

# 等化手続きの違いが項目バンクの パラメタ推定値に及ぼす影響 —等化係数を用いた等化法を同時推定法に併用する効果の実践的検討—

光永悠彦 (島根大学)

キーワード：項目反応理論，等化，項目バンク

## 背景と目的

項目反応理論 (IRT) を用いた等化 (equating) を行い、項目バンクを整備するテストに対する需要が高まりつつある。本試験受験者集団に新作項目 (ユニーク項目) を混ぜて出題し、項目パラメタを推定することで、項目バンクのサイズを増やす方法がしばしば用いられる。その際、本試験の新作項目の項目特性値 (項目パラメタ) は、規準集団を定義するフィールドテストのパラメタに等化される。本研究ではフィールドテストと本試験のデータを一括して多母集団 IRT モデルにより推定する同時推定法において、等化係数を使った等化法 (Arai and Mayekawa, 2011) を併用する手順を用いた場合に推定精度がどう影響を受けるかを検討する。また、テストごとに個別に項目パラメタを推定し、等化係数による等化を行った場合 (個別推定) と結果を比較する。

## シミュレーション方法

**想定したテスト** フィールドテストに続き 5 回の本試験を行った場面を想定した。本試験ではフィールドテストと共通の項目 (アンカー項目) を出題することで、等化を行うと想定した。

**項目パラメタの条件** アンカー項目数、困難度平均と受験者  $\theta$  の乖離度、本試験とフィールドテストとの識別力の乖離度について、それぞれ 3 条件を設定し、条件ごとに以下の手続きを行った。

**手続き** 項目パラメタ及び受験者 (各テスト 1000 人を想定) に真値を乱数により発生させ、それから 2 パラメタ・ロジスティックモデルを仮定し 0-1 データを生成した。次いで、個別推定、同時推定 (等化係数による等化を併用: cu 及び併用せず: curef) の方法で等化済み項目パラメタ推定を行った。同時推定については、5 回の本試験のうち最初の 3 回、4 回分のデータのみ用いた場合についても推定した。以上の手続きを 100 回繰り返した。

**評価指標** DICC (Arai and Mayekawa, 2011) を用

いて真値とのズレを検討した。

## 結果と考察

アンカー項目数：新作項目数 = 1 : 4、困難度平均が  $\theta$  平均と一致、本試験の識別力がフィールドテストより低い場合の結果を Figure 1 に示した。

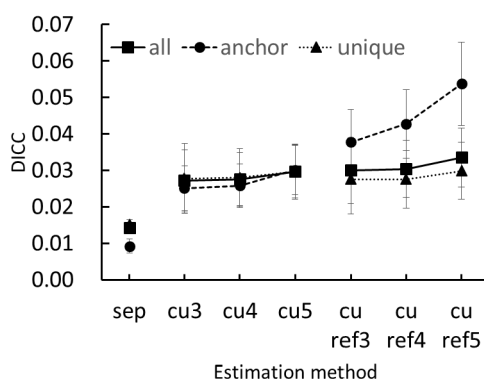


Figure 1: 推定方法の違いによる DICC。

同時推定では curef の方法による推定でアンカー項目の推定値が真値とかけ離れる傾向がみられた。またズレの大きさは本試験実施回が増えるごとに増大する傾向が見られた。テスト実践場面で同時推定を行う場合は、等化係数を用いた等化法を併用したほうがより好ましいことが示唆される。

## 参考文献

Arai, S & Mayekawa, S. (2011). A comparison of equating methods and linking designs for developing an item pool under item response theory. *Behaviormetrika*, **38**, 1-16.

光永悠彦 (2012). 共通項目デザインを用いた大規模テストにおける等化順序の効果の検討  
日本教育心理学会第 54 回大会発表論文集

## 謝 辞

本研究は科学研究費補助金 (若手研究 (B)), 課題番号 16K17312) の助成を受けて行われた。