт данных из системы STATISTICA в виде таблицы предполагает предварительное выделение, а затем, через нажатие правого клика мыши, вызова контекстного меню и выбора команды Copy with Headers. Необходимая информация будет помещена в буфер обмена. Далее нужно открыть программу MS Word и нажать на клавиатуре комбинацию кнопок CTRL и V.

Литература:

1. Ермолаев, О. Ю. Математическая статистика для психологов: Учебник // 2-е изд., испр. / О. Ю. Ермолаев. – Москва : Московский психолого-социальный институт : Флинта. – 336 с. – [Электронный ресурс]. – Режим доступа: <https://e-catalog.nlb.by/Author/Home?author>. – Дата доступа: 15.10.2020

2. Загвязинский, В. И. Методология и методы психолого-педагогического исследования: учеб. пособие для студ. высш. пед. учеб. заведений / В. И. Загвязинский, Р. Атаханов. – 2-е изд., стер. – Москва : Издательский центр «Академия». – 208 с. – [Электронный ресурс]. – Режим доступа: https://preod.nspu.ru/pluginfile.php/45344/mod_resource/content/0/Zagvjazinskii.pdf. – Дата доступа: 15.10.2020.

3. Звонников, В. И. Измерения и шкалирование в образовании (Текст) : (учеб. пособие) / В. И. Звонников. – Москва : Логос. – 134 с. – [Электронный ресурс]. – Режим доступа: http://library.ntspi.ru/CGI/irbis64r_01/cgiirbis_64.exe. – Дата доступа: 15.10.2020.

4. Пашкевич, О. И. Курсовая и дипломная работы: структура, оформление, защита : методические рекомендации для слушателей переподготовки по специальности 1-03 04 72 "Практическая психология" (квалификация "Педагог-психолог") / О. И. Пашкевич, О. М. Савчик. – Минск : РИПО. – [Электронный ресурс]. – Режим доступа: [http://unicat.nlb.by/opac/pls/!search.http_keyword?query=a001a="BY-SEK-277528"&lst_siz=20](http://unicat.nlb.by/opac/pls/!search.http_keyword?query=a001a=). – Дата доступа: 15.10.2020

5. Пашкевич, О. И. Статистическая обработка эмпирических данных в системе STATISTICA : учебно-методическое пособие / О. И. Пашкевич – Электрон. текстовые данные. – Минск : РИПО. – 148 с. – [Электронный ресурс]. – Режим доступа: <http://www.iprbookshop.ru/67607.html>. – Дата доступа: 15.10.2020

1.3 Качественный анализ результатов педагогического исследования

- *Качественный анализ результатов педагогического исследования*
- *Табличное представление результатов педагогического исследования. Баллы, уровни и ранги*

- *Качественный анализ результатов педагогического исследования*

Сущность качественного анализа. Качественный анализ результатов – совокупность процедур и методов описания исследовательских данных на основе теоретических умозаключений и обобщений, индивидуального опыта, интуиции, методов логического вывода.

В ходе качественного анализа результатов педагогического исследования выявляются причины возникновения того или иного явления, вскрываются его существенные свойства, устанавливаются тенденции развития, определяются противоречия функционирования.

Опыт качественной обработки данных педагогического исследования позволяет выделить ее следующие общие компоненты:

- предварительный анализ исследуемых данных;
- группировку однопорядковых результатов исследования;
- анализ зависимостей между различными группировками;
- обобщающие выводы.

Выделяют следующие приемы качественного анализа:

- классификация;
- типологизация;
- систематизация;
- периодизация;
- категоризация;
- казуистика.

Категоризация – систематизация и дифференциация материала исследований по типам, видам, вариантам, построение схем, структур.

Казуистика – системное описание типичных и уникальных случаев, тех, которые представляют типологию, и тех, которые являются исключением.

Методы качественного анализа Методы качественного анализа разделены на две группы:

- анализ по аналогии (при этом используются такие приемы, как обращение к личному опыту, интуиции исследователя, к ссылкам на авторитет, на результаты аналогичных исследований);
- построение таблиц сопряженности;
- непараметрическая статистика.

Первичная обработка «сырых» оценок для создания возможности применения непараметрической статистики производится двумя методами: классификацией (разделение объектов на классы, группы и т.д. по какому-либо критерию) и систематизацией (упорядочивание объектов внутри

классов, групп и т.д., классов между собой, и множеств классов с другими множествами классов).

Таблицы сопряженности — средство представления совместного распределения двух переменных, предназначенное для исследования связи между ними. Таблица сопряженности является наиболее универсальным средством изучения статистических связей, так как в ней могут быть представлены переменные с любым уровнем измерения.

Строки таблиц сопряженности соответствуют значениям одной переменной, столбцы — значениям другой переменной (количественные шкалы предварительно должны быть сгруппированы в интервалы). На пересечении строки и столбца указывается частота совместного появления соответствующих значений двух признаков. Сумма частот по строке называется маргинальной частотой строки; сумма частот по столбцу — маргинальной частотой столбца. Сумма маргинальных частот равна объему выборки N .

В таблицах сопряженности могут быть представлены как абсолютные, так и относительные частоты (в долях или процентах). Относительные частоты могут рассчитываться по отношению: а) к маргинальной частоте по строке; б) к маргинальной частоте по столбцу; в) к объему выборки.

Таблицы сопряженности используются для проверки гипотезы о наличии связи между двумя признаками (критерий χ^2 -квадрат), а также для измерения тесноты связи (коэффициент Фи, коэффициент Чупрова, коэффициент контингенции, коэффициент Крамера).

Кросстабуляция — это процесс объединения двух (или нескольких) таблиц частот так, что каждая ячейка (клетка) в построенной таблице представляется единственной комбинацией значений или уровней табулированных переменных. Таким образом, кросстабуляция позволяет совместить частоты появления наблюдений на разных уровнях рассматриваемых факторов. Исследуя эти частоты, можно определить связи между табулированными переменными. Обычно табулируются категориальные (*номинальные*) переменные или переменные с относительно небольшим числом значений. Если необходимо табулировать непрерывную переменную (например, успеваемость), то вначале ее следует *перекодировать*, разбив диапазон изменения на небольшое число интервалов (например, успеваемость: неудовлетворительная, удовлетворительная, хорошая, отличная).

Простейшая форма кросстабуляции — это таблица сопряженности (2x2), в которой значения двух переменных «пересечены» (сопряжены) на разных уровнях и каждая переменная принимает только два значения, т.е. имеет два уровня (поэтому таблица называется «2 на 2»).

Критерий χ^2 -квадрат Пирсона. Это наиболее простой критерий проверки значимости связи между двумя категоризованными переменными. Критерий Пирсона основывается на том, что в двухходовой таблице *ожидаемые* частоты при гипотезе «между переменными нет зависимости»

можно вычислить непосредственно. Представьте, что 20 мужчин и 20 женщин протестированы относительно типа поведения в заданной ситуации (А или В). Если между типом реагирования и полом нет связи, то естественно *ожидать* равного выбора стиля поведения (доминирование-подчинение) для каждого пола.

Критерий ХИ-квадрат (метод максимального правдоподобия). Максимум правдоподобия ХИ-квадрат предназначен для проверки той же самой гипотезы относительно связей в таблицах сопряженности, что и критерий ХИ-квадрат Пирсона. Однако его вычисление основано на методе максимального правдоподобия (МП). На практике статистика МП ХИ-квадрат очень близка по величине к обычной статистике Пирсона ХИ-квадрат.

Поправка Йетса. Аппроксимация статистики ХИ-квадрат для таблиц (2x2) с малыми числом наблюдений в ячейках может быть улучшена уменьшением абсолютного значения разностей между ожидаемыми и наблюдаемыми частотами на величину 0.5 перед возведением в квадрат (так называемая поправка Йетса). Поправка Йетса, делающая оценку более умеренной, обычно применяется в тех случаях, когда таблицы содержат только малые частоты, например, когда некоторые ожидаемые частоты становятся меньше 10.

Точный критерий Фишера. Этот критерий применим только для таблиц (2x2). Критерий основан на следующем рассуждении. Даны маргинальные частоты в таблице, предположим, что обе табулированные переменные независимы. Зададимся вопросом: какова вероятность получения наблюдаемых в таблице частот, исходя из заданных маргинальных? Оказывается, эта вероятность вычисляется *точно* подсчетом всех таблиц, которые можно построить, исходя из маргинальных. Таким образом, критерий Фишера вычисляет *точную* вероятность появления наблюдаемых частот при нулевой гипотезе (отсутствие связи между табулированными переменными).

Хи-квадрат Макнемара. Этот критерий применяется, когда частоты в таблице (2x2) представляют *зависимые* выборки. Например, наблюдения одних и тех же испытуемых до и после эксперимента. В частности, можно подсчитывать число учащихся, имеющих минимальные успехи, например, по какой-либо спецдисциплине в начале и в конце учебного года. Вычисляются два значения ХИ-квадрат: A/D и B/C . Так, A/D хи-квадрат проверяет гипотезу о том, что частоты в ячейках А и D (верхняя левая, нижняя правая) одинаковы. А B/C хи-квадрат проверяет гипотезу о равенстве частот в ячейках В и С (верхняя правая, нижняя левая).

Коэффициент Фи. Фи-квадрат представляет собой меру связи между двумя переменными в таблице (2x2). Его значения изменяются от 0 (нет зависимости между переменными; ХИ-квадрат = 0.0) до 1 (абсолютная зависимость между двумя факторами в таблице).

Тетрахорическая корреляция. Эта статистика вычисляется (и применяется) только для таблиц сопряженности (2x2). Если таблица (2x2) может рассматриваться как результат (искусственного) разбиения значений

двух непрерывных переменных на два класса, то коэффициент тетраэрической корреляции позволяет оценить зависимость между двумя этими переменными.

Коэффициент сопряженности. Коэффициент сопряженности представляет собой основанную на статистике *XI-квадрат* меру связи признаков в таблице сопряженности (предложенную Пирсоном). Преимущество этого коэффициента перед обычной статистикой *Chi-квадрат* в том, что он легче интерпретируется, т.к. диапазон его изменения находится в интервале от 0 до 1 (где 0 соответствует случаю независимости признаков в таблице, а увеличение коэффициента показывает увеличение степени связи). Недостаток коэффициента сопряженности в том, что его максимальное значение «зависит» от размера таблицы. Этот коэффициент может достигать значения 1 только, если число классов не ограничено.

Сущность количественного анализа. Количественный (математико-статистический) анализ – совокупность процедур, методов описания и преобразования исследовательских данных на основе использования математико-статического аппарата.

Количественный анализ подразумевает возможность обращения с результатами как с числами – применение методов вычислений.

Решившись на количественный анализ, исследователь может сначала провести первичную и вторичную обработку данных и далее обратиться к помощи параметрической или непараметрической статистики.

На этапе первичной обработки решаются две основные задачи:

1) представить полученные данные в наглядной, удобной для предварительного качественного анализа форме в виде упорядоченных рядов, таблиц и гистограмм;

2) подготовить данные для применения специфических методов вторичной обработки.

Упорядочивание (расположение чисел в порядке убывания или возрастания) позволяет выделить максимальное и минимальное количественное значение результатов, оценить, какие результаты встречаются особенно часто и т.д. Набор показателей различных методик, полученных по группе представляют в виде таблицы, в строках которой располагают данные обследования одного испытуемого, представленные в баллах, уровнях, рангах (стенах), а в столбцах – распределение значений одного показателя по выборке.

Гистограмма – это частотное распределение результатов в диапазоне изменения значений.

На этапе вторичной обработки вычисляются характеристики предмета исследования. Анализ результатов вторичной обработки позволяет предпочесть тот набор количественных характеристик, который будет наиболее информативен. Цель этапа вторичной обработки состоит не только в получении информации, но и в подготовке данных к возможной оценке

достоверности сведений. В последнем случае мы обращаемся к помощи параметрической или же непараметрической статистики.

Различают следующие типы методов математико-статистического анализа:

- Методы описательной статистики направлены на описание характеристик исследуемого явления: распределения, особенностей связи и пр.

- Методы статистического анализа служат для установления статистической значимости данных, полученных в ходе экспериментов.

- Методы преобразования данных направлены на преобразование данных с целью оптимизации их представления и анализа.

К количественным методам анализа и интерпретации (преобразования) данных относятся следующие:

1. Дисперсионный анализ позволяет выявить, насколько дисперсия зависимой переменной соотносится с дисперсией независимой переменной, т.е. указывает на ту качественную переменную, которая вызывает изменения исследуемого качества. На практике в рамках дисперсионного анализа чаще всего применяются расчеты при помощи F- критерия Фишера

2. Корреляционный анализ выявляет взаимосвязь и направление между переменными.

3. Факторный анализ выявляет группы референтных переменных через систему факторов.

4. Регрессионный анализ заключается в моделировании системы связей нескольких зависимых переменных в единый фактор, отражающий влияние независимой переменной на исследуемый объект. При помощи регрессионной модели прогнозируется развитие фактора, причем прогноз имеет тенденцию к снижению достоверности через определенный промежуток времени.

5. Кластерный анализ позволяет выявить связь или степени подобия поведенческих реакций различных объектов по «подобию» их переменных (характеристик). В процессе кластерного анализа осуществляется группировка по определенным статистическим критериям различных «сходных» объектов (испытуемых или групп) в один класс (категорию, группу, кластер).

• Табличное представление результатов педагогического исследования. Баллы, уровни и ранги

Сводка опытных данных, полученных на основе проведенного эксперимента, представляет собой их систематизацию и установление качественных и количественных зависимостей между факторами, которые исследовались.

Для обработки результатов исследований, представленных в баллах («сырых» баллах), уровнях (низкий, средний, высокий и др.), рангах, (первый, второй, третий и т.д) чаще всего применяют статистические, табличные и графические методы.

Чтобы количественные показатели результатов исследования дали возможность выявить наличие некоторых зависимостей между исследуемыми факторами, их нужно определенным образом упорядочить. Подходящими для этого могут стать статистические методы ранжирование и группировка.

Ранжирование заключается в распределении количественных показателей в определенном порядке (например, по степени их важности или значимости, или в последовательности рост или наоборот уменьшение). На основе ранжирования исключают все второстепенные и случайные данные, не влияющие на результаты проведенного исследования.

Группировка предполагает разделение исследовательских данных на основе определенных показателей (показатели группировки) на группы из однотипных или близких по значению элементов. Показатели группировки могут быть количественными и качественными.

При группировании по количественными признаками (количественное группировки) за основу берут признаки, которые можно охарактеризовать количественно (например, результаты тестирования или опроса, скорость процессов, производительность труда, точность изготовления изделий и др.). Количественные признаки всегда можно измерить какими-то единицами измерения, а результаты измерения упорядочивать по определенной последовательностью (рост, уменьшение, периодическая повторяемость и тому подобное).

При группировании по качественным признакам (атрибутивные группировки) за основу берут признаки, которые невозможно охарактеризовать количественно, но они могут так повторяться, что это становится возможным определять (например, национальность или социальное происхождение опрошенных). Из качественных признаков невозможно составить некую последовательность.

Опытные данные могут быть сгруппированы по одной или нескольким признакам. По количеству признаков различают простое и комбинированное группирование. Простая группировка происходит по одному признаку (например, всех участников эксперимента можно разделить по стажу работы). Комбинированное группирование конкретной совокупности данных одновременно осуществляют на основе нескольких признаков (когда уже разделенных по стажу работы работников поделить еще и по уровню образования — то это уже будет группировка по двум признакам. Третьим признаком группировки могут стать возрастные категории и др.). Целесообразным считается группировки не более чем по трем признакам.

Полученные при исследовании количественные данные представляются обычно тремя способами: 1) перечисляются в тексте работы; 2) представляются в виде таблиц; 3) представляются в виде графических изображений (диаграммы, полигоны и кривые распределения и тому подобное).

Первый способ используется тогда, когда числовых данных мало.

Представление результатов педагогических исследований в таблицах. Статистически упорядоченные опытные данные для удобства дальнейшего использования заносят в таблицы. Применение таблиц следует считать основным способом представления количественных показателей (числовых данных) в педагогических исследованиях. Таблицы дают возможность отражать количественные признаки исследуемого объекта в сжатом виде, компактно, систематизировано и наглядно. При хорошо составленных таблицах каждое отдельное число можно найти в таблице объяснения для его понимания, так что их не нужно повторять в письменном тексте.

Таблица представляет собой такой способ подачи информации, при котором цифровой или текстовый материал группируется в строки и графы, отделенные одна от другой вертикальными и горизонтальными линиями.

По содержанию таблицы делятся на аналитические и неаналитические. Аналитические таблицы являются результатом обработки и анализа цифровых показателей. Как правило, после таких таблиц делается обобщение о новом (выведено) знании, которое вводится в текст словами: **"таблица позволяет сделать вывод, что..."**, **"из таблицы видно, что..."** и др. Часто такие таблицы способствуют выявлению и формулированию определенных закономерностей.

В неаналитических таблицах помещают, в основном, необработанные статистические данные, необходимые лишь для представления информации или констатации определенного положения вещей.

Приведенные в таблице данные удобно сравнивать и анализировать. Таблица может быть не только средством отображения собранных в ходе эксперимента данных, но и средством их интерпретации. Далеко не все равно, в какой таблице числовые данные представлены. Правильно выбран ее вид помогает лучше раскрывать сущность этих данных и связи между ними. Поэтому важно знать, как правильно составить таблицу, какой ее вид целесообразно избрать.

Для того, чтобы таблицы лучше передавали собранную в ходе исследования информацию, необходимо знать методику составления и использования таблиц, знать, какая таблица в том или ином случае лучше всего подходит для представления данных.

Каждая таблица состоит из двух частей: **текстовой и цифровой**. К текстовой части относятся заголовок таблицы и необходимые словесные пояснения, дополнительные замечания и ссылки. К цифровой части относятся статистические числовые данные, номер таблицы. Цифровая часть в таблицах нередко представляется в виде поля чисел.

Таблицы должны быть простыми, чтобы читатель быстро получал ясный и точный обзор о представленных в таблице числовых данных и связи между ними. Цифры и буквы в таблице не должны быть меньшими, чем цифры и буквы в основном тексте научной работы.

Размещают таблицы непосредственно после ссылки на них в тексте. Когда таблица занимает много места, тогда ее выполняют на отдельной странице и размещают сразу после той страницы, где есть ссылки на нее. Если таблица непосредственно с текстом не связана, то ее можно привести в приложениях.

Обычно таблица состоит из следующих элементов (таблица 1.3.1): порядкового номера и тематического заголовка, боковика, заголовков вертикальных граф (головки), горизонтальных строк и вертикальных граф (основной части, т. е. прографки).

Таблица 1.3.1 - Состояние системы управления воспитанием учащихся (образец простой таблицы)

Критерии оценивания	Экспертная оценка (по десятибалльной шкале)
Иерархия целей по уровням управления	6
Соответствие структуры системы управления целям	7
Систематическое осуществление анализа эффективности	3
Сформированность системы воспитания	4
Наличие и использование четких критериев для оценки существующей воспитательной системы	3

Таблицы нумеруют арабскими цифрами по порядку, с начала работы. **Номер таблицы** пишется на одну строку выше заголовка над правым углом таблицы. Пишут: Таблица 15. Если есть разделы, тогда первой цифрой будет номер раздела, а затем порядковый номер таблицы. Например, 1.11. Читается так: таблица № 11 первого раздела.

Если приведенные в таблице параметры имеют только одну дименцию (например, возраст учащихся в годах), то сокращенное обозначение последней выносится в общий заголовок таблицы. Если же цифры в столбцах таблицы имеют разные дименции, их сокращения отмечаются как подзаголовки каждого столбца.

Параметры, которые находятся в одном столбце, должны иметь одинаковое количество десятичных мест.

Деление головки таблицы по диагонали не допускается.

Информативность таблицы зависит от удачного ее построения. Правильный выбор структуры таблицы помогает лучше раскрыть сущность помещенных в нее данных и связи между ними. По структуре различают **простые, групповые, комбинированные, составные, шахматные** и некоторые другие таблицы.

1. Простая таблица. Содержит перечень данных (таблица 1.3.1), собранных о отдельные факторы изучаемого объекта или об отдельных явлениях рассматриваемой совокупности.

2. Групповая таблица. Это такая таблица, в которой отдельные члены изучаемой совокупности сгруппированы по каким-либо одним признаком (таблица 1.3.2).

Таблица 1.3.2 - Распределение респондентов по стажу педагогической работы (образец групповой таблицы)

Стаж педагогической работы (лет)	Количество участников опроса
До 3	12
От 3 до 5	18
От 5 до 10	27
От 10 до 20	24
Более 20	38

3. Комбинированная таблица. Содержит данные о совокупности, сгруппированные больше, чем по одному признаку (таблица 1.3.3).

Если в головке таблицы дать сначала возраст учеников, а потом разделить в свою очередь каждую возрастную группу на мальчиков и девочек, образуется группировка по полу учащихся разного возраста.

Как видно из приведенного в таблице 7.3 примеру, комбинированная таблица создает более лучшие условия для сравнения и анализа, чем простые и групповые таблицы.

Таблица 1.3.3 - Распределение учащихся экспериментальных школ по полу и возрасту (образец комбинированной таблицы)

Школа	Возраст учащихся (лет)								Всего учащихся
	Мальчики				Девочки				
	8	от 8 до 10	от 10 до 12	Вместе	8	от 8 до 10	от 10 до 12	Вместе	
№ 22	42	36	53	131	53	41	48	142	273
№ 34	39	38	56	133	50	44	45	139	272

4. Составленную таблицу образуют путем объединения нескольких простых или групповых таблиц (таблица 1.3.4). Как видно, 2-й и 3-й столбцы составленной таблицы образуют отдельную таблицу, а вторую

самостоятельную таблицу образуют 4-й, 5-й и 6-й столбцы. По этой таблице нельзя определить сколько среди учащихся до 8 лет имеется мальчиков и девочек. Комбинированная таблица это показывает.

Таблица 1.3.4 - Распределение учащихся экспериментальных школ по полу и возрасту (образец составленной таблицы)

Школа	Пол учащихся		Возраст учащихся (лет)		
	мальчики	девочки	8	от 8 до 10	от 10 до 12
№ 22	131	142	95	77	101
№ 34	133	139	89	82	101

Составленную таблицу следует отличать от комбинированной таблицы. В отличие от комбинированной составлена таблица обычно образуется из двух или более объединенных простых или групповых таблиц (таблица 1.3.4). Сравнение комбинированной и составленной таблицы показывает, что первая с точки зрения представления данных значительно содержательнее. Из нее можно узнать те же данные, что и из составленной таблицы, и, кроме того, данные о распределении мальчиков и девочек по возрасту или распределение по полу учащихся разного возраста.

Выбор комбинированной или составленной таблицы для представления данных зависит от целей исследования. Если таблица нужна как **таблица представления**, то есть когда она должна дать широкое представление о результатах исследования, то ее лучше делать как комбинированную, потому что в составленной таблице часть результатов остается как бы скрытой. При помощи комбинированной таблицы можно более глубоко проанализировать результаты исследования.

Для записи данных наблюдений, их возведения и обработки в качестве так называемых **рабочих таблиц** лучше использовать простые, групповые и составлены таблицы.

5. Шахматная таблица (таблица 1.3.5) является особым видом комбинированной. С помощью такой таблицы становится возможным удобно показать связи между изучаемыми признаками (например, латинский квадрат, корреляционные таблицы и проч.).

Таблица 1.3.5 - Распределение учащихся экспериментальных школ по полу и возрасту (образец шахматной таблицы)

Пол учащихся	Возраст учащихся (лет)			Всего
	8	от 8 до 10	от 10 до 12	
Мальчики	81	74	109	264
Девочки	103	85	93	281
Вместе	184	159	202	545

Как видно из приведенного образца, шахматная таблица дает ответ и на то, сколько в каждой возрастной группе имеется мальчиков и девочек, в то время как из вышеприведенной комбинированной таблицы этого без дополнительных вычислений узнать нельзя.

6. Перекрестная таблица. В шахматной таблице может, конечно, быть столько же столбцов, сколько и строк.

Если, например, исследователя интересует вопрос, сколько выпускников школ из того или иного региона вступило в того или иного учебного заведения и сколько из них после окончания обучения вернулись в свои регионы (при условии, что все поступающие в учебное заведение закончили его), можно составить следующую перекрестную таблицу (таблица 1.3.6).

Таблица 1.3.6 - Сводные сведения о поступающих и распределение выпускников педуниверситета (образец перекрестной таблицы)

Сведения о поступающих в педуниверситет от районов	Районы области, получили выпускников педагогического университета				
	А	Б	В	Другие районы	Всего
А	20	10	15	40	85
Б	31	6	12	15	64
В	15	2	1	5	23
Другие районы	40	6	7	10	63
Вместе	106	24	35	70	235

При суммировании строк таблицы выясняется, сколько отдельные районы дали учеников, а при суммировании столбцов выясняется, сколько тот или иной район получил выпускников.

Из таблицы 1.3.6 видно, что больше всего учеников дал район А, он же получил и больше всего выпускников. Район Б дал 64 учащихся, а выпускников получил почти в три раза меньше. Район дал В 23 учеников, а выпускников получил больше.

В другие районы пошли ученики, выясняется по строке каждого района. Из других районов получили учеников, показывает столбец каждого района. Сколько тот или иной район получил обратно своих учеников, выясняется из граф, которые находятся по диагонали с левого верхнего угла таблицы в правый нижний угол.

Такая таблица позволяет выяснить целый ряд интересных проблем по распределению учащихся и выпускников между отдельными районами.

Кроме рассмотренных часто используют упрощенные таблицы-выводы. Вывод, как правило, содержит боковик, точки и одну или две графы (столбцы)

Вывод дают без заголовка, если он является непосредственным продолжением материала и грамматически связан с вводной фразой текста; с заголовком, если вывод имеет самостоятельное значение. Приведем пример, когда вывод следует давать без заголовка, поскольку он вытекает из текста, предшествующем самому выводу:

По результатам проведенного формирующего этапа эксперимента установлены уровни сформированности мировоззрения учащихся (в процентах):

высокий	38,35
средний	52,11
низкий	9,4

Все приведенные в таблицах данные должны быть достоверными, однородными и могут быть сопоставлены, а в основу их группировки положены лишь существенные признаки.

Приводить в научном тексте следует только те таблицы, которые трудно передать обычным текстом (результаты экспериментальных наблюдений, сопоставление разногласия, подробные справочные данные и др.).

Таблицы являются важной составной частью исследований. Они позволяют представлять числовые данные компактно, систематизировано и наглядно. Приведенные в таблице данные удобно сравнивать и анализировать. Но в педагогических исследованиях нельзя давать слишком много таблиц, передавать с помощью их весь смысл исследований.

Литература:

1. Ермолаев, О. Ю. Математическая статистика для психологов : Учебник // 2-е изд., испр. / О. Ю. Ермолаев. – Москва : Московский психолого-социальный институт : Флинта. – 336 с. – [Электронный ресурс]. – Режим доступа: <https://e-catalog.nlb.by / Author/Home?author.> – Дата доступа: 15.10.2020

2. Пашкевич, О. И. Курсовая и дипломная работы: структура, оформление, защита : методические рекомендации для слушателей переподготовки по специальности 1-03 04 72 "Практическая психология" (квалификация "Педагог-психолог") / О. И. Пашкевич, О. М. Савчик. – Минск : РИПО. – [Электронный ресурс]. – Режим доступа: [http://unicat.nlb.by/opac/pls/!search.http_keyword?query=a001a="BY-SEK-277528"&lst_siz=20.](http://unicat.nlb.by/opac/pls/!search.http_keyword?query=a001a=) – Дата доступа: 15.10.2020

3. Пашкевич, О. И. Статистическая обработка эмпирических данных в системе STATISTICA : учебно-методическое пособие / О. И. Пашкевич – Электрон. текстовые данные. – Минск : РИПО. – 148 с. – [Электронный ресурс]. – Режим доступа: <http://www.iprbookshop.ru/67607.html>. – Дата доступа: 15.10.2020

4. Яковлев, Е. В. Педагогическое исследование: содержание и представление результатов / Е. В. Яковлев, Н. О. Яковлева. – Челябинск : Изд-во РБИУ. – 317 с. – [Электронный ресурс]. – Режим доступа: <http://elib.cspu.ru/xmlui/bitstream/handle/123456789/278/Пед.%20исследование%202010.pdf?sequence=1&isAllowed=y>. – Дата доступа: 15.10.2020

1.4 Определение различий между эмпирической и теоретической таблицами распределения

- **Эмпирическая и теоретическая таблицы распределения**
- **ХИ-квадрат-критерий (χ^2) и определение статистической значимости различия между эмпирической и теоретической таблицами распределения**
- **Таблицы распределения размерности (2x2), (2x3)**

- **Эмпирическая и теоретическая таблицы распределения**

Статистические таблицы подразделяются на таблицы первичных эмпирических данных и таблицы распределения. Задача таблиц первичных эмпирических данных (ПЭД) состоит в том, чтобы все данные, полученные в ходе исследования, представить унифицировано, в одной таблице. На основе таблиц ПЭД строятся так называемые групповые таблицы или ряды распределений, в которых упорядочение осуществляется по одному признаку, и комбинаторные, или таблицы распределения.

Таблица распределения – n-мерная таблица, где n – количество признаков, по которым проводится анализ.

Эмпирическая таблица распределения – это такая таблица распределения, которая строится на основе таблицы первичных эмпирических данных.

Теоретическая таблица распределения – это таблица распределения, построенная при условии, что распределение испытуемых носит случайный характер.

- **ХИ-квадрат-критерий (χ^2) и определение статистической значимости различия между эмпирической и теоретической таблицами распределения**

Различие между эмпирической и теоретической таблицами распределения определяется с помощью ХИ-квадрат-критерия (χ^2) по формуле:

$$\chi^2 = \sum_{i,j=1}^{n,m} \frac{(F_{эij} - F_{тj})^2}{F_{тj}}$$

где i – порядковый номер строки;

j – порядковый номер столбца;

n – число строк;

m – число столбцов таблицы (матрицы) распределения;

$F_{\text{Э}}$ и $F_{\text{Т}}$ – эмпирическая и теоретическая частоты наблюдаемого признака.

Показатель χ^2 , квадратом называется условно: он определяет, насколько эмпирическое распределение отклоняется от теоретического. Так как теоретическое распределение исходит из независимости признаков, то чем больше эмпирическое распределение отклоняется от теоретического (случайного), тем больше значение χ^2 ; и наоборот, чем больше χ^2 , тем больше эмпирическое распределение отклоняется от теоретического. Поэтому показатель χ^2 является оценкой неслучайности распределения. Ниже будут рассмотрены таблицы размерности (1x2), (2x2), (2x3) и проведены расчеты Хи-квадрат критерия с целью установления статистической значимости различия между эмпирической и теоретической таблицами распределения. Следует отметить, что Хи-квадрат критерий работает на номинальных шкалах, и в качестве значений изучаемых признаков рассматриваются частоты распределения.

Распределение (1x2). Таблица распределения размерности один на два записывается как (1x2) и представляется в виде:

Таблица 1.4.1

	j	
i	a	\bar{a}

где число строк $i=1$, а число столбцов $j=2$.

Для расчета $\chi^2_{\text{Э}}$ предварительно проверяется требование Юла и Кендалла, относящееся к теоретическим таблицам распределения.

Требование Юла и Кендалла: каждая теоретическая частота должна быть не меньше 5.

Для распределения (1x2) способ определения теоретических частот не задан, поэтому требование Юла и Кендалла соответствует неравенству:

$$\frac{a + \bar{a}}{2} \geq 5$$

Формула расчета $\chi^2_{\text{Э}}$ для распределения (1x2) является частным случаем формулы (1) и записывается в виде:

$$\chi^2_{\text{Э}}(1x2) = \frac{(a - \bar{a})^2}{2(a + \bar{a})},$$

где a и \bar{a} – эмпирические частоты распределения, выраженные числом.

Статистическая значимость различия между a и \bar{a} устанавливается путем сравнения $\chi^2_{\text{Э}}$ с меньшим значением χ^2_{T} и установления уровня значимости p по таблице:

Таблица 1.4.2 – Значения χ^2_{T} и уровня значимости p

χ^2_{T}	2,71	3,84	6,64	10,83
p	0,10	0,05	0,01	0,001

Статистическое различие считается достоверным, если уровень значимости $p < 0,05$; если $p = 0,1$, то говорят о **наличии тенденции к статистической значимости**, и если $p > 0,05$, то **статистическое различие (статистическая значимость) не имеет места**.

• **Таблицы распределения размерности (2x2), (2x3)**

Таблица распределения размерности (2x2). В отличие от распределения (1x2), где рассматривался только один признак, в распределении (2x2) их два. Пусть в качестве данных признаков используются A и B . Они принадлежат одному элементу измерения и выступают в виде комбинаций: $A + B, A + \bar{B}, \bar{A} + B, \bar{A} + \bar{B}$, т.е. каждый элемент измерения определяется одной из приведенных комбинаций. Каждая комбинация – число элементов измерения (эмпирическая частота встречаемого свойства) – должна отвечать условиям:

$$a + b \approx c + d \quad \text{и} \quad a + c \approx b + d.$$

Если данные условия не выполняются, необходимо изменить значения признаков A и B .

Таблица распределения (2x2) имеет вид:

Таблица 1.4.3 – Эмпирические частоты распределения

	A	\bar{A}
B	a	b
\bar{B}	c	d

Сумма всех эмпирических частот $a, b, c,$ и d должна быть не меньше 30.

Для расчета $\chi^2_{\text{Э}}$ предварительно проверяется требование Юла и Кендалла, относящееся к теоретическим таблицам распределения.

Требование Юла и Кендалла: каждая теоретическая частота в клетках таблицы должна быть не меньше 5. Тогда требование Юла и Кендалла для эмпирической частоты a запишется как:

$$a' = \frac{(a+b)(a+c)}{N} \geq 5,$$

для частоты b :

$$b' = \frac{(a+b)(b+d)}{N} \geq 5,$$

для частоты c :

$$c' = \frac{(a+c)(c+d)}{N} \geq 5,$$

для частоты d :

$$d' = \frac{(b+d)(c+d)}{N} \geq 5,$$

где $N = a + b + c + d$ и a', b', c', d' – теоретические частоты распределения.

Если требование Юла и Кендалла выполняется для каждой теоретической частоты, т.е. $a', b', c', d' \geq 5$, то можно переходить к построению теоретической таблицы распределения и расчету $\chi^2_{(2 \times 2)}$.

Таблица распределения теоретических частот для распределения (2x2) имеет вид:

Таблица 1.4.4 – Теретические частоты распределения

	A'	$\overline{A'}$
B'	a'	b'
$\overline{B'}$	c'	d'

Формула для вычисления $\chi^2_{(2 \times 2)}$:

$$\chi^2_{(2 \times 2)} = \frac{(ad - bc)^2 (a + b + c + d)}{(a + b)(c + d)(a + c)(b + d)}.$$

Статистическая значимость различия устанавливается путем сравнения $\chi^2_{\text{э}}$ с меньшим значением $\chi^2_{\text{т}}$ и установления уровня значимости p по таблице 1.4.5.

Таблица 1.4.5 - Значения $\chi^2_{\text{т}}$ и уровня значимости p

$\chi^2_{\text{т}}$	2,71	3,84	6,64	10,83
p	0,10	0,05	0,01	0,001

Статистическое различие считается достоверным, если уровень значимости $p < 0,05$; если $p = 0,1$, то говорят о **наличии тенденции к**

статистической значимости, и если $p > 0,05$, то *статистическое различие* (статистическая значимость) *не имеет места*.

На основе χ^2 -критерия для распределения (2x2) устанавливается связь между рассматриваемыми признаками A и B , которая относится к столбцам и строкам матрицы. Величина связи определяется с помощью коэффициента сопряженности K (Чупрова) по формуле:

$$K = \sqrt{\frac{\chi^2}{N}}$$

где $N = a + b + c + d$.

Если $0.3 \leq K < 0.5$, то *связь* между изучаемыми признаками *слабая*, если $0.5 \leq K < 0.7$, то *связь умеренная*, и если $K \geq 0.7$, то *связь сильная*.

Примечание. Коэффициент сопряженности K (Чупрова) изменяется в пределах от 0 до 1 и содержит в себе информацию только о величине связи, а не о ее направлении.

Пример вычисления χ^2 -критерия для таблиц распределения размерности (2x2). Пусть признак A – конфликтность, тогда A_1 – низкий уровень конфликтности и A_2 – высокий уровень конфликтности. Пусть признак B – пол, тогда B_1 – женский пол и B_2 – мужской пол. Необходимо установить существует ли статистически значимая связь между признаками A и B . Воспользуемся расчетом χ^2 -критерия.

Предположим, что по результатам исследования для комбинаций признаков (A_1+B_1) имеем 30 испытуемых, т.е. эмпирическая частота $a=30$ (количество женщин с низким уровнем конфликтности); $(A_2+B_1) = 10$, т.е. эмпирическая частота $b=10$ (количество женщин с высоким уровнем конфликтности), $(A_1+B_2) = 15$, т.е. эмпирическая частота $c=15$ (количество мужчин с низким уровнем конфликтности), $(A_2+B_2) = 25$, т.е. эмпирическая частота $d=25$ (количество мужчин с высоким уровнем конфликтности). Учитывая полученные значения эмпирических частот, строим эмпирическую таблицу распределения:

	A_1	A_2
B_1	30	10
B_2	15	25

Проверим выполнимость требования Юла и Кендалла, рассчитав каждую теоретическую частоту, и осуществим переход к теоретической таблице распределения. Получим:

$$a' = \frac{(a+b)(a+c)}{N} = \frac{(30+10)(30+15)}{80} = 22.5 > 5,$$

$$b' = \frac{(a+b)(b+d)}{N} = \frac{(30+10)(10+25)}{80} = 17.5 > 5,$$

$$c' = \frac{(a+c)(c+d)}{N} = \frac{(30+15)(15+25)}{80} = 22.5 > 5,$$

$$d' = \frac{(b+d)(c+d)}{N} = \frac{(10+25)(15+25)}{80} = 17.5 > 5.$$

Итак, требование Юла и Кендалла выполняется для каждой теоретической частоты, и таблица распределения теоретических частот будет иметь вид:

	A_1	A_2
B_1	22,5	17,5
B_2	22,5	17,5

Далее переходим к расчету. Получаем: $\chi_9^2 \approx 11,43$.

Делаем вывод о статистической значимости полученного результата. Для этого сравниваем рассчитанное значение $\chi_9^2 \approx 11,43$ с меньшим значением $\chi_T^2 = 10,83$ и устанавливаем уровень значимости $p=0,001$ (таблица 1.4.3.3). Так как $\chi_9^2 \approx 11,43 > \chi_T^2 = 10,83$, то результат является статистически значимым с уровнем значимости $p=0,001$.

Установим силу связи между изучаемыми признаками, рассчитав коэффициент сопряженности K (Чупрова) по формуле:

$$K = \sqrt{\frac{\chi_9^2}{N}} = \sqrt{\frac{11,43}{80}} \approx 0,3$$

Так как значение $K \approx 0,3$ принадлежит промежутку $[0,3; 0,5)$, то сила связи слабая.

Итак, учитывая полученные результаты, можно заключить, что изучаемые признаки (конфликтность и пол) находятся в статистически значимой зависимости ($\chi_9^2 \approx 11,43$ и $p=0,001$). Это также подтверждается значением коэффициента сопряженности K (Чупрова), указывающем на силу связи между признаками.

Таблица распределения размерности (2x3). В отличие от распределений (1x2) и (2x2), где признаки представлены в номинальной шкале, в распределении (2x3) один из признаков представлен в порядковой шкале, где он принимает значения, имеющие смысл трех уровней (низкий, средний, высокий) или трех рангов (первый, второй, третий).

Таблица для распределения (2x3) имеет вид:

Таблица 1.4.6 – Эмпирические частоты распределения

	A_1	A_2	A_3
B	a	b	c
\bar{B}	d	e	f

Необходимо обратить внимание на то, чтобы эмпирические частоты a, b, c, d, e, f удовлетворяли следующим условиям:

$$(a + d) \approx (b + e) \approx (c + f),$$

$$(a + b + c) \approx (d + e + f)$$

$$(a + b + c + d + e + f) \geq 50.$$

Для расчета χ^2_3 предварительно проверяется требование Юла и Кендалла, относящееся к теоретическим таблицам распределения.

Требование Юла и Кендалла: каждая теоретическая частота в клетках таблицы должна быть не меньше 5 (т.е. $a', b', c', d', e', f' \geq 5$). Тогда a', b', c', d', e', f' – теоретические частоты, определяемые по формулам:

$$a' = (a + d)(a + b + c) / N, \quad b' = (b + e)(a + b + c) / N,$$

$$c' = (c + f)(a + b + c) / N,$$

$$d' = (a + d)(d + e + f) / N, \quad e' = (b + e)(d + e + f) / N,$$

$$f' = (c + f)(d + e + f) / N,$$

где, $N = a + b + c + d + e + f$.

Формула расчета χ^2_3 для распределения (2x3):

$$\chi^2_3 = \frac{(a - a')}{a'} + \frac{(b - b')}{b'} + \frac{(c - c')}{c'} + \frac{(d - d')}{d'} + \frac{(e - e')}{e'} + \frac{(f - f')}{f'}$$

где a', b', c', d', e', f' – теоретические частоты.

Уровень значимости p определяется с помощью таблицы для χ^2_T :

Таблица 1.4.7 - Значения χ^2_T и уровня значимости p

χ^2_T	4,60	5,99	9,21	13,82
p	0,10	0,05	0,01	0,001

Для рассчитанного значения $\chi^2_{\text{э}}$ в таблице находят *меньшее* $\chi^2_{\text{т}}$, устанавливают уровень значимости p и делают вывод о статистической значимости. **Статистическое различие считается достоверным**, если уровень значимости $p < 0,05$; если $p = 0,1$, то говорят о **наличии тенденции к статистической значимости**, и если $p > 0,05$, то **статистическое различие** (статистическая значимость) **не имеет места**.

На основе *XII*-квадрат-критерия рассчитывается **величина связи** между рассматриваемыми признаками. Показатель величины связи – коэффициент сопряженности K (Чупрова), определяемый для распределения (2x3) по формуле:

$$K = 0.84 \sqrt{\frac{\chi^2_{\text{э}}}{N}},$$

где **0.84** – нормирующий коэффициент, позволяющий получить более точные значения связи при малых выборках,

$$N = a + b + c + d + e + f.$$

Если $0.3 \leq K < 0.5$, то между изучаемыми признаками существует **слабая связь**, если $0.5 \leq K < 0.7$, то – **связь умеренная**, и если $K \geq 0.7$, то – **связь сильная**.

Литература:

1. Ермолаев, О. Ю. Математическая статистика для психологов: Учебник // 2-е изд., испр. / О. Ю. Ермолаев. – Москва : Московский психолого-социальный институт : Флинта. – 336 с. – [Электронный ресурс]. – Режим доступа: <https://e-catalog.nlb.by / Author/Home?author>. – Дата доступа: 15.10.2020
2. Килин, П. М. Статистические методы обработки данных : учебное пособие / П. М. Килин, Н. И. Чекмарева. – Тюмень : ТюмГНГУ. – 128 с. – [Электронный ресурс]. – Режим доступа: http://elib.tyuiu.ru/wp-content/uploads/2014/02/Статистические_методы_обработки...110_A5.pdf – Дата доступа: 15.10.2020
3. Пашкевич, О. И. Статистическая обработка эмпирических данных в системе STATISTICA : учебно-методическое пособие / О. И. Пашкевич – Электрон. текстовые данные. – Минск : РИПО. – 148 с. – [Электронный ресурс]. – Режим доступа: <http://www.iprbookshop.ru/67607.html>. – Дата доступа: 15.10.2020

1.5 Описательные статистики

- **Описательные статистики. Показатели уровня ряда**
- **Меры рассеивания**
- **Среднеквадратическое (стандартное) отклонение**
- **Доверительный интервал. Достоверность средней арифметической**

- **Описательные статистики. Показатели уровня ряда**

Информации, содержащейся в таблицах первичных эмпирических данных, бывает недостаточно для глубокого анализа материала, особенно если приходится сравнивать несколько таблиц. Поэтому дополнительно вводятся некоторые обобщающие характеристики, которые вскрывают особенности изучаемых явлений: уровень ряда, т.е. среднее значение признака, вокруг которого варьируют остальные его значения, и амплитуда колебания фиксированных значений вокруг среднего значения. Показателями уровня ряда являются средние величины:

➤ **среднее арифметическое**, вычисляемое по формуле:

$$\bar{x} = \frac{\sum_{i=1}^n x_i P_i}{n},$$

где x_i – i -значение переменной x , P_i – частота, с которой встречается данное значение переменной x , n – число измерений (количество испытуемых).

Среднее арифметическое — очень информативная мера "центрального положения" наблюдаемой переменной, особенно если сообщается ее доверительный интервал. Исследователю нужны такие статистики, которые позволяют сделать вывод относительно популяции в целом. Одной из таких статистик и является среднее;

➤ **медиана** – значение признака, которое находится в середине числового ряда.

Медиана (Median) разбивает выборку на две равные части. Половина значений переменной лежит ниже медианы, половина – выше. Медиана дает общее представление о том, где сосредоточены значения переменной, иными словами, где находится ее центр. В некоторых случаях, например, при описании доходов населения, медиана более удобна, чем среднее.

Медиана делит ряд упорядоченных значений пополам с равным числом этих значений как выше, так и ниже ее (левее и правее медианы на числовой оси).

Вычислить медиану легко, если число наблюдений n нечетное. Это будет наблюдение номер $(n + 1)/2$ в нашем упорядоченном наборе данных.

Например, если $n = 11$, то медиана – это $(11 + 1)/2$, т. е. 6-е наблюдение в упорядоченном наборе данных.

Если n четное, то, строго говоря, медианы нет. Однако обычно мы вычисляем ее как среднее арифметическое двух соседних средних наблюдений в упорядоченном наборе данных (т. е. наблюдений номер $(n/2)$ и $(n/2 + 1)$).

Так, например, если $n = 20$, то медиана – это среднее арифметическое наблюдений номер $20/2 = 10$ и $(20/2 + 1) = 11$ в упорядоченном наборе данных;

➤ **мода** – наиболее часто встречающееся в ряду значение признака.

Мода (Mode) представляет собой максимально часто встречающееся значение переменной (иными словами, наиболее «модное» значение переменной), например, популярная передача на телевидении, модный цвет платья или марка автомобиля и т. д. Мода хорошо описывает, например, типичную реакцию водителей на сигнал светофора о прекращении движения. Классический пример использования моды – выбор размера выпускаемой партии обуви или цвета обоев. Если распределение имеет несколько мод, то говорят, что оно мультимодально или многомодально (имеет два или более «пика»). Мультимодальность распределения дает важную информацию о природе исследуемой переменной. Например, в социологических опросах, если переменная представляет собой предпочтение или отношение к чему-либо, то мультимодальность может означать, что существует несколько определенно различных мнений. Мультимодальность также служит индикатором того, что выборка не является однородной и наблюдения, возможно, порождены двумя или более «наложенными» распределениями.

Основным условием для расчета средней арифметической является требование того, чтобы она отражала положение распределения, которое в свою очередь должно быть *нормальным* (или *гауссовым*). Кривой нормального распределения называется плавная колоколообразная симметричная кривая (рисунок 1.5.1).

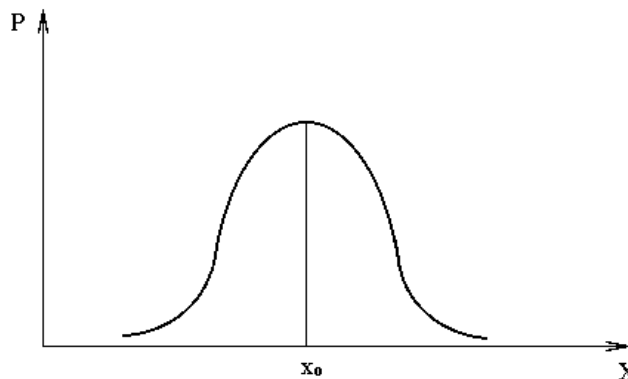


Рисунок 1.5.1 - Кривая нормального распределения

В качестве начальной точки x_0 при построении кривой нормального распределения берется значение средней арифметической \bar{x} . По оси Ox откладывают вправо от x_0 значения x , превышающие величину x_0 и доходящие до максимума, влево – меньшие x_0 и доходящие до минимума.

Отрезок перпендикуляра, восстановленного из x_0 , соответствует

вершине кривой. Отрезки перпендикуляров, восстановленных из других значений x , соответствуют частотам P этих значений. После соединения концов отрезков перпендикуляров линией получают кривую нормального распределения.

• **Меры рассеивания**

Для каждой случайной величины существует свой индивидуальный закон распределения, определяемый:

1) *положением на числовой оси* (на рисунке 1.5.2 кривые 1 и 2 отличаются только положением), где по оси абсцисс откладываются значения переменной x , а по оси ординат – плотность распределения P :

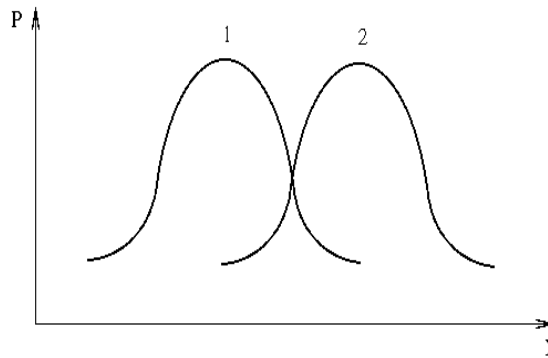


Рисунок 1.5.2 - Положение на числовой оси

2) *рассеиванием значений, или дисперсией D* (на рис.3.3 для кривой 2 дисперсия значений больше, чем для кривой 1, характеризующей степень разброса данных относительно своего центра и определяемой по формуле:

$$D = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n},$$

где x_i – i -значение переменной x ;
 \bar{x} – среднее арифметическое;
 n – число измерений;

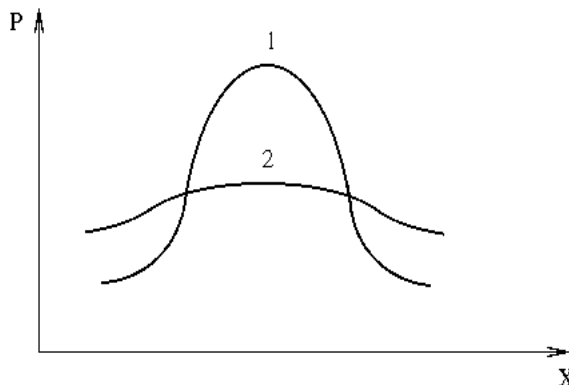


Рисунок 1.5.3 – Дисперсии значений выборочных распределений

Дисперсия (Variance) и стандартное отклонение (Standard deviation) – наиболее часто используемые меры изменчивости переменной. Дисперсия меняется от нуля до бесконечности. Крайнее значение 0 означает отсутствие изменчивости, когда значения переменной постоянны.

Один из способов измерения рассеяния данных заключается в том, чтобы определить степень отклонения каждого наблюдения от средней арифметической. Очевидно, что чем больше отклонение, тем больше изменчивость, вариабельность наблюдений.

Однако мы не можем использовать среднее этих отклонений как меру рассеяния, потому что положительные отклонения компенсируют отрицательные отклонения (их сумма равна нулю). Чтобы решить эту проблему, мы возводим в квадрат каждое отклонение и находим среднее возведенных в квадрат отклонений; эта величина называется вариацией, или дисперсией.

- **Среднеквадратическое (стандартное) отклонение**

Среднее квадратичное (стандартное) отклонение σ есть корень квадратный из среднего квадрата отклонений отдельных значений признака от среднего арифметического:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}},$$

где x_i — i -значение переменной x ;
 \bar{x} — среднее арифметическое;
 n — число измерений (наблюдений).

Величина σ^2 называется дисперсией. Среднее квадратичное отклонение для измерения изменчивости альтернативных признаков определяется по формуле:

$$\sigma = \frac{\sqrt{p_1 p_2}}{n},$$

где p_1 — частота первого взаимоисключающего признака,
 p_2 — частота второго взаимоисключающего признака.

Мы можем представить себе стандартное отклонение как своего рода среднее отклонение наблюдений от среднего. Оно вычисляется в тех же единицах (размерностях), что и исходные данные.

Если разделить стандартное отклонение на среднее арифметическое и выразить результат в процентах, получится *коэффициент вариации*.

Он является мерой рассеяния, не зависит от единиц измерения (безразмерный), но имеет некоторые теоретические неудобства и поэтому не очень одобряется статистиками;

3) *асимметрией, или косо́стью, скошенностью* (на рисунке 1.5.4 кривые 1, 2, 3 отличаются только положением и асимметрией: для кривой 1 имеет место левосторонняя асимметрия, для кривой 2 – нулевая, и для кривой 3 – правосторонняя);

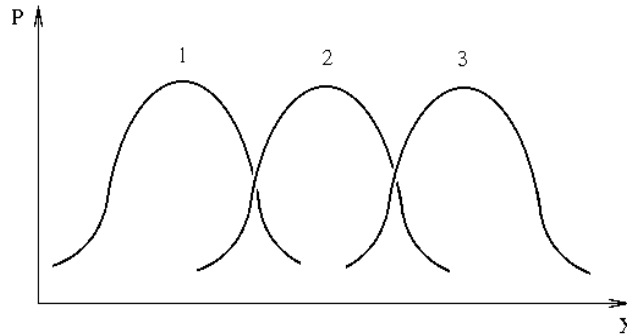


Рисунок 1.5.4 – Асимметрия

Асимметрия (*Skewness*), или коэффициент асимметрии, является мерой несимметричности распределения. Если этот коэффициент значительно отличается от 0, распределение является асимметричным, т.е. несимметричным. Формально имеем:

$$A = \frac{\sum (x - \bar{x})^{-3}}{n \sigma^3},$$

где \bar{x} – среднее, σ – стандартное отклонение (сигма), n – число наблюдений;

4) *эксцессом* (с лат. – выход, отступление, уклонение), или выпуклостью, «кучностью», показывающим, насколько кучно основная масса данных группируется около центра (на рис. 3. 5 кривые 1 и 2 отличаются только по дисперсии и эксцессу: для кривой 1 он больше).

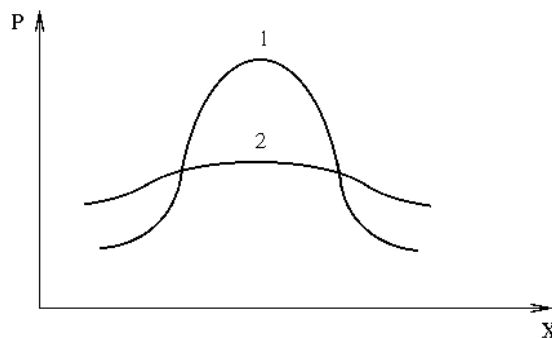


Рисунок 1.5.5 – Эксцесс

Иногда значения моды и медианы, асимметрии и эксцесса используются для проверки гипотезы о том, что наблюдаемые данные (выборка) принадлежат заданному семейству распределений, например, нормальному. Так, для любого нормального распределения мода и медиана совпадают со средним арифметическим \bar{x} , асимметрия равна нулю, а эксцесс — трем.

Экссесс (*Kurtosis*), или коэффициент эксцесса измеряет остроту пика распределения. Оценка эксцесса вычисляется по формуле:

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

Коэффициент эксцесса равен нулю, если наблюдения подчиняются нормальному распределению. Если он значительно отличается от нуля, гипотезу о том, что данные взяты из нормального распределенной генеральной совокупности, следует отвергнуть.

Стандартные ошибки асимметрии и эксцесса (*Standard error of skewness, Standard error of kurtosis*) – это и есть стандартные ошибки асимметрии и эксцесса, аналогичные стандартной ошибке среднего.

- **Доверительный интервал. Достоверность средней арифметической**
Доверительный интервал (95% confidence limits of mean) для среднего представляет интервал значений вокруг оценки, где с данным уровнем доверия, находится "истинное" (неизвестное) среднее популяции. Заметим, что ширина доверительного интервала зависит от объема или размера выборки, а также от разброса (изменчивости) данных. Увеличение размера выборки делает оценку среднего более надежной. Увеличение разброса наблюдаемых значений уменьшает надежность оценки. Вычисление доверительных интервалов основывается на предположении нормальности наблюдаемых величин. Если это предположение не выполнено, то оценка может оказаться плохой, особенно для малых выборок. При увеличении объема выборки, скажем, до 100 или более, качество оценки улучшается и без предположения нормальности выборки.

Например, если среднее выборки равно 23, а нижняя и верхняя границы доверительного интервала с уровнем $p=.95$ равны 19 и 27 соответственно, то можно заключить, что с вероятностью 95% интервал с границами 19 и 27 накрывает среднее популяции.

Если вы установите больший уровень доверия, то интервал станет шире, поэтому возрастает вероятность, с которой он "накрывает" неизвестное среднее популяции, и наоборот.

Хорошо известно, например, что чем "неопределенней" прогноз погоды (т.е. шире доверительный интервал), тем вероятнее он будет верным. Заметим, что ширина доверительного интервала зависит от объема или размера выборки, а также от разброса (изменчивости) данных. Увеличение размера выборки делает оценку среднего более надежной. Увеличение разброса наблюдаемых значений уменьшает надежность оценки.

Вычисление доверительных интервалов основывается на предположении нормальности наблюдаемых величин. Если это предположение не выполнено, то оценка может оказаться плохой, особенно для малых выборок.

При увеличении объема выборки, скажем, до 100 или более, качество оценки улучшается и без предположения нормальности выборки.

Довольно трудно «ощутить» числовые измерения, пока данные не будут содержательно обобщены. Диаграмма часто полезна в качестве отправной точки. Мы можем также сжать информацию, используя важные характеристики данных. В частности, если бы мы знали, из чего состоит представленная величина, или если бы мы знали, насколько широко рассеяны наблюдения, то мы бы смогли сформировать образ этих данных.

Стандартная ошибка среднего значения (*Standard error of mean*) – это стандартное отклонение, деленное на квадратный корень из объема выборки. В интервале шириной, равной удвоенной стандартной ошибке, отложенному вокруг среднего значения, располагается среднее значение генеральной совокупности с вероятностью примерно 67%. Стандартная ошибка, как и стандартное отклонение, может использоваться в качестве меры разброса переменной. По так называемому правилу кулака, в одном диапазоне стандартного отклонения (охватывающем ширину стандартного отклонения в обе стороны от среднего значения) располагается примерно 67% значений, в диапазоне удвоенного стандартного отклонения – примерно 95%, а в диапазоне утроенного стандартного отклонения – примерно 99% значений. С другой стороны, стандартная ошибка позволяет задать доверительный интервал для среднего значения. В диапазоне удвоенной стандартной ошибки по обе стороны от среднего значения с вероятностью примерно 95% находится среднее значение генеральной совокупности. С вероятностью примерно 99% оно лежит в диапазоне утроенной стандартной ошибки. Часто указывают только одну из мер разброса, обычно – стандартную ошибку, так как ее значение меньше.

Арифметическая средняя \bar{x} достоверна с вероятностью ошибки $p=5\%$, если выполняется условие:

$$\bar{x}_{\min} \leq \bar{x} \leq \bar{x}_{\max}.$$

Минимальная допустимая граница:

$$\bar{x}_{\min} = \bar{x} - t_{0.05} \sqrt{\frac{D}{n}}$$

Максимальная допустимая граница:

$$\bar{x}_{\max} = \bar{x} + t_{0.05} \sqrt{\frac{D}{n}},$$

где \bar{x} – среднее арифметическое;

D – дисперсия;

n – число измерений;

$t_{0.05}$ – коэффициент, определяемый по таблице Фишера 1.5.1:

Таблица 1.5.1 – Значения коэффициента Фишера $t_{0.05}$

12.706(2)	4.303(3)	3.182(4)	2.776(5)	2.571(6)	2.447(7)
2.365(8)	2.306(9)	2.262(10)	2.228(11)	2.201(12)	2.179(13)
2.160(14)	2.145(15)	2.131(16)	2.120(17)	2.110(18)	2.101(19)
2.093(20)	2.086(21)	2.080(22)	2.074(23)	2.069(24)	2.064(25)
2.060(26)	2.056(27)	2.052(28)	2.048(29)	2.045(30)	2.042(31)

Квартили представляют собой значения, которые делят две половины выборки (разбитые медианой) еще раз пополам. Таким образом, медиана и квартили делят диапазон значений переменной на четыре равные части. Различают верхний квартиль, который больше медианы и делит пополам верхнюю часть выборки (значения переменной больше медианы), и нижний квартиль, который меньше медианы и делит пополам нижнюю часть выборки (*Lower and Upper quartiles*). Нижний квартиль часто обозначают символом 25%, это означает, что 25% значений переменной меньше нижнего квартиля. Верхний квартиль часто обозначают символом 75%, это означает, что 75% значений переменной меньше верхнего квартиля.

Размах (*Range*) – разница между наибольшим и наименьшим значением переменной

Квартильный (внутриквартильный) размах (*Quartile Range*) – равен разности значений верхнего и нижнего квартиля. Таким образом, это интервал, содержащий медиану, в который попадает 50% наблюдений.

Примечание. В программе STATISTICA (*Statistics Basic Statistics/Tables Descriptive Statistics*) можно посчитать:

- Количество значений (*Valid N*)
- Сумма всех значений (*Sum*)
- Минимум и максимум (*Minimum & Maximum*) – это минимальное и максимальное значения переменной.
- Среднее (*Mean*) – сумма значений переменной, деленная на N (число значений переменной) и еще больше описательных статистик.

Литература:

1. Вуколов, Э. А. Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL : уч. пос. / Э. А. Вуколов – 2 изд., испр. и доп. – Москва : Форум: НИЦ Инфра-М. – 464 с. – [Электронный ресурс]. – Режим доступа: <https://uch-lit.ru/matematika-2/dlya-studentov/vukolov-e-l-osnovyi-statisticheskogo-analiza-praktikum-po-statisticheskim-metodam-i-issledovaniyu-operatsii-s-ispolzovaniem-paketov-statistica-i-excel-onlayn>. – Дата доступа: 15.10.2020

2. Боровиков, В. В. STATISTICA. Искусство анализа данных на компьютере: Для профессионалов. 2-е изд. (+CD) / В. В. Боровиков. – Санкт-Петербург : Питер. – 688 с. – [Электронный ресурс]. – Режим доступа:

<https://www.books.ru/books/statistica-iskusstvo-analiza-dannykh-na-kompyutere-2-e-izdanie-cd-80566>. – Дата доступа: 15.10.2020

3. Пашкевич, О. И. Статистическая обработка эмпирических данных в системе STATISTICA : учебно-методическое пособие / О. И. Пашкевич – Электрон. текстовые данные. – Минск : РИПО. – 148 с. – [Электронный ресурс]. – Режим доступа: <http://www.iprbookshop.ru/67607.html>. – Дата доступа: 15.10.2020

4. Ермолаев, О. Ю. Математическая статистика для психологов : Учебник // 2-е изд., испр. / О. Ю. Ермолаев. – Москва : Московский психолого-социальный институт : Флинта. – 336 с. – [Электронный ресурс]. – Режим доступа: <https://e-catalog.nlb.by/Author/Home?author>. – Дата доступа: 15.10.2020

5. Килин, П. М. Статистические методы обработки данных : учебное пособие / П. М. Килин, Н. И. Чекмарева. – Тюмень : ТюмГНГУ. – 128 с. – [Электронный ресурс]. – Режим доступа: http://elib.tyuiu.ru/wp-content/uploads/2014/02/Статистические_методы_обработки...110_A5.pdf – Дата доступа: 15.10.2020

6. Пашкевич, О. И. Статистическая обработка эмпирических данных в системе STATISTICA : учебно-методическое пособие / О. И. Пашкевич – Электрон. текстовые данные. – Минск : РИПО. – 148 с. – [Электронный ресурс]. – Режим доступа: <http://www.iprbookshop.ru/67607.html>. – Дата доступа: 15.10.2020

7. Тюрин, Ю. Н. Анализ данных на компьютере : учебное пособие / Ю. Н. Тюрин, А. А. Макаров. – Москва : МЦНМО. – 367 с. – [Электронный ресурс]. – Режим доступа: <http://маткнига.рф/wp-content/uploads/2017/02/978-5-4439-3011-4-Tyurin-Makarov-Analiz-dannyh.pdf>. – Дата доступа: 15.10.2020

8. StatSoft. Электронный учебник по статистике. – [Электронный ресурс]. – Режим доступа: <http://statsoft.ru/home/textbook/default.htm>. – Дата доступа: 15.10.2020

1.6 Параметрические методы

- **Параметрические методы. Интервальные и метрические шкалы**
- **Понятие корреляции и корреляционная зависимость. Представления о типе, форме и тесноте (плотности) связи**
- **Коэффициент линейной корреляция Пирсона: расчет, статистическая значимость, связь**
- **t-критерий Стьюдента для зависимых и независимых выборок**
- **Параметрические методы. Интервальные и метрические шкалы**
Статистический критерий — это решающее правило, обеспечивающее принятие истинной и отклонение ложной гипотезы с заданной вероятностью (Г.В.Суходольский). Это правило требуется, чтобы математически обосновать выводы.

Параметрические критерии (методы) основаны на расчете параметров генеральной совокупности (X , σ^2).

Достоинства: более мощные и точные.

Недостатки:

- требуют измерений по шкале интервалов или шкале равных отношений (метрической);
- только нормальное распределение!;
- желательный объем выборки $N > 50$.

Приложения	Параметрический тест	Непараметрический тест	Эффективность
Парные выборки	t-тест или z-тест	Критерий знаков	0,63
		Знако-ранговый критерий	0,95
Две независимые выборки	t-тест или z-тест	Критерий Вилкоксона	0,95
Несколько независимых выборок	Дисперсионный анализ (F-тест)	Критерий Краскела-Уоллиса	0,95
Корреляция	Линейная корреляция	Ранговая корреляция	0,91
Проверка случайности	Нет параметрических тестов	Тест Руна	Нет базы сравнения

Эффективность непараметрических тестов оценивалась в сравнении с параметрическими для нормально распределенной генеральной совокупности.

Рисунок 1.6.1 - Показатели эффективности параметрических и непараметрических критериев

Итак, **критерий называют параметрическим**, если он основан на конкретном типе распределения генеральной совокупности, как правило, нормальном, или использует параметры этой совокупности (средние, дисперсии и т.д.), а также работает на интервальных шкалах и шкалах отношений (метрических)).

При нормальном распределении генеральной совокупности параметрические критерии обладают большей мощностью по сравнению с

непараметрическими. Иными словами, они способны с большей достоверностью отвергать нулевую гипотезу, если последняя неверна. Поэтому в тех случаях, когда выборки взяты из нормального распределения генеральных совокупностей, следует отдавать предпочтение параметрическим критериям.

Однако практика показывает, что подавляющее большинство данных, получаемых в психолого-педагогических экспериментах, не распределены нормально, поэтому применение параметрических критериев в анализе результатов исследований может привести к ошибкам в статистических выводах. В таких случаях непараметрические критерии оказываются более мощными, то есть способными с большей достоверностью отвергать нулевую гипотезу.

Решение о выборе того или иного критерия принимается на основании того, является ли выборка зависимой или независимой, в каких шкалах представлены выборочные распределения (для параметрических критериев это интервальные и шкалы отношений (метрические)), сколько выборок сопоставляется, каков их объем и является ли выборочное распределение нормальным.

• ***Понятие корреляции и корреляционная зависимость. Представления о типе, форме и тесноте (плотности) связи***

Термин "корреляция" впервые применил французский палеонтолог Ж. Кювье, который вывел "закон корреляции частей и органов животных" (этот закон позволяет восстанавливать по найденным частям тела облик всего животного). В статистику указанный термин ввел в 1886 году английский биолог и статистик Френсис Гальтон (не просто связь – relation, а "как бы связь" – co-relation). Однако точную формулу для подсчёта коэффициента корреляции разработал его ученик – математик и биолог — Карл Пирсон (1857 – 1936).

Корреляционным называется исследование, проводимое для подтверждения или опровержения гипотезы о статистической связи между несколькими (двумя и более) переменными. В психологии переменными могут выступать психические свойства, процессы, состояния и др.

"Корреляция" в прямом переводе означает "соотношение". Если изменение одной переменной сопровождается изменением другой, то можно говорить о корреляции этих переменных. Наличие корреляции двух переменных ничего не говорит о причинно-следственных зависимостях между ними, но дает возможность выдвинуть такую гипотезу. Отсутствие же корреляции позволяет отвергнуть гипотезу о причинно-следственной связи переменных. Различают несколько интерпретаций наличия корреляционной связи между двумя измерениями.

Виды корреляций. Виды корреляционной связи между измеренными переменными могут быть различны: так корреляция бывает линейной и нелинейной, положительной и отрицательной. Она линейна, если с

увеличением или уменьшением одной переменной, вторая переменная также растёт, либо убывает. Она нелинейна, если при увеличении одной величины характер изменения второй не линеен, а описывается другими законами (полиномиальная, гиперболическая).

Если повышение уровня одной переменной сопровождается повышением уровня другой, то речь идет о положительной корреляции. Чем выше личностная тревожность, тем больше риск заболеть язвой желудка. Возрастание громкости звука сопровождается ощущением повышения его тона.

Если рост уровня одной переменной сопровождается снижением уровня другой, то мы имеем дело с отрицательной корреляцией. По данным Зайонца, число детей в семье отрицательно коррелирует с уровнем их интеллекта. Чем боязливей особь, тем меньше у нее шансов занять доминирующее положение в группе.

Нулевой называется корреляция при отсутствии связи переменных.

Корреляционная зависимость – взаимосвязь между переменными (признаками), при которой в зависимости от изменения одной переменной меняются значения другой переменной.

В основе корреляционного анализа лежат представления о типе, форме и тесноте (плотности) связи. По типу корреляционные связи дифференцируются на прямые и обратные. При прямой связи (рисунок 1.6.2) с увеличением (уменьшением) значений одной переменной значения другой переменной также возрастают (убывают). При обратной связи (рисунок 1.6.3) увеличение (уменьшение) одной переменной вызывает уменьшение (увеличение) другой.

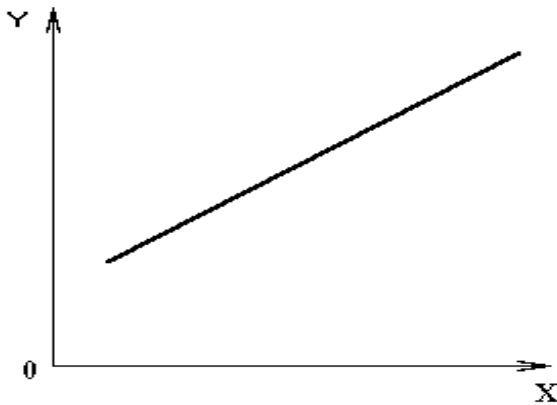


Рисунок 1.6.2

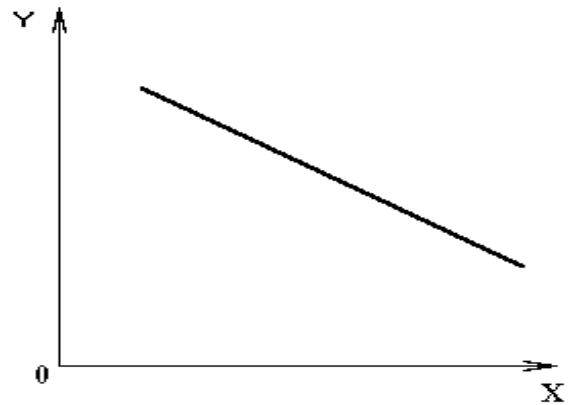


Рисунок 1.6.3

По форме корреляционные связи делятся на линейные и нелинейные. Например, при нелинейной корреляционной связи относительное увеличение (уменьшение) одной переменной влечет за собой неравномерное изменение (увеличение и уменьшение) второй переменной.

Под теснотой (плотностью) понимается степень сопряженности между двумя переменными (признаками, явлениями). На рисунке 1.6.4 корреляционное поле имеет высокую тесноту связи, т.е. числовые значения

близки к своему среднему значению. На рисунке 1.6.5 наблюдается низкая теснота связи, т.е. числовые значения сильно отклоняются от среднего.

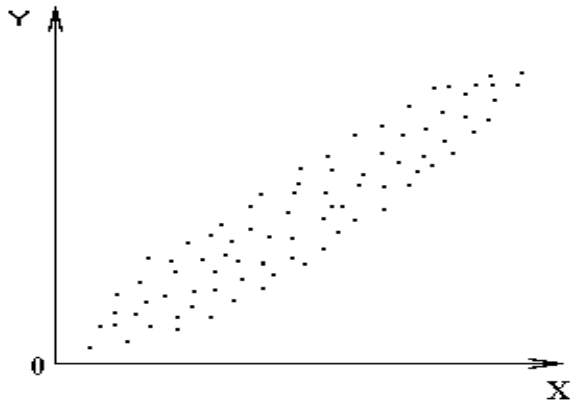


Рисунок 1.6.4

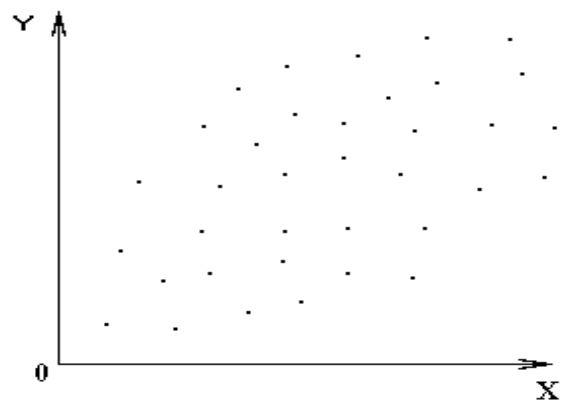


Рисунок 1.6.5

Тип, форма и теснота связи определяются на основе такой статистической характеристики, как коэффициент корреляции.

Переменные X и Y могут быть измерены в разных шкалах, именно это определяет выбор соответствующего коэффициента корреляции. Представим соотношения между типами шкал, в которых могут быть измерены переменные X и Y и соответствующими мерами связи в виде таблицы 1.6.1:

Таблица 1.6.1 - Соотношение типов шкал и мер связи

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 ρ_{xy}
Ранговая	Ранговая	Коэффициент Кендалла τ
Дихотомическая	Дихотомическая	Коэффициент ϕ
Дихотомическая	Ранговая,	Рангово-бисериальный R_{rb}
Дихотомическая	Интервальная или отношений	Бисериальный $R_{бис}$
Интервальная	Ранговая	Не разработан

• **Коэффициент линейной корреляция Пирсона: расчет, статистическая значимость, связь**

В качестве метрической меры парной связи используется коэффициент линейной корреляции r (Пирсона). Исходные данные для его определения – две переменные x и y , принимающие при каждом измерении конкретные числовые значения. Расчет коэффициента линейной корреляции производится по формуле:

$$r_{xy} = \frac{\sum_{i=1}^M (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^M (x_i - \bar{x})^2 \sum_{i=1}^M (y_i - \bar{y})^2}}$$

где индекс xy при коэффициенте r означает, что корреляционная зависимость устанавливается между характеристиками x и y ;

x_i – i -значение переменной x ;

\bar{x} – среднее арифметическое по переменной x ;

y_i – значение переменной y ;

\bar{y} – среднее арифметическое по переменной y ;

M – количество измерений.

Коэффициент Пирсона r_{xy} изменяется от -1 до $+1$ и содержит в себе информацию трех видов:

- об уровне статистической значимости p ;
- о величине связи: при $0.3 \leq |r| < 0.5$ – связь слабая, при $0.5 \leq |r| < 0.7$ – связь умеренная, при $|r| \geq 0.7$ – связь сильная;
- о направлении (знаке) связи: при $r < 0$ переменные x и y являются обратно пропорциональными (чем больше x , тем меньше y , и наоборот), при $r > 0$ переменные x и y являются прямо пропорциональными (чем больше x , тем больше y , и наоборот).

• **t -критерий Стьюдента для зависимых и независимых выборок**

Назначение t -критерия Стьюдента (В.Госсет): сравнение средних значений 2 выборок (есть модификации для зависимых, независимых, эмпирической и теоретической выборок).

Ограничения: нормальное распределение в выборках; предварительное сравнение дисперсий с помощью F -критерия Фишера.

Два случая: при равенстве генеральных дисперсий и при их неравенстве.

Одновыборочный t -критерий: сравнить среднее выборки со средним генеральной совокупности.

Независимый 2-выборочный t -критерий: сравнить средние 2 невязимосвязанные выборки.

t -критерий для 2 зависимых выборок: сравнить изменение среднего в выборке «до» и «после».

t-критерий Стьюдента t-критерий Стьюдента

Дисперсии равны $\sigma_1^2 = \sigma_2^2$ Дисперсии неравны

$$|t| = \frac{|\bar{x}_1 - \bar{x}_2|}{s \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s = \sqrt{\frac{df_1 \cdot s_1^2 + df_2 \cdot s_2^2}{df_1 + df_2}}$$

$$|t| = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Найти df по формуле:

$$\frac{1}{df} = \frac{c^2}{n_1 - 1} + \frac{(1 - c)^2}{n_2 - 1}$$

Где $c = \frac{s_1^2/n_1}{s_1^2/n_1 + s_2^2/n_2}$ и сравнить

Если $t < t_{\text{крит}}$ для $p < 0,01$, то гипотеза H_0 верна

Рисунок 1.6.6 - Расчетные формулы для вычисления t-критерия Стьюдента для зависимых и независимых выборок

Случай зависимых выборок.

$$t = \frac{\bar{d}}{S_d}$$

$$\bar{d} = \frac{\sum d_i}{n} = \frac{\sum (x_{1i} - x_{2i})}{n}$$

$$d_i = x_{1i} - x_{2i}$$

$$df = n - 1$$

$$S_d = \frac{\sigma_d}{\sqrt{n}}$$

Случай независимых выборок.

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$n_1 = n_2 = n$
 $df = n - 1$

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2(n_1 - 1) + \sigma_2^2(n_2 - 1)}{n_1 + n_2 - 2} + \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$n_1 \neq n_2$
 $df = n_1 + n_2 - 2$

Рисунок 1.6.7 - Расчетные формулы для вычисления t-критерия Стьюдента для зависимых и независимых выборок

Пример вычисления t-критерия Стьюдента для 2 зависимых выборок

Пусть в результате проведения методики на выявление уровня профессиональной мотивации учащихся до и после тренинговых мероприятий получены следующие значения:

$x_i = 23, 32, 34, 21, 35, 29, 32, 36, 32, 27, 33, 29, 33, 32, 23, 36, 33, 21, 30, 22$ —

выборочные значения по показателю профессиональной мотивации до

тренинга;

$y_i=33,45,53,28,40,28,33,34,25,41,21,32,50,24,21,43,23,43,30,52$ — выборочные значения по показателю профессиональной мотивации после тренинга.

Тогда:

$\bar{x}=34,95$ — средний показатель уровня профессиональной мотивации учащихся до тренинга, $D_x \approx 106,58$ — значение дисперсии до тренинга, $M_x=20$ — количество испытуемых;

$\bar{y}=29,65$ — средний показатель уровня профессиональной мотивации учащихся после тренинга, $D_y \approx 25,71$ — значение дисперсии после тренинга, $M_y=20$ — количество испытуемых.

Необходимо установить имеет ли место статистическая значимость различия между двумя арифметическими средними \bar{x} и \bar{y} , и каков уровень статистической значимости p . Или, другими словами, отличаются ли средние показатели уровня профессиональной мотивации учащихся до и после тренинговых мероприятий, и с каким уровнем значимости это различие можно считать достоверным?

Воспользуемся расчетом t -критерия Стьюдента и получим значение $t=2,0932$ и коэффициента Фишера $t_{0,05}=2,093$ (для объема выборки $M=20$). Так как $t > t_{0,05}$, то арифметические средние \bar{x} и \bar{y} отличаются с уровнем значимости $p=0,05$. Возвращаясь к условию примера, можно сказать, что средние показатели уровней профессиональной мотивации учащихся до и после тренинга отличаются с указанным уровнем значимости. Или, другими словами, работа, направленная на повышение профессиональной мотивации учащихся, является эффективной.

Литература:

1. Вуколов, Э. А. Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL : уч. пос. / Э. А. Вуколов – 2 изд., испр. и доп. – Москва : Форум: НИЦ Инфра-М. – 464 с. – [Электронный ресурс]. – Режим доступа: <https://uch-lit.ru/matematika-2/dlya-studentov/vukolov-e-l-osnovyi-statisticheskogo-analiza-praktikum-po-statisticheskim-metodam-i-issledovaniyu-operatsii-s-ispolzovaniem-paketov-statistica-i-excel-onlayn>. – Дата доступа: 15.10.2020
2. Боровиков, В. В. STATISTICA. Искусство анализа данных на компьютере: Для профессионалов. 2-е изд. (+CD) / В. В. Боровиков. – Санкт-Петербург : Питер. – 688 с. – [Электронный ресурс]. – Режим доступа: <https://www.books.ru/books/statistica-iskusstvo-analiza-dannykh-na-kompyutere-2-e-izdanie-cd-80566>. – Дата доступа: 15.10.2020
3. Ермолаев, О. Ю. Математическая статистика для психологов : Учебник // 2-е изд., испр. / О. Ю. Ермолаев. – Москва : Московский психолого-социальный институт : Флинта. – 336 с. – [Электронный ресурс]. – Режим

доступа: [https://e-catalog.nlb.by / Author/Home?author](https://e-catalog.nlb.by/Author/Home?author). – Дата доступа: 15.10.2020

4. Килин, П. М. Статистические методы обработки данных : учебное пособие / П. М. Килин, Н. И. Чекмарева. – Тюмень : ТюмГНГУ. – 128 с. – [Электронный ресурс]. – Режим доступа: http://elib.tyuiu.ru/wp-content/uploads/2014/02/Статистические_методы_обработки...110_A5.pdf – Дата доступа: 15.10.2020

5. Пашкевич, О. И. Статистическая обработка эмпирических данных в системе STATISTICA : учебно-методическое пособие / О. И. Пашкевич – Электрон. текстовые данные. – Минск : РИПО. – 148 с. – [Электронный ресурс]. – Режим доступа: <http://www.iprbookshop.ru/67607.html>. – Дата доступа: 15.10.2020

6. Тюрин, Ю. Н. Анализ данных на компьютере : учебное пособие / Ю. Н. Тюрин, А. А. Макаров. – Москва : МЦНМО. – 367 с. – [Электронный ресурс]. – Режим доступа: <http://маткнига.рф/wp-content/uploads/2017/02/978-5-4439-3011-4-Tyurin-Makarov-Analiz-dannyh.pdf>. – Дата доступа: 15.10.2020

7. StatSoft. Электронный учебник по статистике. – [Электронный ресурс] – Режим доступа: <http://statsoft.ru/home/textbook/default.htm>. – Дата доступа: 15.10.2020

1.7 Непараметрические методы

- *Непараметрические методы*
- *Порядковая шкала и коэффициент ранговой корреляции Спирмена. Процедура ранжирования*
- *Расчет ранговой корреляции Спирмена. Статистическая значимость, понятие о связи применительно к коэффициенту ранговой корреляции*

- *Непараметрические методы*
Непараметрические методы (критерии) – методы (критерии) не включающие в формулу расчета параметров распределения, основанные на оперировании частотами или рангами. Работают на номинальных, порядковых шкалах, а также на интервальных и шкалах отношений при несоблюдении требования нормальности выборочного распределения.

Достоинства:

- просты в расчете;
- применимы на малых выборках ($N < 10$);
- не привязаны к характеру распределения.

Недостатки: менее мощные, имеют табличные ограничения по максимальному объему выборки N .

Три преимущества непараметрических методов:

- Они могут использоваться для проверки гипотез о параметрах генеральной совокупности, когда переменная не распределена нормально.

- Они могут использоваться для номинальных и порядковых данных.
- Они могут использоваться для проверки гипотез, которые не связаны с параметрами генеральной совокупности.

Четыре недостатка непараметрических методов:

- Они менее точны, чем соответствующие параметрические методы. Следовательно, требуются более значительные отклонения, чтобы отвергнуть нулевую гипотезу.
- Они менее информативны, чем параметрические критерии. Например, критерий знаков позволяет исследователю определить, превосходит значение данных медиану или нет, но не отвечает — на сколько именно.
- Они менее эффективны, чем соответствующие параметрические критерии. Например, непараметрический критерий знаков дает лишь 60% эффективности от того, что можно получить, используя его параметрическое соответствие — Z-критерий. Требуется больший объем выборки, чтобы компенсировать утрату информации: нужна выборка из 100 человек для критерия знаков, в то время, как для аналогичных результатов при использовании Z-критерия достаточно было бы выборки из 60 человек.
- В большинстве случаев для непараметрических методов подсчеты проще, чем для параметрических. Они более понятны.

Непараметрические статистические критерии и их назначение:

U-критерий Манна-Уитни.

Назначение критерия: оценка достоверности различий между 2 выборками по уровню признака. Выявление различий в показателях исследуемого признака.

Суть критерия: оценивает зону совпадений значений выборок после сплошного ранжирования.

Ограничения критерия: а) $N_1 > 2$, $N_2 > 5$ (или каждая > 3); б) N_1 , N_2 не более 60.

Алгоритм подсчета (Е.В. Сидоренко):

- Перенести все данные на отдельные карточки двух цветов (Например, n_1 - синие, n_2 - красные);
- Разложить все карточки по возрастанию значений;
- Приписать каждому значению ранг, начиная с меньшего (Правила ранжирования!)

• Проверить: для всего ряда рангов $\sum_{рангов} = \frac{N(N+1)}{2}$

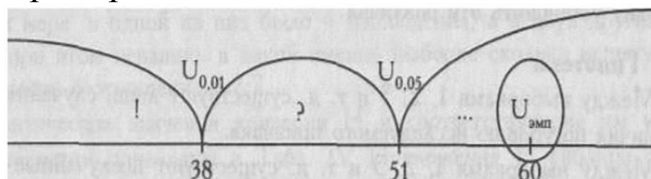
- Для каждой выборки отдельно посчитать сумму рангов
- Наибольшую сумму рангов обозначить как T_x

• Считать $U = n_1 - n_2 + \frac{n_x + 1}{2} - T_x$, где n_x – выборка с наибольшей суммой рангов.

- Сопоставить с табличными критическими значениями $U_{кр}$.

- Если $U < U_{кр.}$ для $p=0,01$, тогда различие значимо.

Пример: Различий нет



Н-критерий Крускала-Уоллеса

Назначение критерия: оценка достоверности различий между 3 и более выборками по уровню признака;

Суть критерия: оценивает различия в суммах рангов, полученных каждой выборкой после сплошного ранжирования всех испытуемых.

Ограничения критерия: а) $N_1 > 2$, N_2 и $N_3 > 4$ (или каждая > 3); б) упускает различия между отдельными парами выборок

Алгоритм подсчета:

- Перенести данные каждой выборки на отдельные карточки определенного цвета;
- Разложить все карточки по возрастанию значений;
- Приписать каждому значению ранг, начиная с меньшего (проверить по Правилам ранжирования!)
- Посчитать сумму рангов каждой выборки, обозначить ее как T_1, T_2, T_3 .

$$H = \left[\frac{12}{N \cdot (N + 1)} \cdot \sum_{j \text{ выборки}} \left(\frac{T_j^2}{n_j} \right) \right] - 3 \cdot (N + 1)$$

- Если хотя бы одна выборка имеет объем $n > 5$, критические значения по таблицам критерия хи-квадрат (χ^2) для $df=N-1$;
- Нарисовать ось значимости, отметить $p=0,05$ и $p=0,01$
- Если рассчитанное значение $H \geq H_{кр.}$ для $p=0,05$, различие значимо и H_0 отвергается

Q-критерий Розенбаума – непараметрическая оценка различий между двумя выборками по уровню какого-либо признака, количественно измеренного (для выборок с $N > 11$);

S - критерий тенденций Джонкира – выявляет тенденции изменения признака при переходе от выборки к выборке при сопоставлении 3 и более выборок (объем выборок одинаков, не более 6 выборок, $N < 10$)

Т-критерий Вилкоксона

Назначение критерия: оценка достоверности изменений показателя выборки в разных условиях, с учетом направления сдвига;

Суть критерия: основан на ранжировании абсолютных разностей пар значений зависимых выборок.

Ограничения критерия:

- а) объем выборки $5 < N < 50$;
- б) нулевые сдвиги из выборки исключаются;
- в) мощнее при значительных сдвигах.

Алгоритм подсчета: $T = \sum Rr$

- Сортировать испытуемых по алфавиту;
- Вычислить разность между показателями «до» и «после»;
- Отдельной колонкой записать модули разностей;
- Ранжировать модули разностей по возрастанию (соблюдать Правила ранжирования!);
- Отдельными колонками выписать ранги для «+» и «-» сдвигов (пометить те, которые являются нетипичными);
- Считать значение T по формуле, где Rr — ранговые значения нетипичных сдвигов;
- По таблице критических значений определить границы значимости. Сделать статистический вывод

G - критерий знаков- установление общего направления сдвига (номинативные и ранговые переменные, незначительные сдвиги; $5 < (N_1 + N_2) < 300$).

Критерий χ^2_r Фридмана- сопоставление показателей, измеренных в 3 или более условиях на одной и той же выборке (не определяет направление изменений; $N > 2$; замеров > 3).

L-критерий тенденций Пейджа- направление изменений 1 выборки от 3 до 6 условий ($N < 12$).

• **Порядковая шкала и коэффициент ранговой корреляции Спирмена.**
Процедура ранжирования

Ранговая корреляция — это метод корреляционного анализа, отражающий отношения переменных, упорядоченных по возрастанию их значения.

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

Для практических целей использование ранговой корреляции весьма полезно. Например, если установлена высокая ранговая корреляция между

двумя качественными признаками, то достаточно осуществлять контроль только по одному из признаков, что, в свою очередь повышает эффективность процесса и ускоряет собственно контроль.

Ранги — это порядковые номера единиц совокупности в ранжированном ряду. Если **проранжировать** совокупность по двум признакам, связь между которыми изучается, то полное совпадение рангов означает максимально тесную прямую связь, а полная противоположность рангов — максимально тесную обратную связь. Ранжировать оба признака необходимо в одном и том же порядке: либо от меньших значений признака к большим, либо наоборот.

В общем случае **ранжирование** — процедура упорядочения любых объектов по возрастанию или убыванию некоторого их свойства при условии, что они этим свойством обладают. Например, можно ранжировать респондентов по степени: их удовлетворенности чем-то, их трудом, политической активности, отношения к чему-то и т. д. Районы города Минска можно ранжировать по их престижности, а государства по качеству жизни, уровню рождаемости и т. д. Можно ранжировать информационные телепередачи по степени их информативности, профессии — по престижности и т.д. Возможно также ранжирование качеств человека по их важности в карьере, ранжирование товаров по предпочтению покупателей.

Объекты ранжирования — это те объекты, которые упорядочиваются. Они могут быть самыми разными. Основание ранжирования — это то свойство, по которому объекты упорядочиваются. В результате упорядочения получаем ранжированный ряд. В нем каждому объекту приписывается ранг — место в этом ряду. Число мест и, соответственно, число рангов равно числу объектов. Существует различие между ранжированием и измерением по порядковой шкале. В последнем случае речь идет о том, что респондент получает оценку по порядковой шкале. При этом оценка получена либо на основе приписывания конкретной цифры вариантам ответа (измерение как кодирование), либо по логическому квадрату, либо по шкале Лайкерта, либо по шкале Гуттмана. Но на основе этих оценок можно провести и ранжирование респондентов.

Ранжирование может выступать как этап анализа данных и как метод измерения. В первом случае ранжирование представляет собой переход от совокупности наблюдений к последовательности их рангов. Даже применительно к шкалам, предположительно высокого уровня измерения, этот этап обработки данных позволяет улучшать результаты анализа. Вторым смыслом термина «ранжирование» подразумевает непосредственное упорядочение объектов респондентом в соответствии с принятыми им внутренними критериями. Метод ранжирования основан на простых, интуитивно доступных и для исследователей, и для респондентов операциях. Поэтому он является одним из старейших и популярнейших методов измерения в социальных науках.

• **Расчет ранговой корреляции Спирмена. Статистическая значимость, понятие о связи применительно к коэффициенту ранговой корреляции**

Для определения взаимосвязи между двумя рассматриваемыми или изучаемыми характеристиками, представленными в порядковой шкале, пользуются ранговой корреляцией. Коэффициент ранговой корреляции ρ (Спирмена) рассчитывается по формуле:

$$\rho_{xy} = 1 - \frac{6 \sum_{i=1}^M (x_i - y_i)^2}{M(M-1)(M+1)},$$

где индекс xy означает, что связь устанавливается между характеристиками x и y ;
 x_i и y_i – ранговые значения рассматриваемых характеристик x и y ;
 M – число измерений.

Коэффициент ранговой корреляции ρ_{xy} изменяется от -1 до 1 и содержит в себе информацию трех видов:

- об уровне статистической значимости p ;
- о величине связи: при $0.3 \leq |\rho| < 0.5$ связь слабая, при $0.5 \leq |\rho| < 0.7$ связь умеренная, при $|\rho| \geq 0.7$ связь сильная;
- о направлении (знаке) связи: при $\rho < 0$ связь обратно пропорциональная, а при $\rho > 0$ – прямо пропорциональная.

Литература:

1. Вуколов, Э. А. Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL : уч. пос. / Э. А. Вуколов – 2 изд., испр. и доп. – Москва : Форум: НИЦ Инфра-М. – 464 с. – [Электронный ресурс]. – Режим доступа: <https://uch-lit.ru/matematika-2/dlya-studentov/vukolov-e-1-osnovyi-statisticheskogo-analiza-praktikum-po-statisticheskim-metodam-i-issledovaniyu-operatsii-s-ispolzovaniem-paketov-statistica-i-excel-onlayn>. – Дата доступа: 15.10.2020
2. Ермолаев, О. Ю. Математическая статистика для психологов: Учебник // 2-е изд., испр. / О. Ю. Ермолаев. – Москва : Московский психолого-социальный институт : Флинта. – [Электронный ресурс]. – Режим доступа: <https://e-catalog.nlb.by/Author/Home?author>. – Дата доступа: 15.10.2020
3. Килин, П. М. Статистические методы обработки данных : учебное пособие / П. М. Килин, Н. И. Чекмарева. – Тюмень : ТюмГНГУ. – 128 с. – [Электронный ресурс]. – Режим доступа: http://elib.tyuiu.ru/wp-content/uploads/2014/02/Статистические_методы_обработки...110_A5.pdf – Дата доступа: 15.10.2020

4. Пашкевич, О. И. Статистическая обработка эмпирических данных в системе STATISTICA : учебно-методическое пособие / О. И. Пашкевич – Электрон. текстовые данные. – Минск : РИПО. – 148 с. – [Электронный ресурс]. – Режим доступа: <http://www.iprbookshop.ru/67607.html>. – Дата доступа: 15.10.2020
5. Сидоренко, Е. В. Методы математической обработки в психологии / Е. В. Сидоренко. – Санкт-Петербург: ООО «Речь». – 350 с., ил. – [Электронный ресурс] – Режим доступа: <https://studfiles.net/preview/5921006>. – Дата доступа: 15.10.2020
6. Тюрин, Ю. Н. Анализ данных на компьютере : учебное пособие / Ю. Н. Тюрин, А. А. Макаров. – Москва : МЦНМО. – 367 с. – [Электронный ресурс]. – Режим доступа: <http://маткнига.рф/wp-content/uploads/2017/02/978-5-4439-3011-4-Tyurin-Makarov-Analiz-dannyh.pdf>. – Дата доступа: 15.10.2020
7. StatSoft. Электронный учебник по статистике. – [Электронный ресурс] – Режим доступа: <http://statsoft.ru/home/textbook/default.htm>. – Дата доступа: 15.10.2020

2 ПРАКТИЧЕСКИЙ РАЗДЕЛ

2.1 Практическая работа 1

Качественный анализ результатов педагогического исследования

Цель: осуществить качественный анализ результатов педагогического исследования с использованием процедуры кросстабулирования (Построение таблиц сопряженности) в системе STATISTICA.

Методическое обеспечение. Инструкция по выполнению практической работы 1 (ПРИЛОЖЕНИЕ А) и пакет статистических программ STATISTICA.

Задания:

1. Представьте результаты педагогического исследования в виде таблицы приведенных данных, содержащей уровни изучаемых показателей в системе STATISTICA.

2. Вычислите абсолютные и относительные величины с использованием процедуры кросстабулирования (Построение таблиц сопряженности) в системе STATISTICA.

3. Представьте результаты кросстабуляции графически в виде категоризованных гистограмм.

4. Результаты кросстабулирования и их качественный анализ (интерпретацию) оформите в виде документа MS WORD.

2.2 Практическая работа 2

Определение различий между эмпирической и теоретической таблицами распределений

Цель: выявить различия между эмпирической и теоретической таблицами распределений посредством вычисления ХИ-квадрат-критерия в системе STATISTICA.

Методическое обеспечение. Инструкция по выполнению практической работы 2 (ПРИЛОЖЕНИЕ Б) и пакет статистических программ STATISTICA.

Также можно воспользоваться ссылкой:

https://www.youtube.com/watch?v=i3piYmVAjQw&feature=emb_rel_err, перейдя по которой осуществится доступ к видео, позволяющему ответить на ряд вопросов (ХИ-квадрат-критерий в системе STATISTICA. Как применять ХИ-квадрат-критерий в программе STATISTICA? Требования к исходным данным при применении критерия ХИ-квадрат, графическое представление, пример, интерпретация, алгоритм работы ХИ-квадрата).

Задания:

1. Вычислите ХИ-квадрат-критерий в системе STATISTICA.

2. Результаты вычислений и их содержательную интерпретацию представьте в виде документа MS WORD.

1.3 Практическая работа 3 Описательные статистики

Цель: вычислить описательные статистики в системе STATISTICA.

Методическое обеспечение. Инструкция по выполнению практической работы 3 (ПРИЛОЖЕНИЕ В) и пакет статистических программ STATISTICA.

Также можно воспользоваться ссылкой:

https://www.youtube.com/watch?v=jrlAGufILL0&feature=emb_rel_pause, перейдя по которой осуществится доступ к видео, позволяющему ответить на ряд вопросов, касающихся вычисления описательных статистик в системе STATISTICA.

Задания:

1. Вычислите описательные статистики (среднюю арифметическую, медиану, моду, дисперсию, асимметрию, эксцесс, среднеквадратическое (стандартное) отклонение в системе STATISTICA.
2. Осуществите проверку выборочного распределения на нормальность с помощью критерия Колмогорова-Смирнова в системе STATISTICA.
3. Результаты вычислений и их интерпретацию представьте в виде документа MS WORD.

3 РАЗДЕЛ КОНТРОЛЯ ЗНАНИЙ

3.1 Вопросы к зачету

1. Рассмотрите этапы становления педагогических измерений в историческом аспекте.
2. Охарактеризуйте современные направления в изучении и оценке педагогических явлений.
3. Обозначьте предмет и задачи математической статистики в педагогике.
4. Дайте характеристику методов и технологий педагогических измерений.
5. Охарактеризуйте культуру измерительной деятельности в образовании.
6. Дайте понятие генеральной и выборочной совокупности.
7. Опишите типы выборок.
8. Дайте характеристику шкалам измерений.
9. Охарактеризуйте номинальную и порядковую шкалы измерения.
10. Охарактеризуйте интервальную шкалу и шкалу отношений.
11. Раскройте понятия зависимые и независимые переменные.
12. Раскройте сущность статистических гипотез и статистических критериев.
13. Дайте определение статистической значимости (p -уровня).
14. Охарактеризуйте оценивание и измерение в педагогическом процессе.
15. Дайте классификацию задач и укажите методы их решения.
16. Дайте характеристику параметрических критериев.
17. Дайте характеристику непараметрических критериев.
18. Опишите процесс принятия решения о выборе метода математико-статистической обработки.
19. Рассмотрите способы упорядоченного представления данных. Таблицы первичных эмпирических данных.
20. Укажите на специфику первичных и вторичных данных педагогического исследования.
21. Дайте характеристику статистических таблиц.
22. Опишите способы упорядоченного представления данных. Таблицы кросстабуляции.
23. Охарактеризуйте эмпирические данные педагогического исследования и проведите их классификацию в математической статистике.
24. Раскройте суть качественного анализа результатов педагогического исследования.
25. Раскройте сущность графического представления результатов педагогического исследования.
26. Определите различия между эмпирической и теоретической таблицами распределения посредством Хи-квадрат-критерия.

27. Раскройте сущность ХИ-квадрат-критерия для таблиц распределения (2x2).
28. Раскройте сущность ХИ-квадрат-критерия для таблиц распределения (2x3).
29. Дайте характеристику описательным статистикам.
30. Раскройте понятие среднего арифметического и опишите его вычисление.
31. Охарактеризуйте средние величины и показатели уровня ряда.
32. Опишите меры рассеивания.
33. Раскройте сущность закона нормального (гауссового) распределения.
34. Раскройте сущность t–критерия Стьюдента. Параметрические методы: t–критерий Стьюдента для зависимых переменных.
35. Охарактеризуйте параметрические методы и, в частности, t–критерий Стьюдента для независимых переменных.
36. Раскройте понятие корреляции. Опишите назначение корреляционного анализа.
37. Охарактеризуйте параметрические методы и, в частности, линейную корреляцию Пирсона.
38. Охарактеризуйте параметрические методы и, в частности, ранговую корреляцию Спирмена.
39. Раскройте сущность непараметрических методов и, в частности, U-критерия Манна-Уитни.
40. Раскройте сущность непараметрических методов и, в частности, T-критерия Вилкоксона.

3.2 Требования к содержанию самостоятельной работы слушателей

РЦПО

Тема программы	Содержание самостоятельной работы	Рекомендуемая литература, с указанием объемов для самостоятельного изучения	Форма предъявления результатов самостоятельной работы
1. Сущность педагогических измерений. Основные понятия математической статистики в педагогике	Выборка и шкалы измерения в педагогическом исследовании. Генеральная и выборочная совокупности. Типы выборок (собственно-случайная, типическая, механическая, серийная, комплексная). Оценивание и измерение в педагогическом процессе. Шкалы измерения в педагогическом исследовании. Параметрические и непараметрические методы	Основная литература: 2, с. 10-11 4, с. 1-21 5, с. 4-10 7, с. 7-22 8, с. 25-39, с. 216-225 Дополнительная литература: 3, с. 61-75, с.243-245 9, с. 12-29	Примеры применения шкал измерения в педагогическом исследовании. Выбор параметрических и непараметрических методов с учетом измерительных шкал
2. Оценивание и измерение в педагогическом процессе	Таблица как способ упорядоченного представления данных педагогического исследования. Вычислительные и аналитические таблицы. Таблицы первичных эмпирических данных. Таблицы распределений. Ряды распределений	Основная литература: 1, с. 14-33 2, с. 11-17 5, с.11-25 Дополнительная литература: 1, с. 8-18 3, с. 217-227 5, с. 4-10, с. 29-33 (MS Word), с. 46-54 (MS Excel) 6, с. 128-135 (MS Word), с. 171-193 (MS Excel) 7, с. 25-30 (MS Excel) 10, с. 99-119	Примеры таблиц первичных эмпирических данных в текстовом процессоре MS Word, в электронных таблицах MS Excel и в системе STATISTICA на электронном носителе. Таблицы и ряды распределений

<p>3. Качественный анализ результатов педагогического исследования</p>	<p>Процедура кросстабулирования в системе STATISTICA. Интерпретация результатов качественного анализа</p>	<p>Основная литература: 1, с. 63-65 5, с. 79-90</p>	<p>Последовательность выполнения процедуры кросстабулирования в системе STATISTICA. Содержательная интерпретация результатов качественного анализа в электронном варианте</p>
<p>4. Определение различий между эмпирической и теоретической таблицами распределений</p>	<p>Непараметрические методы. Последовательность вычисления χ^2-критерия в системе STATISTICA. Представление и содержательная интерпретация результатов вычисления χ^2-критерия в системе STATISTICA</p>	<p>Основная литература: 2, с. 126-159 5, с. 95-102 6, с. 244-253 Дополнительная литература: 1, с. 105-114 4, с. 66-68 8, с. 113-142 10, с. 39-41</p>	<p>Последовательность и результаты расчета χ^2-критерия в системе STATISTICA. Содержательная интерпретация результатов в электронном варианте</p>
<p>5. Описательные статистики</p>	<p>Описательные статистики. Проверка выборочного распределения на нормальность с помощью критерия Колмогорова-Смирнова в системе STATISTICA</p>	<p>Основная литература: 1, с. 33-38 2, с. 43-52 5, с. 26-29 6, с. 44-46 Дополнительная литература: 1, с. 14-18 4, с. 40-57 8, с. 20-24 10, с. 8-17</p>	<p>Результаты вычисления описательных статистик и проверки выборочного распределения на нормальность с помощью критерия Колмогорова-Смирнова в системе STATISTICA на электронном носителе</p>
<p>6. Параметрические методы</p>	<p>Вычисление коэффициента линейной корреляции Пирсона в системе STATISTICA. Представление и содержательная интерпретация результатов</p>	<p>Основная литература: 1, с. 38-40 2, с. 202-207 5, с. 43-55 6, с. 257-260 Дополнительная литература: 4, с. 102-105, с. 105-169 10, с. 43-45</p>	<p>Последовательность и результаты вычислений линейной корреляции Пирсона в системе STATISTICA и их содержательная интерпретация в электронном варианте</p>

<p>7. Непараметрические методы</p>	<p>Выявление различий с помощью U-критерия Манна-Уитни и изменений с помощью T-критерия Вилкоксона. Вычисление U-критерия Манна-Уитни и T-критерия Вилкоксона с использованием системы STATISTICA и содержательная интерпретация результатов</p>	<p>Основная литература: 2, с. 78-82, с. 101-110, с. 212-222 5, с. 104-109, с. 112-114, с. 128-131 6, с. 253-257 Дополнительная литература: 1, с. 116-121 С. 125-127, с. 141-143 4, с. 105-169 8, с. 49-56 10, с. 31-37, с. 87-94, с. 200-224</p>	<p>Последовательность и результаты вычислений U-критерия Манна-Уитни, T-критерия Вилкоксона, в системе STATISTICA Содержательная интерпретация результатов в электронном варианте</p>
------------------------------------	--	--	---

4 ВСПОМОГАТЕЛЬНЫЙ РАЗДЕЛ

4.1 Фрагмент учебно-тематического плана по специальности переподготовки 1-08 01 71 «Педагогическая деятельность специалистов»

Фрагмент учебно-тематического плана переподготовки в соответствии с типовым учебным планом, утвержденным 28.03.2017, регистрационный номер 25-13/82

Специальность переподготовки 1-08 01 71 «Педагогическая деятельность специалистов»

Квалификация: «Преподаватель в соответствии с квалификацией по основному образованию»

Форма получения образования: заочная

Продолжительность обучения: 22 месяца

РПО

№ п/п	Наименования разделов, тем и форм текущей аттестации	Количество учебных часов											Этапы	Кафедра (цикловая комиссия)
		Распределение по видам занятий												
		Аудиторные занятия												
		Всего	лекции	практические занятия	семинарские занятия	круглые столы, тематические занятия	лабораторные занятия	деловые игры	тренинги	консультации	самостоятельная работа			
1	2	3	4	5	6	7	8	9	10	11	12	13	14	
3.4.	Основы педагогических измерений	36	14	6							16	2		
3.4.1	Сущность педагогических измерений. Основные понятия математической статистики в педагогике	4	2								2	2		
3.4.2	Оценивание и измерение в педагогическом процессе	4	2								2	2		
3.4.3	Качественный анализ результатов педагогического исследования	6	2	2							2	2		
3.4.4	Определение различий между эмпирической и теоретической таблицами распределений	8	2	2							4	2		
3.4.5	Описательные статистики	6	2	2							2	2		
3.4.6	Параметрические методы	4	2								2	2		
3.4.7	Непараметрические методы	4	2								2	2		
Форма текущей аттестации		Зачет											2	

РПО

4.2 Содержание учебной программы дисциплины «Основы педагогических измерений»

Тема 1 Сущность педагогических измерений. Основные понятия математической статистики в педагогике (4 ч.)

История становления педагогических измерений. Современные направления в изучении и оценке педагогических явлений. Предмет и задачи математической статистики в педагогике. Зависимые и независимые переменные. Статистические гипотезы. Уровень статистической значимости (p -уровень). Классификация задач и методы их решения. Методы и технологии педагогических измерений.

Выборка и шкалы измерения в педагогическом исследовании. Генеральная и выборочная совокупности. Типы выборок (собственно-случайная, типическая, механическая, серийная, комплексная). Шкалы измерения в педагогическом исследовании. Выбор метода статистического анализа для дальнейшей обработки данных педагогического исследования.

Тема 2 Оценивание и измерение в педагогическом процессе (4 ч.)

Оценивание и измерение в педагогическом процессе. Культура измерительной деятельности в образовании. Таблицы первичных эмпирических данных. Приведенные таблицы.

Таблица как способ упорядоченного представления данных педагогического исследования. Вычислительные и аналитические таблицы. Таблицы первичных эмпирических данных. Таблицы распределений. Ряды распределений.

Тема 3 Качественный анализ результатов педагогического исследования (6 ч.)

Качественный анализ результатов педагогического исследования. Баллы, уровни и ранги. Табличное представление результатов педагогического исследования.

Таблицы сопряженности. Абсолютные и относительные величины. Результаты кросстабулирования и качественный анализ.

Реализация процедуры кросстабулирования в системе STATISTICA. Построение таблиц сопряженности. Графическое представление результатов кросстабуляции. Категоризованные гистограммы. Интерпретация результатов качественного анализа.

Тема 4 Определение различий между эмпирической и теоретической таблицами распределения (8 ч.)

Эмпирическая и теоретическая таблицы распределения. χ^2 -критерий (χ^2) и определение статистической значимости различия между эмпирической и теоретической таблицами распределения. Таблицы распределения размерности (2×2) , (2×3) .

Вычисление $\chi^2_{(2 \times 2)}$ и $\chi^2_{(2 \times 3)}$. Требование Юла и Кендалла. Переход от эмпирической таблицы к теоретической таблице распределения. Определение статистической значимости различия между рассматриваемыми

признаками (*p*-уровень). Понятие связи между признаками. Расчет коэффициента сопряженности К (Чупрова). Определение величины связи. Интерпретация результатов.

Непараметрические методы. Последовательность вычисления *XI*-квadrat-критерия в системе STATISTICA.

Представление и содержательная интерпретация результатов вычисления *XI*-квadrat-критерия в системе STATISTICA.

Тема 5 Описательные статистики (6 ч.)

Описательные статистики. Показатели уровня ряда. Средняя арифметическая. Медиана. Мода. Асимметрия. Экссесс. Среднеквадратическое (стандартное) отклонение. Меры рассеивания. Достоверность средней арифметической. Доверительный интервал.

Вычисление описательных статистик в системе STATISTICA и их содержательная интерпретация.

Описательные статистики. Проверка выборочного распределения на нормальность с помощью критерия Колмогорова-Смирнова в системе STATISTICA.

Тема 6 Параметрические методы (4 ч.)

Параметрические методы. Интервальные и метрические шкалы. Понятие корреляции и корреляционная зависимость. Представления о типе, форме и тесноте (плотности) связи. Расчет коэффициента линейной корреляции Пирсона. Статистическая значимость коэффициента линейной корреляции Пирсона. Понятие о связи применительно к коэффициенту линейной корреляции. *t*-критерий Стьюдента для зависимых и независимых выборок.

Вычисление коэффициента линейной корреляции Пирсона в системе STATISTICA. Представление и содержательная интерпретация результатов.

Тема 7 Непараметрические методы (4 ч.)

Непараметрические методы. Порядковая шкала и коэффициент ранговой корреляции Спирмена. Процедура ранжирования. Расчет коэффициента ранговой корреляции Спирмена. Статистическая значимость коэффициента ранговой корреляции Спирмена. Понятие о связи применительно к коэффициенту ранговой корреляции.

Выявление различий с помощью *U*-критерия Манна-Уитни и изменений с помощью *T*-критерия Вилкоксона. Вычисление *U*-критерия Манна-Уитни и *T*-критерия Вилкоксона с использованием системы STATISTICA и содержательная интерпретация результатов

4.3 Перечень учебных изданий и информационно-аналитических материалов, рекомендуемых для изучения учебной дисциплины

ОСНОВНАЯ

РИПО

Рекомендуемая литература	Объем для самостоятельного изучения по темам / разделам
1. Боровиков, В. В. STATISTICA. Искусство анализа данных на компьютере: Для профессионалов. 2-е изд. (+CD) / В. В. Боровиков. – Санкт-Петербург : Питер. – 688 с. – [Электронный ресурс]. – Режим доступа: https://www.books.ru/books/statistica-iskusstvo-analiza-dannykh-na-kompyutere-2-e-izdanie-cd-80566 . – Дата доступа: 15.10.2020	Тема 2, с. 14-33 Тема 3, с. 63-65 Тема 5, с. 33-38 Тема 6, с. 38-40
2. Ермолаев, О. Ю. Математическая статистика для психологов : Учебник // 2-е изд., испр. / О. Ю. Ермолаев. – Москва : Московский психолого-социальный институт : Флинта. – 336 с. – [Электронный ресурс]. – Режим доступа: https://e-catalog.nlb.by/Author/Home?author . – Дата доступа: 15.10.2020	Тема 1, с. 10-11 Тема 2, с. 12-17 Тема 4, с. 126-159 Тема 5, с. 43-52 Тема 6, с. 202-207 Тема 7, с. 212-227
3. Загвязинский, В. И. Методология и методы психолого-педагогического исследования : учеб. пособие для студ. высш. пед. учеб. заведений. / В. И. Загвязинский, Р. Атаханов. – 2-е изд., стер. – Москва : Издательский центр «Академия». – 208 с. – [Электронный ресурс]. – Режим доступа: https://prepod.nspu.ru/pluginfile.php/45344/mod_resource/content/0/Zagvyjazinskii.pdf . – Дата доступа: 15.10.2020	Тема 2, с. 217-227
4. Математические методы в педагогических исследованиях : учебное пособие / С. И. Осипова, С. М. Бутакова, Т. Г. Дулинец, Т. Б. Шаипова. – Красноярск : Сибирский федеральный университет. – 264 с. – [Электронный ресурс]. – Режим доступа: http://www.studentlibrary.ru/book/ISBN9785763825060.html . – Дата доступа: 15.10.2020.	Тема 1, с. 1-21
5. Пашкевич, О. И. Статистическая обработка эмпирических данных в системе STATISTICA : учебно-методическое пособие / О. И. Пашкевич – Электрон. текстовые данные. – Минск : РИПО. – 148 с. – [Электронный ресурс]. – Режим доступа: http://www.iprbookshop.ru/67607.html . – Дата доступа: 15.10.2020	Тема 1, с. 4-10 Тема 2, с. 11-25 Тема 3, с. 79-90 Тема 4, с. 95-102 Тема 5, с. 26-29 Тема 6, с. 43-55 Тема 7, с. 104-109
6. Тюрин, Ю. Н. Анализ данных на компьютере : учебное пособие / Ю. Н. Тюрин, А. А. Макаров. – Москва : МЦНМО. – 367 с. – [Электронный ресурс]. – Режим доступа: http://маткнига.рф/wp-content/uploads/2017/02/978-5-4439-3011-4-Tyurin-Makarov-Analiz-dannyh.pdf . – Дата доступа: 15.10.2020	Тема 4, с. 244-253 Тема 5, с. 44-46 Тема 6, с. 257-260 Тема 7, с. 253-257
7. Циулина, М. В. Методология психолого-педагогических	Тема 1, с. 7-22

исследований : учебное пособие (Текст) / М. В. Циулина. – Челябинск : Изд-во Челяб. гос. пед. ун-та. – 239 с. – [Электронный ресурс]. – Режим доступа: http://elib.cspu.ru/xmlui/bitstream/handle/123456789/616/Циулина%20М.В.%20Методологияпсих.-пед.%20исследований.pdf . – Дата доступа: 15.10.2020	
8. Яковлев, Е. В. Педагогическое исследование: содержание и представление результатов / Е. В. Яковлев, Н. О. Яковлева. – Челябинск : Изд-во РБИУ. – 317 с. – [Электронный ресурс]. – Режим доступа: http://elib.cspu.ru/xmlui/bitstream/handle/123456789/278/Пед.%20исследование%202010.pdf?sequence=1&isAllowed=y . – Дата доступа: 15.10.2020	Тема 1, с. 25-39, с. 216-225

ДОПОЛНИТЕЛЬНАЯ

1. Вуколов, Э. А. Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL : уч. пос. / Э. А. Вуколов – 2 изд., испр. и доп. – Москва : Форум: НИЦ Инфра-М. – 464 с. – [Электронный ресурс]. – Режим доступа: https://uch-lit.ru/matematika-2/dlya-studentov/vukolov-e-l-osnovyi-statisticheskogo-analiza-praktikum-po-statisticheskim-metodam-i-issledovaniyu-operatsii-s-ispolzovaniem-paketov-statistica-i-excel-onlayn . – Дата доступа: 15.10.2020	Тема 2, с. 8-18 Тема 4, с. 105-114 Тема 5, с. 14-18 Тема 6, с. 38-40 Тема 7, с. 116-122
2. Звонников, В. И. Измерения и шкалирование в образовании (Текст) : (учеб. пособие) / В. И. Звонников. – Москва : Логос. – 134 с. – [Электронный ресурс]. – Режим доступа: http://library.ntspi.ru/CGI/irbis64r_01/cgiirbis_64.exe . – Дата доступа: 15.10.2020	Тема 2, с. 126-159
3. Звонников, В. И. Современные средства оценивания результатов обучения : учебник для студ. учреждений высш. проф. образования / В. И. Звонников, М. Б. Челышкова. – 5-е изд., перераб. – Москва : Издательский центр «Академия». – 304 с. – (Сер. Бакалавриат). – [Электронный ресурс]. – Режим доступа: https://www.studmed.ru/zvonnikov-vi-chelyshkova-mb-sovremennye-sredstva-ocenivaniya-rezultatov-obucheniya_0f8ff339642.html . – Дата доступа: 15.10.2020	Тема 1, с. 61-75 Тема 2, с. 217-227
4. Килин, П. М. Статистические методы обработки данных : учебное пособие / П. М. Килин, Н. И. Чекмарева. – Тюмень : ТюмГНГУ. – 128 с. – [Электронный ресурс]. – Режим доступа: http://elib.tyuiu.ru/wp-content/uploads/2014/02/Статистические_методы_обработки...110_A5.pdf – Дата доступа: 15.10.2020	Тема 4, с. 66-68 Тема 5, с. 40-57 Тема 6, с. 102-105 Тема 7, с. 106-169
5. Кузин, А. В. Основы работы в Microsoft Office 2013 : учебное пособие / А. В. Кузин, Е. В. Чумакова. – Москва : Инфра-М, Форум. – 160 с. – [Электронный ресурс]. – Режим доступа: https://znanium.com/bookread2.php?book=495075 . – Дата доступа:	Тема 2, с. 4-135

15.10.2020	
6. Леонтьев, В. А. Office 2016. Новейший самоучитель / В. А. Леонтьев. – (Компьютерные книги Виталия Леонтьева). – Москва : Эксмо. – 368 с. – [Электронный ресурс]. – Режим доступа: https://avidreaders.ru/book/office-2016-noveyshiy-samouchitel.html . – Дата доступа: 15.10.2020	Тема 2, с. 128-193
7. Пашкевич, О. И. Курсовая и дипломная работы: структура, оформление, защита : методические рекомендации для слушателей переподготовки по специальности 1-03 04 72 "Практическая психология" (квалификация "Педагог-психолог") / О. И. Пашкевич, О. М. Савчик. – Минск : РИПО. – [Электронный ресурс]. – Режим доступа: <a by-sek-277528"&lst_siz='20"' href="http://unicat.nlb.by/opac/pls/!search.http_keyword?query=a001a=">http://unicat.nlb.by/opac/pls/!search.http_keyword?query=a001a="BY-SEK-277528"&lst_siz=20 . – Дата доступа: 15.10.2020	Тема 2, с. 25-119
8. Сидоренко, Е. В. Методы математической обработки в психологии / Е. В. Сидоренко. – Санкт-Петербург: ООО «Речь». – 350 с., ил. – [Электронный ресурс] – Режим доступа: https://studfiles.net/preview/5921006 . – Дата доступа: 15.10.2020	Тема 7, с. 49-56
9. Степанова, Е. А. Основы обработки результатов измерений [Электронный ресурс] : учебное пособие / Е. А. Степанова, Н. А. Скулкина, А. С. Волегов. – Электрон. текстовые данные. – Екатеринбург: Уральский федеральный университет, ЭБС АСВ. – 96 с. – Режим доступа: http://www.iprbookshop.ru/68268.html . – Дата доступа: 15.10.2020	Тема 1, с. 12-29
10. StatSoft. Электронный учебник по статистике. – [Электронный ресурс] – Режим доступа: http://statsoft.ru/home/textbook/default.htm . – Дата доступа: 15.10.2020	Тема 2, с. 99-119 Тема 4, с. 39-41 Тема 5, с. 8-17 Тема 6, с. 43-45 Тема 7, с. 31-224

Инструкция по выполнению практической работы 1

«Качественный анализ результатов педагогического исследования»

- *Ввод данных в системе STATISTICA*
- *Таблицы сопряженности. Процедура кросстабулирования в системе STATISTICA*
- *Графический анализ таблиц сопряженности*

- *Ввод данных в системе STATISTICA*

Работа в системе STATISTICA начинается с ввода данных. Исходные данные в системе STATISTICA организованы в виде таблиц. Если у пользователя имеется опыт работы с электронными таблицами (типа MS Excel), то он быстро привыкнет и к таблицам системы STATISTICA. Следует отметить, что табличная структура данных STATISTICA позволяет отображать большое количество реальных данных.

Электронная таблица состоит из строк и столбцов. Столбцы таблицы STATISTICA называются **Переменные** (Variables), а строки – **Наблюдения** (Cases).

Например, в педагогике наблюдения – это учащиеся (испытуемые), а в качестве переменных могут выступать пол, возраст, название факультета, квалификация специалистов, группы и т. д. Существует несколько вариантов создания таблицы с данными в системе STATISTICA. Рассмотрим один из них.

Ввод данных, содержащих текстовую информацию.

Запустить программу STATISTICA либо выбрать ярлык STATISTICA на рабочем столе.

В строке меню выбрать **Файл** (File) и в выпадающем меню команду **Новый** (New). На экране компьютера появится окно **Создать новый документ**: (CreateNewDocument) (рисунок А.1).

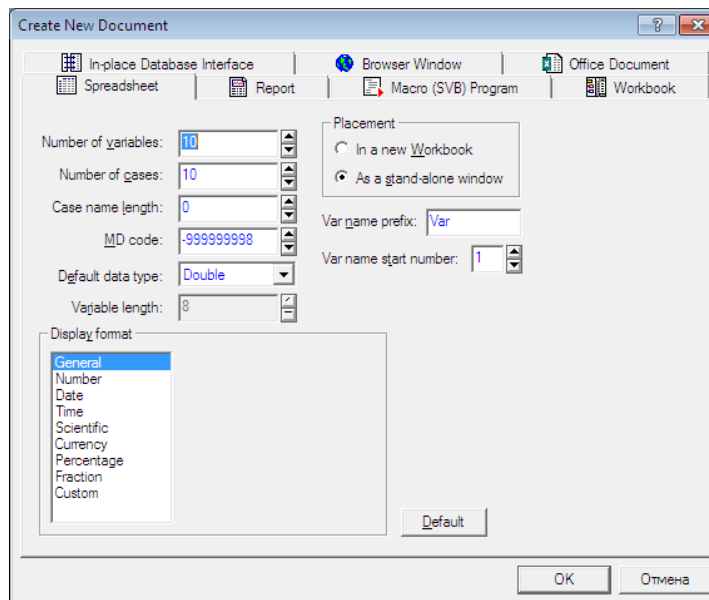


Рисунок А.1 - Окно Создать новый документ (CreateNewDocument)

В окне **Создать новый документ** (CreateNewDocument) можно сразу задать нужное количество **переменных (Number of variables)** и **наблюдений (Number of cases)** (рисунок А.1). Если же этого не сделать, а нажать клавишу **Enter** на клавиатуре или кнопку **OK**, то программа создаст пустую таблицу, содержащую 10 строк и 10 столбцов (рисунок А.2).

	1	2	3	4	5	6	7	8	9	10
	Var1	Var2	Var3	Var4	Var5	Var6	Var7	Var8	Var9	Var10
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										

Рисунок А.2 - Вид электронной таблицы системы STATISTICA

При необходимости можно увеличить или уменьшить количество строк и столбцов таблицы. Для этого используются кнопки **Переменные (Vars)** и **Наблюдения (Cases)** на панели инструментов.

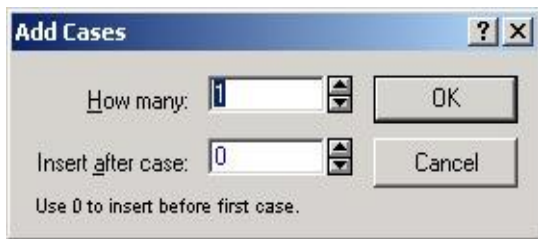


Рисунок А.3 - Окно
Добавить наблюдения

Добавить наблюдения (AddCases) (рисунок А.3), в котором можно задать число наблюдений (испытуемых), добавляемых в таблицу, например 2.

После нажатия на кнопку **OK** количество строк (наблюдений) в таблице увеличится на 2, т. е. станет равным 12.

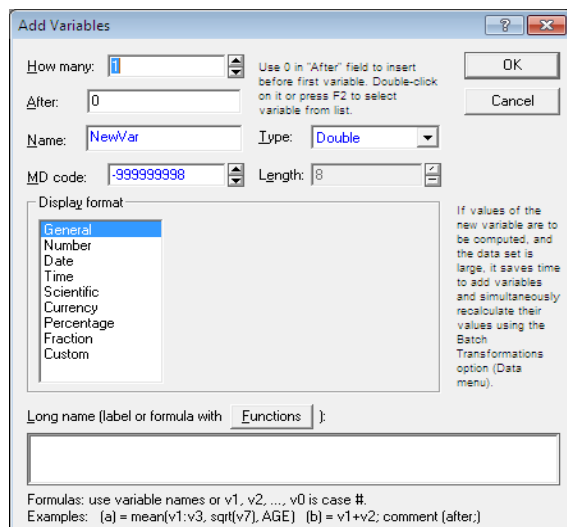


Рисунок А.4 - Окно **Добавить переменные**

Здесь же можно задать и имя переменной (Name).

Кроме описанных выше действий, можно также вводить имена наблюдений (испытуемых). Для этого надо нажать кнопку **Наблюдения** **Cases** и выбрать пункт меню Case NameManager. На экране появится диалоговое окно **Диспетчер имени наблюдения** (Case NameManager) (рисунок А.5), в котором можно определить, сколько символов в таблице будет зарезервировано для имен наблюдений и др.

После нажатия на кнопку **Наблюдения** **Cases** раскроется меню, предлагающее следующий выбор для наблюдений таблицы: **Добавить** (Add), **Переместить** (Move), **Копировать** (Copy), **Удалить** (Delete), **Имена** (Names). Выберем, например, пункт **Добавить** (Add), щелкнув левой кнопкой мыши. Откроется окно

Аналогичным образом можно увеличить количество переменных в таблице. Для этого нужно нажать на кнопку **Переменные** **Vars** на панели инструментов. С помощью курсора мыши в выпадающем меню выбрать пункт **Добавить** (Add). На мониторе появится окно **Добавить переменные** (AddVariables) (рисунок А.4), где надо ввести количество добавляемых переменных (How many), например 1, и номер переменной, после которой будут вставляться добавляемые переменные (After), например Var 2.

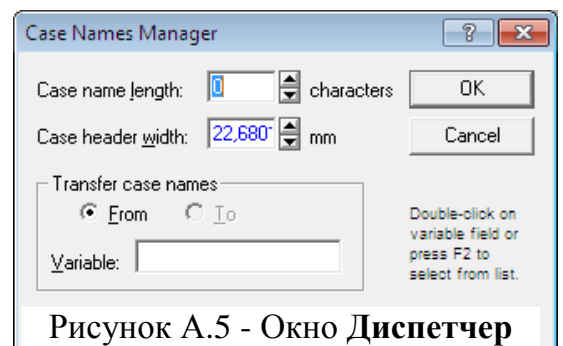


Рисунок А.5 - Окно **Диспетчер имени наблюдения** (Case NameManager)

Поле для имен наблюдений можно раздвинуть с помощью мыши, зафиксировав ее левую кнопку.

Итак, мы имеем электронную таблицу, которая содержит 12 столбцов (переменных) и 10 строк (наблюдений, реализаций, испытуемых), а также место для ввода имен наблюдений (рисунок А.6).

Теперь необходимо ввести наблюдения (значения) и имена переменных. Работать нужно, используя мышь и клавиатуру. Двойной щелчок мыши по полям заголовков открывает диалоговые окна, позволяющие вводить заголовки, описывать переменные, наблюдения (например, имена испытуемых) и т. д.

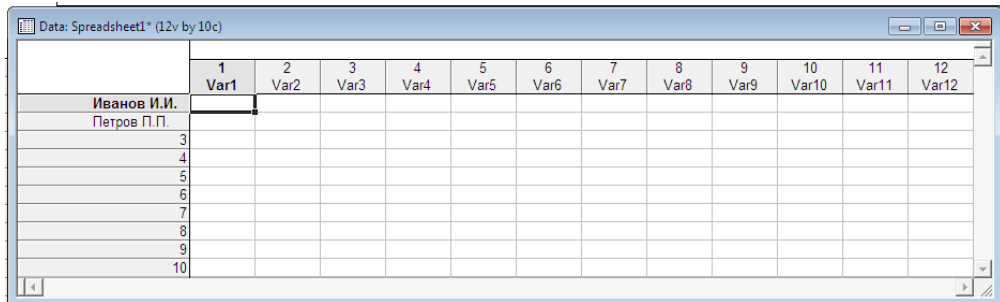


Рисунок А.6 - Вид электронной таблицы системы STATISTICA, содержащей поле ввода имени наблюдения

Для того чтобы описать переменную, необходимо дважды щелкнуть мышью по ее имени. Например, после щелчка по заголовку переменной 1 (VAR1) откроется окно, в котором можно задать ее имя (или переименовать ее), формат, метку, связь и т. д. (рисунок А.7).

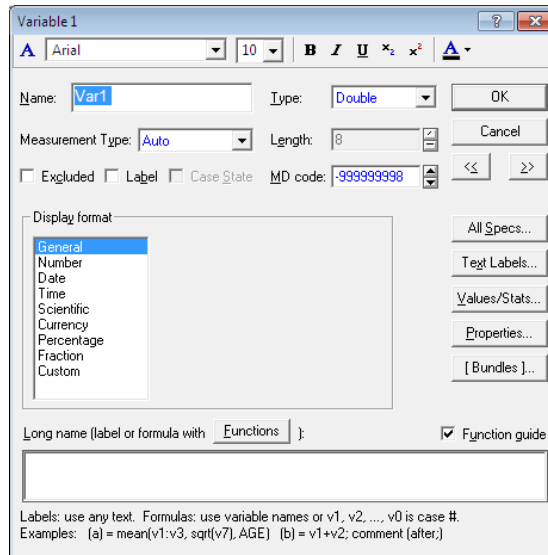


Рисунок А.7 - Окно переменной

Работа с данным окном заканчивается нажатием кнопки **OK**. Далее можно заполнить созданную таблицу данными, которые вводятся непосредственно с клавиатуры.

Для ввода числовых данных используются мышь, клавиатура и стрелки перемещения курсора. Курсор ставится на нужную ячейку таблицы, куда вводятся числовые данные. Текстовые данные вводятся иначе.

Примечание 1. Поскольку система STATISTICA является обычным Window-приложением, можно легко и быстро импортировать данные, полученные в системе STATISTICA, в другое Windows-приложение, например, в MS Word. Импорт данных может осуществляться в виде рисунка или же таблицы. В первом случае необходимо нажать одновременно кнопки **ALT** и **F3**. На экране вместо курсора мыши появится значок «прицел». Используя мышь, надо поместить «прицел» в верхний левый угол таблицы. Затем нажать левую кнопку мыши, зафиксировав «прицел», и, удерживая кнопку мыши, переместить его в правый нижний угол таблицы. Выделенная часть таблицы будет отмечена прямоугольной рамкой. После того как будет отпущена кнопка мыши, отмеченная часть таблицы будет помещена в буфер обмена. Если теперь открыть нужный документ Word и нажать на клавиатуре комбинацию кнопок **CTRL** и **V**, то выбранный сегмент таблицы будет скопирован в документ. При импортировании данных в виде таблицы, ее необходимо предварительно выделить, далее нажать правый клик мыши и выбрать в появившемся контекстном меню команду **Copy with Headers**. Таблица будет помещена в буфер обмена. После этого ее можно вставлять в документ MS Word.

Ввод данных, содержащих текстовую информацию.

TEX VAL	1	2
	СЕМ_ПОЛ	ТРЕВОГА
1	П_семья	Высокая
2	Н_семья	Низкая
3	Н_семья	Высокая
4	Н_семья	Низкая
5	П_семья	Высокая
6	П_семья	Низкая
7	П_семья	Низкая
8	Н_семья	Высокая
9	П_семья	Низкая
10	Н_семья	Высокая

Рисунок А.8 - Рабочий файл данных WOMEN1.STA

Рассмотрим, как можно создать набор данных, который содержит не только числовую, но и текстовую информацию. Создадим рабочий файл данных, как было рассмотрено выше.

В файле **WOMEN1.STA**, представленном в виде электронной таблицы (рисунок А.8), содержатся результаты опроса 10 испытуемых относительно их семейного положения и состояния уровня тревожности. Первая переменная **СЕМ_ПОЛ** описывает семейное положение испытуемых. Эта переменная принимает два значения: **П_семья** – полная семья, **Н_семья** – неполная семья. Вторая переменная, **ТРЕВОГА**, описывает уровень тревожности испытуемых. Она принимает два значения: **низкая**, **высокая**. Известно, что личностная тревожность характеризуется устойчивой склонностью воспринимать жизненную

ситуацию как угрожающую (содержащую в себе тайную угрозу). Как видно из таблицы, первый опрошенный испытуемый (наблюдение № 1 – первая строка в таблице) из полной семьи и его состояние характеризуется как высоко тревожное. Второй опрошенный (наблюдение № 2 – вторая строка таблицы) из неполной семьи и его уровень своей тревожности оценен как низкий и т. д.

Как видно из рисунка А.8, переменные в файле **WOMEN1.STA** принимают текстовые значения. Ввод текста в таблицу занимает слишком много времени. Для удобства лучше вводить численные значения, а затем использовать кодировку значений переменных. Рассмотрим, как это делается, на примере переменной **СЕМ_ПОЛ**. После двойного щелчка левой кнопкой мыши по заголовку указанной переменной на экране отобразится окно переменной (рисунок А.9):

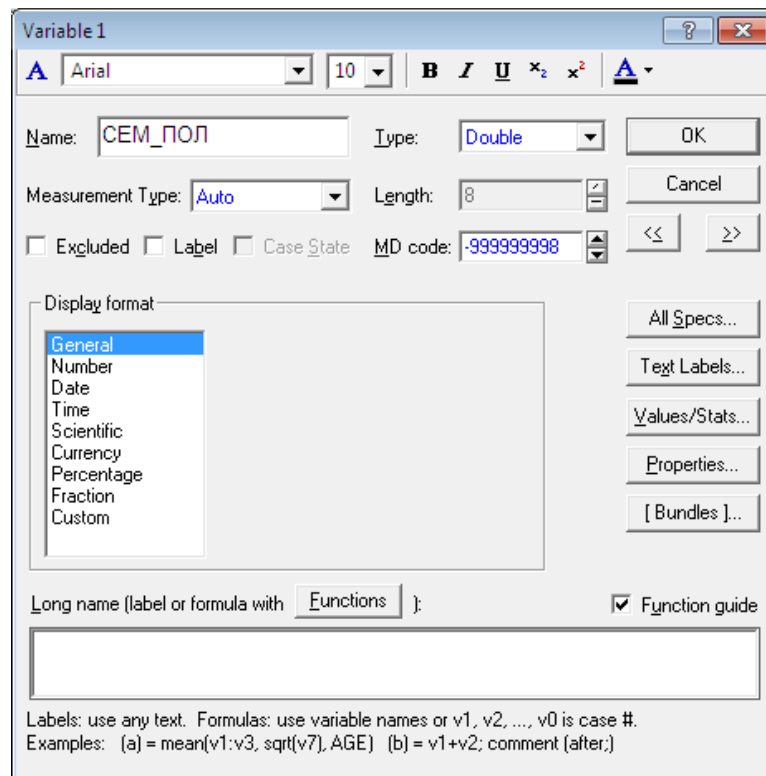


Рисунок А.9 - Окно переменной

Нажатие кнопки **Текстовые Значения** (Text Labels) позволит осуществить ввод текстовой информации с соответствующей кодировкой (рисунок А.10):

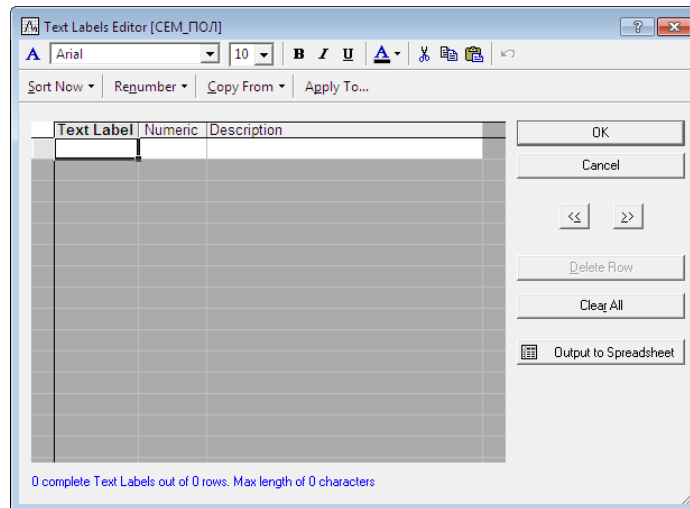


Рисунок А.10 - Окно Редактор текстовых значений (Text Labels Editor)

В этом окне в колонке **Текстовое Значение** (Text Label) необходимо набрать в первой строке **П_семья**, а в колонке **Число** (Numeric) – значение, равное 101 заменить на 1 (!). Таким образом, текстовому значению **П_семья** будет присвоен код 1. Во второй строке колонки **Текстовое Значение** (Text Label) надо набрать **Н_семья**, а в колонке **Число** (Numeric) – числовое значение, равное 101 заменить на 2. Таким образом, текстовому значению **Н_семья** будет присвоен код 2. Далее следует нажать кнопку **OK**.

NUM VAL	СЕМ_ПОЛ
1	1
2	2
3	000
4	000
5	000
6	000
7	000
8	000
9	000
10	000

Рисунок А.11 -
Значения переменной
СЕМ_ПОЛ

После этого можно ввести значения, равные 1, в те ячейки переменной **СЕМ_ПОЛ**, где должно стоять текстовое значение **П_семья**, а значения, равные 2, – текстовое значение **Н_семья**. Таким образом, переменная **СЕМ_ПОЛ** является номинальной, или, другими словами, значения переменной **СЕМ_ПОЛ** отнесены к номинальной шкале (рисунок А.11).

Таким же образом вводятся текстовые значения в ячейки переменной **ТРЕВОГА**, и после этого рабочий файл с именем **WOMEN1.STA** будет создан.

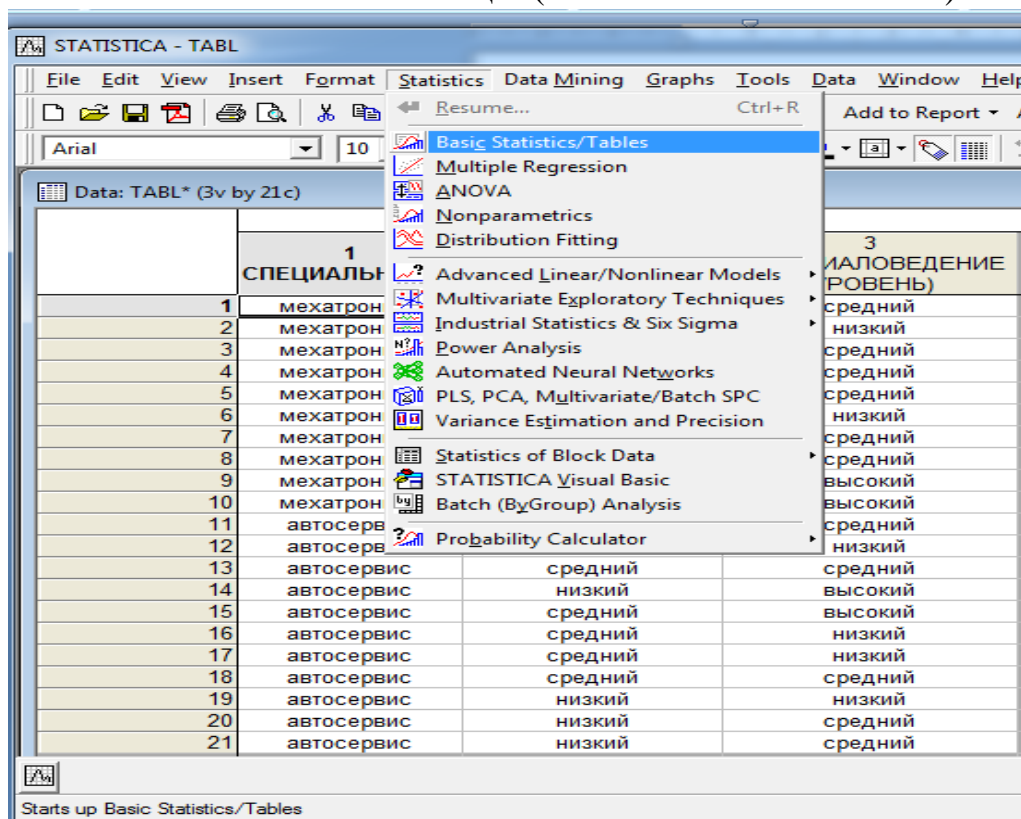
- **Таблицы сопряженности. Процедура кросстабулирования в системе STATISTICA**

Построение таблиц сопряженности будем осуществлять, используя исходные данные файла **TABL.STA** (рисунок А.12).

TABL* (3v by 21c)			
	1 СПЕЦИАЛЬНОСТЬ	2 МЕХАНИКА (УРОВЕНЬ)	3 МАТЕРИАЛОВЕДЕНИЕ (УРОВЕНЬ)
1	мехатроника	низкий	средний
2	мехатроника	низкий	низкий
3	мехатроника	низкий	средний
4	мехатроника	низкий	средний
5	мехатроника	низкий	средний
6	мехатроника	средний	низкий
7	мехатроника	низкий	средний
8	мехатроника	низкий	средний
9	мехатроника	низкий	высокий
10	мехатроника	средний	высокий
11	автосервис	низкий	средний
12	автосервис	низкий	низкий
13	автосервис	средний	средний
14	автосервис	низкий	высокий
15	автосервис	средний	высокий
16	автосервис	средний	низкий
17	автосервис	средний	низкий
18	автосервис	средний	средний
19	автосервис	низкий	низкий
20	автосервис	низкий	средний
21	автосервис	низкий	средний

Рисунок А.12 - Рабочий файл данных TABL.STA

После создания таблицы, содержащей данные педагогического исследования, надо подвести курсор мыши к пункту **Статистики** (Statistics) в строке меню, щелкнуть по нему левой кнопкой мыши. В развернувшемся контекстном меню (рисунок А.13) выбрать с помощью левого клика мыши раздел **Основные статистики и таблицы** (Basic Statistics and Tables).

Рисунок А.13 - Строка меню: пункт **Статистики** (Statistics)

Далее следует выбрать опцию **Таблицы и заголовки** (Tables and banners)

и нажать кнопку  (рисунок А.14).

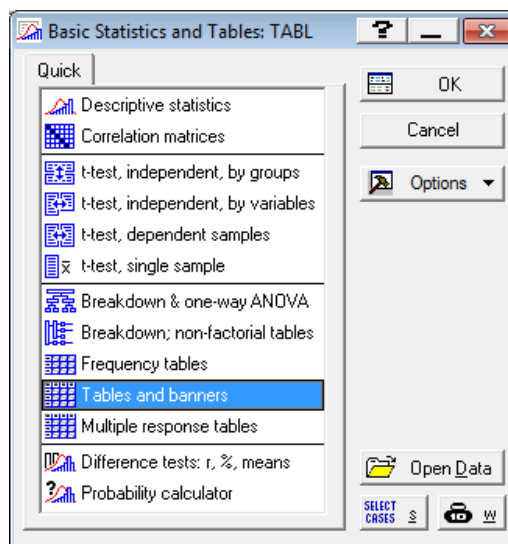


Рисунок А.14 - Окно **Основные статистики и таблицы** (Basic Statistics and Tables)

На экране появится окно **Таблицы кросстабуляции** (Crosstabulation Tables) (рисунок А.15):

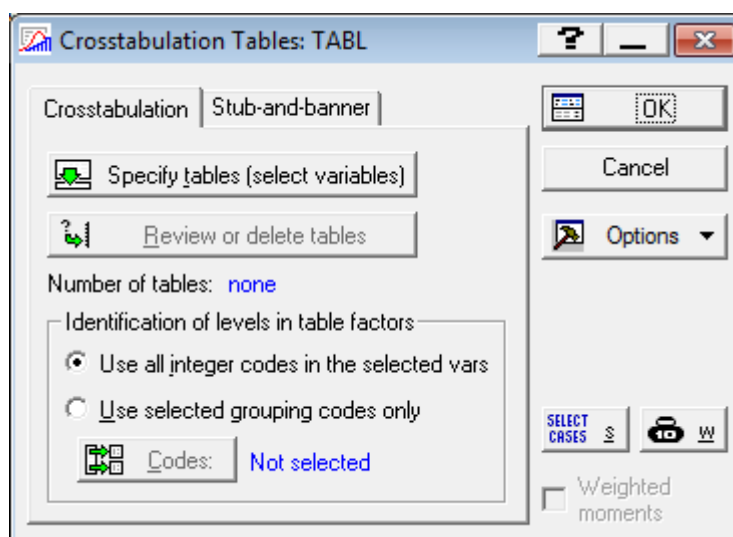
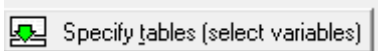


Рисунок А.15 - Окно **Таблицы кросстабуляции** (Crosstabulation Tables)

Здесь следует нажать кнопку **Задать таблицы (выбрать переменные)**



В появившемся окне надо выбрать переменные, которые будут табулированы в таблице. Эти переменные задают разбиение исходных данных на группы, поэтому часто их называют группировочными. Для нашего примера нужно табулировать значения переменных **СПЕЦИАЛЬНОСТЬ** и успеваемость по **УЧЕБНЫМ ДИСЦИПЛИНАМ** (рисунок А.16).

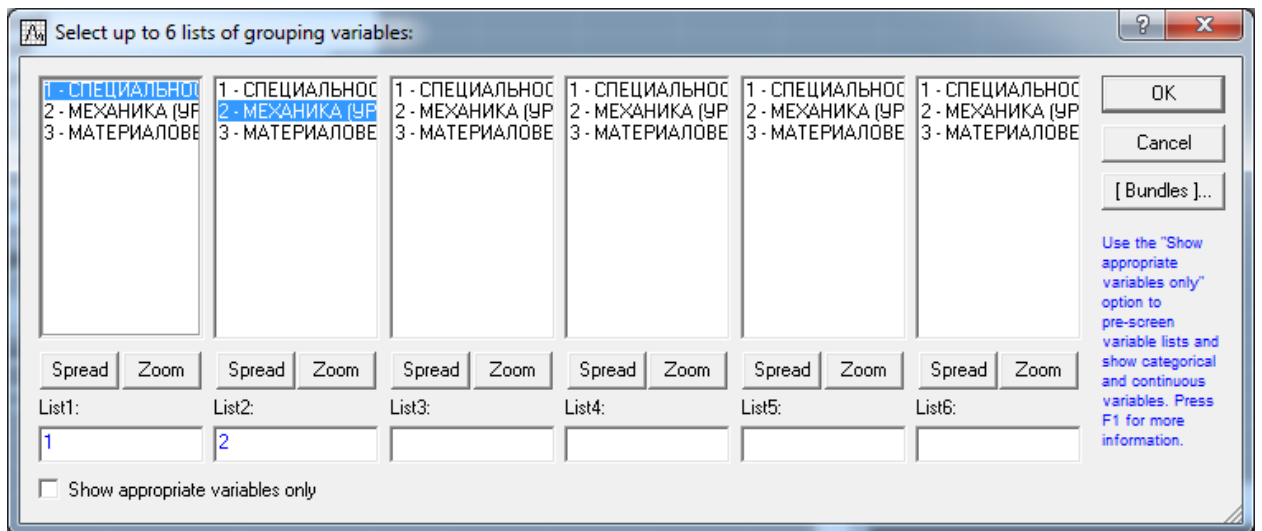




Рисунок А.16 - Окно выбора списков группирующих переменных
(Select up to 6 lists of grouping variables)

Примечание. Можно выбирать до 6 списков группирующих переменных, что, в свою очередь, позволяет построить чрезвычайно сложные таблицы, содержащие гораздо большее число переменных, чем в описываемом примере. Такие таблицы часто возникают при обследовании больших выборок.

После выбора переменных надо нажать кнопку **OK**, и произойдет возврат в диалоговое окно **Таблицы кросстабуляции** (Crosstabulation Tables) (рисунок А.15). Около надписи **Число таблиц** (Number of tables) появится цифра 1, поскольку были выбраны переменные и был запрос системе на построение одной таблицы.

Далее надо нажать клавишу **Enter** на клавиатуре или кнопку  в верхнем правом углу диалогового **Таблицы кросстабуляции** (Crosstabulation Tables) (рисунок А.15).

Система произведет вычисления и предложит посмотреть результат в окне **Результаты кросстабуляции** (Crosstabulation Tables Results) (рисунок А.17) через нажатие кнопки **Суммарный (Итоговый)**  Summary результат.

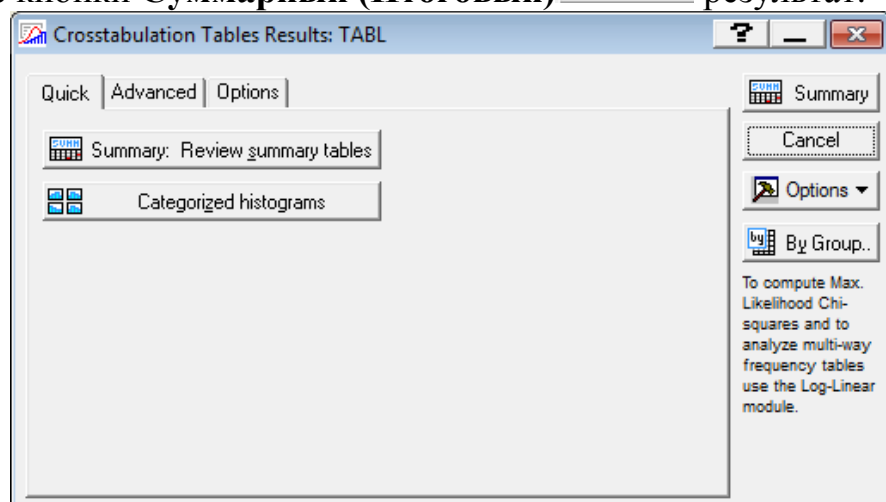


Рисунок А.17 - Окно **Результаты кросстабуляции** (Crosstabulation Tables Results)

На мониторе появится таблица сопряженности (рисунок А.18). В таблице сопряженности табулированы переменные **СПЕЦИАЛЬНОСТЬ** и успеваемость по **УЧЕБНЫМ ДИСЦИПЛИНАМ**: на пересечении строк и столбцов стоят абсолютные (частотные) значения, вычисленные из исходного файла данных **TABL.sta**. Другими словами, совместно табулировались значения двух переменных, **СПЕЦИАЛЬНОСТЬ** и успеваемость по **УЧЕБНЫМ ДИСЦИПЛИНАМ**, и такое действие часто называется кросстабуляцией (от английского cross – пересекать).

Summary Frequency Table (TABL)			
Marked cells have counts > 10 (Marginal summaries are not marked)			
СПЕЦИАЛЬНОСТЬ	МЕХАНИКА (УРОВЕНЬ) низкий	МЕХАНИКА (УРОВЕНЬ) средний	Row Totals
мехатроника	8	2	10
автосервис	6	5	11
All Grps	14	7	21

Рисунок А.18 - Таблица сопряженности (частотные значения)

Как видно из данной таблицы, 8 учащихся, обучающихся по специальности «Мехатроника», имеют низкий уровень успеваемости и 2 — средний. Среди учащихся специальности «Автосервис» 6 имеют низкий уровень успеваемости и 5 — средний. Высокой успеваемости учащихся не выявлено ни по одной из специальностей. Но это только информация по абсолютным величинам. Если же для исследователя важны не только абсолютные, но и относительные величины, то в окне **Результаты кросстабуляции** (Crosstabulation Tables Results) надо выбрать вкладку **Опции (Варианты)** (Options), и далее поставить флажок напротив опции **Проценты от общего числа** (Percentages of total count)(рисунок А.19).

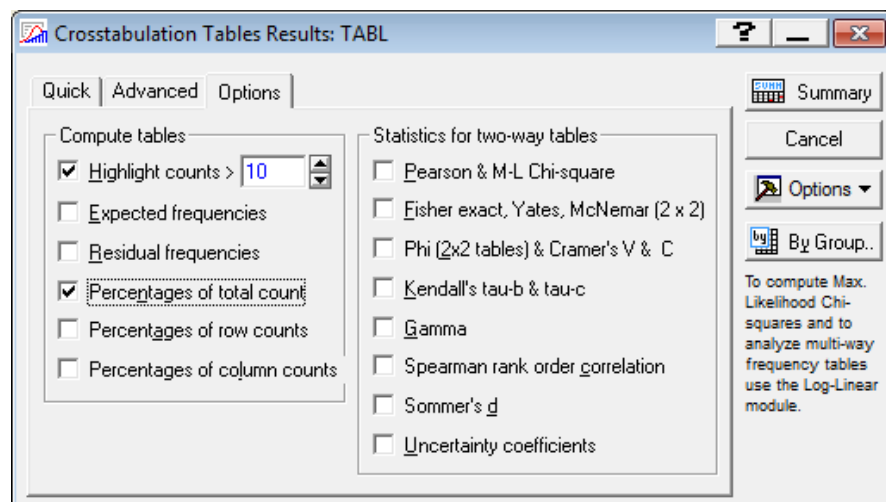



Рисунок А.19 - Окно **Результаты кросстабуляции** (Crosstabulation Tables Results)

Далее нажать кнопку **Суммарный (Итоговый)**  Summary результат. На мониторе отобразится таблица (рисунок А.20), в которой наряду с абсолютными значениями будут приведены относительные величины – проценты, вычисленные от общего числа учащихся, т. е. от 21.

Summary Frequency Table (TABL) Marked cells have counts > 10 (Marginal summaries are not marked)				
	СПЕЦИАЛЬНОСТЬ	МЕХАНИКА (УРОВЕНЬ) низкий	МЕХАНИКА (УРОВЕНЬ) средний	Row Totals
Count	мехатроника	8	2	10
Total Percent		38,10%	9,52%	47,62%
Count	автосервис	6	5	11
Total Percent		28,57%	23,81%	52,38%
Count	All Grps	14	7	21
Total Percent		66,67%	33,33%	

Рисунок А.20 - Вид таблицы сопряженности (абсолютные и относительные значения)

Итак, из таблицы видно, что по группе учащихся, обучающихся по специальности «Мехатроника» 8 человек, что составляет 38,10 % от общей выборки имеют низкий уровень успеваемости и 2 (9,52%) — средний. Среди учащихся специальности «Автосервис» 6 (28,57%) имеют низкий уровень успеваемости и 5 (23,81%) — средний. Высокой успеваемости учащихся не выявлено ни по одной из специальностей.

Построенную таблицу можно отредактировать, изменить ее вид, надписи и т. д., если она импортирована из системы STATISTICA в текстовый процессор MS Word в виде таблицы (через команду **Copy with Headers**).

Таблицы сопряженности позволяют компактно описывать данные. Они удобны и требуют минимум комментариев, поэтому популярны среди педагогов, психологов, социологов, маркетологов, медиков и др.

В системе STATISTICA можно легко построить даже самые сложные таблицы сопряженности. Для нашей таблицы данных мы можем построить таблицу сопряженности, которая будет содержать информацию об учащихся по каждой из специальностей и их успеваемости по обеим дисциплинам.

Покажем, как это выглядит чисто технически, без интерпретации результатов.

После создания таблицы, содержащей данные педагогического исследования (будем использовать данные таблицы TABL.sta), надо подвести курсор мыши к пункту **Статистики** (Statistics) в строке меню, щелкнуть по

нему левой кнопкой мыши. В развернутом контекстном меню (рисунок А.21) выбрать с помощью левого клика мыши раздел **Основные статистики и таблицы** (Basic Statistics and Tables).

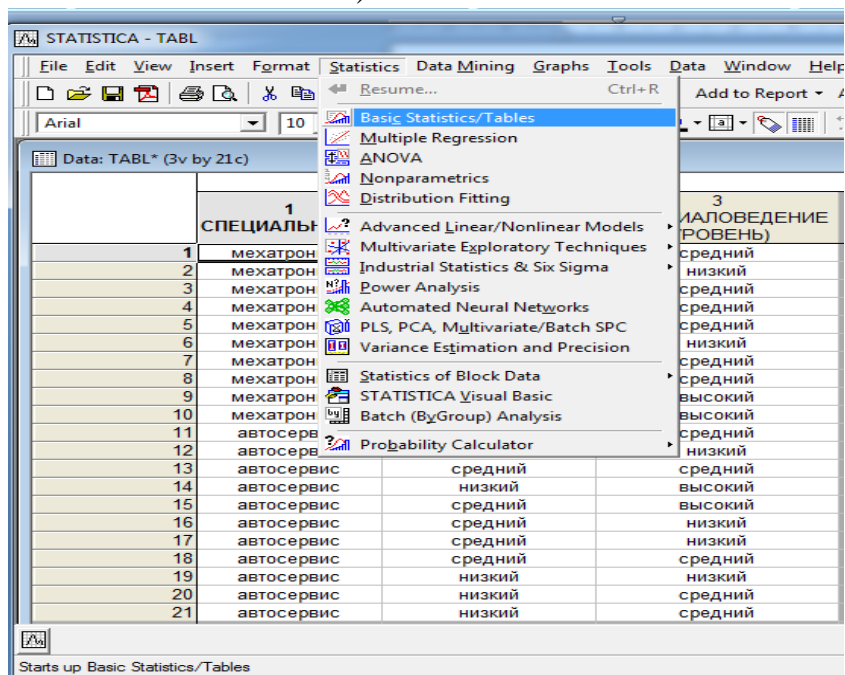


Рисунок А.21 - Строка меню: пункт **Статистики** (Statistics)

Далее следует выбрать опцию **Таблицы и заголовки** (Tables and banners)

и нажать кнопку  (рисунок А.22).

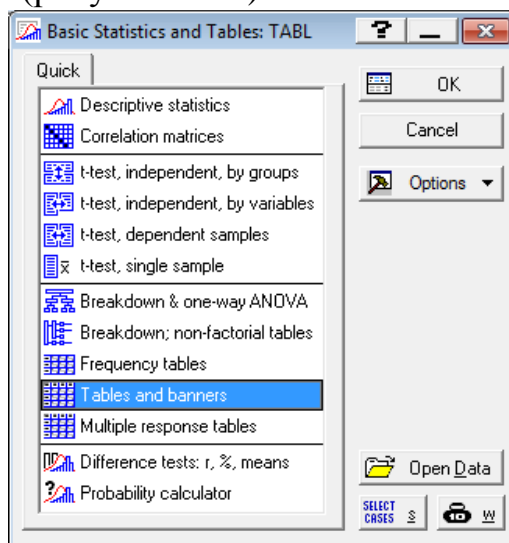


Рисунок А.22 - Окно **Основные статистики и таблицы** (Basic Statistics and Tables)

На экране появится окно **Таблицы кросстабуляции** (Crosstabulation Tables) (рисунок А.23):

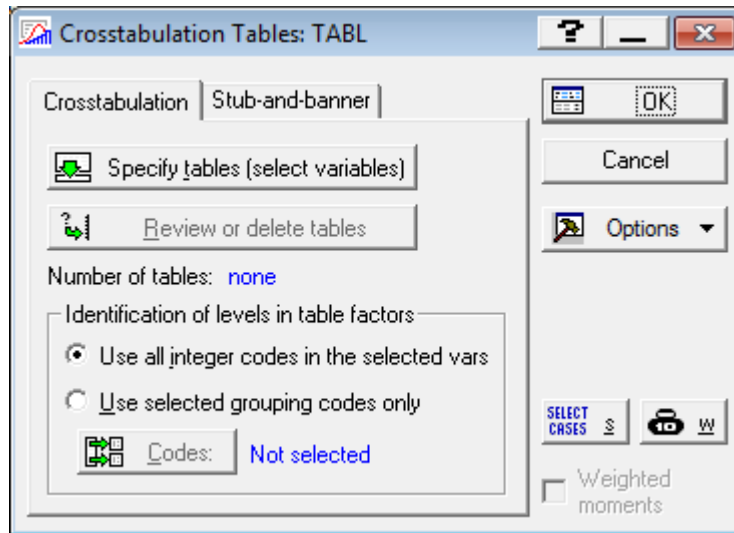
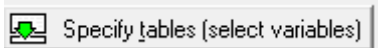


Рисунок А.23 - Окно **Таблицы кросстабуляции** (Crosstabulation Tables)

Здесь следует нажать кнопку **Задать таблицы (выбрать переменные)**



В появившемся окне надо выбрать переменные, которые будут табулированы в таблице в следующем виде (рисунок А.24):

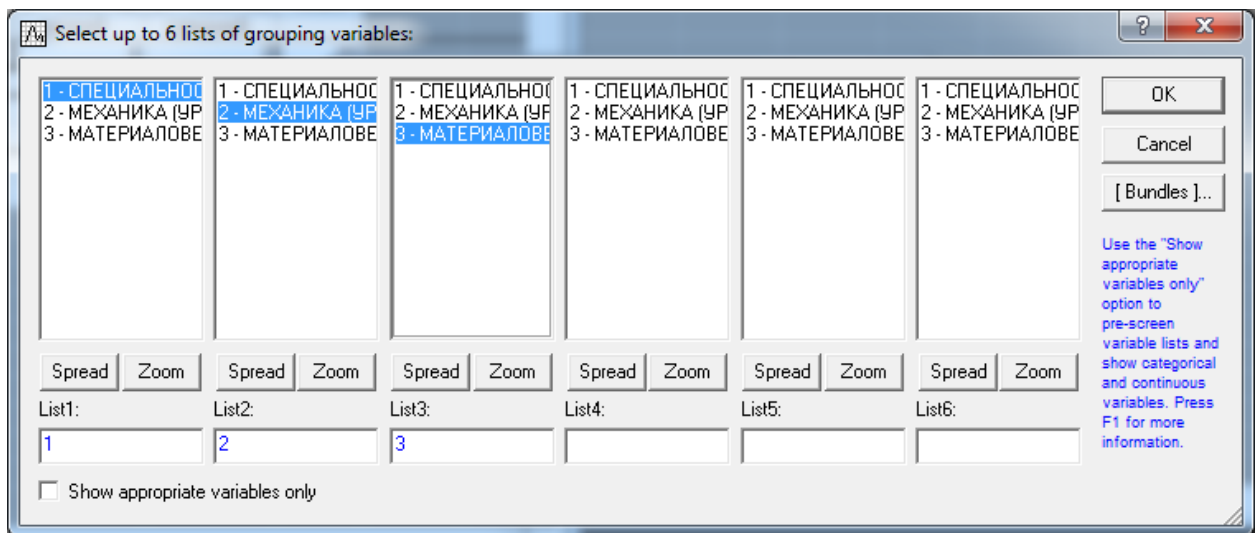



Рисунок А.24 - Окно выбора списков группирующих переменных
(Select up to 6 lists of grouping variables)

После выбора переменных надо нажать кнопку **OK**, и произойдет возврат в диалоговое окно **Таблицы кросстабуляции** (Crosstabulation Tables) (рисунок А.23). Около надписи **Число таблиц** (Number of tables) появится цифра 1, поскольку были выбраны переменные и был запрос системе на построение одной таблицы.

Далее надо нажать клавишу **Enter** на клавиатуре или кнопку **OK** в верхнем правом углу диалогового **Таблицы кросстабуляции** (Crosstabulation Tables) (рисунок А.23).

Система произведет вычисления и предложит посмотреть результат в окне **Результаты кросстабуляции** (Crosstabulation Tables Results) (рисунок А.25) через нажатие кнопки **Суммарный (Итоговый)**  Summary результат.

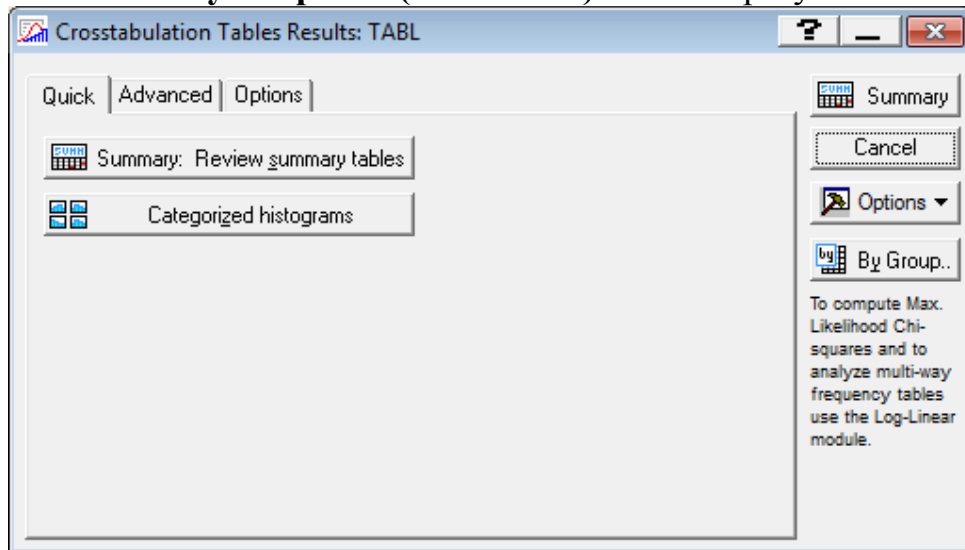


Рисунок А.25 - Окно **Результаты кросстабуляции** (Crosstabulation Tables Results)

Поскольку для нас важны не только абсолютные, но и относительные величины, то в окне **Результаты кросстабуляции** (Crosstabulation Tables Results) надо выбрать вкладку **Опции (Варианты)** (Options), и далее поставить флажок напротив опции **Проценты от общего числа** (Percentages of total count)(рисунок А.26).

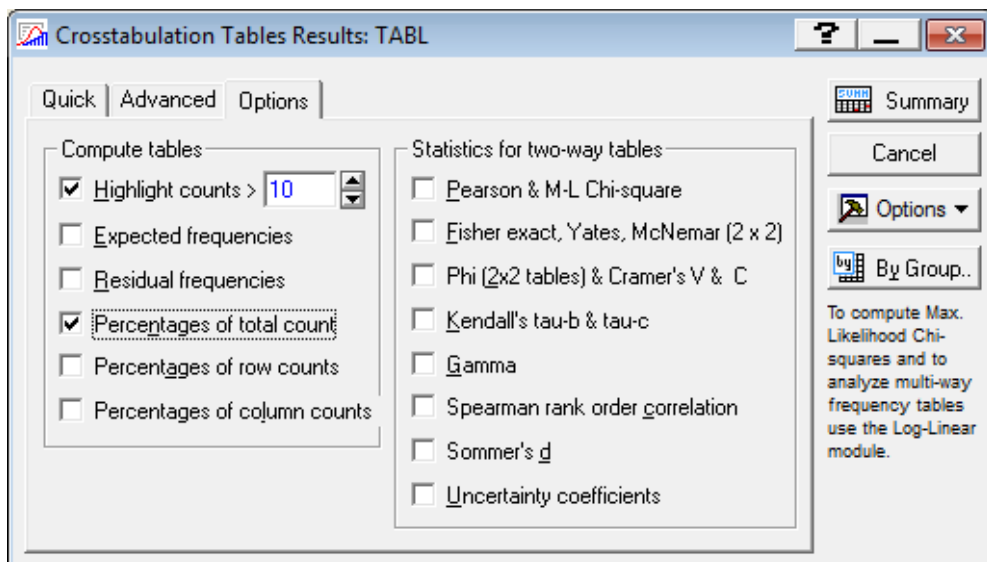



Рисунок А.26 - Окно **Результаты кросстабуляции** (Crosstabulation Tables Results)

Далее нажать кнопку **Суммарный (Итоговый)**  Summary результат. На мониторе отобразится таблица (рисунок А.27), в которой наряду с абсолютными значениями будут приведены относительные величины.

Summary Frequency Table (TABL) Marked cells have counts > 10 (Marginal summaries are not marked)						
СПЕЦИАЛЬНОСТЬ	МЕХАНИКА (УРОВЕНЬ)	МАТЕРИАЛОВЕДЕНИЕ (УРОВЕНЬ) низкий	МАТЕРИАЛОВЕДЕНИЕ (УРОВЕНЬ) средний	МАТЕРИАЛОВЕДЕНИЕ (УРОВЕНЬ) высокий	Row Totals	
Count	мехатроника	низкий	1	6	1	8
Total Percent			4,76%	28,57%	4,76%	38,10%
Count	мехатроника	средний	1	0	1	2
Total Percent			4,76%	0,00%	4,76%	9,52%
Count	Total		2	6	2	10
Total Percent			9,52%	28,57%	9,52%	47,62%
Count	автосервис	низкий	2	3	1	6
Total Percent			9,52%	14,29%	4,76%	28,57%
Count	автосервис	средний	2	2	1	5
Total Percent			9,52%	9,52%	4,76%	23,81%
Count	Total		4	5	2	11
Total Percent			19,05%	23,81%	9,52%	52,38%
Count	Column Total		6	11	4	21
Total Percent			28,57%	52,38%	19,05%	

Рисунок А.27 - Вид таблицы сопряженности (абсолютные и относительные значения)

Результаты, представленные в таблице рисунка А.27 следует проинтерпретировать.

- **Графический анализ таблиц сопряженности**


В данном подразделе будет рассмотрено, как можно визуализировать построенные таблицы, т. е. как использовать средства системы STATISTICA, позволяющие графически проанализировать таблицы данных. Очевиден тот факт, что визуально гораздо проще наблюдать какие-либо закономерности, содержащиеся в таблицах.

Примечание. В рассмотренных ниже примерах будут использоваться данные небольшого объема, чтобы можно было более четко представить основные приемы работы в системе STATISTICA. Таким образом, естественно, что некоторые требования к использованию статистических критериев не будут соблюдены.

Предположим, что система STATISTICA запущена на компьютере, и мы работаем в модуле **Основные статистики** (Basic Statistics) с файлом **TABL.STA** (рисунок А.12). Для данных таблицы осуществлена процедура кросстабуляции и построена итоговая таблица, содержащая абсолютные и относительные величины (рисунок А.28):

Summary Frequency Table (TABL) Marked cells have counts > 10 (Marginal summaries are not marked)				
СПЕЦИАЛЬНОСТЬ	МЕХАНИКА (УРОВЕНЬ) низкий	МЕХАНИКА (УРОВЕНЬ) средний	Row Totals	
Count	мехатроника	8	2	10
Total Percent		38,10%	9,52%	47,62%
Count	автосервис	6	5	11
Total Percent		28,57%	23,81%	52,38%
Count	All Grps	14	7	21
Total Percent		66,67%	33,33%	

Рисунок А.28 - Вид таблицы сопряженности (абсолютные и относительные значения)

Построим для полученных данных (только для абсолютных величин) категоризованную гистограмму. Нажимаем кнопку **Таблица Кросстабуляции**  **Crosstabulation Table...** в левом нижнем углу монитора (рисунок А.29):

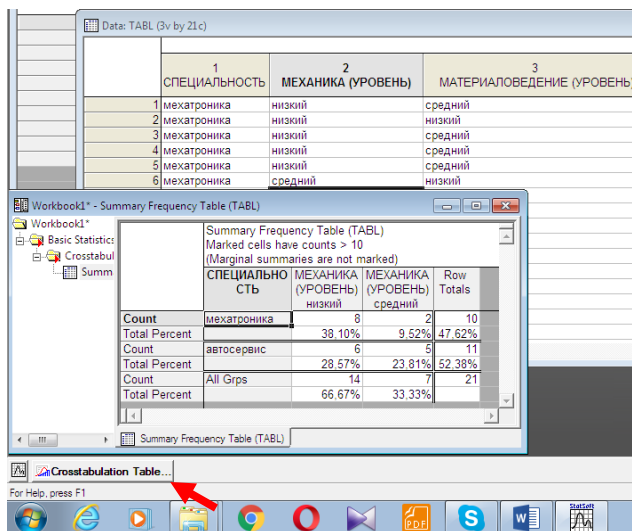


Рисунок А.29 - Окно, содержащее таблицу сопряженности и кнопку, позволяющую продолжить вычисления

Возвращаемся в окно **Результаты кросстабуляции** (Crosstabulation Tables Results) (рисунок А.30):

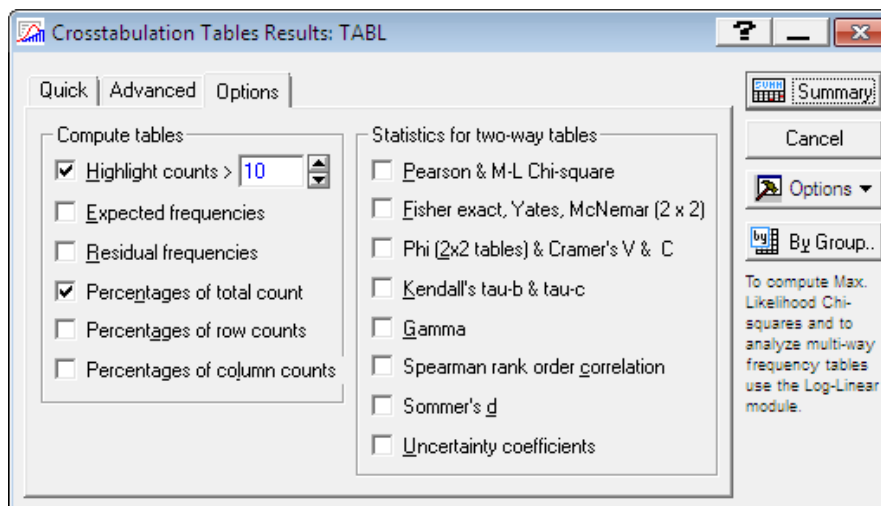



Рисунок А.30 - Окно **Результаты кросстабуляции** (Crosstabulation Tables Results)

Затем выбираем вкладку **Расширенные (Дополнительные)(Advanced) возможности** и нажимаем кнопку **Категоризованные гистограммы**  **Categorized histograms** (рисунок А.31).

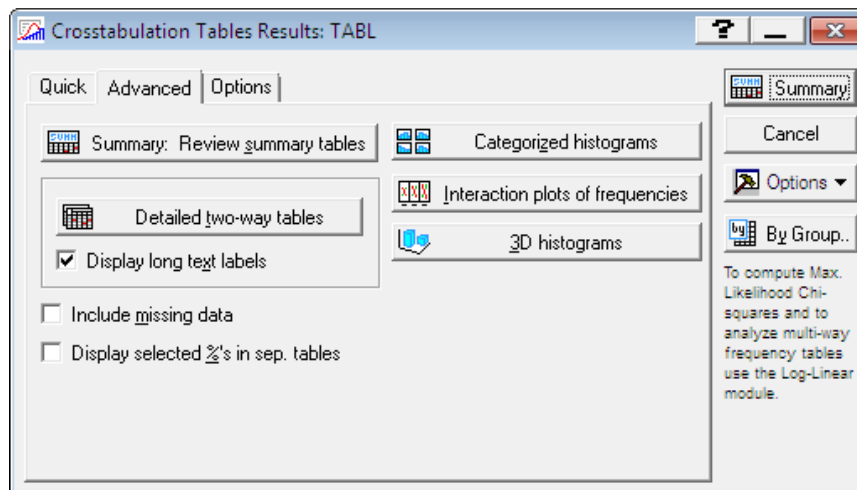


Рисунок А.31 - Окно **Результаты кросстабуляции** (Crosstabulation Tables Results)

На монитор будет выведена категоризованная гистограмма (рисунок А.32).

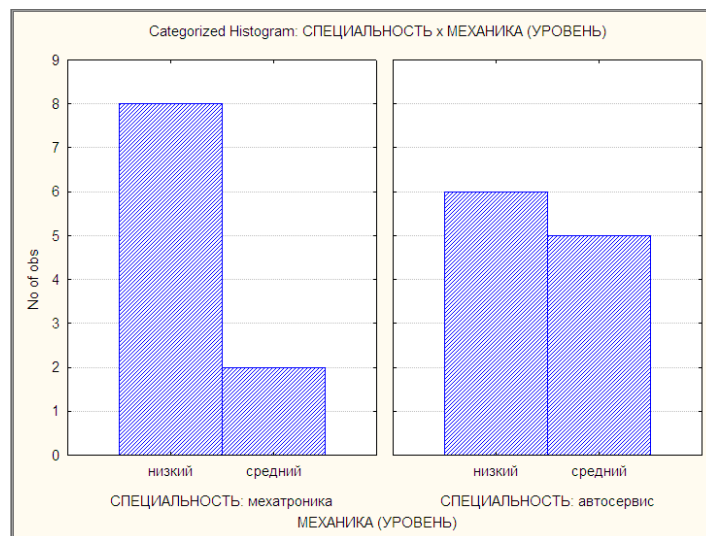



Рисунок А.32 - Категоризованная гистограмма

Как видно из гистограмм, учащиеся разбиты на 2 группы, или категории — по специальностям: мехатроника и автосервис и уровням успеваемости по предмету: механика. Для каждой группы построена отдельная гистограмма, и все они собраны вместе на одном графике, что, в свою очередь, позволяет визуально сравнить группы.

Построение трехмерных гистограмм. В диалоговом окне **Результаты кросстабуляции** (Crosstabulation Tables Results) надо нажать кнопку **3-мерные гистограммы**  **3D histograms** (рисунок А.31).

На мониторе отобразится трехмерная гистограмма (рисунок А.33).

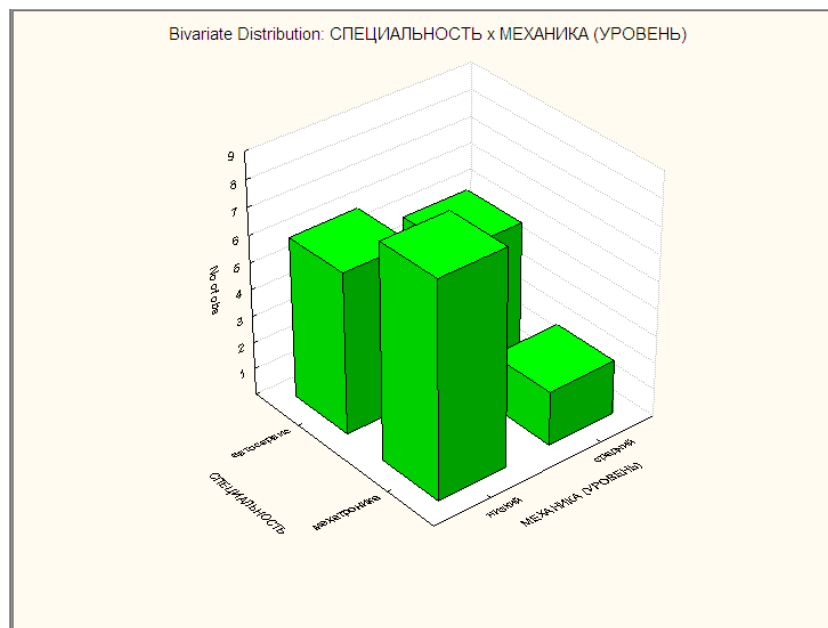



Рисунок А.33 - Трехмерная гистограмма

Трехмерная гистограмма представляет собой всевозможные комбинации значений двух переменных: специальность и уровень успеваемости по предмету.

Трехмерная гистограмма очень наглядно воспроизводит построенную таблицу кросстабуляции (рисунок А.28). Таблица как бы положена на плоскость, и в каждую клетку поставлено по столбцу, высота которого равна количеству наблюдений в клетке таблицы.

Примечание. Если перспектива построенной трехмерной гистограммы не совсем удовлетворительна, то ее можно изменить, воспользовавшись средствами системы STATISTICA. Например, гистограмму можно повернуть.

Для этого надо поставить курсор на область построенной гистограммы и нажать дважды левую кнопку мыши. Развернется окно, где следует выбрать закладку **Точка Наблюдения** (Point of View) и далее с помощью кнопки **Вращение**  подобрать нужную перспективу, например, как на рисунке А.34.

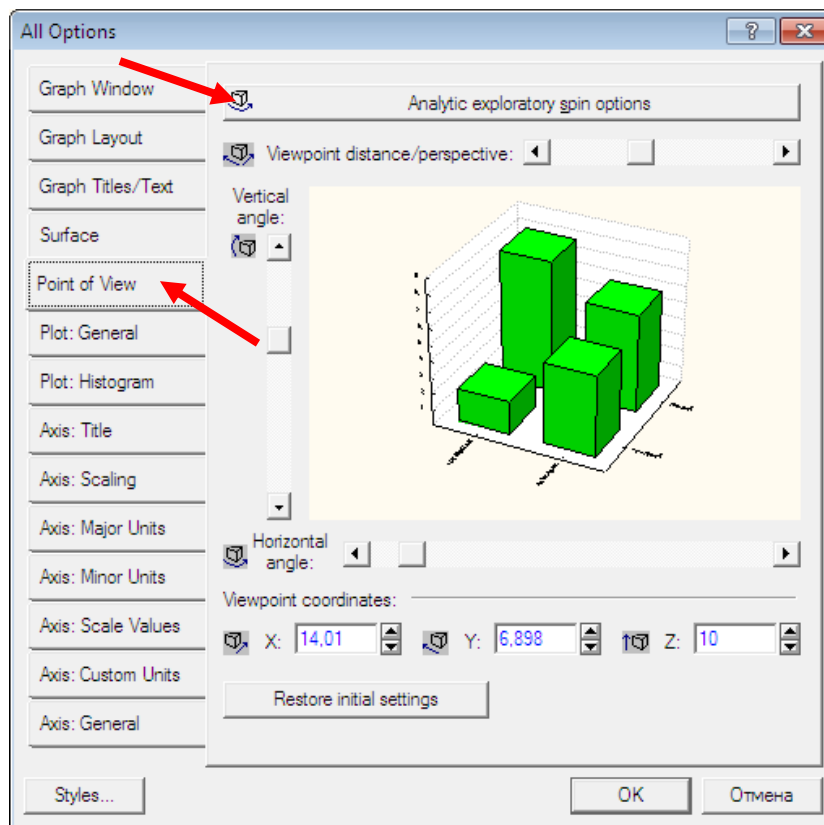


Рисунок А.34 - Окно **Все опции (Варианты)** (All Options)

Для вращения графика используется также и линейка прокрутки. С помощью мыши можно сдвинуть курсор прокрутки в любое положение. Результатом данного действия после нажатия кнопки **OK** будет, например, гистограмма, изображенная на рисунке А.35:

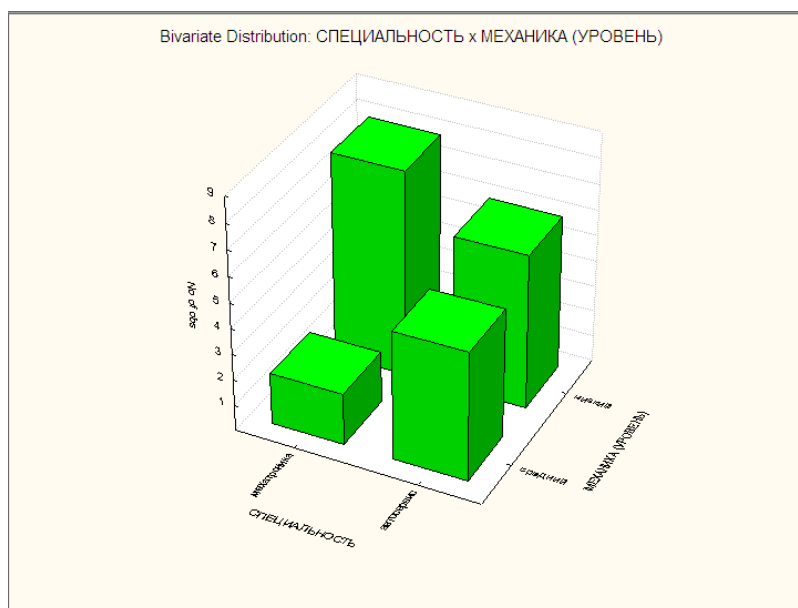



Рисунок А.35 - Трехмерная гистограмма

Построение графиков взаимодействий частот. В диалоговом окне **Результаты кросстабуляции (Crosstabulation Tables Results)** надо нажать кнопку **Графики взаимодействий частот**  **Interaction plots of frequencies** (рисунок А.31). На мониторе появится график взаимодействий частот (рисунок А.36):

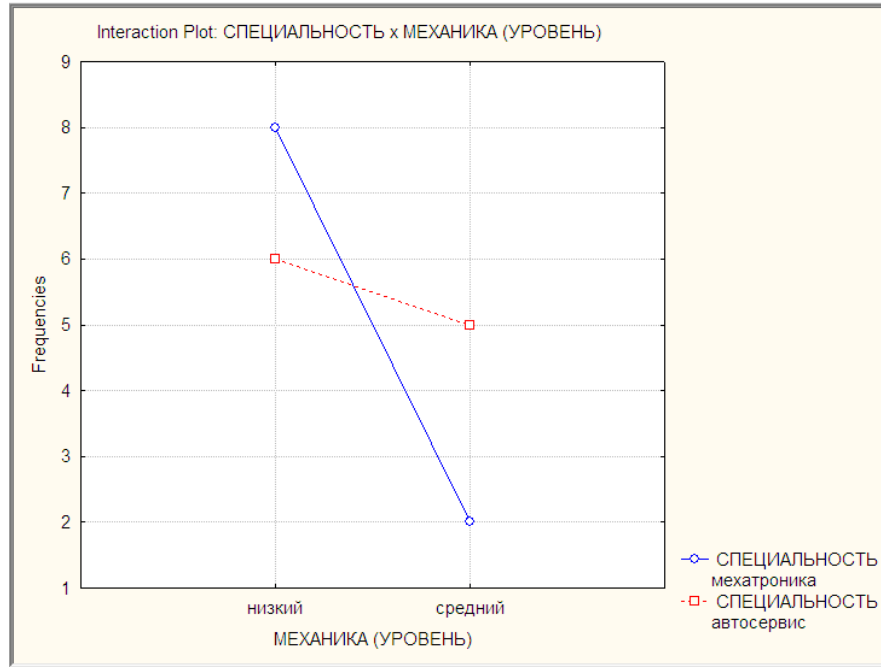


Рисунок А.36 - Графики взаимодействий частот

Изображенные графики показывают, как взаимодействуют или как связаны между собой частоты наблюдений из разных групп. Видно, что количество учащихся, обучающихся по разным специальностям, отличается по уровню успеваемости. Является ли это различие значимым, могут показать статистические тесты (например, ХИ-квадрат критерий).

Инструкция по выполнению практической работы 2

«Определение различий между эмпирической и теоретической таблицами распределений»

Вычисление ХИ-квадрат-критерия в системе STATISTICA

Вычисление ХИ-квадрат-критерия в системе STATISTICA будем осуществлять, используя исходные данные файла **TABL.STA** (рисунок Б.1). В качестве рабочей статистической гипотезы может быть сформулировано следующее предположение: существуют различия в показателях успеваемости по дисциплине «Механика» («Материаловедение») у учащихся, обучающихся по специальностям «Мехатроника» и «Автосервис».

TABL* (3v by 21c)			
	1 СПЕЦИАЛЬНОСТЬ	2 МЕХАНИКА (УРОВЕНЬ)	3 МАТЕРИАЛОВЕДЕНИЕ (УРОВЕНЬ)
1	мехатроника	низкий	средний
2	мехатроника	низкий	низкий
3	мехатроника	низкий	средний
4	мехатроника	низкий	средний
5	мехатроника	низкий	средний
6	мехатроника	средний	низкий
7	мехатроника	низкий	средний
8	мехатроника	низкий	средний
9	мехатроника	низкий	высокий
10	мехатроника	средний	высокий
11	автосервис	низкий	средний
12	автосервис	низкий	низкий
13	автосервис	средний	средний
14	автосервис	низкий	высокий
15	автосервис	средний	высокий
16	автосервис	средний	низкий
17	автосервис	средний	низкий
18	автосервис	средний	средний
19	автосервис	низкий	низкий
20	автосервис	низкий	средний
21	автосервис	низкий	средний

Рисунок Б.1 - Рабочий файл данных **TABL.STA**

После создания таблицы, содержащей данные педагогического исследования, надо подвести курсор мыши к пункту **Статистики** (Statistics) в строке меню, щелкнуть по нему левой кнопкой мыши. В развернувшемся контекстном меню (рисунок Б.2) выбрать с помощью левого клика мыши раздел **Основные статистики и таблицы** (Basic Statistics and Tables).

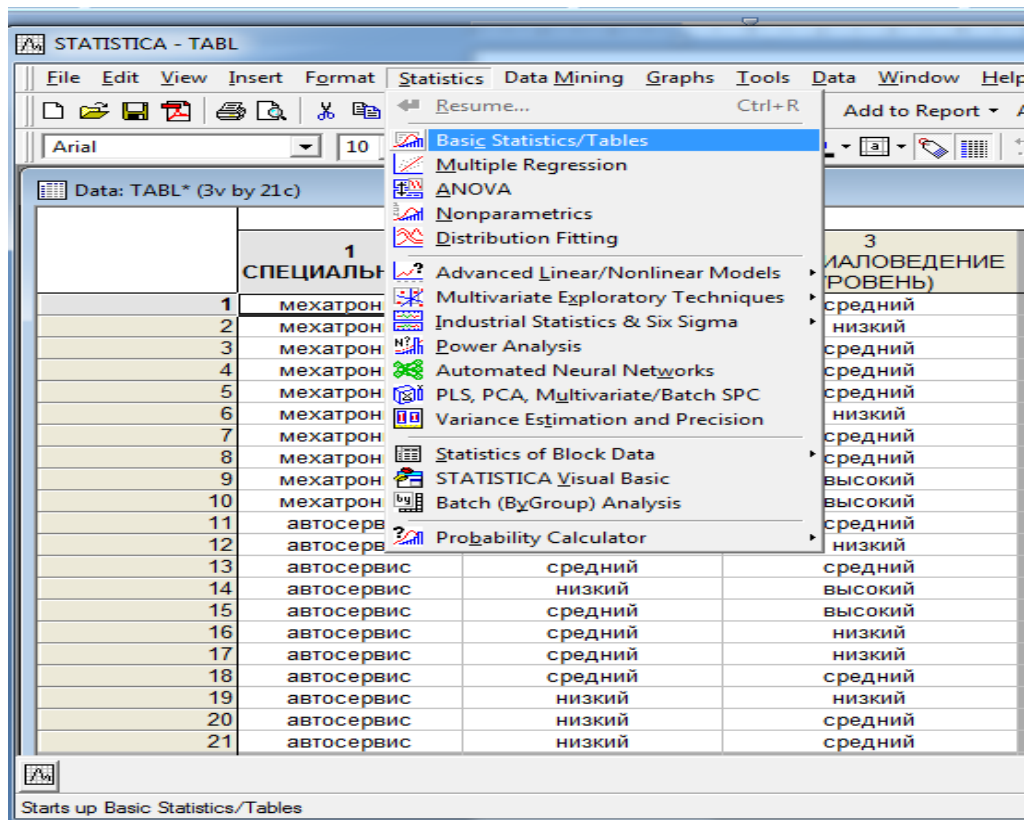


Рисунок Б.2 - Строка меню: пункт **Статистики** (Statistics)

Далее следует выбрать опцию **Таблицы и заголовки** (Tables and banners)

и нажать кнопку  (рисунок Б.3).

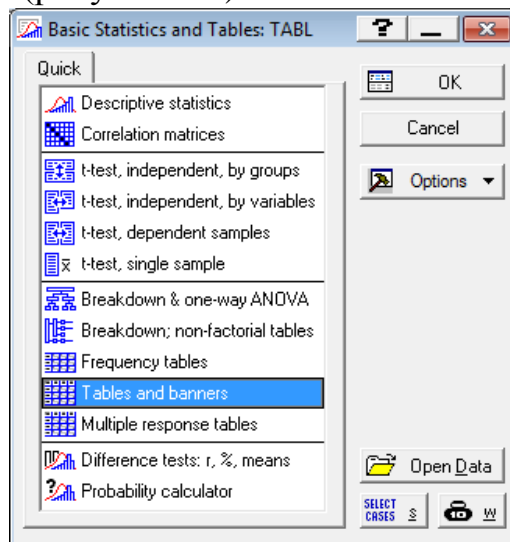


Рисунок Б.3 - Окно **Основные статистики и таблицы** (Basic Statistics and Tables)

На экране появится окно **Таблицы кросстабуляции** (Crosstabulation Tables) (рисунок Б.4):

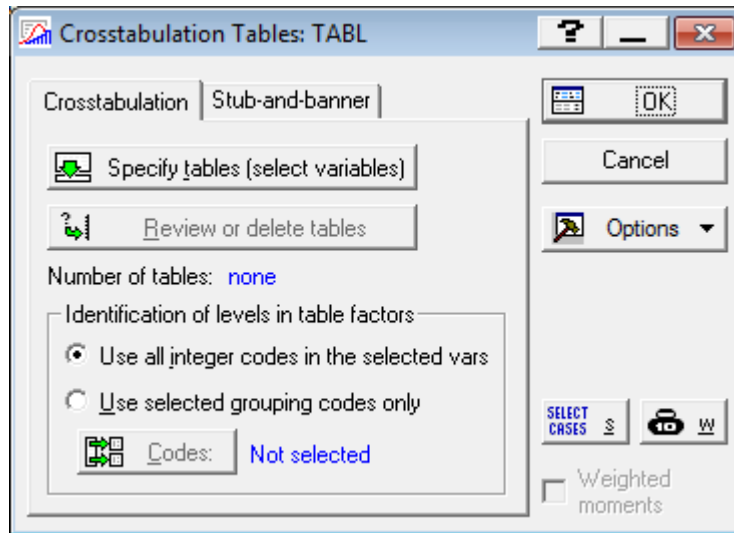
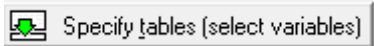


Рисунок Б.4 - Окно **Таблицы кросстабуляции** (Crosstabulation Tables)

Здесь следует нажать кнопку **Задать таблицы (выбрать переменные)**



В появившемся окне надо выбрать переменные, которые будут подвергаться обработке посредством Хи-квадрат критерия.

Для нашего примера нужно выбрать переменные **СПЕЦИАЛЬНОСТЬ** и **успеваемость по УЧЕБНЫМ ДИСЦИПЛИНАМ** (рисунок Б.5).

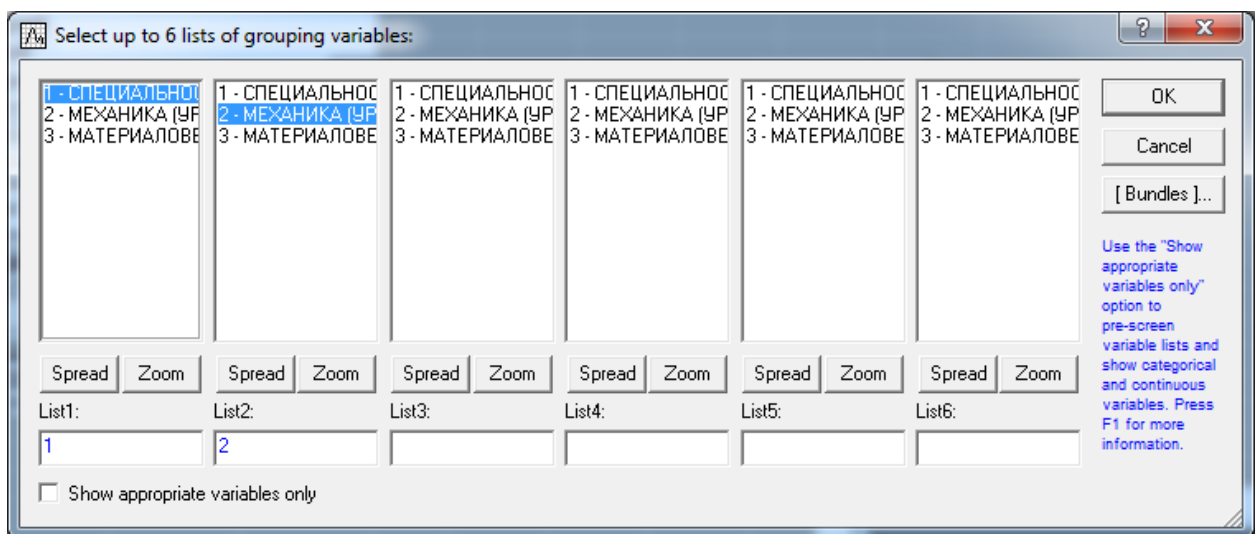



Рисунок Б.5 - Окно выбора списков группирующих переменных
(Select up to 6 lists of grouping variables)

После выбора переменных надо нажать кнопку **OK**, и произойдет возврат в диалоговое окно **Таблицы кросстабуляции** (Crosstabulation Tables) (рисунок Б.4). Около надписи **Число таблиц** (Number of tables) появится цифра 1.

Далее надо нажать клавишу **Enter** на клавиатуре или кнопку **OK** в верхнем правом углу диалогового **Таблицы кросстабуляции** (Crosstabulation Tables) (рисунок Б.4).

Система произведет вычисления и предложит посмотреть результат в окне **Результаты кростабуляции** (Crosstabulation Tables Results) (рисунок Б.6) через нажатие кнопки **Суммарный (Итоговый)**  Summary результат.

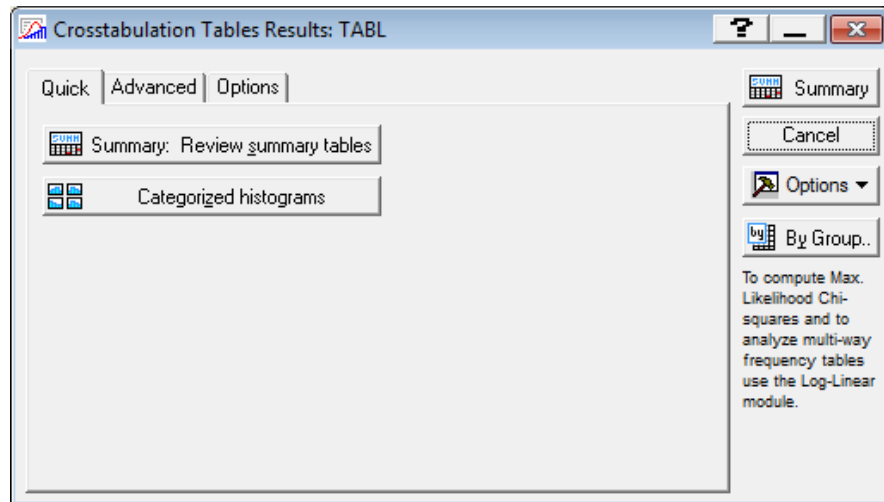


Рисунок Б.6 - Окно **Результаты кростабуляции** (Crosstabulation Tables Results)

Поскольку необходимо вычислить ХИ-квадрат критерий, следует выбрать вкладку **Опции (Варианты)** (Options), и далее поставить флажок напротив опции **Критерий Фи (таблицы 2x2)...** (Phi (2x2 tables)&Cramer's V&C) (рисунок Б.7).

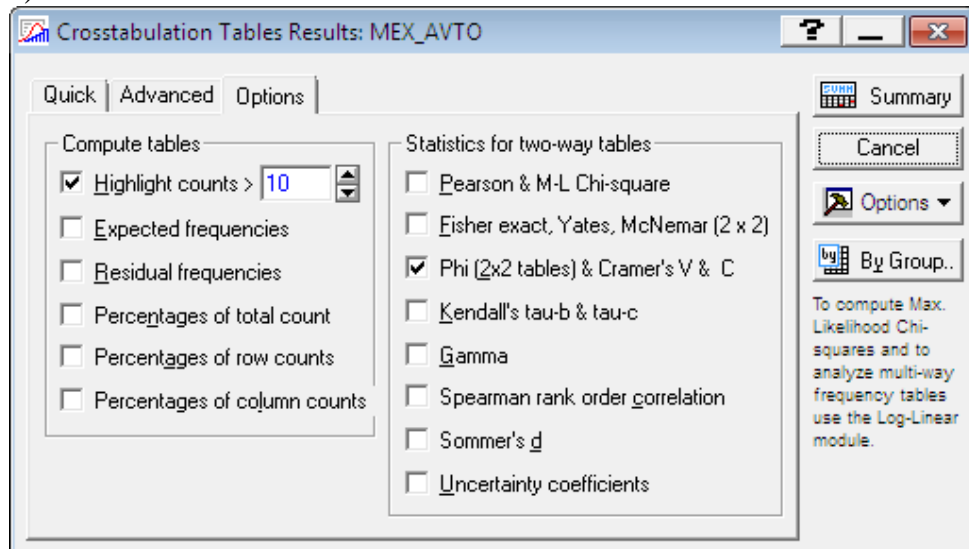



Рисунок Б.7 - Окно **Результаты кростабуляции** (Crosstabulation Tables Results)

Затем выбираем вкладку **Расширенные (Дополнительные)** (Advanced) **возможности** и нажимаем кнопку **Подробные таблицы 2x2**  Detailed two-way tables (рисунок Б.8).

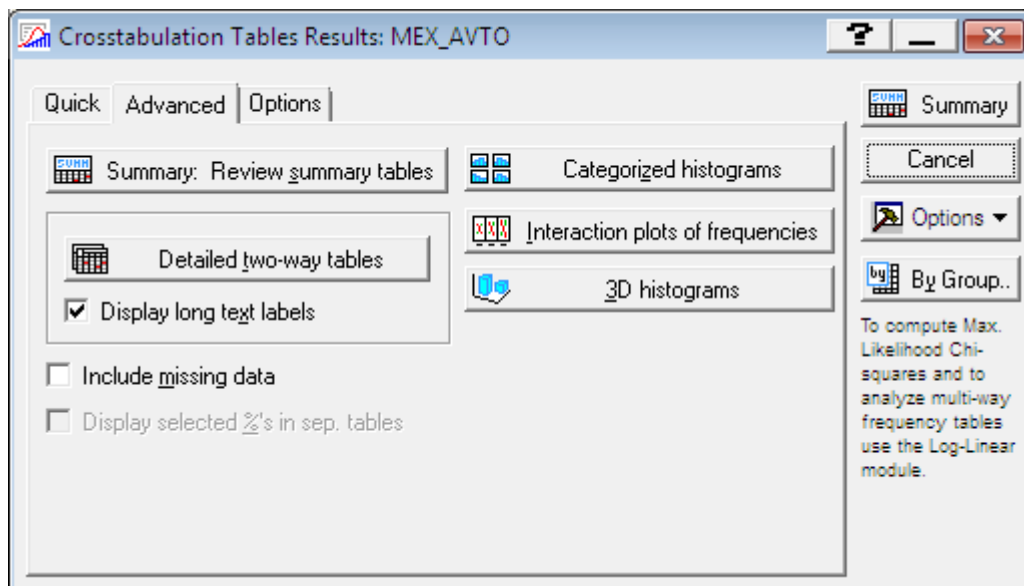



Рисунок Б.8 - Окно **Результаты кросстабуляции** (Crosstabulation Tables Results)

Через нажатие кнопки **Суммарный (Итоговый)**  Summary результат на мониторе появится таблица, содержащая результаты вычислений по ряду критериев, в том числе и по критерию ХИ-квадрат (Pearson Chi-square) (рисунок Б.9):

Statistics: СПЕЦИАЛЬНОСТЬ(2) x МЕХАНИКА (УРОВЕНЬ)(3)			
Statistic	Chi-square	df	p
Pearson Chi-square	6,239256	df=2	p=.04418
M-L Chi-square	7,176225	df=2	p=.02765
Phi	,5450757		
Contingency coefficient	,4785957		
Cramer's V	,5450757		

Рисунок Б.9 – Результаты ХИ-квадрат критерия Пирсона

Как видно из таблицы Б.9, р-уровень равен 0,04418. Таким образом, можно сделать вывод о том, что существуют различия (на уровне средней связи, т.к. коэффициент контингенции (Contingency coefficient) $\approx 0,05$) в показателях успеваемости по дисциплине «Механика» у учащихся, обучающихся по специальностям «Мехатроника» и «Автосервис» ($\chi^2 \approx 6,24$; $p < 0,05$).

Аналогичным образом можно провести вычисления и доказать статистическую гипотезу для тех же групп испытуемых с учетом их успеваемости по дисциплине «Материаловедение».

Инструкция по выполнению практической работы 3

«Описательные статистики»

Вычисление описательных статистик в системе STATISTICA

После создания таблицы, содержащей данные педагогического исследования (успеваемость учащихся специальности «Мехатроника» по учебным дисциплинам «Механика» и «Материаловедение» в БАЛЛАХ!), надо подвести курсор мыши к пункту **Статистики** (Statistics) в строке меню, щелкнуть по нему левой кнопкой мыши. В развернувшемся контекстном меню (рисунок В.1) выбрать с помощью левого клика мыши раздел **Основные статистики и таблицы** (Basic Statistics and Tables).

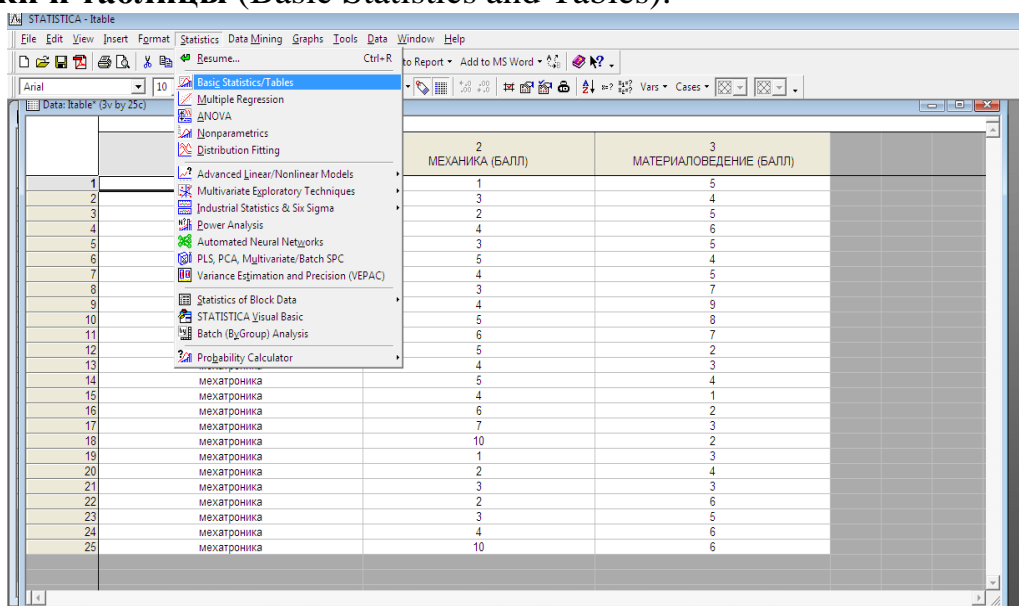



Рисунок В.1 - Строка меню: пункт Статистики (Statistics)

Далее следует выбрать опцию **Описательные статистики** (Descriptive Statistics) и нажать кнопку  (рисунок В.2).

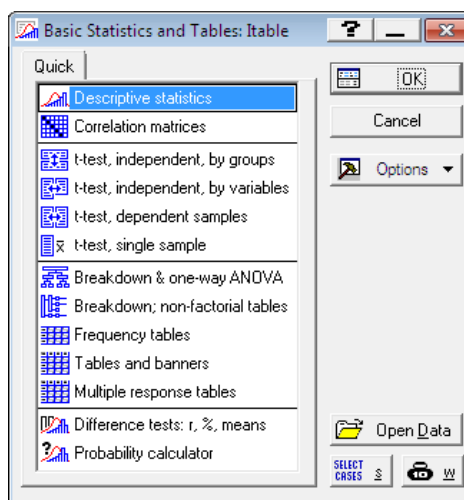


Рисунок В.2 - Окно **Основные статистики и таблицы** (Basic Statistics and Tables)

РИПО

В открывшемся диалоговом окне **Описательные статистики** (Descriptive Statistics) (рисунок В.3) надо нажать кнопку **Переменные**

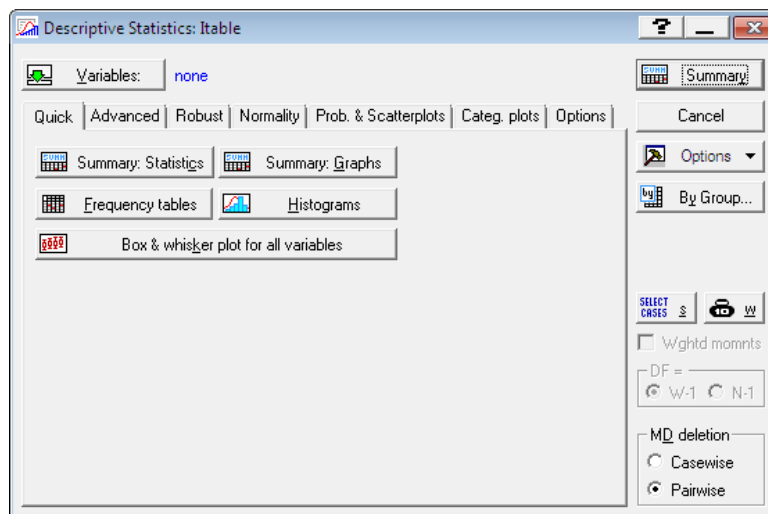



Рисунок В.3 - Окно **Описательные статистики** (Descriptive Statistics)

Далее выбрать переменные для анализа (рисунок В.4) и нажать кнопку  (рисунок В.4).

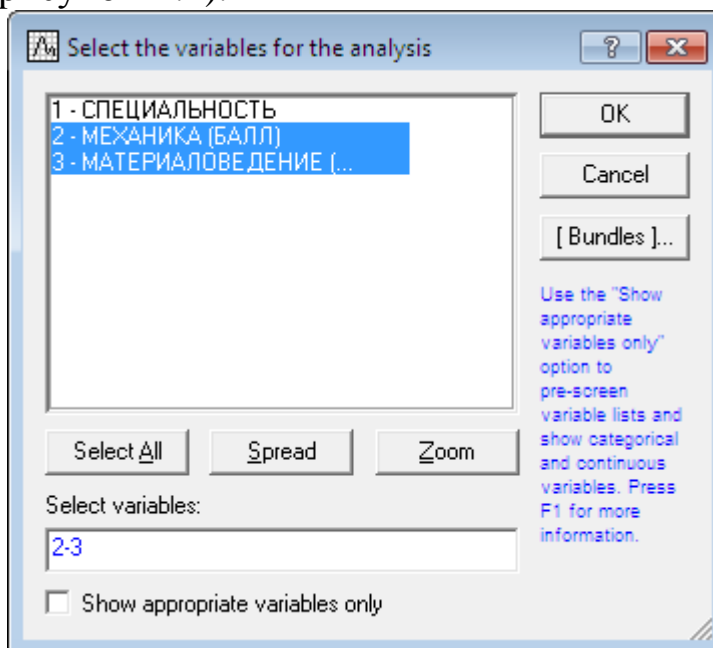



Рисунок В.4 - Окно **Выбор переменных для анализа** (Select the variables for the analysis)

После нажатия кнопки  в окне **Выбор переменных для анализа** (Select the variables for the analysis) снова открывается диалоговое окно **Описательные статистики** (Descriptive Statistics). Затем выбираем вкладку **Расширенные (Дополнительные)** (Advanced) **возможности** и ставим флажки напротив, например, следующих опций

(рисунок В.5):

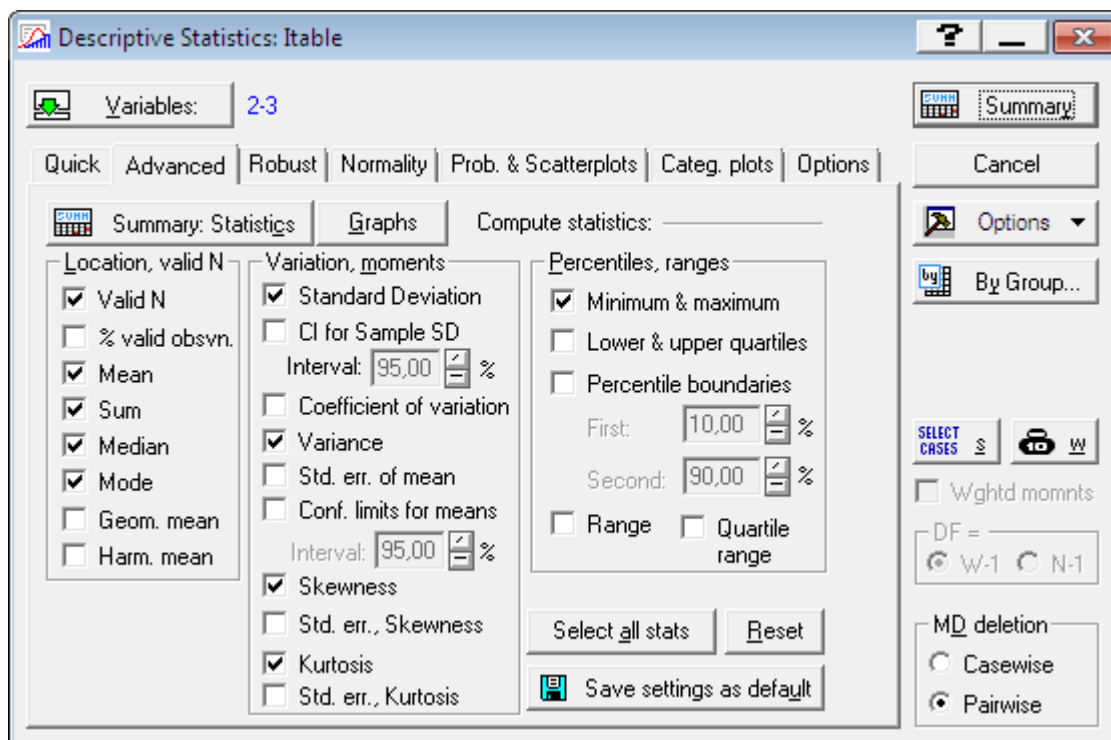


Рисунок В.5 - Окно **Описательные статистики (Descriptive Statistics)**

Через нажатие кнопки **Суммарный (Итоговый)** Summary результат на мониторе появится таблица, содержащая результаты вычислений по ряду описательных статистик (рисунок В.6.):

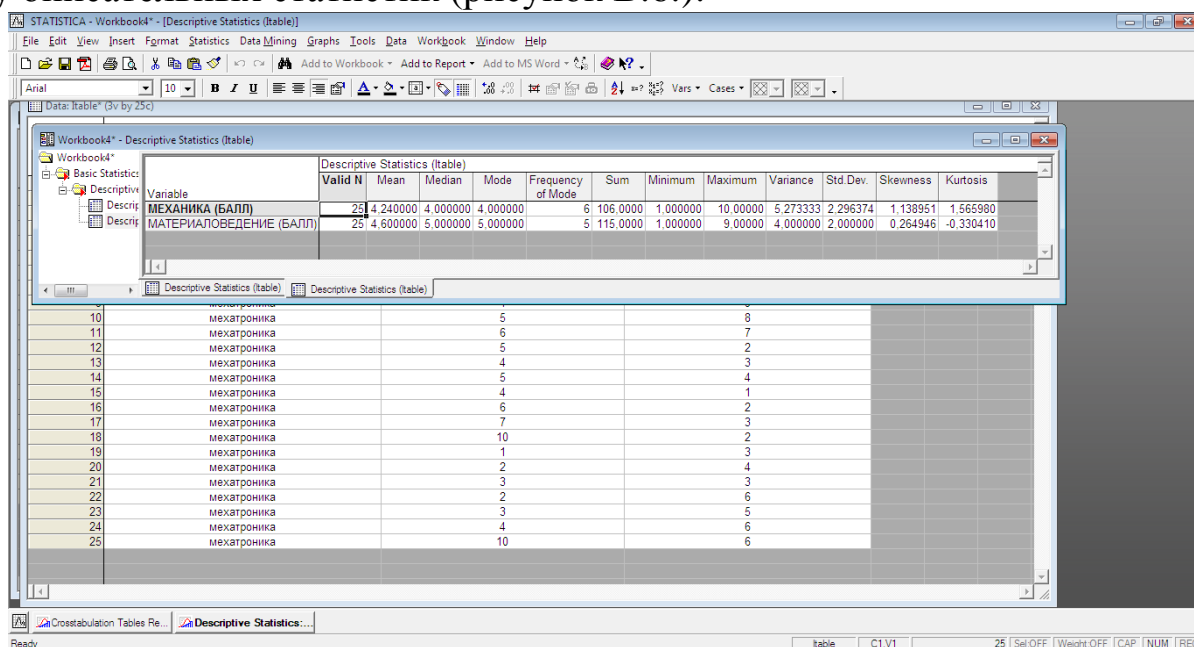


Рисунок В.6 - Результаты вычисления описательных статистик

Итак, Valid N — объем выборки, который равен 25. Это количество учащихся, обучающихся по специальности «Мехатроника».

Mean — среднее значение, и средний балл успеваемости по

дисциплине «Механика» равен 4,24, что меньше, чем по дисциплине «Материаловедение» — 4,6.

Median — медиана – это значение признака, которое находится в середине числового ряда (баллы от 1 до 10) и разбивает выборку на две равные части (1-5 и 6-10). Значение медианы по дисциплине «Механика» – 4, а по дисциплине «Материаловедение» — 5. Так как значение медианы отличается от среднего, то выборочное распределение может быть скошено.

Mode — мода – наиболее часто встречающееся в ряду значение признака (балла) равна по дисциплине «Механика» 4 и частота ее встречаемости (Frequency of Mode) равна 6, а по дисциплине «Материаловедение» — соответственно 5 и 5.

Мода может быть также и множественной (Multiple). Тогда в таком случае говорят, что выборочное распределение мультимодально.

Sum — сумма всех выборочных значений. Суммы баллов по дисциплине «Механика» (106) и по дисциплине «Материаловедение» (115), в принципе, не представляют для данного исследования никакого интереса и их можно было бы не вычислять.

Minimum — минимальное значение. Минимальный балл по обоим дисциплинам равен 1.

Maximum — максимальное значение. Максимальный балл по дисциплине «Механика» – 10, а по дисциплине «Материаловедение» — 9.

Variance – дисперсия, т.е. рассеивание значений, характеризует степень разброса данных относительно своего центра. По дисциплине «Механика» дисперсия значений больше и $\approx 3,27$, чем по дисциплине «Материаловедение», которая равна 4,00.

Std. Dev. — стандартное отклонение, наиболее часто используемая мера изменчивости (вариации) данных, позволяет оценить, насколько значения могут отличаться от среднего значения. Нормальное распределение обладает важной особенностью: чем больше величина признака отклоняется от среднего значения, тем меньше будет частота встречаемости (вероятность) этого признака в распределении. По дисциплине «Механика» отклонение составило $\approx 2,30$, а по дисциплине «Материаловедение» — 2,00, т.е. величина признака находится примерно посередине относительно среднего значения.

Skewness – асимметрия (косость, скошенность) — это характеристика формы распределения. Распределение будет скошено влево, если значение асимметрии отрицательно, вправо, если асимметрия положительна. Асимметрия стандартного нормального распределения равна 0. По

дисциплине «Механика» асимметрия составила $\approx 1,14$, а по дисциплине «Материаловедение» составила $\approx 0,27$, т.е. выборочные распределения скошены вправо и их асимметрии правосторонние.

Kurtosis — эксцесс (выход, отступление, уклонение) или выпуклость, показывает, насколько "кучно" основное количество данных группируется около центра — это характеристика формы распределения, а именно мера остроты его пика (относительно нормального распределения эксцесс равен 0). Если эксцесс больше нуля, то кривая распределения характеризуется островершинностью, т.е. является более крутой по сравнению с нормальной, а случаи более густо группируются вокруг среднего. При отрицательном эксцессе кривая является более плосковершинной, т.е. более пологой по сравнению с нормальным распределением. Отрицательным пределом величины эксцесса является число -2, положительного предела нет.

По дисциплине «Механика» эксцесс составил $\approx 1,57$, т.е. кривая распределения характеризуется островершинностью, а по дисциплине «Материаловедение» — $\approx -0,33$, т.е. кривая является более плосковершинной, более пологой по сравнению с нормальным распределением.