МОДЕЛЬ ОЦЕНИВАНИЯ УРОВНЯ ЗНАНИЙ СТУДЕНТА[1]

Маслякова Ирина Николаевна
старший преподаватель кафедры высшей математики РЭУ им. Г. В. Плеханова.
Адрес: ФГБОУ ВПО «Российский экономический университет имени Г. В. Плеханова», 117997, Москва, Стремянный пер., д. 36.
E-mail: maslyakova@gmail.com

Мочалина Екатерина Павловна
кандидат физико-математических наук, доцент кафедры высшей математики РЭУ им. Г. В. Плеханова.
Адрес: ФГБОУ ВПО «Российский экономический университет имени Г. В. Плеханова», 117997, Москва, Стремянный пер., д. 36.
E-mail: mochalina77@gmail.com

Оценка качества обучения лежит в основе процессов стимулирования учебной деятельности студента, управления качеством обучения, управления обучением как процессом освоения новых знаний и т. д. В общем случае оценка знаний является случайной величиной. Предъявляемые требования современных работодателей к молодому специалисту заставляют задуматься о более дифференцированном подходе к оценке уровня знаний. В статье авторами предложена математическая модель оценивания уровня знаний студента, входными параметрами которой являются объем и сложность тестовых заданий и априорная оценка уровня знаний студента. При этом предполагается, что качество выполнения тестовых заданий характеризуется некоторым непрерывным показателем.
Ключевые слова: тестирование, IRT, оценивание, уровень знаний, непрерывный показатель качества, параметрическая идентификация.

PARAMETRIC IDENTIFICATION IN KNOWLEDGE ESTIMATION MODELING

Maslyakova, Irina N.
Senior lecturer of the Department for Higher Mathematics of the PRUE.
Address: Plekhanov Russian University of Economics, 36 Stremyanny Lane, Moscow, 117997, Russian Federation.
E-mail: maslyakova@gmail.com

Mochalina, Ekaterina P.
PhD, Associate Professor of the Department for Higher Mathematics of the PRUE.

[1] Статья подготовлена по результатам внутреннего гранта Российского экономического университета имени Г. В. Плеханова «Построение процедуры оценивания качества уровня знаний студентов экономических специальностей».

Address: Plekhanov Russian University of Economics, 36 Stremyanny Lane, Moscow, 117997, Russian Federation.
E-mail: mochalina77@gmail.com

The assessment of quality of training is the cornerstone of processes of stimulation of educational activity of the student, quality management of training, management of training as process of development of new knowledge etc. Generally the assessment of knowledge is casual size. Qualifying standards of modern employers to the young expert set thinking on more differentiated approach to an assessment of level of knowledge. In article authors offered mathematical model of estimation of level of knowledge of the student which input parameters are the volume and complexity of test tasks and an aprioristic assessment of level of knowledge of the student. Thus it is supposed that quality of performance of test tasks is characterized by some continuous indicator.
Keywords: testing, IRT, estimation, level of knowledge, continuous indicator of quality, parametrical identification.

Традиционный подход к тестированию уровня знаний студента базируется на допущении, что качество ответа на тестовое задание характеризуется бинарной величиной: правильно – неправильно. В частности, такой подход лежит в основе классической теории тестов и современной теории тестов IRT (Item Response Theory) [6; 7]. Вместе с тем во многих случаях решение задачи (если речь не идет о тестировании, когда правильный ответ определяется однозначно) оценивается развернуто: «решил, но не полностью», «не решил, но идея была правильна» и т. д.

Рассмотрим подход к решению задачи определения уровня знаний студента по результатам выполнения совокупности тестовых заданий, качество выполнения которых оценивается некоторой непрерывной величиной, принимающей значения из отрезка [0; 1]. Предположим, что связь между качеством выполнения задания и уровнем знаний студента определяется некоторой нелинейной зависимостью.

Если принять, что *измерение* – это выполнение обучаемым некоторого тестового задания, а *результат измерений* представляет собой баллы, набранные за это тестовое задание, тогда можно сформулировать следующую задачу: по результатам измерения, т. е. имея оценку качества выполнения тестовых заданий, необходимо определить величину уровня знаний студента.

Введем следующие обозначения:

θ – *уровень знаний обучаемого*, непрерывная случайная величина, распределенная на [0; 1];

ω – *уровень сложности задания*, непрерывная величина, принимающая значения из [0; 1]. Уровень сложности задания характеризует требуемый уровень знаний обучаемого, который необходим для решения сформулированной в задании задачи;

$w = g(\omega, \theta)$ – *качество ответа на тестовое задание.*

Таким образом, это функция двух параметров: уровня сложности задания ω и уровня знаний обучаемого θ. Заметим также, что в непрерывном случае наиболее общей характеристикой качества ответа является плотность распределения уровня знаний θ.

Введенная нами функция w представляет собой предположение, аналогичное гипотезе, лежащей в основе IRT и состоящей в том, что вероятность выполнения задания является функцией латентного параметра способности θ обучаемого и латентного параметра трудности задания ω:

$$P = f(\theta - \omega).$$

Здесь *латентным параметром* считается свойство личности, недоступное для прямого наблюдения.

Пусть по результатам промежуточного или предварительного контроля известна априорная оценка уровня знаний студента θ_0. Поставим следующую задачу: студент выполнил тест, в котором были задания разных уровней сложности. Необходимо по результатам этого теста найти оценку априорного уровня знаний (θ_0) студента. Мы полагаем, что, предлагая студенту тестовое задание с уровнем сложности ω, мы наблюдаем качество ответа w с некоторой ошибкой ξ:

$$w = F(\omega, \theta) + \xi.$$

Пусть $\omega^{(n)} = (\omega_1, \ldots, \omega_n)$ – вектор показателей трудности совокупности тестовых заданий, выполненных студентом.

В результате наблюдения выполнения всех тестовых заданий мы получаем соответственно вектор наблюдений $w^{(n)} = (w_1, \ldots, w_n)$, каждая компонента которого является показателем качества w_j уровня сложности ω_j и ошибкой ξ_j:

$$w_j = F(\omega_j, \theta) + \xi_j.$$

Обозначим необходимую нам оценку параметра θ_0 посредством $\tilde{\theta}$. Согласно информационной теории идентификации соответствие настраиваемой модели объекту исследования, т. е. качество идентификации, оценивается критерием качества идентификации, который представляет собой средние потери. Следовательно, нам необходимо их минимизировать.

Введем *функцию потерь информации* Φ – показатель среднего риска потери информации относительно вектора показателей сложности $(\omega_1, \ldots, \omega_n)$. Основные свойства и алгоритмы получения функции Φ в явном виде рассмотрены в соответствующих работах (см., например, [4]). Следовательно, оценку параметра θ_0 мы получим из условия минимизации средних потерь информации:

$$\tilde{\theta} = \underset{\theta \in (0;1)}{\operatorname{argmin}} \frac{1}{n} \sum_{j=1}^{n} \Phi\left(w_j - F\left(\omega_j, \theta\right)\right).$$

Далее определим *невязку* между качеством w_j выполнения задания сложности ω_j, которое продемонстрировал студент, и ожидаемым качеством ответа следующим образом:

$$\varepsilon_j\left(\tilde{\theta}\right) = w_j - F\left(\omega_j, \tilde{\theta}\right),$$

где $F\left(\omega_j, \tilde{\theta}\right)$ – ожидаемое качество ответа студента на задание с уровнем сложности ω_j. Заметим, что, введя невязку, мы получаем соотношение

$$\varepsilon_j\left(\theta_0\right) = 0. \quad (1)$$

Средние эмпирические потери на основе вектора наблюдений качества выполнения тестового задания заданного объема n (относительно количества различных уровней сложности) определим следующим образом:

$$J_n = \frac{1}{n} \sum_{j=1}^{n} \Phi\left(\varepsilon_j\right) = \frac{1}{n} \sum_{j=1}^{n} \Phi\left(w_j - F\left(\omega_j, \tilde{\theta}\right)\right).$$

В силу необходимого условия экстремума имеем

$$\frac{1}{n} \sum_{j=1}^{n} \Phi'_{\varepsilon_j}\left(\varepsilon_j\right) \cdot \frac{dF\left(\omega_j, \tilde{\theta}\right)}{d\tilde{\theta}} = 0.$$

Полагая, что качество ответа на тестовое задание зависит от уровня знаний, а следовательно, не может быть постоянным, можно сказать, что

$$\frac{dF\left(\omega_j, \tilde{\theta}\right)}{d\tilde{\theta}} \neq 0.$$

Здесь $F\left(\omega_j, \tilde{\theta}\right)$ рассматривается как функция одной переменной при фиксированном параметре ω_j. Следовательно, $\Phi'_{\varepsilon_j}(\varepsilon_j)(\theta_0) = 0.$

Разложим производную функции потерь в окрестности точки θ_0 в ряд Тейлора с точностью до членов первого порядка малости. С учетом (1) получаем:

$$\sum_{j=1}^{n} \Phi''_{\varepsilon_j}\left(\theta_0\right) \cdot \varepsilon_j \cdot \frac{dF\left(\omega_j, \tilde{\theta}\right)}{d\tilde{\theta}} + \bar{o}\left(\tilde{\theta} - \theta_0\right) = 0. \quad (2)$$

Соотношение (2) представляет собой нелинейное уравнение для определения значения оценки $\tilde{\theta}$ уровня знаний обучаемого. В случае если функция потерь имеет вид

$$\Phi(\varepsilon_j) = \frac{1}{2} \cdot (\varepsilon_j)^2 = \frac{1}{2} \cdot \left(w_j - F(\omega_j, \tilde{\theta})\right)^2,$$

имеем

$$\Phi'_{\varepsilon_j} = \varepsilon_j = w_j - F(\omega_j, \tilde{\theta}), \quad \Phi''_{\varepsilon_j \varepsilon_j} = 1.$$

Уравнение (2) для определения оценки $\tilde{\theta}$ примет следующий вид (с точностью до $\bar{o}(\tilde{\theta} - \theta_0)$):

$$\sum_{j=1}^{n} \frac{dF(\omega_j, \tilde{\theta})}{d\tilde{\theta}} \cdot \left(w_j - F(\omega_j, \tilde{\theta})\right) = 0. \qquad (3)$$

В общем случае решение нелинейных уравнений (2) или (3) представляет собой нетривиальную задачу и требует использования соответствующих численных методов.

В случае если априорная оценка выбрана достаточно близко к значению оцениваемого параметра $\tilde{\theta}$, то можно получить решение уравнения (3) для случая квадратичной функции потерь в аналитическом виде в линейном приближении.

Действительно, разложим функцию $F(\omega_j, \tilde{\theta})$ в окрестности точки θ_0 в ряд Тейлора с точностью до членов первого порядка малости:

$$F(\omega_j, \tilde{\theta}) = F(\omega_j, \tilde{\theta}_0) + \frac{dF(\omega_j, \tilde{\theta}_0)}{d\tilde{\theta}} \cdot (\tilde{\theta} - \tilde{\theta}_0) + \bar{o}(\tilde{\theta} - \theta_0).$$

Подставляя это соотношение в уравнение (3), получаем с точностью до $\bar{o}(\tilde{\theta} - \theta_0)$:

$$\sum_{j=1}^{n} \frac{dF(\omega_j, \tilde{\theta}_0)}{d\tilde{\theta}} \left(w_j - F(\omega_j, \tilde{\theta}_0) - \frac{dF(\omega_j, \tilde{\theta}_0)}{d\tilde{\theta}} \cdot (\tilde{\theta} - \tilde{\theta}_0) \right) = 0.$$

Отсюда выражаем оценку $\tilde{\theta}$:

$$\tilde{\theta} = \tilde{\theta}_0 + \frac{\sum_{j=1}^{n} \frac{dF(\omega_j, \tilde{\theta}_0)}{d\tilde{\theta}} \cdot \left(w_j - F(\omega_j, \tilde{\theta}_0)\right)}{\sum_{j=1}^{n} \left(\frac{dF(\omega_j, \tilde{\theta}_0)}{d\tilde{\theta}} \right)^2}. \qquad (4)$$

Соотношение (4) определяет оценку уровня знаний студента в случае наличия априорных данных об уровне его знаний по результатам промежуточного контроля с квадратичным критерием качества. Второе слагаемое в правой части равенства (4) определяет поправку к априорной оценке уровня знаний студента.

В качестве такой априорной оценки θ_0 может рассматриваться некоторая среднестатистическая оценка (например, математическое ожидание уровня знаний), определяемая учебной программой, или оценка, полученная студентом на предыдущем тестировании. Если гипотеза о степени адекватности априорной оценки не является справедливой, то полученное соотношение (4) может быть использовано для построения итерационной процедуры нахождения корня уравнения (3) – оценки уровня знаний обучаемого $\tilde{\theta}$. Итерационный алгоритм определения уровня знаний студента в таком случае будет иметь вид:

$$\tilde{\theta}_{k+1} = \tilde{\theta}_k + \frac{\sum_{j=1}^{n} \frac{dF(\omega_j, \tilde{\theta}_k)}{d\tilde{\theta}} \left(w_j - F(\omega_j, \tilde{\theta}_k)\right)}{\sum_{j=1}^{n} \left(\frac{dF(\omega_j, \tilde{\theta}_k)}{d\tilde{\theta}} \right)^2}, k = 0, 1, \ldots$$

Рассмотрим теперь частный случай определения уровня знаний обучаемого, когда студенту предъявляются тестовые задания одинакового уровня сложности ω. Тогда в случае квадратичной функции потерь уравнение (3) принимает следующий вид:

$$\sum_{j=1}^{n} \frac{dF(\omega, \theta_0)}{d\tilde{\theta}} \cdot F(\omega, \tilde{\theta}) = \sum_{j=1}^{n} \frac{dF(\omega, \theta_0)}{d\tilde{\theta}} \cdot w_j.$$

Откуда, выполняя простейшие преобразования, получаем уравнение для вычисления значения параметра $\tilde{\theta}$:

$$F(\omega, \tilde{\theta}) = \frac{1}{n} \cdot \sum_{j=1}^{n} w_j.$$

Заметим, что правая часть этого уравнения представляет собой *среднее качество* *выполнения* студентом предложенных тестовых заданий.

Принципиальным вопросом реализации рассматриваемого подхода является построение функции ожидаемого качества ответа $F(\omega, \tilde{\theta})$, а в общем случае – построение *функции распределения качества ответа* $P(w, \theta, \omega)$. Здесь $P(w, \theta, \omega)$ – вероятность того, что на задание с уровнем сложности ω студент с уровнем знаний θ даст ответ, качество которого будет не выше w. Решение такой задачи может быть получено на основе статистической обработки субъективной экспертной информации о показателях качества ответа студента в зависимости от уровня сложности заданий и уровня знаний и является предметом дальнейших исследований.

Список литературы

1. *Кадневский В. М.* История тестов : монография. – М. : Народное образование, 2004.
2. *Манахов С. В., Рыжакова А. В.* Основные тенденции развития высшего образования в России: количественные и качественные аспекты // Вестник Российского экономического университета имени Г. В. Плеханова. – 2014. – № 10 (76). – С. 19–28.
3. *Маслякова И. Н.* Рекуррентное оценивание уровня знаний обучаемого при непрерывном показателе качества выполнения заданий // Обозрение прикладной и промышленной математики. – 2009. – Т. 16. – Вып. 5. – С. 893–894.
4. *Цыпкин Я. З.* Информационная теория идентификации. – М. : Наука, 1995.
5. *Andrich D., Sheridan B., Lyne A., Luo G.* RUMM: A Windows-Based Item Analysis Program Employing Rasch Unidimensional Measurement Models. – Perth : Murdoch University, 2000.
6. *Rasch G.* An Item Analysis Which Takes Individual Differences into Account // British Journal of Mathematical and Statistical Psychology. – 1966. – Vol. 19. – Part 1. – P. 49–57.
7. *Rasch G.* Probabilistic Models for Some Intelligence and Attainment Tests. – Copenhagen : Danish Institute for Educational Research, 1960. – Ch. V–VII, X.

References

1. Kadnevskiy V. M. Istoriya testov : monografiya [History of Tests, monograph]. Moscow, Narodnoe obrazovanie, 2004. (In Russ.).
2. Manakhov S. V., Ryzhakova A. V. Osnovnye tendentsii razvitiya vysshego obrazovaniya v Rossii: kolichestvennye i kachestvennye aspekty [Key Trends in the Development of Higher Education in Russia: Quantitative and Qualitative Aspects]. *Vestnik Rossiyskogo ekonomicheskogo universiteta imeni G. V. Plekhanova* [Vestnik of the Plekhanov Russian University of Economics], 2014, No. 10 (76), pp. 19–28. (In Russ.).
3. Maslyakova I. N. Rekurrentnoe ocenivanie urovnya znaniy obuchaemogo pri nepreryvnom pokazatele kachestva vypolneniya zadaniy [Recursive Estimation of the Level of Student's Knowledge in a Continuous Quality Indicator of the Execution of Tasks]. *Obozrenie prikladnoy i promyshlennoy matematiki*, 2009, Vol. 16, No. 5, pp. 893–894. (In Russ.).
4. Tsypkin Ya. Z. Informatsionnaya teoriya identifikatsii [Information Theory of Identification]. Moscow, Nauka, 1995. (In Russ.).
5. Andrich D., Sheridan B., Lyne A., Luo G. RUMM: A Windows-Based Item Analysis Program Employing Rasch Unidimensional Measurement Models. Perth, Murdoch University, 2000.
6. Rasch G. An Item Analysis Which Takes Individual Differences into Account. *British Journal of Mathematical and Statistical Psychology*, 1966, Vol. 19, Part 1, pp. 49–57.
7. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Copenhagen, Danish Institute for Educational Research, 1960, ch. V–VII, X.