

ЧАСТЬ 4. ПРИКЛАДНЫЕ ВОПРОСЫ РАЗРАБОТКИ И ПРИМЕНЕНИЯ ЦИФРОВЫХ ТЕХНОЛОГИЙ

УДК 519.8

СТАТИСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ ТЕСТОВЫХ ЗАДАНИЙ

А.В. Ганичева, А.В. Ганичев

© Ганичева А.В., Ганичев А.В., 2024

Аннотация. В статье рассмотрен вопрос оценки качества тестирования. Отмечены преимущества тестового контроля. Для определения качества тестов рассчитываются его характеристики: надежность, валидность и сложность. Разработана методика определения качества тестов.

Ключевые слова: балл, результаты тестирования, ошибка, доля, точность, вопрос, ответ, вероятность.

При внедрении цифровых образовательных технологий в процесс обучения студентов различным дисциплинам важную роль играет своевременное и полноценное оценивание их знаний о предмете. Наряду с традиционными формами педагогического контроля знаний в последнее время все чаще проводится тестирование обучаемых. Данная форма может быть обучающей, т.е. сочетать обучение и оперативный контроль, а также использоваться при итоговой аттестации. С помощью тестирования можно дать не только объективную, количественную, но и сопоставимую оценку качества подготовки обучаемых по предмету. Важной особенностью процедуры тестирования является возможность организации массового контроля результатов обучения.

Разработка качественных гомогенных тестовых заданий (т. е. заданий, предназначенных для контроля знаний по одной дисциплине) требует использования научных методов: теории педагогических измерений, методов отбора содержания, математических методов, применяемых для определения показателя качества составленных тестов. Тестовые технологии представляют собой методы, средства преобразования исходных (несистематизированных, непроконтролированных) представлений обучаемых по данному предмету, теме, вопросу в итоговую компетенцию в результате осознанного выбора правильных ответов на поставленные вопросы.

При составлении тестов использовались разные формы задания: закрытая (по каждому заданию приводятся ответы), открытая (ответы получают сами обучаемые), на установление соответствия между вопросом и ответом.

Качество теста характеризуется надежностью, валидностью и сложностью. Тест считается надежным, если он обеспечивает достаточно высокую точность измерения знаний обучаемых и устойчив к действию случайных факторов, т.е. при повторном тестировании дает близкие результаты. Если результаты тестирования оцениваются в дихотомической форме, то для вычисления коэффициента надежности r_i , как правило, используется формула Кьюдера – Ричардсона [1]. Форма тестирования является дихотомической, если для результатов оценивания используются функции x_{ij} , причем $x_{ij} = 1$, если i -й тестируемый дал правильный ответ на вопрос j , и $x_{ij} = 0$, если он дал неверный ответ. Формула Кьюдера – Ричардсона имеет вид:

$$r_i = \frac{n}{n-1} \left(1 - \frac{\frac{\sum p_j q_j}{n^2}}{S_x^2} \right), \quad (1)$$

где p_j – доля правильных ответов на задание j ; q_j – доля неправильных ответов на данное задание ($q_j = 1 - p_j$); S_x^2 – несмещенная дисперсия закона распределения наблюдаемых баллов; n – число заданий в тесте.

Несмещенная дисперсия рассчитывается по формуле [3]

$$S_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2, \quad (2)$$

где x_i – индивидуальный балл, т.е. число правильных ответов i -го испытуемого; \bar{x} – средний балл всех испытуемых $\left(\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \right)$; N – число тестируемых.

При разработке теста необходимо учитывать причины, снижающие его надежность. К ним относятся [4] величина теста, некорректная формулировка заданий, случайное угадывание ответов, субъективное оценивание, ошибки в подсчетах. Предсказание истинного балла для i -го обучаемого ($i = 1, N$) осуществляется на основе регрессионной модели Дж. Стенли [5]:

$$T_i = \bar{x} + r_h(x_i - \bar{x}),$$

где T_i – истинный балл i -го студента; x_i – набранный при тестировании балл; r_h – коэффициент надежности теста.

Коэффициент надежности теста

$$r_h = 1 - \frac{S_e^2}{S_x^2},$$

где S_e – стандартная ошибка индивидуального балла студента.

На основе коэффициента надежности теста можно оценить стандартную ошибку индивидуального балла студента:

$$S_e^2 = S_x^2(1 - r_h).$$

Тест считается надежным при $r_h \geq 0,7$. В этом случае стандартная ошибка индивидуального балла студента достаточно мала ($S_e < S_x \cdot \sqrt{0,3}$). Если $r_h < 0,7$, то стандартная ошибка измерений велика и тест надо улучшать. Подозрительным считается и тест, у которого $r_h > 0,9$.

На практике S_e считается статистической величиной, отражающей степень точности измерений знаний студентов, поэтому ее используют для определения границ доверительного интервала, внутри которого с доверительной вероятностью β находится T_i – истинный балл i -го тестируемого студента. С вероятностью β можно считать, что

$$x_i - S_e \cdot t_\beta < T_i < x_i + S_e \cdot t_\beta, \quad (3)$$

где функция t_β находится из условия $\Phi(t_\beta) = \frac{\beta}{2}$; $\Phi(x)$ – интегральная

функция Лапласа, вычисляемая по формуле $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz$.

Другим показателем качества теста является его критериальная валидность, т. е. степень соответствия результатов тестирования некоторому внешнему фактору. Для математических тестов это может быть оценка корреляции между индивидуальным баллом тестирования студента и его оценкой на экзамене или корреляция между индивидуальным баллом и рейтингом текущей успеваемости студента. Критериальная валидность r_v вычисляется по формуле

$$r_a = \frac{M[(x_i - \bar{x})(y_i - \bar{y})]}{S_x S_y} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{S_x S_y}, \quad (4)$$

где x_i – индивидуальный балл i -го студента; y_i – индивидуальная оценка i -го студента (экзаменационная или аттестационная); \bar{x} – средний балл по группе; \bar{y} – средняя оценка по группе.

Сложность теста оценивается как средняя арифметическая долей неправильных ответов для всех обучаемых в группе. Возможно определение сложности теста путем экспертного оценивания.

Приведем методику оценки качества теста.

1. Заполняется таблица обработки результатов теста размером $n \cdot m$ (n – количество студентов; m – число тестовых заданий) следующего вида:

		№ теста					
		№ студента					

Если ответ правильный, то в ячейку таблицы ставится 1, если неправильный – 0. Рассчитывается общая сумма правильных ответов по строке. Это сумма баллов для i -го тестируемого (x_i). В том случае, если на j -е задание правильно ответили все тестируемые (или ни один из тестируемых не дал правильного ответа), то столбец, соответствующий данному заданию, не учитывается в расчетах. В этом случае заполняется дополнительный столбец x_i^* . Для вычисления коэффициента надежности по формуле (1) рассчитывается произведение долей правильных на число долей неправильных ответов. Если доли в таких столбцах окажутся равными нулю, то соответствующие произведения не используются.

2. Рассчитываются суммы долей правильных ответов на j -е задание (общая сумма по j -му столбцу долей правильных ответов) по всем строкам и суммы долей неправильных ответов на j -е задание (общая сумма по j -му столбцу долей неправильных ответов).

3. Для проведения расчетов заполняется следующая таблица:

№	x_i	x_i^2	$x_i - \bar{x}$	y_i	$y_i - \bar{y}$	$(x_i - \bar{x})^2$	$(x_i - \bar{y})^2$	$C_4 \cdot C_6$	T_i
Сумма									

В таблице введены следующие обозначения: 9-й столбец $C_4 \cdot C_6$ есть произведение значений столбцов 4 и 6: $(x_i - \bar{x})(y_i - \bar{y})$.

Рассчитывается несмешенная дисперсия по формуле (2). Для этого используется 7-й столбец таблицы расчетов.

4. Расчет коэффициентов валидности и надежности теста с использованием формул (1) и (4) соответственно.

Пример расчетов по данной методике показан в работе авторов [2].

Статистическая теория тестирования имеет недостаток, заключающийся в том, что оценка знаний обучаемых зависит от характеристик используемых тестовых заданий. Для исправления данного недостатка была разработана технология тестирования, основанная на модели Раша и аналогичных моделях. При использовании моделей данного класса оценка уровня знаний обучаемых не зависит от используемого набора тестовых заданий.

Тестовые технологии могут применяться не только в учебном процессе, но и в других областях, например в промышленности, сельском хозяйстве, военном деле, здравоохранении, социальных проблемах жизни общества, маркетинговых исследованиях и т.д.

Библиографический список

1. Аванесов В.С. Форма тестовых заданий. М.: Центр тестирования, 2005. 156 с.
2. Ганичева А.В. Метод проверки качества тестов // Саморазвивающаяся среда технического университета: материалы Всероссийской научно-практической конференции. 2018. Ч. 2. С. 83–89.
3. Ганичева А.В. Теория вероятностей. СПб.: Лань, 2017. 144 с.
4. Майоров А.Н. Теория и практика создания тестов для системы образования: как выбирать, создавать и использовать тесты для целей образования. М.: Народное образование, 2000. 351 с.
5. Челышкова М.Б. Теория и практика конструирования педагогических тестов: учебное пособие. М.: Логос, 2002. 432 с.

Об авторах:

Ганичева Антонина Валериановна – кандидат физико-математических наук, доцент, доцент кафедры физико-математических

дисциплин и информационных технологий, ФГБОУ ВО «Тверская сельскохозяйственная академия», Тверь. E-mail: tgan55@yandex.ru

Ганичев Алексей Валерианович – старший преподаватель кафедры информатики и прикладной математики, ФГБОУ ВО «Тверской государственный технический университет», Тверь. E-mail: alexej.ganichev@yandex.ru

MODELING THE INVESTOR'S UTILITY FUNCTION

A.V. Ganicheva, A.V. Ganichev

Abstract. The article discusses the issue of evaluating the quality of testing. The advantages of the test control are noted. To determine the quality of tests, its characteristics are calculated: reliability, validity and complexity. An explanatory example showing the methodology for determining the quality of the test is considered.

Keywords: score, test results, error, share, accuracy, question, answer, probability.

About the authors:

Ganicheva Antonina Valerianovna – Ph.D., Associate Professor, Associate Professor of Department of Physical and Mathematical Disciplines and Informational Technologies, Tver Agricultural Academy, Tver. E-mail: tgan55@yandex.ru

Ganichev Alexey Valerianovich – Assistant Professor of Informatics and Applied Mathematics, Tver State Technical University, Tver. E-mail: alexej.ganichev@yandex.ru

УДК 332.142.4

КОНЦЕПЦИЯ ЭЛАСТИЧНОСТИ В ЦИФРОВОЙ ЭКОНОМИКЕ С ПОЗИЦИИ ТЕОРИИ ДИСТОРТНОСТИ

Б.Ф. Зюзин

© Зюзин Б.Ф., 2024

Аннотация. Рассмотрены вопросы концепции эластичности в экономическом анализе. Даны графические модели представления функции эластичности в системе приведенного квадрата на основании теории дистортности. Получено выражение критерия предельного