

透明性の高い達成度テスト運用基盤の開発

岡田 謙介 (東京大学 大学院教育学研究科 准教授)



研究領域「信頼されるAIの基盤技術」
(研究総括: 有村 博紀, 2020年度発足)

研究の概要

社会的役割の大きな達成度テストは、出題する問題項目を秘匿し、繰り返し出題することを基盤として設計され、長年運用されてきた。しかしながら、透明性や情報セキュリティの問題が顕在化しつつある。そこで本研究課題では、基本的に新作問題を出題し、問題項目を逐次公開しながら達成度テストを運用できるようにするための、統計学的方法の開発を行う。これによって社会の公器としての、信頼されるテスト運用を実現することを目指す。

提案研究終了時の達成目標

問題項目を公開する形で達成度テストを運用可能にする方法論の開発。

提案研究の独創性、新規性・優位性

「問題項目を秘匿し、異なる実施回で繰り返し使用する」ことは、項目反応モデルに基づく達成度テスト運用の大前提であった。この体系内で改善を指向する先行研究は少ないが、新作問題を出題しながらの運用を提案した研究は見られない。推定法とテストデザインの開発・工夫によって問題項目を公開しての運用を可能にするアイデアは、統計学理論とテスト運営現場の双方の知見に立脚して生まれた、本提案独自の新規性・優位性である。

提案研究の挑戦性

テストデザインと推定量を一体的に開発し、問題項目公開を実現する本提案研究は、40年に及ぶ達成度テスト運用の常識を覆す大きな挑戦である。

研究の将来展望

(1) 学術研究としての、さきがけ研究成果の将来展開

望ましい統計学的性質を持った項目反応モデルの同時推定法が開発できれば、一般化された潜在変数モデル(GLLAMM等)への拡張が次の理論的展開となる。また実用的な変分ベイズ推定法や自動項目生成法の開発は、心理統計学と機械学習的方法の融合による創発・協働へと繋がる。

(2) さきがけ研究成果と社会との将来の接点(新技術の創出・知的財産権の取得及び活用、又は社会普及・社会受容等)

本研究課題の成果によって問題項目と採点過程がブラックボックスでなくなれば、社会の信頼を得ながら公器としての達成度テストが運営できるようになる。また教育分野はアジアや世界の成長産業であり、開発される技術は公的・民間双方での幅広い活用が見込める。

■ 達成度テストは項目反応モデルに基づいて運用される

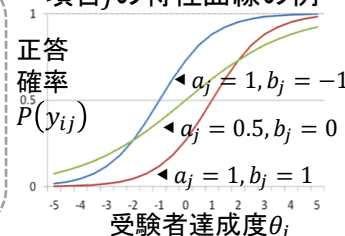


■ その基盤は問題項目の秘匿、繰り返し使用だが...

【標準的な2段階推定法】

1. 予備試験データから
項目パラメータ推定 \hat{a}_j, \hat{b}_j
2. 本番試験データから
受験者パラメータ推定
 $\hat{\theta}_i | a_j, b_j$

項目jの特性曲線の例



- ⚠️ 繰り返し使用するため問題項目が公開できない
- ⚠️ 問題項目の盗難を防ぐことが難しくなっている

■ そこで本研究課題では、

1. 一貫性のある同時推定法
2. 実用的なテストデザイン
3. 自動項目生成と高速化法

を開發

💡 信頼される
テスト運用
の実現

※各種テストのロゴは各社・法人の登録商標等であり、研究提案上の目的で示したものです。