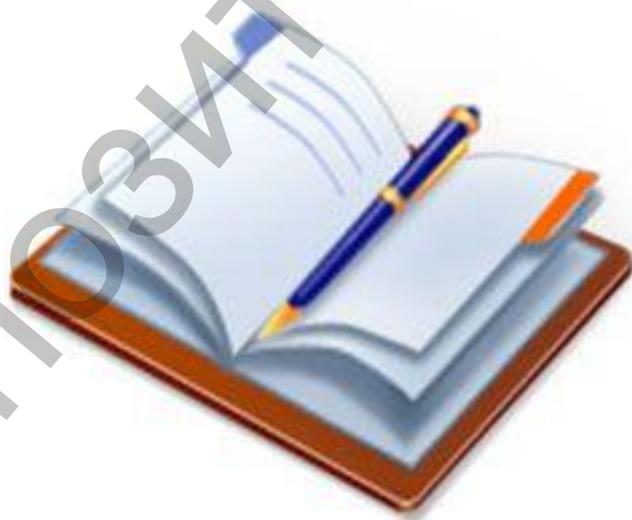


Министерство образования Республики Беларусь  
Учреждение образования «Витебский государственный университет  
имени П.М. Машерова»

# Технологии мониторинга качества учебного процесса

*Практикум*



Витебск

2012

## Содержание

|   |    |
|---|----|
| Введение.....   | 5  |
| Педагогический мониторинг.....  | 6  |
| Система оценки результатов обучения.....                                  | 7  |
| Педагогический контроль .....   | 8  |
| Разработка тестов школьных достижений.....                                | 9  |
| Вопросы и задания для самоконтроля .....                                  | 14 |
| Методические аспекты применения пакета SPSS: описательная статистика..... | 15 |
| Задание .....   | 26 |
| Задание-проект .....  | 30 |
| Литература .....  | 31 |

## Введение

Проблема качества образования в начале XXI века встает со всей остротой в условиях социально-экономических перемен и модернизации системы образования. Отношение к качеству как важнейшему фактору повышения уровня жизни, экономической, социальной, экологической безопасности, как к одной из фундаментальных категорий, определяющих успешное развитие человека и общества, требует поиска новых целевых установок в образовательной политике, кардинального пересмотра традиционных взглядов и представлений о качестве образования и путях его обеспечения.

Под качеством образования понимают соотношение цели и результата, меру достижения цели. Следовательно, необходимо научиться измерять в одинаковых единицах цель, поставленную перед учебным заведением и результат, достигнутый этим учебным заведением.

В настоящее время отмечаются две важнейшие тенденции: современная система образования должна стать «антропоцентричной» (то есть центром всех учебно-воспитательных воздействий должен стать конкретный ученик, студент, слушатель), что требует разработки новых подходов к информационному обеспечению управления и развития новых информационных технологий.

Образовательный мониторинг – это система организации сбора, хранения, обработки и распространения информации о деятельности педагогической системы, обеспечивающая непрерывное слежение за её состоянием и прогнозированием её развития. Образовательный мониторинг – категория педагогическая и управленческая, поскольку он не копирует общие положения теории информации, а переводит их на язык педагогики, психологии и управления. Объектом мониторинга является воздействие процесса обучения на личность обучаемого. Образовательный мониторинг органически связан с управлением качеством образования.

Данная разработка представляет собой подборку материалов для проведения практических занятий в рамках дисциплины «Информационные основы научно-педагогической деятельности» и включает в себя две части: теоретическую и практическую. Содержание теоретической части может быть изучено студентами самостоятельно, для самоконтроля приводятся контрольные вопросы. Практическая часть требует от студентов разработки измерительного инструмента – теста, проведения мониторинга (в учебных целях – по упрощенной схеме) и первичной статистической обработки результатов. В качестве итоговой работы студенты выполняют задание-проект. Работа над проектом может быть организована в групповой форме или индивидуально. На последнем занятии организуется защита проектов с использованием информационных технологий и мультимедийных средств.

## Педагогический мониторинг

Создание системы менеджмента качества образования в Республике Беларусь происходит в период введения образовательных стандартов и появления необходимости оценивать их достижения. В связи с этим актуальной становится задача создания системы получения объективной информации о результатах обучения в соответствии с образовательными стандартами (в том числе определение комплекса критериев, процедур и технологий оценки, организацию педагогического мониторинга и его использование как неотъемлемого инструмента управления качеством образования) на уровне образовательного учреждения, на основе которой можно будет принимать управленческие решения. Современная педагогическая наука и практика поставлены перед необходимостью перехода от традиционных способов сбора сведений об образовательной системе к педагогическому мониторингу.

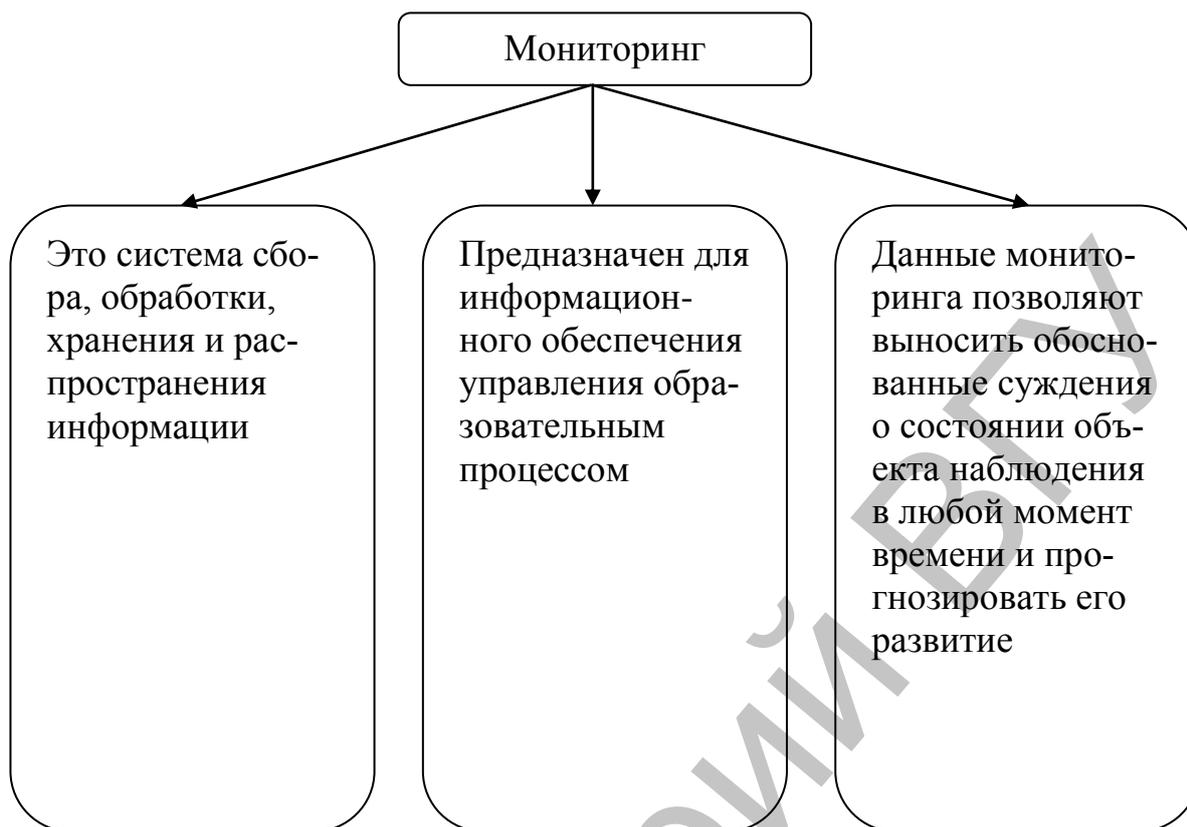
Под педагогическим мониторингом будем понимать **целенаправленное, специально организованное, непрерывное слежение за функционированием и развитием образовательного процесса и/или его отдельных элементов в целях своевременного принятия адекватных управленческих решений на основе анализа собранной информации и педагогического прогноза.**

Понятие «мониторинг» пришло в педагогику из экологии и социологии. Во второй половине XX века процесс слежения за состоянием плодородного слоя почвы стал называться мониторингом (от английского следить – вести наблюдение). В экологии мониторинг – это непрерывное слежение за состоянием окружающей среды с целью предупреждения нежелательных отклонений по важнейшим параметрам. В социологии мониторинг подразумевает определение небольшого числа показателей, отражающих состояние социальной среды.

Традиционная система оценки знаний учащихся имеет богатый опыт в области контроля за результатами обучения. Но она сформировалась ещё в период советской школы, поэтому в силу своих организационных и технологических особенностей не может обеспечить удовлетворение данных потребностей современного общества. Ее результаты неприемлемы для получения объективных количественных и качественных показателей, позволяющих управлять качеством образования. Появление понятия «мониторинг» связано со становлением и развитием информационного общества, которое нуждается в объективных и субъективных сведениях о состоянии тех или иных объектов и структур. Образовательная система оказалась слишком сложной, многоаспектной, чтобы можно было сразу создать такую систему, которая позволила бы объективно судить о состоянии дел.

**Мониторинг в образовании** – это система отбора, обработки, хранения и распространения информации о деятельности образовательной системы или отдельных ее звеньев, ориентированная на информационное обеспечение управления, которая позволяет судить о состоянии объекта в любой момент времени и может обеспечить прогноз его развития и результативности.

Говоря о мониторинге в образовании, следует выделять три основных аспекта (см. рис. 1).



**Рисунок 1.** Основные аспекты системы мониторинга

В настоящее время в нашей стране и в ряде стран СНГ различными научными коллективами создаются подходы к оценке достижения стандартов образования. В связи с этим возникают следующие задачи:

- анализ существующих подходов к оценке достижения стандартов образования;
- оценка перспективности их применения для аттестации обучающихся, а также для построения системы оценки качества образования в образовательном учреждении.

Для этого необходимо иметь ориентиры для сравнения. Одним из таких ориентиров может быть анализ зарубежного опыта, позволяющий выявить тенденции развития систем оценки результатов обучения, преобладающие подходы в оценочной деятельности, технологии оценки, которые можно использовать при разработке системы мониторинга. Предметом анализа могут служить, например, системы оценки результатов обучения в школе, используемые в США на государственном, региональном и школьном уровнях.

### **Система оценки результатов обучения**

Все основные характеристики системы оценки результатов обучения (основные подходы к разработке инструментария, характеристики проверочных работ, процедуры проведения, представление результатов и др.) определяются основными целями и задачами, которые ставятся перед этой системой. В большинстве случаев перед системой оценки результатов обучения ставится **основ-**

**ная цель – получение объективной информации о функционировании системы образования.**

Системы оценки результатов обучения можно разделить на подсистемы в соответствии с целями и объектами оценки:

- *оценка индивидуальных образовательных достижений учащихся* (состояния или динамики роста) для целей аттестации (подтверждения получения определенного уровня образования), коррекции индивидуальных результатов учащихся, перехода на следующую ступень обучения, выбора уровня изучения отдельных учебных предметов;
- *оценка уровня образовательных достижений класса, школы* с целью оценки деятельности учителей или школы, усовершенствования процесса преподавания и обучения;
- *мониторинг образовательных достижений выборочной совокупности учащихся в масштабе отдельных регионов или страны в целом* с целью оценки качества обучения и тенденций развития.

В США образовательные достижения учащихся оцениваются в основном в рамках программ, разработанных на уровне каждого отдельного штата. В стране нет единой системы национальных экзаменов. Каждый штат разрабатывает свои собственные стандарты учебных достижений и систему оценки этих достижений, используя в качестве ориентиров требования к учебным достижениям школьников, сформулированные в национальных образовательных стандартах. На уровне страны существует система мониторинга качества образования NAEP (National Assessment of Educational Progress). Основная цель данных программ получить информацию о результатах образования (обратная связь, отчет школ об их работе), проследить изменения во времени, определить слабые и сильные стороны образования. На основе данной информации планируются любые изменения в школе, включая пересмотр учебных программ.

На Западе считается, что чем выше уровень развития тестового контроля, тем выше рейтинг вуза.

### **Педагогический контроль**

Педагогический контроль, а, следовательно, и педагогическое тестирование, может быть:

- **входным** (или еще говорят «отправным», вводным, начальным, предварительным),
- **текущим**,
- **тематическим**,
- **рубежным** (или периодическим, а если на модульной основе – то суммарным),
- **итоговым**,
- **заключительным** (или выпускным тестированием),
- **отсроченным** (для проведения педагогического эксперимента)

**Метод контроля** – это система последовательных взаимосвязанных диагностических действий педагога и обучаемых, обеспечивающих обратную

связь в процессе обучения с целью получения данных об успешности, эффективности учебного процесса. Методы контроля:

- устный (индивидуальный, фронтальный, уплотненный),
- письменный (контрольные работы, КСР),
- практический (например КСР по методу проектов, проблемные ситуации),
- дидактические тесты,
- наблюдение,
- программированный контроль,
- графический.

**Приемы** письменного контроля – составление таблиц, планов, сочинения, письменные контрольные работы, письменное тестирование; практического – ответ по картине, иллюстрации, обсуждение кадров фильма, описание музейных экспонатов; устного – беседа по теме, рассказ у доски, ответ по цепочке, устное тестирование.

**Формы** контроля подразделяются на фронтальную, групповую (когда только небольшой группе из 4–8 человек предлагается протестироваться на семинарском занятии) и индивидуальную.

В дидактическом компьютерном тесте должны быть реализованы основные **функции** (диагностическая, контролирующая, обучающая, воспитывающая, мотивирующая, развивающая и прогностическая) и принципы (научности и эффективности, иерархической организации, объективности, систематичности, справедливости и всесторонности) педагогического контроля.

## Разработка тестов школьных достижений

**Объединение тестовых заданий (задач) в тест.** Итак, тестовые задачи определены, подобраны, подготовлены и представлены в соответствии с принципами формирования тестовых заданий. Теперь следует проверить их объективность и корректность формулировок. Задачу или вопрос можно считать объективным или корректно сформулированным, если мнения ряда экспертов об их назначении, о правильности формулировки и пригодности вариантов ответов совпадают между собой. Число экспертов, необходимых для работы над тестами, колеблется у разных авторов от 3 до 8. Количество необходимых экспертов зависит от целей создаваемого теста. Для тестов достижений локального применения достаточно и трёх экспертов, для тестов, применение которых планируется в широких масштабах, количество экспертов не должно быть менее пяти. Вопрос оценки качества экспертов и процедура их отбора описаны в работах по технологиям проведения экспертных оценок.

Экспертам предлагается по трёх- или пятибалльной шкале оценить тестовые задания с точки зрения соответствия целям тестирования (отвечает – частично отвечает – не отвечает), однозначности формулировки (однозначна – не совсем однозначна – неоднозначна) и пригодности вариантов ответов (подходят – частично подходят – не подходят). Если хотя бы треть экспертов оценивает тестовое задание как частично соответствующее или несоответствующее, его

исключают. Если процент такой оценки ниже, то тестовое задание пересматривается и корректируется.

**Упорядочивание тестовых заданий для испытания теста.** Существуют правила упорядочивания тестовых заданий при испытаниях теста. Кратко их перечислим.

1. Составьте базу данных для заданий, в которой каждое тестовое задание представлено на отдельном листе (экране), предусмотрев место для занесения экспертных оценок, времени, необходимого для его выполнения, уровня сложности и прочих данных, которые могут его охарактеризовать.
2. Проверьте содержание и формулировку тестовых заданий во взаимосвязи друг с другом.
3. Располагайте задания каждого типа вместе. При этом инструкцию и пояснение к ним необходимо давать один раз для каждой группы заданий. Это даёт возможность испытуемым приспособиться к данному типу заданий, а, следовательно, достаточно понимания одной части инструкции для множества заданий.
4. Располагайте тестовые задания в порядке возрастания предполагаемой трудности. Это предотвратит случаи, когда слишком старательный испытуемый тратит слишком много времени на задания, которые он не может решить и, таким образом, лишает себя возможности попытаться выполнить другие, по которым он мог бы получить баллы. При апробации теста бывает полезным включить в инструкцию пункт о том, что если испытуемому не удаётся справиться с заданием теста, его необходимо пропустить, а после окончания работы над всем тестом вернуться к вызвавшему трудности заданию, если у испытуемого останется время.
5. С учетом пунктов (3) и (4), располагайте материал с разными формами представления данных с чередованием, насколько это возможно. Это снижает монотонность и связанное с ней утомление.
6. Не комплектуйте вместе такое количество заданий, что среднему испытуемому для их выполнения потребуется более получаса – для детей начальной школы; для старшеклассников – более часа. Столько длится период сосредоточения у детей. Если существует необходимость выполнения заданий большей продолжительности, технология проведения должна предусматривать перерыв. Следует отметить, что время появления утомления во многом зависит от мотивации (при этом слишком высокая и слишком низкая мотивация быстрее вызывает утомление), разнообразия материалов тестирования, способов проведения, эмоциональной подготовленности испытуемых.
7. Для того, чтобы определить, какое время требуется для выполнения заданий теста, можно дать испытуемым указание отмечать, какое задание они выполняют в различные моменты времени в процессе тестирования. Но этот способ оценки времени хорош при работе со взрослыми, при работе с детьми эту работу рациональнее будет поручить испытателю, который может

одновременно фиксировать время выполнения заданий у 5-6 испытуемых, отмечая его на специальном бланке.

8. При конструировании бланковых тестов лучше разместить задачи на листах брошюры так, чтобы они были пространственно разделены и легко воспринимались. Задания и варианты ответов к ним должны располагаться на одной странице.
9. Важные части инструкции должны быть подчеркнуты или выделены особым шрифтом.

**Проверка трудности задач. Определение места задачи в тесте.** Важным шагом в конструировании тестов является проверка трудности тестовых заданий. Для этого необходимо провести предварительное тестирование экспериментальной группы (выборки). Отбор участников экспериментальной группы может производиться различными способами: отбор в случайном порядке по алфавиту, по таблице случайных чисел или как-то иначе. Важно, чтобы состав выборки был сбалансированным, т.е. испытуемые должны обладать различными возможностями решения заданий теста. Количественный состав требует отдельного рассмотрения.

После того, как определен состав выборки, испытуемым предлагается пройти тест. Полученные ответы анализируются с целью установления трудности, обоснованности и дискриминативности каждого вопроса, пригодности каждого варианта ответов. Результатом анализа должен явиться отбор и корректировка задач, а также их перераспределение внутри теста.

**Случайное угадывание правильного ответа.** В конструировании тестов достижений существует еще одна проблема – случайное угадывание. Если тесты предназначены для таких целей, как отбор, контрольные или итоговые работы, то очевидно, что испытуемые имеют сильные мотивы для угадывания ответов. Случайное угадывание снижает надежность и валидность тестов. Проблема угадывания ответов стоит особенно остро для альтернативных заданий.

Кроме того, если требуется высокая скорость выполнения тестов, то это подстегнет испытуемых к случайному угадыванию ответов. Еще хуже, однако, проведение тестов в такие сжатые сроки, что испытуемые не могут завершить их выполнения.

В случае с заданиями, содержащими много вариантов выбора и достаточно большим количеством заданий в тесте случайным угадыванием можно пренебречь. Однако здесь возникает еще одна проблема – оценки пропущенных заданий, на которые испытуемые не дали ответа вообще. Может сложиться ситуация, когда одни испытуемые вообще не заполнили часть заданий, а другие заполнили непосильные задания случайным образом. В этом случае необходимо вводить коррекцию на случайное угадывание ответов.

**Дискриминативность** задачи определяется как способность отделять испытуемых с высоким общим баллом по тесту от тех, кто получил низкий балл, или испытуемых с высокой продуктивностью учебной деятельности от испытуемых с низкой продуктивностью. Дискриминативность обозначает различие

тельную способность задачи. Для её определения применяются коэффициент и индекс дискриминативности.

**Надежность** теста является одним из критериев качества теста и показывает, насколько точно измеряет данный тест изучаемое явление, его «помехоустойчивость».

Надежность теста, как правило, определяется после окончания анализа заданий и составления окончательной формы теста.

Как и любая человеческая деятельность, тестирование допускает наличие ошибок. При проведении тестирования возникают следующие виды ошибок: промахи, систематические и случайные ошибки.

**Промахи** возникают при грубых нарушениях процедуры тестирования. Они могут быть легко выявлены и устранены путем анализа резко отклоняющихся значений.

**Систематические ошибки** измерения возникают постоянно или закономерно меняются от измерения к измерению. В силу этих своих особенностей они могут быть предсказаны заранее, а в некоторых случаях и устранены.

**Случайные ошибки** имеют место, когда при последовательных измерениях постоянной характеристики получают различные числовые оценки, т.е. при определении случайных ошибок предполагается, что измеряемая характеристика не изменяется во времени, а все отклонения обусловлены неточностью измерения.

Граница между случайными и систематическими ошибками весьма условна. Суммарная ошибка любого теста состоит из ошибок двух этих типов, но в разном соотношении. Относительный вес каждого типа ошибок зависит от качества теста и условий проведения обследования. Результаты тестирования всегда содержат ошибки, как бы тщательно оно ни проводилось. Однако используя методы математической статистики, можно оценить величину суммарной ошибки тестирования и использовать её для оценки надежности теста. Без статистической оценки надежности теста результаты измерения нельзя считать достоверными.

На практике используются три основных метода оценки надежности тестов:

- повторное тестирование (ретестирование);
- параллельное тестирование (тестирование параллельной формой теста);
- расщипление (метод деления целого на части).

Окончательный вариант теста должен соответствовать требованию внутренней согласованности. Внутренняя согласованность – это характеристика теста, указывающая на степень однородности состава тестовых заданий с точки зрения измеряемого качества (т.е. все задания теста должны определять тестируемое качество). Для её определения применяется процедура установления корреляций между результатами теста в целом и каждым отдельным заданием.

Источниками неудовлетворительной надежности тестов могут быть:

1. Наличие в тесте заданий, которые допускают неоднозначное толкование результатов. Часто к таким заданиям относятся задания со свободно конструируемыми ответами.
2. Угадывание ответов.
3. Величина теста. Чем тест длиннее, тем он надежнее.
4. Инструкции и технология проведения. При помощи инструкции легко изменить трудность задания. Неоднозначные инструкции приводят к снижению надежности теста. Повышение надежности возможно двумя путями - ужесточением инструкций и повышением качества подготовки экспериментаторов.
5. Источники, связанные с испытуемыми. Усталость, скука, невнимательность, жара или холод, самочувствие, различная мотивация и т.д.

**Валидность.** В общем виде понятие валидности раскрывает, насколько полученные в результате тестирования результаты соответствуют объективной реальности. А поскольку отсутствует способ точного определения объективной реальности, то валидность – понятие относительное, поскольку тесты, как и любой измерительный инструмент, никогда не являются абсолютно эквивалентными реальному положению вещей. Соответственно, валидность устанавливается относительно характеристики, признака, величины, принимаемой (а не истинно являющейся) за объективную.

В современной тестологии выделяют следующие виды валидности:

- очевидная валидность;
- валидность по содержанию (содержательная);
- конструктивная (концептуальная);
- валидность по критерию (критериальная или эмпирическая валидность);
- сравнительная или конкурентная валидность.

Валидность и надежность – связанные понятия. ненадежный тест не может быть валидным, и, наоборот, валидный тест всегда надежен.

## Вопросы и задания для самоконтроля

1. Можно ли утверждать, что термины «педагогический мониторинг» и «мониторинг в образовании» – синонимы? Ответ поясните.
2. Каковы подходы к оценке достижения стандартов образования?
3. Каковы основные характеристики системы оценки результатов обучения?
4. В чем отличие методов контроля от приёмов контроля?
5. Перечислите основные формы контроля.
6. Что понимается под педагогическим тестом? Каковы особенности педагогического теста?
7. Почему педагогический тест относят к активным формам контроля?
8. Что такое тестовое задание? Каковы основные формы тестовых педагогических заданий?
9. Перечислите требования к тестовым заданиям.
10. Каковы методы оценки критериев качества тестов?
11. Перечислите известные подходы к классификации тестов. Какие тесты сегодня наиболее применимы в системе образования и почему?
12. Как оцениваются тестовые задания?
13. Как связаны между собой длина теста и оптимальное время тестирования?
14. Все ли категории учебных целей в таксономии Б. Блума (см. табл. 2) могут быть представлены в тестовой форме? Ответ поясните.
15. Как проводится мониторинг инвариантной и вариативной частей образовательного стандарта? Как и где могут быть применены тестовые технологии при таком мониторинге?
16. Перечислите основные характеристики системы мониторинга.
17. На что нацелена система мониторинга качества образования в образовательном учреждении? Каковы субъект и объект мониторинга?
18. Каковы этапы разработки тестов школьных достижений?
19. Может ли один и тот же тест рассматриваться как нормативно-ориентированный и как критериально-ориентированный? Ответ поясните.
20. Какие виды ошибок возникают при проведении тестирования?



## Методические аспекты применения пакета SPSS: описательная статистика

Для обработки результатов мониторинга используют методы математической статистики. В предмете статистики принято различать два основных раздела: – дескриптивную (описательную) статистику и аналитическую статистику (теорию статистического вывода). Эти разделы в различной степени представлены в ряде современных компьютерных программ, предназначенных для обработки данных. Сложные вычислительные процедуры статистического вывода менее автоматизированы и встречаются, главным образом в профессионально-ориентированных пакетах статистических вычислений – Statistica, SPSS, SAS. Многие задачи статистики могут быть решены также с использованием универсальных систем компьютерной математики – MathCAD, Maple, MatLab, Mathematica.

Рассмотрим применение возможностей пакета SPSS для обработки результатов мониторинга. Анализ результатов начинается с описательной статистики.

Целью описательного анализа является систематизация имеющихся данных. В рамках данной задачи происходит построение линейных распределений, а также характеристика переменных в различных статистических аспектах: расчёт **среднего, медианы, моды** и т.д. Линейные (общие) распределения позволяют подсчитать количество респондентов, указавших тот или иной вариант ответа на рассматриваемый вопрос.

Построение линейных распределений обычно является первым шагом в статистическом анализе данных. При помощи линейных распределений становится возможным систематизировать ответы испытуемых.

Пример. Известны результаты контроля в одном из классов общеобразовательной школы: результаты теста по иностранному языку, оценка за сочинение и результаты контрольной работы по предмету «Человек и общество». Они приведены в табл. 3.

Таблица 3.

Результаты мониторинга

| Личный номер | Имя ученика  | Пол | Тест по иностранному языку | Сочинение | Человек и общество |
|--------------|--------------|-----|----------------------------|-----------|--------------------|
| 1            | Агеева Л.    | ж   | 90                         | 4,5       | 4,5                |
| 2            | Агафонова О. | ж   | 66                         | 3,0       | 4,0                |
| 3            | Бегунов С.   | м   | 96                         | 4,5       | 4,5                |
| 4            | Волков В.    | м   | 84                         | 3,5       | 3,5                |
| 5            | Волков П.    | м   | 95                         | 4,5       | 5,0                |
| 6            | Гиржева Е.   | ж   | 83                         | 4,0       | 4,5                |
| 7            | Грицко А.    | ж   | 94                         | 4,0       | 5,0                |
| 8            | Демчук Н.    | ж   | 82                         | 3,5       | 4,0                |

|    |              |   |     |     |     |
|----|--------------|---|-----|-----|-----|
| 9  | Демидова В.  | ж | 97  | 4,5 | 4,5 |
| 10 | Демидович Д. | м | 97  | 3,5 | 5,0 |
| 11 | Еремеев Д.   | м | 59  | 2,0 |     |
| 12 | Жуков А.     | м | 95  | 4,0 | 4,5 |
| 13 | Зуев К.      | м | 78  | 3,5 | 4,0 |
| 14 | Кириллова О. | ж | 90  | 5,0 | 5,0 |
| 15 | Кириянов М.  | м | 47  |     |     |
| 16 | Кощеева В.   | ж | 95  | 4,0 | 4,5 |
| 17 | Леонтьева С. | ж | 100 | 5,0 | 5,0 |
| 18 | Парусова И.  | ж | 69  | 3,0 | 3,5 |
| 19 | Раков О.     | м | 44  |     |     |
| 20 | Рощин В.     | м | 80  | 4,5 | 4,0 |
| 21 | Федорова Р.  | ж | 75  | 3,5 | 2,0 |
| 22 | Шигун Д.     | ж | 75  | 3,5 | 4,5 |
| 23 | Яковлева В.  | ж | 51  | 3,0 | 3,0 |
| 24 | Якубович С.  | м | 89  | 4,5 | 4,5 |
| 25 | Яцук И.      | ж | 82  | 5,0 | 4,5 |

Определим переменные, как показано на рисунке 2. Значения переменной **gender** (Пол) определим следующим образом: 0 – женский, 1 – мужской. Это номинальная переменная, числа 0 и 1 в данном случае представляют собой только код значения этой переменной.

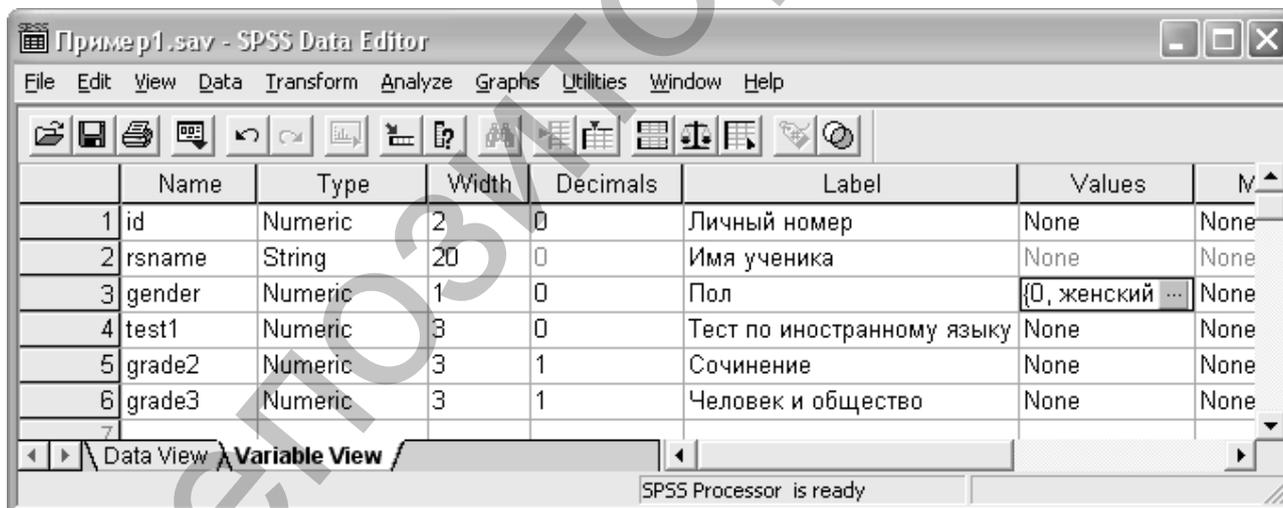


Рисунок 2. Определение переменных в окне Variable View

Введем данные в окно Data View (см. рисунок 3).

Пример1.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

25 :

|    | id | rsname       | gender | test1 | grade2 | grade3 |
|----|----|--------------|--------|-------|--------|--------|
| 1  | 1  | Агеева Л.    | 0      | 90    | 4,5    | 4,5    |
| 2  | 2  | Агафонова О. | 0      | 66    | 3,0    | 4,0    |
| 3  | 3  | Бегунов С.   | 1      | 96    | 4,5    | 4,5    |
| 4  | 4  | Волков В.    | 1      | 84    | 3,5    | 3,5    |
| 5  | 5  | Волков П.    | 1      | 95    | 4,5    | 5,0    |
| 6  | 6  | Гиржева Е.   | 0      | 83    | 4,0    | 4,5    |
| 7  | 7  | Грицко А.    | 0      | 94    | 4,0    | 5,0    |
| 8  | 8  | Демчук Н.    | 0      | 82    | 3,5    | 4,0    |
| 9  | 9  | Демидова В.  | 0      | 97    | 4,5    | 4,5    |
| 10 | 10 | Демидович Д. | 1      | 97    | 3,5    | 5,0    |
| 11 | 11 | Еремеев Д.   | 1      | 59    | 2,0    | .      |
| 12 | 12 | Жуков А.     | 1      | 95    | 4,0    | 4,5    |
| 13 | 13 | Зуев К.      | 1      | 78    | 3,5    | 4,0    |
| 14 | 14 | Кириллова О. | 0      | 90    | 5,0    | 5,0    |
| 15 | 15 | Кириянов М.  | 1      | 47    | .      | .      |
| 16 | 16 | Кощеева В.   | 0      | 95    | 4,0    | 4,5    |
| 17 | 17 | Леонтьева С. | 0      | 100   | 5,0    | 5,0    |
| 18 | 18 | Парусова И.  | 0      | 69    | 3,0    | 3,5    |
| 19 | 19 | Раков О.     | 1      | 44    | .      | .      |
| 20 | 20 | Роцин В.     | 1      | 80    | 4,5    | 4,0    |
| 21 | 21 | Федорова Р.  | 0      | 75    | 3,5    | 2,0    |
| 22 | 22 | Шигун Д.     | 0      | 75    | 3,5    | 4,5    |
| 23 | 23 | Яковлева В.  | 0      | 51    | 3,0    | 3,0    |
| 24 | 24 | Якубович С.  | 1      | 89    | 4,5    | 4,5    |
| 25 | 25 | Яцук И.      | 0      | 82    | 5,0    | 4,5    |

Data View Variable View

SPSS Processor is ready

Рисунок 3. Исходные данные

Получим все характеристики распределения частот для переменных gender, test1, grade2, grade3. Для этого воспользуемся командой **Анализ** → **Описательные статистики** → **Частоты** (Analyze → Descriptive Statistics → Frequencies) (см. рис. 4).

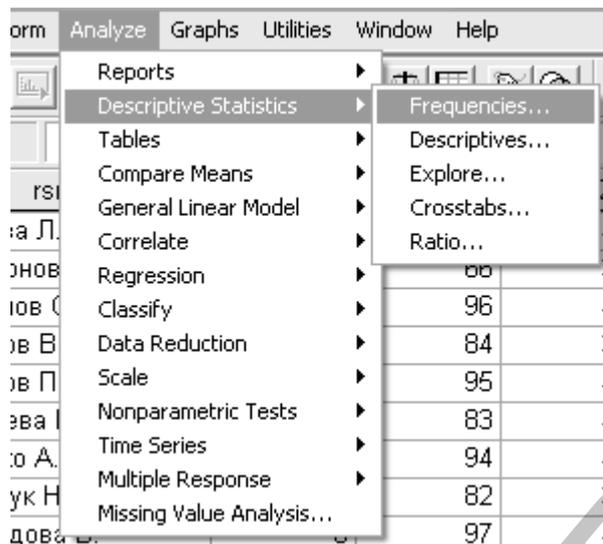


Рисунок 4. Последовательность команд для вызова окна Частоты (Frequencies)

Окно Частоты (Frequencies) наряду с обычными кнопками (Ok, Past и др.) содержит и специфические функциональные, в том числе **Статистика** (Statistics), **Диаграммы** (Charts) и **Формат** (Format) (см. рис. 5). Их названия говорят сами за себя.

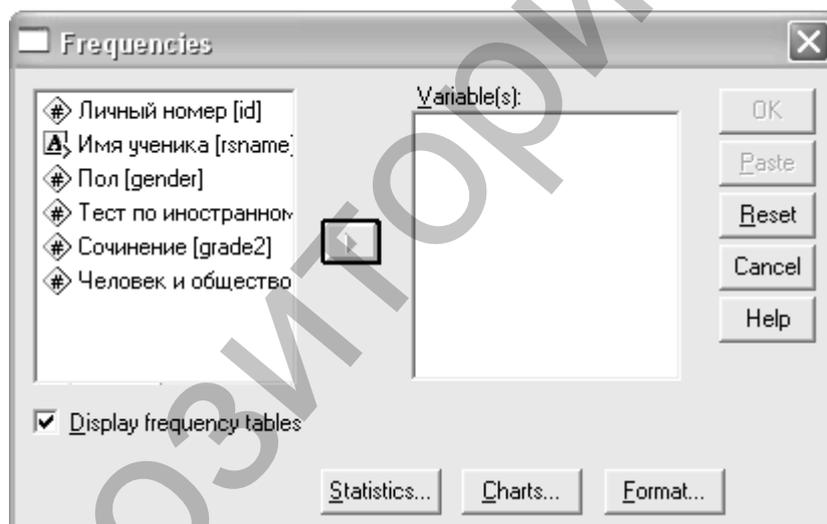


Рисунок 5. Окно Частоты (Frequencies)

Так, с помощью кнопки **Статистика** (Statistics) вызывается диалоговое окно, в котором можно выбрать статистические характеристики распределения случайных величин (см. рис. 6). Эти характеристики разбиты на следующие четыре группы<sup>1</sup>:

1. **Percentile Values** (Процентные значения):

- Квартили – доли функции распределения случайной величины, содержащиеся в интервалах 0-25%, 0-50%, 0-75% и 0-100%;
- **Cut points for ... equal groups** (Разбить точки на ... равных групп) – разделить диапазон изменения случайной величины на одинаковые интервалы (в данном окне указано 10);

<sup>1</sup> Большинство представленных в этом окне статистик подходит только для переменных, имеющих интервальный тип шкалы.

- Процентили – доли функции распределения случайной величины, содержащиеся в интервалах от 0% до 100%, указываемые пользователем в ячейке и затем добавляемые в другую ячейку (должен быть охвачен весь диапазон от 0% до 100%).

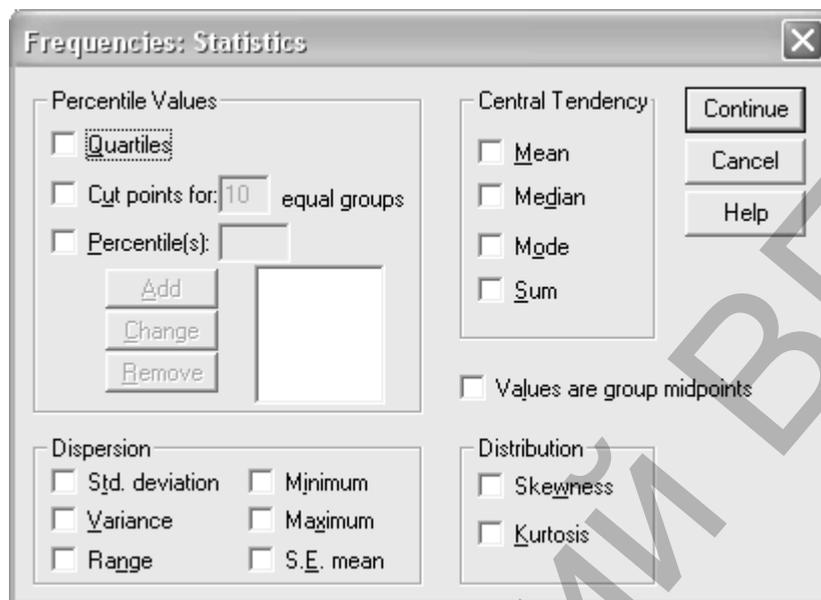


Рисунок 6. Окно Частоты: Статистика

## 2. Дисперсия (Dispersion):

- стандартное отклонение;
- вариация;
- диапазон изменения;
- минимум;
- максимум;
- стандартная ошибка среднего.

## 3. Характеристики положения (Central Tendency)

- мера;
- медиана;
- мода;
- сумма.

## 4. Распространение (Distribution)

- асимметрия (третий центральный момент);
- эксцесс (четвертый центральный момент).

Все эти характеристики указываются установкой флажков в соответствующих ячейках.

Необходимо сказать несколько слов относительно основных описательных статистик, представленных на рисунке 6. Пожалуй, наиболее популярными характеристиками, используемыми для описания переменных, являются показатели группы **Central Tendency** (Центральная тенденция): среднее арифметическое (Mean); медиана, или половина значений отрезка (Median); мода, или наиболее часто встречающееся значение (Mode) а также сумма (Sum). Данные

показатели применяются неодинаково к переменным с различным типом шкалы (см. табл. 4).

Таблица 4.

**Наиболее релевантные показатели центральной тенденции для переменных с различным типом шкалы**

| Тип шкалы    | Наиболее релевантная характеристика | Другие релевантные характеристики |
|--------------|-------------------------------------|-----------------------------------|
| Интервальная | Среднее арифметическое              | Средневзвешенное, мода            |
| Порядковая   | Средневзвешенное                    | Мода                              |
| Номинальная  | Мода                                | –                                 |

Из представленной таблицы видно, что наиболее релевантной описательной статистикой, характеризующей переменные с интервальной шкалой, является среднее арифметическое (Mean). Для переменных с порядковой шкалой данный показатель неприменим, так как он рассчитывается исходя из значений переменной (кодов вариантов ответов), а не самих значений интервалов.

Средняя тенденция переменных с номинальной шкалой не может быть оценена никак, кроме моды, – то есть для таких переменных можно определить только наиболее многочисленную группу. Например, для переменной **gender** (Пол) можно заключить, что в нашей выборке количество представителей женского пола (девушек) больше (см. рис. 9).

Из таблицы 5 видно, что интервальные переменные – наиболее гибкие относительно применения показателей центральной тенденции. Для них можно рассчитать все три рассматриваемые статистики: среднее арифметическое, средневзвешенное<sup>2</sup> и моду. Порядковые переменные находятся на втором месте: с ними могут использоваться только средневзвешенное и мода. И наконец, номинальные переменные являются наименее гибкими: к ним может эффективно применяться только мода.

Нажав кнопку **Далее** (Continue) возвращаемся в окно **Частоты** (Frequencies) (рис. 5). Здесь следует обратить внимание на ячейку **Display frequency tables** (Отображать таблицу частот), наличие или отсутствие флажка в которой определяет, будет ли выводиться на экран или печать таблица частот.

Следующая специфическая кнопка – **Диаграммы** (Charts), нажав на неё откроем тем самым одноименное окно, которое позволяет помимо таблиц вывести диаграммы по выбранным переменным (см. рис. 7). По умолчанию SPSS не выводит диаграмм.

<sup>2</sup> Средневзвешенное значение для переменных с интервальной шкалой равно среднему арифметическому.

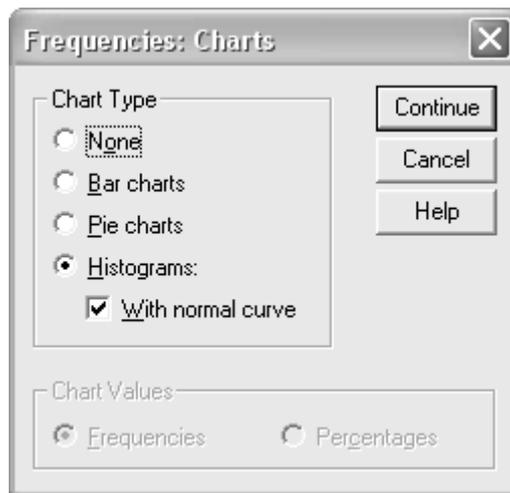


Рисунок 7. Окно Частоты: Диаграммы

В этом окне с помощью переключателя можно задать вид диаграммы, отображающей характер распределения частот. В частности, можно вообще не показывать диаграмму (позиция **None**). В варианте **Bar charts** отображаются только те разряды, в которых есть реальные значения (назовём такую гистограмму первичной или базовой). В варианте **Pie charts** диаграмма показывается в виде круга. Вариант **Histograms** представляет собой стандартизированную гистограмму, где весь диапазон изменения случайной величины (100%) делится на равные интервалы, и числа значений, полученных в ходе наблюдения, приписываются к соответствующим интервалам.

Еще одна специфическая кнопка – **Формат** (Format) (см. рис. 8).

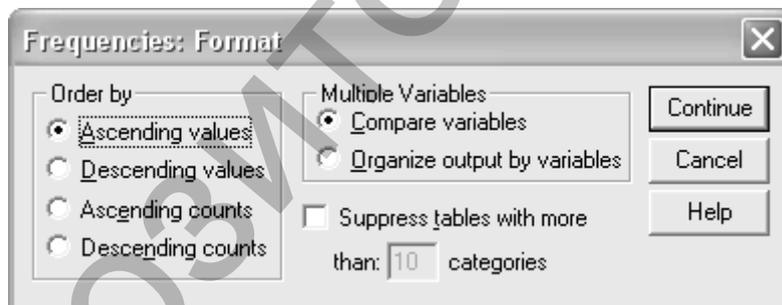


Рисунок 8. Окно Частоты: Формат

Здесь возможен выбор следующих вариантов настройки:

1. **Order by** (Упорядочить по):
  - **Ascending values** (По возрастанию значений);
  - **Descending values** (По значений);
  - **Ascending counts** (По возрастанию частот);
  - **Descending counts** (По убыванию частот);
2. **Multiple variables** (Множественные переменные):
  - **Compare variables** (Все переменные в одной таблице);
  - **Organize output by variables** (Организовать ввод переменных). Опция активизируется только в том случае, когда каждая переменная выводится в своей отдельной таблице. Когда мы ограничиваем рассмотрение одной переменной, это несущественно;

3. **Supress tables with than ... categories** (Не выводить таблицы с числом ре- визитов более ... ). По умолчанию указано значение 10.

После того, как выполнены соответствующие настройки, можно про- смотреть результат. В SPSS для представления результирующей информации, как текстовой, так и графической, предназначена выходная форма, называемая просмотрщиком и имеющая расширение .spo.

Произведем соответствующие настройки для анализа переменной **gender** (Пол). Напомним, что эта переменная принимает значения 0 (женский) или 1 (мужской). В выходной форме содержится информация, представленная на ри- сунке 9.

**Statistics**

| Пол  |         |    |
|------|---------|----|
| N    | Valid   | 25 |
|      | Missing | 0  |
| Mode |         | 0  |

**Пол**

|       |         | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|---------|-----------|---------|---------------|--------------------|
| Valid | женский | 14        | 56,0    | 56,0          | 56,0               |
|       | мужской | 11        | 44,0    | 44,0          | 100,0              |
|       | Total   | 25        | 100,0   | 100,0         |                    |

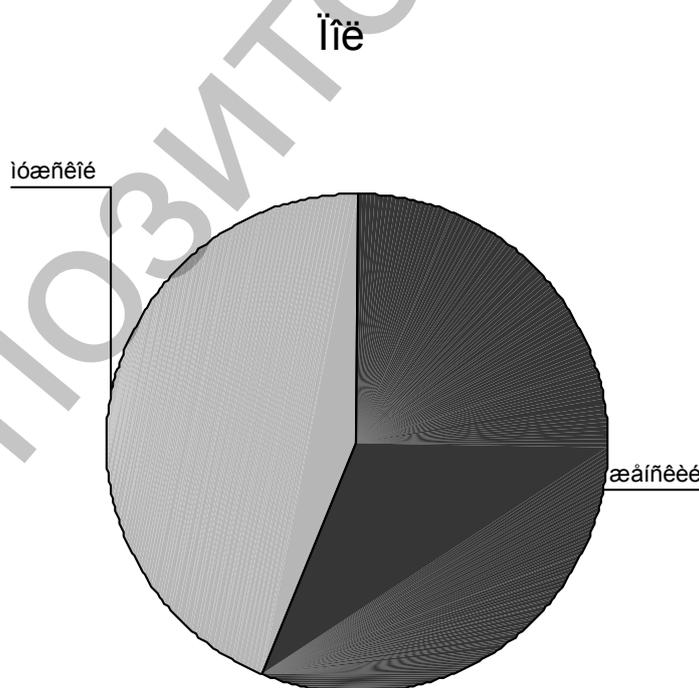


Рисунок 9. Результаты обработки переменной gender (Пол)

Из первой таблицы можем заключить, что общее число обрабатываемых значений – 25, все они корректны. Во второй таблице присутствуют следующие столбцы:

- **Valid** (в буквальном переводе «справедливый») – собственно значения переменной gender (Пол);
- **Frequency** (Частота) повторяемость процентов;
- **Percent** (Процент) – частоты, выраженные в процентах;
- **Valid Percent** (Правильный процент) – теоретически точное значение процента;
- **Cumulative Percent** (Накопительный процент) – процент, суммируемый с начала отсчёта.

Из этой таблицы заключаем, что девушек в списке – 14, юношей – 11, что составляет соответственно 56% и 44%.

Затем эта таблица приводится в графическом виде **Pie charts** (Круговая диаграмма). На ней отображено процентное соотношение юношей и девушек в обрабатываемой выборке.

Произведем настройки для переменной **test1** (Тест по иностранному языку). В выходной форме получим результат, представленный на рисунках 10-12.

В таблице **Statistics** (рис.10) отражаются общие параметры линейного распределения. Здесь представлены:

- количество испытуемых, проходивших тестирование по иностранному языку – 25 человек;
- количество испытуемых, не проходивших тестирование, равно 0, т.е. тестирование прошли все участники выборки;
- медиана (средневзвешенное значение) равна 83, т.е.средний балл по результатам теста в данной выборке равен 83;
- мода, т.е. результат теста (в баллах), который чаще всего встречается в выборке, равен 95.

| Statistics                 |         |       |
|----------------------------|---------|-------|
| Тест по иностранному языку |         |       |
| N                          | Valid   | 25    |
|                            | Missing | 0     |
| Median                     |         | 83,00 |
| Mode                       |         | 95    |

**Рисунок 10. Основные показатели описательной статистики по переменной test1 (Тест по иностранному языку)**

Следующая таблица, озаглавленная меткой анализируемой переменной (Тест по иностранному языку), отражает количество респондентов, которые набрали то или иное количество баллов по тесту (столбец **Frequency**), отсортированное по возрастанию. Также в этой таблице представлен процент лиц, набравших такие баллы от общего числа респондентов (столбец **Percent**). Последний столбец (**Cumulative Percent**) отражает кумулятивные (накопленные) проценты, т.е. вклад каждого значения в общую сумму. Также как и в таблице

**Statistics**, здесь указано общее количество испытуемых, прошедших тестирование.

**Тест по иностранному языку**

|          | Frequency | Percent | Valid Percent | Cumulative Percent |
|----------|-----------|---------|---------------|--------------------|
| Valid 44 | 1         | 4,0     | 4,0           | 4,0                |
| 47       | 1         | 4,0     | 4,0           | 8,0                |
| 51       | 1         | 4,0     | 4,0           | 12,0               |
| 59       | 1         | 4,0     | 4,0           | 16,0               |
| 66       | 1         | 4,0     | 4,0           | 20,0               |
| 69       | 1         | 4,0     | 4,0           | 24,0               |
| 78       | 1         | 4,0     | 4,0           | 28,0               |
| 80       | 1         | 4,0     | 4,0           | 32,0               |
| 83       | 1         | 4,0     | 4,0           | 36,0               |
| 84       | 1         | 4,0     | 4,0           | 40,0               |
| 89       | 1         | 4,0     | 4,0           | 44,0               |
| 94       | 1         | 4,0     | 4,0           | 48,0               |
| 96       | 1         | 4,0     | 4,0           | 52,0               |
| 100      | 1         | 4,0     | 4,0           | 56,0               |
| 75       | 2         | 8,0     | 8,0           | 64,0               |
| 82       | 2         | 8,0     | 8,0           | 72,0               |
| 90       | 2         | 8,0     | 8,0           | 80,0               |
| 97       | 2         | 8,0     | 8,0           | 88,0               |
| 95       | 3         | 12,0    | 12,0          | 100,0              |
| Total    | 25        | 100,0   | 100,0         |                    |

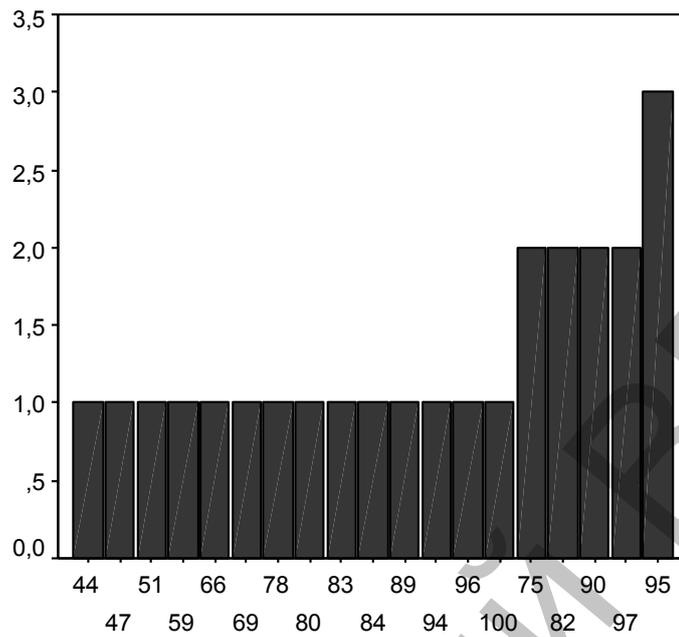
**Рисунок 11.** Данные частот по переменной test1 (Тест по иностранному языку) в виде таблицы

Поскольку в настройках было указано на необходимость построения гистограммы по рассматриваемой переменной (был выбран вариант **Bar charts**), она представлена в результатах линейных распределений после таблицы **Тест по иностранному языку** (см. рис. 12).

К сожалению, графическая подсистема программы SPSS весьма слаба и не выдерживает сравнения со средствами Micrisift Office или Statistica, поэтому рекомендуется ею пользоваться, только когда это действительно оправдано. В остальных случаях предпочтительнее копировать вводимые таблицы в Microsoft Excel и уже там строить по полученным данным диаграммы.

Для сравнения полученных результатов на рисунке 13 приведены результаты использования пакета Анализ данных программы Microsoft Excel.

### Όαño ïï èíïñòðàíίίίο ÿçûέó



### Όαño ïï èíïñòðàíίίίο ÿçûέó

Рисунок 12. Гистограмма частот по переменной test1 (Тест по иностранному языку)

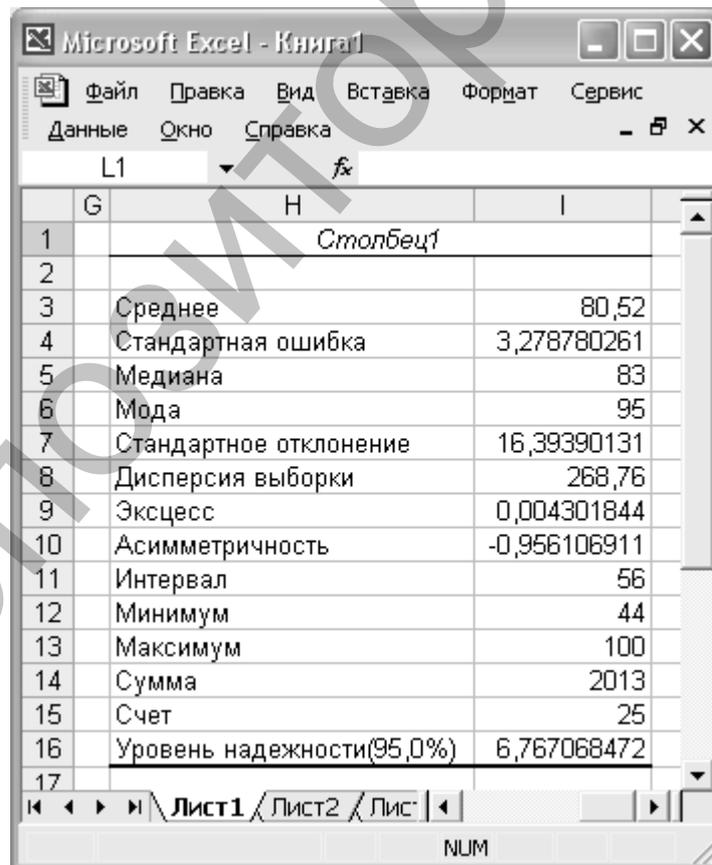


Рисунок 13. Результаты описательной статистики пакета Анализ данных программы Microsoft Excel

## Задание

1. Прокомментируйте результаты линейных распределений по переменной **grade2** (Сочинение), представленных на рисунке 26.

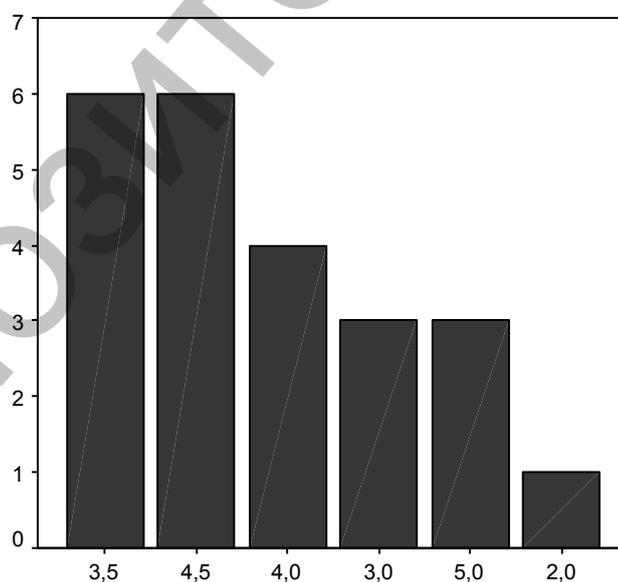
**Statistics**

| Сочинение |         |       |
|-----------|---------|-------|
| N         | Valid   | 23    |
|           | Missing | 2     |
| Mean      |         | 3,913 |
| Variance  |         | ,5830 |

**Сочинение**

|         |        | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|--------|-----------|---------|---------------|--------------------|
| Valid   | 3,5    | 6         | 24,0    | 26,1          | 26,1               |
|         | 4,5    | 6         | 24,0    | 26,1          | 52,2               |
|         | 4,0    | 4         | 16,0    | 17,4          | 69,6               |
|         | 3,0    | 3         | 12,0    | 13,0          | 82,6               |
|         | 5,0    | 3         | 12,0    | 13,0          | 95,7               |
|         | 2,0    | 1         | 4,0     | 4,3           | 100,0              |
|         | Total  |           | 23      | 92,0          | 100,0              |
| Missing | System | 2         | 8,0     |               |                    |
| Total   |        | 25        | 100,0   |               |                    |

Ñî÷-èíáíèà



Ñî÷-èíáíèà

**Рисунок 14. Результаты линейных распределений по переменной grade2 (Сочинение)**

2. Прокомментируйте результаты анализа, представленные на рис. 25. Можно ли на основании этих результатов сделать вывод о нормальности распределения переменной test1? Сформулируйте нулевую и альтернативную гипотезы. Какую из них следует отклонить?

3. Самостоятельно проведите описательный анализ переменной **grade3** (Человек и общество), используя любой из пакетов (Statistica или SPSS).

4. Вам предлагается анкета для родителей детей, участвующих в эксперименте по полилингвальной модели поликультурного образования одной из школ России. Создайте матрицу данных с помощью пакетов Statistica или SPSS. Для этого определите переменные из анкеты № 2: задайте для каждой из них имя, тип, формат, метку переменной, метки значений (где это требуется), код пропущенных значений, уровень измерения. (**Методическое указание:** Не забудьте создать переменную *Личный номер*.)

Заполните матрицу данных случайным образом или опросите студентов своей группы.

Постройте одномерные распределения для всех номинальных и порядковых переменных.

Проверьте по распределениям качество ввода данных (отсутствие значений, выходящих за допустимые пределы); корректность группировок; правильность использования пропущенных значений.

Для количественных переменных вычислите описательные статистики (моду, медиану, среднее арифметическое, среднее квадратическое отклонение, минимум, максимум, эксцесс и асимметрию).

## Анкета № 2

## **Уважаемые родители!**

Просим высказать Ваше мнение об участии учебного заведения, в котором обучается Ваш ребенок, в эксперименте по полилингвальной модели поликультурного образования (ПМПО). Анкета анонимная. Результаты анкетирования будут использованы в анализе эффективности данной модели

1. Была ли организована работа по ознакомлению Вас с целями, задачами и содержанием эксперимента?

да                      нет

2. Насколько полно Вы были ознакомлены с целями, задачами и содержанием эксперимента?

1                      2                      3                      4                      5

3. Довольны ли Вы тем, что Ваш ребенок стал участником эксперимента?

1                      2                      3                      4                      5

не доволен

очень доволен

4. На Ваш взгляд, доволен ли Ваш ребенок учебной деятельностью, которой он стал заниматься в школе?

1                      2                      3                      4                      5

не доволен

очень доволен

5. Как Вы оцениваете степень получения информации о достижениях Вашего ребенка? Выберите один из предложенных вариантов.

- информация о достижениях моего ребенка исчерпывающая
- информируют только о неудачах моего ребенка
- получаю информацию только об оценках моего ребенка
- практически нет информации о достижениях моего ребенка
- другой вариант

6. Как Вы оцениваете свое влияние на ход эксперимента?

1                      2                      3                      4                      5

не имею влияния

значительное влияние

7. Знает ли Ваш ребенок, что он обучается в особых образовательных условиях?

да                      нет                      частично                      не знаю

8. Знает ли Ваш ребенок значение таких понятий, как ?

|           |    |     |
|-----------|----|-----|
| Родина    | да | нет |
| патриот   | да | нет |
| гражданин | да | нет |
| страна    | да | нет |
| дружба    | да | нет |

|          |    |     |
|----------|----|-----|
| мир      | да | нет |
| согласие | да | нет |

9. Каких специальных результатов в развитии Вашего ребенка в условиях ПМПО Вы ждете?

- привития любви к родному языку
- привития любви к родной истории
- привития любви к родному краю
- формирования потребности в изучении родного языка
- формирования потребности в изучении традиций и обычаев родного народа
- привития интереса к изучению истории, обычаев, традиций двух народов
- более высокого уровня формирования общей культуры ребенка
- формирования чувства дружелюбия, коллективизма, оказания помощи другим

10. Какие, по Вашему мнению, возможности поликультурного воспитания остались в эксперименте без внимания?

- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_

11. Если вы хотите что-нибудь дополнить, просим написать ниже.

---

---

---

.....

## Задание-проект

1. В соответствии с теорией самостоятельно разработайте тест для диагностики результатов образовательного процесса. Определите, какого подхода в оценке образовательных достижений учащихся Вы придерживались? При создании теста максимально придерживайтесь требований к педагогическим тестам. Организуйте работу теста в электронном виде. Протестируйте студентов Вашей группы (подгруппы), сохраните результаты для дальнейшей обработки.
2. Представьте свой тест, опишите типы заданий, которые в него входят, способы и критерии оценивания ответов на задания теста, подведение итогов (как рассчитывается результирующая оценка), как предьявляется тест (например, можно ли чередовать задания случайным образом или порядок предьявления заданий однозначно определен, чередуются ли сложные задания и простые), предлагается ли подробная инструкция перед началом тестирования, как фиксируются и сохраняются результаты тестирования и т.д.
3. Произведите анализ сложности тестовых заданий (всего теста) математическими методами. (См. формулу (1) на стр. 16)
4. Приведите описательную статистику по результатам Вашего теста (используйте пакет анализа Excel, программы Ststistica или SPSS). Поясните показатели. Постройте гистограммы распределения частот.
5. Проведите анализ переменных на нормальность распределения. Сформулируйте нулевую и альтернативную гипотезы. Какую из них следует отклонить по результатам анализа Вашего теста? Аргументируйте выводы, подтвердите их расчетами.
6. На основе сравнения результатов по подготовленному Вами тесту и результатов других тестов в подгруппе опишите педагогические гипотезы, которые могут быть подтверждены или опровергнуты с помощью пакета анализа Excel, программ Ststistica или SPSS. Свои выводы обоснуйте.
7. Свои выводы оформите в виде отчёта – презентации Power Point. Настройте переходы между слайдами презентации, выполните её оформление в едином стиле, на титульном листе укажите своё авторство. Представьте, что ваша презентация будет демонстрироваться на педсовете во время Вашего выступления с докладом (время выступления – не менее 5 минут). Сохраните презентацию с расширением pps, т.е. как самостоятельную демонстрацию.
8. Предьявите проект для защиты.



## Литература

1. Аванесов В.С. Научные проблемы тестового контроля знаний – М., 1994
2. Балыкина Е.Н., Бузун Д.Н. Компьютерное педагогическое тестирование: Теория и практика – Минск: РИВШ, 2010
3. Беспалько В.М. Мониторинг качества обучения – средство управления образованием. – М., 1996
4. Вовна В.И., Короченцев В.В. и др. Мониторинг качества обучения как важнейший инструмент управления образованием // Инновации в образовании, 2005, № 5
5. Горб В.Г. Педагогический мониторинг образовательного процесса как фактор повышения его уровня и результатов // Стандарты и мониторинг, 2000, № 5
6. Гутник И.Ю. Организация педагогической диагностики в профильном обучении. – СПб., 2005
7. Кайнова Э.Б. Критерии качества образования: основные характеристики и способы измерения. – М., 2005
8. Лизинский В.М. Критерии оценки и стимулирование педагогической деятельности учителей как один из ресурсов управления образовательным процессом // Завуч, 2005, № 5
9. Майоров А.Н. Мониторинг в образовании. – СПб., 1998
10. Майоров А.Н. Теория и практика создания тестов для системы образования. – М., 2001
11. Поташник М.М. Управление качеством образования в школе. – М., 1996
12. Применение статистических методов анализа данных в научно-исследовательской работе: учебно-методическое пособие / Авт.-сост.: А.И. Бочкин, Л.П. Колбасич. – Витебск: Изд-во УО «ВГУ им. П.М. Машерова», 2006
13. Субетто А.И. Качество образования в России: состояние, тенденции, перспективы. – М., 2001
14. Третьяков П.И. Школа: управление по результатам. – М., 2001
15. Трубина И.И. Мониторинг качества образования: проблемы и подходы // Информатика и образование, 2005, № 5