

*Елена Сергеевна ОХОТНИКОВА —  
ассистент кафедры программного обеспечения*

УДК 37.001.5

## **МАТЕМАТИЧЕСКИЕ МЕТОДЫ И МОДЕЛИ В СОВРЕМЕННЫХ СИСТЕМАХ УПРАВЛЕНИЯ ОБУЧЕНИЕМ**

*АННОТАЦИЯ. В статье рассматриваются основные возможности систем управления обучением, математические модели, используемые в таких системах. Ставится вопрос о применимости e-learning систем в образовательном процессе вуза.*

*The article subject is learning management systems (LMS) potentialities and math models. LMS usage in universities also is discussed.*

Развитие европейских университетов проходит в последние годы под знаком так называемого Болонского процесса, основной целью которого стало преодоление кризиса существующей системы высшего образования. В сентябре 2003 г. Россия официально присоединилась к Болонскому процессу, взяв на себя обязательства соответствовать европейским стандартам и принципам в сфере высшего образования. Чем же вызвана необходимость реформ в образовании? Современное общество характеризуется, прежде всего, быстрым развитием, мобильностью, большим количеством информации, поэтому специалисты должны отличаться способностью быстро воспринимать и применять на практике новые знания. В противоречие с этим вступает традиционная система образования, которая дает в первую очередь фундаментальные, энциклопедические знания и не вполне нацелена на подготовку специалистов, отвечающих требованиям рынка труда.

В настоящее время большую ценность имеет специалист, способный при необходимости получить дополнительную квалификацию (причем без отрыва от производственного процесса), поэтому очень востребована дистанционная форма обучения. Директор Российского государственного института открытого образования В. Солдаткин считает, что «...большинство отечественных вузов работают так же, как и полвека назад, то есть без учета радикальных изменений жизни людей и возможностей новых информационных технологий... В сложившихся условиях вопросы развития дистанционного образования становятся наиболее актуальными. Одним из важных профессиональных качеств преподавателя становится способность разрабатывать и использовать открытые методы преподавания...» [2].

В качестве одного из вариантов решения этой проблемы рассматривается интенсивное применение информационных технологий в образовательном процессе. В последние годы в зарубежных странах часто используется термин «e-learning» (электронное обучение), означающий процесс обучения в электронной форме через сеть Интернет или Интранет с использованием систем управления обучением. Понятие «электронное обучение» сегодня является расширением термина «дистанционное обучение». Электронное обучение — более широкое понятие, включающее разные формы и способы обучения на основе информационных и коммуникационных технологий.

Технически система электронного обучения (e-learning решение) строится на основе целого ряда программных продуктов [3]:

- средство разработки учебного контента (Authoring Tool);
- система управления обучением (Learning Management System);
- система обмена информацией;
- система управления контентом/учебным порталом.

С помощью средств разработки учебного контента создаются учебные материалы и тестовые задания, которые затем помещаются в базу данных системы управления обучением. Через нее слушатели получают доступ к учебным материалам. Учебный контент может быть пассивным (HTML-страницы, тексты) и интерактивным, с элементами анимации и голосовым сопровождением. Для создания пассивного контента можно использовать стандартные редакторы, такие как Microsoft Word. Интерактивный контент создается с помощью специальных программных продуктов.

Система управления обучением реализует следующие функции [3]:

- учет слушателей, персонализация, разграничение прав доступа;
- управление процессом обучения, учет результатов обучения и тестирования;
- интеграция с механизмами синхронного и асинхронного общения;
- подготовка аналитической отчетности;
- интеграция с внешними информационными системами.

Система обмена информацией позволяет учащимся, преподавателям, экспертам и другим участникам учебного процесса обмениваться информацией между собой как в режиме реального времени (голосовой и компьютерный чат, видеоконференции, совместное использование программных продуктов, виртуальная аудитория), так и асинхронно (форумы, доски объявлений, электронная почта).

Вэб-интерфейс системы управления обучением, как правило, строится на основе средств управления контентом.

Популярность e-learning формы обучения объясняется тем, что она имеет ряд преимуществ по сравнению с традиционной. При электронной форме обучения не требуется обязательного посещения лекций и практических занятий. Кроме того, у людей появляется возможность учиться в удобное для них время и в удобном темпе. При использовании традиционных форм заочного обучения многие сотрудники организаций и фирм не могут присутствовать на занятиях из-за командировок, встреч с клиентами или других обстоятельств. Но обучение является обязательным, оно необходимо для успешной работы. Еще одним достоинством электронного обучения является то, что оно позволяет быстрее вывести знания на рынок. Известно, что актуальность знаний уменьшается со временем, и именно такой способ помогает в кратчайшие сроки получить к ним доступ.

В настоящее время электронное обучение, как правило, применяется для получения второго высшего образования, дополнительного образования или для переподготовки специалистов. Как в мировой практике, так и в российской, e-learning используется именно как дополнение к очному образованию.

Ряд авторов делает вывод, что применение электронного обучения может решить ряд основных проблем (при условии разработки и использования качественного контента) в сфере послевузовского дополнительного образования [3]. В учебных процессах очного, заочного и дополнительного образования существуют значительные отличия, поэтому перед внедрением в процесс очного обучения e-learning систем следует провести тщательный анализ возможностей, предоставляемых данными системами, поскольку даже на первый взгляд заметны некоторые проблемы. Например, предоставление учебно-методических

материалов в электронном виде, с одной стороны, дает студентам больше времени на занятиях для восприятия нового материала, а также возможность неоднократного обращения к нему, но с другой стороны — может повлечь за собой снижение посещаемости лекционных и практических занятий.

Другой трудностью применения систем управления обучением является необходимость формирования современных дидактических материалов, на разработку которых затрачивается большое количество ресурсов (временных, материальных и т.д.). При отсутствии у преподавателей опыта подготовки материалов в требуемой форме составление качественного контента может стать настоящей проблемой, так как простой перенос учебников в электронную форму не сможет обеспечить реализацию потенциала, заложенного в e-learning системах.

Можно сделать вывод, что при попытке внедрения систем управления обучением в образовательный процесс вуза следует искать разумное соотношение между требованиями, которые выдвигают особенности учебного процесса, и возможностями, которые предоставляют современные системы управления обучением, то есть следует принять интересные и полезные инновации в адаптированном виде и отказаться от неподходящих.

Поскольку при использовании систем электронного обучения увеличивается доля самостоятельной работы обучаемого, следовательно, повышается роль контроля в образовательном процессе. Большинство систем управления обучением предоставляют возможность оценки знаний и успеваемости студентов в виде тестирования. Контрольно-измерительные мероприятия в форме тестов обеспечивают реализацию всех выделяемых функций контроля при соответствующем уровне качества тестовых материалов [5].

В настоящее время тестовые формы контроля знаний получили достаточно широкое применение в образовательном процессе учебных заведений нашей страны, но довольно часто встречается и негативное к ним отношение.

В первую очередь, тестовая форма контрольно-измерительных мероприятий предназначена для повышения объективности оценки знаний учащегося. При выставлении оценок по результатам тестов используются стандартизированные средства измерения, четкие критерии, а участие преподавателя сведено к минимуму [2]. Но в отличие от традиционных форм контроля знаний, при которых преподаватель имеет возможность пояснить смысл вопроса, вызвавшего затруднение, задать наводящие и дополнительные вопросы, в тестовых заданиях большое значение имеет четкость, ясность и однозначность формулировок. По мнению М. Б. Челышковой, в этом случае близость тестовых оценок к объективным зависит только от того, насколько профессионально сделаны тесты [5].

Широкое распространение в нашей стране тесты получили совсем недавно, и хотя существует достаточное количество соответствующей методической литературы, многие преподаватели, стремящиеся использовать в учебном процессе тестовые формы контроля, при составлении последних действуют скорее интуитивно, чем опираясь на четкие, математически обоснованные правила.

При должной квалификации составителя тестовых заданий, по результатам проведения тестирования можно получить целостную картину знаний обучаемого или группы обучаемых по отдельному предмету или тематическому разделу, а при отсутствии опыта в создании тестов можно добиться только необъективных оценок знаний учащихся и негативного отношения к тестовым формам контроля.

Грамотная разработка теста должна начинаться с определения целей его создания, т.е. преподаватель должен ответить на вопрос: какие результаты уча-

щихся он хочет оценить с помощью теста. Следующий этап — это отбор содержания. В отличие от отбора содержания традиционных средств контроля, который производится в основном интуитивно на основании практического опыта педагога, отбор содержания теста имеет четкую целевую направленность, а это при условии правильной постановки целей является залогом высокого качества теста. Далее разрабатывается план теста, то есть устанавливается примерное процентное соотношение между содержанием разделов, и определяется необходимое количество заданий по каждому разделу дисциплины (исходя из важности раздела и числа часов, отведенных на его изучение). В идеале следующим этапом разработки теста должна стать экспертная проверка его качества. На практике, к сожалению, этот этап в подавляющем большинстве случаев пропускается.

Если качество заданий традиционных форм контроля оценить достаточно сложно, т.к. не существует разработанных методов такой оценки, то для заданий тестовой формы разработаны соответствующие теории проверки качества заданий [5].

Процесс научного обоснования качества теста предполагает оценку его надежности и валидности. В зарубежных исследованиях показателей качества тестов существует множество методов и подходов, которые не конкурируют между собой, а взаимно дополняют друг друга [4]. Каждый из методов имеет свои достоинства и недостатки, свою область применения и особенности интерпретации оценок качества тестов.

Для оценки валидности, т.е. характеристики качества теста, ориентированной на проверку адекватности теста поставленной цели его создания, обычно используют корреляцию между показателями теста и некоторым внешним критерием. Для педагогических тестов в качестве критерия обычно берутся оценки экспертов, выставленные ими при традиционной проверке знаний учеников.

Показатели валидности и надежности тестовых материалов, а также другие характеристики можно вычислить с помощью математических методов классической и современной тестовых теорий [4].

В классической тестовой теории для определения качества теста и интерпретации результатов тестирования применяются методы теории вероятности и математической статистики [5].

Дифференциация испытуемых по уровню подготовки определяется с помощью дисперсии и стандартного отклонения:

$$S_x^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N - 1};$$

$$S_x = \sqrt{S_x^2}.$$

За норму принимается  $\bar{X} = 3S_x$ . Низкая же дисперсия соответствует слабой дифференциации испытуемых по уровню подготовки.

Валидность тестовых материалов оценивается с помощью точечно-бисериального коэффициента корреляции:

$$r_{pbis} = \frac{\bar{X} - (X_0)_j}{S_x} * \sqrt{\frac{(N_0)_j * N}{(N_1)_j * (N - 1)}};$$

где  $(\overline{X_0})_j$  — среднее значение индивидуальных баллов испытуемых, выполнивших неправильно  $j$ -е задание теста;

$(N_1)_j$  — число испытуемых, выполнивших правильно  $j$ -е задание теста;

$(N_0)_j$  — число испытуемых, выполнивших неправильно  $j$ -е задание теста;

$S_x$  — стандартное отклонение по множеству значений индивидуальных баллов;

$N$  — общее число испытуемых;

$\overline{X}$  — среднее значение всех индивидуальных баллов по выборке учеников.

Сложность теста в целом определяется знаком коэффициента асимметрии:

$$A = \frac{\sum_{i=1}^N (X_i - \overline{X})^3}{N * S_x^3}.$$

Положительная асимметрия характерна для излишне легких тестов, отрицательная — для сложных или неверно сбалансированных по трудности тестов [5].

Тест считается надежным, если он обеспечивает высокую точность измерений и дает при повторном проведении близкие результаты (при условии, что подготовка обучаемого не изменилась за время, прошедшее до повторного проведения теста).

Для оценки надежности теста обычно применяются следующие методы:

- ретестовый (двукратное тестирование);
- параллельных форм (двукратное тестирование);
- расщепления теста (однократное тестирование).

Выбор метода проверки надежности теста в основном зависит от цели тестирования и перечисленные методы не противоречат друг другу, а скорее друг друга дополняют.

Эффективность тестовых оценок зависит не только от качества тестовых материалов, но и от методов сравнения и интерпретации первичных (сырых) баллов испытуемых.

Необходимость разработки методов интерпретации вызвана стремлением выявить истинные различия в уровне знаний испытуемых, сравнить результаты между собой даже в том случае, когда они получены по различным тестам. Большинство авторов [1, 5] отмечают, что сырые баллы не дают информации о реальном уровне знаний. Один и тот же испытуемый может выглядеть по-разному на фоне более сильной или более слабой группы.

Шкала — средство фиксации результатов измерения свойств объектов путем упорядочения их в определенную систему чисел, в которой отношение между отдельными результатами выражено в соответствующих величинах. В процессе упорядочения каждому элементу совокупности наблюдаемых эмпирических данных — результату выполнения теста — ставится в соответствие определенный балл (шкальный индекс), устанавливающий положение результата на шкале [1].

Согласно классификации шкал, предложенной С. Стивенсом [1], для обработки результатов тестирования:

- по системе «зачет/незачет» используется номинальная шкала (шкала наименований);
- в пятибалльной системе применяется порядковая (ранговая) шкала;
- в соответствии с принципами современной тестовой теории используется интервальная шкала.

В классической тестовой теории преимущественно используются шкалы:

- процентильных рангов;
- z-шкала.

Шкала процентильных рангов относится к порядковому типу. Данная шкала позволяет установить ранг первичного показателя испытуемого в нормативной группе. Ранг показателя в процентилях определяется процентной долей испытуемых из выборки, результаты которых не выше данного первичного показателя. Процесс построения шкалы процентилях состоит в определении процентильных рангов первичных показателей группы испытуемых [1].

Z-шкала также является порядковой, ее показатели отражают отличие индивидуального балла испытуемого от среднего балла по выборке в единицах стандартного отклонения для полученного при тестировании распределения. Эти производные показатели служат для определения места первичного балла каждого испытуемого в сравнении с результатами других испытуемых на основе подсчета нормированных отклонений. Такой перевод не меняет точности измерений, но позволяет значительно повысить сопоставимость результатов тестируемых и добиться благодаря этому более высокой обоснованности принятия решений [1].

Современная тестовая теория (в зарубежных источниках Item Response Theory, или IRT) предназначена для оценки латентных (т.е. скрытых от непосредственного наблюдения) параметров испытуемых и параметров заданий теста с помощью математико-статистических моделей измерения.

В IRT вводится основное предположение о существовании некоторой взаимосвязи между наблюдаемыми результатами тестирования и латентными качествами испытуемых, выполняющих тест. Обычно под этими латентными качествами понимаются способности испытуемых или уровни подготовки по предмету в зависимости от целей создания педагогического теста [5].

В отличие от классической теории, где индивидуальный балл тестируемого рассматривается как постоянное число, в IRT латентный параметр трактуется как некоторая переменная. Начальное значение параметра получается непосредственно из эмпирических данных тестирования. Переменный характер измеряемой величины указывает на возможность последовательного приближения к объективным оценкам параметра с помощью тех или иных итерационных методов.

Наиболее значимым преимуществам IRT являются [5]:

- Устойчивость и объективность оценок параметра, характеризующего уровень подготовки испытуемых. Устойчивость можно считать наиболее важным преимуществом IRT. Источником ее является относительная инвариантность оценок параметра испытуемых от трудности заданий теста.

- Устойчивость и объективность оценок параметра трудности заданий, их независимость от свойств выборки испытуемых, выполняющих тест.

- Возможность измерения значений параметров испытуемых и заданий теста в одной и той же шкале, имеющей свойства интервальной.

Наиболее важным является последнее преимущество, т.к. преобразование исходных величин разного происхождения в одну стандартную шкалу позволяет соотнести уровень знаний любого испытуемого с мерой трудности каждого задания теста. Введение единой шкалы приобретает большую важность в последние годы, поскольку на этом основана организация современного адаптивного автоматизированного контроля знаний, который на сегодняшний день является наиболее эффективной формой оценки знаний школьников или студентов.

С помощью IRT можно предсказать вероятность правильного выполнения заданий теста любым испытуемым в выборке до предъявления теста группе учеников, выявить эффективность различных по трудности заданий, используемых для оценки знаний, отличающихся по подготовке учеников тестируемой группы.

В IRT рассматривают два множества латентных параметров [1]:

- уровень подготовки  $N$  испытуемых  $\theta_i, i = 1..N$ ;
- трудность  $n$  заданий теста  $\beta_j, j = 1..n$ .

Взаимодействие двух множеств латентных параметров порождает наблюдаемые результаты выполнения теста. Но на практике всегда ставится обратная задача: по ответам испытуемых на задания теста оценить значения латентных параметров  $\theta$  и  $\beta$ .

Идея установления соотношения между латентными параметрами принадлежит Г. Рашу [4], который предложил ввести его в виде разности  $\theta - \beta$ , предполагая, что параметры  $\theta$  и  $\beta$  оцениваются в одной и той же шкале, могут располагаться на одной и той же оси переменных. В этом случае  $|\theta_i - \beta_j|$  можно интерпретировать как расстояние, на котором находится испытуемый с уровнем подготовки  $\theta_i$ , от задания с трудностью  $\beta_j$ .

Важнейшими в IRT являются следующие величины [5]:

- условная вероятность правильного выполнения  $i$ -м испытуемым с уровнем подготовки  $\theta_i$  различных по трудности заданий теста, считая  $\theta_i$  параметром  $i$ -го ученика,  $\beta$  — независимой переменной:

$$P_i \{x_{ij} = 1 | \theta_i\} = f(\theta_i - \beta), i = 1..N;$$

- условная вероятность правильного выполнения  $j$ -го задания трудностью  $\beta_j$  различными испытуемыми группы; здесь независимой переменной является  $\theta$ , а  $\beta_j$  — параметр, определяющий трудность  $j$ -го задания теста:

$$P_j \{x_{ij} = 1 | \beta_j\} = f(\theta - \beta_j), j = 1..n.$$

Основная проблема в IRT связана с выбором математической модели для описания связи между латентными параметрами и наблюдаемыми результатами выполнения теста.

Наиболее используемыми в IRT являются следующие математические модели [5]:

- однопараметрическая модель Г. Раша

$$P_j(\theta) = \frac{e^{1,7(\theta - \beta_j)}}{1 + e^{1,7(\theta - \beta_j)}},$$

$$P_i(\beta) = \frac{e^{1,7(\theta_i - \beta)}}{1 + e^{1,7(\theta_i - \beta)}};$$

- двухпараметрическая модель А. Бирнбаума

$$P_j(\theta) = \frac{e^{1,7a_j(\theta - \beta_j)}}{1 + e^{1,7(\theta - \beta_j)}},$$

$$P_i(\beta) = \frac{e^{1,7a_i(\theta_i - \beta)}}{1 + e^{1,7(\theta_i - \beta)}},$$

- коэффициент  $a_j$  называется дифференцирующей способностью заданий теста, параметр  $a_j$  указывает меру структурированности знаний ученика;
- трехпараметрическая модель А. Бирнбаума

$$P_j \{x_{ij} = 1 | \beta_j\} = c_j + (1 - c_j) \frac{e^{1,7 a_j (\theta - \beta_j)}}{1 + e^{1,7 (\theta - \beta_j)}},$$

коэффициент  $c_j$  является вероятностью угадывания правильного ответа на  $j$ -е задание теста.

На практике чаще всего используется однопараметрическая модель, которая получила название простой логистической модели.

Начальная оценка уровня знаний испытуемого определяется по формуле  $\theta_i^0 = \ln \frac{p_i}{q_i}$ , где  $p_i$  — доля правильных ответов  $i$ -го испытуемого,  $q_i$  — доля неправильных ответов, причем  $q_i = 1 - p_i$ . Аналогично определяется начальная оценка уровня трудности задания теста  $\beta_j^0 = \ln \frac{q_j}{p_j}$  где  $p_j$  — доля правильных ответов на  $j$ -е задание теста,  $q_j$  — доля неправильных ответов.

С помощью соответствующих преобразований полученных характеристик для заданий теста, а также при сопоставлении их с характеристиками испытуемых разработчик теста получает информацию о надежности и валидности тестовых материалов.

В современной тестовой теории латентные оценки параметров испытуемых и заданий располагаются вдоль одной шкалы, которая получила название шкалы логитов. Благодаря этому каждую точку шкалы, соответствующую оценке уровня подготовки испытуемого, можно соотнести с трудностью заданий, лежащих на шкале логитов левее и правее этой точки.

Условно процесс шкалирования можно подразделить на три этапа. Первый предполагает построение шкалы логитов уровня знаний, второй — шкалы логитов трудности заданий и третий этап позволяет свести две шкалы в общую шкалу стандартных оценок для уровня подготовки испытуемых и трудности заданий теста.

Процедура построения шкалы латентных переменных связана с процедурой шкалирования по Гуттману, когда задания отбираются в порядке нарастания их трудности по определенным, тщательно структурированным элементам содержания дисциплины. При этом предполагается, что любой испытуемый с правильной структурой знаний, справившись с каким-либо заданием, может успешно выполнить все предыдущие, более легкие задания теста. Это обстоятельство делает чрезвычайно привлекательной шкалу Гуттмана для педагогов, хотя ее реализация на практике довольно сложна.

Преобразование в единую интервальную шкалу дает возможность объективного сравнения достижений испытуемых по различным тестам и предметам, что, конечно, особенно важно в процессе экзаменов. Равные приращения сырых баллов испытуемых не соответствуют равным приращениям шкалированных оценок параметров, зато последние разности приобретают вполне интерпретируемый смысл, поскольку их можно считать мерой отличия уровня подготовки испытуемых по предмету. Кроме того, получаемые с помощью математических моделей IRT статистические оценки параметров обладают относительной независимостью друг от друга.

Рассмотренные выше математические методы тестовых теорий позволяют провести качественный анализ тестовых материалов и обработку данных, полученных в результате тестирования, при этом:

- методы классической тестовой теории являются более доступными и проработанными, следовательно, их применение на практике более оправдано для анализа результатов мероприятий текущего контроля знаний;
- методы современной тестовой теории предоставляют больше возможностей для оценки и уровня знаний учащихся и качества тестов и обладают большим потенциалом развития и применения в будущем.

Во многих системах управления обучением (например, ATutor, Moodle, Sakai) в той или иной мере используются математические модели классической тестовой теории для обработки результатов различных видов тестирования и проверки качества тестовых материалов. К сожалению, в e-learning системах слабо проработаны методы IRT, возможно, причиной этому послужила относительная «молодость» данной теории. Однако существуют системы управления обучением, имеющие расширяемую модульную структуру, что позволяет включить в их состав дополнительные аналитические функции.

На основании изучения предметной области и литературных источников представляется необходимым:

- определить, в каких областях и процессах образовательной деятельности возможно эффективное применение методов и концепций электронного обучения;
- оценить трудоемкость практического применения методов современной тестовой теории;
- проверить целесообразность применения математических моделей IRT для различных видов тестирования (итоговое, диагностирующее, самотестирование и т.п.).

Без проведения этого анализа простой перенос методов и технологий e-learning из сферы дистанционного обучения в очный образовательный процесс высшей школы вряд ли способен дать ожидаемый положительный эффект.

#### СПИСОК ЛИТЕРАТУРЫ

1. Звонников В. И. Измерения и шкалирование в образовании. М.: Логос, 2006. 136 с.
2. Талызина Н. Ф. Теоретические основы контроля в учебном процессе. М.: Знание, 1983 г. 253 с.
3. Хортон У. Электронное обучение: инструменты и технологии. КУДИЦ-ОБРАЗ, 2005. 640 с.
4. Чельшкова М. Б. Разработка педагогических тестов на основе современных математических моделей. М.: МИСИС, 1995. 320 с.
5. Чельшкова М. Б. Теория и практика конструирования педагогических тестов: учебное пособие. М.: Логос, 2002. 432 с.