Lordによる多肢選択テストの欠損値処理方法の偏りの補正

岡山大学 藤森進¹

Biss reduction of Lord's procedure of dealing with omitted responses in multiple choice tests

Susumu Fujimori (Laboratory of Educational Psychology, Faculty of Education, Okayama University, Tsushima-Naka, Okayama 700)

This article is concerned with determining the bias in Lord's (1974) procedure for estimating the ability parameter in item response theory model. Lord's procedure is one of the methods of dealing with omitted responses for dichotomously scored multiple choice test data. This study highlights the bias inherent in Lord's procedure and how to reduce it. Theoretical justification and a simulation study are presented for the estimates obtained by the bias reduction method of this study. The results of the simulation study show that the bias reduction method has a higher efficiency than Lord's procedure in mean squared error. This study also takes notice of the appropriateness of Lord's formula of asymptotic variance of estimates of ability parameter.

Key words: omitted response, item response theory, bias, multiple choice test.

項目反応理論 (Lord & Novick, 1968) は、テストデータの分析に用いられる理論である。テストでは被験者の能力値を推定する必要が常に生じる。その際にテストデータによく見られる無答の処理をいかに行うかという問題がある。項目反応理論では、この無答を能力母数の推定のデータから省いたり、誤答扱いにするなどの処理方法が採用されることが多い。特に多肢選択形式のテストデータにおいて、Lord (1974) の方法 (後述) によって無答が処理されることがある。この研究は、Lord の無答処理による能力推定値の偏りの検討と、その軽減方法、および Lord により導出されている漸近分散の推定上の問題点について論じるものである。

項目反応理論における被験者能力の推定

項目反応モデルにはいくつかの種類がある。本研究では、その中から被験者の項目に対する正答確率を式(1)で表す3母数ロジスティックモデルを取り上げる。

$$P_j(x|\theta) = c_j + (1 - c_j) \frac{1}{1 + \exp(-Da_j(\theta - b_j))} \quad (1)$$

ここで j は項目番号、 $D=1.7$ の定数である。 a_j は項目の識別力、 b_j は項目の困難度、 c_j は偶然的正答確率を表す項目母数である。 θ は被験者の能力を表す母数である (以下能力母数と呼ぶ)。 x は正答のとき1、誤答のとき0となるダミー変数である。

この研究で3母数ロジスティックモデルを取り上げた理由は、同モデルがあて推量をモデル化していることによる。研究対象とする多肢選択形式のテストデータでは、正答を知らなくてもあて推量によって偶然に正答する現象が見られるからである。もっとも、 c_j を0とすれば本研究で述べる結果は、そのまま2母数ロジスティックモデルでも適用できると考えられる。

さて、テストのすべての項目母数が所与であり、無答による欠損値がないとき、能力母数 θ を推定するための尤度は式(2)である。

$$L = \prod_{j=1}^n P_j^{x_j} Q_j^{1-x_j} \quad (2)$$

ここで n はテストの項目数、 P_j は式(1)の正答確率、 $Q_j = 1 - P_j$ である。能力母数の最尤推定値は式(2)を θ の関数と見たとき、式(2)の最大値を与える $\hat{\theta}$ である。通常は式(2)そのものではなく、その対数をとった対数尤度式(3)を最大にする $\hat{\theta}$ を求めることが多い。

$$l = \sum_{j=1}^n \{x_j \log(P_j) + (1 - x_j) \log(Q_j)\} \quad (3)$$

¹ 本研究において語彙理解力検査の利用をお許し頂いた東京大学名誉教授芝祐順先生に心から感謝致します。また本研究の費用の一部の補助を福武教育振興財団より受けました。担当の白石光史さんに心から感謝致します。

式(3)を θ に関して微分し 0 とすることにより最尤推定値 $\hat{\theta}$ を得るための対数尤度方程式(4)を得る。

$$\frac{\partial l}{\partial \theta} = \frac{\partial}{\partial \theta} \sum_{j=1}^n \{x_j \log(P_j) + (1-x_j) \log(Q_j)\} = 0 \quad (4)$$

能力母数の最尤推定値は、この式(4)の解として与えられる $\hat{\theta}$ である。

項目反応理論における無答処理の問題

テストで正答がわからないとき解答することは、恥をかくだけだと被験者は思うかもしれない。その場合被験者は無答する可能性がある。実際に被験者が無答するかどうかはさまざまな条件が関与する。例えば、被験者が無答を採点から外すという情報を知ったとしたらどうなるだろう。被験者は、確実に正答がわかっているものだけに解答し残りは無答とすることにより、望むがままの高い能力推定値を得ることになるだろう (Lord, 1980)。

たとえ被験者が採点方法を知らない場合でも、無答を採点しなければどうなるだろう。多肢選択形式のテストであれば、無答にしないであて推量により解答することもできる。全くわからない項目に対して、あて推量で解答する被験者と、あて推量せずに無答にしてしまう被験者が生じらるだろう。この場合、両者の真の能力が等しくても無答の処理方法により異なった得点となってしまう。項目反応理論では、慎重な後者の方が高い推定値を得ることになるだろう。

このような観点からは、無答を採点の対象から外すことは適当ではなさそうである。すなわち無答にもペナルティが必要ということになる。この最も強硬な立場が無答を誤答として扱う方法である。この方法は、必然的に被験者に、あて推量などの無理な解答を強いることになる。もう少し穏やかな立場に、無答に適当な重み付けをして能力母数の推定に利用しようとする考えがあり、Lord の方法はこの一種である。彼は、無答に対する重みとして、被験者が多肢選択形式のテスト項目にランダムな反応をしたと想定した場合の重みを与えることを提案している。すなわち当該項目の選択肢数分の 1 の重みを無答に与えるものである。次節でその概要を述べよう。

Lord の無答の処理方法の概要

Lord (1974) は、被験者の能力の推定に利用する尤度関数において、無答を選択肢数分の 1 に置き換えることによって無答を処理している。

テストの s 項目が無答であるとするとき、Lord は、式(2)の代わりに式(5)の疑似的な尤度を考える。

$$L^* = \prod_{j=1}^s P_j^{k_j} Q_j^{1-k_j} \prod_{j=s+1}^n P_j^{x_j} Q_j^{1-x_j} \quad (5)$$

ここで k_j は、被験者が項目 j に無答したことに對する重みである。Lord は、この重みについて偶然正

答する確率をあてるべきだと考え、多肢選択データでは $k_j = 1/\text{選択肢数}$ とした。このような重みを与えた式(5)は通常の意味での尤度とはならず疑似的な尤度となる。Lord は、この疑似的な尤度と、その s 個の無答の重み k_j に選択肢数分の 1 の確率で正答 1、残余の確率で誤答 0 を埋め込んだ式(6)の尤度との比較を行っている。

$$L = \prod_{j=1}^s P_j^{x_j} Q_j^{1-x_j} \prod_{j=s+1}^n P_j^{x_j} Q_j^{1-x_j} \quad (6)$$

式(5)の対数をとって式(7)を得る。

$$l^* = \sum_{j=1}^s \{k_j \log(P_j) + (1-k_j) \log(Q_j)\} + \sum_{j=s+1}^n \{x_j \log(P_j) + (1-x_j) \log(Q_j)\} \quad (7)$$

続いて式(6)の対数 l と式(7)の差をとって式(8)を得る。

$$l^* - l = \sum_{j=1}^s \{(k_j - x_j) \log(P_j) + (1 - k_j - 1 + x_j) \log(Q_j)\} = \sum_{j=1}^s (k_j - x_j) \log\left(\frac{P_j}{Q_j}\right) \quad (8)$$

次に式(8)の両辺をテストの項目数 n で割った式(9)を考える。

$$\frac{l^*}{n} - \frac{l}{n} = \frac{1}{n} \sum_{j=1}^s (k_j - x_j) \log\left(\frac{P_j}{Q_j}\right) \quad (9)$$

式(9)の右辺は、Lord によると $s \rightarrow \infty$ で 0 に近づく。続いて式(8)を θ に関して微分した式(10)の疑似的な対数尤度方程式を考える。

$$\frac{\partial l^*}{\partial \theta} = \frac{\partial l}{\partial \theta} + \frac{\partial}{\partial \theta} \left\{ \sum_{j=1}^s (k_j - x_j) \log\left(\frac{P_j}{Q_j}\right) \right\} = 0 \quad (10)$$

Lord の方法の θ の推定値は式(10)の解として与えられる θ^* である。

式(10)を n で割って、 $\{ \}$ 内の微分を行うと式(11)を得る。

$$\begin{aligned} \frac{1}{n} \frac{\partial l^*}{\partial \theta} &= \frac{1}{n} \frac{\partial l}{\partial \theta} + \frac{1}{n} \sum_{j=1}^s \frac{\partial P_j}{\partial \theta} \left(\frac{k_j - x_j}{P_j Q_j} \right) \\ &= \frac{1}{n} \left[\left\{ \sum_{j=1}^s \frac{\partial P_j}{\partial \theta} \left(\frac{k_j - x_j}{P_j Q_j} \right) \right\} + \left\{ \sum_{j=s+1}^n \frac{\partial P_j}{\partial \theta} \frac{x_j - P_j}{P_j Q_j} \right\} \right] \\ &= 0 \end{aligned} \quad (11)$$

式(6)の尤度の対数を θ_0^* でテイラー展開し式(11)に代入して θ を疑似的な尤度方程式の解 θ^* で置き換えることにより次式を得る。

$$\begin{aligned} \frac{1}{n} \frac{\partial l^*}{\partial \theta} \Big|_{\theta_0^*} &= \frac{1}{n} \frac{\partial l}{\partial \theta} \Big|_{\theta_0^*} + \frac{1}{n} \frac{\partial^2 l}{\partial \theta^2} \Big|_{\theta_0^*} (\theta^* - \theta_0^*) \\ &\quad + \frac{1}{2n} \frac{\partial^3 l}{\partial \theta^3} \Big|_{\theta_1^*} (\theta^* - \theta_0^*)^2 \\ &\quad + \frac{1}{n} \sum_{j=1}^s \frac{\partial P_j}{\partial \theta} \left(\frac{k_j - x_j}{P_j Q_j} \right) \\ &= 0 \end{aligned} \quad (12)$$

ここで θ_1^* は θ_0^* と θ^* の間にある。式(12)の各成分を以下のように置く。

$$LA = \frac{\partial I^*}{\partial \theta} \Big|_{\theta^*} \quad (13)$$

$$A = \frac{\partial I}{\partial \theta} \Big|_{\theta_0^*} \quad (14)$$

$$B = \frac{\partial^2 I}{\partial \theta^2} \Big|_{\theta_0^*} \quad (15)$$

$$C = \frac{1}{2} \frac{\partial^3 I}{\partial \theta^3} \Big|_{\theta_1} \quad (16)$$

$$Z = \sum_{j=1}^s \frac{\partial P_j}{\partial \theta} \left(\frac{k_j - x_j}{P_j Q_j} \right) \quad (17)$$

式(12)に $n^{3/2}$ を乗じて式(13)～式(17)を代入すると式(18)を得る。

$$\begin{aligned} \sqrt{n} LA &= \sqrt{n} A + \sqrt{n} B (\theta^* - \theta_0^*) \\ &\quad + \sqrt{n} C (\theta^* - \theta_0^*)^2 + \sqrt{n} Z \\ &= 0 \end{aligned} \quad (18)$$

これを整理することにより式(19)を得る。

$$\sqrt{n} (\theta^* - \theta_0^*) = \frac{\frac{1}{\sqrt{n}} (-A - Z)}{\frac{1}{n} B + \frac{1}{n} C (\theta^* - \theta_0^*)} \quad (19)$$

$n \rightarrow \infty$ のとき、 $\theta^* \rightarrow \theta_0^*$ となり式(19)の右辺分母第2項は、確率1で0に収束する。また式(19)の右辺分母第1項は、その期待値が次式で与えられる。

$$E \left(\frac{1}{n} B \right) = \frac{1}{n} E \left(\frac{\partial^2 I}{\partial \theta^2} \Big|_{\theta_0^*} \right) = - \frac{I(\theta_0^*)}{n} \quad (20)$$

ここで $I(\theta)$ は次式のテスト情報関数である。

$$I(\theta) = \sum_{j=1}^n D^2 \alpha_j^2 \frac{Q_j}{P_j} \left(\frac{P_j - c_j}{1 - c_j} \right)^2 \quad (21)$$

以上により Lord は、 θ^* の漸近分散が次式で与えられるとしている。

$$V(\theta^*) = V(\hat{\theta}) \{1 - k_j(1 - k_j) V(\hat{\theta}) S\} \quad (22)$$

ここで $V(\hat{\theta})$ は、確率的に正誤データを無答に埋め込んだ式(6)の尤度の最尤推定値 $\hat{\theta}$ の漸近分散であり、テスト情報関数の逆数に等しい。また S は、式(23)で表される量である。

$$S = E \left(\sum_{j=1}^s \frac{1}{P_j^2 Q_j^2} \left(\frac{\partial P_j}{\partial \theta} \right)^2 \right) = \sum_{j=1}^n \frac{1}{P_j^2 Q_j^2} \left(\frac{\partial P_j}{\partial \theta} \right)^2 w_j \quad (23)$$

ここで w_j は被験者が項目 j に無答する確率である。Lord は、式(22)の $V(\theta^*)$ は、0, 1 を埋め込んだ尤度から得られる $V(\hat{\theta})$ より小さいので良い推定値が得られたとしている。

ところで、式(22)の漸近分散の導出にあたって、Lord (1974) の Appendix の式 (A13) では、 s 個の無答に 1 または 0 を埋め込んだ項目についても $E(x - P_j)^2 = P_j Q_j$ 、つまり $E(x) = P_j$ としている。その後の式 (A17) では $E(x - k_j)^2 = k_j(1 - k_j)$ 、つまり $E(x) = k_j$ としている。この両式が同時に成立することは困難なのではないだろうか。もちろん、たまたま $P_j = k_j$ であれば両立するが、一般に、これを想定するのは困難であろう。もし両立しない事態では、 θ^* の漸近分散式(22)中の $V(\hat{\theta})$ が、通常のテスト情報関数の逆数の形にならないので注意が必要となる。

Lordの方法の偏りとその除去方法

続いてここでは、Lordの方法がもつ偏りの問題とその除去方法について述べよう。

Lordは、項目反応モデルの式(1)の正答確率について被験者が当該項目と同じ項目特性曲線をもつ項目に正答できる確率であると言っている。あるいは当該の項目に同じ能力母数の人からランダムに選ばれた被験者が正答できる確率であると言っている。しかし彼は1974年の論文で扱うのは、ある特定の被験者が特定の項目に正答する確率であり、通常の項目反応モデルの正答確率の解釈は適当でないとして議論を展開している。

無答に対する正答確率を k_j とする Lord の見解は、結局、無答は無知であるという見解にほかならない。その点からは、無答は誤答とするのが自然であろう。Lordの方法の狙いは、テストで無答しないであて推量した者と無答とした者の採点上の公平さを期すために前述のような処理を行ったと考えられる。無答を誤答とすると、あて推量した者が偶然に正答した部分で得をしてしまうからである。両者の公平を図るという点はよいのであるが、その反面で、Lordの処理は能力母数の過小推定、あるいは過大推定になっているのではないかという疑問がある。

Lordの方法では $n \rightarrow \infty$ のとき、 $\theta^* \rightarrow \theta_0^*$ であり(漸近的に一致推定量)、そのとき θ^* の期待値は θ_0^* となり不偏となる。しかし Lord は無答に対して確率的に正誤データを埋め込んだ実際には存在しない疑似尤度式(6)の θ_0^* と比較していることに注意したい。Lordの方法の不偏性は、この θ_0^* に対して成立するだけである。だがこの θ_0^* がどのような特徴をもつものなのかが明確でない。

そこで、不偏性の対象となった 0, 1 データを埋め込んだ尤度を検討してみよう。被験者が無答した項目に選択肢数分の 1 で正答 1 を埋め込むことは、結局のところ、当該被験者がその項目に確率 k_j で正答すると考えることになる。すなわち式(6)の s 個の無答した項目の正答確率が k_j であるということになるだろう。だが、その確率は、項目反応モデルに従うものではない。項目反応モデルでは、 θ が所与のときに、正答あるいは誤答が生じる確率を記述している。無答もその延長上で記述されるのが自然であろう。しかし、Lordの処理では、どのような θ の被験者がどのような項目に無答しても正答確率は k_j とされてしまう。従来の項目反応モデルの枠組みで Lordの方法をとることが、この点において困難になっているのである。

このような問題は、重み k_j を無答に与えた疑似尤度と、0, 1 データを埋め込んだ尤度を比較することによって生じている。その代わりに本研究では式(2)の

尤度と式(5)の比較をしてみよう。そうするとLordの方法の概要と全く同様のプロセスで θ^* の期待値が式(24)で与えられる。但しLordの方法の概要と異なり θ_0 は真の値である。

$$\begin{aligned} E(\theta^*) &\doteq \theta_0 + \frac{\frac{1}{n}E(-Z)}{-\frac{I(\theta_0)}{n}} \\ &= \theta_0 + \frac{E(Z)}{I(\theta_0)} \\ &= \theta_0 + \frac{1}{I(\theta_0)} \sum_{j=1}^s \frac{\partial P}{\partial \theta} \bigg|_{\theta_0} \left(\frac{k_j - P_0}{P_0 Q_0} \right) \quad (24) \end{aligned}$$

さて、たまたま

$$P(\theta_0) = k_j \quad (25)$$

でない限り式(24)の右辺第2項は0とならない。それ故に一般にはLordの方法による推定値 θ^* は、真の値 θ_0 に対して不偏とならないと考えられる。あるいはLordの方法が適切であるのは、式(25)が成立するときの無答に限ると言ってもよい。しかし、現実には式(25)が成立するときだけ無答が発生するとは期待できない。このため、一般的にLordの方法によって無答を処理することは、 $\hat{\theta}$ に偏りをもたらすと考えてよいだろう。無答項目の k_j (=選択肢数分の1)と式(1)の正答確率の大小により、偏りは正負いずれにもなり得る。この偏りを除去するには

$$\theta_x^* = \theta^* - \frac{1}{I(\theta_0)} \sum_{j=1}^s \frac{\partial P}{\partial \theta} \bigg|_{\theta_0} \left(\frac{k_j - P_0}{P_0 Q_0} \right) \quad (26)$$

とすればよいわけである。本研究が提案するLordの方法の偏りの改善は式(26)の θ_x^* によって能力母数の推定を行おうとするものである。

では、次にLordの方法および本研究で提案する偏りの除去方法をシミュレーションによって検討してみよう。

シミュレーションデータの作成と分析の方法

シミュレーションデータの作成は以下のようにして行った。まず被験者の能力母数の真値は、-2から2までを0.5刻みして、検討する能力母数 θ の対象とした。テスト項目の困難度母数の分布は、-2から2までの一様分布に従うとする。識別力母数は、平均0.85、標準偏差0.25とする下限0.3、上限2.0の切断正規分布に従うとした。また偶然正答確率の母数はすべて0.2で固定した。偶然正答確率の母数についても何らかの分布に従うと考えてシミュレーションデータを作成することももちろん可能である。しかし、偶然正答確率の母数の値と、Lordが述べるところの選択肢のランダムな選択による正答確率の値が同じであると理解しやすくなる。このためテスト項目の選択肢数を一般によく利用される5として、偶然正答確率の母数の値を0.2にそろえることにした。

このような項目母数を、各能力母数の真値の点ごと

に80項目分作成した。この後、能力母数の真の θ と各項目母数の組合せによって決まる3母数ロジスティックモデルの式(1)の正答確率 P_j を計算する。そして P_j と、0から1の一様乱数 R を比較し、 R が P_j より小さいとき正答1、大きいとき誤答0を80項目分発生させた。この80項目の正誤パターンを1ケース(被験者)とする。

次に無答の発生率 P_{omit} は、 $0.3 \times (1 - P_j)$ とした。そして、0から1の一様乱数を発生させ、 P_{omit} より小さい乱数値を得たときに当該項目に対する被験者の反応を無答と置き換えた。誤答確率 $(1 - P_j)$ に比例する形で P_{omit} を定めたのは、無答の発生が被験者の能力より項目の困難度が大きい場合に無答が発生しやすいという実態を考慮したためである。なお、 $0.3 \times (1 - P_j)$ という P_{omit} の大きさは、通常は無答の発生率としてはやや高めのものであるかもしれない。しかし、 P_{omit} の大きさは、 θ が大きくなるにつれて小さくなる。このため、 θ が大きくなるとシミュレーションで無答がほとんど生じなくなる。このこと自体は現実ともよく当てはまり自然である。しかし、たとえ少数ではあっても θ の大きい場合にも無答が生じるのであり、その無答処理の影響を明確にするには、ある程度の無答の発生も必要と考え P_{omit} の大きさを本研究のように設定した。以上のような手続きで無答を含む正誤パターンを θ の各真値ごとに500ケース作成した。

また、テスト項目が5肢選択でできているとして、前記の方法で作成したテストデータの無答をLordの考えに従って0あるいは1で置き換えたデータを作成した。

このようにして得られたデータに関して、次の5種類の能力母数の推定値を求め分析を行った。

1. 無答に、正答1を確率0.2、誤答0を確率0.8で埋め込んだデータから得られた最尤推定値
2. 無答をLordの方法で処理して得られた推定値
3. 無答をLordの方法で処理し、更に本研究の提案する偏りの補正を行って得られた推定値
4. 無答を推定から除外し、残余の正誤データから得られた最尤推定値
5. 無答をすべて誤答として扱い、得られた最尤推定値

シミュレーションの結果と考察

Table 1は、0あるいは1が無答の代わりに埋め込まれたデータに関する結果である。表中には能力母数の推定値を通常的最尤推定法により求めた結果の平均、分散、および平均2乗誤差(MSE)を示してある。Table 1より明らかに、いずれの真の θ に関しても、無答に0あるいは1を埋め込んだデータの最尤推定値に偏りが存在することがわかる。偏りがすべて負

Table 1
無答に 0, 1 を埋め込んだデータの最尤推定値 $\hat{\theta}$ の平均, 分散, MSE

真値 θ	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0
平均	-2.241	-1.701	-1.147	-.654	-.164	.338	.858	1.358	1.878
分散	.231	.156	.077	.060	.063	.067	.057	.078	.097
MSE	.289	.197	.098	.083	.090	.094	.077	.098	.112

Table 2
無答を含むデータの Lord の方法による θ^* の平均, 分散, および MSE

真値 θ	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0
平均	-2.254	-1.710	-1.148	-.654	-.164	.335	.853	1.351	1.854
分散	.219	.134	.070	.055	.058	.061	.053	.073	.088
MSE	.283	.178	.092	.079	.085	.088	.074	.095	.110

Table 3
無答を含むデータの Lord の方法の偏りを補正した θ_a^* の平均, 分散, および MSE

真値 θ	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0
平均	-2.089	-1.564	-1.000	-.520	-.024	.478	.996	1.503	2.013
分散	.224	.138	.077	.064	.065	.072	.060	.091	.112
MSE	.232	.142	.077	.064	.066	.072	.060	.091	.112

Table 4
無答を除外したデータの最尤推定値 $\hat{\theta}$ の平均, 分散, および MSE

真値 θ	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0
平均	-2.080	-1.542	-.977	-.498	-.002	.499	1.018	1.528	2.046
分散	.261	.146	.081	.068	.069	.076	.063	.100	.119
MSE	.267	.148	.082	.068	.069	.076	.063	.101	.121

になっているのは, 各項目とも $k_j < P_j$ としたので式 (24) の右辺第 2 項が必ず負になるためである。

続いて Table 2 を見てみよう。Table 2 は Table 1 と同じデータ (但し無答は無答のまま) に対し, Lord の方法によって能力母数の推定値を求めたものである。Table 1, 2 より, Lord の無答処理方法による推定値が, 無答に 0 あるいは 1 を埋め込んだ尤度の最尤推定値とほぼ一致していることがわかる。

また無答に 0 あるいは 1 を埋め込んだ尤度に基づく能力母数の推定値の分散に比較して Lord の方法による分散が, θ のどの真値の場合でも小さいことが示されている。この点は, Lord が述べているとおりである。最小 2 乗誤差 (MSE) で比較すると, やはり Lord の方法による推定値が, 無答に 0 あるいは 1 を埋め込んだ尤度の最尤推定値よりよいことが示されている。

次に Lord の方法による推定値に対して本研究の補正を行った Table 3 を見てみよう。Table 2 と Table 3 の比較より, 偏りを補正した効果が明らかである。能力母数の真値がいずれの場合も, 補正した推定値の平

均は, 真値とほぼ同じ値になっている。また補正した推定値の分散は, Lord の方法よりやや大きくなっているものの, MSE の観点においては偏りの補正を行った方がよいことが示唆される。

続いてデータの無答を無視し, 残余の正答あるいは誤答のみのデータから得られた能力母数の最尤推定値を見てみよう (Table 4)。Table 4 より明らかに推定値の平均は真値に近いものとなっていることがわかる。能力母数の推定値の平均と真値との差だけで考えれば Lord の方法で得た推定値の偏りを補正した Table 3 の結果とほとんど同じ成績を示すと言えるだろう。推定値の分散の大きさに関しては, Table 2 あるいは Table 3 の結果の方が Table 4 より小さくなっている。MSE の観点からは Table 4 の結果は, Table 2 の Lord の方法よりはよく, Table 3 の Lord の方法の補正值よりは劣ると言える。

最後に, データの無答をすべて誤答としたデータから得られた能力母数の最尤推定値を見てみよう (Table 5)。Table 5 では, Table 2 の Lord の方法の場合よりも更に偏りが大きくなっていることが示されて

Table 5
無答を誤答としたデータの最尤推定値 $\hat{\theta}$ の平均、分散、および MSE

真値 θ	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0
平均	-2.487	-1.892	-1.270	-.743	-.253	.249	.774	1.292	1.822
分散	.203	.166	.075	.056	.059	.061	.055	.072	.092
MSE	.440	.319	.148	.115	.122	.123	.107	.115	.124

いる。つまり、少なくともこのシミュレーションでは、無答を誤答とすることは明らかに推定の成績を悪化させていることがわかる。

芝の語彙理解力検査データを用いた分析

続いて実際の多肢選択形成のテストデータの分析により無答処理の各方法を検討してみよう。ここでは、藤森・中野（1994）で取り上げられた芝の語彙理解力検査（芝，1978）の5肢選択形式のU2版のデータの一部を分析対象とする。藤森・中野では、テストの教示として、採点方法について言及しない“中立的教示”，誤答数を正答数から差し引くという“修正教示”，誤答数には関係なく正答数で採点するという“正答数教示”の3種類を取り上げている。修正教示を受けた群のデータには、大量の無答が発生した（無答数の平均24.3，標準偏差14.7）。中立教示群では少々無答が発生した（無答数の平均2.8，標準偏差7.8）。正答数教示条件では、無答がほとんど発生しない（無答数の平均0.9，標準偏差3.0）ためここでは取り上げない。

本研究では、中立的教示と修正教示の2群の中で、1問以上無答のあったケースを分析対象とした。これは、無答が存在しなければどの方法も同じ推定値となるので見かけ上、類似性が高まることを避けたものである。更に20問以上の有効データ（正答あるいは誤答）が存在することも必要とした。これは修正教示群では、無答が異常に多いことがあるためである。無答が多くなれば有効データが少なくなり、必然的にテスト情報関数が小さくなるため推定値の漸近的性質が満たされないという問題が発生する。経験的には20項目程度テストデータが存在すれば能力母数の推定に問題がないことが多いため、このような処理を行った。

さて両群の被験者は、入学年度などにばらつきはあるものの、〇大学教育学部の“教育心理学”などの受講生であり、能力差があるとは考えにくい。もちろん、たまたま両群の能力に差が生じてしまうという可能性はないわけではないが事前にそのように考えるべき理由はない。

このため両群について、Lordの方法などにより能力母数の推定値を求めて中立的教示群と修正教示群の平均値の差の t 検定を行い、2群の差を見ることにした。但し、能力母数の推定値を求める際に θ の推定

Table 6
修正教示群と中立的教示群の能力母数の各推定値の t 検定

	修正教示群の 平均(SD)	中立的教示群の 平均(SD)	t 値
無答に0, 1を埋め込んだ $\hat{\theta}$	3.58(.66) N=169	3.81(.63) N=114	-2.83*
Lordの方法による θ^*	3.57(.66) N=173	3.80(.62) N=114	-2.92*
Lordの偏りを補正した θ_{adj}^*	3.76(.68) N=172	3.86(.64) N=114	-1.22
無答を除外した $\hat{\theta}$	3.89(.72) N=171	3.89(.70) N=114	.0
無答を誤答とした $\hat{\theta}$	3.25(.71) N=157	3.71(.65) N=113	-4.16*

*5%水準で有意差あり

値が大学生の能力母数の分布から大きく外れるものは分析から除外した。このような推定値が生じてしまう原因の一つは、当該被験者の有効なデータが無答ばかりで少ないことによる。他の理由には、有効データがあっても正答か誤答どちらかに偏っており、適当な最尤推定値が得られないということがある。U2版の項目母数の推定に利用した大学生の $\hat{\theta}$ の平均は3.9，標準偏差0.94であった。このため、本研究では、分析対象とする $\hat{\theta}$ の下限を2.0，上限を6.0と設定した（上限にはプログラムの制限という事情もある）。

このようなプロセスによって得られた結果がTable 6である。Table 6よりわかるように、有意でない t 値を得ているのは、無答を無視して能力母数の推定をした場合とLordの推定値の偏りを補正した θ_{adj}^* である。すなわち、この二つの処理方法では、両群の能力母数の推定値の母平均に差があるとは言えない。もともと二つの群には、教示の違いのほかに差があるとは考えにくいことから、この二つの処理方法がよさそうなことが支持される。なお、中立的教示群で無答が一つもなかったため分析対象から外した被験者253人の $\hat{\theta}$ の平均は3.82，標準偏差は0.68であった。この値は、Table 6の無答を無視して能力母数の推定をした結果とほぼ同じである。

Table 6で最も両群の差が大きくなったのは、無答を誤答とする処理方法である。二つの群間に差はない

と考えられる以上、無答を誤答とみなすことは、無答の処理方法としては適当とは考え難いだろう。また、無答に0, 1を埋め込んで $\hat{\theta}$ を求めた結果とLordの方法による結果は、予期されたとおり、かなり類似している。どちらの結果も、無答を誤答とする処理方法ほどの大きさではないものの両群間に有意な差が認められる。

以上の結果から、多肢選択形式のテストデータの無答をLordの方法で処理するときには本研究の示したように偏りを補正した方がよいことが示唆される。

引用文献

藤森 進・中野和代 1994 テストの教示が被験者の

解答行動と成績に与える影響 教育心理学研究, **42**, 455-462.

Lord, F. M. 1974 Estimation of latent ability and item parameters when there are omitted responses. *Psychometrika*, **39**, 247-264.

Lord, F. M. 1980 *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lord, F. M., & Novick, M. R. 1968 *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

芝 祐順 1978 語彙理解力尺度作成の試み 東京大学教育学部紀要, **17**, 47-58.

——1995. 12. 15. 受稿, 1996. 8. 31. 受理——