外国語教育メディア学会（LET）関西支部 メソドロジー研究部会 2014 年度 第 6 号報告論集
住 政二郎 （pp. 96–116）

PROX 法と同時最尤推定法の概説

住 政二郎

流通科学大学

概要

本稿の目的は，PROX 法（the Normal Approximation Algorithm Method）と同時最尤推定法について概説することである。PROX 法と同時最尤推定法は，応答データから受験者能力と項目困難度を推定するために使われる。著者は，これまでラッシュモデルの導出（住, 2013）と項目反応理論の各モデル（住, 2014）についてまとめてきた。その後，プレイスメント・テストや教材開発に各モデルを利用してきた（住, 2014）。その際，受験者能力と項目困難度の推定には *R* と *Winstep* を利用してきた。しかし，どのような計算過程を経て推定結果が出力されているのかについては十二分に理解していなかった。幸いにも大友（1996）には PROX 法について，静（2007）には同時最尤推定法について詳細な概説がある。しかし，わずかではあるが記載内容に誤りと紙面の制約から説明が十分とはいえない箇所がある。この 2 冊は外国語教育研究にとって財産ともいえる貴重な書籍である。本稿は著者が一人の学習者となり，この 2 冊を通読し，その理解をまとめたものである。内容は研究ノート程度のものであるが，必要だと思われる箇所に解説を加え，計算過程が再現できるようにデータを公開した。これまで実践現場への応用には敷居の高かった項目反応理論ではあるが，*R*，そして，関西大学の水本先生が開発された lang.test のおかげで身近に使えるものになった。本稿がこれから項目反応理論を学ぶ読者の助けになれば幸いである。

Keywords: 項目反応理論，PROX 法，同時最尤推定法

1.　受験者能力と項目困難度の推定

テストの結果などの反応データから受験者能力（θ）と項目困難度（δ または b）を推定する場合，以下の 3 つのケースが考えられる。

1. すでに受験者能力が分かっており，項目困難度を推定する場合
2. すでに項目困難度が分かっており，受験者能力を推定する場合
3. 受験者能力と項目困難度のいずれも分かっていない場合

教育現場では上記3のケースが一般的ではないだろう。つまり，テストを実施し，その応答データがあり，そこから受験者能力と項目困難度を推定する，という場合である。表1は，あるテストを実施した結果の応答データである。我々は日常的にこのような応答データを目にしている。こうした結果から受験者能力と項目困難度を推定するためには，どのような準備が必要なのだろうか。

表1
受験者応答データ1

	項目1（δ_1 =?）	項目2（δ_2 =?）	項目3（δ_3 =?）	項目4（δ_4 =?）	合計
受験者1（θ_1 =?）	1	1	0	1	3
受験者2（θ_2 =?）	1	0	0	1	2
受験者3（θ_3 =?）	0	0	1	1	2
合計	2	1	1	3	7

表2は，表1を符号化したものである。表2の x_{ni} は，受験者 n が，問題項目 i に解答した応答データである。各受験者の合計点は r_n で，各問題項目の合計点は s_i である。応答データとは，受験者 n が問題項目 i を受験して得られた観測得点（observed score）である。

観測得点は誤差得点（erro score）を含む。各受験者および問題項目は，反応データとして得られた観測得点の他に，期待得点（expected score）を潜在的に有している。受験者の期待得点とは，受験者 n が真の力を発揮した時に期待される得点である。問題項目の期待得点とは，直感的には理解しにくいが，問題項目 i が真の難しさを発揮した時に期待される得点である（参照：靜, 2007, pp. 108–110）。

表2
受験者応答データ2

	項目1（δ_1 =?）	項目2（δ_2 =?）	項目3（δ_3 =?）	項目4（δ_4 =?）	合計
受験者1（θ_1 =?）	x_{11}	x_{12}	x_{13}	x_{14}	r_1
受験者2（θ_2 =?）	x_{21}	x_{22}	x_{23}	x_{24}	r_2
受験者3（θ_3 =?）	x_{31}	x_{32}	x_{33}	x_{34}	r_3
受験者4（θ_4 =?）	x_{41}	x_{42}	x_{43}	x_{44}	r_4
受験者5（θ_5 =?）	x_{51}	x_{52}	x_{53}	x_{54}	r_5
合計	s_1	s_2	s_3	s_4	

誤差を含む観測得点から受験者能力と項目困難度を推定する，という作業は，つまり，応答データの観測得点と潜在的な期待得点を限りなく近づけることができる受験者能力と項目困難度の組み合わせを推定する，という作業に等しい。この数学的証明は本稿の付録に添えた。この作業の終了条件について靜（2007, p. 250）は，以下のようにまとめている。

1. 各受験者の期待得点が，各受験者の実際の得点に等しい
2. 各項目の期待得点が，各項目の実際の得点に等しい

応答データの観測得点と潜在的な期待得点を限りなく近づけることができる受験者能力と項目困難度の組み合わせを推定する作業は「本質的には試行錯誤するしかない」(靜, 2007, p. 251）が，効率的にその作業を実施する手法は整備されている。大友（1996, p. 120）は，以下 6 つの推定方法を紹介している。

a. Maximum Likelihood Estimation：項目パラメーターの値（a, b, c）がわかっているとき，それを用いて能力パラメータ（θ）を最尤推定する方法

b. Maximum Likelihood Estimation：能力パラメータ（θ）がわかっているとき，それを用いて項目パラメータ（a，b，c）を最尤推定する方法

c. Joint Maximum Likelihood Estimation：項目パラメーターと能力パラメータを同時に最尤推定する方法

d. Marginal Maximum Likelihood Estimation：能力の周辺分布を利用して，項目パラメーターを最尤推定する方法

e. Joint and Marginal Baysian Estimation：同時および周辺分布によるベイス推定方法

f. Heuristic Method：ある仮定を必要とするが，計算の速い簡便法

上記 a，b に関しては靜（2007, pp. 223–239）に詳しい解説がある。大友（1996）は PROX 法を詳細に解説している。本稿では PROX 法に加え，最も一般的な上記 c の同時最尤推定法（Joint Maximum Likelihood Estimation）について概説する。

2. PROX 法

PROX 法は計算過程がシンプルであり，表計算ソフトを使って受験者能力と項目困難度の推定ができる。一般的に受験者能力と項目困難度を同時に推定するためには，複雑な連立方程式を終了条件に至るまで繰り返し計算しなければいけない。この作業にはコンピュータが欠かせない。一方，PROX 法は，受験者能力と項目困難度が正規分布に近似であると仮定することで計算過程を省略することができる（大友, 1996，p. 99）。以下，大友（1996, pp. 98–109）に沿って PROX 法を概説する。大友（1996）にはすでに PROX 法の詳細な解説があるが，紙面だけでは PROX 法の簡便さを体験するには不十分である。そこで本稿では，計算過程を再現できるようにデータを公開した（http://goo.gl/nEbtGH)。以下，このファイルを DATA 1 とする。また，大友（1996）には不偏分散の代入式にわずかではあるが誤りがある。本稿ではその部分を指摘し，その他必要と思われる解説を加えた。

2.1　項目困難度と受験者能力の線形化

受験者 5 名が，問題項目 5 問のテストを受験し，表 3 の応答データが得られた。

表 3
受験者応答データ 3

受験者番号	項目	1	2	3	4	5	計
20001		1	0	0	0	1	2
20002		1	1	1	1	0	4
20003		1	1	1	1	0	4
20004		1	1	0	1	0	3
20005		1	1	0	0	1	3
		5	4	2	3	2	

表 3 の内，項目 1 は全員が正解をしているために除外する。これは全員が正解し，項目 1 は推定に必要な情報量を持っていないと考えられるからである。表 4 は項目 1 を除外した応答データである（参照：DATA 1, STEP 0）。

表 4
受験者応答データ 4

受験者番号	項目	2	3	4	5	計
20001		0	0	0	1	1
20002		1	1	1	0	3
20003		1	1	1	0	3
20004		1	0	1	0	2
20005		1	0	0	1	2
		4	2	3	2	

応答データから受験者能力と項目困難度を同時に推定するためには，まず各項目の正答数と誤答数を使い，受験者能力と項目困難度の線型化を行う必要がある。その理由は，素点に基づく応答データは，受験者能力や項目困難度の高低に関する正確な情報を有していないからである。例えば，受験者 20001 の正答数が 1 で，受験者 20002 の正答数が 3 の場合，感覚的に受験者 20002 の方が能力が高いことは分かる。しかし，受験者 20001 と受験者 20002 の能力の差と，受験者 20001 と受験者 20004 の能力の差には，どのくらい能力の差があるのかということは，素点に基づく応答データからでは本質的には分からない。同様に，正答数が 2 の項目 5 の方が，正答数が 4 の項目 2 よりも難しいことは感覚的に分かる。しかし，項目 5 と項目 3 の難しさの差と，項目 5 と項目 2 の難しさの差には，どのくらい難しさの差があるのかも，素点に基づく応答データからでは本質的には分からない。これは正答数が順序尺度上の応答データであるためである。受験者能力と項目困難度を推定するためには，受験者能力と項目困難度を連続的に変化する指標に変形し，間隔尺度上の数値におきかえなくてはいけない。そのために必要な作業が線型化である。正答数に基づく得点の問題点と素点の非線形性については，静（2007, pp. 151–156）と大友（1996, pp.93–94）が詳しい。

線型化の作業は，正答と誤答のオッズを対数変換したログ・オッズを使う。単位は logits である。オッズとは，標本空間において，ある事象が起こる確率と起こらない確率の比である。*Probability* は正答率である。1 − *Probability* は誤答率である。ログ・オッズに関しては静（2007, p. 23）と住（2013, pp. 90–93）が詳しい。正答数を確率変数に変換し，さらに対数変換することによって，ある値を $-\infty$ から $+\infty$ まで理論的には連続的に無限に変化する値に変化させることができる（参照：住, 2014, p. 38）。

$$\text{正答の } Odds = \frac{Probability}{1 - Probability} \tag{1}$$

$$\text{正答の } Odds \text{ の対数} = \log\left(\frac{Probability}{1 - Probability}\right) \tag{2}$$

$$\text{誤答の } Odds = \frac{1 - Probability}{Probability} \tag{3}$$

$$\text{誤答の } Odds \text{ の対数} = \log\left(\frac{1 - Probability}{Probability}\right) \tag{4}$$

2.2　項目困難度の線形化

大友（1996, p. 100）は項目困難度の線型化作業から行っている。表 5 は，その結果である。受験者能力の線型化作業を最初に行っても問題はないが，おそらくこれは後述する項目困難度の基準設定についての理解を深めるための配慮と考えられる。項目困難度の線型化作業には，誤答のログ・オッズを用いる（式 4）（参照：DATA 1, STEP 1）。

表 5
誤答のログオッズ

項目番号	正答数	正答率	誤答率	誤答のログ・オッズ
2	4	0.80	0.20	-1.386
3	2	0.40	0.60	0.405
4	3	0.60	0.40	-0.405
5	2	0.40	0.60	0.405
	合計	2.20	1.80	-0.981
	平均	0.55	0.45	-0.245

2.3　項目困難度の基準設定と初期項目困難度の計算

各項目の誤答のログ・オッズ値は，項目 2：-1.386, 項目 3：0.405，項目 4：-0.405，項目 5：0.405 である。合計は-0.981 であり，平均値は-0.245 である。次に，すべての誤答のログ・オッズ値から，平均値の-0.245 を引き，初期項目困難度を計算する。これは項目困難度の原点を定める重要な調整作業である。後述する同時最尤推定法でも同様の作業が行われる。この作業によって受験者集団から項目を切り離し，独自の水準を定めることができる（大友, 1996, pp. 101–102）。表 6 は初期項目困難度を計算したものである。初期項目困難度の合計と平均が 0 になっていることに注目して欲しい（参照：DATA 1, STEP 2）。

2.4　受験者能力の線形化と初期受験者能力の計算

受験者能力の線型化は正答のログ・オッズを使う（式 2）。受験者能力では項目困難度のような基準設定作業は行わない。これは初期項目困難度を計算するために，項目の側ですでに原点を定める作業を行っているからである。この点は後述する靜（2007）の同時最尤推定法と大きく異なる点である。受験者能力に関しては，正答のログ・オッズがそのまま初期受験者能力となる（表 7）（参照：DATA 1, STEP 3）。

表 6
初期項目困難度

項目番号	正答数	正答率	誤答率	誤答のログ・オッズ	初期項目困難度
2	4	0.80	0.20	-1.386	-1.141
3	2	0.40	0.60	0.405	0.651
4	3	0.60	0.40	-0.405	-0.160
5	2	0.40	0.60	0.405	0.651
	合計	2.20	1.80	-0.981	0.00
	平均	0.55	0.45	-0.245	0.00

表 7
初期受験者能力

項目番号	正答数	正答率	誤答率	正答のログ・オッズ	初期受験者能力
20001	1	0.25	0.75	-1.099	-1.099
20002	3	0.75	0.25	1.099	1.099
20003	3	0.75	0.25	1.099	1.099
20004	2	0.50	0.50	0.000	0.000
20005	2	0.50	0.50	0.000	0.000
	合計	2.75	2.25	1.099	1.099
	平均	0.55	0.45	0.220	0.220

2.5　項目困難度と受験者能力の不偏分散

応答データは誤差を含む。誤差はデータのばらつきとなる。項目困難度と受験者能力の不偏分散を計算すると，応答データのばらつき度合い調べることができる。不偏分散は，以下の公式のとおりである。

$$\text{不偏分散} = \frac{\Sigma(X-\overline{X})^2}{n-1} \tag{5}$$

分子の $\Sigma(X-\overline{X})^2$ は，各項目の値から平均値を引いて 2 乗し，その値をすべて合計するという意味である。項目困難度では，誤答のログ・オッズから平均値を引き 2 乗し（表 8)，受験者能力では，正答のログ・オッズから平均値を引き 2 乗する（表 9)。分母の $n-1$ は，データ数から 1 を引いたものを意味する。項目数は 4 問なので 3，受験者数は 5 人なので 4 になる。故に，項目困難度の不偏分散は 2.175/3 = 0.725 で，これを U とする。受験者能力

の普遍分散は 3.379/4 = 0.845 で，これを V とする。尚，大友（1996, p. 106）の受験者能力の不偏分散を求める代入式で，ロジット・コレクトの 2 乗の総和が，正しくは 3.621 の所が 3.625 で代入されている（参照：DATA 1, STEP 4 & 5）。

表 8
項目困難度 $\Sigma(X - \overline{X})^2$

項目番号	正答数	正答率	誤答率	ログ・オッズ	$(X - \overline{X})^2$
2	4	0.80	0.20	-1.386	1.302
3	2	0.40	0.60	0.405	0.423
4	3	0.60	0.40	-0.405	0.026
5	2	0.40	0.60	0.405	0.423
合計		2.20	1.80	-0.981	2.175
平均		0.55	0.45	-0.245	0.544

表 9
受験者能力 $\Sigma(X - \overline{X})^2$

項目番号	正答数	正答率	誤答率	ログ・オッズ	$(X - \overline{X})^2$
20001	1	0.25	0.75	-1.099	1.738
20002	3	0.75	0.25	1.099	0.772
20003	3	0.75	0.25	1.099	0.772
20004	2	0.50	0.50	0.000	0.048
20005	2	0.50	0.50	0.000	0.048
合計		2.75	2.25	1.099	3.379
平均		0.55	0.45	0.220	0.676

2.6　最終項目困難度と最終受験者能力

項目困難度と受験者能力の不偏分散を使い拡張要素を計算することで，応答データの標本による分散の違いを取り除くことができる。これは受験者能力と項目困難度が正規分布に近似であると仮定することで可能となり，PROX 法が繰り返しの計算を省略できる理由にもなっている。

項目困難度の拡張要素は，以下のとおりである。

$$\text{項目困難度の拡張要素} = \sqrt{\frac{\left(\dfrac{1+V}{2.89}\right)}{\left(\dfrac{1-UV}{8.35}\right)}} \tag{6}$$

受験者能力の拡張要素は，以下のとおりである。

$$\text{受験者能力の拡張要素} = \sqrt{\frac{\left(\dfrac{1+U}{2.89}\right)}{\left(\dfrac{1-UV}{8.35}\right)}} \tag{7}$$

拡張要素の 2.89 は 1.7 を 2 乗した値で，8.35 は 1.7 を 4 乗した値である。1.7 は，標準正規分布の正規累積モデル（オージブモデル）をロジスティックモデルに変換する際に，近似値になるように導入されたものである（参照：住, 2014, p. 45）。

$U = 0.725$，$V = 0.845$ であることから

$$\begin{aligned}\text{項目困難度の拡張要素} &= \sqrt{\frac{\left(\frac{1+V}{2.89}\right)}{\left(\frac{1-UV}{8.35}\right)}} \\ &= \sqrt{\frac{\left(\frac{1+0.845}{2.89}\right)}{\left(\frac{1-0.725\times 0.845}{8.35}\right)}} \\ &= 1.181\end{aligned}$$

$$受験者能力の拡張要素 = \sqrt{\frac{\left(\frac{1+U}{2.89}\right)}{\left(\frac{1-UV}{8.35}\right)}}$$

$$= \sqrt{\frac{\left(\frac{1+0.725}{2.89}\right)}{\left(\frac{1-.725 \times 0.845}{8.35}\right)}}$$

$$= 1.162$$

最後に，項目困難困難度の拡張要素 1.181 を初期項目困難度に掛けると，最終項目困難度が得られる（表 10）。同様に，受験者能力の拡張要素 1.162 を初期受験者能力に掛けると，最終受験者能力が得られる（表 11）。最終項目困難度の合計と平均が共に 0 になっていることに注目して欲しい（参照：DATA 1, STEP 6 & 7）。

表 10
最終項目困難度

項目番号	誤答のログ・オッズ	初期項目困難度	拡張要素	最終項目困難度
2	-1.386	-1.141	1.181	-1.348
3	0.405	0.651	1.181	0.768
4	-0.405	-0.160	1.181	-0.189
5	0.405	0.651	1.181	0.768
合計	-0.981	0.00		0.00
平均	-0.245	0.00		0.00

表 11
最終受験者能力

項目番号	正答のログ・オッズ	初期受験者能力	拡張要素	最終受験者能力
20001	-1.099	-1.099	1.162	-1.276
20002	1.099	1.099	1.162	1.276
20003	1.099	1.099	1.162	1.276
20004	0.000	0.000	1.162	0.000
20005	0.000	0.000	1.162	0.000
合計	1.099	1.099		1.276
平均	0.220	0.220		0.255

3. 同時最尤推定法

同時最尤推定法については，静（2007, pp.240–250）が詳しい。静（2007）は，なぜ応答データの観測得点から潜在的な受験者能力と項目困難度を推定することが原理的に可能なのか，数式を丁寧に展開し文系読者にも分かりやすく解説している。しかし，紙面だけで同時最尤推定法の実例を理解することは難しい。また，同時最尤推定法を体験するには不十分である。そこで本稿では計算過程を再現できるようにデータを公開した（http://goo.gl/UE9cze）。以下，このファイルを DATA 2 とする。本稿に合わせて参考にして頂きたい。本稿文末には静（2007, p. 240–249）に記載されている展開式に，著者が理解に苦しんだ部分に解説を加えた付録を用意した。同時最尤推定法の原理と導出に興味のある読者には参考にして頂きたい。

3.1 同時最尤推定法の手順

静（2007）では，同時最尤推定法の手順として，以下 6 つのステップを定めている。

ステップ 1

すべての項目の難度が 0.0 だと仮定したうえで，受験者能力の初期値を計算する

ステップ 2

すべての受験者の能力が等しく 0.0 であると仮定した上で，各項目難度の初期値を計算する

ステップ 3

ステップ 1 で得られた各受験者の初期能力値と，ステップ 2 で得られた各項目の初期難度値を用いて，各受験者および各項目の期待得点と観測得点を比較する

ステップ 4

各受験者について期待得点と実際の得点を比較し，ニュートン・ラフソン法によってより適切な能力値にアップデートする

ステップ 5

各項目について期待得点と実際の得点を比較し，ニュートン・ラフソン法によってより適切な難度値にアップデートする

ステップ 6

全項目の難度の平均値が 0.00 になるよう，全項目および全受験者の推定値にある一定の値を加える，もしくは引く

以下，6 つのステップに従い，以下の表 12 を実例に同時最尤推定法の再現を行う（参照：靜, 2007, p. 260; DATA 2, STEP 0）。

表 12
受験者応答データ 5

Observed	item 1	item 2	item 3	item 4	item 5	Correct Total
person 1	1	1	1	1	0	4
person 2	1	1	0	0	1	3
person 3	1	1	0	1	0	3
person 4	1	0	1	0	0	2
person 5	0	1	0	0	0	1
Correct Total	4	4	2	2	2	

3.2　ステップ 1–3

受験者能力と項目困難度の初期値を計算するために，以下の計算（ステップ 1–3）を行う。

1. 正答数と誤答数を使って，正答のログ・オッズと誤答のログ・オッズを計算（ステップ 1 & 2）
2. 誤答のログ・オッズの平均値を計算
3. 正答のログ・オッズと誤答のログ・オッズから上記 2 の誤答のログ・オッズの平均値を引き，受験者能力と項目困難度の初期値を計算
4. 上記 3 で得られた受験者能力と項目困難度の初期値をラッシュ・モデルに代入し，期待得点を計算（ステップ 3）

表 13 は上記 3 までの結果である（参照：DATA 2, STEP 1-1）。誤答のログ・オッズの平均値は-0.115 であった。正答のログ・オッズと誤答のログオッズから誤答のログ・オッズの平均値-0.115 を引く。これは，PROX 法の「2.3 項目困難度の基準設定と初期項目困難度の計算」での処理と同様のものである。靜（2007）の同時最尤推定法では，正答のログ・オッズからも誤答のログ・オッズの平均値を引き初期受験者能力を計算している。初期項目困難度の平均が 0 になっていることに注目して欲しい。

表 13
初期受験者能力と初期項目困難度

Observed	item 1	item 2	item 3	item 4	item 5	Correct Total	Incorrect Total	Logit Correct	Init. Logit Correct
person 1	1	1	1	1	0	4	1	1.386	1.501
person 2	1	1	0	0	1	3	2	0.405	0.521
person 3	1	1	0	1	0	3	2	0.405	0.521
person 4	1	0	1	0	0	2	3	-0.405	-0.290
person 5	0	1	0	0	0	1	4	-1.386	-1.271
Correct Total	4	4	2	2	1				
Incorrect Total	1	1	3	3	4				
Logit Incorrect	-1.386	-1.386	0.405	0.405	1.386	-0.115			
Init. Logit Incorrect	-1.271	-1.271	0.521	0.521	1.501	0.00			

表 14 は，受験者能力と項目困難度の初期値をラッシュ・モデルに代入し期待得点を計算したものである（ステップ 3）。静（2007, p. 259）の図 10.32 と同じ結果になっている（参照：DATA 2, STEP 1-2）。

表 14
期待得点の計算

Expected	item 1	item 2	item 3	item 4	item 5	Total	Logit Correct	Init. Logit Correct
person 1	0.941	0.941	0.727	0.727	0.500	3.837	1.386	1.501
person 2	0.857	0.857	0.500	0.500	0.273	2.987	0.405	0.521
person 3	0.857	0.857	0.500	0.500	0.273	2.987	0.405	0.521
person 4	0.727	0.727	0.308	0.308	0.143	2.213	-0.405	-0.290
person 5	0.500	0.500	0.143	0.143	0.059	1.345	-1.386	-1.271
Total	3.883	3.883	2.178	2.178	1.247			
Logit Incorrect	-1.386	-1.386	0.405	0.405	1.386	-0.115		
Init. Logit Incorrect	-1.271	-1.271	0.521	0.521	1.501	0.00		

3.3　ニュートン・ラフソン

ここまでの作業で，受験者能力の初期値，項目困難度の初期値，観測得点，期待得点の 4 つが得られた。ここから観測得点と期待得点の差を限りなく 0 に近づける作業を行う。これは「(受験者能力と項目難度の）観測得点」–「(受験者能力と項目難度の）期待得点」によって得られる「(受験者能力と項目難度の）残差」(residuals）を限りなく 0 に近づける作業と同義である。この作業は，同じ作業を繰り返す試行錯誤であるが，残差を 0 に近づける効率的な方法にニュートン・ラフソン法がある。

受験者能力の期待得点を段階的に更新し，観測得点との残差を限りなく 0 に近づける
ニュートン・ラフソン法の式は以下のとおりである。

$$\theta'_n = \theta_n + \frac{r_n - \sum_{i=1}^{I} E(x_{ni})}{\sum_{i=1}^{I} E(x_{ni})} \tag{8}$$

式（8）の意味は，以下のとおりである。

$$最新の受験者\ n\ の能力 = 更新された受験者\ n\ の能力 + \frac{(観測得点 - 更新された受験者\ n\ の期待得点の合計)}{更新された受験者\ n\ の期待得点の分散}$$

第 2 項の分母の「更新された受験者 n の期待得点の分散」は，「(期待）正答確率」×「(期待）誤答確率」，つまり $P(1-P)$ で得られる（靜, 2007, p. 92)。

項目困難度の期待得点を段階的に更新し，観測得点との残差を限りなく 0 に近づける
ニュートン・ラフソン法の式は以下のとおりである。

$$\delta'_i = \delta_i - \frac{s_i - \sum_{n=1}^{N} E(x_{ni})}{\sum_{n=1}^{N} E(x_{ni})} \tag{9}$$

式（9）の意味は，以下のとおりである。

$$最新の項目\ i\ の困難度 = 更新された項目\ i\ の困難度 - \frac{(観測得点 - 更新された項目\ i\ の期待得点の合計)}{更新された項目\ i\ の期待得点の分散}$$

3.4　ステップ 4–6

表 15 は，表 14 を再掲したもので，受験者能力と項目困難度の初期値，そして期待得点の値を示している（参照：DATA 2, STEP 1-2）。

表 15
受験者能力・項目困難度・期待得点

Expected	item 1	item 2	item 3	item 4	item 5	Total	Logit Correct	Init. Logit Correct
person 1	0.941	0.941	0.727	0.727	0.500	3.837	1.386	1.501
person 2	0.857	0.857	0.500	0.500	0.273	2.987	0.405	0.521
person 3	0.857	0.857	0.500	0.500	0.273	2.987	0.405	0.521
person 4	0.727	0.727	0.308	0.308	0.143	2.213	-0.405	-0.290
person 5	0.500	0.500	0.143	0.143	0.059	1.345	-1.386	-1.271
Total	3.883	3.883	2.178	2.178	1.247			
Logit Incorrect	-1.386	-1.386	0.405	0.405	1.386	-0.115		
Init. Logit Incorrect	-1.271	-1.271	0.521	0.521	1.501	0.00		

表 15 の期待得点から「期待得点の分散」を計算すると，以下の表 16 になる。期待得点の分散は $P(1 - P)$ で得られる（参照：DATA2, STEP 2-1）。

表 16
期待得点の分散 1

Variance of Expected	item 1	item 2	item 3	item 4	item 5	Total
person 1	0.055	0.055	0.198	0.198	0.250	0.757
person 2	0.122	0.122	0.250	0.250	0.198	0.943
person 3	0.122	0.122	0.250	0.250	0.198	0.943
person 4	0.198	0.198	0.213	0.213	0.122	0.945
person 5	0.250	0.250	0.122	0.122	0.055	0.800
Total	0.749	0.749	1.034	1.034	0.825	

次に，ニュートン・ラフソン法を使い受験者能力と項目困難度を更新する（参照：DATA 2, STEP 2-2, J44–48 & C51–G51）。更新された項目困難度は，平均値を計算し，更新された項目困難度のすべてから引く（参照：DATA 2, STEP 2-2, C52–G52）。最後に，更新された受験者能力と項目困難度をラッシュモデルに代入し期待得点を更新する（参照：DATA 2, STEP 2-2）。表 17 は更新された期待得点を表している。受験者能力の残差が最大で-0.105，項目困難度の残差が最大で 0.096 になっている。これで 1 度目の反復計算は終わりである。尚，期待得点の調整方法が靜（2007）とは異なるため，出力結果が書籍のものとわずかに異なる。

表 17
反復 1：期待得点

Expected	item 1	item 2	item 3	item 4	item 5	Total	Residuals	Updated Ability
person 1	0.961	0.961	0.748	0.748	0.495	3.915	0.085	1.717
person 2	0.884	0.884	0.477	0.477	0.231	2.953	0.047	0.534
person 3	0.884	0.884	0.477	0.477	0.231	2.953	0.047	0.534
person 4	0.727	0.727	0.242	0.242	0.095	2.033	-0.033	-0.516
person 5	0.448	0.448	0.089	0.089	0.031	1.105	-0.105	-1.702
Total	3.904	3.904	2.033	2.033	1.085			
Residuals	0.096	0.096	-0.033	-0.033	-0.085			
Updated Difficulty	-1.428	-1.428	0.693	0.693	1.801	0.066		
Adjusted Difficulty	-1.494	-1.494	0.626	0.626	1.735	0.000		

表 17 の結果から再び期待得点の分散を計算し，ニュートン・ラフソン法を使って受験者能力と項目困難度を更新する。更新された項目困難度から平均値を計算し，更新された項目困難度のすべてから引く。そして, 更新された受験者能力と項目困難度をラッシュモデルに代入し，再び期待得点を更新する。この反復計算を 8 回行った結果が表 18 である。受験者能力と項目困難度の残差は共に最大で ±0.001 になっている。静（2007, p. 263, 図 10.38）の結果ともほぼ一致する。計算過程については DATA 2, STEP 3–9 を参照にして頂きたい。興味のある読者には計算過程が再現できるように R のコードも公開した（http://rpubs.com/seisumi/jmle）。尚，結果は反復回数と推定方法の違いからわずかに異なる。

表 18
反復 8：期待得点

Expected	item 1	item 2	item 3	item 4	item 5	Total	Residuals	Updated Ability
person 1	0.976	0.976	0.775	0.775	0.496	3.999	0.001	1.979
person 2	0.917	0.917	0.479	0.479	0.208	2.999	0.001	0.658
person 3	0.917	0.917	0.479	0.479	0.208	2.999	0.001	0.658
person 4	0.758	0.758	0.208	0.208	0.070	2.000	0.000	-0.598
person 5	0.432	0.432	0.060	0.060	0.018	1.001	-0.001	-2.013
Total	3.999	3.999	2.000	2.000	1.001			
Residuals	0.001	0.001	0.000	0.000	-0.001			
Updated Difficulty	-1.739	-1.739	0.741	0.741	1.992	-0.001		
Adjusted Difficulty	-1.738	-1.738	0.742	0.742	1.993	0.000		

4. まとめ

本稿では，項目反応理論で受験者能力と項目困難度の推定に使われる PROX 法と同時最尤推定法について概説をした。PROX 法については大友（1996）を，同時最尤推定法については靜（2007）を参考にした。この 2 冊は外国語教育学研究にとって大変貴重な書籍である。項目反応理論についてのみならず，日常的な教育実践に深く関わる測定と評価についても多くを学ぶことができる。しかし，著者の理解力の不足も重なり，紙面だけでは十二分に計算過程を理解したり，項目反応理論の醍醐味を体感することができないことを感じた。項目反応理論を手軽に試すことができる環境が整いつつある中で残念なことだと感じた。そこで本稿では，著者が一人の学習者となり，大友（1996）と靜（2007）を通読し理解に苦しんだ箇所に解説を加え，また再現可能なようにデータを公開した。通常はソフトウェアを使って済ませてしまう推定作業だが，じっくりと計算過程を追いかけることで理解は一段と深まる。項目反応理論は，順序尺度に過ぎない応答データを間隔尺度におきかけ，受験者能力と項目困難度を分離し，それぞれを客観的な指標で検証可能にするとても魅力的なものである。本稿がこれから項目反応理論を学ぶ読者の一助となれば幸いである。そして誤りや不十分な点は，新たな読者に是非とも指摘して頂きたい。

参考文献

大友 賢二（1996）.『項目応答理論入門』大修館書店.

靜 哲人（2007）.『基礎から深く理解するラッシュモデリング：項目応答理論とは似て非なる測定のパラダイム』関西大学出版.

住 政二郎（2013）.「ラッシュモデルの導出」『メソドロジー研究部会 2012 年度報告論集』83–101. Retrieved from http://www.mizumot.com/method/2012-07_Sumi.pdf

住 政二郎（2014）.「項目反応理論：1PLM, 2PLM, 3PLM」『メソドロジー研究部会 2013 年度報告論集』34–62. Retrieved from http://www.mizumot.com/method/04-04_Sumi.pdf

住 政二郎（2014）.「ICT を活用した 2 年間を一貫した大学英語教育の取り組み」吉田 晴世・野澤 和典（編著）『最新 ICT を活用した私の外国語授業』(pp. 42—52.) 丸善プラネット.

付録

以下は，静（2007, pp. 242–250）にある同時確率の展開式に解説を加えたものである。

$$L = \prod_{n=1}^{N}\prod_{i=1}^{I}\frac{\exp[x_{ni}(\theta_n - \delta_i)]}{1+\exp(\theta_n - \delta_i)} \tag{10}$$

$\prod$ の要素は分母および分子に掛かることから，式（10）は以下のように変形できる。

$$= \frac{\prod_{n=1}^{N}\prod_{i=1}^{I}\exp[x_{ni}(\theta_n - \delta_i)]}{\prod_{n=1}^{N}\prod_{i=1}^{I}[1+\exp(\theta_n - \delta_i)]}$$

一般に，$\prod_{n=1}^{N} a^n = a^1 \times a^2 \times ... \times a^N = a^{1+2+...N} = a^{\sum_{n=1}^{N} n}$ なので，

$$= \frac{\exp\sum_{n=1}^{N}\sum_{i=1}^{I}[x_{ni}(\theta_n - \delta_i)]}{\prod_{n=1}^{N}\prod_{i=1}^{I}[1+\exp(\theta_n - \delta_i)]}$$

分子を展開して，

$$= \frac{\exp\sum_{n=1}^{N}\sum_{i=1}^{I}(x_{ni}\theta_n - x_{ni}\delta_i)}{\prod_{n=1}^{N}\prod_{i=1}^{I}[1+\exp(\theta_n - \delta_i)]}$$

分子の $\sum$ を分配する。

$$= \frac{\exp(\sum_{n=1}^{N}\sum_{i=1}^{I}x_{ni}\theta_n - \sum_{n=1}^{N}\sum_{i=1}^{I}x_{ni}\delta_i)}{\prod_{n=1}^{N}\prod_{i=1}^{I}[1+\exp(\theta_n - \delta_i)]}$$

分子の $\sum$ の順番を整理する。

$$= \frac{\exp(\sum_{n=1}^{N}\sum_{i=1}^{I}x_{ni}\theta_n - \sum_{i=1}^{I}\sum_{n=1}^{N}x_{ni}\delta_i)}{\prod_{n=1}^{N}\prod_{i=1}^{I}[1+\exp(\theta_n - \delta_i)]}$$

分子の θ_n は n 番目の受験者の能力を表す定数で，δ_i は i 番目の項目の困難度を表す定数であることから，共に $\sum$ の外に出すことができる（参考：静, 2007, p. 5）。

$$= \frac{\exp\left[\sum_{n=1}^{N}\left(\theta_n \sum_{i=1}^{I} x_{ni}\right) - \sum_{i=1}^{I}\left(\delta_i \sum_{n=1}^{N} x_{ni}\right)\right]}{\prod_{n=1}^{N}\prod_{i=1}^{I}[1 + \exp(\theta_n - \delta_i)]}$$

$\sum_{i=1}^{I} x_{ni}$ は n 番目の受験者の項目 I 番目までの観測得点の合計 r_n であり，$\sum_{n=1}^{N} x_{ni}$ は i 番目の項目の受験者 N 番目までの観測得点の合計 s_i であることから，

$$L = \frac{\exp\left(\sum_{n=1}^{N} r_n\theta_n - \sum_{i=1}^{I} s_i\delta_i\right)}{\prod_{n=1}^{N}\prod_{i=1}^{I}[1 + \exp(\theta_n - \delta_i)]} \tag{11}$$

観測得点である応答データから未知の潜在的な受験者能力と項目困難度を同時に推定する作業は，観測得点と期待得点との差を限りなく 0 に近づける値を推定する作業と同義である。その作業は，N 人 $\times$ I 個からなる応答データの 1 つ x_{ni} において，2 変数の情報量を最大化する値を推定することである。つまり，それは 2 変数からなる関数の瞬間的な傾きが 0 になる値を推定することであり，その値は偏微分によって得られる。受験者能力と項目困難度の同時確率を表す式（11）を偏微分しやすくするために対数変換する。

$$\log(L) = \log\left\{\frac{\exp\left(\sum_{n=1}^{N} r_n\theta_n - \sum_{i=1}^{I} s_i\delta_i\right)}{\prod_{n=1}^{N}\prod_{i=1}^{I}[1 + \exp(\theta_n - \delta_i)]}\right\} \tag{12}$$

$\log\left(\frac{A}{B}\right) = \log(A) - \log(B)$ なので，

$$= \log\left[\exp\left(\sum_{n=1}^{N} r_n\theta_n - \sum_{i=1}^{I} s_i\delta_i\right)\right] - \log\prod_{n=1}^{N}\prod_{i=1}^{I}[1 + \exp(\theta_n - \delta_i)]$$

log の底が e の時，$\log(\exp(x)) = x$ となり，$\log(A \times B) = \log(A) + \log(B)$ であるならば $\log\prod = \sum\log$ となることから，

$$= \sum_{n=1}^{N} r_n\theta_n - \sum_{i=1}^{I} s_i\delta_i - \sum_{n=1}^{N}\sum_{i=1}^{I}\log[1 + \exp(\theta_n - \delta_i)] \tag{13}$$

式（13）を使い，受験者能力 θ_1〜θ_N までの値から，θ_t に関して偏微分をする。

$$\frac{\partial}{\partial\theta_t}\log(L)=\frac{\partial}{\partial\theta_t}\left\{\sum_{n=1}^{N}r_n\theta_n-\sum_{i=1}^{I}s_i\delta_i-\sum_{n=1}^{N}\sum_{i=1}^{I}\log[1+\exp(\theta_n-\delta_i)]\right\}$$

偏微分記号を分配する。

$$=\frac{\partial}{\partial\theta_t}\sum_{n=1}^{N}r_n\theta_n-\frac{\partial}{\partial\theta_t}\sum_{i=1}^{I}s_i\delta_i-\frac{\partial}{\partial\theta_t}\sum_{n=1}^{N}\sum_{i=1}^{I}\log[1+\exp(\theta_n-\delta_i)]$$

$n=t$ の時, 第1項は $r_t\theta_t$ となり, θ_t に関して偏微分すると r_t になる。第2項には θ_t は含まれないため, 定数項とみなし偏微分すると0になる。第3項は, $n=t$ の時, $\sum_{i-1}^{I}\log\left[1+\exp(\theta_t-\delta_i)\right]$ となる。

$$=r_i-\frac{\partial}{\partial\theta_i}\sum_{i-1}^{I}\log\left[1+\exp(\theta_t-\delta_i)\right]$$

静（2007, p. 62）にある和のルール（和の微分係数は微分係数の和）を使って,

$$=r_i-\sum_{i-1}^{I}\frac{\partial}{\partial\theta_i}\log\left[1+\exp(\theta_t-\delta_i)\right]$$

静（2007, p. 66）の例題 $y=\log(2x+3)$ を微分するケースを参考に,

$$=r_i-\sum_{i-1}^{I}\left\{\left(\frac{1}{1+\exp(\theta_t-\delta_i)}\right)\frac{\partial}{\partial\theta_t}[1+\exp(\theta_t-\delta_i)]\right\}$$

以下の展開式が分かりにくいので，静（2007）に式を加えた。

$$=r_i-\sum_{i-1}^{I}\left\{\left[\frac{1}{1+\exp(\theta_t-\delta_i)}\right]\cdot\left[\frac{\partial}{\partial\theta_t}\cdot1+\frac{\partial}{\partial\theta_t}\cdot\exp(\theta_t-\delta_i)\right]\right\}$$

$$=r_i-\sum_{i-1}^{I}\left\{\left(\frac{1}{1+\exp(\theta_t-\delta_i)}\right)\cdot\left[0+\frac{\partial}{\partial\theta_t}\cdot\exp(\theta_t-\delta_i)\right]\right\}$$

一般に指数関数 $y=e^{f(x)}$ を微分すると，$y'=e^{f(x)}\cdot f'(x)$ となることから,

$$=r_i-\sum_{i-1}^{I}\left\{\left(\frac{1}{1+\exp(\theta_t-\delta_i)}\right)\cdot\left\{0+\exp(\theta_t-\delta_i)\left[\frac{\partial}{\partial\theta_t}\cdot(\theta_t-\delta_i)\right]\right\}\right\}$$

$\theta_t-\delta_i$ を θ_t で微分すると 1 になることから,

$$=r_i-\sum_{i-1}^{I}\left\{\left[\frac{1}{1+\exp(\theta_t-\delta_i)}\right]\cdot[\exp(\theta_t-\delta_i)\cdot1]\right\}$$

$$=r_i-\sum_{i-1}^{I}\frac{\exp(\theta_t-\delta_i)}{1+\exp(\theta_t-\delta_i)}$$

$\log(L)$ は，傾きが 0 の時に最大になることから，

$$= r_i - \sum_{i-1}^{I} \frac{\exp(\theta_t - \delta_i)}{1 + \exp(\theta_t - \delta_i)} = 0$$

つまり

$$r_i = \sum_{i-1}^{I} \frac{\exp(\theta_t - \delta_i)}{1 + \exp(\theta_t - \delta_i)} \tag{14}$$

式（14）の左辺（r_i）は，t 番目の受験者の観測得点の合計である。右辺は，t 番目の受験者の期待得点の合計である。つまり，式（14）は，t 番目の受験者が応答データのように解答する最もらしい θ_t の値は，観測得点と期待得点が一致する θ_t であることを証明している。項目困難度に関しても同様の作業を行うことによって，同様の結果を得ることができる（参照：靜, 2007, pp. 248–249）。