

УДК 519.2

doi:10.15217/issn1684-8853.2016.1.102

АНАЛИЗ ДОСТОВЕРНОСТИ ТЕСТОВОЙ ФОРМЫ КОНТРОЛЯ ЗНАНИЙ

Р. С. Кударов^а, канд. техн. наук, доцент

Ю. Сутугинене^а, PhD, старший преподаватель

^аПетербургский государственный университет путей сообщения Императора Александра I, Санкт-Петербург, РФ

Цель: разработать методику статистического анализа достоверности тестовой формы контроля знаний, отличающуюся от известных подходов наличием правила вычисления критериальных баллов теста по эмпирическим данным экзаменационных оценок. **Методы:** предложен вероятностно-статистический подход к проверке достоверности теста в виде методики последовательного выборочного анализа, позволяющей выносить суждения на заданном уровне значимости. **Результаты:** разработана методика оценки достоверности теста, включающая в себя корреляционный анализ результатов тестирования и экзаменационных оценок, проверку гипотезы о виде закона распределения результатов тестирования и построение правила вычисления критериальных баллов на основе эмпирических частот экзаменационных оценок. **Практическая значимость:** разработанная методика способствует научно обоснованному внедрению тестов в учебный процесс, показано ее успешное применение для оценки достоверности тестов по дисциплине «Методика преподавания русского языка как иностранного (РКИ)».

Ключевые слова — достоверность теста, метод выборочного исследования, корреляционный анализ, критерий хи-квадрат.

Введение

В настоящий момент в образовании сложилась ситуация, обеспечивающая сосуществование многообразных форм и методов контроля. Это устные опросы, письменные контрольные и самостоятельные работы, рефераты, курсовые, лабораторные и проектные работы, дневниковые записи, журналы наблюдений, коллоквиумы, семинары и индивидуальные собеседования, зачеты и экзамены. Каждая из форм имеет свои особенности, достоинства и недостатки, в заметной степени влияющие на результаты контроля. С начала XX в. к этим формам добавились тестовые методы, которые в разных странах то применялись, то отвергались, в зависимости от общей ситуации и числа сторонников этого метода. Дополнительно отметим, что широкое распространение компьютерных технологий привело к формированию методов оценки знаний, используемых в различных дистанционных (электронных) системах образования, например в системе Moodle.

В системе высшей школы на сегодняшний день все более востребованными и актуальными средствами контроля знаний учащихся выступают педагогические тесты. В отличие от традиционных средств контроля (устных и письменных экзаменов и зачетов по билетам, контрольных работ и пр.), именно педагогические тесты имеют научно обоснованные критерии качества [1, 2]. Кроме того, тесты обладают очевидными преимуществами перед другими методами педагогического контроля. В первую очередь, это эффективность и мобильность контроля при помощи

емких тестовых заданий, что выражается в минимальных затратах времени, усилий и средств, при этом индивидуализация контроля позволяет справедливо оценить каждого отдельного испытуемого по единым критериям [3]. Объективность и качество оценивания достигаются за счет стандартизации процедуры тестирования, а также стандартизации и проверки показателей качества заданий и тестов в целом. Отметим, что использование тестовых методик исключает возможность возникновения конфликтов из-за несогласия с экзаменационной оценкой, способствуя тем самым установлению доброжелательных отношений как между студентами, так и между преподавателем и студентами.

Однако внедрение тестов вместо контроля по экзаменационным билетам приобретает закономерный характер лишь при установленной высокой связи результатов теста с экзаменационными оценками. Следовательно, вопрос проверки достоверности такой замены становится особенно актуальным.

В настоящем исследовании сопоставление результатов двух контрольных процедур осуществлено на примере итогового экзамена по дисциплине «Методика преподавания русского языка как иностранного (РКИ)», проводимого на кафедрах русского языка в вузах Литвы (Литовского эдукологического университета, Вильнюсского университета), выпускающих бакалавров, будущих учителей РКИ. Всего в эксперименте приняли участие 100 студентов.

В преддверии данного экзамена, предполагающего ответы испытуемых на экзаменационные

вопросы, студенты прошли процедуру контроля по апробированным тестам. Таким образом был достигнут одинаковый уровень подготовки учащихся и к экзамену, и к тестированию, что в итоге позволило получить более точные результаты для дальнейшего их анализа.

Вероятностно-статистический подход к проверке достоверности теста

С математической точки зрения выстроим понятие достоверности разработанного теста в виде двух взаимодополняющих утверждений:

1. Результаты тестирования и экзаменационные оценки одной и той же группы студентов связаны прямой линейной связью.

2. Правило преобразования результатов тестирования в пятибалльную (или десятибалльную) систему учитывает реальное распределение экзаменационных (зачетных) оценок.

Первое утверждение позволяет избежать абсурдной ситуации, когда успешно выполнившие тест студенты получают низкие оценки на экзамене/зачете или наоборот. Следует добиться такой структуры и содержания теста, чтобы отклонение результатов тестирования от экзаменационных оценок в этом смысле стало незаметным.

Второе утверждение дает возможность избежать неопределенности и искажения информации при переходе от шкалы результатов тестирования к десятибалльной системе оценок. Правило преобразования баллов в оценки, сконструированное с учетом экспериментальных данных, максимально приближает идею тестирования к классическим процедурам проведения экзамена.

Если относительно разработанного теста имеют место оба утверждения, то этот тест будем считать достоверным: количественная констатация уровня знаний испытуемых (ранжирование), проводимая с помощью тестирования, аналогична работе экзаменатора. Конечно же, в этом случае полностью исключаются влияния случайности выбора билета с вопросами и «человеческого фактора» как в отношении студента, так и самого преподавателя.

Далее в работе будем полагать, что результат тестирования студента есть случайная величина X , а экзаменационная оценка того же студента по той же дисциплине есть случайная величина Y . Такое допущение оправдано тем, что на результаты тестирования и на экзаменационные оценки студентов оказывает влияние множество факторов, и ни при каких условиях точно предугадать результаты заранее нельзя.

На традиционном экзамене фактические знания оцениваются по умению их преподнести. Здесь сказывается уровень развития устной ре-

чи, активный словарный запас, кругозор, коммуникабельность, стрессоустойчивость и многие другие личностные качества экзаменуемого. Кроме того, влияние оказывает сложившееся впечатление экзаменатора о работоспособности обучающегося в течение учебного семестра. Не следует исключать и случайность выбора теоретических вопросов и выдачи практических задач на экзамене.

К факторам, влияющим на результаты тестирования, относятся следующие умения экзаменуемого: способность четко выполнять задания, понимать их формулировку, строго и грамотно оформлять свою работу; сосредоточенность и внимание (невозможность исправлять ошибки). Только при хорошей самоорганизации обучающегося возможна длительная работа над тестом, содержащим определенное количество заданий по разным темам.

Вместе с тем, независимо от формы контроля знаний, результат аттестационной процедуры всегда зависит от ряда обстоятельств, не относящихся к учебному процессу, но влияющих на психофизиологическое состояние студента (ситуация в личной жизни).

Случайный характер результатов теста и экзаменационных оценок дает возможность применять классические методы теории вероятностей и математической статистики.

Методы теории вероятностей и математической статистики, применяемые для обоснования достоверности разработанного теста

При исследовании связи между результатами тестирования и экзаменационными оценками целесообразно применять следующие элементы корреляционного анализа: вычисление выборочного коэффициента корреляции [4], оценку значимости выборочного коэффициента корреляции [4], вычисление коэффициента детерминации [5].

Вид закона распределения результатов тестирования рекомендуется устанавливать по критерию хи-квадрат Пирсона [6–9].

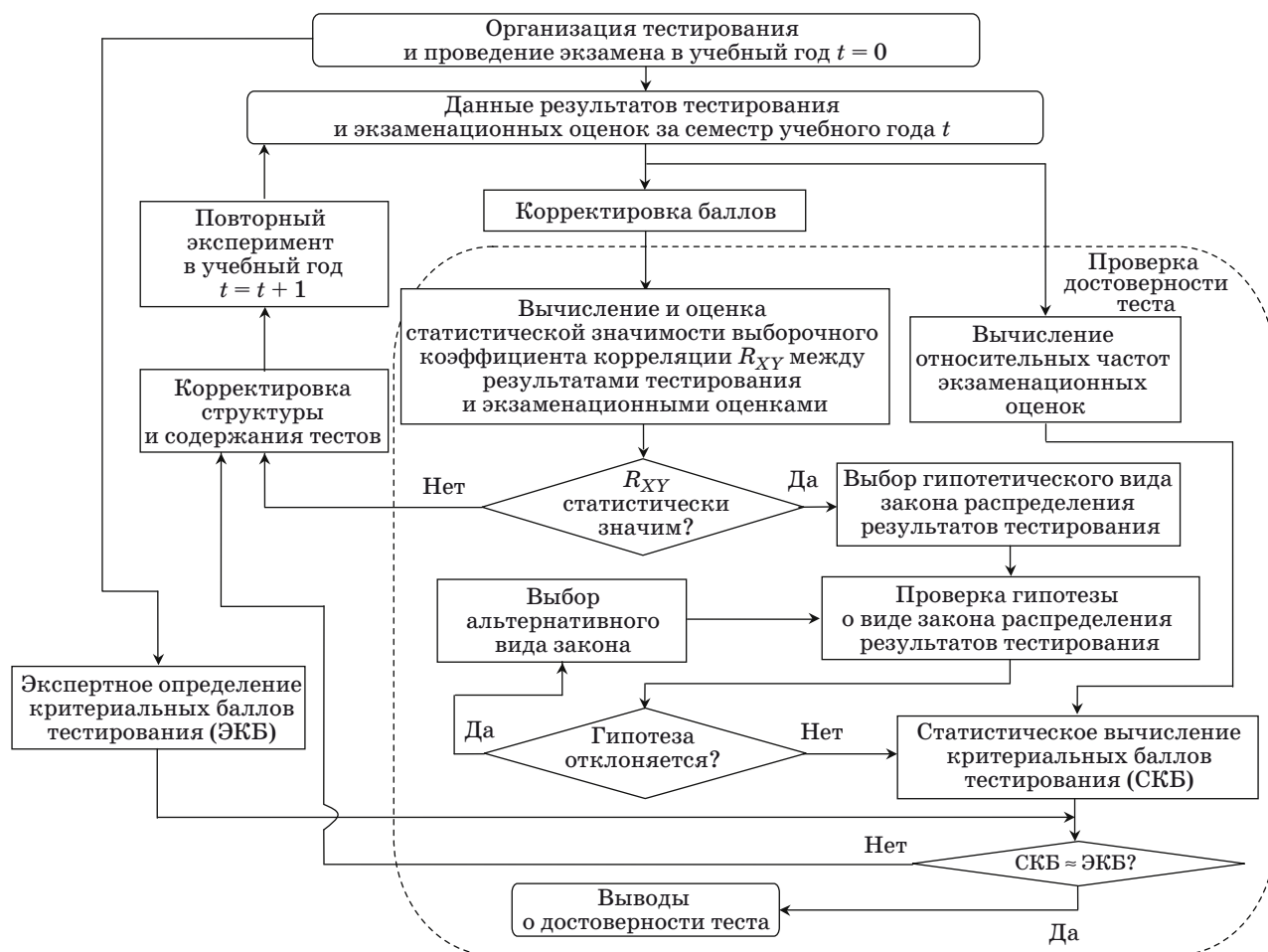
Удовлетворительные результаты использования описанных выше статистических методов позволяют построить правило вычисления критериальных баллов на основе сопоставления экспериментальных частот экзаменационных оценок с установленным теоретическим законом результатов теста.

Частоты экзаменационных оценок n_i представляют собой количество студентов, сдавших экзамен на i баллов. Следовательно, относительная частота n_i/n (где n — количество студентов) есть доля сдавших экзамен на i баллов студентов от общего их числа.

■ Таблица 1. Эмпирические критериальные баллы

| ЭО* <i>i</i> | Частота ЭО | Относительная частота ЭО | Накопленная относительная частота ЭО | Квантиль закона | Критериальный балл |
|--|------------|--------------------------|---------------------------------------|-----------------|--------------------|
| Пятибалльная экзаменационная оценка | | | | | |
| 2 | n_2 | n_2/n | $k_2 = n_2/n$ | x_{k2} | — |
| 3 | n_3 | n_3/n | $k_3 = n_2/n + n_3/n$ | x_{k3} | $[x_{k3}]$ |
| 4 | n_4 | n_4/n | $k_4 = n_2/n + n_3/n + n_4/n$ | x_{k4} | $[x_{k4}]$ |
| 5 | n_5 | n_5/n | $k_5 = n_2/n + n_3/n + n_4/n + n_5/n$ | x_{k5} | $[x_{k5}]$ |
| Десятибалльная экзаменационная оценка | | | | | |
| 4 | n_4 | n_4/n | $k_4 = n_4/n$ | x_{k4} | — |
| 5 | n_5 | n_5/n | $k_5 = n_4/n + n_5/n$ | x_{k5} | $[x_{k5}]$ |
| ... | ... | ... | ... | ... | ... |
| 9 | n_9 | n_9/n | $k_9 = n_4/n + n_5/n + \dots + n_9/n$ | x_9 | $[x_{k9}]$ |
| 10 | n_{10} | n_{10}/n | 1 | — | $[x_{k10}]$ |

* ЭО — экзаменационная оценка.



■ Рис. 1. Блок-схема статистической проверки достоверности тестовой формы контроля знаний

Используя накопленные эмпирические относительные частоты $k_i = \sum_{r=4}^i n_r / n$, эту закономерность распределения экзаменационных оценок перенесем на теоретический закон результатов тестирования с помощью понятия квантиля x_{ki} [10]:

$$P(X < x_{ki}) = k_i,$$

где X — случайная величина результата теста (в баллах); x_{ki} — квантиль порядка k_i .

Тогда округленный до целого значения квантиль $[x_{ki}]$ будет представлять собой нижний критериальный балл для оценки i . То есть предлагается правило, которое работает следующим образом: вероятность того, что случайно выбранный результат теста будет меньше критериального балла $[x_{ki}]$, равна накопленной частоте k_i . Тем самым ранжирование испытуемых по знаниям, произведенное с помощью тестирования, оказывается близким к аналогичному ранжированию на экзамене.

Вычисление критериальных баллов на основе эмпирических данных экзаменационных оценок с округлением до целого числа по двум экзаменационным оценкам представляется в виде табл. 1.

В настоящем подходе рекомендуется согласовывать статистически полученные критериальные баллы $[x_{ki}]$ с экспертными критериальными баллами (ЭКБ_{*i*}) согласно следующему соотношению:

$$[x_{ki}] \in (\text{ЭКБ}_i - 10\% \cdot x_{\max}; \text{ЭКБ}_i + 10\% \cdot x_{\max}),$$

где $[x_{ki}]$ — округленный до целого значения нижний критериальный балл для оценки i ; (ЭКБ_{*i*}) — нижний экспертный критериальный балл для оценки i ; x_{\max} — максимальный скорректированный результат теста.

Представим вероятностно-статистический подход к проверке достоверности теста с помощью описанных выше статистических процедур в виде блок-схемы (рис. 1).

Реализация разработанной методики к тестам по дисциплине «Методика преподавания РКИ»

Исходные данные скорректированных баллов тестирования и экзаменационных оценок сведем в табл. 2.

Вычислив средние выборочные величины $\bar{x} = 18,7$ и $\bar{y} = 6,63$, получим значение выборочного коэффициента корреляции: $R_{XY} = 0,87$.

Найденное значение $R_{XY} = 0,87$ близко к единице и свидетельствует о высокой степени прямой линейной связи между скорректированными баллами и экзаменационными оценками студентов, прошедших контрольные испытания.

Поскольку выборочное значение $R_{XY} = 0,87$ больше критического $R_{0,05} = 0,80$, то принимается предположение о статистической значимости выборочного коэффициента корреляции: можно утверждать, что связь между скорректированными баллами и экзаменационной оценкой близка к прямой линейной.

Учитывая очевидную тематическую связь между процедурой тестирования и экзаменом, а также их непосредственную последовательность во времени, имеет смысл проведенный анализ дополнить вычислением коэффициента детерминации: $R_{XY}^2 = (R_{XY})^2 = (0,87)^2 = 0,76$.

Коэффициент детерминации показывает, что вариация экзаменационных оценок на 76 % обуславливается результатами тестирования, что может свидетельствовать о высокой степени аналогичности двух аттестационных процедур.

При использовании критерия χ^2 выдвинем нулевую гипотезу H_0 , состоящую в том, что статистические данные соответствуют нормальному закону распределения. В качестве альтернативной гипотезы H_1 примем предположение о том, что статистические данные не соответствуют нормальному закону распределения.

Произведем группировку вариационного ряда таким образом, чтобы в каждой группе оказались примерно равные частоты. Результаты равномерной группировки скорректированных баллов приведем в табл. 3.

По таблице интегральной функции $\Phi_0(x)$ Лапласа восстановим теоретические значения

вероятностей $p_j = \Phi_0\left(\frac{x_{j,\max} - \bar{x}}{s}\right) - \frac{1}{2}$ нормального распределения $N(18,70;159,71)$. Обозначив кратко $b_j = \frac{x_{j,\max} - \bar{x}}{s}$, сведем результаты расчетов в табл. 4.

Вычисленные частоты n_j и приведенные вероятности np_j нормального закона запишем в табл. 5.

По данным табл. 5 вычислим расчетную величину хи-квадрат:

$$\begin{aligned} \chi_{\text{расчет}}^2 &= \sum_{j=1}^8 \frac{(n_j - np_j)^2}{np_j} = \\ &= \frac{(14 - 13,392)^2}{13,392} + \dots + \frac{(12 - 6,339)^2}{6,339} = 9,72. \end{aligned}$$

На уровне статистической значимости $\alpha = 0,05$, при количестве групп $k = 8$ и числе параметров гипотетического закона $l = 2$ по таблице [9] квантилей распределения хи-квадрат найдем критическую величину $\chi_{1-\alpha}^2(k-l-1)$:

$$\chi_{1-0,05}^2(8-2-1) = \chi_{0,95}^2(5) = 11,07.$$

Поскольку расчетное $\chi_{\text{расчет}}^2 = 9,72$ не превосходит критическое $\chi_{0,95}^2(5) = 11,07$, нет основа-

■ **Таблица 2.** Исходные данные скорректированных баллов (СБ) и экзаменационных оценок (ЭО)

| № п/п <i>i</i> | СБ x_i | ЭО y_i | № п/п <i>i</i> | СБ x_i | ЭО y_i | № п/п <i>i</i> | СБ x_i | ЭО y_i | № п/п <i>i</i> | СБ x_i | ЭО y_i |
|-------------------|-------------|-------------|-------------------|-------------|-------------|-------------------|-------------|-------------|-------------------|-------------|-------------|
| 1 | -8,75 | 4 | 26 | 8,8 | 6 | 51 | 18 | 5 | 76 | 26,5 | 8 |
| 2 | -8 | 4 | 27 | 11,3 | 4 | 52 | 18 | 7 | 77 | 27 | 8 |
| 3 | -8 | 4 | 28 | 11,7 | 6 | 53 | 18,2 | 7 | 78 | 28,5 | 7 |
| 4 | -5,5 | 4 | 29 | 11,4 | 6 | 54 | 18,3 | 5 | 79 | 28,8 | 8 |
| 5 | -4 | 4 | 30 | 11,5 | 5 | 55 | 18,3 | 7 | 80 | 30,8 | 8 |
| 6 | -3 | 4 | 31 | 11,5 | 6 | 56 | 18,8 | 7 | 81 | 31,8 | 8 |
| 7 | -2,7 | 5 | 32 | 11,7 | 6 | 57 | 19,7 | 7 | 82 | 31,8 | 7 |
| 8 | 3,7 | 4 | 33 | 11,5 | 6 | 58 | 19,7 | 7 | 83 | 32,6 | 8 |
| 9 | 4,1 | 5 | 34 | 11,8 | 6 | 59 | 20 | 6 | 84 | 32,7 | 8 |
| 10 | 4,3 | 6 | 35 | 12,3 | 6 | 60 | 20,6 | 7 | 85 | 32,8 | 8 |
| 11 | 4,5 | 4 | 36 | 12,5 | 7 | 61 | 21,6 | 7 | 86 | 32,9 | 9 |
| 12 | 4,5 | 5 | 37 | 13,3 | 6 | 62 | 21,7 | 7 | 87 | 37,55 | 8 |
| 13 | 4,7 | 5 | 38 | 14 | 6 | 63 | 22,3 | 7 | 88 | 37,8 | 9 |
| 14 | 4,7 | 6 | 39 | 14,3 | 6 | 64 | 22,3 | 6 | 89 | 38,2 | 6 |
| 15 | 4,8 | 5 | 40 | 14,5 | 7 | 65 | 22,8 | 7 | 90 | 38,7 | 8 |
| 16 | 5 | 5 | 41 | 14,7 | 6 | 66 | 23 | 7 | 91 | 39,2 | 9 |
| 17 | 5,3 | 4 | 42 | 15 | 6 | 67 | 23,6 | 7 | 92 | 39,3 | 9 |
| 18 | 5,5 | 5 | 43 | 15 | 9 | 68 | 23,9 | 7 | 93 | 39,5 | 10 |
| 19 | 5,6 | 5 | 44 | 16,2 | 6 | 69 | 24,7 | 7 | 94 | 40,3 | 9 |
| 20 | 6,2 | 5 | 45 | 16,2 | 6 | 70 | 25 | 5 | 95 | 40,5 | 9 |
| 21 | 6,5 | 5 | 46 | 16,8 | 8 | 71 | 25,3 | 8 | 96 | 41 | 9 |
| 22 | 6,7 | 5 | 47 | 17,2 | 7 | 72 | 25,5 | 8 | 97 | 41,8 | 10 |
| 23 | 6,7 | 8 | 48 | 17,2 | 7 | 73 | 25,6 | 8 | 98 | 50,5 | 9 |
| 24 | 7,7 | 5 | 49 | 17,5 | 7 | 74 | 25,7 | 8 | 99 | 50,7 | 10 |
| 25 | 7,8 | 6 | 50 | 17,8 | 7 | 75 | 26,5 | 7 | 100 | 54 | 10 |

■ **Таблица 3.** Частоты равномерно сгруппированных скорректированных баллов

| Показатель | <i>j</i> | | | | | | | |
|---|-----------|------|-------|-------|-------|-------|-------|-----------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| $x_{j,\min}$, нижняя граница интервала | $-\infty$ | 4,70 | 9,50 | 14,50 | 18,20 | 23,00 | 28,50 | 38,00 |
| $x_{j,\max}$, верхняя граница интервала | 4,70 | 9,50 | 14,50 | 18,20 | 23,00 | 28,50 | 38,00 | $+\infty$ |
| n_j , частота скорректированного балла | 14 | 13 | 13 | 13 | 13 | 12 | 10 | 12 |

■ **Таблица 4.** Вероятности нормального закона распределения

| Показатель | <i>j</i> | | | | | | | |
|---------------------------------------|----------|--------|--------|--------|-------|-------|-------|-----------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| b_j | -1,108 | -0,728 | -0,340 | -0,056 | 0,300 | 0,617 | 1,527 | $+\infty$ |
| $\Phi_0(b_j)$ | -0,366 | -0,267 | -0,133 | -0,022 | 0,118 | 0,231 | 0,437 | 0,5 |
| $p_j = \Phi_0(b_{j-1}) - \Phi_0(b_j)$ | 0,134 | 0,099 | 0,134 | 0,111 | 0,140 | 0,113 | 0,205 | 0,063 |

■ Таблица 5. Приведенные вероятности кривой нормального закона распределения

| Показатель | <i>j</i> | | | | | | | |
|------------|----------|-------|--------|--------|--------|--------|--------|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| n_j | 14 | 13 | 13 | 13 | 13 | 12 | 10 | 12 |
| np_j | 13,392 | 9,931 | 13,350 | 11,107 | 14,026 | 11,329 | 20,524 | 6,339 |

■ Таблица 6. Частоты и относительные частоты вариантов скорректированных баллов

| Показатель | <i>j</i> | | | | | | |
|---|----------|------|------|------|------|------|------|
| | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| n_j , частота экзаменационной оценки | 10 | 16 | 21 | 24 | 16 | 9 | 4 |
| n_j/n , относительная частота экзаменационной оценки | 0,1 | 0,16 | 0,21 | 0,24 | 0,16 | 0,09 | 0,04 |

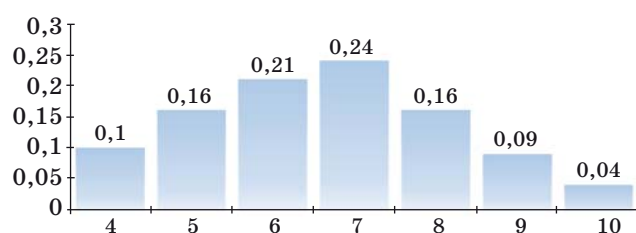
■ Таблица 7. Критериальные баллы

| ЭО | Частота ЭО | Относительная частота ЭО | Накопленная относительная частота ЭО | Квантиль нормального закона $N(18,90;159,71)$ | Критериальный балл |
|----|------------|--------------------------|--------------------------------------|---|--------------------|
| 4 | 10 | 0,1 | 0,1 | 2,51 | — |
| 5 | 16 | 0,16 | 0,26 | 10,57 | 3 |
| 6 | 21 | 0,21 | 0,47 | 17,75 | 11 |
| 7 | 24 | 0,24 | 0,71 | 25,70 | 18 |
| 8 | 16 | 0,16 | 0,87 | 32,94 | 26 |
| 9 | 9 | 0,09 | 0,96 | 40,83 | 33 |
| 10 | 4 | 0,04 | 1 | — | 41 |

ния отклонять гипотезу о соответствии распределения скорректированных баллов нормальному закону $N(18,70;159,71)$.

По данным экзаменационных оценок, содержащихся в табл. 2, построим гистограмму относительных частот. Для этого вычислим частоту n_j каждой экзаменационной оценки (j принимает целые значения от 4 до 10) и вычислим их относительные частоты n_j/n (табл. 6).

Полученные относительные частоты n_j/n экзаменационных оценок изобразим на рис. 2 в виде гистограммы.



■ Рис. 2. Относительные частоты экзаменационных оценок

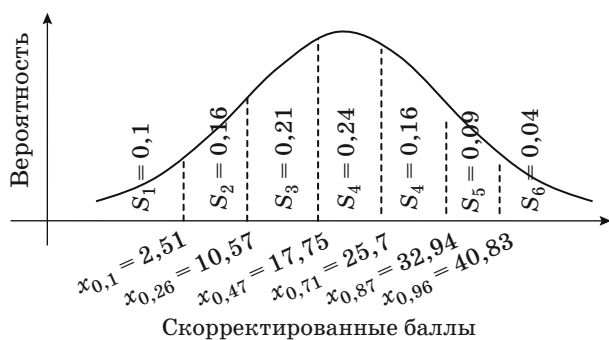
Зная относительные частоты экзаменационных оценок, получим критериальные баллы для результатов тестирования. Критериальные баллы вычислим как квантили закона распределения $N(18,70;159,71)$ скорректированных баллов с округлением до целого числа. Графически такие квантили разбивают фигуру, заключенную под кривой нормального закона скорректированных баллов, на части, площади которых совпадают с соответствующими относительными частотами экзаменационных оценок (рис. 3).

Вычисление критериальных баллов на основе эмпирических данных экзаменационных оценок с округлением до целого числа представим в табл. 7.

Нижний критериальный балл для оценки 5, посчитанный статистически, равен 3 (т. е. $[x_{k5}]$), а экспертный критериальный балл для этой оценки равен 7,82. Максимальный результат теста $x_{\max} = 54$:

$$[x_{k5}] \in (7,82 - 10\% \cdot 54; 7,82 + 10\% \cdot 54),$$

$$[x_{k5}] \in (2,42; 13,22).$$



■ **Рис. 3.** Квантили нормального закона $N(18,70;159,71)$



■ **Рис. 4.** Частоты экзаменационных оценок и результатов теста

Нижний критериальный балл для оценки 10, посчитанный статистически, равен 41 (т. е. $[x_{k10}]$), а экспертный критериальный балл для этой оценки равен 40,74:

$$[x_{k10}] \in (40,74 - 10\% \cdot 54; 40,74 + 10\% \cdot 54),$$

$$[x_{k10}] \in (35,34; 46,14).$$

Следовательно, можно сделать вывод, что критериальные баллы, полученные статистически, согласуются с экспертными критериальными баллами ЭКБ_i.

Воспользуемся полученными критериальными баллами для преобразования скорректированных баллов тестирования в десятибалльную шкалу и сравним полученные результаты графически с гистограммой относительных частот экзаменационных оценок на рис. 4.

Построенное авторами правило обработки результатов тестирования демонстрирует вполне приемлемые показатели — очевидна схожесть гистограммы преобразованных баллов тестирования с гистограммой эмпирических частот экзаменационных оценок.

Таким образом, тесты по дисциплине «Методика преподавания РКИ» рекомендуются к внедрению в учебный процесс в аттестационных целях вместо экзамена по билетам, а выстроенный подход к обоснованию адекватности теста, наглядно представленный выше в виде схемы с указанием последовательных действий, успешно можно применять для разработки новых тестовых батарей.

Заключение

Разработанная методика статистического анализа достоверности тестовой формы контроля знаний значительно упрощает работу по внедрению тестов в учебный процесс. Отметим, что построенный подход согласуется с известной теорией IRT и является относительно этой теории более простым алгоритмом, позволяющим с заданной точностью и минимальными затратами добиться поставленной цели. В качестве преимуществ предлагаемого подхода относительно IRT можно отметить сравнительно низкие затраты на его введение в учебный процесс: формирование команды специалистов по внедрению подхода, ожидание вычислительных процедур обработки результатов, наглядность результатов внедрения подхода и др.

Сама форма представления методики допускает автоматизацию. В перспективе предполагается создание соответствующего компьютерного программного обеспечения. Отличие такого программного обеспечения от существующих прикладных статистических пакетов (SPSS, Statistica др.) заключается в легкости его освоения, а также в минимальных требованиях к пользователям и технике. Умение по входным данным (результатам тестирования и экзаменационным оценкам) оперативно получать информацию о достоверности теста, вычислять критериальные баллы теста, полностью согласующиеся с границами экзаменационных оценок, представляет неограниченную пользу при внедрении любого банка тестов в различные учебные заведения.

Литература

1. Аванесов В. С. Основы научной организации педагогического контроля в высшей школе. — М.: НИТУ «МИСиС», 1989. — 167 с.
2. Анастаси А., Урбина С. Психологическое тестирование. — СПб.: Питер, 2006. — 688 с.
3. Ким В. С. Тестирование учебных достижений: монография. — Уссурийск: Изд-во УГПИ, 2007. — 214 с.
4. Кобзарь А. И. Прикладная математическая статистика. Для инженеров и научных работников. — М.: Физматлит, 2006. — 816 с.
5. Айвазян С. А., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: Исследование зависимос-

- тей: справ. изд. / под ред. С. А. Айвазяна. — М.: Финансы и статистика, 1985. — 487 с.
6. Вадзинский Р. Н. Справочник по вероятностным распределениям. — СПб.: Наука, 2001. — 295 с.
 7. Handbook of Applicable Mathematics. Vol. VI: Statistics. Part A. /Ch. Ed. Walter Ledermann, Ed. Emlyn Lloyd. — John Wiley & Sons, 1984. — 366 p.
 8. Handbook of Applicable Mathematics. Vol. VI: Statistics. Part B. /Ch. Ed. Walter Ledermann, Ed. Emlyn Lloyd. — John Wiley & Sons, 1984. — 522 p.
 9. Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. — М.: Наука. Гл. ред. физ.-мат. лит., 1983. — 416 с.
 10. Вероятность и математическая статистика: энциклопедия / под ред. Ю. В. Прохорова. — М.: Большая Российская энциклопедия, 2003. Репр. изд. — 912 с.

UDC 519.2

doi:10.15217/issn1684-8853.2016.1.102

Analysis of Validity of Test Forms of Knowledge ControlKudarov R. S.^a, PhD, Tech., Associate Professor, krs.83@mail.ruSutuginiene J.^a, PhD, Senior Lecturer, sutuginiene@mail.ru^aPetersburg State Transport University, 9, Moskovsky Pr., 190031, Saint-Petersburg, Russian Federation

Purpose: The goal is to develop a technique for statistical analysis of validity of the test form of knowledge control which would differ from the available approaches by having a rule for cut-off score calculation according to the empirical data of the exam scores. **Methods:** A statistical-probabilistic approach was proposed for test validity control as a method of selective analysis which allows you to make judgements according to the specified level of significance. **Results:** A technique has been developed for test validity evaluation. It includes correlation analysis of the test results and exam scores, verification of the hypothesis of the test results distribution law and the construction of the rule of calculating cut-off scores based on the empirical frequency of the exam scores. **Practical relevance:** The developed technique promotes scientifically grounded introduction of test forms into the educational process. The article demonstrates the successful usage of this technique for the evaluation of test validity in the course 'Didactics and Teaching Russian as a Foreign Language'.

Keywords — Test Validity, Selective Research, Correlation Analysis, Chi-Square Criterion.

References

1. Avanesov V. S. *Osnovy nauchnoi organizatsii pedagogicheskogo kontrolya v vysshei shkole* [The Basis of Scientific Organization of Pedagogical Control in Higher School]. Moscow, Natsional'nyi issledovatel'skii tekhnologicheskii universitet «MISiS» Publ., 1989. 167 p. (In Russian).
2. Anastazi A., Urbina S. *Psikhologicheskoe testirovanie* [Psychological Testing]. Saint-Petersburg, Piter Publ., 2006. 688 p. (In Russian).
3. Kim V. S. *Testirovanie uchebnykh dostizhenii* [Testing of the Education Achievements]. Ussuriysk, UGPI Publ., 2007. 214 p. (In Russian).
4. Kobzar A. I. *Prikladnaia matematicheskaia statistika* [Applied Mathematical Statistics]. Moscow, Fizmatlit Publ., 2006. 816 p. (In Russian).
5. Aivazian S. A., Eniukov I. S., Meshalkin L. D. *Prikladnaia statistika: Issledovanie zavisimostei* [Applied Statistics: Dependence Research]. Moscow, Finansy i statistika Publ., 1985. 487 p. (In Russian).
6. Vadzinskii R. N. *Spravochnik po veroiatnostnym raspredeleniyam* [Reference Book on Probabilistic Distribution]. Saint-Petersburg, Nauka Publ., 2001. 295 p. (In Russian).
7. *Handbook of Applicable Mathematics. Vol. VI. Statistics. Part A.* Ch. Ed. Walter Ledermann, Ed. Emlyn Lloyd. John Wiley & Sons, 1984. 366 p.
8. *Handbook of Applicable Mathematics. Vol. VI. Statistics. Part B.* Ch. Ed. Walter Ledermann, Ed. Emlyn Lloyd. John Wiley & Sons, 1984. 522 p.
9. Bol'shev L. N., Smirnov N. V. *Tablitsy matematicheskoi statistiki* [Tables of Mathematical Statistics]. Moscow, Nauka Publ., 1983. 416 p. (In Russian).
10. *Veroiatnost' i matematicheskaia statistika: Entsiklopediia* [Probability and Mathematical Statistics: Encyclopaedia]. Ed. by Iu. V. Prokhorov. Moscow, Bol'shaia Rossiiskaia entsiklopediia Publ., 2003. 912 p. (In Russian).