

個人正答確率関数の推定精度

—項目数と受験者数を要因としたシミュレーション—

○登藤直弥*, 孫媛*, Yi Sun**

*国立情報学研究所, *University of China Academy of Science

問題と目的

ある受験者が同じ項目に繰り返し解答した場合の正答数の割合を個人正答確率という。この個人正答確率を項目の困難度の関数であると考え、この個人正答確率関数からは受験者の潜在特性値以外の情報が得られるようになる。したがって、学習者へのフィードバックや学習支援という観点からは、項目反応関数よりも有用であろうと考えられる。個人正答確率に関しては、主にモデルの適合度という観点から検討が行われてきており (Conijn, Emons, van Assen, & Sijtsma, 2011; Emons, 2003; Emons, Sijtsma, & Meijer, 2004, 2005; Nering, Bay, & Meijer, 2002; Nering & Meijer, 1998; Sijtsma & Meijer, 2001), 母数の推定精度という観点からの検討はあまり行われてきていない。

そこで本研究では、個人正答確率関数の推定精度に関して項目数と受験者数を要因とするシミュレーションを行い、特定の応用場面を想定した場合に項目数によってその推定精度がどのように変動するのか、検討を行った。

方法

個人正答確率関数に関しては、その定義・推定法が上述の先行研究において様々に提案されているのだが、本発表では、孫 (1998) や Sun, Todo, & Sun (2015) で提案された個人正答確率関数に基づきシミュレーションを行う。孫 (1998) や Sun et al. (2015) では、受験者*i*の項目*j*に対する正答確率の関数 P_{ij} が

$$P_{ij} = \frac{1}{1 + e^{-\alpha_i(\gamma_i - \beta_j)}} = \frac{1}{1 + e^{-\alpha_i((- \beta_j) - (- \gamma_i))}}$$

と表わされる。ここで、 α_i は受験者*i*に関する傾きを表しており、この値が正の方向に大きい場合、項目困難度が上がるにつれて受験者*i*の正答率は急激に低下することになる。また、 γ_i と β_j はそれぞれ受験者*i*の潜在特性値と項目*j*の困難度を表わしており、項目反応関数に含まれる母数と同様に解釈することが可能である。一方で、項目反応関数では、 P_{ij} が

$$P_{ij} = \frac{1}{1 + e^{-a_j(\theta_i - b_j)}}$$

と表わされる。ここで、 a_j と b_j はそれぞれ項目*j*の識別力と困難度を表しており、 θ_i は受験者*i*の潜在特性値を表している。したがって、受験者と項目をそれぞれ項目と受験者に見立てて項目反応

関数を推定し、得られた母数を

$$\begin{aligned}\gamma_i &= -b_j \\ \beta_j &= -\theta_i\end{aligned}$$

と変換することで、個人正答確率関数の母数を得ることができる。

本発表では、この関係を利用して項目数と受験者数を要因とするシミュレーションを行った。ここでは、その例として、1 クラス相当 30 人を分析の対象とした場合 ($N = 30$) の結果を報告する。本シミュレーションでは、 $N = 30$ という条件において α_i や γ_i , β_j を精度良く推定するためにはどの程度の項目数 (J) が必要となるのか調べるため、項目数を 10, 30, 50, 70, 100 と変化させて検討を行った。具体的には、各条件において P_{ij} の真値を固定した上で $N \times J$ の項目反応行列 U を乱数を利用して 100 個発生させ、この 100 個の U のそれぞれから α_i や γ_i , β_j を推定した。そして、これら 100 個の推定値 $\hat{\lambda}^r$ ($r = 1, \dots, 100$) を用いて、以下の式より、各推定量 $\hat{\lambda}$ のバイアス ($Bias(\hat{\lambda})$)、標準誤差 ($SE(\hat{\lambda})$) を算出し、これらの値を α_i , γ_i , そして β_j ごとに平均して、各母数の平均的な推定量のバイアス ($\overline{Bias(\hat{\lambda}_i)}$)、標準誤差 ($\overline{SE(\hat{\lambda}_i)}$) とみなすことにした (ただし、以下では、 λ が該当する母数の真値を、 $\bar{\lambda}^r$ が 100 個の推定値の平均値を、それぞれ表すものとする)。

$$\begin{aligned}Bias(\hat{\lambda}) &= \frac{1}{100} \sum_{r=1}^{100} \hat{\lambda}^r - \lambda \\ SE(\hat{\lambda}) &= \frac{1}{100} \sum_{r=1}^{100} (\hat{\lambda}^r - \bar{\lambda}^r)^2\end{aligned}$$

なお、本シミュレーションの実行にあたっては、R 3.1.2 (R Core Team, 2014), irtoys パッケージ (Partchev, 2014), IRT Command Language (Hanson, 2002) を使用した。

結果と考察

上述のシミュレーションを行った結果が以下の表 1 になる。

表 1. α_i , γ_i , そして β_j ごとのバイアス, 標準誤差の平均値 ($N = 30$)

項目数	$\overline{Bias(\hat{\alpha}_i)}$	$\overline{SE(\hat{\alpha}_i)}$	$\overline{Bias(\hat{\gamma}_i)}$	$\overline{SE(\hat{\gamma}_i)}$	$\overline{Bias(\hat{\beta}_j)}$	$\overline{SE(\hat{\beta}_j)}$
10	-0.105	0.375	-0.260	1.322	-0.033	0.674
30	-0.105	0.359	-0.322	0.778	-0.143	0.600
50	0.008	0.310	0.027	0.491	0.042	0.511
70	0.001	0.292	0.025	0.429	0.051	0.491
100	-0.027	0.260	-0.016	0.358	0.018	0.491

上記の結果からは、1 クラス 30 人程度を分析の対象とした場合、50 項目ほどのテストを実施すれば、平均的には、個人の特性値 (傾き, 潜在特性値) に関してバイアスなく推定を行うことができるといえる。また、本シミュレーションの結果からは、テストに含まれる項目数を増やすことにより、推定値のばらつきを抑えられることも明らかとなった。

謝辞

本研究は、JSPS 科研費 25280121 の助成を受けたものです。