Proceedings of the 3rd International Conference on Mathematical Physics and Computational Simulation
DOI: 10.54254/2753-8818/100/2025.21960

Evaluation of Students' Ability Through the Continuous Rasch Model

Tengyi Liu

Institute of School of Mathematical Science, Peking University, BeiJing, China
2200010836@stu.pku.edu.cn

Abstract: This paper proposes an extension of the traditional Rasch model to evaluate student abilities and question difficulties in educational settings that utilize continuous scoring systems. This enhanced model incorporates maximum likelihood estimation (MLE) to calculate ability and difficulty parameters with demonstrated consistency and asymptotic normality. Traditional assessments often fall short by not accounting for errors from high-performing students and providing limited insights into student abilities beyond right or wrong answers. The continuous Rasch model developed here accommodates the complexities of continuous data and broadens the applicability of the Rasch model beyond dichotomous data types, making it suitable for modern educational evaluations. The paper introduces a continuous version of the Rasch model, leveraging it to derive a more comprehensive understanding of the interactions between student performance and test difficulty. This model views each scored response as a sum of independent Bernoulli trials, thus enabling a nuanced analysis of scores within a continuous range. Such a framework is particularly beneficial in scenarios where traditional dichotomous scoring fails to capture the depth of student understanding and question complexity.

Keywords: Rasch Model, Maximum Likelihood Estimation, Consistency, Asymptotic Normality

1. Introduction

Educational assessment has traditionally relied on test scores to evaluate student abilities. Yet, this method often does not fully capture the complexities of student performance, particularly for high-achieving students who may still make errors [1]. Conventional assessment techniques, such as multiple-choice questions, provide a limited perspective on student capabilities and struggle to capture the intricacies of continuous scoring systems. Introduced by Georg Rasch in 1960, the Rasch model has been fundamental in item response theory (IRT) for analyzing both the difficulty of questions and the abilities of students [1]. However, its utility is predominantly confined to dichotomous data, rendering it less effective in contemporary educational assessments where continuous scoring is more common.

Recent statistical advancements, including the sparse Rasch model and Bayesian nonparametric approaches, strive to overcome these limitations by introducing sparsity and heterogeneity into item response data [2,3]. These innovations underscore the necessity for more sophisticated models capable of navigating the complexities of modern educational assessments [4].

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/100/2025.21960

This study presents a continuous Rasch model that expands the traditional framework to support continuous scoring systems, thus offering a more detailed evaluation of student abilities and question difficulties. Utilizing maximum likelihood estimation (MLE), this model focuses on deriving precise ability and difficulty parameters while ensuring the consistency and asymptotic normality of these estimates. It integrates recent progress in sparse Rasch models and Bayesian nonparametric methods to bolster the robustness and relevance of the proposed model.

Specifically, the study sets forth to:

- Develop a continuous Rasch model that effectively assesses student abilities and question difficulties in educational assessments with continuous scoring.
- Establish a solid theoretical foundation for the consistency and asymptotic normality of MLE within this new model framework.
- Demonstrate the practical utility of the model through real data analysis, validating its effectiveness in educational settings.

By achieving these objectives, this research enhances the field of educational assessment by providing a more adaptable and comprehensive framework for evaluating student performance, especially in contexts where continuous scoring is pivotal.

2. Model Development

2.1. Rasch Model

In 1960, Georg Rasch introduced a probabilistic model to characterize difficulties of questions and abilities of students [1]:

$$\Pr(\text{student A correctly answers question B}) = \frac{e^{\alpha-\beta}}{1+e^{\alpha-\beta}} \tag{1}$$

where α denotes the ability parameter of student A and β denotes the difficulty parameter of question B.

However, an exam in the university always includes more than one question, so a more advanced model is needed.

2.2. A Continuous Version of Rasch Model

When there are more, supposed, more than 10 questions, in an exam or an assignment, Rasch model can be used to evaluate the probability of the student correctly answering each problem, so, combining the answer status, right or wrong, of each question, can get the total score expectation of the student [5]. Using question amount set to be 10 as an example. The following is a combination of binomial distribution and Rasch model:

$$\Pr(\text{student A correctly answers question B}) = \binom{10}{t} \cdot \left(\frac{e^{\alpha-\beta}}{1+e^{\alpha-\beta}}\right)^{t} \cdot \left(\frac{e^{\alpha-\beta}}{1+e^{\alpha-\beta}}\right)^{10-t} \tag{2}$$

In this formula, α denotes the ability parameter of student A again, but β denotes the total difficulty parameter of the whole exam, not the difficulty of one specific question once.

However, an exam can include some questions that may score continuously in the range 0 to the full score. In this case, the classical Rasch model or the combination with binomial distribution can be not so efficient, so generalize the following continuous version.

Viewing this kind of questions as mixtures of several steps with each step being regarded as independent Bernoulli variables and recalling the asymptotic normality of binomial variables, the sum of independent Bernoulli variables, develop the following formula:

DOI: 10.54254/2753-8818/100/2025.21960

$$K \sim N(mp, mpk(1-p)) \tag{3}$$

where K denotes student A's score of a specific question with continuous distributed score, m denotes the total score of the question, k denotes the total score of each step, set to be all the same and small enough with respect to the total score of the question, and p is defined relating to Rasch model as below:

$$p = \frac{e^{\alpha-\beta}}{1+e^{\alpha-\beta}} \tag{4}$$

in which α denotes the ability parameter of student A, and β denotes the difficulty parameter of the question.

In practical trials, need to cut the outliers, that is the possibility when K is less than 0 or greater than full score m, but to simplify the calculation in real data analysis, assume K just follows normal distribution.

After considering students' expected score of a specific question, can use the same method to estimate students' total score in a quiz.

$$K_{total} \sim N(M\hat{p}, M\hat{p}k(1-\hat{p})) \tag{5}$$

$$\hat{p} = \frac{e^{\alpha-\hat{\beta}}}{1+e^{\alpha-\hat{\beta}}} \tag{6}$$

Similarly, K_{total} denotes student A's total score in this quiz, M denotes the total score of this quiz, k denotes the total score of each step which is small enough with respect to the total score of the quiz, and $\hat{\beta}$ denotes the average difficulty parameter of this quiz.

3. Methodology

3.1. Maximum Likelihood Estimation

In this section, use the following notations.

Let N_1, N_2 denote the number of students and quizzes respectively, M_i denote the total score of thei^{th}quiz where $i = 1,2,\cdots,N_2$, $\alpha_i$denote the ability parameter of the i^{th}student where $i = 1,2,\cdots,N_1$, β_i denote the difficulty parameter of the ith quiz where $i = 1,2,\cdots,N_2$, $X_{i,j}$ denote the score student i get in the jth quiz.

Using the continuous model and setting $k = 1$, derive the negative log-likelihood function with parameter $\vec{\theta} = (\theta_1, \theta_2, \cdots, \theta_{N_1+N_2}) = (\alpha_1, \alpha_2, \cdots, \alpha_{N_1}, \beta_1, \beta_2, \cdots, \beta_{N_2})$ in $R^{N_1+N_2}$:

$$L(\vec{\theta}) = \frac{1}{2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \log\left(2\pi M_j \mu(\alpha_i - \beta_j)\left(1-\mu(\alpha_i-\beta_j)\right)\right) + \frac{\left(X_{I,j}-M_j\mu(\alpha_i-\beta_j)\right)^2}{M_j\mu(\alpha_i-\beta_j)\left(1-\mu(\alpha_i-\beta_j)\right)} =$$

$$\frac{1}{2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \log\left(2\pi M_j \mu(\theta_i - \theta_{j+N_1})\left(1-\mu(\theta_i-\theta_{j+N_1})\right)\right) + \frac{\left(X_{i,j}-M_j\mu(\theta_i-\theta_{j+N_1})\right)^2}{M_j\mu(\theta_i-\theta_{j+N_1})\left(1-\mu(\theta_i-\theta_{j+N_1})\right)} \tag{7}$$

where $\mu(x) = \frac{e^x}{1+e^x}$.

DOI: 10.54254/2753-8818/100/2025.21960

To maximize the likelihood function or to minimize the negative log-likelihood function, consider the partial derivatives of $l(\vec{\theta})$, and assume the maximum likelihood estimator $\hat{\theta}$ satisfies the following equation set, if exists:

$$\frac{\partial l(\hat{\theta})}{\partial \hat{\theta}_i} = 0, \text{ for all } i \in \{1,2,\cdots,N_1+N_2\} \tag{8}$$

After rearranging the system of equations, obtain the following expressions:

$$\sum_{J=1}^{N_2}\left(1-2\mu\left(\hat{\theta}_i-\widehat{\theta_{j+N_1}}\right)\right)+\frac{\left(X_{I,j}-M_j\mu\left(\hat{\theta}_i-\widehat{\theta_{j+N_1}}\right)\right)\left(2X_{I,j}\mu\left(\hat{\theta}_i-\widehat{\theta_{j+N_1}}\right)-M_j\mu\left(\hat{\theta}_i-\widehat{\theta_{j+N_1}}\right)-Xi,j\right)}{M_j\mu\left(\hat{\theta}_i-\widehat{\theta_{j+N_1}}\right)\left(1-\mu\left(\hat{\theta}_i-\widehat{\theta_{j+N_1}}\right)\right)}=0, \text{ for } i \in \{1,2,\cdots,N_1\} \tag{9}$$

$$\sum_{i=1}^{N_1}\left(1-2\mu\left(\hat{\theta}_i-\widehat{\theta_{j+N_1}}\right)\right)+\frac{\left(X_{i,j}-M_j\mu\left(\hat{\theta}_i-\widehat{\theta_{j+N_1}}\right)\right)\left(2X_{i,j}\mu\left(\hat{\theta}_i-\widehat{\theta_{j+N_1}}\right)-M_j\mu\left(\hat{\theta}_i-\widehat{\theta_{j+N_1}}\right)-Xi,j\right)}{M_j\mu\left(\hat{\theta}_i-\widehat{\theta_{j+N_1}}\right)\left(1-\mu\left(\hat{\theta}_i-\widehat{\theta_{j+N_1}}\right)\right)}=0, \text{ for } j \in \{1,2,\cdots,N_2\} \tag{10}$$

Noticing that use $\hat{\theta}+c$ to substitute $\hat{\theta}$ where c is a constant that can be arbitrarily chosen, so in the following two sections may assume $\hat{\theta}_i = 0$ to better consider the characteristic of $\hat{\theta}$.

Please remember that all the papers must be in English and without orthographic errors [6].

Do not add any text to the headers (do not set running heads) and footers, not even page numbers, because text will be added electronically.

For a best viewing experience the used font must be Times New Roman, on a Macintosh use the font named times, except on special occasions, such as program code.

3.2. Consistency

In Peng, P.'s work, they considered the combination of the Erdős–Rényi graph and the sparse Rasch model, and proved the existence and consistency, or to say uniqueness, of the MLE [6,7]. In Pan, T.'s work, they proved the consistency of the MLE in binomial distribution using classical Rasch model. Noticing the normal distribution is an asymptotic binomial distribution, the same method to prove the existence and uniqueness of the MLE under mild assumption [8,9].

In real data analysis, do not need to consider the existence or the uniqueness if there is exact one solution generated by the computer when solving the non-linear equation set [10,11].

3.3. Asymptotic Normality

Under the assumption of $\hat{\theta_1} = 0$, let $I\left(\hat{\theta}'\right)$ be the Fisher information matrix of parameter $\hat{\theta}' = \left(0,\hat{\theta_2},\hat{\theta_3},\cdots,\widehat{\theta_{N_1+N_2}}\right)$, then the inverse of the Fisher information matrix $I^{-1}\left(\hat{\theta}'\right) = \left(v_{i,j}\right)_{i,j\in\{1,2,\cdots,N_1+N_2\}}$ equals to the Hessian matrix of $l\left(\hat{\theta}'\right) = l\left(0,\hat{\theta_2},\hat{\theta_3},\cdots,\widehat{\theta_{N_1+N_2}}\right)$, in which still view $\hat{\theta_1}$ as a parameter first.

According to Peng, P.'s work, under mild assumption, vector $\left(\hat{\theta_2}-\theta_2,\hat{\theta_3}-\theta_3,\cdots,\widehat{\theta_{N_1+N_2}}-\theta_{N_1+N_2}\right)$ asymptotically follows a multivariate normal distribution with zero mean and covariance

DOI: 10.54254/2753-8818/100/2025.21960

matrix $\Sigma = (\sigma_{ij})$, where $\sigma_{ij} = \frac{\delta_{ij}}{v_{ii}} + \frac{1}{v_{11}}$, δ_{ij} denotes the Kronecker delta function in which $\delta_{ij} = 1$ if $i = j$ and otherwise $\delta_{ij} = 0$, and $\theta_i, i = 2,3,\cdots,N_1 + N_2$ denotes the true value of the parameter [2].

4. Real Data Analysis

In the application of the continuous Rasch model, the University Students Marks Sheet is utilized to illustrate its effectiveness. The analysis focuses on the third sheet of this dataset, which contains two columns of student grades. The first column represents the sessional score of students, capped at 75 points, encompassing various small quizzes and assignments consolidated into one composite test score. The second column reflects the terminal score of students, also totaling 75 points. This treatment simplifies the data, effectively reducing the assessment components to two main tests. Data cleansing involved removing entries with incomplete data, resulting in a sample size of 202 students, each with scores from the two consolidated tests. Under the assumptions detailed in the methodology section, where the score increment for each evaluation step is considered to be 1, numerical methods are employed to solve the nonlinear equations necessary for calculating the maximum likelihood estimation. This approach sets the initial difficulty parameters of the tests to zero, facilitating the estimation process. The estimations yield the ability parameters, denoted as $(\widehat{\theta_i})$for students indexed from 1 to 202. These parameters provide insights into the relative abilities of students as derived from their performance across the sessional and terminal tests, using the continuous Rasch model framework to interpret the data. As shown in Figure 1, the process underscores the model's capability to handle and analyze educational assessment data, accommodating its complexities and continuous nature.

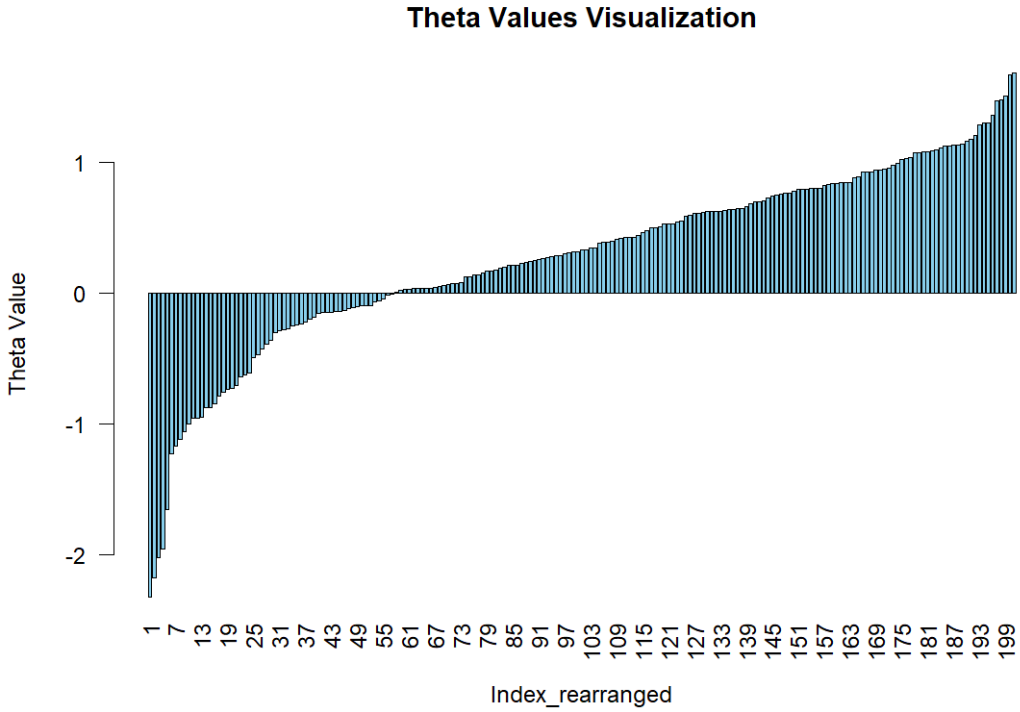

Figure 1: Theta Values Visualization. (Picture credit : Original)

And the difficulty parameter of the two tests are almost the same, As shown in Table 1:

DOI: 10.54254/2753-8818/100/2025.21960

Table 1: The difficulty parameter of the two tests.

$\widehat{\theta_{203}}$ or $\widehat{\beta_1}$	$9.585042763702969482 \times 10^{-2}$
$\widehat{\theta_{204}}$ or $\widehat{\beta_2}$	$9.585042763703258140 \times 10^{-2}$

And plot the difficulty parameters onto the previous graph, using one red line (they are too close):

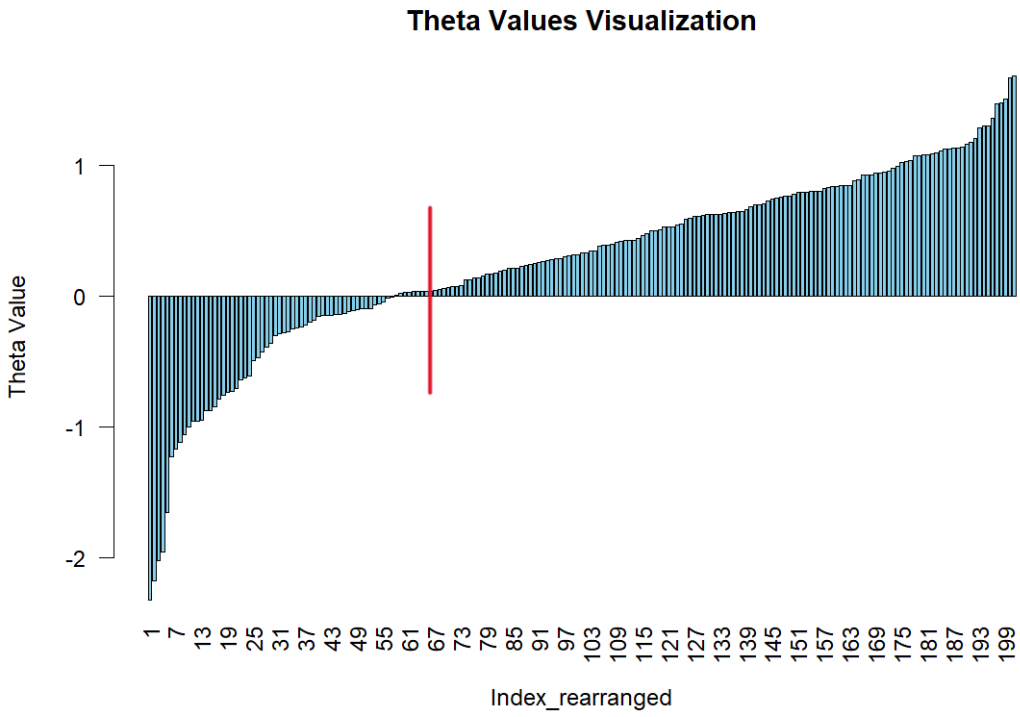

Figure 2: Theta Values Visualization. (Picture credit : Original)

As shown in Figure 2, analyzing the graph that about 33% students' ability parameters are less than the difficulty parameters. As can see in the original data set, excluding the students that are screened out in the first data-cleaning step, there are 80 out of 202 students that failed in this course, and that's is a failing percentage of 40%.

5. Conclusion

This study explores the application of a continuous Rasch model for evaluating students' abilities and the difficulties of exam questions beyond traditional multiple-choice formats. The development and application of this model demonstrate its potential to offer a more nuanced and comprehensive assessment of student performance, effectively accounting for the complexities and variability inherent in continuous scoring systems.

The continuous Rasch model, as introduced in this paper, extends the traditional Rasch model by incorporating characteristics of continuous scoring questions. This allows for a more accurate estimation of student abilities and question difficulties. Maximum likelihood estimation (MLE) is used for parameter estimation, proving to be effective with consistency and asymptotic normality, providing a solid theoretical foundation for the model's reliability and validity.

In real data analysis, the continuous Rasch model was applied to a university students' marks dataset. The analysis reveals that approximately 33% of students had ability parameters lower than

DOI: 10.54254/2753-8818/100/2025.21960

the difficulty parameters of the tests. This finding aligns with an observed failure rate of 40% in the original dataset, suggesting that the model can effectively identify students who may struggle with course material. These results also highlight the importance of considering the difficulty of exams and assignments in evaluating student performance, as traditional methods that rely solely on test scores may not fully capture the underlying abilities and challenges faced by students.

References

[1] *Rasch, G. (1981). Probabilistic Models for Some Intelligence and Attainment Tests. The SAGE Encyclopedia of Research Design.*

[2] *Peng, P., Qu, L., Wang, Q., Wang, S., & Yan, T. (2025). Maximum likelihood estimation in the sparse Rasch model. arXiv preprint arXiv:2501.07770.*

[3] *Pan, T., Shen, W., Davis-Stober, C. P., & Hu, G. (2022). Precision education: A Bayesian nonparametric approach for handling item and examinee heterogeneity in assessment data. arXiv preprint arXiv:2211.11888.*

[4] *KHOSRONEJAD A. (2019). Optimization of the Sefid-Roud Dam desiltation process using a sophisticated one-dimensional numerical model. International Journal of Sediment Research, 24(02), 189-200.*

[5] *Bond, T. G., & Fox, C. M. (2019). Applying the Rasch model: Fundamental measurement in the human sciences. Routledge.*

[6] *Erdos, P.L., & Rényi, A. (1984). On the evolution of random graphs. Transactions of the American Mathematical Society, 286, 257-257.*

[7] *Yang, Y., & Ma, C. (2024). Random pairing MLE for estimation of item parameters in Rasch model. arXiv preprint arXiv:2406.13989.*

[8] *Tarima, S., & Flournoy, N. (2019). Asymptotic properties of maximum likelihood estimators with sample size recalculation. Statistical Papers, 60(2), 373-394.*

[9] *Suryandari, K. C., & Wijayanti, M. D. (2024). Using The Rasch Model to Measure Prospective Elementary Teachers' Perception toward Sciencepreneurship, Creative Thinking, Career Adaptability, and Personal Goals. Jurnal Penelitian Pendidikan IPA, 10(7), 3975-3982.*

[10] *Amin, M. L., Supriyadi, S., & Susilaningsih, E. (2024). Development of Minimum Competency Evaluation Instrument to Analyze Conceptual and Daily Life Abilities Based on The Rasch Model. Journal of Research and Educational Research Evaluation, 13(1), 73-84.*

[11] *Dewanti, S. S., Islamiyah, A. M., Dewi, I. P., & Ashabulabib, A. (2025). Development and Validation of A Self-Confidence Measurement Instrument In Mathematics Learning Using The Rasch Model. Mathline: Jurnal Matematika dan Pendidikan Matematika, 10(1), 55-69.*