

【個人研究】

非対称事前分布を利用した能力母数の推定 成長を前提としたテスト場面への適用

藤 森 進*

Ability Parameter Estimation Utilizing Asymmetric Prior Distribution: Application to a Testing Situation in Which Ability Shows Improvement

Susumu FUJIMORI

The present study simulated ability estimation based on item response theory, where several tests are conducted during a period in which the academic ability of learners shows improvement, in environments such as language schools. In such a situation, favorable ability estimates can be obtained by developing a prior distribution for the ability of learners based on past measurement results and performing Bayes estimation using the prior distribution. A normal distribution is generally used as a prior distribution; however, in the present study, ability estimation was done using an asymmetric lognormal distribution, which better suits a situation in which ability shows improvement. The results of the present simulation study showed that using a lognormal distribution as a prior distribution, rather than a normal distribution, enabled more favorable ability estimates to be obtained.

Key words: item response theory, Bayes estimation, prior distribution, lognormal distribution

1. 研究の目的

この研究では、成長期のような学力が向上し、しかも複数回のテストによって能力測定が行われる場面に項目反応理論を適用することを想定する。また成長そのものは、質的な変化を伴わずに同一の能力次元上での変化に限るものとする。例えば一般の学校教育の場面であり、また語学学校のような、比較的短期間とは言っても数ヶ月から1年程度となるであろうがで教育を行い、しかも期間内にテストを繰り返すケースである。このような場合、初回の測定の際はともかくとして、2

回目以降のテストでは被験者の能力についてかなり明確な情報が存在する。そのような場合、一般的に行われることの多い母数の最尤推定ではなくて、事前情報を利用するベイズ推定を行うことが可能である。ベイズ推定では、被験者の能力に関する事前情報を事前分布に表現することにより安定した推定を得ることが期待できるためである。もちろん、成長期のような学力が伸長する場合に、適用可能な分析方法は事前情報を利用したベイズ推定だけではない。たとえば一般の学校であれば経過時間や、語学学校などではレッスン回数関数として学力の伸びを時系列的にモデル化することも可能であろう。また、能力の質的变化を想定することもできる。この種の研究としては、多次元ラッシュモデルを利用

* ふじもり すすむ 文教大学人間科学部人間科学科

した Embretson (1991) や Willson (1989) の SALTUS モデルなどがある。このタイプのアプローチは、発達過程を明らかに出来るという点で魅力的ではあるが、発達現象によるパターンの違いが存在することや、能力の違い及び発達スピードの個人差のため、モデル化とその母数の推定のためにはかなりしっかりとした基礎データを集積する必要があるなどの問題も考えると、実用場面に容易に適用できるものではない。その点ベイズ推定によるアプローチは、事前分布を準備するだけで済むので相対的に適用は容易である。事前分布を利用する場合にも色々な工夫があり得る。例えば、事前分布の位置を発達データに基づいて決定するような方法である。具体的には、1学年進んだ学生や所定のレッスンを受講した者の平均的な学力の伸びをあてはめる等が考えられる。これも一種の発達モデルであろうが、発達現象そのもののモデル化と比較すればデータの集積だけで済むので、その分容易であろう。しかし、この場合でも基礎データの事前の集積は必要であり、それだけの準備をしたにもかかわらず、平均的な発達速度以外の者にとっては、事前分布は場合により有利になったり不利に働いたりする。すなわち、能力の伸びの大きかった者にとっては不利となり、伸びの少なかった者にとっては有利に働くこととなるだろう。その点からは、やや適用しにくいと考えられる。

このような事情により本研究では、項目反応理論を前提として、既存の測定結果に基づく能力の推定値とその推定精度を利用して事前分布を作成することにする。ただし、事前分布として通常良く利用される正規分布ではなく非対称な分布を用いることにする。多くのベイズ統計学のテキストで事前分布として推奨されるのは、次式のベータ分布である。

$$\frac{x^{p-1}(1-x)^{q-1}}{B(p,q)} \quad (1)$$

ここで $B(p, q)$ は次式のベータ関数である。

$$B(p,q) = \int_0^1 u^{p-1}(1-u)^{q-1} du \quad (2)$$

ベータ分布が推奨されるのは、事前分布として多様な形態を表現し得ること、ベータ分布は2つの母数 p, q を持つだけでシンプル、データがベルヌーイ分布に従うとき、事前分布がベータ分布であるならば事後分布もベータ分布になるというような自然共役事前分布の性質を持つなどの便利な点が指摘されるためである。このような長所は指摘されるものの、本研究で採用した確率分布は、ベータ分布でなく対数正規分布である。この理由は、理論的なものではなく、次の実用上の理由による。対数正規分布は、次式で定義されるように2つの母数 μ と σ を持ち、ベータ分布と母数の数での違いは無い。

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\log x - \mu)^2}{2\sigma^2}\right\} \quad (3)$$

本研究は、実用上の観点に立っており、その点から μ を見ると、事後分布を求めた後、これを再度事前分布として利用するようなことは想定していないので（例えば適応型テストではこのような可能性も考慮されるべきであろうが）自然共役性の有無は意味を持たない。これに対して σ の特徴は、実は長所とされる一方で問題を持っている。すなわち、ベータ分布が、事前分布の形状として、実にさまざまなものをたった2つの母数で表現できるということは、母数のほんの僅かな違いが分布の形状にかなりの相違を生じることになることを意味する。また、母数のベータ分布の母数の値の持つ意味合いは、正規分布のような母数の持つ意味合いとは異なる。正規分布では、各母数は、分布の位置と散らばり度合いという、分布の分かりやすい1つの特性と結びついているのに対して、ベータ分布では、そのような単純な把握は出来ない。すなわちベータ分布は、母数の値を見ただけでは不慣れな者にとって形状等について分かりにくい問題点を持つ。これに対して、対数正規

分布は、関数形も比較的簡単なものであり、母数の変化に対して形状がベータ分布に比し安定的である。母数の意味合いは、直接的にはベータ分布と同様に解釈困難であるが、(3)式の対数正規分布に従う変数 X の対数 $\log(X)$ は、 $N(\mu, \sigma^2)$ に従うため、母数の性質は、ある程度正規分布との関係から理解できるという利点を持つ。このことは、実際の適用場面での事前分布の具体的な設定の際に意味がある。また、本研究で想定しているような、学力の伸びを想定している場面では、初回のテストによる能力より2回目以降のテスト結果の能力の方が上回っており、初回の能力より下回することは基本的に生じないと考えられる。これは真値に関してだけでなく推定値にも同様のことが言える。対数正規分布は、右すそ部分が重く、このような場面の事前分布として用いるのに適当である。これらの理由により、事前分布として余り利用されることの無い対数正規分布を事前分布として利用することにした。ベイズ統計学に関しては例えば渡部(1999)を参考にされたい。

以上により、本研究は、成長期のような学力が向上し、しかも複数回能力測定がテストによって行われる場面に項目反応理論を適用するという条件の下で、能力母数のベイズ推定のための事前分布として対数正規分布を利用し、一般的に利用されることの多い正規分布との比較をシミュレーションによって検討する。

2. 方法

2.1. 項目反応モデル

項目反応モデルに属するものは数多くあるが、ここではよく利用される(4)式の2母数ロジスティックモデル(Birnbaum, 1968)を採用する。

$$p(x_{ij} = 1 | \theta_i, a_j, b_j) = \frac{1}{1 + \exp(-Da_j(\theta_i - b_j))} \quad (4)$$

ここで i は被験者、 θ はその能力を表す母数、 $D = 1.7$ の定数、 j は項目番号、 a_j はその識別力、 b_j は困難度を表す母数である。また x_{ij} は、被験者 i の項目 j に対する正誤を表し、正答のとき1、誤答のとき0をとる。モデルや母数の意味の詳細は、藤森(2002)などを参考にされたい。

2.2. 能力母数の推定

項目反応理論では、母数の推定を最尤法あるいはベイズ法によるのが一般的である。本研究では、項目母数を所与としたとき能力母数を推定するため(5)式の対数尤度を最大にする、すなわち最大事後確率推定法 *maximum a posteriori*(MAP) 推定法による θ を能力母数の推定値とする。

$$\ell_i = \sum_{j=1}^n \{x_{ij} \ln(P_j(\theta_i)) + (1 - x_{ij}) \ln(Q_j(\theta_i))\} + \varphi_i(\theta_i) \quad (5)$$

ここで n は項目数、 P は(4)式の2母数ロジスティックモデルであり、 $Q = 1 - P$ である。また、 $\varphi_i(\theta_i)$ は、被験者 i に関する事前分布の対数である。

母数の推定は、自作のpascalプログラム(delphi6)によった。

2.3. 事前分布としての対数正規分布の設定

本研究の前提とする成長期のような学力が向上し、しかも複数回能力測定がテストによって行われる場面では、2回目以降のテストでは、前回のテスト結果が被験者の事前情報を与えてくれる。ベイズ推定では、これを事前分布にするわけであるが、ここでは次のようにして事前分布を作成した。

まず被験者の初回のテスト結果に基づく能力母数のベイズ推定値 $\hat{\theta}_i$ を求める。項目母数が所与のとき、能力母数の最尤推定値の分布は、真値を平均とし、分散はテスト情報関数 $I(\theta)$ によって(6)式により定まる正規分布で近似される。ベイズ推定の場合、推定値の分散には事前分布の影響があるが本研究では考慮していない。(6)式による評価が便利で

あるのは、項目母数のみによって分散が評価できるところにあるのであり、能力母数の推定方法による相違が生じるのは利用の際に煩雑となるため、これを回避するためである。

$$V(\hat{\theta}) = \frac{1}{I(\theta)} = \frac{1}{\sum_{j=1}^n D^2 a_j^2 P_j(\theta) Q_j(\theta)} \quad (6)$$

さて、最尤推定値の分布が正規分布で近似されるわけであるから、その正規分布を事前分布とするならば、平均をベイズ推定値 $\hat{\theta}_1$ とし、また分散を(6)式により定めるのが自然であるし、能力の伸びを考えない場合は、これで一向に差し支えないと思われる。再度のテストを受けて能力推定を行ったとしても推定値には良い影響があると考えられる。しかし能力の伸びがある場合は、このような事前分布では $\hat{\theta}_1$ が低いため、以後の能力推定に下方圧力がかかるという問題点が生じる。この研究では、その1つの簡便な対策として、事前分布に非対称を持ち込み(3)式の対数正規分布を利用することにしたわけであるが、その母数は以下のように決定している。

初回のベイズ推定値 $\hat{\theta}_1$ を平均とする2つの正規分布を考える。 $\hat{\theta}_1$ より低い、すなわち数直線上では左側部分となる正規分布と、右側部分となる正規分布の平均は、どちらも $\hat{\theta}_1$ であるが、分散は異なっても良いことにして目的により適宜指定できることとし、この2つの分布に最も近い対数正規分布を事前分布として定めることにした。図1は、 $\hat{\theta}_1 = 0$ 、すなわち左右の正規分布の平均を0、左の正規分布の標準偏差を0.5、右の正規分布の標準偏差

を1.0とした場合の、最も近い対数正規分布を示したものである。このときの対数正規分布の母数 μ は、0.55、母数 σ は0.38となる。また平均は1.86であり、標準偏差は0.73である。もちろん対数正規分布は、0よりおおい範囲の x で定義されるため、図1は対数正規分布を平行移動させて、そのモードと左右の正規分布の平均値の位置を一致させている。2つの正規分布に「最も近い」対数正規分布の導出では、図1の場合と逆に、正規分布の方を左正規分布の標準偏差の3倍だけ平行移動し(図1の場合は $0.5 \times 3 = 1.5$ となる) 2つの分布の差の面積を数値積分で求め、これを最小とする母数を求める値とし、事後的に対数正規分布のモードと正規分布の平均値を一致させている。ただし、これは多少計算時間がかかるため実用性を考慮して、対数正規分布のモードと左右の正規分布の平均値 (= モード) の位置を一致させるという条件の下で簡易推定を行うこととした。対数正規分布の2つの母数の間にこの制約式が増えるため、分布の差の面積の最小値を実質的に1つの母数に関してのみ探索すれば良いことになる。ちなみに簡易推定による対数正規分布の母数 μ の値は、0.56、母数 σ は0.40となる。また、その平均は1.90であり、標準偏差は0.79である。先の結果と大差なく、実際図的に見比べても差異はそれほど感じられない。以下のシミュレーションによる検討も簡易推定による対数正規分布を利用している。対数正規分布の母数を直接指定せず(もちろんこれも可能ではあるが)、このような方法により分布形を定めるのは、推定値の分布が正規分布で近似されることと、学力の伸びの可能性を右側の正規分布の分散に反映することで容易に表現できること、そして誰にも理解しやすいことによる。

以上をまとめると、ここでは初回テストの能力推定値 $\hat{\theta}_1$ を左右2つの正規分布の平均とし、テスト情報関数 $K(\hat{\theta}_1)$ の正の平方根を左側の正規分布の標準偏差とし、右側の正規分布の標準偏差はその2倍の値として、これに近くなるように対数正規分布の母数を定めて

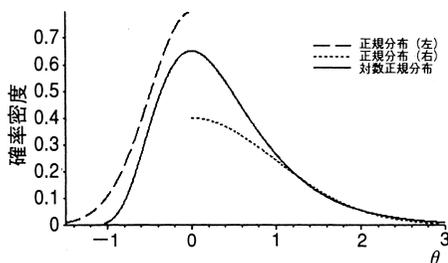


図1 2つの正規分布による対数正規分布の作成

いる。右側正規分布の標準偏差の大きさをどのくらいとするべきかは、学力の伸びの程度及びテスト情報関数との関係の中で決定できるが、集団を対象とする場合は別としても個人の学力の伸びの大きさについては差があり、現時点で厳密に決めることにそれほどの意味があるわけではない。標準偏差の2倍という違いは、対数正規分布による事前分布が、能力分布の0.5標準偏差程度の伸びをカバーできることを意図して設定されたものである。

2.4. シミュレーションデータ

2母数ロジスティックモデルを前提としたシミュレーションデータは、以下のようにして作成した。被験者数はいずれも3000人とする。

データAでは、被験者の能力分布は、標準正規分布に従うと仮定した。テスト項目数は40項目あるいは20項目の2通りとした。テスト項目の2母数ロジスティックモデルの母数の分布型は以下のように定めた。識別力母数は、平均0.85、標準偏差0.25、下限0.3、上限2.0の切断正規分布、また困難度母数は、平均0、標準偏差0.5の正規分布に従うと仮定した。能力母数 θ の被験者のある項目に対する正誤は、2母数ロジスティックモデルから予想される正答確率を、範囲0～1の一樣乱数と比較し、前者が下回る場合被験者の反応を正答1、上回る場合誤答0とする。2母数ロジスティックモデルに従う、この2値データパターンを、項目数40、及び20として各5回繰り返し作成し（データ1～5）、初回テストデータとした。このデータを利用して能力母数と項目母数の同時推定を行い、その際得られた θ の推定値を2.3節の2つの正規分布の平均とし、テスト情報関数 $I(\theta)$ の正の平方根を左側の正規分布の標準偏差とし、右側の正規分布の標準偏差はその2倍の値として、対応する対数正規分布を事前分布とすることにした。また比較のためテスト情報関数の正の平方根を標準偏差とする正規分布を事前分布とするものも作成した。

続いて、 θ 上での能力の伸びを0.25、及び0.5として、初めの θ に加算し、初回とは別のテストを受験したとして、正誤パターンを作成した。0.25から0.5標準偏差の伸びは、小学校3年生から中学校3年生までの算数学力の差からみると約1年分に相当すると思われる（藤森, 1999）。学校などで年に1度の大規模な実力テストなどを行う例が多いことを考えれば、前年度比での能力の伸びを問題にするケースに相当し（もちろん、絶対尺度的な意味での学力の伸びを問題にするのであろうから項目反応理論の利用は必須となる）、本研究が想定する範囲と言えよう。また、この場合の項目母数は、識別力の分布は同一としたが、困難度母数の分布の平均は θ の上昇に見合う分だけ上昇させている。

データBでは、被験者の能力分布は、 $\theta = 0.0$ の1点分布に従うと仮定した。これは、データAでは、 θ の分布を標準正規分布としたため全体的評価は可能であるとしても、ある能力水準 θ_0 の者の能力が向上した場合の推定値がどのような挙動をしているのか、必ずしも明確ではなく、これを検討するために行うものである。初回のテスト正誤データで能力母数と項目母数の推定を行い、その能力推定値で事前分布を決定することは、 θ が1点分布であることを考えれば適当でないため、真の項目母数を利用して能力推定値を求め、これを利用して事前分布を決定している。データBは被験者の分布が異なる点と、前述の事前分布の決定に関する細かい相違を除けば、その他の点でのデータAとシミュレーションの行い方に違いはない。

以上を整理すると、シミュレーションでは、事前分布の違い（正規分布と対数正規分布）が、 θ の分布形（正規分布と1点分布）、項目数（40項目と20項目）、能力の伸び（0.25と0.5）による推定値の違いとどのような関係を持つかを検討する。

3. 結果と考察

表1は、データAの能力分布が正規分布、項目数20問で能力が各被験者の真値から+0.25上伸したときの結果であるが、データ1~5のいずれを見ても、事前分布として正規分布を設定した場合は、平均が0.1前後であるのに対し、対数正規分布を事前分布としたときは、平均が0.2に近くなっており、平均二乗誤差MSEの大きさもやや小さいことがわかる。すなわち対数正規分布の方が良い成績を示している。表2は、表1の各データについて初回真値の水準ごとにグループ化して、誤差(=0.25上伸時真値-推定値)の大きさを見たものである。表より、真値が-1.5より小さい場合を除き、いずれの場合も正規分布より対数正規分布の成績が良いことが分かる。このような現象が起こるのは以下のような理由によると思われる。初回推定は標準正規分布を事前分布とするベイズ推定であり、このことは分布の両端に近い被験者の能力推定値に対して平均0への接近傾向をもたらすことになる。そして、この推定値を2回目の事前分布の平均としているのであるが、対数正規分布は、事前分布として正規分布を採用した場合よりも、能力の伸びに対して推定値がすみや

かに上昇するため、全体的に、その推定値が正規分布の推定値よりプラス側に位置することになる。このため、真値の能力分布の左部分では、平均的には「真値、正規分布の推定値、対数正規分布の推定値」の順になるのに対し、能力分布の右部分では、「正規分布の推定値、対数正規分布の推定値、真値」の順になり、表2の結果がもたらされるのであろう。実際、表3に、表1のデータ1のみであるが、表1の推定値を事前分布として利用した場合の0.5上伸時の推定値の誤差(=0.5上伸時真値-推定値)の大きさを求めているが、能力分布の両端部分に関する傾向は表2の様とかなり異なっており、先の考察が妥当なことを示している。以上より、表2の真値が-1.5より小さい場合の対数正規分布と正規分布の差は極わずかであり、その原因が初期推定値にあることを考慮すれば、対数正規分布の良さを認めてよいだろう。

続いて表4は、データAの項目数20、0.5上伸時の結果である。ただし、事前分布は、初回テスト時の推定値を利用している。0.25上伸時の推定値を利用せず、元の推定値を利用するのは、事前分布の情報に比較して0.5上伸したときの影響を評価するためである。表4より、事前分布として正規分布を利用した場合、その推定値の平均は0.2ほどであり、真の平均0.5に比較して、大幅に下回っている。このためMSEも0.170と表1に比較して大きく悪化している。一方対数正規分布の場合は、同様に悪化しているものの、平均は0.345であり、

表1 データA、20項目の能力0.25上伸時の推定値の平均、SD及びMSE

事前分布	データ1	データ2	データ3	データ4	データ5	平均	
正規分布	θ の平均	0.088	0.114	0.095	0.111	0.115	0.105
	θ のSD	0.980	0.992	1.003	0.983	1.002	0.992
	MSE	0.106	0.101	0.099	0.114	0.111	0.106
対数正規分布	θ の平均	0.170	0.197	0.175	0.196	0.201	0.188
	θ のSD	1.006	1.024	1.032	1.018	1.031	1.022
	MSE	0.103	0.099	0.100	0.109	0.104	0.103

表2 真値の水準ごとの誤差の大きさ(データA、能力0.25上伸、20項目条件)

事前分布	真値範囲	データ1		データ2		データ3		データ4		データ5		全体	
		差の平均	差のSD										
正規分布	~-1.5	-0.091	0.361	-0.058	0.448	-0.051	0.401	-0.017	0.402	0.025	0.401	-0.038	0.403
	-1.5~-0.5	0.129	0.286	0.120	0.279	0.141	0.284	0.099	0.279	0.136	0.304	0.125	0.286
	-0.5~0.5	0.153	0.269	0.143	0.235	0.141	0.235	0.151	0.249	0.167	0.244	0.151	0.246
	0.5~1.5	0.143	0.280	0.144	0.271	0.133	0.269	0.175	0.312	0.175	0.283	0.154	0.283
	1.5~	0.245	0.362	0.254	0.374	0.252	0.366	0.314	0.417	0.238	0.383	0.260	0.380
	全体	0.134	0.297	0.132	0.289	0.134	0.285	0.144	0.305	0.158	0.293	0.130	0.320
対数正規分布	~-1.5	-0.131	0.390	-0.096	0.475	-0.086	0.436	-0.044	0.422	-0.008	0.435	-0.073	0.432
	-1.5~-0.5	0.062	0.314	0.057	0.302	0.079	0.312	0.039	0.305	0.072	0.330	0.062	0.312
	-0.5~0.5	0.067	0.283	0.061	0.250	0.055	0.258	0.062	0.264	0.078	0.259	0.064	0.263
	0.5~1.5	0.058	0.298	0.048	0.298	0.051	0.296	0.074	0.342	0.076	0.296	0.061	0.306
	1.5~	0.098	0.420	0.102	0.435	0.106	0.438	0.161	0.472	0.101	0.441	0.114	0.441
	全体	0.052	0.316	0.050	0.310	0.054	0.312	0.059	0.325	0.072	0.314	0.046	0.351

MSEも0.131と相対的には良好な成績を示している。

表5と表6は、データAの40項目の、0.25及び0.5上伸時の結果である。表5より推定値の平均値は表1と大きく違わないが、事前分布として正規分布を利用した場合も対数正規分布の場合も、MSEは項目数の増加

を反映して改善されている。明らかに表5より、対数正規分布の方が良い成績を示していることが分かる。表6は、0.5上伸時の結果であるが、どちらも表4の20項目の場合と比較して、改善されているが、改善幅は対数正規分布の場合が大きくなっていることが分かる。ベイズ推定では、データ量が増えるにつれて、すなわちここではテスト項目数が増加するにつれて、事前分布の影響力が低下する。しかし、40項目という、テスト利用場面で比較的あり得るケース程度でも、能力の上伸がある場合には、事前分布の影響は無視できない大きさであることが分かる。特に、初回テストの測定値とテスト情報量に基づいて事前分布として正規分布を採用すると、その影響は大きい事が示されたと言えよう。

表7～表10は、データBの1点分布の結果である。データAに示されたように、ここでも事前分布として対数正規分布の方が良い結果を示している。項目数による違いや、上伸幅の違いによる影響もデータAと同様の傾向が示されている。1つだけデータAの結果と比較して特徴的なのは、事前分布として正規分布を採用したときMSEは改善されるものの、20項目の場合と比較して40項目では、平均は逆に真値から遠ざかっていることである。対数正規分布の場合も、0.25上伸時には同様の傾向がある。項目数の増加が能力推定値の分散の縮小を促し、事前分布の影響力が増したことがこの現象の原因と思われる。

表3 真値の水準ごとの誤差の大きさ(データAのデータ1、能力0.5上伸、20項目条件)

真値範囲	人数	正規分布		対数正規分布	
		差の平均	差のSD	差の平均	差のSD
~-1.5	204	0.123	0.459	0.070	0.502
-1.5~-0.5	737	0.259	0.299	0.162	0.345
-0.5~0.5	1177	0.257	0.260	0.126	0.306
0.5~1.5	688	0.212	0.313	0.077	0.395
1.5~	194	0.171	0.454	-0.012	0.552
全体	3000	0.232	0.316	0.111	0.375

注意:初期値は0.25上伸時の推定値

表4 データA、20項目の能力0.5上伸時の推定値の平均、SD及びMSE

事前分布	データ1	データ2	データ3	データ4	データ5	平均	
正規分布	θ の平均	0.190	0.206	0.193	0.227	0.224	0.208
	θ のSD	0.994	0.992	1.002	0.994	0.994	0.995
	MSE	0.171	0.174	0.164	0.165	0.177	0.170
対数正規分布	θ の平均	0.324	0.346	0.322	0.366	0.368	0.345
	θ のSD	1.029	1.029	1.040	1.033	1.033	1.033
	MSE	0.134	0.135	0.132	0.124	0.132	0.131

表5 データA、40項目の能力0.25上伸時の推定値の平均、SD及びMSE

事前分布	データ1	データ2	データ3	データ4	データ5	平均	
正規分布	θ の平均	0.100	0.105	0.126	0.113	0.121	0.113
	θ のSD	1.000	0.987	0.981	0.986	0.984	0.988
	MSE	0.069	0.063	0.071	0.068	0.062	0.067
対数正規分布	θ の平均	0.174	0.180	0.208	0.192	0.198	0.190
	θ のSD	1.017	1.004	0.996	1.003	1.004	1.005
	MSE	0.060	0.056	0.058	0.059	0.053	0.057

表6 データA、40項目の能力0.5上伸時の推定値の平均、SD及びMSE

事前分布	データ1	データ2	データ3	データ4	データ5	平均	
正規分布	θ の平均	0.223	0.233	0.231	0.229	0.219	0.227
	θ のSD	1.000	1.001	0.986	0.990	0.989	0.993
	MSE	0.120	0.106	0.140	0.126	0.131	0.125
対数正規分布	θ の平均	0.361	0.370	0.380	0.375	0.364	0.370
	θ のSD	1.024	1.025	1.006	1.012	1.013	1.016
	MSE	0.076	0.066	0.083	0.076	0.080	0.076

表7 データB、20項目の能力0.25上伸時の推定値の平均、SD及びMSE

事前分布	データ1	データ2	データ3	データ4	データ5	平均	
正規分布	θ の平均	0.097	0.091	0.102	0.110	0.102	0.100
	θ のSD	0.220	0.229	0.246	0.223	0.225	0.229
	MSE	0.072	0.078	0.083	0.070	0.072	0.075
対数正規分布	θ の平均	0.174	0.175	0.186	0.194	0.184	0.183
	θ のSD	0.239	0.253	0.262	0.236	0.247	0.247
	MSE	0.063	0.070	0.073	0.059	0.065	0.066

表8 データB、20項目の能力0.5上伸時の推定値の平均、SD及びMSE

事前分布	データ1	データ2	データ3	データ4	データ5	平均	
正規分布	θ の平均	0.192	0.217	0.194	0.211	0.203	0.203
	θ のSD	0.225	0.213	0.244	0.227	0.225	0.227
	MSE	0.145	0.125	0.153	0.135	0.139	0.140
対数正規分布	θ の平均	0.334	0.357	0.336	0.357	0.348	0.346
	θ のSD	0.263	0.243	0.273	0.256	0.258	0.259
	MSE	0.097	0.079	0.101	0.086	0.090	0.091

以上の結果より、能力分布の0.25～0.5標準偏差の上伸が期待され、その間に複数回の能力測定が行われるときには、正規分布を事前分布とするよりも対数正規分布を事前分布とすることが適当であることが示されたといえよう。

本研究では、能力の上伸が期待できる場面

表9 データB、40項目の能力0.25上伸時の推定値の平均、SD及びMSE

事前分布		データ1	データ2	データ3	データ4	データ5	平均
正規分布	θ の平均	0.101	0.096	0.090	0.090	0.099	0.095
	θ のSD	0.156	0.169	0.165	0.174	0.163	0.165
	MSE	0.047	0.052	0.053	0.056	0.049	0.051
対数正規分布	θ の平均	0.177	0.172	0.166	0.161	0.177	0.171
	θ のSD	0.174	0.178	0.182	0.188	0.172	0.179
	MSE	0.036	0.038	0.040	0.043	0.035	0.038

表10 データB、40項目の能力0.5上伸時の推定値の平均、SD及びMSE

事前分布		データ1	データ2	データ3	データ4	データ5	平均
正規分布	θ の平均	0.175	0.194	0.189	0.193	0.210	0.192
	θ のSD	0.161	0.173	0.164	0.166	0.162	0.165
	MSE	0.131	0.124	0.123	0.122	0.110	0.122
対数正規分布	θ の平均	0.332	0.349	0.344	0.344	0.369	0.348
	θ のSD	0.200	0.197	0.196	0.196	0.184	0.195
	MSE	0.068	0.062	0.063	0.063	0.051	0.061

で、事前情報を利用して能力の推定を行うことを目指した。ソフトウェアも含めて実用性を考慮してのものであるだけに、シミュレーションでなく実データへの適用を通じて、その評価を行うことが今後の課題と言えよう。

文献

- Birnbaum, A. 1968 Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick(Eds.), *Statistical theories of mental test scores*(pp. 395-479). Reading, MA : Addison-Wesley .
- Embretson, S. E. 1991 A multidimensional latent trait model for measuring learning and change . *Psychometrika*, 56, 495-516 .
- 藤森進 1999 算数・数学学力の到達度水準に関する発達的研究(研究課題番号08610130)平成8年度～平成10年度科学研究費補助金(基盤研究(C)(2))研究成果報告書 .
- 藤森進 2002 テスト得点を統計的枠組みで分析する 項目反応理論 渡部洋編「心理統計の技法」第7章 福村出版 .
- 渡部洋 1999 ベイズ統計学入門 福村出版 .
- Wilson, M. 1989 Saltus : A psychometric model of discontinuity in cognitive development. *Psychological Bulletin*, 105, 276 - 289 .