

МИНИСТЕРСТВО ОБЩЕГО И ПРОФЕССИОНАЛЬНОГО
ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ЧЕЛЯБИНСКИЙ ГОСУДАРСТВЕННЫЙ
ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ

Е. В. Яковлев

Педагогический эксперимент:
квалиметрический аспект

ЧЕЛЯБИНСК — 1998

УДК 371
ББК 74.00
Я 47

Яковлев Е. В. Педагогический эксперимент: квалиметрический аспект: Монография: – Челябинск: Издательство ЧГПУ, 1998. – 136 с.

В монографии рассматриваются основные идеи, технология и методы организации и оценки качества результатов педагогического эксперимента с позиций квалиметрического подхода.

Работа рассчитана на научных работников в области педагогики, преподавателей и студентов педагогических вузов, учителей-исследователей.

Рецензенты: Н. А. Томин, доктор педагогических наук,
профессор ЧГПУ
М. М. Кипнис, доктор физико-математических
наук, профессор ЧГПУ

ISBN 5–85716–213–0

© Е. В. Яковлев, 1998
© Издательство ЧГПУ, 1998

Введение

Повышение эффективности и качества педагогического исследования является актуальной проблемой. Её решение во многом связано с уровнем формализации педагогических знаний и их теоретическим обобщением. Большие эвристические возможности в этом плане содержатся в квалиметрическом подходе, базирующемся на концептуальных положениях и методах педагогики, теории измерения, теории моделирования и математической статистики. Использование математических методов в педагогике позволит внести необходимую строгость, чёткость в понимание исходных данных, в постановку исследовательских задач, их решение, интерпретацию полученных результатов, осуществить прогноз. И, что немаловажно, достижения методологии и методики исследования за счёт математических методов расширяют область доступных изучению научных проблем.

Квалиметрический подход позволяет по-новому подойти к осмыслению проблем из области методологии педагогики и методики педагогических исследований и решить ряд актуальных задач, к числу которых следует отнести: 1) технологию планирования и организации эксперимента, включая разработку технологии моделирования, сравнимости экспериментальных и контрольных объектов; 2) разработку технологии оценки результатов педагогического эксперимента, в том числе доказательства необходимости и достаточности условий эффективного протекания педагогического процесса, разработку технологии прогнозирования результатов исследования.

Цель настоящей работы — рассмотреть пути и методы решения поставленных задач.

Глава I. Теоретико-методологические основы организации и проведения педагогического эксперимента

§1. Сущность и особенности педагогического исследования

Любое научное исследование — это специально организованный процесс познания, в котором происходит выработка и теоретическая систематизация знаний о действительности. В случае педагогического исследования речь идёт о педагогической действительности.

Педагогическое исследование подчиняется общим требованиям, выработанным теорией и практикой научных исследований в области общественных наук. К таким требованиям относят объективность, точность, воспроизводимость, доказательность (89, с. 510). Педагогическое исследование обладает рядом особенностей, присущих только ему, а, именно, оно всегда требует особого внимания, чуткости исследователя, поскольку объектом его являются люди; оно, как правило, ориентировано на максимальную практическую пользу при изучении самого сложного, иногда чисто теоретического вопроса; педагогическое исследование — это совместное творчество исследователя, учителя и учащихся (воспитанников) (49). Эти особенности, на наш взгляд, отражают гуманистическую направленность педагогического исследования.

Особенности педагогического исследования можно рассмотреть и в иной плоскости. Так, в педагогическом исследовании мы имеем дело с изменчивостью протекающих здесь процессов, что определяет уникальность каждой педагогической ситуации; включением человека как элемента системы, что порождает определённый

ные трудности на всех этапах исследования; невозможностью исследования данной системы без учёта, а иногда и изменения других смежных с нею систем.

Далее, в педагогическом исследовании мы имеем дело с педагогическим процессом, который принадлежит к нелинейным системам (при изменении одного из элементов нелинейной структуры другие изменяются не пропорционально, а по более сложному закону). Поэтому исследование её структуры не может быть осуществлено изучением её отдельных элементов, так как сумма действий компонентных причин, действующих отдельно, не равна тому следствию, которое получается при их совместном действии. В этом заключается важная методологическая особенность педагогического исследования.

В педагогической литературе делались попытки классифицировать педагогические исследования. Наибольшее распространение получили следующие классификации педагогических исследований: 1) фундаментальные, прикладные разработки (учебники, учебные пособия, программы и т.п.); 2) теоретические, теоретико-практические, практические; 3) теоретические, экспериментальные. Первые две классификации отражают цель и направленность педагогического исследования, третья — метод.*

* Существует классификация научно-педагогических исследований, в основу которой положен фасетный метод, т.е. деление исследований на независимые классификационные группировки, которые характеризуют определённые свойства объектов классификации. Выделяют четыре самостоятельных фасета (задачи исследования, результаты исследования, адрес пользователя, вид публикации), которые раскрывают свойства исследований с точки зрения их теоретической и практической направленности. Подробно эта классификация изложена в монографии В. М. Полонского (78).

В философии выделяют два взаимосвязанных уровня научного исследования: эмпирический и теоретический. На первом уровне устанавливаются новые факты науки и на основе их обобщения формулируются эмпирические закономерности. На втором уровне выдвигаются и формулируются общие для данной предметной области закономерности, позволяющие объяснить ранее открытые факты и эмпирические закономерности, а также предсказать и предвидеть будущие события и факты (82, с. 542).

В педагогике (18 и др.) выделяют ещё один уровень — методологический. На этом уровне на базе эмпирических и теоретических исследований формируются общие принципы и методы исследования педагогических явлений, построение теории.

Проблема эмпирического и теоретического в отечественной педагогике является на протяжении последних двадцати лет одной из ведущих. Этой проблеме посвящены работы Б. С. Гершунского, В. В. Давыдова, В. В. Краевского, И. Я. Лернера, М. Н. Скаткина, В. М. Розина, В. С. Швырёва, Г. П. Щедровицкого и др.

В статье А. Я. Данилюк (31) выделены и проанализированы два основных противоположных направления, которые в новых условиях получили иное общественное звучание. Первое направление — методолого-педагогический нигилизм (Б. С. Гершунский, В. М. Розин и др.), второе — методолого-педагогический традиционализм (В. В. Краевский, И. Я. Лернер, М. Н. Скаткин и др.). Каждое из этих направлений внесло свой вклад в понимание изучаемой проблемы и наметило новую методологическую парадигму. Сегодня проблема эмпирического и теоретического концентрируется в плоскости методов педагогического исследования. Тем не менее, оба уровня педагогического исследования, и теоретический,

и эмпирический, как правило, отражают ступени одного конкретного исследования.

В соответствии с этим общеметодологическим положением педагогическое исследование включает два крупных блока — теоретический и экспериментальный. Если теоретические основы педагогического исследования и их оценка достаточно полно представлены в ряде научных публикаций (70; 78; 112), то этого нельзя сказать об его экспериментальной части. Поэтому сосредоточим своё внимание именно в этой области.

В научной литературе главным в этой части считают эксперимент. Прежде чем перейти к анализу этого метода научного познания, остановимся на основных положениях методологии и методики педагогического исследования.

В современных философских работах (В. А. Лекторский, В. С. Швырёв, Э. Г. Юдин и др.) выделяют четыре уровня методологии. Первый уровень — философская методология — анализ общих принципов познания и категориального строя науки в целом. Второй уровень — общенаучные (наукovedческие) знания. Основная их функция состоит в исследовании общенаучных концепций, воздействующих на все или большую часть научных дисциплин. К числу этих знаний относят, в том числе математические и кибернетические методы. Третий уровень — конкретно-научные методологические знания (например, методология педагогики). В задачу конкретно-научной методологии входит описание методов, принципов исследования и процедур, применяемых в той или иной специальной научной дисциплине. Четвёртый уровень методологии — методика и техника научного исследования. Этот уровень предусматривает разработку выбора процедур, обеспечивающих получение эмпирического материала и его первичную обработку.

Каждый из вышеназванных уровней выполняет определённые функции в научном познании, а в совокупности они образуют сложную систему, в которой между ними существует определённое соподчинение.

Отметим, что методы специальных наук соотносятся с общенаучными методами как особенное с общим. В каждом конкретном методе проявляются в той или иной мере общие принципы научного мышления.

В педагогической литературе (38; 49; 65) при планировании и организации исследовательского процесса выделяют следующие познавательные процедуры: мотивировку темы, формулирование её проблемных аспектов, выявление в объекте научного познания явлений, составляющих предмет исследования, построение гипотез, определение цели и задач исследования, выбор методов исследования, адекватных исследовательским задачам, внедрение в практику результатов исследования.

Важное место в этом перечне отводится выбору комплекса методов исследования и преобразования их в конкретные исследовательские методики.

Понятие «метод научного исследования» по сути дела категория, объединяющая и формы научного мышления, и общие модели исследовательских процедур, и способы (приёмы) выполнения исследовательских действий.

Под методами педагогического исследования понимают определённые совокупности приёмов и операций, направленных на изучение педагогических явлений и решение разнообразных научных проблем учебно-воспитательного характера.

В научной литературе встречается несколько классификаций методов исследования. В зависимости от цели исследования выде-

ляют методы теоретического поиска и методы выявления путей совершенствования практики, т.е. теоретические и практические методы. В зависимости от источников накопления информации методы подразделяют на методы изучения теоретических источников и методы анализа реального педагогического процесса. В зависимости от развития исследования выделяют методы изучения состояния проблемы, экспериментального поиска новых решений проблемы, обработки данных эксперимента и т.п. По способу обработки и анализу данных исследования — методы качественного анализа и методы количественной обработки результатов исследования (статистические и нестатистические). В зависимости от форм причинности, используемых при анализе полученных результатов, различают детерминистские и вероятностные методы.

Наиболее «рабочей» является классификация методов, в основе которой заложено деление на теоретические и эмпирические методы. В эту классификацию удачно вписываются так называемые непедagogические методы, т.е. методы математики, кибернетики и т.п., которые широко используются педагогией, но пока ещё не настолько адаптированные ею, чтобы приобрести статус педагогических.

В настоящей работе мы не ставили перед собой задачу подробного раскрытия классификаций методов исследования и их педагогических характеристик, поскольку по данному вопросу изда ны обстоятельные работы (13; 14; 37; 49; 51; 87 и др.).

Очевидно, что успех педагогического исследования определяется состоянием разработанности исследовательских методик, обеспечивающих получение достоверного фактического материала.

Разработанные методики должны отвечать таким требованиям, как информативность, полнота, экономичность, прогностичность.

Естественно, что методы педагогического исследования совершенствуются под воздействием других наук, а также областей самой педагогической науки. В том числе и поэтому разработка исследовательских методик является процессом творческим. Идёт постоянный процесс совершенствования методов, а, следовательно, и методик педагогического исследования. Этот процесс, как отмечается в коллективной монографии (94), идёт по следующим направлениям: 1) совершенствование методик в зависимости от содержания и цели научного поиска; 2) периодическое обобщение исследовательского опыта, сложившегося за определённый период развития педагогической науки; 3) выявление новых возможностей известных методов.

Мы выделяем ещё такие направления, как: 1) «обогащение» известных методов за счёт приращения знаний из других наук; 2) оптимальный выбор и сочетание методов, их комплексное использование.

В. И. Загвязинским (37, с. 148) выделены некоторые критерии выбора методов исследования: 1) адекватность объекту, предмету, общим задачам исследования, накопленному материалу; 2) соответствие современным принципам научного исследования; 3) научная перспективность, т.е. обоснованное предположение о том, что выбранный метод даст новые и надёжные результаты; 4) соответствие логической структуре (этапу) исследования; 5) направленность на всестороннее и гармоническое развитие личности; 6) гармоническая взаимосвязь с другими методами в единой методической системе.

Несколько с других позиций к этому вопросу подходит Ю. К. Бабанский. При выборе методов исследования он предлагает руководствоваться следующими требованиями: применять такое сочетание методов, которое позволяет получить разносторонние сведения о развитии личности, коллектива или другого объекта воспитания и обучения; применяемые методы должны обеспечить одновременно изучение деятельности; общения и информированности личности; методы должны отражать динамику развития определённых качеств, как в возрастном плане, так и в течение определённого промежутка времени; применять надо методы, которые позволяют получить сведения из возможно большего числа источников; методы должны позволять анализировать не только ход процесса, его результаты, но и условия, в которых он функционирует (13, с. 69).

Подводя итоги вышесказанному, отметим, что нет плохих или хороших методов, а есть адекватное или неадекватное их применение. Учёные отмечают (83), что исследователю-гуманитарию зачастую приходится использовать тот или иной метод анализа данных не потому, что он наиболее целесообразен, а лишь в силу его доступности. Это объясняется тем, что педагог-исследователь не всегда может найти информацию относительно оснований для такого выбора. Устранению такого положения дел, на наш взгляд, будет способствовать наша работа.

Определившись с основными положениями методологии и методики педагогического исследования, сконцентрируем своё внимание на эксперименте. Как же определяется это научное понятие? Обратимся к справочным и нормативным источникам, в которых делается попытка обобщить имеющиеся по этому поводу высказывания. В энциклопедическом словаре эксперимент определяется

как «чувственно-предметная деятельность в науке; в более узком смысле — опыт, воспроизведение объекта познания, проверка гипотез и т.п.» (89, с. 1545). В философском словаре это понятие трактуется как «исследование каких-либо явлений путём активного воздействия на них при помощи создания новых условий, соответствующих целям исследования, или же через изменение течения процесса в нужном направлении» (101, с. 558).

Вообще говоря, даже в этих определениях, в целом верно отображающих этот феномен, нет согласованности. Анализ вышеприведённых определений понятия «эксперимент», а также определений, сформулированных в ряде других работ (13; 17; 18; 49; 94 и др.), заставляет нас констатировать тот факт, что учёные не пришли к единому пониманию того, что такое эксперимент. Есть даже группа учёных (71), которые полагают, что лучше вовсе не пытаться давать определения того, что есть эксперимент, полагая, что это понятие не поддаётся компактному определению.

Мы не сторонники такого подхода. Нам представляется удачным определение эксперимента, данное В. Оконем, в основу которого он положил свойства, присущие этому методу познания. «Эксперимент, — отмечает он, — это метод научного исследования, основанный на инициировании некоторого процесса (явления) или воздействия на данный процесс, на таком регулировании этого процесса, которое позволяет его контролировать и измерять, а также верифицировать принятые гипотезы» (74, с. 27). Причём, эксперимент, в котором проявляются все четыре свойства, называется лабораторным, отсутствие первого означает, что это естественный эксперимент.

У учёных нет единства и по таким вопросам, как: «Какие требования предъявлять этому методу научного познания?», «Каков

удельный вес кибернетики, логики и математики при организации и оценке педагогического исследования в целом и её экспериментальной части, в частности?»).

Следует отметить, что в некоторых публикациях происходит смешение требований, которые предъявляются естественно-научному и психолого-педагогическому эксперименту, который назван формирующим. Однако, концептуальное содержание формирующего эксперимента принципиально отличается от естественно-научного по следующим пунктам: 1) естественнонаучный эксперимент строится исходя из представления, что воспроизводимый в его структуре объект существует сам по себе. Формирующий — предполагает построение объекта внутри его структуры, после чего он и становится предметом изучения; 2) естественно-научный — строится на монопредметной идеализации, а формирующий — на полидисциплинарной; 3) естественно-научный является средством организации и технологии исследовательского поиска, а формирующий — помимо исследования включает элементы проектирования; 4) естественно-научный опирается на критерии истинности, а результаты формирующего оцениваются с точки зрения реализуемости проектной идеи и анализа последствий; 5) естественно-научный эксперимент обеспечивает исследовательскую деятельность внутри социального института науки, а формирующий, кроме того, обеспечивает организацию и связь управленцев, учителей, учащихся и др. (27, с. 32).

Педагогический эксперимент рассматривался в работах Ю. К. Бабанского, Н. К. Гончарова, М. А. Данилова, Л. В. Занкова, В. Оконя и др. В этих работах сущность эксперимента и его основные функции заключаются в проверке гипотезы или гипотез о

взаимоотношениях между отдельными факторами педагогического воздействия и его результатами.

Основные характеристики эксперимента требуют: «1) изолировать изучаемое явление от влияния побочных, несущественных и затемняющих его сущность влияний и изучить его в чистом виде; 2) многократно воспроизводить ход процесса в строго фиксированных, поддающихся контролю и учёту условиях; 3) планомерно изменять, варьировать, комбинировать различные условия в целях получения искомого результата» (111, с. 89).

В научном эксперименте выделяют объективную и субъективную стороны. И если объективная сторона находит своё отражение в сущности и содержании, то субъективная — проявляется прежде всего в деятельности познающего субъекта. Так, к субъективной стороне эксперимента относят: «1) особенности органов чувств человека, воспринимающих информацию, полученную в процессе наблюдения; 2) теоретические способности и вообще деятельность логического мышления; 3) уровень научных знаний и духовной культуры; квалификация и способности экспериментатора; 4) поставленные и сформулированные цели и задачи эксперимента; 5) сама деятельность экспериментатора» (111, с. 87).

В целом соглашаясь с перечисленными характеристиками научного эксперимента и проецируя их на педагогический эксперимент, отметим, что нецелесообразно сводить эксперимент к наблюдению.

К значимым проблемам повышения эффективности педагогического исследования следует отнести проблему организации и оценки его качества. Этой проблеме посвящён ряд работ (8; 13; 19; 50; 78; 88; 109). Так, в монографии В. М. Полонского (78) нашли освещение такие вопросы, как критерии и методы определения ка-

чества научно-педагогических исследований с выделением общенаучных, типовых и конкретно-научных критериев. Особый акцент в работе сделан на таких критериях, как актуальность, научная новизна, теоретическая и практическая значимость, перспективность научно-педагогического исследования. Причём каждая характеристика оценивается с позиции: а) установления нового; б) развития и углубления известного; в) уточнения отдельных положений педагогической теории. Для оценивания качества работы с позиций её новизны В. М. Полонский предлагает использовать метод эталонного сравнения и метод антиципации (предвосхищения). Метод эталонного сравнения состоит в анализе и сопоставлении результатов выполненной работы с имеющимся эталоном. Этот метод включает такие этапы, как: составление эталона (базового варианта); анализ и классификация полученных результатов; сравнение базового варианта и результатов анализируемой работы. Метод антиципации (предвосхищения) заключается в том, что результаты анализируемой работы формируются в виде вопросов, которые задаются экспертам. Отвечая на заданные вопросы, эксперты как бы предвосхищают выводы исследования. При совпадении ответов результаты считаются известным*.

В случае ярко выраженного прикладного характера исследования вышеупомянутые критерии дополняются ещё одним — внедрение результатов, под которым понимают разработку системы мер, целенаправленной планомерной деятельности, подчинённой перестройке педагогического процесса на основе теоретических разработок.

* Примеры использования этих методов приведены в работах (35; 78).

Методы оценивания качества и эффективности педагогического исследования приобретают иное звучание в связи с новыми научными направлениями.

Так в настоящее время активно разрабатывается новое научное направление — методология и методика применения кибернетики, логики и математики в педагогике.

В. И. Михеев (67) отмечает, что в рамках этого направления разработка и внедрение стандартизированных методик с применением методов формализации и ЭВМ выступает как основная цель исследования, для достижения которой требуется решение трёх важных задач: 1) переход от субъективных качественных описаний к описанию строго количественному; 2) разработка специальных математических методов и процедур по организации и планированию педагогического эксперимента и проведения выборочного обследования; 3) совершенствование языка педагогической науки.

Современным, на наш взгляд, является высказывание Д. И. Менделеева по поводу задач, стоящих перед учёным: а) определять и выражать качество неизвестного при помощи известного; б) изолировать всё то, что может подлежать измерению, показывать численное отношение изучаемого к известному; в) определить место изучаемого в системе известного, пользуясь как качественными, так и количественными сведениями (64, с. 353).

В настоящее время остро стоит вопрос о соотношении качественных и количественных подходов в педагогическом исследовании. В этом плане интересны мысли, высказанные В. И. Загвязинским (37, с. 52) о соотношении понятий «содержательный анализ» и «качественный анализ». В. И. Загвязинский полагает, что понятие содержательного подхода включает не только качественные, но и неформализованные количественные методы. Поэтому автор дела-

ет вывод, что во всяком педагогическом исследовании нужно стремиться к органическому единству качественных и количественных подходов. Далее, он утверждает, что, во-первых, качественный анализ предшествует количественному*, и, во-вторых, качественный подход пока ограничен вследствие недостаточной строгости качественных определений и несовершенства измерительных инструментов и процедур. Мы полагаем, что проникновение количественных методов в гуманитарные науки задерживалось также и из-за невозможности полного применения классических количественных методов. Однако в последнее время эта ограниченность снимается разработкой «неклассических» количественных методов типа математическое планирование эксперимента, математическое моделирование и т.п.

По сути дела вышеприведённые рассуждения связаны с проблемой соотношения содержательных и формализованных, качественных и количественных элементов исследования. Поскольку в формализованных теориях объекты рассматриваются вне их внутренних закономерностей и вне изменения, то формальные подходы более пригодны для аспектного изучения отдельных связей и зависимостей, они полезны при моделировании и математической обработке результатов обучения.

Резюмируя вышеизложенное, отметим, что педагогическое исследование, подчиняясь общим требованиям, выработанным тео-

* Это положение убедительно аргументировано философами. Они утверждают, что процесс познания реального мира как исторически, так и логически совершается таким образом, что познание качества предшествует познанию количественных отношений. Познание количественной стороны системы есть ступень к углублению знаний всей системы. Наука движется от общих качественных оценок и описаний явлений к установлению точных количественных закономерностей (91, с. 236).

рией и практикой научных исследований в области общественных наук, имеет свои специфические особенности, которые накладывают определённый отпечаток на организацию и оценку педагогического эксперимента. Неиспользованные возможности повышения эффективности организации и оценка качества педагогического исследования и особенно его экспериментальной части содержится в целесообразном сочетании количественных и качественных методов, что, по сути, и составляет методологическую основу использования математических методов в педагогических исследованиях.

§ 2. Квалиметрический подход к организации и оценке педагогического исследования

Для решения любой исследовательской задачи определяющим является выбор теоретико-методологической стратегии, которая наиболее полно находит своё отражение в подходах к познанию. Теоретико-методологический подход в ряде работ определяется как принципиальная ориентация исследования, как точка зрения, с которой рассматривается объект изучения (способ определения объекта), как понятие или принцип, руководящей общей стратегией исследования (114, с. 46–47), и как совокупность (система) принципов, которые определяют общую цель и стратегию соответствующей деятельности, независимо от того, является ли она теоретической или практической (92, с. 27).

Исследовательские задачи решаются с позиций системно-структурного, программно-целевого, оптимизационного, кибернетико-математического и других подходов. Анализ подходов, проведённый учёными (112; 115), во-первых, позволяет утверждать, что между ними существует глубокая взаимосвязь, внутреннее единство. Это объясняется имманентной взаимосвязью (в типологическом, содержательном и генетическом аспектах) тех категорий, которые лежат в основе каждого из них. Во-вторых, каждый из этих подходов представляет собой качественно новый, единый тип средств научного познания, имеющий существенные положительные стороны. В-третьих, каждый из подходов содержит в себе не только научный аппарат исследования, но косвенно и идеи реализации научной проблемы. В-четвёртых, важной особенностью внутренней взаимосвязи общенаучных подходов является их взаимная комплементарность, дополнительность по отношению друг к другу. И, наконец, мы отмечаем такую особенность этих подходов

как неравносильность, поскольку они в разной степени точности решают поставленные проблемы.

Остановимся на кибернетико-математическом подходе (по терминологии Г. П. Щедровицкого). В его разработку внесли существенный вклад Б. П. Битинас, М. И. Грабарь, Дж. Гласс, Л. Б. Ительсон, К. А. Краснянская, Л. Н. Ланда, В. И. Михеев, Н. М. Розенберг, Дж. Стэнли, Г. Франк и др.

Очевидно, что новая идея, заложенная в этом подходе, а именно, перенос понятий и знаковых средств из математики и кибернетики в педагогику является весьма продуктивной. В то же время всякая система формальных средств и понятий любой науки связана с особым расчленением и представлением объекта изучения. Поэтому, утверждает Г. П. Щедровицкий (112, с. 76), реальный объект при исследовании будет «видеться» сквозь призму его системного представления. Схематически это можно представить следующим образом: реальный объект \leftarrow системное представление \leftarrow математическая форма описания.

Если системное представление кибернетики или математики переносится в педагогику, уже обладающую своими системными представлениями, то они объединяются друг с другом в рамках третьего, более общего системного представления, так называемого конфигулятора. После того как, конфигуратор построен, можно считать перенос нового системного представления оправданным и обоснованным (Г. П. Щедровицкий).

Анализ научной литературы убеждает нас в том, что частным случаем кибернетико-математического подхода является квалиметрический подход, обогащённый концептуальными положениями и идеями педагогики.

В настоящее время активно развивается целое научное направление, ориентированное на количественное описание качества предметов, т. е. квалиметрия. Квалиметрия (от латинского «*квали*» — «качество» и древнегреческого «*метро*» — «измерять») представляет собой область научного знания, изучающую методологию и проблематику разработки комплексных, а в некоторых случаях и системных количественных оценок качества любых объектов (предметов, явлений, процессов).

Квалиметрия, как научная дисциплина, имеет две ветви: теоретическую и прикладную. В прикладной квалиметрии сформировались различные разделы: географическая квалиметрия, строительная квалиметрия, квалиметрия механизмов и т. п. В качестве самостоятельного направления выделилась педагогическая квалиметрия, под которой понимают применение методов квалиметрии к оценке психолого-педагогических и дидактических объектов. Её основы заложены в работах отечественных (С. И. Архангельский (9; 10), И. Б. Ительсон (40), В. П. Мизинцев (66), Н. М. Розенберг (84), А. М. Сохор (90), и др.) и зарубежных (Р. Аткинсон, Г. Бауэр, Э. Кроттерс (11; 12), Дж. Гласс, Дж. Стэнли (24) и др.) учёных.

Вначале была подвергнута сомнению правильность приложения понятия «измерение» к понятию «качество», а также оспаривалась принципиальная возможность получения комплексных количественных показателей качества. Эти возражения были сняты и об этом подробно изложено в публикациях, в частности, в работах Г. Г. Азгальдова (4), В. С. Черепанова (107; 108).

Педагогическая квалиметрия базируется на таких науках, как педагогика, психология, социология, математика и кибернетика. Сама педагогика использует из математики методы многомерного

статистического, факторного и корреляционного анализа. Использование этих методов вызывает необходимость дополнительной разработки разделов прикладной математики, а именно, теории систем, таксономии и др. В педагогической квалиметрии широко используются методы социологической квалиметрии. Интерес в этом плане представляют работы В. С. Аванесова (1), Э. П. Андреева, Г. В. Осипова (76), И. А. Филипповой (100) и др. В работах С. И. Архангельского (9, 10), В. П. Беспалько (14), Б. П. Битинаса (15), Н. В. Кузьминой (51), Е. К. Марченко (61), В. П. Мизинцева (66), В. И. Михеева (10; 67) и др. обосновывается использование кибернетических и математических методов в педагогике. При этом используются такие разделы кибернетики и математики, как теория управления, теория математического моделирования, теория алгоритмов и др.

Прежде чем перейти к конкретике квалиметрического подхода, остановимся на категориях «качество» и «количество». Их анализу посвящены многочисленные философские работы, в которых эти категории рассматриваются в различных аспектах, чаще всего в связи с законом перехода количественных изменений в качественные.

Существует достаточно много определений этих категорий.* Так, качество понимается и как функциональное единство существенных свойств предмета, и как множество свойств, органически связанных между собой законом взаимодействия, и как целостная характеристика вещи как системы с определённой структурой.

Под количеством понимается и мощность какого-либо множества, и качество, рассматриваемое в его пространственно-

* Подробно о методологическом значении категорий «качество» и «количество» изложено в монографии И. С. Тимофеева (95).

временном аспекте, и отношения частей к целому или части, и элементы целого.

Таким образом, мы видим, что и «качество» и «количество» отражает объективные стороны и отношения предметов, но содержание их разное. Разумеется, любое качество выражается в специфической для него системе количественных характеристик. Качество и количество выступают как нечто раздельное лишь в абстракции, реально они пребывают в нерасторжимом единстве, образующем их *меру*. Где *мера* — это зона, в пределах которой данное качество модифицируется, варьируется в силу изменения количества и отдельных несущественных свойств, сохраняя при этом свои существенные характеристики (91, с. 237).

Анализируя категории «качество» и «количество», мы вынуждены касаться и таких понятий, как «число», «величина», «оценка», которые являются фундаментальными в теории измерения, которая, в свою очередь, является основой квалиметрического подхода. Масштаб затруднений виден из того, что и эти понятия трактуются по-разному. Мы в своём исследовании будем придерживаться следующей трактовки этих понятий, данной с позиции математики.

Величина — «1) обобщение конкретных понятий: длины, площади, веса и т.п. Выбрав одну из величин данного рода за единицу измерения, можно выразить числом отклонение любой другой величины того же рода к единице измерения; 2) в более общем смысле скалярной величиной, или скаляром называется объект, полностью характеризующийся заданием одного числа. Обобщени-

ем скалярных величин являются векторные величины и тензорные величины» (89, с. 205).

Число — форма отражения количества.

Оценка в общем виде — это «суждения об определённом феномене» (77, с. 91), представленные в качественной или количественной форме.

Оцениванием в статистике называется «указание приближённого значения интересующего нас параметра (или функции от некоторых параметров) на основе наблюдаемых данных. Оценка — это правило вычисления приближённого значения параметра (или функции от некоторых параметров) по наблюдаемым данным» (98, с. 140).

Основными требованиями, предъявляемыми к оценкам, являются точность (близость к истинному значению), несмещённость (математическое ожидание оценки должно быть равно истинному значению параметра), состоятельность (при увеличении числа наблюдений оценка должна стремиться к истинному значению).

Главной проблемой количественного анализа в педагогических исследованиях являются процедуры измерения. Истоки использования таких процедур в отечественной науке мы находим в работах П. П. Блонского, Л. С. Выготского и др. В 20-е – 30-е годы дидактические исследования связывали с тестовыми и математико-статистическими методами. К сожалению, этот опыт не был закреплён, и к проблеме измерения вернулись лишь в начале 60-х годов. К этому времени изменилась ситуация, возникли новые разделы математики (многомерный статистический анализ, непараметрические статистики, теория информации, кибернетические разделы математики и др.), которые сделали возможным их использование в педагогических исследованиях.

В отечественных философских исследованиях процедуры измерения рассматриваются с той стороны, с которой они раскрываются как своеобразный переход от качества к количеству в познании. Причём процедура измерения предполагает: а) наличие знаний о качестве исследуемой предметной области; б) сущность процедуры измерения и основные абстракции, лежащие в её фундаменте, определены переходом от качества к количеству в познании; в) основные гносеологические свойства процедуры измерения и её результаты также определены переходом от качества к количеству в познании.

«Измерение в педагогике — познавательный процесс, состоящий в том, что на основании ранее полученной числовой системы (или системы классов), изоморфной эмпирической системе с отношениями, экспериментально определяют числовые значения величин, характеризующих некоторые признаки педагогических объектов или явлений, или указывают на класс, к которому они относятся» (84, с. 15).*

Н. М. Розенберг предлагает следующую упрощённую схему первоначального представления о процессах измерения и шкалирования: выявление и качественное описание предмета измерения — характеристик, признаков педагогических объектов и явлений → составление измерителей — тестов, контрольных заданий, критериев → шкалирование — конструирование шкал, которым соответствуют измерители → педагогический эксперимент или наблюдение, определение первичных, «сырых» данных измерения → математико-статистическая обработка первичных данных →

* Гегель употреблял понятие «измерения» в смысле «количественного сравнения (сравнения с эталоном)» (21, с. 182).

содержательная интерпретация результатов измерения и обработки (84, с. 16).

Рассматривая эту проблему с позиций моделирования, В. И. Михеев выделяет два основных направления, связанных между собой единой логикой исследования: 1) построение модели эмпирического объекта путём выделения его свойств и описания их с учётом психических процессов и свойств личности; 2) разработка методов, средств и процедур измерений свойств эмпирического объекта и приложения их в практике решения различных педагогических задач обучения (67, с. 15).

Очевидно, что если первое направление находит отражение во многих психолого-педагогических исследованиях, то второе только определяется в своём конструктивном решении.

Зададимся вопросом: что должно составлять основу квалиметрического подхода и что способно сделать его концептуально единым и непротиворечивым? Очевидно, что такой основой является система принципов и требований.

Попытка выделить принципы, лежащие в основе квалиметрического подхода, была предпринята Н. В. Акинфиевой (6) на основе работ Г. Г. Азгальдова, Э. П. Райхмана и др. К числу принципов автор отнёс: 1) рассмотрение исследуемого качества как иерархическую совокупность свойств на уровневой основе; 2) выделение простых свойств и соотнесение им специфических шкал измерения; 3) присвоение каждому свойству двух параметров (характеристик) — весомость и значимость; 4) весомость всех свойств одного уровня считается величиной постоянной; 5) весомость свойства определяется как среднее арифметическое оценок весомости, полученных отдельными экспертами; 6) в итоге определяется комплексная оценка качества в целом.

Как нетрудно убедиться, предложенные принципы не являются таковыми в строгом смысле этого слова. Скорее — это последовательность действий, отражающих процедуры измерения. Кроме того, они охватывают лишь одну предметную область — область теории измерения — в то время как квалиметрический подход, как мы уже отмечали, базируется ещё и на концептуальных положениях и методах педагогики, теории моделирования и математической статистики.

Мы полагаем, что систему принципов, лежащую в основе квалиметрического подхода, целесообразно строить, во-первых, на основе анализа методов исследования гуманитарных наук, во-вторых, на основе эвристических возможностей математики, логики, кибернетики, в-третьих, на основе теоретико-методологических положений педагогической квалиметрии, в-четвёртых, на основе идей информационной технологии.*

В результате мы пришли к системе принципов, включающей информативность, интегративность, оптимальность, точность, доказательность, технологичность, распределённость, унифицированность, доступность.

Информативность означает всесторонность представлений (качественная и количественная определённость) об исследуемом объекте (явлении, процессе).

Интегративность требует: а) межнаучных связей педагогики с философией, социологией, экономикой, психологией, математи-

* Образование с полным правом может быть рассмотрено в качестве одной из важных информационных систем общества. Информационную технологию определяют и как систему, транслирующую (трансформирующую) определённую информацию (92), и как совокупность и описание процессов циркуляции и переработки информации (30).

кой, кибернетикой и другими науками; б) междисциплинарных связей (общих основ педагогики с дидактикой, методикой, управлением и др.); в) внутридисциплинарных связей, предусматривающих синтез результатов исследования в рамках одной дисциплины; г) взаимной связи, приспособления и взаимодействия методов исследования, сочетающих количественные и качественные методы.

Оптимальность выражается в минимизации времени, усилий и средств при планировании, организации и проведении педагогического исследования; в оптимальном подборе методик исследования, сочетающих количественные и качественные методы.

Точность означает степень истинного соответствия полученных знаний изучаемому объекту (явлению, процессу).^{*} Точность научных знаний складывается из нескольких компонентов: точности метрической, логико-математической и точности семантической.[†]

Доказательность проявляется: в обоснованной не только с педагогической, но и с математической точки зрения логики оценки педагогического исследования; в комплексной методике исследова-

^{*} В. И. Журавлёв отмечает, что современное представление о продуктах научного труда как о специфической ценности, как о «товаре» принуждает вводить понятие *качество* знаний как результат добытой исследователем новой информации (36, с. 124).

[†] Метрическая точность достигается там, где есть возможность количественного выражения свойств изучаемого объекта. Логико-математическая точность устанавливается при оперировании процедурами умственной деятельности, т.е. высказываниями, понятиями, схемами, моделями, символами и т.д. Она устанавливается благодаря сопоставлению абстрактно-логических объектов, их связей и отношений. Семантическая точность устанавливается благодаря операциям с научными понятиями, причём они должны быть однозначными. Результатом метрической, логико-математической и семантической точности является интегральный результат — *научная истина*.

дования, базирующемся на гармонизации качественных и количественных методов исследования: в надёжности и валидности используемых измерителей; в репрезентативности полученных экспериментальных данных; в проверке результатов исследования на различных этапах экспериментальной работы, количественным и качественным их анализом; обработкой результатов эксперимента методами математической статистики с использованием вычислительной техники.

Технологичность, с одной стороны, означает необходимость проведения содержательного описания исследуемого объекта, его формализации и выражения на языке математической логики, с другой стороны, разделения исследуемого процесса на систему последовательных взаимосвязанных процедур и операций, которые выполняются более или менее однозначно, и имеют целью достижение высокой эффективности.

Распределённость предусматривает применение широкого использования методов вычислительной техники и обеспечения переработки и хранения данных исследования.

Унифицированность предусматривает рациональное сокращение номенклатуры средств и приведения её к единой структуре, расширяет масштабы внедрения результатов педагогического исследования при относительном сокращении затрат на его проведение.

Доступность предполагает использование методик педагогом-экспериментатором без специальной математической подготовки.*

* Тем не менее, бóльший эффект достигается при творческом содружестве педагога-исследователя и математика-консультанта. Подробно об этом изложено в работе Ю. П. Адлера (3).

Требования, которым должен удовлетворять квалиметрический подход:

1. Установить критерии оценки возможных методов проведения измерений, исходя из поставленной цели педагогического эксперимента.
2. Обеспечить выбор методов исследования и преобразования их в конкретные исследовательские методики, адекватные исследовательским задачам.
3. Обеспечить предварительную оценку точности, достоверности и надёжности измерения, исходя из цели, задач и формы проведения педагогического эксперимента.
4. Обеспечить такое сочетание методов исследования, которое позволит получить разносторонние сведения об объекте и «отследить» динамику изменения объекта.
5. Обеспечить подбор методов исследования, позволяющих анализировать не только ход исследовательского процесса, его результаты, но и осуществить прогноз.
6. Выразить измерения в числовой форме или в условных показателях.

Таким образом, квалиметрический подход позволяет по-новому подойти к осмыслению проблем из области методологии и методики педагогических исследований; решить ряд актуальных задач, связанных с теорией и практикой эксперимента; избежать некоторого формализма в использовании методов математики, кибернетики, логики.

§ 3. Теория измерений как основа квалиметрического подхода

Мы ранее отмечали, что квалиметрический подход базируется на концептуальных положениях и методах педагогики, теории измерения, теории моделирования и математической статистики. Однако, прежде чем использовать эти положения и методы, во-первых, необходимо провести содержательный анализ объекта изучения, во-вторых, анализ содержания объекта изучения необходимо рассматривать в связи с конкретной педагогической областью описания экспериментального объекта, в-третьих, методы качественного описания объекта изучения должны дополняться методами количественными.

Введём некоторые понятия теории измерения. Поскольку значительная часть работ по теории измерения посвящена шкалированию, остановимся, прежде всего, на типах шкал и их характеристиках с учётом особенностей педагогического исследования.

Прежде всего, дадим общее определение шкалы.

Пусть \mathbf{A} — эмпирическая система с отношениями, \mathbf{R} — полная числовая система с отношениями, \mathbf{f} — функция, изоморфно отображающая \mathbf{A} в подсистему \mathbf{R} , \mathbf{G} — группа допустимых преобразований на \mathbf{R} . Тогда упорядоченный набор $\langle \mathbf{A}, \mathbf{R}, \mathbf{G} \rangle$ называется **шкалой**, а набор $\langle \mathbf{A}, \mathbf{R}, \mathbf{G}, \mathbf{f} \rangle$ — **реализацией шкалы**.

Зафиксируем следующие группы преобразований на множестве действительных чисел \mathbf{R} :

\mathbf{G}_1 — группа всевозможных взаимно однозначных отображений;

\mathbf{G}_2 — группа отображений, сохраняющих отношение порядка;

\mathbf{G}_3 — группа линейных отображений вида $\varphi(x) = \alpha x + \beta$, где $\alpha > 0$;

\mathbf{G}_4 — группа растяжений вида $\varphi(x) = \alpha x$, где $\alpha > 0$;

G_5 — группа сдвигов вида $\varphi(x) = x + \beta$.

Легко увидеть, что введённые группы образуют следующую цепочку включений:

$$G_1 > G_2 > G_3 > G_4 \text{ и } G_3 > G_5.$$

Из теории измерений известна следующая основная теорема классификации шкал:

Теорема. 1) Имеется пять типов шкал измерений признаков: шкала наименований S_1 , шкала порядка S_2 , шкала интервалов S_3 , шкала отношений S_4 и шкала разностей S_5 ; 2) каждая шкала S_i имеет в качестве группы допустимых преобразований группу G_i и полностью ею определяется; 3) группы G_i частично упорядочены по включению, причём, если $G_i > G_j$, то шкала S_j реализует более точный способ измерений, чем шкала S_i .

Рассмотрим подробнее перечисленные выше шкалы.

Шкала наименований — это любое взаимно однозначное отображение изучаемой эмпирической системы во множество чисел.

Данная шкала позволяет лишь различать объекты. Например, эмпирическому признаку пол ученика можно сопоставить одно из чисел 0 или 1. Скажем, девочка — 0, мальчик — 1. Поскольку группа допустимых преобразований G_1 состоит из любых взаимно однозначных отображений чисел друг в друга, то можно было бы выбрать другие два различных числа для нашей шкалы, например, девочка — 100, мальчик — 50. Числа, сопоставленные полу ученика, не несут в себе никакого смысла присущего собственно числам. Их нельзя, скажем, складывать или сравнивать. Так для второй шкалы $50 + 50 = 100$ не означает, что два мальчика равны одной девочке. А из сравнения $100 > 50$, не следует, что девочка (100) лучше, чем мальчик (50).

Шкала наименований не предоставляет никакой количественной информации, она лишь даёт объектам имена, отсюда и название — шкала наименований.

Таким образом, при использовании шкалы наименований объекты измерения распадаются на непересекающиеся классы, охватывающие все изучаемые объекты. Каждому классу даётся наименование, числовое обозначение которого является одним из шкальных значений. Один класс — это совокупность объектов, имеющих одно и то же шкальное значение. Другими словами, при осуществлении измерения моделируются только отношения равенства и неравенства.

В более общей трактовке можно использовать для шкалы наименований любые объекты, а не только числа, скажем, мальчики — «+», девочки — «-», или мальчики — «А», девочки — «Б», или мальчики — ♂, девочки — ♀. Более того, для практических целей бывает удобнее использовать именно буквенные или знаковые обозначения, чтобы не возникало ненужных ассоциаций и соблазнов делать поспешные выводы (два мальчика равны одной девочке, девочки лучше мальчиков и т.д.). Отметим попутно, что шкала наименований, состоящая из двух значений, активно используется в различного рода исследованиях под названием дихотомическая шкала.

Однако, несмотря на кажущуюся простоту шкалы наименований, она при грамотном использовании позволяет получить важные данные. Мы ещё вернёмся к статистической обработке таких шкал в главе III.

Шкала порядка — это любое отображение упорядоченной эмпирической системы в числовую систему с отношениями, сохраняющее порядок. Примером такой шкалы служит хорошо извест-

ная шкала балльных оценок: 2, 3, 4, 5. Выбранные числа сохраняют в себе упорядоченность всей числовой системы, т.е. их всегда можно сравнить, например, $2 < 3$, $3 < 5$, $3 = 3$. Такая шкала не только позволяет различать объекты, но и сравнивать их. Например, 2 и 3: ученик, оцененный двойкой, знает или умеет меньше, чем ученик, оцененный тройкой. Но такая шкала также не позволяет выполнять арифметические операции. Хотя $2 + 2 = 4$, две «двойки» не равны одной «четвёрке», так же как сумма знаний «двоечника» и «троечника» не равна знаниям одного «отличника».

Группа допустимых преобразований данной шкалы состоит из преобразований, сохраняющих порядок. Следовательно, мы могли бы вместо шкалы 2, 3, 4, 5 взять шкалу 1, 15, 16, 40. Она также позволяет сравнивать объекты. Заметим здесь, что «расстояние» между оценками не несёт в данном случае никакой смысловой нагрузки. Нельзя сказать, что ученик, получивший 16 баллов, знает на 1 балл больше ученика, получившего 15 баллов, а ученик, имеющий 15 баллов, знает на 14 баллов больше ученика, имеющего 1 балл. Шкала порядка указывает только на порядок расположения объектов, но ничего не говорит о расстояниях между ними.

Таким образом, порядковая шкала моделирует не только отношения равенства и неравенства, но и отношения порядка между ними.

Если опять вернуться к более общей ситуации, забыв про числа, то в качестве примера порядковой шкалы можно взять шкалу «неудовлетворительно», «удовлетворительно», «хорошо», «отлично», или шкалу «высший сорт», «первый сорт», «второй сорт», «третий сорт», или же шкалу «верно», «неверно». Важно, чтобы была возможность упорядочить эти оценки также как и числа.

Порядковые шкалы несут в себе, очевидно, гораздо больше информации, чем шкалы наименований, и в то же время не требуют от исследователя точности количественных оценок. Оценивая качество объектов, исследователь расставляет их в некотором порядке, не обращая внимания на величину различий между ними.

Шкала интервалов — это шкала порядка, для которой задана операция, позволяющая сравнивать интервалы между делениями на шкале. Типичным примером является температурная шкала Цельсия. Расстояния между точками на шкале всегда можно сравнить между собой. Используя допустимые преобразования, можно получить другие температурные шкалы. Например, если **F** и **C** — температуры по Цельсию и Фаренгейту соответственно, то $F = 1,8C + 32$.

Отметим, что допустимое в данной шкале преобразование $\varphi(x) = \alpha x + \beta$ действует следующим образом: оно растягивает шкалу в α раз и сдвигает её на β единиц. Отсюда легко следует, что отношения длин одних и тех же интервалов в разных реализациях одной шкалы равны. Основной характерной чертой шкалы интервалов является наличие масштабной единицы, что позволяет выяснить не только, в каком из сопоставляемых объектов признак выражен сильнее, но и насколько сильнее выражен. Как мы видим, основную нагрузку в данном случае несут на себе величины интервалов, а числовые отметки можно задавать произвольным образом (0° по Цельсию равны 32° по Фаренгейту и т.д.). Отсюда и название шкалы.

Примером шкалы интервалов в педагогических исследованиях может служить процентное выражение каких-либо данных, скажем, качественная успеваемость — процент учащихся, не имеющих «двоек» и «троек» и т.д.

Таким образом, интервальные шкалы помимо отношений равенства и порядка изучаемых объектов моделируют отношения равенства «интервалов» между объектами, имея фиксированную масштабную единицу.

Понятно, что любую порядковую шкалу можно превратить в интервальную, описав правило сравнения интервалов. Но именно в обосновании равенства и разности интервалов между объектами и состоит основная трудность.

Шкала отношений — это шкала интервалов, в которой естественным образом задана точка отсчёта. Например, шкала измерений роста или веса. Допустимыми преобразованиями этой шкалы являются только растяжения $\varphi(x) = \alpha x$, где $\alpha > 0$. Например, вес можно мерить в сантиметрах, метрах и т.д. Основным требованием для шкалы отношений является отображение в ноль одного и того же объекта. Отсюда следует, что в различных реализациях одной шкалы сохраняются отношения не только между интервалами, но и между самими шкальными значениями. А это означает, что можно ответить на вопрос, во сколько раз сильнее выражен признак в одном из сравниваемых объектов.

В виду достаточно жестких требований к системе шкала интервалов редко используется в педагогических исследованиях. В качестве примера можно привести разве что затраченное время как меру трудности учебного задания.

Переход от шкалы интервалов к шкале отношений сопряжён с большими трудностями при незначительном увеличении числа способов статистической обработки, поэтому с точки зрения педагогических исследований такой переход малоэффективен.

Шкала разностей — это шкала интервалов с жёстко закреплённым масштабом. Допустимые преобразования для этой шкалы

— только сдвиги $\varphi(x) = x + \beta$. В качестве примера такой шкалы можно привести возраст. В силу достаточно серьёзных ограничений на систему шкала разностей практически не получила распространения в педагогических исследованиях. Она используется, пожалуй, лишь в методе парных сравнений.

Отметим, в заключение, что по традиции шкалы наименования и порядка относят к качественным шкалам, а шкалы интервалов, отношений и разностей — к количественным. Такое деление естественным образом вытекает из свойств перечисленных шкал. С позиций статистической обработки данных чаще выделяют три группы шкал: номинальные, порядковые и количественные. Это объясняется тем, что разработанные статистические методы недостаточно точны для выделения особенностей, например, шкалы интервалов по отношению к шкале разностей.

Для большей наглядности и удобства дальнейшего использования сведём всё вышеизложенное в таблицу.

Таблица 1

Типы шкал и их основные характеристики

Шкала	Свойства	Группа допустимых преобразований	Примеры	Статистический аппарат
Шкала наименований S_1	Различает предметы по наличию свойства. Не различает уровней проявления свойства	G_1 — группа всевозможных взаимно однозначных отображений	Порядковый номер, пол, результат сдачи зачёта	Частота Мода
Шкала порядка S_2	Различает уровень проявления	G_2 — группа отображений, сохра-	Школьные оценки, год обучения,	Частота Мода Медиана

	свойств объекта. Не определяет величину различия в проявлении свойств	няющих отношение порядка	стаж	Коэффициент Кендэлла Размах Коэффициент Спирмена
Шкала интервалов S_3	Определяет величину различия проявления свойства. Имеет масштабную единицу. Сравнивает, на сколько больше проявляется свойство. Не определяет уровень исчезновения свойства (естественный ноль шкалы)	G_3 — группа линейных отображений вида $\varphi(x) = ax + \beta$, где $a > 0$	Температура, рейтинг, семестры	Частота Относительная частота Квантили Мода Медиана Среднее Дисперсия Коэффициент корреляции
Шкала отношений S_4	Определяет любые отношения между уровнями проявления свойств. Имеет масштабную единицу и фиксированное начало шкалы	G_4 — группа растяжений вида $\varphi(x) = ax$, где $a > 0$	Масса, длина, скорость выполнения задания, процент учащихся, объём часов по предметам	Частота Относительная частота Квантили. Мода Медиана Среднее Дисперсия Коэффициент корреляции
Шкала разностей S_5	Определяет накопление свойства.	G_5 — группа сдвигов вида $\varphi(x) = x + \beta$	Номера учебных недель,	Частота Относительная частота

	Имеет масштабную единицу. Не имеет фиксированного начала шкалы		хронометрия	Квантили Мода Медиана Среднее Дисперсия Коэффициент корреляции
--	---	--	-------------	---

При измерении качественных показателей, скажем, успеваемости, исследователь, естественно использует качественные шкалы, а именно шкалу наименований или шкалу порядка. Пользуясь шкалой наименований, можно лишь отметить различия, но невозможно сравнить проявление описываемых качеств. Следовательно, с точки зрения дальнейшей обработки информации предпочтительнее шкала порядка. Использование количественных шкал для первоначальных измерений в педагогике физически невозможно в силу объективных причин. Не существует прибора, который позволил бы измерить качество знаний, также как мы можем измерить вес или температуру. Следовательно, проблема измерения и оценки педагогического объекта исследования состоит в количественной оценке информации, заключённой в качественных первичных оценках.

Для использования более точных шкал требуется больше информации, а, следовательно, лучшее понимание качественных закономерностей свойств исследуемой системы. Попытки построить сразу интервальную шкалу или шкалу отношений при недостатке информации приводит к использованию субъективных, а зачастую и неверных данных. Хотя как было отмечено выше, это не даёт значительных преимуществ для дальнейшей статистической обработки.

Как мы уже выяснили, измерение каких-либо показателей представляет собой выбор шкалы, которая отражала бы имеющиеся в исследуемой системе отношения. Этот выбор всегда ложится на исследователя. Правильнее было бы говорить не о выборе, а о построении шкалы, поскольку недостаточно решить использовать, к примеру, пятибалльную шкалу. Самое трудное — определиться, за что выставлять конкретную оценку.

Михеев (67, с. 76) приводит следующий общий алгоритм построения *качественной* шкалы.

Этап 1 — Формирование экспертной группы педагогов.

Этап 2 — Подбор тестовых заданий с предварительным их анализом.

Этап 3 — Формирование нормативной выборки учащихся.

Этап 4 — Проведение первичных измерений.

Этап 5 — Проведение производных измерений.

Этап 6 — Корректировка функции ранга, матрицы парных сравнений и экспертной группы.

Этап 7 — Построение классов эквивалентности.

Этап 8 — Формулирование правила предписания.

Этап 9 — Проверка правила предписания на устойчивость.

Этап 10 — Выработка конечных рекомендаций.

Построение интервальной шкалы требует выполнения дополнительно следующих действий:

Этап 11 — Определение величины классового интервала.

Этап 12 — Последовательное вычисление нижних границ классов.

Этап 13 — Разграничение классовых интервалов на величину, равную точности оценки значений изучаемого признака.

Этап 14 — Определение середины для каждого классового интервала и соответствующих для них частот.

Построенная шкала должна быть надёжной и валидной, т.е. точно измерять именно то, для чего она предназначена. Разберём подробнее эти две характеристики.

Прежде всего, разберёмся с понятием *надёжности*.

Целью любого измерения является установление истинного значения измеряемой величины, т.е. существующей в действительности неискажённой величины признака, присущего данному индивиду. Этот изучаемый признак проявляется довольно устойчиво в тестах, подготовленных для его измерения.

Однако в философии хорошо известен постулат о неизбежности погрешности измерения. Как отмечает Л. Яноши, «результат эксперимента всегда содержит ошибку, как бы тщательно не проводились измерения» (117, с. 15). Принятие этого постулата неизбежно приводит к одному из основных положений теории измерения — к тезису о невозможности знания абсолютного значения измеряемой величины. Следовательно, измерениям подвергается только наблюдаемый результат измерения, искажённый под влиянием различных факторов. Можно выделить множество факторов, искажающих истинное значение измеряемой величины, например, влияние цели, задач и характера исследования, условия тестового опроса и т.д.

Итак, всякое измерение всегда содержит ошибку. Значит, задача исследователя — понизить эту ошибку до приемлемого уровня.

Но точность измерения составляет только одну сторону понятия надёжности. Другой важной характеристикой надёжности является возможность повторить измерения с тем же результатом.

Перейдём теперь к более обширному понятию *валидности*. Б. П. Битинас (15) выделяет содержательную, конструктивную, конкурентную, прогностическую и перекрёстную валидность.

Содержательная валидность методики измерений определяется тем, что её составляющие признаки охватывают те области информации и деятельности, в которых проявляется измеряемое свойство. Содержательная валидность не оценивается количественно и достигается путём составления качественной схемы сбора необходимой информации.

Конструкционная валидность указывает на то, в какой степени диагностическая методика отражает то составное свойство, для определения которого оно предназначено, и определяется, во-первых, теоретическим содержанием признаков, во-вторых, их внутренней согласованностью. Конструкционная валидность требует, возможно, более полного накопления информации об измеряемом свойстве.

Конкурентная валидность определяется с помощью внешнего критерия, по которому информация собирается одновременно с проверяемой методикой. Она показывает возможность замены одной диагностической методики другой методикой.

Прогностическая валидность определяется также по внешнему критерию, но информация по этому критерию собирается некоторое время спустя после применения проверяемой методики.

Перекрёстная валидность выявляется сопоставлением результатов измерения, проверенных на некоторой выборке, с результатами измерения в новой выборке испытуемых из этой же генеральной совокупности.

Приведём в заключение пример из монографии В. И. Михеева (67, с. 20). «Соотношение между надёжностью и валидностью

лучше представить аналогией с часами: часы могут иметь точный (надёжный) ход, но, будучи поставленными на неверный час, они непригодны (невалидны) для получения ответа на главный вопрос — который час?».

Отметим ещё две характеристики шкалы, важных для дальнейшей интерпретации результатов измерения. Во-первых, это *однородность шкалы*, которая означает, что все признаки отражают один общий фактор — измеряемое свойство. И, во-вторых, — *одномерность шкалы*, которая означает, что она измеряет только одну переменную. Может оказаться, что все признаки включают, кроме измеряемого, и некоторое другое свойство. Такая шкала уже не будет одномерной.

Подводя итог вышесказанному, ещё раз подчеркнём, что теория измерений занимается вопросами численного описания характеристик объекта. Её привлечение в целях педагогического исследования позволяет решить одну из самых сложных проблем — проблему количественного описания качественных показателей. Корректное измерение уровней проявления качественных показателей позволяет привлекать мощные методы статистического анализа. Всё это выводит педагога-исследователя на новый уровень доказательной строгости. Появляется возможность от нечётких и часто субъективных оценок перейти к математически обоснованным выводам.

Выводы по первой главе

1. Педагогическое исследование в общем виде представляет собой специально организованный процесс познания, в котором происходит выработка и теоретическая систематизация знаний о педагогической действительности. Общенаучные требования, предъявляемые к педагогическому исследованию, являются следующие: объективность, точность, воспроизводимость, доказательность. Особенности педагогического исследования являются его гуманистическая направленность, а также гармонизация теоретической и экспериментальной составляющих.
2. Методология — учение о методах научного познания и преобразования мира. Выделяют четыре уровня методологии: философская методология, общенаучная (научоведческая), конкретно-научная, методика и техника научного исследования

Метод научного исследования — категория, объединяющая и формы научного мышления, и общие модели исследовательских процедур, и способы (приёмы) выполнения исследовательских действий.

Методика педагогического исследования — определённая совокупность приёмов и операций, направленных на изучение педагогических явлений и решение разнообразных научных проблем учебно-воспитательного характера.

3. Значительные эвристические возможности содержатся в квалиметрическом подходе, под которым в общем смысле понимается теоретико-методологическая стратегия исследования, ориентированная на количественное описание качества объектов.

Он содержит иной теоретико-методологический ключ к пониманию проблемы организации и оценки результатов качества педагогического эксперимента: процесс организации и оценивания педагогического эксперимента должен быть, с одной стороны, доказательным, а с другой — технологичным.

В основе квалиметрического подхода заложена система принципов (информативность, интегративность, оптимальность, точность, доказательность, технологичность, распределённость, унифицированность, доступность) и требований теоретико-методологического и практико-методического характера, вытекающих из этих принципов.

4. Важной теоретико-методологической основой квалиметрического подхода является теория измерений, значительная часть которой представлена шкалированием. Существуют пять типов шкал измерений: шкала наименований, шкала порядка, шкала интервалов, шкала отношений и шкала разностей. Каждая из них имеет свои ограничения по применению к реальным объектам и по дальнейшей статистической обработке полученных данных.
5. Основными характеристиками шкалы измерений являются её надёжность и валидность. Надёжность включает в себя величину неизбежной ошибки измерения и возможность повторения измерения с тем же результатом. Под валидностью в наиболее широком смысле понимается пригодность данной шкалы для измерения конкретного признака. Валидность подразделяется на содержательную, конструкционную, конкурентную, прогнозную и перекрёстную.

Глава II. Технология планирования и организации педагогического эксперимента с позиций квалиметрического подхода

§1. Технология моделирования в педагогическом исследовании

Логически продуманная постановка любого исследования включает в себя два важных вопроса: выбор модели, описывающей исследуемый объект, и выбор плана эксперимента, оптимального для данной модели.

Решение первой задачи связано с глубоким знанием объекта исследования. Необходимо уметь выделить те факторы, которыми исследователь может реально управлять, и определить предполагаемые взаимоотношения между ними. Решение же второй задачи совершенно не зависит от реально изучаемого объекта и относится только к его математической модели.

Обсудим подробнее поставленные выше задачи и начнём с изучения технологии моделирования.

Обратимся, прежде всего, к общему понятию «технология». До недавнего времени этим понятием пользовались в узком смысле, ограниченном преимущественно сферой материального производства. Это объясняется, прежде всего, тем, что технология возникла и начала развиваться в основном в этой сфере. В последние годы стала активно развиваться социальная технология. М. Марков пишет: «Не только производство, но и любой вид социальной деятельности имеет свою внутреннюю логику развития и функционирования. Технология как научная дисциплина не только раскрывает суть этой логики и описывает её, но и указывает пути её имманент-

ных принципов и императивные требования для социального субъекта» (60, с.40).

Н. Стефанов в связи с этим отмечает: «Какими бы сложными ни были социальные процессы, каждый из них поддаётся технологизации. Реальная проблема не в том, возможно ли в принципе технологизировать социальные процессы, а в том, как это сделать» (92, с. 183).

Доказательность суждений учёных убеждает нас в том, что можно говорить не только о технологии процессов, исследуемых в педагогике, но и об их моделировании.

Понятие «технология» многогранно, отсюда существует и достаточно много определений этого феномена. Тем не менее, условно их можно подразделить на два вида: один из них состоит из определений, данных в теоретическом аспекте, а другой — в практическом. Причём в определениях теоретического вида «отражены средства деятельности, их структурно-функциональные блоки (механизмы), взятые в активном деятельностном состоянии, предлагающем органическое сочетание объективизированных средств деятельности, способностей и умений» (59, с. 42).

В определениях, данных в практическом ключе, технология рассматривается как способ реализации людьми конкретного сложного процесса путём разделения его на систему последовательных взаимосвязанных процедур и операций, которые выполняются более или менее однозначно, и имеет целью достижение высокой эффективности.

В данной работе мы будем опираться на определения сущности понятия «технология», которые, в свою очередь, связаны с понятиями процедуры и операции. При этом под процедурой понимается совокупность действий (операций), посредством которых

осуществляется процесс (или его отдельный этап), выражающий суть конкретной технологии, а под операцией — непосредственное практическое решение задачи в рамках данной процедуры. «Под операцией мы подразумеваем способ выполнения действия», — пишет А. Н. Леонтьев (57, с. 517).

Отмеченные моменты принимались нами во внимание при разработке как технологии моделирования, рассматриваемой в данном параграфе, так и при организации и оценке экспериментальной части педагогического исследования, рассматриваемых в работе в целом.

Известно, что образование, будучи системой, относится к сложноорганизованным объектам.* Оно состоит из ряда подсистем, которые по отношению к более частным образованиям сами являются системами. При изучении сложноорганизованных объектов, рассматриваемых как система, возникает необходимость использования системного подхода. В философии системный подход характеризуется как общеметодологический принцип. Его гносеологический аспект базируется на идеях и основных положениях теории систем, а методический аспект (рецепты, приёмы, методы, технологии и т.п.) связывается с конкретными науками.

Инструментом использования системного подхода является системный анализ, который представляет собой совокупность методов и приёмов для изучения сложных объектов. В эту совокупность входят методы моделирования, декомпозиции, графические и ряд других. Выделяют четыре аспекта осуществления системного анализа, которые могут быть применены к педагогическим объек-

* В монографии Г. Н. Серикова (86) с позиций методологии системного подхода изложены теоретические представления о существенных аспектах образования как дидактически подвижной целостности различных компонентов.

там и явлениям: морфологический, структурный, функциональный и генетический. Ю. А. Конаржевский (47) отмечает, что все аспекты системного анализа выступают в органическом единстве, и это единство обусловлено единством свойств целостной системы. Наиболее полно это единство проявляется в моделировании. В. В. Давыдов и А. У. Варданян (29) определяют моделирование как метод познания интересующих нас качеств объекта через модель. А. И. Уёмов определяет модель как систему, исследование которой служит средством для получения информации о другой системе (99, с. 48). Несмотря на то, что модель многозначна, она сводится к условному образу (изображению, схеме, описанию) некоторого объекта или системы объектов. Моделирование включает в качестве важного компонента экспериментирование.

Широкое использование метода моделирования в педагогических исследованиях объясняется многообразием его функций, а, следовательно, и возможностей. Главное заключается в том, что модель является промежуточным звеном между субъектом — педагогом-исследователем и предметом исследования.

В соответствии с лидирующей функцией выделяют виды моделей — модель-заместитель, модель-представление, модель-интерпретация, исследовательская модель. Модель-заместитель оригинала используется тогда, когда она более удобна для определённых действий в конкретных условиях. Модель-представление, — когда требуется составить первоначальное представление об оригинале. Модель-интерпретация, — когда требуется истолкование оригинала. Исследовательская модель, — когда необходимо исследовать (изучить) свойства и закономерности оригинала.

Л. М. Фридман отмечает, что с психологической точки зрения *модельные отношения* должны рассматриваться иначе, чем в тео-

рии познания. Поскольку там отношение рассматривается как бинарное, в которое входят оригинал и его модель, а при рассмотрении моделирования в широком смысле модельное отношение будет тернарным или даже квартернионным, когда наряду с моделируемым объектом и его моделью рассматривается ещё и субъект, который конструирует или выбирает модель, и та цель, ради которой он это делает (102, с. 91). В этом проявляется некоторый субъективизм моделирования в педагогических исследованиях, который, на наш взгляд, можно «снять» или, по крайней мере, уменьшить за счёт математических методов.

Концепцию математического моделирования в педагогике теоретически обосновал И. Г. Кулль (52), который создал «предметную модель» на базе одного из разделов преподавания физики. Не давая строгого определения математической модели, исследователи отмечают, что математической моделью реального объекта называется такое его отображение, которое позволяет описать существенные стороны объекта языком математической логики и математических формул. Причём математическая модель, отображая объект исследования в идеальный образ, даёт новую, хотя возможно и не полную информацию об этом объекте (75, с. 35).

Таким образом, при построении математической модели исходными являются только те свойства объекта, которые могут быть описаны количественно, и только те связи между свойствами, которые поддаются описанию языком математики.

Свойства объекта, поддающиеся числовой оценке, называются его параметрами или характеристиками.

В зависимости от характера изменения параметров (характеристик) различают детерминированные и стохастические, непрерывные и дискретные, стационарные и динамические модели. Сто-

хастическая модель в отличие от динамической строится тогда, когда параметры (характеристики) объекта подвергаются воздействию случайных неконтролируемых факторов. В непрерывных моделях в отличие от дискретных область допустимых значений параметров (характеристик) непрерывна. И, наконец, модель называется стационарной в отличие от динамической, если её параметры (характеристики) меняются во времени.

Следовательно, математические модели в педагогике — это модели стохастические, дискретные и динамические.

К началу 80-х годов в дидактику были введены теоретико-множественные, функциональные, информационно-детерминированные, вероятностно-статистические, эвристические и другие модели. К настоящему времени наибольшее распространение получили следующие дидактические модели: модель периодичности контроля знаний учащихся (73), модель управления процессом усвоения информации (2), модель уровня психолого-дидактических возможностей учащихся (41), информационно-семантическая модель обучения (97), модель оптимизации самостоятельной работы студентов (22), модель эвристического поиска решений (53), интегративно-модульная модель подготовки будущего учителя (116) и другие.

В настоящее время популяризируются два подхода к разработке моделей: формально-аксиоматический (дедуктивный) и содержательно-логический (индуктивный). При формально-аксиоматическом подходе основу моделирования составляет набор аксиом, из которых формальным путём по определённым правилам вывода получаются все допустимые в данной системе математические конструкции. При содержательно-логическом подходе математиче-

ская модель разрабатывается как идеализированный образ уже имеющейся содержательной системы.

Таким образом, при формально-аксиоматическом подходе получается одна формальная система, которая допускает множество применений (интерпретаций), а при содержательно-логическом подходе один содержательный объект может быть представлен многими моделями (61, с. 9).

Отметим, однако, что при любом подходе выполняется общая процедура построения модели, включающая такие этапы, как: создание модели; вывод теоретических соотношений и аналитических представлений и зависимостей; оценка параметров модели; получение численных предсказаний; уточнение самой модели.

Построение математической модели, т.е. формальная постановка задачи на языке математики, является на сегодняшний день самым слабым звеном в теории математического эксперимента. Как отмечает Ю. П. Адлер, «всякая формализация, по существу, представляет собой насилие над задачей, втискивание её в прокрустово ложе некоторой априорной формальной схемы» (3, с. 7). Однако, другого пути на сегодняшний день нет. Для использования богатых возможностей, предоставляемых теорией вероятности и математической статистикой, мы должны формализовать нашу задачу. Разрабатывать свой математический аппарат для каждой конкретной задачи не представляется возможным в связи с несоизмеримостью затраченных сил и получаемой при этом разовой выгоды. Следовательно, мы вынуждены подбирать формальную схему, наиболее точно характеризующую изучаемый объект, и попытаться с минимальными потерями представить с её помощью нашу проблему.

Всё вышесказанное приводит нас к необходимости очень тщательного предварительного изучения объекта исследования, чтобы неизбежная неточность приближения не привела к неправильным выводам в результате некорректно выбранной модели.

Как правило, выделяют два типа моделирования — *содержательное* и *статистическое* моделирование. Статистическое моделирование также часто называют *моделированием типа «Монте-Карло»*. Основным различием этих методов является тип используемых для построения модели данных. С этой точки зрения выделяют *априорную информацию* о природе и характере исследуемых соотношений и совокупность *исходных статистических данных*, характеризующих процесс и результат функционирования анализируемой системы.

Если исследователь располагает информацией обоих типов, то, из априорной информации, предварительно математически формализованной, выводится общий вид аналитических уравнений, описывающих исследуемые соотношения, после чего с помощью обработки исходных статистических данных оцениваются численные значения неизвестных параметров этих уравнений. Такой приём носит название *содержательного* моделирования.

Если же исследователь располагает только априорной информацией, то он пытается симитировать поведение анализируемой реальной системы при варьировании численных значений параметров, входящих в аналитическую запись модели, чтобы получить дополнительные статистические данные. Такое моделирование чаще всего осуществляется с использованием ЭВМ и носит название *статистического* моделирования или *метода Монте-Карло*.

Обратимся теперь к самому процессу построения математической модели.

На первом этапе необходимо перейти от реального объекта к его схеме, называемой моделью «чёрного ящика». В основе такой модели лежит представление об объекте как о некоторой системе, функционирующей по неизвестному нам принципу. Не зная, как этот механизм работает, мы можем наблюдать результат его работы. Наблюдая за функционированием нашего объекта, мы замечаем, что он изменяет своё состояние под действием каких-то внешних причин. Наша задача попытаться выделить эти причины в виде отдельных воздействий, которые принято называть факторами. То есть *факторы* — это переменные воздействия на объект, влиянию которых приписывается переход объекта из одного состояния в другое. При этом переменные, характеризующие состояние объекта, называются *откликами*.

После выделения факторов и откликов заканчивается первый этап в построении модели. Следующий этап состоит в выборе уровней варьирования факторов, т.е., другими словами, в выборе шкалы измерения. Несмотря на кажущуюся простоту, этот этап очень важен и достаточно труден. Выбор определённой шкалы ограничивает круг статистических методов, применимых для дальнейшего анализа данных. Следовательно, слишком грубая шкала приведёт к невозможности использования более тонких критериев, что в свою очередь даст очень приблизительные (а, значит, и малоубедительные) результаты. С другой стороны, введение чрезмерно точных шкал может оказаться неоправданным из-за отсутствия способов проведения измерений с выбранной точностью. В результате приближённость измерений вместе с достаточно тонкими ста-

тистическими критериями может дать гораздо худшие результаты при больших затратах сил и средств.

После выбора шкалы измерения, т.е. сопоставления уровням проявления факторов некоторых числовых значений можно перейти к следующему этапу, который состоит в установлении соответствия между факторами и откликами. То есть в нахождении *функции отклика*, определяющей связь между значениями факторов и значениями откликов. Невозможно сразу же точно указать функцию отклика. Если бы это было так просто, то не возникало бы никаких проблем при изучении нашего объекта. Зная, как связаны определённые действия с изменением состояния объекта, мы знаем об этом объекте всё. Он не является для нас загадкой, он «прозрачен» перед нашим взором. В таком случае часто говорят о построении модели «белого ящика». К сожалению, а может быть и к счастью, реальные объекты намного сложнее любых моделей, которые мы можем для них построить, а, значит, задача построения функции отклика остаётся одной из наиболее сложных.

Откуда же в таком случае исследователь может найти требуемую связь? Только из внимательного изучения свойств объекта. Длительное и тщательное наблюдение за объектом исследования позволяет сделать некоторые предположения о законах функционирования этого «черного ящика» и выдвинуть некоторые предположения, гипотезы.

Выдвижением правдоподобной гипотезы и завершается второй этап построения модели. Далее в дело вступает эксперимент, направленный на проверку выдвинутой гипотезы. Однако этим дело может и не ограничиться. Где гарантия того, что наша гипотеза окажется верной? Данные эксперимента могут убедить нас в необоснованности предположений. И хотя отрицательный результат,

безусловно, тоже результат, приходится возвращаться к этапу выдвижения гипотезы с учётом полученных новых знаний. Недостаток информации об объекте или поспешность исследователя может превратить этот путь в заколдованный круг, но иного пути исследования не существует.

Изложенная выше схема построения модели носит несколько упрощённый характер. В теории математического моделирования процесс построения модели включает в себя шесть основных этапов:

1. Определение конечных целей моделирования, набора участвующих в модели факторов и показателей, их роли.
2. Предмодельный анализ физической сущности изучаемого явления, формирование и формализация априорной информации.
3. Собственно моделирование, т.е. вывод общего вида модели.
4. Статистический анализ модели, т.е. оценка неизвестных значений участвующих в описании модели параметров.
5. Верификация модели.
6. Уточнение модели, в частности возвращение ко второму этапу.

Подробно эти этапы с точки зрения математического моделирования изложены, например, в справочном издании по прикладной статистике под редакцией С. А. Айвазяна (5).

Особенности педагогического исследования накладывают свои ограничения на схему построения модели, что вынуждает прибегать к упрощённой схеме «чёрного ящика». Тем не менее,

привлечение методов математического моделирования приносит несомненную пользу педагогу-исследователю. Даже первый этап — выделение факторов, влияющих на характеристики объекта, и откликов на их изменение вносит ясность в дальнейшее исследование и позволяет избежать ошибок, связанных с непониманием структурных связей изучаемого явления. Применение методов математического моделирования открывает новые возможности для использования мощнейшего аппарата математики.

§ 2. Технология планирования и организации эксперимента

Следуя классической работе Ч. Хикса (104), мы выделяем три этапа в проведении исследовательской работы:

I. Формализация проблемы^{*}

- A. Постановка задачи.
- B. Выбор отклика.
- C. Выбор варьируемых факторов.
- D. Выбор уровней для этих факторов.
 - 1. Количественные или качественные.
 - 2. Фиксированные или случайные.
- E. Подбор сочетаний уровней факторов.

II. Планирование

- A. Необходимое число наблюдений.
- B. Порядок проведения эксперимента.
- C. Используемый метод рандомизации.[†]
- D. Математическая модель для описания эксперимента.

III. Анализ

- A. Сбор и обработка данных.
- B. Вычисление статистик для проверки гипотез.
- C. Интерпретация результатов.

^{*} У Хикса этот этап называется «Эксперимент», при этом сам термин эксперимент в данном случае понимается в самом узком смысле — как перевод проблемы на язык математики. Следует отметить, что в дальнейшем понятие эксперимент используется в привычном смысле.

[†] «Рандомизация — статистическая процедура, в которой решение принимается случайным образом» (63, с. 517). Применяется для усреднения влияния случайных факторов.

Первый этап этой работы подробно обсуждался в первом параграфе второй главы. Рассмотрение третьего этапа составляет содержание третьей главы. Проблемы, возникающие на втором этапе, мы рассмотрим в данном параграфе.

Как уже отмечалось, педагогический эксперимент в целом представляет собой комплекс методов исследования, предназначенный для объективной и доказательной проверки достоверности гипотезы (гипотез). В первом параграфе первой главы мы уже отмечали особенности педагогического исследования, которые проецируются на педагогический эксперимент. Кроме этого педагогический эксперимент отличается и рядом других, присущих ему особенностей. К их числу следует отнести создание специальных экспериментальных ситуаций для формирования заданного качества, в моделировании педагогического процесса как образца для педагогической практики, в повторении эксперимента в различных условиях, в апробации полученных данных в массовом опыте обучения и воспитания (94, с. 32).

Одной из важных проблем, требующих своего решения, является планирование эксперимента. В двадцатые годы XX века Фишер показал, что если эксперимент плохо спланирован и проведён, то его уже нельзя поправить никакими методами обработки данных. С другой стороны, рациональное планирование эксперимента даёт не менее существенный выигрыш в точности оценок, чем оптимальная обработка результатов измерений. В практическом отношении планирование эксперимента является мощным инструментом математической статистики. Планирование позволяет сократить число экспериментов, увеличить объём получаемой новой информации, уменьшить вероятность принятия ошибочных решений по результатам эксперимента.

Следует отметить, что вести речь о планировании эксперимента можно только в том случае, когда ещё до начала исследования можно сформулировать предварительные знания в виде математической модели. Именно запись изучаемой проблемы на языке математики позволяет полностью абстрагироваться от физического содержания задачи и рассматривать общие закономерности.

И ещё один важный момент. Само по себе планирование эксперимента не может улучшить качества построенной нами модели, оно лишь позволяет улучшить её статистические свойства, т.е. даёт возможность точнее оценить параметры имеющейся модели. Следовательно, даже самое тщательное планирование не принесёт желаемых результатов при неверно осуществлённой формализации задачи.

Ещё в конце 70-х – начале 80-х годов отмечалось, что теоретические основания планирования эксперимента и практическая реализация конкретных планов эксперимента шли как бы параллельным курсом. Это отставание теоретического осмысления метода от его практического применения учёные объясняют особенностями его исторического развития. Со временем появилась возможность осмыслить накопленный материал с единых позиций. В основу таких позиций положена идея о критерии оптимальности эксперимента.

Критерии оптимальности подразделяются на две группы — группа статистических критериев и группа динамических критериев. К статистическим критериям относятся высказывания, формулирующие требования, которым должно удовлетворять некоторое, задаваемое планом эксперимента расположение экспериментальных точек в пространстве факторов (независимых переменных), подлежащих варьированию. К динамическим критериям относят

высказывания, выраженные математическим языком, которые представляются правомерными на уровне интуитивных представлений (71, с. 7–8).

Предположим, что математическая модель исследуемого явления (объекта) выбрана правильно, т.е. в измерениях отсутствует систематическая ошибка. Это значит, что при выборе критерия оптимальности мы должны сосредоточить своё внимание на случайных ошибках. То есть нас интересует точность самих оценок. Для более наглядной интерпретации свойств полученных оценок часто прибегают к их геометрическому истолкованию, используя свойства *эллипсоида рассеяния*. В качестве параметров, определяющих эллипсоид рассеяния, используют математическое ожидание и дисперсию.*

Перечислим наиболее распространённые критерии оптимальности.

Первая группа критериев относится к статистическим критериям, связанным с точностью оценок параметров.

D–оптимальность направлена на уменьшение рассеяния оценок параметра. *D–оптимальному* плану соответствует эллипсоид рассеяния минимального объёма.

A–оптимальность направлена на уменьшение средней дисперсии оценок.

E–оптимальность направлена на ограничение дисперсий отдельных параметров. Таким планам соответствует эллипсоид по возможности наиболее близкий к шару.

* С некоторой долей неточности математическое ожидание можно определить как величину, определяющую среднее значение, оцениваемого параметра, а дисперсию — как величину, характеризующую разброс вокруг этого среднего значения. Подробнее эти понятия мы рассмотрим в §1 третьей главы (см. стр. 80).

Вторая группа критериев представляет собой статистические критерии, связанные с ошибкой в оценке модели.

G-оптимальность направлена на минимизацию максимального значения дисперсии оценки. В *G-оптимальных* планах нет точек, в которых точность оценки слишком низкая.

Q-оптимизация направлена на минимизацию средней дисперсии оценки.

Ротатабельность направлена на некоторое упорядочение плана эксперимента, в результате которого все направления становятся равнозначными с точки зрения точности оценок.

Униформность требует, чтобы дисперсия оценки была постоянной в некоторой области вокруг центра эксперимента.

К важнейшим характеристикам оптимальных планов относится следующая группа показателей, которые с некоторой натяжкой также можно отнести к критериям оптимальности. Эти характеристики в меньшей степени затрагивают специфические понятия математической статистики и более удобны в работе.

Рандомизация состоит в случайном порядке проведения эксперимента. Рандомизация производится для исключения влияния переменных, неконтролируемым образом изменяющихся во времени или пространстве. Этот критерий традиционно имеет большое значение в реальных экспериментах. Поэтому многие авторы строят теорию планирования эксперимента, отталкиваясь от идеи рандомизации (см., например, 104).

Насыщенность плана характеризуется количеством проведённых измерений. Если число измерений равно числу неизвестных параметров, которые нужно оценить, то план называется насыщенным. Очевидно, планы с меньшим числом измерений не позволяют найти единственные оценки всех параметров.

Композиционность отражает структуру проведения эксперимента. Композиционный план позволяет разделить эксперимент на несколько этапов и постепенно переходить от простых моделей к более сложным.

Простота обработки относится к косвенным характеристикам и отражает свойства математической модели. Чем она проще, тем меньше вычислений и меньше их сложность. При отсутствии специальных средств обработки числовой информации этот показатель часто становится одним из самых важных.

Мы не будем дальше продолжать список критериев оптимальности. Развитие теории планирования эксперимента происходит быстрыми темпами и, как следствие, появляются всё новые критерии, предназначенные для решения более частных проблем. Остановимся лишь на связи различных критериев оптимальности. Даже беглое изучение нашего списка показывает, что невозможно совместить различные критерии. План оптимальный с точки зрения одного критерия может оказаться неудовлетворительным с точки зрения другого. Построить планы, удовлетворяющие одновременно многим критериям оптимальности, удаётся только для отдельных моделей. Обычно приходится идти на компромисс, выбирая план близкий по оптимальности по нескольким критериям. Проблемы построения таких планов в общем случае довольно сложны и пока недостаточно разработаны.

Исследователя в первую очередь интересует не сама по себе теоретическая модель, а те выводы, которые позволяет сделать её изучение. Отсюда большой интерес не к абстрактным характеристикам планов, а к перспективам их применения. В этой связи на первое место выходят те планы, которые позволяют уменьшить вероятность ошибки при проверке гипотез на следующем этапе ра-

боты — на этапе анализа. Развивая эту идею, Дж. Кифер (119) ввёл понятие *процедуры* как совокупности плана эксперимента и теста для проверки гипотезы. Наиболее полно свойства процедуры раскрываются в задачах дисперсионного анализа.

Планирование педагогического эксперимента, имея свои особенности, в целом подчиняется общим принципам планирования.

Ю. К. Бабанский (13) отмечает, что эксперимент следует проводить с минимально необходимыми расходами времени, усилий и средств экспериментатора, т.е., другими словами, оптимально.

Отсюда возникают вопросы: «Каково оптимальное количество экспериментальных и контрольных объектов?», «Какова должна быть длительность проведения эксперимента?» и т.д.

Общие положения, на которые следует ориентироваться исследователю-практику при организации эксперимента, в определённой степени разработаны в конкретных приложениях. Так, для социологических исследований пределы выборочной совокупности колеблются от 350 до 3 тысяч единиц наблюдения (39, с. 99). В педагогических исследованиях, как считает Ю. К. Бабанский, если выделять типичные по определённым показателям контрольные и экспериментальные объекты, то можно свести их число до минимума (от 250 до 350). Такую процедуру можно осуществить, используя математический аппарат (например, методы построения репрезентативных выборок или различные способы рандомизации). Такие же процедуры можно осуществить при доказательстве минимально необходимой продолжительности эксперимента (от 2–3 месяцев до 2–3 лет).

Эти же вопросы, но несколько под другим углом рассмотрены В. Филлипсом (120). Всю совокупность вопросов, связанных с изучением особенностей педагогического эксперимента, он рассматривает как проблему пространства, времени и взаимодействия. Где проблема пространства — это проблема контроля, управления и измерения фактов, которые надо учитывать при изучении экспериментального объекта, проблема времени связана с продолжительностью эксперимента, проблема взаимодействия касается оценки действий экзаменатора и их влияние на результаты эксперимента.

В методологической литературе общий ход эксперимента определяется следующими процедурами:

- установление выборки, репрезентативной для данной популяции; подбор однородных групп или пар испытуемых;
- выбор конкретных методик, способов и параметров измерения экспериментальных данных;
- проверка доступности и эффективности методик на небольшом числе испытуемых;
- определение признаков, по которым можно судить об изменениях в экспериментальном объекте под влиянием соответствующих воздействий;
- определение времени и длительности проведения эксперимента;
- проведение эксперимента;
- качественный и количественный анализ результатов эксперимента;
- интерпретация результатов эксперимента;

- указание границ применения проверенной в ходе эксперимента системы мер.

Педагогический эксперимент как метод реализации научного замысла включает два основных этапа — *констатирующий* и *формирующий*. Эти этапы отличаются друг от друга задачами, содержанием, способами проведения. Так, если задачами констатирующего этапа является констатация состояния данного объекта, то задачей формирующего — создание новых условий (содержания, форм, методов) педагогических явлений. Констатирующий этап эксперимента проводится до формирующего этапа и после него.

На констатирующем этапе эксперимента, предшествующем формирующему, необходимо выявить: а) состояние объекта (качества личности, знания, умения и навыки, познавательный интерес, мотивы и т.п.); б) педагогические средства и условия, обусловившие исходное состояние; в) педагогические средства и условия, наличествующие к началу формирующего этапа эксперимента.

Формирующий этап эксперимента строится: а) на основе введения новой или модернизированной системы (технологии, ...), б) на основе создания в вариационном режиме различных педагогических условий, при которых эта система (технология, ...) эффективно функционирует и развивается.

Констатирующий этап эксперимента, следующий за формирующим этапом, имеет своей целью: а) констатацию состояния объекта; б) констатацию наличия происшедших изменений; в) качественный и количественный анализ происшедших изменений; г) интерпретацию результатов эксперимента.

В эксперименте существенное значение имеет выделение и определение переменных: зависимых, независимых и опосредующих. В этом случае опираются на двойное правило различия и со-

гласия (Дж. С. Милль): если за группой событий **А, Б** следует событие *a*, но за **Б, В** не следует *a*, то **А** есть причина *a*. Первая из двух групп событий рассматривается как экспериментальная, вторая — как контрольная. Единственное отличающее их событие называется *независимой* переменной, которая в ходе эксперимента может подвергаться различным изменениям. Изменение зависимой переменной рассматривается как следствие воздействий, оказываемых одной или несколькими независимыми переменными. Если исследуется зависимость скорости чтения от уровня развития интеллекта, то уровень развития интеллекта является независимой переменной, а время, затрачиваемое на осмысленное прочтение текста, — зависимой (пример приведён В. Оконею). Среди нескольких независимых переменных выделяют главные и второстепенные.

В ходе экспериментальной работы могут появляться *опосредующие* переменные, которые усиливают воздействие независимых переменных или ослабляют его воздействие. В этом случае говорят об искажающих переменных.*

Известно, что схема типичного педагогического эксперимента строится на сравнении экспериментальных и контрольных объектов (классов, групп, учебно-образовательных учреждений).

* Принимая во внимание выбор переменных, учёные под экспериментальным процессом понимают такую процедуру, которая позволяет: а) манипулировать, по крайней мере, одной главной независимой переменной; б) контролировать второстепенные независимые переменные; в) сводить к минимуму влияние искажающих независимых переменных на зависимую переменную; г) измерять изменения зависимой переменной, вызванные главными независимыми переменными.

Д. Кэмбелл, Дж. Стэнли (118) и др. считают эксперименты, не отвечающие этим условиям квазиэкспериментами. К квазиэкспериментам относят классический эксперимент Дж. Милля.

Таблица 2

Группа	Характеристика группы до эксперимента	Характеристика группы после эксперимента	Происшедшие изменения
Экспериментальная	x_1	x_2	$x_1 - x_2$
Контрольная	y_1	y_2	$y_1 - y_2$

Таким образом, результатом эксперимента являются изменения, которые произошли в экспериментальном объекте по сравнению с контрольным. А. Я. Найн (69) отмечает, что значение правильного выбора контрольных объектов явно недооценивается во многих диссертационных исследованиях, а между тем их наличие — главное условие чистоты эксперимента. Поэтому важно, чтобы контрольные объекты были типичными для отрасли и пользовались всеми преимуществами, которые имеют экспериментальные площадки.

Очень важным, на наш взгляд, является проведение статистических процедур, оценивающих сравнимость объектов. В том случае, когда экспериментальные и контрольные объекты несравнимы, необходимо применять процедуры уравнивания объектов. Эти процедуры (метод случайного отбора, предварительное обучение всех воспитуемых и т.д.) достаточно подробно изложены в ряде публикаций (13; 87 и др.).

Технологии доказательства сравнимости контрольных и экспериментальных объектов представлены в третьей главе.

Как показывает проведённый нами контент-анализ, в педагогических исследованиях в большинстве случаев выявляются, опре-

деляются и создаются педагогические условия, влияющие на исследуемый процесс.

В методологической литературе условия трактуют как философскую категорию, выражающую отношение предмета к окружающим его явлениям, без которых он существовать не может. Более того, условия составляют ту среду, обстановку, в которой явления возникают, существуют и развиваются.

Очевидно, что такая трактовка этой категории, верная в целом, требует педагогического осмысления. Дело в том, что рассмотрение категории «условие» как видового по отношению к понятию «среда» («обстановка») неоправданно расширяет совокупность объектов, необходимых для возникновения, осуществления или изменения, обуславливаемой педагогической конструкции.

В работе Н. М. Яковлевой (115) предлагается технология выявления комплекса педагогических условий, способствующих подготовке будущих специалистов к профессионально-творческой деятельности: фиксация социального заказа общества средней и высшей школе; выявление особенностей учебно-воспитательного процесса средней и высшей школы; моделирование комплекса педагогических условий, исходя из сущности педагогической деятельности как творческого процесса и системно-функциональной характеристики подготовки будущих специалистов к творческому решению воспитательных задач.

Предложенная технология широко используется в диссертационных исследованиях (46; 62; 72 и др.) с учётом специфики конкретного исследования.

В связи с вышеизложенным возникает задача доказательства необходимости и достаточности выявленных педагогических условий. В этом плане для нас является методологически важным

утверждение известного философа В. Н. Сагатовского: «В сознании субъекта, поднимающегося на уровень системной деятельности, происходит принципиальная переориентация: он стремится теперь не к тому, чтобы описать или сделать что-либо по принципу «чем больше, тем лучше», но переходит во власть совершенно иной парадигмы: «сделать то и только то, что необходимо и достаточно» (85, с. 56).

В математическом энциклопедическом словаре под необходимыми и достаточными условиями понимаются условия правильности утверждения, без выполнения которых утверждение заведомо не может быть верным (*необходимые условия*) и, соответственно, при выполнении которых утверждение заведомо верно (*достаточные условия*) (63, с. 403).

Из приведённого выше определения следует, что необходимые условия эффективного функционирования какой-либо системы — это условия, без которых система не может работать в полной мере, а достаточные условия — это условия, которых достаточно для нормальной работы системы.

Необходимость введённых условий вытекает из анализа психолого-педагогической литературы, опыта работы школы (вуза, ...), путей построения предлагаемой системы (технологии, ...) и результатов констатирующего этапа эксперимента. Внимательный анализ показывает невозможность получения желаемых результатов без обеспечения предлагаемых условий, а значит, их необходимость для эффективности системы.

Достаточность выводится из результатов опытно-экспериментальной работы. Данные срезов, проведённых в ходе эксперимента, отражают стабильный качественный рост формируемых знаний (умений, навыков, ...) и значительное превосходст-

во учащих (студентов, ...) той экспериментальной группы, в которой осуществлялись все условия, над учащимися (студентами, ...) остальных групп и в особенности над учащимися (студентами, ...) контрольной группы. Полученные результаты свидетельствуют об эффективном функционировании системы (технологии, ...) и, следовательно, позволяют сделать вывод о достаточности предложенных условий.

Формирующий этап эксперимента протекает в естественных условиях образовательного процесса высшей или средней школы по типу вариативного, для которого характерно целенаправленное варьирование в различных группах с выровненными начальными параметрами отдельных условий, подвергающихся экспериментальной проверке, и сравнению результатов обучения (воспитания). Экспериментальные группы отличаются ориентацией на различные педагогические условия.

Для проверки, допустим, трёх условий необходимо как минимум четыре экспериментальных и одна контрольная группы. В одной группе (ЭГ-1) проверяется эффективность первого условия, во второй группе (ЭГ-2) — второго условия, в третьей группе (ЭГ-3) — третьего условия, в четвёртой группе (ЭГ-4) проверяется эффективность комплекса условий.

В контрольной группе обучение (воспитание) ведётся без выделенных исследователем педагогических условий. В противном случае было бы невозможным сравнение результатов.

Анализ результатов экспериментальной работы приводит исследователя к выводу, что каждое из условий «работает» на повышение качества учебно-воспитательного процесса, но только их комплекс решает эту задачу в полной мере.

Подводя итог вышесказанному, ещё раз подчеркнём, что планирование любого эксперимента является важнейшим этапом исследовательской работы в любой области знаний. Составление рационального плана эксперимента позволяет, прежде всего, получить более точные и достоверные знания об интересующем нас объекте или явлении. Кроме того, это позволяет значительно сократить затраты времени и средств на проведение самого эксперимента. С другой стороны, неудачно спланированный эксперимент, потребовав много сил и времени исследователя, в результате может дать весьма неточные, а иногда и сомнительные результаты.

Достижения теории оптимального эксперимента носят достаточно общий характер, а поэтому несложно переносятся в отдельные области знаний. В частности, её основные положения можно применять для проведения педагогических исследований, сохранив при этом все достоинства общей теории.

Выводы по второй главе

1. Понятие «технология» многогранно. Существующие определения этого феномена можно подразделить на теоретические и практические. В первом случае технологию раскрывают через понятия «деятельность» и «средства её осуществления», а во втором — через процедуры и операции. Наиболее продуктивным в аспекте нашего исследования являются определения, в основе которых заложены процедуры и операции. При этом под процедурой понимается совокупность действий (операций) посредством которых осуществляется процесс (или его отдельный этап), выражающий суть конкретной технологии, а под операцией — непосредственное практическое решение задачи в рамках данной процедуры.
2. Образование, являясь сложноорганизованным объектом, относится к разряду системных, и поэтому возникает необходимость использования системного подхода, инструментом которого является системный анализ. Моделирование, в свою очередь, является инструментом системного анализа и представляет собой метод познания исследуемых качеств объекта через модель, где модель — система, исследование которой служит средством для получения информации о другой системе (А. И. Уёмов).
3. В проведении исследовательской работы выделяют три этапа: формализацию проблемы (построение математической модели), планирование эксперимента и анализ полученных данных.
4. Выделяют два типа моделирования — содержательное и статистическое моделирование (метод Монте-Карло). Первый метод, опираясь на априорную информацию об объекте, позволяет вывести общий вид аналитических уравнений, описывающий ис-

следуемую зависимость, а затем найти неизвестные параметры этих уравнений на основе исходных статистических данных. Второй метод состоит в имитации поведения реальной системы для получения дополнительных статистических данных.

5. Широко используемый на практике метод построения модели по принципу «черного ящика» состоит из трёх этапов: 1) выделение факторов, действующих на объект, и откликов на эти действия; 2) выбор шкалы измерения уровней факторов и откликов; 3) выдвижение гипотезы о виде функции отклика, определяющей связь между значениями факторов и откликов.
6. Планирование эксперимента проводится для получения оптимального в некотором смысле эксперимента. Оптимальность определяется с помощью специально разработанных критериев оптимальности. Наиболее важными с точки зрения педагогического эксперимента характеристиками оптимальности эксперимента являются рандомизация, насыщенность, композиционность и простота обработки.
7. Педагогический эксперимент представляет собой комплекс методов исследования, предназначенный для объективной проверки достоверности гипотезы или гипотез.

Технология организации констатирующего этапа эксперимента, предшествующего формирующему этапу, включает следующие компоненты: выявление состояния объекта; выявление педагогических средств и условий, обусловившие исходное состояние; выявление педагогических средств и условий, наличествующих к началу формирующего этапа эксперимента.

Технология организации формирующего этапа эксперимента включает введение новой или модернизированной системы (технологии, ...); создание в вариативном режиме различ-

ных педагогических условий, при которых эта система (технология, ...) эффективно функционирует и развивается.

Технология организации констатирующего этапа эксперимента, следующего за формирующим этапом, состоит в выявлении состояния объекта; в выявлении наличия происшедших изменений; в качественном и количественном анализе происшедших изменений; интерпретация результатов эксперимента.

Глава III. Технология квалиметрической оценки результатов педагогического эксперимента

§ 1. Основные понятия математической статистики

Математическая статистика, как научная дисциплина, представляет собой «раздел математики, посвящённый математическим методам систематизации, обработки и использования статистических данных для научных и практических выводов» (63, с. 344). При этом под статистическими данными понимается любая числовая информация, характеризующая некоторую совокупность объектов, обладающих теми или иными общими признаками. Например, количество учащихся в классе, имеющих оценку «5» по геометрии, или процент студентов, получающих стипендию, и т.д. Основная ценность методов математической статистики для педагога-исследователя состоит в возможности делать некоторые выводы обо всей совокупности объектов, используя данные о сравнительно небольшой группе из этой совокупности.

Статистический подход к изучению каких-либо процессов или явлений состоит в мысленном расчленении происходящих изменений на две части — закономерную и случайную, и выявлении закономерной изменчивости на фоне случайной. Методы математической статистики позволяют оценивать параметры имеющихся закономерностей, проверять те или иные гипотезы об этих закономерностях и т.д. При этом статистика существенно опирается на теорию вероятностей, которая изучает закономерности, порождённые случайными событиями.

Введём несколько понятий теории вероятности и математической статистики, необходимых для дальнейшего изложения материала.

Прежде всего, определимся с понятием вероятности. В математическом энциклопедическом словаре *вероятность* определяется как «числовая характеристика степени возможности наступления какого-либо определённого события в тех или иных определённых, могущих повторяться неограниченное число раз условиях» (63, с. 118). Часто численное значение вероятности определяется как отношение числа «благоприятных» для нашего события исходов к общему числу исходов (это так называемое классическое определение вероятности). Например, выпадение чётного числа очков при бросании игральной кости возможно с вероятностью $\frac{1}{2}$, поскольку благоприятных исходов 3 (выпадение 2, 4 или 6 очков), а всего исходов — 6. Разумеется, такое правило действует только, если мы имеем дело с конечным числом возможных исходов. В общем случае, действует следующая закономерность. Предположим, что мы хотим измерить вероятность наступления какого-либо события в некотором эксперименте. Для этого мы осуществим этот эксперимент некоторое число раз, например, n раз. Подсчитаем число тех опытов, в которых произошло интересующее нас событие, пусть это будет m . Тогда частота $V = m/n$ нашего события в n повторениях приблизительно равна вероятности наступления этого события. Причём эта оценка тем точнее, чем больше опытов мы провели. Это несложное, на первый взгляд, замечание составляет основное содержание исторически одной из первых теорем теории вероятностей — теоремы Бернулли, доказанной в конце семнадцатого века.

Поскольку в статистических расчётах мы имеем дело с данными, подверженными случайной изменчивости, невозможно точно определить характеристики исследуемых случайных процессов. Это вынуждает нас прибегать к некоторым их оценкам, т.е. к приближённым значениям. Поэтому для статистики большую роль играют так называемые *выборочные* или *эмпирические* значения характеристик. В частности, в качестве оценки для вероятности некоторого события берётся частота наступления этого события.

Руководствуясь вышеизложенным, мы не будем давать строгое аксиоматическое определение вероятности, отметив лишь её основные свойства.

1. Значение вероятности лежит в промежутке от 0 до 1.
2. Вероятность достоверного события (т.е. события, которое обязательно произойдёт при случайном испытании) равна 1, а вероятность невозможного события (т.е. такого, которое не может произойти) равна 0.
3. Если события несовместимы (т.е. не могут произойти одновременно), то вероятность наступления одного из этих событий равна сумме вероятностей наступления каждого из них.

Традиционно вероятность наступления события A обозначается через $P(A)$.

Следующим важнейшим понятием теории вероятности является понятие случайной величины.

Под *случайной величиной* будем понимать некоторую переменную величину, принимающую в зависимости от случая те или иные числовые значения. Типичным примером случайной величины является число очков, выпавших при бросании игральной кости. Понятно, что выпадение конкретного числа определяется случай-

ными причинами. Другим примером может служить время, затраченное учеником на выполнение задания. Обычно выделяют *дискретные* и *непрерывные* случайные величины. Дискретные величины характеризуются тем, что можно перенумеровать все их возможные значения. Число очков при бросании игральной кости является как раз примером дискретной случайной величины. Непрерывные же случайные величины могут принимать любые значения из некоторого промежутка на множестве действительных чисел. Так, затраченное время является непрерывной случайной величиной.

Основной характеристикой случайной величины является её функция распределения.

Пусть ξ — некоторая случайная величина, принимающая действительные значения. Тогда *функцией распределения* случайной величины ξ называется функция $F(x)=P(\xi < x)$, т.е. значением функции распределения является вероятность того, что случайная величина ξ принимает значение меньше x .

Учитывая то, что вероятность оценивается при помощи частоты наступления события, *выборочную функцию распределения* можно найти по формуле

$$\hat{F}(x) = \frac{f(x)}{n},$$

где $f(x)$ — число наблюденных значений случайной величины ξ меньших x , и n — число испытаний.*

* Отметим попутно, что более строгое определение непрерывной случайной величины гласит: «Случайную величину, принимающую вещественные значения, называют непрерывной, если непрерывна её функция распределения» (98,

Если функция распределения случайной величины непрерывна и дифференцируема, то закон распределения описывают также при помощи *плотности вероятности* $p(x)$, которая может быть определена как производная функции распределения, т.е. $p(x) = F'(x)$.

Функция распределения даёт исчерпывающую информацию о поведении случайной величины. Однако для практических целей зачастую бывает достаточно знания нескольких числовых характеристик распределения, позволяющих оценить такие его свойства как центр группирования значений, разброс значений вокруг этого центра и т.д. Важнейшими из таких характеристик являются *моменты* распределения. Мы обсудим только два из них: математическое ожидание (первый момент) и дисперсию (второй момент).

Для дискретной случайной величины ξ , принимающей значения $x_1, x_2, x_3, \dots, x_n$, с вероятностями $p_1, p_2, p_3, \dots, p_n$, *математическое ожидание* определяется по формуле

$$M\xi = \sum_{k=1}^n x_k p_k.$$

То есть математическое ожидание — это некоторое «осреднение» случайной величины с учётом вероятности получения отдельных значений. Для непрерывной случайной величины такое

с. 26). Отсюда следует, что для непрерывных распределений вероятность каждого отдельного значения случайной величины равна нулю. Для дискретных же распределений все такие вероятности принимают некоторые положительные значения. Именно в этом и кроется глубинный смысл противопоставления непрерывных и дискретных случайных величин.

осреднение производится при помощи интеграла.* Заменяв в определении математического ожидания вероятности их оценками через частоту наступления событий, получим *выборочное математическое ожидание*:

$$\hat{M}\xi = \bar{x} = \frac{1}{n} \sum_{k=1}^n x_k.$$

Таким образом, математическое ожидание можно оценить при помощи среднего арифметического \bar{x} значений случайной величины ξ .

Величина разброса случайной величины вокруг среднего значения характеризуется дисперсией.

Дисперсия случайной величины ξ определяется по формуле

$$D\xi = M(\xi - M\xi)^2 \text{ или } D\xi = M\xi - (M\xi)^2.$$

Следовательно, *выборочная дисперсия* может быть найдена по формуле

$$\hat{D}\xi = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2,$$

где \bar{x} — среднее значение наблюдений.

Из приведённых формул видно, что чем больше разброс вокруг среднего значения, тем больше значение дисперсии.

Дисперсию часто обозначают через σ^2 (сигма в квадрате), где σ — означает *среднеквадратичное отклонение*. То есть среднеквадратичное отклонение — это корень квадратный из дисперсии.

* Если быть более точным, то математическое ожидание непрерывной случайной величины равно $M\xi = \int_{-\infty}^{+\infty} xp(x)dx$ где $p(x)$ — плотность вероятности.

Эту характеристику используют, когда нужно, чтобы показатель разброса случайной величины выражался в тех же единицах, что и сами значения случайной величины.

Выборочное среднеквадратичное отклонение вычисляется по формуле*

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^2}.$$

Вернёмся теперь от числовых характеристик распределений случайных величин к законам их распределения. Очевидно, что существует бесконечное число различных функций распределений, а значит и бесконечное число самих вероятностных распределений. Однако среди этого великого многообразия можно выделить такие распределения, которые встречаются на практике довольно часто. Многие из них лежат в основе целых областей знания — таких как теория массового обслуживания, теория надёжности, теория измерений, теория игр и т.п. Мы бегло рассмотрим те распределения, которые понадобятся нам для целей педагогического эксперимента.

Одним из самых распространённых дискретных распределений является *биномиальное распределение*. Оно возникает в тех случаях, когда нас интересует, сколько раз происходит некоторое событие в серии из определённого числа независимых наблюдений

* Несколько более точные оценки дисперсии и среднеквадратичного отклонения получаются по формулам $\hat{\sigma}^2 = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2$, и

$$\hat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2}.$$

Такие оценки являются *несмещёнными* в том смысле, что математическое ожидание от полученных по ним значений равно дисперсии и среднеквадратичному отклонению соответственно.

(опытов), выполняемых в одинаковых условиях. Последовательность независимых испытаний, в которых результатом каждого из испытаний может быть один из двух исходов (например, «успех» и «неуспех»), и вероятность «успеха» в каждом из испытаний одна и та же, называется *схемой Бернулли*.

Итак, предположим, что случайная величина ξ равна числу «удач» в n испытаниях Бернулли, причём вероятность успеха в каждом испытании равна p . Тогда ξ принимает значения $0, 1, 2, \dots, n$ с вероятностями

$$P(\xi = k) = C_n^k p^k (1 - p)^{n-k},$$

где C_n^k — число сочетаний из n элементов по k элементов. Эта формула и описывает закон *биномиального распределения* с параметрами n и p .

Математическое ожидание и дисперсия случайной величины, имеющей биномиальное распределение, равны соответственно $M\xi = np$ и $D\xi = p(1-p)$.

На рисунке 1 показаны вероятности $P(\xi = k)$ при $n = 10$ для различных значений p .

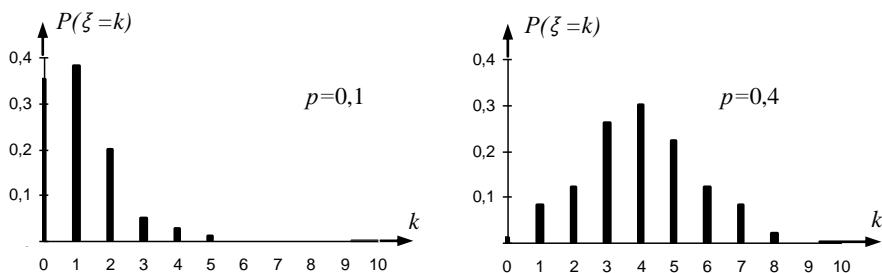


Рис. 1. Вид биномиального распределения для $p = 0,1$ и $p = 0,4$ при $n = 10$

Для биномиального распределения существуют два типа таблиц. В таблицах первого типа приводятся вероятности $P(\xi = k)$, а в таблицах второго типа вероятности $P(\xi \leq k)$ при различных значениях n и p . Достаточно полные таблицы с подробными комментариями можно найти, например, в (16) или (68).

Важнейшим среди непрерывных распределений является *нормальное (гауссовское) распределение*. Это распределение возникает, когда «значение непрерывной случайной величины формируется под воздействием очень большого числа независимых случайных факторов, причём сила воздействия каждого отдельного фактора мала и не может превалировать среди остальных, а характер воздействия — аддитивный (т.е. при воздействии случайного фактора F на величину a получается величина $a + \Delta F$, где случайная «добавка» ΔF мала и равновероятна по знаку)» (5, с. 169).

Если функция распределения случайной величины ξ задаётся формулой

$$F(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \int_{-\infty}^x e^{-\frac{(t-a)^2}{2\sigma^2}} dt,$$

то говорят, что случайная величина ξ имеет *нормальное распределение* с параметрами a и σ^2 . Функцию $F(x)$ часто называют функцией Лапласа. Понятно, что плотность нормального распределения определяется формулой

$$p(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{(t-a)^2}{2\sigma^2}}.$$

Математическое ожидание и дисперсия в этом случае равны соответственно $M\xi = a$ и $D\xi = \sigma^2$. Если $a = 0$ и $\sigma = 1$, то распределение называют *стандартным нормальным распределением*.

Смысл параметров нормального распределения хорошо виден из рисунка 2.

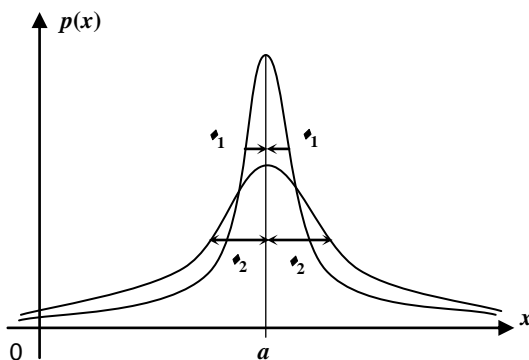


Рис. 2. Плотность нормального распределения со средним a и различными значениями дисперсии σ^2

Для функции распределения и для плотности вероятности нормального распределения существует множество таблиц разной степени точности (см., например, (16; 68; 105)). В практических вычислениях более удобны таблицы для плотности и таблицы квантилей.

При использовании нормального закона распределения следует учитывать следующее важное обстоятельство. Нормальное распределение не является универсальным распределением, описывающим любые явления. Это лишь один из многих типов распределения, имеющих в природе, хотя и наиболее распространённый. Однако нормальный закон распределения имеет понятные математические свойства и хорошо изучен, что делает его очень

популярным. Кроме того, с помощью нормального закона несложно описываются многие другие распределения. В одних случаях точно, а в других — с достаточной точностью приближения.

Использованию нормального распределения для приближённого описания распределений случайных величин не препятствует и то обстоятельство, что эти величины могут принимать значения из некоторого ограниченного интервала (например, возраст от 0 до 100 лет), а нормальное распределение не может быть сосредоточено ни на каком конкретном интервале. Дело в том, что вероятность больших отклонений нормальной случайной величины от центра распределения практически равна нулю и ею можно пренебречь. Этот факт хорошо иллюстрирует график плотности нормального распределения (см. рис. 2).

Но даже когда экспериментальные данные существенно отклоняются от нормального закона, существует, по крайней мере, два варианта его использования. Во-первых, использовать его в качестве первого приближения. При этом часто оказывается, что точность такого приближения нас вполне устраивает. Во-вторых, произвести некоторое преобразование исследуемой случайной величины, которое даст новую случайную величину с нормальным законом распределения. На этом пути получено много эффективных критериев для проверки гипотез.

При операциях со случайными величинами возникает несколько новых видов распределений. В первую очередь это распределение Стьюдента, распределение χ -квадрат и F –распределение. Эти распределения играют важную роль в статистическом анализе и широко используются в дальнейшем. Мы не будем приводить здесь описание перечисленных распределений, так как для практического использования статистических критериев нам понадобятся

лишь конкретные значения соответствующих функций распределения, которые можно найти в многочисленных таблицах (см., например, 16; 58; 68).

Как мы уже отмечали в начале параграфа, одной из задач статистики является проверка гипотез о свойствах всей совокупности объектов на основании данных о некоторой её части. Решению этой важной задачи посвящён целый раздел математической статистики, называемый *теорией статистического вывода*.

Основная идея этой теории заключается в том, что некоторое множество объектов, выбранных случайным образом из общей совокупности объектов, имеет те же свойства, что и вся исходная совокупность. При этом естественно такое рассуждение имеет вероятностный характер. А, значит, оно неверно также с некоторой вероятностью. Если вероятность справедливости некоторого утверждения близка к единице, то оно практически *достоверно*, и мы можем принять данное утверждение. Например, вероятность 0,99 означает, что лишь для одного объекта из ста наш вывод может оказаться ошибочным. Следовательно, определив для себя приемлемый уровень ошибки, мы можем делать какие-либо выводы относительно исследуемых объектов.

Понятно, что для проверки некоторой гипотезы о реальных объектах средствами математической статистики мы должны, прежде всего, перевести наши знания об объекте на язык математики, т.е. другими словами, построить математическую модель. После этого гипотезу о свойствах объекта следует сформулировать как гипотезу о свойствах распределения некоторой случайной величины. Например, гипотеза об одинаковом уровне подготовленности учащихся двух классов может быть сформулирована как гипотеза о

принадлежности оценок учащихся этих классов одному закону распределения.

Опишем общую схему применения статистических критериев для проверки гипотез.

Итак, предположим, что мы построили математическую модель нашего объекта.*

Прежде всего, необходимо перевести проверяемую гипотезу на язык математики, т.е. выдвинуть *статистическую гипотезу*. Это может быть гипотеза о законе распределения некоторой случайной величины, или гипотеза о числовых значениях параметров исследуемой совокупности объектов, или гипотеза об однородности нескольких групп данных, и т.д. Если мы можем непосредственно проверить выдвигаемую гипотезу, то не возникает никаких проблем. Однако чаще всего приходится проверять не саму гипотезу, а некоторые обязательные следствия из неё. При этом если следствия не выполняются, мы делаем вывод о ложности гипотезы. Однако, если следствия выполняются, мы не можем утверждать, что гипотеза справедлива. То есть косвенным образом гипотезу *доказать* нельзя, а можно лишь *опровергнуть*. Поэтому, чтобы расширить круг применяемых критериев, принято формулировать так называемую «нулевую» гипотезу, ложность которой означала бы истинность нашего предположения. Например, вместо проверки неоднородности групп по какому-либо показателю, мы будем проверять их однородность. Тогда отвержение второй гипотезы будет означать справедливость первой. Нулевая гипотеза, как правило, носит более общий характер и поэтому отвергнуть её бывает легче, чем подтвердить альтернативную гипотезу.

* Вопросы построения модели подробно обсуждались в §1 второй главы.

Итак, будем считать, что мы выдвинули статистическую гипотезу. Традиционно проверяемую гипотезу обозначают через H_0 , а конкурирующую с ней альтернативную гипотезу — через H_1 . Поскольку при проверке любой статической гипотезы решение всегда принимается с некоторой долей вероятности, мы должны определить для себя приемлемую точность вывода, то есть задать степень риска получить неверный вывод. Этот риск, представленный как вероятность, называется *уровнем значимости* и обычно обозначается через α . Другими словами, уровень значимости — это вероятность отвержения выдвинутой гипотезы в случае, когда она на самом деле верна. Выбор величины уровня значимости α зависит от величины потерь, которые мы понесём в результате неверно принятого решения. Однако, поскольку такое сопоставление в практических задачах бывает весьма затруднительным, прибегают к некоторым стандартным уровням значимости. Например, α берут равным 0,1; 0,05; 0,025 и т.д. Так уровень значимости 0,01 означает, что в одном случае из ста мы будем ошибочно отвергать высказанную гипотезу при использовании данного статистического критерия. Наиболее распространённым является уровень значимости 0,05, так как, с одной стороны, его достаточно для большинства задач, а с другой стороны, такая точность согласуется с потерями при моделировании реальных объектов.

После выбора уровня значимости мы должны составить некоторую функцию от результатов наблюдения, которая сама является случайной величиной и при справедливости гипотезы H_0 должна иметь известное нам распределение. Например, если мы предполагаем, что средние значения в двух группах данных равны, то отношение оценок дисперсий, вычисленных двумя способами, имеет

распределение Стьюдента. На этом факте основаны некоторые методы дисперсионного анализа.

Для выбранного уровня значимости мы находим так называемое *критическое* значение выбранной функции. Эти значения приведены в таблицах соответствующих распределений и их нахождение, как правило, не требует никаких дополнительных вычислений. Найденное критическое значение выделяет *критическую область* (или *область отбрасывания*).

Наконец, вычисляем значение выбранной функции для полученных экспериментальных данных и сравниваем его с критическим. Если полученное значение попадает в критическую область, то мы считаем результаты наблюдений несовместимыми с гипотезой и поэтому отвергаем её и принимаем альтернативную гипотезу с вероятностью $1 - \alpha$. При этом мы предполагаем, что альтернативная гипотеза H_1 является отрицанием основной гипотезы H_0 . Этот случай достаточен для целей педагогического эксперимента.*

Из приведённой схемы применения статистического критерия следует возможность совершения ошибок двух видов:

* Поясним более общую постановку проблемы на следующем примере. Предположим, что гипотеза H_0 состоит в том, что средний возраст студентов вуза равен 19 годам. Отрицанием этой гипотезы будет гипотеза о том, что средний возраст не равен 19 годам. Однако в качестве альтернативной гипотезы можно выдвинуть предположение о том, что средний возраст равен 20 годам. Никак не сказываясь на общей схеме применения критерия, это различие выявится лишь при вычислении вероятности справедливости альтернативной гипотезы. Примеры вычисления таких вероятностей можно найти, скажем, в (24, с. 258). Отметим лишь, что вероятность зависит от конкретного вида альтернативной гипотезы. Так если альтернативная гипотеза состоит в том, что средний возраст равен 19,5 лет, вероятность будет уже другой.

- отвержение гипотезы в случае, когда она на самом деле верна;
- принятие гипотезы в случае, когда она неверна.

Эти возможности называются соответственно ошибками *первого* и *второго рода*. Вероятность ошибки первого рода равна выбранному уровню значимости α . Вероятность ошибки второго рода обозначают через β , и в общем случае $1 - \beta$ не обязательно равно α . Величина $1 - \beta$ называется *мощностью критерия* и зависит от альтернативной гипотезы. Понятно, что чем ближе к единице мощность критерия, тем он эффективнее.

Все статистические критерии проверки гипотез строятся по изложенной выше схеме и различаются только конкретным видом выбираемой функции от значений наблюдения. Конкретные статистические критерии будут изложены в §2,3 данной главы.

§ 2. Методы определения связи признаков

Во многих практических задачах при исследовании объектов, обладающих несколькими признаками, необходимо бывает выяснить, насколько эти признаки связаны между собой. Например, как связаны оценки по математике и оценки по физике, или успеваемость в вузе с типом темперамента и т.д. Поскольку мы ведём речь о статистических критериях, прежде всего мы должны выразить экспериментальные данные в числовой форме, т.е. измерить их. Как мы выяснили в §3 первой главы, существуют разные типы шкал измерений, и значения на этих шкалах несут в себе различный запас информации. Следовательно, методы определения связи признаков существенно зависят от шкалы измерения этих признаков. Поэтому первым этапом анализа является классификация типа данных, т.е. отнесение их к той или иной шкале измерений. Вторым этапом является проверка гипотезы об отсутствии связи признаков. На третьем этапе производится оценка силы связи признаков, если гипотеза об их независимости была отвергнута.

Как было отмечено выше, для определения связи признаков, измеренных в разных шкалах, применяются разные критерии. Для данных, измеренных в шкале наименований, чаще всего используются таблицы сопряжённости и статистика Фишера-Пирсона χ -квадрат. Для данных, измеренных в шкале порядка, используют ранжирование и коэффициенты корреляции Спирмена и Кендэлла. Для данных, измеренных в количественных шкалах, применяют коэффициент корреляции Пирсона и средства регрессионного анализа. Остановимся подробнее на этих критериях, ограничившись случаем двух признаков.

До конца этого параграфа под «нулевой» гипотезой H_0 будем понимать гипотезу о независимости двух признаков A и B . При

этом признаки A и B будем называть *независимыми*, если значение, принятое признаком A , не влияет на вероятности возможных значений признака B и наоборот.

Таблицы сопряжённости

Предположим, что в нашем распоряжении имеется выборка, состоящая из n объектов. Каждый объект характеризуется двумя признаками, скажем, A и B . Пусть признак A может принимать r значений A_1, A_2, \dots, A_r , а признак B принимает s значений B_1, B_2, \dots, B_s . Составим таблицу сопряжённости данных признаков.

Таблица 3

Таблица сопряжённости признаков A и B

	B_1		B_2		...		B_s		Сумма по столбцам
A_1	n_{11}	v_{11}	n_{12}	v_{12}	n_{1s}	v_{1s}	P_1
A_2	n_{21}	v_{21}	n_{22}	v_{22}	n_{2s}	v_{2s}	P_2
...
A_r	n_{r1}	v_{r1}	n_{r2}	v_{r2}	n_{rs}	v_{rs}	P_r
Сумма по строкам	Q_1		Q_2		...		Q_s		n

Здесь через n_{ij} обозначено количество объектов выборки, обладающих комбинацией уровней A_i и B_j , через $P_i = n_{i1} + n_{i2} + \dots + n_{is}$ — число появлений значений признака A_i и через $Q_j = n_{j1} + n_{j2} + \dots + n_{jr}$ — число появлений значений признака B_j . Через v_{ij} в таблице обозначена ожидаемая частота совместного появления значений A_i и B_j , найденная по формуле

$$v_{ij} = \frac{P_i Q_j}{n}.$$

Статистика Пирсона-Фишера

Критерий Пирсона-Фишера позволяет выяснить, существует ли какая-либо связь между признаками A и B , или же они ведут себя независимо друг от друга? Этот критерий основан на свойствах распределения χ^2 .

Вычислим значение статистики Пирсона-Фишера, пользуясь формулой

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - v_{ij})^2}{v_{ij}}.$$

Теперь нужно сравнить полученное значение с критическим для выбранного уровня значимости значением, найденным по таблицам распределения χ^2 с $(r - 1)(s - 1)$ степенями свободы. Если полученное значение превосходит табличное, то гипотеза о независимости признаков A и B отвергается на выбранном уровне значимости.

Следует отметить, что критерий Пирсона-Фишера не следует использовать при малом числе наблюдений. Для его корректного применения считается достаточным, чтобы все ожидаемые частоты v_{ij} были бы не меньше 5.

Проиллюстрируем применение критерия Пирсона-Фишера на следующем примере из (93). У каждого члена группы из 195 человек по методике Чижковской был определён тип нервной деятельности (признак B). Затем каждому испытуемому было предложено выбрать инструкцию, регламентирующую его дальнейшую деятельность (признак A). Каждый из признаков имеет два уровня. Для признака A — это высокая реактивность (+) и низкая реактивность (–) нервной деятельности. Для признака B — детальная инструкция

и краткая инструкция. Результаты измерений с уже вычисленными значениями P_i , Q_j и V_{ij} приведены в таблице 4.

Таблица 4

Предпочтение различных видов инструкции высокорективными и низкорективными индивидами

Тип инструкции	+		–		Сумма по столбцам
Детальная	63	52,2	42	52,7	105
Краткая	34	44,8	56	45,2	90
Сумма по строкам	97		98		195

Вычисление статистики хи-квадрат по приведённой выше формуле даёт значение $\chi^2 = 9,58$. Число степеней свободы в нашем случае равно произведению $(2 - 1)(2 - 1) = 1$. По таблице распределения χ^2 с одной степенью свободы находим, что вероятность получить значение 7,88 равна 0,005, а вероятность получить значение 10,83 равна 0,001. Следовательно, мы можем сделать вывод: вероятность случайно получить для независимых признаков наше значение 9,58 не превышает 0,005. Можно говорить поэтому, что гипотеза о независимости признаков A и B противоречит результатам эксперимента и мы должны её отвергнуть, т.е. связь между этими признаками присутствует.

Как показывает приведённый выше пример, практическое применение критерия Пирсона-Фишера не представляет никакой сложности и не содержит больших вычислений. Излагаемые дальше методики также просты в работе, поэтому мы не будем иллюстрировать их конкретными вычислениями. Достаточное количество примеров по большинству предлагаемых в дальнейшем методик можно найти в (24).

Ранжирование

Как известно, в шкалах порядка основную нагрузку несут на себе не сами числа, сопоставленные проявлению признака, а их взаимное расположение. Следовательно, для нас важны не значения этих чисел, а их порядок. Поэтому для дальнейшей обработки данных, измеренных в порядковой шкале, полезнее перейти к рангам.

Рангом наблюдения называется тот номер, который получит это наблюдение после упорядочения всех данных по возрастанию (или убыванию) силы проявления признака.

Процедура перехода от совокупности наблюдений к последовательности их рангов называется *ранжированием*.

Если все данные наблюдений различны, не возникает никакой трудности при их упорядочении. Если же есть повторяющиеся данные, то переходят к так называемым *средним рангам*. Они вычисляются следующим образом. Предположим, что мы имеем n одинаковых наблюдений, называемых *связкой*. Если бы они были различными, то занимали бы после ранжирования места, допустим, с i -того по $i+n$ -тое. Чтобы посчитать средний ранг связки, надо сложить все эти ранги и разделить их на число элементов в связке, т.е. вычислить выражение $[i+(i+1)+(i+2)+\dots+(i+n)]/n$. Каждому элементу в связке сопоставляется этот средний ранг. Например, после ранжирования выборки 3, 4, 5, 9, 10, 17, 42 мы получим последовательность 1, 2, 3, 4, 5, 6, 7, а после ранжирования выборки 2, 3, 3, 7, 15, 15, 15 — последовательность 1, $2\frac{1}{2}$, $2\frac{1}{2}$, 3, 6, 6, 6.

Перейдём теперь непосредственно к анализу связи признаков. Пусть нам даны два признака A и B , измеренных в шкале порядка. Требуется выяснить, есть ли связь между этими признаками.

Прежде всего, осуществим ранжирование данных. В результате каждому объекту будет сопоставлена пара чисел (r_i, s_i) , где r_i — ранг объекта относительно признака A и s_i — ранг объекта относительно признака B .

Критерий Спирмена

Сначала надо вычислить статистику S , используя формулу

$$S = \sum_{i=1}^n (r_i - s_i)^2.$$

Затем, чтобы сгладить действие численности группы n , перейдём к коэффициенту ранговой корреляции Спирмена:

$$\rho = 1 - \frac{6S}{n^3 - n}.$$

Коэффициент ρ лежит на отрезке от -1 до 1 . Свои крайние значения он принимает в случаях, когда признаки A и B полностью противоположны ($\rho = -1$) или когда они полностью совпадают ($\rho = 1$). Чем ближе значение ρ к нулю, тем меньше связь между признаками. Следует отметить, что если связь между признаками более сложная, чем просто монотонная, то коэффициент Спирмена не всегда может отличить зависимость от независимости.

Для проверки гипотезы об отсутствии связи признаков надо воспользоваться таблицами распределения коэффициента Спирмена. Если полученное значение превосходит табличное по модулю, то предположение о независимости признаков отвергается на выбранном уровне значимости.

Отметим в заключение, что наличие связей, т.е. повторяющихся данных, снижает точность изложенного метода.

Критерий Кендэлла

Данный критерий основан на понятии инверсии.

Говорят, что числа x_i и x_j образуют *инверсию* в последовательности натуральных чисел (x_1, x_2, \dots, x_n) , если большее из них стоит левее меньшего.

Число инверсий находится следующим образом:

- 1) Подсчитываем количество чисел, стоящих в последовательности левее 1. Обозначим его через k_1 . Вычёркиваем единицу в нашей последовательности.
- 2) Подсчитываем количество не зачёркнутых чисел, стоящих в последовательности левее 2. Обозначим его через k_2 . Вычёркиваем двойку. И так далее.

Число инверсий в нашей последовательности равно $k_1 + k_2 + \dots + k_n$.

Например, число инверсий в последовательности (4, 3, 1, 2) равно $2+2+1=5$.

Вернёмся теперь к методу Кендэлла. Отметим сразу, что этот метод используется при отсутствии связей в исходных данных. Применяется он следующим образом. Прежде всего, упорядочим наши объекты по рангу первого признака, т.е. перейдём к набору данных $(1, s_1), (2, s_2), \dots, (n, s_n)$. Подсчитав число инверсий в последовательности (s_1, s_2, \dots, s_n) , мы найдём значение статистики K . Как и прежде для ослабления влияния числа объектов в группе переходят к *коэффициенту ранговой корреляции Кендэлла*, рассчитанному по формуле

$$\tau = 1 - \frac{4K}{n(n-1)}.$$

Коэффициент Кендэлла τ , как и коэффициент Спирмена ρ , может изменяться только в пределах от -1 до 1 , принимая крайние значения при полном совпадении ($\tau = 1$), либо при полной противоположности ($\tau = -1$) исследуемых признаков.

Пользуясь таблицами распределения, найдём критические для выбранного уровня значимости значения τ . Если полученное нами значение превосходит табличное по модулю, то гипотеза о независимости признаков отвергается на данном уровне значимости.

Коэффициент корреляции Пирсона

Предположим, что нам дана выборка из n объектов, каждому из которых сопоставлены два значения, характеризующих признаки A и B . Т.е. с каждым объектом связаны две последовательности оценок a_1, a_2, \dots, a_n и b_1, b_2, \dots, b_n . Вычислим значение коэффициента корреляции по формуле

$$r = \frac{\sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^n (a_i - \bar{a})^2} \sqrt{\sum_{i=1}^n (b_i - \bar{b})^2}}, \text{ где } \bar{a} = \frac{1}{n} \sum_{i=1}^n a_i.$$

Коэффициент r принимает значения от -1 до 1 . Причём, он равен -1 или 1 , если связь между признаками линейная. Если признаки независимы и коэффициент корреляции существует, то он равен 0 . С другой стороны, равенство нулю коэффициента r не означает статистической независимости исследуемых признаков. Однако, если совместное распределение пар (a_i, b_i) является двумерным нормальным распределением, то из равенства $r = 0$, безусловно, следует независимость признаков A и B . Для этого случая построены таблицы распределений, пользуясь которыми мы можем сделать вывод о справедливости нашей гипотезы.

Итак, если полученное значение r по модулю превосходит критическое значение для выбранного уровня значимости, то гипотеза о независимости исследуемых признаков отвергается на данном уровне значимости.

Отметим попутно, что в случае двумерного нормального распределения признаков A и B коэффициент корреляции позволяет не только решить вопрос о зависимости признаков, но и определить степень их связи. Это легко сделать, определив доверительный интервал для истинного значения корреляции на данном уровне значимости. Значения этих интервалов также можно найти в соответствующих таблицах.

Если гипотеза о независимости признаков отвергается, то обычно выясняют степень силы связи. Для этого используют различные *меры связи*, т.е. некоторые числовые характеристики, указывающие на эту связь. Наиболее распространёнными среди них являются уже знакомые нам коэффициенты корреляции Кендэлла и Спирмена. Используются также коэффициенты Крамера, Чупрова, Гудмена-Краскела и др. Поскольку описанные выше методы определения связи признаков позволяют определять и силу этой связи, мы не будем приводить здесь другие коэффициенты. Более подробную информацию по этой теме можно найти в (7; 43).

Нередко приходится устанавливать связь признаков, измеренных в разных шкалах. Для таких случаев разработаны специальные методики, которые, безусловно, гораздо сложнее приведённых выше. Многие из них описаны в (44). Однако проще всего в такой ситуации одну из шкал понизить до уровня другой и проводить исследования при помощи стандартных методик. Понятно, что при этом происходит некоторая потеря информации, зато последующий анализ становится более понятным.

Существует ещё много статистических критериев для определения связи признаков и установления силы этой связи. Мы попытались представить в этом параграфе наиболее универсальные из них. Эта универсальность означает минимум предположений о характере изучаемых данных, а, значит, и минимум дополнительных проверок, которые, как правило, и представляют основную трудность для неспециалиста. Применение же какого-либо критерия для несвойственной ему ситуации в лучшем случае даст очень неточный, а в худшем — просто неверный результат. Кроме того, отбирались те критерии, которые не требуют каких-либо специальных знаний и осуществляются с минимальными вычислениями. Для применения вышеизложенных критериев достаточно иметь элементарные вычислительные навыки и набор статистических таблиц в рамках любого справочника.

§ 3. Факторный анализ

Если можно явно выделить несколько факторов, влияющих на конечный результат, причём эти факторы принимают конечное число значений, то возникает задача *факторного анализа*. Типичный пример — сравнение по достигаемым результатам нескольких различных способов действия, например, методик преподавания, стимулирования и т.д. То, что, по нашему мнению, оказывает влияние на конечный результат, называется *фактором*, конкретную реализацию фактора (например, конкретную методику) называют *уровнем фактора* или *способом обработки*, а значение измеряемого признака (т.е. величину результата) называют откликом.

Если мы выделяем несколько факторов, то получаем задачу *многофакторного анализа*, а если один — то *однофакторного анализа*.

Бывает, что в рамках однофакторной модели влияние интересующего нас фактора не проявляется, хотя достаточно убедительные соображения указывают, что такое влияние должно быть. Причиной этого может быть большой внутригрупповой разброс, на фоне которого действие фактора становится малозаметным, либо наличие других факторов, сглаживающих это влияние. В первом случае дело можно поправить более тщательным отбором экспериментальных данных, во втором же надо вводить в рассмотрение дополнительные факторы. Один из методов борьбы с нежелательными воздействиями «мешающих» факторов основан на специальном планировании эксперимента. Цель такого планирования — свести к нулю влияние «мешающих» факторов. А именно, при фиксированном уровне фактора проводят испытания на такой группе, внутри которой действия мешающих факторов уравновешивают друг друга. Однако если информация о характере влияния

таких факторов отсутствует, такой подбор становится невозможным. Тогда можно поступить следующим образом. Из большого числа потенциально пригодных объектов случайным образом выбираются те, которые и образуют требуемую группу. Методики такого выбора можно найти в (33). Если же этими методами не удаётся избавиться от мешающего фактора, надо переходить к схемам многофакторного анализа. Желательно выделить один главный «мешающий» фактор и перейти к методам двухфакторного анализа, которые наиболее эффективны и просты. Мы не будем рассматривать эти методы, а остановимся лишь на задачах однофакторного анализа. Более подробно эти вопросы изложены в (103).

Для сравнения влияния факторов на результат необходимо собрать определённый статистический материал. Обычно его получают следующим образом: каждый из k способов обработки применяют несколько раз (не обязательно одно и то же число раз) к исследуемому объекту и регистрируют результаты. Наиболее удобным способом представления полученных данных является таблица.

Таблица 5

Таблица однофакторного анализа

	Способ обработки			
	1	2	...	k
Результаты измерений	x_{11}	x_{12}	...	x_{1k}
	
	$x_{n(1)1}$	$x_{n(2)2}$...	$x_{n(k)k}$

Здесь $n(i)$ — количество измерений, соответствующих i -тому способу обработки.

Первый вопрос, на который необходимо ответить при изучении экспериментальных данных следующий: Есть ли влияние, вы-

деленного фактора на полученный результат? Другими словами может ли различие наблюдаемых в опыте значений быть вызвано случайными причинами? То есть надлежит проверить справедливость следующего предположения: все данные таблицы принадлежат одному и тому же распределению, а, значит, отсутствуют эффекты обработки. Такое предположение обычно называют *нулевой гипотезой*.

Если гипотеза об отсутствии эффектов обработки отвергается, то проводится оценка действия этих эффектов или различий между ними и строятся доверительные интервалы для этих характеристик.

Ранговый однофакторный анализ

В том случае, когда мы ничего не знаем о законе распределения наблюдений, то для проверки нулевой гипотезы лучше использовать не сами значения наблюдений x_{ij} , а их ранги r_{ij} во всей совокупности данных. Используемые для этого критерии называются *ранговыми*. Их целесообразно использовать также для данных, измеренных в порядковой шкале.

Итак, проведём ранжирование данных таблицы 5 и перейдём к таблице 6.

Таблица 6

Ранговая таблица однофакторного анализа

	Способ обработки			
	1	2	...	k
Ранги результатов измерений	r_{11}	r_{12}	...	r_{1k}
	
	$r_{n(1)1}$	$r_{n(2)2}$...	$r_{n(k)k}$

Критерий Краскела-Уоллиса

Одним из наиболее универсальных ранговых критериев является критерий Краскела-Уоллиса. Он применяется, если нельзя ска-

зать ничего определённого об альтернативах к нашей нулевой гипотезе.

Итак, пусть мы имеем $N = n(1) + n(2) + \dots + n(k)$ наблюдений, ранги которых занесены в таблицу 5. Статистика Краскела-Уоллиса вычисляется по формуле

$$H = \frac{12}{N(N+1)} \sum_{j=1}^k n(j) \left(R_j - \frac{N+1}{2} \right)^2, \text{ где } R_j = \frac{1}{n(j)} \sum_{i=1}^{n(j)} r_{ij}.$$

Небольшие таблицы распределения данной статистики можно найти в сборниках статистических таблиц. При больших объёмах N , находящихся за пределами таблиц, распределение H приблизительно равно распределению хи-квадрат с $(k - 1)$ степенями свободы.

Итак, мы отвергаем нулевую гипотезу на данном уровне значимости, если наблюдаемое значение H превышает соответствующее значение для распределения хи-квадрат.

Если в исходной таблице 5 много совпадающих значений, то выводы будут носить приближённый характер. Для повышения точности можно использовать модифицированную форму статистики. Подробные сведения по этому поводу можно найти, например, в (105). Там же можно ознакомиться с другими ранговыми критериями, предназначенными для более частных ситуаций. Эти критерии являются более мощными в каждом конкретном случае, но, применённые некорректно, дают гораздо худшие результаты. Мы не будем касаться этих методик, чтобы не усложнять изложение. Отметим, что для целей педагогических исследований вполне достаточны универсальные методики.

Дисперсионный анализ

Если мы можем считать, что ошибки наблюдений носят случайный характер и имеют нормальное распределение с нулевым

средним и общей дисперсией, то можно использовать методы однофакторного дисперсионного анализа. Мы опишем один из таких методов.

F–тест для дисперсий

Данный метод проверяет справедливость гипотезы о равенстве средних значений при различных способах обработки и основан на сопоставлении двух оценок для дисперсии.

Осуществляется эта проверка следующим образом.

Предположим, что данные, полученные в эксперименте, измерены в количественной шкале и помещены в таблицу 5. Отметим ещё раз, что данный метод основан на предположении о том, что ошибки наблюдений имеют стандартное нормальное распределение. Сначала необходимо вычислить средние значения X_j по каждому столбцу и общее среднее X . Для этого воспользуемся формулами

$$X_j = \frac{1}{n(j)} \sum_{i=1}^{n(j)} x_{ij}, \text{ где } j = 1, \dots, k, \text{ и } X = \frac{1}{N} \sum_{j=1}^k \sum_{i=1}^{n(j)} x_{ij}.$$

Затем найдём две оценки σ^2 и σ_*^2 для дисперсии:

$$\sigma^2 = \frac{1}{N - k} \sum_{j=1}^k \sum_{i=1}^{n(j)} (x_{ij} - X_j)^2 \text{ и } \sigma_*^2 = \frac{1}{k - 1} \sum_{j=1}^k n(j)(X_j - X)^2.$$

И в заключение вычислим отношение

$$F = \frac{\sigma_*^2}{\sigma^2}.$$

Это отношение должно иметь F -распределение с $(k - 1, N - k)$ степенями свободы. Если полученное значение F превышает табличное значение, то гипотеза отвергается на выбранном уровне значимости.

Если гипотеза оказалась несовместимой с результатами наблюдений, то можно вычислить доверительные интервалы для средних значений в каждой группе наблюдений, т.е. для результатов действия каждого способа обработки. Для коэффициента доверия $(1 - \alpha)$ доверительный интервал в i -той группе равен

$$\left(X_j - \frac{\sigma_*}{\sqrt{n(j)}} t_{1-\alpha}, X_j + \frac{\sigma_*}{\sqrt{n(j)}} t_{1-\alpha} \right),$$

где $t_{1-\alpha}$ — квантиль уровня $(1 - \alpha)$ соответствующего распределению Стьюдента с $(N - k)$ степенями свободы.

Классификация типов данных (критерии согласия)

Во многих статистических задачах мы предполагаем, что некоторые случайные величины имеют заданное распределение (нормальное, экспоненциальное и т.д.) и, исходя из этого предположения, применяем те или иные методы дальнейшей обработки. При этом естественно возникает вопрос: насколько наши предположения отвечают экспериментальным данным? Для ответа на этот вопрос применяются специально разработанные критерии, которые называются *критериями согласия*.

Наиболее распространённые из них это: глазомерный метод проверки нормальности, критерий Колмогорова, критерии омега-квадрат и хи-квадрат, критерии асимметрии, эксцесса и т.д.

При проверке правильности нашего предположения о законе распределения следует выделять *простые* и *сложные* гипотезы:

- простая гипотеза указывает определённый закон распределения с конкретными параметрами, например, нормальное распределение с математическим ожиданием 0 и дисперсией 1 (стандартное нормальное распределение)

- сложная гипотеза указывает семейство распределений, например, нормальное распределение с неизвестными математическим ожиданием и дисперсией.

Начнём с критериев согласия, предназначенных для проверки только простых гипотез.

Глазомерный метод проверки нормальности

Пусть даны значения n наблюдений x_1, \dots, x_n , выполненных в количественной шкале, и мы хотим проверить гипотезу о нормальном характере закона её распределения. Для этого, прежде всего, упорядочим имеющиеся значения по возрастанию: $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$. Затем для всех i из отрезка $[1, n]$ найдём значение обратной функции Лапласа $\Phi^{-1}\left(\frac{2i-1}{2n}\right)$, воспользовавшись, например, таблицами квантилей функции Лапласа. Нанесём на координатной плоскости точки $\left(x_{(i)}, \Phi^{-1}\left(\frac{2i-1}{2n}\right)\right)$ и посмотрим, насколько хорошо эти точки «ложатся» на прямую линию. Чем ближе расположение этих точек к некоторой прямой линии, тем вероятнее наша гипотеза о нормальном распределении исходной выборки.

Следует заметить, что этот метод не даёт чёткого критерия отклонения гипотезы. Но, тем не менее, для первого приближения такой точности вполне достаточно.

Кроме того, глазомерный метод позволяет обнаружить «выбросы», т.е. значения, резко отличающиеся от остальных. Дополнительная проверка и корректировка таких значений значительно улучшает исходные данные, что приводит к большей точности в дальнейших вычислениях.

Проверим с помощью изложенного метода гипотезу о том, что время реакции на свет распределено по нормальному закону.

Данные для этого примера взяты из (25) и приведены в следующей таблице.

Таблица 7

Время реакции на свет в миллисекундах

Номер	1	2	3	4	5	6	7	8	9
Время	181	194	173	153	168	176	163	152	155
Номер	10	11	12	13	14	15	16	17	
Время	156	178	160	164	169	155	122	144	

Упорядочим наши данные по возрастанию и сведём их в таблицу, там же отметим значения обратной функции Лапласа вычисленные в соответствующих точках.

Таблица 8

**Время реакции на свет в миллисекундах
(упорядоченные данные)**

Номер (i)	(1)	(2)	(3)	(4)	(5)	(6)
Время	122	144	152	153	155	155
$\Phi^{-1}((2i-1)/34)$	-1,89	-1,35	-1,05	-0,82	-0,63	-0,46
Номер (i)	(7)	(8)	(9)	(10)	(11)	(12)
Время	156	160	163	164	168	169
$\Phi^{-1}((2i-1)/34)$	-0,30	-0,15	0	0,15	0,30	0,46
Номер (i)	(13)	(14)	(15)	(16)	(17)	
Время	173	176	178	181	194	
$\Phi^{-1}((2i-1)/34)$	0,63	0,82	1,05	1,35	1,89	

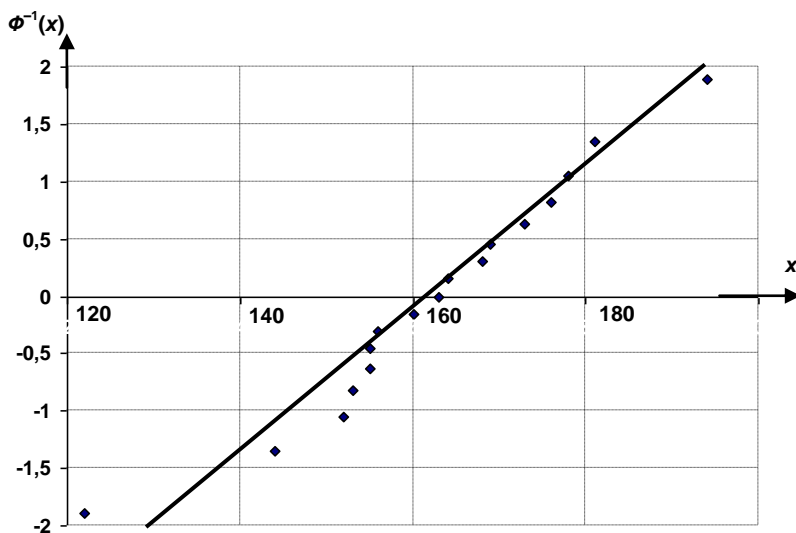


Рис. 3. Значения скачков эмпирической функции распределения

Глазомерный метод позволяет нам судить, насколько правдоподобна гипотеза о нормальности распределения имеющихся данных.

Критерий Колмогорова

Пусть мы имеем n значений x_1, \dots, x_n некоторой характеристики, измеренной в количественной шкале. Прежде всего, перейдём к *вариационному ряду* $x_{(1)}, \dots, x_{(n)}$, т.е. расположим значения в нашей выборке по возрастанию. Обозначим через $F(x)$ предполагаемую функцию распределения. Необходимым условием применимости критерия Колмогорова является непрерывность функции $F(x)$, т.е. использовать его можно лишь для непрерывных распределений.

Разберём сначала *случай* простой гипотезы. То есть, предположим, что нам известны все параметры предполагаемого распределения. Например, если речь идёт о нормальном распределении

$N(a, \sigma)$, то мы должны знать математическое ожидание a и дисперсию σ .

Для проверки гипотезы о совпадении эмпирической и гипотетической функций распределения надо вычислить значение статистики Колмогорова D_n , пользуясь следующей формулой:

$$D_n = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - F(x_{(i)}), F(x_{(i)}) - \frac{i-1}{n} \right\}.$$

Значения функции распределения $F(x)$ можно найти в соответствующих таблицах. Полученную величину D_n следует сравнить с найденными по таблицам критическими значениями. Гипотеза отвергается на выбранном уровне значимости, если полученное значение превосходит критическое значение для данного уровня значимости. Таблицы для статистики Колмогорова можно найти, например, в (16).

Перейдём теперь к рассмотрению сложной гипотезы, т.е. к предположению о том, что мы имеем дело с конкретным распределением, но с неизвестными параметрами. Например, мы хотим проверить, что наши данные подчиняются нормальному закону распределения $N(a, \sigma)$, но мы не знаем его параметры. Итак, пусть $F(x, \theta)$ — предполагаемая функция распределения, где θ — неизвестный нам параметр или группа параметров. Прежде всего, мы должны найти для θ приближённое значение, основываясь на имеющейся выборке x_1, \dots, x_n . Для этого можно использовать любые методы оценивания, например, метод наибольшего правдоподобия. Итак, пусть $\hat{\theta}$ — полученная оценка для нашего неизвестного параметра. Вычислим модифицированную статистику Колмогорова:

$$\hat{D}_n = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - F(x_{(i)}, \hat{\theta}), F(x_{(i)}, \hat{\theta}) - \frac{i-1}{n} \right\}.$$

Если полученное значение больше табличного, мы отвергаем гипотезу на данном уровне значимости. Следует обратить внимание на то, что при всём своём внешнем сходстве статистики \hat{D}_n и D_n имеют *разное* распределение, поэтому для них составлены *разные* таблицы. Таблицы для модифицированной статистики Колмогорова можно найти, например, в (44).

Критерий омега-квадрат

В тех же предположениях, что и критерий Колмогорова применяется критерий омега-квадрат. Пусть опять даны n значений x_1, \dots, x_n , измеряемой величины. Обозначим через $F(x)$ предполагаемую функцию распределения. Также как и для критерия Колмогорова требуется, чтобы функция распределения была непрерывна.

В случае простой гипотезы статистика омега-квадрат (ω^2) вычисляется по следующей формуле:

$$\omega^2 = \frac{1}{12n^2} + \frac{1}{n} \sum_{i=1}^n \left(F(x_{(i)}) - \frac{2i-1}{2n} \right)^2.$$

Если полученное значение превосходит найденное по таблице критическое значение, то гипотеза отвергается на данном уровне значимости. Таблицы для статистики омега-квадрат можно найти в (16).

В случае сложной гипотезы вычисляется модифицированная статистика омега-квадрат:

$$\hat{\omega}^2 = \frac{1}{12n^2} + \frac{1}{n} \sum_{i=1}^n \left(F(x_{(i)}, \hat{\theta}) - \frac{2i-1}{2n} \right)^2.$$

И в этом случае гипотеза отвергается на данном уровне значимости, если полученное значение статистики превосходит табличное значение. Как и для критерия Колмогорова, модифициро-

ванная статистика омега-квадрат имеет другое распределение, а значит надо использовать другие таблицы (см., например, 44).

Критерий хи-квадрат К. Пирсона

Данный критерий относится к независимым испытаниям с конечным числом исходов и предназначен для простых гипотез. Пусть мы имеем n независимых повторений некоего опыта, который заканчивается одним из r исходов. Критерий хи-квадрат предназначен для проверки гипотезы о том, что вероятности этих исходов равны p_1, \dots, p_r . Обозначим через m_1, m_2, \dots, m_n количества опытов, заканчивающихся, соответственно, первым, вторым, \dots , n -ым исходами и вычислим значение статистики хи-квадрат (χ^2) с $(r - 1)$ степенью свободы по следующей формуле:

$$\chi^2 = \sum_{i=1}^r \frac{(m_i - np_i)^2}{np_i}.$$

Если полученное значение превосходит табличное, то гипотеза должна быть отвергнута на выбранном уровне значимости. Значение распределения χ^2 можно найти в соответствующих таблицах.

Следует отметить, что критерий хи-квадрат достаточно точен только при больших значениях n . Его рекомендуется применять, если все значения произведений np_i не меньше 10.

Описанный критерий можно использовать и для произвольных выборок, после предварительной группировки данных. Для этого пространство наблюдений разбивают на конечное число непесекающихся областей, а затем для каждой области подсчитывают наблюдаемую частоту и гипотетическую вероятность. Но следует отметить, что при этом приходится преодолевать значительные трудности и действовать очень аккуратно. Поэтому в условиях другого эксперимента, например, для проверки гипотез о

нормальном характере распределения правильнее применять специально разработанные для таких случаев критерии. Учитывая вышесказанное, мы не будем останавливаться на методиках использования критерия хи-квадрат для проверки других гипотез. Отметим, однако, что использование современных статистических компьютерных программ позволяет автоматизировать этот процесс, а значит, без особенных усилий применять критерий хи-квадрат для проверки, например, нормальности распределения.

Критерий хи-квадрат Р. Фишера

Существует модификация критерия хи-квадрат для проверки сложных гипотез. В литературе она встречается под названием критерий хи-квадрат Р. Фишера. Воспользуемся обозначениями, сделанными при изложении критерия К. Пирсона (см. стр. 113). Если вероятности p_1, \dots, p_r известны нам с точностью до какого-либо неопределённого k -мерного параметра $\theta=(\theta_1, \theta_2, \dots, \theta_r)$, то мы имеем дело со сложной гипотезой. Для её проверки будем действовать уже знакомым нам способом. А именно, каким-либо методом получим оценку $\hat{\theta}$ для параметра θ . Затем вычислим статистику хи-квадрат с $(r - k - 1)$ степенями свободы:

$$\chi^2 = \sum_{i=1}^r \frac{(m_i - np_i(\hat{\theta}))^2}{np_i(\hat{\theta})}.$$

Если полученное значение χ^2 больше табличного, мы вынуждены отвергнуть нашу гипотезу на принятом уровне значимости.

В заключение сделаем несколько общих замечаний по поводу использования критериев согласия.

1. Теоретическое представление о законе распределения, которому должна подчиняться выборка, всегда имеет характер математической модели, т.е. является в какой-то мере приближён-

ным. Следовательно, точность статистических проверок должна быть сопоставима с точностью, которую мы ожидаем от математической модели в целом.

2. Критерии согласия очень чувствительны к «выбросам», т.е. к значениям резко отличающимся от остальных. А, значит, учитывая первое замечание, мы можем отбросить такие значения, чтобы выровнять распределение для дальнейших исследований.

Практическое применение методов факторного анализа и критериев согласия также не составит большого труда, но, тем не менее, принесёт несомненную пользу исследователю. Строгое доказательство выдвигаемых гипотез позволяет двигаться вперёд, сохраняя уверенность в том, что на предыдущих этапах не было допущено ошибок, способных перечеркнуть всю дальнейшую работу.

Выводы по третьей главе

1. Статистический подход к изучению каких-либо процессов или явлений состоит в мысленном расчленении происходящих изменений на две части — закономерную и случайную, и выявлении закономерной изменчивости на фоне случайной. Методы математической статистики позволяют оценивать параметры имеющихся закономерностей, проверять те или иные гипотезы об этих закономерностях и т.д. При этом статистика существенно опирается на теорию вероятностей, которая изучает закономерности, порождённые случайными событиями.
2. Основные методы статистической обработки данных основаны на свойствах нормального распределения. Использование этого распределения для приближённого описания случайных величин, сосредоточенных на определённом интервале, оправдано тем, что вероятность больших отклонений нормальной случайной величины от центра распределения практически равна нулю, и ею можно пренебречь.
3. Основная идея теории статистического вывода заключается в том, что некоторое множество объектов, выбранных случайным образом из общей совокупности объектов, имеет те же свойства, что и вся исходная совокупность. При этом, естественно, такое рассуждение имеет вероятностный характер. То есть всякое утверждение, полученное на основании конечного числа экспериментальных данных, верно или неверно лишь с некоторой вероятностью. Вероятность принятия гипотезы в случае, когда она неверна, называется уровнем значимости и чаще всего выбирается на уровне 0,05.
4. Среди статистических методов выделяют методы определения связи признаков, методы факторного анализа и критерии согла-

сия. К основным методам определения связи признаков относятся таблицы сопряжённости, статистика Фишера-Пирсона χ^2 -квадрат, коэффициенты корреляции Спирмена и Кендэлла, коэффициент корреляции Пирсона. Эти методы предназначены для установления наличия зависимости между несколькими признаками. К наиболее распространённым методам факторного анализа относятся критерий Краскела-Уоллиса и F-тест для дисперсий. Они предназначены для выявления влияния на конечный результат нескольких способов воздействия на объект (факторов). Критерии согласия предназначены для проверки гипотез о свойствах распределения некоторой случайной величины и носят, зачастую, вспомогательный характер. Наиболее распространёнными среди них являются критерий Колмогорова, критерии омега-квадрат и хи-квадрат.

Заключение

На современном этапе развития гуманитарного знания, к которому с полным правом относится и педагогическое знание, актуальной является задача систематизации данных, полученных в ходе эксперимента, их обобщение, количественный анализ и интерпретация. Важным средством решения этой задачи являются математические методы. Стало уже очевидным фактом, что в принципе нет «нематематизируемых» дисциплин даже в области гуманитарного знания. Речь может идти только о степени математизации конкретного педагогического исследования.

В настоящей работе мы определили теоретико-методологические основы организации, планирования и оценки результатов педагогического эксперимента с позиций квалиметрического подхода. Квалиметрический подход мы рассматриваем как теоретико-методологическую стратегию исследования, ориентированную на количественное описание качества объектов.

Исходя из такого понимания квалиметрического подхода, мы определили теоретические основы создания инструментального обеспечения организации и оценки качества педагогического эксперимента на основе принципов информативности, интегративности, оптимальности, точности, доказательности, технологичности, распределённости, унифицированности, доступности.

Мы доказали, что повышение эффективности качества педагогического исследования обеспечивается, в том числе, и созданием технологий организации и оценивания результатов педагогического эксперимента на основе квалиметрического подхода.

Преимущества квалиметрического подхода и основанных на нём технологий заключаются в экономичности, технологичности, достоверности и объективности.

Усиление эффективности квалиметрического подхода к педагогическому эксперименту возможно за счёт использования ЭВМ.

Внедрение ЭВМ в сферу деятельности педагогов-исследователей, во-первых, расширяет возможности решения актуальных задач в области теории и практики педагогики, во-вторых, освобождает педагога-исследователя от непроизводительных затрат времени на обработку полученных данных, в-третьих, позволяет использовать более мощный математический аппарат; в-четвёртых, проводить анализ параллельно несколькими методами, и, наконец, хранить полученную информацию и по необходимости обращаться к ней.

Таким образом, настоящее исследование может явиться базой для проведения дальнейших исследований по разработке инструментального обеспечения организации и оценивания качества педагогического исследования и особенно его экспериментальной составляющей с использованием ЭВМ.

Понятийный аппарат исследования

Альтернатива (альтернативное распределение) — распределение которое может возникнуть в случае нарушения проверяемой гипотезы.

Анализ — (от греч. analysis — разложение) — расчленение (мысленное или реальное) объекта на элементы. Анализ неразрывно связан с синтезом (соединение элементов в единое целое) (23; 89).

Биномиальное распределение — это распределение вероятностей числа наступлений некоторого события («удачи») в серии из определённого числа независимых наблюдений, выполняемых в одинаковых условиях, если вероятность наступления этого события в каждом из испытаний одна и та же.

Вариационный ряд — выборка, перенумерованная в порядке возрастания.

Величина (матем.) — «1) обобщение конкретных понятий: длины, площади, веса и т.п. Выбрав одну из величин данного рода за единицу измерения, можно выразить числом отношение любой другой величины того же рода к единице измерения; 2) в более общем смысле скалярной величиной, или скаляром, называется объект, полностью характеризующийся заданием одного числа. Обобщением скалярных величин являются векторные величины, тензорные величины» (89, с. 205).

Вероятность — «числовая характеристика степени возможности наступления какого-либо события в тех или иных определённых, могущих повторяться неограниченное число раз условиях» (63, с. 118).

Выборка — понятие математической статистики, объединяющее результаты каких-либо однородных наблюдений. Выборкой в широком смысле называется конечная совокупность результатов наблюдений x_1, x_2, \dots, x_n , представляющих собой независимые одинаково распределённые случайные величины.

Выброс — грубое (ошибочное) наблюдение, сильно отличающееся от основной массы наблюдений.

Дисперсия — одна из характеристик распределения вероятностей случайной величины, наиболее употребительная мера рассеяния её значений.

Измерение — «совокупность действий, выполняемых при помощи средств измерений с целью нахождения числового значения измеряемой величины в принятых единицах измерения» (89, с. 479).

Качество — «философская категория, выражающая существенную определённую объекта, благодаря которой он является именно этим, а не иным. Качество — объективная и всеобщая характеристика объектов, обнаруживающаяся в совокупности их свойств» (89, с. 561).

Кластерный анализ — «математическая процедура многомерного анализа, позволяющая на основе множества показателей, характеризующих ряд объектов (например, испытуемых), сгруппировать их в классы (кластеры) таким образом, чтобы объекты, входящие в один класс, были более однородными, сходными по сравнению с объектами, входящими в другие классы. На основе численно выраженных параметров объектов вычисляются расстояния между ними, которые могут выражаться как в евклидовой метрике (наиболее употребляемой), так и в других метриках. Метод кластерного

анализа нашёл широкое применение в психолингвистике» (80, с. 162).

Количество — «философская категория, выражающая внешнюю определённую объект: его величину, объём, степень развития свойств и т.д.; изменение количественной определённости объекта, достигнув определённой меры, ведёт к изменению качества» (89, с. 601).

Контент-анализ — (от англ. contents — содержание) — метод анализа и оценки информации (содержащейся в документах, видеозаписях, радиопередачах, интервью и т.д.) путём выделения в формализованном виде смысловых единиц информации и замера частоты упоминания этих единиц в выборочной совокупности.

Корреляция — (в матем. статистике) — «вероятностная или статистическая зависимость. В отличие от функциональной зависимости корреляция возникает тогда, когда зависимость одного из признаков от другого осложняется наличием ряда случайных факторов» (89, с. 634).

Критерий — (от греч. kriterium — средство для суждения) — «признак, на основании которого производится оценка, определение или классификация чего-либо, мерило оценки» (89, с. 656).

Критерий оптимальности — система целей эксперимента, состоящая в улучшении некоторых параметров эксперимента.

Математическая статистика — как научная дисциплина, представляет собой «раздел математики, посвящённый математическим методам систематизации, обработки и использования статистических данных для научных и практических выводов» (63, с. 344). При этом под статистическими данными понимается любая числовая информация, характеризующая некоторую совокупность объектов, обладающих теми или иными общими признаками.

Математическое ожидание — это некоторое «осреднение» случайной величины с учётом вероятности получения отдельных значений. В случае конечного числа возможных исходов испытаний математическое ожидание равно среднему арифметическому значений случайной величины.

Мера — «философская категория, выражающая диалектическое единство качества и количества объекта; указывает предел, за которым изменение количества влечёт за собой изменение качества объекта и наоборот» (89, с. 791).

Моделирование — «исследование каких-либо явлений, процессов или систем объектов путём построения и изучения их моделей; использование моделей для определения или уточнения характеристик и рационализации способов построения вновь конструируемых объектов» (89, с. 819).

Нормальное (гауссовское) распределение — это распределение вероятностей в случае, когда «значение непрерывной случайной величины формируется под воздействием очень большого числа независимых случайных факторов, причём сила воздействия каждого отдельного фактора мала и не может превалировать среди остальных, а характер воздействия — аддитивный (т.е. при воздействии случайного фактора F на величину a получается величина $a + \Delta F$, где случайная «добавка» ΔF мала и равновероятна по знаку)» (5, с. 169).

Отклик — переменная, характеризующая состояние объекта, изменяющееся под воздействием некоторых *факторов*.

Плотность вероятности — функция, которая описывает закон распределения вероятности и может быть определена как производная функции распределения.

Процедура — (франц. *procedure*, от лат. *procedo* — продвигаюсь) — «установленный порядок ведения, рассмотрения какого-либо дела» (89, с. 1075).

Процесс — (от лат. *processus* — продвижение) — «1) последовательная смена явлений, состояний в развитии чего-нибудь; 2) совокупность последовательных действий для достижения какого-либо результата» (89, с. 1075).

Ранг наблюдения — номер, который получит это наблюдение после упорядочения всех данных по возрастанию (или убыванию) силы проявления признака. Процедура перехода от совокупности наблюдений к последовательности их рангов называется *ранжированием*.

Рандомизация — «статистическая процедура, в которой решение принимается случайным образом» (63, с. 517). Применяется для усреднения влияния случайных факторов.

Регрессионный анализ — «раздел математической статистики, объединяющий практические методы исследования, регрессионной зависимости величин по статистическим данным» (89, с. 1112).

Случайная величина — некоторая переменная величина, принимающая в зависимости от случая те или иные числовые значения. Типичным примером случайной величины является число очков, выпавших при бросании игральной кости. Обычно выделяют *дискретные* и *непрерывные* случайные величины.

Статистическая гипотеза — «предположение о вероятностных закономерностях, которым подчиняется изучаемое случайное явление. Как правило, статистическая гипотеза определяет значения параметров распределения вероятностей или непосредственно его вид и свойства» (63, с. 559).

Уровень значимости — вероятность отвержения выдвинутой гипотезы в случае, когда она на самом деле верна.

Фактор — это переменная воздействия на объект, влиянию которой приписывается переход объекта из одного состояния в другое.

Функция отклика — закон, описывающий связь между значениями факторов и значениями откликов.

Функция распределения случайной величины ξ — эта функция $F(x)=P(\xi < x)$, т.е. значением функции распределения является вероятность того, что случайная величина ξ принимает значение меньше x .

Литература

1. Аванесов В. С. Тесты в социологическом исследовании. – М.: Наука, 1982. – 199 с.
2. Авчукова Р. Э. Некоторые модельные представления при исследовании системы управления процессом обучения. Автореф. дис. ... канд. пед. наук. – Рига, 1974. – 21 с.
3. Адлер Ю. П. Предпланирование эксперимента. – М.: Знание, 1978. – 72 с.
4. Азгальдов Г. Г. Теория и практика оценки качества товаров (основы квалиметрии). – М.: Экономика, 1982. – 256 с.
5. Айвазян С. А., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: Основы моделирования и первичная обработка данных. Справочное изд. Под ред. С. А. Айвазяна. – М.: Финансы и статистика, 1983. – 471 с.
6. Акинфиева Н. В. Квалиметрический инструментарий педагогических исследований // Педагогика, 1998, № 4. – С. 30–35.
7. Аптон Г. Анализ таблиц сопряжённости. – М.: Финансы и статистика, 1982. – 144 с.
8. Аристер Н. И., Загузов Н. И. Процедура подготовки и защиты диссертаций: АОЗТ «Икар», 1995. – 200 с.
9. Архангельский С. И. Лекции по теории обучения в высшей школе. – М.: Высшая школа, 1974. – 384 с.
10. Архангельский С. И., Михеев В. И., Перельцвайг Ю. М. Вопросы измерения, анализа и оценки результатов в практике педагогических исследований. М.: Знание, 1975. – 42 с.
11. Аткинсон Р. Человеческая память и процесс обучения. – М.: Прогресс, 1980. – 528 с.

12. Аткинсон Р., Бауэр Г., Кроттерс Э. Введение в математическую теорию обучения. – М.: Мир, 1969. – 486 с.
13. Бабанский Ю. К. Проблемы повышения эффективности педагогических исследований: Дидактический аспект. – М.: Педагогика, 1982. – 182 с.
14. Беспалько В. П. Основы теории педагогических систем. Воронеж: Изд-во Воронежск. ун-та, 1983. – 175 с.
15. Битинас Б. П. Многомерный анализ в педагогике и педагогической психологии. Вильнюс, 1971. – 374 с.
16. Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. – М.: Наука, 1983. – 416 с.
17. Быков В. В. Научный эксперимент. – М.: Наука, 1989. – 176 с.
18. Введение в научное исследование по педагогике: Учебное пособие для студентов пед. ин-тов / Ю. К. Бабанский, В. И. Журавлёв, В. К. Розов и др.: Под ред. В. И. Журавлёва. – М.: Просвещение, 1988. – 239 с.
19. Внедрение достижений педагогики в практику школы / Под ред. В. Е. Гмурмана. – М.: Педагогика, 1981. – 144 с.
20. Вопросы прогнозирования и планирования развития личности. – Красноярск, 1985.
21. Гегель. Сочинение. Т. 1. М. – Л. 1930.
22. Гершунский Б. С. Педагогическая прогностика. – Киев, 1986.
23. Гильманов Р. А. Проблемы конструктивной дидактики. – Казань: Изд-во Казан. ун-та, 1994. – 147 с.
24. Гласс Дж., Стэнли Дж. Статистические методы в педагогике и психологии. – М.: Прогресс, 1976. – 495 с.
25. Готтсданкер Р. Основы психологического эксперимента. – М.: МГУ, 1982. – 463 с.

26. Грабарь М. И., Краснянская К. А. Применение математической статистики в педагогических исследованиях. Непараметрические методы. – М.: Педагогика, 1977. – 136 с.
27. Громько Ю. В., Давыдов В. В. Концепция экспериментальной работы в сфере образования // Педагогика. 1994, № 6. – С. 31-37.
28. Гузеев В. В. Системные обоснования образовательной технологии. – М.: Знание, 1975. – 64 с.
29. Давыдов В. В., Варданян А. У. Учебная деятельность и моделирование. – Ереван, 1981. – 213 с.
30. Данилевский Ю. Г., Петухов И. А., Шибанов В. С. Информационная технология в промышленности. – Л.: Машиностроение, 1988. – 283 с.
31. Данилюк А. Я. Проблема эмпирического и теоретического в отечественной педагогике // Педагогика, 1997, № 5. – С. 42–46.
32. Демиденко Е. З. Линейная и нелинейная регрессия. – М.: Финансы и статистика, 1981. – 302 с.
33. Джонсон Н., Лион Ф. Статистика и планирование эксперимента в технике и науке. – М.: Мир. Т. 1, 1980, – 610 с., Т. 2, 1981, – 520 с.
34. Дрейпер Н., Смит. Г. Прикладной дисперсионный анализ: В 2 кн., Кн. 1. – М.: Финансы и статистика, 1986. – 366 с. Кн. 2. – М.: Финансы и статистика, 1987. – 351 с.
35. Дуранов И. М., Дуранов М. Е., Жернов В. И., Лешер О. В. Педагогика воспитания и развития личности учащихся. – Магнитогорск, МГПИ, 1996. – 315 с.
36. Журавлёв В. И. Педагогика в системе наук о человеке. – М.: Педагогика, 1990. – 168 с.

37. Загвязинский В. И. Методология и методика дидактического исследования. – М.: Педагогика, 1982. –160 с.
38. Загузов Н. И. Теоретико-методологические аспекты подготовки и защиты кандидатской диссертации по педагогическим наукам: Метод. рекомендации. Урал. гос. пед. ун-т. Екатеринбург, 1995. – 34 с.
39. Здравомыслов А. Г. Методология и процедура социологических исследований. – М.: Мысль, 1969. – 205 с.
40. Ительсон Л. Б. Математические и кибернетические методы в педагогике. – М.: Просвещение, 1964. – 248 с.
41. Касимов Р. Я. Модель как средство научной организации обучения. Дис. ... канд. пед. наук. – М.: МГПИ им. В. И. Ленина, 1973.
42. Кедровский О. И., Соловей Л. А. Алгоритмичность практики, мышления, творчества. – Киев: Вища школа, 1980. –184 с.
43. Кендэлл М. Ранговые корреляции. – М.: Статистика, 1975. – 212 с.
44. Кендэлл М., Стьюарт А. Статистические выводы и связи. – М.: Наука, 1973. – 899 с.
45. Китаев Н. Н. Групповые экспертные оценки. – М.: Знание, 1975. – 64 с.
46. Климова Т. Е. Подготовка будущего учителя к профессионально-творческой самообразовательной деятельности: Автореф. дис. ... канд. пед. наук. – Челябинск, 1995. – 22 с.
47. Конаржевский Ю. А. Педагогический анализ учебно-воспитательного процесса как фактор повышения эффективности управления общеобразовательной школы: Дис. ... д-ра пед. наук. – Челябинск, 1980. – 492 с.

48. Концептуальные основы экспертизы качества образования. – Екатеринбург, 1992. – 33 с.
49. Кочетов А. И. Педагогическое исследование: Учебное пособие. – Рязань, 1975. – 178 с.
50. Краевский В. В. Педагогическая теория: Что это такое? Зачем она нужна? Как она делается? – Волгоград: Перемена, 1996. – 86 с.
51. Кузьмина Н. В. Методы системного педагогического исследования. Л.: Изд-во ЛГУ, 1980. – 172 с.
52. Куль И. Г. Модели в учебном процессе // Материалы к Всесоюзной конференции по ПО и применению ТС в учебном процессе. Февраль 1966. – Тарту, 1966.
53. Кулюткин Ю. Н. Эвристические методы в структуре решений. – М.: Педагогика, 1970. – 231 с.
54. Кураев В. И., Лазарев Ф. В. Точность, истина и рост знаний. – М., 1988.
55. Кутьев В. О. Педагогическая прогностика: научные подходы и мифы // Педагогика, 1995, № 3. – С. 7–14.
56. Ланда Л. Н. Алгоритмизация в обучении. – М.: Просвещение, 1966. – 523 с.
57. Леонтьев А. Н. Проблемы развития психики. 3-е изд. М.: МГУ, 1972. – 575 с.
58. Ликеш И., Ляга И. Основные таблицы математической статистики. – М.: Финансы и статистика, 1985. – 356 с.
59. Маркарян Э. С. Теория культуры и современная наука. – М.: Мысль, 1983. – 284 с.
60. Марков М. Технология и эффективность социального управления. – М.: Прогресс, 1982. – 267 с.

61. Марченко Е. К. Методы квалиметрии в педагогике. – М.: Знание, 1979. – 33 с.
62. Марьенко Т. Ю. Подготовка будущих учителей к работе в педагогических классах в условиях многопрофильной гимназии: Автореф. дис. ... канд. пед. наук. – Челябинск, 1997. – 24 с.
63. Математический энциклопедический словарь. – М.: Советская энциклопедия, 1988. – 847 с.
64. Менделеев Д. И. Основы химии. Соч.: В 2 т. – Т. 1. – М.: Наука, 1934.
65. Методы педагогических исследований / Под ред. А. И. Пискунова, Г. В. Воробьёва. – М.: Педагогика, 1979. – 256 с.
66. Мизинцев В. П. Применение моделей и методов моделирования в дидактике. М.: Знание, 1977. – 52 с.
67. Михеев В. И. Моделирование и методы измерений в педагогике: Научн.-метод. пособие. – М.: Высшая школа, 1987. – 200 с.
68. Мюллер П., Нойман П., Шторм Р. Таблицы по математической статистике. – М.: Финансы и статистика, 1982. – 278 с.
69. Найн А. Я. Педагогические инновации и научный эксперимент // Педагогика, 1966, № 5. – С. 10–15.
70. Найн А. Я. Технология работы над кандидатской диссертацией по педагогике. – Челябинск: УралГАФК, 1996. – 144 с.
71. Налимов В. В., Голикова Т. И. Логические основания планирования эксперимента. – М.: Металлургия, 1976. – 128 с.
72. Никитина Е. Ю. Педагогические условия подготовки будущего учителя к осуществлению дифференцированного обучения в школе: Автореф. дис. ... канд. пед. наук. – Челябинск, 1995. – 24 с.

73. Овакимян Ю. О. Формирование приёмов самоконтроля в процессе обучения. Дис. ... канд. пед. наук. – М.: МГПИ им. В. И. Ленина, 1969.
74. Оконь В. Введение в общую дидактику: Пер. с польск. – М.: Высшая шк., 1990. – 382 с.
75. Осипов Б. В., Мировская Е. А. Математические методы и ЭВМ в стандартизации и управлении качеством. – М.: Изд-во стандартов. 1990. – 168 с.
76. Осипов Г. В., Андреев Э. П. Методы измерений в социологии. – М.: Наука, 1974.
77. Платонов К. К. Краткий словарь системы психологических понятий. – М.: Высшая школа, 1981. – 175 с.
78. Полонский В. М. Оценка качества научно-педагогических исследований. – М.: Педагогика, 1987. – 144 с.
79. Проблемы методологии и педагогики исследований / Под. ред. М. А. Данилова, Н. И. Болдырева. – М., 1971.
80. Психология. Словарь / Под общ. ред. А. В. Петровского, М. Г. Ярошевского. – М.: Политиздат, 1990.
81. Рабочая книга социолога / Отв. ред. Г. В. Осипов. – М.: Наука, 1983. – 477 с.
82. Ракитов А. И. Исследование научное. БСЭ. – 3-е изд. – М.: Советская энциклопедия, 1972. – Т. 10. – С. 542.
83. Репин С. В., Шеин С. А. Математические методы обработки статистической информации с помощью ЭВМ. – Минск: Университетское, 1990. – 128 с.
84. Розенберг Н. М. Проблемы измерений в дидактике. Киев: Вища школа, 1979. – 175 с.

85. Сагатовский В. Н. Системная деятельность и её философское осмысление / Системные исследования. М.: Наука, 1981. С. 52-68.
86. Сериков Г. Н. Образование: аспекты системного отражения. – Курган: Изд-во «Зауралье», 1997. – 464 с.
87. Скалкова Я. и др. Методология и методы педагогического исследования: Пер. с чешск. – М.: Педагогика, 1989. – 224 с.
88. Скаткин М. Н. Методология и методика педагогических исследований. – М.: Педагогика, 1986. – 150 с.
89. Советский энциклопедический словарь. – М.: Советская энциклопедия, 1986. – 1599 с.
90. Сохор А. М. Логическая структура учебного материала. – М.: Педагогика, 1974. – 192 с.
91. Спиркин А. Г. Основы философии: Учеб. пособие для вузов. – М.: Политиздат, 1988. – 592 с.
92. Стефанов Н. Общественные науки и социальная технология. – М.: Прогресс, 1976. – 251 с.
93. Стреляу Я. Роль темперамента в психическом развитии. – М.: Прогресс, 1982. – 231 с.
94. Теория и практика педагогического эксперимента / Под ред. А. И. Пискунова, Г. В. Воробьёва. – М.: Педагогика, 1979. – 208 с.
95. Тимофеев И. С. Методологическое значение категорий «качество» и «количество». – М.: Наука, 1972. – 216 с.
96. Тулькибаева Н. Н., Яковлева Н. М., Большакова З. М. Теоретико-методологическая концепция экспертизы качества образования на основе стандартизации: Монография. – Челябинск: Изд-во ЧГПУ «Факел», 1998. – 143 с.

97. Турбович Л. Т. Информационно-семантическая модель обучения. – Л.: ЛГУ, 1970.
98. Тюрин Ю. Н., Макаров А. А. Статистический анализ данных на компьютере / Под ред. Б. Э. Фигурнова. – М.: ИНФРА-М, 1998. – 528 с.
99. Уёмов А. И. Логические основы метода моделирования. – М.: Мысль, 1971. – 312 с.
100. Филиппова И. А. Методика социально-педагогического исследования: Учебное пособие. – Челябинск, Челяб. гос. пед. ин-т, 1989. – 81 с.
101. Философский словарь / Под ред. И. Т. Фролова. – 5-е изд. – М.: Политиздат, 1987. – 590 с.
102. Фридман Л. М. Использование моделирования в обучении // Вестник Челябинского государственного педагогического университета. Сер. 2. Педагогика. Психология. Методика преподавания. – Челябинск: Изд-во «Факел», 1995, № 1. – С. 88-93.
103. Хартман Г. Современный факторный анализ. – М.: Статистика, 1972.
104. Хикс Ч. Основные принципы планирования эксперимента. – М.: Мир, 1967. – 406 с.
105. Холлендер М., Вульф Д. Непараметрические методы статистики. – М.: Финансы и статистика, 1983. – 518 с.
106. Христева А. В. Подготовка будущего учителя к анализу и самоанализу профессионально-педагогической деятельности: Автореф. дис. ... канд. пед. наук. – Челябинск, 1996. – 24 с.
107. Черепанов В. С. Экспертные методы в педагогике: Учебное пособие. Пермь: ПГПИ, 1988. – 84 с.
108. Черепанов В. С. Экспертные оценки в педагогических исследованиях. – М.: Педагогика, 1989. – 152 с.

109. Черкасов В. А., Тротт К. Г., Черкасова Э. С. Внедрение достижений педагогической науки в практику образовательных учреждений: функции проблемы (аспект оптимизации образования). – Челябинск: ЧелГУ, 1995. – 184 с.
110. Штофф В. А. Моделирование в философии. – М., 1966.
111. Штофф В. А. Проблемы методологии научного познания. М., 1978.
112. Щедровицкий Г. П. Система педагогических исследований // Педагогика и логика. – М.: Касталь, 1993. – С. 3–200.
113. Щукина Г. И. К вопросу о критериях ценности педагогических исследований // Сов. педагогика, № 9, 1967. – С. 11–16.
114. Юдин Э. Г. Системный подход и принцип деятельности: Методологические проблемы современной науки. – М.: Наука, 1978. – 391 с.
115. Яковлева Н. М. Подготовка студентов к творческой воспитательной деятельности. – Челябинск: ЧГПИ, 1992. – 128 с.
116. Яковлева Н. М. Теория и практика подготовки будущего учителя к творческому решению воспитательных задач: Дис. ... д-ра пед. наук. – Челябинск, 1992. – 403 с.
117. Яноши Л. Теория и практика обработки результатов измерений. – М.: Мир, 1965.
118. Campbell D. T., Stanley J. C. Experimental and quasi-experimental designs for research on teaching. W.: Gage N. L. Handbook of Research on Teaching. Chicago, 1967.
119. Kiefer J. – Ann. of Mathematical Statistics, 1958. V. 29. № 3. P. 675–699.
120. Phillips B. S. Social researches. Strategy and Tactics. N.-Y., 1976.

Оглавление

ВВЕДЕНИЕ	3
ГЛАВА I. ТЕОРЕТИКО-МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ ОРГАНИЗАЦИИ И ПРОВЕДЕНИЯ ПЕДАГОГИЧЕСКОГО ЭКСПЕРИМЕНТА	
§ 1. Сущность и особенности педагогического исследования....	4
§ 2. Квалиметрический подход к организации и оценке педагогического исследования	19
§ 3. Теория измерений как основа квалиметрического подхода	31
Выводы по первой главе.....	44
ГЛАВА II. ТЕХНОЛОГИЯ ПЛАНИРОВАНИЯ И ОРГАНИЗАЦИИ ПЕДАГОГИЧЕСКОГО ЭКСПЕРИМЕНТА С ПОЗИЦИЙ КВАЛИМЕТРИЧЕСКОГО ПОДХОДА	
§ 1. Технология моделирования в педагогическом исследовании	46
§ 2. Технология планирования и организации эксперимента ..	58
Выводы по второй главе	73
ГЛАВА III. ТЕХНОЛОГИЯ КВАЛИМЕТРИЧЕСКОЙ ОЦЕНКИ РЕЗУЛЬТАТОВ ПЕДАГОГИЧЕСКОГО ЭКСПЕРИМЕНТА	
§ 1. Основные понятия математической статистики	76
§ 2. Методы определения связи признаков.....	92
§ 3. Факторный анализ.....	102
Выводы по третьей главе.....	116
ЗАКЛЮЧЕНИЕ	118
ПОНЯТИЙНЫЙ АППАРАТ ИССЛЕДОВАНИЯ	120
ЛИТЕРАТУРА	126
ОГЛАВЛЕНИЕ	136

Научное издание

Евгений Владимирович Яковлев

Педагогический эксперимент: квалиметрический аспект

Редактор Ю. В. Тихонова

ISBN 5–85716–213–0

Лицензия ЛР № 040277 от 17 августа 1997 г.

Формат 60x84/16

Заказ 267

Тираж 500 экз.

Объём 7 уч. изд. л.

454080 г. Челябинск, пр. им. В. И. Ленина, 69

Издательство ЧГПУ

Отпечатано на ризографе типографии ЧГПУ

454080 г. Челябинск, пр. им. В. И. Ленина, 69