

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА
И ПРОДОВОЛЬСТВИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

ГЛАВНОЕ УПРАВЛЕНИЕ ОБРАЗОВАНИЯ, НАУКИ И КАДРОВ

УЧРЕЖДЕНИЕ ОБРАЗОВАНИЯ
«БЕЛОРУССКАЯ ГОСУДАРСТВЕННАЯ
СЕЛЬСКОХОЗЯЙСТВЕННАЯ АКАДЕМИЯ»

Н. Г. Трапянок

ОСНОВЫ ПЕДАГОГИЧЕСКИХ ИЗМЕРЕНИЙ

Курс лекций

*для слушателей Института повышения квалификации
и переподготовки кадров, обучающихся по специальности
1-08 01 71 Педагогическая деятельность специалистов*

Горки
БГСХА
2018

УДК 37.013(075.8)

ББК 74.00

Т65

Рекомендовано научно-методическим советом ИПКиПК

30.05.2016 (протокол № 4)

и Научно-методическим советом БГСХА

29.06.2016 (протокол № 10)

Автор:

кандидат социологических наук, доцент *Н. Г. Трапьянок*

Рецензенты:

доктор социологических наук, доцент *О. В. Кобяк*;

кандидат социологических наук, доцент *И. А. Андрос*

Трапьянок, Н. Г.

Т65 Основы педагогических измерений : курс лекций /
Н. Г. Трапьянок. – Горки : БГСХА, 2018. – 76 с.
ISBN 978-985-467-761-3.

Курс лекций подготовлен на основе программы по учебной дисциплине «Основы педагогических измерений». Целью его является обеспечение слушателей лекционным материалом для самостоятельной подготовки и совместной с преподавателем работы по изучению дисциплины. Представлены материал лекционных занятий, вопросы для промежуточного контроля, задания для самостоятельной работы, глоссарий и рекомендуемая литература по каждой теме.

Для слушателей Института повышения квалификации и переподготовки кадров, обучающихся по специальности 1-08 01 71 Педагогическая деятельность специалистов.

УДК 37.013(075.8)

ББК 74.00

ISBN 978-985-467-761-3

© УО «Белорусская государственная
сельскохозяйственная академия», 2018

ПРЕДИСЛОВИЕ

Наука начинается там,
где начинают измерять.

Д. И. Менделеев

Одним из важнейших факторов повышения качества профессионального образования является системное использование измерений. Педагогические измерения в условиях современного образовательного процесса направлены на оптимизацию структуры и содержания учебных дисциплин, тематического планирования, модернизации всей учебно-воспитательной работы, тесно связанной с внедрением блочно-модульной технологии и рейтинговой системы ее организации, повышением мотивации обучающихся к обучению и т. д.

Современная педагогическая практика невозможна без знаний, навыков и умений в области образовательной метрологии, квалиметрии и тестологии, базирующихся на применении современных информационных технологий в обучении и исследовательской работе.

Роль педагогических измерений в образовании не сводится к констатации учебных достижений испытуемых. Область применения результатов интегративного педагогического измерения распространяется на определение возможности формирования компетентности обучающихся, структурирование и синхронизацию взаимосвязей между элементами учебных дисциплин, корректировку тематических учебных планов.

Всего на изучение дисциплины «Основы педагогических измерений» отводится 20 аудиторных часов: лекционных – 8, семинарских занятий – 12 и 16 часов – на самостоятельную работу.

Курс лекций разработан с учетом практической значимости материала для педагогической деятельности преподавателей, методистов и научных работников учреждений образования, а также его доступности для понимания самих обучающихся и ориентирован на самостоятельную подготовку слушателей, не имеющих специальных знаний в области математической статистики и квалиметрии. В этой связи особое внимание уделено методологическим проблемам измерений в образовании при оценке достижений обучающихся, проведении педагогических исследований, обработке и анализе их результатов.

Курс лекций предназначен для слушателей системы переподготовки педагогических кадров, но может быть использован в практической работе методистов, преподавателей и студентов (учащихся) учреждений высшего и среднего специального образования, а также полезен всем тем, кто интересуется проблемами педагогических измерений.



Л е к ц и я 1. МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ ПЕДАГОГИЧЕСКИХ ИЗМЕРЕНИЙ

Человек есть мера всех вещей – сущих в
их бытии и несущих в их небытии.

Протагор

- 1.1. История становления и развития измерений.
- 1.2. Теоретико-методологические основы педагогических измерений.
- 1.3. Основные компоненты и критерии педагогических измерений.
- 1.4. Роль измерений в современной педагогической практике и значение учебного курса в системе педагогического образования.

1.1. История становления и развития измерений

Потребность человека в осуществлении измерений является для него практически вынужденной и появилась еще в незапамятные времена.

Измерение возникло из необходимости правильно ориентироваться на местности, различать предметы и явления, сравнивая их между собой и обозначая символами (так появилась шкала наименований); из потребности постоянного слежения за сменой сезонов, дней и ночей и упорядочивания их в своем сознании, отдавая приоритет одним вещам и событиям над другими (шкала порядка); из-за зависимости человека от их протяженности и продолжительности он был вынужден производить их отсчет через равные интервалы равной мерой (метрическая шкала).

В античной мифологии символом меры была богиня Немезида (в римской – Фемида была заимствована с греческой), изображавшаяся с весами в руках, поскольку выполняла функции кары и возмездия за преступные деяния человека. Она же почиталась как богиня справедливости, равенства, закона и была устроительницей космического порядка. Так, согласно Еврипиду, Немезида ответственна за чередование ночи и дня, более того, «И меру нам и вес она дает и числа образует» (Финикиянки 541–545). Мера имела и в античной мифологии, и у досократиков этическое и космическое значение. Согласно Гесиоду, «Меру в словах соблюдешь – и всякому будешь приятен» (Работы и дни, 694, пер. В. Вересаева). Мера, понимаемая как умеренность в поступках и словах, – одно из центральных понятий первых античных философов (Клеобул: «Мера лучше всего»; Солон: «Ничего слишком»; Питтак: «Знай меру»; Фалес: «Блюдай меру». – В кн.: Фрагменты ранних греческих философов. М., 1989. Ч. 1. С. 92–93). Одновременно мера имела и космическо-онтологическое значение, характеризуя способ бытия первоначал и их пропорции.

История становления и развития измерений говорит нам о том, что представления об измерении вначале формировались как об абсолютной оценке качества. Сначала люди определяли качество предметов, а затем уже фиксировали их количественные отличия. Измерить означало применить к предметам равную меру. Сократ, Платон и Аристотель рассматривали меру как одну из высших добродетелей. У Пифагора мера приобретает космический характер, она управляет гармонией чисел (количества). У поздних философских школ античности категория меры выступает как категория эстетическая, этическая и правовая.

Основоположнику движения софистов Протагору принадлежит знаменитое положение «человек есть мера всех вещей», в котором он говорил об относительности всякого знания, доказывая, что каждому утверждению может быть противопоставлено противоречащее ему утверждение. Протагор написал законы для города Фурий, определявшие демократический образ правления, и обосновал равенство свободных людей.

Представления о мере также трансформировались в представления о порядке как об определенной последовательности и субординации. Измерить означало определить порядок предпочтений. Следующий этап формирования представлений об измерении – это установление величины различий. До этого рассматривалась величина как качественная определенность, например, один предмет маленький, другой – большой. Здесь же устанавливаются различия, насколько он стал больше или меньше в количественном отношении.

За тысячи лет до нашей эры Древний Египет стал свидетелем появления «Книги мертвых», которую можно считать первой документированной системой измерения на Земле. Она содержала подробное описание процедур, которым необходимо было следовать для подготовки тела и души к загробной жизни. Открытие гробницы Тутанхамона показало, что эти процедуры были безукоризненно выполнены, а состояние многих мумий доказало эффективность системы в обеспечении соответствия требованиям.

Примерно в XI веке было введено клеймение для золотых и серебряных изделий с указанием минимального содержания благородных металлов в сплаве.

Начиная со средневековья, измерение изделий стало целью квалифицированных ремесленников. Они приобретали репутацию лучших производителей за свое мастерство, которого достигали годами вынужденного и сурового обучения и стремлением добиться успеха в торговле. Чтобы обеспечить высокий уровень качества продукции, ремесленники объединялись в монопольные гильдии, совершенствуя торговые отношения.

В индустриальном обществе, где стандартизация, спецификация и сертификация являются характерными чертами массового потребления и производства, измерения превратились в мощную отрасль, из которой выделился ряд направлений, обслуживающих промышленность, сельское хозяйство, управление, банки, сферу услуг, торговлю, здравоохранение, образование, науку и культуру.

Проблема измерений и оценки их результатов на сегодняшний день является одной из самых важных в педагогической теории и практике, а педагогические измерения стали неотъемлемой составной частью учебного процесса и образовательного пространства.

1.2. Теоретико-методологические основы педагогических измерений

Философские основания измерения. Процесс познания реального мира как исторически, так и логически совершается таким образом, что познание качества вещей предшествует познанию количественных отношений. Наука движется от качественных оценок и описаний явлений к установлению количественных закономерностей. Количество находится в единстве с качественной определенностью явлений, вещей, процессов; это единство составляет их меру. Изменение количественной определенности вещей до известного предела не затрагивает их качества. За этими пределами количественные изменения сопровождаются изменением качества.

Что означает слово «измерить»? Измерить – исходить из меры.

Мера же – понятие не математическое, не физическое и не юридическое, а категория философского дискурса, фиксирующая и обобщающая результаты и процедуры измерения качественно-количественной определенности предметов, процессов и их взаимоотношений с теми средствами измерения, которые выступают в роли эталона.

В противовес *квалитативизму* («квали» – *качество*), когда речь идет только о качественном описании, и *квантификационизму* («кванти» – *количество*), где абсолютизируются количественные параметры и величины, Гегель рассматривал все во взаимосвязи и определял меру как «третью ступень бытия», «единство первых двух, качественного и количественного».

Анализ меры и мерных отношений исходит из единства качественных и количественных характеристик, из важности, с одной стороны, количественных изменений для существования качественной определенности, для возникновения нового качества и, с другой стороны, качественной определенности для порогового значения количественных изменений, интервала изменений количественных величин, в рамках которого можно говорить о сохранении качества исследуемого предмета или процесса. Категория меры тесно связана с такими понятиями философского тезауруса, как гармония, порядок, умеренность, ритм и др.

В теории оценочной функции психики (Н. А. Батулин, И. В. Выбойщик, Е. В. Гудкова) глобальными отношениями называется особая группа оценочных явлений, представленных в виде свойств личности, отражающих пристрастную позицию по отношению к внутренней и внешней реальности. В структуре межличностных отношений В. Н. Мясичев рассматривает оценку, определяя класс «оценочных отношений» как те, которые формируются на основе этических, эстетических, правовых и других общественных критериев поведения и жизнедеятельности людей. Через оценку определяется нормативность этих взаимоотношений в различных формах их проявления.

Оценка предполагает наличие критериев, эталонов, единиц измерения оцениваемых свойств, которые в процессе сравнения, анализа и синтеза служат отправной точкой для вынесения суждения о качестве и уровне развития явления, характере его связей и позволяют соотносить процессы и явления. Оценка – это форма проявления отношения, его сознательная объективация. Оценочные отношения определяются в контексте мыслительной деятельности человека, включающей эмоциональный компонент в виде одобрения (неодобрения), и могут в

этой связи оказаться проекцией самооценки, что часто случается в оценке людьми друг друга.

В оценках себя и других людей проявляются отношения самоуважения и уважения других, которые трансформируются во взаимоотношения авторитетности или авторитарности руководства (подчинения) и др. Оценочные отношения в этом процессе выполняют функцию преобразователя субъективных отношений в субъект-субъектные и субъект-объектные.

Конвенциональность измерений. Измерение – это процедура присвоения символов наблюдаемым объектам в соответствии с некоторым правилом. Символы могут быть просто метками (маркёрами), представляющими собой классы или категории объектов в популяции, или числами, характеризующими степень выраженности у объекта измеряемого свойства. Символы-метки могут также представлять собой числа, но при этом не обязательно нести в себе характерную «числовую» информацию.

Существует значительное число классификаций знаков, основанных на различиях формы, содержания, связи формы с содержанием и других параметров. До настоящего времени сохраняет свое значение классическое (введенное Ч. С. Пирсом) разбиение знаков на три группы: **иконы**, **индексы** и **символы**. Эта классификация основана на типологии соотношения формы и содержания.

Иконами называются знаки, чьи форма и содержание сходны качественно или структурно. Например, батальное полотно или план сражения являются знаками-иконами, если считать их содержанием само сражение.

Индексы – это знаки, форма и содержание которых смежны в пространстве или во времени. Следы на песке, позволяющие предположить о том, что ранее в этом месте кто-то прошел, дым, предполагающий наличие огня, симптомы болезни, предполагающие саму болезнь, – все это индексальные знаки. По-видимому, точнее было бы говорить не о смежности формы и содержания, как это принято, а о наличии между ними причинно-следственных связей.

Символами в узком значении принято называть знаки, для которых связь между формой и содержанием устанавливается произвольно, по соглашению, касающемуся именно данного знака. Для иконических и индексальных знаков форма позволяет догадаться о содержании знака даже не знакомому с ним адресату. Что же касается символических знаков, то их форма сама по себе, т. е. вне специального договора, не дает никакого представления о содержании. Ф. де Соссюр говорил в

этом случае о немотивированности выбора означающего или отсутствия естественной связи между означаемым и означающим.

Например, знак сложения (+) никак не связан с самой этой арифметической операцией: ни сходством, ни смежностью, ни причинно-следственными связями. Их связь произвольна в том смысле, что определяется особым соглашением, или конвенцией, обуславливающей использование соответствующего значка для передачи данного смысла.

Аксиологические (ценностные) основания измерений. Измерение тесно связано с оценкой в силу следующих причин:

оно имеет своей основой пристрастно-оценочное отношение человека к миру;

в основе измерения лежит оценка – форма проявления отношения субъекта к объекту, его сознательная и эмоциональная объективация;

в результате такой оценки выявляется положительное (отрицательное) значение объекта для субъекта;

эти значения закрепляются не биологическим, а социальным кодом, выраженным символом или суждением;

измерение – это формализованная процедура присвоения символов наблюдаемым объектам в соответствии с некоторым правилом;

оно предполагает наличие меры – критериев, эталонов, единиц отсчета оцениваемых свойств, которые служат отправной точкой для вынесения суждения о их качестве и количестве.

1.3. Основные компоненты и критерии педагогических измерений

Основные понятия и определения. Как категория диалектики мера есть то, что «снимает» противоречие между количественной и качественной определенностями и воплощает их единство.

Измерение – это процедура формализованного оценивания какого-либо объекта путем соотнесения с заданной адекватной мерой и определения его значения (количественного или качественного) путем приписывания ему по определенным правилам условного обозначения (числа, буквы, слова или какого-либо другого символа).

Измерение основывается на соответствии чего-то чему-то. Для установления соответствия необходимо выполнение следующих условий:

1. Должен быть определен (ясен) эталон (образец) соответствия.
2. Должна быть установлена мера достижения соответствия.
3. Всегда должен быть понятен, ясен механизм, позволяющий установить и проверить это соответствие.
4. Соответствие должно быть измерено, охарактеризовано и представлено в одних единицах, в одних параметрах как соответствие чему-то.

Педагогические измерения (в зарубежной литературе – *Educational Measurements*) – научная теория и практика, сформировавшаяся в течение последнего века на стыке педагогики, психологии, общей (фундаментальной) теории измерений, статистики, математики, логики и философии. Их можно также рассматривать как практическую образовательную деятельность, нацеленную на получение объективных, а точнее, объективированных оценок уровня текущей и итоговой подготовленности обучающихся.

Измерение в образовании – формализованная процедура соотнесения образовательного процесса или явления (системы) с установленным стандартом (нормой) по заданной шкале и определение (истолкование) его значения.

Квалиметрия – формальная теория меры качества, выражающая ее количественно. Квалиметрия в образовании строится на различении понятий «измерение» и «оценка», «оценка» и «отметка», «количественное измерение» и «качественная оценка».

Оценка – способ установления значимости чего-либо для действующего и познающего субъекта, а также его результат и суждение о ценности, значении, качестве, признаках какого-либо объекта.

Отметка – выраженная словесно либо представленная графически, в виде какого-либо знака (символа) оценка.

Критерии педагогических измерений: объективность, достоверность, валидность, надежность, правильность, точность, сравнимость (рис. 1.1).

Объективность	первичные показатели и исход измерения не должны зависеть от наблюдателя
Достоверность	это характеристика, определяющая степень доверия к полученным результатам измерений
Валидность	характеризует соответствие измерения его цели – обоснованной и ясной концептуализации теоретических представлений
Надежность	воспроизводимость результатов измерения, повторяемого при идентичных условиях
Правильность	измерительное действие по определенным правилам и близость к нулю систематических погрешностей результата измерений
Точность	насколько близки результаты измерений к истинному значению измеряемой величины
Сравнимость	результаты измерений должны быть сопоставимыми и сравнительными

Рис. 1.1. Критерии педагогических измерений

Принцип измерений – это некое явление или комплекс явлений, на которых базируется измерение.

Важно отметить, что надежность инструмента измерения не отражает его точность и правильность. Оценка надежности-повторяемости – это корреляция между результатами повторных измерений. В случае единичного вопроса-индикатора достаточно просто сопоставить ответы одной и той же группы людей, полученные в первом опросе, с ответами, данными двумя неделями или месяцем позднее, и вычислить коэффициент корреляции. На практике хорошей можно считать корреляцию 0,8 и выше.

Средства измерения – совокупность методических и технических средств, используемых при измерении и имеющих нормированные свойства, в частности, по единицам и точности измерений, надежности и воспроизводимости результатов, а также требованиям к «входным» и «выходным» параметрам и др.

Результаты измерений в обязательном порядке подвергаются оценке на точность, результативность и соответствие поставленным целям.

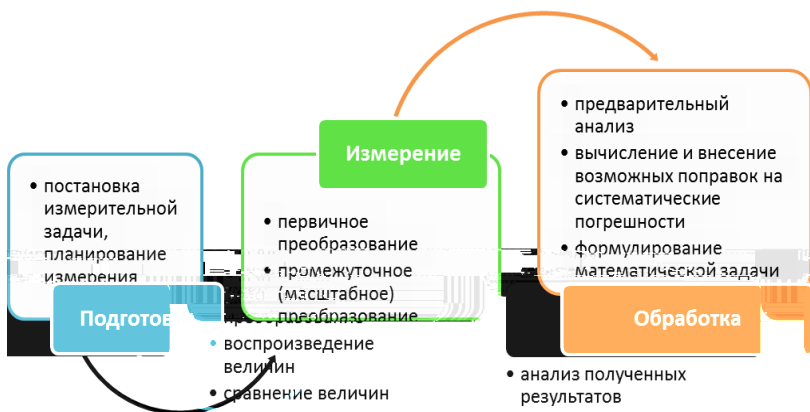


Рис. 1.2. Этапы педагогических измерений

Этапы педагогических измерений. На рис. 1.2 представлены основные этапы измерений. К ним относятся следующие:

1. Подготовка: постановка измерительной задачи, планирование измерительного эксперимента.

2. Измерительный эксперимент.

3. Обработка экспериментальных данных.

Постановка задачи в полном объеме включает в себя:

- уточнение данных об условиях измерения;
- формирование модели объекта и определение измеряемой величины;

- постановку измерительной задачи на основе принятой модели объекта;

- выбор конкретных величин, на основе которых находится значение измеряемой величины;

- создание алгоритма измерений.

Планирование измерений в полном объеме включает в себя:

- выбор метода измерений;
- априорную оценку погрешностей измерений («предположение»);
- формулирование требований к метрологическим характеристикам средств измерений и условиям измерения;

- выбор средств измерения в соответствии с требованиями;
- выбор параметров измерительной процедуры (точки измерений);
- подготовку средств измерений;
- обеспечение требуемых условий. Для многих измерений разрабатывают стандартизованные методики.

Измерительный эксперимент в полном объеме включает в себя:

- первичное преобразование;
- промежуточное (масштабное) преобразование;
- воспроизведение физических величин;
- сравнение величин.

Обработка данных в полном объеме включает в себя:

- предварительный анализ информации, полученной на предыдущих этапах;

- вычисление и внесение возможных поправок на систематические погрешности;

- формулирование математической задачи обработки данных;

- алгоритмы обработки данных (результатов измерения и погрешности);

- проведение вычислений;

- анализ полученных результатов.

Виды измерений. Измерения различают по способу получения информации, характеру изменений измеряемой величины, количеству измерительной информации, отношению к основным единицам.

По способу получения информации измерения подразделяют на *прямые, косвенные, совокупные и совместные*.

Прямые измерения – это непосредственное сравнение физической величины с ее мерой. Например, измерение массы на весах, длины с помощью линейных мер.

Косвенные измерения – измерения, при которых искомое значение величины находят на основании известной зависимости между этой величиной и величинами, подвергаемыми прямым измерениям. Например, нахождение плотности тела по массе и размерам.

Совокупные измерения сопряжены с решением системы уравнений, составляемых по результатам одновременных измерений нескольких однородных величин. Решение системы уравнений дает возможность вычислить искомую величину.

Совместные измерения – это измерения двух и более неоднородных величин для определения зависимости между ними.

По характеру изменения измеряемой величины выделяют *статические, динамические и статистические* измерения.

Статические измерения имеют место тогда, когда измеряемая величина практически постоянна.

Динамические измерения связаны с такими величинами, которые в процессе измерений претерпевают те или иные изменения.

Статистические измерения связаны с определением характеристик случайных процессов, звуковых сигналов, уровня шумов.

По отношению к основным единицам измерения подразделяют на *абсолютные и относительные*.

Абсолютными измерениями называют такие, при которых используются прямое измерение одной (иногда нескольких) основной величины и физическая константа. Так, в формуле $E = mc^2$ масса (m) – основная физическая величина, которая может быть измерена прямым путем (взвешиванием), а скорость света (c) – физическая константа.

Относительные измерения основаны на установлении отношения измеряемой величины к применяемой в качестве единицы измерения.

Выделяют также *однократные и многократные* измерения.

Однократные измерения – это одно измерение одной величины. Практическое применение такого вида измерений всегда сопряжено с большими погрешностями, поэтому следует проводить не менее трех однократных измерений и находить конечный результат как среднее арифметическое значение.

Многократные измерения характеризуются повышением числа измерений количества измеряемых величин. Обычно минимальное число измерений в данном случае больше трех. Преимущество многократных измерений заключается в значительном снижении влияния случайных факторов на погрешность измерений.

Метод измерений – это совокупность приемов использования средств измерений. В педагогике находят применение следующие методы измерений:

метод непосредственной оценки, при котором значение величины определяют непосредственно по отсчетному устройству (давление – манометром, характеристика тока – амперметром, вольтметром);

метод сравнения с мерой, когда измеряемую величину сравнивают с величиной, воспроизводимой мерой (сравнение масс на весах с гирями);

метод совпадений, при котором разность между измеряемой величиной и величиной, воспроизводимой мерой, измеряют по совпадению отметок шкал;

нулевой метод, в котором результирующий эффект воздействия величин на прибор сравнения равен нулю;

метод прямого ранжирования является приемом измерения, в результате применения которого получается порядковый уровень измерения;

метод парных сравнений (шкала Терстоуна) основан на попарном сравнении объектов ранжирования по заданному основанию;

метод замещения, в котором измеряемую величину замещают известной величиной, воспроизводимой мерой, например, взвешивание с поочередным помещением измеряемой массы и гирь на одну и ту же чашку весов.

1.4. Роль измерений в современной педагогической практике и значение учебного курса в системе педагогического образования

Практика показывает, что измерения затрагивают практически все сферы и уровни образования, всех его субъектов.

Современный педагог должен владеть как методикой и техникой, так и теоретическими основами педагогических измерений. Однако, несмотря на их востребованность, проблема далека от своего решения,

что во многом объясняется отсутствием единства между исследователями по принципиальным вопросам и должного уровня разработок по теоретико-методологическим аспектам педагогических измерений. Этим, в частности, объясняется и то, что большинство учебников и пособий по их основам в основном ограничивается вопросами квалитметрии и тестологии.

Активно развивающиеся технологии массового тестирования находят широкое применение в профессиональном образовании. В перспективе тестовый контроль как базовый станет основой для получения обоснованной и надежной информации о качестве всей системы подготовки высококвалифицированных кадров в стране. В то же время необходим системный мониторинг образования, объединяющий результаты качественных и количественных методов, тестирования и традиционного оценивания.

В настоящее время для организации такого мониторинга только появляются необходимые условия. К ним относятся: проведение массового тестирования (централизованное тестирование, другие формы независимой аттестации); формирование банка стандартизированных педагогических измерителей, баз данных по образовательной статистике и результатам тестового контроля; использование информационных технологий накопления, представления и передачи данных педагогических измерений для различных категорий пользователей.

Таким образом, в условиях стандартизации и интенсивного развития национальной системы образования происходит внедрение и развитие непрерывной, целостной и комплексной системы многоуровневого измерения и оценивания как средства управления качеством образования.

Вопросы для промежуточного контроля

1. Что представляет собой измерение и какова его природа?
2. Можно ли измерение отделить от оценки?
3. Чем отличается отметка от оценки?
4. Всегда ли отметки в измерении имеют числовую (математическую) природу? Всегда ли число выражает количество?
5. Что представляет собой валидность измерительных средств, и почему измерительные средства и результаты измерений нуждаются в валидации (валидизации)?

6. Перечислите и охарактеризуйте основные этапы измерения с момента его организации и подготовки до завершения.

7. Чем вызван повышенный интерес в обществе к проблеме педагогических измерений и что обостряет ее, делая более актуальной?

Вопросы и задания для самостоятельной работы

Операционализация понятий, определение индикаторов, методы построения индексов. Понятие переменной и ее значения. Латентная и наблюдаемая переменные.

Задание 1. Осуществите факторную операционализацию понятия «успеваемость студентов»; сконструируйте систему эмпирически наблюдаемых и латентных переменных; укажите на методы получения данных и их анализа.

Задание 2. Проведите структурную операционализацию понятия «дисциплина студентов»; постройте систему индексов и индикаторов для измерения.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. А н а н ь е в, Б. Г. Психология педагогической оценки / Б. Г. Ананьев // Избранные психологические труды: в 2 т. – Санкт-Петербург: Изд-во С.-Петерб. ун-та, 2007. – Т. 2. – 389 с.
2. Б о р ы т к о, Н. М. Методология и методы психолого-педагогических исследований: учеб. пособие для вузов / Н. М. Борытко, А. В. Моложавенко, И. А. Соловцова; под ред. Н. М. Борытко. – Москва: Академия, 2008. – 289 с.
3. В о л к о в, Б. С. Методология и методы психологического исследования: учеб. пособие для вузов / Б. С. Волков, Н. В. Волкова, А. В. Губанов. – Изд. 4-е, испр. и доп. – Москва: Академический Проект, 2005. – 256 с.
4. Г у с е в, А. Н. Измерение в психологии. Общий психологический практикум / А. Н. Гусев, Ч. А. Измайлов, М. Б. Михалевская. – Москва, 1997. – 433 с.
5. О б р а з ц о в, П. И. Методы и методология психолого-педагогического исследования: учеб. пособие / П. И. Образцов. – Санкт-Петербург: Питер, 2004. – 390 с.
6. Т а т а р и н о в а, Г. Г. Методология анализа данных в социологии / Г. Г. Татарина. – Москва: NOTA BENE, 1999. – 224 с.



Л е к ц и я 2. МЕТОДЫ КОЛИЧЕСТВЕННОЙ ОЦЕНКИ И ШКАЛИРОВАНИЕ

Лучшее количество всегда там, где
отсутствует расхождение с качеством.

Франсуа де Ларошфуко

2.1. Количество. Метрические шкалы, используемые в педагогическом измерении.

2.2. Элементарные понятия анализа данных измерения.

2.3. Сопряженность признаков и теснота связи.

2.4. Методы количественной оценки образовательного процесса.

2.1. Количество. Метрические шкалы, используемые в педагогическом измерении

Количество – категория, выражающая внешнее, формальное взаимоотношение предметов или их частей, а также свойств, связей: их величину, число, степень проявления того или иного свойства.

Как особую категорию количество рассматривал Аристотель: «Количеством называется то, что может быть разделено на составные части, каждая из которых, будет ли их две или несколько, является чем-то одним, данным налицо. То или другое количество есть множество, если его можно счесть, это – величина, если ее можно измерить».

Каждая совокупность предметов есть некоторое множество. Если оно конечное, то его можно сосчитать. Всякий счет состоит в повторенном полагании единицы. Например, число 40 является количественной характеристикой любого множества из 40 предметов, будут ли это люди или деревья. Следовательно, числа и величины оказываются

формальной, внешней, по Гегелю, «равнодушной» стороной качественных отношений. Есть вещи большие и маленькие, длинные и короткие, есть движения быстрые и медленные, есть степень развития чего-либо высокая и низкая и т. д. Все это можно измерить с помощью определенного эталона: метра, секунды и т. д.

С целью установления количественной определенности предмета сравниваются составляющие его элементы – пространственные размеры, скорость изменения, степень развития – с определенным *эталон*ом как единицей счета и измерения. Чем сложнее явление, тем труднее его подвергать изучению с помощью количественных методов (например, явления в сфере нравственности, политики, эстетического восприятия мира и т. п.); в этих случаях прибегают к различного рода шкалам.

Процедура шкалирования. Измерение – процедура присвоения рубрикационных символов (знаков) наблюдаемым объектам в соответствии с некоторым правилом (*О. В. Терещенко*). Измерение осуществляется с помощью шкал.

Шкала измерения (оценки) – принятая система градации присвоения символа объекту, представляющая собой упорядоченный ряд отметок, позволяющий количественно или качественно определить величину показателей, значение параметров объекта.

В образовании применяются следующие типы простых шкал: *номинальная* шкала (наименований), шкала *порядка* (ранговая), шкала *интервалов* (метрическая шкала, на которой начало отсчета произвольно, а единица измерения задана) и шкала *отношений* (метрическая шкала с естественным и однозначным присутствием нулевой точки отсчета).

Шкалы подразделяются на *простые* и *сложные* (комбинированные из нескольких шкал), *одномерные* и *многомерные*.

Простые шкалы измерений подразделяются на два типа: *неметрические* (качественные) и *метрические* (количественные), а также на четыре класса: наименований, порядковая, интервальная и отношений.

К количественным шкалам измерений относятся две последние: интервальная и отношений (нередко называемая «абсолютная шкала» из-за ее универсальных свойств).

Шкала интервалов – реже применяющаяся в педагогике и менее мощная. Примером ее являются шкалы температур по *Цельсию*, *Реомюру* или *Фаренгейту*. Шкала Цельсия, как известно, установлена следующим образом: за ноль принималась точка замерзания воды, за 100 °С – точка ее кипения, и, соответственно, интервал температур

между замерзанием и кипением воды поделен на 100 равных частей – градусов. Здесь уже утверждение, что температура в 30 °С в три раза больше, чем 10 °С, будет неверным. Справедливо говорить лишь об интервалах температур – температура в 30 °С на 20 °С больше, чем температура в 10 °С.

Шкала отношений – это самая мощная шкала. Она позволяет оценивать, во сколько раз один измеряемый объект больше (меньше) другого объекта, принимаемого за эталон, единицу измерения. Одновременно здесь возможно и сравнение: на сколько один объект больше (меньше) другого. Шкалами отношений измеряются практически все физические величины – время, размеры, площадь, объем, напряжение, сила тока, мощность и т. д. В педагогических измерениях шкала отношений будет иметь место, например, когда измеряется время выполнения того или иного задания, количество ошибок или количество правильно решенных задач между контрольными и экспериментальными группами, классами.

2.2. Элементарные понятия анализа данных измерения

Что такое *переменная*?

Переменная (английский термин *variable*) – это то, что можно измерять, контролировать. Иными словами, переменная (от англ. корня *var*) – это то, что варьируется, изменяется, а не является постоянным.

Например, измеряя давление или содержание лейкоцитов в крови, получают различные значения у разных пациентов или значения для одного и того же пациента в разное время суток. Измеряя уровень осадков, получают различные значения в разные дни недели, а также различные значения в одни и те же дни в разных точках географической карты.

Другие примеры переменных из разных областей: анкетные данные, систолическое давление пациентов, количество лейкоцитов в крови, цена акций, товаров, услуг, объемы инвестиций, доходов, государственных закупок; рейтинг телевизионных программ, доля зрителей, количество посещений сайта и т. д.

Переменная – некоторое общее для всех объектов качество (признак), конкретные проявления которого могут меняться.

Значения переменной – это те показания переменной, которые могут изменяться.

В зависимости от типа и класса шкал, применяемых для измерения переменных, различают *номинальную, порядковую, интервальную и относительную* переменные.

Латентная переменная – это переменная, которая не поддается непосредственному наблюдению или измерению. Например, любая социальная установка носит *латентный* (скрытый) характер.

Непосредственному наблюдению и измерению с помощью прямо поставленных вопросов она не поддается. Возникает необходимость в построении моделей изучения этой установки посредством доступных наблюдению признаков, из которой конструируется **индекс** латентной переменной.

Индекс – это обобщенная оценка латентной переменной, полученная путем агрегирования результатов измерения наблюдаемых переменных.

Примерами применения латентных переменных в образовании могут служить измерения уровней успеваемости обучающихся, сформированности у них компетенций, общеучебных и предметных умений и навыков.

Очевидно, что это очень разные по своим свойствам переменные, и поэтому можно сказать, что переменные отличаются характеристиками, в частности, типом измерений, той ролью, которую они призваны играть в исследованиях и т. д.

Поскольку значения переменных не постоянны, то нужно научиться описывать их изменчивость. Для этого существуют *простейшие дескриптивные (описательные) статистики*: минимум, максимум, среднее, дисперсия, стандартное отклонение, медиана, квартили, мода и т. д.

Среднее – сумма значений переменной, деленная на n (число значений переменной).

Среднее арифметическое. Для любой совокупности значений признака это сумма всех значений, деленная на их число. Возьмем, например, признак продолжительности затрат времени на учебу (табл. 2.1).

Обозначим число студентов-гуманитариев через n (для данного случая $n = 1000$), а через X_i – значение этой продолжительности для i -го студента. Тогда средняя арифметическая продолжительности будет равна:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Таблица 2.1. Частотные распределения

Показатель	Продолжительность учебы, лет									Итого
	0–1	1–2	2–3	3–4	4–5	5–6	6–7	7–8	8–9	
Абсолютная частота	27	50	75	100	48	100	200	250	150	1000
Относительная частота, %	2,7	5,0	7,5	10	4,8	10	20	25	15	100
Накопленная частота	2,7	7,7	15,2	25,2	30	40	60	85	100	

Среднее взвешенное. Формула для вычисления средней взвешенной выглядит для k интервалов следующим образом:

$$\bar{X} = \frac{\sum_{j=1}^k p_j x_j}{\sum_{j=1}^k p_j}, \text{ где } p_j - \text{коэф. весомости.}$$

Минимум и максимум – это минимальное и максимальное значения переменной.

Медиана разбивает выборку на две равные части. Половина значений переменной лежит ниже медианы, половина – выше.

Медиана менее чувствительна к разбросам значений, чем среднее. В официальной статистике США не среднее, а именно медиана используется в оценке центральной точки доходов населения.

Мода – наиболее часто встречающееся значение переменной.

Удельный вес

Процент (от лат. *per cent* – на сотню) – одна сотая доля. Обозначается знаком «%». Используется для обозначения доли чего-либо по отношению к целому. Например, 17 % от 500 кг означает 17 частей по 5 кг каждая, т. е. 85 кг.

Кумулятивный процент – накапливаемая величина процента.

Валидный процент – рассчитывается от числа фактических значений (от числа ответивших).

рассчитается как для фактических, так и для пропущенных значений от числа опрошенных).

Промилле (от лат. *pro mille* – на тысячу) – одна тысячная доля, 1/10 процента. Обозначается «‰». Используется для обозначения количества тысячных долей чего-либо в целом. Происходит (как и процент) от написания простой дробью: $27/1000 \rightarrow 27 \text{ ‰}$.

Меры вариации и рассеяния

Совокупность значений, характеризующая изменение показателя во времени, называется **динамическим рядом**.

Вариационный ряд – изменение показателя в пространстве.

Тренд (от англ. *trend*) – это основная тенденция динамики показателя, очищенная от случайных влияний и индивидуальных особенностей отдельных периодов.

Одним из показателей вариации количественной переменной является **дисперсия**. Она же является самой распространенной оценкой рассеивания случайной величины.

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

Дисперсия вычисляется как средняя из квадратов отклонений вариантов значений признака от их средней величины:

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

В статистическом анализе вариацию признака принято оценивать чаще всего с помощью *среднего квадратического (стандартного) отклонения* (σ). Среднее квадратическое отклонение представляет собой корень квадратный из дисперсии (σ^2), или корень квадратный из суммы квадратов разниц между элементами выборки и средним, деленной на количество элементов в выборке. Чем выше дисперсия или стандартное отклонение, тем сильнее разбросаны значения переменной относительно среднего. Часто стандартное отклонение – более удобная характеристика, так как измерена в тех же единицах, что исходная величина.

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

Коэффициент вариации – наиболее часто применяемый показатель относительной колеблемости, характеризующий однородность совокупности. Совокупность считается однородной, если коэффициент вариации не превышает 33 % для распределений, близких к нормальному.

$$V = \frac{\sigma}{\bar{x}} \times 100\%$$

Нормальное распределение представляет собой одну из эмпирически проверенных истин относительно общей природы действительности, и его положение может рассматриваться как один из фундаментальных законов природы.

Точная форма нормального распределения (характерная «колоколообразная кривая») определяется только двумя параметрами: средним и стандартным отклонением.

Кривая нормального распределения (К.-Ф. Гаусса) выглядит следующим образом (рис. 2.1).

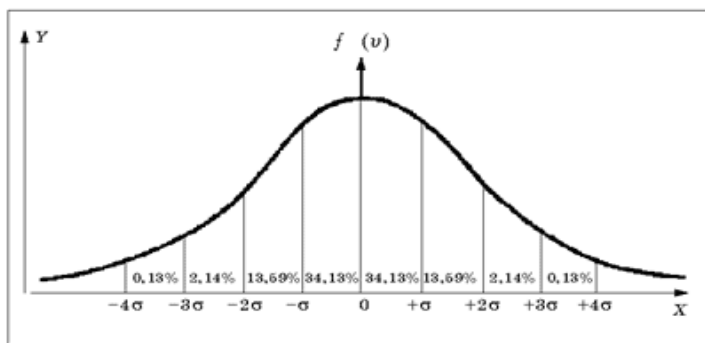


Рис. 2.1. Кривая нормального распределения Гаусса

Так, 68 % из всех его наблюдений лежат в диапазоне $-\sigma \leq 0 \leq +\sigma$ стандартного отклонения от среднего, а диапазон $-2\sigma \leq 0 \leq +2\sigma$ стандартных отклонений включает 95 % значений. Другими словами, при нормальном распределении стандартизованные наблюдения, меньшие -2σ или большие $+2\sigma$, имеют относительную частоту менее 5 % (стандартизованное наблюдение означает, что из исходного значения вычтено среднее и результат поделен на стандартное отклонение).

Это и есть знаменитое правило двух сигм (2σ), или двух стандартных отклонений, вместе с правилом трех сигм (3σ) чрезвычайно популярное на практике.

2.3. Сопряженность признаков и теснота связи

Таблица 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 (табл. 2.2).

Т а б л и ц а 2.2. Таблица сопряженности

Будущая профессия студента	Степени удовлетворенности учебой					
	1	2	3	4	5	n_{10}
1. Политолог	$n_{11} = 14$	$n_{12} = 20$	$n_{13} = 31$	$n_{14} = 30$	$n_{15} = 5$	$n_{10} = 100$
2. Социолог	$n_{21} = 30$	$n_{22} = 40$	$n_{23} = 60$	$n_{24} = 60$	$n_{25} = 10$	$n_{20} = 200$
3. Культуролог	$n_{31} = 90$	$n_{32} = 90$	$n_{33} = 60$	$n_{34} = 45$	$n_{35} = 15$	$n_{30} = 300$
4. Филолог	$n_{41} = 31$	$n_{42} = 30$	$n_{43} = 19$	$n_{44} = 15$	$n_{45} = 5$	$n_{40} = 100$
5. Психолог	$n_{51} = 8$	$n_{52} = 10$	$n_{53} = 15$	$n_{54} = 15$	$n_{55} = 2$	$n_{50} = 50$
6. Историк	$n_{61} = 27$	$n_{62} = 110$	$n_{63} = 15$	$n_{64} = 85$	$n_{65} = 13$	$n_{60} = 250$
n_{0j}	$n_{01} = 200$	$n_{02} = 300$	$n_{03} = 200$	$n_{04} = 250$	$n_{05} = 50$	$n_{00} = 1000$

Для определения взаимосвязи двух переменных применяются уравнения *регрессии* и парной *корреляции*.

Корреляция – зависимость, при которой данные об одной переменной ведут к определенной информации о другой и эти переменные оказываются связанными друг с другом.

Коэффициент корреляции Пирсона применяется для переменных как с *интервальной* шкалой, так и с *номинальной*. Он вычисляется по формуле

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}},$$

где x и y – значения двух переменных;
и – их средние значения.

Величина коэффициента линейной корреляции Пирсона не может превышать +1 и быть меньше, чем –1. Эти два числа (+1 и –1) являются границами для коэффициента корреляции. Когда при расчете полу-

чается величина, большая +1 или меньшая -1, – произошла ошибка в вычислениях.

Пример расчета коэффициента корреляции приведен ниже.

x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})^2$	$(y - \bar{y})^2$	$(x - \bar{x})(y - \bar{y})$
9	51	-3	1	9	1	-3
13	40	1	-10	1	100	-10
16	64	4	14	16	196	56
11	54	-1	4	1	16	-4
20	60	8	10	64	100	80
8	51	-4	1	16	1	-4
12	31	0	-19	0	361	0
9	55	-3	5	9	25	-15
7	43	-5	-7	25	49	35
15	51	3	1	9	1	3

$$\Sigma = 850 \quad \Sigma = 150 \quad \Sigma = 138$$

$$r = \frac{138}{\sqrt{850 \times 150}} = 0,386$$

Итак, в результате вычислений получено значение коэффициента корреляции, равное 0,386. Как интерпретировать данное значение?

На основе свойств коэффициента можно сделать заключение, что его значение является положительным, т. е. изменения переменных имеют однонаправленный характер – по мере увеличения (уменьшения) величины одной переменной растет (уменьшается) и величина другой переменной.

Для определения силы связи необходимо интерпретировать величину коэффициента корреляции. В этом случае следует использовать табл. 2.3, в которой приведена шкала определения тесноты связи (силы корреляции) по таким критериям, как очень слабая, которой можно пренебречь, слабая, средняя и высокая.

Таблица 2.3. Шкала интерпретации значения коэффициента

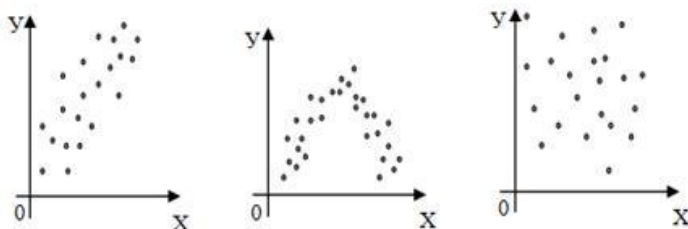
Значение	Интерпретация
До 0,2	Очень слабая корреляция
До 0,5	Слабая корреляция
До 0,7	Средняя корреляция
До 0,9	Высокая корреляция

Значение коэффициента Пирсона, равное 0,386, говорит о невысокой взаимосвязи между переменными.

Если, по меньшей мере, одна из двух переменных имеет **порядковую** шкалу либо не является нормально распределенной, то рассчитывается коэффициент ранговой корреляции r_s по *Спирмену* или τ (тау) *Кендала*.

В большинстве реальных задач исследователь не должен следовать ограничению круга используемых математических методов только известными коэффициентами парной связи.

Регрессия – статистический термин, представляющий изменение одной переменной, которая определяется как частично зависящая от другой, плюс фактор ошибки.



Уравнение линейной регрессии $Y = aX + b$ определяется двумя параметрами a и b , которые находятся по формулам:

$$a = \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{X})^2}; \quad b = \bar{Y} - a\bar{X}$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i \quad \bar{Y} = \frac{1}{n} \sum_{i=1}^n y_i$$

2.4. Методы количественной оценки образовательного процесса

Квалиметрический педагогический мониторинг. Образование представляет собой единую функциональную систему, что дает возможность стандартизировать системные параметры контроля его качества, предоставив субъектам образовательного процесса стандартизованные, соизмеримые и сопоставимые показатели. Речь идет о мони-

торинге и управлении качеством образования в учреждении высшего образования.

В большинстве случаев мониторинг качества проводится на двух уровнях, по двум критериям и двум основаниям – количественными и качественными методами.

Среди основополагающих оценочных технологий в настоящее время широко используются экспертные оценки, метрологические наблюдения, аудиторский контроль, опросы и тестирование.

Инструментарием таких исследований могут быть анкеты, тесты, опросные листы, протоколы наблюдений, акты проверки, статистические отчеты, различного рода формы, таблицы, графики и т. д.

Инструментами мониторинговой деятельности также могут выступать технические, электронные, вычислительные и информационные средства, программные продукты, каналы телекоммуникаций и др.

Ведущую роль в оценке и измерении образовательного процесса призван играть квалитетрический мониторинг. В широком смысле термин «квалитетрия» (от лат. *qualis* – какой по качеству и *metros* – измерять) означает количественное измерение качества.

Под **квалитетрическим мониторингом** в образовании понимается специально организованная, систематически осуществляемая деятельность по отслеживанию и измерению количественных параметров качества образовательного процесса. Измерение качества образовательного процесса предполагает оценку удовлетворенности потребителей качеством образовательных услуг и научной продукцией УВО.

Образовательный мониторинг в вузе включает организацию системы сбора, хранения, обработки и распределения информации об образовательной деятельности, обеспечивающую непрерывное слежение за ее состоянием и прогнозирование ее развития.

Состояние качества образовательных услуг контролируется в мониторинге с помощью разнообразных методов.

Организованный таким образом мониторинг позволяет отследить параметры, характеризующие, с одной стороны, качество образовательного процесса и уровень подготовки его участников, а, с другой, – «объективные», основанные на установлении соответствия нормативным требованиям, и «субъективные», представляющие собой оценки и суждения людей, показатели (приложение).

Ниже в табл. 2.4 приведены некоторые показатели, с помощью которых измеряют качество образования в вузе.

Т а б л и ц а 2.4. Показатели качества образования в вузе

№ п/п	Цель направления деятельности	Оценочный показатель			Методы количествен- ной оценки
		Наименование, ед. изм.	Базовое значение	Планируе- мое значе- ние	
3	Обеспечение высого каче- ства образо- вательных услуг и внутриву- зовского контроля	Подготовка специа- листов на I ступени высшего образова- ния: 1. Удовлетворен- ность студентов проведением учеб- ного процесса, НИРС, воспитатель- ного процесса, %: на очном отделе- нии 63 заочном отделе- нии 65		65 67	Социологиче- ский опрос
		2. Абсолютная успеваемость, %: на очном отделе- нии (бюджет) 96,3 заочном отделе- нии 82,4		96,5 83	Анализ документов
		3. Качественная успеваемость В т. ч.: удельный вес сту- дентов, получивших экзаменационную оценку автоматом (7 и более баллов) по обучению МРТ, % 39		43	Анализ документов
		4. Удельный вес выпускников, полу- чивших диплом с отличием, % 7,1		7,5	Анализ документов

Среди основополагающих исследовательских методов в настоящее время используются различные – экспертные оценки, социологические опросы, психологическое тестирование, анализ документов, педагогическое обследование и эксперимент. Важным направлением измерения качества образования в вузе является внутренний аудит, который проводится специально подготовленными и компетентными аудиторами.

Инструментарием мониторинговых измерений могут служить анкеты, тесты, опросные листы, статистические отчеты, аттестационные

ведомости, разного рода формы, таблицы, графики и др. Инструментами мониторинговой деятельности также могут выступать технические и электронные средства, программные артефакты, средства телекоммуникаций и т. д.

Вопросы для промежуточного контроля

1. Почему подход к измерению в педагогике существенно отличается от традиционного, сложившегося в метрологии?
2. В чем состоит главная особенность педагогических измерений?
3. Чем вызвано появление такой теории измерения, как квалиметрия? Почему квалиметрические измерения нередко называют *квазиметрическими* (от лат. *квази* – около, рядом)?
4. Перечислите основные компоненты и критерии измерений.
5. Чем отличается достоверность измерения от его надежности, а точность измерения от его правильности?
6. Перечислите и охарактеризуйте основные показатели качества образования в вузе (ссузе).

Вопросы и задания для самостоятельной работы

1. Сплошной и выборочный методы. Генеральная и выборочная совокупности.
2. Случайная и неслучайная выборки и их виды. Правило «левой руки» для определения объема выборки. Причины возникновения случайных и систематических ошибок выборки.
3. Расчет и качество выборочной совокупности. Проблема репрезентативности выборки. Типичные ошибки выборочного исследования.

Задание 1. Рассчитайте среднее арифметическое и среднее квадратическое (стандартное) отклонение для двух выборок:

1, 2, 3, 4, 5;

–235, –103, 3, 100, 250.

Статистически обоснуйте их сходство либо различие.

Задание 2. При обследовании 900 человек – лиц трудоспособного возраста – определен их средний возраст. Для вероятности p ($1 - \alpha$) = 0,90 необходимо найти доверительный интервал, в котором содержится генеральное среднее. Поскольку дисперсия признака неизвестна, оцените ее по значению размаха для генеральной совокупности.

Задание 3. Используя критерий Пирсона, при уровне значимости 0,05 проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности X по результатам выборки:

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

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Г л а с с, Дж. Статистические методы в педагогике и психологии / Дж. Гласс, Дж. Стэнли. – Москва: Прогресс, 1976. – 432 с.
2. Д е м ч у к, А. Математические методы в психологических и педагогических исследованиях: учеб. пособие / А. Демчук. – Красноярск: КГПУ, 2013. – 347 с.
3. Е л и с е е в а, И. И. Общая теория статистики: учебник для вузов / И. И. Елисеева, М. М. Юзбашев; под ред. И. И. Елисеевой. – Москва: Финансы и статистика, 2009. – 656 с.
4. К р и ч е в е ц, А. Н. Математическая статистика для психологов: учебник / А. Н. Кричевец, А. А. Корнеев, Е. И. Рассказова. – Москва: ИЦ «Академия», 2012. – 400 с.
5. К у з н е ц о в, В. А. Основы метрологии: учеб. пособие / В. А. Кузнецов, Г. В. Ялунина. – Москва: Изд-во стандартов, 2008. – 280 с.
6. Практикум по общей теории статистики: учеб. пособие для вузов / М. Р. Ефимова [и др.]. – Москва: Финансы и статистика, 2007. – 368 с.
7. Т а т а р и н о в а, Г. Г. Методология анализа данных в социологии / Г. Г. Татарина. – Москва: NOTA BENE, 1999. – 224 с.
8. Ч а ш к и н, Ю. Р. Математическая статистика. Анализ и обработка данных: учеб. пособие / Ю. Р. Чашкин. – Ростов-на-Дону: Феникс, 2010. – 236 с.
9. Ч е к м а р е в, А. Н. Квалиметрия и управление качеством: учеб. пособие: в 2 ч. / А. Н. Чекмарев. – Самара: Изд-во Самар. гос. аэрокосм. ун-та, 2010. – Ч. 1: Квалиметрия. – 172 с.



Лекция 3. МЕТОДЫ ОЦЕНКИ КАЧЕСТВЕННЫХ ПРИЗНАКОВ И ПРОБЛЕМЫ НАДЕЖНОСТИ ИЗМЕРЕНИЯ

Самое главное процентами
не измеришь...

Харуки Мураками

- 3.1. Проблема измерения качественных признаков в педагогике.
- 3.2. Процедура ранжирования: прямое ранжирование и метод парных сравнений.
- 3.3. Тестирование как квалиметрический подход к измерению.
- 3.4. Экспертные оценки.

3.1. Проблема измерения качественных признаков в педагогике

Для измерения качественных признаков применяются *неметрические* (качественные) шкалы – номинальная (наименований) и порядковая (ранговая, ординальная).

В метрологии качественные шкалы принято называть не шкалами измерений, а шкалами оценки, поскольку считается, что в них отсутствуют главные критерии измерительных шкал – интервальность, нулевая точка и др., а также потому, что с ними не допустимы операции с числами.

Однако практика свидетельствует об обратном – эффективности применения математических операций для таких шкал, например, при расчете рейтингового значения (суммы накопленных баллов или очков

за ранговые места) для определения показателей достижений (средняя балльная оценка успеваемости обучающихся), качества (продукции, услуг), интенсивности (например, труда как произведения балльных оценок его напряженности и сложности).

Возникновение квалитметрии было связано с переосмыслением проблемы измерения в связи с широким практическим использованием математики и количественных методов при оценке качества.

Номинальная шкала используется для того, чтобы идентифицировать объект, сравнить с эталоном или чтобы отличить один объект от другого. Разновидностью номинальной шкалы является *дихотомическая шкала* (рис. 3.1), которая имеет только два значения: да – нет, наличие признака – его отсутствие, соответствие – несоответствие, мальчик – девочка. К таким шкалам можно отнести *полихотомическую (поли-номинальную) шкалу*, которая имеет более двух значений, например, род существительных – мужской, средний и женский, которые можно обозначить как м. р., ср. р., ж. р., а можно им приписать значения 1, 2, 3. Тем не менее нельзя сказать, что один из этих родов более предпочтителен, чем остальные.



Рис. 3.1. Неметрические шкалы и их виды

На *порядковой шкале* объекты упорядочиваются, располагаются в зависимости от предпочтений и им приписываются те или иные значения (ранги, баллы и т. д.). Поэтому может возникнуть иллюзия, что мы имеем дело с количеством. С таким же успехом можно было бы заменить цифры на буквы, сочетания букв и цифр, другие символы, и тогда становится ясно, что речь идет о качественной шкале.

Шкала порядка может быть критериально ориентированной и нормативно ориентированной в зависимости от того, что принимается за норму.

Нормативно ориентированная шкала используется, если норма не известна, например, при оценке знаний обучающихся, и устанавливается эмпирически, *апостериори* (после опыта). Первые места (высшие баллы) присуждают тем, кто добился наибольших достижений. Если критерии оценки известны *априори* (до опыта) и сравнивают объекты на их соответствие им и дают более высокие баллы тем, которые в большей степени отвечают этим критериям, то применяется *критериально ориентированная шкала*.

Измерение качества может осуществляться через непосредственную оценку на соответствие критериям отбора, определенным требованиям и представлять собой процедуру идентификации объекта с эталоном, типичным объектом, совокупностью необходимых свойств, которыми он должен обладать, и т. д. Эквивалентным объектам приписывается одно и то же значение (число или любой другой символ). Этот метод может быть осуществлен при полной информированности о свойствах объектов по номинальной шкале.

Объекты могут располагаться в определенном порядке предпочтения, руководствуясь одним или несколькими критериями, показателями сравнения. Такая процедура осуществляется путем ранжирования по шкале порядка.

3.2. Процедура ранжирования: прямое ранжирование и метод парных сравнений

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 политической активности, отношения к чему-то и т. д. Районы г. Минск можно ранжировать по их престижности, а государства – по качеству жизни, уровню рождаемости и т. д. Информационные телепередачи можно ранжировать по степени их информативности, профессии – по

престижности, политических лидеров – по их влиянию на принятие решений, качества человека – по их важности в карьере, товары – 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 получают *ранжированный ряд*. В нем каждому объекту приписывается *ранг* или место в этом ряду. Число мест и, соответственно, число рангов равно числу объектов.

Следует обратить внимание на 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 «цифр» вариантам ответа (измерение как кодирование), либо по логическому квадрату, либо по одной из шкал (*Лайкерта* или *Гуттмана*), которые будут рассмотрены ниже. Но на основе этих оценок можно провести и ранжирование респондентов. В табл. 3.1 представлена именно такая ситуация.

Т а б л и ц а 3.1. Результаты ранжирования по качеству жизни

Государства	А	Б	В	Г	Д	Е	К	Л	М
Показатели качества жизни	6,5	7,0	6,5	5,9	4,6	5,9	4,5	5,9	4,5
Ранг	2,5	1	2,5	5	7	5	8,5	5	8,5

В первой строке данной таблицы приведены показатели качества жизни для произвольных девяти государств, обозначенных буквами А, Б, В, Г, Д, Е, К, Л, М; во второй строке – результаты ранжирования, т. е. ранжированный ряд.

Ранжирование является приемом измерения, в результате применения которого получается порядковый уровень измерения. Этот прием для отличия от других можно обозначить как *прямое, простое ранжирование*.

Прямое ранжирование. В табл. 3.2 приведены результаты прямого ранжирования восьми конкурсных работ группой экспертов, состоящей из пяти человек.

Т а б л и ц а 3.2. Результаты прямого ранжирования конкурсных работ

Названия конкурсных работ студентов	Члены жюри					Мода	Медиана
	А	Б	В	Г	Д		
«А»	8	5	6	4	5	6	6
«В»	1	6	2	8	6	—	6
«С»	7	1	3	1	1	1	1
«D»	6	4	4	2	2	2 и 4	4
«Е»	2	3	1	3	3	3	3
«F»	3	2	5	5	4	5	4
«G»	5	8	8	6	8	8	8
«H»	4	7	7	7	5	7	7

Следует обратить внимание на то, что где мода совпадает с медианой, там респонденты достаточно единодушны в своих оценках.

Существуют и другие приемы ранжирования, отличные от прямого.

Метод парных сравнений. Метод основан на попарном сравнении объектов ранжирования по заданному основанию.

Метод парных сравнений Терстоуна. Этот метод был разработан Л. Терстоуном и впервые использован для ранжирования преступлений по степени серьезности и для ранжирования различных национальностей по предпочтительности с точки зрения дружеских отношений. Метод парных сравнений основан на попарном сравнении объектов ранжирования по заданному основанию. Процедура сбора данных происходит следующим образом. На отдельные карточки заносятся названия объектов ранжирования. (Пусть речь идет о тех же конкурсных работах студентов.) Карточки перетасовываются, и респонденту предъявляется первая пара карточек с вопросом: «Будьте любезны, какая из этих конкурсных работ Вам предпочтительнее?» Затем предъявляются вторая, третья пара и т. д. Результаты парных сравнений отдельно взятого респондента заносятся в табл. 3.3.

Т а б л и ц а 3.3. Результаты парных сравнений

	п1	п2	п3	п4	п5	п6	п7	п8	Число предпочтений
п1	—	1	0	0	0	0	1	1	3
п2	0	—	0	0	0	0	0	1	1
п3	1	1	—	1	1	1	1	1	7
п4	1	1	0	—	1	1	1	1	6
п5	1	1	0	0	—	1	1	1	5
п6	1	1	0	0	0	—	1	1	4
п7	0	0	0	0	0	0	—	0	0
п8	0	1	0	0	0	0	1	—	2

В таблице приведены результаты парных сравнений для последнего (пятого) респондента. Число различных между собой пар будет равно 28. Легко понять, что в общем случае, когда число объектов ранжирования равно ft , число сравнений, или число различных пар, будет равно $N(N - 1)/2$. Необходимо запомнить эту формулу, так как она в дальнейшем пригодится. Получается данная формула просто. Число клеток в таблице равно $N \cdot N$. Это число всех возможных сравнений. Диагональ нас не интересует, так как сравнение объекта с самим собой не имеет смысла. Поэтому число сравнений сокращается и становится равным $N/N \cdot N = N(N - 1)$. Но правая верхняя часть таблицы есть зеркальное отражение левой нижней. Если $p1$ сравнили с $p2$, то нет никакой необходимости в сравнении $p2$ с $p1$. Поэтому нас интересует количество сравнений или необходимость заполнения клеток, число которых равно $N(N - 1)/2$. Тогда, когда $N = 8$, число сравнений равно $8 \cdot 7/2 = 28$.

В каждую клетку таблицы заносится результат сравнения пар, обозначенных в строке и в столбце. В клетку ставится единица, если конкурсная работа по строке предпочтительнее, чем работа, соответствующая столбцу. В данном случае первая конкурсная работа предпочтительнее эксперту, чем вторая, седьмая и восьмая. Следует особо отметить случай, когда отдельно взятый эксперт не может предпочесть одну работу другой. Тогда в клеточку сносится 0,5.

Вторая строка – результаты сравнения второй конкурсной работы с другими. Эта работа предпочтительнее респонденту только по сравнению с работой $p8$. В третьей строке одни единицы, т. е. третья работа предпочтительнее по сравнению со всеми остальными. Таким образом заполняется вся таблица. После этого по каждой строке подсчитывают число единиц, т. е., сколько раз эта конкурсная работа была предпочтена респондентом при сравнении с другими. По этому числу предпочтений конкурсные работы расположились для пятого респондента следующим образом: $p3 > p4 > p5 > p6 > p1 > p8 > p2 > p7$.

Шкала Лайкерта (Ликерта). Заслугой Лайкерта явилось то, что он: а) предложил некий критерий, который показывает, насколько правдоподобно предположение о самом существовании измеряемой однородной латентной переменной; б) дал основания отобрать именно те наблюдаемые признаки (суждения), которые имеют отношение к тому, что мы измеряем; в) показал, что пятибалльная шкала приемлема для измерения этих признаков; г) дал обоснование применению *суммы значений наблюдаемых признаков* в качестве значения латентной переменной и определил порядковый характер получающейся в результате шкалы.

Процедура построения шкалы. Формируется большое число суждений, около 100. Данные суждения не имеют нейтральных значений и они не обязательно должны охватывать весь смысловой континуум, но они должны иметь ярко выраженное положительное либо отрицательное отношение к объекту; каждое суждение предъявляется экспертами, которые выражают свое к нему отношение по 5-балльной шкале (полностью согласен, согласен, нейтрален, не согласен, полностью не согласен); далее для каждого эксперта социолог рассчитывает общий балл (за каждое суждение эксперт получает балл от 1 до 5, которые затем складываются по всем суждениям). Затем каждое суждение сравнивается с остальными суждениями. Для этого строится таблица сопряженности, причем на ее основе рассчитывается коэффициент корреляции, который показывает тесноту связи данного вопроса с итоговым показателем. В основе таблицы сопряженности лежат два показателя: балл эксперта по проверяемому вопросу (от 1 до 5), разность между общим баллом эксперта и баллом проверяемого вопроса. Вопросы с отрицательными и малыми корреляциями «выбрасываются». В общей сложности должно остаться около 20 суждений; получается порядковая шкала. В процессе массового опроса респондент отвечает на эти 20 вопросов и получает суммированный балл.

Шкалограммный анализ Гуттмана. Этот метод измерения используется исходя из предположения о существовании одномерной шкалы, континуума значений социальной установки. При разработке шкалы Гуттмана исходят из существования иерархической совокупности суждений. Согласие (несогласие) с ними говорит об установке респондента.

Шкала Гуттмана строится в три этапа. Рассмотрим их для измерения отношения к феномену «новые русские».

Набираются суждения, определенным образом упорядоченные. Практически их можно придумать для данного случая не так много. Это в равной степени относится и к другим социальным установкам. Обозначим суждения через А, Б, В, Г, Д, Е. Например, таковыми являются:

А. «Новые русские» умные, деловые люди:
согласен (1), не согласен (0).

Б. «Новые русские» самые непорядочные люди:
согласен (0), не согласен (1).

В. У «Новых русских» нет моральных принципов:
согласен (0), не согласен (1).

Г. Без «Новых русских» нет будущего России:

согласен (1), не согласен (0).

Д. России было нужно появление «новых русских»:

согласен (1), не согласен (0).

Е. «Новые русские» состоят из одних бандитов и жуликов:

согласен (0), не согласен (1).

Перечень можно продолжить. Как и в случае со шкалой Лайкерта, в этой совокупности есть два типа суждений. О позитивном отношении к «новым русским» говорит согласие с суждениями первого типа А, Г, Д и несогласие с суждениями второго типа Б, В, Е. В том и другом случае респондент получает как бы один балл. Максимальное число баллов, которое может набрать респондент, равно числу суждений (6), минимальное – 0. При этом необязательно, чтобы число суждений первого типа совпадало с числом суждений второго типа. Рассуждения можно провести и исходя из негатива, т. е. измерять степень негативного отношения к «новым русским».

В этом примере нет ярко выраженной упорядоченности суждений.

Последующие два этапа построения шкалы позволяют проверить существование упорядоченности в совокупности суждений.

Семантический дифференциал Ч. Осгуда. Был задуман как методика, позволяющая определять признак как с качественной стороны, так и с количественной. Суть методики состоит в том, что испытуемый оценивает интересующие понятия по набору шкал. Каждая шкала – это отрезок прямой с семью делениями (табл. 3.4). Полярные точки обычно представлены прилагательными-антонимами: красивый – уродливый; глубокий – мелкий и т. д. Серединная, или нейтральная, позиция означает либо признание, либо отрицание обеих характеристик («и...и» или «ни...ни»).

Таблица 3.4. Шкала семантического дифференциала Ч. Осгуда

слабый	– 3	– 2	– 1	0	+ 1	+ 2	+ 3	сильный
женский	– 3	– 2	– 1	0	+ 1	+ 2	+ 3	мужской
пассивный	– 3	– 2	– 1	0	+ 1	+ 2	+ 3	активный
медленный	– 3	– 2	– 1	0	+ 1	+ 2	+ 3	быстрый
необычный	– 3	– 2	– 1	0	+ 1	+ 2	+ 3	обычный
ложный	– 3	– 2	– 1	0	+ 1	+ 2	+ 3	правдивый
плохой	– 3	– 2	– 1	0	+ 1	+ 2	+ 3	хороший

3.3. Тестирование как квалиметрический подход к измерению

Использование тестирования в образовании интересно и актуально. Несмотря на ограничения применения тестовой процедуры, есть основное достоинство, которое привлекает многих преподавателей – это оперативность получения обратной связи с учащимися и возможность быстрого оценивания результатов.

Основная задача преподавателя заключается в том, чтобы использовать различные методы контроля знаний путем измерения, в том числе и тестовые. Что же касается ограничений, то любые средства для проверки уровня достижений имеют свои границы применимости.

Под **тестом** в педагогическом измерении понимается стандартизированная квалиметрическая процедура, служащая определению степени выраженности (обычно количественной) у индивида тех или иных свойств. Обычно состоит из ряда относительно коротких испытаний, в качестве которых могут выступать различные задачи, вопросы, ситуации.

Виды тестов. Существуют тестовые задания открытого и закрытого типов, которые объединяют шесть видов (рис. 3.3).



Рис. 3.3. Виды тестовых заданий

Задания закрытого типа.

К ним относятся задания альтернативных ответов, множественного выбора, восстановления соответствия и последовательности. Такие задания предусматривают различные варианты ответа на поставленный вопрос: из ряда предлагаемых выбираются один или несколько пра-

вильных ответов, выбираются правильные (или неправильные) элементы списка и др. Они предполагают наличие предварительно разработанных вариантов ответа на заданный вопрос.

1. Задания альтернативных ответов.

К каждой задаче альтернативных ответов дается только два варианта ответов. Испытуемый должен выбрать один из них – «да – нет», «правильно – неправильно» и др.

Инструкция для задания альтернативных ответов:

Вам необходимо выбрать один вариант ответа, который Вы считаете правильным.

Задания альтернативных ответов в большей степени подходят для выявления уровня овладения сложными определениями, знания достаточно сложных графиков, диаграмм, схем и др.

Особенностью заданий альтернативных ответов является то, что вопрос должен быть сформулирован в форме утверждения, поскольку он предполагает согласие или несогласие, которое можно отнести к утверждению.

Задания с альтернативными ответами (как и при использовании метода парных сравнений) в наибольшей степени соответствуют задаче выявления того, в какой степени испытуемый понимает материал в целом. С помощью этих заданий можно проверить умения работать с графиками, навыки приближенного вычисления. Любая другая форма представления заданий будет гораздо более громоздкой и менее удобной.

2. Задания множественного выбора.

Это основной вид заданий, применяемый в тестах достижений. Задачи с множественным выбором предполагают наличие вариативности в выборе. Испытуемый должен выбрать один из предложенных вариантов, среди которых чаще всего только один правильный.

3. Задания на восстановление соответствия.

К заданиям данного типа относятся задания на восстановление соответствия между элементами двух списков, порядка ряда.

Главными преимуществами заданий этого вида являются возможность быстрой оценки достижений в конкретной области знаний и экономичность размещения задач в тесте.

4. Задания на восстановление последовательности.

Задания на восстановление последовательности можно рассматривать как вариант задания на восстановление соответствия, когда одним из рядов является время, расстояние или иной континуальный конструкт, который подразумевается в виде ряда.

Задания на восстановление последовательности – это довольно

качественная форма тестовых заданий, обладающая значительными преимуществами: краткостью, простотой проверки.

Преимущества заданий закрытого типа:

задания могут быть надежны, поскольку отсутствуют факторы, связанные с субъективными оценками, которые снижают надежность;

оценивание заданий полностью объективно: между оценками различных проверяющих не может быть различий;

не учитывается умение испытуемых формулировать ответы;

задания этого типа легко обрабатываются, тестирование быстро проводится;

простой алгоритм заполнения снижает количество случайных ошибок и описок;

эти задания позволяют охватить большую область знаний, что для тестов достижений особенно важно;

возможна машинная обработка ответов;

низкая вероятность угадывания правильных ответов;

возможно получение точной оценки содержательности теста, что особенно важно для определения соответствия теста целям исследования.

Задания открытого типа.

К ним относятся задания двух видов:

1) *дополнения* (задачи с ограничением на ответы). В этих заданиях испытуемые должны самостоятельно давать ответы на вопросы, однако их возможности ограничены.

Ограничения обеспечивают объективность оценивания результата выполнения задания, а формулировка ответа должна дать возможность однозначного оценивания;

2) *свободного изложения, или свободного конструирования*. Они предполагают свободные ответы испытуемых по сути задания. На ответы не налагаются ограничения. Однако формулировки заданий должны обеспечивать наличие только одного правильного ответа.

Трудность в применении этого вида задач заключается в следующем: сложность формализации ответов, затруднение стандартизации в связи с необходимостью подготовки оценочных схем, громоздкость процедуры и большие затраты времени на проведение.

Основная трудность при составлении заданий открытого типа – соблюдение основного требования к тестовым заданиям (наличие однозначного правильного ответа).

Положительными сторонами хорошо составленных заданий дополнения и свободного изложения являются:

1) невозможность угадывания ответа;

- 2) краткость и однозначность ответов;
- 3) необходимость воспроизведения ответа по памяти;
- 4) отсутствие необходимости нахождения нескольких вариантов;
- 5) простота формулировки вопроса;
- 6) простота проверки.

Особую группу качественных методов измерения составляют проективные тесты и методики, служащие определению индивидуальных особенностей обследуемых с помощью фиксации их реакций на неопределенные и многозначные ситуации.

Впервые проективная диагностическая стратегия была использована Э. Роршелином в его методике свободных ассоциаций.

В качестве неопределенной ситуации могут выступать сюжетные картинки, которые надо проинтерпретировать (ТАТ), тест незаконченных предложений (тест С. Розенцвейга), метод «чернильных пятен» (тест Г. Роршаха).

Ответы на задания проективных тестов допускают очень широкий диапазон значений, поэтому очень большую роль при интерпретации играют качественные заключения.

Надежность измерений качественных признаков. Различают два метода оценки – субъективный и объективный.

Субъективный метод означает выявление, измерение и оценку объекта, исходящую из личных представлений человека. Этот метод оценки знаний пригоден для итогового контроля, так как не обладает необходимой точностью и воспроизводимостью результатов.

Под *объективным методом* подразумевают оценку, которая обладает необходимой точностью, воспроизводимостью результатов.

Инструментом, позволяющим объективно оценивать качество усвоения знаний, является *критериально ориентированный тест*, сочетающий в себе контрольное задание и эталон, по которому можно судить о качестве усвоения.

Критериально ориентированный тест – особый метод оценки. Его особенность состоит в том, что на первом этапе осуществляется выяснение целей, в отличие от традиционных тестов.

Результаты критериально ориентированных тестов представляются не столько в условных баллах, сколько в конкретных показателях усвоенности и сформированности определенных знаний, навыков, умений.

Главным при его выполнении является не количество выполненных заданий, а то, что человек знает, усвоил и какие операции умеет делать.

3.4. Экспертные оценки

Суть экспертной оценки состоит в проведении компетентными лицами в данной предметной области (экспертами) анализа проблемы с качественной и количественной оценкой суждений и формальной обработкой результатов индивидуальных экспертных оценок.

Значимость экспертных оценок в образовании, необходимость научно-педагогической экспертизы существенно возросли в последние годы. Это связано с теми изменениями, которые произошли в системе образования и развитии педагогической системы. Данный метод обеспечивает оценку альтернативных решений и выбор предпочтительных вариантов решения.

Метод экспертных оценок подразделяется на ряд логически взаимосвязанных этапов.

Организационный этап включает: определение цели и задач экспертизы, постановку проблемы; выбор технологии экспертизы, методов и критериев оценки; определение меры ответственности, прав и полномочий рабочей группы экспертов; установление сроков проведения экспертизы; подбор экспертов, формирование экспертных групп (при необходимости определение их компетентности). Для педагогической экспертизы чаще всего требуется разработка анкет, бланков интервью, опросников, тестов, контрольно-диагностических заданий. Экспертная оценка является результатом аналитической деятельности экспертов, основана на их умении видеть и разрешать противоречия, прогнозировать, предвидеть и находить способы решения проблемы. Эксперт также несет ответственность (правовую, моральную) за принятое им решение.

Основной этап экспертизы связан со сбором данных, проведением исследовательской работы и экспертной оценкой, анализом имеющегося материала. Он включает опрос экспертов (индивидуальный или групповой; личный, очный или заочный; устный или письменный), оформление документа (отчет, справка, рецензия и пр.), а также сбор и обработку результатов экспертного оценивания. В повторных исследованиях процедура экспертного опроса может осуществляться в несколько туров (например, при использовании Дельфийского метода).

Заключительным этапом экспертной оценки является принятие экспертного заключения – основы для компетентного управленческого решения, которое принимает руководитель.

Понятия «экспертный опрос», «экспертная оценка» и «экспертное заключение» близкие, но не тождественные.

Экспертный опрос – это метод или процедура получения первичной информации от экспертов.

Экспертная оценка – это процесс и результат оценивания экспертами того или иного процесса и явления.

Экспертное заключение – это итоговый документ, оформленный в соответствии с установленными требованиями и содержащий мотивированную экспертную оценку, т. е. мнение эксперта о предмете.

В зависимости от области применения экспертных методов в педагогике возможны индивидуальные, групповые и коллективные экспертные оценки.

Для принятия важных экспертных заключений о перспективах и концепциях развития системы образования, государственных стандартах образования, новых образовательных программах используется *метод коллективных экспертных оценок*. Четко регламентирован количественный состав членов экспертной группы при проведении аттестации образовательных учреждений.

Методы коллективной работы. Они разнообразны, это – совещания, дискуссии, «мозговая атака», деловые игры, а также метод Дельфи, который является формой индивидуальной работы, но согласованной с деятельностью группы. Смысл коллективных методов работы экспертов состоит в активизации совместного поиска решения проблемы путем сбора предложений и идей, возникающих в процессе дискуссии.

Метод «мозговой атаки» может быть рекомендован для нахождения нетрадиционного, оригинального решения в предельно сжатые сроки. Возможны разные способы организации «мозгового штурма». Например, после определения проблемы и задач для поиска решения, знакомства с правилами, формирования рабочих групп (3–5 человек) дают сигнал к «мозговому штурму» и начинается генерирование идей, которые фиксируются без обсуждения и критических замечаний (возможны только уточняющие вопросы). После завершения «мозговой атаки» все высказанные идеи анализируют на основе ранее определенных критериев и подвергают экспертной оценке. Для получения обобщенного суждения экспертов относительно прост и удобен *метод ранговых оценок*. Он заключается в том, что любую полученную информацию располагают в порядке возрастания или убывания величины признака. Каждому признаку приписывают число, обозначающее его ранг. Результаты ранжирования легко поддаются математической обработке. Проверку согласованности ранжировок определяют вычислением коэффициента ранговой корреляции Спирмена (r_s) или коэффициента конкордации (W).

Индивидуальные экспертные оценки могут быть получены путем анкетирования, интервью, свободной беседы. Для проведения педагогической экспертизы целесообразно использовать уже разработанные, апробированные, доказавшие свою валидность и объективность методики: они, как правило, являются результатом психолого-педагогических исследований. Если же таких методик нет, их приходится разрабатывать самим экспертам. Такая работа предполагает ряд этапов.

1-й этап: деление экспертами объектов оценивания на заданное число групп;

2-й этап: фиксация результатов экспертного оценивания;

3-й этап: анализ единодушия экспертов в своих суждениях;

4-й этап: отбор качественных суждений для шкалы;

5-й этап: подведение итогов экспертной оценки.

Метод Дельфи позволяет учесть независимое мнение всех участников группы экспертов по обсуждаемому вопросу путем последовательного объединения идей, выводов и предложений и прийти к согласию. Метод основан на итерациях – многократных анонимных групповых интервью и подведении их итогов (рис. 3.4).

Известно, что использование коллективных знаний ведет к возможности нахождения сильных решений, однако в процессе обмена мнениями между участниками может сказаться влияние авторитета коллег и все сведется к появлению популярных ответов. Метод Дельфи позволяет разрешить это противоречие. Для этого прямые дискуссии экспертов заменяются индивидуальными опросами. Собранные варианты ответов подвергаются статистической обработке. Полученные обобщенные ответы передаются каждому эксперту путем личного общения либо по обычной или электронной почте с просьбой пересмотреть и уточнить свое мнение, если он сочтет необходимым. Эта процедура может повторяться несколько раз.

К достоинствам метода можно отнести то, что он способствует выработке независимости мышления членов группы, обеспечивает спокойное и объективное изучение проблем, которые требуют оценки.

Недостатки метода связаны с чрезмерной субъективностью оценок; метод требует достаточно много времени и организационных усилий для проведения множества итераций, обработки полученных результатов и выработки согласованного экспертного заключения.

Ожидаемый результат: согласованный список идей с их сопутствующими сильными и слабыми сторонами.

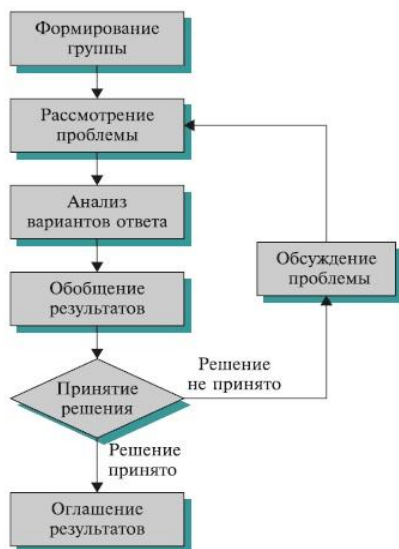


Рис. 3.4. Алгоритм экспертной оценки методом Дельфи

Метод фокус-группы. Он относится к качественным, или «гибким», методам измерения, поскольку ориентирован не на количественные результаты, а на то, чтобы определить глубинные мотивы и причины тех или иных действий, определить отношение к поставленной проблеме.

Преимуществом данного метода является и то, что помимо анализа фактического материала фиксирует мотивы в различного рода высказываниях участников, в нем используется анализ невербального поведения (выразительные движения, жесты, пантомимика). Основным способом получения результатов групповых дискуссий является контент-анализ, которому могут быть подвергнуты аудио- и видеоматериалы, стенограммы групповых дискуссий, отчеты модераторов и наблюдателей. Качество получаемых результатов зависит от множества факторов: разработанного плана исследования, сценария; профессиональной подготовленности исследователей, модератора (ведущего фокус-групп); выбранных критериев отбора участников групповых дискуссий, количества участников; организации места проведения дискуссии, подготовленности аналитика и др.

Оптимальное число участников фокус-группы составляет 6–10 человек при минимуме 2–3 человека. Критерии подбора участников групповой дискуссии зависят от целей конкретного исследования и могут учитывать социально-демографические признаки (пол, возраст, образование, уровень дохода), а также такие специфические особенности, как, например, частота, объем потребления определенной категории или определенного вида услуг и др.

Особые требования предъявляются к модератору. Он должен быть в высшей степени *коммуникабельным* человеком, тонко чувствующим психологию группы, с тем, чтобы в управлении дискуссией умело балансировать между естественной беседой на заданную тему и уходом от нее. Модератор должен чувствовать, когда ему следует вмешаться в ход дискуссии, а когда не нужно этого делать. Он должен понимать, что некоторые отклонения от темы вполне естественны и не только допустимы, но и полезны, поскольку позволяют участникам дискуссии проводить аналогии, сравнения и т. п., что способствует генерированию неординарных идей, предложений. В процессе беседы он должен проявлять интерес к каждому высказыванию, мотивируя коллективный мыслительный процесс, делая его более результативным. К работе с фокус-группой модератор должен тщательно готовиться: как можно больше узнать об обсуждаемой проблеме, составить план проведения дискуссии, т. е. разделить проблему на составляющие элементы, определить очередность их обсуждения и т. д.

Для обеспечения результативного обсуждения изучаемых проблем состав фокус-групп целесообразно формировать, исходя из принципа гомогенности ее участников (по возрасту, полу, виду профессиональной деятельности, семейному положению и т. п.).

Вопросы для промежуточного контроля

1. Как называется шкала измерения, имеющая только два возможных значения? Приведите примеры таких шкал.
2. Назовите один из наиболее известных способов измерения социальной установки, который базируется на согласии или несогласии респондента с заранее прошкалированными суждениями.
3. Какой способ выявления присущих отдельным респондентам субъективных смыслов объектов был предложен Осгудом?
4. Что такое ранжирование? Приведите пример применения данной процедуры в педагогическом измерении.

5. Как называется шкала, отображающая только отношение «равенства – неравенства» между явлениями, которые включены в один класс?

6. На конкретном примере продемонстрируйте познавательные возможности качественного подхода к измерению.

Задания для самостоятельной работы

Задание 1. Создайте шкалу измерения удовлетворенности студентов прочитанной лекцией методом Лайкерта. Осуществите факторную операционализацию данного понятия. Постройте систему индексов для измерения этой установки.

Задание 2. Опишите особенности и различия исследовательских стратегий при количественном и качественном подходах к измерению, заполнив соответствующие графы таблицы:

Исследовательские стратегии	Подходы к измерению	
	количественный	качественный
Особенности и различия подходов		

Задание 3. Приведите аргументированные опровержения или подтверждения утверждения: «с качественными шкалами математические действия недопустимы».

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Докторов, Б. З. О надежности измерения в социологическом исследовании / Б. З. Докторов. – Ленинград: Наука, 1979. – 173 с.
2. Лакутин, О. В. Принципы оценки качества и сравнения коэффициентов связи номинальных признаков / О. В. Лакутин, Ю. Н. Толстова. – Москва: ИСАН, 2003. – 288 с.
3. Майоров, А.Н. Теория и практика создания тестов для системы образования / А. Н. Майоров. – Москва: Академия, 2000. – 188 с.
4. Патругин, Ю. А. Об измерении качественных признаков / Ю. А. Патругин // Методы моделирования социальных процессов. – Москва: ЮНИТИ-ДАНА, 2005. – 240 с.



Лекция 4. СОВРЕМЕННЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ОБРАБОТКИ ДАННЫХ, ПРЕДСТАВЛЕНИЯ И ТРАНСЛЯЦИИ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ

Машины должны работать.

Люди должны думать.

Девиз компании IBM

4.1. Современные информационные технологии обработки данных.

4.2. Основные формы представления данных.

4.3. Значение эффективной презентации результатов измерения в повышении качества информации.

4.1. Современные информационные технологии обработки данных

Процесс автоматизации обработки статистических данных в настоящее время активно внедряется в систему образования, предоставляя качественно новые возможности. Самыми распространенными на сегодняшний день средствами компьютерной обработки данных являются электронные таблицы.

Информационная технология обработки данных в среде табличного процессора Excel. Это прикладное программное обеспечение, входящее в пакет Microsoft Office.

Microsoft Excel удобен для проведения математических и статистических расчетов. Помимо работы с числами, в Excel хорошо развиты возможности работы с текстом, позволяющие создавать документы практически любой сложности. Excel располагает большим количест-

вом шаблонов, с помощью которых можно быстро оформить самые разные, в том числе финансовые документы.

Окно приложения табличного процессора Excel имеет вид, представленный на рис. 4.1.

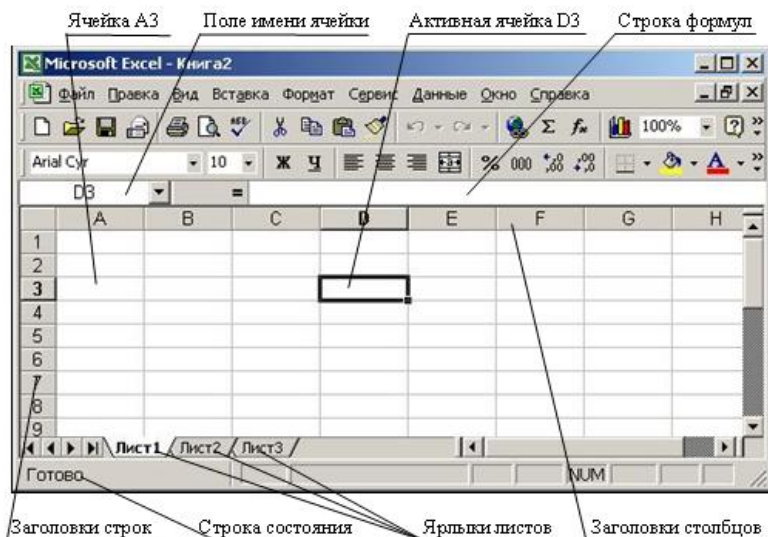


Рис. 4.1. Окно Microsoft Excel

Microsoft Excel имеет ряд встроенных функций для математических, статистических, финансовых и других вычислений. Одной из важнейших функций программы является возможность представления данных в виде разнообразных графиков, диаграмм и карт.

Средства, которые включены в пакет анализа данных, описаны ниже. Они доступны через команду *Анализ данных* меню *Сервис*. Если этой команды нет в меню, необходимо загрузить [настройку](#) пакета анализа, включающего:

- дисперсионный анализ;
- корреляционный анализ;
- ковариационный анализ;
- описательную статистику;
- экспоненциальное сглаживание;
- двухвыборочный f-тест для дисперсии;
- анализ Фурье;
- гистограмму;

скользящее среднее;
генерацию случайных чисел;
ранг и перцентиль;
регрессию;
выборку;
t-тест;
Z-тест.

В пакет Microsoft Office входит Microsoft Access – программа, предназначенная для формирования баз данных. **База данных** представляет собой список связанных данных, в котором строки данных являются записями, а столбцы – полями. Верхняя строка списка содержит названия всех столбцов. Microsoft Access обеспечивает многоаспектный доступ к данным и использование одних и тех же данных в различных задачах и различными приложениями пользователей.

4.2. Основные формы представления данных

Эмпирические данные в табличном процессоре могут храниться как в числовом виде, так и в графическом.

Построение диаграмм, графиков и поверхностей в Microsoft Excel. Microsoft Excel удобен для хранения, сортировки и представления данных.

Создадим таблицу Microsoft Excel по заданию (рис. 4.2).

	A	B	C	D	E	F
1	Квартал	I	II	III	IV	
2	Выработка, %	11	13	15	9	
3						
4						
5						
6						

Рис. 4.2. Пример построения таблицы в Microsoft Excel

В таблице Microsoft Excel на рис. 4.2 приведены данные о выработке предприятием продукции в миллионах тонн по кварталам за год.

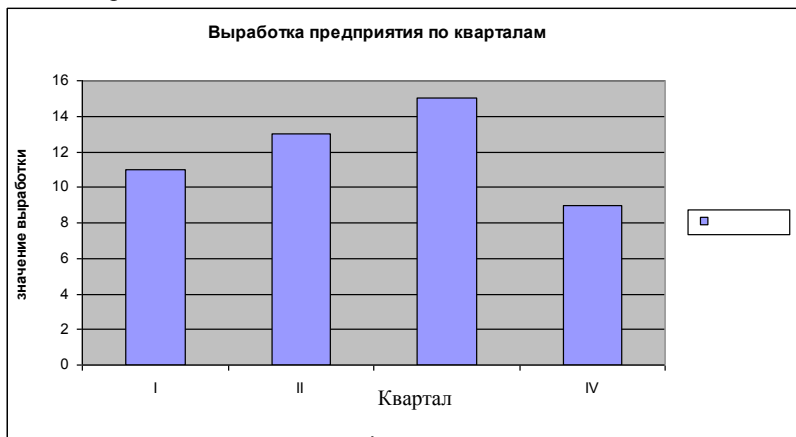
Представим эти данные в виде:

- а) гистограммы;
- б) объемной гистограммы;
- в) кольцевой диаграммы;
- г) круговой диаграммы;
- д) объемной круговой диаграммы.

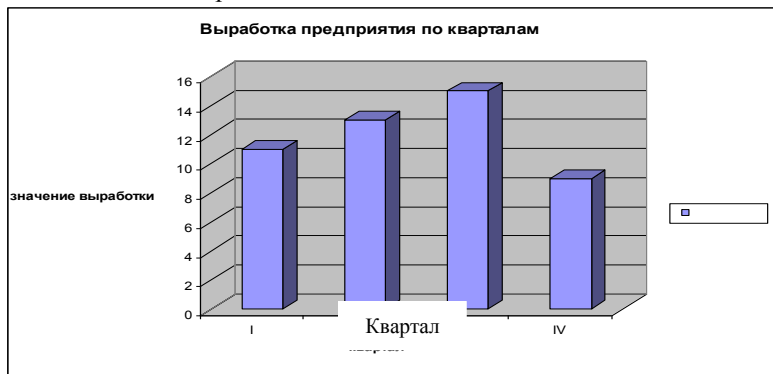
Построим графики.

Для этого выделим таблицу и с помощью мастера диаграмм получим следующие виды графиков.

Гистограмма.



Объемная гистограмма.



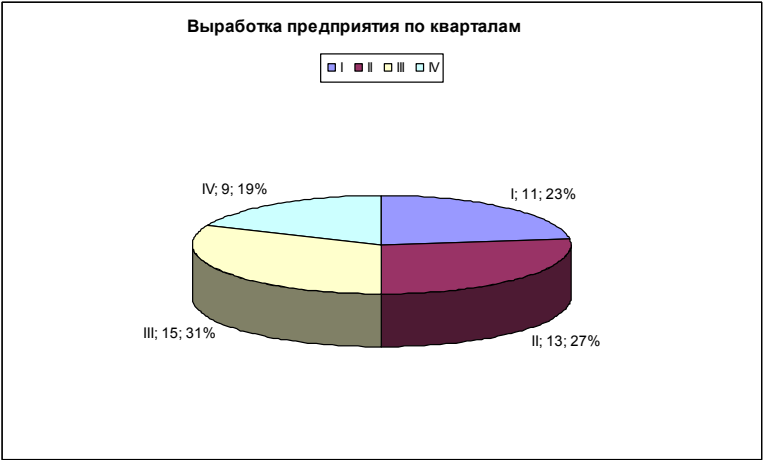
Кольцевая диаграмма.



Круговая диаграмма.



Объемная круговая диаграмма.

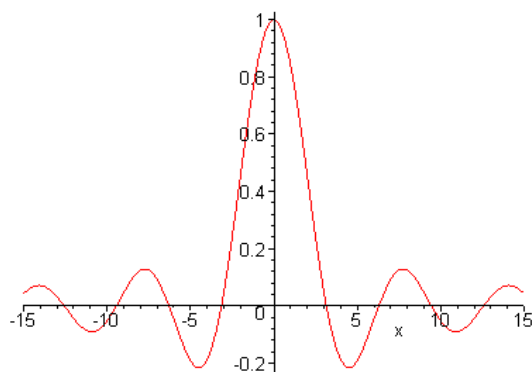


Результат измерения имеет ценность лишь тогда, когда его можно оценить будучи представленным.

Результаты измерения могут быть представлены следующим образом в виде таблицы:

Социометрическая матрица											
Кто выбирает		Кого выбирают							Итого		Σ выборов
Номер позиции	Ф.И.О. (шифр)	А	Б	В	Г	Д	Е	К	+	–	
1	А	–	+	+	0	0	0	0	2	0	2
2	Б	+	–	0	0	0	0	0	1	0	1
3	В	+	+	–	–	–	0	0	2	2	4
4	Г	+	0	+	–	0	–	0	2	1	3
5	Д	+	+	0	0	–	–	0	2	1	3
6	Е	+	0	0	0	–	–	0	1	1	2
7	К	0	0	0	0	0	0	–	0	0	0
Итого	+	5	3	2	0	0	0	0	10		
Итого	–	0	0	0	1	2	2	0		5	
Σ выборов		5	3	2	1	2	2	0			15

В виде графика:



В виде аналитической формулы:

$$V = \frac{\sigma}{x} \times 100\%$$

В пакет Microsoft Office также входит Microsoft Power Point – программа презентаций. Позволяет быстро создавать динамические презентации, диаграммы с иерархиями и связями на основе пользовательских макетов и встроенной библиотеки слайдов. Последняя версия отличается практичным и приятным интерфейсом, что ускорило процесс создания презентационных материалов.

Помимо создания целевых, радиальных, пирамидальных диаграмм, использования кругов Эйлера, возможности смешивания и сравнения изменений в презентациях и защиты готовых материалов паролем, появились полезные новшества: добавление оттенков, теней, освещения, сглаживания и пр., а для удобства пользователей – панели облегченного форматирования («быстрые стили»), предварительный просмотр результатов.

Особенности русской версии Microsoft Power Point (2010):

мощные диаграммы и новые эффекты рисунков визуального компонента SmartArt;

продвинутая библиотека слайдов;

согласованное форматирование материалов с помощью тем;

поиск в материалах скрытых метаданных, содержимого и личных сведений;

нахождение и удаление лишней информации;

возможность преобразования файлов в форматы *. PDF, *. XML;

использование презентаций на любой платформе программы;

поддержка режима отправки презентаций на рецензирование;

наличие процесса утверждения;

возможность уменьшения размеров документов и восстановления файлов;

защита созданных материалов паролем;

удобный интерфейс на русском языке.

Применение компьютерных технологий оправдано тогда, когда они эффективны, т. е. позволяют при одних и тех же затратах субъектов образовательного процесса получать более высокий образовательный результат или получать тот же результат при меньших затратах участников образовательного процесса.

4.3. Значение эффективной презентации результатов измерения в повышении качества информации

Под **презентацией** обычно понимают представление заинтересованной аудитории некоторой еще неизвестной или малоизвестной информации, продукции, услуг и их создателей. Компьютерная презентация на сегодняшний день является неременным атрибутом деятельности любой организации, в том числе и учреждения образования.

Презентация помогает специалисту, автору, преподавателю, студенту, работодателю получить благоприятные условия во время доклада и привлечь внимание аудитории для достижения цели презентации. Если для учащихся и студентов это этап обучения и получения важных навыков делать доклады и сообщения, то для преподавателей – важнейший элемент информационной культуры и квалификации, необходимый в профессиональной деятельности.

Основные признаки эффективной презентации:

1. *Увлекательность*. Презентация должна с самого начала увлечь аудиторию и удерживать ее внимание до ее логического завершения.

2. *Содержательность*. Презентация должна содержать значительные сведения для лиц, которым она предназначена.

3. *Запоминаемость*. Презентация эффективна, когда она действует и после ее завершения, а аудитория продолжает размышлять о предмете ее содержания.

4. *Активизация.* Презентация эффективна, когда побуждает людей к действию в направлении, желательном для презентатора.

5. *Сбалансированность.* Презентация эффективна, когда в ней соблюдено равновесие всех ее элементов: эстетическая привлекательность, грамотная композиция, стиль, длительность, структурированность, понятность.

Цель презентации – убедить аудиторию своим сообщением, чтобы она действовала в нужном докладчику направлении. Люди должны ответить на сообщение, которое принесла им презентация. Во мнениях, поведении и убеждениях людей должны произойти изменения.

Нужно искать «точку входа» в душу аудитории, т. е. сделать акцент на точке зрения или действии, которые люди охотно воспримут. Убеждение должно происходить на благоприятном и благожелательном уровне и вызвать согласие с их стороны. Презентация не должна делаться слишком напористо и назидательно, что может вызвать отторжение аудитории. Передний фронт технических средств и методик проведения эффективных презентаций в настоящее время сосредоточен на оптимизации контакта презентатора с аудиторией.

Речь идет о том, что многие технические и методические вопросы проведения презентаций в настоящее время достаточно хорошо проработаны. Поэтому, приняв их во внимание, нужно делать акцент на технике контакта.

Обратим внимание на точки входа в контакт с аудиторией:

1. Понять, что нужно аудитории показать и что она может получить.
2. Дать аудитории новую, полезную информацию.
3. Указать аудитории новый путь достижения результатов.
4. Заинтересовать аудиторию, чтобы она одобрила предложения, содержащиеся в презентации.

5. Творчески мобилизовать аудиторию. Дать людям возможность давать свои оценки внутри информационного поля презентации, что сделает их соучастниками процесса.

Контакт презентатора с аудиторией в значительной степени зависит от понятийного пространства и языка мышления аудитории, от ее национальных, культурных и социальных особенностей. Это наиболее сильные факторы влияния на отношение аудитории к презентатору и его презентации. Аудитория прежде всего оценивает презентатора и его сообщение сквозь фильтр категорий и установок своего мышления, традиций и привычек, сложившихся в данном социуме.

Ниже приведены причины неэффективной презентации, описывающие действия презентатора, которые он не должен допускать.

Выглядит или говорит взволнованно. Нервозность может проявляться по-разному: у презентатора скованы движения, или, наоборот, он постоянно ходит взад и вперед, не знает, куда деть руки, или безостановочно ими что-то делает. Иногда у него дрожит голос, или, еще хуже, он сам признается в том, что слишком волнуется.

Говорит монотонно. Абсолютно не имеет значения то, что пытается сообщить презентатор: если его голос монотонен – аудитория засыпает.

Говорит слишком тихо. Если оратора почти не слышно дальше первого ряда, обычно никто и не прилагает особых усилий, чтобы хоть что-то воспринять.

Плохо поддерживает зрительный контакт. Если во время разговора говорящий смотрит вниз, в сторону или вверх аудитории, получатели информации часто отключаются и перестают слушать.

Считывает информацию со слайдов и раздаточных материалов. Наглядные пособия могут намного улучшить впечатление от презентации. Но когда говорящий считывает почти все со своих слайдов, он обижает слушателей, своими действиями демонстрируя, будто представители аудитории не в состоянии разобраться во всем сами.

Использует слишком мелкий шрифт на визуальных материалах. Если люди не видят мелкие детали на слайдах или плакатах, то перестают обращать на них внимание и могут вообще перестать воспринимать речь говорящего, если презентация во многом зависит от неудачно составленных наглядных пособий.

Употребляет слишком много специальной терминологии. Не все знакомы с такой терминологией, и если говорящий употребляет ее слишком часто, он и аудитория перестают понимать друг друга.

Перегружает данными. Если презентация состоит исключительно из данных и других подробностей, слушатели забывают, что же было главным в сообщении.

Не умеет обращаться с мультимедийной техникой. Если адресант, решив прибегнуть к помощи техники, совсем не умеет с нею обращаться и своими неловкими действиями нарушает ход презентации, слушатели сосредоточатся на чем угодно, но только не на его выступлении.

Нечетко формулирует свои мысли. Это обратная сторона промаха, когда выступающий предоставляет слишком много малозначимых подробностей или говорит туманно, вы слышите лишь общие, банальные фразы и невозможно уловить смысл высказывания. Перескакивает с предмета на предмет. В этом случае становится непонятным, что именно хотел сказать выступающий, потому что он все продолжает и продолжает говорить, а слушающие теряют нить и идею речи.

Слишком часто пытаются казаться забавным. Юмор в официальной презентации может оказаться весьма кстати, но если говорящий шутит через силу или чересчур много, его неудачные остроты вызывают неприятие у аудитории.

Не делает никаких выводов в конце разговора. На некоторых презентациях единственным подтверждением их завершения является факт, что адресант прекратил говорить. В его сообщении нет заключительной части, и он вообще не прилагает никаких усилий сделать какие-то выводы из своего высказывания.

Помните, что даже самое современное программное обеспечение презентаций имеет не только сильные, но и слабые стороны. Если вдруг погаснет свет или выйдет из строя техника – имейте запасной вариант.

Таким образом, эффективная презентация повышает доверие аудитории к качеству информации, достоверности и надежности полученных данных, валидности методик и соблюдению процедур измерения.

Вопросы для промежуточного контроля

1. Каково назначение и основные функциональные возможности табличного процессора MS Excel?
2. Какие типы данных можно ввести в ячейки электронной таблицы?
3. Как выполнить проверку вводимых в ячейки MS Excel значений?
4. Что такое сортировка списка и обязательно ли выделение списка при выполнении сортировки?
5. Что такое база данных и каково основное предназначение процессора MS Access?
6. Что представляет собой пакет прикладных компьютерных программ SPSS и какие задачи он может решать в измерении?
7. Какие существуют возможности у пакета компьютерных программ MS Office для трансляции результатов измерений?
8. Что такое компьютерная презентация и какие существуют режимы ее просмотра?
9. С каким расширением по умолчанию сохраняется файл презентации в MS Power Point 2007?
10. Что делает презентацию эффективной и неэффективной (каких типичных ошибок в работе с презентацией следует избегать)?
11. Как отражается эффективность презентации на качестве информации о результатах педагогических измерений?

Задания для самостоятельной работы

Задание 1. Построение таблиц, графика и расчет элементарных статистик в MS Excel.

1. Создайте таблицу по приведенному образцу.

	А	В
1	Месяц	Средний балл
2	Сентябрь	4,6
3	Октябрь	4,05
4	Ноябрь	4,3
5	Декабрь	4,54
6	Январь	4,07
7	Февраль	4,95
8	Март	4,59
9	Апрель	4,36
10	Май	4,84
11	Июнь	5,58

2. Постройте диаграмму типа графика и отформатируйте ее (для создания подписей по осям и подписей данных используйте вкладку *Работа с диаграммами – Макет*).

Примерный вид диаграммы:



3. Рассчитайте для приведенных данных среднее арифметическое, стандартное отклонение, коэффициент вариации.

Задание 2. Создание прогнозной модели в MS Excel.

Microsoft Excel дает возможность отображать данные в виде графика, а затем добавлять одну или несколько линий тренда к ней, чтобы прогнозировать появление данных в будущем.

1. Введите самостоятельно данные в таблицу Excel. Вы должны иметь достаточно данных, чтобы охватить допустимый период времени, например 2 года или более, а также иметь последовательные интервалы, например еженедельные или ежемесячные записи.

2. Выберите данные, которые необходимо включить в диаграмму.

3. Постройте график (щелкните вкладку *Вставка*, затем нажмите кнопку выпадающего меню *График* в разделе *Диаграммы* меню ленты *Вставка*. Выберите нужный график из отображенных вариантов).

4. Выберите ряд данных, для которых хотите иметь линию тренда.

5. Определите нужный тип линии тренда (выберите тип линии тренда из выпадающего списка *Линия тренда*. Если не находите нужный тип линии тренда, нажмите опцию *Дополнительные параметры* линии тренда, чтобы выбрать их в разделе *Параметры* линии тренда диалогового окна *Формат линии тренда*).

6. Установите диапазон, для которого хотите составить прогноз.

7. Выберите линию тренда (используйте вначале линейную линию тренда с прогнозом на 2–3 шага, чтобы затем перейти к более сложным).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Васильев, А. В. Работа в электронных таблицах: практикум / А. В. Васильев, О. Б. Богомолова. – Москва: БИНОМ. Лаборатория знаний, 2007. – 160 с.
2. Велихов, А. В. Основы информатики и компьютерной техники: учеб. пособие / А. В. Велихов. – Москва: СОЛОН-ПРЕСС, 2007. – 326 с.
3. Дубина, А. Excel для экономистов и менеджеров / А. Дубина. – Санкт-Петербург: Питер, 2004. – 295 с.
4. Златопольский, Д. М. 1700 заданий по Microsoft Excel / Д. М. Златопольский. – Санкт-Петербург: БХВ-Петербург, 2003. – 544 с.
5. Информатика для юристов и экономистов: учебник для вузов / под ред. С. В. Симоновича. – Санкт-Петербург: Питер, 2006. – 688 с.
6. Окишева, Т. Н. Информатика: учеб. пособие: в 2 ч. / Т. Н. Окишева, Е. Ю. Вардомацкая. – Санкт-Петербург: Питер, 2007. – Ч. 2: Excel. – 237 с.
7. Трещалин, М. Ю. Основы информационных технологий: учеб. пособие для студентов вузов / М. Ю. Трещалин. – Москва: Элит, 2007. – 108 с.
8. Шарстнев, В. Л. Компьютерные информационные технологии: курс лекций / В. Л. Шарстнев. – Витебск: УО ВГТУ, 2008. – 350 с.
9. Анализ данных с применением статистического пакета SPSS [Электронный ресурс] – Режим доступа: <http://sociolane.ru/analiz-dannyh-s-primeneniem-statisticheskogo-paketa-SPSS>. – Дата доступа: 10.05.2016.

ГЛОССАРИЙ

Анкета (questionnaire) – специально разработанный документ для проведения опроса, содержащий определенную последовательность вопросов, позволяющих унифицированно (формализованно) выяснить у респондентов их отношение к интересующей теме.

Бесповторная выборка (nonrepeated sample) – выборка, состоящая из единиц наблюдения, которые отобраны из генеральной совокупности и после регистрации в нее не возвращаются.

Валидность (validity) – характеристика адекватности результатов измерения поставленной цели их проведения; удостоверение соответствия измерительного средства своему назначению.

Вопросник (questionnaire) – см. «Анкета».

Выборка (sample) – ограниченная по численности группа единиц наблюдения, отобранных из генеральной совокупности для изучения ее свойств.

Выбросы (outliers) – нетипичные или редкие значения, которые существенно отклоняются от распределения остальных выборочных данных. Эти данные могут отражать истинные свойства изучаемого явления (переменной), а могут быть связаны с ошибками измерения или аномальными явлениями.

Генеральная совокупность (population) – это все множество объектов, в отношении которых формулируется исследовательская гипотеза.

Генеральный фактор (general factor) – фактор, который влияет на все измеряемые переменные.

Гипотеза (hypothesis) – научное предположение, формулируемое в виде обобщений или утверждений, которые позволяют по-новому рассуждать об изучаемом предмете, и предполагающее доказательство.

Гистограмма (histogram, bar graph) – один из видов графического изображения статистического распределения величины по определенному признаку. Представляет собой совокупность столбчатых диаграмм, построенных на одной прямой. Высота столбца пропорциональна частоте значения или нахождения данной величины в интервале. Столбиковая диаграмма частотного распределения.

Глубинное интервью (in-depth interview) – один из качественных методов измерения, когда по ходу беседы интервьюируемому в произвольной форме задается ряд зондирующих вопросов с тем, чтобы выяснить суть проблемы.

Группировка (grouping) – разделение множества объектов выборки по определенному признаку, объединение нескольких значений переменной в диапазон значений. Проводится для получения лучшей информации о случайной величине, нахождения наилучшей модели ее описания.

Дисперсия (variance) – мера разброса случайной величины, отклонение значений от среднего значения, пропорциональная сумме квадратов отклонений измеренных значений от их арифметического среднего.

Доверительный интервал (confidence interval) – интервал, в который с определенной вероятностью попадает оценка статистического показателя генеральной совокупности. Показывает диапазон вокруг значения статистики, в котором находится истинное значение этой статистики, с определенным уровнем надежности.

Достоверность (reliability) – критерий, определяющий степень доверия к результатам измерений. Характеризуется тем, что всегда можно удостовериться в их устойчивости и стабильности.

Единица измерения (unit of measure) – стандартное значение, относительно которого производится сравнение измеряемого объекта.

Единица наблюдения (unit of observat) – элемент объекта выборочной совокупности, носитель признаков, подлежащих измерению.

Единичность (uniqueness) – доля дисперсии наблюдаемой переменной, соответствующая специфичному фактору, которая не объясняется общими факторами.

Зависимые выборки (dependent samples) – выборки, в которых каждому объекту одной выборки поставлен в соответствие по определенному критерию объект другой выборки. Имеют одинаковый объем.

Значимость результатов (significance) – статистическая достоверность полученных результатов.

Измерение (measurement) – процедура формализованного оценивания какого-либо объекта путем соотнесения с заданной адекватной мерой и определения его значения (количественного или качественного) путем приписывания ему по определенным правилам условного обозначения (числа, буквы, слова или какого-либо другого символа). Осуществляется путем установления взаимно-однозначного соответствия между множеством объектов и множеством символов. В классическом понимании измерение – это процедура сравнения измеряемого объекта с эталоном, при котором он получает численное выражение в определенном масштабе или шкале.

Индекс – это обобщенная оценка латентной переменной, полученная путем агрегирования результатов измерения наблюдаемых переменных.

Итерация (iteration) – результат применения какой-либо математической операции, полученный в серии аналогичных операций.

Квалиметрия (qualimetry) – формальная теория меры качества, выражающая ее количественно.

Квантиль (quantile) – это точка на числовой оси измеренного признака, которая делит всю совокупность упорядоченных измерений на группы с известным соотношением их численности. Значение, которое измеренный признак не превышает с фиксированной вероятностью. Среди квантилей различают медиану, проценти и квартили.

Квартиль (quartile) – это 3 точки – значения признака, которые делят упорядоченное множество значений на 4 равные по численности части. Первый квартиль соответствует 25-му процентилу, второй – 50-му процентилу или медиане, третий – 75-му процентилу.

Кластер (cluster) – группа схожих, однородных объектов, которая может рассматриваться как самостоятельная единица, обладающая определенными свойствами. Результат кластерного анализа.

Кластерный анализ (cluster analysis) – процедура упорядочивания объектов в сравнительно однородные классы на основе попарного сравнения этих объектов по предварительно определенным и измеренным критериям. Кластерный анализ решает задачу построения классификации.

Ковариация (covariation) – мера линейной связи двух переменных, измеряется как сумма попарных произведений отклонений переменных от их среднего.

Корреляционная матрица (correlation matrix, cross-laggedpanel correlation) – это результат вычисления корреляций одного типа для каждой пары из множества переменных, измеренных на одной выборке. Применяется при корреляционных расчетах и в факторном анализе.

Корреляция (correlation) – соотношение нескольких переменных, когда изменчивость одной переменной соотносится с изменчивостью другой переменной.

Корреляционный анализ (correlation analysis) – метод, направленный на проверку статистической связи между несколькими переменными.

Коэффициент детерминации (coefficient of determination) – степень зависимости одной переменной от изменчивости второй переменной. Вычисляется путем возведения в квадрат коэффициента корреляции.

Коэффициент корреляции (coefficient of correlation) – количественная мера взаимосвязи двух переменных.

Латентная переменная (latent variable) – переменная, которая не поддается непосредственному наблюдению или измерению. Например, любая социальная установка носит латентный (скрытый) характер.

Ложная связь (false correlation) – наличие наблюдаемой связи переменных, между которыми нет никаких отношений.

Математическое ожидание (average of distribution, expectation) – числовая характеристика случайной величины, вокруг которой сосредоточены значения случайной величины, координата центра тяжести распределения.

Матрица (matrix) – всякая прямоугольная таблица чисел, состоящая из m строк и n столбцов.

Медиана (median) – такое значение признака, которое делит упорядоченное множество данных пополам так, что одна половина всех значений оказывается меньше медианы, а другая – больше.

Меры разброса (measure of spread) – статистические показатели разброса значений переменной относительно меры центральной тенденции. Основные меры – размах, дисперсия, среднеквадратическое отклонение, коэффициент вариации.

Меры центральной тенденции (central measure) – статистические показатели, характеризующие наиболее выраженное, репрезентативное значение переменной. Основные из них: мода, медиана, среднее значение.

Многомерные шкалы (multidimensional scales) – шкалы, используемые для измерения свойств объекта, характеризуемого двумя и более параметрами. Результаты измерения выражаются двумя и более значениями.

Множественная регрессия (multiple regression) – обобщение парной регрессии на несколько независимых переменных; построение модели связи между несколькими независимыми и зависимой переменными для изучения возможности предсказания некоторого результата по ряду предварительно измеренных характеристик.

Мода (mode) – это такое значение из множества измерений, которое встречается наиболее часто. Моде, или модальному интервалу признака, соответствует наибольший подъем (вершина) графика

распределения частот. Если график распределения частот имеет одну вершину, то такое распределение называется унимодальным.

Мощность критерия (power of test) – величина, характеризующая способность критерия отклонять нулевую гипотезу, когда она неверна.

Надежность (questionnaire) – воспроизводимость результатов измерения в повторных и других измерениях. Обычно определяется вычислением коэффициента корреляции результатов измерений случайной подвыборки с выборкой, при сравнении двух половин выборки, а в случае, если это различные измерения одного и того же объекта, выполненные различными субъектами примерно в одно и то же время, – взаимной согласованностью их оценок. Надежной является также мера, дающая те же результаты при многократном измерении одних и тех же объектов.

Независимые выборки (independent samples) – выборки, объекты которых не поставлены ни в какое соответствие друг с другом. Вероятность отбора объектов в одну выборку не зависит от вероятности отбора объекта в другую выборку. Могут иметь разные объемы.

Непараметрические критерии (nonparametric tests) – критерии, свободные от распределения, не базируются на предположении о типе распределения генеральной совокупности и не используют параметры этой совокупности.

Нольмерное шкалирование (threshold scale) – построение шкалы, имеющей единственное (пороговое) значение и нулевую размерность.

Нормальное распределение (normal distribution) – вид распределения, который симметричен относительно среднего значения. Крайние значения встречаются редко, и частота постепенно повышается к серединным значениям признака. Меры центральной тенденции совпадают. Асимметрия и эксцесс равны нулю.

Нормирование (normalization) – преобразование данных, при котором начало координат помещается в точку, соответствующую среднему арифметическому всех объектов, а цена деления соответствует единице стандартного отклонения значений.

Нулевая гипотеза (null-hypothesis) – основная гипотеза, проверяемая на выборочной совокупности о неизвестных параметрах случайной величины. Рассматривается как наиболее вероятное решение о случайной величине на основе априорной модели, утверждает о сходстве параметров выборочной совокупности и модели случайной величины.

Общий фактор (common factor) – неизмеримая скрытая величина, которая определяет изменчивость многих переменных.

Общность (communality) – доля дисперсии наблюдаемых переменных, обусловленная общими факторами.

Объективность (objectivity) – независимость результатов измерения от субъекта измерения (его воли, чувств, желания, ожиданий и т. д.).

Объем выборки (amount of sampling) – количество объектов (единиц наблюдения) в выборке.

Одномерные шкалы (one-dimensional scales) – шкалы, используемые для измерения свойств объекта, характеризующиеся одним числовым значением. Объекты отображаются в числовой системе одним-единственным числом.

Операциональная валидность (operational validity) – соответствие операций, через которые фиксируется изучаемый признак, теоретическому описанию переменных.

Описательные статистики (descriptive statistics) – критерии, обобщающие выборочные значения; дают представление об анализируемых данных. К ним относят меры центральной тенденции, меры разброса, минимальное и максимальное значения, квантили. Также для обобщения результатов используют графические способы и таблицы.

Отношение (relation) – математическая структура, которая формально определяет свойства различных объектов и их взаимосвязи. Примерами могут быть равенство, делимость, транзитивность и пр.

Ошибка измерения (measurement error) – ошибка, вызванная любым фактором, вводящим погрешность в измерение некоторой переменной.

Параметрические критерии (nonparametric tests) – критерии, основанные на определенном типе распределения генеральной совокупности (как правило, нормальном); используют в расчетах параметры этой совокупности (среднее, дисперсии и т. д.).

Параметры распределения (population moments) – числовые характеристики распределения, которые описывают данный вид распределения. Выделяют следующие параметры: математическое ожидание или среднее арифметическое, дисперсию, асимметрию и эксцесс. Нормальное распределение является двухпараметрическим, так как его можно описать через два параметра: среднее значение и среднеквадратическое отклонение.

Переменная (variable) – любое свойство объектов, которое может изменяться.

Повторная выборка (replicate sample) – выборка, состоящая из объектов, которые отобраны из генеральной совокупности и после регистрации снова в нее возвращаются.

Полигон частот (polygon of frequency) – один из способов графического представления плотности вероятности случайной величины; представляет ломаную линию, соединяющую значения частот по градациям признака.

Правильность (correctness) – соответствие определенным правилам.

Простая структура (simple structure) – специальный термин, относящийся к факторной структуре, которая обладает определенными свойствами простоты: переменные должны иметь максимальную нагрузку на минимальное число общих факторов, каждый фактор должен нагружать некоторые переменные и не нагружать остальные. Простота факторного решения заключается в том, что переменная имеет наиболее простое факторное объяснение, т. е. преобладающее влияние одного фактора, и в меньшей степени связана с другими.

Процентиль (percentile) – 99 значений признака, делящих упорядоченное множество значений на 100 равных частей.

Психометрия (psychometrics) – область математической психологии, разрабатывающая математические основы психологических измерений, психометрические показатели тестов.

Размах (range of sample) – разность между максимальным и минимальным значением признака.

Ранг (rank) – место значения в упорядоченном ряду значений, величина значения относительно других значений в совокупности.

Рандомизация (randomization) – стратегия случайного отбора или распределения объектов, при котором все объекты имеют равные шансы попасть в выборку. Обеспечивает репрезентативность выборки.

Ранжирование (ranking) – определение порядка согласно рангу.

Регрессионный анализ (regression analysis) – статистический метод исследования влияния одной или нескольких независимых переменных на зависимую переменную.

Репрезентативность выборки (representativeness of sample) – степень адекватности отражения свойств и характеристик изучаемой генеральной совокупности; достигается случайным отбором.

Случайная величина (variate) – величина, которая в результате испытания принимает то или иное числовое значение, заранее

неизвестное, т. е. случайно принимает одно значение из множества возможных.

Среднее значение (mean) – значение, полученное суммированием всех значений определенного набора данных и делением суммы на общее количество значений; относится к мерам центральной тенденции.

Стандартизация (standardization) – преобразование данных в стандартную шкалу на основе нормального распределения, дающее возможность их корректного сравнения и проведения анализа.

Стандартное (среднеквадратическое) отклонение (standard deviation) – величина разброса значений относительно среднего; находится как квадратный корень дисперсии.

Статистическая гипотеза (statistical hypothesis) – это утверждение относительно неизвестного параметра. Различают нулевую гипотезу и альтернативную.

Степени свободы df (degrees of freedom) – число свободно варьирующих единиц в составе выборки.

Тренд (trend) – это основная тенденция динамики показателя, очищенная от случайных влияний и индивидуальных особенностей отдельных периодов. Может описываться различными уравнениями – линейными, логарифмическими, степенными и т. д. Фактический тип тренда устанавливается на основе графического изображения реальных данных временного ряда, путем осреднения показателей динамики ряда, на основе статистической проверки гипотезы о постоянстве параметров тренда.

Уровень значимости p (p-level) – это вероятность отклонения основной гипотезы, когда она верна. Обычно используют три уровня статистической значимости: $p < 0,05$; $p < 0,01$; $p < 0,001$.

Фактор (factor) – гипотетические, непосредственно не измеряемые, скрытые переменные, которые обуславливают изменчивость других переменных. Параметр внешних условий либо особенностей объекта, влияющий на изменение зависимой переменной.

Факторная матрица (factorial matrix) – таблица, описывающая факторные нагрузки для анализируемых переменных относительно выделенных факторов.

Факторная нагрузка (factorial loading) – значение, выражающее степень, в которой данный фактор объясняет общую вариативность корреляций, на которых был проведен факторный анализ. Степень взаимосвязи соответствующих переменных и факторов: чем больше

абсолютная величина факторной нагрузки, тем больше данная переменная обусловлена действием соответствующего фактора.

Факторный анализ (factor analysis) – многомерный метод, применяемый для изучения взаимосвязи множества переменных для выявления причин совместной изменчивости нескольких переменных – факторов. Основной целью является уменьшение размерности исходных данных с целью их экономного описания при условии минимальных потерь исходной информации. Результатом факторного анализа является переход от множества исходных переменных к меньшему числу новых переменных – факторов.

Фокус-группа (focus group) – метод неструктурированного опроса представителей различных социальных позиций, образующих группу от 8 до 12 человек, в которой под руководством модератора (ведущего) обсуждаются конкретные проблемы. Требуется большого опыта со стороны модератора, который предлагает тему дискуссии, стимулирует группу к ее обсуждению и поддерживает групповую динамику.

Характерный (специфичный) фактор (unique factor) – фактор, влияющий только на данную переменную.

Цензурирование (sample censoring) – процесс удаления из выборки выбросов.

Частная корреляция (partial correlation) – процедура оценки эффекта, вызываемого третьей переменной; взаимосвязь двух переменных за счет того, что обе они согласованно меняются под влиянием некоторой третьей переменной.

Шкала измерения (scale of measurement) – измерительный инструмент, правила приписывания каждому объекту определенного значения (числа или другого символа), определяющие допустимые преобразования и типы отношений, отображаемых соответствующей шкалой. Выделяют следующие типы шкал: номинальная (номинативная, наименований); порядковая (ранговая, ординарная); интервальная (шкала вычитания); шкала отношений (равных отношений). Интервальную и отношений нередко называют *шкалами измерения*, а номинальную и порядковую – *шкалами оценки*.

Номинальная шкала. Используется для оценки объектов и их свойств, обозначенных именами (пол, семейное положение, род деятельности и т. д.) для их идентификации и классификации. Отношения между шкальными значениями – отношения тождества и различия. Допустимые расчеты – процент, доля, мода.

Порядковая шкала. Измеряет степень проявления чего-либо или его значимости в ряду других. Отношения между шкальными значениями –

иерархия признаков, сравнение, отношение неравенства (больше, меньше, равно, не равно). Допустимые расчеты – процент, доля, мода, медиана.

Интервальная шкала. Классическим примером измерений по шкале интервалов является измерение температур по шкале Цельсия с двумя реперными точками – замерзания и кипения воды. Отношения между шкальными значениями – равенство, неравенство, больше, меньше, больше на..., меньше на..., отношения между интервалами. Допустимые статистические расчеты – процент, доля, мода, медиана, среднее арифметическое, дисперсия, среднеквадратическое (стандартное) отклонение.

Шкала отношений. Измеряет стаж работы, возраст, доход. Отношения между шкальными значениями – равенство, неравенство, больше, меньше, больше на..., меньше на..., больше в..., меньше в... Допустимые статистические расчеты – процент, доля, мода, медиана, среднее арифметическое, дисперсия, среднеквадратическое отклонение.

Шкала оценки (scale of rating) – см. «Шкала измерения».

Эксперт (expert) – специалист, компетентный в данном вопросе (от лат. *expertus* – опытный) и способный дать квалифицированную оценку. Компетентность эксперта в отношении объекта исследования – это профессиональная компетентность, а в отношении методологии принятия экспертного решения – экспертная компетентность. Эксперт должен быть не только беспристрастным и объективным в своей оценке, но и способным к такому оцениванию.

Экспертный опрос (expert poll) – метод или процедура получения первичной информации от экспертов в устной или письменной форме.

Экспертное заключение (expert opinion) – итоговый документ, оформленный в соответствии с установленными требованиями и содержащий мотивированную экспертную оценку, т. е. мнение эксперта о предмете.

Экспертная оценка (expert judgement) – один из методов измерения и сбора первичных данных, основанный на использовании опыта и знаний компетентных лиц (экспертов), с возможностью последующей обработки собранной информации. Экспертные оценки используют для прогнозирования событий будущего, если статистические данные отсутствуют или их недостаточно. Они применяются также при оценке событий настоящего, для которых нет других методов измерения, например оценка важности педагогических целей или предпочтительность отдельных методов обучения.

ПРИЛОЖЕНИЕ

Сводные данные об объектах и методах измерения качества образования в вузе

Задачи измерения	Объекты измерения	Нормируемые требования	Методы измерения
1	2	3	4
Контроль и отбраковка продукции	Учебные программы и планы	Соответствие заданиям, полученным от ректората, квалификационным требованиям, аналогам и нормам законодательства	Самоаккредитация, экспертиза учебных документов
	Методическая литература	Соответствие учебным программам. Актуальность содержания. Грамотность и логичность изложения	Рецензирование, редактирование, нормоконтроль
	Расписание занятий	Соответствие нормативам, учебным программам. Выполнение договорных обязательств вузом	Проверка соответствия по запланированным часам. Согласование с заинтересованными сторонами
	Учебное оборудование	Соответствие учебным программам. Техническая исправность. Безопасность	Инспектирование. Анкетирование студентов и преподавателей
	Знания рефлектантов	Соответствие установленным для учебной программы требованиям	Централизованные вступительные экзамены
	Текущая успеваемость	Выполнение графика индивидуальной работы обучающегося	Зачеты по лабораторным, домашним и прочим работам. Рассмотрение на заседаниях кафедр
	Знания по учебной дисциплине	Усвоение теоретического минимума. Наличие предусмотренных программой практических навыков	Экзамены, зачеты, защита отчетов по практикам
	Готовность выпускной работы к защите	Решение задач выпускной работы. Соблюдение графика работы. Качество оформления	Нормоконтроль, предзащита

Продолжение

1	2	3	4
	Квалификация выпускников	Соответствие квалификационным требованиям. Решение задач, сформулированных в задании на выпускную работу	Защита выпускной работы и выпускной экзамен
Управление процессами и ресурсами	Квалификация преподавателей	Соответствие квалификационным требованиям	Выборы на должность, ежегодная аттестация, открытые занятия
	Готовность аудиторий и лабораторий к проведению занятий	Техническое состояние. Соответствие нормативам	Приемка специалистами технической службы
	Выполнение учебных планов	Соответствие годовому рабочему учебному плану и методическому плану	Проверка по записям в журналах занятий. Сравнение ведомостей с учебным планом
	Соблюдение расписания занятий	Соответствие правилам внутреннего распорядка	Выборочные проверки, учет срывов и переносов занятий
	Посещаемость обучающихся	Выполнение графика самостоятельной работы	Ведение журналов посещаемости
	Методика преподавания	Правила проведения аудиторных занятий	Открытые занятия с обсуждением на заседании кафедры
	Соблюдение учебных планов	Соответствие аккредитованным учебным программам и условиям аккредитации	Самоаккредитация
Обеспечение эффективности системы управления	Организационная структура и правила функционирования	Соблюдение норм законодательства и условий аккредитации	Экспертиза проектов руководящих документов
	Процессы управления и взаимодействия между уровнями управления	Выполнение внутренних руководящих документов	Внутренний аудит качества
	Повседневная работа персонала и подразделений	Выполнение организационных планов (от плана работы вуза до индивидуальных планов преподавателей). Выполнение распоряжений руководителей	Ежегодная аттестация персонала. Ежегодные отчеты кафедр, факультетов, ректората

1	2	3	4
	Эффективность системы менеджмента качества	Соответствие требованиям стандарта ISO 9001	Внутренний аудит качества
Определение целей развития и улучшения качества	Потребности рынка и запросы потенциальных потребителей (абитуриентов)	Соответствие предложений вуза актуальному спросу и перспективным потребностям рынка	Анкетирование учащихся выпускных классов в школах и абитуриентов
	Содержание образовательных услуг конкурентов и ведущих учебных заведений	Соответствие учебных программ и уровня качества услуг вуза передовым образцам	Сравнение программ обучения в процессе самоаккредитации
	Удовлетворенность обучающихся	Повышение удовлетворенности качеством услуг. Своевременное выявление проблем	Анкетирование обучающихся
	Удовлетворенность выпускников и их работодателей	Конкурентоспособность и востребованность выпускников. Признаваемость дипломов вуза	Анкетирование выпускников и работодателей
	Достижение целей института	Выполнение поставленных целей. Своевременное выявление препятствий	Экспертная оценка достижений и анализ на уровне правления АО
	Достижение целей подразделений	Выполнение поставленных целей. Своевременное выявление препятствий	Экспертная оценка достижений и анализ со стороны ректората
	Эффективность корректирующих действий	Выполнение запланированных мероприятий. Устранение причин несоответствий	Экспертная оценка достижений и анализ со стороны ректората

Примечания: 1. Продукция, подлежащая контролю и отбраковке (первый этаж пирамиды качества) – это результаты деятельности персонала и подразделений вуза на различных этапах жизненного цикла предоставления услуг.

2. Учебные пособия и лабораторное оборудование предоставляются в распоряжение обучающихся и, таким образом, входят в состав предоставляемой потребителям продукции. (Можно отнести эти объекты контроля и к ресурсам.)

3. Управление процессами предоставления образовательных услуг предусматривает контроль условий проведения занятий, состояния ресурсов (человеческих, материальных, информационных), контроль соблюдения программ, планов, расписания занятий.

4. На уровне обеспечения качества решается задача управляемости вуза с целью поддержания достигнутого уровня качества. Для этого контролю подлежат процессы управления и система управления в целом.

5. На уровне развития и улучшения качества оцениваются характеристики среды, в которой функционирует вуз. На основе этих оценок принимаются соответствующие решения об улучшении продукции (образовательных услуг) и процессов управления.

СОДЕРЖАНИЕ

Предисловие.....	3
Л е к ц и я 1. Методологические основы педагогических измерений	4
Л е к ц и я 2. Методы количественной оценки и шкалирование	17
Л е к ц и я 3. Методы оценки качественных признаков и проблемы надежности измерения	31
Л е к ц и я 4. Современные информационные технологии обработки данных, представления и трансляции результатов измерений	49
Глоссарий	62
Приложение	74

Учебное издание

Трапянок Николай Григорьевич

ОСНОВЫ ПЕДАГОГИЧЕСКИХ ИЗМЕРЕНИЙ

Курс лекций

Редактор *Е. В. Ширалиева*

Технический редактор *Н. Л. Якубовская*

Компьютерная верстка *Е. А. Радченко*

Подписано в печать 18.04.2018. Формат 60×84 ¹/₁₆. Бумага офсетная.

Ризография. Гарнитура «Таймс». Усл. печ. л. 4,42. Уч.-изд. л. 3,45.

Тираж 75 экз. Заказ .

УО «Белорусская государственная сельскохозяйственная академия».

Свидетельство о ГРИИРПИ № 1/52 от 09.10.2013.

Ул. Мичурина, 13, 213407, г. Горки.

Отпечатано в УО «Белорусская государственная сельскохозяйственная академия».

Ул. Мичурина, 5, 213407, г. Горки.