

【個人研究】

項目反応理論における無答処理とニューラルネットワークによる能力母数の推定

藤 森 進*

Treatment of Omitted Responses in Item Response Theory and Estimation of Ability Parameters in a Neural Network

Susumu FUJIMORI

In the item response theory, if omitted responses (no answers) are detected in the analysis data of a test, the no answers must be handled appropriately. The omitted responses are usually processed by "ignoring them" or "handling them as wrong answers," or "handling them according to the Lord Method" if the data is composed of answers from a multiple choice test. In this study, some test data sets were simulated, and they were analyzed using the no-answer handling method proposed by Fujimori (1994). The analysis results aren't overall if the ability parameters are estimated using a neural network. Findings suggest that the method may be more appropriate for a long test, but that it isn't useful for a short test. In addition, this study also includes an analysis of no-answer processing by the method when the ability parameters are estimated by the maximum likelihood method. Results from this analysis show that the method using the maximum likelihood estimates is effective for any test length.

Key words: item response theory, omitted response, neural network

1. 序

項目反応理論を適用してテストデータを分析する場合に、欠損値がデータの中にあると何らかの方法でこれを処理しなければならない

い。本研究では後に(1・2節)欠損値と無答を区別し、研究が直接対象とするのは無答であるが、ここでは欠損値と表現しておく。さて、被験者の能力を知るためには、その能力を表す能力母数を推定しなければならない。最も一般的な推定方法である最尤推定でも欠

*ふじもり すすむ 文教大学人間科学部人間科学科

損値をどのように処理するかは常に問題となるが、藤森(2000a)で取り上げられたニューラルネットワークによる推定でも、どのようにこれを処理するかは実用上大きな問題となる。ニューラルネットワークでは、項目反応理論の最尤推定でしばしば行われている方法、すなわち欠損値を生じた項目が初めから存在しなかったように無視するということが構造上行きにくい。しかし欠損値を適宜部分得点として補完すれば、ニューラルネットワークで処理することが可能であろう。項目反応理論で部分得点を能力母数の推定に利用する場合は、能力母数の推定値に関して、異なった部分得点を与えた場合の順序性は最低限満足される必要があると思われるが、藤森(2000b)によれば学習が十分であればニューラルネットワークによる能力母数の推定は、この条件を満足している。そこで本研究では、藤森(1994)で取り上げられた、被験者の欠損値以外の有効な項目を利用して当該項目の正答確率を求めて部分得点とする欠損値処理方法の有効性を、ニューラルネットワークによる能力母数の推定についても検討する。

1.1 項目反応理論

項目反応理論は、テスト項目への正答確率を、被験者の能力を表す母数 θ_i と、項目の特徴を表す項目母数によって表現しようとするものである。項目反応モデルには幾つかのものがあるが本研究では、次の2母数ロジスティックモデル(Birnbaum, 1968)をとりあげる(項目反応理論についてはLord & Novick(1968)、芝(1991)などを参照のこと)。

$$P_j(x_{ij} = 1 | \theta_i) = \frac{1}{1 + \exp(-Da_i(\theta_i - b_j))} \quad (1)$$

ここで j は項目番号、 i は被験者、 θ_i は被験者の能力母数、 $D = 1.7$ 、 a は項目の識別力、 b は項目の困難度を表す母数である。また x は正答のとき 1、誤答のとき 0 であり、(1)式は能力 θ_i の被験者が項目 j に正答する

確率を表す。また誤答する確率を

$$Q = 1 - P \quad (2)$$

で定義する。ただし P は(1)式を表す。被験者の能力母数の最尤推定値は

$$L = \prod_{j=1}^n P_j^{x_j} Q_j^{1-x_j} \quad (3)$$

を最大にする θ_i として得られる。ここで n はテスト項目数である。被験者の全ての解答が与えられている場合は、最尤推定により能力母数の推定値を求めることが多い。

1.2 欠損値の色々

まず初めにここでは、どんな事情にせよデータに欠けた部分があるととき欠損値と一般に呼ぶことにしよう。これに対して無答とは被験者が解答(質問紙では回答)しなかったことを意味することにしよう。データに欠損値が発生する原因には様々なものがある。たとえば被験者が項目に反応しない場合もあれば、二重解答による欠損、データ入力に伴う欠損など様々なものがある。データとして存在しないという意味ではいずれも同じであるが、その発生の原因により、また測定の方法によりどの処理方法が適しているかも異なってくる可能性がある。またデータが質問紙であるか学力検査であるかによっても異なる。本研究が対象とする欠損値がどのようなものなのかを明確にするため、まず学力検査データの場合を考えてみよう。

表1では、問題に接したが解答しなかった場合を ~ の無答、問題に接しなくて解答できなかったものを の未解答、被験者の学力に係わり無く発生した欠損値を の偶然的な欠損値とした。表1の中で誤答として問題がないのは だけであろう。測定目的にもよるが、他は多少なりとも部分得点を与える可能性について考慮する必要があるものと思われる。

続いて質問紙データの欠損値はどのように生ずるのかを考えてみよう。表2は、質問紙

表 1 学力検査において生じる欠損値

敢えて白紙で提出する、面倒などの学力とは無関係の無答
項目を読んだが後回しにしたため時間切れ無答
項目を読んで後回しにしたが解答し忘れて無答
正答が部分的にはわかっているが完全には分からないため無答
正答が全く分からないことによる無答
項目に到達しなくて時間切れ未解答
ページをとばすなどの事情による未解答
二重解答など被験者の記入ミスによる狭義の欠損値
採点時やデータ入力時に発生するような被験者とは無関係に生じる狭義の欠損値

表 2 質問紙において生じる欠損値

回答が面倒などの理由による無答
項目を読んだが後回しにしたため時間切れ無答
項目を読んで後回しにしたが回答し忘れて無答
項目に到達しなくて時間切れ未回答
ページをとばすなどの事情による未回答
二重回答など被験者の記入ミスによる狭義の欠損値
採点時やデータ入力時に発生するような被験者とは無関係に生じる狭義の欠損値
社会的望ましさなどを考慮して回答を回避する無答
項目の意味が理解できないための無答

データの欠損値の色々であるが、学力検査の場合とほぼ同様の理由により生じているものは表 1 と同じ番号を付けてある。表 1、表 2 に共通していないものは、学力検査あるいは質問紙データに特有の欠損値と思われるものである。

本研究で検討するのは無答タイプの欠損値に限ることとした。もちろんこのことは、後に述べる(1.3.4 節)欠損値処理方法が無答に限られるということではなくて、論文の題目にも示したように本研究の報告する範囲が表 1、表 2 の ~ の範囲であるということである。

1.3 項目反応理論での欠損値処理

項目反応理論で能力母数の推定に際して一般的に利用されている欠損値の処理には、「無視する」「誤答扱い」「(多肢選択テストデータの場合) Lord の方法」などの方法が

あり、これらについて以下で検討しておこう。

1.3.1 その 1 「無視する」

これは欠損値がある場合、あたかもその被験者と項目の組み合わせが存在しなかったかのごとくに処理してしまう方法である。つまり能力母数の推定の際に尤度関数から無視してしまう方法である。欠損値からは正誤データが得られず、通常の項目反応モデルでは欠損値の発生をモデル化していない以上やむを得ない選択ではある。この処理では、能力母数の推定は正誤が得られているデータで行えばよいと考えることになるわけであるが、欠損値をこのような方針で採点すると(そして被験者がその採点方法を知れば)、確実に正答できるものだけに解答することにより被験者は望むがままの高いの推定値を得ることができてしまうとの指摘がある(Lord, 1980, p.228)。「無視」したとしても、被験者とそ

の項目の組み合わせがあったという事実がなくなるわけではないので、本来無視すべき性格の欠損値であるかどうかが問題となる。無視しても良さそうな欠損値の型は、被験者が問題を見なかったという意味では、表1、表2の の未解(回)答であり、テストの測定する学力や質問紙の測定する特性とは無関係な要因により欠損が生じたという意味では、表1及び表2の と表2の となる。本研究が問題とする表1 ~ の無答などに関しては「無視」は必ずしも本当の意味で適当な処理方法とは考えられない。

1.3.2 その2「誤答扱い」

無答を誤答として処理してしまう方法である。これは被験者が学力検査で無答したということは、「分からないので無答したのだから誤答と処理してもよい」と考えたものである。表1の のケースであり、この場合には妥当な処理方法といえる。しかし学力検査において、テストの教示、緊張や不安、あるいは本人の性格などの影響により被験者の解答行動が抑制された場合などには、前記の構造にならず「完全な自信がなく不安になり、もし解答していたら正答する可能性が高かったのに無答してしまう」というような現象が生じる。表1の ~ のような無答である。本来は幾分かは正答する可能性を持つ、この種の無答を誤答と処理するのは必ずしも適当ではない。この種の無答発生を抑制するためにテストの教示などで解答を強要すると、あて推量などが助長される問題もある。教示などにより解答行動が大きく影響されることは、たとえば Bliss(1980)、Angoff & Schrader(1984)、藤森・中野(1994)などに示されている。また表1の のような未解答にも、この処理は適切とは言えない。問題が困難度の順に並んでいて被験者の未解答前の解答結果で誤答が連続しているならば妥当化できるかもしれないが常にこのような条件が満たされるとは限らない。たとえこのような条件が満たされたとしても未解答項目に正答する可能

性は残ってしまうだろう。表1の ~ に対処するのも困難である。また誤答として処理するのが許されるのは学力検査のときだけであり、質問紙データでは被験者の反応を0とすることは妥当化されないことにも注意する必要がある。

1.3.3 その3「(多肢選択テストデータの場合) Lord(1974)の方法」

項目反応理論では、学力検査が多肢選択形式である場合に Lord(1974)の方法によって無答を処理する場合がある。この方法は、表1の のケース、すなわち全く項目の正答がわからない状態を想定したモデルである。テストデータが多肢選択形式で与えられている場合、Lordの方法は(3)式において被験者の反応が無答なら x_j を1/選択肢数で置き換えて を推定するものである。すなわち、 s を無答数、 m を選択肢数とすれば

$$L = \prod_{j=1}^s P_j^m Q_j^{1-\frac{1}{m}} \prod_{j=s+1}^n P_j^{x_j} Q_j^{1-x_j} \quad (4)$$

を最大化する を求めるものである。多肢選択テストのためランダムなあて推量が行われるとして x を期待される正答確率 $1/m$ に置き換えることによって の推定値を得ている訳である。この方法は、多肢選択形式の学力検査にしか適用できず、しかも表1の のケースだけというきわめて限られた適用可能性しかないことに注意すべきである。無答をすべて のケースとしてしまうことには無理があるといわざるを得ない。もう一つの問題は藤森(1997)の指摘にあるようにバイアスがあることである。

1.3.4 藤森(1994)で取り上げられた欠損値処理方法

前記した既存の欠損値処理方法の中で多肢選択形式の学力検査だけでなく、記述式の学力検査の欠損値や質問紙への適用も含め、最も一般的な処理方法は「無視する」であろう。しかしそれでも表1、表2の ~ などに

「無視」で対処するのは困難である。

さて藤森（1994）で取り上げられた欠損値処理方法は、以下のようなものである。無答が s 個存在するときの推定のために次の疑似尤度 L^* を考える。

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

この疑似尤度 L^* を最大とする θ を推定値とする。これは Lord の方法の延長上にあるが、 $u_j = 1 / m$ とすれば(5)式の表現に含めることができる。さて実際に(5)式を最大化するには s 個の欠損値に対する重み u を求めなければならない。

u を求める第 1 の方法では、当該項目から予想される被験者の正誤の確率に応じて P_j と Q_j を重みづけることを考える。すなわち被験者が当該項目に正答する可能性を考慮に入れて

$$u_j = P_j \quad (6)$$

と重みづけることが考えられる。この考え方なら、表 1、表 2 の ~ 、表 2 の のいずれに適用しても大きな問題はないと思われる。すなわち、この方法は学力検査が多肢選択でなくても、また質問紙のデータであっても適用できるという利点がある。なお表 2 の は「無視」する方が適当と思われる。

さて(6)式のような重みを利用して(5)式を最大化する θ を求めることが必要であるが、(6)式右辺 P_j は、 θ が不明の時は定まらないので、何らかの方法で決定して(5)式を θ に関して最大化することが必要である。これには(5)式右辺の有効解答部分だけを利用することで θ を求めて(6)式右辺を推定することができるだろう。すなわち部分的尤度

$$L^{**} = \prod_{j=s+1}^n P_j^{x_j} Q_j^{1-x_j} \quad (7)$$

から θ を求めて(6)式右辺 P_j を推定する方法である。

この方法の一つの変形であるが、第 2 の方法として(5)の繰り返しを利用することが考えられる。つまり(7)の θ から出発し(6)右辺を経て(5)式から新たな θ を得て、これを(6)式右辺に代入し新たな重みを得て(5)の θ を推定するプロセスを繰り返すものである。最終的な推定値を得るには、(5)式と(6)式の繰り返し計算を(5)式の θ の変化が所定の値より小さくなるまで行うことになる。本研究では主に第 1、2 の方法を検討した。

さて、この他には(5)式に代入し、式に関して直接最大化することも考えられた。

$$L^* = \prod_{j=1}^s P_j^{P_j} Q_j^{Q_j} \prod_{j=s+1}^n P_j^{x_j} Q_j^{1-x_j} \quad (8)$$

つまり(8)式を最大化する θ を求める方法になるが、これは θ に関して(8)が単調な変化とはならないため適当ではないようである。この他にも、 u を 0 ~ 1 の値をとるパラメータと見なして u と θ に関して(5)式を最大にする最尤推定を行うことが考えられたが本研究では試みていない。

2. 方法

2.1 シミュレーションデータ

2 母数ロジスティックモデルに基づき以下のようにしてシミュレーションデータを作成し、無答の処理と能力母数の推定方法の関係を見ることにした。項目の識別力 a は平均 0.65、SD が 0.8、下限 -.8、上限 2.0 の切断正規分布に従い、困難度 b は平均 0、SD 1.0 の正規分布に従い、 θ は標準正規分布に従う。全ての母数を確定した後に 0 から 1 の一様乱数 R と 2 母数ロジスティックモデルの正答確率 P を比較し、 R が大きいとき誤答、小さいとき正答として正誤パターンを生成している。ネットワークの学習に利用するデータとして、テスト項目数は 20、40 および 60 項目であり、各項目数ごとにデータ A ~ J までの 10 組を作成した。学習のための人数は、いずれの場合も藤森（2000 a）の結果に従い 10000 人としている。以上のデータに基づきニューラル

ネットワークの学習を行った。

この他に最尤法による推定成績と比較するため、学習用のデータの項目母数をそのまま利用して3000名分の能力母数の真値を追加し、新規に作成した正誤パターンをクロスバリデーションデータとすることにした。このクロスバリデーションデータに関して無答を2.2節の方法に従って発生させ、無答の処理の影響を最尤推定とニューラルネットワークによる推定に関して比較している。

2.2 欠損値の生成のモデル化について

欠損値のあるテストデータについて能力母数の最尤推定とニューラルネットワークによる推定の比較検討をシミュレーションで行うためには、欠損値の発生についてモデル化する必要が生じる。ここでは藤森(1997)と同様の欠損値の生成を行うことにした。すなわち一度作成した各正誤データについて

$$M(\theta) = Q \times 0.1 \quad (9)$$

の確率で欠損値として置き換えたデータを作成した。本研究の欠損値発生モデルは、表1、表2の ~ の無答の発生モデルである。残余の型の欠損値、たとえば表1の は全くランダムに生じさせることで生成できる。表1の 、表2の などは別の発生モデルが必要となるが、ここでは試みていない。

2.3. 能力母数の推定

2.3.1. 最尤法による推定

能力母数の最尤推定値を求める際には項目母数が必要となるが、これにはシミュレーションの真値を充てた。もちろん一般には項目母数の真値を知ることができないが、推定値を用いるのでは、その推定精度も検討の対象にしなければならず、研究が複雑化する。これを避け、ここでは真値を用いることにした。この推定には自作のFORTRANプログラムを利用した。項目母数の真値を用いて能力母数の推定をしているわけであるから能力母数

の最尤推定として成績はその可能性の上限にあると考えてよいだろう。

クロスバリデーションデータの無答の処理方法としては1.3.1節の「無視」、1.3.2節の「誤答」、及び1.3.4節の式を最大にする \hat{u} から式の u を正答確率 P で置き換える方法1を適用する形で3種類の最尤推定値を得ている。

2.3.2 ニューラルネットワークによる推定

神経細胞の仕組みをコンピュータにより再現する試みが、パターン認識などを主な適用分野として1960年代より行なわれてきた。神経細胞が多数結合したものをニューラルネットワークと呼ぶ。ラメルハートら(Rumelhart et al,1986)のニューラルネットPDP(Parallel Distributed Processing)モデルは、入力層、中間層、出力層の3層を持つもので、その代表的なものとして知られている。3層モデルのネットワークの重みを変更する方法の代表的なものはバックプロパゲーション学習方式である。これは、出力層の信号を、外部の教師信号と最小二乗的に近づけるようにネットワークの重みを変更するものである(ニューラルネットワークについては豊田(1996)などを参照のこと)。

この研究で採用したニューラルネットワークモデルもやはり入力層、中間層、出力層を持つ3層のPDPモデルである(図1)。入力層のユニット数は、テスト項目数とし、各項目の解答は、正答=1あるいは誤答=0の入力信号として処理される。中間層の数は1であり、ここでは藤森(2000a)に従い、そのユニット数は2として検討した。出力層のユニット数は1であり、教師信号はシミュレーションで作成された能力母数の真値である。ネットワークの重みの推定はバックプロパゲーション法によることとし、このためのFORTRANプログラムを作成した。誤差関数の収束判定の基準は、最尤推定値によるMSEの0.9倍より小さく、同時に誤差関数の減少が $1.0E-07$ より小さくなったときに収

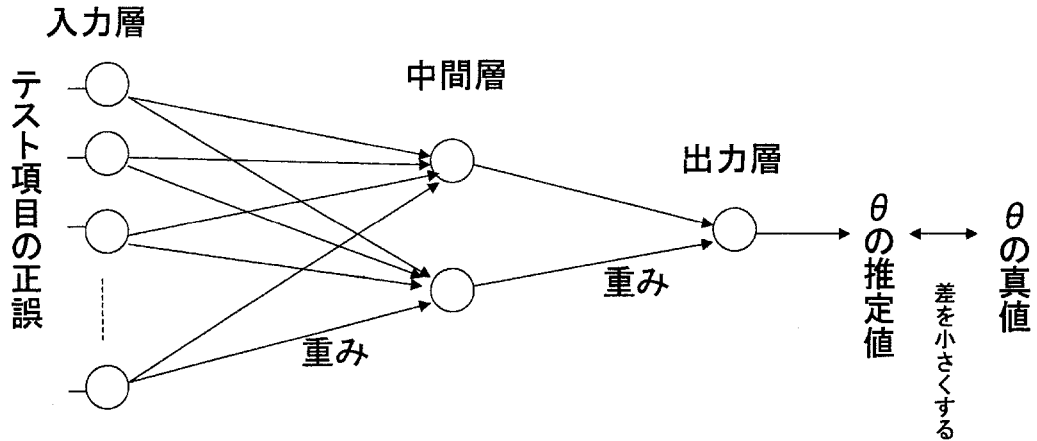


図1 3層のニューラルネットワークモデル

束としている。なお学習回数は上限10000回としてある。能力母数の真値とニューラルネットワークの能力母数の推定値との差の二乗の平均を、最小化するための誤差関数としている。これは、平均二乗誤差 (MSE) であるが本研究ではクロスバリデーションデータにおけるMSEと区別するため、誤差関数と表現する。ネットワークの重みの初期値は藤森 (2000a) に従い、 $-1 \sim 1$ の一様乱数を充てている。これらのデータはいずれも正答あるいは誤答のみからなるのでネットワークの学習には部分得点や欠損値は用いられていない。

学習が済んだネットワークを用いて無答を含んだクロスバリデーションデータの能力母数の推定値 $\hat{\theta}$ を求めるが、本研究では、無答は1.3.2節の「誤答」処理の他に、1.3.4節の方法1の考えを適用し、一旦無答を「誤答」処理して得られた $\hat{\theta}$ に基づき(1)式のPを求めて、無答をこのPで置き換えてネットに再度入力して出力層に現れた値を能力母数の推定値としている。2.3.1節の最尤法による能力母数の推定と異なり、ニューラルネットワークでは「無視」するは適用出来ないため、1.3.1節の「無視」は行っていない。方法1では、無答のPによる置き換えは1

回だけであったが、ここでは1.3.4節の方法2と同じ考えで一度得られたニューロ推定値 $\hat{\theta}$ からPを計算し無答をこれで置き換え、再度ネットの入力値として $\hat{\theta}$ を求めるといふ繰り返し計算も行ってみた。すなわちニューラルネットワークでも合計3種類の無答処理に伴う $\hat{\theta}$ を検討している。

3. 結果と考察

最尤推定値 (MLE) 及びニューラルネットによる推定値の成績比較は、クロスバリデーションデータにおける能力母数の真値と推定値に基づく平均二乗誤差 (MSE) によることとした。ニューラルネットワークモデルにおいては、学習の誤差関数の最小値を与えた重みをもとにして、クロスバリデーションデータについての能力母数の推定値を得ている。

表3～表5は、クロスバリデーションデータに関する20項目～60項目の最尤推定の無答なし 無視 誤答扱い 正答確率Pで置き換えの各ケースの平均、SD、MSEと、同じくニューラルネットワークの 無答なし 誤答扱い 正答確率Pで置き換え 正答確率Pで置き換え (繰り返し) の各ケー

スのMSEをまとめたものである(最尤推定と番号を統一するためを省いている)。クロスバリデーションデータで「無答なし」の場合のMSEが、能力母数推定のための情報が多いのだから、他と比較して最も小さくなる筈である。ニューラルネットワークではそのような結果が得られているが、最尤推定ではのPで置き換えた場合のMSEがどのデータでも明らかに最も小さくなっている。これは(6)式による置き換えが項目反応モデルをその根拠に置いており、そのことがシミュレーションというモデルの当てはまりが良い状況下で、項目反応モデルに基づいて能力母数を最尤推定する際に有利に働いているものと考えられる。ニューラルネットワークによる能力母数の推定では、正答確率Pの計算以外は項目反応モデルを利用していないため、このような現象が生じていないものと思われる。すなわち、(6)式による置き換えはモデルの当てはまりが良いときに有効であるものと考えておくのが適当であろう。

最尤法においての「誤答扱い」は項目数が20項目と少ない場合は「無視する」より成績がよいが、60項目では逆にになっている。無答を「無視する」場合は、より少ない項目数で推定しなければならず、20項目ではその影響が強いものと考えられる。20項目での「誤答扱い」の成績のよさは、(9)式による無答発生によれば元の正誤データの中で正答が無答に置き換わる確率が相対的に低いことが幸いし(この無答を誤答扱いすれば誤った処置となる)「無視する」の項目数の減少の悪影響よりも推定成績の悪化の程度が抑えられたものと思われる。

なお最尤推定もニューラルネットワークによる推定にしても、「誤答扱い」の平均値が絶対値は小さいものの負となっている。能力母数の真値は標準正規分布に従うとしているから、「誤答扱い」は負のバイアスが掛かっていることになる。無答の中に正答度のあるものが含まれれば含まれるほどこのバ

イアスの影響は強くなり成績が悪化する可能性があるのは注意しておきたい。

続いてニューラルネットワークによる推定では、の「誤答扱い」は項目数が20項目と少ない場合は「正答確率Pで置き換え」より成績がよいが、60項目では逆にになっている。しかしとの成績の差は大きいものではない。またの「正答確率Pで置き換え(繰り返しあり)」は、単純なよりも全体的に成績が悪く、繰り返しを行う意味のないことが分かった。

なお表1～表5をみると全体的に最尤推定はSDが1.0より大きく、すなわち真のSDより大きく、逆にニューラルネットワークでは小さいことが分かる。どちらも平均の位置は0付近で大差ないので、最尤推定のMSEが小さいのは能力母数の絶対値が大きい点での最尤推定値特有のずれの大きさが主因ではないかと思う人もいるかもしれない。そこで本論より少し外れるがニューラルネットワークによる能力母数の推定値の分布がどのような状態にあるのかを最尤推定と比較することも無意味ではないだろう。表6～表8は、項目20～60の中の10組のデータの中から代表としてデータAを選び、真のが-1.5, -1.0, -0.5, 0.0, 0.5, 1.0, 1.5の各点であるとして、各点ごとに1000人分の(欠損値はない)正誤パターンをクロスバリデーションデータの要領と同様にして作成し、の最尤推定値とニューラルネットワークによる推定値を元に、その平均、SD、MSEを示したものである。表8の=-0.5を除いて表6～表8のいずれの点においてもニューラルネットワークの成績が最尤推定に比較して良く、能力母数の絶対値が大きい点だけでニューラルネットワークの成績が良いわけではないことが示された。すなわちの真値の大小に係わらずニューラルネットワークの推定が良さそうなことが示されたといえよう。

以上をまとめると、無答の処理方法としてはの「正答確率Pで置き換え」による

項目反応理論における無答処理とニューラルネットワークによる能力母数の推定

表3 クロスバリデーションデータでのMSE (項目数20)

	方法	データA			データB			データC			データD			データE		
		MSE	平均	SD	MSE	平均	SD	MSE	平均	SD	MSE	平均	SD	MSE	平均	SD
最尤法		0.2619	0.0242	1.1549	0.2520	0.0627	1.1837	0.2589	0.0179	1.1887	0.2635	0.0121	1.1981	0.2452	-0.0257	1.1924
		0.2843	0.0246	1.1693	0.2657	-0.0384	0.8951	0.2731	0.0160	1.1996	0.2817	0.0081	1.2085	0.2614	-0.0238	1.2044
		0.2777	-0.0734	1.1608	0.2503	-0.0266	1.1827	0.2632	-0.0756	1.1877	0.2701	-0.0815	1.1963	0.2601	-0.1147	1.2002
		0.2521	0.0667	1.0823	0.2346	0.0837	1.1212	0.2371	0.0498	1.1146	0.2403	0.0453	1.1159	0.2212	0.0295	1.0973
ニューロ		0.1950	0.0393	0.8572	0.1687	0.0291	0.8942	0.1784	0.0111	0.8639	0.1838	0.0070	0.8909	0.1964	-0.0015	0.8752
		0.2002	-0.0326	0.8579	0.1703	-0.0384	0.8951	0.1794	-0.0561	0.8652	0.1880	-0.0614	0.8917	0.1978	-0.0657	0.8773
		0.2037	0.0372	0.8557	0.1732	0.0262	0.8927	0.1816	0.0063	0.8610	0.1908	0.0008	0.8905	0.2023	-0.0030	0.8748
		0.2058	0.0439	0.8557	0.1753	0.0325	0.8917	0.1836	0.0129	0.8602	0.1925	0.0071	0.8902	0.2046	0.0033	0.8746
	方法	データF			データG			データH			データI			データJ		
		MSE	平均	SD	MSE	平均	SD	MSE	平均	SD	MSE	平均	SD	MSE	平均	SD
最尤法		0.2518	-0.0021	1.1793	0.2397	0.0304	1.1549	0.2275	-0.0173	1.1519	0.2294	0.0487	1.1690	0.2404	-0.0065	1.1722
		0.2599	-0.0013	1.1835	0.2592	0.0024	1.1716	0.2451	-0.0167	1.1604	0.2434	0.0479	1.1775	0.2587	-0.0091	1.1852
		0.2562	-0.0915	1.1754	0.2538	-0.0646	1.1658	0.2367	-0.1022	1.1543	0.2367	-0.0345	1.1699	0.2528	-0.0997	1.1758
		0.2217	0.0448	1.0872	0.2156	0.0808	1.0623	0.2146	0.0414	1.0582	0.2121	0.0840	1.0884	0.2173	0.0441	1.0799
ニューロ		0.1807	0.0083	0.8797	0.1861	0.0528	0.8770	0.1706	-0.0012	0.8679	0.1607	0.0278	0.8758	0.1732	0.0075	0.8787
		0.1840	-0.0585	0.8775	0.1865	-0.0157	0.8795	0.1729	-0.0655	0.8692	0.1646	-0.0336	0.8791	0.1751	-0.0611	0.8801
		0.1842	0.0059	0.8748	0.1917	0.0471	0.8793	0.1778	-0.0033	0.8629	0.1656	0.0253	0.8734	0.1793	0.0028	0.8800
		0.1858	0.0127	0.8742	0.1935	0.0537	0.8793	0.1800	0.0036	0.8622	0.1672	0.0310	0.8723	0.1814	0.0095	0.8801

注意： は無答なし 無視 誤答扱い 正答確率Pで置き換え 正答確率Pで置き換え(繰り返しあり)

表4 クロスバリデーションデータでのMSE (項目数40)

	方法	データA			データB			データC			データD			データE		
		MSE	平均	SD	MSE	平均	SD	MSE	平均	SD	MSE	平均	SD	MSE	平均	SD
最尤法		0.1097	-0.0302	1.0706	0.1293	-0.0143	1.1110	0.1135	-0.0318	1.0692	0.1300	0.0276	1.1185	0.1301	-0.0041	1.0649
		0.1156	-0.0324	1.0737	0.1335	-0.0122	1.1119	0.1195	-0.0297	1.0727	0.1387	0.0014	1.1269	0.1381	-0.0015	1.0682
		0.1188	-0.1174	1.0715	0.1355	-0.0995	1.1104	0.1235	-0.1124	1.0737	0.1415	-0.0844	1.1237	0.1369	-0.0932	1.0675
		0.1055	0.0089	0.9866	0.1226	0.0383	1.0161	0.1099	0.0179	0.9875	0.1193	0.0541	1.0257	0.1240	0.0360	0.9868
ニューロ		0.0922	-0.0235	0.9107	0.1097	-0.0051	0.9402	0.0960	-0.0229	0.9135	0.1028	0.0027	0.9507	0.1039	-0.0037	0.9006
		0.0980	-0.0977	0.9121	0.1134	-0.0777	0.9400	0.1009	-0.0917	0.9150	0.1099	-0.0708	0.9535	0.1080	-0.0795	0.9024
		0.0965	-0.0279	0.9096	0.1123	-0.0064	0.9358	0.0992	-0.0241	0.9114	0.1068	-0.0013	0.9514	0.1085	-0.0050	0.8984
		0.0973	-0.0225	0.9085	0.1133	-0.0006	0.9348	0.1000	-0.0186	0.9105	0.1076	0.0043	0.9505	0.1097	0.0009	0.8972
	方法	データF			データG			データH			データI			データJ		
		MSE	平均	SD	MSE	平均	SD	MSE	平均	SD	MSE	平均	SD	MSE	平均	SD
最尤法		0.1150	-0.0143	1.1081	0.1141	-0.0308	1.0932	0.1191	-0.0431	1.0968	0.1149	0.0124	1.0994	0.1228	0.0290	1.0563
		0.1248	-0.0124	1.1140	0.1208	-0.0327	1.0974	0.1297	-0.0449	1.1026	0.1211	0.0136	1.1045	0.1308	0.0284	1.0632
		0.1256	-0.1010	1.1095	0.1221	-0.1120	1.0955	0.1335	-0.1311	1.1006	0.1189	-0.0655	1.0981	0.1312	-0.0622	1.0619
		0.1134	0.0519	1.0043	0.1111	0.0320	0.9883	0.1190	0.0129	1.0029	0.1098	0.0509	1.0151	0.1194	0.0628	0.9839
ニューロ		0.0953	-0.0004	0.9356	0.9428	-0.0162	0.9364	0.1030	-0.0292	0.9281	0.0889	0.0012	0.9351	0.1046	0.0208	0.8913
		0.1007	-0.0725	0.9334	0.0982	-0.0857	0.9365	0.1099	-0.1034	0.9276	0.0946	-0.0661	0.9362	0.1108	-0.0556	0.8963
		0.1004	-0.0015	0.9328	0.0985	-0.0204	0.9350	0.1085	-0.0329	0.9259	0.0927	-0.0012	0.9350	0.1086	0.0168	0.8915
		0.1014	0.0044	0.9319	0.0995	-0.0150	0.9340	0.1096	-0.0271	0.9250	0.0935	0.0038	0.9339	0.1094	0.0224	0.8902

注意： は無答なし 無視 誤答扱い 正答確率Pで置き換え 正答確率Pで置き換え(繰り返しあり)

表5 クロスバリデーションデータでのMSE(項目数60)

	方法	データA			データB			データC			データD			データE		
		MSE	平均	SD	MSE	平均	SD	MSE	平均	SD	MSE	平均	SD	MSE	平均	SD
最尤法		0.0805	-0.0104	1.0800	0.0738	0.0231	1.0661	0.0781	-0.0046	1.0692	0.0771	0.0213	0.9576	0.0760	0.0017	1.0449
		0.0859	-0.0099	1.0840	0.0780	0.0239	1.0678	0.0830	-0.0058	1.0736	0.0799	0.0198	1.0769	0.0789	0.0027	1.0445
		0.0921	-0.0936	1.0847	0.0797	-0.0566	1.0661	0.0869	-0.0891	1.0724	0.0800	-0.0612	1.0730	0.0817	-0.0765	1.0415
		0.0844	0.0535	0.9774	0.0752	0.0647	0.9810	0.0784	0.0371	0.9845	0.0768	0.0656	0.9812	0.0761	0.0339	0.9719
ニューロ		0.0731	-0.0029	0.9591	0.0662	0.0250	0.9585	0.0701	-0.0032	0.9478	0.0677	0.0229	0.9592	0.0688	0.0045	0.9340
		0.0808	-0.0774	0.9616	0.0706	-0.0470	0.9587	0.0760	-0.0783	0.9515	0.0704	-0.0485	0.9587	0.0733	-0.0659	0.9323
		0.0768	-0.0058	0.9594	0.0691	0.0225	0.9581	0.0731	-0.0074	0.9490	0.0700	0.0213	0.9576	0.0712	0.0213	0.9318
		0.0773	-0.0006	0.9584	0.0698	0.0273	0.9569	0.0736	-0.0026	0.9477	0.0707	0.0262	0.9566	0.0718	0.0068	0.9310
	方法	データF			データG			データH			データI			データJ		
		MSE	平均	SD	MSE	平均	SD	MSE	平均	SD	MSE	平均	SD	MSE	平均	SD
最尤法		0.0762	-0.0062	1.0364	0.0823	0.4320	1.0546	0.0708	0.0056	1.0644	0.0798	0.0019	1.0685	0.0690	-0.0417	1.0449
		0.0819	-0.0045	1.0394	0.0865	0.0432	1.0597	0.0750	0.0066	1.0642	0.0848	0.0022	1.0718	0.0739	-0.0414	1.0468
		0.0866	-0.0900	0.9274	0.0900	-0.0471	1.0583	0.0768	-0.0700	1.0628	0.0882	-0.0822	1.0701	0.0784	-0.1207	1.0437
		0.0768	0.0242	0.9677	0.0799	0.0793	0.9776	0.0741	0.0573	0.9728	0.0773	0.0611	0.9668	0.0732	0.0147	0.9479
ニューロ		0.0713	-0.0068	0.9240	0.0723	0.0452	0.9405	0.0653	0.0141	0.9543	0.0669	0.0091	0.9533	0.0638	-0.0321	0.9360
		0.0791	-0.0811	0.9274	0.0770	-0.0347	0.9424	0.0698	-0.0549	0.9535	0.0731	-0.0664	0.9542	0.0703	-0.1021	0.9361
		0.0747	-0.0083	0.9233	0.0743	0.0418	0.9409	0.0686	0.0113	0.9526	0.0700	0.0055	0.9526	0.0674	-0.0347	0.9347
		0.0752	-0.0034	0.9218	0.0748	0.0469	0.9397	0.0692	0.0162	0.9518	0.0705	0.0108	0.9516	0.0679	-0.0299	0.9336

注意： は無答なし 無視 誤答扱い 正答確率Pで置き換え 正答確率Pで置き換え(繰り返しあり)

表6 を固定したときの $\hat{\theta}$ の分布(20項目のデータA)

真の	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5
	最 尤 推 定 値						
平均	-1.6471	-1.1180	-0.5772	-0.0351	0.4989	1.0639	1.6404
SD	0.6430	0.5798	0.5004	0.4887	0.4670	0.5280	0.6104
MSE	0.4351	0.3501	0.2564	0.2401	0.2181	0.2829	0.3923
	ニューラルネットワークによる推定値						
平均	-1.1842	-0.8593	-0.4642	-0.0209	0.4267	0.8710	1.2522
SD	0.3513	0.3813	0.3864	0.4042	0.3860	0.3786	0.3491
MSE	0.2231	0.1652	0.1506	0.1638	0.1544	0.1600	0.1833

注意： の各点での被験者数はいずれも1000名(以下同様)

表7 を固定したときの $\hat{\theta}$ の分布(40項目のデータA)

真の	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5
	最 尤 推 定 値						
平均	-1.5982	-1.0358	-0.5061	0.0157	0.5189	1.0441	1.5508
SD	0.4424	0.3412	0.2967	0.2802	0.2946	0.3429	0.4376
MSE	0.2053	0.1177	0.0881	0.0788	0.0871	0.1195	0.1941
	ニューラルネットワークによる推定値						
平均	-1.3901	-0.9529	-0.4596	0.0300	0.4841	0.9422	1.3285
SD	0.2856	0.2989	0.2862	0.2561	0.2640	0.2855	0.2884
MSE	0.0937	0.0915	0.0835	0.0665	0.0699	0.0847	0.1126

表8 を固定したときの $\hat{\theta}$ の分布 (60項目のデータ A)

真の	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5
最 尤 推 定 値							
平均	-1.5760	-1.0512	-0.5013	0.0027	0.5016	1.0071	1.5412
SD	0.3850	0.2972	0.2571	0.2342	0.2250	0.2529	0.3318
MSE	0.1540	0.0910	0.0661	0.0548	0.0506	0.0640	0.1118
ニューラルネットワークによる推定値							
平均	-1.4260	-1.0051	-0.4705	0.0146	0.4708	0.9572	1.4419
SD	0.2591	0.2730	0.2564	0.2167	0.2105	0.2473	0.2643
MSE	0.1540	0.0745	0.0666	0.0472	0.0452	0.0630	0.0732

のが最尤推定ではいずれの場合も成績が良かった。項目反応モデルの当てはまりが良い場合には、 θ の「正答確率 P で置き換え」を行うのが適当であろう。項目数が少ない場合のニューラルネットワークによる推定では、 $\hat{\theta}$ の処理は θ の誤答処理に比較して、その差はわずかであるものの悪い成績なので「誤答扱い」が可能な学力検査では現時点では必ずしも妥当化出来ない。ニューラルネットワークによる推定でも項目数が多い場合や質問紙データなどでは $\hat{\theta}$ による処理が適当と思われる。最尤推定の成績は、ニューラルネットワークによる推定の成績に全体的にはかなり劣っているのでニューラルネットワークによる推定が可能な場合にはこれを用いる方が良い。なお本研究では、表1、表2の θ の無答をシミュレーションで発生させたが、適用可能としながらも、本研究では取り上げなかった他のタイプの欠損値の検討も、今後の課題として残っていることを指摘しておきたい。

文 献

Angoff, W. H. & Schrader, W.B. 1984 A study of hypotheses basic to the use of right and formula scores. *Journal of Educational Measurement*, 21,1,1-17.

Birnbaum, A. 1968 Some latent trait models and their uses in inferring an examinee's ability. In F.M.Lord and M.R.Novick, *Statistical theories of mental test scores* (pp.397-479). Reading MA : Addison-Wesley.

Bliss, L.B. 1980 A test of Lord's assumption regarding examinee guessing behavior on

multiple-choice tests using elementary school students. *Journal of Educational Measurement*, 17, 2, 147-153.

藤森 進 1994 項目反応理論における無答処理について — 算数・数学問題解決能力尺度への新無答処理方式の適用 — 第2回項目反応理論研究会発表資料.

藤森 進 1997 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 *心理学研究*,67,6,429-435.

藤森 進 2000a ニューラルネットワークによる項目反応モデルの能力母数の推定について 日本教育心理学会第42回総会発表論文集 (当日の追加資料含む)

藤森 進 2000b 項目反応モデルにおける部分得点への対応 — ニューラルネットワークの利用 — 日本心理学会第64回大会発表論文集 (当日の追加資料含む)

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*,39,2,247-264.

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.

Rumelhart, D.E., McClelland, J.L. and the PDP Research Group. 1986 *Parallel distributed processing: Explorations in the microstructure of cognition*. Vol.1,2. The MIT Press.

芝 祐順編 1991 項目反応理論 東京大学出版会.

豊田秀樹 1996 非線形多変量解析: ニューラルネットワークによるアプローチ 朝倉書店.

引用文献

- (1) Immanuel Kant 「判断力批判」篠田英雄訳岩波書店、1979(1790)年、p29
- (2) D. Jeans, 'Changing formulation of man-environment relationship in Anglo-American geograph', 'Journal of Gography, 1974.73(3), pp36-40
- (3) Rachel Carson 「失われた森」Linda Lear編、古草秀子訳、集英社、2000(1998)、p181
- (4) Friedrich Von Schiller 「人間の美的教育について」小栗孝則訳、法政大学出版局、1972(1795)、p99
- (5) Dabid Dickson 「オルターナティブ・テクノロジー」田窪雅文、里深文彦訳、時事通信社、1980(1974)、p12
- (6) Robin Clarke, ' Technolojy for an alternative society , ' New Scientist'(Jarnuary 11,1973)
- (7) Wald Emerson 「自然について」斉藤光訳、日本教文社、1996(1836),p56
- (8) Rene Descartes 「方法序説」落合太郎訳、岩波書店、1967(1637)、p76
- (9) Benedictus de Spinoza 「エチカ 倫理学」畠中尚志訳、岩波書店、1951(1677)、p86
- (10) 河井徳治 「スピノザ哲学論攷 自然の生命統一について」創文社、1994,p84
- (11) 河井、前掲出、p7
- (12) 高木仁三郎 『エコロジーの考え方』「自然とは」岩波書店、1989、p302
- (13) William Ashworth " The Encyclopedia of Environmental Studies ", Facts on File, New York, 1991, p116l
- (14) Ashworth, ibid,116r
- (15) Edward Goldsmith "「エコロジーへの道」大熊 昭信訳、法政大学出版局、1998(1992)、p42
- (16) James Lovelock 「地球生命圏 ガイアの科学」スワミ・プレム・ブラブダ訳、工作舎、1998(1979)
- (17) 井上民二 『生物多様性 —— その意義と現状』「生物多様性とその保全」岩波書店、1998、p6
- (18) 鬼頭秀一 「自然保護を問いなおす —— 環境倫理とネットワーク」筑摩書房、1996、p254
- (19) Fritjof Capra 「ディープ・エコロジー考」つる田栄作、校成出版、1995、pp35-36
- (20) Albert Schweitzer 「わが生活と思想より」竹山道雄訳、白水社、1959(1931)、p218
- (21) Aldo Leopold 「野生のうたが聞こえる」新島義昭訳、森林書房、1986(1949)、p313
- (22) Arne Naess 「ディープ・エコロジーとは何か」斉藤直輔、関龍美訳、文化書房博文社、1997(1989)、pp48-49
- (23) Naess、前掲出、p50
- (24) Murray Boockchin 「エコロジーと社会」藤堂麻理子他訳、白水社、1996(1990)
- (25) Carolyn Merchant 「ラディカルエコロジー」川本隆史他訳、産業図書、1994(1992)
- (26) Luc Ferry 「エコロジーの新秩序」加藤 宏幸訳、法政大学出版局、1994(1992)、p216
- (27) 井上有一 『ホーリスティックな世界観と 民主的・市民的価値』「環境の豊か さをもとめて」昭和堂、1999、p92
- (28) Jost Hermand 「森なしには生きられない」山縣光晶訳、築地書館、1999(1993)、p140
- (29) Roger Sidaway, 'Sport, Recreation and Nature Conservation ', Leisure and the Enviroment, Sue Glyptis Belhaven Press, 1993, p163
- (30) 朝日新聞、2000、7、2 日刊、経済p 8、14版
- (31) James Lvelock 「ガイアの時代」スワミ・ブラブダ訳、工作舎、1989(1988)、p337
- (32) Hargrove Eugene, ' Foundation of Environmental Hall Ethics', Prentice, 1989
- (33) Roderick Frazier Nash 「自然の権利」松野弘訳、TBSブリタニカ、1993(1990)、pp168-169
- (34) 高辻正基 『エントロピーとアメニティ』「アメニティを考える」AMR編修、未来社、1989
- (35) 沼田真 「景相生態学」朝倉書、1996、p 2
- (36) Marrtin .J.Elson, 'Sport and Recreation in the Green Bert Countryside'; Leisure and the Environment, Sue Glyptis, 1993, p131
- (37) 岸由二 「いるか丘陵の自然観察ガイド」山と渓谷社、1997、p14
- (38) 鷺谷いづみ 「生物保全の生態学」共立出版、1994、p33
- (39) Peter Bromley, ' Cuntoryside Recreation, A Handbook For Manager, E & FNSPON 1994, p85
- (21) Aldo Leopold 「野生のうたが聞こえる」新島義