

連続反応との統合を目指した段階反応のモデリング手法 Modeling Techniques for Graded Response Aimed at Integration with Continuous Response

塚村 祐希, 植田 一博
Yuki Tsukamura, Kazuhiro Ueda

東京大学
The University of Tokyo
tsukkacogsci@gmail.com

概要

本発表では、段階反応データに対し、連続的な反応分布を仮定したモデルを適用し、閾値に対称制約を置くことによってモデルを識別可能にする方法を説明する。さらに、この方法を用いた分析の例として、因果的説明における潜在スコープバイアスに関する先行研究のデータをモデリングした結果を報告する。本手法は、モデリングを想定せず取得された段階データをモデルベースで分析しやすくするものであり、二次分析の積極的な実施を促進すると考えられる。

キーワード：段階反応, バイズ統計モデリング, 二次分析, 潜在スコープバイアス

1. 問題と目的

近年、計算機や統計手法の発展により、実験データや回答データのモデリングが可能になっている。数理モデルをデータにフィッティングし、推定されたパラメータを解釈したり、どのモデルがデータをよりよく説明できるかを判断することは、心理学・認知科学における理論を発展させる上で重要である。

一方、心理学・認知科学で用いられる回答形式には様々なものがあり、統一的にモデルをフィッティングするのが難しい場合が少なくない。特に、モデリングすることを想定せずに取得されたデータの中には、5段階や7段階など、必ずしも段階数の多くないものが存在する。このようなデータに対して、連続的な反応分布を仮定して構成されたモデルを直接適用することは難しい。

このような問題に対して、主にテストの作成や評価などに用いられる項目反応理論 (IRT) の枠組みでは、段階反応モデル (graded response model; Samejima, 1969), およびその拡張としての連続反応モデル (continuous response model; Samejima, 1973) が提案されており、連続回答と段階回答に対して、同一の枠組みのもとで項目の識別力や困難度のパラメータを推定す

ることが可能となっている。前述のケースにおいても、同一の枠組みの中で、何らかの仮定を置いて連続回答と段階回答を対応づける必要がある。本発表では、特に連続反応と段階反応の「中点」が対応しているケースについて、閾値に対称制約を課し、連続的な反応分布から段階反応のモデルを構成する方法を提案する。

2. 提案モデル

以下、個人 $i = 1, \dots, N$ が項目 $j = 1, \dots, J$ に回答する状況を考え、このモデルの詳細を述べる。回答が連続値、あるいは離散値であっても十分段階数が多く、連続値として扱える場合 (例えば 101 段階など) 「連続回答」と呼び、それ以外の場合「段階回答」と呼ぶことにする。

ここで主要な仮定として、連続回答と段階回答の「中点」が対応しているとする。例えば、意思決定課題では、事象が起きる確率や尤もらしさを判断させることがしばしば存在する。主観確率を 0 から 100% で問うような連続回答と、尤もらしさを「非常に尤もらしくない (= 1)」から「非常に尤もらしい (= 7)」の 7 段階で問うような段階回答を考えると、概ね「50%」と「どちらとも言えない (= 4)」が対応すると考えられる。ここでは、このように仮定できる状況を考える。

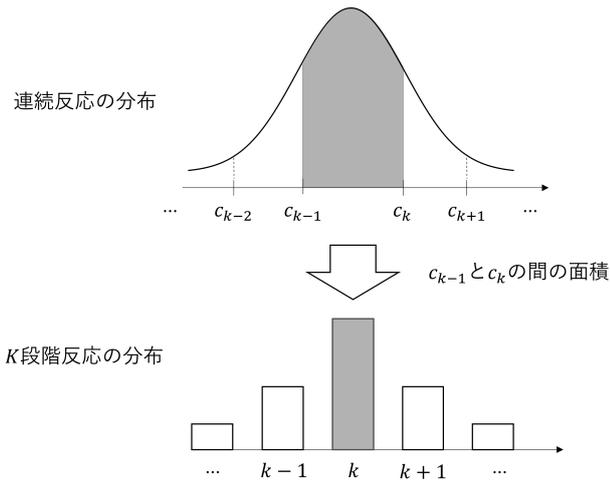
また、ここでは連続回答のモデルを用いて段階回答のモデルを構成することを説明するため、連続回答のモデルを既知として考える。すなわち、連続回答 y_{ij} の累積分布関数 $F_{ij}(y)$ が (何らかのパラメータを用いて) 定まっているとき、各項目 j ごとに閾値パラメータ

$$c_j = (c_{j,1}, \dots, c_{j,K-1}) \quad (1)$$

$$(c_{j,k} < c_{j,k+1}, k = 1, \dots, K-2)$$

を導入することによって、 $K (> 2)$ 段階回答のモデルを定めることができる。図 1 のように、個人 i が各項

Figure 1 提案モデルの図的表現



目 j における回答が段階 k_{ij} となる確率 $P(k_{ij}|c_j)$ は,

$$P(k_{ij}|c_j) = \begin{cases} F_{ij}(c_{j,1}) & (k_{ij} = 1) \\ F_{ij}(c_{j,k_{ij}}) - F_{ij}(c_{j,k_{ij}-1}) & (1 < k_{ij} < K) \\ 1 - F_{ij}(c_{j,K-1}) & (k_{ij} = K) \end{cases} \quad (2)$$

と定義される。

ただし、これだけでは c_j の任意性が大きく、 $F_{ij}(y)$ のパラメータと同時に推定しようとすると識別不能になる場合が多い。そのため、

$$c_{jk} = \begin{cases} 0 & (K : \text{even}, k = K/2) \\ c_{j,K-k} & (\text{otherwise}) \end{cases} \quad (3)$$

という対称制約を置く。連続回答と段階回答の「中点」が対応していない限り、このような制約を置く正当性が保証されない¹。このような制約を置くことにより、閾値パラメータと $F_{ij}(y)$ のパラメータを同時に推定することが可能になる。

3. 実データへの適用

前節で述べた手法を用いて、潜在スコープバイアスの大きさを分析した。潜在スコープバイアス (latent scope bias) は、因果的説明における重要な現象の一つであり、観測できない事象のうち、説明されているもの (の集合) が狭い説明の方が好まれる傾向のことを指す (Johnson et al., 2016; Khemlani et al., 2011)。ここでは潜在スコープバイアスを初めて指摘した論文であり、かつデータがオープンアクセスとなっている

Khemlani et al. (2011) の Exp.1b のデータのうち、実験条件のデータを用いてモデリングを実施した²。

Khemlani et al. (2011; Exp.1b) では、Amazon Mechanical Turk を通じて集められた 31 名の参加者に対し、2 要因参加者内要因計画で実験を行い、統制条件と実験条件それぞれ 4 項目の課題に対する回答を得た。なお、8 項目すべてを含めて提示順序はランダム化された。実験課題は、魔法によって生じた現象に関するものであった。各課題において、いくつかの現象と、それを説明する 2 つの魔法が提示された。実験条件では、一方の魔法からのみ予測される現象は観測できなかった。統制条件では、一方からのみ予測される現象は観測されており、2 つの魔法から共通して予測される現象が観測できないという状況であった。参加者は、これらの条件において、2 つの魔法それぞれがどの程度尤もらしいかを 7 段階のリッカート尺度 (+3 が very likely, -3 が very unlikely を表すものとした) により回答した。

このうち、実験条件の 2 つの魔法における尤もらしさの差が、潜在スコープバイアスを意味する。したがって、ここでは実験条件のデータのみを用いてモデリングした。実験条件の課題における因果構造を図 2 に示す。実験条件の各課題においては、観測できない現象を引き起こす魔法 H_W の尤もらしさ L_{H_W} と、観測できない現象を引き起こさない魔法 H_N の尤もらしさ L_{H_N} の 2 つが得られている ($-3 \leq L_{H_W}, L_{H_N} \leq 3$) ため、この 2 つをもとに、 H_W の、 H_N と比較した相対的な尤もらしさを

$$L_{H_W} - L_{H_N} + 7 \quad (4)$$

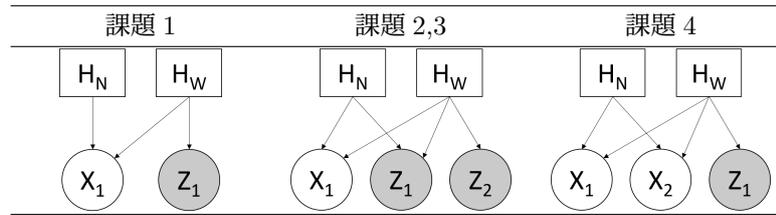
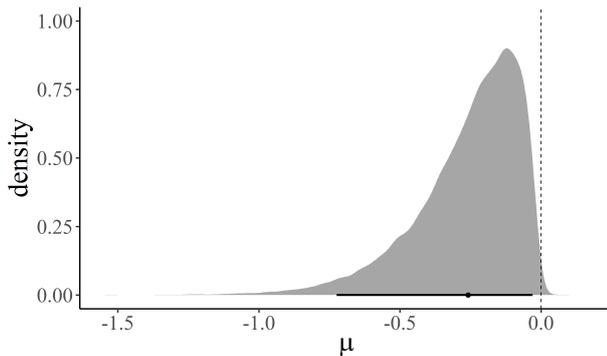
により求めた。これは 13 段階のデータであり、前項で定義したモデルを適用することができる。また、(4) 式で得られる値が小さいほど、大きな潜在スコープバイアスが生じていると言える。

そこで、前項で定義したモデルにおける $F_{ij}(y)$ を $\text{Normal}(\mu, \sigma^2)$ の累積分布関数としたものに基づいて、事前分布 $\mu \sim \text{Normal}(0, 5^2)$, $\sigma \sim \text{Normal}^+(0, 5^2)$ を用いたベイズ推定を行った。結果のうち、 μ の事後分布を図 3 に示す。特に、EAP 推定値は $\mu \approx -0.259$ となり、ロジスティック関数を用いて確率に直したところ、 $p \approx 0.436$ となった。規範解 $p = 0.5$ と比較すると、潜在スコープバイアスの効果量は大きくないと言える。

¹対応していれば必ずこのような制約が置けるわけではないが、正当化の根拠にはなりうる。

²この論文における最初の実験は Exp.1a だが、ここでは 2 肢選択課題を行っており、上記モデルの適用条件 $K > 2$ を満たさないと考えられた。

Figure 2 Khemlani et al. (2011, Exp.1b) の実験条件の各課題における因果構造のグラフ表現

Figure 3 μ の事後分布

Note: 点は EAP を, 水平の実線は 95%信用区間を, 垂直の破線は規範解である $\mu = 0$ をそれぞれ表す.

soning. *Memory & Cognition*, 39(3), 527–535. <http://doi.org/10.3758/s13421-010-0028-1>

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika*, 34(Suppl 1), 1–97. <https://doi.org/10.1007/BF03372160>

Samejima, F. (1973). Homogeneous case of the continuous response model. *Psychometrika*, 38(2), 203–219. <https://doi.org/10.1007/BF02291114>

4. まとめ

このように, 回答が離散値の場合でも, 連続値と見なせる場合であっても, 同じ枠組みで分析が可能となる. また, 段階回答データに対してモデルをフィッティングすることにより, リッカート尺度そのものよりも解釈しやすいパラメータを推定することが可能である.

本手法は, モデリングすることを想定せず取得されたデータに関するモデルベースでの分析を容易にするものであり, 実験データのオープン化が盛んになっている近年において, 二次分析の積極的な実施を促進すると考えられる. 今後は, 様々なデータおよびモデルに対する適用可能性を調べ, 本手法をより妥当かつ利用可能性の高いものに改善していくことが求められる.

文献

- Johnson, S. G. B., Rajeev-Kumar, G., & Keil, F. C. (2016). Sense-making under ignorance. *Cognitive Psychology*, 89, 39–70. <http://doi.org/10.1016/j.cogpsych.2016.06.004>
- Khemlani, S. S., Sussman, A. B., & Oppenheimer, D. M. (2011). Harry potter and the sorcerer's scope: Latent scope biases in explanatory rea-