Владимир Бодряков, заведующий кафедрой математического анализа математического факультета, доцент, доктор физико-математических наук Уральского государственного педагогического университета
Нина Фомина, старший преподаватель кафедры математического анализа

ВЕРОЯТНОСТНО-СТАТИСТИЧЕСКАЯ МОДЕЛЬ КОЛИЧЕСТВЕННОЙ ОЦЕНКИ ЗНАНИЙ СТУДЕНТОВ В НЕОДНОРОДНОЙ ГРУППЕ

Обсуждается вероятностно-статистическая модель количественной оценки уровня знаний студентов, обучающихся в группе, в общем случае, неоднородной по уровню подготовленности. Согласно модели, каждый студент получает $n = 5$ примерно одинаковых по сложности экзаменационных заданий (вопросов) по оцениваемой дисциплине, каждое из которых решает с вероятностью p (или не решает с вероятностью $q = 1 - p$). Набранная в итоге сумма баллов $k = 0, 1, ..., n$ распределена по близкому к нормальному биномиальному закону, если для всех студентов $p = \text{const}$. В неоднородной группе с $p \neq \text{const}$ распределение баллов k значимо отлично от нормального, что требует изменений в стиле обучения и оценивания знаний студентов.

Не вызывает сомнений актуальность проблемы адекватной количественной оценки качества образовательного процесса в любом образовательном учреждении, последующая обработка и интерпретация результатов наблюдений, выработка плана необходимых корректирующих мероприятий, их реализация и отслеживание их эффективности. Несомненна также актуальность и эффективность применения современного аппарата математической статистики для обработки информации на всех этапах мониторинга. Адекватная количественная модель оценки знаний позволяет дать реалистичный вероятностный прогноз результатов контроля и своевременно внести необходимые корректировки в учебный процесс.

Особую важность приобретает проблема адекватной количественной оценки знаний учащихся как в течение семестра (текущий контроль), так и при итоговом контроле, например, на экзаменационных сессиях. Если в первом случае могут применяться самые разнообразные, в том числе и неформальные виды контроля, то официальная оценка знаний имеет регламентированный формальный характер. Для определённости будем говорить далее именно об экзаменационной оценке знаний студента в ходе экзаменационной сессии, хотя сказанное ниже полностью применимо и к любому другому «поводу» для оценивания знаний учащихся любого образовательного учреждения.

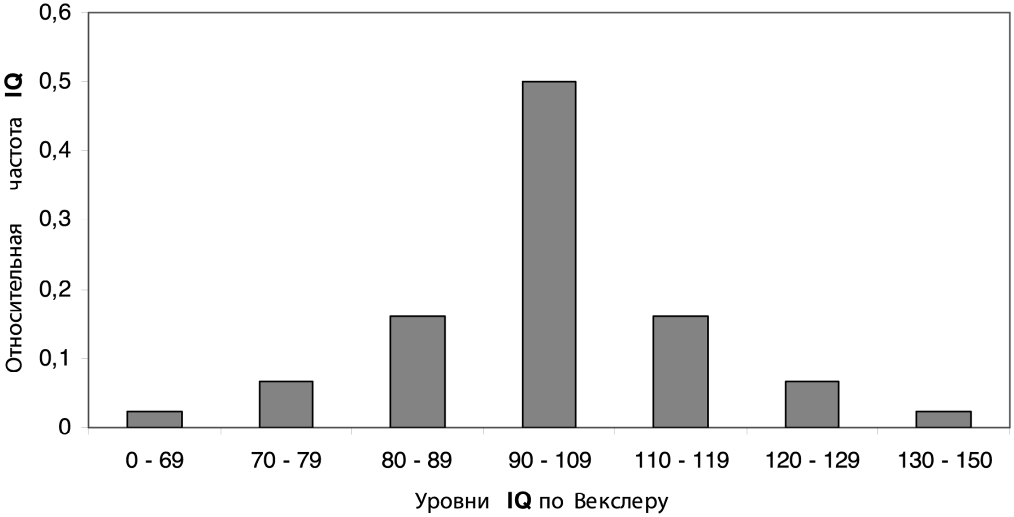

Рис. 1. Гистограмма распределения уровней интеллекта IQ подростков по Векслеру

Знания студентов по экзаменуемой дисциплине должны быть оценены одной из четырёх оценок «2» (неудовлетворительно), «3» (удовлетворительно), «4» (хорошо), «5» (отлично). При таком предельно ограниченном наборе значений признака (рангов) без применения достаточно тонких статистических технологий не обойтись: скажем, при оценке значения изменений, которые произошли в результате тех или иных педагогических воздействий. Традиционно применяемый анализ с заключениями вида «стало лучше», «стало хуже» «по сравнению с аналогичным периодом прошлого года» вряд ли может быть признан удовлетворительным в смысле содержательного количественного анализа качества учебного процесса.

Эмпирические распределения относительных частот оценок по профильным математическим дисциплинам удовлетворительно описываются нормальными[1] частотами, даваемыми нормальным законом распределения $N(a; \sigma^2; x) = N(3{,}5; 1{,}0; k)$:

$$N(a; \sigma^2; k) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(k-a)^2}{2\sigma^2}\right\} \Delta k, \qquad (1)$$

где k = 0, 1, 2, 3, 4, 5 — возможное суммарное количество баллов, набранных студентами при ответах на пять заданий по экзаменуемой дисциплине; $\Delta k = 1$ — «ширина» оценочных интервалов. Это нормальное распределение оценок, для типичного экзамена, который принимают опытные педагоги — эксперты, мы назвали опорным или базовым. При таком распределении на долю (относительную частоту) неудов-

[1] Бодряков В.Ю., Торопов А.П., Фомина Н.Г., Статистические методы исследования и оценки результатов экзаменационной сессии в вузе: + Сб. научных статей «Математика в образовании». Вып. 4. Чебоксары: ЧувГУ, 2008. + С. 18-21. Бодряков В.Ю. ., Торопов А.П., Фомина Н.Г. Динамика успеваемости студентов - математиков при обучении в УрГПУ в период поступления 2000 - выпуска 2005 г. Новые технологии в образовании. 2008. №3. С.71-75.

летворительных оценок приходится примерно 15% (здесь объединены нормальные вероятности для k = 0, 1, 2), на долю отличных оценок приходится 13%, на долю удовлетворительных и хороших оценок приходится по ~ 35%.

Полученные далее модельные вероятностные распределения оценок сопостовляются с нормальными кривыми вида (1). По-видимому, нормальное распределение оценок является частным отражением общей нормальности распределений учащихся по уровню интеллекта IQ и другим способностям, отмечаемых многими авторами. Сказанное иллюстрирует рис. 1, где приведено близкое к нормальному распределение уровней интеллекта IQ подростков по Векслеру[2]. Вполне естественно считать, что в целом уровень подготовленности студентов, выявляемый при контроле, будет соответствовать их уровням IQ.

ОПИСАНИЕ МОДЕЛИ

Примем следующую вероятностно-статистическую модель (ВСМ) оценки знаний студента. Будем считать, что весь экзаменуемый материал разбит на некоторое количество (обычно несколько десятков) примерно равных по сложности вопросов и содержит достаточно представительную базу (желательно не менее нескольких сотен) практических заданий (задач). Будем для простоты считать, что задачи имеют приблизительно равный с теоретическими вопросами уровень сложности. Теоретические вопросы заранее сообщаются студентам для подготовки; база задач также может быть доступна студентам, например, через компьютерную сеть образовательных учреждений.

Согласно модели, каждый экзаменуемый получает n = 5 примерно одинаковых по сложности теоретических и практических заданий по оцениваемой дисциплине, каждое из которых решает с вероятностью p (или не решает с вероятностью q = 1 — p). Такое относительно небольшое число заданий характерно для традиционного устного или письменного экзамена. Для экзамена в виде теста число заданий может быть больше. Задания должны соответствовать требованиям стандарта[3]. Фактически экзаменационное испытание происходит по схеме Бернулли. Техническое удобство (хотя и необязательность) выбора именно n = 5 видно из дальнейшего. Набранная учащимся полная сумма баллов k = 0, 1, ..., 5 распределена по биномиальному закону. Вероятность исхода с k баллами определяется законом биномиального распределения:

$$B_n(k) = C_n^k \times p^k \times {}^{n-k}. \qquad (2)$$

Само экзаменационное испытание начинается с выдачи студентам n = 5 экзаменационных заданий в виде подготовленных экзаменационных комплектов. Эти задания должны включать теоретические вопросы и практические задачи из различных частей учебного курса. Процедура формирования экзаменационных комплектов может быть организована как в традиционном виде, так и со случайным, но контролируемым, выбором заданий из компьютерной базы заданий персонально для каждого экзаменуемого с применением переносного персонального компьютера. Инновационный подход к реорганизации традиционного экзамена в вузе с применением переносного персонального компьютера подробно описан нами ранее[4]. Преимущест-

---

2 Поташник М.М. Требования к современному уроку. М.: Центр педагогического образования, 2008.

3 Гмурман В.Е. Теория вероятностей и математическая статистика. + М.: Высшее образование, 2006. Орлов А.И. Прикладная статистика. М.: Экзамен, 2006.

4 Бодряков В.Ю., Форма Н.Г. Инновационный подход к реорганизации традиционного экзамена в вузе с применением переносного персонального компьютера: + Сб. научных статей «Математический вестник педвузов и университетов Вольго-Вятского региона». Вып. 10. + Киров: ВытГГУ, 2008. + С. 136-139.

вом «компьютерного» подхода к организации экзамена является более широкий охват курса, чем это возможно при ограниченном наборе заранее подготовленных билетов, большая объективность, а при необходимости и «протоколизация» экзаменационного процесса. Экзамен можно проводить непосредственно в тестовой форме. Однако мы рассмотрим нетестовый вариант экзамена.

Официальная итоговая оценка в общепринятой системе «2» — «3» — «4» — «5» формируется путём перевода набранной суммы баллов k в собственно оценку по одной из двух схем. По одной схеме, удобной именно для n = 5, оценка в точности соответствует набранному числу баллов при k = 3, 4, 5; а при k = 0, 1, 2 ставится оценка «неудовлетворительно». По другой схеме, пригодной для любого n, перевод набранной суммы баллов k, выраженной в процентах от kmax = n, в традиционную систему оценок «2» — «3» — «4» — «5» осуществляется с помощью таблицы, типа табл. 1. По завершении экзамена желательно проводить собеседование с обсуждением ошибок учащихся.

Таблица 1
Таблица перевода набранной студентом на экзамене суммы баллов k, выраженной в процентах от максимальной kmax = n, в систему оценок «2» — «3» — «4» — «5».

$\kappa = k/k_{max}$, %	Оценка
$\kappa \leqslant 40\%$	2 (неудовлетворительно)
$40\% < \kappa \leqslant 67\%$	3 (удовлетворительно)
$67\% < \kappa \leqslant 87\%$	4 (хорошо)
$87\% < \kappa$	5 (отлично)

Результаты оценивания знаний студентов для однородной средней студенческой группы (для всех студентов p = 0,7) приведены

Таблица 2
Распределение вероятностей π правильных ответов студентов на экзамене в корреляции с распределением уровней IQ по Векслеру[5].

ι	IQ	Характеристика	Относительная частота П(Aι)	π_ι
1	0 ÷ 69	Дефективный	0,022	0,5
2	70 ÷ 79	Пограничный	0,067	0,55
3	80 ÷ 89	Нижесредний	0,161	0,6
4	90 ÷ 109	Средний	0,5	0,7
5	110 ÷ 119	Вышесредний	0,161	0,8
6	120 ÷ 129	Превосходный	0,067	0,9
7	130 ÷ 150	Сверхпревосходный	0,022	0,95

5 *Поташник М.М.* Требования к современному уроку. М.: Центр педагогического образования, 2008.

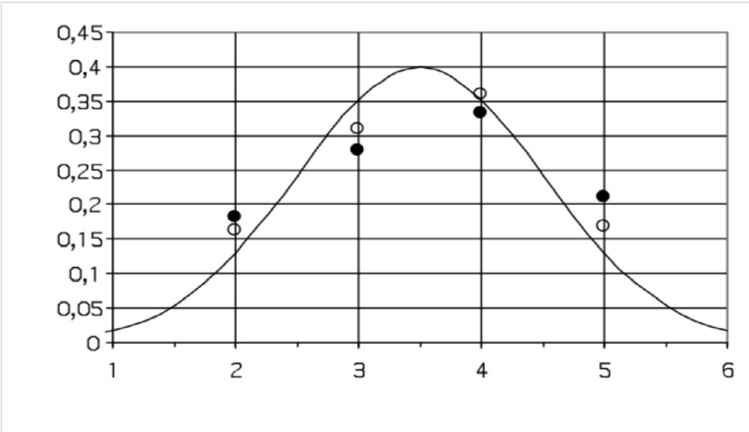

Рис. 2. Частотные распределения оценок знаний студентов. Белые кружки — согласно биномиальному распределению B5(k) для однородной студенческой группы с одинаковыми вероятностями правильных ответов p = 0,7; чёрные кружки — для неоднородной студенческой группы с распределёнными по Векслеру[6] вероятностями правильных ответов p от 0,5 до 0,95; сплошная линия — нормальная кривая N(3,5; 1,0; k)

в виде распределения относительных частот на рис. 2. В этом случае относительные частоты близки к нормальным N(3,5; 1; k) из выражения (1). Например, для коллектива студентов из N = 100 человек величина критерия Пирсона χ^2 согласия биномиального распределения B5(k) с нормальным (1) составляет $\chi^2_{\text{одн}} = 2{,}65$ при критической величине $\chi^2_{\text{крит}} = 7{,}81$ для уровня значимости $\alpha = 0{,}05$ и числа степеней свободы $r = 5-3 = 2$.

Между тем, реальное распределение вероятностей правильных ответов студентов на экзамене, конечно, неравномерно. В нашей модели мы полагаем, что распределение вероятностей p соответствует распределению уровней IQ (табл. 2). Тогда вероятность того, что при ответе на экзамене в неоднородной по уровню подготовленности студенческой группе будет набрано ровно k баллов из n = 5, определится по формуле полной вероятности[7]:

$$P_5(k) = \sum_{i=1}^{7} P(A_i)\cdot C_5^k\, p_i^k\, q_i^{5-k}\ , \qquad (3)$$

объединяющей в себе биномиальное распределение для вероятностей pi с частотным распределением самих pi в неоднородной группе. Соответствующее частотное распределение оценок показано на рис. 2.

Из рис. 2 видно, что при рассмотренном в модели распределении вероятностей pi в неоднородной группе происходит заметная трансформация распределения оценок по сравнению с однородной группой. С одной стороны, увеличивается доля неудовлетворительных оценок при заметном сокращении удовлетворительных. С другой — уменьшается доля хороших оценок при заметном возраста-

[6] *Поташник М.М.* Требования е современному уроку. М.: Центр педагогического образования, 2008.

[7] *Гмурман В.Е.* Теория вероятностей и математическая статистика. + М.: Высшее образование, 2006. Орлов А.И. Прикладная статистика. М.: Экзамен, 2006.

нии доли отличных. Иными словами, отчётливо проявляется расслоение студенческого коллектива на слабую и сильную подгруппы при относительном сокращении средней «прослойки». Несомненно, подобное положение вещей значительно усложняет педагогический процесс и требует серьёзной педагогической корректировки, без которой ситуация со временем будет лишь усугубляться. Формами такой корректировки могут быть индивидуальная педагогическая работа как с отстающими, так и с очень сильными студентами. На первых этапах важную роль может сыграть также разработка для таких студентов индивидуальных проверочных заданий, уровень которых соответствовал бы уровням их подготовленности. Сказанное подтверждает и статистически значимое несоответствие модельного распределения оценок в неоднородной группе нормальному распределению N(3,5; 1; k). Вычисленный для неоднородной группы в N = 100 человек критерий Пирсона χ^2 составляет $\chi^2_{неодн} = 11{,}07$ при критической величине $\chi^2_{крит} = 7{,}81$.

ОСНОВНЫЕ ИТОГИ РАБОТЫ

На примере экзамена в вузе детально разработана удобная вероятностно-статистическая модель количественного оценивания уровня знаний учащихся и интерпретации наблюдаемых результатов. В модели экзаменуемый получает n = 5 примерно одинаковых по сложности экзаменационных заданий, каждое из которых решает с вероятностью p (или не решает с вероятностью q = 1 — p), так что итоговая оценка определяется простой суммой k = 0, ..., n набранных баллов за решённые задания. Вероятность реализации k баллов из n может быть найдена по закону биномиального распределения $B_n(k) = C_n^k \times p^k \times q^{n-k}$ для однородной студенческой группы и по формуле полной вероятности для неоднородной по IQ и по уровню подготовленности студенческой группы, учитывая неравномерное распределение по группе самих вероятностей p; проведены модельные расчёты итогов сдачи экзаменов для студентов в однородной (с $p = 0{,}7$) и неоднородной (с p от 0,5 до 0,95) студенческой группе. Показано наличие выраженной тенденции к усилению расслоения неоднородной по уровню IQ студенческой группы на подгруппы слабых и сильных студентов. Даны рекомендации по организации педагогического процесса в этом случае; разработанная модель универсальна, эффективна и делает процесс количественного оценивания знаний студентов объективным, открытым, наглядным и убедительным. Обеспечиваемое моделью надёжное вероятностное прогнозирование результатов контроля может значительно повысить эффективность управления учебным процессом.