

ニューラルテスト理論における DIF 指標の提案

○熊谷 龍一*

*新潟大学全学教育機構

1. はじめに

本発表では、ニューラルテスト理論における DIF 検出の方法を提案する。DIF 検出方法としては、これまで IRT の枠組みで提案されてきた熊谷・山口・小林 (2005) および熊谷 (2007) を、ニューラルテスト理論に応用することとする。

2. DIF 指標の計算

今回提案する方法は、以下の手続きにより DIF 指標が計算される。

1) データ行列の分解

Thissen, Steinberg, & Wainer(1993)による尤度比を用いた方法と同様に、データ行列において DIF を検討すべき項目を下位集団ごとに分解し (図 1)、個別の項目として扱う。

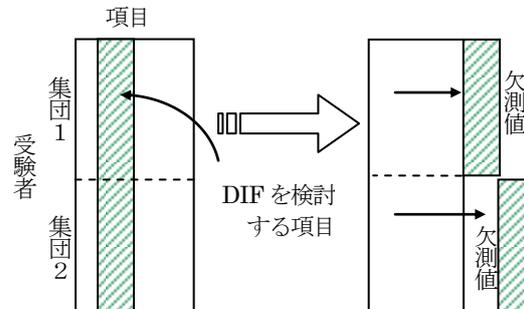


図 1 データ行列の分解

2) 項目参照プロファイルの推定

1) で作成したデータ行列において、項目参照プロファイルの推定を行う。この時、DIF を検討する項目については、集団 1 と集団 2 の部分として、2 つの項目参照プロファイルが推定されることとなる。

3) DIF 指標の計算

2) で推定された DIF を検討する項目に対する項目参照プロファイル (IRP) および、潜在ランク分布 (latent rank distribution, LRD) もしくはランク・メンバーシップ分布 (rank membership distribution, RMD) を用いて、以下の式により、DIF 指標 K を算出する。

$$K = \sum_{q=1}^Q |P_1(q) - P_2(q)| \cdot g_q \quad (1)$$

ここで Q は潜在ランク数、 $P_1(q)$ 、 $P_2(q)$ はそれぞれ集団 1 および集団 2 として推定した項目参照プロファイルでの潜在ランク q における項目正答確率、 g_q は潜在ランク q における相対 LRD もしくは相対 RMD の値とする。

3. 指標について

(1) 式で示された DIF 指標 K は、2つの IRP から得られる、各潜在ランクにおける正答確率の差 (図 1 の上段参照) について、母集団全体 (図 1 の下段) での期待値を算出していることになる。

これにより、指標 K は、0 から 1 の数値で表されることになり、正答確率の差として理解が容易になる。

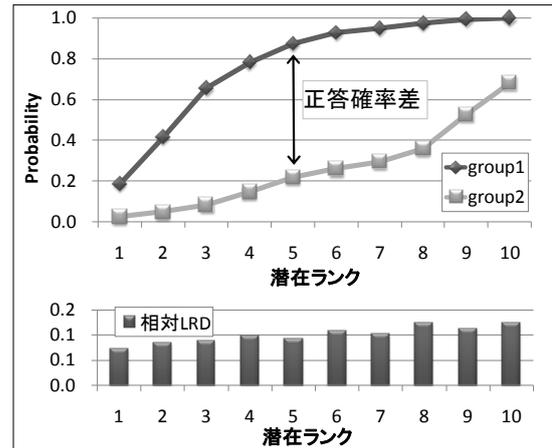


図 1 DIF 指標の意味

4. 指標 K についての検討

指標 K の性能を検討するために、2母数項目反応モデルと一様乱数を利用して、DIF シミュレーションデータ*を作成し、Mantel-Haenszel 法 (以下 M-H 法) との比較を行った。

指標 K と M-H 法における Δ_{MH} との相関は .92 となり、非常に高いことから、指標 K も M-H 法と同程度の性能があることが示された。

また、指標 K と Δ_{MH} との散布図を図 2 に示す。Roussos and Stout (1996) による Large or C-level DIF の基準である $\Delta_{MH} \geq 1.5$ に対して、指標 K では、0.1 を超える部分において、ほぼ対応していることが示された。

*15 項目を DIF が無い項目、50 項目を DIF が生じている項目としたデータセットを 2 セット作成

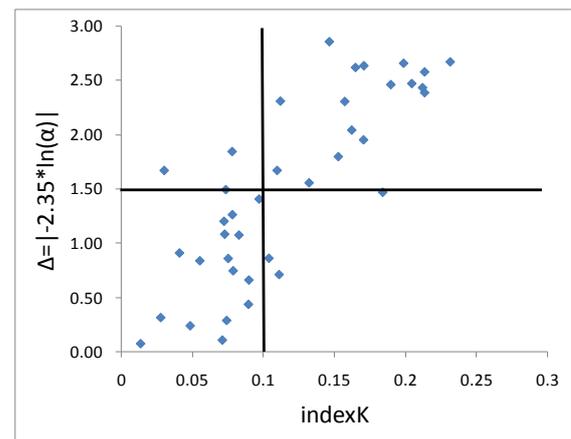


図 2 指標 K と Δ_{MH} の散布図

文 献

熊谷龍一・山口大輔・小林万里子 2005 母集団分布を考慮した項目反応モデルによる DIF 検出法の提案 日本心理学会第 69 回大会発表論文集, 414

熊谷龍一 2007 母集団分布情報を加味した 3 母集団以上の DIF 検出方法の提案 日本テスト学会第 5 回大会発表論文抄録集, 232-233.

Roussos, L. A., & Stout, W. F. (1996). Simulation studies of the effects of small sample size and studied item parameters on SIBTEST and Mantel-Haenszel type I error performance. *Journal of Educational Measurement*, 33, 215-230.

Thissen, D., Steinberg, L., & Wainer, H. 1993 "Detection of differential item functioning using the parameters of item response models." in P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 67-113). Hillsdale, NJ: Lawrence Erlbaum.