

**Л.М. БУХМАН, доцент
Н.С. БУХМАН, профессор
Самарский государственный
архитектурно-строительный
университет**

О некоторых недостатках централизованного интернет-тестирования

Предложена математическая модель централизованного интернет-тестирования по методике ФЭПО. Показано, что результаты тестирования зависят не только от уровня подготовки студентов, но и от количества дидактических единиц в изучаемом учебном курсе. Сделан вывод о неоптимальности используемой системы интернет-тестирования и необходимости отказа от требования усвоения всех дидактических единиц учебного курса.

Ключевые слова: интернет-тестирование, дидактическая единица, математическая модель централизованного тестирования

В настоящее время одним из важнейших показателей, влияющих на аккредитацию вуза, являются результаты интернет-тестирования студентов этого вуза по дисциплинам, входящим в циклы «ГСЭ», «ЕН» и «ОПД». Это вполне понятно и естественно: если студенты не в состоянии нормально пройти тестирование, значит, вуз не способен подготовить их на должном уровне.

Тестирование по методике ФЭПО организовано достаточно логично и имеет, как и любое компьютерное тестирование вообще, целый ряд хорошо известных достоинств [1–3]. Весь курс (например, курс физики) делится на несколько относительно самостоятельных дидактических единиц (ДЕ), при этом отдельный студент считается успешно прошедшим тестирование в том случае, если он продемонстрировал знание *всех* ДЕ курса. Это требование выглядит вполне естественно – ведь незнание хотя бы одной из ДЕ свидетельствует о том, что студент не овладел учебным курсом полностью (или преподаватель не прочитал курс полностью). Студент считается знающим (и успешно сдавшим) соответствующую ДЕ в том случае, если он правильно ответил не менее чем на половину вопросов, относящихся к данной ДЕ.

Группа считается успешно прошедшей тестирование, если не менее 50% (иногда – 60%) студентов от ее списочного состава

успешно прошли тестирование. Именно успешное прохождение группой студентов централизованного интернет-тестирования по той или иной учебной дисциплине является жизненно важным для сохранения вуза «на плаву». Цель данной работы – создать математическую модель централизованного тестирования группы студентов и выяснить, от каких именно параметров и как зависит вероятность успешного прохождения группой студентов интернет-тестирования.

В качестве исходных параметров модели будем использовать следующие:

n – количество ДЕ в курсе данной дисциплины;

N – количество студентов в группе;

p – вероятность успешной сдачи *одним* студентом *одной* ДЕ;

d – минимальная доля группы, которая должна успешно сдать тест, для того чтобы тестирование считалось успешно пройденным всей группой. Приведенные в данной работе расчеты сделаны в предположении $d=0,5$; ясно, что качественные выводы сохраняются и при других значениях этого параметра (например, при $d=0,6$).

Мы предполагаем, что события, заключающиеся в успешной сдаче той или иной ДЕ разными студентами, независимы (в вероятностном смысле [4]). Также мы предполагаем, что независимыми (в вероятност-

ном смысле) являются события, заключающиеся в успешной сдаче различных ДЕ одной и той же учебной дисциплины одним и тем же студентом. Если бы это было не так, то было бы совершенно непонятно, зачем нужен независимый контроль знания студентами разных ДЕ одной учебной дисциплины, и деление теста на дидактические единицы потеряло бы всякий смысл¹. Поскольку для успешной сдачи теста студент должен сдать n независимых ДЕ, вероятность успешной сдачи теста одним студентом (вероятность успеха) равна p^n , а вероятность того, что студент провалит тест (вероятность неудачи), равна $1 - p^n$.

Рассматривая успешную сдачу теста различными студентами группы как серию независимых испытаний, приходим к схеме Бернулли [4], в которой число студентов группы, успешно сдавших тест, – случайная величина μ , имеющая распределение Бернулли с вероятностями успеха и неудачи в одном испытании p^n и $1 - p^n$ соответственно:

$$P(\mu = m) = C_N^m p^{nm} (1 - p^n)^{N-m}. \quad (1)$$

Поскольку для успешной сдачи теста группой студентов необходимо, чтобы количество успешно сдавших тест студентов было не менее Nd , где минимально допустимая доля сдавших тест студентов d обычно принимается за 0,5 (50%) или 0,6 (60%), то для вероятности успешной сдачи теста группой студентов P_{test} имеем:

$$P_{test}(p, n, N, d) = \sum_{m=[Nd]}^N C_N^m p^{nm} (1 - p^n)^{N-m}, \quad (2)$$

где $[x]$ – ближайшее к x целое, не меньшее x .

В случае достаточно большого (свыше 20) количества студентов в группе N происходит нормализация распределения количества студентов, успешно прошедших

тест; в этом случае в нормальном приближении (на основе центральной предельной теоремы [4]) имеем:

$$P_{test}(p, n, N, d) = \frac{1}{2} + \Phi \left(\frac{\sqrt{N}}{\sqrt{p^n(1-p^n)}} (p^n - d) \right), \quad (3)$$

где $\Phi(x)$ – интеграл вероятностей [4].

При еще большем количестве студентов в группе (или при еще меньших требованиях к точности математической модели) можно пренебречь дисперсией количества студентов, успешно сдавших тест, и получить в приближении закона больших чисел [4] следующий результат:

$$P_{test}(p, n, d) = \begin{cases} 0, & \text{если } p < p_{min} \\ 1, & \text{если } p > p_{min} \end{cases}, \quad (4)$$

где $p_{min} = d^{1/n}$ – минимальная вероятность успешной сдачи отдельным студентом одной ДЕ, необходимая для успешной сдачи группой студентов теста из n ДЕ.

Нетрудно заметить, что в этом приближении результаты тестирования группы студентов перестают быть случайными – как и должно быть в полном соответствии с духом и буквой закона больших чисел. Это же, кстати, одновременно означает, что обсуждаемая методика тестирования ФЭПО является состоятельной в том смысле, который обычно вкладывается в это понятие в математической статистике: при неограниченном росте объема выборки (в данном случае – количества студентов в группе) результат тестирования перестает быть случайным.

В случае конечного количества студентов в группе при любом уровне подготовки студентов (то есть при любой вероятности сдачи одной ДЕ одним студентом p существует некоторая ненулевая вероятность сдачи теста плохо подготовленной группой и ненулевая же вероятность провала теста хорошо подготовленной группой. Это естественно – при конечном количестве про-

¹ Впрочем, как станет ясно из дальнейшего, и в предположении реальной независимости различных ДЕ от дробления теста на ДЕ вреда в конечном счете больше, чем пользы.

веряемых роль случайности в принципе не может быть устранена полностью. В случае же бесконечного количества студентов в группе (бесконечный объем выборки) результат тестирования группы перестает быть случайным: в случае плохой подготовки студентов ($p < p_{min}$) группа с гарантией проваливает тестирование, а в случае хорошей ($p > p_{min}$) – с гарантией его проходит.

Для иллюстрации характера полученных зависимостей на рис. 1 приведены графики зависимости вероятности успешной сдачи теста из $n=4$ и $n=8$ ДЕ группой студентов из $N=15$ (пунктир), 30 (штриховая линия) и 300 (штрих-пунктир) человек² (кривые получены путем численного счета по формуле (2)) и для группы бесконечного объема $N=\infty$ (сплошные ступенькообразные кривые, получены по формуле (4)). Предполагалось, что для успешной сдачи теста группой достаточно, чтобы тест успешно сдала хотя бы половина группы ($d=0,5$).

Из рисунка видно, что случайность результатов тестирования действительно сильнее проявляется в небольших группах. Тем не менее даже в небольших группах (15 человек) она достаточно иллюзорна – вероятность успешного прохождения теста слабоподготовленной группой (например, при $p = 0,6$), равно как и вероятность провала тестирования прекрасно подготовленной группой ($p = 0,97$) пренебрежимо мала. Случайность играет заметную роль только в «пороговой» ситуации, когда уровень подготовки группы к тестированию таков, что она сдает тест «еле-еле», то есть когда $p \approx p_{min}$.

Кроме того, видно, что случайность проявляется лишь в некотором «размытии» пороговой вероятности $p \approx p_{min}$, причем это размытие даже в случае небольших групп невелико. Практически оказывается, что для успешного прохождения тестирования группой студентов достаточно обеспечить уровень подготовки студентов не ниже некоторого предельного. Именно: ве-

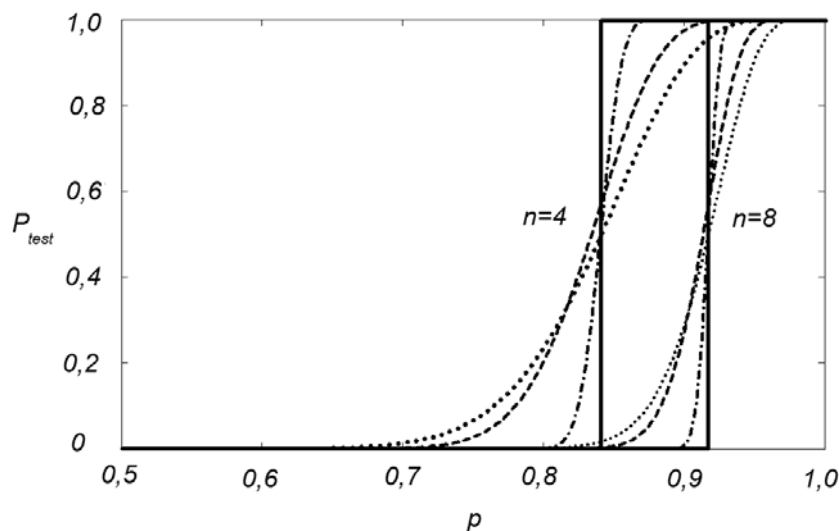


Рис. 1. Зависимость вероятности успешной сдачи теста из $n=4$ и $n=8$ ДЕ группой из 15, 30, 300 и бесконечного количества студентов

² Разумеется, авторы не предполагают, что кто-нибудь подвергнет тестированию группу из 300 студентов; мы просто хотели показать, как происходит предельный переход от реального объема группы (15–30 человек) к бесконечно большому.

роятность успешной сдачи одним студентом одной ДЕ должна быть не ниже значения $p_{\min} = d^{1/n}$. Если группа студентов невелика, то для исключения случайных провалов при групповом тестировании необходимо обеспечить чуть более высокий уровень подготовки студентов. График зависимости функции $p_{\min} = d^{1/n}$ от количества ДЕ в данном учебном курсе n при различных значениях параметра d приведен на рис. 2.

хождения теста вероятность успешной сдачи одной ДЕ одним студентом существенно зависит от количества ДЕ в данном курсе (см. рис. 2). Так, например, в случае $d=0,5$ для курса из трех ДЕ допустимая вероятность несдачи³ одним студентом одной ДЕ составляет 20,6%, для курса из четырех ДЕ – 15,9%, для курса из восьми ДЕ – только 8,3%, а для курса из 12 ДЕ – 5,6%. Практически это означает, что при сдаче «коротких» курсов (три-четыре ДЕ) студенты име-

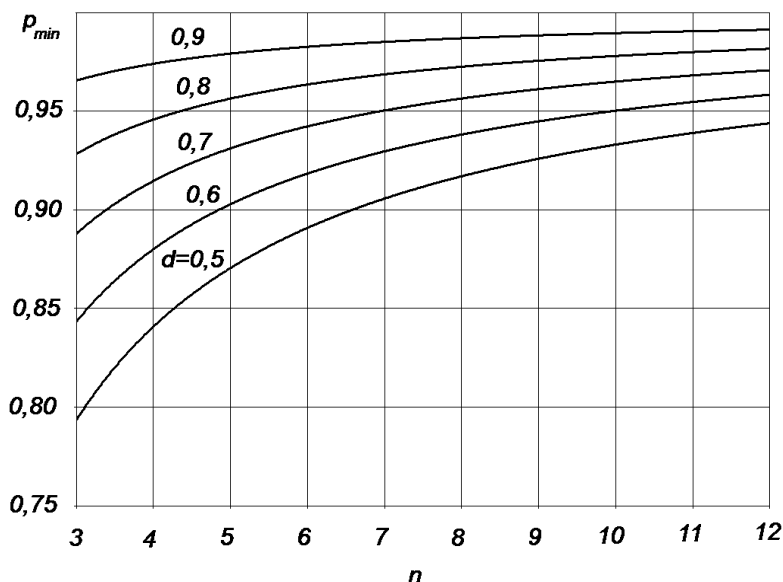


Рис. 2. Минимально допустимая вероятность сдачи одной ДЕ одним студентом при различных значениях параметра $d = 0,5; 0,6; 0,7; 0,8; 0,9$ и различном количестве ДЕ в учебной дисциплине ($n = 3 \dots 12$). Разумеется, физический смысл имеют только целочисленные значения параметра n

В приближении закона больших чисел (то есть при бесконечном объеме группы студентов) результаты успешного прохождения тестирования не являются случайными: группа либо с гарантией проходит тестирование (в случае $p > p_{\min(n)}$), либо с гарантией не проходит его (в случае $p < p_{\min(n)}$).

Существенно и неприятно то, что необходимая и достаточная для успешного про-

ют возможность ошибаться гораздо чаще (в разы!), чем при сдаче «длинных» (8–12 ДЕ).

По мнению авторов, это обстоятельство является существенным недостатком существующей системы тестирования и свидетельствует о необходимости изменения критериев успешного прохождения тестирования группой студентов.

³ На рис. 2 показаны вероятности успешной сдачи. Поэтому вероятность несдачи в 20,6% соответствует вероятности сдачи 79,4%, и так далее.

Действительно, приемлемая система тестирования должна обеспечивать контроль уровня подготовки студентов в определенной предметной области, причем результаты этого контроля не должны зависеть от того, как называются части этой предметной области и на сколько именно частей делится эта предметная область. (Качество партии металла не может и не должно зависеть от того, в килограммах, фунтах или слитках измеряют его количество).

Возьмем в качестве примера курс физики. Тесты ФЭПО по этой дисциплине состоят из 8 ДЕ (у строителей; нас по вполне понятным причинам в основном интересует эта специальность). В приближении закона больших чисел для успешной сдачи теста из 8 ДЕ преподаватель должен добиться того, чтобы вероятность провала при сдаче одной ДЕ одним студентом не превышала 8,3%. Пусть теперь курс физики разделили на два курса – например, «Физика – 1 часть» и «Физика – 2 часть», причем ни содержание образовательной программы, ни уровень подготовки студентов не изменились. Изменилась только вероятность успешного прохождения тестирования: теперь для подтверждения достаточного уровня подготовки вместо одного теста из 8 ДЕ студенты должны будут сдать два независимых теста из 4 ДЕ (если разделить курс пополам). В приближении закона больших чисел для успешной сдачи одного, двух и вообще любого количества тестов⁴ по 4 ДЕ преподаватель должен добиться того, чтобы вероятность провала при сдаче одной ДЕ одним студентом не превышала 15,9% (вместо 8,3% при сдаче курса целиком). Таким образом, чисто словесное изменение «вывески» приводит к двукратному изменению реального уровня требований к знаниям студентов, что является со-

вершенно недопустимым. Сложившаяся практика является причиной того, что «под ударом» в результате провалов студентов при централизованном тестировании неизбежно оказываются такие «тяжеловесы», как математика (12 ДЕ), физика (8 ДЕ), начертательная геометрия (15 ДЕ). В привилегированном же положении оказываются «мелкие» дисциплины типа материаловедения (3 ДЕ), гидравлики (4 ДЕ), механики грунтов (4 ДЕ) и т.д.

Может быть, ситуация изменится, если учесть конечное количество студентов в группе и случайность результатов группового тестирования? Для проверки этого предположения с использованием точной формулы (2) для $P_{test}(p, n, N, d)$ был проведен расчет вероятности успешной сдачи теста на 8 ДЕ $P_8 = P_{test}(p, 8, N)$ и вероятности успешной последовательной сдачи двух тестов на 4 ДЕ каждый и $P_{4+4} = P_{test}(p, 4, N, d) \times P_{test}(p, 4, N, d)$ в зависимости от вероятности успешной сдачи одним студентом одной ДЕ c . Результаты этих расчетов приведены на рис. 3. Расчеты были проведены для групп, состоящих из $N=15$ студентов (сплошная линия) и из $N=30$ студентов (штриховая линия).

Видно, что теперь успех или неудача группы при тестировании случайны. Тем не менее основной вывод, сделанный в приближении закона больших чисел, сохраняется: при одном и том же уровне подготовки студентов *вероятность провала группы при сдаче одного теста на 8 ДЕ существенно выше, чем при сдаче двух тестов на 4 ДЕ каждый*. Так, например, при вероятности успешной сдачи одним студентом одной ДЕ 0,9 (вероятность провала при сдаче одной ДЕ равна 0,1) вероятность успешной сдачи группой двух тестов по 4 ДЕ каждый составляет 80–90% (в зависимости от числа студентов в группе), а вероятность

⁴ В приближении закона больших чисел случайность «исчезает» и тест либо сдается (вероятность сдачи 1), либо не сдается (вероятность сдачи 0). В первом случае вероятность последовательной сдачи любого количества тестов равна 1 ($1*1*1*...*1=1$), во втором – 0 ($0*0*0*...*0=0$).

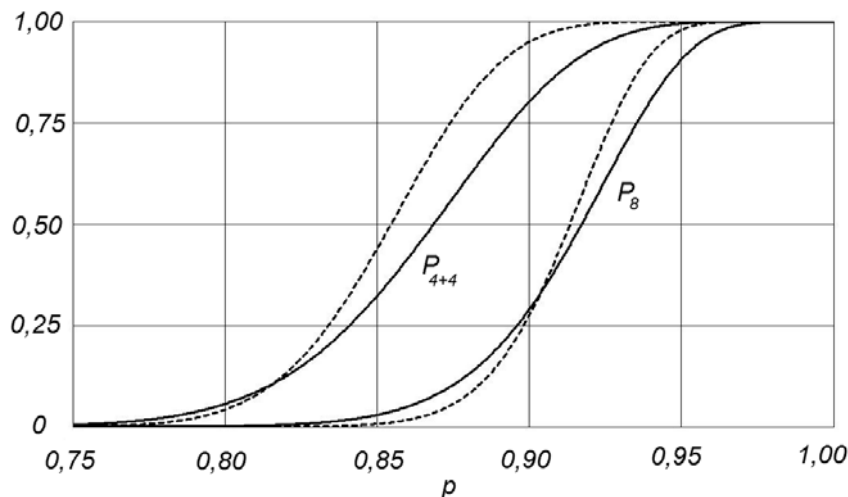


Рис. 3. Сравнение вероятностей успешной сдачи одного теста на 8 ДЕ P_8 и двух тестов по 4 ДЕ в каждом P_{4+4} для группы из 15 студентов (сплошная линия) и 30 студентов (штриховая линия)

успешной сдачи одного теста на 8 ДЕ — всего около 25%.

Из приведенного на рис. 3 графика видно, что в практически важной области параметров задачи ($N=15-30$, вероятность сдачи теста группой более 50%) вероятность успешной сдачи теста группой студентов при одном и том же уровне их подготовки растет с ростом числа студентов в группе. Поэтому тестирование «малокомплектных» групп вузам невыгодно: при том же уровне подготовки более многочисленные группы показывают лучшие результаты. Например, при $\rho=0,95$ в случае сдачи теста из 8 ДЕ вероятность провала группы из 15 человек равна 9,3%, а группы из 30 человек — только 2,1%.

В качестве вывода проведенного рассмотрения можно констатировать, что *используемая ФЭПО методика тестирования имеет серьезный недостаток*, а именно: результаты тестирования зависят не только от уровня подготовки студентов по данному кругу вопросов, но и от чисто формального и сложившегося исторически распределения этого круга вопросов по учебным дисциплинам. Авторы убеждены в том, что ситуация, при кото-

рой в случае одного и того же уровня знания студентами физики результат тестирования существенно зависит от того, называют этот предмет предметом «Физика» (двухсеместровым) или совокупностью двух предметов, например, «Классическая физика» (первый семестр) и «Современная физика» (второй семестр) — совершенно ненормальна.

Поэтому мы считаем, что при обработке результатов тестирования группы студентов необходимо вернуться к «процентной» оценке успеваемости, т.е. считать тест успешно пройденным группой студентов в том случае, когда суммарный процент правильных ответов на все вопросы теста достаточно высок и составляет, скажем, 60, 80 или 90% — без искусственного распределения этих вопросов по ДЕ и надуманного требования непременно сдать все ДЕ без исключения. Разумеется, само понятие ДЕ никуда при этом деваться не должно, но его следует оставить «для внутреннего употребления» составителя теста, который должен следить за тем, чтобы в тесте были представлены примерно в равной мере все ДЕ курса и не возникало «перекосов» в сторону его любимых тем.

Литература

1. Прахова М.Ю., Коловертов Г.Ю., Шаловников Э.А. О месте тестирования как инструмента оценки знаний в вузе // Высшее образование в России. 2012. № 7. С. 113–116.
2. Косухин В., Логинова Г., Логинова И. Роль и место тестирования в деятельности вуза // Высшее образование в России. 2008. № 1. С. 94–97.
3. Бухман А.М. Проблемы тестового контроля экологических знаний и их решение // Изв. Самарского научного центра РАН. 2010. Т. 12. № 5. С. 21–24.
4. Розанов Ю.А. Теория вероятностей, случайные процессы и математическая статистика. М.: Наука, 1985. 320 с.

С.В. КАЗНАЧЕЕВ, профессор
О.В. ЛОПАТИНА, зав. кафедрой
Ж.Ю. БОГОЛЮБОВА, зам. зав. кафедрой
Сибирский институт управления –
филиал РАНХ и ГП при Президенте РФ
В.П. КАЗНАЧЕЕВ, академик РАН

Опыт преподавания физической культуры в нефизкультурном вузе

Физическая культура как составная часть системы образования имеет воспитательное, образовательное, оздоровительное, экономическое и общекультурное значение. Она способствует возникновению такого социального течения, как физкультурное движение – совместная деятельность людей по использованию, распространению, приумножению ценностей физической культуры. Управление системой физического воспитания – сложный процесс, требующий выбора наилучшей стратегии и тактики, эффективных методов достижения общественного здоровья и формирования здорового образа жизни населения. Особенно сложен подход к управлению системой физического воспитания в настоящее время, т.к. с начала XX в. занятия физической культурой перестали быть органически связаны с повседневным трудом и бытом россиян и стали развиваться в качестве хобби отдельных граждан общества.

Ключевые слова: физическая культура, организация учебного процесса на кафедре физической культуры, нефизкультурный вуз, здоровье студентов, телесно-ориентированная педагогика, здоровьесберегающие технологии

Современная система массовой физической культуры требует своего концептуального реформирования, свидетельством чему являются многочисленные публикации, указывающие на то, что школьники, учащиеся нефизкультурных инновационных школ, студенты вузов не считают данный предмет важной учебной дисциплиной и при удобном случае находят повод для отказа им заниматься [1–3].

Физическую культуру необходимо вновь сделать органической частью системы образования. Она давно зарекомендовала себя в виде особой самостоятельной

области, а физическое воспитание – одно из направлений общего воспитания личности, представленного в виде социально-педагогического процесса. Этот процесс в своём развитии ориентирован не только на укрепление и совершенствование физических качеств человека и его личностных свойств, но и на жизненно важные стороны социального поведения индивида, полученные в виде задатков от родителей и сформированные в условиях социально-семейного воспитания. В результате объединения в одном человеке умений, навыков, практик физической деятельности, на ос-