「テスト情報関数を利用した潜在特性尺度の同時等化法」

豊　田　秀　樹

A Simultaneous Equating Method by Using Test Informations

Hideki TOYODA

An equating method for two latent ability scales by using common subjects and test informations was proposed in the previous paper (Toyoda, 1986). The purpose of this paper is to expand the method and to propose a simultaneous equating method for two or more latent ability scales. An advantage of this method is that treatment of missing data is very easy.

I　目的

尺度の等化とは異なった項目で構成された複数のテストの結果を直接比較可能なものにするために同一の尺度上の値に変換して表す操作のことをいう。等化は，

(1)　テストを構成する項目の困難度レベルが異なり，適切に能力を測定することが可能な被験者の能力レベルがそれぞれ異なるテストに行う垂直的等化（vertical equating），

(2)　テストを構成する項目の困難度レベルがほとんど同じで，適切に能力を測定することが可能な被験者の能力レベルも同じテストに行う水平的等化（horizontal equating）

の２種類に大別される。垂直的等化を行うことにより，例えば発達データから構成された異なる年齢向けの複数のテストの結果が比較可能になるので，個人の発道的変化を単一の尺度上で表現することができる。また水平的等化を行うことにより，例えば本試験と追試験の結果が比較可能になるので，どちらの試験を受験したかに影響されない公平な合否の判定に行うことができる。

本論文の目的は，共通被験者の推定尺度値とテスト情報関数を利用した２つの潜在特性尺度の等化法（豊田，1986）を多数の潜在特性尺度の同時等化法に拡張することである。

１つの等化される潜在特性尺度 So と m 個の等化する潜在特性尺度 S_j $(j=1,\cdots,m)$ が存在し，尺度 S_0 における被験者 i の尺度値を θ_{io} とし，尺度 S_j におけるそれを θ_{ij} とすると θ_{io} と θ_{ij} の間には

$$\theta_{io}=k_j\theta_{ij}+l_j \tag{1}$$

なる関係がある（Lord & Novick, 1968)。k_j, l_j は潜在特性尺度 S_j を S_o に等化するための等化係数（Lore, 1980）といわれる。m 個の潜在特性尺度を等化するためには $2\times m$ 個の等化係数を推定すればよい。

II　共通被験者の尺度値の推定

本等化法を実施するためには以下の手順でデータを集収し，共通被験者の尺度値を推定する。

[i]共通被験者をN人用意する。共通被験者を最低限何人集めれば充分であるか，という問題は等化する（される）テストの項目の数やパラメータの分布，共通被験者の能力尺度値の分布によって異なる。しかし $m=1$, テストの項目数が25，35であり，それ等のテストで適切に能力が推定され得る被験者が用いられる状況であるならば，N＝50という比較的少数の被験者で，等化係数の推定誤差分散が実際場面で使することが可能な程度に小さくなることがシミュレーション（豊田，1986）で示されている。

[ii]各尺度から n_j 個 $(j=0,\cdots m)$ の項目を選び $m+1$個のテストを構成する。n_0 は等化されるテストの項目数である。本等化法は等化係数の推定の際に項目数が充分に大きい場合の共通被験者の推定尺度値の分布を利用するので項目数 n_j は大きいほうが望ましい。

具体的には後述する理由により $m+1$ 個のテストの項目数がそれぞれ全て20以上であることが１つの目安である。

[iii] N 人の共通被験者に $m+1$ 個のテストを実施す

る。共通被験者は全てのテストを受験することが望ましいが，(1)垂直的等化を行う場合など，全てのテストを全ての被験者に実施することが実際的でない状況，(2)水平的等化を行う場合であっても時間的制約などから全てのテストを必ずしも全ての被験者に実施できない状況などで生じる欠測値は存在してもさしつかえない。欠測値の処理については後述する。

[iv] 共通被験者 i のテスト j に対する 0-1 反応パタンベクトル $\boldsymbol{u}_{ij}$ による θ_{ij} の最尤推定値 $\hat{\theta}_{ij}$ を求める (Lord, 1980)。$\hat{\theta}_{ij}$ は共通被験者 i の尺度 S_j における推定尺度値である。

Ⅲ 最尤推定法

A 尤度関数の構成

θ_{ij} の最尤推定量 $\hat{\theta}_{ij}$ は漸近的に平均 θ_{ij}，分散 $I(\theta_{ij})^{-1}$ の正規分布に従う (Hambleton & Swaminathan, 1985)。ただし

$$I(\theta_{ij})=\sum_{h=1}^{n_j}\left[(P_h(\theta_{ij})Q_h(\theta_{ij}))^{-1}\left\{\frac{\partial}{\partial\theta_{ij}}P_h(\theta_{ij})\right\}^2\right] \tag{2}$$

はテスト情報関数と呼ばれ，潜在特性モデルにおいて尺度値を独立変数としてテストの精度を表し，フィッシャー情報量に一致する関数である。ただし $P_h(\theta_{ij})$ は被験者 i がテスト j の h 番目の項目に正答する確率を表す項目特性関数 (Lord, 1980) であり，2パラメータロジスティックモデルの場合は

$$P_h(\theta_{ij})=\frac{1}{1+\exp(-Da_h(\theta_{ij}-b_h))} \tag{3}$$

と表され，3パラメータロジスティックモデルの場合は

$$P_h(\theta_{ij})=c_h+\frac{(1-c_h)}{1+\exp(-Da_h(\theta_{ij}-b_h))} \tag{4}$$

と表される。ここで a_h, b_h, c_h はそれぞれ項目 h の識別力パラメータ，困難度パラメータ，偶然正答確率のパラメータといい，$D=1.7$ の定数である。また

$$Q_h(\theta_{ij})=1-P_h(\theta_{ij}) \tag{5}$$

である。$I(\theta_{ij})$，$P_h(\theta_{ij})$，$Q_h(\theta_{ij})$ は厳密にはそれぞれ $I_j(\theta_{ij})$，$P_{jh}(\theta_{ij})$，$Q_{jh}(\theta_{ij})$ と表すべきであるが，添字 i が重複するので消略する。

ここで項目数が比較的少数の場合にも $\hat{\theta}_{ij}$ が平均 θ_{ij}・分散 $I(\theta_{ij})^{-1}$ の正規分布に従うと仮定すると，θ_{ij} が与えられたときの $\hat{\theta}_{ij}$ の確率密度は

$$f(\hat{\theta}_{ij}\mid\theta_{ij})=\sqrt{I(\theta_{ij})/2\pi}\times\exp\left\{\frac{-1}{2}(\hat{\theta}_{ij}-\theta_{ij})^2I(\theta_{ij})\right\} \tag{6}$$

と表される。この近似は $n_j\geqq20$ のときに十分よいものであることが報告されている (Samejima, 1977)。ここで(1)式より

$$\theta_{ij}=\frac{\theta_{i0}-l_j}{k_j} \tag{7}$$

の関係があるから θ_{i0}, k_j, l_j が与えられたときの $\hat{\theta}_{ij}$ の確率密度は

$$f(\hat{\theta}_{ij}\mid\theta_{i0},k_j,l_j)=\sqrt{I(\theta_{ij})/2\pi}\times\exp\left\{\frac{-1}{2}(\hat{\theta}_{ij}-\theta_{ij})^2\ I(\theta_{ij})\right\} \tag{8}$$

と表される。$j=0$ の場合は，定数 $k_0=1$，$l_0=0$ とすることによって(8)式は $j=0$ の場合をも含んだ式となる。潜在特性モデルの仮定より被験者 i についての $m+1$ 個の推定尺度値の分布はそれぞれ独立であるから，$\theta_{i0}, K=(k_0=1, k_1, k_2, \cdots k_m)$，$L=(l_0=0, l_1, l_2, \cdots l_m)$ が与えられたときの $\hat{\boldsymbol{\theta}}_i=(\hat{\theta}_{i0}, \hat{\theta}_{i1}, \cdots l_m)$の同時分布は

$$f(\hat{\boldsymbol{\theta}}_i\mid\theta_{i0},K,L)=\prod_{j=0}^{m}f(\hat{\theta}_{ij}\mid\theta_{i0},k_j,l_j) \tag{9}$$

と表される。同様に，N 人の被験者の $\boldsymbol{\theta}_i$ の分布もそれぞれ独立であるから，$\hat{\boldsymbol{\Theta}}=(\hat{\boldsymbol{\theta}}_1, \cdots\hat{\boldsymbol{\theta}}_N)$ が与えられたときの $\boldsymbol{\theta}_0=(\theta_{10}, \theta_{20}, \cdots\theta_{N0})$，$K, L$ の尤度関数は

$$L(\hat{\Theta}\mid\boldsymbol{\theta}_0,K,L)=\prod_{i=1}^{N}f(\hat{\boldsymbol{\theta}}_i\mid\theta_{i0},K,L) \tag{10}$$

と表される。(10)式を $\boldsymbol{\theta}_0, K, L$ について最大化することにより等化係数の最尤推定値を得る。

(8)式の対数尤度関数は，定数部分を無視すると

$$\log L=\sum_{i=1}^{N}\sum_{j=0}^{m}G_{ij} \tag{11}$$

$$G_{ij}=\log I(\theta_{ij})-(\hat{\theta}_{ij}-\theta_{ij})^2\ I(\theta_{ij}) \tag{12}$$

と表される。

B 偏導関数

最尤推定を行うために必要となる k_j, l_j, θ_{i0} による対数尤度関数の偏導関数をそれぞれ(13)，(14)，(15)式に示す。

$$\frac{\partial}{\partial k_j}\log L=\sum_{i=1}^{N}\theta_{ij}\ g_{ij} \tag{13}$$

$$\frac{\partial}{\partial l_j}\log L=\sum_{i=1}^{N}g_{ij} \tag{14}$$

$$\frac{\partial}{\partial\theta_{i0}}\log L=-1\times\sum_{j=0}^{m}g_{ij} \tag{15}$$

ただし

$$g_{ij}=\frac{1}{k_j}\left\{((\hat{\theta}_{ij}-\theta_{ij})^2-I(\theta_{ij})^{-1})\right.$$

$$\times \frac{\partial}{\partial \theta_{ij}} I(\theta_{ij}) - 2(\hat{\theta}_{ij} - \theta_{ij}) I(\theta_{ij}) \Big\} \tag{16}$$

である。潜在特性尺度が2パラメータロジスティックモデルで表現されている場合は，

$$\frac{\partial}{\partial \theta_{ij}} I(\theta_{ij}) = D^3 \sum_{h=1}^{n_j} \{a^3{}_h P_h(\theta_{ij}) Q_h(\theta_{ij}) \times (Q_h(\theta_{ij})\, P_h(\theta_{ij}))\} \tag{17}$$

となり，3パラメータロジスティックモデルの場合は

$$\frac{\partial}{\partial \theta_{ij}} I(\theta_{ij}) = D^3 \sum_{h=1}^{n_j} \{a^3{}_h (P_h(\theta_{ij}) - c_h)^2 \times (2P_h(\theta_{ij}) Q_h(\theta_{ij}) - P_h(\theta_{ij}) + c_h) \times Q_h(\theta_{ij})\, P_h(\theta_{ij})^{-2} (1-c_h)^{-3}\} \tag{18}$$

となる。

IV 周辺最尤推定法

最尤推定法によって等化係数を推定するためには $\boldsymbol{\theta}_0$ を同時に推定しなくてはならない。$\boldsymbol{\theta}_0$ はテストを等化するためには値を推定する必要のない局外母数であるから，モデルを θ_0 で周辺化し，積分消去して K, L を周辺最尤推定する。

等化のために用いられた共通被験者は，等化される尺度 S_0 上で平均 μ 分散 σ^2 の正規分布 $g(\theta_0|\mu.\sigma^2)$ する母集団からのN人の無作為標本であるという仮定を新たに加える。すると θ_0 で周辺化した i 番目の無作為標本の $\hat{\boldsymbol{\theta}}_i$ の分布は

$$f(\hat{\boldsymbol{\theta}}_i | K, L, \mu, \sigma^2) = \int_{-\infty}^{+\infty} f(\hat{\boldsymbol{\theta}}_i | \theta_0, K, L) \times g(\theta_0 | \mu, \sigma^2)\, d\theta_0 \tag{19}$$

となるので，$\hat{\Theta}$ が与えられたときの K, L, μ, σ^2 の周辺尤度関数は

$$L(\hat{\Theta} | K, L, \mu\sigma^2) = \prod_{i=1}^{N} f(\hat{\boldsymbol{\theta}}_i | K, L, \mu, \sigma^2) \tag{20}$$

となる。(20)式を K, L, μ, σ^2 について最大化することにより等化係数の周辺最尤推定値を得る。

V 欠測値の処理

N 人の共通被験者が必ずしも $m+1$ 個のテスト全てを受験しなかった場合に生じる欠測値の処理の方法を述べる。

最尤推定法・周辺最尤推定法における尤度関数の構成に共通して必要とされる(10)式は K, L が与えられたときの $\hat{\boldsymbol{\theta}}_i$ の同時分布であるが，仮りに被験者 i がテスト t を受験しなかったとすると $\hat{\boldsymbol{\theta}}_i = (\hat{\theta}_{i0}, \cdots \hat{\theta}_{it-1}, \hat{\theta}_{it+1}, \cdots \hat{\theta}_{im})$ となり，K, L が与えられたときの $\hat{\boldsymbol{\theta}}_i$ の同時分布は

$$f(\hat{\boldsymbol{\theta}}_i | \theta_{i0}, K, L) = \prod_{j=0, j\neq t}^{m} f(\hat{\theta}_{ij} | \theta_{i0}, k_j, l_j) \tag{21}$$

と表される。1人の被験者について欠測値が複数あっても同様に考えることができる。そこで被験者 i がテスト j を受験したならば $\delta_{ij}=1$，しなかったならば $\delta_{ij}=0$ として(10)式の代りに

$$f(\hat{\boldsymbol{\theta}}_i | \theta_{i0}, K, L) = \prod_{j=1}^{m} f(\hat{\theta}_{ij} | \theta_{i0}, k_j, l_j)^{\delta_{ij}} \tag{22}$$

を用いる。以後の与式については(11)式中の G_{ij} を $\delta_{ij} \times G_{ij}$ に，(13)・(14)・(15)式中の g_{ij} を $\delta_{ij} \times g_{ij}$ に置き換えればよい。

VI 今後の課題

本論文は先に提案された2つの潜在特性尺度の等化法（豊田，1986）を多数の尺度を同時に等化するためのモデルに拡張し，そのモデルが尤度モデルであるために欠測値の処理が比較的容易であることを示した。豊田（1986）では「等化法の精度を確認すること」が目的の1つであったから，目的関数最大化の方法としては収束する確率が最も高い最急降下法を用いた。本方法も同様に解を得ることができるが，最急降下法は解の近傍で収束が特に遅く，効率が悪い。今後の課題は，実際のテスト等化の場面で使用した場合に解の「発散」の危険がほとんどなく，かつ手軽に使用することが可能な程度に収束が速い計算手続による実用的なプログラムを作成することである。

（指導教官　芝　祐順教授）

引用文献

Hambleton, R.K. & Swaminathan, H. 1985 Item Response Theory Principles and Applications. Kluwer Nijhoff Publishing.

Lord, F.M. 1980 Applications of Item Response Theory to Practical Testing Problems. Lawrence Erlbaum Associate.

Lord, F.M. & Novick, M.R. 1968 Statistical Theories of Mental Test Scores, Addison-Wesley.

Samejima, F. 1977 A use of the information function in tailored testing. Applied Psychological Measurement, 233-247.

豊田秀樹　1986　被験者の推定尺度値とテスト情報関数を利用した潜在特性尺度の等化法，教育心理学研究，34，163-167.