

# 項目反応理論概説

前川眞一

040416,050527,0714, 060414, 0911, 080923,1031,1119,20,091020,120214  
,151013,170908,221219,20230508

## 目次

1	はじめに	2
2	基本的仮定	3
2.1	2値データの場合	3
2.2	多値データの場合	4
3	項目特性曲線の導出	6
3.1	体重測定の例	6
3.2	2値データの場合	7
3.3	多値データの場合	8
3.4	プロビット・ロジスティック回帰、カテゴリカルデータの因子分析モデルモデルからの導出	9
4	項目反応理論が示すテストの特徴	11
4.1	テスト特性曲線	11
4.2	テスト得点の条件付き分布	12
4.3	得点分布の予測	13
4.4	テストの等化	14
5	古典的テスト理論との関係	15
6	統計的推定	16
6.1	正答数得点からの $\theta$ の推定	16
6.2	重み付き得点からの $\theta$ を推定する際の信頼区間と情報関数	17
6.2.1	信頼区間	17
6.2.2	情報関数	18
6.2.3	局所的最適重み	19
6.3	項目反応パターンからの特性値の推定	20
6.3.1	重み付き合計得点を基にした特性値の推定	22
6.4	項目パラメタの推定：2値項目の場合	23
6.4.1	同時最尤法	23
6.4.2	周辺最尤法	23
6.4.3	EM アルゴリズムによるパラメタの推定（一般形）	24
6.5	項目パラメタの推定：多値項目の場合	29
7	他の方法との比較	29
7.1	因子分析モデルとの比較	29
7.2	ロジスティック回帰分析との比較	30

<b>8</b>	<b>項目パラメタの等化</b>	<b>30</b>
8.1	原点と単位の不定性 . . . . .	30
8.2	共通項目を用いた項目パラメタの等化 . . . . .	31
8.3	共通尺度と個別尺度との関係 . . . . .	32
8.4	共通尺度への等化法 . . . . .	34
8.4.1	パラメタを用いた最小2乗解 . . . . .	34
8.4.2	反応確率を用いた最小2乗解 . . . . .	34
<b>1</b>	<b>最尤解</b>	<b>37</b>
1.1	正規分布の場合 . . . . .	37
1.2	ベルヌーイ分布の場合 . . . . .	38
<b>2</b>	<b>回帰分析</b>	<b>39</b>
2.1	正規分布の場合：回帰分析 . . . . .	39
2.2	2項分布の場合：ロジスティック回帰分析 . . . . .	39
<b>1</b>	<b>コマンドファイル</b>	<b>42</b>
<b>2</b>	<b>例</b>	<b>44</b>
2.1	単一フォーム、1 集団の場合 (test1L.dat) . . . . .	44
2.2	単一フォーム、複数集団の場合 (test1.dat) . . . . .	48
2.3	複数フォーム、複数集団の場合 1 compressed format (test21c.dat) . . . . .	50
2.4	複数フォーム、複数集団の場合 2 expanded format (test21e.dat) . . . . .	51
2.5	項目パラメタが既知の場合 . . . . .	54

## 1 はじめに

項目反応理論では、テストの最小単位のことを項目と呼ぶ。項目数を  $p$  とし、これらの項目への被験者の反応を、正答は 1、誤答は 0 で記録すると、

$$U_j = \{0, 1\}, \quad j = 1, 2, \dots, p \quad (1)$$

となるが、これを項目反応、または項目得点と呼ぶ。従って、テスト得点は

$$X = \sum_{j=1}^p U_j \quad (2)$$

となる。項目反応理論に於いては、この項目反応を直接に取り扱う。(古典的テスト理論と呼ばれるものは、どちらかという、項目反応の合計としての、テスト得点を中心に扱うことが多い。)

項目反応理論の最大の特徴は、項目の難易度や精度などに代表される項目の特性(項目パラメタ)と、テストを受験する人々の特性(すなわち、学力、能力等)とを分離して考えることが基本となっている。

項目パラメタには主なものが二つあり、ある項目がどの程度難しいか、ということは項目の困難度パラメタと呼ばれるもので表現され、項目の精度を表すためには識別力パラメタと呼ばれるものが利用される。

項目反応理論においては、各項目が測ろうとする被験者の特徴(能力、学力、性格等)のことを一般に特性(潜在特性、もしくは特性値  $\theta$ )と呼ぶが、これも、テスト得点とは別に、項目に依存しない形で定義がなされている。たとえば、Aさんが小学校3年生用のテストで50点、Bさんが中学生用のテストで40点をとったとしても、それだけではどちらの成績が良いかは比べられないが、項目反応理論

を用いてその人々の特性値を求めれば、その比較が可能となる。(ただし、等化が必要)たとえば、項目 a の正答率(通過率)が 70% であり、項目 b の正答率が 50% であることが分かっているとしても、これらの項目正答率が、同一集団(生徒)を用いて計算されたのでなければ項目の難易度の指標とはならないが、項目反応理論の項目パラメタは、比較可能である。(ただし、等化が必要)

## 2 基本的仮定

### 2.1 2値データの場合

1. 特性値を与えた場合の項目反応の条件付き分布  
 $\theta$  が既知の場合、各項目への条件付き正答確率を

$$\text{Prob}(U_j = 1 | \theta) = P_j(\theta) \quad (3)$$

とすると、各項目への項目反応は、ベルヌーイ分布 (n=1 の2項分布)

$$f_j(u_j | \theta) = P_j(\theta)^{u_j} (1 - P_j(\theta))^{1-u_j} \quad (4)$$

をする。(正答確率が  $\theta$  に依存することに注意されたい。)

2. 項目特性曲線 (Item Response Function, Item Characteristic Curve)

この条件付き正答確率を特性値  $\theta$  の関数として見たものを項目特性曲線と呼ぶが、これは、項目反応の特性値への回帰曲線(条件付き期待値のプロット)を表している。項目特性曲線としては、

$$0 \leq P_j(\theta) \leq 1 \quad (5)$$

であり、最低限、 $\theta$  の単調増加関数であることがのぞましいが、通常は、何らかのパラメトリックな関数を仮定したものが用いられる。そのよく用いられる形としては以下のようなものがある。

累積正規モデル

$$P_j(\theta) = \Phi(a_j(\theta - b_j)) \quad (6)$$

ただし、 $\Phi$  は標準正規分布の分布関数である。

3 パラメタロジスティックモデル (3PLM)

$$P_j(\theta) = c_j + (1 - c_j) \times \frac{1}{1 + \exp(-1.7a_j(\theta - b_j))} \quad (7)$$

ただし、 $a_j, b_j, c_j$  は項目固有のパラメタであり、それぞれ、識別力パラメタ、困難度パラメタ、チャンスレベルパラメタと呼ばれている。

2 パラメタロジスティックモデル (2PLM)

$$P_j(\theta) = \frac{1}{1 + \exp(-1.7a_j(\theta - b_j))} \quad (8)$$

これは、3PLM の全ての  $c_j$  を 0 とおいたものである。

1 パラメタロジスティックモデル (1PLM)

$$P_j(\theta) = \frac{1}{1 + \exp(-1.7(\theta - b_j))} \quad (9)$$

これは、2PLM のすべての  $a$  パラメタを 1 とおいたものであり、ラッシュモデルと呼ばれる場合もある。

### 3. 局所独立

$\theta$  が既知の場合の各項目への項目反応は互いに独立である。したがって項目反応の同時確率は以下のように書ける。

$$\text{Prob}(U_1 = u_1 \& U_2 = u_2 \& \dots | \theta) = f_1(u_1 | \theta) \times f_2(u_2 | \theta) \times \dots \quad (10)$$

$$= \prod_{j=1}^p f_j(u_j | \theta) \quad (11)$$

$$= f(\mathbf{u} | \theta) \quad (12)$$

なお、この仮定は、個人ごとに項目反応が独立であることを仮定すると解釈するよりは、 $\theta$  が同じ人を集めて来たときに独立になると解釈する場合が多い。

### 4. 母集団における特性値の分布

対象としている母集団において、特性値は密度関数  $h(\theta | \boldsymbol{\mu})$  にしたがって分布している

$$\Theta \sim h(\theta | \boldsymbol{\mu}) \quad (13)$$

とする。ただし、 $\boldsymbol{\mu}$  はこの分布を定めるパラメタであり、例えば、 $h(\theta | \boldsymbol{\mu})$  として正規分布の密度関数を用いれば、 $\boldsymbol{\mu} = \{\mu, \sigma^2\}$  となる。なお、この仮定は、必ずしも必要な仮定ではないが、 $\theta$  の原点と単位を決めておく必要がある。

## 2.2 多値データの場合

項目が順序が付いた多値項目である場合、それらの点数を低い方から 0 点 1 点と数え、項目  $j$  の満点を  $m_j$  とすると

$$U_j = \{0, 1, 2, \dots, m_j\}, \quad j = 1, 2, \dots, p \quad (14)$$

が各項目の得点となる。(カテゴリの数は 0 も加えて  $m_j + 1$  個) 従って、テスト得点は

$$\mathbf{X} = \sum_{j=1}^p U_j \quad (15)$$

なお、 $U_j$  をダミー変数を用いて表すと

$$U_j = k \quad \leftarrow \rightarrow \quad U_{jk} = 1 \quad (16)$$

と書ける。

#### 1. 特性値を与えた場合の項目反応の条件付き分布

$\theta$  が既知の場合、各項目への条件付き正答確率を

$$\text{Prob}(U_j = k | \theta) = \text{Prob}(U_{jk} = 1 | \theta) = P_{jk}(\theta), \quad k = 0, 1, 2, \dots, m_j \quad (17)$$

とすると、各項目への項目反応は、多項分布 ( $n=1$ )

$$f_j(\mathbf{u}_j | \theta) = \prod_{k=0}^{m_j} P_{jk}(\theta)^{u_{jk}}, \quad \sum_{k=0}^{m_j} P_{jk}(\theta) = 1 \quad (18)$$

をする。(正答確率が  $\theta$  に依存することに注意されたい。)

## 2. 項目特性曲線 (Categori Response Function and Cumulative CRF)

この条件付きの各カテゴリへの正答確率  $\text{Prob}(U_{jk} = 1 | \theta)$  を特性値  $\theta$  の関数として見たもの  $P_{jk}(\theta)$  を項目カテゴリ特性曲線 (Item Category Response Function, Item Category Characteristic Curve) と呼ぶ。られる。良く用いられる多値項目のモデルとしては、以下の二つが存在する。

段階反応モデル Graded Response Model

カテゴリ  $k$  への反応確率は、カテゴリ  $k$  以上のカテゴリ ( $k, k+1, \dots, m_j$ ) に反応する確率  $P_{jk}(\theta)$ , *ItemCategoryBoundaryResponseFunction*, とカテゴリ  $k+1$  以上のカテゴリ ( $k+1, k+2, \dots, m_j$ ) に反応する確率との差として定義される。

$$P_{jk}(\theta) = P_{jk}^+(\theta) - P_{j(k+1)}^+(\theta), \quad k = 0, 1, 2, \dots, m_j \quad (19)$$

ただし

$$P_{j0}^+(\theta) = 1 \quad \text{and} \quad P_{jm_j+1}^+(\theta) = 0 \quad (20)$$

$$P_{jk}^+(\theta) = \frac{1}{1 + \exp(-1.7a_j(\theta - b_{jk}))}, \quad k = 1, 2, \dots, m_j \quad (21)$$

または (22)

$$P_{jk}^+(\theta) = \frac{1}{1 + \exp(-1.7a_j(\theta - b_j + c_{jk}))}, \quad k = 1, 2, \dots, m_j, \quad \sum_{k=1}^{m_j} c_{jk} = 0 \quad (23)$$

部分採点モデル (Generalized) Partial Credit Model

部分採点モデルは分子の  $\theta$  の関数をその総和で除した形をしている。

$$P_{jk}(\theta) = \frac{\exp\left(\sum_{h=0}^k (a_j^*(\theta - b_{jh}^*))\right)}{\sum_{k'=0}^{m_j} \exp\left(\sum_{h=0}^{k'} (a_j^*(\theta - b_{jh}^*))\right)} \quad (24)$$

$$= \frac{\exp(\alpha_j k \theta + c_{jk})}{\sum_{q=0}^{m_j} \exp(\alpha_j q \theta + c_{jq})}, \quad k = 0, 1, 2, \dots, m_j \quad (25)$$

ただし

$$c_{j0} = b_{j0} = 0 \quad \text{and} \quad b_{j0}^* = 0, \quad j = 1, 2, \dots, p \quad (26)$$

であり、これは第 0 番目のカテゴリへの反応をベースラインとしている。

なお、各カテゴリの重みを  $v_{jk}$ ,  $k = 0, 1, 2, \dots, m_j$  (通常は  $v_{jk} = k$ ,  $k = 0, 1, 2, \dots, m_j$ ) とした時の重み付き項目得点

$$U_j = \sum_{k=0}^{m_j} v_{jk} U_{jk} \quad (27)$$

の条件付き期待値

$$P_j(\theta) = \sum_{k=0}^{m_j} v_{jk} P_{jk}(\theta) \quad (28)$$

を項目特性曲線と呼ぶが、多値項目の場合、これはもはや確率とはならない。

### 3. 局所独立

$\theta$  が既知の場合の各項目への項目反応は互いに独立である。したがって項目反応の同時確率は以下のように書ける。

$$\text{Prob}(U_1 = u_1 \& U_2 = u_2 \& \dots | \theta) = f_1(\mathbf{u}_1 | \theta) \times f_2(\mathbf{u}_2 | \theta) \times \dots \quad (29)$$

$$= \prod_{j=1}^p f_j(\mathbf{u}_j | \theta) \quad (30)$$

### 4. 母集団における特性値の分布

対象としている母集団において、特性値は密度関数  $h(\theta | \mu)$  にしたがって分布している

$$\Theta \sim h(\theta | \mu) \quad (31)$$

とする。ただし、 $\mu$  はこの分布を定めるパラメタであり、例えば、 $h(\theta | \mu)$  として正規分布の密度関数を用いれば、 $\mu = \{\mu, \sigma^2\}$  となる。なお、この仮定は、必ずしも必要な仮定ではないが、 $\theta$  の原点と単位を決めておく必要がある。(通常は  $\mu = 0, \sigma = 1$  と置く。)

## 3 項目特性曲線の導出

### 3.1 体重測定の例

2 値項目からなるテストを利用した学力の測定のモデルの中で一番簡単なものは、各項目を、性能が異なるが同一の特性である学力を測定するための測定器具であると考え、被験者と項目との間に定義される 2 項関係を利用し項目反応をモデル化する方法である。

以下に、天秤と重りを用いて個人の体重測定を行う例を用いてこのモデルを解説するが、これは、項目が重りに、体重が学力に相当する。いま、体重が 50 キロということはどういうことかを考えてみる。この場合、上記の 2 項関係を、天秤である人の体重と重りの重さを比較することと考える。たとえば、様々な重りを乗せたシーソーに被験者を乗せ、1 秒後にどちらが下になっているかを記録する。重りが正確であれば、体重が 50 キロの人の場合、49 キロというラベルが付いた重りまでは被験者が下になり、50 キロで平行、51 キロかかれた重りからは重りが下になるであろう。このような実験を何人かに実施した結果を、被験者が下を 1、重りが下を 0 で記録すると、およそ以下の表の左側ようになる。

項目の精度が高い場合 (被験者・項目ともに軽い順)													項目の精度が低い場合 (被験者・項目ともに軽い順)														
...	45	46	47	48	49	50	51	52	53	54	55	56	...	...	45	46	47	48	49	50	51	52	53	54	55	56	...
a...	1	1	0	0	0	0	0	0	0	0	0	0	...	a...	0	1	0	1	1	0	0	0	0	0	0	0	...
b...	1	1	1	0	0	0	0	0	0	0	0	0	...	b...	1	1	1	0	0	1	0	0	0	0	0	0	...
c...	1	1	1	0	0	0	0	0	0	0	0	0	...	c...	1	1	1	0	0	1	1	0	0	0	0	0	...
d...	1	1	1	1	0	0	0	0	0	0	0	0	...	d...	1	1	1	1	1	1	0	1	0	0	0	0	...
e...	1	1	1	1	0	0	0	0	0	0	0	0	...	e...	1	1	1	1	0	1	0	1	0	0	0	0	...
f...	1	1	1	1	1	0	0	0	0	0	0	0	...	f...	1	1	1	1	0	1	1	0	0	0	0	0	...
g...	1	1	1	1	1	1	0	0	0	0	0	0	...	g...	1	1	1	1	1	1	1	0	0	0	0	0	...
h...	1	1	1	1	1	1	1	0	0	0	0	0	...	h...	1	1	1	1	0	1	1	1	0	0	0	0	...
i...	1	1	1	1	1	1	1	1	0	0	0	0	...	i...	1	1	1	1	1	0	1	0	1	0	1	0	...
j...	1	1	1	1	1	1	1	1	0	0	0	0	...	j...	1	1	1	1	1	1	0	1	0	0	0	0	...
k...	1	1	1	1	1	1	1	1	0	0	0	0	...	k...	1	1	1	1	0	1	0	1	0	0	0	0	...
l...	1	1	1	1	1	1	1	1	1	0	0	0	...	l...	1	1	1	1	0	1	1	1	1	1	0	0	...
m...	1	1	1	1	1	1	1	1	1	1	1	1	...	m...	1	0	1	1	1	1	1	0	0	1	1	1	...

この表は、各行に被験者を、各列に重りを配置しているが、その特徴は、うまく被験者を並べ替えることにより、きれいに 1 と 0 が並び、0 右には決して 1 がこないことである。このような特徴を持った測定結果をガットマンスケールということがある。

次に重りの精度が悪い場合を考える。たとえば、重りが様々な形の容器に様々な液体が入ったものから出来ているとする。その場合、人が乗ることにより、大きな容器に入った液体は揺れ、重心が移動して、1秒後の位置は必ずしも体重と重り本来の重さを正確には反映するとは限らない。従って、その結果は上記の表の右側のようになるであろう。この表の特徴は、どのように行と列を並べ替えても、だいたい左下に1が、右上に0が多い様には出来るが、きれいに1と0が並び、0の右には決して1がない様には出来ない、ということである。この、ガットマンスケールからのずれは、重りが不正確であることに由来しているが、項目反応理論においては、重りの本来の重さが項目困難度当たり、重りの精度に当たる概念を項目の識別力と呼んでいる。

次節では、Thurstoneの比較判断の法則やTorgersonの範疇判断の法則を使って、項目特性曲線を導出する。

### 3.2 2値データの場合

測ろうとする1次元の特性を考える。そして、その上に各被験者がその特性値の大きさ（これを $\theta$ と記す）により位置づけられているとする。また、各項目（ $j = 1, 2, \dots, p$ ）もその難易度（これを $b_j$ とする）に応じてその上に並んでいるとする。また、項目には、精度の善し悪しがあるとし、その精度の悪さを $\psi_j$ で表すとする。そして、直接的に $b_j$ と $\theta$ を比較するのではなく、 $\theta$ と比較されるのは、平均的に $b_j$ という値を持つが、必ずしも $b_j$ という値をとらず、ある一定の幅（標準偏差） $\psi$ をもって変動する確率変数 $Y_j$ であるとする。従って、

$$\theta \geq Y_j \quad \text{ならば} \quad U_j = 1 \quad (\text{項目 } j \text{ に正答}) \quad (32)$$

$$\theta < Y_j \quad \text{ならば} \quad U_j = 0 \quad (\text{項目 } j \text{ に誤答}) \quad (33)$$

となる。これを図示したものが図1と図4である。

この場合、変動幅 $\psi_j$ が大きければ $Y_j$ は $b_j$ と異なる値をとることがあるため、その大きさが、その項目の不正確さを表していることになる。このモデルによると、たとえ、本来は $\theta \leq b_j$ である場合でも、そのときの $Y_j$ の値により、その項目に正答する場合もあり得ることになるため、先の表の様な結果を得ることになる。しかし、上記の仮定から、特性値が $\theta$ の人が困難度 $b_j$ の項目に正答する確率というものを計算することが出来る。これは、 $Y_j$ の分布として平均が $b_j$ 、標準偏差が $\psi_j$ の正規分布を考えれば、 $\Phi$ を標準正規分布の累積分布関数として、

$$\begin{aligned} \text{Prob}(U_j = 1|\theta) &= \text{Prob}(Y_j < \theta) \\ &= \text{Prob}\left(\frac{Y_j - b_j}{\psi_j} < \frac{\theta - b_j}{\psi_j}\right) \\ &= \text{Prob}\left(Z_j < \frac{\theta - b_j}{\psi_j}\right) \\ &= \Phi\left(\frac{\theta - b_j}{\psi_j}\right) \end{aligned} \quad \begin{aligned} Y_j &\sim N(b_j, \psi_j^2) \\ Z_j &= \frac{Y_j - b_j}{\psi_j} \sim N(0, 1) \end{aligned} \quad (34)$$

となる。

しかし、通常は、項目の精度を表すために $a_j = 1/\psi_j$ を用い、累積正規分布関数をロジスティック曲線で近似したもの

$$\text{Prob}(U_j = 1|\theta) = P_j(\theta) = \Phi(a_j(\theta - b_j)) \approx \frac{1}{1 + \exp(-1.7a_j(\theta - b_j))} \quad (35)$$

を用いることが多い。この $P_j(\theta)$ は2パラメタロジスティック項目特性曲線 (Item Characteristics Curve) と呼ばれているが、これは、特性値が $\theta$ の人が項目 $j$ に正答する確率を与える関数である。(特性値 $\theta$ を持つ人が、その項目に答えたということを忘れながら、何度もその項目に答えた場合の平均正答率に当たる。)

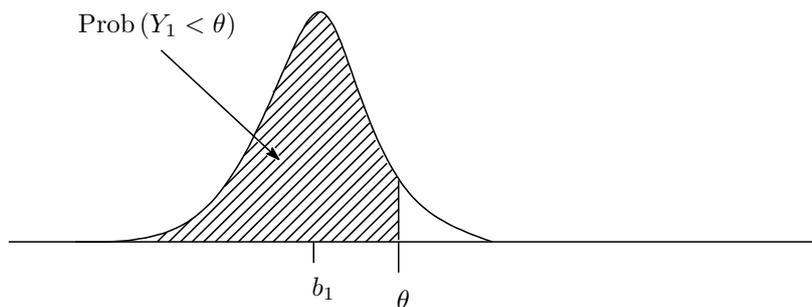


図 1: Thurstone Model of IRT (easy item)

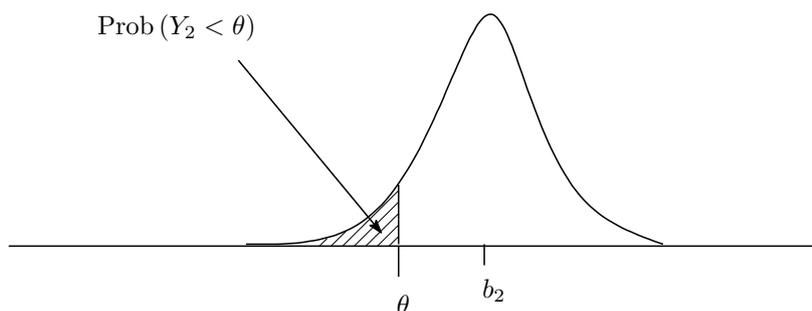


図 2: Thurstone Model of IRT (difficult item)

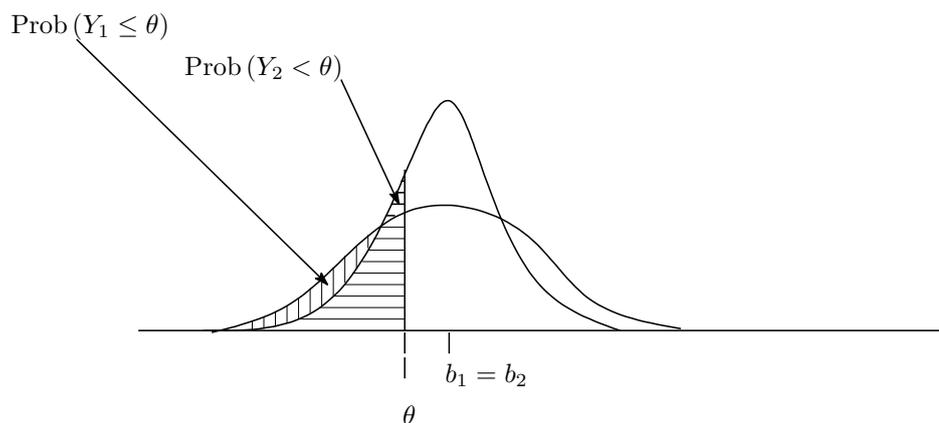


図 3: Thurstone Model of IRT ( $\psi_1 > \psi_2$ )

また、 $a_j$  を項目  $j$  の識別力パラメタ (discrimination/slope)、 $b_j$  を困難度パラメタ (difficulty/threshold) と呼ぶ。なお、これを拡張した3パラメタロジスティック項目特性曲線

$$P_j(\theta) = c_j + (1 - c_j) \times \frac{1}{1 + \exp(-1.7a_j(\theta - b_j))} \quad (36)$$

が用いられることもある。ただし、 $c_j$  はチャンスレベルパラメタ (gussing/pseudo-chance level/lower asymptote) と呼ばれるもので特性値が低い場合にも偶然に項目に正答する確率を表すパラメタである。

### 3.3 多値データの場合

2値データのモデルは以下のように評定尺度の数量化に拡張することが出来る。いま、それぞれ  $0, 1, 2, \dots, m_j$  の  $m_j + 1$  のカテゴリを持つ評定尺度が  $p$  個あるとする。特性値  $\theta$  を持つ個人が第

$j$  番目の尺度で第  $k$  番目のカテゴリで評定されたとする。(正答か誤答で採点される 2 値データの場合は  $m_j = 1$  である。) これは、各尺度項目ごとに、先ほど考えた刺激の尺度上にカテゴリの境界  $b_{jk}, k = 0, 1, 2, \dots, m_j$  を考え、それに誤差が加わったもの

$$Y_{jk} \sim N(b_{jk}, \psi_j^2) \quad (37)$$

と  $\theta$  との比較において

$$Y_{jk+1} \geq \theta \geq Y_{jk} \rightarrow \text{特性値 } \theta \text{ を持つ個人が第 } j \text{ 番目の項目の第 } k \text{ 番目のカテゴリで評定される} \quad (38)$$

であるとすれば、比較判断の法則と同じフレームワークで取り扱うことが出来る。

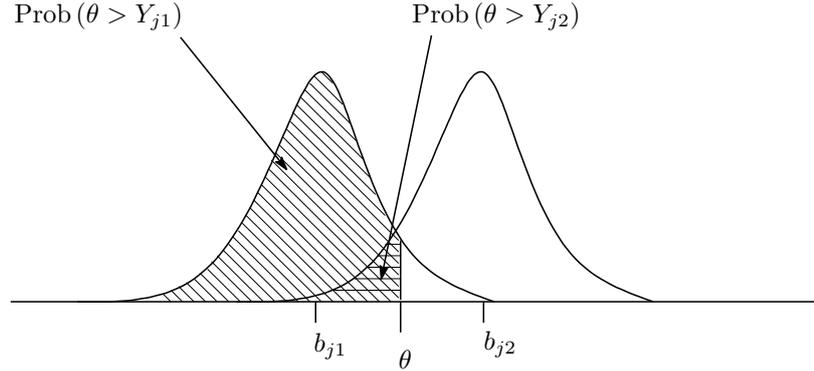


図 4: Thurstone Model of IRT (Graded Response Model)

この確率は

$$\text{Prob}(Y_{jk+1} \geq \theta \geq Y_{jk}) = P_{jk}(\theta) \quad (39)$$

$$= \text{Prob}(\theta \geq Y_{jk}) - \text{Prob}(\theta \geq Y_{jk+1}) \quad (40)$$

と書けるが、項目境界との比較に 2 値データと同じ形を用いれば

$$P_{jk}(\theta) = P_{jk}^+(\theta) - P_{j(k+1)}^+(\theta), \quad k = 0, 1, 2, \dots, m_j \quad (41)$$

$$P_{j0}^+(\theta) = 1 (b_{j0} = -\infty) \quad \text{and} \quad P_{jm_j+1}^+(\theta) = 0 (b_{jm_j+1} = +\infty) \quad (42)$$

$$P_{jk}^+(\theta) = \Phi\left(\frac{\theta - b_{jk}}{\psi_j}\right) \quad k = 0, 1, 2, \dots, m_j \quad (43)$$

という形で用いられる。また、項目反応理論の文脈では、

$$P_{jk}^+(\theta) = \Phi\left(\frac{\theta - b_{jk}}{\psi_j}\right) = \Phi(a_j(\theta - b_{jk})), \quad k = 0, 1, \dots, m_j \quad (44)$$

という記す。また、累積正規分布関数の代わりに、2PLM を用いることが多い。

カテゴリが 2 値のばあい、すなわち、 $m_j = 1$  の場合は

$$P_j(\theta) = P_{j1}^+(\theta) = \Phi\left(\frac{\theta - b_{j1}}{\psi_j}\right) \quad (45)$$

となる。

### 3.4 プロビット・ロジスティック回帰、カテゴリカルデータの因子分析モデルからの導出

#### 1 因子モデル

$$R_j = \mu_j + \lambda_j \theta + e_j, \quad e_j \sim N(0, \psi_j^2) \quad (46)$$

$$R_j | \theta \sim N(\mu_j + \lambda_j \theta, \psi_j^2) \quad (47)$$

を考える。カテゴリカルデータ  $U_j$  の発生は

$$R_j \geq \gamma_j \quad \text{ならば} \quad U_j = 1 \quad (48)$$

$$R_j < \gamma_j \quad \text{ならば} \quad U_j = 0 \quad (49)$$

であるとする。これを図示したものが 図 5 である。

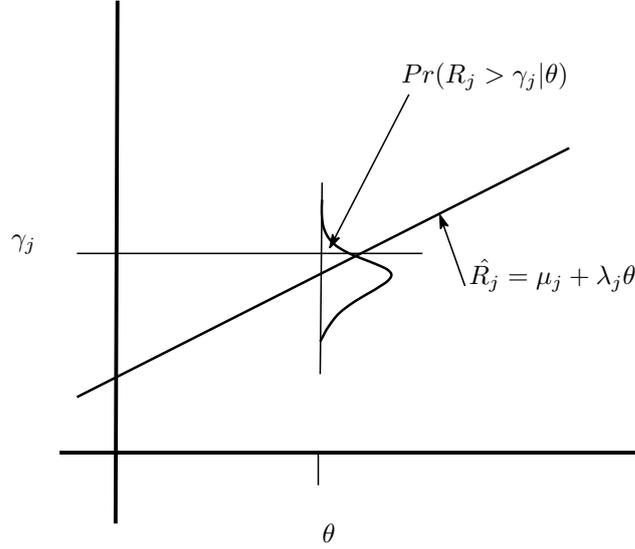


図 5: Categorical FA Model

このとき正反応データの発生確率は

$$\begin{aligned} \text{Prob}(U_j = 1|\theta) &= \text{Prob}(R_j \geq \gamma_j) \\ &= \text{Prob}\left(\frac{R_j - \mu_j - \lambda_j \theta}{\psi_j} > \frac{\gamma_j - \mu_j - \lambda_j \theta}{\psi_j}\right) \\ &= \text{Prob}\left(Z_j > \frac{\gamma_j - \mu_j - \lambda_j \theta}{\psi_j}\right) \\ &= \text{Prob}\left(Z_j < -\frac{\gamma_j - \mu_j - \lambda_j \theta}{\psi_j}\right) \\ &= \Phi\left(\frac{\lambda_j \theta - (\gamma_j - \mu_j)}{\psi_j}\right) \end{aligned} \quad \begin{array}{l} R_j \sim N(\mu_j - \lambda_j \theta, \psi_j^2) \\ Z_j = \frac{R_j - \mu_j - \lambda_j \theta}{\psi_j} \sim N(0, 1) \end{array} \quad (50)$$

対称性

となるが、パラメタは識別されない。

$$a_j = \frac{\lambda_j}{\psi_j} \quad (51)$$

$$b_j = \frac{(\gamma_j - \mu_j)}{\lambda_j} \quad (52)$$

を用いれば、項目反応理論の形となる。

多値データの場合を図示したものが 図 6 である。

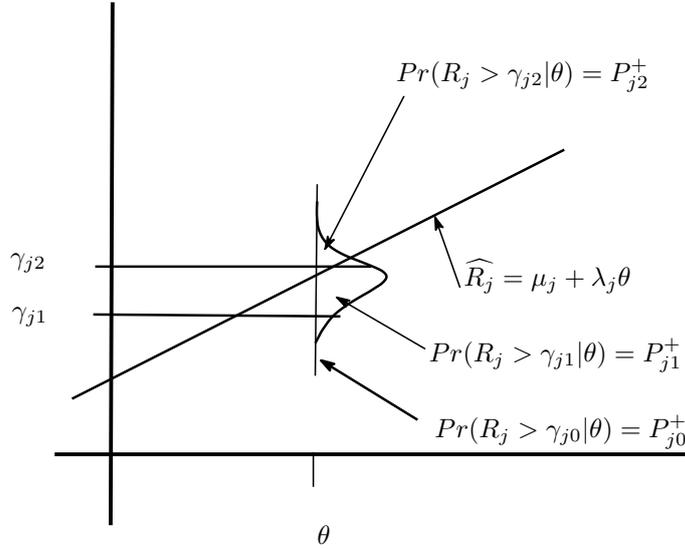


図 6: Graded Response Model

## 4 項目反応理論が示すテストの特徴

### 4.1 テスト特性曲線

いま、 $p$  個の 2 値項目を含むテストを考える。項目反応理論においては、項目特性曲線  $P_j(\theta)$  が特性値が  $\theta$  の人の各項目への正答確率、すなわち条件付き期待得点を表すため、その和

$$\mathcal{E}(X|\theta) = T(\theta) = \sum_{j=1}^p P_j(\theta) \quad (53)$$

が特性値が  $\theta$  の人のそのテスト得点の条件付き期待値（平均値）となる。これをテスト特性曲線（Test Characteristics Curve）もしくはテスト反応関数（Test Response Function）と呼ぶ。なお、テスト特性曲線の範囲は 0 から  $p$ （項目数＝満点）である。

2 値項目の場合、各項目に重み（配点） $w_j$  を掛けて足し合わせた重み付き合計得点

$$X = \sum_{j=1}^p w_j U_j \quad (54)$$

の条件付き期待値

$$\mathcal{E}(X|\theta) = T(\theta) = \sum_{j=1}^p w_j P_j(\theta) \quad (55)$$

を、重み付き得点のテスト特性曲線と呼ぶ。

項目が多値反応項目を含む場合、項目  $j$  のカテゴリ  $k$  にかかる重みを  $v_{jk}$  とし、各項目得点にかかる重みを  $w_j$  として計算される重み付き合計点  $X$ （(27) 式に項目の重み  $w_j$  を掛けて足したもの）

$$X = \sum_{j=1}^p w_j \sum_{k=0}^{m_j} v_{jk} \times U_{jk} \quad (56)$$

の  $\theta$  を与えた時の期待値、すなわちテスト特性曲線は

$$\mathcal{E}(X|\theta) = T(\theta) = \sum_{j=1}^p w_j \sum_{k=0}^{m_j} v_{jk} \times P_{jk}(\theta) \quad (57)$$

で与えられる。

## 4.2 テスト得点の条件付き分布

項目  $j$  への反応をダミー変数ベクトルで表したものを

$$\text{確率変数: } \mathbf{U}_j = (U_{j0}, U_{j1}, U_{j2}, \dots, U_{jm_j}) \quad (58)$$

$$\text{確率の実現値: } \mathbf{u}_j = (u_{j0}, u_{j1}, u_{j2}, \dots, u_{jm_j}) \quad (59)$$

$$\text{母確率: } \mathbf{p}_j(\theta) = (P_{j0}(\theta), P_{j1}(\theta), P_{j2}(\theta), \dots, P_{jm_j}(\theta)) \quad (60)$$

とすると、 $U_{jk}, k = 0, 1, 2, \dots, m_j$  の中のどれかひとつが 1 となり、残りが 0 となるため、 $\mathbf{U}_j$  はカテゴリ確率  $\mathbf{p}(\theta)$  を持つ  $m_j + 1$  次元の多項分布であると考えられる。したがって、多項分布の性質から

$$\mathcal{E}(U_{jk}|\theta) = P_{jk}(\theta) \quad (61)$$

$$\text{var}(U_{jk}|\theta) = P_{jk}(\theta)(1 - P_{jk}(\theta)) \quad (62)$$

$$\text{cov}(U_{jk}, U_{jk'}|\theta) = -P_{jk}(\theta)P_{jk'}(\theta) \quad (63)$$

であり、まとめると

$$\mathcal{E}(\mathbf{U}_j|\theta) = \mathbf{p}_j(\theta) \quad \text{and} \quad \mathcal{D}(\mathbf{U}_j|\theta) = \text{Diag}(\mathbf{p}_j(\theta)) - \mathbf{p}_j(\theta)\mathbf{p}_j(\theta)' \quad (64)$$

である。したがって、項目得点が (27) 式により、項目の重み  $v_{jk}, k = 0, 1, 2, \dots, m_j$ , を縦に並べた重みベクトル  $\mathbf{v}_j$  を用いて

$$U_j = \mathbf{v}_j' \mathbf{U}_j \quad (65)$$

と書けるため、項目得点の条件付き期待値は、

$$\mathcal{E}(U_j|\theta) = \mathcal{E}(\mathbf{v}_j' \mathbf{U}_j|\theta) = \mathbf{v}_j' \mathcal{E}(\mathbf{U}_j|\theta) = \mathbf{v}_j' \mathbf{p}_j(\theta) = \sum_{k=0}^{m_j} v_{jk} P_{jk}(\theta) = P_j(\theta) \quad (66)$$

となるが、これを項目反応関数 (Item Response Function, IRF) と呼び  $P_j(\theta)$  と記す。(多値モデルの場合、 $P_j(\theta)$  はもはや確率では無い。) また、 $\theta$  を与えたときの項目得点の条件付き分散は、

$$\text{var}(U_j|\theta) = \mathbf{v}_j' \mathcal{D}(\mathbf{U}_j|\theta) \mathbf{v}_j \quad (67)$$

$$= \mathbf{v}_j' (\text{Diag}(\mathbf{p}_j(\theta)) - \mathbf{p}_j(\theta)\mathbf{p}_j(\theta)') \mathbf{v}_j \quad (68)$$

$$= \mathbf{v}_j' \text{Diag}(\mathbf{p}_j(\theta)) \mathbf{v}_j - \mathbf{v}_j' \mathbf{p}_j(\theta) \mathbf{p}_j(\theta)' \mathbf{v}_j \quad (69)$$

$$\begin{aligned} &= \sum_{k=0}^{m_j} v_{jk}^2 P_{jk}(\theta) - \left( \sum_{k=0}^{m_j} v_{jk} P_{jk}(\theta) \right)^2 \\ &= \sum_{k=0}^{m_j} v_{jk}^2 P_{jk}(\theta) - P_j(\theta)^2 \end{aligned} \quad (70)$$

となる。なお、上記の平均や分散は、数値  $(v_{j0}, v_{j1}, v_{j2}, \dots, v_{jm_j})$  を離散値として持つ項目得点  $U_j$  の度数分布を考え、各カテゴリの確率が  $(P_{j0}(\theta), P_{j1}(\theta), P_{j2}(\theta), \dots, P_{jm_j}(\theta))$  で与えられていると考えれば、

$$\mathcal{E}(U_j|\theta) = \sum_{k=0}^{m_j} v_{jk} P_{jk}(\theta) = P_j(\theta) \quad (71)$$

$$\text{var}(U_j|\theta) = \sum_{k=0}^{m_j} (v_{jk} - P_j(\theta))^2 P_{jk}(\theta) = \sum_{k=0}^{m_j} v_{jk}^2 P_{jk}(\theta) - P_j(\theta)^2 \quad (72)$$

として計算することが出来る。

また、2値項目は、第0番目のカテゴリへの反応を誤答、第1番目のカテゴリへの反応を正答とし、重みを  $v_{j0} = 0, v_{j1} = 1$  とした  $m_j = 1$  の多値項目と解釈することが出来る。したがって、

$$\mathcal{E}(U_j|\theta) = P_{j0}(\theta) = P_j(\theta) \quad (73)$$

$$\text{var}(U_j|\theta) = P_{j1}(\theta) - P_{j1}(\theta)^2 = P_j(\theta)(1 - P_j(\theta)) \quad (74)$$

となる。

次に、局所独立の仮定から、 $u_j = \{v_{j0}|v_{j1}|\dots|v_{jm_j}\}$ ,  $j = 1, 2, \dots, p$  を第  $j$  番目の項目への項目反応であるとする項目反応の同時確率は以下のように書ける。

$$\begin{aligned} \text{Prob}(U_1 = u_1 \& U_2 = u_2 \& \dots \& U_p = u_{m_p}|\theta) &= \text{Prob}(U_1 = u_1|\theta) \times \text{Prob}(U_2 = u_2|\theta) \times \dots \\ &= \prod_{j=1}^p \text{Prob}(U_j = u_j|\theta) \end{aligned} \quad (75)$$

$$= \prod_{j=1}^p \prod_{k=0}^{m_j} P_{jk}(\theta)^{u_{jk}} \quad (76)$$

したがって、(54) 式 や (56) 式 に示す重み付き合計点に関してはその条件付き平均が

$$\mathcal{E}(X|\theta) = \mathcal{E}\left(\sum_{j=1}^p w_j U_j|\theta\right) = \sum_{j=1}^p w_j \mathcal{E}(U_j|\theta) \quad (77)$$

$$= \sum_{j=1}^p w_j P_j(\theta) \quad (78)$$

また、条件付き分散が、局所独立の仮定により、

$$\text{var}(X|\theta) = \text{var}\left(\sum_{j=1}^p w_j U_j|\theta\right) = \sum_{j=1}^p w_j^2 \text{var}(U_j|\theta) \quad (79)$$

$$= \sum_{j=1}^p w_j^2 \left( \sum_{k=0}^{m_j} v_{jk}^2 P_{jk}(\theta) - P_j(\theta)^2 \right) \quad (80)$$

となる。なお、この  $\mathcal{E}(X|\theta)$  のことをテスト反応関数 (Test Response Function, TRF) またはテスト特性曲線 (Test Characteristic Curve, TCC) と呼び

$$TCC(\theta) = \mathcal{E}(X|\theta) \quad (81)$$

と書く。また、条件付き分散を

$$SEM(\theta)^2 = \text{var}(X|\theta) \quad (82)$$

と書くが、この  $SEM(\theta)$  が古典的テスト理論でいう測定の標準誤差 (SEM) を  $\theta$  の関数として表したものととなっている。なお、この関数をプロットすると、上に凸の曲線となる。すなわち、特性値の中央部分における標準誤差が、両端部分に比べて大きくなっている。

### 4.3 得点分布の予測

一般的に、項目  $j$  のカテゴリ  $k$  にかかる重みを  $v_{jk}$  とし、それで計算される項目得点を

$$U_j = \sum_{k=0}^{m_j} v_{jk} U_{jk} \quad (83)$$

項目にかかる重みを  $w_j$  として合計点を

$$X = \sum_{j=1}^p w_j U_j \quad (84)$$

$$= \sum_{j=1}^p w_j \sum_{k=0}^{m_j} v_{jk} U_{jk} \quad (85)$$

として計算される重み付き得点の分布を考える。

$\theta$  を与えたときに、各項目が  $k$  という値を取る確率は (19) 式 の  $P_{jk}(\theta)$  で与えられる。(2値データの場合は  $m_j = 1$  の場合となる。)

従って、 $\theta$  を与えた場合のテスト得点の分布は、 $P_{jk}(\theta)$  の各列で与えられる、異なる確率を持つ独立な多項分布変数の重み付き和の分布となる。(独立性は、局所独立の仮定から来ている。)

たとえば、 $p = 2$  項目で、 $m_1 = 1, m_2 = 2$  であり、カテゴリの重みが  $v_{10} = 0, v_{11} = 1, v_{20} = 0, v_{21} = 2, v_{22} = 3$  の場合、第1項目の得点  $X_1$  はカテゴリ数が2の多項分布でその値域は  $X_1 = 0|1$  であり、その生起確立は  $P_{10}(\theta)$  と  $P_{11}(\theta)$ 、第2項目  $X_2$  の得点は、カテゴリ数が3の多項分布でその値域は  $X_2 = 0|2|3$  であり、その生起確立は  $P_{20}(\theta)$ 、 $P_{21}(\theta)$  と  $P_{23}(\theta)$  となる。したがって  $X$  の分布は、上記のふたつの多項分布をする変数  $X_1$  と  $X_2$  の  $w$  で重み付けられた和の分布となる。たとえば、 $w_1 = 2, w_2 = 1$  とすると、 $X$  の値域は  $(0, 2, 3, 4, 5)$  となる。実際に  $Pr(X = x|\theta)$  を計算するためには、値域のそれぞれの値をもたらすふたつの多項分布のカテゴリの組み合わせを考え、その確率を加えればよい。

2値データの場合、この分布を複合2項分布 Compound Binomial、多値データの場合、この分布を複合多項分布 Compound Multinomial と呼ぶ。これを

$$Pr(X = x|\theta) = f(x|\theta) \quad (86)$$

と書く。従って、 $X$  の周辺分布 (これが観測される) は、その  $\theta$  の分布に関する積分

$$Pr(X = x) = \int_{\Theta} f(x|\theta)h(\theta|\mu)d\theta \quad (87)$$

で与えられる。

なお、これらの分布の計算の仕方に関しては Mayekawa & Arai (2008) を参照のこと。

項目反応理論を利用してテストを編集する場合、望ましい得点分布を持つテストを作成することが出来る。たとえば、資格試験を作るような場合、被験者を合格者群と不合格者群に分けることが目的であるから、得点分布に二つの山があれば、その谷の部分で2群を分ければ良いことになって便利である。また、被験者の順位を問題とする場合には、いわゆる正規分布型の得点分布が望まれることが多い。

対象とする集団における特性値の分布がわかっているならば、テスト特性曲線を利用することにより、そのテストの得点分布の形の見当をつけることが出来る。特性値が正規分布をすることを仮定した場合、たとえば、ふたつの山の分布がほしければ識別力が高く難易度に差がある項目を集めてテストを編集すればよいことが、正規分布に近い分布がほしければ、識別力のあまり高くない項目を集めれば良いことがわかる。ここで注意してほしいことは、これらの得点分布の元になる特性値の分布は全て同一であるということである。すなわち、テスト得点の分布は、そのテストに含まれる項目の特性により、様々な形を取る可能性があるということである。

#### 4.4 テストの等化

同一の潜在特性  $\theta$  を測定する複数のテスト  $X_a$ ,  $a = 1, 2, \dots, A$  があり、それらのテストのテスト特性曲線を

$$\mathcal{E}(X_a|\theta) = TCC_a(\theta) \quad (88)$$

とする。いま、 $\theta$  を媒介変数として、TCC の逆関数間の関係

$$TCC_a^{-1}(x_a) = \theta \quad \text{and} \quad TCC_b^{-1}(x_b) = \theta \quad (89)$$

を満たす  $(x_a, x_b)$  のプロットを作れば、それは、同一能力に対応するテスト  $a$  と  $b$  の得点の関係を示すプロットとなる。これを項目反応理論による真値を用いた得点の等化 (IRT True Score Equating) と呼ぶ。すなわち、(89) 式を  $x_b$  について解いたもの

$$x_{b.a} = TCC_b(TCC_a^{-1}(x_a)) \quad (90)$$

がテスト  $a$  に等しいテスト  $b$  の得点となる。

## 5 古典的テスト理論との関係

$\theta$  を与えたときのテスト得点の条件付き期待値と分散は (78) 式と (80) 式で与えられる。また条件付きの標準偏差、すなわち、測定の標準誤差 (SEM) は (82) 式に示す  $SEM(\theta)$  である。しかし、これらは  $\theta$  の関数であり、古典的テスト理論の真値の関数ではない。

そこで、(81) 式に示されたテスト得点  $X$  とその条件付き期待値の関係は一对一の関数関係であることに注意して  $T = TCC(\theta)$  と置けば、(78) 式と (80) 式を  $T$  の関数と見なすことが出来、その結果

$$\mathcal{E}(X|T) = T \quad \text{and} \quad SEM(T)^2 = var(X|T) = var(X|TCC^{-1}(T)) \quad (91)$$

を得る。ただし、 $TCC^{-1}(T)$  は  $TCC(\theta)$  の逆関数である。

これを多値項目を含む 10 項目からなる 19 点満点のテストの項目パラメタを用いて計算した結果を図 7 に示した。

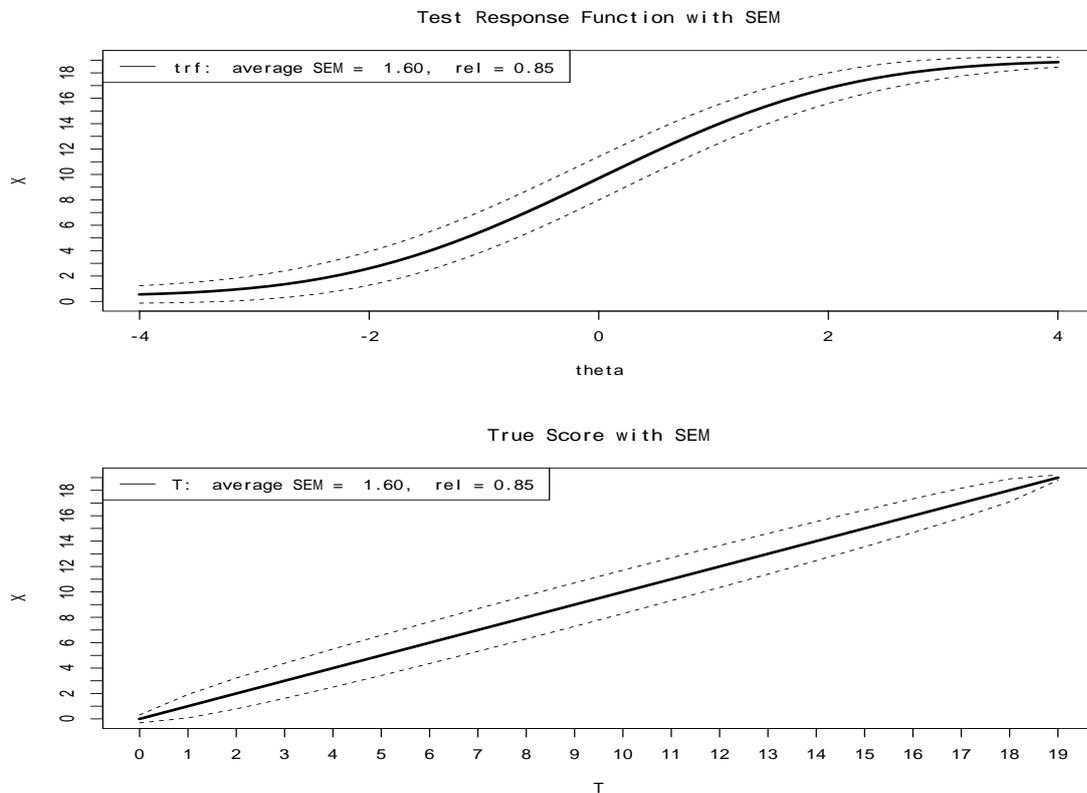


図 7: 測定の標準誤差

## 6 統計的推定

### 6.1 正答数得点からの $\theta$ の推定

項目パラメタが既知の項目からなるテストを受けた被験者の特性値を推定するためには、項目パラメタの推定ほどではないにしても、かなり複雑な統計的な手法を用いなければならない。先に述べた、テスト特性曲線を逆に読みテスト得点から特性値の値を推定する方法は、その簡便法であるということが出来る。テスト特性曲線を逆読みすということは

$$\sum_{j=1}^p u_j = \sum_{j=1}^p P_j(\theta) \quad (92)$$

という方程式を  $\theta$  に関して解くということである。すなわち、得られた得点をその期待値に等しいと置いた方程式である。

しかし、たとえ特性値は比較可能でも、その精度はテストに依存することが知られている。たとえば、小学6年生の算数の学力を測定するのに、小学3年生用のテストを用いたのでは、大部分の人が満点を取ってしまうことになり、結果として、その人の学力が小学校3年相当以上であるということしかわからない。(テスト得点が満点の0点や満点の人の特性値はテスト特性曲線を逆読みすることによっては求めることは出来ない。この方法は、テスト特性曲線がある程度の傾きを持った部分でのみ利用することが出来る。なお、統計的に特性値を推定する方法を用いても同様である。) このことは、ある個人の特性値のきめの細かい測定を行うためには、その人の特性値付近で大きな傾きを持つテスト特性曲線を持ったテストを使うべきであること示している。上記のテスト特性曲線を逆読みする例から明らかなように、各項目やその集合としてのテストは  $\theta$  の全ての範囲で同じ精度を持つわけではない。項目反応理論では、この個人の特性値の推定の精度を以下に示すテスト情報量 (情報関数 Test Information Function  $I(\theta)$ ) として表すが、これは、通常、項目情報量 (Item Information)  $I_j(\theta)$  の和

$$I(\theta) = \sum_{j=1}^p I_j(\theta) \quad (93)$$

として表される。この項目情報量は、2 値項目の場合には

$$I_j(\theta) = 1.7^2 a_j^2 \frac{(P_j(\theta) - c_j)^2 (1 - P_j(\theta))}{(1 - c_j)^2 P_j(\theta)} \quad (94)$$

多値項目の場合には

$$I_j(\theta) = \sum_{k=0}^{m_j} A_{jk}(\theta) \quad (95)$$

$$A_{jk}(\theta) = 1.7^2 a_j^2 \frac{[P_{jk}^+(\theta)(1 - P_{jk}^+(\theta)) - P_{j,k+1}^+(\theta)(1 - P_{j,k+1}^+(\theta))]^2}{P_{jk}(\theta)} \quad (96)$$

となる。一般に、情報関数はある特定の  $\theta$  の付近でピークを持つ関数となる。

したがって、2 値項目の場合、測定すべき人の特性値がおおよそ  $\theta_0$  付近であることがわかっている場合には、項目特性曲線が  $P_j(\theta_0) = 0.5$  となり、しかも識別力の高い項目を選べば  $\theta = \theta_0$  付近で精度の良い測定が可能となることを示している。すなわち、 $b_j \approx \theta_0$  でありかつ  $a_j$  の大きな項目を選べばよい。

## 6.2 重み付き得点からの $\theta$ を推定する際の信頼区間と情報関数

### 6.2.1 信頼区間

2 値項目の重み付きテスト得点

$$X = \sum_{j=1}^p w_j U_j \quad (97)$$

から  $\theta$  を推定することを考える。 $\theta$  を与えたときの得点  $X$  の分布は、局所独立の仮定から、独立なベルヌーイ分布（ただし、その値域が 0 と  $w_j$ ）の和として表現することが出来、結果として得られる分布は、値域が  $(0, \sum_{j=1}^p w_j)$  の多項分布となる。したがって、 $\theta$  の値ごとに、

$$\text{Prob}(X_L < X < X_U | \theta) \approx 1 - \alpha \quad (98)$$

を満たすような上限と下限  $X_U, X_L$  を計算することが出来る。なお、 $\approx$  を用いるのは、 $X$  が離散量であるからである。なお、この分布の条件付き平均は、定義から、テスト特性曲線

$$\mathcal{E}(X|\theta) = \sum_{j=1}^p w_j P_j(\theta) = T_{cc}(\theta) \quad (99)$$

となる。重み付き得点からの  $\theta$  を推定するということは、

$$x = \left( = \sum_{j=1}^p w_j u_j \right) = \sum_{j=1}^p w_j P_j(\theta) \quad (100)$$

という、観測値を期待値と等しいと置いた方程式を  $\theta$  に関して解き、 $\theta$  を求めることを考える。

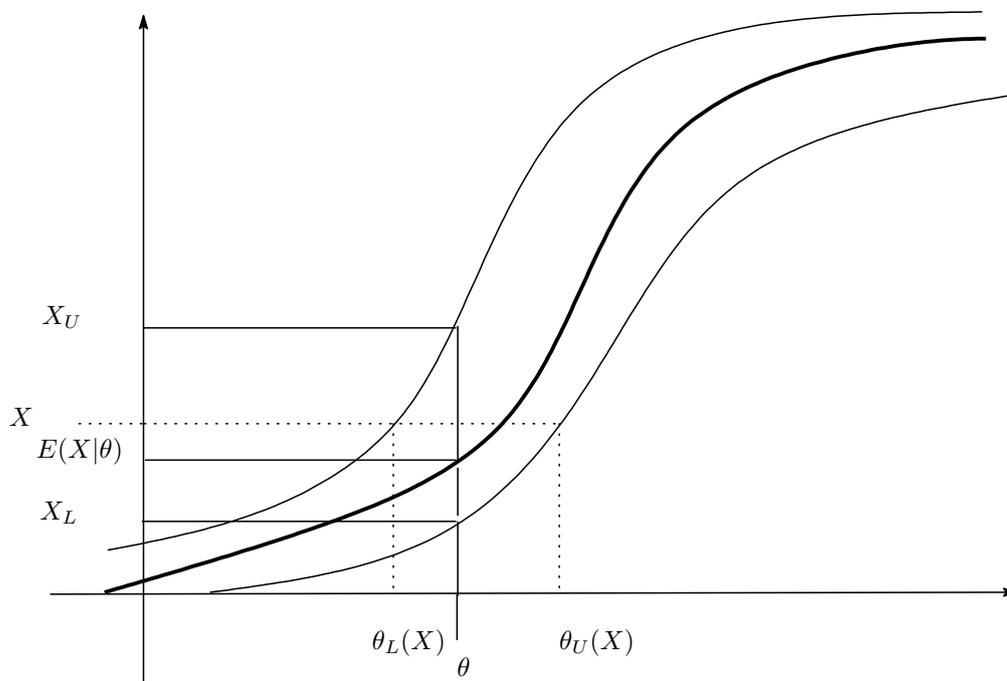


図 8: Conditional Distribution of X given  $\theta$  and Confidence Interval

これを図示したのものが図 8 であるが、テスト特性曲線を太い実線で、上限と下限を繋いだものを細い実線で表している。

また、この図から、 $X$  から横軸に平行な線を引き、それが  $X_U$  を繋いだと交わった点を  $\theta_L(X)$ 、 $X_L$  を繋いだと交わった点を  $\theta_U(X)$  とすれば、 $X$  が区間  $(X_L, X_U)$  に含まれるという事象と、区間  $(\theta_L(X), \theta_U(X))$  が真の  $\theta$  を含むという事象が一対一に対応していることが分かる。したがって、区間  $(\theta_L(X), \theta_U(X))$  が真の  $\theta$  を含む確率は、

$$\text{Prob}(\theta_L(X) < \theta < \theta_U(X)) \approx 1 - \alpha \quad (101)$$

となる。すなわち、これが、重み付き得点  $X$  から  $\theta$  を推定する際の  $100(1 - \alpha)\%$  の信頼区間である。

### 6.2.2 情報関数

この信頼区間の幅  $\theta_U(X) - \theta_L(X)$  は、テスト特性曲線の傾き

$$\frac{dTcc(\theta)}{d\theta} = \sum_{j=1}^p w_j P_j'(\theta) \quad (102)$$

に反比例し、 $X$  の条件付き分布の幅  $X_U - X_L$  に比例すると考えられる。また、局所独立の仮定から、 $\theta$  を与えたときの  $X$  の分布の分散は

$$\text{var}(X|\theta) = \text{var}\left(\sum_{j=1}^p w_j U_j|\theta\right) = \sum_{j=1}^p w_j^2 \text{var}(U_j|\theta) \quad (103)$$

$$= \sum_{j=1}^p w_j^2 P_j(\theta)(1 - P_j(\theta)) \quad (104)$$

で与えられるが、 $X_U - X_L$  は、およそ、上記の標準偏差に比例する

$$X_U - X_L \propto \sqrt{\text{var}(X|\theta)} = \sqrt{\sum_{j=1}^p w_j^2 P_j(\theta)(1 - P_j(\theta))} \quad (105)$$

と考えられる。

そこで、信頼区間の幅の短さ、すなわち推定の精度を表す指標として、これらの量の商の2乗の逆数を情報関と数して定義する。(Birnbaum, 1968, sec. 17.7, p418)

$$I(\theta, \sum_{j=1}^p w_j U_j) = \frac{\left(\frac{dTcc(\theta)}{d\theta}\right)^2}{\text{var}(X|\theta)} = \frac{\left(\sum_{j=1}^p w_j P_j'(\theta)\right)^2}{\sum_{j=1}^p w_j^2 P_j(\theta)(1 - P_j(\theta))} \quad (106)$$

すなわち、情報関数  $I(\theta, \sum_{j=1}^p w_j U_j)$  を与えられた場合、 $x$  から推定される  $\theta$  の信頼区間の幅は、 $1/\sqrt{I(\theta, \sum_{j=1}^p w_j U_j)}$  に比例することになる。

情報関数を  $\theta$  の関数として見た場合、その値域は正であり、 $\theta$  の絶対値が大きくなるにしたがって0に近づく。したがって、 $\theta$  の定義域の中で、最大値や極大値を持つ関数である。また、情報関数は、各項目に掛かる重みを変化させることによりその値を変化させるため、これを重みの多変数関数としてみることが出来る。

この情報関数の最大値を与える重みに関しては、情報関数を重み  $w_g$  で偏微分すると、

$$Q_j(\theta) = 1 - P_j(\theta) \quad (107)$$

$$(108)$$

として、

$$\begin{aligned}
\frac{\partial I(\theta, \sum_{j=1}^p w_j U_j)}{\partial w_g} &= \frac{2}{\left(\sum_{j=1}^p w_j^2 P_j(\theta) Q_j(\theta)\right)^2} \times \\
&\quad \left( P'_g(\theta) \left(\sum_{j=1}^p w_j P'_j(\theta)\right) \sum_{j=1}^p w_j^2 P_j(\theta) Q_j(\theta) - w_g P_g(\theta) Q_g(\theta) \left(\sum_{j=1}^p w_j P'_j(\theta)\right)^2 \right) \\
&= 2P'_g(\theta) \frac{\sum_{j=1}^p w_j P'_j(\theta)}{\sum_{j=1}^p w_j^2 P_j(\theta) Q_j(\theta)} - 2w_g P_g(\theta) Q_g(\theta) \frac{\left(\sum_{j=1}^p w_j P'_j(\theta)\right)^2}{\left(\sum_{j=1}^p w_j^2 P_j(\theta) Q_j(\theta)\right)^2}
\end{aligned} \tag{109}$$

となることに注意すると、いわゆる局所的最適重み (Locally Best Weight)

$$w_j = w_j(\theta) = \frac{P'_j(\theta)}{P_j(\theta)(1 - P_j(\theta))} \tag{110}$$

もしくはその定数倍を用いると、この偏微分は 0 となる。したがって、この重みを用いると、情報関数が全ての  $\theta$  の値において最大値、すなわち、テスト情報量

$$I(\theta) = \sum_{j=1}^p \frac{P'_j(\theta)^2}{P_j(\theta)(1 - P_j(\theta))} = \sum_{j=1}^p I_j(\theta) \tag{111}$$

を取り、したがって、信頼区間の幅が最小になると考えられる。また、このことは、ベイズ推定における事後標準偏差の大きさにも対応しているはずである。

ただし、この局所的最適重みは、その名が示すとおり、 $\theta$  に依存する重みとなっているため、実際にこの重みを利用することは難しい。

### 6.2.3 局所的最適重み

(110) 式 の形は、これまでに、以下のような文脈で定義されている。

- Lord and Novick(1968) の Birnbaum の節による定義  
まず、 $\theta$  が  $\theta_1, \theta_2$  の二つの値しか取らない場合の十分統計量としての best weight を定義し、(Birnbaum, 1968, sec. 18.3)、その  $\theta_2 \rightarrow \theta_1$  の場合の極限として定義されたもの。(Birnbaum, 1968, sec.19.3)
- 情報関数の最大化  
Cauchy の不等式を用いることにより、テスト情報関数に等しくなる重みとして定義されている。(Birnbaum, 1968, sec. 20.2)
- 最尤方程式の解  
項目反応パターンを

$$u_{kj} = \begin{cases} 1 & \text{if a person's response to the } j\text{-th item is at the } k\text{-th category} \\ 0 & \text{otherwise} \end{cases} \tag{112}$$

とする。  $u_{kj}$  をもとに  $\theta$  を推定する際に対数尤度関数は

$$\ln L = \sum_{j=1}^p \sum_{k=0}^{m_j} u_{kj} \log P_{kj}(\theta) \tag{113}$$

であるが、これを  $\theta$  で微分すると

$$d \ln L = \sum_{j=1}^p \sum_{k=0}^{m_j} u_{kj} \frac{P'_{kj}(\theta)}{P_{kj}(\theta)} \quad (114)$$

となるが、この項目反応に掛かる係数

$$w_{kj}(\theta) = \frac{P'_{kj}(\theta)}{P_{kj}(\theta)} \quad (115)$$

を最適重みと定義する。これは Samejima(1969) が basic function として定義しているものである。したがって、最適重みで重みを付けた合計得点が 0 となる様な  $\theta$  の値が最尤推定値であり、このことから、最尤推定値は最適重みで重みを付けた合計得点の関数であると解釈できる。

3PLM の場合、 $u_{0j} = 1 - u_{1j}$  となるため上式の  $u_{j1}$  に掛かる係数は

$$w_j(\theta) = \frac{P'_{1j}(\theta)}{P_{1j}(\theta)(1 - P_{1j}(\theta))} \quad (116)$$

となるが、最尤推定値は

$$\sum_{j=1}^p w_j(\theta) u_{1j} = \sum_{j=1}^p \frac{P'_{1j}(\theta)}{1 - P_{1j}(\theta)} \quad (117)$$

の解となる。(Birnbbaum, 1968, sec. 20.3, p456)

### 6.3 項目反応パターンからの特性値の推定

項目反応理論では、項目特性曲線もしくは項目オペレーティング特性曲線でその項目への正答確率もしくは各カテゴリへの反応確率が与えられるが、項目が複数ある場合、特性値  $\theta$  を固定したときに、異なる項目への反応が互いに独立になる、

$$\text{Prob}(U_1 = k_1 \& U_2 = k_2 \dots \& | \theta) = P_{1k_1}(\theta) \times P_{2k_2}(\theta) \times \dots \quad (118)$$

という局所独立性 (Local Independence) が仮定されている。この仮定はかなり強い仮定であり、実際にこの仮定が満たされるためには、テストに含まれる項目への反応が、先に示された項目の反応に依存しないようにテストが作られているのみならず、全ての項目がある次元の特性値  $\theta$  を測定している必要がある。

局所独立性を仮定できる場合、ある被験者の  $p$  個の項目への反応を

$$u_{jk} = \begin{cases} 1 & \text{第 } j \text{ 番目の項目得点が } k \text{ 点} \\ 0 & \text{それ以外} \end{cases}, \quad k = 0, 1, 2, \dots, m_j, \quad j = 1, 2, \dots, p \quad (119)$$

とすると、この特定の項目反応を得る確率は

$$\text{Prob}(\mathbf{U} = \mathbf{u} | \theta) = \prod_{j=1}^p \prod_{k=0}^{m_j} P_{jk}^{u_{jk}}(\theta) \quad (120)$$

と書き表せる。これは、2値項目の場合、

$$P_j(\theta) = P_{j1}(\theta), \quad 1 - P_j(\theta) = P_{j0}(\theta), \quad u_j = u_{j1}, \quad 1 - u_j = u_{j0} \quad (121)$$

として

$$\text{Prob}(\mathbf{U} = \mathbf{u} | \theta) = \prod_{j=1}^p (1 - P_j(\theta))^{1-u_j} P_j(\theta)^{u_j} \quad (122)$$

となる。したがって、これをパラメタ  $\theta$  の尤度  $L$  と見なせば、これを最大とする  $\theta$  の値が反応  $\mathbf{u}$  を持つ被験者の特性値の推定値となる。(この場合、推定値の精度は、次節に示す、テスト情報量として与えられる。)

特に、2PLM の場合、対数尤度関数は以下のようになるが、

$$\ln L = \ln \left( \prod_{j=1}^p (1 - P_j(\theta))^{1-u_j} P_j(\theta)^{u_j} \right) \quad (123)$$

$$= \sum_{j=1}^p u_j \ln P_j(\theta) + (1 - u_j) \ln(1 - P_j(\theta)) \quad (124)$$

$$= \sum_{j=1}^p u_j (\ln P_j(\theta) - \ln(1 - P_j(\theta))) + \ln(1 - P_j(\theta)) \quad (125)$$

$$= \sum_{j=1}^p u_j \ln \frac{P_j(\theta)}{1 - P_j(\theta)} + \ln(1 - P_j(\theta)) \quad (126)$$

$$= \sum_{j=1}^p u_j \times Da_j(\theta - b_j) + \ln(1 - P_j(\theta)) \quad (127)$$

$$= D\left(\sum_{j=1}^p a_j u_j\right)\theta + \sum_{j=1}^p \ln(1 - P_j(\theta)) - D\left(\sum_{j=1}^p a_j b_j u_j\right) \quad (128)$$

この形から、対数尤度関数の極値は、識別力で重みを付けた重み付き正答数得点

$$x^* = \sum_{j=1}^p a_j u_j \quad (129)$$

のみに依存していることがわかる。したがって、 $x^*$  が同じ値の人の  $\theta$  の推定量は同じとなる。これを  $\theta$  の推定に際する十分統計量と呼ぶ。

ただし、尤度関数の値そのものは項目反応パターンに依存し、たとえば、1PLM の場合、難しい項目 5 題のみに正答した人の尤度は、易しい項目反応理論のみ 5 題正答した人の尤度より低くなる。

また、 $\theta$  の事前分布を  $h(\theta)$  としベイズ推定法を用いれば、 $\theta$  の事後分布は

$$h(\theta|\mathbf{u}) \propto \prod_{j=1}^p \prod_{k=0}^{m_j} P_{jk}(\theta)^{u_{jk}} \times h(\theta) \quad (130)$$

または

$$h(\theta|\mathbf{u}) \propto \prod_{j=1}^p (1 - P_j(\theta))^{1-u_j} P_j(\theta)^{u_j} \times h(\theta) \quad (131)$$

となり、その事後平均が  $\theta$  の推定値、その事後標準偏差が推定値の精度の悪さを表すことになる。なお、母集団が一つしかない場合、通常は標準正規分布を仮定するため、 $h(\theta|\boldsymbol{\mu}) = h(\theta)$  としている。

2PLM の場合、最尤解と同様に、 $x^*$  が十分統計量となる。また、同じ  $x^*$  を持つ人の事後分布は同一のものとなる。(尤度関数の項目反応パターンによる値の変化は、事後分布を計算する時の比例定数に吸収されてしまう。)

以下に、特性値の推定のための尤度関数、対数尤度関数、事前分布、事後分布を示す。2PLM を用い、項目パラメタは  $a_1 = 0.8, b_1 = 0.0, a_2 = 1.0, b_2 = 0.8, a_3 = 0.7, b_3 = 0.5, a_4 = 1.1, b_4 = -0.5$  であった。この図は 1011 という項目反応に対応している。最尤推定値は 0.8814、EAP は 0.5920、事後標準偏差は 0.6162 であった。また、事後分布の最頻値は 0.5579 であった。左側の二つの図から、尤度関数と対数尤度関数が、ともに、0.8814 付近で最大値を持つことが分かる。右の図では、上段に示した事前分布(標準正規分布)が、最下段に示す事後分布に於いては位置、標準偏差ともに変化していることが分かる。

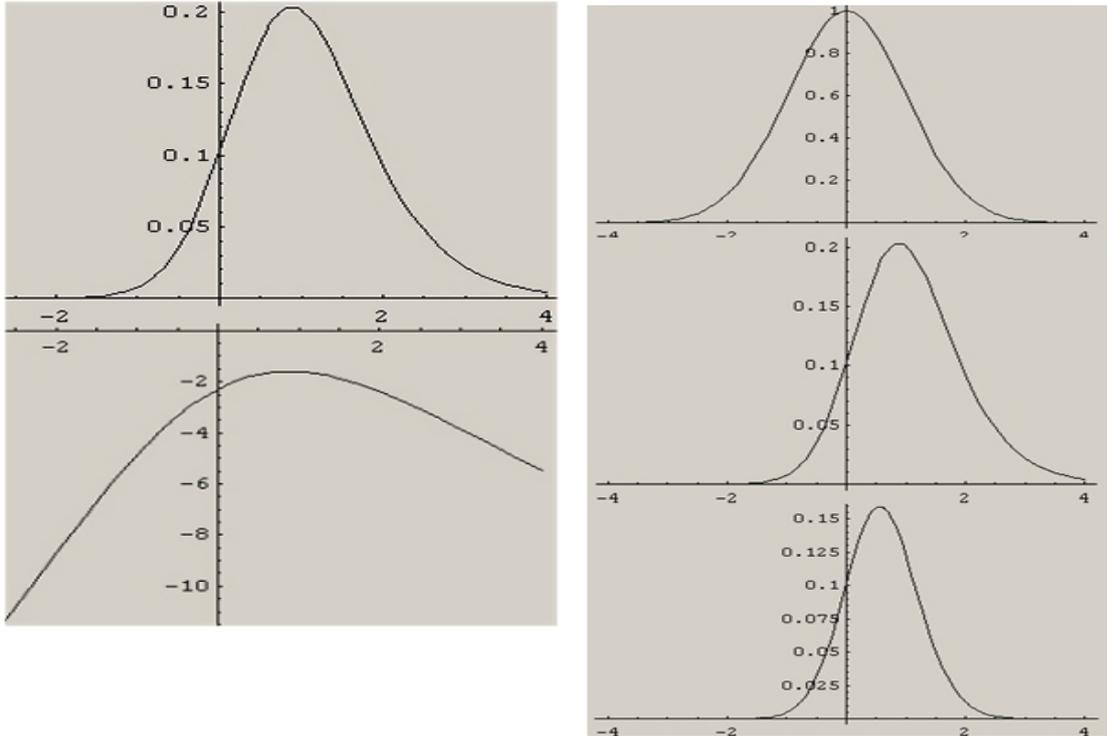


図 9: 特性値の推定

### 6.3.1 重み付き合計得点を基にした特性値の推定

テスト得点

$$X = \sum_{j=1}^p w_j U_j \quad (132)$$

を与えたときの特性値  $\theta$  の（事後）分布は

$$h(\theta|X) = \frac{f(x|\theta)h(\theta)}{\int_{\Theta} f(x|\theta)h(\theta)d\theta} \quad (133)$$

で与えられる。ただし、 $h(\theta)$  は  $\theta$  の事前分布、 $f(x|\theta)$  は  $\theta$  を与えたときのテスト得点の分布（確率関数）で、多項分布の重み付きの和として与えられるものである。（全ての重みが等しい2値データの場合には複合2項分布となる。なお、これらの分布の計算の仕方に関しては Mayekawa & Arai (2008) を参照のこと。）

したがって、この分布の平均や標準偏差

$$\mu_{\theta|x} = \int_{\Theta} \theta h(\theta|x) d\theta \quad (134)$$

$$\sigma_{\theta|x} = \sqrt{\int_{\Theta} (\theta - \mathcal{E}(\theta|x))^2 h(\theta|x) d\theta} \quad (135)$$

を用いて、特性値  $\theta$  の推測を行うことが出来る。

なお、2PL の場合、合計点を計算する際の重みを識別力に比例するように取れば

$$w_j \propto a_j \quad (136)$$

この重み付き合計点からの  $\theta$  の推定は項目反応からの推定と一致し、 $\theta$  の事後標準偏差が最小となる。また、3PL の場合は、重みとして、

$$w_j = 2^{j-1} \quad (137)$$

を用いると、項目反応からの  $\theta$  の推定と一致する。

## 6.4 項目パラメタの推定：2値項目の場合

### 6.4.1 同時最尤法

項目パラメタの集合を  $\xi$  と書く。個人  $i$  のデータを  $u_i$ 、また個人  $i$  の特性値の値を  $\theta_i$  とすると、 $n$  人のデータの同時分布は特性値  $\theta$  を与えたときの項目反応の同時分布、すなわち、パラメタ  $\theta_i, i = 1, 2, \dots, n, \xi_j, j = 1, 2, \dots, p$  の尤度は

$$L(\theta, \xi) = f(U|\theta, \xi) = \prod_{i=1}^n \prod_{j=1}^p P_j(\theta_i)^{u_{ij}} (1 - P_j(\theta_i))^{(1-u_{ij})} \quad (138)$$

である。これを以下のようにして  $\theta_i, i = 1, 2, \dots, n, \xi_j, j = 1, 2, \dots, p$  に関して最大化する。 $\xi$  が与えられた時には、 $\theta_i, i = 1, 2, \dots, n$  が互いに無関係であるため、 $\theta_i$  は

$$L(\theta_i, \xi) = \prod_{j=1}^p P_j(\theta_i)^{u_{ij}} (1 - P_j(\theta_i))^{(1-u_{ij})} \quad (139)$$

を最大にするように求められる。(前節の方法)

また、 $\theta$  が与えられた時には、 $\xi_j, j = 1, 2, \dots, p$  が互いに無関係であるため、 $\xi_j$  は

$$L(\theta, \xi_j) = \prod_{i=1}^n P_j(\theta_i)^{u_{ij}} (1 - P_j(\theta_i))^{(1-u_{ij})} \quad (140)$$

を最大とするように求められる。(ロジスティック回帰)

したがって、適当な初期値から、上記のふたつの条件付き最尤解を求めるステップを繰り返せば、その各繰り返しごとに、 $L(\theta, \xi)$  の値は増大していき、最終的に極大値に達するものと考えられる。

なお、 $\theta$  に関しては、その原点と単位が不定であるため、各繰り返しごとに、たとえば、

$$\sum_{i=1}^n \theta_i = 0, \quad \text{and} \quad \sum_{i=1}^n \theta_i^2 = n \quad (141)$$

という制約を掛ける必要がある。

### 6.4.2 周辺最尤法

項目パラメタの集合を  $\xi$  と書く。特性値  $\theta$  を与えたときの項目反応の分布(確率関数)は

$$f(u|\theta, \xi) = \prod_{j=1}^p P_j(\theta)^{u_j} (1 - P_j(\theta))^{(1-u_j)} \quad (142)$$

である。従って、 $\theta$  の事前分布(母集団における分布)の密度関数を  $h(\theta)$  とすれば、項目反応の周辺分布( $\theta$  に依存しない分布)は

$$f(u|\xi) = \int_{-\infty}^{+\infty} \prod_{j=1}^p P_j(\theta)^{u_j} (1 - P_j(\theta))^{(1-u_j)} h(\theta) d\theta \quad (143)$$

となる。実際のデータは、 $\theta$  が観測されないので、この分布からのサンプルであると考えられる。  
従って、尤度関数（周辺尤度関数）、すなわち、 $N$  人分のデータの同時分布は

$$mL(\boldsymbol{\xi}) = \prod_{i=1}^N f(\mathbf{u}_i|\boldsymbol{\xi}) = \prod_{i=1}^N \int_{-\infty}^{+\infty} \prod_{j=1}^p P_j(\theta)^{u_{ij}} (1 - P_j(\theta))^{(1-u_{ij})} h(\theta) d\theta \quad (144)$$

となる。この  $mL(\boldsymbol{\xi})$  を最大とするような  $\boldsymbol{\xi}$  を求める。

全ての項目に答えてはいない場合（欠測値がある場合）

$$d_{ij} = \begin{cases} 0 & \text{答えていない} \\ 1 & \text{答えている} \end{cases} \quad (145)$$

を用いれば、

$$\text{Prob}(\mathbf{U} = \mathbf{u}|\theta, \mathbf{d}) = \prod_{j=1}^p P_j(\theta)^{d_j u_j} (1 - P_j(\theta))^{d_j (1-u_j)} \quad (146)$$

と書けるため、尤度関数は

$$mL(\boldsymbol{\xi}) = \prod_{i=1}^N f(\mathbf{u}_i|\boldsymbol{\xi}) = \prod_{i=1}^N \int_{-\infty}^{+\infty} \prod_{j=1}^p P_j(\theta)^{d_{ij} u_{ij}} (1 - P_j(\theta))^{d_{ij} (1-u_{ij})} h(\theta) d\theta \quad (147)$$

と書き表される。

また、 $\theta$  の分布を規定する密度関数  $h(\theta)$  に、集団の分布を規定するパラメタ  $\boldsymbol{\mu}$  が含まれる場合は、それを  $h(\theta|\boldsymbol{\mu})$  と書き、M step で  $\boldsymbol{\mu}$  に関する最大化も行う。

(144) 式 の様な、積分を含む形の尤度関数を最大とするようなパラメタの値を求めるためには EM アルゴリズムを用いるのが簡単である。EM アルゴリズムは、 $\theta$  を欠測データと見なし、その、観測値を与えた時の条件付き分布を計算し (E-step)、それに関する対数完全データ尤度の期待値を最大化する (M-step) ことを繰り返す方法である。

### 6.4.3 EM アルゴリズムによるパラメタの推定（一般形）

EM アルゴリズム は以下のような積分を含む形をした尤度関数の最大化に際して用いられる。

$$mL = \prod_{i=1}^N f(\mathbf{u}_i|\boldsymbol{\Xi}, \boldsymbol{\mu}) \quad (148)$$

$$= \prod_{i=1}^N \int_{\Theta} f(\mathbf{u}_i, \theta|\boldsymbol{\Xi}, \boldsymbol{\mu}) d\theta \quad (149)$$

$$= \prod_{i=1}^N \int_{\Theta} f(\mathbf{u}_i|\theta, \boldsymbol{\Xi}) h(\theta|\boldsymbol{\mu}) d\theta \quad (150)$$

ただし、

$$\left\{ \begin{array}{l} \theta \\ \mathbf{u}_i \\ \boldsymbol{\Xi} \\ \boldsymbol{\mu} \\ f(\mathbf{u}_i|\theta, \boldsymbol{\Xi}) \end{array} \right. \begin{array}{l} = \text{潜在変数 (} r \text{ 次元)} \\ = \text{個人 } i \text{ のデータ (} p \text{ 次元)} \\ = \text{潜在変数を与えたときの } \mathbf{u} \text{ の分布を規定するパラメタの集合} \\ = \theta \text{ の分布のパラメタ} \\ = \theta \text{ を与えたときの } \mathbf{u}_i \text{ の密度関数または確率関数} \\ \text{IRT の場合は } \prod_{j=1}^p P(\theta_i|\boldsymbol{\xi}_j)^{u_{ij}} (1 - P(\theta_i|\boldsymbol{\xi}_j))^{1-u_{ij}} \end{array} \quad (151)$$

である。この形は尤度関数  $f$  の  $h$  に関する期待値であると解釈できるため、周辺尤度関数と呼ばれる。対数尤度は

$$\ln mL = \sum_{i=1}^N \ln f(\mathbf{u}_i | \Xi, \mu) \quad (152)$$

$$= \sum_{i=1}^N \ln \left( \int_{\Theta} f(\mathbf{u}_i, \theta | \Xi, \mu) d\theta \right) \quad (153)$$

である。(積分を和に置き換えてもあまり綺麗な形ではない。)

EM アルゴリズムのアウトライン

1. パラメタ  $\Xi$  と  $\mu$  の初期値  $\Xi_0, \mu_0$  を求める。
2. 対数完全データ尤度  $\ln L(\Xi, \mu)$  を計算する。
3. データ  $\mathbf{u}$  を与えたときの潜在変数  $\theta$  の条件付き分布または確率  $h(\theta | \mathbf{u}, \Xi_0, \mu_0)$  を計算する。
4. 現在のパラメタの値を使って、期待対数完全データ尤度  $E \ln L(\Xi, \mu)$  を計算する。
5. 期待対数完全データ尤度  $E \ln L(\Xi, \mu)$  を最大とするパラメタ  $\Xi$  と  $\mu$  の値を求める。
6. 周辺尤度の値が大きく改善したなら step 2 へ、そうでない場合は収束したと見なす。

潜在変数が連続的な場合の詳細

記法:

$\mathbf{u} = p \times 1$  観測されるデータ ( $p$  次元ベクトル)

$\theta = r \times 1$  連続的な潜在変数

$\mathbf{u}_i, i = 1 \dots N$ , 個人  $i$  のデータ

$\theta_i, i = 1 \dots N$ , 個人  $i$  の潜在変数の値

$\Xi, \mu =$  推定すべきパラメタ

1. 尤度関数 ( $\theta$  を与えたときのデータの同時分布)

$$f(\mathbf{u} | \theta, \Xi) = \prod_{j=1}^p f_j(u_j | \theta, \Xi) \quad \text{if locally independent.} \quad (154)$$

2.  $\theta$  の事前 (周辺) 分布

$$h(\theta | \mu) \quad (155)$$

3.  $\mathbf{U}, \theta$  の同時分布

$$f(\mathbf{u}, \theta | \Xi, \mu) = f(\mathbf{u} | \theta, \Xi) h(\theta | \mu) \quad (156)$$

4. データの周辺分布  $U$  (これが周辺尤度関数の基本となるもの)

$$f(\mathbf{u} | \Xi, \mu) = \int_{\Theta} f(\mathbf{u}, \theta | \Xi, \mu) d\theta \quad (157)$$

5. データを与えたときの  $\theta$  の条件付き (事後) 分布

$$h(\theta | \mathbf{u}, \Xi, \mu) = \frac{f(\mathbf{u}, \theta | \Xi, \mu)}{f(\mathbf{u} | \Xi, \mu)} \quad (158)$$

$$\propto f(\mathbf{u}, \theta | \Xi, \mu) = f(\mathbf{u} | \theta, \Xi) h(\theta | \mu) \quad (159)$$

6. 周辺尤度関数 (これを最大化したい)

$$mL(\Xi, \boldsymbol{\mu}) = \prod_{i=1}^N f(\mathbf{u}_i | \Xi, \boldsymbol{\mu}) \quad (160)$$

$$= \prod_{i=1}^N \int_{\Theta} f(\mathbf{u}_i, \theta | \Xi, \boldsymbol{\mu}) d\theta \quad (161)$$

7. 完全データ尤度

$$L(\Xi, \boldsymbol{\mu}) = \prod_{i=1}^N f(\mathbf{u}_i, \theta_i | \Xi, \boldsymbol{\mu}) \quad (162)$$

$$= \prod_{i=1}^N f(\mathbf{u}_i | \theta_i, \Xi) h(\theta_i | \boldsymbol{\mu}) \quad (163)$$

8. 対数完全データ尤度

$$\ln L(\Xi, \boldsymbol{\mu}) = \sum_{i=1}^N \ln f(\mathbf{u}_i, \theta_i | \Xi, \boldsymbol{\mu}) \quad (164)$$

$$= \sum_{i=1}^N \ln f(\mathbf{u}_i | \theta_i, \Xi) + \sum_{i=1}^N \ln h(\theta_i | \boldsymbol{\mu}) \quad (165)$$

9. 期待対数完全データ尤度

これは  $\ln L$  の期待値を  $\theta$  の  $U$  を与えたときの条件付き (事後) 分布に関して計算したもの。ただし、条件付き (事後) 分布の中に出てくる  $\Xi$  と  $\boldsymbol{\mu}$  は  $\Xi_0$  と  $\boldsymbol{\mu}_0$  と記、既知とする。

$$\mathcal{E} \ln L(\Xi, \boldsymbol{\mu}) = \mathcal{E}(\ln L(\Xi, \boldsymbol{\mu}) | U, \Xi_0, \boldsymbol{\mu}_0) \quad (166)$$

$$= \sum_{i=1}^N \mathcal{E}(\ln f(\mathbf{u}_i, \theta | \Xi, \boldsymbol{\mu}) | \mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0) \quad (167)$$

$$= \sum_{i=1}^N \mathcal{E}(\ln f(\mathbf{u}_i | \theta, \Xi) | \mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0) + \sum_{i=1}^N \mathcal{E}(\ln h(\theta | \boldsymbol{\mu}) | \mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0) \quad (168)$$

$$= \sum_{i=1}^N \int_{\Theta} \ln f(\mathbf{u}_i | \theta, \Xi) h(\theta | \mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0) d\theta \\ + \sum_{i=1}^N \int_{\Theta} \ln h(\theta | \boldsymbol{\mu}) h(\theta | \mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0) d\theta \quad (169)$$

ただし、積分に関しては、 $Q$  個の離散点  $A_q, q = 1, 2, \dots, Q$  を用い、

$$\mathcal{E} \ln L(\Xi, \boldsymbol{\mu}) \approx \sum_{i=1}^N \sum_{q=1}^Q \ln f(\mathbf{u}_i | A_q, \Xi) h^*(A_q | \mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0) + \sum_{i=1}^N \sum_{q=1}^Q \ln h(A_q | \boldsymbol{\mu}) h^*(A_q | \mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0) \quad (170)$$

$$= \sum_{q=1}^Q \sum_{i=1}^N h^*(A_q | \mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0) \ln f(\mathbf{u}_i | A_q, \Xi) \\ + \sum_{q=1}^Q \left( \sum_{i=1}^N h^*(A_q | \mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0) \right) \ln h(A_q | \boldsymbol{\mu}) \quad (171)$$

として計算する。なお、この中に出てくる  $h^*(A_q|\mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0)$  は

$$h^*(A_q|\mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0) = \frac{h(A_q|\mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0)}{\sum_{q=1}^Q h(A_q|\mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0)} \quad (172)$$

で定義されるが、これは  $h(A_q|\mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0)$  を全ての離散点に関して加えると 1 になるように基準化したものである。

M step では上記の  $\varepsilon \ln L(\Xi, \boldsymbol{\mu})$  を最大とするようなパラメタ  $\Xi, \boldsymbol{\mu}$  を求めることになる。ただし、 $\Xi_0, \boldsymbol{\mu}_0$  は定数として取り扱う。

項目反応理論への応用

- E-step

$\theta$  の事後分布は (159) 式 より

$$h(\theta|\mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0) \propto f(\mathbf{u}_i|\theta, \Xi_0)h(\theta|\boldsymbol{\mu}_0) \quad (173)$$

$$= \prod_{j=1}^p P(\theta|\xi_{0j})^{u_{ij}} (1 - P(\theta|\xi_{0j}))^{1-u_{ij}} h(\theta|\boldsymbol{\mu}_0) \quad (174)$$

$$= \prod_{j=1}^p P(\theta|\xi_{0j})^{u_{ij}} (1 - P(\theta|\xi_{0j}))^{1-u_{ij}} \times \frac{1}{\sigma_0} \exp(-\frac{1}{2}(\theta - \mu_0)^2/\sigma_0^2) \quad (175)$$

ただし、最後の行は事前分布に正規分布を仮定した形であり、 $\boldsymbol{\mu}_0 = \{\mu_0, \sigma_0\}$  としている。要は事前分布の密度関数を書けばよいだけなので、(174) 式の形でよい。

したがって、これを  $Q$  個の  $\theta$  に関して加えて 1 と正規化したものは、(172) 式 より、

$$h^*(A_q|\mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0) = \frac{h(A_q|\mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0)}{\sum_{q=1}^Q h(A_q|\mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0)} \quad (176)$$

$$= \frac{\prod_{j=1}^p P(A_q|\xi_{0j})^{u_{ij}} (1 - P(A_q|\xi_{0j}))^{1-u_{ij}} h(A_q|\boldsymbol{\mu}_0)}{\sum_{q=1}^Q \prod_{j=1}^p P(A_q|\xi_{0j})^{u_{ij}} (1 - P(A_q|\xi_{0j}))^{1-u_{ij}} h(A_q|\boldsymbol{\mu}_0)} \quad (177)$$

となる。なお、この値は、パラメタを含まないため (添字 0 がついているものは別) 以降、以下の記号を用いる。

$$h_{iq} = h^*(A_q|\mathbf{u}_i, \Xi_0, \boldsymbol{\mu}_0) \quad (178)$$

この記法を用いると期待対数完全データ尤度 (171) 式 は以下のように書ける。

$$\mathcal{E} \ln L(\Xi, \boldsymbol{\mu}) \approx \sum_{q=1}^Q \sum_{i=1}^N h_{iq} \ln f(\mathbf{u}_i | A_q, \Xi) + \sum_{q=1}^Q \left( \sum_{i=1}^N h_{iq} \right) \ln h(A_q | \boldsymbol{\mu}) \quad (179)$$

$$= \sum_{q=1}^Q \sum_{i=1}^N h_{iq} \sum_{j=1}^p (u_{ij} \ln P(A_q | \boldsymbol{\xi}_j) + (1 - u_{ij}) \ln(1 - P(A_q | \boldsymbol{\xi}_j))) + \sum_{q=1}^Q \left( \sum_{i=1}^N h_{iq} \right) \ln h(A_q | \boldsymbol{\mu}) \quad (180)$$

$$= \sum_{q=1}^Q \sum_{j=1}^p \left[ \left( \sum_{i=1}^N h_{iq} u_{ij} \right) \ln P(A_q | \boldsymbol{\xi}_j) + \left( \sum_{i=1}^N h_{iq} - \sum_{i=1}^N h_{iq} u_{ij} \right) \ln(1 - P(A_q | \boldsymbol{\xi}_j)) \right] + \sum_{q=1}^Q \left( \sum_{i=1}^N h_{iq} \right) \ln h(A_q | \boldsymbol{\mu}) \quad (181)$$

$$= \sum_{j=1}^p \sum_{q=1}^Q [r_{qj} \ln P(A_q | \boldsymbol{\xi}_j) + (n_q - r_{qj}) \ln(1 - P(A_q | \boldsymbol{\xi}_j))] + \sum_{q=1}^Q n_q \ln h(A_q | \boldsymbol{\mu}) \quad (182)$$

ただし、

$$r_{qj} = \sum_{i=1}^N h_{iq} u_{ij} \quad (183)$$

$$h_q = \sum_{i=1}^N h_{iq} \quad (184)$$

である。当然であるがこれも添字 0 無しのパラメタを含まない。

- M-step

M-step では、(182) 式 を最大とする、 $\Xi$  と  $\boldsymbol{\mu}$  を求めることになる。(添字 0 付きのパラメタは定数として扱う。) 上記の期待対数完全尤度の形は、その第 1 項は  $(r_{qj}, n_q, A_q)$ ,  $q = 1, 2, \dots, Q$  をデータとして与えたときの、 $j = 1, 2, \dots, p$  個の独立なロジスティック回帰の形をしている。(  $A_q$  という説明変数を持った個体が項目  $j$  に  $n_q$  回回答し、そのうち  $n_{jq}$  回正反応をした、というデータ。) したがって、その最大化は簡単である。

また、第 2 項は、密度関数  $h(A_q | \boldsymbol{\mu})$  を持つデータの  $(A_q, n_q)$ ,  $q = 1, 2, \dots, Q$  という度数分布から、そのパラメタを推定する際の尤度

$$\prod_{q=1}^Q h(A_q | \boldsymbol{\mu})^{n_q} \quad (185)$$

の対数であるから、最尤解を用いれば  $\boldsymbol{\mu}$  が求まる。

また、通常、 $\Xi$  と  $\boldsymbol{\mu}$  は数学的に無関係であるため、これらの二つの最大化は独立に行うことが出来る。

なお、通常、 $\theta$  の原点と単位に関する不定性を取り除くため、 $\theta$  の平均と分散を、それぞれ、1 と 0 に固定することが行われる。従って、1 集団の場合に、 $\theta$  の分布として正規分布を仮定する場合には、そのパラメタ  $\boldsymbol{\mu} = \{\mu, \sigma^2\}$  を推定する必要はない。しかし、多集団 (multi-group) の場合は、一つの集団の  $\theta$  の平均と分散は固定し、それ以外の集団の  $\boldsymbol{\mu}$  を M-step に於いて推定する事になる。

## 6.5 項目パラメタの推定：多値項目の場合

項目パラメタの集合を  $\xi$  と書く。特性値  $\theta$  を与えたときの項目反応の分布（確率関数）は

$$f(u|\theta, \xi) = \prod_{j=1}^p \prod_{k=0}^{m_j} P_{jk}(\theta)^{u_{jk}} \quad (186)$$

である。従って、 $\theta$  の事前分布（母集団における分布）の密度関数を  $h(\theta)$  とすれば、項目反応の周辺分布（ $\theta$  に依存しない分布）は

$$f(u|\xi) = \int_{-\infty}^{+\infty} \prod_{j=1}^p \prod_{k=0}^{m_j} P_{jk}(\theta)^{u_{jk}} h(\theta) d\theta \quad (187)$$

となる。実際のデータは、 $\theta$  が観測されないので、この分布からのサンプルであると考えられる。

従って、尤度関数、すなわち、 $N$  人分のデータの同時分布は

$$mL(\xi) = \prod_{i=1}^N f(u_i|\xi) = \prod_{i=1}^N \int_{-\infty}^{+\infty} \prod_{j=1}^p \prod_{k=0}^{m_j} P_{jk}(\theta)^{u_{ijk}} h(\theta) d\theta \quad (188)$$

となる。この  $mL(\xi)$  を最大とするような  $\xi$  を求める。

欠測値がある場合には、2 値項目と同様の取り扱いが可能である。

実際には 2 値項目の場合と同様に EM アルゴリズムを用いてパラメタの推定を行う。

## 7 他の方法との比較

### 7.1 因子分析モデルとの比較

項目反応理論を、同じ局所独立を仮定する潜在特性モデルである因子分析モデル（同族テストモデル）と比較すると、因子分析の場合は (4) 式において  $\theta$  を多次元ベクトルとし、

$$f_j(y_j|\theta) = f_N(m_j + \lambda'_j \theta, \psi_j) \quad (189)$$

ただし、 $y_j$  は第  $j$  番目の観測される変数（連続量）であり、 $f_N(a, b)$  は平均と分散が  $a, b$  である正規分布の密度関数を表す。また、(146) 式 の事前分布においては、

$$h(\theta | \mu) = f_N(\mu, \Phi) \quad (190)$$

とした場合に相当する。なお、通常は、原点と単位の不定性を除くため、

$$\mu = \mathbf{0}, \quad \Phi = I \quad (191)$$

とおいている。

従って、(143) 式 もしくは (187) 式 が

$$f(\mathbf{y}|\xi) = \int_{-\infty}^{+\infty} \left( \prod_{j=1}^p f_j(y_j | \theta) \right) h(\theta | \mu) d\theta \quad (192)$$

$$= \int_{-\infty}^{+\infty} \left( \prod_{j=1}^p f_N(m_j + \lambda'_j \theta, \psi_j) \right) f_N(\mu, \Phi) d\theta \quad (193)$$

$$= \int_{-\infty}^{+\infty} f_N(\mathbf{m} + \Lambda \theta, \Psi) f_N(\mu, \Phi) d\theta \quad (194)$$

$$= f_N(\mathbf{m} + \Lambda \mu, \Lambda \Phi \Lambda' + \Psi) \quad (195)$$

という形になる。

## 7.2 ロジスティック回帰分析との比較

因子分析は、観測される連続な基準変数群  $Y$  を、連続であるが観測されない説明変数群  $X$  に線形に回帰する方法であると考えられる。これと同様に、項目反応理論は、観測されるカテゴリカルな基準変数群  $U$  を、連続であるが観測されない説明変数  $\theta$  に回帰する方法（ロジスティック回帰分析）であると考えられる。2値データの場合は通常のロジスティック回帰分析が、多値データの場合にはプロポーショナルオッズモデルが相当する。

## 8 項目パラメタの等化

### 8.1 原点と単位の不定性

項目特性曲線を

$$P_j(\theta) = \frac{1}{1 + \exp(-1.7a_j(\theta - b_j))} \quad (196)$$

とする。ここで  $\theta$  を線形変換したものを

$$\theta^* = u_* + v_* \theta \quad (197)$$

と定義する。従って、

$$\theta = \frac{1}{v_*} \theta^* - \frac{u_*}{v_*} \quad (198)$$

であるが、これを項目特性曲線に代入すると

$$P_j(\theta) = \frac{1}{1 + \exp(-1.7a_j((\frac{1}{v_*}\theta^* - \frac{u_*}{v_*}) - b_j))} \quad (199)$$

$$= \frac{1}{1 + \exp(-1.7a_j \frac{1}{v_*} ((\theta^* - u_*) - v_* b_j))} \quad (200)$$

$$= \frac{1}{1 + \exp(-1.7a_j \frac{1}{v_*} (\theta^* - (u_* + v_* b_j)))} \quad (201)$$

$$= \frac{1}{1 + \exp(-1.7a_j^*(\theta^* - b_j^*))} \quad (202)$$

ただし

$$a_j^* = \frac{1}{v_*} a_j \quad \text{and} \quad b_j^* = u_* + v_* b_j \quad (203)$$

となる。

このことは、 $\theta$  を線形変換した場合、困難度パラメタを同じ線形変換し、識別力パラメタを  $v_*$  で割った新しいパラメタのセットを用いても、項目反応曲線が書き表せることを示している。（ $c$  パラメタは不変）これを  $\theta$  の原点と単位に関する不定性と呼ぶ。なお、同様の不定性が、因子分析モデルや一対比較法に関しても存在する。

通常はこの不定性をさけるため、パラメタの推定に用いられた集団に於ける  $\theta$  の平均を 0, 分散を 1 と制約するが、この制約は尤度の大きさに影響を与えるものではない。

ちなみに  $n_g, g = 1, 2, \dots, G, n$  で構成される  $G$  群があり、それぞれの群に於ける特性値の平均と分散が  $\bar{\theta}_g, \sigma_{\theta_g}^2$  の場合、全群をまとめた時の特性値の平均と分散は

$$\bar{\theta} = \frac{1}{\sum_{g=1}^G n_g} \sum_{g=1}^G n_g \bar{\theta}_g \quad (204)$$

$$\text{var}(\theta) = \frac{1}{\sum_{g=1}^G n_g} \sum_{g=1}^G n_g (\sigma_{\theta_g}^2 + \bar{\theta}_g^2) - \bar{\theta}^2 \quad (205)$$

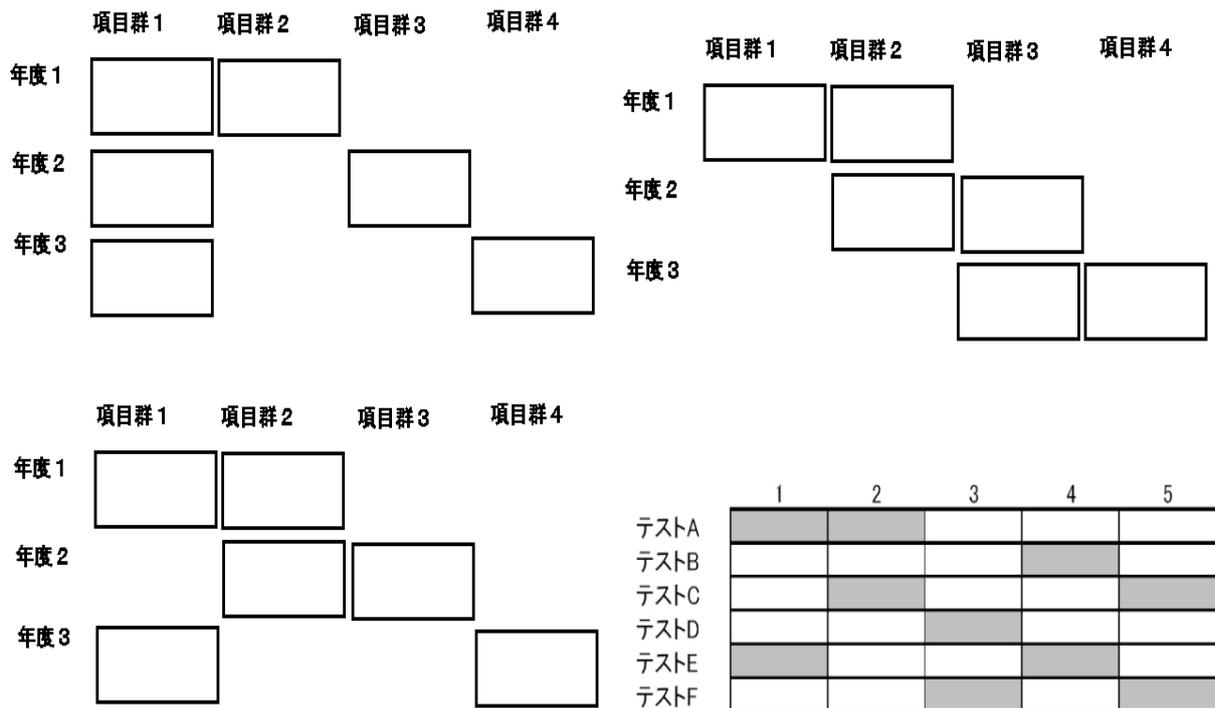


図 10: 項目パラメタの等化デザイン

となる。

原点と単位の不定性は、異なる集団を用いて独立に推定された項目パラメタや特性値の値は、そのままの形では比較可能でないことを示している。

## 8.2 共通項目を用いた項目パラメタの等化

独立に推定された項目パラメタを比較可能にする作業を項目パラメタの等化と呼ぶ。異なるテスト・フォームに含まれる項目パラメタの等化が可能となる条件としては

1. 異なるフォームに共通に含まれる項目（共通項目・アンカー項目）が存在する。
2. 異なるフォームを同時に受験している受験生が存在する。
3. 上記 2 項が同時に成り立っている。

の 3 つが考えられるが、本節では、項目バンクの構築の際に良く用いられる共通項目を利用したデザインに基づく等化の方法に関して説明する。

等化のデザインとしては 図 10 に示したようなものがあるが、左上のデザインでは項目群 1 が 3 フォーム全てに於いて共通となっている。右上の図では項目群 2 がフォーム 1 と 2 の間で、また、項目群 3 がフォーム 2 と 3 の間で共通となっている。下段の図においても各フォームは少なくとも他の一つのフォームと共通項目群を持っている。

このようなデザインの下では、各フォームごとにパラメタを推定すると、それぞれのフォームを受験した集団の能力が異なる場合には、推定された項目パラメタや特性値は比較出来ない。従って、各フォームから独立にパラメタを推定した場合には、以下に述べる等化の作業が必要となる。また、これら 3 フォームをまとめて、欠損値を含む大きなデータセットとして取り扱い、3 フォームに含まれる全ての項目パラメタを推定することも可能である。その場合、各フォームを受験した集団の能力が異なると考えられ

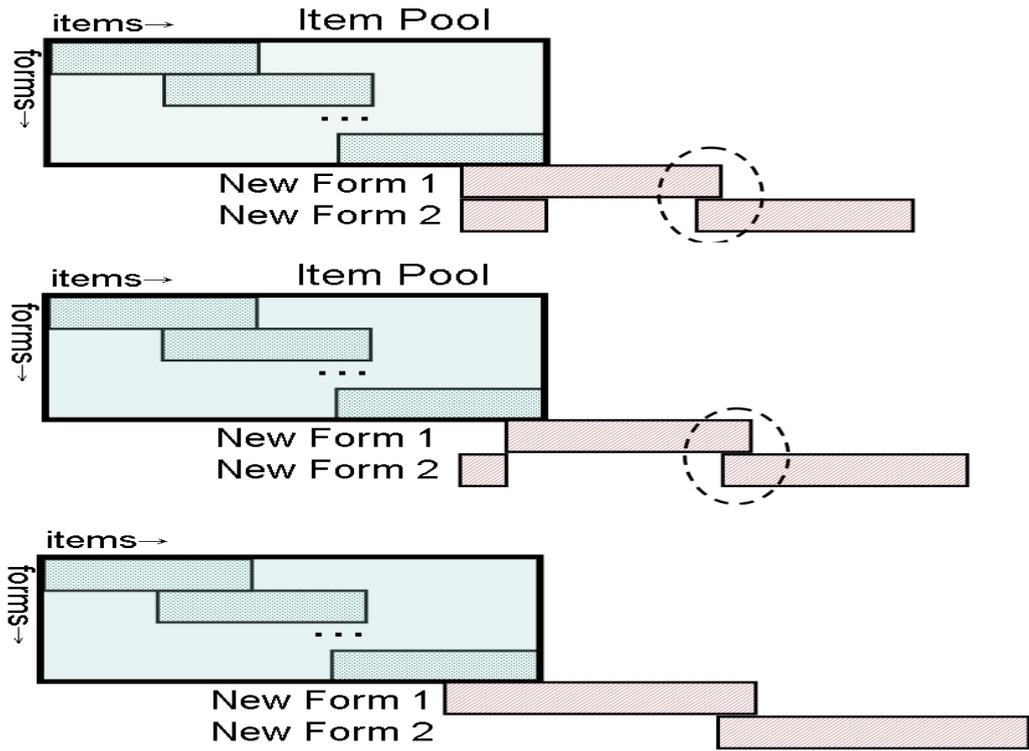


図 11: 目バンクを利用した項目パラメタの等化デザイン

るばあいには、所謂、多群の分析 (multi-group analysis) が必要となり、集団ごとの特性値の分布のパラメタ  $\mu$  を推定する必要が生じるが、それさえ行えば、比較可能な項目パラメタが得られる。これを同時推定法と呼ぶ。

図 11 には、項目バンクが存在する場合に、そこに新たに項目を追加するためのデザインを示している。一番上のデザインは、フォーム間に共通項目 (丸で囲まれた部分) があり、また、項目バンクとも共通の項目を共有している。真ん中のデザインは、上のデザインの中で、項目バンクとの共通項目を独自に設定した場合である。下段のデザインは、フォーム間に共通項目は存在するが、フォーム 2 は項目バンクと共通項目を持たない場合である。なお、上段と中段のデザインに於いては、丸で囲まれた部分、すなわちフォーム間の共通項目は存在しなくても、等化は可能である。

このような、項目バンクへの等化の際には、新しいフォームに含まれる項目パラメタを、項目バンクの尺度へ変換する必要が生じるが、先に説明した、フォームごとに独立に項目パラメタを推定し、その後、項目バンクの尺度へ変換する方法 (分離等化法) や、項目バンクに含まれる項目も込みで同時推定法を行うことも出来る。

以下に 図 10 の様なデザインを念頭に、共通項目を用いてテスト・フォームごとに独立に推定された項目パラメタの等化方法に関して説明する。

### 8.3 共通尺度と個別尺度との関係

各テスト・フォームは異なる集団に実施されているとし、集団 = テスト・フォームの総数を  $G$ 、項目の総数を  $p$  とし、

$$\hat{\xi}_j^{(g)} = \{\hat{a}_j^{(g)}, \hat{b}_j^{(g)}\}, j = 1, 2, \dots, p, g = 1, 2, \dots, G \quad (206)$$

を、第  $g$  群のデータから推定された  $j$  番目の項目のパラメタの推定値であるとする。ただし、項目  $j$  が第  $g$  群が受けたテストに含まれていない場合、これらは観測されていない。なお、これらの計算の際

しては、各群における  $\theta$  の平均と分散をそれぞれ 0 と 1 に設定してあるとする。また、これの期待値 (に相当するもの) を

$$\xi_j^{(g)} = \{a_j^{(g)}, b_j^{(g)}\}, j = 1, 2, \dots, p, g = 1, 2, \dots, G \quad (207)$$

とする。項目反応理論では、 $\theta$  はその原点と単位をのぞいて一意に定まる。従って、各群ごとの平均と分散に制約をおいた解の尺度を  $\theta^{(g)}$  と呼べば、それらはもとの尺度と以下のような線形の関係を持つ。

$$\theta^{(g)} = u_g + v_g \theta \quad (208)$$

ただし、 $u_g$  と  $v_g > 0$  は各尺度の原点と単位を定めるための定数である。また、

$$\xi_j = \{a_j, b_j\}, j = 1, 2, \dots, p \quad (209)$$

を、この共通尺度上でのパラメタの値とすれば、 $\theta^{(g)}$  上で定義される項目パラメタ  $\xi^{(g)}$  は、 $\xi$  と以下のような関係にある。

$$a_j^{(g)} = \frac{1}{v_g} a_j \quad \text{and} \quad b_j^{(g)} = u_g + v_g b_j, j = 1, 2, \dots, p, \quad (210)$$

なお、次節では、この共通尺度から各群の尺度への項目パラメタの変換をまとめて

$$\xi_j^{(g)} = B_g(\xi_j) \quad (211)$$

と表すことにする。

逆に、

$$q_g = -\frac{u_g}{v_g}, r_g = \frac{1}{v_g} \quad (212)$$

とにおいて、

$$\theta = q_g + r_g \theta^{(g)} \quad (213)$$

$$a_j = v_g a_j^{(g)}, \quad \text{and} \quad b_j = q_g + r_g b_j^{(g)} \quad (214)$$

である。従って、二つのフォーム  $g$  と  $h$  で独立に推定された項目パラメタの間には

$$a_j : v_g a_j^{(g)} = v_h a_j^{(h)} \quad (215)$$

$$b_j : q_g + r_g b_j^{(g)} = q_h + r_h b_j^{(h)} \quad (216)$$

の関係から

$$a_j^{(g)} = \frac{r_g}{r_h} a_j^{(h)} \quad (217)$$

$$b_j^{(g)} = \frac{1}{r_g} (q_h - q_g) + \frac{r_h}{r_g} b_j^{(h)} \quad (218)$$

という関係がある。すなわち、二つのフォームから独立に項目パラメタを推定し、二つのフォームに共通に含まれる項目の  $a$  パラメタを取り出してプロットすれば、それは原点を通る直線上に載るはずであり、 $b$  パラメタのプロットは、直線上に載るはずである。また、これらの直線の傾きや切片から、等化係数を求めることも出来る。しかし、実際にこれらのプロットを作ってみると、共通項目は必ずしも直線上には散布しない。このプロットが直線に近いほどデータに項目反応理論が適合していると言える。

## 8.4 共通尺度への等化法

### 8.4.1 パラメタを用いた最小2乗解

データ  $\hat{\xi}_j^{(g)}$  とモデル  $\xi_j^{(g)}$  の間で

$$\hat{\xi}_j^{(g)} \approx \xi_j^{(g)} = B_g(\xi_j), j = 1, 2, \dots, p \quad (219)$$

の近似がなるべくよくなるように係数  $u_g, v_g, g = 1, 2, \dots, G$  と  $\theta$ —尺度上のパラメタの値  $\xi_j, j = 1, 2, \dots, p$  を求める。具体的には、 $w_{Aj}, w_{Bj}$  を事前に付与した a パラメタと b パラメタの重要度を表す重みとして、

$$RSS_p = \sum_{g=1}^G \sum_{j \in g} \left[ w_{Aj} (\hat{a}_j^{(g)} - \frac{1}{v_g} a_j)^2 + w_{Bj} (\hat{b}_j^{(g)} - (u_g + v_g b_j))^2 \right] \quad (220)$$

を最小とするような解を求めればよい。これは、基本的には主成分分析である。

### 8.4.2 反応確率を用いた最小2乗解

データ  $\hat{\xi}_j^{(g)}$  とモデル  $\xi_j^{(g)}$  の間でそれぞれから計算される反応確率がうまく一致するように、すなわち

$$P(\theta | \hat{\xi}_j^{(g)}) \approx P(\theta | \xi_j^{(g)}) = P(\theta | B_g(\xi_j)), j = 1, 2, \dots, p, \quad (221)$$

の近似がなるべくよくなるように係数  $u_g, v_g, g = 1, 2, \dots, G$  と  $\theta$ —尺度上のパラメタの値  $\xi_j, j = 1, 2, \dots, p$  を求める。ただし、項目パラメタ  $\xi$  を持つ項目特性曲線を

$$P(\theta | \hat{\xi}) = \frac{1}{(1 + \exp(-1.7a_j(\theta - b_j)))} \quad (222)$$

と書く。

いま、

$$P(\theta | \xi_j^{(g)}) = \frac{1}{(1 + \exp(-1.7a_j^{(g)}(\theta - b_j^{(g)})))} \quad (223)$$

$$= \frac{1}{(1 + \exp(-1.7(\frac{1}{v_g})a_j(\theta - (u_g + v_g b_j))))} \quad (224)$$

$$= P\left(-\frac{u_g}{v_g} + \frac{1}{v_g}\theta | \xi_j\right) \quad (225)$$

$$= P(q_g + r_g\theta | \xi_j) \quad (226)$$

$$(227)$$

であるから、

$$P_k(\theta | \xi_j^{(g)}) = P_k(q_g + r_g\theta | \xi_j) \quad (228)$$

である。同様に

$$P(u_g + v_g\theta | \xi_j^{(g)}) = P(\theta | \xi_j) \quad (229)$$

である。

これらのいずれかの関係を用いれば以下のような最小2乗基準を考えることができる。

$$RSS_1 = \sum_{g=1}^G \sum_{j \in g} w_{gj} \int_{\Theta_g} \left( P(\theta | \hat{\xi}_j^{(g)}) - P(q_g + r_g\theta | \xi_j) \right)^2 h_g(\theta) d\theta \quad (230)$$

ただし、 $h_g(\theta)$  は各集団における  $\theta$  の分布の密度関数（既知）で、パラメタの推定時の制約条件どおり、その平均は 0 分散は 1 となるものであり、通常は正規分布が用いられる。（これは、項目パラメタ推定時の  $\theta$  の周辺分布にあたるものである。）また、 $\Theta_g$  はその定義域である。なお、実際の計算に際しては、積分を適当な範囲の和で置き換える。 $w_{gjk}$  は事前に定めた重みである。詳細は前川 (1991) を参照のこと。

この最小 2 乗基準の利点は、 $P(\theta | \hat{\xi}_j^{(g)})$  が未知パラメタを含まないため、通常最小 2 乗法を用いることが出来る点である。やり方としては、交互最小 2 乗法を用い、

- $q_g, r_g, g = 1, 2, \dots, G$  が所与として  $\xi_j, j = 1, 2, \dots, p$  を求める。
- $\Xi$  が所与で  $q_g, r_g, g = 1, 2, \dots, G$  を求める。

を繰り返す。この操作により毎回  $RSS_1$  を減少させることが可能となる。

他方、

$$RSS_2 = \sum_{g=1}^G \sum_{j \in g} w_{gj} \int_{\Theta} \left( P(u_g + v_g \theta | \hat{\xi}_j^{(g)}) - P_k(\theta | \xi_j) \right)^2 h(\theta) d\theta \quad (231)$$

を最小 2 乗基準として用いることも可能である。ただし、 $h(\theta)$  は全集団における  $\theta$  の分布の密度関数（既知）で、通常は、共通尺度の平均は 0 分散は 1 とする。また、 $\Theta$  はその定義域である。なお、実際の計算に際しては、積分を適当な範囲の和で置き換える。

この二つの最小 2 乗基準の利点は、 $h(\theta)$  をうまく用いることにより、共通尺度上で重要な部分の精度をコントロールする事が出来ることである。

いずれにしても、必要となるのは  $P(q_g + r_g \theta | \xi_j)$  のパラメタ  $q, r, \xi$  に関する 1 次微分、もしくは  $P(u_g + v_g \theta | \xi_j^{(g)})$  のパラメタ  $u, v, \xi$  に関する 1 次微分だけである。全て最小 2 乗解を求める問題であるから Gauss-Newton 法を用いればよく 2 次微分は不要である。

項目反応理論の基礎となるのは、各項目への反応確率であるため、パラメタの値を用いて最小 2 乗基準を設定した方法よりも、反応確率を用いたものの方が、理論的に優れていると考えられる。 $RSS_1$  と  $RSS_2$  の比較では、 $h_g(\theta)$  は実在するため、前者の最小化が適当であると考えている。

## 参考文献

### テスト一般に関する参考書

日本テスト学会 (2007) AERA, APA & NCME (編) テスト・スタンダードー日本のテストの将来に向けて 金子書房

池田央・藤田恵壘・柳井晴夫・繁樹算男 (監訳) (1992) R.L. Linn (編) 教育測定学 原著第 3 版 みに出版 (項目反応理論の章と、等化の章を含む。)

AERA, APA, & NCME (1999) *The Standards for Educational and Psychological Testing*. American Psychological Association

### 等化法一般に関する参考書

Braun, H.I. & Holland, P.W. (1982) Observed-score test equating: A mathematical analysis of some ETS equating procedures. In Holland, P.W. & Rubin, D.B. Eds. *Test Equating*. Academic Press, Inc. New York.

Kolen M.J. & Brennan, R.L. (1995) *Test Equating. Methods and Practices*. New York, N.Y.: Springer-Verlag New York, Inc.

Kolen M.J. & Brennan, R.L. (2004) *Test Equating, Scaling, and Linking. Methods and Practices* New York, N.Y.: Springer-Verlag New York, Inc.

#### 古典的テスト理論を用いた等化法の参考書

芝祐順・渡部洋 (1998) 入試データの解析 新曜社

前川眞一 (1999) 得点調整の方法について

柳井・前川 (編) (1999) 大学入試データの解析 [理論と応用] 第 II-3 章 p88-109 現代数学社

Angoff, W. H. (1971) Scales, norms, and equivalent scores. In R.L. Thorndike(Ed.) *Educational Measurement* (2nd ed). Washington, D.C., American Council on Education, 508-600.

Angoff, W. H. (1984) *Scales, norms, and equivalent scores*. Princeton, NJ: Educational Testing Service. (Reprint of the one in Thorndike above)

Gulliksen, H. (1950). *Theory of Mental Tests*. New York: Wiley.

#### 項目反応理論の参考書

大友賢二 (1996) 項目応答理論入門 大修館書店

芝祐順 (1991) 項目反応理論. 東大出版会

高橋正視 (2002) 項目反応理論入門～新しい絶対評価～ イデア出版局

豊田秀樹 (2002) 項目反応理論—テストと測定の科学【入門編】，【事例編】 朝倉書店

渡辺直登・野口裕之 (1999) 組織心理測定論 — 項目反応理論のフロンティア 白桃書房

#### 複合二項分布や多項分布の確率の計算に関して

Mayekawa, S. & Arai, S. (2008) Distribution of the Sum of Scored Multinomial Random Variables and Its Application to the Item Response Theory. In K. Shigemasu, A. Okada, T. Imaizumi, & T. Hoshino (Eds.) *New Trends in Psychometrics*. Tokyo: University Academic Press.

# 付録 A

## 最尤法について

### 1 最尤解

$Y_i, i = 1, 2, \dots, n$  を独立な観測値とする。 $f(y_i|\theta)$  をデータの確率関数（密度関数）とする。データの同時分布の確率関数は

$$f(y_1, y_2, \dots, y_n|\theta) = \prod_{i=1}^n f(y_i|\theta) \quad (232)$$

である。これは、母数の値が  $\theta$  の時に、 $y_1, y_2, \dots, y_n$  というデータを観測する確率（密度）であり、その値が高いほど、それらのデータ系列（パターン）が出現する確率が高い。

この同時確率関数を母数  $\theta$  の関数と見たものが尤度関数である。

$$L(\theta) \propto f(y_1, y_2, \dots, y_n|\theta) \quad (233)$$

これは、データ系列を与えられた時に、母数の値がどの程度尤もらしいかを表す値と解釈することが出来る。ただし、 $\theta$  の確率を与えるものではない。（ $\theta$  は確率変数ではないので確率は付与できない。）

尤度関数の対数を対数尤度関数という。

$$\ln L(\theta) = \log L(\theta) = \sum_{i=1}^n \log f(y_i|\theta) \quad (234)$$

尤度関数もしくは対数尤度関数の最大値を与える母数  $\theta$  の値、 $\hat{\theta}$ 、をその母数の推定値とする方法を、最尤解と呼ぶ。最尤解を求めるためには対数尤度関数の微分を 0 とおいた方程式

$$d \ln L(\theta) = \frac{d \ln L(\theta)}{d \theta} = 0 \quad (235)$$

を解けばよい。

#### 1.1 正規分布の場合

$$f(y_1, y_2, \dots, y_n|\theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y_i - \theta)^2\right) \quad (236)$$

$$L(\theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y_i - \theta)^2\right) \quad (237)$$

$$\ln L(\theta) = \sum_{i=1}^n \left[-\frac{1}{2}(y_i - \theta)^2 + \text{const}\right] \quad (238)$$

$$d \ln L(\theta) = \sum_{i=1}^n (y_i - \theta) = 0 \quad (239)$$

例

$\mathbf{y} = \{1, 2, 3, 4, 5\}$  の場合

THETA	L	LL	DLL
2.5	0.000019	-10.8236	2.5000
2.6	0.000025	-10.5986	2.0000
2.7	0.000029	-10.4236	1.5000
2.8	0.000033	-10.2986	1.0000
2.9	0.000036	-10.2236	0.5000
3	0.000037	-10.1986	0.0000
3.1	0.000036	-10.2236	0.5000
3.2	0.000033	-10.2986	1.0000
3.3	0.000029	-10.4236	1.5000
3.4	0.000025	-10.5986	2.0000
3.5	0.000019	-10.8236	2.5000

## 1.2 ベルヌーイ分布の場合

$$f(y_1, y_2, \dots, y_n | \theta) = \prod_{i=1}^n \theta^{y_i} (1 - \theta)^{1 - y_i} \quad (240)$$

$$L(\theta) = \prod_{i=1}^n \theta^{y_i} (1 - \theta)^{1 - y_i} \quad (241)$$

$$\ln L(\theta) = \sum_{i=1}^n (y_i \log \theta + (1 - y_i) \log(1 - \theta)) \quad (242)$$

$$d \ln L(\theta) = \sum_{i=1}^n \left( \frac{y_i}{\theta} - \frac{(1 - y_i)}{(1 - \theta)} \right) = 0 \quad (243)$$

例

$\mathbf{y} = \{0, 1, 1, 0, 1\}$  の場合

THETA	L	LL	DLL
0.1	0.0008	-7.1185	27.7778
0.2	0.0051	-5.2746	12.5000
0.3	0.0132	-4.3253	7.1429
0.4	0.0230	-3.7705	4.1667
0.5	0.0313	-3.4657	2.0000
0.6	0.0346	-3.3651	-0.0000
0.7	0.0309	-3.4780	-2.3810
0.8	0.0205	-3.8883	-6.2500
0.9	0.0073	-4.9213	-16.6667

## 2 回帰分析

### 2.1 正規分布の場合：回帰分析

データ

$$(x_i, y_i), i = 1, 2, \dots, n \quad \text{ただし } y_i \text{ は連続量} \quad (244)$$

モデル

$$y_i = \alpha x_i + \beta + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2) \quad (245)$$

または

$$\mathcal{E}(y_i|x_i) = \alpha x_i + \beta \quad (246)$$

モデルのパラメタは  $\xi = \{\alpha, \beta, (\sigma^2)\}$

(教室では  $\theta$  としていたが、混乱を避けるため  $\xi$  に変更)

尤度関数

$$L(\xi) = \prod_{i=1}^n \frac{1}{\sqrt{2\sigma^2}} \exp\left(-\frac{1}{2}(x_i - (\alpha x_i + \beta))^2/\sigma^2\right) \quad (247)$$

対数尤度関数

$$\ln L(\xi) = -n\sigma^2 - \frac{1}{2} \sum_{i=1}^n (x_i - (\alpha x_i + \beta))^2 \quad (248)$$

これを最大にするように  $\xi$  を求める。すなわち、連立方程式

$$\frac{\partial \ln L}{\partial \alpha} = 0, \quad \frac{\partial \ln L}{\partial \beta} = 0, \quad \frac{\partial \ln L}{\partial \sigma} = 0 \quad (249)$$

を解く。答えは

$$\hat{\alpha} = \frac{\text{cov}(x, y)}{\text{var}(x)}, \quad \hat{\beta} = \bar{y} - \hat{\alpha}\bar{x}, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - (\hat{\alpha}x_i + \hat{\beta}))^2 \quad (250)$$

### 2.2 2項分布の場合：ロジスティック回帰分析

データ

$$(x_i, y_i), i = 1, 2, \dots, n \quad \text{ただし } y_i \text{ は } 0 \text{ か } 1 \quad (251)$$

これを  $x_i$  の異なる値ごとに度数分布の形にしたもの

$$(x_i, f_i, r_i), i = 1, 2, \dots, N \quad \text{ただし } f_i \text{ は試行数 } r_i \text{ は成功数} \quad (252)$$

モデル

$$\text{Prob}(y_i = 1|x_i) = \frac{1}{1 + \exp(-(\alpha x_i + \beta))} = P(x_i|\alpha, \beta) \quad (253)$$

または

$$\mathcal{E}(y_i|x_i) = P(x_i|\alpha, \beta) \quad (254)$$

$$\mathcal{E}(r_i|x_i) = f_i \times P(x_i|\alpha, \beta) \quad (255)$$

モデルのパラメタは  $\xi = \{\alpha, \beta\}$

尤度関数

$$L(\xi) = \prod_{i=1}^n P(x_i|\alpha, \beta)^{y_i} (1 - P(x_i|\alpha, \beta))^{1-y_i} \quad (256)$$

$$L(\xi) = \prod_{i=1}^n P(x_i|\alpha, \beta)^{r_i} (1 - P(x_i|\alpha, \beta))^{f_i-r_i} \quad (257)$$

対数尤度関数

$$\ln L(\xi) = \sum_{i=1}^n y_i \log P(x_i|\alpha, \beta) + (1 - y_i) \log(1 - P(x_i|\alpha, \beta)) \quad (258)$$

$$\ln L(\xi) = \sum_{i=1}^n r_i \log P(x_i|\alpha, \beta) + (f_i - r_i) \log(1 - P(x_i|\alpha, \beta)) \quad (259)$$

これを最大にするように  $x_i$  を求める。すなわち、連立方程式

$$\frac{\partial \ln L}{\partial \alpha} = - \sum_{i=1}^n (r_i - f_i P(x_i|\alpha, \beta))(x_i - \beta) = 0 \quad (260)$$

$$\frac{\partial \ln L}{\partial \beta} = \alpha \sum_{i=1}^n (r_i - f_i P(x_i|\alpha, \beta)) = 0 \quad (261)$$

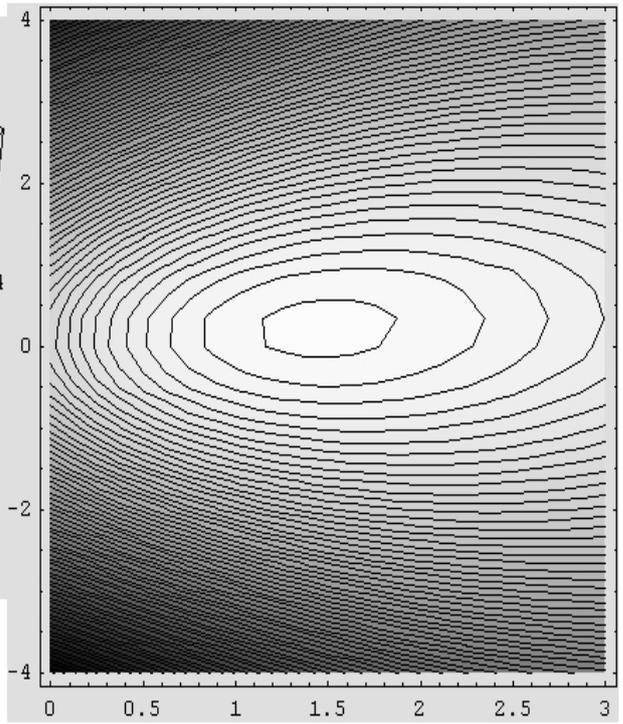
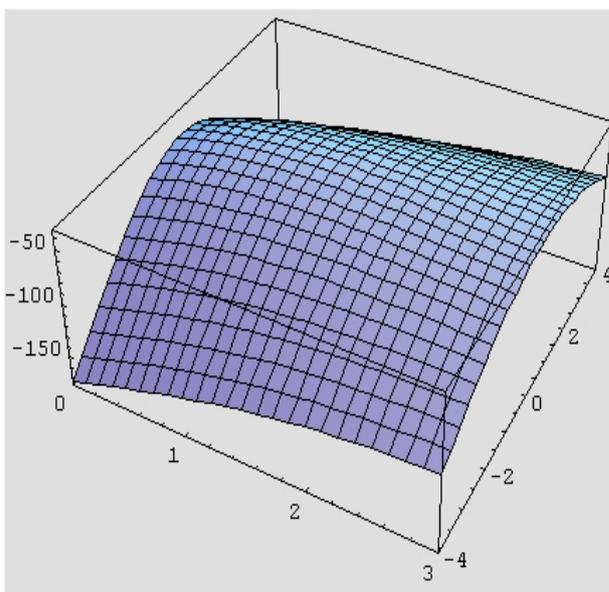
を解けばよいが、解析的には解けないため、数値解を求める。

IRT のパラメタとの関係

指数関数の中が、IRT:  $-1.7a(x - b)$ , LR:  $-(\alpha x + \beta)$  であるから、 $a = \alpha/1.7$ ,  $b = -\beta/\alpha$  となる。

例

Data1				Data2			
Obs	X	F	R	Obs	X	F	R
1	-3.0	1	0	1	-3.0	1	0
2	-2.6	2	0	2	-2.6	2	0
3	-2.3	3	0	3	-2.3	3	0
4	-1.9	4	0	4	-1.9	4	0
5	-1.5	5	1	5	-1.5	5	0
6	-1.1	6	1	6	-1.1	6	0
7	-0.8	7	2	7	-0.8	7	0
8	-0.4	8	1	8	-0.4	8	0
9	0.0	9	7	9	0.0	9	2
10	0.4	8	7	10	0.4	8	2
11	0.8	7	5	11	0.8	7	1
12	1.1	6	6	12	1.1	6	5
13	1.5	5	4	13	1.5	5	4
14	1.9	4	4	14	1.9	4	4
15	2.3	3	2	15	2.3	3	2
16	2.6	2	2	16	2.6	2	2
17	3.0	1	1	17	3.0	1	1
$\beta$	0.2127	(-0.14)		$\beta$	-2.0829	(0.95)	
$\alpha$	1.4642	(0.861)		$\alpha$	2.1759	(1.28)	



## 付録 C

### BILOG-MG の使い方

#### 1 コマンドファイル

BILOG-MG はコマンドファイル（バッチファイル）名を引数に渡して実行する。コマンドファイル名の拡張子は .BLM とする。たとえば test.blm にかかれたコマンドを実行したい場合には DOS 窓から

```
pathname\bilogmg test
```

とする。これで通常は、少なくとも test.ph1, test.ph2, test.ph3 というファイルが出来る。phase1 は古典的項目統計量とマスターファイルの作成、 phase2 は周辺最尤解による項目パラメタ及び特性値の母集団分布のパラメタの推定、 phase3 は個人の特性値の推定に対応している。以下の記述では CALIB コマンドが phase2 の指定を、 SCORE コマンドが phase3 の指定を行っている。

BILOG-MG 3 は Windows の GUI を用いたインターフェースプログラムが用意されている。

Command File

2行分のタイトル

```
>COMMENT comment
      comment
>GLOBAL DFNAME=' データファイル名', NPARM=モデルのパラメタ数, SAVE;
>SAVE SCORE=' 特性値出力ファイル名.sco', PARM=' パラメタ出力ファイル名.par',
      ISTAT=' 項目統計量出力ファイル名.ist';
>LENGTH NITEMS=(項目総数);
>INPUT NTOTAL=総項目数, SAMPLE=利用観測値数, NGROUP=集団数, NIDCH=受験 ID 長,
      KFNAME=' 正答反応識別値ファイル名',
      NFNAME=' 欠測反応識別値ファイル名',
      NALT=選択枝数, NFORM=版数, TYPE=データタイプ ( 1 );
>ITEMS INUMBERS=項目番号のリスト ( (1(1)ntotal) ),
      INAME=項目名のリスト ( (Q01(1)Qntotal) );
>TEST TNAME=テスト名, INUM=項目番号のリスト;
>FORM1 LENGTH=版 1 に含まれる項目数, INUM=項目番号のリスト (データの順);
>FORM2 LENGTH=版 2 に含まれる項目数, INUM=項目番号のリスト (データの順);
>GROUP1 GNAME=集団名 1, LENGTH=集団 1 が受ける最大項目数, INUM=項目番号のリスト ;
>GROUP2 GNAME=集団名 2, LENGTH=集団 2 が受ける最大項目数, INUM=項目番号のリスト;
(8A1, T11, I1, T21, 40A1) ID, 集団番号, 項目反応
>CALIB CYCLE=EM 繰り返し数, NEWTON=NR 繰り返し数, CRIT=0.001, PLOT=0.05,
      normal, 母集団分布の形 ( normal or empirical )
      idist=0, NQPT=求積点数,
      REF=平均と分散を固定する集団番号,
      tprior, sprior, gprior, nofloat 項目パラメタの事前分布の指定
      ;
>SCORE IDIST=3, METHOD=2, NOPRINT;
```

## FORM と GROUP コマンドについて

グループが複数ある場合には GROUP コマンドを複数書く。一つの場合は書かない  
版が複数ある場合には FORM コマンドを複数書く。一つの場合は書かない

## FORMAT について

データファイルのレコード長は 256 byte 以下。

データファイルの各項目の順序は、FORM コマンドが一つの場合には

受験者 ID      集団番号      項目反応

であり、

複数の FORM コマンドがある場合には

受験者 ID      版番号      集団番号      項目反応

となる。

データファイルに様々な key（正答反応や欠測等の識別値）が含まれる場合

版1の正答識別値

版1の欠測識別値

      :

版2の正答識別値

版2の欠測識別値

      :

の順でなければいけない。そうでない場合は版の順に並べる。

## Extended Format と Compressed Format（参照：P110）

Extended Format とは、すべての集団があたかもすべての項目を受験したかのような  
データのことであり、受験されなかった項目は not presented としてコード化されている。

これに対して、Compressed Format では、各集団は異なる版を受験し、  
各版が異なる項目を含む形となる。その場合には、複数の FORM コマンドが必要となる。

## 出力

各 phase の結果は、それぞれ .ph1, .ph2, .ph3 という拡張子が付いたファイルとして出力される。  
また、古典的項目統計量、項目パラメタや特性値  $\theta$  の値は、それぞれ SAVE コマンド  
で指定した .ist, .par ならびに .sco というファイルに出力される。

## 2 例

### 2.1 単一フォーム、1 集団の場合 (test1L.dat)

データ

G1_1	-2.497	1	1	111011000001000	13
G1_2	-2.386	1	1	001010000100100	9
G1_3	-2.337	1	1	110111000000010	11
G1_4	-2.314	1	1	010100101000001	8
G1_5	-2.280	1	1	110001011000000	10
G1_6	-2.256	1	1	110000010100000	10
G1_7	-2.255	1	1	011010001101010	13
G1_8	-2.213	1	1	011000001100000	7
G1_9	-2.186	1	1	100100001011000	8
G1_10	-2.179	1	1	111101110101000	13

コマンドファイル

```
test1L.blm
One Test: Low Ability Group
>COMMENT
Simple Test
061117
>GLOBAL DFName = 'test1L.DAT',
      NPArm = 2,
      SAve;
>SAVE PARM = 'test1L.PAR',
      SCOrE = 'test1L.SCO',
      ISTat = 'test1L.list';
>LENGTH NITems = (15);
>INPUT NTOTal = 15,
      SAMple = 99999,
      NALt = 5,
      NIDchar = 8,
      NFName = 'miss.dat';
>ITEMS INames = (Q01(1)Q15);
>TEST1 TName = 'Simple',
      INumber = (1(1)15);
(8A1,T25,60A1)
>CALIB ACCel = 1.0000,
      NORmal;
>SCORE NQPt = (31),
      NOPrint;
```

結果 test1L.ph2 の部分

-2 LOG LIKELIHOOD: 1673.3895

CYCLE 6; LARGEST CHANGE= 0.00428

1

SUBTEST SIMPLE ; ITEM PARAMETERS AFTER CYCLE 6

ITEM	INTERCEPT	SLOPE	THRESHOLD	LOADING	ASYMPTOTE
Q01	1.161	0.498	-2.333	0.446	0.000
	0.224*	0.162*	0.769*	0.145*	0.000*
Q02	1.361	0.647	-2.104	0.543	0.000
	0.270*	0.216*	0.610*	0.182*	0.000*
Q03	0.482	0.363	-1.326	0.341	0.000
	0.161*	0.113*	0.584*	0.106*	0.000*
Q04	0.674	0.387	-1.739	0.361	0.000
	0.174*	0.124*	0.668*	0.116*	0.000*
Q05	0.487	0.387	-1.257	0.361	0.000
	0.169*	0.120*	0.569*	0.112*	0.000*
Q06	0.027	0.557	-0.048	0.486	0.000
	0.182*	0.183*	0.328*	0.160*	0.000*
Q07	0.104	0.421	-0.248	0.388	0.000
	0.159*	0.132*	0.382*	0.122*	0.000*
Q08	0.103	0.381	-0.271	0.356	0.000
	0.152*	0.115*	0.404*	0.108*	0.000*
Q10	-0.074	0.327	0.227	0.311	0.000
	0.153*	0.102*	0.470*	0.097*	0.000*
Q11	-0.709	0.385	1.842	0.359	0.000
	0.183*	0.124*	0.716*	0.116*	0.000*
Q12	-0.257	0.325	0.791	0.309	0.000
	0.149*	0.101*	0.520*	0.097*	0.000*
Q13	-0.621	0.425	1.460	0.391	0.000
	0.175*	0.135*	0.571*	0.124*	0.000*
Q14	-0.477	0.331	1.438	0.315	0.000
	0.155*	0.105*	0.642*	0.100*	0.000*
Q15	-1.238	0.633	1.956	0.535	0.000
	0.257*	0.217*	0.627*	0.183*	0.000*

\* STANDARD ERROR

LARGEST CHANGE = 0.004284

PARAMETER	MEAN	STN DEV
SLOPE	0.433	0.108
LOG(SLOPE)	-0.862	0.232
THRESHOLD	-0.115	1.475

QUADRATURE POINTS, POSTERIOR WEIGHTS, MEAN AND S.D.:

	1	2	3	4	5
POINT	-0.5071E+01	-0.4347E+01	-0.3623E+01	-0.2898E+01	-0.2174E+01
POSTERIOR	0.2253E-05	0.5137E-04	0.6933E-03	0.5802E-02	0.3014E-01
	6	7	8	9	10
POINT	-0.1450E+01	-0.7258E+00	-0.1595E-02	0.7226E+00	0.1447E+01
POSTERIOR	0.1004E+00	0.2166E+00	0.2869E+00	0.2225E+00	0.1022E+00
	11	12	13	14	15
POINT	0.2171E+01	0.2895E+01	0.3620E+01	0.4344E+01	0.5068E+01
POSTERIOR	0.2899E-01	0.5076E-02	0.5367E-03	0.3403E-04	0.0000E+00
MEAN	0.00000				
S.D.	1.00000				

項目統計量出力ファイル test1L.list

test1L.blm

One Test: Low Ability Group

ITEM STATISTICS FOR SUBTEST SIMPLE

ITEM	NAME	#TRIED	#RIGHT	PCT	ITEM*TEST CORRELATION		
					LOGIT/1.7	PEARSON	BISERIAL
1	Q01	100.0	85.0	85.0	-1.02	0.060	0.092
2	Q02	100.0	87.0	87.0	-1.12	0.179	0.284
3	Q03	100.0	68.0	68.0	-0.44	0.038	0.049
4	Q04	100.0	74.0	74.0	-0.62	0.038	0.051
5	Q05	100.0	68.0	68.0	-0.44	0.038	0.049
6	Q06	100.0	51.0	51.0	-0.02	0.151	0.190
7	Q07	100.0	54.0	54.0	-0.09	0.126	0.158
8	Q08	100.0	54.0	54.0	-0.09	0.023	0.028
9	Q09	100.0	50.0	50.0	0.00	-0.137	-0.172
10	Q10	100.0	47.0	47.0	0.07	0.016	0.020
11	Q11	100.0	25.0	25.0	0.65	0.009	0.013
12	Q12	100.0	40.0	40.0	0.24	-0.057	-0.073
13	Q13	100.0	28.0	28.0	0.56	0.062	0.082
14	Q14	100.0	32.0	32.0	0.44	-0.020	-0.026
15	Q15	100.0	15.0	15.0	1.02	0.164	0.252

パラメタ出力ファイル test1L.par

test1L.blm

One Test: Low Ability Group

1 15

15

Q01	SIMPLE	1.16127	0.22449	0.49769	0.16240	-2.33333	0.76880	0.44556	0.14539	0.00000	0.00000	0.00000	0.00000	110
Q02	SIMPLE	1.36062	0.26997	0.64674	0.21650	-2.10382	0.61017	0.54306	0.18179	0.00000	0.00000	0.00000	0.00000	210
Q03	SIMPLE	0.48161	0.16095	0.36322	0.11277	-1.32595	0.58352	0.34140	0.10600	0.00000	0.00000	0.00000	0.00000	310
Q04	SIMPLE	0.67376	0.17362	0.38749	0.12410	-1.73878	0.66844	0.36131	0.11571	0.00000	0.00000	0.00000	0.00000	410
Q05	SIMPLE	0.48684	0.16945	0.38736	0.12042	-1.25679	0.56943	0.36121	0.11229	0.00000	0.00000	0.00000	0.00000	510
Q06	SIMPLE	0.02674	0.18218	0.55663	0.18317	-0.04803	0.32834	0.48636	0.16004	0.00000	0.00000	0.00000	0.00000	610
Q07	SIMPLE	0.10437	0.15908	0.42146	0.13221	-0.24764	0.38163	0.38838	0.12183	0.00000	0.00000	0.00000	0.00000	710
Q08	SIMPLE	0.10313	0.15158	0.38086	0.11542	-0.27077	0.40449	0.35592	0.10786	0.00000	0.00000	0.00000	0.00000	810
Q09	SIMPLE	0.00000	0.00000	0.00100	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	9 0
Q10	SIMPLE	-0.07409	0.15322	0.32709	0.10179	0.22653	0.46956	0.31088	0.09675	0.00000	0.00000	0.00000	0.00000	1010
Q11	SIMPLE	-0.70851	0.18261	0.38459	0.12398	1.84225	0.71619	0.35896	0.11572	0.00000	0.00000	0.00000	0.00000	1110
Q12	SIMPLE	-0.25685	0.14894	0.32453	0.10148	0.79147	0.52011	0.30868	0.09653	0.00000	0.00000	0.00000	0.00000	1210
Q13	SIMPLE	-0.62104	0.17502	0.42536	0.13522	1.46003	0.57064	0.39142	0.12443	0.00000	0.00000	0.00000	0.00000	1310
Q14	SIMPLE	-0.47671	0.15452	0.33142	0.10491	1.43841	0.64241	0.31459	0.09958	0.00000	0.00000	0.00000	0.00000	1410
Q15	SIMPLE	-1.23763	0.25712	0.63261	0.21654	1.95640	0.62742	0.53461	0.18300	0.00000	0.00000	0.00000	0.00000	1510

特性値出力ファイル test1L.sco

test1L.blm

One Test: Low Ability Group

1 G1\_1

1.00 SIMPLE 14 6 42.86 -0.307975 0.648266 0.000000 0.000528

1 G1\_2

1.00 SIMPLE 14 4 28.57 -1.220745 0.652933 0.000000 0.000017

1 G1\_3

1.00 SIMPLE 14 6 42.86 -0.285712 0.648317 0.000000 0.000496

1 G1\_4

1.00 SIMPLE 14 4 28.57 -0.799497 0.649001 0.000000 0.000018

1 G1\_5

1.00 SIMPLE 14 4 28.57 -0.804066 0.649026 0.000000 0.000362

1 G1\_6

1.00 SIMPLE 14 4 28.57 -0.968708 0.650171 0.000000 0.000430

1 G1\_7

1.00 SIMPLE 14 6 42.86 -0.590654 0.648237 0.000000 0.000037

1 G1\_8

1.00 SIMPLE 14 3 21.43 -1.341358 0.654713 0.000000 0.000313

別のデータ (test1H.dat)

```
G2_101  -0.543  1  1  111111111010111  19
G2_102  -0.539  1  1  110101011101011  18
G2_103  -0.536  1  1  111111111111101  22
G2_104  -0.535  1  1  111111111011101  22
G2_105  -0.535  1  1  111111111000100  16
G2_106  -0.526  1  1  1111111111111001  20
G2_107  -0.524  1  1  111110110001000  17
G2_108  -0.508  1  1  111111111111010  26
G2_109  -0.507  1  1  111111111101110  22
G2_110  -0.503  1  1  1111111110010111  28
```

結果 test1H.ph2

-2 LOG LIKELIHOOD: 908.0535

CYCLE 9; LARGEST CHANGE= 0.00980

1

```
SUBTEST SIMPLE ; ITEM PARAMETERS AFTER CYCLE 9
ITEM INTERCEPT SLOPE THRESHOLD LOADING ASYMPOTOTE
-----
Q02 | 3.165 | 0.755 | -4.191 | 0.603 | 0.000
    | 0.681* | 0.348* | 1.853* | 0.278* | 0.000*
Q03 | 3.490 | 1.009 | -3.458 | 0.710 | 0.000
    | 1.335* | 0.537* | 1.342* | 0.378* | 0.000*
Q05 | 2.156 | 0.903 | -2.387 | 0.670 | 0.000
    | 0.483* | 0.377* | 0.741* | 0.280* | 0.000*
Q06 | 2.753 | 1.524 | -1.806 | 0.836 | 0.000
    | 1.427* | 1.002* | 0.410* | 0.550* | 0.000*
Q07 | 2.242 | 0.712 | -3.149 | 0.580 | 0.000
    | 0.426* | 0.285* | 1.286* | 0.232* | 0.000*
Q08 | 1.857 | 0.579 | -3.208 | 0.501 | 0.000
    | 0.299* | 0.219* | 1.205* | 0.189* | 0.000*
Q09 | 1.025 | 0.601 | -1.704 | 0.515 | 0.000
    | 0.213* | 0.201* | 0.578* | 0.173* | 0.000*
Q10 | 1.123 | 0.425 | -2.642 | 0.391 | 0.000
    | 0.191* | 0.145* | 0.961* | 0.134* | 0.000*
Q11 | 1.431 | 0.920 | -1.556 | 0.677 | 0.000
    | 0.318* | 0.328* | 0.459* | 0.242* | 0.000*
Q12 | 0.886 | 0.381 | -2.328 | 0.356 | 0.000
    | 0.171* | 0.125* | 0.864* | 0.116* | 0.000*
Q13 | 0.775 | 0.581 | -1.335 | 0.502 | 0.000
    | 0.184* | 0.187* | 0.465* | 0.162* | 0.000*
Q14 | 0.721 | 0.432 | -1.669 | 0.397 | 0.000
    | 0.167* | 0.141* | 0.634* | 0.129* | 0.000*
Q15 | 0.287 | 0.366 | -0.785 | 0.344 | 0.000
    | 0.149* | 0.117* | 0.488* | 0.110* | 0.000*
```

\* STANDARD ERROR

LARGEST CHANGE = 0.009805

```
-----
PARAMETER MEAN STN DEV
-----
SLOPE 0.707 0.325
LOG(SLOPE) -0.434 0.426
THRESHOLD -2.325 0.973
```

```
MEAN 0.00000
S.D. 1.00000
```

## 2.2 単一フォーム、複数集団の場合 (test1.dat)

データ

G1_1	-2.497	1	1	111011000001000	13
G1_2	-2.386	1	1	001010000100100	9
G1_3	-2.337	1	1	110111000000010	11
G1_4	-2.314	1	1	010100101000001	8
G1_5	-2.280	1	1	110001011000000	10
G1_6	-2.256	1	1	110000010100000	10
G1_7	-2.255	1	1	011010001101010	13
G1_8	-2.213	1	1	011000001100000	7
G1_9	-2.186	1	1	100100001011000	8
G1_10	-2.179	1	1	1111011110101000	13
G2_101	-0.543	1	2	111111111010111	19
G2_102	-0.539	1	2	1101010111101011	18
G2_103	-0.536	1	2	111111111111101	22
G2_104	-0.535	1	2	111111111011101	22
G2_105	-0.535	1	2	111111111000100	16
G2_106	-0.526	1	2	111111111111001	20
G2_107	-0.524	1	2	111110110001000	17
G2_108	-0.508	1	2	111111111111010	26
G2_109	-0.507	1	2	111111111101110	22
G2_110	-0.503	1	2	111111110010111	28

コマンドファイル

```
test1g.blm
One test, Low and High Ability Groups
>COMMENT
Simple Test
061117
>GLOBAL DFName = 'test1.DAT', NPArm = 2, SAVe;
>SAVE PARM = 'test1g.PAR', SCORe = 'test1g.SCO', ISTat = 'test1g.ist';
>LENGTH NITems = (15);
>INPUT NTotAl = 15,
    SAMple = 99999,
    NALt = 5,
    NIDchar = 8,
    NGRoup = 2,
    NFName = 'miss.dat';
>ITEMS INAmes = (Q01(1)Q15);
>TEST1 TName = 'Simple',
    INUmber = (1(1)15);
>GROUP1 GName = 'L',
    LENgth = 15,
    INUmbers = (1(1)15);
>GROUP2 GName = 'H',
    LENgth = 15,
    INUmbers = (1(1)15);
(8A1,T22,I1,T25,60A1)
>CALIB NQPt = 15, CYCles = 10, REFerence = 0, NOFfloat;
>SCORE NQPt = (31), NOPrint;
```

結果

-2 LOG LIKELIHOOD: 2717.1968

CYCLE 10; LARGEST CHANGE= 0.03459

====> NOTE: CONVERGENCE HAS NOT BEEN REACHED TO CRITERION = 0.01000

1

SUBTEST SIMPLE ; ITEM PARAMETERS AFTER CYCLE 10

ITEM	INTERCEPT	SLOPE	THRESHOLD	LOADING	ASYMPTOTE
Q01	1.923	0.917	-2.097	0.676	0.000
	0.271*	0.247*	0.390*	0.182*	0.000*
Q02	2.078	0.987	-2.106	0.702	0.000
	0.314*	0.286*	0.398*	0.204*	0.000*
Q03	1.626	1.221	-1.331	0.774	0.000
	0.459*	0.465*	0.190*	0.295*	0.000*
Q04	1.442	0.868	-1.660	0.656	0.000
	0.393*	0.374*	0.340*	0.282*	0.000*
Q05	1.129	0.750	-1.506	0.600	0.000
	0.167*	0.173*	0.265*	0.138*	0.000*
Q06	0.943	1.041	-0.906	0.721	0.000
	0.202*	0.231*	0.132*	0.160*	0.000*
Q07	1.084	1.055	-1.028	0.726	0.000
	0.196*	0.205*	0.131*	0.141*	0.000*
Q08	0.918	0.875	-1.049	0.659	0.000
	0.154*	0.160*	0.158*	0.121*	0.000*
Q09	0.454	0.522	-0.869	0.463	0.000
	0.106*	0.102*	0.223*	0.090*	0.000*
Q10	0.500	0.615	-0.813	0.524	0.000
	0.109*	0.109*	0.194*	0.093*	0.000*
Q11	0.243	1.007	-0.241	0.709	0.000
	0.133*	0.151*	0.124*	0.107*	0.000*
Q12	0.305	0.578	-0.527	0.501	0.000
	0.102*	0.099*	0.183*	0.086*	0.000*
Q13	0.066	0.720	-0.091	0.584	0.000
	0.104*	0.111*	0.144*	0.090*	0.000*
Q14	0.116	0.623	-0.186	0.529	0.000
	0.101*	0.102*	0.162*	0.087*	0.000*
Q15	-0.381	0.716	0.533	0.582	0.000
	0.114*	0.117*	0.158*	0.095*	0.000*

\* STANDARD ERROR

LARGEST CHANGE = 0.034590

PARAMETER	MEAN	STN DEV
SLOPE	0.833	0.206
LOG(SLOPE)	-0.212	0.254
THRESHOLD	-0.925	0.749

GROUP: 1	L	GROUP: 2	H
MEAN	-0.91798	MEAN	0.91798
S.E.	0.08487	S.E.	0.08396
S.D.	0.25516	S.D.	0.49950
S.E.	0.00109	S.E.	0.00504

COMBINED GROUPS:

MEAN	0.00000
S.D.	1.00000

## 2.3 複数フォーム、複数集団の場合 1 compressed format (test21c.dat)

データ

G1_1	-2.497	1	1	111011000001000	13
G1_2	-2.386	1	1	001010000100100	9
G1_3	-2.337	1	1	110111000000010	11
G1_4	-2.314	1	1	010100101000001	8
G1_5	-2.280	1	1	110001011000000	10
G1_6	-2.256	1	1	110000010100000	10
G1_7	-2.255	1	1	011010001101010	13
G1_8	-2.213	1	1	011000001100000	7
G1_9	-2.186	1	1	100100001011000	8
G1_10	-2.179	1	1	111101110101000	13
G2_1	-1.863	2	2	01000000111010000	12
G2_2	-1.803	2	2	00001000010000000	12
G2_3	-1.774	2	2	101000100010010011	13
G2_4	-1.630	2	2	001000001100010000	8
G2_5	-1.617	2	2	110000000010010000	11
G2_6	-1.550	2	2	000001000010101000	15
G2_7	-1.385	2	2	101101001100110000	19
G2_8	-1.379	2	2	110000001000001010	15
G2_9	-1.361	2	2	100110000010000011	17
G2_10	-1.347	2	2	011001011000010000	15
G3_1	-1.042	3	3	1000010000000000	15
G3_2	-0.865	3	3	101001000001000	20
G3_3	-0.846	3	3	1000000000000001	19
G3_4	-0.795	3	3	000100100100000	16
G3_5	-0.776	3	3	000100010100010	18
G3_6	-0.730	3	3	1000001010100001	20
G3_7	-0.684	3	3	000001000111101	19
G3_8	-0.672	3	3	010000010000000	21
G3_9	-0.645	3	3	001010100010001	21
G3_10	-0.620	3	3	000100001101000	17

コマンドファイル

```
test21c.blm
Nonequivalent Groups Equating: compressed format
>COMMENT
Using Compressed Format and Not Presented
990630,061117
>GLOBAL DFName = 'test21c.DAT', NPArm = 2, SAVE;
>SAVE PARM = 'test21c.PAR', SCORe = 'test21c.SCO', IStat = 'test21c.ist';
>LENGTH NITems = (40);
>INPUT NTotAl = 40, SAMple = 99999, NALt = 5, NIDchar = 8,
    NGRoup = 3, NForm = 3, NFName = 'missc.dat';
>ITEMS INames = (Q01(1)Q40);
>TEST1 TName = 'CmprsdFo', INUmber = (1(1)40);
>FORM1 LENgth = 15, INUmbers = (1(1)15);
>FORM2 LENgth = 18, INUmbers = (12(1)29);
>FORM3 LENgth = 15, INUmbers = (26(1)40);
>GROUP1 GName = 'g1', LENgth = 15, INUmbers = (1(1)15);
>GROUP2 GName = 'g2', LENgth = 18, INUmbers = (12(1)29);
>GROUP3 GName = 'g3', LENgth = 15, INUmbers = (26(1)40);
(8A1,T19,I1,T22,I1,T25,60A1)
>CALIB NQPt = 15, REFerence = 0, NOFloat;
>SCORE NQPt = (31), NOPrint;
```

## 2.4 複数フォーム、複数集団の場合 2 expanded format (test21e.dat)

### データ

G1_1	-2.497	1	111011000001000.....	13
G1_2	-2.386	1	001010000100100.....	9
G1_3	-2.337	1	110111000000010.....	11
G1_4	-2.314	1	010100101000001.....	8
G1_5	-2.280	1	110001011000000.....	10
G1_6	-2.256	1	110000010100000.....	10
G1_7	-2.255	1	011010001101010.....	13
G1_8	-2.213	1	011000001100000.....	7
G1_9	-2.186	1	100100001011000.....	8
G1_10	-2.179	1	111101110101000.....	13
G2_1	-1.863	2	.....010000001110100000.....	12
G2_2	-1.803	2	.....000010000100000000.....	12
G2_3	-1.774	2	.....101000100010010011.....	13
G2_4	-1.630	2	.....001000001100010000.....	8
G2_5	-1.617	2	.....110000000010010000.....	11
G2_6	-1.550	2	.....000001000010101000.....	15
G2_7	-1.385	2	.....101101001100110000.....	19
G2_8	-1.379	2	.....110000001000001010.....	15
G2_9	-1.361	2	.....100110000010000011.....	17
G2_10	-1.347	2	.....011001011000010000.....	15
G3_1	-1.042	3	.....1000010000000000.....	15
G3_2	-0.865	3	.....101001000001000.....	20
G3_3	-0.846	3	.....1000000000000001.....	19
G3_4	-0.795	3	.....000100100100000.....	16
G3_5	-0.776	3	.....0001000101000010.....	18
G3_6	-0.730	3	.....1000001010100001.....	20
G3_7	-0.684	3	.....000001000111101.....	19
G3_8	-0.672	3	.....010000010000000.....	21
G3_9	-0.645	3	.....001010100010001.....	21
G3_10	-0.620	3	.....000100001101000.....	17

### コマンドファイル

```

test21e.blm
Nonequivalent Groups Equating: expanded format
>COMMENT
Using Expanded Format and Not Presented
990630,050701,061117
>GLOBAL DFName = 'test21e.DAT', NPArm = 2, SAVe;
>SAVE PARM = 'test21e.PAR', SCORe = 'test21e.SCO', ISTat = 'test21e.ist';
>LENGTH NITems = (40);
>INPUT NTOtal = 40, SAMple = 99999, NALt = 5, NIDchar = 8,
    NGRoup = 3, NFName = 'misse.dat';
>ITEMS INAmes = (Q01(1)Q40);
>TEST1 TNAme = 'Exp', INUmber = (1(1)40);
>GROUP1 GNAme = 'g1', LENgth = 40, INUmbers = (1(1)40);
>GROUP2 GNAme = 'g2', LENgth = 40, INUmbers = (1(1)40);
>GROUP3 GNAme = 'g3', LENgth = 40, INUmbers = (1(1)40);
(8A1,T19,I1,T21,60A1)
>CALIB NQPt = 15, CYCLe=20, NEWton=20, REFerence = 0, NOFfloat;
>SCORE NQPt = (31), NOPrint;

```

### 結果

```

-2 LOG LIKELIHOOD:          16607.1790

CYCLE    27;  LARGEST CHANGE=  0.00999

```

SUBTEST EXP ; ITEM PARAMETERS AFTER CYCLE 27

ITEM	INTERCEPT S.E.	SLOPE S.E.	THRESHOLD S.E.	LOADING S.E.	ASYMPTOTE S.E.	CHISQ (PROB)	DF
Q01	2.756 0.399*	0.922 0.237*	-2.990 0.397*	0.678 0.175*	0.000 0.000*	2.5 (0.6377)	4.0
Q02	2.956 0.488*	1.002 0.287*	-2.952 0.412*	0.708 0.203*	0.000 0.000*	3.9 (0.2706)	3.0
Q03	2.826 0.566*	1.259 0.337*	-2.246 0.187*	0.783 0.210*	0.000 0.000*	1.0 (0.8093)	3.0
Q04	2.541 0.409*	0.996 0.251*	-2.552 0.278*	0.706 0.178*	0.000 0.000*	3.6 (0.4658)	4.0
Q05	1.925 0.272*	0.803 0.180*	-2.398 0.254*	0.626 0.140*	0.000 0.000*	9.9 (0.0775)	5.0
Q06	1.933 0.336*	1.135 0.229*	-1.703 0.101*	0.750 0.151*	0.000 0.000*	2.9 (0.7099)	5.0
Q07	2.058 0.313*	1.106 0.207*	-1.861 0.117*	0.742 0.139*	0.000 0.000*	5.5 (0.3590)	5.0
Q08	1.899 0.266*	0.944 0.178*	-2.012 0.153*	0.686 0.129*	0.000 0.000*	7.0 (0.2211)	5.0
Q09	0.730 0.122*	0.436 0.093*	-1.673 0.212*	0.400 0.085*	0.000 0.000*	2.9 (0.8236)	6.0
Q10	0.933 0.140*	0.546 0.106*	-1.709 0.182*	0.479 0.093*	0.000 0.000*	1.1 (0.9809)	6.0
Q11	1.264 0.208*	0.998 0.163*	-1.266 0.087*	0.707 0.116*	0.000 0.000*	6.5 (0.2603)	5.0
Q12	0.937 0.083*	0.618 0.080*	-1.516 0.139*	0.526 0.068*	0.000 0.000*	4.5 (0.7226)	7.0
Q13	0.684 0.074*	0.724 0.080*	-0.945 0.088*	0.587 0.065*	0.000 0.000*	6.3 (0.6166)	8.0
Q14	0.585 0.071*	0.641 0.076*	-0.912 0.094*	0.540 0.064*	0.000 0.000*	6.9 (0.5489)	8.0
Q15	0.246 0.063*	0.690 0.082*	-0.356 0.079*	0.568 0.067*	0.000 0.000*	6.3 (0.4996)	7.0
Q16	0.143 0.080*	1.001 0.192*	-0.143 0.081*	0.707 0.136*	0.000 0.000*	6.5 (0.2632)	5.0
Q17	0.564 0.084*	0.842 0.171*	-0.670 0.150*	0.644 0.131*	0.000 0.000*	3.8 (0.5721)	5.0
Q18	0.671 0.103*	1.365 0.276*	-0.491 0.096*	0.807 0.163*	0.000 0.000*	2.2 (0.5343)	3.0
Q19	0.681 0.091*	0.860 0.184*	-0.792 0.166*	0.652 0.140*	0.000 0.000*	5.1 (0.4069)	5.0
Q20	0.821 0.100*	0.937 0.201*	-0.876 0.171*	0.684 0.147*	0.000 0.000*	3.5 (0.6163)	5.0
Q21	-0.094 0.070*	0.408 0.096*	0.230 0.173*	0.378 0.089*	0.000 0.000*	3.2 (0.7781)	6.0
Q22	-0.049 0.073*	0.586 0.122*	0.084 0.123*	0.506 0.105*	0.000 0.000*	6.0 (0.4208)	6.0
Q23	-0.230 0.078*	0.893 0.171*	0.257 0.095*	0.666 0.128*	0.000 0.000*	1.0 (0.9634)	5.0
Q24	-0.250 0.075*	0.659 0.134*	0.379 0.127*	0.551 0.112*	0.000 0.000*	2.8 (0.8310)	6.0
Q25	-0.255 0.072*	0.427 0.098*	0.598 0.200*	0.393 0.091*	0.000 0.000*	2.2 (0.8986)	6.0
Q26	0.289 0.061*	0.726 0.103*	-0.399 0.116*	0.587 0.083*	0.000 0.000*	8.5 (0.2044)	6.0
Q27	-0.110 0.063*	0.662 0.093*	0.166 0.085*	0.552 0.077*	0.000 0.000*	8.1 (0.2288)	6.0
Q28	-0.148 0.065*	0.793 0.103*	0.186 0.072*	0.621 0.081*	0.000 0.000*	5.9 (0.4327)	6.0

Q29	-0.373	0.812	0.459	0.630	0.000	4.5	5.0
	0.070*	0.103*	0.068*	0.080*	0.000*	(0.4862)	
Q30	-0.355	0.763	0.466	0.607	0.000	7.7	4.0
	0.138*	0.157*	0.114*	0.125*	0.000*	(0.1032)	
Q31	-0.562	0.803	0.699	0.626	0.000	4.9	4.0
	0.146*	0.163*	0.095*	0.127*	0.000*	(0.2942)	
Q32	-0.545	0.866	0.629	0.655	0.000	2.6	3.0
	0.150*	0.171*	0.093*	0.129*	0.000*	(0.4508)	
Q33	-0.470	0.535	0.878	0.472	0.000	1.4	3.0
	0.126*	0.125*	0.133*	0.111*	0.000*	(0.7101)	
Q34	-0.576	0.580	0.993	0.502	0.000	2.2	4.0
	0.131*	0.130*	0.129*	0.112*	0.000*	(0.6964)	
Q35	-0.821	0.882	0.932	0.661	0.000	3.8	3.0
	0.162*	0.173*	0.091*	0.130*	0.000*	(0.2871)	
Q36	-0.800	0.473	1.691	0.428	0.000	1.0	3.0
	0.128*	0.116*	0.248*	0.105*	0.000*	(0.7918)	
Q37	-0.941	0.727	1.294	0.588	0.000	1.7	4.0
	0.156*	0.152*	0.135*	0.123*	0.000*	(0.7891)	
Q38	-0.846	0.503	1.682	0.449	0.000	9.5	4.0
	0.125*	0.116*	0.239*	0.103*	0.000*	(0.0498)	
Q39	-1.057	0.515	2.051	0.458	0.000	2.0	4.0
	0.139*	0.123*	0.309*	0.110*	0.000*	(0.7341)	
Q40	-1.164	0.367	3.175	0.344	0.000	2.8	4.0
	0.128*	0.102*	0.668*	0.096*	0.000*	(0.5983)	

-----  
LARGEST CHANGE = 0.009986 173.9 196.0  
(0.8702)  
-----

PARAMETER	MEAN	STN DEV
SLOPE	0.770	0.237
LOG(SLOPE)	-0.309	0.320
THRESHOLD	-0.340	1.439

GROUP: 1 G1 QUADRATURE POINTS, POSTERIOR WEIGHTS, MEAN AND S.D. :  
MEAN -0.92511  
S.E. 0.07168  
S.D. 0.79165  
S.E. 0.05974

GROUP: 2 G2 QUADRATURE POINTS, POSTERIOR WEIGHTS, MEAN AND S.D. :  
MEAN 0.09440  
S.E. 0.05619  
S.D. 0.65315  
S.E. 0.03648

GROUP: 3 G3 QUADRATURE POINTS, POSTERIOR WEIGHTS, MEAN AND S.D. :  
MEAN 0.83071  
S.E. 0.07100  
S.D. 0.62600  
S.E. 0.05141

COMBINED GROUPS: QUADRATURE POINTS AND POSTERIOR WEIGHTS, MEAN AND S.D. :  
MEAN 0.00000  
S.D. 1.00000

## 2.5 項目パラメタが既知の場合

```
test1Ln.blm
One Test: Low Ability Group Item Parameters Given
>COMMENT
Simple Test
Item Parameters Given

>GLOBAL DFName = 'test1L.DAT',
      IFName = 'test1L.PAR',
      NPArm = 2,
      SAVe;
>SAVE PARm = 'test1Ln.PAR',
      SCOrE = 'test1Ln.SCO',
      ISTat = 'test1Ln.ist';
>LENGTH NITems = (15);
>INPUT NTOTal = 15,
      SAMple = 99999,
      NALt = 5,
      NIDchar = 8,
      NFName = 'miss.dat';
>ITEMS INAmes = (Q01(1)Q15);
>TEST1 TNAme = 'Simple',
      INUंबर = (1(1)15);
(8A1,T25,60A1)
>CALIB SElect = (0);
>SCORE NQPt = (31);
```