УДК 519.24

В. В. Братищенко
Байкальский государственный университет экономики и права,
г. Иркутск, Российская Федерация

ЛАТЕНТНАЯ МОДЕЛЬ СТОБАЛЛЬНОЙ ОЦЕНКИ НА ОСНОВЕ БЕТА-РАСПРЕДЕЛЕНИЯ

АННОТАЦИЯ. Для измерения влияния на стобалльную оценку подготовленности студента и сложности экзамена в статье предлагаются соответствующие латентные параметры, являющиеся параметрами распределения вероятностей оценок, а в качестве закона распределения вероятностей — бета-распределение. При этом как и в модели Раша, применяемой для обработки результатов тестирования, параметр подготовленности увеличивает вероятность получить высокую оценку, а параметр сложности уменьшает. Разработаны численные методы определения латентных параметров по множеству оценок. Выполнена проверка гипотезы о соответствии модели имеющимся оценкам по критерию хи-квадрат. На основе неравенства Рао-Крамера выведены формулы для расчета дисперсии оценок латентных параметров. Результаты статистических проверок свидетельствуют о возможности применения модели на основе бета-распределения для исследования стобалльных оценок. Модель может использоваться для выявления слишком трудных или слишком легких экзаменов.

КЛЮЧЕВЫЕ СЛОВА. Модель Раша; латентный параметр; бета-распределение; критерий хи-квадрат; неравенство Рао-Крамера; дисперсия оценки.

ИНФОРМАЦИЯ О СТАТЬЕ. Дата поступления 7 сентября 2015 г.; дата принятия к печати 20 сентября 2015 г.; дата онлайн-размещения 30 сентября 2015 г.

V. V. Bratishchenko
Baikal State University of Economics and Law,
Irkutsk, Russian Federation

LATENT MODEL OF A 100-POINT GRADE ON THE BASIS OF BETA DISTRIBUTION

ABSTRACT. For measurement of influence of the student's preparedness and examination difficulty on a 100-point grade, the article offers corresponding latent parameters which are the parameters of grade probabilities distribution, while as a law of probabilities distribution it offers to use the beta distribution. At that, like in the Rasch model used for processing the results of testing, the preparedness ability parameter increases the chance to get a high grade, while the difficulty parameter decreases it. The article develops numerical methods for determination of latent parameters by a set of grades, as well as checks a hypothesis of compliance of the model to the available grades by the chi-square criterion. It draws formulas for calculating dispersion of estimating the latent parameters on the basis of the Rao-Kramer's inequality. The results of statistical checks testify the possibility of using the model on the beta distribution basis for investigating hundred-grades. The model can be used to identify too difficult or too easy examinations.

KEYWORDS. Rasch model; latent parameter; beta distribution; criterion of chi-square; Rao-Kramer inequality; dispersion of assessment.

ARTICLE INFO. Received September 7, 2015; accepted September 20, 2015; available online September 30, 2015.

Латентные модели нашли широкое применение для анализа результатов тестирования. В частности, в параметрической теории тестирования (Item Response Theory) применяется модель Раша [3; 6; 11; 12], в которой вероятность

$$P(\delta, \theta) = \frac{e^{\theta}}{e^{\theta} + e^{\delta}}$$

правильного ответа на тестовое задание зависит от «трудности» задания, характеризуемой параметром δ, уровня «подготовленности» тестируемого, задаваемого параметром θ.

Область применения идей Раша непрерывно расширяется на многие другие области, в которых влияние нескольких факторов можно описать латентными параметрами [2, с. 35; 5]. В частности, продемонстрирована возможность применения таких моделей для анализа экзаменационных оценок в высшей школе [1] и сформированности компетенций студентов [7; 8].

Достаточно часто наблюдаемым результатом в таких случаях является не дихотомическая (правильно / неправильно), а политомическая величина (три и более градаций). Это приводит к применению политомических моделей [9; 10]. Количество параметров в этих моделях увеличивается в несколько раз по количеству градаций, что существенно снижает точность оценивания каждого параметра. Это может стать непреодолимым препятствием при построении политомических моделей для распространенных в современном обучении стобалльных оценок. Таким образом, актуальной становится задача построения латентной модели для политомических оценок с большим количеством градаций.

Модель стобалльной оценки можно получить, используя непрерывное распределение, просто считая наблюдаемые оценки непрерывными величинами. Основаниями для такой модели выступают следующие допущения. Во-первых, предполагается, что оценка будет не только порядковой (чем выше оценка, тем лучше знания), но и метрической (разница в знаниях с оценками 55 и 50 одинакова с разницей в знаниях с оценками 95 и 90). Во-вторых, оценка — непрерывная величина, измеряемая с определенной точностью. Оба допущения являются весьма спорными. Обоснованием первого может служить психологическое отношение к оценкам как к метрическим измерениям, которое только усиливается при увеличении градаций. Второе допущение можно связать с ошибкой моделирования, которая также должна уменьшаться с увеличением градаций.

Из известных непрерывных распределений наиболее подходящим для моделирования стобалльных оценок представляется бета-распределение. Благодаря наличию двух параметров оно позволяет менять числовые характеристики случайной величины в широких пределах. Поделив стобалльную оценку на сто, получаем величину из интервала от 0 до 1, что соответствует области значений бета-распределения. Далее по тексту под оценкой понимается именно такое значение.

Пусть X_{ij} — оценка i-го студента ($i = 1, \ldots, n$) на j-м экзамене ($j = 1, \ldots, m$). Будем считать, что все оценки независимым в совокупности и подчиняются бета-распределению, а плотность X_{ij}

$$f_{ij}(x) = \frac{\Gamma(\theta_i + \delta_j)}{\Gamma(\theta_i)\Gamma(\delta_j)} x^{\theta_i - 1}(1-x)^{\delta_j - 1}, 0 \le x \le 1$$

зависит от параметров $\theta > 0$ «подготовленности» i-го студента и $\delta_j > 0$ «сложности» j-го экзамена.

Данная модель близка по свойствам к модели Раша. Чем выше подготовленность студента, тем больше вероятность сдать экзамен с высокой оценкой. Например, на графиках вероятностей (рис. 1) представлена возможность получения оценки выше 0,5 в зависимости от уровня подготовленности при разных уровнях сложности экзамена. При любой сложности экзамена больше шансов заработать высокий балл у лучше подготовленного студента. В модели Раша вероятность получить оценку выше 0,5 (сдать экзамен) также равна 0,5 при совпадении «подготовленности» студента и «сложности» экзамена. Аналогично, на графиках вероятностей (рис. 2) показана возможность получения оценки выше 0,5 в зависимости от сложности при разных уровнях подготовленности студентов.

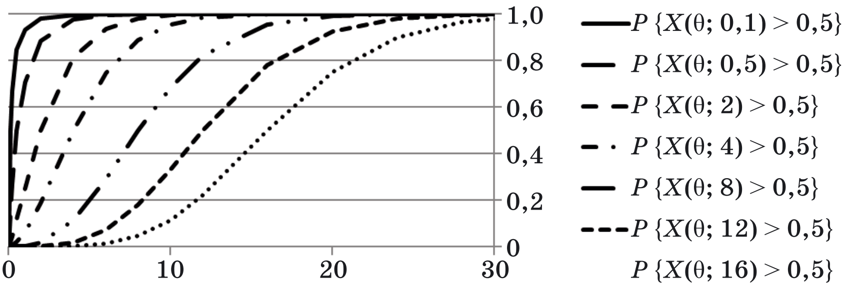

Рис. 1. Зависимость вероятности получить оценку выше 0,5 от уровня подготовленности θ для экзаменов разной сложности

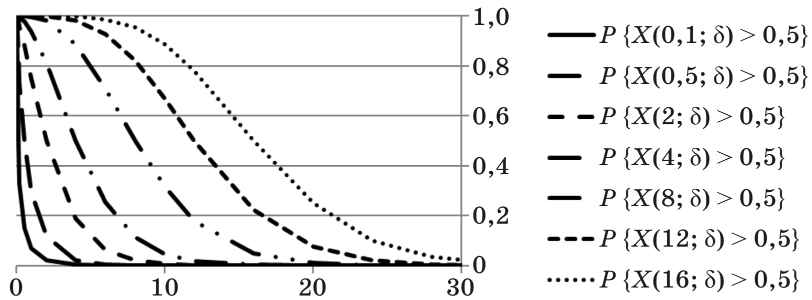

Рис. 2. Зависимость вероятности получить оценку выше 0,5 от уровня сложности экзамена δ для студентов разной подготовленности

Для оценки параметров модели можно воспользоваться методом моментов. Используя формулу

$$M[X_{ij}] = \frac{\theta_i}{\theta_i + \delta_j}$$

для математического ожидания случайной величины X_{ij} и выполняя суммирования по i и j, получаем уравнения для сумм оценок каждого студента и сумм оценок каждого экзамена:

$$\begin{aligned} &\sum_{j=1}^{m} \frac{\theta_i}{\theta_i + \delta_j} - \sum_{j=1}^{m} x_{ij} = 0, \quad i = 1, \ldots, n; \\ &\sum_{i=1}^{n} \frac{\theta_i}{\theta_i + \delta_j} - \sum_{i=1}^{n} x_{ij} = 0, \quad j = 1, \ldots, m. \end{aligned} \tag{1}$$

Для поиска решения применим метод Ньютона

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})}.$$

Подставляем в эту формулу производные

$$\frac{\partial}{\partial \theta_i}\left(\sum_{j=1}^{m} \frac{\theta_i}{\theta_i + \delta_j} - \sum_{j=1}^{m} x_{ij}\right) = \sum_{j=1}^{m} \frac{\delta_j}{(\theta_i + \delta_j)^2},$$

$$\frac{\partial}{\delta_j}\left(\sum_{i=1}^{n} \frac{\theta_i}{\theta_i + \delta_j} - \sum_{i=1}^{n} x_{ij}\right) = -\sum_{i=1}^{n} \frac{\theta_i}{(\theta_i + \delta_j)^2}$$

левых частей системы (1) и получаем выражения для итерационных вычислений

$$\theta_i^{(k+1)} = \theta_i^{(k)} - \frac{\sum_{j=1}^{m} \frac{\theta_i^{(k)}}{\theta_i^{(k)} + \delta_j^{(k)}} - \sum_{j=1}^{m} x_{ij}}{\sum_{j=1}^{m} \frac{\delta_j^{(k)}}{(\theta_i^{(k)} + \delta_j^{(k)})^2}};$$

$$\delta_j^{(k+1)} = \delta_j^{(k)} + \frac{\sum_{i=1}^{n} \frac{\theta_i^{(k)}}{\theta_i^{(k)} + \delta_j^{(k)}} - \sum_{i=1}^{n} x_{ij}}{\sum_{i=1}^{n} \frac{\theta_i^{(k)}}{(\theta_i^{(k)} + \delta_j^{(k)})^2}}.$$

Для вычисления начальных значений предположим, что в первом уравнении системы (1) все параметры экзаменов одинаковые $\delta_1 = \ldots = \delta_m = \bar{\delta}$. Уравнение

$$\sum_{j=1}^{m} \frac{\theta_i^{(0)}}{\theta_i^{(0)} + \bar{\delta}} = \sum_{j=1}^{m} x_{ij}$$

позволяет найти начальные значения параметров подготовленности

$$\theta_i^{(0)} = \frac{\bar{\delta} \sum_{j=1}^{m} x_{ij}}{m - \sum_{j=1}^{m} x_{ij}}.$$

Аналогично в предположении одинаковых значений параметров подготовленности решение второго уравнения системы (1)

$$\sum_{i=1}^{n} \frac{\bar{\theta}}{\bar{\theta} + \delta_j^{(0)}} = \sum_{i=1}^{n} x_{ij}$$

позволяет найти начальные значения параметров сложности экзаменов

$$\delta_j^{(0)} = \frac{\bar{\theta}\left(n - \sum_{i=1}^{n} x_{ij}\right)}{\sum_{i=1}^{n} x_{ij}}.$$

Оценки параметров $\bar{\delta}$, $\bar{\theta}$ можно найти методом моментов в предположении, что все наблюдения имеют бета-распределение с данными параметрами. Решение системы

$$\bar{x} = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} = \frac{\bar{\theta}}{\bar{\theta} + \bar{\delta}},$$

$$\bar{\sigma}^2 = \frac{1}{nm-1} \sum_{i=1}^{n} \sum_{j=1}^{m} (x_{ij} - \bar{x})^2 = \frac{\bar{\theta}\bar{\delta}}{(\bar{\theta} + \bar{\delta})^2 (\bar{\theta} + \bar{\delta} + 1)}$$

позволяет вычислить искомые значения

$$\bar{\theta} = \frac{(1 - \bar{x})\bar{x}^2 - \bar{x}\bar{\sigma}^2}{\bar{\sigma}^2};$$

$$\bar{\delta} = \frac{(1 - \bar{x})}{\bar{x}} \bar{\theta} = \frac{(1 - \bar{x})\bar{x} - \bar{\sigma}^2}{\bar{\sigma}^2} (1 - \bar{x}).$$

Для проверки найденных оценок параметров применим статистику хи-квадрат, построенную следующим образом. Все оценки $\{X_{ij}, j = 1, \ldots, m\}$ разделим на k-классов: l-й класс характеризуется множеством индексов J_l. Для каждого класса определим случайную величину

$$Y_l = \sum_{j \in J_l} X_{ij}.$$

В условиях применения модели (а именно независимости в совокупности величин X_{ij}, имеющих бета-распределение с параметрами θ_i и δ_j, Y_l) будет иметь асимптотически нормальное распределение со следующими параметрами:

$$a_l = M[Y_l] = \sum_{j \in J_l} \frac{\theta_i}{\theta_i + \delta_j};$$

$$\sigma_l^2 = D[Y_l] = \sum_{j \in J_l} \frac{\theta_i \delta_j}{(\theta_i + \delta_j)^2 (\theta_i + \delta_j + 1)}.$$

Статистика

$$\chi^2 = \sum_{l=1}^{k} \left(\frac{Y_l - a_l}{\sigma_l} \right)^2$$

будет иметь хи-квадрат распределение с k-степенями свободы. Предложенная статистика позволяет проверять по критерию хи-квадрат гипотезы о равенстве истинных значений параметров модели найденным значениям $\hat{\theta}_i = \lim_{k \to \infty} \theta_i^{(k)}$ и $\hat{\delta}_j = \lim_{k \to \infty} \delta_j^{(k)}$. Статистическая обработка одного потока оценок подтвердила гипотезы о применимости предложенной модели статистическим данным.

Для оценки точности оценок параметров можно воспользоваться неравенством Рао-Краммера [4]:

$$D[\hat{\theta}_i] \geq \frac{1}{I(\theta_i; \{X_{ij}\})}, \quad D[\hat{\delta}_j] \geq \frac{1}{I(\delta_j; \{X_{ij}\})},$$

где $I(\theta_i; \{X_{ij}\}) = M\left[\left(\frac{\partial \ln(L)}{\partial \theta_i} \right)^2 \right]$ — количество информации о параметре θ_i;

$L = \prod_{i=1}^{n} \prod_{j=1}^{m} \frac{\Gamma(\theta_i + \delta_j)}{\Gamma(\theta_i)\Gamma(\delta_j)} x_{ij}^{\theta_i - 1} (1 - x_{ij})^{\delta_j - 1}$ — функция правдоподобия.

Частная производная будет иметь вид

$$\frac{\partial \ln(L)}{\partial \theta_i} = \sum_{j=1}^{m} \left(\frac{1}{\Gamma(\theta_i + \delta_j)} \frac{\partial \Gamma(\theta_i + \delta_j)}{\partial \theta_i} - \frac{1}{\Gamma(\theta_i)} \frac{\partial \Gamma(\theta_i)}{\partial \theta_i} + \ln(x_{ij}) \right) = \\ = \sum_{j=1}^{m} (\psi(\theta_i + \delta_j) - \psi(\theta_i) + \ln(x_{ij})), \tag{2}$$

где $\psi(x)$ — дигамма-функция.

Используя независимость в совокупности величин X_{ij}, можно выразить количество информации в виде сумм математических ожиданий

$$I\left(\theta_i; \{X_{ij}\}\right) = M\left[\left(\frac{d \ln(L)}{d\theta_i} \right)^2 \right] = M\left[\left(\sum_{j=1}^{m} (\psi(\theta_i + \delta_j) - \psi(\theta_i) + \ln(X_{ij})) \right)^2 \right] =$$

$$= \sum_{j,l=1}^{m} M[(\psi(\theta_i + \delta_j) - \psi(\theta_i) + \ln(X_{ij})) \cdot (\psi(\theta_i + \delta_l) - \psi(\theta_i) + \ln(X_{il}))] =$$

$$= \sum_{\substack{j,l=1 \\ j \neq l}}^{m} M[(\psi(\theta_i + \delta_j) - \psi(\theta_i) + \ln(X_{ij}))] \cdot M[(\psi(\theta_i + \delta_l) - \psi(\theta_i) + \ln(X_{il}))] +$$

$$+ \sum_{j=1}^{m} M[(\psi(\theta_i + \delta_j) - \psi(\theta_i) + \ln(X_{ij}))^2].$$

Приравнивая производные (2) к нулю, можно получить уравнения для оценок параметров методом максимального правдоподобия. Поскольку оценки, полученные методом максимального правдоподобия и методом моментов, сходятся к одним и тем же истинным значениям параметров, то в последней сумме слагаемые для $j \neq l$ близки к нулю, и ими можно пренебречь. Таким образом,

$$I(\theta_i; \{X_{ij}\}) = M\left[\left(\frac{d\ln(L)}{d\theta_i}\right)^2\right] \approx \sum_{j=1}^{m} M[(\psi(\theta_i + \delta_j) - \psi(\theta_i) + \ln(X_{ij}))^2].$$

Аналогичные выкладки можно привести для параметров экзаменов

$$\frac{\partial \ln(L)}{\partial \delta_j} = \sum_{i=1}^{n}\left(\frac{1}{\Gamma(\theta_i + \delta_j)}\frac{\partial \Gamma(\theta_i + \delta_j)}{\partial \theta_i} - \frac{1}{\Gamma(\delta_j)}\frac{\partial \Gamma(\delta_j)}{\partial \delta_j} + \ln(1 - x_{ij})\right) =$$

$$= \sum_{i=1}^{n}(\psi(\theta_i + \delta_j) - \psi(\delta_j) + \ln(1 - x_{ij}));$$

$$I(\delta_j; \{X_{ij}\}) = M\left[\left(\frac{d\ln(L)}{d\delta_j}\right)^2\right] = M\left[\left(\sum_{i=1}^{n}(\psi(\theta_i + \delta_j) - \psi(\delta_j) + \ln(1 - X_{ij}))\right)^2\right] =$$

$$= \sum_{i,h=1}^{n} M[(\psi(\theta_i + \delta_j) - \psi(\delta_j) + \ln(1 - X_{ij})) \cdot (\psi(\theta_h + \delta_j) - \psi(\delta_j) + \ln(1 - X_{il}))] \approx$$

$$\approx \sum_{i=1}^{n} M[(\psi(\theta_i + \delta_j) - \psi(\delta_j) + \ln(1 - X_{ij}))^2].$$

Зависимость нижней границы дисперсии параметра экзамена от величины параметра для оценок одного потока студентов (рис. 3) демонстрирует рост нижней границы дисперсии при увеличении значения параметра. Тем не менее, нижняя граница дисперсии будет уменьшаться при увеличении количества оценок студента. Аналогично ведет себя нижняя граница дисперсии оценки параметра экзамена — повышается при увеличении параметра и уменьшается при увеличении оценок одного экзамена.

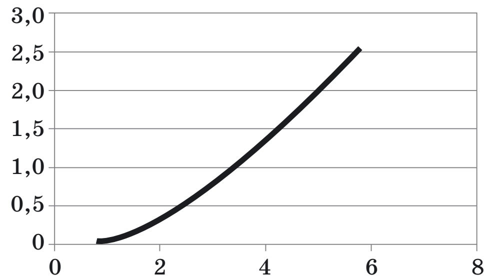

Рис. 3. Зависимость нижней границы дисперсии оценки параметра θ_i *от величины параметра*

В целом, приведенные формулы и обработка оценок одного потока студентов демонстрируют возможность применения данного метода определения параметров подготовленности студентов и сложности экзаменов. Предложенная модель стобалльной оценки на основе бета-распределения может использоваться для исследования подготовленности студентов и сложности экзаменов, в частности выделения неоправданно легких или сложных экзаменов.

Список использованной литературы

1. Братищенко В. В. Параметрическая модель экзаменационных оценок / В. В. Братищенко // Качество. Инновации. Образование. — 2012. — № 3. — С. 32–35.

2. Дубина И. Н. Математические основы эмпирических социально-экономических исследований : учеб. пособие / И. Н. Дубина. — Барнаул : Изд-во Алт. ун-та, 2006. — 263 с.

3. Ким В. С. Тестирование учебных достижений : монография / В. С. Ким. — Уссурийск : Изд-во Уссур. гос. пед. ин-т, 2007. — 214 с.

4. Крамер Г. Математические методы статистики : пер. с англ. / Г. Крамер ; ред.: А. С. Монин, А. А. Петров, А. Н. Колмогоров. — 2-е изд. — М. : Мир, 1975. — 648 с.

5. Маслак А. А. Измерение латентных переменных в образовании и других социально-экономических системах: теория и практика / А. А. Маслак. — Славянск н/Кубани : Изд. центр Славян.-на-Кубани гос. пед. ин-та, 2007. — 424 с.

6. Нейман Ю. М. Введение в теорию моделирования и параметризации педагогических тестов / Ю. М. Нейман, В. А. Хлебников. — М. : Прометей, 2000. — 168 с.

7. Родионов А. В. Модификация рейтинговой параметрической модели оценки латентных факторов для измерения уровня сформированности компетенций / А. В. Родионов // Известия Иркутской государственной экономической академии. — 2014. — № 6 (98). — С. 168–174. — DOI : 10.17150/1993-3541.2014.24(6).168-174.

8. Родионов А. В. Применение IRT-моделей для анализа результатов обучения в рамках компетентностного подхода / А. В. Родионов, В. В. Братищенко // Современные проблемы науки и образования. — 2014. — № 4. — URL : www.science-education.ru/118-13858.

9. Andrich D. A rating formulation for ordered response categories / D. Andrich // Psychometrika. — 1978. — Vol. 43, iss. 4. — P. 561–573.

10. Masters G. N. A Rasch model for partial credit scoring / G. N. Masters // Psychometrika. — 1982. — Vol. 47, no. 2. — P. 149–174.

11. Rasch G. Probabilistic models for some intelligence and attainment tests / G. Rasch. — Chicago : The University of Chicago Press, 1980. — 199 p.

12. Wright B. D. Rating Scale Analysis. Rasch Measurement / B. D. Wright, G. N. Masters. — Chicago : Mesa Press, 1979. — 206 p.

References

1. Bratishenko V. V. Parametric model of examination grades. *Kachestvo. Innovatsii. Obrazovanie = Quality. Innovations. Education*, 2012, no. 3, pp. 32–35. (In Russian).

2. Dubina I. N. *Matematicheskie osnovy jempiricheskih social'no-jekonomicheskih issledovanij* [Mathematical bases of empiric socio-economic studies]. Barnaul, Altai State University Publ., 2006. 263 p.

3. Kim V. S. *Testirovanie uchebnyh dostizhenij* [Testing academic achievements]. Ussuriysk State Pedagogical Institute Publ., 2007. 214 p.

4. Cramer H. *Mathematical Methods of Statistics*. Princeton University Press, 1962. 590 p. (Russ. ed.: Cramer H. *Matematicheskie metody statistiki*. Moscow, Mir Publ., 1975. 648 p.

5. Maslak A. A. *Izmerenie latentnyh peremennyh v obrazovanii i drugih social'no-jekonomicheskih sistemah: teorija i praktika* [Measuring latent variables in education and other socio-economic systems: theory and practice]. Slavyansk-on-Kuban' State Pedagogical University Publ., 2007. 424 p.

6. Neyman Ju. M., Khlebnikov V. A. *Vvedenie v teoriju modelirovanija i parametrizacii pedagogicheskih testov* [Introduction to theory of modelling and parametrization of training tests]. Moscow, Prometey Publ., 2000. 168 p.

7. Rodionov A. V. Modification of the rating parametric model of latent factors estimation for assessing the level of competence formation. *Izvestiya Irkutskoy gosudarstvennoy ekonomicheskoy akademii = Izvestiya of Irkutsk State Economics Academy*, 2014, no. 6 (98), pp. 168–174. (In Russian). DOI: 10.17150/1993-3541.2014.24(6).168-174.

8. Rodionov A. V., Bratischenko V. V. Application of irt-model for analysis training results within the competence approach. *Sovremennye problemy nauki i obrazovaniya = Modern Problems of Science and Education*, 2014, no. 4. (In Russian). Available at: www.science-education.ru/118-13858.

9. Andrich D. A rating formulation for ordered response categories. *Psychometrika*, 1978, vol. 43, iss. 4, pp. 561–573.

10. Masters G. N. A Rasch model for partial credit scoring. *Psychometrika*, 1982, vol. 47, no. 2, pp. 149–174.

11. Rasch G. *Probabilistic models for some intelligence and attainment tests*. Chicago, The University of Chicago Press, 1980. 199 p.

12. Wright B. D., Masters G. N. *Rating Scale Analysis. Rasch Measurement*. Chicago, Mesa Press, 1979. 206 p.

Информация об авторе

Братищенко Владимир Владимирович — кандидат физико-математических наук, доцент, заведующий кафедрой информатики и кибернетики, Байкальский государственный университет экономики и права, 664003, г. Иркутск, ул. Ленина, 11, e-mail: bvv@isea.ru.

Author

Vladimir V. Bratishchenko — PhD in Physics and Mathematics, Assistant Professor, Head of Chair of Computer Science Cybernetics department, Baikal State University of Economics and Law, 11 Lenin St., 664003, Irkutsk, Russian Federation; e-mail: bvv@isea.ru.

Библиографическое описание статьи

Братищенко В. В. Латентная модель стобалльной оценки на основе бета-распределения / В. В. Братищенко // Baikal Research Journal. — 2015. — Т. 6, № 5. — DOI : 10.17150/2411-6262.2015.6(5).9.

Reference to article

Bratishchenko V.V. Latent model of a 100-point grade on the basis of beta distribution. *Baikal Research Journal*, 2015, vol. 6, no. 5. DOI : 10.17150/2411-6262.2015.6(5).9. (In Russian).