

## 平成29年度 情報工学コース卒業研究報告要旨

関 研究室	氏 名	早 川 今 日 子
卒業研究題目	CBTにおける複数軸の学力推定法の提案と評価	

計算機に基づくテスト法（Computer-Based Testing, CBTと略）とは、答案用紙や筆記用具を使用せず、コンピュータを利用して行うテスト法をいう。CBTの理論の一つとして、項目応答理論（Item Response Theory, IRTと略）が知られている。IRTとは、ある受験者がある特定のテスト項目に应答するプロセスを数学モデルに表したものである。IRTでは、与えられた問題（項目ともいう）に対して、受験者の学力 $\theta$ を引数としたときの正答確率を表す関数を仮定する。この関数はその問題の難易度、識別度、偶然正答率等をパラメータとするロジスティクス関数として与えることが多く、パラメータの個数により、1-PL, 2-PL, 3-PLモデルと呼ばれる。正答確率関数は、試験結果からの受験者の能力の推定に用いられる他に、受験者の現在までの解答状況に応じて次に出题する問題を決定する項目選択法、数多くの問題群から実際に出題する問題集合を決定する項目プール設計、同一の問題をできるだけ繰返し出題しないようにする項目提示制御などを実現するアルゴリズム設計の基礎となっている。これらの中で最も基本的な問題である能力推定では、受験者の能力 $\theta$ を未知とし、受験者の解答の正誤を観測データとしたとき、仮定した正答確率関数のもとで、尤度を最大化する $\theta$ を求める。 $\theta$ の事前確率が一樣ならば、これは事後確率最大化問題と等価となる。

さて、高大接続システム改革会議等における議論では、現在のセンター試験を改革し新しい大学入試方式において、記述式問題を取り入れるとともに、CBTの導入も検討している。またそこでの議論では、学力の3要素である、「知識・技能」、「思考力・判断力・表現力」、「学習意欲」を多面的・総合的に評価することを前提としている。したがって、CBTにおいて、複数の観点から受験者の学力を推定する手法の確立が必要である。

本研究報告ではまず、我が国における入試改革の状況を調査整理し、特に、学力の3要素の測定が重要視されている経緯を述べる。次にこれを受け、IRTに基づいて、受験者に複数軸の能力を仮定したモデルを設定する。さらに、能力推定における提案モデルの有効性を実験的に評価した結果を述べる。具体的に、2種類の能力 $\theta_1, \theta_2$ を仮定し、それらの正答確率関数を2-PLロジスティクス関数とした。そして、2種の能力ごとに正答であったかどうかを観測し、そこから $\theta_1, \theta_2$ の推定を行うという手法を用いた。実験では、それぞれの能力に対する50問の問題の2つのパラメータ及び正答確率関数と、それぞれの能力値を0.25刻みでもつ受験者561人の仮想データを作成し、2種の能力値の推定を行った。比較対象として、2種の能力を区別しないような標準的なIRTを用いた場合の推定実験を行った。比較モデルでは、 $\theta_1$ と $\theta_2$ に対する個別の問題の正誤は観測することができず、 $\theta = \theta_1 + \theta_2$ という単一の能力のみを推定する。問題は、 $\theta_1, \theta_2$ のそれぞれを測定する小問からなり、それらの両方に正解したときかつそのときのみ、その問題に正解するとし、その正誤のみが観測可能であると仮定した。このとき、 $\theta$ の正答確率関数は、 $\theta_1$ と $\theta_2$ の正答確率関数の積となる。これらの前提で実験を行い、 $-4 \leq \theta \leq 4$ の範囲で能力 $\theta$ を変化させ、両モデルにおける推定学力 $\hat{\theta}$ と $\theta$ の誤差の絶対値（ $L_1$ 誤差）を比較したところ、提案モデルでは平均で0.56、比較モデルでは平均で1.26となり、提案手法の基本的な有用性を実験によって確認することができた。