УДК 681.3

К ВОПРОСУ О МОДЕЛИРОВАНИИ РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ

БАТЫРХАНОВ Ж.И., ТАЖИЕВА Ш.Ж.
izvestiya@ktu.aknet.kg

Для управления в сфере образования необходима объективная оценка качества подготовки лиц, освоивших образовательные программы. Это качество (уровень подготовленности) является латентной характеристикой и оценивается по данным тестирования с использованием заданий стандартизованной формы – контрольных измерительных материалов.

В основе технологии обработки результатов тестирования лежит математическое моделирование на основе математических моделей Раша. Поэтому предварительно следует обосновать выбор этих моделей, а именно рассмотреть три вопроса. Во – первых, как в рамках моделей Раша могут быть получены оценки латентных характеристик испытуемых и заданий теста. Во – вторых, в чём преимущества моделей Раша перед другими аналогичными по назначению моделями. В-третьих, как при использовании моделей Раша решается задача обнаружения испытуемых, измерение которых в силу искажения следует считать недостоверными.

Георг Раш [1] предложил модель, в которой вероятность $\rho(\theta,\delta)$ успешного выполнения испытуемым с уровнем подготовленности θ задания трудности δ вычисляется по формуле:

$$\rho(\theta,\delta) = (1+\exp(-(\theta-\delta)))^{-1}. \quad (1.)$$

Эта модель называется дихотомической, поскольку подразумевает, что мы различаем только два случая: задание выполнено правильно (1 балл) или нет (0 баллов). Вывод формулы (1) дан в [2]. Модель также называется логистической, так как в результате введения в формулу экспоненты формула (1) оказывается частным случаем функции плотности логистического распределения.

В дальнейшем основная, дихотомическая модель Раша была обобщена на случай оценивания с произвольными внутренними категориями выполнения задания, выражаемыми в матрице ответов баллом из множества $\{0,1,\dots,B.\}$ В этом случае вероятность $\rho(\theta,\{\delta_j\},k)$ получения k баллов за выполнение задания испытуемым с уровнем подготовленности θ вычисляется по формуле:

$$\rho(\theta,\{\delta_j\},k) = \frac{\exp(k\theta - \sum_{j=0}^{k}\delta_j)}{\sum_{l=0}^{B}\exp(l\theta - \sum_{j=0}^{l}\delta_j)}, \qquad (2)$$

где δ_j - трудность перехода при выполнении задания с категории $(j-1)$ на категорию $j,$ то есть получения j баллов при условии, что $(j-1)$ балл уже заработан, $j = 1..B$, причём полагается $\delta_0 = 0$; B - количество категорий оценивания в задании, то есть максимально возможный балл. Любое политомическое задание может быть представлено как многошаговое: чтобы достичь категории k (и, таким образом, получить k баллов за выполнение задания), испытуемый должен последовательно преодолеть k шагов, $k = 0,1,\dots,B.$ Трудности δ_j выполнения каждого шага в общем случае различны и не зависят от трудности выполнения остальных шагов. Для краткости модель с произвольными категориями выполнения задания будем называть политомической моделью Раша (Partial Credit Model). При выполнении некоторых дополнительных условий политомическое задание эквивалентно множеству независимых дихотомических заданий и может быть на них разложено (в остальных случаях остаётся неразложимое задание с тремя категориями).

Предположим, что тест состоит из I заданий $(i = 1..I)$, каждое из которых имеет максимальный балл (количество внутренних категорий) B_i и трудности шагов $\{\delta_{ij}\}$, $j = 1..B_i$. Пусть тест выполняли N испытуемых $(n = 1..N)$, каждый со своим уровнем подготовленности θ_n. Величины θ_n и δ_{ij} в совокупности являются параметрами модели Раша (латентными). Наблюдаемыми являются элементы матрицы ответов, то есть баллы x_{ni}, полученные каждым из испытуемых n за выполнение каждого задания i. Строка матрицы называется профилем испытуемого, столбец – профилем задания. Суммы баллов по строкам (столбцам) называются первичными баллами испытуемого (задания):

$$X_n = \sum_{i=1}^{I} x_{ni}. \qquad (3)$$

$$Z_i = \sum_{n=1}^{N} x_{ni}. \qquad (4)$$

Первая группа вопросов, связанных с оцениванием параметров модели Раша, состоит в том, как отобразить на метрической шкале испытуемых и задания теста, и какова точность этого отображения (погрешность модели). Известно много способов приближённого нахождения (оценивания) параметров модели Раша по реальной матрице ответов [3,4]. Это может быть метод моментов или метод наибольшего (максимального) правдоподобия (условного или безусловного) с разными подходами к численному решению соответствующих систем уравнений, а также метод симметрических функций.

Указанные методы основаны на замечательном свойстве модели Раша: первичные баллы (3-4) являются достаточными статистиками для параметров модели Раша. Это значит, что результат решения обратной задачи моделирования по заданной матрице ответов зависит только от первичных баллов $X_n(n = 1..N)$ и $Z_i(i = 1..I)$ (в политомической модели используются первичные баллы категорий), но не зависит от того, как именно расставлены составляющие эти суммы баллы внутри профилей испытуемых и заданий.

Математические модели Раша обладают рядом существенных преимуществ перед другими описанными в литературе моделями тестирования.

Важнейшим преимуществом модели Раша является объективность измерения в смысле его инвариантности как по отношению к объекту измерения (испытуемому), так и по отношению к средству измерения (тесту). Более того, модель выводится, исходя из этого свойства объективности.

Отсюда вытекает и большая, чем в рамках некоторых других моделей, определенность терминов «уровень подготовленности испытуемого» и «уровень трудности задания». Эти понятия имеют строгий математический смысл как для дихотомической модели Раша, так и для политомической: не возникает разночтений в определении того, у кого из испытуемых уровень подготовленности выше (какое из заданий труднее), даже при измерении их разными заданиями (на разных испытуемых).

Ещё одно существенное преимущество модели Раша состоит в том, что результаты измерения получаются на линейной метрической шкале логитов, причём интерпретация единицы измерения «логит» тоже, строго формально определена. Метрический характер шкалы вытекает непосредственно из вывода формул модели Раша. При этом, поскольку формула вероятности в рамках модели Раша (2) зависит только от разности уровней подготовленности и трудности, то «ноль» метрической шкалы не фиксирован и может выбираться произвольно (нас интересуют не абсолютные значения параметров модели, а их взаимное расположение и расстояние между ними). Это открывает возможность выравнивать на единой шкале результаты измерения разных групп испытуемых разными вариантами теста.

Литература

1. Карданова Е.Ю. Обнаружение искажений при тестировании с использованием модели Раша. // Обозрение прикладной и промышленной математики, Т.14, вып. 4. М., 2007.
2. Карданова Е.Ю. Технология обработки информации в многокритериальном мониторинге на основе политомической модели Раша. // Системы управления и информационные технологии, №3. М. 2007.
3. Нейман Ю.М. Введение в теорию моделирования и параметризации педагогических тестов. //- М.:Прометей, 2000 г.
4. Нейман Ю.М., Хлебников В.А. Педагогическое тестирование как измерение. // Центр тестирования МО РФ. М., 2000.