

項目反応理論 (Item Response Theory) に関する研究

1. 研究背景

通常行われる一斉テストの得点は、採点者の意図する配点で算出されるため、学生の能力を正しく測定できていない可能性がある。そこで、数理的な根拠に基づいたIRT(項目反応理論)を用いて、学生の能力を正しく測定しようとする試みが考えられている。IRTは受験する場所・時間の制約を受けないことから、コンピュータに組み込んだテスト環境を整えることが検討されている。



3. 研究目的

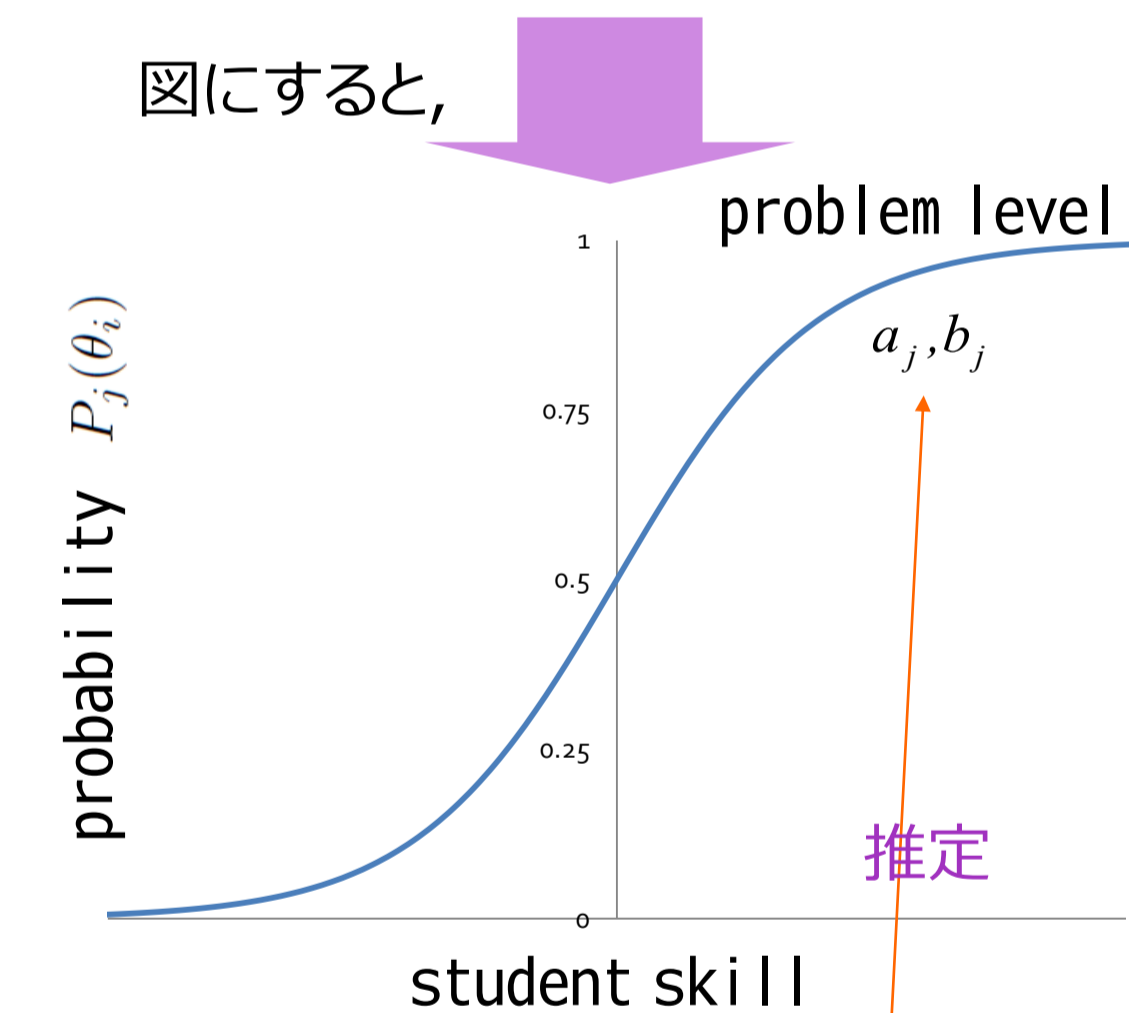
本研究室では、IRT評価システムを e-learning システムへ実装することで、学生の習熟度を正確に把握し、またそれを学生にフィードバックさせることによって学生の学習意欲を向上させるためのシステム構築を図る。また、Stress-Strength モデルと up-down 法(昇降法)を組み合わせた最適テスト法を提案し、少ない問題数で受験者能力を評価する可能性を検証し、より高い精度による能力測定を試みている。

2. IRTって?

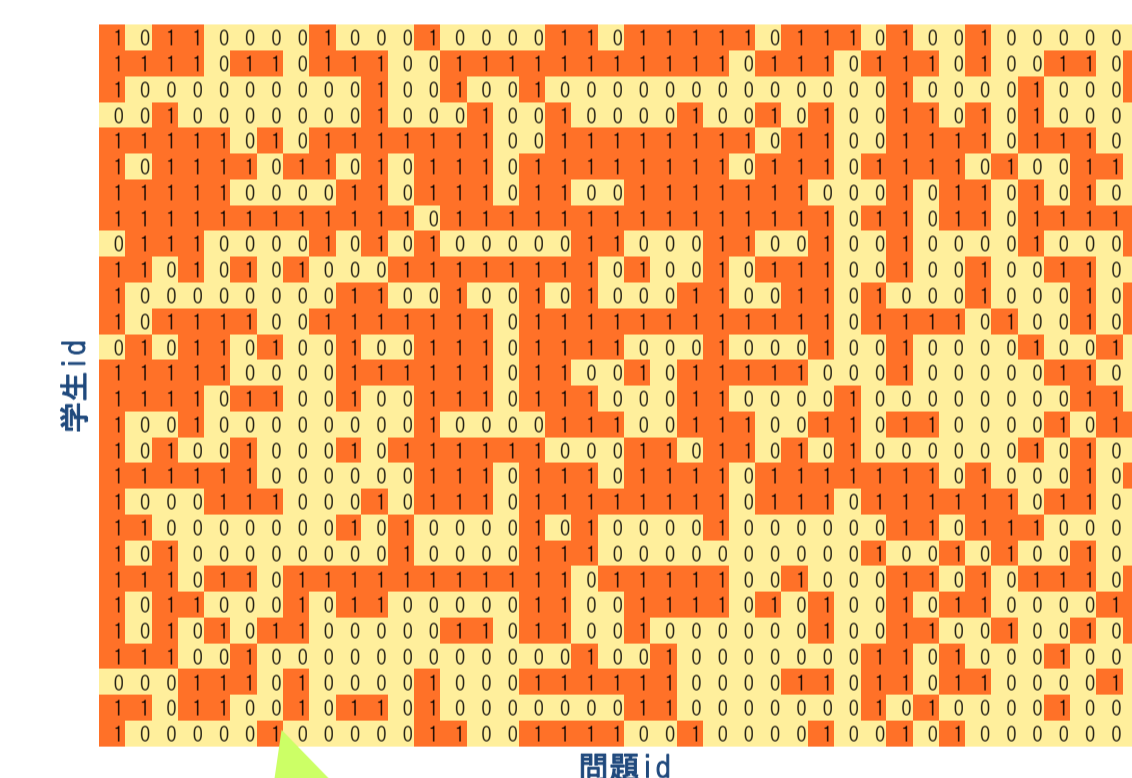
IRT(項目反応理論)は、数理モデルに基づいたテストを作成・実施・評価する為の理論である。実際に、IRTを適用している例として、TOEFLやTOEICなどが挙げられる。

学力 θ_i の学生が問題 j に正解する確率

$$P_j(\theta_i) = \frac{1}{1 + \exp\{-1.7a_j(\theta_i - b_j)\}}$$



試験 試験結果
反応パターン 正答=赤色, 誤答=黄色



拡大

問題①	問題②	問題③	問題④	問題⑤	問題⑥	問題⑦	問題⑧	問題⑨	問題⑩
1	0	0	0	0	0	1	0	0	0

ある学生の試験結果

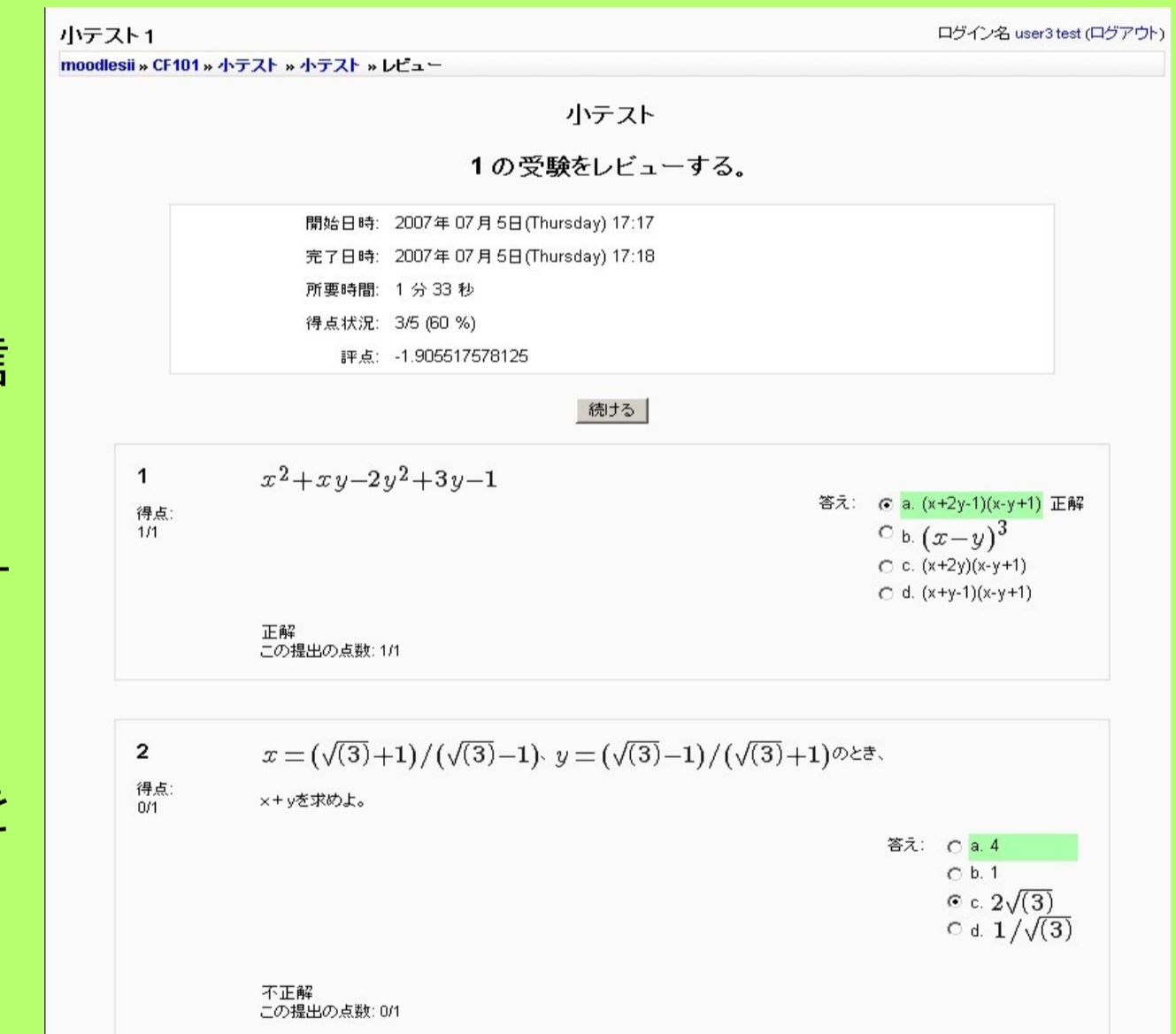
入力データ

$$L = \prod_{i=1}^N \prod_{j=1}^n P_j(\theta_i)^{\delta_{i,j}} (1 - P_j(\theta_i))^{1 - \delta_{i,j}}$$

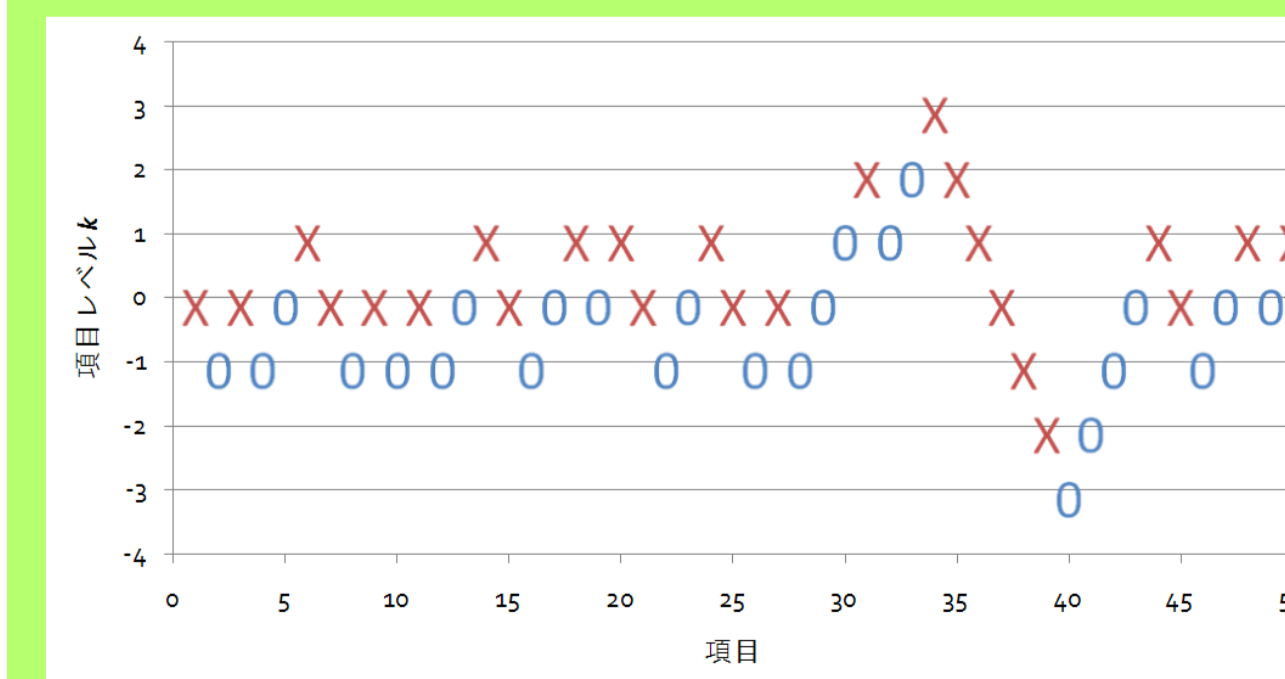
4. e-learningシステムによるテスト環境

IRTによるテスト環境

1. コンピュータのe-learningシステムを利用して、右図のように、問題を出題する。
2. 各受験者は選択式の回答を行い、答案を送信する。
3. テスト結果をIRTによって分析し、能力を測定する。
4. 受験者(もしくはテスト出題者)に能力評価値を送信する。



Stress-Strengthモデルとup-down法による最適テスト法



左図のように、学生の能力に応じたレベルの問題が出題されるようになっており、この間に答えていくことで、受験者の能力を判定することができる。