

# IMPS 2013 参加報告

ベネッセ教育総合研究所 高等教育研究室

野澤 雄樹・堀 一輝

Psychometric Society の第 87 回年次大会である International Meeting of Psychometric Society 2013 (IMPS 2013) は、オランダのアーネムで、7 月 22 日から 26 日までの日程で開催された。アーネムはアムステルダムから鉄道で 1 時間半ほど南東に移動した場所にある。オランダは夏でもそれほど暑くならないということであったが、大会期間中に熱波が到来したため、連日 30°C を超える暑さとなった。

## 23 日の報告

開会式の直後に行われた Wim van der Linden 氏 (CTB/McGraw-Hill) の基調講演では、多次元項目反応理論 (multidimensional item response theory; 多次元 IRT) において単調性と補償性が等価であることを示した上で、補償型モデル (compensatory model) と非補償型モデルの 1 つである連結モデル (conjunctive model) の比較、連結モデルの構成、その他の関連するトピックについての議論が行われた。

基調講演の後、一般セッション、招待講演、ポスター発表に参加した。その中で特に印象に残ったのは、多次元 IRT に関するセッションである。不完全なデータ収集デザインが母集団パラメータの推定に与える影響を調べた研究や、多次元の高次 IRT モデルを多母集団に拡張する試みなど、多次元 IRT を実際に使用するという段階を超えて、状況に応じたさまざまな応用・拡張が行われており、これらの研究を追っていく必要性が感じられた。

また、この日は 2015 年の PISA で導入される協調的課題解決 (Collaborative Problem Solving; CPS) に関するセッションも開かれた。実際に開発を行っている Educational Testing Service の Alina A. von Davier 氏からは測定アプローチと測定に関する諸問題について、マサチューセッツ工科大学の Yoav Bergner 氏からは Massive open online courses (MOOCs) の掲示板から得られたログデータの解析結果について報告があった。これからの社会を生き抜くうえで CPS が重要であるという認識は日本でも広まりつつあり、その測定方法の開発・整備は我々の課題の 1 つである。CPS の測定に関する研究には今後も注目していきたい。

## 24 日の報告

24 日に参加したセッションの中では、多段階テスト (multistage testing; MST) に関する招待講演が興味深かった。MST は 2011 年にリリースされた新版の GRE で採用されたことから注目が高まったテスト方式で、項目レベルのコンピュータ適応型テスト (item-level computerized adaptive testing; item-level CAT) と比べて効率は落ちるものの、ステージ内でなら前の項目に戻って自由に解答を変更できるので、受験者のストレスを軽減できると考えられている。また、作成されたテストを専門家が事前にレビューできるため、テストの品質管理の面で望ましいとされている。この招待セッションでは、MST の一般的な解説のほか、MST におけるテストの自動作成、テスト情報の保管理、ルーティングの方法、信頼性推定方法など、最新の研究動向が紹介された。

## 25 日の報告

Psychometric Society の現会長である Hua-Hua Chang 氏 (University of Illinois, Champaign) の講演では、自身が行った item-level CAT の研究の紹介のほか、MST や認知診断に関連した話題も出された。一般セッションでは、因子分析、多次元 IRT、IRT モデリング、反応データと反応時間のモデリングに関する各セッションに参加した。多次元 IRT のセッションでは、多次元 IRT のパラドックス (受験者がある項目に正答することで、能力の最尤推定値が減少してしまう次元が存在する可能性がある) に関する Mark Reckase 氏 (Michigan State University) の見解が述べられた。

## 26 日の報告

26 日のセッションの中では強制選択データ (forced choice data) に関するセッションが小規模ながら非常に興味を引いた。強制選択データとは、その名の通り複数の異なる性格特性などを表す選択肢から、回答者に自分に合っていると思うものを強制的に選択させて得られるデータのことである。いわゆる一対比較法は強制選択法の 1 つの特殊ケースであり、性格検査などで回答者が自分をよく見せようとする行動 (フェイキング) への対処法として広く知られている。しかし強制選択データは各個人の性格特性のバランスを見るのには適しているものの、合計得点が回答者間で同じになるため (このようなデータをイプサティブデータという)、回答者間の違いを見る (回答者を弁別する) のには適していない。このようなデータの特徴のためこれまであまり研究が進んでいなかったが、ここ数年 Thurstonian IRT と呼ばれるモデルを強制選択データに適用することで前述の問題の解決を図る研究が増えており、今回の IMPS ではセッションが 1 つできるまでになっている。フェイキングへの対処法として有効であると考えられるため、企業を巻き込んで今後ますます研究が加速する分野であると思われる。

## まとめ

今回の IMPS はテストの開発・運用会社である Cito がホストであったこともあり、例年以上にテスト・アセスメント関係の発表が多かったように思われる。その中でも多次元 IRT の話題が特に多かったのではないだろうか。多次元 IRT 自体は 1980 年半ばに提唱されていたが、最近になって市販のコンピュータプログラムでも信頼度の高いパラメータ推定ができるようになり、実践的・応用的な研究が増えてきたという印象である。IMPS2014 はアメリカ・ウィスコンシン州で開催予定である。