

## 研究委員会企画チュートリアルセミナー

### 項目反応理論の考え方と実践

#### ー測定の高い質のテストや尺度を作成するための技術ー

##### 【企画趣旨】

言語能力試験など、様々な大規模能力試験の作成・運用においてグローバル・スタンダードとなっている方法論に項目反応（応答）理論（Item Response Theory : IRT）がある。IRT は、ある受験者がある項目（問題）に正答する確率を、項目反応モデルと呼ばれる統計モデルで表現する。これにより、各受験者がそのときに受けた試験問題の難易度や、受験者集団の能力分布の違いに極力左右されないという意味で公平な評価が可能になる。また、IRT はいわゆるテスト（試験）だけでなく、測定の性能評価や効率化等を目的として、（教育）心理学研究における心理尺度の作成・改良のためにも既に広く活用されている。加えて、昨今の高大接続改革において導入が検討されている英語の民間試験にも関わる話題として、異なるテストに基づく得点（スコア）を同一尺度上で比較可能にするための等化(equating) またはリンキング(linking)と呼ばれる方法があるが、これも IRT のもとで体系化された技術である。

本チュートリアルの目的は、IRT について初めて学ぶ方、またテスト（心理尺度）を作成・利用する立場にある研究者や教育関係者の方などを対象に、IRT の基本的な考え方とともにその適用上の利点や注意点、および実際の心理学研究や実践例を基にした統計ソフト R の実行方法について平易に解説することである。具体的に、本チュートリアルでは、IRT の利点についてまず俯瞰した後に、幾つかの基本的な項目反応モデルを紹介する。そして、項目反応モデルを用いた研究・実践に密接に関わっている等化について、その基本的な考え方や方法を解説し、そして統計ソフト R を用いた IRT の実行例を幾つか紹介する。統計ソフト R については、その操作方法についての解説も本チュートリアル内で行うが、例えば回帰分析の実行などのごく基本的な操作経験は既にありますが、本チュートリアルの内容をスムーズに理解する上で望ましい。また、本チュートリアルで使用するデータ等は事前にアップロードし、当日参加者が自身の PC で同じ分析を実行できるようにする予定であるため、R を予めインストールした PC を持参することを推奨する。