

潜在ランク理論に基づくコンピュータ適応型テストの 能力推定アルゴリズムの開発と評価

秋山 實^{†1}

本報告では、潜在ランク理論に基づくコンピュータ適応型テストの能力推定アルゴリズムを開発し、評価した結果を報告する。項目応答理論では、最尤推定法(Maximum Likelihood estimation)、重みづけ最尤推定法(Weighted Maximum Likelihood estimation)などが提案されている。潜在ランク理論において同様のアルゴリズムを必要な統計量を定義し、アルゴリズムを提案・評価した。

Development and Evaluation on Ability Estimation Algorithm for Computerized Adaptive Test Based on Latent Rank Theory

MINORU AKIYAMA^{†1}

There are estimation algorithms, for example, Maximum Likelihood estimator and Weighted Maximum Likelihood estimator for Computerized Adaptive Test for Item Response Theory. In this report, author proposed new estimation algorithms, Maximum Likelihood estimator and Weighted Maximum Likelihood estimator for Computerized Adaptive Test based on Latent Rank Theory, and evaluated them by Monte-Carlo simulation and Real-data simulation.

1. はじめに

著者は、潜在ランク理論(Latent Rank Theory, LRT) [1] は、「項目応答理論が適用できない、回答データのアイテム数や受験者数が小さい場合でも適用可能である」[2]ことに着目し、テストの時間を大幅に短縮するコンピュータ適応型テストを実行する Moodle プラグインを開発した[3]。

新しいテスト理論である潜在ランク理論は、項目応答理論(Item Response Theory, IRT)においては、受験者の能力特性やアイテム特性を連続値の間隔尺度で表わすのに対して、ランクという離散値の順序尺度で表わす。これによって能力測定が粗くなるように見えるが、受験者の能力特性を表すランクメンバープロファイル(Rank Membership Profile, RMP)、アイテム特性を表すアイテム参照プロファイル(Item Reference Profile, IRP)はランク数と同じ要素数を持つベクトルで、IRT よりもきめ細かく特性を表現できる。

コンピュータ適応型テスト(Computerized Adaptive Test, CAT)は、テストの精度を維持しながら、短時間にテストを実施する方法の一つである。CATは、従来の、受験者全員が同じアイテムを受験するテスト(以降、リニアテストと呼ぶ)の半分以下のアイテムで同じ程度の精度が得られ、実施時間が約半分に短縮できること、アイテムの露出が少なくなるため、必ずしも一斉受験の必要がないことなど多くのメリットがある。しかし、従来のCATはIRTに基づいているもの(以降、IRT-CATと呼ぶ)がほとんどで、

大きなアイテムバンクを必要としている。LRTに基づくCAT(以降、LRT-CATと呼ぶ)は、小規模なアイテムバンクしかない場合には、テストの時間を短縮するための有効な手段となると思われる。しかし、LRT-CATのアルゴリズムは、木村(2012)が唯一の先行研究であり、研究の余地がある。

2. 目的

本研究の目的は、LRT-CATの新しい能力推定アルゴリズムを提案し、小規模なアイテムバンクを使用する場合の性能を明らかにすることである。

^{†1} 東北大学大学院教育情報学教育部
Tohoku University Graduate School Educational Informatics Education
Division