

論文の内容の要旨

論文題目 New parameter estimation procedure for heterogeneous population in Item Response Model

(項目反応理論における、非等質母集団のための新たなパラメータ推定法)

氏名 宮崎 慧

本博士論文の主旨は、近年新しいテスト理論として注目を集めている項目反応理論 (item response theory, IRT) について、非等質母集団のための新たな母数推定法を提案することである。

IRT は心理テストを作成・実施・評価・運用するための実践的な数理モデルであり、既存のテスト理論には無い特長を幾つか持つ。例えば、複数の異なるテスト間の結果を容易に比較し、また項目レベルで受験者能力値測定の精度を確認する、受験者ごとに最適な問題を瞬時に選び、その場で出題するといったことが可能になる (豊田, 2003)。

現在では多様化しているテスト実施場面に対応するために、IRT に関する様々な拡張モデルが提案されている。例えば項目群が 1 次元以上の潜在特性を測定している可能性を考慮した多次元 IRT モデル (multidimensional IRT, Rechase & Mackinly, 1983; Reckase, 1985; Fraser & McDonald, 1986)、受験者が複数の母集団からサンプリングされている状況を想定した多母集団モデル (Bock & Zimowski, 1997)、パラメトリック IRT の当てはまりが悪いときに、項目母数を仮定せず、能力パラメータを和得点などで置き換える nonparametric IRT モデル (NIRT, Mokken, 1971) などがある。

項目反応理論の効果的な運用において重要な事項は、(1) 等化法、及び (2) 項目特性曲線 (ICC: Item Characteristic Curve) の 2 つである。等化とはそれぞれ異なるテストデータから得られた項目パラメータ値を同一尺度上に並べる統計操作のことである。等化により、テストを実施する前に困難度・識別力などの項目特性が既に判明している項目群、即ち「項目プール」

を充実させることができる。項目プールが豊富であるほど、受検者の特性を高い精度で調べるテストを構成することが可能である。そのため、等化は項目反応理論を運用する上で最も重要な統計操作の1つであると言ってよい。

項目特性曲線とは、横軸にそのテストで測ろうとしている能力特性を、縦軸に正答確率を配して描く曲線のことであり、項目の特性（困難度や識別力）は、この項目特性曲線に全て反映される。従って ICC は作成したテストの性質・性能を評価する際の重要な指標となる。また ICC の形状を変えることにより、様々な状況に対応したモデルを構成できる。

さて、心理学を含め行動科学全般においては、観察研究によりデータを取得解析する場面がしばしば存在する。その場合、実験的操作が不可能な状況の下でデータを収集するために、母集団の等質性が保証されないことが多い。項目反応理論を用いる場合においても、非等質母集団から得られたデータを解析せねばならない場面にしばしば遭遇し、その際に先述の等化と項目特性曲線についてそれぞれ問題が浮上する。

等化においては、最も標準的な等化計画として共通項目デザインというものがある。これはテスト実施者がいくつかの共通項目を含んだテストを用意し、各受験者はランダムにそれぞれのテスト形式に振り分けられるというものである (Hanson & Beguin, 2002; Kim & Cohen, 2002)。ところが調査観察研究の状況において共通項目デザインを考えると、母集団が非等質となる原因に、受験者がテスト形式を選択する行動が存在してしまう。実データ解析を例にして説明すると、アメリカにおける AP (Advanced Placement) Examination では同じ能力を測っていると考えられる小論文の題目を受検者が自由選択できるようになっている。他に日本におけるビジネス日本語能力試験などでは、受検する地域によって受けるテスト形式が異なる。日本在住の受検者の方が日本語学習に対する動機付けが高いことを考えると、受検者がテスト形式へ無作為に割り当てられているとは考えにくい。そのような場面に対して、既存の等化法ではテスト形式選択行動を無視するものであり、そのまま適用するとバイアスを含んだ推定値が得られてしまう。

そこで、受験者が受けるテストを自由選択できる際のモデルを構築し、実データとしてビジネス日本語能力試験の回答パターンデータに本手法を適用、モデル選択を行ったところ、既存の方法より開発手法がより正確に項目特性を測定できることが示された (Miyazaki et al., 2009)。本研究により、心理テスト作成場面において受検者の正確な能力推定のために誤った項目を選択する可能性を軽減することに成功した。

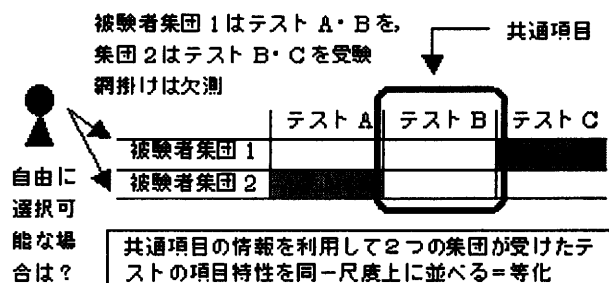


図 1. 共通項目等化デザイン

より、心理テスト作成場面において受検者の正確な能力推定のために誤った項目を選択する可能性を軽減することに成功した。

次に項目特性曲線について、母集団が非等質である場合の問題を述べる。既述の通り、IRT において項目の特性は項目特性曲線で表される。母集団の等質性が保証されない状況では様々な形状の ICC が考えられるが、通常のリジスティックモデルでは表現できる ICC の形状が限られるため、仮定されたモデルがデータに合致しない可能性がある。対処法として項目の特性や受験

者の能力値をパラメータ化しない、項目分析を精緻化した Nonparametric IRT (NIRT)がしばしば適用される。しかし項目反応理論を用いる場面では受験者の回答データから項目の特性と被験者の特性を求めることに興味がある場合が多く、そもそもモデル中に項目パラメータや能力パラメータを設定しない NIRT は、本来の研究関心に合致しない (Mokken, 1971; Junker & Sijtsma, 2001)。

そこで本研究では母集団の非等質性を部分的に表現可能にする混合モデリングと、さらに混合要素数の上限設定を解除できる方法である有限ディリクレ過程事前分布 (Ishwaran & Zarepour, 2000; Ishwaran & James, 2001) という一種のセミパラメトリック法を項目反応モデルに導入することで、ロジスティックモデルで表現しきれない、能力値と正答率間の複雑な関係を、複数の項目特性曲線を混合することにより表現可能にし、母集団の異質性をモデル化でき、能力パラメータをモデル内に導入し、推定する方法を開発した (Miyazaki & Hoshino, 2009)。

以上、本論文では IRT における代表的な事項である等化と項目特性曲線において生じる問題とその解決法を提示した。そしてどちらの研究も根本的には、実験的統制が不可能な状況 (調査・観察状況) において異質な母集団から得られたデータを解析するための手法という意味において共通している。

今後の研究の方向性は、そもそも心理データの解析において、非等質な母集団がなぜ発生するかを考えることで明らかになる。研究デザインを組み立てる際に重要なのは、独立変数の各水準にサンプルを無作為に割り当てることであるが、心理学を含め行動科学では、無作為割り当てを行わない研究、つまり観察研究が行われることが多い。しかし先述のように無作為割当などの実験的操作を取らない場合、データの内容に偏りが生じてしまう可能性があるため、従来の解析法をそのまま適用すると誤った結論を導き出してしまう場合がある (Hill et al., 2002; Condrón, 2008)。観察研究においては母集団の非等質性は不可避であるため、それに対応できるパラメータ推定法の開発は必須であり、事実ここ数十年間で着実に調査観察データの解析法が開発が進んでいる。星野 (2009) は観察研究において使用できる解析法として「共変量の利用」「セミパラメトリック推定」「欠測データモデリング」の三つを挙げている。「共変量の利用」とは従属変数と説明変数の双方に関係する背景情報を積極的に利用することで関心のあるパラメータの推定精度を高める方法である。「セミパラメトリック推定」とはモデル仮定をなるべく置かずに関心のある変数を推定する方法である。「欠測データモデリング」とは調査観察研究においてデータが偏りうる状況を、データ欠測の状況に読み替えることで、仮定すべきモデルを統一的に表現する枠組みのことをいう。

本研究は、上述の方法のうち「欠測データモデリング」と「セミパラメトリック推定」をそれぞれ IRT における等化法と項目特性曲線に適用したものに当たる。等化法では実際には受けていないテスト形式を欠測データと見なすことでテスト形式選択行動をモデル化し、項目特性曲線においてはセミパラメトリック法の一つであるディリクレ過程を用いて関心のあるパラメータを推定する方法を提案した。

上述の調査観察データへの適用法を概観することで、今後の研究の展望が明らかとなる。例

えば等化においてはテスト形式選択行動を説明する共変量(例えば別のテストの得点や知能テスト得点など)をモデル内に組み込めば欠損データを無視可能と見なせるかもしれない。また現実には、複数の受検者集団間に相関構造を仮定する方が自然な場合が多く、従来の推定法では相関構造を誤ってモデル化した場合にロバストでなく、実際のテストデータ解析場面において、能力測定に不適切な項目を選択しかねない。この問題に対し、分布形を仮定しないセミパラメトリック法の一つである M 推定法 (Huber, 1964) を用いることにより、モデルの誤設定に対しロバストな、バイアスを最小限に抑えることのできる推定法を開発できる。あるいはディリクレ過程についても、これを誤差変数に適用し、能力パラメータだけでなく項目パラメータを関心母数と見なし推定する方法も考えられる。