Best-Worst 尺度法を利用した展開型項目反応モデル¹

池原 一哉² 早稲田大学

Unfolding item response model using best-worst scaling

Kazuya Ikehara (Waseda University)

In attitude measurement and sensory tests, the unfolding model is typically used. In this model, response probability is formulated by the distance between the person and the stimulus. In this study, we proposed an unfolding item response model using best-worst scaling (BWU model), in which a person chooses the best and worst stimulus among repeatedly presented subsets of stimuli. We also formulated an unfolding model using best scaling (BU model), and compared the accuracy of estimates between the BU and BWU models. A simulation experiment showed that the BWU model performed much better than the BU model in terms of bias and root mean square errors of estimates. With reference to Usami (2011), the proposed models were applied to actual data to measure attitudes toward tardiness. Results indicated high similarity between stimuli estimates generated with the proposed models and those of Usami (2011).

Key words: unfolding model, best-worst scaling, attitude measurement, Hamiltonian Monte Carlo method.

The Japanese Journal of Psychology

2015, Vol. 85, No. 6, pp. 560–570

J-STAGE Advanced published date: November 11, 2014, doi.org/10.4992/jjpsy.85.13077

項目反応理論 (Item Response Theory: IRT, Lord, 1980) は、テストデータや質問紙データを分析して、項目の特徴を捉え、回答者の能力や性格などの潜在特性を測定するための数理モデルである。学力テストでは、特性値の高い回答者ほど難しい項目に正答する確率が高く、また性格検査では、潜在特性の値が高いほど上位カテゴリを選択する確率が高くなる。このように、潜在特性 θ の値が高くなるにつれて、反応 (正答) 確率が高くなるようなモデルは、累積型 (cumulative) と呼ばれる。一方、意見・政策に対する態度の測定や味覚の官能検査では、自分の考えにより近い意見・政策に賛成と反応する確率が高く、また自分の好みにより合った製品にのびと反応する確率が高くなる。

このように、潜在特性 θ と位置母数 b (意見や製品の位置) との距離によって反応確率が定式化されるモデルは、展開型 (unfolding) と呼ばれる。

これまで、様々な展開型項目反応モデルが提案されてきた。2 値データに適用されるモデル (Andrich, 1988; Andrich & Luo, 1993) や、順序カテゴリカルデータに適用されるモデル (Andrich, 1996; Roberts, Donoghue, & Laughlin, 2000; Roberts & Laughlin, 1996) の他にも、意見や製品に関する一対比較データのためのモデル (Andrich, 1989, 1995) も提案されている。また、多次元展開法の観点から多肢選択データに対するモデル化も行われている (Takane, 1996)。しかし、調査対象となる複数の意見や製品の集合から、一部を取り出して部分集合を作成し、その中に含まれる意見や製品から最も自分に合うものを選択させるモデルは提案されていない。Takane (1996) も多肢選択型のモデルであるが、名義反応モデルの拡張であり、各項目のカテゴリに関する結果から項目ごとに解釈が行われる。

複数の選択肢の中から最も好ましいものを選ぶ方法は Best 尺度法 (Best Scaling) と呼ばれる。談・高野・岸野 (2000) は、20 個の花の嗜好テストに関して、不完備デザインを利用した 4 肢択一モデルと一対比較

Correspondence concerning this article should be sent to: Kazuya Ikehara, Global Education Center, Waseda University, Totsuka-cho, Shinjuku-ku, Tokyo 169-8050, Japan (e-mail: k.ikehara@ruri.waseda.jp)

¹ 分析プログラムおよびシミュレーションに関する情報については著者までご連絡ください。

² 本論文の執筆にあたり、豊田 秀樹先生 (早稲田大学) には有益な助言を頂きました。また、宇佐美 慧先生 (筑波大学) には、調査項目に関する情報を提供して頂きました。ここに記して謝意を表します。

モデルの比較を行い、嗜好度の高い対象に興味がある場合には、Best 尺度法が有効であることを示している。

一方、選択肢の中から、自分の考えや好みに最も合うもの (Best) と最も合わないもの (Worst) を両方とも選ぶ方法は Best-Worst 尺度法 (Best-Worst Scaling) と呼ばれる³。Best-Worst 尺度法は、Thurstone (1927) の一対比較モデルの多肢選択への拡張として Finn & Louviere (1992) により提案され、様々な分野で利用されている (Auger, Devinney, & Louviere, 2007; Cohen, 2009; 澤田・合崎・佐藤, 2010)。

Marley & Louviere (2005) は、Best-Worst 尺度法の利点として、Best 尺度法と比べて回答者の対象に対する順位付けに関する情報をより多く含んでいることや、Best と Worst の判断が比較的容易であることを挙げている。また、対象の数が多く設問数に限りがある場合に、対象に対する好みの情報をより多く収集するためには、対象を二つずつ評価する一対比較法よりも、複数個を同時に評価する Best-Worst 尺度法の方が効率的である。よって、対象に関する尺度を構成する場面において、各設問の情報量が多く、判断が容易な Best-Worst 尺度法は、推定精度と回答の負担の観点から、有益な手法であると考えられる。しかし、態度測定や官能検査において、Best-Worst 尺度法を利用し、かつ、展開型のメカニズムを考慮したモデルはこれまでに提案されていない。

以上より、本研究では、Best-Worst 尺度法を利用した展開型項目反応モデル (Best-Worst Unfolding Model: 以下 BWU モデルとする) を提案する。これにより、態度測定や官能検査場面において、意見や製品などの位置母数 b と個人の態度や嗜好を表す潜在特性 θ を、効率的にかつ精度よく推定することが期待できる。また、Best 尺度法を利用したモデル (Best Unfolding Model: 以下 BU モデルとする) の定式化も行い、BU モデルと BWU モデルにおける情報関数を導出する。そして、提案モデルの推定精度を検証するためにシミュレーション研究を行い、さらに遅刻に対する態度の調査を実施し、提案モデルを適用した分析例を示す。

モ デ ル

Best 尺度法と Best-Worst 尺度法

Best 尺度法および Best-Worst 尺度法は、多項ロジットモデル (multinomial logit model) を利用して表現される。Best 尺度法では、複数の対象を含む集合 S' において対象 k を選択する確率は、対象 k の効用 V_k を利用して、

$$P(k|S') = \frac{\exp(V_k)}{\sum_{k' \in S'} \exp(V_{k'})} \quad (1)$$

と表される。一方、Best-Worst 尺度法では、集合 S' において Best として対象 k を、Worst として対象 l を選択する確率は、それぞれの効用を用いて、

$$P(k, l|S') = \frac{\exp(V_k - V_l)}{\sum_{k' \in S'} \sum_{l' \in S', l' \neq k'} \exp(V_{k'} - V_{l'})} \quad (2)$$

と表現される。効用の差が大きいほど、対象 k を Best、対象 l を Worst として選ぶ確率が高くなることが分かる。

モデルの定式化

BU モデル 以下では、回答者数を I 、対象 (意見や製品など) の数を H 、設問数を J 、各設問で提示される対象の数を M とする。Best 尺度法を利用した展開型項目反応モデルでは、(1) 式より、回答者 i ($i = 1, \dots, I$) が設問 j ($j = 1, \dots, J$) において対象 k ($k = 1, \dots, M$) を選択する確率 $p_{j(k)}(\theta_i)$ を以下のように表現する。

$$p_{j(k)}(\theta_i) = \frac{\exp(z_{ij(k)})}{\sum_{k'=1}^M \exp(z_{ij(k')})}, \quad k' \in S_j \quad (3)$$

$$z_{ij(k)} = -(\theta_i - b_{j(k)})^2 \quad (4)$$

θ_i は回答者 i の潜在特性値、 $b_{j(k)}$ は設問 j における対象 k の位置母数である。 S_j は設問 j に含まれる対象の集合であり、 H 個から任意に選ばれた M 個の対象が含まれる。また、対象は設問ごとに配置される場所が異なることに注意が必要である。 $z_{ij(k)}$ が (1) 式の効用 V_k に対応しており、(3) 式と (4) 式より、潜在特性値 θ_i と対象の位置母数 b_h ($h = 1, \dots, H$) との距離が近いほど、選択確率が高くなる。

回答者 i の設問 j への反応ベクトルを $\mathbf{u}_{ij}^B$ とし、 $\mathbf{u}_{ij(k)}$ は反応ベクトル $\mathbf{u}_{ij}^B$ の k 番目の要素を表すとする。 $M = 4$ の場合、反応ベクトル $\mathbf{u}_{ij}^B$ は、以下のいずれかとなる。

$$\mathbf{u}_{ij}^B = \begin{cases} [1, 0, 0, 0]' \\ [0, 1, 0, 0]' \\ [0, 0, 1, 0]' \\ [0, 0, 0, 1]' \end{cases}$$

(3) 式より、 θ_i と $\mathbf{b} = [b_1, b_2, \dots, b_M]'$ が与えられたときに反応ベクトル $\mathbf{u}_{ij}^B$ が得られる確率は以下となる。

$$p(\mathbf{u}_{ij}^B | \theta_i, \mathbf{b}) = \prod_{k'=1}^M p_{j(k')}(\theta_i)^{u_{ij(k')}} \quad (5)$$

また、(3) 式は Andrich の展開型一対比較モデルの拡張であり、 $M = 2$ の時に Andrich (1989) に一致する。なお、以下では、 $p_{j(k)}(\theta_i)$ を $p_{j(k)}$ と略記する。

³ Best-Worst Scaling は、Maximum Difference Scaling とも呼ばれる。

BWU モデル (2) 式を拡張して、Best-Worst 尺度法を利用した展開型項目反応モデルでは、回答者 i ($i = 1, \dots, I$) が設問 j ($j = 1, \dots, J$) において、対象 k ($k = 1, \dots, M$) を Best とし、対象 l ($l = 1, \dots, M$) を Worst として選択する確率 $p_{j(k,l)}(\theta_i)$ を以下のように表現する。

$$p_{j(k,l)}(\theta_i) = \frac{\exp(z_{ij(k,l)})}{\sum_{k'=1}^M \sum_{l'=1}^M \exp(z_{ij(k',l')})}, \quad k', l' \in S_j \quad (6)$$

$$z_{ij(k,l)} = -(\theta_i - b_{j(k)})^2 + (\theta_i - b_{j(l)})^2 \quad (7)$$

$$= 2(b_{j(k)} - b_{j(l)}) \left(\theta_i - \frac{b_{j(k)} + b_{j(l)}}{2} \right) \quad (8)$$

θ_i は回答者 i の潜在特性値、 $b_{j(k)}$, $b_{j(l)}$ は位置母数であり、 S_j は設問 j に含まれる対象の集合を表す。(7) 式は効用の差を表しており、整理すると (8) 式が導かれる。(8) 式より、 $b_{j(k)} > b_{j(l)}$ の場合には、 $b_{j(k)}$ と $b_{j(l)}$ の差が大きく、かつ、 $b_{j(k)}$ と $b_{j(l)}$ の平均よりも θ_i が正に大きい場合に、選択確率が高くなる。

ここで、行に Best を、列に Worst を対応させた $M \times M$ 行列 U_{ij}^{BW} を考える。回答者 i が設問 j において対象 k を Best、対象 l を Worst として選択した場合には、行列 U_{ij}^{BW} の (k, l) 要素に 1 を配し、その他は 0 を配するとする。 $M=4$ の場合には、行列 U_{ij}^{BW} は、

$$U_{ij}^{BW} = \begin{bmatrix} -- & u_{ij(1,2)} & u_{ij(1,3)} & u_{ij(1,4)} \\ u_{ij(2,1)} & -- & u_{ij(2,3)} & u_{ij(2,4)} \\ u_{ij(3,1)} & u_{ij(3,2)} & -- & u_{ij(3,4)} \\ u_{ij(4,1)} & u_{ij(4,2)} & u_{ij(4,3)} & -- \end{bmatrix}$$

となり、対角要素には何も割り当てないとする。この行列 U_{ij}^{BW} を、対角要素を除いて行を優先してベクトル化したものを、回答者 i の設問 j への反応ベクトル $u_{ij}^{BW} = [u_{ij(1,2)}, u_{ij(1,3)}, \dots, u_{ij(1,M)}, u_{ij(2,1)}, u_{ij(2,3)}, \dots, u_{ij(M,1)}, u_{ij(M,2)}, \dots, u_{ij(M,M-1)}]^T$ とする。(6) 式より、 θ_i と $\mathbf{b}$ が与えられたときに反応ベクトル u_{ij}^{BW} が得られる確率は、

$$p(u_{ij}^{BW} | \theta_i, \mathbf{b}) = \prod_{k=1}^M \prod_{l'=1}^M p_{j(k',l')}(\theta_i)^{u_{ij(k',l')}} \quad (9)$$

となる。 $u_{ij(k,l)}$ は行列 U_{ij}^{BW} の (k, l) 要素である。以下では、 $p_{j(k,l)}(\theta_i)$ を $p_{j(k,l)}$ と略記する。

S_j の作成 設問 j で提示する M 個の対象が含まれる集合 S_j は、談他 (2000) を参考に、全設問を通して各対象の出現回数が同じになるよう不完備型計画により作成する。以下では、各設問の S_j を全設問分まとめて設問パターン S と呼ぶ。設問パターン S は統計解析環境 R のパッケージ “crossdes” (Sailer, 2013) に含まれる関数 find.BIB を利用して作成する。例として、対象の数が 8、設問数が 8、各設問で提示される対象の数が 4 の場合に、関数 find.BIB を利用して作成した設問

Table 1
設問パターン S の例 ($H=8, J=8, M=4$ の場合)

S_j	提示される対象			
設問 1 (S_1)	1	2	3	7
設問 2 (S_2)	2	3	5	8
設問 3 (S_3)	1	4	5	8
設問 4 (S_4)	3	5	6	7
設問 5 (S_5)	1	6	7	8
設問 6 (S_6)	1	3	4	6
設問 7 (S_7)	2	4	5	6
設問 8 (S_8)	2	4	7	8

パターン S の一つを Table 1 に示す。設問 1 では、対象 1, 2, 3, 7 の四つが提示され、設問 2 では対象 2, 3, 5, 8 の四つが提示される。八つの設問を通して、各対象が出現する回数は 4 回であり、関数 find.BIB により、各対象の出現回数が公平になるように設問パターン S を作成することができる。

母数推定と情報関数

本研究では、マルコフ連鎖モンテカルロ法 (Markov Chain Monte Carlo Method: MCMC 法) によるベイズ推定で母数推定を行い、サンプリング手法にはハミルトニアンモンテカルロ法 (Hamiltonian Monte Carlo Method: HMC 法, Duane, Kennedy, Pendleton, & Roweth, 1987) を利用する。HMC 法は、ハミルトン力学系による確定的な状態遷移と、メトロポリス法による確率的な状態遷移を融合した方法であり (松本, 2004)、母数による対数事後分布の偏導関数が容易に求められる場合に適用可能である。HMC 法は、ハイブリッド法とも呼ばれ、仮想的な運動量を導入して、推定対象となる母数と運動量変数の結合空間上でのエルゴード的なマルコフ連鎖を考える。

HMC 法の利点として、(a) 棄却確率を小さく保ったままで現在の状態から新しい状態へ大きく移動できる、(b) 自己相関が低い、(c) すべての母数を同時に更新できることが挙げられる。HMC 法については、Neal (2011) や Bishop (2006 元田・栗田・樋口・松本・村田監訳 2012) が詳しい。本研究では、村木 (2011) を参考に、はじめに MCMC 法により位置母数 $\mathbf{b}$ を推定し、位置母数の推定値 $\hat{\mathbf{b}}$ (事後平均) を利用して MAP (maximum a posteriori) 法 や EAP (expected a posteriori) 法によって潜在特性値 θ を求める。

事前分布

ベイズ推定では、母数に関する事前情報として事前分布を指定する。BU モデルおよび BWU モデルにおいて異なる事前分布を設定する理由が特段ないため、両モデルにおいて同一のものを利用する。de la Torre,

Stark, & Chernyshenko (2006) や Usami (2011) を参考に、潜在特性値 θ の事前分布 $p(\theta)$ には標準正規分布 $(N(0,1))$ を仮定する。一方、対象の位置母数 b_h ($h=1, \dots, H$) の事前分布 $p(b_h|\tau)$ については、平均 0、分散 τ^{-1} の正規分布 $(N(0, \tau^{-1}))$ を設定する。本研究では、BU モデルおよび BWU モデルで位置母数の散らばりに違いがあるかを検討するために、階層ベイズを利用して位置母数の分散 τ^{-1} も推定する。また、超母数 τ の事前分布 $p(\tau)$ には、ガンマ分布 $(G(\alpha_0, \beta_0))$ を仮定する。

尤度関数と事後分布

BU モデル 回答者 i ($i=1, \dots, I$) のすべての設問に対する反応を $U_i^B = [u_{i1}^B, u_{i2}^B, \dots, u_{iJ}^B]'$ とすると、 U_i^B が得られる確率は、(5) 式より局所独立を仮定すること、

$$p(U_i^B|\theta_i, \mathbf{b}) = \prod_{j=1}^J p(u_{ij}^B|\theta_i, \mathbf{b}) = \prod_{j=1}^J \prod_{k'=1}^M p_{j(k')}^{u_{ij(k')}} \quad (10)$$

となる。よって、全回答者の潜在特性値を $\theta = [\theta_1, \dots, \theta_I]'$ とすると、全回答者に関する尤度関数 $L_B(U^B|\theta, \mathbf{b})$ および対数尤度関数 $\log L_B(U^B|\theta, \mathbf{b})$ は、

$$L_B = \prod_{i=1}^I \prod_{j=1}^J \prod_{k'=1}^M p_{j(k')}^{u_{ij(k')}} \quad (11)$$

$$\log L_B = \sum_{i=1}^I \sum_{j=1}^J \sum_{k'=1}^M u_{ij(k')} \log p_{j(k')} \quad (12)$$

となる。一方、母数の事後分布 $p_B(\theta, \mathbf{b}, \tau|U^B)$ および対数事後分布 $\log p_B(\theta, \mathbf{b}, \tau|U^B)$ は以下となる。

$$p_B \propto L_B(U^B|\theta, \mathbf{b})p(\theta)p(\mathbf{b}|\tau)p(\tau) \quad (13)$$

$$\begin{aligned} \log p_B &= \log L_B(U^B|\theta, \mathbf{b}) + \log p(\theta) \\ &\quad + \log p(\mathbf{b}|\tau) + \log p(\tau) + \text{const.} \end{aligned} \quad (14)$$

BWU モデル 回答者 i ($i=1, \dots, I$) のすべての設問に対する反応を $U_i^{BW} = [u_{i1}^{BW}, u_{i2}^{BW}, \dots, u_{iJ}^{BW}]'$ とすると、 U_i^{BW} が得られる確率は、(9) 式より局所独立を仮定すること、

$$p(U_i^{BW}|\theta_i, \mathbf{b}) = \prod_{j=1}^J p(u_{ij}^{BW}|\theta_i, \mathbf{b}) = \prod_{j=1}^J \prod_{k'=1}^M \prod_{k' \neq l'}^M p_{j(k', l')}^{u_{ij(k', l')}} \quad (15)$$

となる。よって、全回答者に関する尤度関数 $L_{BW}(U^{BW}|\theta, \mathbf{b})$ および対数尤度関数 $\log L_{BW}(U^{BW}|\theta, \mathbf{b})$ は、

$$L_{BW} = \prod_{i=1}^I \prod_{j=1}^J \prod_{k'=1}^M \prod_{k' \neq l'}^M p_{j(k', l')}^{u_{ij(k', l')}} \quad (16)$$

$$\log L_{BW} = \sum_{i=1}^I \sum_{j=1}^J \sum_{k'=1}^M \sum_{k' \neq l'}^M u_{ij(k', l')} \log p_{j(k', l')} \quad (17)$$

となる。母数の事後分布 p_{BW} および対数事後分布 $\log p_{BW}$ は、(13) 式と (14) 式の尤度関数 L_B および対数尤度関数 $\log L_B$ を、尤度関数 L_{BW} および対数尤度関数 $\log L_{BW}$ に置き換えることで得られる。

分析の実行と初期値

それぞれのモデルにおいて、HMC 法を利用して事後分布からのマルコフ標本を発生させ、位置母数 $\mathbf{b}$ に関する事後統計量を算出する。実際の母数推定には、HMC 法を実装しているソフトウェア Stan (Stan Development Team, 2013) と統計解析環境 R のパッケージ “rstan” (Guo, Betancourt, Brubaker, Carpenter, Goodrich, Hoffman, Lee, Malecki, & Gelman, 2014) を利用する。なお、付録に Stan のプログラムを掲載する。

潜在特性値 θ_i ($i=1, \dots, I$) の初期値は 0、位置母数の分散の逆数 τ の初期値は 1 とする。 $\mathbf{b}$ の初期値には de la Torre et al. (2006) を参考に、-2.58 から 2.58 を H 個に等間隔に区切る値を用いる。なお、 τ の事前分布には無情報事前分布を仮定するため、 $\alpha_0 = \beta_0 = 0.001$ とする。

潜在特性値の推定

推定された位置母数 $\hat{\mathbf{b}}$ (事後平均) を利用して、 θ_i ($i=1, \dots, I$) に関する事後分布から各回答者の潜在特性値を推定する。MAP 法は、潜在特性値 θ_i に関する事後分布を最大にする θ_i を推定値とする方法であり、

$$\frac{\partial}{\partial \theta_i} \log L(U_i|\theta_i, \hat{\mathbf{b}}) + \frac{\partial}{\partial \theta_i} \log p(\theta_i) = 0 \quad (18)$$

を θ_i について解くことで、推定値 $\hat{\theta}_i^{MAP}$ を得る。ここで、 $L(U_i|\theta_i, \hat{\mathbf{b}})$ は各回答者の尤度であり (U_i は U_i^B または U_i^{BW})、 $p(\theta_i)$ は事前分布である。BU モデルの場合には尤度 L_B と BU モデルにおける推定値 $\hat{\mathbf{b}}_B$ を、BWU モデルの場合には尤度 L_{BW} と BWU モデルにおける推定値 $\hat{\mathbf{b}}_{BW}$ を用いる。

一方、EAP 法は、潜在特性値 θ_i の事後分布の θ_i に関する平均を推定値とする方法である。よって、推定値 $\hat{\theta}_i^{EAP}$ は、以下によって得られる。

$$\hat{\theta}_i^{EAP} = \int_{-\infty}^{\infty} \frac{\theta_i L(U_i|\theta_i, \hat{\mathbf{b}}) p(\theta_i)}{\int_{-\infty}^{\infty} L(U_i|\theta_i, \hat{\mathbf{b}}) p(\theta_i)} d\theta_i \quad (19)$$

実際には、区分求積法を利用して近似計算を行う。

情報関数

項目反応理論では、潜在特性 θ の推定の精度を表す指標として情報量 $I(\theta)$ を用いる。位置母数 $\mathbf{b}$ が既知のもとで、BU モデルにおける情報関数 $I_B(\theta)$ および BWU モデルにおける情報関数 $I_{BW}(\theta)$ は次式で表される。

$$I_B(\theta) = 4 \sum_{j=1}^J \sum_{k'=1}^M (\theta - b_{j(k')})^2 p_{j(k')} - 4 \sum_{j=1}^J \left\{ \sum_{k'=1}^M (\theta - b_{j(k')}) p_{j(k')} \right\}^2 + 1 \quad (20)$$

$$I_{BW}(\theta) = 4 \sum_{j=1}^J \sum_{k'=1}^M \sum_{l'=1}^M (b_{j(k')} - b_{j(l')})^2 p_{j(k',l')} - 4 \sum_{j=1}^J \left\{ \sum_{k'=1}^M \sum_{l'=1}^M (b_{j(k')} - b_{j(l')}) p_{j(k',l')} \right\}^2 + 1 \quad (21)$$

本研究では、潜在特性 θ の事前分布に標準正規分布を仮定しているため、 $I_B(\theta)$ および $I_{BW}(\theta)$ の最終項に分散の逆数 1 が足される。なお、付録には情報関数の導出を掲載した。情報関数の逆数の平方根を利用することで、推定値の標準偏差 $SD(\theta) = 1/\sqrt{I(\theta)}$ を、 θ のレベルごとに見積もることが可能となる。

シミュレーション

シミュレーションデータを用いて、提案モデルにおける母数の推定精度を検討する。具体的には、回答者数、対象の数、設問数の条件を変化させ、シミュレーションデータから求められる推定値と真値の違いを評価する。

方法

シミュレーションデータを以下の手順で発生させた。

step 1 θ_j を標準正規乱数を利用して I 個発生させる。

step 2 任意に設定した母数 b の真値と設問パターン S を利用して、設問ごとに回答者の選択確率 ((3) 式および (6) 式) を計算する。

step 3 選択確率をもとに多項分布から反応データ (U_i^B および U_i^{BW}) を I 個作成する。

各設問で提示される対象の数を $M=4$ と固定し、回答者数を $I=100, 200, 400$ の 3 通り、対象の数を $H=10, 20$ の 2 通り、設問数を $J=10, 20$ の 2 通りとした。よって、BU モデルと BWU モデルのそれぞれで 12 条件 ($= 3 \times 2 \times 2$) のシミュレーションを実施した。母数 b は $H=10, 20$ の各条件において標準正規分布から発生させ、設問パターン S は関数 find.BIB により作成した。

上述の手順で各条件においてシミュレーションデータを 50 個発生させ、BU モデルと BWU モデルを適用した。推定精度を検証するために、位置母数 b_h ($h=1, \dots, H$) に関してはバイアスと RMSE (Root Mean Square Error) を、潜在特性値 θ に関しては相関の平均を算出した。

$$b_{\text{bias}} = \frac{1}{50} \sum_{c=1}^{50} \hat{b}_{hc} - b_{\text{true}} \quad (22)$$

$$b_{h\text{RMSE}} = \sqrt{\frac{1}{50} \sum_{c=1}^{50} (\hat{b}_{hc} - b_{\text{true}})^2} \quad (23)$$

$$\theta_{\text{cor}} = \frac{1}{50} \sum_{c=1}^{50} \text{cor}(\hat{\theta}_c, \theta_{\text{true}}) \quad (24)$$

$\hat{b}_{hc}$ は c 回目 ($c=1, \dots, 50$) のシミュレーションにおける位置母数 b_h の推定値 (事後平均) であり、 b_{true} はその真値である。また $\hat{\theta}_c$ は c 回目のシミュレーションにおける全回答者の潜在特性値 θ の EAP 推定値であり、 θ_{true} はその真値を表す。なお、紙面の都合上、各位置母数 b_h のバイアスと RMSE を報告するのではなく、バイアスの絶対値に関する平均と RMSE の平均を算出する。

$$b_{\text{bias}} = \frac{1}{H} \sum_{h=1}^H |b_{h\text{bias}}| \quad b_{\text{RMSE}} = \frac{1}{H} \sum_{h=1}^H b_{h\text{RMSE}} \quad (25)$$

R のパッケージ “rstan” を利用して MCMC 法を実行し、15,000 回のサンプリングを行った。また、初期値の影響を除くために、バーンイン期間として前半の 5,000 回分を破棄し、残りの 10,000 個の MCMC 標本を母数推定に利用した。なお、(19) 式の近似にはガウス・エルミート求積法 (分割数 21) を利用した。

結果

位置母数 b のバイアスの絶対値に関する平均 b_{bias} 、RMSE の平均 b_{RMSE} 、潜在特性値 θ に関する相関の平均 θ_{cor} を算出した結果を Table 2 にまとめる。Table 2 より、BU モデルおよび BWU モデルに共通して、回答者数 I が多いほど、バイアスの値や RMSE の値が小さくなる傾向にあることが分かる。また、対象の総数 H が同じであれば、設問数 J が多いほど RMSE の値が小さく、精度よく推定できることが分かる。なお、潜在特性値に関する相関は、どの条件においても 0.95 以上であり、提案モデルにより真値を復元できることが示された。

BU モデルと BWU モデルを比べると、BWU モデルの方がバイアスや RMSE の値が小さく、相関の値が大きいたことが示された。さらに、 $H=10$ の場合に注目し、設問数 20 の BU モデルと設問数 10 の BWU モデルを比較すると、位置母数に関する RMSE や潜在特性値に関する相関は同程度であることが確認できる。つまり、設問数を減らしても、BWU モデルを適用することで同程度の精度で母数推定を行えることが示唆される。なお、同じような傾向が $H=20$ の場合においても見られた。

適用例

遅刻に対する態度を測定するために、Best-Worst 尺度法による質問紙を作成してデータを収集し、提案モ

Table 2
各母数におけるシミュレーション結果

		b_{bias}			
I		H10J10	H10J20	H20J10	H20J20
100	BU	0.0147	0.0125	0.0214	0.0154
	BWU	0.0142	0.0119	0.0158	0.0117
200	BU	0.0078	0.0080	0.0150	0.0073
	BWU	0.0051	0.0053	0.0103	0.0072
400	BU	0.0089	0.0085	0.0095	0.0052
	BWU	0.0061	0.0073	0.0064	0.0043
		b_{RMSE}			
100	BU	0.1334	0.1086	0.1584	0.1194
	BWU	0.1097	0.0972	0.1221	0.1050
200	BU	0.0882	0.0771	0.1115	0.0838
	BWU	0.0783	0.0679	0.0840	0.0687
400	BU	0.0688	0.0583	0.0832	0.0638
	BWU	0.0566	0.0532	0.0637	0.0542
		θ_{cor}			
100	BU	0.9606	0.9767	0.9540	0.9758
	BWU	0.9739	0.9825	0.9693	0.9815
200	BU	0.9590	0.9766	0.9547	0.9752
	BWU	0.9726	0.9824	0.9702	0.9812
400	BU	0.9600	0.9768	0.9549	0.9750
	BWU	0.9724	0.9824	0.9703	0.9813

注) $H=10$ と $J=10$ の組合せは, H10J10 のように略記して示す。

デルを適用する。態度測定に用いる意見には、順序カテゴリカルデータを分析した Usami (2011) の 20 項目を利用し、BU モデルと BWU モデルを適用した結果の比較を行う。

方 法

調査時期 2013 年 11 月から 12 月にかけて行った。

回答者 回答者は、東京都の私立大学に通う大学生・大学院生 172 名（男性 110 人・女性 62 人）である。

質問紙 Usami (2011) の遅刻に対する態度の測定に利用された 20 個の意見を、各設問において四つずつ提示し、最も自分の考えに合うものと最も合わないものを選択してもらった。Figure 1 に設問に対する回

Table 3
態度測定の調査で利用した設問パターン S

S_j	提示される意見				S_j	提示される意見			
S_1	5	7	8	18	S_{11}	1	8	13	14
S_2	2	3	10	15	S_{12}	7	11	13	18
S_3	1	5	11	14	S_{13}	5	12	14	17
S_4	6	13	19	20	S_{14}	2	3	10	15
S_5	4	7	9	16	S_{15}	4	6	16	19
S_6	10	14	16	19	S_{16}	9	12	18	20
S_7	1	2	6	9	S_{17}	1	2	9	16
S_8	3	13	17	18	S_{18}	3	6	11	17
S_9	8	12	15	20	S_{19}	8	10	19	20
S_{10}	4	11	12	17	S_{20}	4	5	7	15

答形式の例を示す。

設問パターン 関数 find.BIB を利用して作成した設問パターン S を Table 3 に示す。設問数 (J) は 20 であり、各意見が 4 回ずつ出現するように設問パターン S を作成した。Table 3 に示した設問パターン S では、前半 10 個の設問において各意見が 2 回ずつ出現し、後半の 10 個の設問において、同様に各意見が 2 回ずつ出現する。

教示文例 四つの意見のうち、“遅刻”に対するあなたの考えに最も合う意見と最も合わない意見をそれぞれ一つずつ選んで下さい。最も合う意見は Best の欄に、最も合わない意見は Worst の欄に○印を入れて下さい。

母数の推定 シミュレーション研究と同様に、Stan と統計解析環境 R を利用して、MCMC 法を実行した。R のパッケージ “rstan” を利用して、間引き間隔を 5 とし、11,000 回のサンプリングを行った（実際には 55,000 個の標本を得た）。初期値の影響を除去するために、前半の 1,000 回分をバーンイン期間として破棄し、後半の 10,000 個の MCMC 標本を利用して母数の推定を行った。

情報関数 推定された位置母数を利用して両モデルにおける潜在特性の情報量を算出する。そして、情報量の逆数の平方根を利用して、潜在特性のレベルごとの $SD(\theta)$ を図示する。また、前半 10 個の設問と 20 個すべての設問を利用した場合の $SD(\theta)$ の違いも検討す

Best	意見	Worst
	理由が何であっても、遅刻をするような人には信頼を置くことはできない	
	遅刻をしないように、我々は常に最善の努力をすべきである	
	本人に直接的な原因が無い場合には、遅刻が許されることはある	
	時間を守ることや遅刻することは、その人の人間性とは無関係であると思う	

Figure 1. Best-Worst 尺度法の設問の例

Table 4
意見の位置母数に関する推定結果（括弧内は事後標準偏差）

意見	BU モデル	BWU モデル	Usami
1. いかなる理由であっても、大事な人に会うかどうかに限らず、遅刻をすることは絶対に許されることではない	-2.249 (0.098)	-1.888 (0.083)	-2.480
2. 突然の事故など、本当に特別な理由がない限り遅刻を許すことはできない	-1.659 (0.092)	-1.491 (0.086)	-2.160
3. いかなる理由であっても、遅刻をすることはその人の人間的な価値を減じる	-1.832 (0.089)	-1.686 (0.084)	-1.960
4. 遅刻をすることは人として恥ずかしいことである	-1.533 (0.090)	-1.433 (0.085)	-1.810
5. 理由が何であっても、遅刻をするような人には信頼を置くことはできない	-2.059 (0.096)	-1.746 (0.083)	-1.520
6. 基本的に遅刻は許される行為ではない	-0.770 (0.109)	-0.818 (0.103)	-1.120
7. 遅刻をしないように、我々は常に最善の努力をすべきである	-0.210 (0.119)	-0.336 (0.117)	-1.090
8. 本人に直接的な原因がない場合には、遅刻が許されることはある	0.602 (0.116)	0.658 (0.114)	-0.230
9. 遅刻をしても許される場面はいくらかある	1.039 (0.095)	1.038 (0.093)	0.360
10. 遅刻は時に誰でもしてしまう、ある意味仕方のない行為である	1.148 (0.093)	1.138 (0.091)	0.560
11. 遅刻をするかしないかで、その人の価値を判断することはない	1.427 (0.089)	1.372 (0.085)	0.600
12. 多くの場合、遅刻は望ましくないが、遅刻は必ずしも責められるべきものではない	0.593 (0.108)	0.603 (0.114)	0.830
13. 遅刻が良いことなのか悪いことなのかは判断しかねる	1.796 (0.090)	1.616 (0.083)	1.010
14. 大した用件ではない場合などでは、遅刻はしてしまうものである	1.516 (0.089)	1.523 (0.084)	1.210
15. 決められた時間に縛られて行動する必要はないと思う	2.217 (0.105)	1.887 (0.082)	1.260
16. 時間については少しくらいルーズな方が良いと思う	2.068 (0.097)	1.823 (0.083)	1.390
17. 時間にとらわれず、自分のペースに基づいて行動すべきである	2.186 (0.098)	1.885 (0.082)	1.500
18. 時間を守ることや遅刻することは、その人の人間性とは無関係であると思う	1.906 (0.092)	1.792 (0.082)	1.600
19. 決められた時間通りに着こうとする努力はしなくて良い	2.509 (0.127)	2.168 (0.082)	2.110
20. どのような場合であっても、時間を守って行動する必要はない	2.402 (0.133)	2.295 (0.081)	2.920
位置母数 b の平均（括弧内は事後標準偏差の平均）	0.555 (0.101)	0.520 (0.090)	0.149
位置母数の分散 τ^2	3.268 (1.169)	2.695 (0.956)	2.379

注) Usami (2011) の位置母数の分散は、推定値を利用して計算した。

る。

結果・考察

はじめに、BU モデルおよび BWU モデルにおいて推定された遅刻に対する 20 個の意見の位置母数 b と位置母数の分散 τ^2 の推定値（事後平均）と事後標準偏差を Table 4 に示す⁴。同時に、位置母数の推定値の平均と事後標準偏差の平均も計算した。Usami (2011) の推定値とのピアソンの積率相関係数は、BU モデルでは 0.965、BWU モデルでは 0.970 と高い値となった（ケンドールの順位相関係数はそれぞれ 0.842 と 0.853 であった）。なお、BU モデルと BWU モデルの推定値間の相関は 0.998 であった。Table 4 には、Usami (2011)

の推定値の順番で意見を並べた。位置母数 b の値が小さいほど遅刻に対して厳しい意見であり、値が大きいほど遅刻に対して寛容な意見であることを表す。同様に、潜在特性値 θ の値が小さいほど遅刻に対して厳しい回答者であり、値が大きいほど遅刻に対して寛容な回答者であることを表す。

Usami (2011) の推定値による意見の順番と BU モデルおよび BWU モデルの推定値による意見の順番を比較すると、意見 5, 12, 15 において違いがみられた。意見 5 “理由が何であっても、遅刻をするような人には信頼を置くことはできない” と意見 12 “多くの場合、遅刻は望ましくないが、遅刻は必ずしも責められるべきものではない” は、Usami (2011) と比べて、相対的に厳しい意見として推定された。一方、意見 15 “決められた時間に縛られて行動する必要はないと思う” は、相対的に寛容な意見であることが示された。Usami (2011) の分析結果と順番が異なる点もあるが、

⁴ 両モデルにおいて、位置母数 b および位置母数の分散 τ^2 の Geweke 指標の絶対値は 1.96 を下回っており、マルコフ連鎖の収束が示唆されたため、分析結果には載せなかった。

推定値間の相関係数は高く、BU モデルおよび BWU モデルによって分析を行っても同様な結果が得られることが確認できた。

BU モデルと BWU モデルの推定結果を比べると、BU モデルの方が位置母数の分散 σ^2 が大きく、広範囲に散らばっていることが分かる。また、事後標準偏差の平均は BWU モデルの方が小さく、Best-Worst 尺度法を利用することで推定精度が向上することが示唆される。

次に、潜在特性値 θ の EAP 推定値のヒストグラムを Figure 2 に示す。Figure 2 の上図が BU モデル (平均: -0.028 , 標準偏差: 0.388)、下図が BWU モデル (平均: -0.028 , 標準偏差: 0.331) であり、ヒストグラムには濃い黒線で位置母数も同時に示した。Figure 2 より、BU モデルの潜在特性値 $\hat{\theta}$ の方が BWU モデルよりも散らばりが大きく、BWU モデルでは 0 付近の度数が多くなっていることが分かる。潜在特性値は、両極端

Figure 2. 潜在特性値のヒストグラムと位置母数

注) 上図が BU モデルのヒストグラム、下図が BWU モデルのヒストグラムである。

の意見の位置母数の内側に分布しており、 -1 から 1 の間にほとんどの回答者が含まれている。

最後に、BWU モデルの推定値を利用して (20) 式と (21) 式の情報量を計算し、潜在特性のレベルごとの $SD(\theta)$ をプロットした図を Figure 3 に示す。Figure 3 には、両モデルにおいて、20 個すべての設問を利用した場合の $SD(\theta)$ の曲線 (BU20 と BWU20) と前半 10 個の設問を利用した場合の $SD(\theta)$ の曲線 (BU10 と BWU10) をプロットした。Figure 3 より、設問数によらず、BWU モデルの方が BU モデルよりも潜在特性のレベルごとの $SD(\theta)$ が小さく、精度がよいことが分かる。また、設問数が増えることで $SD(\theta)$ が小さくなり、推定精度が向上することも確認できる。さらに、潜在特性値が -0.3 から 0.4 の間では、20 個の設問を用いて BU モデルを適用するよりも、10 個の設問を用いて BWU モデルを利用した方が精度がよく、 0.4 から 1.5 の範囲においても同程度の誤差で潜在特性値を推定できることが示唆された。ただし、情報関数は調査で利用した設問パターン S により決まるものであり、別の設問パターンを利用することで、 $SD(\theta)$ の曲線は変わることには注意が必要である。

考 察

本研究では、Best 尺度法および Best-Worst 尺度法を利用した展開型項目反応モデルの提案を行った。また、シミュレーション研究を行って提案モデルの推定精度について確認し、遅刻に対する態度の調査を行って提案モデルを適用した。

シミュレーション研究より、両モデルにおいて、適切に母数の推定ができることが確認された。適用例では、遅刻に対する態度データを分析し、位置母数に関

Figure 3. BU モデルおよび BWU モデルにおける潜在特性 θ の標準偏差 $SD(\theta)$

して、BU モデルおよび BWU モデルのいずれにおいても、Usami (2011) の分析結果と高い相関があることが分かった。従来の手法とは異なる定式化により提案したモデルにおいて、先行研究と同様の結果が得られたことは、提案手法が妥当であることの傍証と言えるだろう。

また、シミュレーション研究と適用例に共通して、BU モデルよりも BWU モデルの方が精度よく推定できることが示された。これは、Best 尺度法よりも Best-Worst 尺度法の方が、対象に対する順位付けの情報をより多く有しているためであり、Marley & Louviere (2005) で言及に一致する。さらに、シミュレーションでは、設問数を半分に減らした場合でも、BWU モデルは BU モデルと同程度の精度で母数推定を行うことが確認された。加えて、適用例で示したように、提案モデルでは潜在特性のレベルごとの $SD(\theta)$ を評価することができる。

以上より、Best-Worst 尺度法の利点を生かした BWU モデルは、展開型のメカニズムが仮定される態度測定や官能検査場面において、有用な手法であると考えられる。BWU モデルでは、設問数を減らす代わりに、Best と Worst を選択させることで、推定精度を低下させることなく分析することが可能である。特に、対象の数が多く、設問数に制限がある場合には、推定精度や回答負担の観点から、Best 尺度法よりも Best-Worst 尺度法を用いる方が有効であろう。

ただし、今後検討すべき課題も幾つか残っている。提案モデルは、設問パタンの作成によって推定の精度が変化するという特徴があるため、設問パターンが精度に及ぼす影響について検証する必要がある。また、今回の研究だけでは、回答者の負担の程度を十分に知ることはできない。回答時間の測定や自由記述での意見収集など、様々な観点から回答負担に関して検討を行う必要がある。

本研究では、MCMC 法によるベイズ推定によって母数推定を行ったが、推定方法には EM アルゴリズムを利用した周辺最尤推定や同時最尤推定などもある。提案モデルに関して、様々な推定方法間の比較やノイズを含んだシミュレーションデータの推定精度の検証を行うことも重要である。加えて、実施方法についても議論すべきであろう。本研究の適用例では紙ベースの調査を行ったが、IRT の分野で研究が進んでいる適応型テスト (Computerized Adaptive Testing: CAT) の利用により、実施時間や設問数に制約がある場合でも提案手法を応用することができる。その点についても今後の検討が望まれる。

引用文献

- Andrich, D. (1988). The application of an unfolding model of the PIRT type to the measurement of attitude. *Applied Psychological Measurement*, **12**, 33–51.
- Andrich, D. (1989). A probabilistic IRT model for unfolding preference data. *Applied Psychological Measurement*, **13**, 193–216.
- Andrich, D. (1995). Hyperbolic cosine latent trait models for unfolding direct responses and pairwise preferences. *Applied Psychological Measurement*, **19**, 269–290.
- Andrich, D. (1996). A hyperbolic cosine latent trait model for unfolding polytomous responses: Reconciling Thurstone and Likert methodologies. *British Journal of Mathematical and Statistical Psychology*, **49**, 347–365.
- Andrich, D., & Luo, G. (1993). A hyperbolic cosine latent trait model for unfolding dichotomous single-stimulus responses. *Applied Psychological Measurement*, **17**, 253–276.
- Auger, P., Devinney, T. M., & Louviere, J. J. (2007). Using best-worst scaling methodology to investigate consumer ethical beliefs across countries. *Journal of Business Ethics*, **70**, 299–326.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer.
- (ビショップ, C. M. 元田 浩・栗田 多喜夫・樋口 知之・松本 裕治・村田 昇 (監訳) (2012). *パターン認識と機械学習 下——ベイズ理論による統計的予測——* 丸善出版)
- Cohen, E. (2009). Applying best-worst scaling to wine marketing. *International Journal of Wine Business Research*, **21**, 8–23.
- Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. (1987). Hybrid monte carlo. *Physics Letters B*, **195**, 216–222.
- Finn, A., & Louviere, J. J. (1992). Determining the appropriate response to evidence of public concern: The case of food safety. *Journal of Public Policy and Marketing*, **11**, 12–25.
- Guo, J., Betancourt, M., Brubaker, M., Carpenter, B., Goodrich, B., Hoffman, M., Lee, D., Malecki, M., & Gelman, A. (2014). *RStan: The R interface to Stan*. <http://mc-stan.org/rstan.html> (April 22, 2014)
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Marley, A. A. J., & Louviere, J. J. (2005). Some probabilistic models of best, worst, and best-worst choices. *Journal of Mathematical Psychology*, **49**, 464–480.
- 松本 隆 (2004). 非線形ダイナミカルシステムの再構成と予測 石黒 真木夫・松本 隆・乾 敏郎・田辺 国土・伊庭 幸人・甘利 俊一・竹内 啓・竹村 彰通 (編) *階層ベイズモデルとその周辺——時系列・画像・認知への応用——* 岩波書店 pp. 89–170.
- (Matsumoto, T.)
- 村木 英治 (2011). *項目反応理論* 朝倉書店 (Muraki, E.)
- Neal, R. M. (2011). MCMC using Hamiltonian dynamics.
- Andrich, D. (1988). The application of an unfolding model

- In S. Brooks, A. Gelman, G. L. Jones, & X. L. Meng (Eds.), *Handbook of Markov Chain Monte Carlo*. Boca Raton, FL: Chapman & Hall/CRC. pp. 113–162.
- Roberts, J. S., Donoghue, J. R., & Laughlin, J. E. (2000). A general item response theory model for unfolding unidimensional polytomous responses. *Applied Psychological Measurement*, **24**, 3–32.
- Roberts, J. S., & Laughlin, J. E. (1996). A unidimensional item response model for unfolding responses from a graded disagree-agree response scale. *Applied Psychological Measurement*, **20**, 231–255.
- Sailer, M. O. (2013). *Crossdes: Construction of crossover designs*. R package version 1.1-1. <http://CRAN.R-project.org/package=crossdes> (April 22, 2014)
- 澤田 学・合崎 英男・佐藤 和夫 (2010). 牛肉生産における飼料自給率向上の利点に関する消費者評価 帯広畜産大学学術研究報告, **31**, 18–24.
(Sawada, M., Aizaki, H., & Sato, K. (2010). Consumers' evaluation on the advantages of improving feed self-sufficiency in beef production. *Research Bulletin of Obihiro University*, **31**, 18–24.)
- Stan Development Team. (2013). *Stan modeling language user's guide and reference manual*, Version 1.3.
- Takane, Y. (1996). An item response model for multidimensional analysis of multiple-choice data. *Behaviormetrika*, **23**, 153–167.
- 談 小健・高野 泰・岸野 洋久 (2000). 不完備型嗜好テストにおける四者択一方式の有効性 計量生物学, **20**, 167–179.
(Tan, X., Takano, Y., & Kishino, H. (2000). Efficiency of the method of four-fold choice in incomplete preference trials. *Japanese Journal of Biometrics*, **20**, 167–179.)
- Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review*, **34**, 273–286.
- de la Torre, J., Stark, S., & Chernyshenko, O. S. (2006). Markov Chain Monte Carlo estimation of item parameters for the generalized graded unfolding model. *Applied Psychological Measurement*, **30**, 216–232.
- Usami, S. (2011). Generalized graded unfolding model with structural equation for subject parameters. *Japanese Psychological Research*, **53**, 221–232.

—— 2014. 2. 4 受稿, 2014. 7. 12 受理 ——

付 録

偏導関数と情報関数の導出

BU モデル $\theta_i (i = 1, \dots, I)$ による対数尤度 $\log L_B$ の 1 次偏導関数および 2 次偏導関数は

$$\begin{aligned} \frac{\partial \log L_B}{\partial \theta_i} &= \sum_{j=1}^J \sum_{k'=1}^M u_{ij}(k') \frac{1}{p_j(k')} \frac{\partial p_j(k')}{\partial \theta_i} \\ &= -2 \sum_{j=1}^J \sum_{k'=1}^M (\theta_i - b_j(k')) (u_{ij}(k') - p_j(k')) \\ \frac{\partial^2 \log L_B}{\partial \theta_i^2} &= -2 \sum_{j=1}^J \sum_{k'=1}^M \left\{ (u_{ij}(k') - p_j(k')) - (\theta_i - b_j(k')) \frac{\partial p_j(k')}{\partial \theta_i} \right\} \\ &= -2 \sum_{j=1}^J \sum_{k'=1}^M (u_{ij}(k') - p_j(k')) - 4 \sum_{j=1}^J \sum_{k'=1}^M (\theta_i - b_j(k'))^2 p_j(k') \\ &\quad + 4 \sum_{j=1}^J \left\{ \sum_{k'=1}^M (\theta_i - b_j(k')) p_j(k') \right\}^2 \end{aligned}$$

となる。ただし $\frac{\partial p_j(k')}{\partial \theta_i} = -2p_j(k') \{(\theta_i - b_j(k')) - \sum_{k'=1}^M p_j(k') (\theta_i - b_j(k'))\}$ である。情報関数は $I_B(\theta) = -E \left[\frac{\partial^2 \log L_B}{\partial \theta_i^2} + \frac{\partial^2 \log p(\theta)}{\partial \theta_i^2} \right]$ により算出され、 $E[u_{ij}(k') - p_j(k')] = 0$ および $\partial^2 \log p(\theta) / \partial \theta_i^2 = -1$ より (20) 式が導かれる。

BWU モデル $\theta_i (i = 1, \dots, I)$ による対数尤度 $\log L_{BW}$ の 1 次偏導関数および 2 次偏導関数は

$$\begin{aligned} \frac{\partial \log L_{BW}}{\partial \theta_i} &= \sum_{j=1}^J \sum_{k'=1}^M \sum_{l'=1}^M u_{ij}(k', l') \frac{1}{p_j(k', l')} \frac{\partial p_j(k', l')}{\partial \theta_i} \\ &= 2 \sum_{j=1}^J \sum_{k'=1}^M \sum_{l'=1}^M (b_j(k') - b_j(l')) (u_{ij}(k', l') - p_j(k', l')) \\ \frac{\partial^2 \log L_{BW}}{\partial \theta_i^2} &= -2 \sum_{j=1}^J \sum_{k'=1}^M \sum_{l'=1}^M (b_j(k') - b_j(l')) \frac{\partial p_j(k', l')}{\partial \theta_i} \\ &= -4 \sum_{j=1}^J \sum_{k'=1}^M \sum_{l'=1}^M (b_j(k') - b_j(l'))^2 p_j(k', l') \\ &\quad + 4 \sum_{j=1}^J \left\{ \sum_{k'=1}^M \sum_{l'=1}^M (b_j(k') - b_j(l')) p_j(k', l') \right\}^2 \end{aligned}$$

となる。ただし $\frac{\partial p_j(k', l')}{\partial \theta_i} = (b_j(k') - b_j(l')) - \sum_{k'=1}^M \sum_{l'=1}^M p_j(k', l') (b_j(k') - b_j(l'))$ である。よって、 $-E \left[\frac{\partial^2 \log L_{BW}}{\partial \theta_i^2} + \frac{\partial^2 \log p(\theta)}{\partial \theta_i^2} \right]$ より (21) 式の情報関数 $I_{BW}(\theta)$ が導かれる。

Stan プログラム

BWU モデルの分析プログラムを以下に示す。

#I: 回答者数, H: 対象数, J: 設問数, J2: 2*J, M: 各設問で提示される対象数, MM: 各設問の対象の組合せ数, res: データ, Comb: 各設問の対象の組合せ行列

```
data{
  int<lower=0> I; int<lower=0> H;
  int<lower=0> J; int<lower=0> J2;
  int<lower=0> MM;
  int<lower=1, upper=MM> res[I, J];
  int<lower=1, upper=H> Comb[J2, MM];
}
parameters{
  real theta[I]; real b[H];
  real<lower=0> tau_b;
}
transformed parameters{
  vector[MM] b_p[J]; vector[MM] b_m[J];
  vector[MM] p[I, J]; real sig_b;
  sig_b<-1/tau_b;
  # (8) 式の位置母数の和と差を計算
  for(j in 1:J){for(kk in 1:MM){
    b_p[j, kk]<-(b[Comb[(2*j-1), kk]]+
      b[Comb[(2*j), kk]])/2;
    b_m[j, kk]<-2*(b[Comb[(2*j-1), kk]]-
      b[Comb[(2*j), kk]]);}}
  # (6) 式の各設問における対象の選択確率
  for(i in 1:I){for(j in 1:J){
    p[i, j]<-softmax(b_m[j] .*
      (theta[i]-b_p[j]));}}
}
model{
  #事前分布
  theta~normal(0, 1);
  b~normal(0, sqrt(sig_b));
  tau_b~gamma(1.0E-3, 1.0E-3);
  #モデル
  for(i in 1:I){for(j in 1:J){
    res[i, j]~categorical(p[i, j]);}}}
```

設問パタンの作成

設問パターン s および各設問における対象の組合せ行列 Comb を作成するプログラムを以下に示す。

```
#設問パターン s の作成 (パッケージ"crossdes")
S<-find.BIB(trt=H, b=J, k=M)
#Comb の作成 (パッケージ"gregmisc")
Comb <- matrix(0, J2, MM)
for(j in 1:J){Comb[(2*j-1):(2*j), ]<-
  t(permutations(n=M, r=2, v=S[j, ],
    repeats=FALSE))}
```