流通科学大学論集－人間・社会・自然編―第 20 巻第 1 号, 1-25（2007）

Rasch 評定尺度モデル解題とそのモデル上での閾値順序性

Interpretation of the Polytomous Rasch Models and the Ordering of the Thresholds

井澤　廣行*
Hiroyuki Izawa

To the author's knowledge, there has been no availability of explication written in Japanese about the polytomous Rasch models. This paper has attempted to interpret Andrich's Rating Scale, Masters' Partial Credit, and Linacre's Many-Facet Rasch Measurement. It has been notified that the first two models are identical in the structure at the fundamental level of one person responding to one item, in which the latent thresholds, i.e., points partitioning a continuum, are defined to be conditionally ordered. This paper has also referred to the significance of confirming the natural ordering of the thresholds' estimates for the categorically divisional construct validation.

Key words: Polytomous Rasch models, Andrich, Masters, Linacre, Thresholds.

1.　はじめに

筆者が Rasch 項目分析モデル(Rasch[1], 1960)に併せて Andrich[2] (1978a)の Rating Scale モデル及び Masters[3] (1982)の Partial Credit モデルに文献上で初めて出会ったのは、McNamara[4] (1996)によって著された *Measuring Second Language Performance* においてであった。それは英国 Reading 大学での統計学 Diploma 修了後の 2001 年 7 月のことであったが、Rating Scale モデルと Partial Credit モデルそれぞれにおける閾値の理解に悩んだことを鮮明に記憶している。McNamara[4] (1996)自身がいみじくも『概して、[Partial Credit モデル上での閾値母数推定値]デルタは解釈するのに難しい』(p. 290)と述べている。又、『[Rating Scale モデル上での閾値母数推定値]タウは[その値が大きくなる]順序に従う必要はない』(p. 290)とも McNamara[4] (1996)により指摘されているが、それは誤解である(Andrich[5], 2005、参照)。

Rating Scale モデル上での閾値母数は Partial Credit モデル上での閾値母数を単純化したものとして導出され得る(Masters[3], 1982, pp. 162−163; Wright and Masters[6], 1982, p. 49)ことにより、両モデルは一人の被験者の一つの項目への応答焦点化の上では同一モデルである(Andrich[5], 2005, pp. 30−31; Luo[7], 2005)。上記の McNamara[4] (1996, p. 290)のみならず他の著名な心理測

*流通科学大学サービス産業学部、〒651-2188　神戸市西区学園西町 3-1

（2007 年 3 月 29 日受理）

定学者によっても Rating Scale モデルと Partial Credit モデルを併せた Rasch 評定尺度モデル上での閾値母数推定値の解釈困難性と順序不要性が指摘されている(Andrich[8), 2004, p. 191; Luo[7), 2005, pp. 451−452、参照)。Andrich[5) (2005)による閾値順序性への洞察は順序尺度データ分析研究者を納得させるものであり、順序尺度データへの Rasch 評定尺度モデル適用上での閾値母数推定値順序性は設定項目群によるカテゴリー順序段階尺度化妥当性を見る上での最優先指標であると思われる。

筆者の知り得る限りにおいて、Andrich[2) (1978a)による Rating Scale モデルについての日本での詳述は存在していない。又、Andrich・Rasch モデル・評定尺度に関するインターネット検索上での Rating Scale Model との名目のみによる出現は藤森[9)(2002, p. 21)と村木[10)(2005, p. 10)に限られている。一方、Rating Scale モデルに評定者を変数として加えた Linacre[11) (1989)考案による Many-Facet Rasch Measurement の日本における適用上での研究報告増加は秋山[12)(2000)を含めて周知の事実である。順序尺度データへの Rating Scale モデル適用上での上記の閾値順序性は多相 Rasch 測定上でも必然であり(Linacre[11), 1989, pp. 57−59、参照)、その類の研究報告においても評定者による被験者評価項目群に関する評定基準カテゴリー順序段階尺度化妥当性を見る上での最優先必須事項である。その確認とデータ潜在特性一次元性充足度の検証が上記の秋山[12)(2000)においてはなされていない。日本でのその Rasch 評定尺度モデル適用状況を鑑みて、その解題に併せてモデル属性規定要件とされる閾値順序性(Andrich, 1978a[2), 2004[8), and 2005[5))を考察することが本稿の目的である。

2. Rating Scale モデル解題

Rasch 項目分析モデル(Rasch[1), 1960, pp. 62−125)は各項目が 0 と 1 から成る二つのカテゴリーへの応答データについて受験者群能力と項目群困難度の母数結合一次元性規定の上で 0 と 1 それぞれの生起確率をモデル化したものである。1961 年に Rasch[13) (1961)は三つ以上の順序カテゴリーから成る項目群への被験者群応答に関する多次元ベクトルを仮定した確率モデルの数学的考察で以って一次元順序段階尺度データ分析モデルを発表した。そのモデル化は、母数 θ を持つ被験者 v のカテゴリー数 m から成り刺激母数 σ を持つ項目 i に対するカテゴリー x への応答確率が次式により表されるものであった。

$$P\{x \mid \theta_v + \sigma_i\} = \exp[\phi(x)(\theta_v + \sigma_i) + \rho(x)] / \gamma(\theta_v, \sigma_i)$$

(Rasch[13), 1961, p. 333)

上式右辺分母の $\gamma(\theta_v, \sigma_i)$ は、$P\{x \mid \theta_v + \sigma_i\}$, $x \in \{1, 2, \cdots, m\}$ の確率合計を 1 とする基準化係数 $\sum_{u=1}^{m} \exp[\phi(\mu)(\theta_v + \sigma_i) + \rho(\mu)]$ と理解される(Andersen[14), 1972, p. 43)。ϕ と ρ はそれぞれ得点母数(scoring parameter)とカテゴリー母数(category parameter)との名称が Andersen (1972[14), p. 43)によって付されている。それは、各被験者総点一群と各項目総点一群が

それぞれ十分統計量となって各被験者位置母数と各項目位置母数の分離推定を可能として、両母数の結合一次元性(‘a single invariant conjoint order of item and person parameters’in Wright[15], 1991, p. 158)を規定する必要十分なモデルとして演繹されたものであった(Masters[3], 1982, p. 152、参照)。この母数結合一次元性は Rasch 項目分析モデル規定に準ずるものである。それは、例えば、ある一つの質問紙調査潜在特性に関して、すべての被験者が賛成度の高い項目群に対してその高い程度に順ずる強い態度を示し、賛成態度のより強い被験者はいずれの項目に対しても賛成する程度が高くなることを意味する規定である。更に、Andersen[16] (1977) により、各被験者総点一群が各被験者位置母数推定のための十分統計量となるためには、$\phi_1, \phi_2, \cdots, \phi_m$ は必然的に等間隔になる(p. 76)、と示された。これらが以降の Rasch 評定尺度モデル考案の端緒となり、Rating Scale モデル(Andrich, 1978a[2] and b[17])、Partial Credit モデル(Masters[3], 1982)、そして Many-Facet Rasch Measurement (Linacre, 1989[11] and 1989−2001[18])への結実となった。以上が本節の導入部である。

Rasch (1961[13], p. 333)によって与えられた上記の確率式 $P\{x \mid \theta_v + \sigma_i\} = \exp[\phi(x)(\theta_v + \sigma_i) + \rho(x)] / \gamma(\theta_v, \sigma_i)$について 1977 年の時点ではその明確な解釈が不首尾であると断じた(Andrich[2], 1978a, p. 564)上で、オーストラリア人である David Andrich[2] (1978a)によりその確率モデルを項目群が三つの順序カテゴリーから成る以下のものとしてその解釈が思考される。

$$p\{X = 0 \mid \beta, \delta, \kappa, \phi\} = 1 / \{1 + \sum_{k=1}^{2} \exp[\kappa_k + \phi_k(\beta - \delta)]\}$$

$$p\{X = x \mid \beta, \delta, \kappa, \phi\} = \exp[\kappa_x + \phi_x(\beta - \delta)] / \{1 + \sum_{k=1}^{2} \exp[\kappa_k + \phi_k(\beta - \delta)]\}$$

$x = 1, 2$

Andrich[2] (1978a, p. 564)

始めに、項目を構成する三つの順序カテゴリーの潜在生起境界値としての閾値母数 τ_1 と τ_2 が項目位置母数 δ を限定する(qualify)変数として導入されて、潜在閾値母数順序性 $\tau_1 < \tau_2$ と規定される(Andrich[2], 1978a, p. 565)。τ_1 と τ_2 それぞれの閾値母数において独立に分離して該当カテゴリーが生起するか否かの 1・0 が仮定されて、更に、τ_1 と τ_2 それぞれにおいて異なるカテゴリー弁別母数 α_1 と α_2 が導入される(Andrich[2], 1978a, pp. 565−566)。なお、Andrich[2] (1978a)による α_1 と α_2 の導入は、以降に示される様に、上に参照された Rasch (1961[13], p. 333)提示確率モデルの明確な解釈の意図の下であることに留意される。その結果として、τ_1 と τ_2 それぞれの閾値母数における独立分離した二値反応としての 0・1 生起確率が次のものとして示される。

$$p\{X_1 = 0 \mid \beta, \delta, \alpha_1, \tau_1\} = 1 / \eta_1$$

$$p\{X_1 = 1 \mid \beta, \delta, \alpha_1, \tau_1\} = \exp[\alpha_1\{\beta - (\delta + \tau_1)\}] / \eta_1$$

なお、基準化係数としての $\eta_1(\beta, \delta, \alpha_1, \tau_1) = 1 + \exp[\alpha_1\{\beta - (\delta + \tau_1)\}]$

$$p\{X_2 = 0 \mid \beta, \delta, \alpha_2, \tau_2\} = 1 / \eta_2$$

$p\{X_2 = 1 \mid \beta, \delta, \alpha_2, \tau_2\} = \exp[\alpha_2\{\beta - (\delta + \tau_2)\}] / \eta_2$

なお、基準化係数としての $\eta_2(\beta, \delta, \alpha_2, \tau_2) = 1 + \exp[\alpha_2\{\beta - (\delta + \tau_2)\}]$

Andrich[2] (1978a, p. 566)

それぞれ 0 と 1 の要素事象から成る二つの独立事象を集合する全事象 Ω が閾値母数 τ_1と τ_2での事象生起順序の上で Ω = {(0, 0), (1, 0), (1, 1), (0, 1)}と表現される(Andrich[2], 1978a, p. 566)。独立事象の乗法定理により、各事象の生起確率が次のものとして与えられる。

$p\{(0, 0)\} = 1 / \eta_1\eta_2$

$p\{(1, 0)\} = \exp[\alpha_1\{\beta - (\delta + \tau_1)\}] / \eta_1\eta_2$

$p\{(1, 1)\} = \exp[\alpha_1\{\beta - (\delta + \tau_1)\} + \alpha_2\{\beta - (\delta + \tau_2)\}] / \eta_1\eta_2$

$p\{(0, 1)\} = \exp[\alpha_2\{\beta - (\delta + \tau_2)\}] / \eta_1\eta_2$

Andrich[2] (1978a, p. 567)

ここで、Guttman (1950, cited in Andrich[5], 2005, p. 34)の完全順序尺度モデルに依拠する潜在閾値母数順序性 $\tau_1 < \tau_2$が規定されている(Andrich[2], 1978a, p. 565)故に、事象(0, 1)の生起が排除される(Andrich[2], 1978a, p. 567)。従って、三つのカテゴリーから成る順序尺度の上で、潜在閾値母数順序性に基づく正当な全事象 Ω’ は Ω’ = {(0, 0), (1, 0), (1, 1)}と限定されて、p{(0, 0)} + p{(1, 0)} + p{(1, 1)} = 1 と必然化される(Andrich[2], 1978a, p. 567)。そのために、三つの事象の生起確率合計を 1 とする基準化係数 $\gamma(\beta, \delta, \alpha, \tau) = 1 + \exp[\alpha_1\{\beta - (\delta + \tau_1)\}] + \exp[\alpha_1\{\beta - (\delta + \tau_1)\} + \alpha_2\{\beta - (\delta + \tau_2)\}]$が導入されて、三つの事象それぞれの生起確率が次のものとして示される。

$p\{(0, 0) \mid \beta, \delta, \alpha, \tau\} = 1 / \gamma$

$p\{(1, 0) \mid \beta, \delta, \alpha, \tau\} = \exp[\alpha_1\{\beta - (\delta + \tau_1)\}] / \gamma$

$p\{(1, 1) \mid \beta, \delta, \alpha, \tau\} = \exp[\alpha_1\{\beta - (\delta + \tau_1)\} + \alpha_2\{\beta - (\delta + \tau_2)\}] / \gamma$

Andrich[2] (1978a, p. 567)

素データ上での項目群についての三つの順序カテゴリー変数 X (= 0, 1, 2)それぞれの生起が上記三つの事象それぞれに対応するものとして、各順序カテゴリー変数の生起確率が次のものとして表現される。それは、$p\{X = 0\}$、$p\{X = 1\}$、$p\{X = 2\}$がそれぞれ τ_1を超えていない確率、τ_1のみを超えて τ_2を超えていない確率、τ_1と τ_2のいずれをも超えている確率であることを意味している。

$p\{X = 0\} = p\{(0, 0) \mid \beta, \delta, \alpha, \tau\} = 1 / \gamma$

$p\{X = 1\} = p\{(1, 0) \mid \beta, \delta, \alpha, \tau\} = \exp[\alpha_1\{\beta - (\delta + \tau_1)\}] / \gamma$

$p\{X = 2\} = p\{(1, 1) \mid \beta, \delta, \alpha, \tau\}$

$= \exp[\alpha_1\{\beta - (\delta + \tau_1)\} + \alpha_2\{\beta - (\delta + \tau_2)\}] / \gamma$

Andrich[2] (1978a, pp. 567–568)

上式において、$\alpha_1\{\beta - (\delta + \tau_1)\} = -\alpha_1\tau_1 + \alpha_1(\beta - \delta)$、並びに、$\alpha_1\{\beta - (\delta + \tau_1)\} + \alpha_2\{\beta - (\delta + \tau_2)\} = -\alpha_1\tau_1 - \alpha_2\tau_2 + (\alpha_1 + \alpha_2)(\beta - \delta)$と変換される(Andrich[2], 1978a, p. 568)。従って、次の様な記号変換が可能となる。

$\phi_1 = \alpha_1;\ \kappa_1 = -\alpha_1\tau_1$

$\phi_2 = \phi_1 + \alpha_2 = \alpha_1 + \alpha_2;\ \kappa_2 = \kappa_1 - \alpha_2\tau_2 = -\alpha_1\tau_1 - \alpha_2\tau_2$

Andrich[2] (1978a, p. 568)

上記の新たな使用記号によって、項目群についての三つの順序カテゴリー変数 X (= 0, 1, 2)それぞれの生起確率が次のものとして表現される。

$p\{X = 0 \mid \beta, \delta, \kappa, \phi\} = 1 / \gamma$

$p\{X = 1 \mid \beta, \delta, \kappa, \phi\} = \exp[\kappa_1 + \phi_1(\beta - \delta)] / \gamma$

$p\{X = 2 \mid \beta, \delta, \kappa, \phi\} = \exp[\kappa_2 + \phi_2(\beta - \delta)] / \gamma$

なお、$\gamma = \gamma(\beta, \delta, \kappa, \phi) = 1 + \sum_{k=1}^{2} \exp[\kappa_k + \phi_k(\beta - \delta)]$

Andrich[2] (1978a, p. 568)

上式は、本節冒頭に参照された Rasch (1961[13], p. 333)のモデル確率式 $P\{x \mid \theta_v + \sigma_i\} = \exp[\phi(x)(\theta_v + \sigma_i) + \rho(x)] / \gamma(\theta_v, \sigma_i)$をその確率モデルについて項目群が三つの順序カテゴリーから成るものとして Andrich[2] (1978a)によりその解釈思考の糧とされたもの(p. 564)と正に同一である。上式は、更に、(m + 1)個の順序カテゴリーから構成される項目について次の様に一般化される。

$p\{X = x \mid \beta, \delta, \kappa, \phi\} = \exp[\kappa_x + \phi_x(\beta - \delta)] / \sum_{k=0}^{m} \exp[\kappa_k + \phi_k(\beta - \delta)]$

なお、$\kappa_x = -\sum_{k=1}^{x} \alpha_k\tau_k;\ \phi_x = \sum_{k=1}^{x} \alpha_k$　　$x = 1, 2, \cdots, m$

但し、$x = 0$ の場合、$\kappa_x = \phi_x = 0$

Andrich[2] (1978a, p. 569)

従って、Andrich[2] (1978a)は Rasch[13] (1961, p. 333)により与えられたそのモデル確率式の解釈提示を完遂したことになる。なお、前述された様に、Andrich[2] (1978a)によるカテゴリー弁別母数 α の導入は Rasch[13] (1961, p. 333)案出確率モデルの明確な解釈の意図の下になされたものである。α はその母数推定十分統計量を持たない故に、上式は母数推定十分統計量の存在を本質的な特徴とする Rasch モデルではないことが顕示されている(Andrich[19], 1995, p. 13)。それは、又、Rasch[13] (1961, p. 333)により与えられたそのモデル確率式が母数推定十分統計量の存在を可能とするには $\phi_{x+1} - \phi_x = \phi_x - \phi_{x-1}$ が必要条件であると Andersen[16] (1977, p. 76)によって示された(Andrich[8], 2004, p. 177、参照)ことと不可分の関係にある。

Rasch 項目分析モデル(Rasch[1], 1960, pp. 62−125)上では、母数推定十分統計量の必然性により項目弁別母数の概念が所与されておらず、その母数が項目群全体に渡って同値 1 であるとの演繹的規定とみなされる。それを踏まえて、項目群全体に対してのみならずカテゴリー間において

も弁別母数を同値1であると規定することにより上式に $\alpha = 1$ が挿入されて次式が導出される。

$$p\{X = x \mid \beta, \delta, \kappa\} = \exp[\kappa_x + x(\beta - \delta)] / \sum_{k=0}^{m} \exp[\kappa_k + \kappa(\beta - \delta)]$$

なお、$\kappa_x = -\sum_{k=1}^{x} \tau_k;\ \kappa_0 = 0$　　$x = 0, 1, 2, \cdots, m$

Andrich[2] (1978a, p. 569)

上式が Andrich の Rating Scale モデルである。得点母数(scoring parameter)が 0 ないしは正の整数となる x で表現されており、それは順序カテゴリー変数 X と同値である x との表示簡易利便性を有している(Andrich[2], 1978a, p. 569)。得点母数のこの整数表示は、閾値母数間の等間隔性ではなくカテゴリー間での弁別母数同等性として示されたモデル規定に由来することに留意される(Andrich[2], 1978a, p. 570)。又、$(x + 1) - x = x - (x - 1) = 1$ となり、上記の Andersen[16] (1977)によって示された母数推定十分統計量の存在を可能とせしめる必要条件を満たしており、項目群全体についての被験者得点合計 Σx が被験者位置母数推定のための十分統計量となる(Andrich[2], 1978a, p. 570)。同様に、受験者群全体に関する項目得点合計が項目位置母数推定のための十分統計量となる(Andrich[17], 1978b, p. 584)。更に、各カテゴリー出現度数がカテゴリー母数 κ_x 推定の十分統計量となり(Andrich[17], 1978b, 584)、第二番目以降の各カテゴリー出現度数一群が一体となって十分統計量となる(Wright and Masters[6], 1982, p. 59)ことにより各閾値母数 τ_k が推定される。全項目に渡って共通の閾値母数推定値が各カテゴリー順序段階に異なるものとして被験者位置母数推定値分布とは分離独立して与えられる(Andrich[20], 1998, p. 648)。各閾値母数推定値はその一点上で隣接する二つのカテゴリーに対する被験者選択同一確率点としての被験者位置母数推定値尺度上での各項目位置母数推定値との相対関係にある隣接順序カテゴリー二者択一境界指標値である(Andrich[20], 1998, p. 648; Linacre[21], 2001, p. 794)。

ここで本質的に留意されるべきことは、Rating Scale モデル適用母数推定上で付与される閾値の順序性はデータの属性である(Andrich[5], 2005, p. 49)との理解である。モデル規定としての潜在閾値母数順序性に対する閾値母数推定結果としてのその規定違反データは、質問紙調査設定項目群全体に係わる一元的構成潜在概念上でのカテゴリー順序段階尺度化妥当性に欠けることを示している(Andrich[8], 2004, p. 184)。データ属性としてのこの閾値母数推定値順序性については第五節において実例上で観察する。

3. Partial Credit モデル解題

オーストラリア人である Geofferey N. Masters が 1982 年に Partial Credit モデルを発表した。正答 1 あるいは誤答 0 という配点項目群に加えて、例えば、0、1、2、3 点により部分得点として 2 点以上の配点項目群を含むテスト、ないしは、リッカート順序尺度上での同一カテゴリー数を持つ項目群から成る質問紙調査データのいずれに対しても適用可能なモデルである。Masters[3] (1982)の着想は、Rasch 項目分析モデルにおける項目個別困難度母数を項目個別一段階通過困難

度母数とみなして、それに基づく項目個別各順序段階通過困難度母数推定への拡張モデル化である。その数理基盤理解のために、Masters[3)] (1982, pp. 157−158 and p. 163)、Masters and Wright[22)] (1984, pp. 542−543)、並びに、Wright and Masters[6)] (1982, p. 42)により提示された Rasch 項目分析モデルに基づく Partial Credit モデルへの数理展開をまとめて以下に参照記述する。

0・1 データに対する Rasch 項目分析モデルは次式で表され、能力母数 β_n を持つ受験者 n が困難度母数 δ_i を持つ項目 i に正答する確率が P であると定義される(Rasch[1)], 1960, p. 75; Wright[23)], 1980, p. 187)。

$$P = \exp(\beta_n - \delta_i) / [1 + \exp(\beta_n - \delta_i)]$$

ここで、項目 i に正答することを項目 i について 0 ではなく 1 が選択されて一段階通過とみなせば、被験者 n により項目 i に対して 0 ではなく 1 が選択される確率 ϕ_{ni1} は次式で表される。

$$\phi_{ni1} = \pi_{ni1} / (\pi_{ni0} + \pi_{ni1}) = \exp(\beta_n - \delta_{i1}) / [1 + \exp(\beta_n - \delta_{i1})]$$

Masters[3)] (1982, p. 157)

上式において、π_{ni1} と π_{ni0} はそれぞれ被験者 n が項目 i について一段階通過境界点で 1 あるいは 0 を選択する確率であり、δ_{i1} は項目 i の一段階通過困難度母数である。0 と 1 のみが表れる二値反応データに関しては、当然に$(\pi_{ni0} + \pi_{ni1}) = 1$ であるから、上式において $\pi_{ni1} / (\pi_{ni0} + \pi_{ni1})$ は不要である。

次に、各項目が三つの順序カテゴリー0・1・2 から成るデータを考えて、項目 i の第一段階通過境界点で 0 ではなく 1 が選択される確率は上式で表現されたものとする。更に、第二カテゴリー1 を選んだ被験者群において、1 に止まらずに第三カテゴリー2 に進行する被験者を想定する。被験者 β_n が項目 i に対して第二段階通過境界点で 1 ではなく 2 を選択する確率 ϕ_{ni2} は次式で表される。

$$\phi_{ni2} = \pi_{ni2} / (\pi_{ni1} + \pi_{ni2}) = \exp(\beta_n - \delta_{i2}) / [1 + \exp(\beta_n - \delta_{i2})]$$

Masters[3)] (1982, p. 158)

各項目が四つの順序カテゴリー0・1・2・3 から成るデータについても同様に、項目 i に対して第三カテゴリー2 を選んだ被験者群において、2 に止まらずに第四カテゴリー3 に進行する被験者 β_n を想定する。被験者 β_n が項目 i に対して第三段階通過境界点で 2 ではなく 3 を選択する確率 ϕ_{ni3} は次式で表される。

$$\phi_{ni3} = \pi_{ni3} / (\pi_{ni2} + \pi_{ni3}) = \exp(\beta_n - \delta_{i3}) / [1 + \exp(\beta_n - \delta_{i3})]$$

Masters[3)] (1982, p. 157)

従って、各項目が(m + 1)の順序カテゴリー0・1・・・m から成る順序尺度データ上でのカテゴリー段階(k − 1)を選んだ被験者群において、(k − 1)に止まらずにカテゴリー段階 k に進行する被験者 β_n の(k − 1)を超えて k を選択する確率 ϕ_{nik} は次式で表される。

$$\phi_{nik} = \pi_{nik} / (\pi_{nik-1} + \pi_{nik}) = \exp(\beta_n - \delta_{ik}) / [1 + \exp(\beta_n - \delta_{ik})]$$

$k = 1, 2, \cdots, m_i$

Wright and Masters[6] (1982, p. 42)

上式を変更して、

$\pi_{nik} + \pi_{nik}*\exp(\beta_n - \delta_{ik}) = \pi_{nik-1}*\exp(\beta_n - \delta_{ik}) + \pi_{nik}*\exp(\beta_n - \delta_{ik})$

両辺から $\pi_{nik}*\exp(\beta_n - \delta_{ik})$ を消去して、$\pi_{nik} = \pi_{nik-1}*\exp(\beta_n - \delta_{ik})$

ここで、例えば、π_{ni3} を考えれば、

$\pi_{ni3} = \pi_{ni2}*\exp(\beta_n - \delta_{i3}) = \pi_{ni1}*\exp(\beta_n - \delta_{i2})*\exp(\beta_n - \delta_{i3})$

$= \pi_{ni0}*\exp(\beta_n - \delta_{i1})*\exp(\beta_n - \delta_{i2})*\exp(\beta_n - \delta_{i3}) = \pi_{ni0}*\exp \sum_{j=1}^{3} (\beta_n - \delta_{ij})$

従って、$\pi_{nik} = \pi_{nik-1}*\exp(\beta_n - \delta_{ik}) = \pi_{ni0}*\exp \sum_{j=1}^{k} (\beta_n - \delta_{ij})$ と一般化される。

カテゴリー段階 k = 1, 2, …, m により、各カテゴリー段階が選択される確率を合計すれば、

$\sum_{k=1}^{m} \pi_{nik} = \sum_{k=1}^{m} [\pi_{ni0}*\exp \sum_{j=1}^{k} (\beta_n - \delta_{ij})] = \pi_{ni0}* \sum_{k=1}^{m} [\exp \sum_{j=1}^{k} (\beta_n - \delta_{ij})]$

カテゴリー順序段階 k = 1, 2, …, m は項目群についての順序カテゴリーが 0・1・2・・・m であることを意味しているからカテゴリー選択確率の上で k = 0 をも含めて考慮すれば、$\sum_{k=0}^{m} \pi_{nik} = 1$ である。従って、第一カテゴリー0 が選択される確率は $\pi_{ni0} = 1 - \sum_{k=1}^{m} \pi_{nik}$ である。$\sum_{k=1}^{m} \pi_{nik} = 1 - \pi_{ni0}$ となるから、これを上式に代入すれば、

$1 - \pi_{ni0} = \sum_{k=1}^{m} \pi_{nik} = \pi_{ni0}* \sum_{k=1}^{m} [\exp \sum_{j=1}^{k} (\beta_n - \delta_{ij})]$

従って、$1 = \pi_{ni0}* \{1 + \sum_{k=1}^{m} [\exp \sum_{j=1}^{k} (\beta_n - \delta_{ij})]\}$

更に、$\pi_{ni0} = 1 / \{1 + \sum_{k=1}^{m} [\exp \sum_{j=1}^{k} (\beta_n - \delta_{ij})]\}$

ここで、$\pi_{nik} = \pi_{ni0}*\exp \sum_{j=1}^{k} (\beta_n - \delta_{ij})$ であるから、

$\pi_{nik} = \exp \sum_{j=1}^{k} (\beta_n - \delta_{ij}) / \{1 + \sum_{k=1}^{m} [\exp \sum_{j=1}^{k} (\beta_n - \delta_{ij})]\}$ と導かれる。

π_{nik} はカテゴリー順序段階 k = 1, 2, …, m の上での潜在的第 k 段階出現確率である。

順序カテゴリー変数 x = 0, 1, 2, …,m が選択される確率の上で、$\sum_{j=0}^{0} (\beta_n - \delta_{ij}) \equiv 0$ とみなされる(Masters[3], 1982, p. 158)。$\exp(\beta_n - \delta_{i0}) = 1$ となるから、カテゴリー変数選択確率上では、π_{nix} が次の様に導出される。

$$\pi_{nix} = \exp \sum_{j=0}^{x} (\beta_n - \delta_{ij}) / \{\exp(\beta_n - \delta_{i0}) + \sum_{k=1}^{m} [\exp \sum_{j=1}^{k} (\beta_n - \delta_{ij})]\}$$

$$= \exp \sum_{j=0}^{x} (\beta_n - \delta_{ij}) / \sum_{k=0}^{m} [\exp \sum_{j=0}^{k} (\beta_n - \delta_{ij})] \qquad x = 0, 1, \cdots, m_i$$

上式が Masters[3] (1982, p. 158)により案出された Partial Credit モデルである。以上が、Rasch 項目分析モデルに基づいて、項目個別各段階通過困難度母数推定が考慮された Partial Credit モデルへの拡張導出過程である。

各被験者得点が各被験者位置母数推定のための十分統計量になる(Masters[3], 1982, p. 159)ことは Rating Scale モデルにおいてと同様である。一方、各項目における各カテゴリー出現度数一群が一体となって各項目に関する各段階通過困難度母数推定のための十分統計量となる

(Masters[3], 1982, pp. 160−161; Wright and Masters[6], 1982, p. 59)。従って、各項目個別に各段階通過困難度推定値が与えられる点において Rating Scale モデルとは異なっている。被験者群位置母数推定値と項目個別各段階通過困難度推定値が分離独立していることは Rasch 項目分析モデルに準じている(Masters[3], 1982, p. 161)。

なお、上記の Partial Credit モデルの確率式に $\delta_{ij} = \delta_i + \tau_j$ を代入すれば、

$$\pi_{nix} = \exp\sum_{j=0}^{x}(\beta_n - \delta_{ij}) \Big/ \sum_{k=0}^{m}\left[\exp\sum_{j=0}^{k}(\beta_n - \delta_{ij})\right]$$
$$= \exp\sum_{j=0}^{x}[\beta_n - (\delta_i + \tau_j)] \Big/ \sum_{k=0}^{m}\left\{\exp\sum_{j=0}^{k}[\beta_n - (\delta_i + \tau_j)]\right\}$$
$$x = 0, 1, 2, \cdots, m$$

となり、Andrich[2] (1978a)が考案した Rating Scale モデルに単純化される(Masters[3], 1982, p. 163)。Partial Credit モデルと Rating Scale モデルのこの同質性は明確な理解の上で以下の様に示される。

前節で参照された Andrich[2] (1978a)による Rating Scale モデルの確率式は次式である。

$$p\{X = x \mid \beta, \delta, \kappa\} = \exp[\kappa_x + x(\beta - \delta)] \Big/ \sum_{k=0}^{m}\exp[\kappa_k + \kappa(\beta - \delta)]$$

なお、$\kappa_x = -\sum_{k=1}^{x}\tau_k$; $\kappa_0 = 0$　　$x = 0, 1, 2, \cdots, m$

Andrich[2] (1978a, p. 569)

上式においては、一人の被験者の一つの項目に対する一つの順序カテゴリーへの応答確率が想定されている故に、被験者と項目に添え字が付されていない。被験者 n 並びに項目 i と明記し、カテゴリー第 j 段階の潜在閾値母数を τ_j で表示する。$\kappa_0 = 0$ により存在しないカテゴリー第 0 段階での閾値母数 $\tau_0 = 0$ とみなされる。従って、上式は次式での表示に改められる。

$$p\{X = x\} = \exp\left[\tau_0 - \sum_{j=1}^{x}\tau_j + x(\beta_n - \delta_i)\right] \Big/ \sum_{k=0}^{m}\exp\left[\tau_0 - \sum_{j=1}^{k}\tau_j + \kappa(\beta_n - \delta_i)\right]$$
$$= \exp\left[-\sum_{j=0}^{x}\tau_j + x(\beta_n - \delta_i)\right] \Big/ \sum_{k=0}^{m}\exp\left[-\sum_{j=0}^{k}\tau_j + \kappa(\beta_n - \delta_i)\right]$$
$$x = 0, 1, 2, \cdots, m$$

Masters[3] (1982, p. 163)により示された $\delta_{ij} = \delta_i + \tau_j$ の変換式 $\tau_j = \delta_{ij} - \delta_i$ を上式左辺の対数項に挿入すれば、以下の様に展開される。なお、第 0 段階での $\beta_0 = \delta_0 = 0$ とみなされている。

$$-\sum_{j=0}^{x}\tau_j + x(\beta_n - \delta_i) = x\beta_n - x\delta_i - \sum_{j=0}^{x}(\delta_{ij} - \delta_i)$$
$$= \sum_{j=0}^{x}\beta_n - \sum_{j=0}^{x}\delta_i - \sum_{j=0}^{x}\delta_{ij} + \sum_{j=0}^{x}\delta_i$$
$$= \sum_{j=0}^{x}(\beta_n - \delta_{ij})$$

(Andrich[5], 2005, p. 31、参照)

同様に、右辺対数項についても $-\sum_{j=0}^{k}\tau_j + \kappa(\beta_n - \delta_i) = \sum_{j=0}^{k}(\beta_n - \delta_{ij})$ となる。又、Andrich[5] (2005)により Partial Credit モデルにおいても順序カテゴリー選択の上で 1 と 0 の生起事象が Guttman の完全順序尺度モデルに基づいていると数理展開の上でその導出が提示されている (pp. 41−44)。従って、Andrich[2] (1978a, p. 569)の Rating Scale モデルと Masters[3] (1982, p. 158)

の Partial Credit モデルは、次式に示される通りに一人の被験者 n の一つの項目 i に対する一つの順序カテゴリーx への応答確率の上では同一物である(Andrich[5], 2005, pp. 30−31; Masters[3], 1982, p. 163)。

$$p\{X = x\} = \exp[-\sum_{j=0}^{x}\tau_j + x(\beta_n - \delta_i)] \Big/ \sum_{k=0}^{m}\exp[-\sum_{j=0}^{k}\tau_j + \kappa(\beta_n - \delta_i)]$$
$$= \exp\sum_{j=0}^{x}(\beta_n - \delta_{ij}) \Big/ \sum_{k=0}^{m}[\exp\sum_{j=0}^{k}(\beta_n - \delta_{ij})] \qquad x = 0, 1, \cdots, m_i$$

4. Many-Facet Rasch Measurement 解題

アメリカ人である John M. Linacre が 1987 年に‘An Extension of the Rasch Model to Multi-facet Situations’と題する論文をシカゴ大学の教育学部で発表して、1988 年にそのコンピュータプログラムを開発した(Linacre[11], 1989, p. 135、参照)。Linacre の基本的着想は Andrich[2] (1978a)によって考案された Rating Scale モデルにおける被験者位置母数と項目位置母数に加えて別の母数変数挿入という Rasch 拡張モデル化であり、その数理展開は以下の通りである。

第二節で解題された如く、Andrich の Rating Scale モデルは次式で表わされる。

$$p\{X = x \mid \beta, \delta, \kappa\} = \exp[\kappa_x + x(\beta - \delta)] \Big/ \sum_{k=0}^{m}\exp[\kappa_k + \kappa(\beta - \delta)]$$

なお、$\kappa_x = -\sum_{k=1}^{x}\tau_k$; $\kappa_0 = 0$ $\qquad x = 0, 1, 2, \cdots, m$

Andrich[2] (1978a, p. 569)

上式において(β−δ)に修正項を加えることは可能である(Andrich[24], 1997, p. 879)故に、添え字が付された$(\beta_n - \delta_i)$に$(-\lambda_j)$を加える。

$$\pi_{xnij} = \exp[\kappa_x + x(\beta_n - \delta_i - \lambda_j)] \Big/ \sum_{k=0}^{m}\exp[\kappa_k + \kappa(\beta_n - \delta_i - \lambda_j)]$$
$$= \exp[-\sum_{k=0}^{x}\tau_k + x(\beta_n - \delta_i - \lambda_j)] \Big/ \sum_{k=0}^{m}\exp[-\sum_{k=0}^{k}\tau_k + \kappa(\beta_n - \delta_i - \lambda_j)]$$

但し、$\kappa_0 = 0$ $\qquad x = 0, 1, 2, \cdots, m$

Linacre[11] (1989, p. 54)が与える表示に変更する。

$$P_{nijk} = \exp[-\sum_{s=0}^{k}F_s + k(B_n - D_i - C_j)] \Big/ \sum_{h=0}^{K}\exp[-\sum_{s=0}^{h}F_s + h(B_n - D_i - C_j)]$$
$$= \exp[k(B_n - D_i - C_j) - \sum_{s=0}^{k}F_s] \Big/ \sum_{h=0}^{K}\exp[h(B_n - D_i - C_j) - \sum_{s=0}^{h}F_s]$$

但し、$F_0 \equiv 0$ $\qquad k = 0, 1, 2, \cdots, K$

更に、プログラム出力尺度原点局所化必要性で以って一般的には受験者位置母数推定値が尺度固定化上での固有対象とされて、$\Sigma D_i = 0$、$\Sigma C_j = 0$、$\Sigma F_k = 0$ と設定される(Linacre[11], 1989, pp. 54−55)。

これが、Many-Facet Rasch Measurement を代表する共通段階(common step)モデル(Linacre[11], 1989, p. 58)であり、帰納的観点の上で母数推定値の安定性と再現性において最も普遍度の高いモデル(Linacre[11], 1989, p. 98)とされている。C_jを、例えば、面接官 j の評定難度とすれば、上式は、面接テストにおいて受験者 n による項目 i への応答に対して面接官 j による評価として評定段階 k が選択される確率をモデル化したものである。この確率式が、Linacre[11]

(1989)により Rasch モデルの根幹規定である構成要素単位上での生起度数比率同一性と尺度基準化概念を特徴とする「固有客観性」に基づくものとして導出されている(pp. 47–52)。各受験者総点一群、各項目総点一群、各面接官総点一群がそれぞれ独立に各受験者位置母数、各項目位置母数、各面接官位置母数を推定する上での十分統計量であり(Linacre[11], 1989, p. 4)、それら三種類の母数推定値が独立に同一間隔尺度ロジット単位上で推定される(Linacre[11], 1989, pp. 1–2)。更に、上掲の共通段階モデルについては、各カテゴリー出現度数一群によりカテゴリー数より一つ少ない個数の閾値が推定されて、全項目に渡って共通の閾値が各カテゴリー進行段階に異なるものとして与えられ、Rating Scale モデルについてと同様に隣接カテゴリー間での生起確率境界値となる。なお、この閾値は Many-Facet Rasch Measurement においては段階困難度推定値(step difficulty estimate)と呼ばれ、他の母数推定値とは推定種類上での性質が異なっており、各受験者能力推定値、各項目困難度推定値、各面接官評定難度推定値のすべてから独立しているというものではない(Linacre[11], 1989, p. 77)が、他の母数推定値との関連依存性は無視されて構わない程度に微小であると Linacre[11] (1989, p. 77)は述べている。上記説明便宜上 C_j を面接官 j の評定難度としたが、それに限られるものではない。共通段階モデルは、あらゆる種類の受験者課題達成度に対して複数評定者人員により与えられる順序段階評定、又は、一人の評定者による数種類の課題範疇別に基づく項目群分類設定上での順序段階評定に適用され得るものである。

以下に、Many-Facet Rasch Measurement の上掲共通段階モデル指数関数表示の対数オッズを参照する。

$$P_{nijk} = \exp[k(B_n - D_i - C_j) - \sum_{s=0}^{k} F_s] \;/\; \sum_{h=0}^{K} \exp[h(B_n - D_i - C_j) - \sum_{s=0}^{h} F_s]$$

但し、$F_0 \equiv 0$ $\qquad k = 0, 1, 2, \cdots, K$

その P_{nijk-1} との対数オッズをとれば、

$$\ln(P_{nijk} / P_{nijk-1}) = \ln P_{nijk} - \ln P_{nijk-1}$$

$$= \ln\{\exp[k(B_n - D_i - C_j) - \sum_{s=0}^{k} F_s] \;/\; \sum_{h=0}^{K} \exp[h(B_n - D_i - C_j) - \sum_{s=0}^{h} F_s]\}$$
$$- \ln\{\exp[(k-1)(B_n - D_i - C_j) - \sum_{s=0}^{k-1} F_s] \;/\; \sum_{h=0}^{K} \exp[h(B_n - D_i - C_j) - \sum_{s=0}^{h} F_s]\}$$
$$= \{\ln \exp[k(B_n - D_i - C_j) - \sum_{s=0}^{k} F_s] - \ln \sum_{h=0}^{K} \exp[h(B_n - D_i - C_j) - \sum_{s=0}^{h} F_s]\}$$
$$- \{\ln \exp[(k-1)(B_n - D_i - C_j) - \sum_{s=0}^{k-1} F_s] - \ln \sum_{h=0}^{K} \exp[h(B_n - D_i - C_j) - \sum_{s=0}^{h} F_s]\}$$
$$= \ln \exp[k(B_n - D_i - C_j) - \sum_{s=0}^{k} F_s] - \ln \exp[(k-1)(B_n - D_i - C_j) - \sum_{s=0}^{k-1} F_s]$$
$$= [k(B_n - D_i - C_j) - \sum_{s=0}^{k} F_s] - [(k-1)(B_n - D_i - C_j) - \sum_{s=0}^{k-1} F_s]$$
$$= B_n - D_i - C_j - F_k$$

これが Linacre[11] (1989)の対数オッズ表現式による共通段階モデル(p. 1)である。C_j を面接官 j の評定難度とすれば、面接テストにおいて受験者 n による項目 i への応答に対して面接官 j による評定として評定値(k − 1)ではなく評定値 k が選択される確率の対数オッズとして示されている。それはロジット単位上で(受験者 n の能力位置母数 − 項目 i の困難度位置母数 − 面接官 j の評

定難度位置母数 − 評定値k− 1 から評定値kへの評定到達困難度母数)として表現されることになる。Many-Facet Rasch Measurement でのその対数オッズ表現式がロジット単位上で構成要素変数の和により示されることはそれらの変数が全体として一つの間隔尺度次元を構成していることになり(Linacre[11], 1989, p. 2)、Rasch モデルが普遍的に規定する母数結合一次元性を意味する。データの共通段階モデルへの低い適合度は、受験者群による項目群への応答に対して面接官全員により付与された順序評定値の全体的増加が受験者群能力高位度と項目群困難度を全体的にそれぞれの一方向性で以って反映していないことになり、データの潜在特性一次元性充足度欠如との解釈になる(Linacre[11], 1989, p. 43)。それは、潜在特性に関する項目順序評定値の上で受験者群、項目群、面接官群の三要素全体として一つの面接評価構成概念妥当性が低いとの明示でもある。

なお、以下の単純操作により、Many-Facet Rasch Measurement を代表する上掲の共通段階モデル(Linacre[11], 1989, p. 58)が Rasch モデルの一般規定式であることが分かる。

1) 対数オッズ表現式による共通段階モデル式から C_j を削除して、項目 i が被験者群によって、あるいは、被験者群の項目 i への応答に対して一人の評定者によって、順序評定 k ＝0, 1, 2, …, kから一つ選択されるとすれば、Andrich[2] (1978a)によって考案された Rating Scale モデルの対数オッズ表現式となる。

$$\ln(P_{nik} / P_{nik-1}) = B_n - D_i - F_k$$

2) 更に、項目 i が 0 又は 1 で応答される場合には $F_1 = 0$ と定義されている(Linacre[11], 1989, p. 77)ことにより、Rasch 項目分析モデルの対数オッズ表現式となる。

$$\ln(P_{ni1} / P_{ni0}) = B_n - D_i$$

共通段階モデルにおいては全項目に渡って共通段階 k への同一到達困難度推定値が各段階に付与されるが、Many-Facet Rasch Measurement 上で以下の様な異種モデル設定が可能とされている(Linacre[11], 1989, p. 58–59)。

1) 前節で解題された Partial Credit モデルの様に項目間で段階 k への到達困難度構造が異なると想定されれば、次式で表現される項目段階(item-step)モデルとなる。

$$\ln(P_{nijk} / P_{nijk-1}) = B_n - D_i - C_j - F_{ik}$$

2) 面接官の間で段階 k への到達困難度構造が異なると想定されれば、次式で表現される面接官段階(judge-step)モデルとなる。

$$\ln(P_{nijk} / P_{nijk-1}) = B_n - D_i - C_j - F_{jk}$$

3) 各面接官が各項目について段階 k への異なる到達困難度構造を付与すると想定されれば、次式で表現される面接官項目段階(judge-item-step)モデルとなる。

$$\ln(P_{nijk} / P_{nijk-1}) = B_n - D_i - C_j - F_{ijk}$$

4) 何らかの基準設定の上で項目群あるいは面接官群に関する段階評定尺度が数種類に範

疇化される(g = 1, 2, …, G)と想定されれば、次式で表現されるグループ段階(group-step)モデルとなる。

$$\ln(P_{nijk} / P_{nijk-1}) = B_n - D_i - C_j - F_{gk}$$

5) 共通段階モデルに前もって配慮・準備された数種類の課題範疇別項目群分類設定により課題範疇別困難度変数をも付け加えることが可能であり、A_m が課題 m (= 1, 2, …, M)の困難度母数推定値を指すものとすれば、そのモデルは次式で表現される(Linacre[11], 1989, p. 3)。

$$\ln(P_{nmijk} / P_{nmijk-1}) = B_n - A_m - D_i - C_j - F_k$$

適用モデル選択は研究対象への理論的考察に基づく先験的に妥当な調査実施要領考査の上で事前に決定されるのが常ではあるが、順序尺度データへの Many-Facet Rasch Measurement 適用上での測定各構成要素のモデル適合度事後検証により研究上の推察に最も益する分析モデルの決定となることもある。又、調査実施要領の不備に伴う諸要因考慮の上で不適合項目削除、不適合評定者除去等、あるいは、順序評定カテゴリー最適構成(Fox and Jones[25], 1998, pp. 38−41)により変更を加えたデータの測定結果分析をも参照された上での最良分析モデル選択になることも考えられる。但し、Linacre[11] (1989)は、母数推定値の安定性と再現性において普遍度の最も高い共通段階モデル(p. 98)の元データへの適用に比べて、不適合項目削除等のデータ改変による再度の共通段階モデル測定ないしは Many-Facet Rasch Measurement 上での他のモデル適用がデータのそのモデル適合度を増大するとしてもそれにより得られる母数推定値の恒常不変的再現度は減少する(p. 2)、と指摘している。Linacre[11] (1989, p. 2)によるこの見解は、推定母数変数の少ない Rating Scale 測定が推定母数変数の多い Partial Credit 測定より以上に安定した再現性の高い母数推定値と質問紙調査構成概念について有意義な推察をもたらす(Linacre[26], 2000, p. 768)との彼の指摘に呼応するものである。更に、データ改変については、『モデル不適合とみなされた個々のはずれ値を削除することは、Rasch モデル推定にとって本質的に重要であるデータ生起の無作為性構造を変質させる結果となる』との Linacre[27] (1990, p. 80)による言及に留意される。なお、Rating Scale と Partial Credit との間での測定モデル選択に関して Linacre (2000[26], p. 768)と同様な見解の持ち主である Wright[28] (1998)により引用された Ockham's Razor は分析に使用する変数の数への思慮の上で銘記に値する。それは『より少ない前提により説明され得ることはより多い前提により説明不能となる』(Wright[28], 1998, p. 642)というものである。

評定尺度構成において本質的に重要なことは、評定者間での何らかの順序評定識別合意内容に基づいて面接官等の受験者能力順序評定者各個により意識峻別された順序識別基準に沿って設定評定数内で個人的に一貫して適切に可能な限り多くの異なる評定値が使用される必然性である。各受験者による各項目への応答に対して現実に履行され得ないことであるとしても評定者全員が同一の順序評定結果を目指すことではなく、各評定者が個別独立に受験者全員による各項目への

応答に対して納得合意された順序評定識別基準に徹底一貫して沿う各自評定結果を目指すことに評定尺度構成の本質的有用性が存在する。Linacre[11] (1989)による上記の観点(pp. 10−12)に基づき、『評定者間での評定値変動は、Many-Facet Rasch Measurement の上では除去されるべきものとはならず、むしろ[普遍性のある]評定尺度構成への必要成分とみなされる』と Linacre[11] (1989, p. 21)は述べている。Rasch モデルの属性である「固有客観性」が受験者能力、項目困難度、評定者難度を分離することにより、Many-Facet Rasch Measurement が各構成要素の母集団において一般普遍性の高い評定尺度構成に貢献する(Linacre[11], 1989, p. 20)とされている。

受験者能力測定精度の観点からは、潜在特性一次元性と局所独立性の充足度の高い出来得る限り多くの項目について可能な限り多くの信頼度の高い評定値多使用・評定一貫性保持評定者による順序評定機会の確保が期待されるが、現実には、時間、労力、そして対価が係わる。例えば、受験者 32 名各自により書かれた三種類の小論文すべてに対して評定者 12 名が順序評定するならば、各評定者は 96 篇の小論文を順序評定することになり、評定実施要領の点でこの現況下での最も精度の高い受験者能力測定となる。現実的にこの余裕がない場合として、測定精度は当然に下がるけれども Many-Facet Rasch Measurement の適用要領上で正当に可能とする評定者 12 名による受験者 32 名の小論文 96 篇に対する次の様な評定配置が Linacre (1989)により提案されている(pp. 14−15)。

1） 各受験者の三つの小論文それぞれが異なった評定者によって順序評定される。

2） 各評定者は異なった受験者 8 名による小論文 8 篇を順序評定する。

3） 各評定者が順序評定する小論文 8 篇は三種類の論文数比として 3:3:2 となる。

これが「連鎖」(linkage)の一例であり、この連鎖への配慮は Linacre[11] (1989)によって開発された Facets (Linacre[18], 1989-2001, Winsteps)へのデータ入力必須重要事項であり、各受験者、各項目、各評定者、並びに、各評定段階閾値に関する母数推定値が確定されるためには必要最小限以上の連鎖設定が不可欠とされる(Linacre[11], 1989, p. 13)。なお、上記の様な Linacre (1989[11], pp. 14−15)による効率の良い評定者・被験者・項目評定配置案を用いた Many-Facet Rasch Measurement の成功例が Lunz, Wright, and Linacre[29] (1990)によって報告されている。

5. Rasch 評定尺度モデル閾値順序性

第二節での解題から理解される様に、Andrich[2] (1978a)の Rating Scale モデルはガットマン完全順序尺度を構成要素としており、閾値母数の順序性 $\tau_1 < \tau_2 < \tau_3$, …, $< \tau_m$ がそのモデル尺度上で規定されている。第三節解題での Masters[3] (1982)の Partial Credit モデルにおいても同様であり、各項目 i に関する閾値母数の順序性 $\delta_{i1} < \delta_{i2} < \delta_{i3}$, …, $< \delta_{im}$ がその尺度上で規定されており(p. 157、参照)、それは第三節最後尾に参照された Partial Credit モデルの Rating Scale モデルとの同質性(Andrich[5], 2005, pp. 30−31; Masters[3], 1982, p. 163)によっても明らか

である。更に、第四節解題での Linacre[11] (1989)の Many-Facet Rasch Measurement においても、筆者によって Linacre[11] (1989)にその言明は見出されないけれども、その原理は Andrich[2] (1978a)の Rating Scale モデルに基づく故に、又、Linacre[30] (2002)への参照で以って、閾値母数順序性規定が同様であると理解される。

Rasch 評定尺度モデル規定としての潜在閾値母数のこの順序性は人間科学研究者の項目群順序尺度設定への意識上の基にあることが容易に察知される。例えば、1. 大反対　2. 少々反対　3. 少々賛成　4. 大賛成　という設定尺度上での調査質問紙集計時における項目群順序カテゴリー1、2、3、4 それぞれへの配点を 1・2・3・4 あるいは 0・1・2・3 とすること自体にその閾値順序性への潜在認識が存在する。これは、何らかの構成概念について統計的に分析検証する順序尺度設定において、カテゴリー1 から 4 への項目群全体としてのその一元的潜在特性上での強度増大性に関して先験性に基づく研究考察仮定が内在していることを意味する(Andrich[5], 2005, pp. 27–28、参照)。従って、順序尺度データへの Rasch 評定尺度モデル測定適用後の閾値母数推定値順序性は設定項目群全体としての一元的構成概念上でのカテゴリー順序段階尺度化妥当性についての最優先検証事項とされる(Andrich[8], 2004, pp. 189−192; Andrich[5], 2005, p. 49、参照)。

図 1 は Rasch 評定尺度モデル測定の上でカテゴリー特性曲線(Category Characteristic Curves, e.g., in Andrich[2], 1978a, p. 571)ないしは尺度構造確率曲線(Scale Structure Probability Curves, e.g., in Linacre, 1989−2001[18], p. 87)と呼ばれるものの例示である。それは、右脳と左脳の機能区別化に基づく筆者作成による性格二極分別性を一元的潜在特性とする 15 項目から成る四段階順序尺度質問紙(*付録)調査データの Facets (Linacre[18], 1989-2001, Winsteps)による Rating Scale モデル測定出力である。そのデータは 2004 年度と 2005 年度での流通科学大学外国語センターによる特別講義「言葉と文化」上での筆者担当授業受講者(調査研究承諾者)合計 709 名への質問紙調査結果である。以降の主成分分析適用をも踏まえて、質問紙調査研究承諾者総数 709 名の中には満点被験者 1 名、及び、少なくとも一項目への無応答存在被験者 16 名が含まれている故に、欠損値個数を 0 とする完全有効被験者数 692 名から成るデータを分析対象としている。

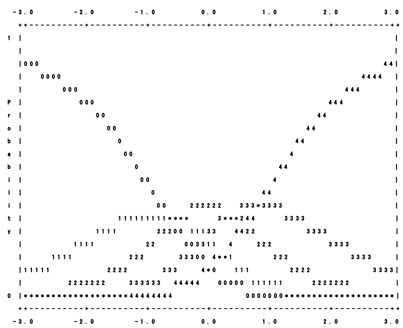

図 1　　15 項目から成る四段階順序尺度質問紙への 692 名応答データの Rating Scale モデル測定 Facets 出力上でのカテゴリー特性曲線

上図からは判別し難いけれども、Facets により Rating Scale モデル測定上での閾値母数推定値として $\tau_1 = -0.68$、$\tau_2 = -0.48$、$\tau_3 = 0.37$、$\tau_4 = 0.79$ が各数値順次増大性で以って出力されている。その測定出力上での最優先確認事項として筆者作成当該調査質問紙に関する項目群全体としての先験的カテゴリー順序四段階尺度化設定について低いとは言えない妥当性が窺われる。上記四つの閾値母数推定値により Rating Scale モデル測定 Facets 出力上でその平均を 0 とする設定(Andrich[5), 2005, p. 29、参照)が確認される。ちなみに、上記四つの閾値母数推定値と 15 項目位置母数推定値及び被験者位置母数推定値との関係は次の様に解釈される(Masters and Wright[31)], 1982, p. 24、参照)。

１） 各項目位置母数推定値マイナス 0.68 未満の被験者位置母数推定値を持つ当該データ母集団被験者群は当該質問紙調査各対応項目に対して第一カテゴリーを選ぶ確率が最も高い。

2） 各項目位置母数推定値マイナス 0.68 から各項目位置母数推定値マイナス 0.48 までの範囲内にある被験者位置母数推定値を持つ当該データ母集団被験者群は当該質問紙調査各対応項目に対して第二カテゴリーを選ぶ確率が最も高い。

3） 各項目位置母数推定値マイナス 0.48 から各項目位置母数推定値プラス 0.37 までの範囲内にある被験者位置母数推定値を持つ当該データ母集団被験者群は当該質問紙調査各対応項目に対して第三カテゴリーを選ぶ確率が最も高い。

4） 各項目位置母数推定値プラス 0.37 から各項目位置母数推定値プラス 0.79 までの範囲内にある被験者位置母数推定値を持つ当該データ母集団被験者群は当該質問紙調査各対応項目に対して第四カテゴリーを選ぶ確率が最も高い。

５） 各項目位置母数推定値プラス 0.79 を超える被験者位置母数推定値を持つ当該データ母集団被験者群は当該質問紙調査各対応項目に対して第五カテゴリーを選ぶ確率が最も高い。

表 1 当該データへの Partial Credit モデル測定 Facets 出力上での閾値母数推定値とその順序性

	項目番号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	平均値
	Rating 位置母数推定値	0.37	−0.13	0.19	0.53	−0.12	−0.05	0.58	−0.87	0.06	0.26	0.67	−0.64	−0.27	0.04	−0.61	0.00
	Partial 位置母数推定値	0.41	−0.26	0.14	0.48	−0.20	−0.05	0.51	−0.67	0.08	0.21	0.62	−0.53	−0.24	0.00	−0.50	0.00
		閾値母数推定値															
	閾値順序性	○	○	○	X	○	○	○	X	X	X	X	X	X	X	X	*
項目カテゴリー段階	1	−1.05	−1.13	−0.93	−0.62	−0.86	−0.76	−0.48	−0.29	−0.41	−0.94	−0.74	−0.19	−0.31	−0.44	−0.11	−0.62
	2	−0.57	−0.28	−0.69	0.03	−0.49	−0.42	−0.21	0.08	−0.71	−0.88	0.02	−0.51	−0.59	−1.09	−0.67	−0.47
	3	0.30	0.65	0.66	−0.22	0.62	0.08	0.08	−0.23	0.31	0.95	−0.19	0.23	0.35	0.86	−0.12	0.29
	4	1.32	0.76	0.96	0.80	0.73	1.10	0.61	0.43	0.82	0.87	0.91	0.47	0.54	0.67	0.90	0.79
	点双列相関係数	0.26	0.28	0.21	0.15	0.46	0.19	0.23	0.25	0.33	0.15	0.18	0.26	0.46	0.18	0.28	0.26

表 1 は、当該データへの Partial Credit モデル測定適用上での Facets 出力による 15 項目それぞれについての各閾値母数推定値とその順序性を示すものである。記号○はモデル規定としての閾値母数順次増大性に適合している項目であり、×は閾値母数推定値が少なくとも一つのカテゴリー順序段階においてモデル規定としての閾値母数順次増大性に反している項目である。なお、参照までに、Rating Scale モデルと Partial Credit モデルそれぞれによる測定出力上での各項目位置母数推定値、及び、簡易な項目弁別力指標として素点上での SPSS (Version 15.0, SPSS Inc., 2006)出力に基づく各項目点双列相関係数をも併せて表提示に含めている。

Facets による Partial Credit モデル閾値母数推定値出力は、第三節で参照された項目 i についてのカテゴリー第 j 段階での δ_{ij} そのものではなく、$\delta_{ij} = \delta_i + \tau_{ij}$ として項目 i の位置母数推定値 δ_i と項目 i についてのカテゴリー第 j 段階での閾値母数推定値 τ_{ij} への分離出力であり、そして、δ_i は δ_{ij} の平均値として求められている(Andrich[5), 2005, p. 31、参照)。Partial Credit モデルと Rating Scale モデルそれぞれによる測定出力上での項目群位置母数推定値同等性が観察されて、15 項目に関する両者の値におけるピアソン相関係数とケンドール順位相関係数タウ b はそれぞれ 0.988、0.981 である。又、Partial Credit モデルと Rating Scale モデルによる測定出力上での被験者群位置母数推定値の平均はそれぞれ 0.196、0.224、その標準偏差はそれぞれ 0.435、0.437、そして、両者推定値群の間での相関係数は 1 である。これにより、両モデル測定出力間での各項目と各被験者に関するいずれもの位置母数推定値はほぼ同等であると判断される。

更に、Partial Credit モデル測定上での 15 項目各閾値母数推定値平均が上に参照した Rating Scale モデル測定上での閾値母数推定値 $\tau_1 = -0.68$、$\tau_2 = -0.48$、$\tau_3 = 0.37$、$\tau_4 = 0.79$ のそれぞれにほぼ相応していることが表 1 から窺い知り得る。これは、一人の被験者の一つの項目に対する一つの順序カテゴリーへの応答確率の上では両モデルは同一物であり(Andrich[5)], 2005, pp. 30−31; Masters[3)], 1982, p. 163)、それぞれのモデル上での閾値母数推定値総計が 0 (Andrich[5)], 2005, p 29−30、参照)とのモデル測定出力設定からも容易に察知される。従って、両モデル測定出力間での各項目と各被験者に関して上述されたいずれもの位置母数推定値同等性をも踏まえれば、両モデル測定 Facets 出力上での違いは閾値母数推定値付与の出力対象が項目群全体か各個別項目かの唯一点のみに限られている。これは、すでに解題された両モデルそれぞれの特徴を端的に示すものである。なお、『同一の順序カテゴリー構成から成る項目群を持つ質問紙調査において Partial Credit モデルを適用することは Rating Scale モデルによるものとは異なる推察に導く上での項目間での潜在的閾値異同尺度設定への統計的及び実質的な強い根拠を必要とする』との Wright[25)] (1998, p. 642)による現実適用モデル選択上での先験的理念立脚点意識化重要性への指摘には留意される。

Rating Scale モデル測定上では 15 項目全体としての同一四段階順序尺度化妥当性は否定されなかったけれども、Partial Credit モデル測定上で当該データに含まれた 15 項目中の 9 項目が四

段階順序尺度化妥当性に欠けるとの表 1 による示唆である。Wright[25] (1998, p. 642)による上記の両モデル適用間での潜在的順序カテゴリー尺度化設定上での先験理念的相違があるとしても、Rating Scale のみならず Partial Credit によるモデル測定出力上での項目個別閾値母数推定値順序性への参照有意義性が顕示されている。要するに、当該質問紙調査における項目群の意味内容と四段階順序尺度化設定については潜在特性一次元性ないしは構成概念妥当性の上で大きな問題が存在すると意識される。

順序尺度素点データへの因子分析はその適用妥当性に欠ける(Linacre[32], 1998a, p. 603; Wright[33], 1996, p. 509)とされた上で、データ局所独立性充足程度の確認法として Wright[33] (1996)により Rasch 測定ロジット残差主成分分析の有用性が示された。Wright[33] (1996)が唱えたその Rasch 測定ロジット残差主成分分析よりも分析結果明瞭性の点で僅かにしても優る Rasch 測定標準化残差主成分分析適用が Linacre[34] (1998b)により実証され、データ一次元性検知法としてのその有効性(Linacre[34], 1998b)と共にその理論的適用根拠が更に言及されている(Linacre[35], 1998c, p. 636)。その方法は、現在では識者によりデータ局所独立性充足度確認法というよりもむしろ特異項目機能(DIF)察知を含めた全般的データ一次元性検証法として位置付けられている(Wright and Stone[36], 2004, p. 21)。但し、0・1 データ Rasch 測定の上で局所独立性と潜在特性一次元性の間には表示実体としての異質性が存在しており、それぞれについての充足度の間に大きな乖離がデータ次第で起こり得ることは井澤[37] (2006)によって検証された通りである。然しながら、局所独立性と潜在特性一次元性のデータ充足度を観る上でその方法の有用性は今のところ識者によっては否定されていない(Tennant and Pallant[38], 2006、参照)。従って、Rating Scale モデル測定上での閾値母数推定値順序性が認められる当該データについての潜在特性一次元性と局所独立性の充足度を Linacre (1998b[34] and 1998c[35])が推奨する Rasch 測定標準化残差主成分分析によって以下に査察する。

表 2　Rasch 測定標準化残差主成分分析固有値情報

Component	Initial Eigenvalues Total	% of Variance	Cumulative %
1	1.905	12.702	12.702
2	1.512	10.080	22.782
3	1.389	9.259	32.041
4	1.258	8.390	40.431
5	1.157	7.715	48.146
6	1.045	6.965	55.111

Extraction Method: Principal Component Analysis.

表 3　第 1 主成分負荷量

	Component 1
1	-.452
2	.181
3	.298
4	-.176
5	-.322
6	.257
7	-.537
8	.367
9	-.624
10	.169
11	-.165
12	.371
13	-.118
14	.475
15	.366

Extraction Method: Principal Component Analysis.

表 2 と 3 が当該データ Rasch 測定標準化残差主成分分析上でのそれぞれ第 6 主成分までの固有値情報と 15 項目第 1 主成分負荷量一覧であり、その出力は Rasch 測定標準化残差上での SPSS による項目間ピアソン相関係数に基づくものである。線形関係にはない素点データから Rasch 一次元性が抽出された残量としての Rasch 次元では説明されないデータ構成成分が少なくとも素点データ以上に線形関係を有する標準化残差上で Rasch モデル適合理想として誤差変動内で提示されるとの考え方である(Linacre[35], 1998c, p. 636; Wright[33], 1996, pp. 509−510)。当該データに関しては、1 以上の固有値を持つ Rasch 測定標準化残差主成分が六つもあり、第 3 主成分までの累積寄与率だけでも 32%に達している。要するに、当該データ Rasch 測定標準化残差が誤差変動を大きく超える構成成分残量を保持しており、その成分残量が多次元性を有していると判断される。これは、表 3 での 15 項目第 1 主成分負荷量の Rasch モデル適合理想値としての 0 (Linacre[39], 1999a, p. 710、参照)から大きく遊離したかなり顕著な正負二極分化傾向によっても明示されている。従って、当該データのモデル適合度は低いと帰結され、Rasch 一次元尺度測定に不備がある故に、潜在特性一次元性に欠けるデータであると判断される。なお、本稿の考察目的ではなく、即断は禁物であるとしても、調査・分析方法をより精密にすれば、俗に言う「右脳型」・「左脳型」人間への二極分別には疑問が呈されることになる筈である。

表 4　被験者群 690 名位置母数推定値不変性成立程度（項目群一次元性成立程度）

項目群折半法	有効被験者数	Pearson	Kendall's tau-b
正の負荷量8項目　対　負の負荷量7項目	690	0.242	0.161

表 5　15 項目位置母数推定値不変性成立程度（被験者群一次元性成立程度）

受験者群折半法	有効項目数	Pearson	Kendall's tau-b
正の負荷量360名　対　負の負荷量332名	15	0.722	0.505

当該素データ上での項目群に関する Cronbach のアルファ係数は SPSS により 0.639 と出力されて、15 項目の内的一貫性程度は高いものではないと示唆されている。又、表 3 での第 1 主成分負荷量記号正負分別に基づく項目群折半法の上での被験者群位置母数推定値不変性成立程度を示すものが表 4 である。なお、負の負荷量 7 項目データにおいて満点と 0 点の被験者がそれぞれ一名ずつ存在することから、有効被験者数が 690 名となっている。表 4 におけるいずれの相関係数も大変に低い値であり、これによっても当該データ 15 項目の潜在特性一次元性充足度の低さが顕示されている。更に、当該データ Rasch 測定標準化残差の被験者群についての主成分分析第 1 主成分負荷量記号正負分別に基づく項目群位置母数推定値不変性成立程度を示すものが表 5 である。表 5 からは当該データ被験者群の質問紙応答傾向一次元性充足度も高いとは言えないと判断される。

潜在特性一次元性充足度の低さと不可分の関係にある 15 項目局所従属性の形状と程度を俯瞰するものが図 2 である。縦軸と横軸それぞれに当該データの Rating Scale モデル測定上での標準化残差項目群主成分分析第一主成分負荷量と項目群位置母数推定値がとられており、Wright[33] (1996)により与えられた図(p. 509)に対応するものである。

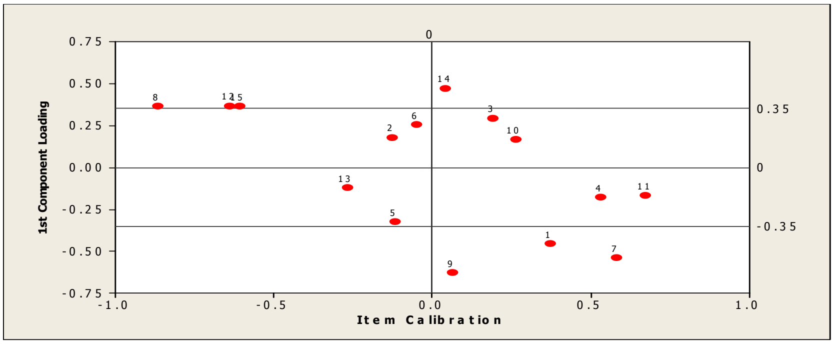

図 2　Rating Scale モデル測定標準化残差項目群主成分分析
第 1 主成分負荷量 対 項目群位置母数推定値

図 2 に引かれた参照線 y = 0.35 と y = − 0.35 それぞれの境界線の上部と下部にある一群がそれぞれ相対的に強い局所従属関係にある項目群であると解釈される(Wright[33], 1996)。その基準に従えば、項目番号 8、12、14、15 が相対的に強い局所従属関係にあり、又、項目番号 1、7、9 が前者と反対傾向軸上での相対的に強い局所従属関係にある。前者の一群に含まれている項目は前掲表 1 に与えられた閾値母数推定値順序性違反 9 項目群に属しており、後者の一群においては項目番号 9 以外の 1 と 7 が表 1 に与えられた閾値母数推定値順序性準拠 6 項目群に属している。項目番号 9 の第 1 主成分負荷量は前掲表 3 に−0.624 と与えられており、閾値母数推定値順序性違反 9 項目群における項目番号 9 の異端性が観察される。但し、表 3 あるいは図 2 での第 1 主成分負荷量正負記号分別と表 1 での Partial Credit モデル測定上での閾値母数推定値順序性準拠・違反のそれぞれに基づく項目群構成の間には大きな異同があることから、両者が与える情報は同一のものではないと認識される。

本節考察により帰結される最重要事項は、Rating Scale モデル閾値母数推定値順序性は質問紙項目群カテゴリー順序段階尺度化妥当性検証への必要条件ではある(Andrich, 1978a[2], 2004[8], and 2005[5])が、潜在特性一次元性と局所独立性を併せた構成概念妥当性を保証する充分条件ではないということである。又、Rating Scale モデル適用上での閾値母数推定値順序性準拠データの Partial Credit モデル測定で以って、過半数の項目群に閾値母数推定値順序性への違反が観察されることはあり得る。前掲表 1 について理解される様に、Partial Credit モデル測定 Facets 出力上での項目群各閾値母数推定値平均が Rating Scale モデル測定上での各閾値母数推定値に相当する。従って、Rating Scale モデル測定上での隣接閾値母数推定値間の差が小さくなればなる程、

Partial Credit モデル測定上での各項目閾値母数推定値順序性違反の生起確率が高くなると推測される。

Rating Scale モデル測定上での隣接閾値母数推定値間の差については、Linacre[30] (2002)により、三つのカテゴリーから成る二段階順序尺度構成データと五つのカテゴリーから成る四段階順序尺度構成データにおける隣接閾値母数推定値間での理想最小値はそれぞれ 1.4 ロジット値、1.0 ロジット値である(pp. 102−103)と指摘されている。それぞれの段階順序尺度上でのその値は、Rating Scale モデルにおける一つ手前のカテゴリーが選択された結果として次のカテゴリーが選択されるか否かという一連の二者択一における理論上の隣接閾値間理想最小差である(pp. 102−103)と説明されており、5.0 ロジット値を超えないものとの指摘(p. 103)も付加されている。又、『Rating Scale モデル測定上での閾値母数推定値順序性違反は、調査質問紙上で提示された順序カテゴリーの順序評定定義の逆転を意味するものではなく、隣接順序カテゴリー間での識別不明瞭性あるいは解釈定義狭小性の故にモデルが期待する程には被験者群に選択されていないカテゴリーの存在を示唆している』と Linacre[30] (2002, p. 98)により説明されている。

Rating Scale モデルにおいて項目弁別力は全項目に渡って同一であると規定されており、閾値母数での同一弁別性で以って順序カテゴリー間における弁別母数を同値 1 として各閾値が全項目に共通のものとして付与される(Andrich[2], 1978a, pp. 568−569)。一方、順序カテゴリー間における弁別力同一性は同様である(豊田[40]、2002, p. 105、参照)けれども、Partial Credit モデルにおいては項目間での弁別力同一性が規定されていない(Andrich[6], 1988, pp. 369−370)。その結果として、Partial Credit モデル測定出力上で各閾値が各項目にそれぞれ異なるものとして付与される。『閾値母数推定値間での差の数値大小性が項目弁別力に影響する』と Andrich[41] (1988, p. 370)によって指摘され、項目弁別力の異同程度が各項目に付される各閾値母数推定値の数値異同程度に反映される。Andrich[41](1988, p. 370)によるこの事象についての更なる言及を以下に引用する。

1)『閾値母数推定値の数値増大性が保持されていることを以ってして、その値の差が大きくなればなる程、その項目弁別力は順じて小さくなる。』

2)『閾値母数推定値間の差が小さくなればなる程、その項目弁別力は順じて大きくなる。』

3)『閾値母数推定値順序性に違反があれば、その項目弁別力は更に大きくなる。』

Linacre[42] (1999b)も、同様に、『閾値順序性への違反は、その項目が潜在特性に関する限られた範囲の上で過度に弁別的であることの示唆である』(p. 675)と述べている。なお、筆者による今後の考察課題の一つとして、上記引用 3) 閾値母数推定値順序性違反が表 1 に観察される素点上での点双列相関係数値大小との整合性に欠けることに関心が向く。

6. おわりに

本稿での筆者による Rasch 評定尺度モデル解題は当然ながら完璧ではない。唯、Andrich (1978a[2], 2004[8], and 2005[5])により洞察されたそのモデル規定としての潜在閾値母数順序性がモデル測定出力上での閾値母数推定値順序性を項目群カテゴリー順序段階尺度化妥当性の上で必然としていることは明確に理解される。Andrich[5] (2005)によってデータの統計的モデル適合度指標は最優先確認事項としての閾値母数推定値順序性とは無連関である(p. 53)とも指摘されている。これは、閾値母数推定値順序性が Rasch 評定尺度モデル測定 Facets 出力上で各項目、各被験者、並びに、各順序カテゴリーに対してモデル適合度指標として付される Infit・Outfit 統計量からは分離独立しているとの示唆である。

Linacre[11] (1989)の Many-Facet Rasch Measurement も含めた Rasch 評定尺度モデル測定出力上での閾値母数推定値順序性参照に基づく項目群カテゴリー順序段階尺度化妥当性へのその仮定先験性を踏まえた研究考察が順序尺度データ分析上での最優先事項として真に望まれる。但し、Rating Scale モデル閾値母数推定値順序性は質問紙項目群カテゴリー順序段階尺度化妥当性検証への必要条件ではある(Andrich, 1978a[2], 2004[8], and 2005[5])が、潜在特性一次元性と局所独立性を併せた構成概念妥当性を保証する充分条件ではないということは前節で査察された通りである。最後に、Wright[28] (1998)により引用された『より少ない前提により説明され得ることはより多い前提により説明不能となる』(p. 642)との Ockham's Razor が Rasch 評定尺度モデル測定上で筆者に思惟される。

【参照文献】

1) G. Rasch.1960.*Probabilistic models for some intelligence and attainment tests*. The Danish Institute for Educational Research. (Reprinted in 1980 by the University of Chicago Press with a Foreword and Afterword by B. D. Wright)

2) D. Andrich. 1978a. A rating formulation for ordered response categories. *Psychometrika*, 43, 4, pp. 561–573.

3) G. N. Masters. 1982. A Rasch model for partial credit scoring. *Psychometrika*, 47, 2, pp. 149–174.

4) T. McNamara. 1996. *Measuring second language performance*. London: Longman.

5) D. Andrich. 2005. The Rasch model explained. In Alagumalai, S, Curtis, D. D., & Hungi, N (Eds.), *Applied Rasch measurement: A book of exemplars* (pp. 27–59). The Netherlands: Springer.

6) B. D. Wright & G. N. Masters. 1982. *Rating scale analysis*. Chicago: MESA Press.

7) G. Luo. 2005. The relationship between the rating scale and partial credit models and the implication of disordered thresholds of the Rasch models for polytomous responses. *Journal of Applied Measurement*, 6, 4, pp. 443–455.

8) D. Andrich. 2004. Understanding resistance to the data-model relationship in Rasch's paradigm: A

reflection for the next generation. In Smith, Jr., E. V. & Smith, R. M. (Eds.), *Introduction to Rasch Measurement* (pp. 167−200). Maple Grove, Minnesota: JAM Press.

9) 藤森　進　2002. 「項目反応理論による多値データの分析について」 文教大学人間科学部、『人間科学研究』第 24 号、pp. 21−31.

10) 村木英治　2005. 「全米学力調査(NAEP)概説」　東京大学大学院教育学研究科　教育測定・カリキュラム開発講座公開研究会、2005 年 11 月 24 日　於：東京大学赤門総合研究棟

11) J. M. Linacre.　1989.　*Many-facet Rasch measurement*.　Chicago:　MESA Press.

12) 秋山朝康　2000.　「スピーキングテストの分析と評価 ― 項目応答理論を使っての研究」*STEP* (The Society for Testing English Proficiency, Inc.) *Bulletin*, 12, pp. 67−78.

13) G. Rasch. 1961. On general laws and the meaning of measurement in psychology. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Theory of Probability, Vol. IV* (pp. 321−333). Berkeley: University of California Press.

14) E. B. Andersen. 1972. The numerical solution of a set of conditional estimation equations. *Journal of the Royal Statistic Society, Series B*, 34, pp. 42−54.

15) B. D. Wright. 1991. Scores, reliabilities and assumptions. *Rasch Measurement Transactions*, 5, 3 in Linacre, J. M. (Ed.), 1995, *Rasch Measurement Transactions, Part 1* (pp.157−158). Chicago: MESA Press.

16) E. B. Andersen. 1977. Sufficient statistics and latent trait models. *Psychometrica*, 42, 1, pp. 69−81.

17) D. Andrich. 1978b. Application of a psychometric rating model to ordered categories which are scored with successive integers. *Applied Psychological Measurement*, 2, 4, pp. 581−594.

18) J. M. Linacre. 1989−2001. *A user's guide to FACETS: Rasch measurement computer program*. Chicago: Winsteps.com.

19) D. Andrich. 1995. Models for measurement, precision, and the nondichotomization of graded responses. *Psychometrika*, 60, 1, pp. 7−26.

20) D. Andrich. 1998. Thresholds, steps and rating scale conceptualization. *Rasch Measurement Transactions*, 12, 3, pp. 648−649.

21) J. M. Linacre. 2001. Category, step and threshold: Definitions and disordering. *Rasch Measurement Transactions*, 15, 1, p. 794.

22) G. N. Masters & B. D. Wright. 1984. The essential process in a family of measurement models. *Psychometrika*, 49, 4, pp. 529−544.

23) B. D. Wright. 1980. Foreword and afterword. In Rasch, G., *Probabilistic models for some intelligence and attainment tests* (Reprinted in 1980, pp. ix−xix and pp. 185−190). The University of Chicago Press.

24) D. Andrich. 1997. Rating scale analysis. In Keeves, J. P. (Ed.), *Educational research, methodology, and measurement:　An international handbook (2nd. Edition)* (pp. 874−880).　Oxford:　Elsevier Science Ltd.

25) C. M. Fox & J. A. Jones. 1998. Uses of Rasch modeling in counseling psychology research. *Journal of Counseling Psychology*, 45, 1, pp. 30−45.

26) J. M. Linacre. 2000. Comparing "partial credit" and "rating scale" models. *Rasch Measurement*

Transactions, 14, 3, p. 768.

27) J. M. Linacre. 1990. Where does misfit begin? *Rasch Measurement Transactions*, 3, 4 in Linacre, J. M. (Ed.), 1995, *Rasch Measurement Transactions, Part 1* (pp. 80–81). Chicago: MESA Press.

28) B.D.Wright. 1998. Model selection: Rating scale or partial credit? *Rasch Measurement Transactions*, 12, 3, pp. 641–642.

29) M. E. Lunz, B. D. Wright, & J. M. Linacre. 1990. Measuring the impact of judge severity on examination scores. *Applied Measurement in Education*, 3, 4, pp. 331–345.

30) J. M. Linacre.2002. Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3, 1, pp. 85–106.

31) G. N. Masters & B. D. Wright. 1982. Defining a 'fear-of-crime' variable: A comparison of two Rasch models. *Educational Research and Perspectives*, 9, pp. 18–31.

32) J. M. Linacre. 1998a. Rasch first or factor first? *Rasch Measurement Transactions*, 11, 4, p. 603.

33) B. D. Wright. 1996. Local dependency, correlations and principal components. *Rasch Measurement Transactions*, 10, 3, pp. 509–511.

34) J. M. Linacre. 1998b. Detecting multidimensionality: Which residual data-type works best? *Journal of Outcome Measurement*, 2, 3, pp. 266–283.

35) J. M. Linacre. 1998c. Structure in Rasch residuals: Why principal components analysis? *Rasch Measurement Transactions*, 12, 2, p. 636.

36) B. D. Wright & M. H. Stone. 2004. *Making measures*. Chicago: The Phaneron Press.

37) 井澤廣行 2006.「局所独立性と潜在特性一次元性に関する表示実体としての異質性」『流通科学大学論集―人間・社会・自然編』第 19 巻、第 1 号、pp. 1–14.

38) A. Tennant & J. F. Pallant. 2006. Unidimensionality matters. *Rasch Measurement Transactions*, 20, 1, pp. 1048–1051.

39) J. M. Linacre. 1999a. Explorations into local independence with T-Rasch. *Rasch Measurement Transactions*, 13, 3, p. 710.

40) 豊田秀樹 2002.『項目反応理論[入門編]』東京：朝倉書店

41) D. Andrich. 1988. A general form of Rasch's extended logistic model for partial credit scoring. *Applied Measurement in Education*, 1, 4, pp. 363–378.

42) J. M. Linacre. 1999b. Category disordering vs. step (threshold) disordering. *Rasch Measurement Transactions*, 13, 1, pp. 675–678.

*【付録】

このアンケート用紙を提出して下さい。　学年 ____　性別 ____　この質問紙調査の井澤研究

自分自身について、それぞれの項目上で該当する□にレ印を挿入して下さい。使用への 承諾 ・ 不承諾

1. 計画を立ててそれに従う	□	□	□	□	□	計画には固執しない
	強	弱	中間	弱	強	
2. 感情表現・情緒発露を嫌う	□	□	□	□	□	感情表現・情緒発露を好む
	強	弱	中間	弱	強	
3. 個々の事象に注視する	□	□	□	□	□	全体像を把握する
	強	弱	中間	弱	強	
4. あいまいさを嫌う	□	□	□	□	□	あいまいさを認める
	強	弱	中間	弱	強	
5. 知性を重視する	□	□	□	□	□	感性を重視する
	強	弱	中間	弱	強	
6. 言葉のみで説明する	□	□	□	□	□	図を加えて説明する
	強	弱	中間	弱	強	
7. いろいろ考えて判断する	□	□	□	□	□	瞬時に判断する
	強	弱	中間	弱	強	
8. 人の表情に無意識である	□	□	□	□	□	人の表情を意識する
	強	弱	中間	弱	強	
9. 論理的に理解する	□	□	□	□	□	直観的に理解する
	強	弱	中間	弱	強	
10. 従属・追随的である	□	□	□	□	□	率先・進取的である
	強	弱	中間	弱	強	
11. 一つずつ課題を処理する	□	□	□	□	□	同時に複数の課題に携る
	強	弱	中間	弱	強	
12. 技術・職人芸に無関心	□	□	□	□	□	技術・職人芸に惹かれる
	強	弱	中間	弱	強	
13. 規律性を好む	□	□	□	□	□	柔軟性を好む
	強	弱	中間	弱	強	
14. 試行錯誤を嫌う	□	□	□	□	□	試行錯誤を好む
	強	弱	中間	弱	強	
15. 説明にたとえ話や人から聞いたことを加えない	□	□	□	□	□	説明にたとえ話や人から聞いたことを加える
	強	弱	中間	弱	強	