

Применение нормированного отклонения для шкалирования тестовых баллов испытуемых

Альмира Искакова,
Национальный центр тестирования,
Министерство образования Республики Казахстан
iskakova_am@mail.ru

В статье рассмотрена процедура шкалирования исходных тестовых баллов. Произведён расчёт основных характеристик теста. Шкалирование на научной основе предполагает перевод первичных тестовых баллов в баллы истинных значений результатов испытуемых.

Ключевые слова: шкалирование, первичный балл, дисперсия, отклонение

Определение

Шкалирование — это процедура перевода исходных баллов в тестовые, процесс формирования правил начисления тестовых баллов по результатам тестирования на основе статистических данных.

Шкалируемое свойство личности

Главное интересующее свойство личности, измеряемое в процессе педагогических измерений — это уровень подготовленности испытуемых. В определение этого свойства включаются знание учебной дисциплины, умения, навыки, представления и компетенции. Свойство — это внешнее выражение непосредственно не наблюдаемого, латентного качества личности. Взятые вместе, все перечисленные свойства можно представить как обобщённую совокупность свойств, образующих уровень подготовленности. Уровень подготовленности испытуемых — измеряемое свойство. Этот уровень измеряется посредством педагогического теста.

Сложение условно назначенных (т.е. заведомо неточных) баллов за выполнение заданий по каждому предмету не корректно без специального обоснования.

Шкалирование тестовых результатов рассматривается как метод трансформации тестовых баллов испытуемых в педагогические измерения, как главный завершающий этап процесса педагогических измерений. За шкалированием следуют лишь интерпретация и практическое применение результатов.

В данном исследовании представлено шкалирование теста, который состоит из 3-х форм тестовых заданий. Тестовые задания по формам распределены следующим образом: 20 заданий с выбором одного ответа (оцениваются 0 и 1), 15 тестовых заданий с выбором одного или нескольких ответов (оцениваются 0, 1, 2) и 5 тестовых заданий на установление соответствия (оцениваются 0, 1, 2). Максимально по каждому предмету

испытуемые могут набрать 60 баллов.

При выборе метода шкалирования наиболее часто обращаются к стандартным показателям, которые в свою очередь могут быть определены с помощью линейного и нелинейного преобразований первичных баллов, полученных испытуемыми по результатам выполнения теста. В силу линейного характера преобразований все свойства исходного распределения первичных баллов переносятся на множество шкалированных оценок. Стандартные показатели, найденные таким путём, обычно называют z -оценками, а результат отображения их на числовую ось Z -шкалой.

Для анализа z -оценок обычно рассматривается разность между первичным баллом и средним значением баллов по группе тестируемых студентов. Если их разность < 0 , то индивидуальный балл i -го испытуемого ниже среднего по группе, в противном случае его индивидуальный результат превышает средний по тесту.

Нормирование тестовых результатов

Из-за того, что среднее арифметическое по различным тестам в отличающихся группах студентов существенно различается между собой, возникает проблема сопоставимости отклонений. Один и тот же балл в слабой группе может оказаться выше среднего, а в сильной — значительно ниже. Решением данной проблемы может стать нормирование. Нормирование позволяет выразить отклонение индивидуальных результатов от среднего значения первичных баллов в единицах, пропорциональных стандартному отклонению.

Для перевода в Z -шкалу находят шкалированный результат i -го испытуемого:

$$Z_i = \frac{x_i - \bar{x}}{s_x},$$

x_i — исходный балл i -го испытуемого;
 \bar{x} — среднее значение индивидуальных баллов испытуемых группы;
 s_x — стандартное отклонение по множеству исходных баллов.

Вычисление z -оценок требует подсчёта стандартного отклонения s_x , которое равно квадратному корню

из дисперсии: $s_x = \sqrt{s_x^2}$. Дисперсия обычно обозначается символом s_x^2 и вычисляется по формуле [1]:

$$s_x^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}.$$

В табл. 1 произведён расчёт z -оценок по первому профильному предмету. Как свидетельствуют данные во втором столбце, имея исходные баллы, можно судить об уровне подготовленности испытуемых. Исходные баллы испытуемых разнятся. Следовательно отличается и разность между средним баллом по группе и баллом испытуемого. По итогам вычислений стандартное отклонение в приведённом примере составило 15,4.

В табл. 2 произведены аналогичные расчёты по второму профильному предмету. Во втором столбце представлены первичные баллы испытуемых, средний балл по тесту составил 45. Уменьшение разности между первичным и средним баллом испытуемых в группе привело к снижению стандартного отклонения и дисперсии соответственно.

Перевод z -оценок в область положительных целых чисел обычно осуществляется с помощью новых значений среднего арифметического (M) и стандартного отклонения (σ). Эти значения выбираются с таким расчётом, чтобы сохранить все различия между баллами испытуемых, выявленных в z -шкале. Это

¹ Аванесов В.С. Применение статистических методов в педагогических измерениях // Педагогические измерения. — 2011. — № 1. — С. 49.

Таблица 1

Расчет z-оценок (первый профильный предмет)

ФИО	Первичный балл (x_i)	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	z
Иванов С.	58	16,8	282,24	1,089
Павлов М.	45	3,8	14,44	0,246
Шукенова Б.	51	9,8	96,04	0,635
Таиров Т.	33	-8,2	67,24	-0,531
Бекишев А.	19	-22,2	492,84	-1,438
Средний балл	41,2			
$\sum_{i=1}^N (x_i - \bar{x})^2$	952,8			
s_x^2	238,2			
Стандартное отклонение ($s_x = \sqrt{s_x^2}$)	15,4			

Таблица 2

Расчет z-оценок (второй профильный предмет)

ФИО	Первичный балл (x_i)	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	z
Иванов С.	44	-1	1	-0,108
Павлов М.	49	4	16	0,430
Шукенова Б.	55	10	100	1,075
Таиров Т.	47	2	4	0,215
Бекишев А.	30	-15	225	-1,613
Средний балл	45			
$\sum_{i=1}^N (x_i - \bar{x})^2$	346			
s_x^2	86,5			
Стандартное отклонение ($s_x = \sqrt{s_x^2}$)	9,3			

Таблица 3

Сложение баллов и z-оценок по двум предметам

ФИО	Исходный балл (x_i)	z	Исходный балл (x_i)	z	Сумма исходных баллов	Сумма z-оценок	T-шкала
Иванов С.	58	1,089	44	-0,108	102	0,981	598
Павлов М.	45	0,246	49	0,430	94	0,676	568
Шукенова Б.	51	0,635	55	1,075	106	1,710	671
Таиров Т.	33	-0,531	47	0,215	80	-0,316	468
Бекишев А.	19	-1,438	30	-1,613	49	-3,051	195

можно сделать, избавившись от отрицательных и дробных значений z путём умножения каждой z -оценки на одно и то же число, прибавления общей константы и последующего округления. Для преобразования z -оценок используется формула $z = M + \sigma z$. (Шкала СЕЕВ, ETS для сообщения абитуриентам о результатах приёмных экзаменов в колледжи).

В нашем случае планируется пересчёт в 1000-бальную шкалу со средним $M = 500$ и $\sigma = 100$ по формуле $Z = 500 + 100z$.

Таким образом, согласно произведённому шкалированию результатов тестирования, можно утверждать, что применение Z -шкалы даёт возможность корректного сравнения тестовых баллов испытуемых по нескольким тестам.