

# Rating Scale モデル適合データでの異なる閾値による 点双列相関係数、位置母数推定値標準誤差、並びに、Outfit $\chi^2$ 平均平方統計量の分布特徴

## Distributions of Point-Biserials, Model Standard Errors, and Outfit Mean-Squares along 3-Category Items' Different First Threshold Values on the Rating Scale Model

平越裕之\*, 井澤廣行†

Hiroyuki Hirakoshi, Hiroyuki Izawa

This paper investigates the distributions of point-biserials, model standard errors, and Outfit Mean-Squares, which are output by applying FACETS to Rating Scale model-fitted raw data produced along 20 different first threshold set values in 3-category data. The key features of the RSM-fitted data are that 1) the standard deviation of the uniformly set items' (25, 50, 100, 150) location values is 1.0; and 2) the distribution of the set persons' (500) location values follows  $N(0, 1)$ . All the 3 indices' standard deviations are found the smallest around the first threshold values of -1.0, -1.5, -2.0, and -2.5.

**Key words:** Rating Scale model; Category thresholds; Point-Biserials; Standard errors; Outfit Mean-Squares.

### I. はじめに

本論への導入として、平越・井澤<sup>1)</sup>(2012)における図 2.2(p. 20)を下に再掲する。その図は、Rating Scale Model (Andrich<sup>2)</sup>, 1978、以降、RSM)上での 3 カテゴリーデータに関して、項目群と被験者群の Outfit  $\chi^2$  平均平方統計量(以降、Outfit MnSq)平均標準偏差の第 1 閾値異同に伴う理論値変動を俯瞰したものである。閾値設定値データ番号 12( $\tau_1 = -0.7$ )、13( $\tau_1 = -1.0$ )、14( $\tau_1 = -1.5$ )、15( $\tau_1 = -2.0$ )、16( $\tau_1 = -2.5$ )の範囲内にて、項目群と被験者群いずれにも関する Outfit MnSq 分布標準偏差の最小性が示された。

本稿の目的は、上記理論値上での Outfit MnSq 平均標準偏差の最小性を与える 3 カテゴリーデータ第 1 閾値範囲が、点双列相関係数分布、並びに、RSM 適用 FACETS<sup>3)</sup>出力による位置母数推定値標準誤

\*流通科学大学総合政策学部、〒651-2188 神戸市西区学園西町 3-1

†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

差分布と Outfit MnSq 分布においても同様であることを査察することにある。

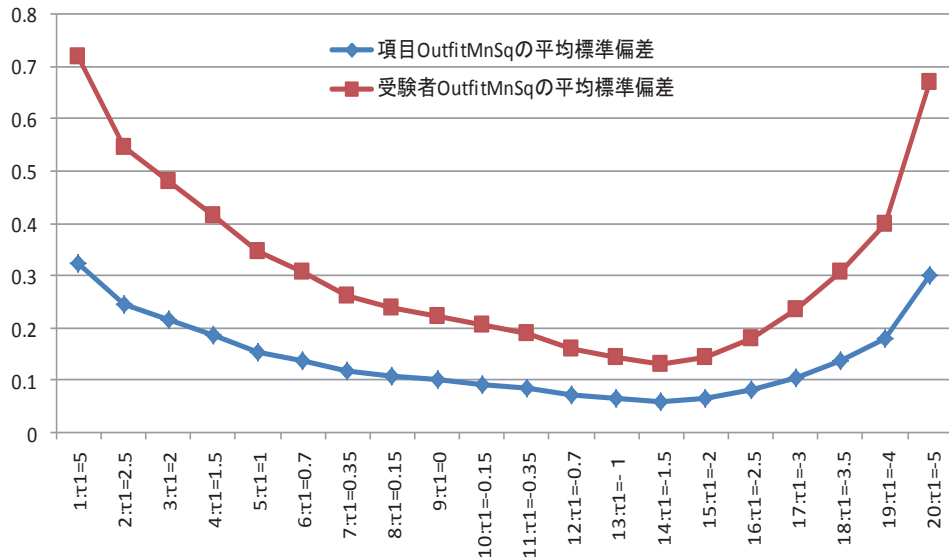


図 1.1 100 項目(標準偏差を 1.0 とする一様分布として $-1.71 \sim 1.71$  を範囲とする位置母数設定値)と 500 名(標準正規分布  $N(0, 1)$  に従う位置母数設定値)それぞれの Outfit MnSq 平均標準偏差<sup>4)</sup>

なお、本稿で査察する 3 指標値の理解は以下の通りである。

- 点双列相関係数は、素点上での弁別力指標値(discrimination index)である(静・竹内・吉澤<sup>5)</sup>、2002, p. 132)と共に、「妥当性の指標」(豊田<sup>6)</sup>、2002, p. 67)でもある。又、項目群・被験者群点双列相関係数分布の平均値は、データ次元充足度の視認簡易な指標値でもある(平越・井澤<sup>1)</sup>、2012, 図 3.8, p. 26、参照)。なお、FACETS により出力される点双列相関係数は、SPSS<sup>7)</sup>出力と同じく、「修正」(Corrected)点双列相関係数である。それは、各項目列と各被験者行の得点並びそれぞれと、被験者得点列と項目得点行のそれぞれから各項目列と各被験者行の得点並びそれぞれを差し引いたもののピアソン相関係数である。
- 位置母数推定値標準誤差は母数推定精度(precision)の指標値である(Linacre<sup>8)</sup>、1989-2001, p. 52)。
- Outfit MnSq は、モデルが規定する特性(期待得点)曲線形状同一性からの乖離度指標値である(Linacre<sup>9)</sup>、2000; 平越・井澤<sup>10)</sup>、2010; 平越・井澤<sup>1)</sup>、2012)。

## II. 被験者群と項目群に関する点双列相関係数

多値 Rasch モデルである RSM における被験者点双列相関係数(以下、単に相関係数とも記す)について考察する。

$m$  段階 ( $m+1$  カテゴリー) の RSM では、被験者  $i$  が項目  $j$  に対してカテゴリ  $k$  の得点となる確率を次のように表している(Andrich<sup>2)</sup>, 1978, モデル式 13, p. 569、参照)。

$$P_i(x=k|\delta_j) = \frac{\exp\{\kappa_k + k(\beta_i - \delta_j)\}}{\sum_{l=0}^m \exp\{\kappa_l + l(\beta_i - \delta_j)\}} \quad (2.1)$$

$\beta_i$ : 被験者  $i$  の能力値

$\delta_j$ : 項目  $j$  の困難度

$\tau_n$ : カテゴリ  $n$  の閾値 (全項目共通),  $\sum \tau_n = 0, n \in (1, 2, \dots, m)$

$\kappa_l = -\sum_{n=1}^l \tau_n, \kappa_0 = 0, l \in (0, 1, 2, \dots, m)$

本節での考察を目的として、以下のようなシミュレーション設定とする。

- カテゴリ数: 0, 1, 2 の 3 つ。
- 被験者群: 500 人で能力は正規分布  $N(0, 1)$  に従う。
- 項目群: A) 25 問 B) 50 問 C) 100 問 D) 150 問として以下の設定値とする。
  - A) 25 問の項目困難度範囲を  $-1.63 \sim 1.63$  として等間隔に配置する(設定値分布標準偏差 1.0)。
  - B) 50 問の項目困難度範囲を  $-1.68 \sim 1.68$  として等間隔に配置する(設定値分布標準偏差 1.0)。
  - C) 100 問の項目困難度範囲を  $-1.71 \sim 1.71$  として等間隔に配置する(設定値分布標準偏差 1.0)。
  - D) 150 問の項目困難度範囲を  $-1.714 \sim 1.714$  として等間隔に配置する(設定値分布標準偏差 1.0)。
- 閾値母数:  $\tau_1 (= -\tau_2)$ 
  - 1)  $\tau_1 = 5.0$  2)  $\tau_1 = 2.5$  3)  $\tau_1 = 2.0$  3)  $\tau_1 = 1.5$  5)  $\tau_1 = 1.0$
  - 6)  $\tau_1 = 0.7$  7)  $\tau_1 = 0.35$  8)  $\tau_1 = 0.15$  9)  $\tau_1 = 0.0$  10)  $\tau_1 = -0.15$
  - 11)  $\tau_1 = -0.35$  12)  $\tau_1 = -0.7$  13)  $\tau_1 = -1.0$  14)  $\tau_1 = -1.5$  15)  $\tau_1 = -2.0$
  - 16)  $\tau_1 = -2.5$  17)  $\tau_1 = -3.0$  18)  $\tau_1 = -3.5$  19)  $\tau_1 = -4.0$  20)  $\tau_1 = -5.0$

RSM の反応行列は、本来その要素が各被験者、各項目への得点(反応)を表し、0, 1, 2 のいずれかの値をとる行列である。ここでは、理解を深めるために、まず、各反応の期待得点(実数)行列を基にして求めた被験者点双列相関係数分布の平均と標準偏差を、それぞれ図 2.1 と 図 2.2 に与える。

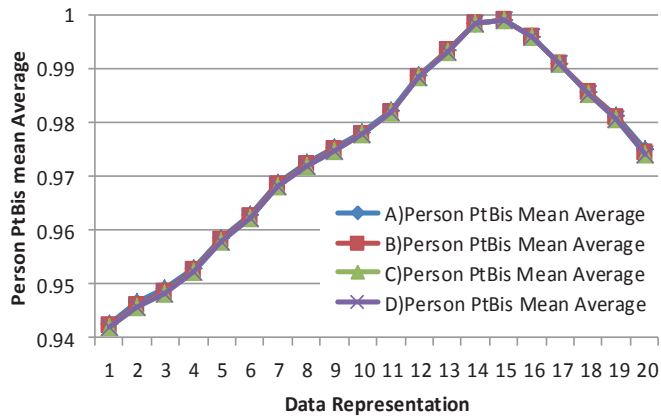


図 2.1 期待得点から計算した被験者点双列相関係数分布平均値

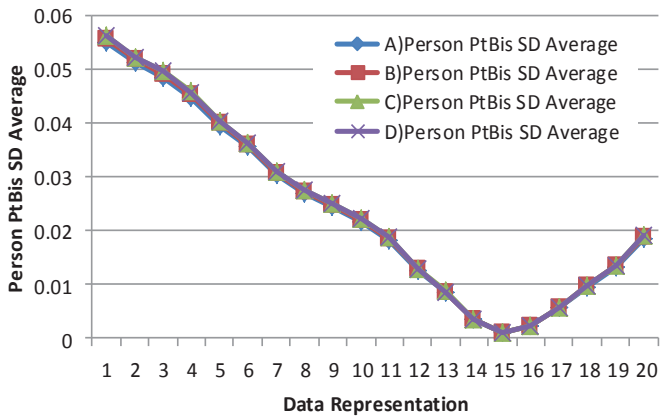


図 2.2 期待得点から計算した被験者点双列相関係数分布標準偏差

なお、上記の相関係数は、以下のようにして算出したものである。

- 1) 項目困難度の配置は、前述の A)~D)に従う。
- 2) 500 人の被験者能力  $\beta_i$  を、次のように設定する。

$$\beta_i = x_i, \quad x_i \text{は正規分布} N(0,1) \text{において} f(x_i) = \left(\frac{i}{501}\right) \text{の確率密度を持つ} x_i \text{とする}$$

$$(i = 1, 2, \dots, 500)$$

- 3) 前述の 1)~20)の閾値設定により、式(2.1)のカテゴリ選択確率にカテゴリ(0,1,2)を掛けて期待得点を求め、行列にする。
- 4) 3)の行列の列方向の総和を求め、「各被験者の期待得点行」と、「総和から各被験者の期待得点行を引いたもの」のピアソン相関係数を求める。

図 2.1 と図 2.2 は、実際の反応では起こりえない期待得点(実数値)を並べた反応行列の点双列相関係数であるから、0, 1, 2 のカテゴリが生成される前のモデルそのものの特徴が表れる。図 2.1 では、すべての場合で相関係数は 0.94 を超え、特にカテゴリ閾値設定値 13)と 14)ではほぼ 1 となり、13)と 14)を頂点として左右に徐々に小さくなっていく傾向が見て取れる。この際の反応行列は理論値である期待値なので、1)~20)のそれぞれにおいて、相関係数に影響を及ぼすのは、(能力  $\beta$  - 困難度  $\delta$ ) の変化に対して、期待得点が線形であるかどうかである。

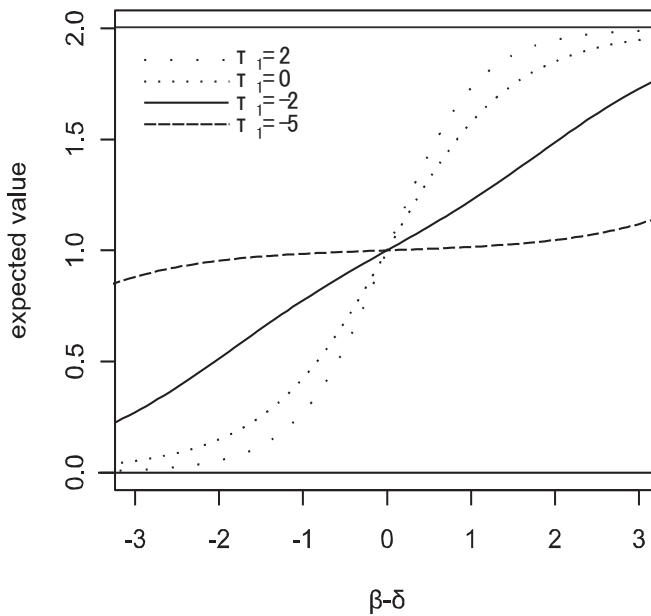


図 2.3 期待得点の変化

ここで、図 2.3 に、(能力  $\beta$  - 困難度  $\delta$ ) が  $-3 \sim 3$  の場合の期待得点の変化を示す。(能力  $\beta$  - 困難度  $\delta$ ) を  $-3 \sim 3$  としているのは、本稿での項目群と被験者群の位置母数設定において、被験者すべての各項目への反応が概ねこの範囲に入っているからである。図 2.3 において観察されるように、 $\tau_1 = 2$ 、 $\tau_1 = 0$ 、 $\tau_1 = -2$ 、 $\tau_1 = -5$  となるにつれて、傾斜の強い S 字曲線形から直線を経由して逆 S 字曲線形に変化していくことがわかる。このことから、図 2.1 に示されているように、 $\tau_1 = -2$  の時、つまり  $\tau_1$  設定 15) の時を頂点として、左右に少しずつ相関係数が下がっていくことが説明できる。つまり、期待得点の線形性が高いほど相関係数が高くなるということである。標準偏差については、相関係数平均値の逆の形状となっているが、ばらつきを表すことを考えれば、期待得点の線形性が大きくかわることは明らかである。つまり、相関係数が低い要因である非線形性は、各被験者能力の違いによる期待得点にも当然に影響を及ぼすから、能力の違いによる相関係数の異同が発生する要因になると考えられる。従って、線形性が高いときには相関係数標準偏差が小さくなり、逆に線形性が低くなればなるほど相関係数標準偏差は大きくなる。

次に、A)~D)、1)~20)において、それぞれ 300 の RSM 適合データを乱数により生成し、シミュレーションを行った結果から、被験者点双列相関係数の平均、及び、被験者点双列相関係数標準偏差の平均を求めたものを図 2.4 と図 2.5 に示す。

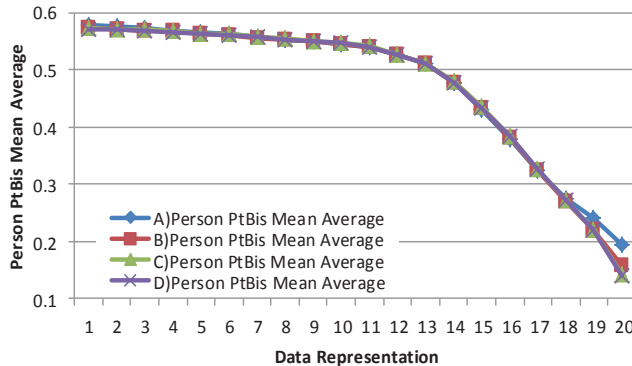


図 2.4 シミュレーションによる被験者点双列相関係数分布平均値の平均値  
(各値は RSM 適合 300 データの平均値)

RSM 適合データからシミュレーションによって求めた図 2.4 の相関係数の平均値と図 2.5 の標準偏差については、実数値である期待得点とは異なり、反応は 0, 1, 2 の 3 値に限られるから、相関係数の値は当然に減少する。又、期待得点のように実数値をとらないことから、実際のカテゴリ生起確率から議論することが必要である。

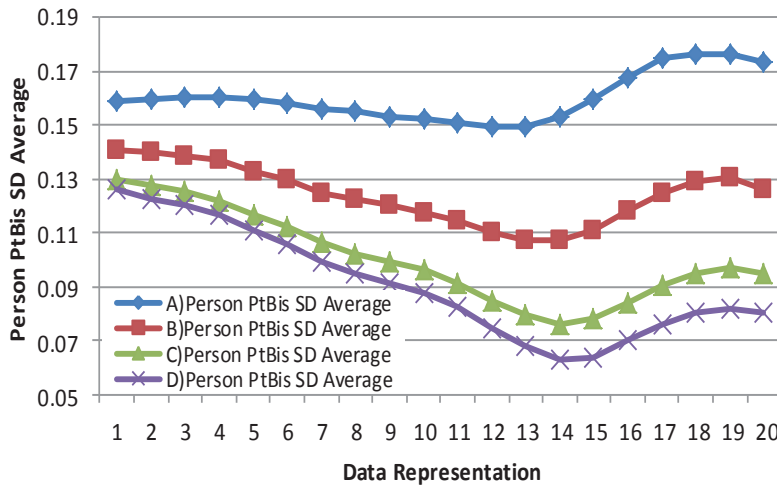


図 2.5 シミュレーションによる被験者点双列相関係数分布標準偏差の平均値  
(各値は RSM 適合 300 データの平均値)

図 2.4 の相関係数平均値は、カテゴリ閾値設定値 13)と 14)あたりを境として 20)の方向に急激に減少している。これは、図 2.6～図 2.8 に示す各カテゴリ生起確率と関連している。図 2.6～図 2.8 は、それぞれ  $\tau_1 = 2$ 、 $\tau_1 = -2$ 、 $\tau_1 = -5$  の時のカテゴリ 0, 1, 2 の生起確率を示している。

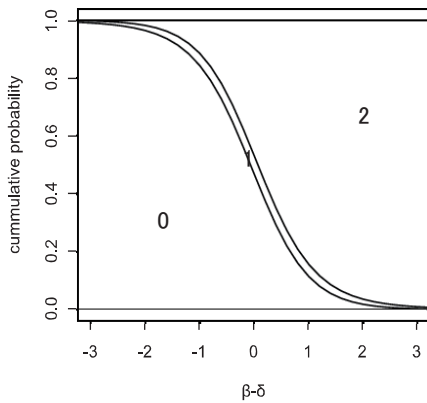


図 2.6 カテゴリ生起確率( $\tau_1 = 2$ )

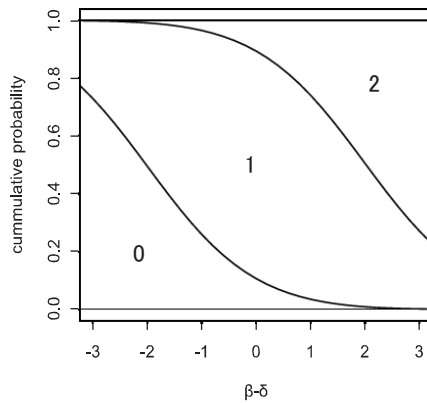


図 2.7 カテゴリ生起確率( $\tau_1 = -2$ )

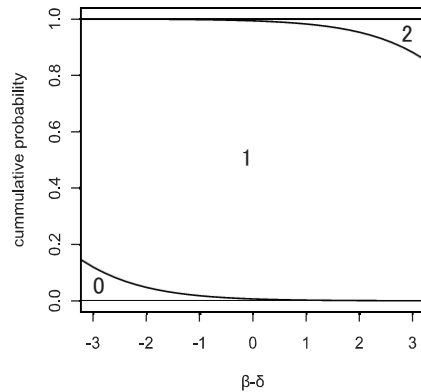


図 2.8 カテゴリ生起確率( $\tau_1 = -5$ )

$\tau_1 = 2$ の時にはカテゴリ 0 とカテゴリ 2 にはっきり分かれて生起する特徴があり、 $\tau_1 = -2$ の時にはカテゴリ 1 にやや偏ったものとなる。又、 $\tau_1 = -5$ の時にはほとんどの場合でカテゴリ 1 が生起する。相関係数はあくまで二対データの線形の関連を調べる指標であるから、片方のデータ値が変化に乏しい場合には、相関係数値は低くなる。これにより、図 2.4 のように相関係数平均値が変化すると考えられる。

図 2.5 の相関係数標準偏差については、期待得点から求めた図 2.2 と傾向は同じになるが、相関係数平均値と同様に、実数値ではなく 0, 1, 2 の反応しかしないため、全体的に値が増大していることがわかる。又、シミュレーションによって計算される相関係数標準偏差の場合にも、期待得点の線形性が大きく影響しており、非線形になればなるほどその影響でばらつきが大きくなると考えられる。これは、図 2.2 の期待得点から計算した標準偏差の傾向を比較的に反映している。なお、図 2.5 において、A)~D)における差異は項目数の違いによるものであり、被験者あたりの項目数が大きくなればなる程、各被験者の相関係数値は安定する。その結果として、標準偏差が小さくなっていると考えられる。

次に、被験者点双列相関係数と同様に、項目点双列相関係数についても、期待得点から計算した相関係数平均を図 2.9 に、相関係数標準偏差を図 2.10 にそれぞれ示す。更に、シミュレーションによって計算した相関係数平均を図 2.11 に、相関係数標準偏差を図 2.12 に示す。

図 2.9 と図 2.10 における期待得点からの項目相関係数の平均値及び標準偏差については、項目数(=25, 50, 100, 150)と被験者数(=500)の違いによる若干の差異は認められるが、概ね図 2.1 と図 2.2 のそれぞれにおける被験者相関係数の平均値及び標準偏差と同傾向である。



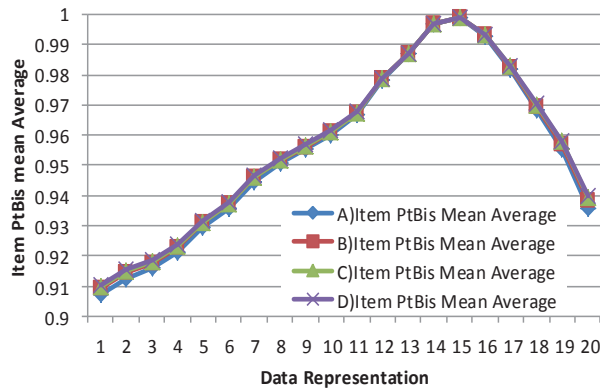


図 2.9 期待得点から計算した項目点双列相関係数分布平均値

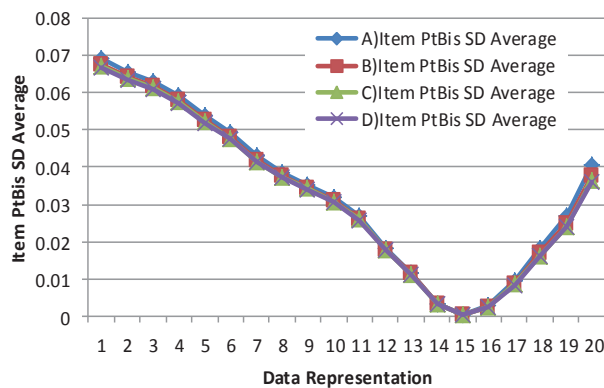


図 2.10 期待得点から計算した項目点双列相関係数分布標準偏差

図 2.11 と図 2.12 のシミュレーションによる項目相関係数については、図 2.11 での項目相関係数平均値は、図 2.4 のシミュレーションによる被験者相関係数平均値とほぼ同様な値を示している。この理由は、前述の被験者相関係数に関する言及に準ずる。一方、図 2.12 のシミュレーションによる項目相関係数標準偏差については、図 2.10 の期待得点から計算した項目相関係数標準偏差に酷似しており、最小値が若干大きい点を除いてほぼ同じ形状となっている。これは、項目数(=25, 50, 100, 150)に比較して被験者数の多大な 500 名による影響である。図 2.5 でのシミュレーションによる被験者相関係数標準偏差においても、項目数が増えれば期待得点から計算した被験者相関係数標準偏差に近づいていくこと、並びに、0, 1, 2 という 3 値反応であることを考えれば、シミュレーションによる項目相関係数標準偏差は、期待得点からの相関係数標準偏差ほど小さくはならない。従って、図 2.12 の傾向は

合理的なものと考えられる。

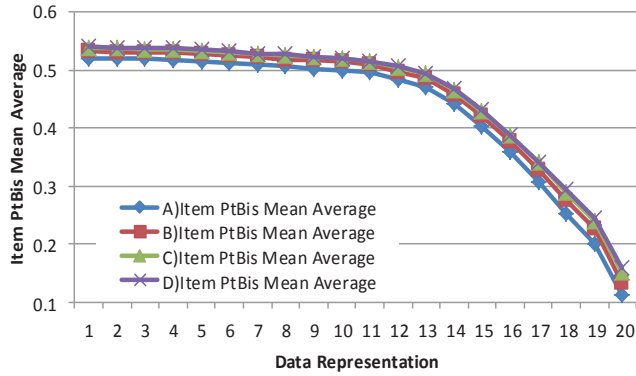


図 2.11 シミュレーションによる項目点双列相関係数分布平均値の平均値  
(各値は RSM 適合 300 データの平均値)

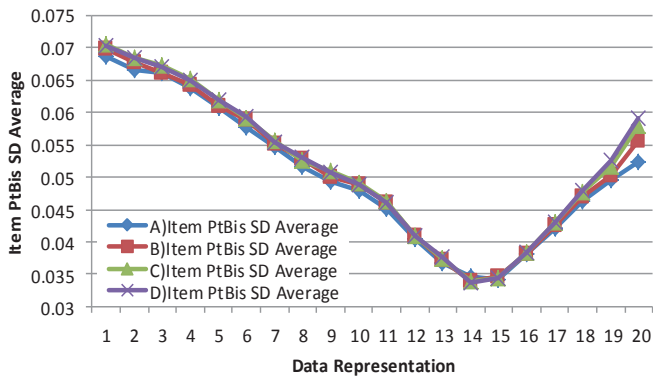


図 2.12 シミュレーションによる項目点双列相関係数分布標準偏差の平均値  
(各値は RSM 適合 300 データの平均値)

### Ⅲ. RSM 適合 3 カテゴリ・閾値 20 種データの FACETS 出力による位置母数推定値標準誤差と Outfit MnSq の分布査察

Ⅱ節での RSM 適合 3 カテゴリ・閾値 20 種データに関する図 2.4、2.5、2.11、2.12 のそれぞれに照応して、その傾向線上での相対関係をほぼ同様とする FACETS 出力による位置母数推定値標準誤差と Outfit MnSq の分布を本節にて査察する。

なお、以下に与える各図は、II節での困難度設定値分布の標準偏差を 1.0 として等間隔に配置された A) 25 項目、B) 50 項目、C) 100 項目、D) 150 項目、及び、標準正規分布  $N(0, 1)$  に従う被験者群 500 名位置母数分布、並びに、1)~20) の  $\tau_1$  設定値に関するものである。上記各母数設定値に基づいて、フリーソフト R 上で著者により作成された RSM 適合多値カテゴリ順序データ出力プログラム<sup>11)</sup> 使用による項目数 A)、B)、C)、D) に関する 3 カテゴリ・閾値 20 種データ各 10 標本取得の後に、RSM 適用 FACETS 出力による各指標値分布の平均値と標準偏差の 10 標本平均値である。

項目群：

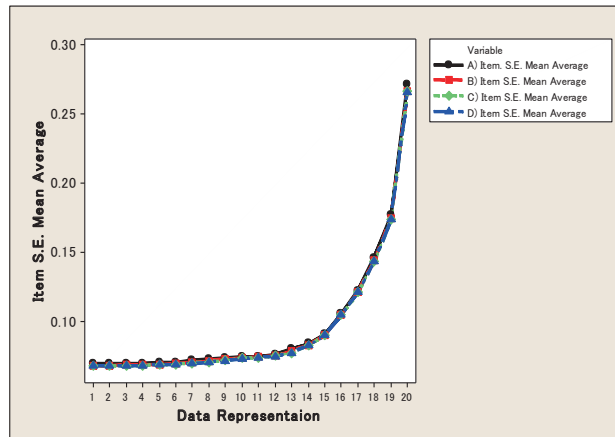


図 3.1 A)、B)、C)、D)に関するカテゴリ閾値 20 種データ各 10 標本項目群位置母数推定値標準誤差分布平均値の平均値

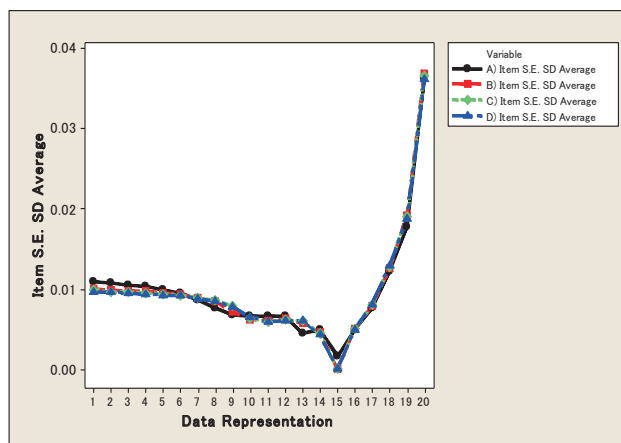


図 3.2 A)、B)、C)、D)に関するカテゴリ閾値 20 種データ各 10 標本項目群位置母数推定値標準誤差分布標準偏差の平均値

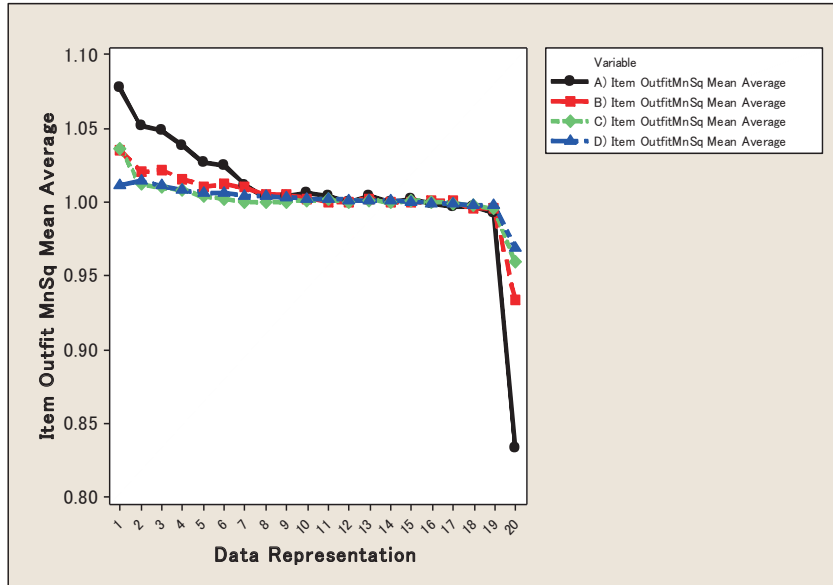


図 3.3 A)、B)、C)、D)に関するカテゴリ閾値 20 種データ各 10 標本  
項目群 Outfit MnSq 分布平均値の平均値

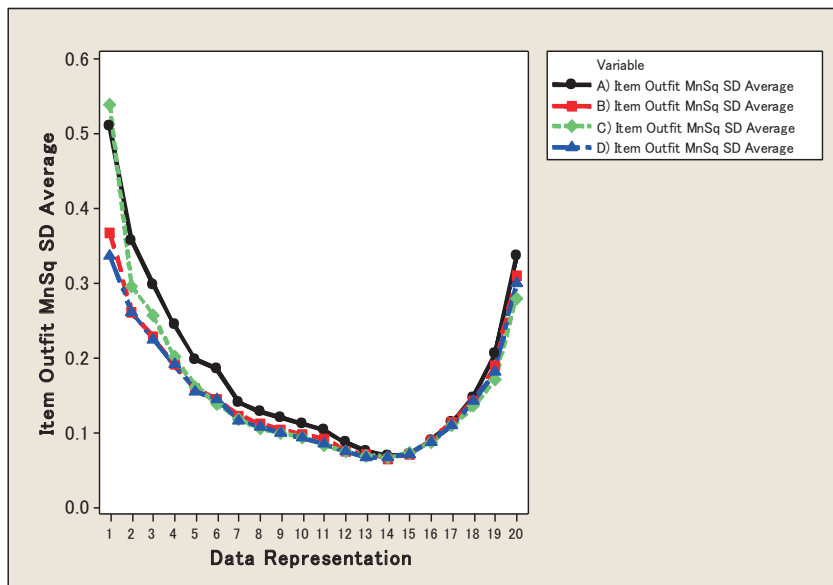


図 3.4 A)、B)、C)、D)に関するカテゴリ閾値 20 種データ各 10 標本  
項目群 Outfit MnSq 分布標準偏差の平均値

被験者群：

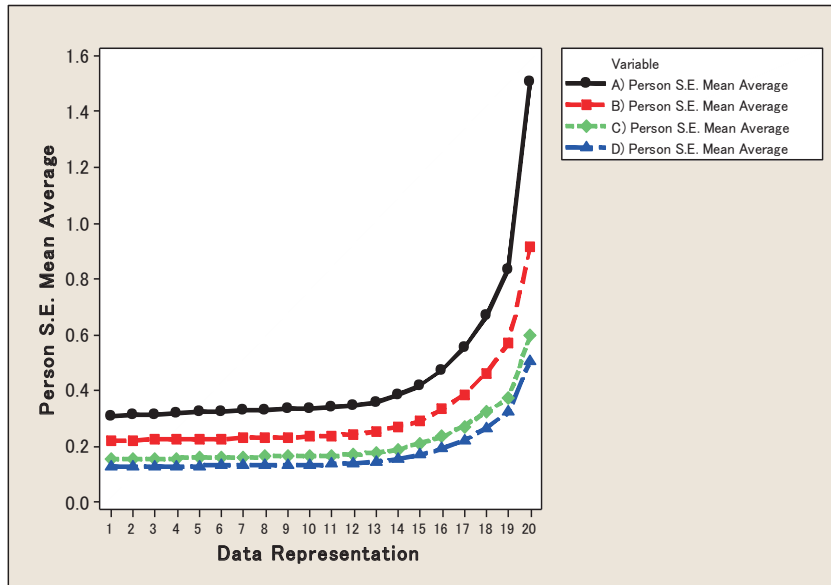


図 3.5 A)、B)、C)、D)に関するカテゴリ閾値 20 種データ各 10 標本  
被験者群位置母数推定値標準誤差分布平均値の平均値

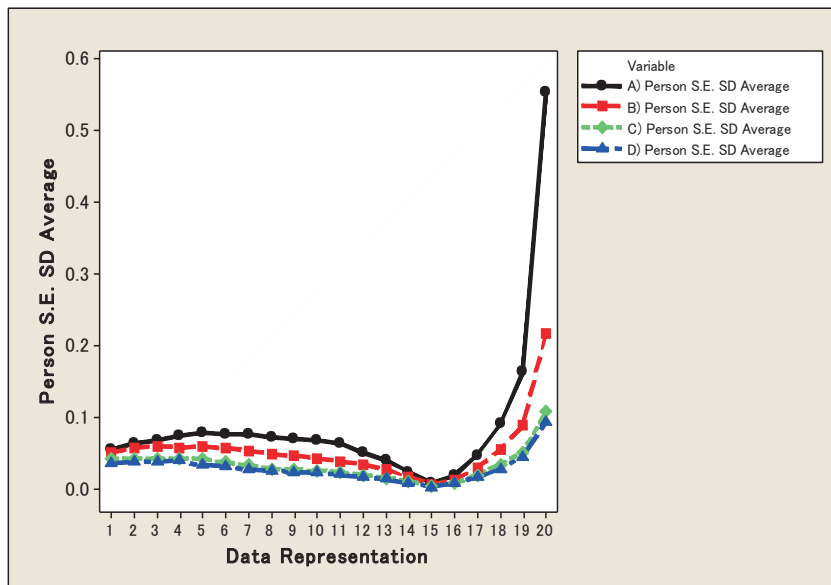


図 3.6 A)、B)、C)、D)に関するカテゴリ閾値 20 種データ各 10 標本  
被験者群位置母数推定値標準誤差分布標準偏差の平均値

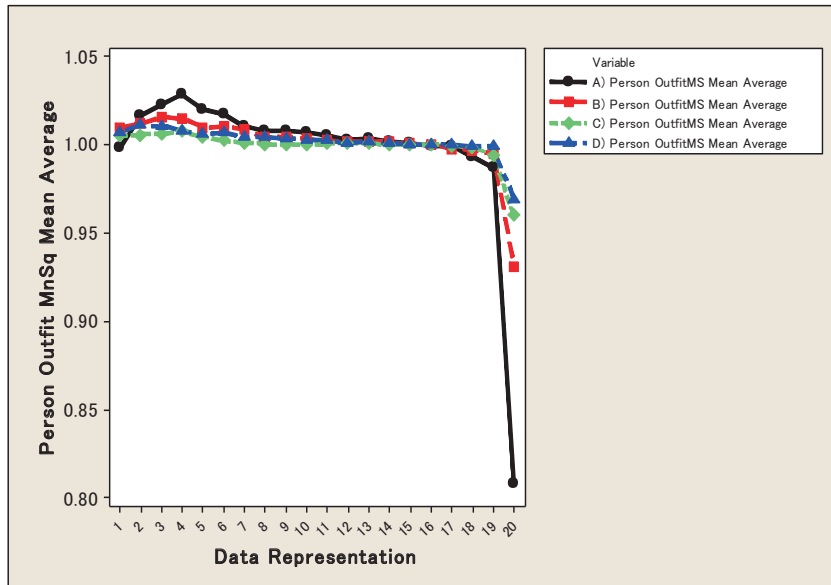


図 3.7 A)、B)、C)、D)に関するカテゴリ閾値 20 種データ各 10 標本  
被験者群 Outfit MnSq 分布平均値の平均値

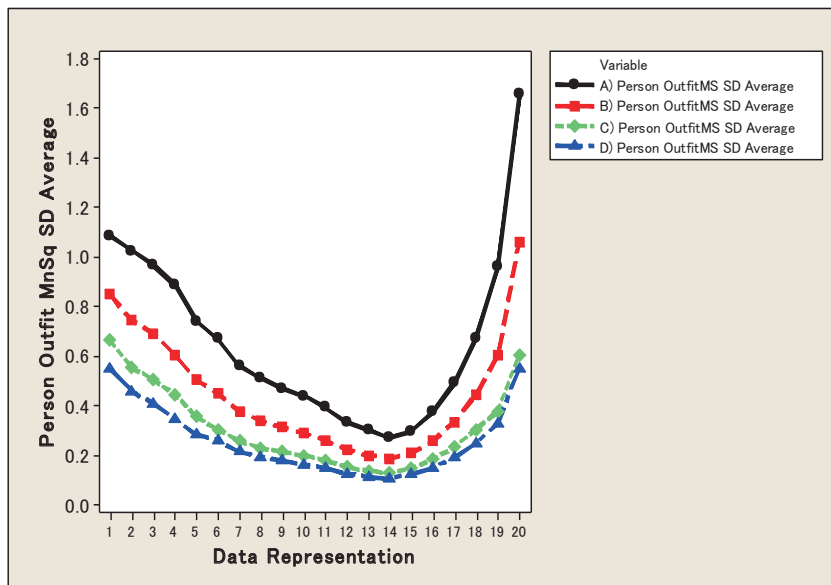


図 3.8 A)、B)、C)、D)に関するカテゴリ閾値 20 種データ各 10 標本  
被験者群 Outfit MnSq 分布標準偏差の平均値

上記 3 指標値に関して、II 節での提示図 2.5 と図 2.12、及び、図 3.2、3.4、3.6、3.8 における観察要点は、端的に、閾値設定値データ番号 13( $\tau_1 = -1.0$ )、14( $\tau_1 = -1.5$ )、15( $\tau_1 = -2.0$ )、16( $\tau_1 = -2.5$ ) の範囲内にて項目群と被験者群いずれにも関する指標値分布標準偏差が最小になっている。その重要な示唆は、本稿での項目群・被験者群位置母数設定値分布標準偏差(=1.0)と 3 カテゴリデータ第 1 閾値の設定値 1)~20)における RSM 適用 FACETS 出力指標値に関しては、データ番号 13)、14)、15)、16)の範囲内にて項目群と被験者群いずれにも関する各指標値の均質近似性(出力値安定性)が最も高いということである。指標値分布平均値については、II 節での提示図 2.4 及び図 2.11 との相対関係として、図 3.1、3.3、3.5、3.7 においてデータ番号 1)から 20)に渡り漸増ないしは漸減の傾向が示されて、概ねデータ番号 16 辺りにて各指標値平均としての許容性高低臨界点が観察される。各指標値分布に関する上述の傾向は、被験者群の設定位置母数分布を標準正規分布  $N(0, 1)$  とする被験者数 5,000 名と 10,000 名に関する各設定項目数(位置母数設定値分布標準偏差=1.0)と各閾値への対応 1 標本においても同様と観察される。500 名から 5,000 名と 10,000 名への被験者数増大による項目群指標値精度の大きな上昇に違いがあるのみである。従って、上述の傾向を生起させている主要因は、項目群と被験者群の位置母数設定値分布標準偏差を同一値 1.0 としていることにあると推測される。Rasch 測定における位置母数推定値分布標準偏差におけるこの両者均衡性の重要性は留意されて然るべきである(平越・井澤<sup>12)</sup>、2008, p. 32; 平越・井澤<sup>1)</sup>、2012, p. 27)。本考察においては、II 節提示図 2.3 におけるデータ番号 15( $\tau_1 = -2.0$ )の項目特性(期待得点)曲線の線形性にその根源重要性が察知される。要するに、データ番号 13( $\tau_1 = -1.0$ )、14( $\tau_1 = -1.5$ )、15( $\tau_1 = -2.0$ )、16( $\tau_1 = -2.5$ )の範囲内にて観察される項目群と被験者群いずれにも関する指標値分布標準偏差の最小性は、その  $\tau_1$  設定値(=  $\tau_1$  推定値)辺りにおけるデータ群項目・被験者特性曲線いずれもの強い線形性に関係していると思われる。被験者群と項目群に関する点双列相関係数分布におけるその関係については II 節に詳しく言及されている。

更に、データ番号 13( $\tau_1 = -1.0$ )、14( $\tau_1 = -1.5$ )、15( $\tau_1 = -2.0$ )、16( $\tau_1 = -2.5$ )の範囲内にて観察される項目群と被験者群いずれにも関する上記 3 種指標値分布標準偏差の最小性は、Linacre<sup>13)</sup> (2002) による 3 カテゴリデータにおける理論上の好ましい隣接閾値推定値間正順差の妥当な範囲は 1.4~5.0 との指摘(pp. 102-104)に合致している。Linacre<sup>13)</sup> (2002)によるその詳しい説明は以下の通りである。

- カテゴリ閾値推定値は、正順で少なくとも 1.4 ロジット値の差[ $\tau_1$  推定値=-0.7]で以って大きくなることが望ましい。RSM における一つ手前のカテゴリが選択された結果として、当該カテゴリが選択されるか否かという一連の二者択一に関する理論上の隣接閾値間最小差が 1.4 ロジット値である(p. 102)。
- カテゴリ閾値推定値は、最大で 5 ロジット値の差[ $\tau_1$  推定値=-2.5]を越えることなく、カテゴリの進行に伴って正順で大きくなることが望ましい。隣り合う二つの閾値の推定値における正順差が極端に大きくなれば、被験者群による一つのカテゴリの認識幅が広くなり過ぎることにより、そのカテゴリの中間部は認識不鮮明なものとなる。従って、そ

のカテゴリ自体の情報力と尺度安定性が極端に低くなる(pp. 103-104)。

図 3.9 は、Linacre<sup>13)</sup> (2002)の提示図(Figure 8, p. 103)に対応する 3 カテゴリ項目群に関する RSM 第 1 閾値の違いによる項目情報関数値の異同を示すものである。上記に言及された II 節提示図 2.3 におけるデータ番号 15( $\tau_1 = -2.0$ )の項目特性(期待得点)曲線の線形弁別力 0.25 が、 $-3 \sim 3$  を( $\beta - \delta$ )値とする範囲内でその項目情報関数値の範囲がほぼ 0.21~0.27 と示されて、2 値項目における最大弁別力 0.25 に相応している。データ番号 13( $\tau_1 = -1.0$ )、14( $\tau_1 = -1.5$ )、15( $\tau_1 = -2.0$ )、16( $\tau_1 = -2.5$ )におけるデータ群項目・被験者特性曲線いずれものこの強い線形性が、項目群と被験者群いずれにも関する上記 3 種指標値分布標準偏差の最小性と関係しているとの推測である。

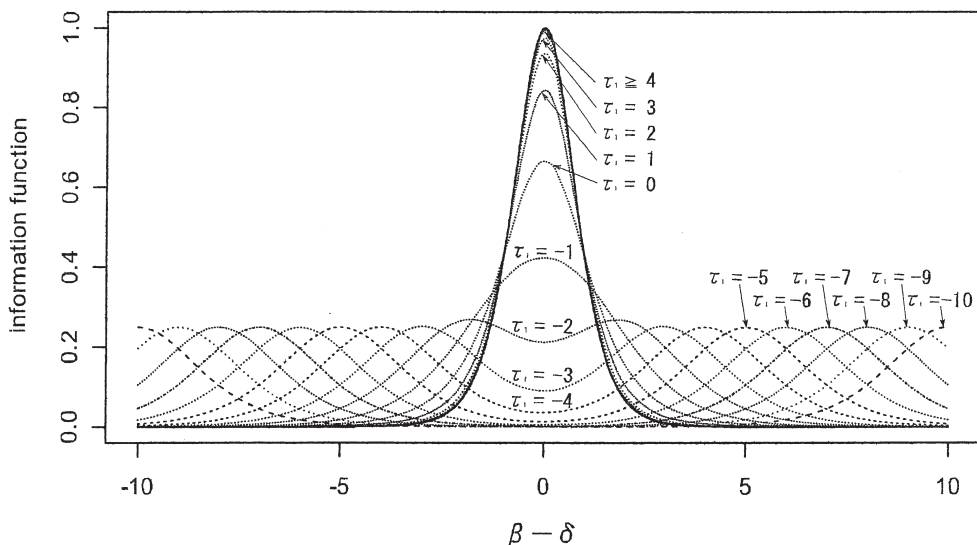


図 3.9 3 カテゴリ項目(満点 2)の情報関数値<sup>14)</sup>

本節補遺として、図 3.10~図 3.13 に A)、B)、C)、D)の FACETS 出力による項目群と被験者群それぞれに関するカテゴリ閾値 20 種データ各 10 標本における位置母数推定値分布の平均値と標準偏差平均値を与える。その参照意図は、異なる  $\tau_1$  設定値(=  $\tau_1$  推定値)における両群位置母数設定値分布の設定平均値(=0)と設定標準偏差(=1.0)からの出力位置母数推定値分布の乖離異同程度の観察である。

項目群については、同一被験者数 500 名によるその位置母数の推定であり、FACETS 上での項目群位置母数推定値平均が 0.0 と設定されることにより、異なる  $\tau_1$  設定値(=  $\tau_1$  推定値)における項目群位置母数推定値分布平均値の異同は微小である。その分布標準偏差は、項目数増大につれて、A)、B)、C)、D)の順に、僅少なながらも推定値精度の上昇が窺われる。一方、被験者群については、項目数増大(25, 50, 100, 150)につれて被験者群位置母数の推定精度が全般的に上昇しているが、データ番号 18( $\tau_1 = -3.5$ )、19( $\tau_1 = -4.0$ )、20( $\tau_1 = -5.0$ )における推定精度は、特に A)と B)において、劣悪である。



これにより、3 値項目群から成る現実質問紙調査の設計・取得に際しては、RSM 母数推定精度の上で、100 項目と 500 名が標本の大きさの上で望まれる目安最小数との示唆である。又、 $\tau_1$  設定値(≒  $\tau_1$  推定値)が  $-3.5$  より小さな値となるにつれて、被験者群の位置母数推定精度は顕著に悪化する傾向が観察される。これは、前掲図 3.1 と 3.2、並びに、図 3.5 と 3.6 での項目群・被験者群位置母数の推定精度を示す標準誤差分布の  $\tau_1$  設定値異同における 10 標本の平均値と標準偏差の平均値傾向曲線に準じている。

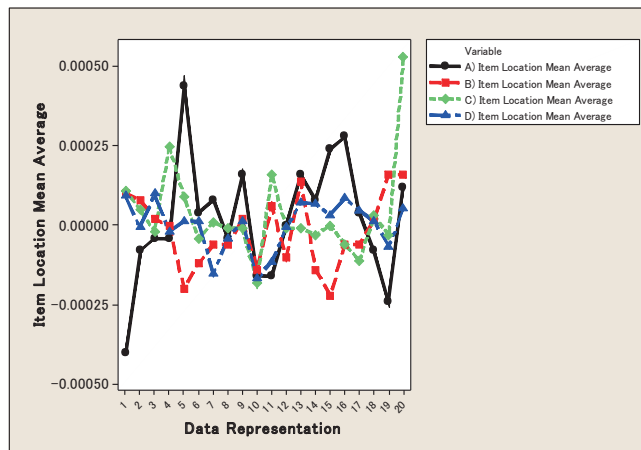


図 3.10 A)、B)、C)、D)に関するカテゴリ閾値 20 種データ各 10 標本  
項目群位置母数推定値分布平均値の平均値

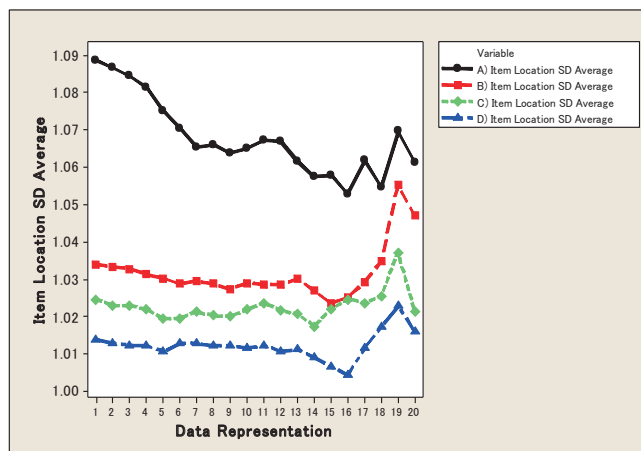


図 3.11 A)、B)、C)、D)に関するカテゴリ閾値 20 種データ各 10 標本  
項目群位置母数推定値分布標準偏差の平均値

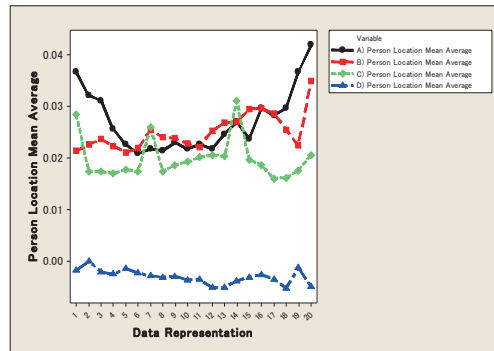


図 3.12 A)、B)、C)、D)に関するカテゴリ閾値 20 種データ各 10 標本  
被験者群位置母数推定値分布平均値の平均値

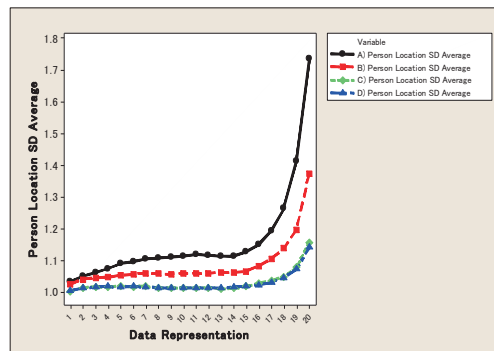


図 3.13 A)、B)、C)、D)に関するカテゴリ閾値 20 種データ各 10 標本  
被験者群位置母数推定値分布標準偏差の平均値

#### IV. おわりに

RSM の考案者 Andrich は次の様に述べている(Andrich<sup>15)</sup>, 2005)。

『RSM 式 [本稿での式 2.1、ないしは、Andrich<sup>2)</sup>, 1978, 式 13, p. 569、参照]に従って、シミュレーションにより潜在閾値母数の推定値が逆順となるデータの模擬作成は可能である。その作成された[潜在閾値逆順]データは Rasch モデルに完全に適合している』(p. 53)。

これにより、本稿査察のために作成された全ての 3 カテゴリ・シミュレーションデータは RSM に完全に適合していると言える。その意味は、本稿での 3 カテゴリ・RSM 適合シミュレーションデータにおいては、0・1・2 生起の上で「質の良悪」は存在せず、0・1・2 の生起に「0→1→2」の強度増大性(潜在カテゴリ閾値を想定するガットマンパタン<sup>16)</sup>の生起性)は仮定されていないということである。それが、現実に設計・取得される 3 値順序データとの違いであり、その旨の理解は大変に重要である。

本査察は、1) 現実取得データが RSM に完全に適合していること、及び、2) 項目群と被験者群い

ずれもの位置母数推定値分布標準偏差が 1.0 であることを想定条件としている。その想定の下で、項目数(=25, 50, 100, 150)と被験者数(=500)に近似する現実標本が与えられるならば、各項目数・被験者数における現実取得標本の RSM への適合度に関して、本稿での設定第 1 閾値の異同による点双列相関係数、位置母数推定値標準誤差、並びに、**Outfit MnSq** の分布における平均値と標準偏差が参照値となる。その本稿提示参照値と現実取得標本での各対応値との同等・乖離程度により、現実データの RSM への適合度が概ね判断される。

3 カテゴリ項目群から成る現実質問紙調査の設計・取得においては、リカート尺度(Likert Scale, 静・竹内・吉澤<sup>9)</sup>, 2002, p. 104, 参照)が一般的に使用される。それは、例えば、2 値の「反対・賛成」に「中立」を加えて、「反対・中立・賛成」の 3 カテゴリ構成から成る順序尺度である。その場合には、暗に、「0→1→2」の賛成強度増大性が仮定されている。従って、RSM の本質的なカテゴリ潜在閾値として  $\tau_1 < \tau_2$  が想定されて(Andrich<sup>2)</sup>, 1978, p. 565)、カテゴリの「0」・「1」・「2」それぞれが、ガットマンパターンとしての(0, 0)・(1, 0)・(1, 1)の生起と想定される(Andrich<sup>2)</sup>, 1978, pp. 567-568)。この理解の上で、第 II 節での点双列相関係数、並びに、第 III 節での RSM 適用 FACETS 出力上での位置母数推定値標準誤差と **Outfit MnSq** の分布において、 $\tau_1 = -1.0$ 、 $\tau_1 = -1.5$ 、 $\tau_1 = -2.0$ 、 $\tau_1 = -2.5$  の範囲内での項目群と被験者群いずれにも関する両者指標値分布標準偏差の最小性が観察された。これにより、現実の 3 カテゴリ順序データにおいて、潜在特性の上で「0→1→2」の強度増大性が仮定されているならば(Andrich<sup>15)</sup>, 2005, p. 53, 参照)、RSM 適用の上で出力される第 1 閾値推定値が上記  $\tau_1$  推定値の範囲内にあることは、項目群と被験者群に関する上記 3 種出力指標値それぞれの高い均質近似性、つまりは、安定性を示唆していることになる。それは、第 II 節にて参照した Linacre<sup>13)</sup> (2002) による 3 カテゴリデータにおける理論上の好ましい隣接閾値推定値間正順差の妥当な範囲は 1.4~5.0 との指摘(pp. 102-104)に合致している。

なお、本稿査察における  $\tau_1 = -1.0$ 、 $\tau_1 = -1.5$ 、 $\tau_1 = -2.0$ 、 $\tau_1 = -2.5$  の範囲内での項目群と被験者群いずれにも関する 3 種指標値分布標準偏差の最小性と上記  $\tau_1$  値辺りでの項目・被験者特性曲線いずれもの強い線形性(第 II 節提示図 2.3, 参照)との関連性が、第 II 節での点双列相関係数に関する論説に加えて、今後の研究課題として意識される。それは、RSM 適用上での現実取得データ良質性(RSM 適用妥当性)において、項目群と被験者群に関する均衡度の高い両者程良い次元充足度、及び、線形的特性曲線の同形性(弁別力同一性)上での両者程良い傾きの本質的必然性を示唆しているからである。シミュレーションデータにおける項目群と被験者群の位置母数設定値分布標準偏差を同一値 1.0 としていることが、本稿査察結果の主要起因であるとの推測である。RSM 測定における位置母数推定値分布標準偏差におけるこの両者均衡性の重要性は、2 値データに適用される Rasch 項目分析モデル(Rasch<sup>16)</sup>, 1960, 再刊 1980)に準ずる(平越・井澤<sup>12)</sup>, 2008; 井澤・平越<sup>11)</sup>, 2011, 参照)と思われる。

#### 【参考文献及び注】

- 1) 平越裕之・井澤廣行 2012. 「Rating Scale モデル測定項目群 Outfit  $\chi^2$  平均平方統計量の本質性に関する一考察」 『流通科学大学論集 — 人間・社会・自然編』第 24 巻、第 2 号、pp. 17-31.
- 2) D. Andrich. 1978. A rating formulation for ordered response categories. *Psychometrika*, 43, 4, pp. 561-573.
- 3) J. M. Linacre. 1989-2001. *FACETS* (Rasch measurement computer program). Chicago: Winsteps.com.
- 4) 平越裕之・井澤廣行 2012. 「Rating Scale モデル測定項目群 Outfit  $\chi^2$  平均平方統計量の本質性に関する一考察」 『流通科学大学論集 — 人間・社会・自然編』第 24 巻、第 2 号における図 2.2(p. 20)の再掲.
- 5) 静哲人・竹内理・吉澤清美共編著 2002. 『外国語教育リサーチとテストングの基礎概念』 大阪： 関西大学出版部
- 6) 豊田秀樹 2002. 『項目反応理論[事例編]』 東京： 朝倉書店
- 7) 統計解析ソフトウェア IBM SPSS Statistics.
- 8) J. M. Linacre. 1989-2001. *A user's guide to FACETS* (Rasch measurement computer program). Chicago: Winsteps.com.
- 9) J. M. Linacre. 2000. Item discrimination and infit mean-square. *Rasch Measurement Transactions*, 14, 2, p. 743.
- 10) 平越裕之・井澤廣行 2010. 「Rash 測定  $\chi^2$  平均平方統計量、項目反応理論二母数モデル項目弁別力、及び、データ次元性の連関性一考察」 『流通科学大学論集 — 人間・社会・自然編』第 23 巻、第 1 号、pp. 97-112.
- 11) 井澤廣行・平越裕之 2011. 『項目応答分析 Rasch モデル精察』における補遺 B) (pp. 296-299) 神奈川県： 現代図書
- 12) 平越裕之・井澤廣行 2008. 「テストの正誤反応行列から得られる次元性情報に関する考察」 『流通科学大学論集 — 経済・経営情報編』第 17 巻、第 1 号、pp. 9-33.
- 13) J. M. Linacre. 2002. Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3, 1, pp. 85-106.
- 14) 平越裕之 2009. 「多値 Rasch モデルのカテゴリ閾値に関する一考察」 『流通科学大学論集 — 経済・経営情報編』第 17 巻、第 2 号における図 3-2(p. 95)の再掲.
- 15) 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.
- 16) 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.)