УДК 378

КОНТРОЛЬ ОБРАЗОВАТЕЛЬНОЙ ДЕЯТЕЛЬНОСТИ НА ОСНОВЕ МЕТОДОВ МЕНЕДЖМЕНТА КАЧЕСТВА

Левина Е.Ю.

Институт педагогики и психологии профессионального образования, Казань, e-mail: frau.levina2010@yandex.ru)

Автоматизированное тестирование применяется при текущем или итоговом контроле знаний студентов достаточно часто. Данные контроля знаний студентов были оценены с точки зрения статистических зависимостей. Анализ показал тип распределения, приближенный к биномиальному. Следовательно, существует возможность использования в компьютерном тестировании студентов схемы последовательного плана выборочного контроля качества. Размер частной выборки по числу вопросов теста заранее не устанавливается. При этом ориентировочная оценка считается известной уже после малого количества вопросов, с каждым новым вопросом (испытанием) оценка уточняется. Решение о прекращении тестирования принимается исходя из заданных минимальных рисков преподавателя и студента. Применение данной схемы в автоматизированном тестировании обеспечило оценки достоверности наряду с существенным сокращением трудоемкости текущего контроля знаний.

Ключевые слова: контроль обучения, тестирование студентов, статистические методы контроля качества

CONTROL OF LEARNING ACTIVITIES BASED ON THE QUALITY MANAGEMENT METHODS

Levina E.U.

Institute of Pedagogy and Psychology of Professional Education, Kazan, e-mail: frau.levina2010@yandex.ru

Automated testing is used in the current or final control of knowledge students often. data control knowledge were evaluated in terms of statistical dependencies. Analysis showed the type of distribution close to the binomial. Consequently, there is the possibility of using computer testing students circuit sequential sampling plan quality. The sample size of (number of questions) not installed in advance. In this case, approximate estimate is known already after a small number of questions, with each new question becomes more precisely estimate. The decision to terminate testing is conducted with minimal risk of teacher and student. The application of this scheme in automated testing provided the validation along with a significant reduction in the complexity of the current control knowledge.

Keywords: control training, student testing, statistical methods for quality control

Непосредственно в процессе обучения осуществляется контроль уровня усвоения знаний студентами – текущий, тематический, рубежный, итоговый контроль. Вовремя обнаруженные отклонения параметров качества позволяют внести коррективы в технологию обучения и контроля, что особенно важно при освоении новых дисциплин или специальностей. Качество обучения считается приемлемым, если студент может выполнить свои функции в полном объеме, соответственно уровню требований, предъявляемых преподавателем на данном уровне обученности [4].

Текущий контроль знаний – операция, имеющая определенную трудоемкость, а при использовании компьютерных средств еще и повышающая стоимость образования. Даже при хорошо отлаженной системе контроля знаний в вузах отклонения регистрируемых оценок от реальных уровней существуют практически всегда хотя бы потому, что контроль имеет выборочный характер.

В соответствии со стандартами государственной приемки в промышленности заключение о качестве изделий массового выпуска выводится из результатов исследований случайной выборки из больших партий, в которой объем и допустимый процент брака определяется в согласованных документах, формулирующих вид и условия приемки [3]. При этом предполагается, что в определенном числе случаев дешевле допустить присутствие нескольких процентов брака, чем стремиться к полному исключению некачественных изделий, повышая тем самым цену изделий в несколько раз. По ГОСТ Р50779.71–99 [2] партия изделий считается качественной в том случае, если она включает в себя процент бракованных изделий, не превышающий определенную норму, задаваемую потребителем.

Применяя этот стандарт к процессу контроля качества учебной деятельности студентов, реализуемой в вузе, можно сказать следующее: ввиду различных способностей учащихся, сформированных в одну группу, довольно часто существует процент студентов, плохо освоивших данную дисциплину (брак). Однако по основной массе учащихся, удовлетворительно прошедших контрольные процедуры, можно делать выводы о качестве технологии обучения данному предмету, поскольку дешевле допустить присутствие малого количества отстающих

студентов в группе, чем стремиться к полному исключению брака в процессе обучения, вводя корректирующие курсы, повышая тем самым стоимость обучения. При всем том, если количество отстающих студентов значительно (процент брака выше допустимого), то необходимо изменение технологии преподавания или контроля. Аналогично получается и в случае с одним студентом, когда идет его опрос по большому объему материала. Стопроцентный уровень знаний редко является достигнутым в силу разных обстоятельств, поэтому для удовлетворительной сдачи экзамена или зачета допускается некоторое количество ошибок в ответе обучаемого, при этом, если количество неправильных ответов значительно, экзамен (зачет) не принимается.

При исследовании качества обучения службами мониторинга может использоваться как сплошной контроль, так и выборочный контроль, когда заключение о качестве обучения выносится из результатов исследования случайной выборки студентов. При этом психолого-педагогические факторы, влияющие на качество обучения, многообразны и малопредсказуемы, что делает неизбежным привлечение специальных методов математической обработки результатов контроля.

Контроль качества товаров [2] включает в себя несколько планов выборочного контроля: одноступенчатый, двухступенчатый, многоступенчатый, последовательный план, при этом основной характеристикой партии изделий является генеральная доля дефектных изделий $Q = D/N$, где N – объем партии, а D – количество брака во всей партии. На практике генеральная доля Q (или число D) неизвестны и оцениваются по результатам контроля случайной выборки объемом n изделий из той же партии. Доля дефектных изделий в выборке:

$$q = d/n,$$

где d – количество дефектных изделий в выборке n.

Таким образом, проблема состоит в достоверном определении неизвестного D исходя из известных числа изделий в партии N, числа изделий в выборке n и числа дефектных изделий в выборке d.

Наибольший интерес представляют минимальные объемы выборки n, при которых D надежно определяется. Для того чтобы минимизировать n, используются различные планы контроля. Под планом статистического контроля понимается система правил, устанавливающих методы случайного выбора для проверки изделий из партии, а также условия на N, n, d, при выполнении которых партию следует принять, забраковать или продолжить контроль (увеличить значение n) [5].

Для процесса компьютерного контроля знаний минимизация объема n случайной выборки означает уменьшение трудоемкости и сокращение экономических затрат, либо при проверке остаточных знаний группы студентов по определенной дисциплине, либо при проверке одного студента по большому числу вопросов теста.

Тесты, используемые в вузах, реализуют, как правило, одноступенчатый план выборочного контроля, в котором оценка выставляется в жесткой связи с количеством (или процентом) правильных ответов на n вопросов, представляющих из себя частную выборку из их большого числа $Q >> n$.

В последовательном плане контроля знаний, который предлагается применить в системе проверки остаточных знаний, размер n частной выборки заранее не устанавливается. При этом ориентировочная оценка качества считается известной уже при небольших значениях n, с каждым новым вопросом (испытанием) оценка уточняется. Решение о прекращении тестирования принимается исходя из заданных максимальных рисков преподавателя (вероятности того, что выставлена завышенная оценка) и студента (вероятности того, что оценка занижена).

Таким образом, при последовательном контроле число заданных вопросов n вычисляется исходя из статистических критериев достоверности. При этом n оказывается случайной величиной, неодинаковой для различных студентов, однако полученный результат контроля получает статистическое обоснование. Задача правильного выбора плана контроля состоит в том, чтобы одновременно минимизировать риски приемщика и поставщика, то есть риска преподавателя выставить оценку уровня знаний выше заслуженной и риск студента получить оценку ниже существующего уровня знаний.

Несмотря на то, что используемые в настоящее время одноступенчатые планы весьма просты, тем не менее более сложные последовательные планы контроля обеспечивают при меньших объемах выборок большую достоверность принимаемых решений [1]. Такие планы требуют больших объемов вычислений, однако этот факт не оказывается решающим при наличии соответствующего программного обеспечения.

При реализации планов выборочного контроля существует вероятность получения завышенного или заниженного уровня знаний в связи со случайным выбором,

либо групп студентов из потока учащихся (например, факультета), либо более успевающих (или неуспевающих) студентов из группы, либо вопросов из всего множества вопросов по дисциплине, поэтому в соответствии с ГОСТ Р 50779.71–99 существует оперативная характеристика контроля, одновременно оценивающая уровни риска как преподавателя, так и студента.

Автором разработано несколько программных средств, реализующих стандартную методику планов выборочного контроля: одноступенчатого, двухступенчатого, многоступенчатого и программы последовательного контроля качества, осуществляющую расчет рисков выставления ошибочных оценок в процессе оценки знаний.

Работу программы последовательного контроля проиллюстрируем следующим примером. Предположим, необходимо оценить знание студентом 150 иностранных слов в виде оценки, определяемой в соответствии с таблицей.

Критерии выставления оценки по результатам тестирования

Оценка	Доли z выученных слов
Отлично	$z \geq 0{,}89$
Хорошо	$0{,}89 > z \geq 0{,}62$
Удовлетворительно	$0{,}62 > z \geq 0{,}34$
Неудовлетворительно	$0{,}34 > z \geq 0{,}10$
Единица	$0{,}10 > z$

В процессе контроля знаний вопросы задаются студенту по одному, на каждом шаге подсчитывается количество правильных ответов и рассчитываются оперативные характеристики контроля, из которых интерполированием вычисляются для каждой оценки, с использованием границ интервалов из таблицы, риски преподавателя и студента. Например, если из 150 вопросов задано к настоящему моменту всего вопросов $n = 10$, из которых на 7 вопросов получены от студента верные ответы, итог проверки знаний может быть предположен в любом виде, т.е. все еще можно предположить в принципе любую итоговую оценку. При этом наиболее вероятной кажется оценка «хорошо», т.к. доля правильных ответов составляет 0,7.

С другой стороны, если принять оценку «хорошо» за данность, можно оценить вероятности того, что в процессе тестирования итоговая, реальная оценка окажется все-таки выше (риск студента) или ниже (риск преподавателя). Здесь риск студента определен как вероятность того, что студент в результате выборочного ($M < 150$) контроля получит заниженную оценку, а риск приемки – как вероятность того, что оценка окажется выше реальной – той, которая в точности соответствует результатам проверки знаний всех 150 слов. В результате расчета для рассматриваемого примера $Q = 150$ и этапа $n = 10$, $z = 0{,}7$ получено, что для оценки «хорошо» упомянутые риски ошибок составляют соответственно величины около 0,1 и 0,2. При этом для оценки «отлично» риск студента и приемки соответственно близки к нулю и к единице, а для оценки «неудовлетворительно» – наоборот, единица и ноль.

Этапы процедуры последовательного плана выборочного контроля знаний студента представлены на рис. 1.

На рис. 1, а видно, что при определении оценки «хорошо» на этапе $n = 10$ имеют место минимальные риски, однако остается достаточно вероятной и оценка «удовлетворительно». При продолжении тестирования эта возможность постепенно все более реализуется: при $n = 20$ риск приемки становится уже достаточно большим (70 %). При $n = 30$ для «удовлетворительно» повторяется примерно то же, что для «хорошо» при $n = 10$, а при $n = 50$ оценить знания на «удовлетворительно» уже можно достаточно определенно, т.е. с рисками ошибки и в большую, и в меньшую сторону – около 1 %.

В результате тестирования, проведенного до конца ($n = 150$), получено значение $z = 0{,}500$, действительно близкое к значению $z = 0{,}527$ этапа $n = 50$. Кстати, близкие к итоговой доле верных ответов 0,527 получены значения $z = 0{,}55$ при еще при $n = 20$, однако доверять этому значению, в силу больших рисков завысить или занизить оценку, не имело смысла.

Развернутая процедура контроля знаний представлена на рис. 2.

Видно, что в ходе контроля текущее значение z оказалось в рамках, определенных для оценки «удовлетворительно», уже при $n = 12$ и при дальнейшем тестировании (при $n > 12$) из этих рамок уже не выходило. При этом вероятности того, что реальная оценка окажется другой, стали менее 10 % при $n = 30$, когда студенту было задано всего около одной четвертой части вопросов (т.е. при $n > 35$).

Такие ситуации являются достаточно типичными. При более благоприятных стечениях обстоятельств на уровне рисков около 1 % тестирование оказывается возможным прекратить и ранее, когда доля заданных вопросов менее одной пятой.

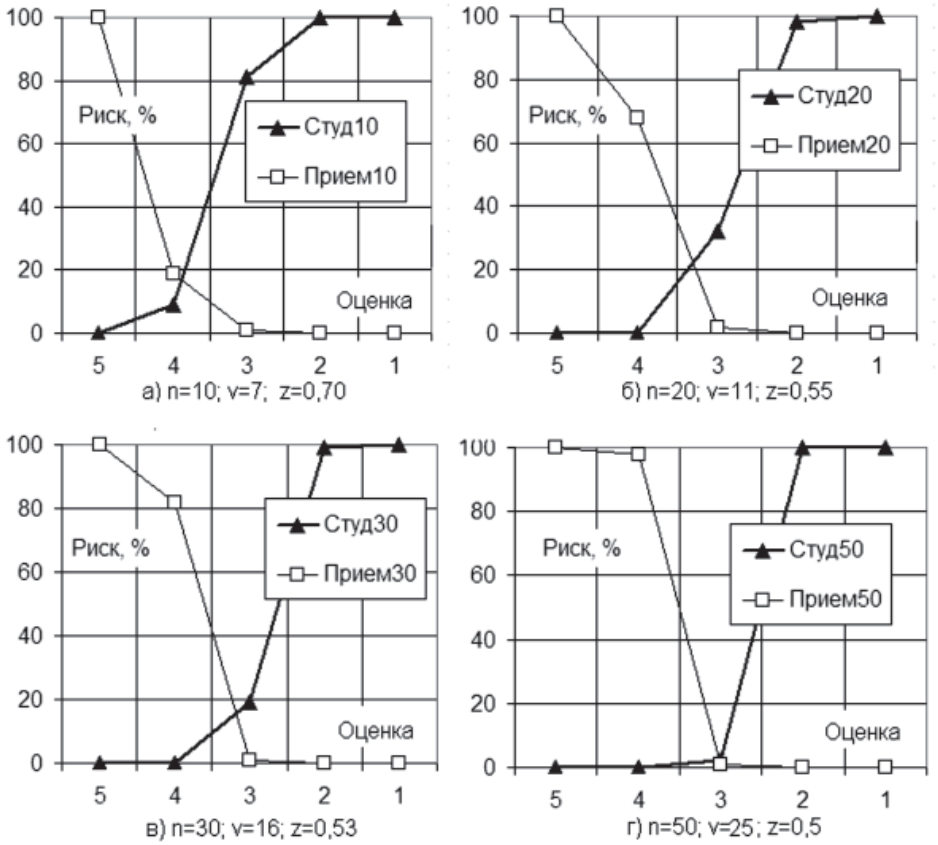

Рис. 1. Результаты расчета вероятностей ошибок при определении оценок 5 (отл), 4 (хор) и т.д. в процессе выборочного контроля для количеств заданных вопросов n = 10 (а), 20 (б), 30 (в), 50 (г). Величина v – это текущее значение количество правильных ответов на n вопросов, z = v/n, относительный уровень знаний

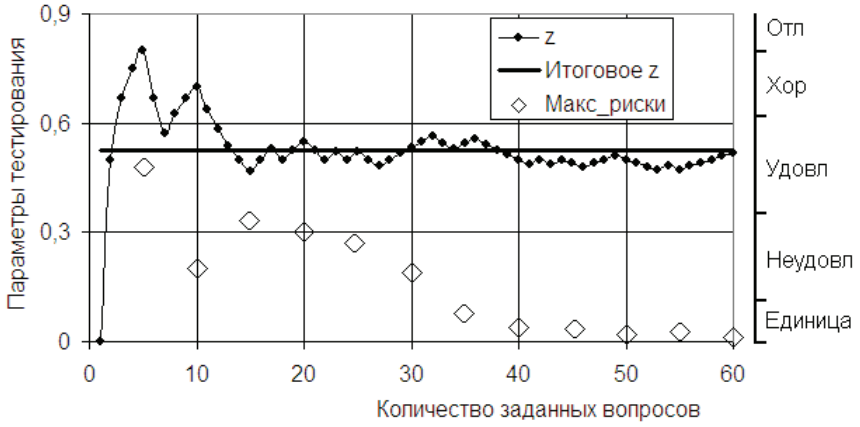

Рис. 2. Зависимости текущего значения зарегистрированного уровня знаний z = v/n и риска, определенного как максимум среди рисков студента и приемки, от n на этапе тестирования до n = 60

Наименее благоприятны с точки зрения последовательного контроля случаи, когда знания студента находятся у границы интервалов, для которых определены оценки с точностью до одного или нескольких правильных ответов. При этом процедура выборочного контроля затягивается, и хотя такие случаи не слишком часты, для достижения малых рисков приходится затягивать контроль до значений n, меньших максимального значения Q всего на несколько вопросов. Пример приведен на рис. 3 для описанной выше процедуры при $Q = 150$ и студента, уровень знаний которого $z = 0{,}34$ находится на нижней границе оценки «удовлетворительно».

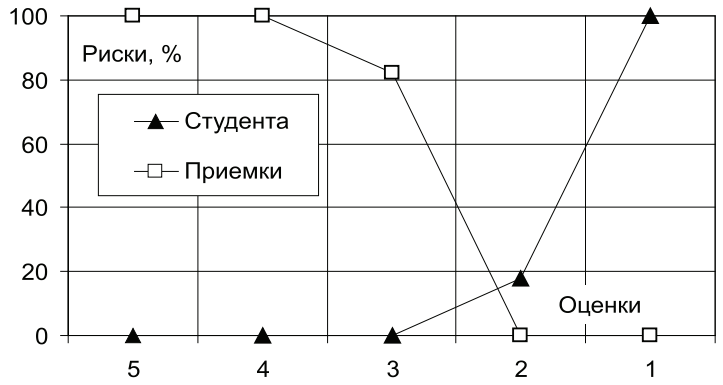

Рис. 3. Вероятности оценить неверно знания студента для оценок от 5 (отлично) до 1 (экстремально низкие знания) на этапе n = 100 с текущим значением z = 0,31, не соответствующим реальному

Для данной ситуации, несмотря на то, что большая часть вопросов студенту задана, полученному значению доле правильных ответов $z = 0{,}31$ (оценка 2 или «неудовлетворительно») доверять все еще нельзя, так как риск сделать ошибку в меньшую сторону еще достаточно велик (18%). При продолжении данной процедуры тестирования риски удалось свести к нулю только при $n = 148$, т.е. всего за два вопроса до окончания процедуры полной проверки всего набора иностранных слов.

Практика применения (проведено более 5000 испытаний) разработанного программного блока в учебном процессе показала, что при $24 \le Q \le 150$ более чем в 75% случаев удается сократить процедуру проверки уровней усвоения материала в два раза и более.

Список литературы

1. Горбань А.Н. Нейронные сети на персональном компьютере / А.Н. Горбань, Д. Н. Россиев. – Новосибирск: Наука. Сибирская издательская фирма РАН, 1996. – 276 с.

2. ГОСТ Р-50779.71-99 Статистические методы. Основные положения / Госстандарт России. – М.: ИПК Изд-во стандартов, 2001.

3. Ильенкова С.Д. Управление качеством: учебник для вузов / С.Д. Ильенкова, Н.Д. Ильенков, В.С. Мхитарян и др. / под ред. С.Д. Ильенковой. – М.: Банки и биржи, ЮНИТИ, 1998. – 199 с.

4. Качалов В.А. Стандарты ИСО 9000 и проблемы управления качеством в вузах (записки менеджера качества). – М., 2001. – 128 с.

5. Мишин В.М. Управление качеством: учебное пособие для вузов / В.М. Мишин. – М.: Юнити-дана, 2002. – 303 с.

References:

1. Gorban A. Neural network on a personal computer. Novosibirsk: Science (RAN), 1996. 276 p.

2. GOST-50779.71-99 Statistical methods. The main provisions: State Standard of Russia. – Moscow: Publishing IEC standards, 2001.

3. Ilyenkova S. Quality management. aid for students. Moscow: Banks and exchange units, 1998. 199 p.

4. Katchalov V. ISO 9000 quality management and issues in higher education (note manager of quality). Moscow: AT, 2001. 128 p.

5. Mishin V. Quality management. aid for students. MoscowUnity-dana, 2002. 303 p.

Рецензенты:

Данилаев П.Г., д.ф-м.н., профессор, заведующий кафедрой «Вычислительная математика» ФГБОУ ВПО «Казанский национальный исследовательский технический университет», г. Казань;

Муравьева Е.В., д.п.н., заведующая кафедрой «Промышленная экология» ФГБОУ ВПО «Казанский национальный исследовательский технический университет», г. Казань.

Работа поступила в редакцию 26.10.2012.