多値型 IRT の近似ベイズ推定量を用いた項目等化について

森一将　大森拓哉　新谷真由
東京大学大学院　多摩大学　筑波大学

1　概要とモデル

本研究では多値型 IRT モデルの 1 つである Rating Scale Model(RSM;Rasch, 1960; Andrich,1978a, 1978b; Andersen,1996) の母数の条件付き尤度を基にしたベイズ推定量 (予測量）を近似的に導出し、この推定量を基にした等化法を提案する. k 個の項目がそれぞれ m カテゴリを持ち,n 人の受験者から成るテストデータを考える. このとき,i 番目の受験者が j 番目の項目の h 番目のカテゴリに対する反応を $X_{ijh} = \{0,1\}(i = 1,\cdots,n, j = 1,\cdots,k, h = \cdots,m)$ とする. このとき,RSM によって対応する確率は以下のようになる.

$$P_{ijh}(\theta_i,\alpha_j) = P(X_{ijh} = 1;\boldsymbol{\theta},\boldsymbol{\alpha}) = \frac{\exp(w_h\theta_i + \alpha_{jh})}{\sum_{h=1}^{m}\exp(w_h\theta_i + \alpha_{jh})}.$$

ここで,$\boldsymbol{\theta} = (\theta_1,\cdots,\theta_n)$ は受験者の能力母数を示し,$\boldsymbol{\alpha} = (\alpha_{11},\cdots,\alpha_{1m},\cdots,\alpha_{k1},\cdots,\alpha_{km})$ はテストの項目母数を示す. また,$w_h(h = 1,\cdots,m)$ は重みづけ係数 (配点）であり, 所与とする. この式を基にして項目母数の条件付き最尤推定値を考える. この時の結合尤度関数は以下のようになる.

$$L(\boldsymbol{\theta},\boldsymbol{\alpha}|X) = \frac{\exp\left(\sum_{i=1}^{n}\theta_i t_i + \sum_{j=1}^{k}\sum_{h=1}^{m}\alpha_{jh}r_{jh}\right)}{.}C(\boldsymbol{\theta},\boldsymbol{\alpha}).$$

2　尤度関数の近似とベイズ推定量

ここで,$t_i = \sum_{j=1}^{k}\sum_{h=1}^{m} w_h x_{ijh}$ は i 番目の受験者の得点を示し,$r_{jh} = \sum_{i=1}^{n} x_{ijh}$ は j 番目の項目の h 番目のカテゴリに反応した受験者数を示す. また,$C(\boldsymbol{\theta},\boldsymbol{\alpha}) = \prod_{i=1}^{n}\prod_{j=1}^{k}\sum_{h=1}^{m}\exp(w_h\theta_i + \alpha_{jh})$ とする. この尤度関数の一部に以下のような近似を考える.

$$C(\boldsymbol{\theta},\boldsymbol{\alpha}) \simeq m^{nk}\prod_{i=1}^{n}\prod_{j=1}^{k}\sqrt[m]{\prod_{h=1}^{m}\exp(w_h\theta_i + \alpha_{jh})} = \tilde{C}(\boldsymbol{\theta},\boldsymbol{\alpha}).$$

すると, 尤度関数は以下のようになる.

$$L(\boldsymbol{\theta},\boldsymbol{\alpha}|X) = \frac{\exp(\sum_{i=1}^{n}\theta_i t_i + \sum_{j=1}^{k}\sum_{h=1}^{m}\alpha_{jh}r_{jh})}{\tilde{C}(\boldsymbol{\theta},\boldsymbol{\alpha})} = \tilde{L}(\boldsymbol{\theta},\boldsymbol{\alpha}|X).$$

この近似は「相加平均 $\geq$ 相乗平均」の関係を応用したものであり,Spect(1960) によって近似をした場合の比の上限が明らかになっている. つまり, この近似による条件付き尤度関数の近似誤差は一定の範囲にとどまる. したがって, この近似を用いた条件付き尤度関数について, ベイズ推定量を考える. ベイ

ズ推定量を導出するにあたって必要な損失関数，および事前分布は以下のとおりとする。

$$\mathcal{L} = (\vartheta - \delta)^2,\ \theta_i \sim N(\mu, \sigma^2),\ \alpha_{jh} \sim N(\tau, \phi^2)$$

ここで、ϑ,δ はそれぞれ推定対象の母数と対応するベイズ推定量,N は正規分布、μ, σ, τ, ϕ はそれぞれ所与の母数を示す。これらを用いると, 母数 θ_i,α_{jh} に対応するベイズ推定量は以下の通りになる。

$$\tilde{\theta}_i = \mu - \frac{1}{2\sigma^2}\left(t_i - \frac{kw}{m}\right),\ \tilde{\alpha}_{jh} = \tau - \frac{1}{2\phi^2}\left(r_{jh} - \frac{n}{m}\right).$$

3　等化法の提案

この推定量を用いた１種の等化について考える. 上式を見ればわかる通り近似尤度を用いた場合, ベイズ推定量 $\tilde{\theta}_i$,$\tilde{\alpha}_{jh}$ はそれぞれ項目反応 X より計算された統計量 t_i,r_{jh} の線形推定量になっている. ここで, 古典的テスト理論の考え方を用いると,t_i,r_{jh} は誤差要因を含み, それぞれの「真値」に対する精度は, 例えばクロンバックの α 係数などによって評価される. したがって，この考え方を応用して，等化に用いる推定量の数があらかじめ決まっているとき，対応する統計量の精度がもっとも高くなる組み合わせを選択し，これを用いて等化を行えばよい.

等化法の詳細および数値例は当日示す.

参考文献

[1] Andrich, D. (1978a). A rating formulation for ordered response categories. *Psychometrika*, **43**, 561-573.

[2] Andrich, D. (1978b). Application of a psychometric rating model to ordered categories which are scored with successive integers. *Applied psychological measurement*, **2**, 581-594.

[3] Andersen, E.B. (1996). The rating scale model. In van der Linden, W.J. and Hambleton, R.K. (Eds.) *Handbook of modern item response theory*. 67-84. Springer.

[4] Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Copenhagen, Denmark: Danish Institute for Educational Research.

[5] Specht, W. (1960). Zur theorie der elementaren Mittel. *Math. Z.*, **74**,91-98.