DOI 10.17150/2411-6262.2020.11(4).3 http://brj-bguep.ru

УДК 519.24

В.В. Братищенко
Байкальский государственный университет,
г. Иркутск, Российская Федерация

ПРИМЕНЕНИЕ БЕТА-БИНОМИАЛЬНОГО РАСПРЕДЕЛЕНИЯ ДЛЯ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ ОЦЕНОК СТУДЕНТОВ

АННОТАЦИЯ. Статистическая обработка данных студентов применяется для решения различных задач в управлении учебным процессом. Традиционные методы математической статистики не соответствуют свойствам оценок студентов. В работе предлагается адаптировать модели IRT для описания оценок с целью устранения отмеченного недостатка. Предлагается модель оценки на основе бета-биномиального распределения, в которой учитывается подготовленность студента и трудность аттестации в виде соответствующих параметров. Применение этого распределения вероятностей обосновывается возможностью учитывать разброс оценок. Для поиска параметров предлагается использовать два подхода. Первый подход основан на применении численных методов поиска параметров для решения системы уравнений. Второй — использует байесовский подход, что позволяет достаточно эффективно уточнять параметры при получении новых оценок. По результатам обработки массива оценок адекватность модели подтверждена по нескольким статистическим критериям.

КЛЮЧЕВЫЕ СЛОВА. Статистическая обработка оценок, современная теория тестирования, латентные параметры, бета-биномиальное распределение, байесовские оценки.

ИНФОРМАЦИЯ О СТАТЬЕ. Дата поступления 5 ноября 2020 г.; дата принятия к печати 15 декабря 2020 г.; дата онлайн-размещения 31 декабря 2020 г.

V.V. Bratischenko
Baikal State University,
Irkutsk, Russian Federation

APPLICATION OF THE BETA-BINOMIAL DISTRIBUTION FOR STATISTICAL PROCESSING OF STUDENTS' GRADES

ABSTRACT. Statistical processing of students' grades is used to solve various problems in managing the educational process. Traditional methods of mathematical statistics do not correspond to the properties of student grades. The paper proposes to adapt the IRT models for describing grades to improve the whole process. A model based on beta-binomial distribution is proposed, in which the parameters take into account the student's ability and the difficulty of getting a passing grade. The use of this probability distribution is justified by the ability to take into account the variation of grades. It is proposed to use two approaches to identify the parameters. The first approach is based on the use of numerical methods for finding parameters for solving a system of equations. The second one uses the Bayesian approach, which makes it possible to rather effectively refine the parameters when obtaining new estimates. Based on the results of processing an array of estimates, the adequacy of the model was confirmed by several statistical criteria.

KEYWORDS. Statistics of grades, Item Response Theory, latent parameters, Beta-Binomial Distribution, Bayesian estimates.

ARTICLE INFO. Received November 5, 2020; accepted December 15, 2020; available online December 31, 2020.

Статистическая обработка данных успеваемости может быть использована для решения широкого круга задач от подведения итогов обучения до управления качеством учебного процесса и прогнозирования академических успехов студентов. Вычисление по оценкам широко применяемых статистических характеристик, таких как математическое ожидание и дисперсия, не вполне правомерно, так как оценки студентов измеряются не в метрических, а в ординальных шкалах, которые в общем отличаются для разных дисциплин и преподавателей. Традиционные методы обработки оценок студентов основаны на модели генеральной совокупности. Применение таких методов продемонстрировало статистически значимую разницу в характеристиках множеств оценок, выставленных разными преподавателями, также как множеств оценок, полученных разными студентами [1]. Это вынуждает искать другие модели для описания статистических характеристик оценок.

Такой инструмент предлагает современная теория тестирования (IRT — Item Response Theory) [2] для обработки результатов выполнения тестовых заданий разной степени сложности студентами с разными уровнями подготовки. В IRT вероятность

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

правильного ответа на тестовое задание зависит от «трудности» задания, характеризуемого параметром δ, уровня «подготовленности» тестируемого, задаваемого параметром θ. Так как эти переменные не наблюдаются непосредственно, то их называют латентными. Вероятность правильного ответа увеличивается при возрастании подготовленности и уменьшается при увеличении трудности. Максимум дифференцирующей способности теста наступает при совпадении уровней трудности и согласованности.

Привлекательной особенностью IRT является то, что она переводит оценки, которые измеряются в ординальных шкалах, в уровни, измеряемые в метрической шкале, которая получила название шкалы логитов. Недостатком является применение численных методов для определения латентных параметров.

Первоначально IRT разрабатывалась для обработки дихотомических заданий с двумя результатами выполнения: 1 — «правильно» и 0 — «неправильно». Впоследствии этот подход был распространен [3; 4] на политомические задания с несколькими вариантами ответа различной степени правильности (точности). При этом два основных латентных параметра (трудность задания и подготовленность тестируемого) дополняются параметрами вариантов ответа. Такая универсальная модель позволяет разделить по трудности варианты ответа. Однако, для надежности оценивания большого количества параметров требуется больше наблюдений.

В работах [5; 6] предлагается применить это подход для оценок преподавателей. Прямое применение политомических моделей оценок затруднительно, так как на экзаменах, зачетах и в текущей успеваемости применяются разнообразные оценочные средства. Кроме этого, применяемые оценочные средства в меньшей степени формализованы по сравнению с тестами. Для исследования оценок преподавателей предлагается [5] использовать биномиальную модель, в которой вероятность успеха, также как в IRT, зависит от «трудности» аттестации и «подготовленности» студента, а второй параметр — от количества градаций в шкале оценивания.

В биномиальной модели вероятность того, что i-й студент получит оценку $X_{ij} = k \in \{0, \dots, N\}$ за j-ое задание, имеет биномиальное распределение

$$P\left(X_{ij} = k\right) = C_N^k p_{ij}^k q_{ij}^{N-k},$$

где $p_{ij} = \frac{e^{\theta_i}}{e^{\theta_i} + e^{\delta_j}}$,

$$q_{ij} = 1 - p_{ij} = \frac{e^{\delta_j}}{e^{\theta_i} + e^{\delta_j}},$$

θ_i — параметр «подготовленности» i-го студента ($i = 1, \ldots, n$), δ_j – параметр «трудности» j-го задания ($j = 1, \ldots, m$).

Все оценки полагаются независимыми в совокупности. Биномиальное распределение соответствует дискретной природе оценок и предположению, что оценка преподавателя складывается под действием многостороннего исследования выполнения задания студентом и должна быть близкой к сумме бернуллиевских случайных величин.

Исследования оценок показали, что модель в целом проходит статистические проверки по критерию Фишера и традиционные для IRT проверки с применением Infit и Outfit статистик [2; 5]. Все перечисленные критерии используют математические ожидания, вычисленные в соответствии с предложенной моделью. В то же время обнаружилось, что оценки имеют дисперсию большую, чем это предсказывает модель. Для устранения такого недостатка можно использовать распределение близкое к биномиальному, но имеющее дополнительные параметры, влияющие на дисперсию. Таким распределением является бета-биномиальное.

Для описания оценки предлагается использовать следующее Бета-биномиальное распределение

$$P\{X_{ij} = k\} = C_{k_j}^{k} p_{ij}^{k} \left(1 - p_{ij}\right)^{k_j - k}$$

где k_j — количество градаций оценки j -го задания, p_{ij} — случайная величина, имеющая бета-распределение

$$f(x) = \frac{x^{\theta_i - 1}(1 - x)^{\delta_j - 1}}{\mathrm{B}\left(\theta_i, \delta_j\right)},$$

которое зависит от параметров $\theta_i > 0$ «подготовленности» i-го студента и $\delta_j > 0$ «трудности» j-го задания.

В результате усреднения по распределению p_{ij} получаем вероятности оценок

$$P\{X_{ij} = k\} = \frac{\Gamma\left(k_j + 1\right)}{\Gamma(k + 1)\Gamma\left(k_j - k + 1\right)} \frac{\Gamma(k + \theta_i)\Gamma\left(k_j - k + \delta_j\right)}{\Gamma\left(k_j + \theta_i + \delta_j\right)} \frac{\Gamma\left(\theta_i + \delta_j\right)}{\Gamma(\theta_i)\Gamma\left(\delta_j\right)}$$

и числовые характеристики оценки

$$M\left[X_{ij}\right] = \frac{k_j \theta_i}{\theta_i + \delta_j},$$

$$D\left[X_{ij}\right] = \frac{k_j \theta_i \delta_j \left(k_j + \theta_i + \delta_j\right)}{\left(\theta_i + \delta_j\right)^2 \left(1 + \theta_i + \delta_j\right)}.$$

Особенностью параметров θ_i и δ_j в данной модели является то, что при умножении их на некоторый положительный коэффициент γ математическое ожидание не изменится, а дисперсия будет варьироваться в широких пределах от

$k_j p_{ij}(1-p_{ij})$ при $\gamma \to \infty$, что соответствует дисперсии биномиального распределения, до $k_j^2 p_{ij}(1-p_{ij})$ при $\gamma \to 0$.

При обработке оценок (особенно оценок текущей успеваемости) достаточно часто встречаются неполные наборы, если некоторые студенты по каким-либо причинам не сдают экзамены или не выполняют учебные задания. Для описания неполного набора оценок введем индикаторы

$$I_{ij} = \begin{cases} 1, \text{если есть оценка } i-\text{го студента за } j-\text{е задание}, \\ 0, \text{если нет оценки } i-\text{го студента за } j-\text{е задание}. \end{cases}$$

Для оценки параметров модели можно воспользоваться методом моментов. Используя формулу для математического ожидания случайной величины X_{ij} и выполняя суммирования по i и j, получаем уравнения

$$\sum_{j=1}^{m} I_{ij} \frac{k_j \theta_i}{\theta_i + \delta_j} - \sum_{j=1}^{m} I_{ij} x_{ij} = 0,$$

$$\sum_{i=1}^{n} I_{ij} \frac{k_j \theta_i}{\theta_i + \delta_j} - \sum_{i=1}^{n} I_{ij} x_{ij} = 0,$$

которые можно использовать для подбора параметров. Традиционно в IRT для этого используется метод Ньютона (касательных). В этом методе итерационная формула

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

использует производные по параметрам

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

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

В итоге получаем формулы

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

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

для определения оценок параметров численными методами. Для вычисления начальных значений положим, что в первом уравнении все параметры экзаменов одинаковы $\delta_j^{(0)} = \bar{\delta}$, в этом случае

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

Аналогично полагая $\theta_i^{(0)} = \bar{\theta}$, получаем

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

Значения $\bar{\theta} = 7{,}5$ и $\bar{\delta} = 2{,}5$ обеспечивают начальные значения оценок, соответствующие самой массовой оценке — «хорошо». Вычислительный процесс обладает хорошей сходимостью при соблюдении условия вариативности оценок: отсутствуют или исключены ситуации, когда все оценки некоторого задания минимальны или максимальны. Можно показать, что поиск параметров сводится к минимизации неотрицательной квадратичной формы, что обеспечивает сходимость итерационных вычислений.

Для оценки латентных параметров бета-биномиального распределения можно применить байесовский подход. Привлекательность подхода обусловлена возможностью скорректировать уже известные оценки с учетом новых наблюдений по известной формуле Байеса. Существенную роль в реализации байесовского подхода играют распределения, сопряженные с наблюдаемой генеральной совокупностью. Если априорное распределение принадлежит семейству сопряженных распределений, то и апостериорное принадлежит этому семейству.

В [7] показано что для биномиального распределения с неизвестным параметром p (вероятность успеха), семейство сопряженных распределений принадлежит классу бета-распределений. Для применения байесовского подхода к оценке латентных параметров предположим, что априорным распределением является бета распределение:

$$f(x) = \frac{x^{\theta_i - 1}(1 - x)^{\delta_j - 1}}{\mathrm{B}\left(\theta_i, \delta_j\right)},$$

Пусть в результате очередного измерения была получена оценка $X_{ij} = x_{ij}$. В соответствии с байесовским подходом апостериорное распределение

$$f\left(x \middle| x_{ij}\right) = \frac{f(x) P\left\{X_{ij} = x_{ij} \middle| \theta_i, \delta_j\right\}}{P\left\{X_{ij} = x_{ij}\right\}} =$$

$$= \frac{C_{k_j}^{x_{ij}} x^{x_{ij}} (1-x)^{k_j - x_{ij}} \frac{x^{\theta_i - 1}(1-x)^{\delta_j - 1}}{\mathrm{B}(\theta_i, \delta_j)}}{\frac{\Gamma(k_j+1)}{\Gamma(x_{ij}+1)\Gamma(k_j - x_{ij}+1)} \frac{\Gamma(x_{ij}+\theta_i)\Gamma(k_j - x_{ij} + \delta_j)}{\Gamma(k_j+\theta_i+\delta_j)} \frac{\Gamma(\theta_i+\delta_j)}{\Gamma(\theta_i)\Gamma(\delta_j)}} =$$

$$= \frac{x^{\theta_i + x_{ij} - 1}(1-x)^{\delta_j + (k_j - x_{ij}) - 1}}{\mathrm{B}\left(\theta_i + x_{ij}, \delta_j + (k_j - x_{ij})\right)}$$

также является бета-распределением с параметрами $\theta_i + x_{ij}, \delta_j + (k_j - x_{ij})$.

Рассматривая апостериорные распределения для всего множества наблюдаемых оценок, получаем следующие параметры бета-распределения латентных переменных

$$\theta_i = \theta_{i0} + \sum_{j=1}^{m} I_{ij} x_{ij}, i = 1, \ldots, n$$

$$\delta_j = \delta_{j0} + \sum_{i=1}^{n} I_{ij}(k_j - x_{ij}), j = 1, \ldots, m$$

В качестве начальных значений θ_{i0} и δ_{j0} можно выбрать предварительно вычисленные оценки или выбрать небольшие положительные значения, чтобы исключить ситуацию нулевого значения параметров, если все оценки будут нулевыми или максимальными.

Обработка данных показала, что совпадают математические ожидания оценок биномиальной и бета-биномиальной моделей, а оценка дисперсии остатков совпадает с вычисленным значением бета-биномиальной модели.

Список использованной литературы

1. Братищенко В.В. Статистический анализ экзаменационных оценок / В.В. Братищенко // Известия Иркутской государственной экономической академии (Байкальский государственный университет экономики и права). — 2011. — № 3. — URL: http://brj-bguep.ru/reader/article.aspx?id=8014.

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

3. Masters G.N. A Rasch model for partial credit scoring / G.N. Masters // Psychometrika. — 1982. — Vol. 47, iss. 2. — P. 149–174.

4. Andrich D. A Rating Formulation for Ordered Response Categories / D. Andrich // Psychometrika. — 1978. — Vol. 43, iss. 4. — P. 561–573.

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

6. 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.

7. Айвазян С.А. Байесовский подход в эконометрическом анализе / С.А. Айвазян // Прикладная эконометрика. — 2008. — № 1 (9). — С. 93–130.

References

1. Bratishenko V.V. Statistic Analysis of Examination Grades. *Izvestiya Irkutskoy gosudarstvennoy ekonomicheskoy akademii (Baykalskiy gosudarstvennyy universitet ekonomiki i prava) = Izvestiya of Irkutsk State Economics Academy (Baikal State University of Economics and Law)*, 2011, no. 3. Available at: http://brj-bguep.ru/reader/article.aspx?id=8014.

2. 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.

3. Masters G.N. A Rasch Model for Partial Credit Scoring. *Psychometrika*, 1982, vol. 47, iss. 2, pp. 149–174.

4. Andrich D. A Rating Formulation for Ordered Response Categories. *Psychometrika*, 1978, vol. 43, iss. 4, pp. 561–573.

5. Bratishenko V.V. Parametric Model of Examination Grades. *Kachestvo. Innovatsii. Obrazovanie = Quality. Innovations. Education*, 2012, no. 3, pp. 32–35. (In Russian).

6. 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. Available at: www.science-education.ru/118-13858. (In Russian).

7. Aivazian S.A. Bayesian Methods in Econometrics. *Prikladnaya ekonometrika = Applied econometrics*, 2008, no. 1 (9), pp. 93–130. (In Russian).

Информация об авторе

Братищенко Владимир Владимирович — кандидат физико-математических наук, доцент, кафедра математических методов и цифровых технологий, Байкальский государственный университет, г. Иркутск, Российская Федерация, e-mail: vvb@bgu.ru.

Author

Vladimir V. Bratischenko — PhD in Physics and Mathematical Sciences, Associate Professor, Department of Mathematical Methods and Digital Technologies, Baikal State University, Irkutsk, Russian Federation, e-mail: vvb@bgu.ru.

Для цитирования

Братищенко В.В. Применение бета-биномиального распределения для статистической обработки оценок студентов / В.В. Братищенко. — DOI: 10.17150/2411-6262.2020.11(4).3 // Baikal Research Journal. — 2020. — Т. 11, № 4.

For Citation

Bratischenko V.V. Application of the Beta-Binomial Distribution for Statistical Processing of Student Grades. *Baikal Research Journal*, 2020, vol. 11, no. 4. DOI: 10.17150/2411-6262.2020.11(4).3. (In Russian).