

УДК 004.62:519.2

С.А. РАКОВ¹, М.С. МАЗОРЧУК², Е.О. БОНДАРЕНКО²¹ Український центр оцінювання якості освіти, Харків² Національний аерокосмічний університет ім. Н.Е. Жуковського «ХАІ»

УТОЧНЕНИЕ РЕЗУЛЬТИРУЮЩИХ БАЛЛОВ ТЕСТИРОВАНИЯ НА ОСНОВЕ МОДИФИЦИРОВАННОГО АЛГОРИТМА ЭКВИПРОЦЕНТИЛЬНОЙ НОРМАЛИЗАЦИИ

Целью работы является разработка метода уточнения баллов, полученных в ходе тестирования учащихся по нормативно-ориентированным тестам, которые применяются при массовом оценивании знаний. В процессе шкалирования возникают проблемы, связанные с неоднозначным переводом баллов в новую шкалу, что обусловлено дискретностью значений и малой популяцией на концах распределения. В работе предлагается алгоритм формирования модифицированной выборки, что позволит выровнять пороговые значения и сократит разрыв в пунктах новой шкалы. Для проведения расчетов используются пакеты для анализа данных, такие как SPSS и MS Excel.

Ключевые слова: пороговые баллы; эквипроцентильная нормализация; модифицированная выборка.

Введение

На сегодняшний день оценка знаний в учебном процессе часто проводится в форме тестов. Тестирование – это оценка учебных достижений в форме тестов, т.е. системы заданий специфической формы, позволяющих оценить знания, умения и навыки по изучаемой дисциплине. Педагогические тесты классифицируются по различным признакам: по целям использования, по методам разработки тестовых заданий, по степени однородности заданий и т.д. [1, 2]. Для оценки учебных достижений учащихся при большой популяции чаще всего применяется нормативно-ориентированный подход к разработке тестов. Тест, созданный в рамках этого подхода, позволяет сравнивать уровень знаний и умений испытуемых друг с другом. Данный вид теста направлен на сопоставление результатов каждого испытуемого с результатами других, принимающих участие в том же тестировании. Таким образом, основной задачей нормативно-ориентированного теста является дифференциация испытуемых по уровню учебных достижений. Результаты тестирования, в этом случае, носят относительный характер.

Нормативно-ориентированные тесты предполагают установку нормативных требований к уровню подготовки на основе анализа тестирования представительной выборки учащихся. Оценка результатов тестирования проводится в установленной нормативной шкале, так называемых «сырых» баллах, а затем с помощью одного из методов шкалирования переводится в рейтинговую процентильную шкалу, например в шкалу 100-200. Метод эквипроцентиль-

ной нормализации (ЭПН) предполагает преобразование эмпирического распределения плотности частот «сырых» баллов в «эталонное» распределение, т.е. нормальное распределение с заданным математическим ожиданием и дисперсией [3, 4]. Как правило – это нормированное и центрированное нормальное распределение. Пороговые баллы, как правило, определяются экспертным путем [5].

В процессе анализа результатов тестирования возникают проблемы, связанные с дифференцированием учащихся, находящихся ниже и выше пороговых значений, отражающих «самых слабых» и «самых сильных» учащихся. При этом распределение тестовых баллов, даже в нормальной процентильной шкале, не дает возможность дифференцировать испытуемых и может давать довольно большой разброс процентильных баллов по сравнению с «сырым» баллом. Поэтому, актуальным является поиск эффективных методов шкалирования результатов, позволяющих соотнести пороговые баллы, полученные с помощью экспертов, и в шкале процентильных рангов (шкале 100-200).

1. Постановка задачи

Рассмотрим пример шкалирования результатов нормативно-ориентированного теста. На рис. 1 и рис.2 приведены диаграммы распределения тестовых «сырых» баллов (от 0 до 62) и результаты тестирования в шкале 100-200 по одной из нормативных дисциплин студентов университета. В тестировании приняло участие 1048 человек. Шкалирование результатов проводилось по эквипроцентильному методу с параметрами распределения $N(150, 20)$.

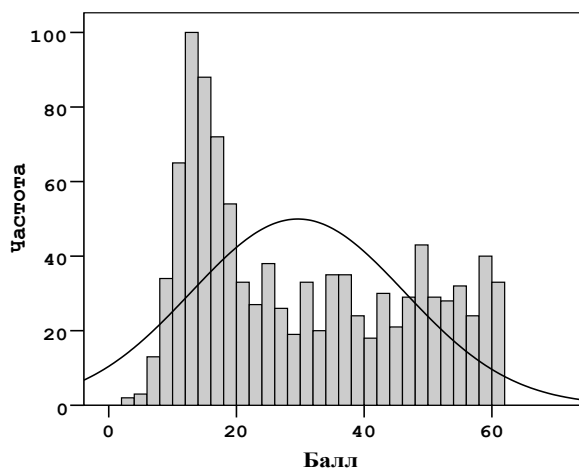


Рис. 1. Распределение «сырых» баллов

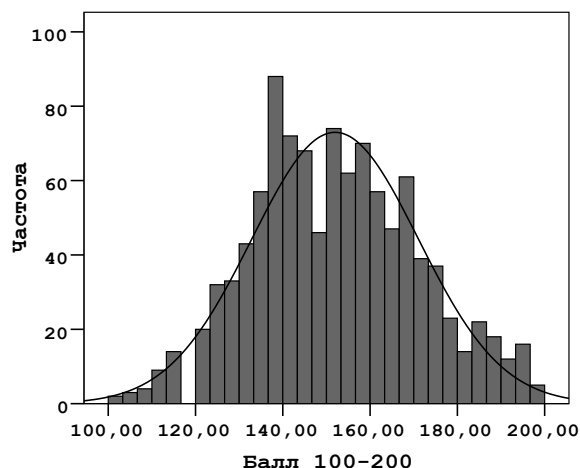


Рис. 2. Распределение баллов в шкале 100-200

При переводе в шкалу 100-200 на концах интервала один «сырой» балл соответствовал от полутора до 3,5 пунктов, тогда как в середине совокупно-

сти разница между баллами не превышала 0,5 пунктов. В таблице 1 приведена таблица перевода результатов тестирования.

Таблица 1

Перевод баллов по эквипроцентильному методу

«Сырой» балл	Балл 100-200	«Сырой» балл	Балл 100-200	«Сырой» балл	Балл 100-200	«Сырой» балл	Балл 100-200
0	100,0	17	143,0	34	157,0	51	172,0
1	100,5	18	144,5	35	157,5	52	173,0
2	101,0	19	146,0	36	158,5	53	174,0
3	102,0	20	146,5	37	159,0	54	175,5
4	103,5	21	147,5	38	160,0	55	177,5
5	105,5	22	148,5	39	160,5	56	179,0
6	108,0	23	149,0	40	161,0	57	181,5
7	111,5	24	150,0	41	162,0	58	184,0
8	116,0	25	150,5	42	162,5	59	187,5
9	120,5	26	151,5	43	163,5	60	191,5
10	124,0	27	152,0	44	164,5	61	196,0
11	127,5	28	153,0	45	165,5	62	200,0
12	131,0	29	153,5	46	166,0		
13	134,5	30	154,5	47	167,0		
14	137,5	31	155,0	48	168,5		
15	139,5	32	155,5	49	169,5		
16	141,5	33	156,0	50	170,5		

Такая ситуация может быть обусловлена прежде всего тем, что тест был недостаточно сложным, и большой процент участников набрал высокие баллы, а низкие баллы набрало меньшее число испытуемых. В большинстве тестов ниже и выше установленных экспертами пороговых значений доля участников не должна быть ниже или превышать, соответственно, 10%. А в случае с данным тестом, процент «слабых» составил около 3%, а сильных – порядка 16% (рис. 3).

Целью данной работы является разработка подхода, позволяющего минимизировать разницу между пунктами шкалы 100-200 на концах распределения.

Основными задачами исследования является разработка алгоритма формирования модифицированной выборки и проведение расчетов, позволяющих выявить соответствие между пороговыми значениями, полученными экспертами, и пороговыми точками в шкале 100-200, что позволит дифференцировать испытуемых, набравших низкие или высокие баллы.

Значение	Частота	Процент	Процент значений	Накопленный процент
2	1	,1	,1	,1
3	1	,1	,1	,2
5	3	,3	,3	,5
6	4	,4	,4	,9
7	9	,9	,9	1,7
8	14	1,3	1,3	3,1
9	20	1,9	1,9	5,0
10	32	3,1	3,1	8,0
11	33	3,1	3,1	11,2
12	12	1,1	1,1	12,3
13	17	1,6	1,6	13,9
14	12	1,1	1,1	15,0
15	10	1,0	1,0	16,0
16	18	1,7	1,7	17,7
17	19	1,8	1,8	19,5
18	13	1,2	1,2	20,7
19	10	1,0	1,0	21,7
20	14	1,3	1,3	23,0
21	22	2,1	2,1	25,1
22	18	1,7	1,7	26,9
23	12	1,1	1,1	28,0
24	16	1,5	1,5	29,5
25	5	,5	,5	30,0
Total	1048	100,0	100,0	

Рис. 3. Элементы таблицы частот распределения «сырых» баллов испытуемых

2. Алгоритм формирования модифицированной выборки

Для решения поставленных задач предлагается сформировать модифицированную выборку участников таким образом, чтобы соблюдались требования к пороговым значениям. Процент участников, которые набрали меньше нижнего порога (например, 8 «сырых» баллов) и больше верхнего порога (например, 50 «сырых» баллов), не превышал 10%. Формирование модифицированной выборки будет осуществляться по методу «квот», который широко применяется в социологических исследованиях [6]. Метод «квот» относится к детерминированным методам, при котором обеспечивается пропорциональное представительство носителей существенных признаков (пол, образование, место жительства и т.д.) генеральной совокупности в выборке. В случае с популяцией участников тестирования, существенным признаком является пороговый балл, который разбивает всю популяцию на пороговые группы. Размеры пороговых групп (фактически, распределение квот) устанавливаются экспертным путем на основе анализа предыдущих исследований или экспериментов. В каждую пороговую группу должно входить определенное число участников, составляющих определенный процент от всей популяции испытуемых. Если эмпирические данные не соответствуют желаемому распределению квот, то количество участников пороговых групп предлагается привести к требуемому количеству путем добавления или удаления испытуемых. Если речь идет только о крайних интервалах распределения, то участники добавляются или удаляются из популяции пропорционально набранным баллам. Таким образом, незначительное изменение количества участников популяции на концах распределения позволит выровнять пороговые значения и сократит разрыв в пунктах шкалы 100-200, особенно, если выборка является достаточно большой и репрезентативной.

Алгоритм формирования выборки заключается в следующем:

1. Определяют нижнюю и верхнюю критериальные точки – количество тестовых баллов, которые соответствуют верхней границе шкалированных баллов (175 баллов) и нижней границе шкалированных баллов (124 балла).

2. Формируется модифицированная выборка участников тестирования:

- 2.1. 200-балльники извлекаются из процедуры шкалирования;

- 2.2. Количество участников, которые набрали меньше или ровно 124 бала искусственно выравниваются до 10% популяции:

- 2.2.1. Если число участников пороговой группы (до 124 баллов или больше 175 баллов) меньше 10%, то следует добавить фиктивных участников, которые имеют тестовые баллы такие, чтобы отношение набранных «сырых» тестовых баллов соответствовало первичной выборке (например, 5 участников получили 1 балл, 6 – 2 балла, 10 – 3 балла и т.д. Добавляем соответственно до требуемого процента участников с такими же набранными тестовыми баллами в пропорциональной зависимости).

- 2.2.2. Если участников пороговой группы больше 10%, то следует удалить необходимое количество участников, которые имеют такие же баллы, как участники данной группы в пропорциональной зависимости.

Далее применяется метод ЭПН к модифицированной выборке и формируется таблица пересчета тестовых баллов. На основе построенной таблицы пересчета определяются баллы участников тестирования по исходной популяции.

3. Результаты расчетов

Осуществляем эксперимент, используя результаты тестирования, рассмотренные в постановке задачи данной статьи. Предположим, что эксперты определили нижнюю и верхнюю критериальные границы для «сырых» баллов – 8 и 50 соответствен-

но. Добавляем и удаляем пропорционально в крайних пороговых группах участников таким образом, чтобы число испытуемых в этих пороговых группах было равным на уровне $\approx 10\%$. При этом, процентное распределение остальных участников не должно существенно измениться. На рис. 4 и рис. 5

изображены частотные распределения тестовых баллов до и после модификации выборки.

Применяя метод ЭПН к данной выборке, строим новую таблицу пересчета «сырых» баллов (табл. 2). На последнем этапе осуществляем перевод результатов тестирования исходной популяции в шкалу 100-200 по новой таблице перевода.

	Частота	Процент	Процент значений	Накопленный процент
Значение 2	1	,1	,1	,1
3	1	,1	,1	,2
5	3	,3	,3	,5
6	4	,4	,4	,9
7	9	,9	,9	1,7
8	14	1,3	1,3	3,1
9	20	1,9	1,9	5,0
10	32	3,1	3,1	8,0
11	33	3,1	3,1	11,2
12	12	1,1	1,1	12,3
13	10	1,0	1,0	13,3
14	18	1,7	1,7	15,0
15	19	1,8	1,8	16,8
16	13	1,2	1,2	18,0
17	10	1,0	1,0	19,0
18	14	1,3	1,3	20,3
19	22	2,1	2,1	22,4
20	18	1,7	1,7	24,1
21	12	1,1	1,1	25,2
22	10	1,0	1,0	26,2
23	18	1,7	1,7	27,9
24	19	1,8	1,8	29,7
25	13	1,2	1,2	30,9
26	10	1,0	1,0	31,9
27	14	1,3	1,3	33,2
28	22	2,1	2,1	35,3
29	18	1,7	1,7	37,0
30	12	1,1	1,1	38,1
31	16	1,5	1,5	39,6
32	5	,5	,5	40,1
Total	1048	100,0	100,0	

Рис. 4. Частотные распределения тестовых баллов исходной популяции

	Частота	Процент	Процент значений	Накопленный процент
Значения 2	5	,5	,5	,5
3	6	,6	,6	1,0
5	18	1,7	1,7	2,8
6	16	1,5	1,5	4,3
7	28	2,7	2,7	7,0
8	42	4,0	4,0	11,0
9	20	1,9	1,9	12,9
10	32	3,0	3,0	15,9
...				
48	19	1,8	1,8	87,7
49	24	2,3	2,3	90,0
50	17	1,6	1,6	91,6
51	6	,6	,6	92,2
52	5	,5	,5	92,7
53	9	,9	,9	93,5
54	10	1,0	1,0	94,5
55	6	,6	,6	95,0
56	5	,5	,5	95,5
57	7	,7	,7	96,2
58	11	1,0	1,0	97,2
59	10	1,0	1,0	98,2
60	6	,6	,6	98,8
61	8	,8	,8	99,5
62	5	,5	,5	100,0
Total	1050	100,0	100,0	

Рис. 5. Частотные распределения тестовых баллов после модификации

Таблица 2

Перевод баллов по методу ЭПН по модифицированной выборке							
«Сырой» балл	Балл 100-200	«Сырой» балл	Балл 100-200	«Сырой» балл	Балл 100-200	«Сырой» балл	Балл 100-200
0	100,0	18	148,43	34	160,62	50	178,43
2	103,15	19	149,31	35	161,49	51	179,18
3	106,02	20	150,01	36	162,69	52	179,84
5	111,64	21	150,95	37	163,62	53	181,06
6	115,00	22	151,74	38	164,36	54	182,59
7	119,50	23	152,34	39	165,14	55	183,63
8	124,46	24	153,19	40	165,70	56	184,51
9	126,37	25	154,26	41	166,30	57	185,93
10	129,05	26	154,83	42	167,32	58	188,32
11	131,54	27	155,63	43	168,35	59	191,10
12	134,51	28	156,16	44	169,03	60	193,11
13	138,00	29	156,60	45	169,96	61	196,73
14	140,91	30	157,66	46	170,83	62	200,00
15	142,86	31	158,39	47	172,35		
16	145,14	32	158,44	48	174,08		
17	146,60	33	159,48	49	176,51		

Результаты перевода, полученные по модифицированной выборке, существенно не отличаются, от результатов, полученных по исходной выборке, однако сокращают разницу для участников, которые набрали меньше нижней критериальной границы и больше верхней критериальной границы. Результаты сравнения выборок в шкале 100-200 по исходной популяции и модифицированной выборки по критерию χ^2 - Пирсона (Chi-Square) существенно не отличаются (табл. 3). Как видно из таблицы, значения χ^2 одинаковы для исходного распределения и для модифицированного. Асимптотическая значимость (Asymp. Sig.) равна нулю. Это говорит о том, что различия между двумя распределениями незначительны.

Нет существенных различий и по дескриптивным статистикам (табл. 4). Произошел сдвиг по процентильным значениям, что и необходимо было для уточнения критериальных значений пороговых групп. Среднее значение (Mean) и среднее квадратичное отклонение (Std. Deviation) незначительно отличаются друг от друга, однако они не оказывают существенного влияния на значения баллов в сере-

дине распределения. Все расчеты проводились в SPSS, что позволило оперативно обработать данные и получить наглядные результаты.

Таблица 3

Статистика теста (Test Statistics)

	Балл 100-200	Балл 100-200 (New)
χ^2 (Chi-Square)	433,679	433,679
Число степеней свободы (df)	59	59
Асимптотическая значимость (Asymp. Sig.)	0,00	0,00

На рис. 6 и рис. 7 приведены диаграммы частотных распределений до и после модификации выборки. Из рисунков также видно, что различия в распределении частот незначительны.

Таблица 4

Дескриптивные статистики распределения (Descriptive Statistics)

	N	Среднее (Mean)	Среднее квадратичное отклонение (Std. Deviation)	Минимум (Minimum)	Максимум (Maximum)	Процентили (Percentiles)		
						25th	50th (Медиана (Median))	75th
Балл 100-200	1048	151,9814	19,09783	101,00	200,00	137,5000	150,5000	165,5000
Балл 100-200 (New)	1048	156,2809	19,17621	103,15	200,00	140,9100	154,2600	169,9600

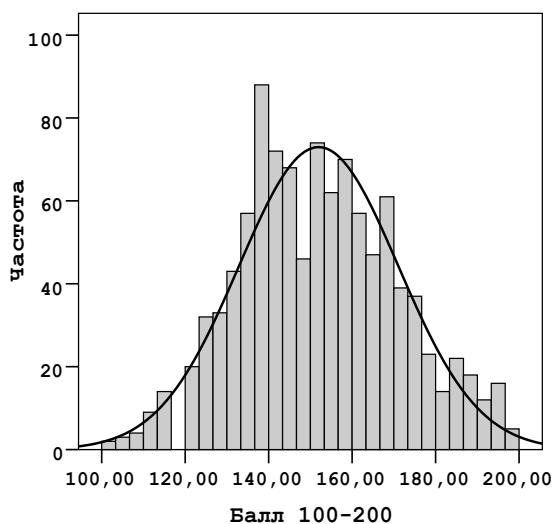


Рис. 6. Распределение частот по шкале 100-200 до модификации выборки

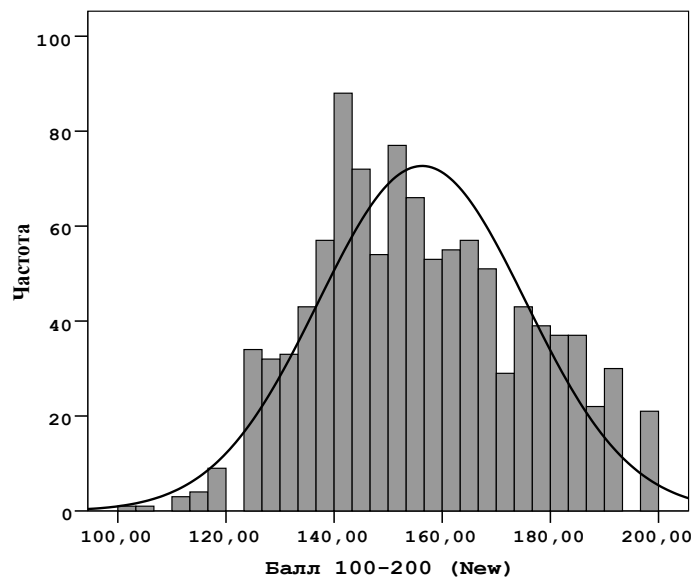


Рис. 7. Распределение частот по шкале 100-200 после модификации выборки

Выводы

Таким образом, предлагаемый подход позволяет уточнить результирующие баллы в соответствии с определенными экспертами пороговыми значениями «сырых» баллов. Данный подход также позволяет минимизировать разницу пунктов шкалы в крайних пороговых группах распределения популяции испытуемых.

Литература

1. Гронлунд Норман, Е. Оцінювання студентської успішності [Текст]: практ. посіб. / Е. Норман Гронлунд. – К.: Навчально-методичний центр «Консорціум із удосконалення менеджмент-освіти в Україні», 2005. – 312 с.
2. Чельщикова, М.Б. Теория и практика конструирования педагогических тестов [Текст]: учеб.

пособие / М.Б. Чельщикова. – М.: Логос, 2002. – 432 с.

3. Нейман, Ю.М. Введение в теорию моделирования и параметризации педагогических тестов [Текст] / Ю.М. Нейман, В.А. Хлебников. – М.: Прометей, 2000. – 169 с.

4. Введение в классическую и современную теорию тестов [Текст]: учебник / Л. Крокер, Дж. Алгина; под общей ред. В.И. Звонникова и М.Б. Чельщиковой. – М.: Логос, 2012. – 668 с.

5. Раков, С.А. Методи встановлення критеріальних порогів (теоретичні засади) [Текст] / С.А. Раков, Т.С. Вакуленко // Вісник. Тестування і моніторинг в освіті. 2012. – №7-8. – С. 72-79.

6. Горшков, М.К. Прикладная социология: методология и методы [Текст]: интерактивное учеб. пособие / М.К. Горшков, Ф.Э. Шереги – М.: ФГАНУ «Центр социологических исследований», Институт социологии РАН, 2012. – 404 с.

Поступила в редакцию 10.12.2012

Рецензент: д-р техн. наук, зав. каф. охраны труда, стандартизации и сертификации Р.М. Трищ, Украинская инженерно-педагогическая академия, Харьков.

УТОЧНЕННЯ РЕЗУЛЬТУЮЧИХ БАЛІВ ТЕСТУВАННЯ НА ОСНОВІ МОДИФІКОВАНОГО АЛГОРИТМУ ЕКВІПРОЦЕНТІЛЬНОЇ НОРМАЛІЗАЦІЇ

С.А. Раков, М.С. Мазорчук, О.О. Бондаренко

Метою роботи є розробка методу уточнення балів, отриманих в ході тестування учнів за нормативно-орієнтованими тестами, які застосовуються при масовому оцінюванні знань. У процесі шкалювання виникають проблеми, пов'язані з неоднозначним переводом балів у нову шкалу, що обумовлено дискретністю значень і малою популяцією на кінцях розподілу. У роботі пропонується алгоритм формування модифікованої вибірки, що дозволить вирівняти порогові значення і скоротить розрив у пунктах нової шкали. Для проведення розрахунків використовуються пакети для аналізу даних, такі як SPSS та MS Excel.

Ключові слова: порогові бали; еквіпроцентільна нормалізація; модифікована вибірка.

UPDATE RESULTANT POINTS TESTING ON BASE MODIFIED ALGORITHM EQUIPERCENTILE NORMALIZATION

S.A. Rakov, M.S. Mazorchuk, O.O. Bondarenko

The goal of this work is to develop a method for clarifying points obtained during the testing of students in the normative-oriented tests, which are used in estimating massive of knowledge. In the process of scaling problems associated with ambiguous transfer points in the new scale, due to discrete values and the small population at the ends of the distribution. We propose a modified algorithm of forming a sample that will align the cut-off points and close the gap in the points of the new scale. For the calculations used for data analysis packages such as SPSS and MS Excel.

Key words: cut-off points; equipercetile normalization; modified sample.

Раков Сергей Анатольевич – доктор педагогических наук, кандидат физико-математических наук, доцент, заведующий отделом научного обеспечения, Украинский центр оценивания качества образования, Харьков, e-mail: rakov_s@ukr.net

Мазорчук Мария Сергеевна – кандидат технических наук, доцент, доцент кафедры информатики Национального аэрокосмического университета им. Н.Е. Жуковского «ХАИ», Харьков, Украина, e-mail: mazorchuk_mary@inbox.ru

Бондаренко Елена Олеговна – аспирант кафедры информатики Национального аэрокосмического университета им. Н.Е. Жуковского «ХАИ», Харьков, Украина, e-mail: mango26.88@mail.ru