

# IRTのソフトウェア PARSCALEの使い方

室橋 弘人

(お茶の水女子大学)

# PARSCALEとは

- IRTに基づく分析を行うためのソフトウェア
  - 発売元: SSI(<http://www.ssicentral.com/>)
  - 価格: 250\$
- 扱うことの出来る分析
  - 元々, 順序カテゴリカルデータのためのソフト
    - 中でも一般化部分採点モデル(GPCM)がメイン
    - 段階反応モデル(GRM)も推定可能
  - 2値データも扱うことができる
    - 1~3母数のロジスティックモデル
  - 大規模かつ複雑なテストデータを扱うことが可能
    - 異なる回答形式が混在するテストの分析
    - 複数の集団に対して複数のテストを実施したデータの分析・等化
    - DIFの分析
    - 評定者効果の分析

# 本発表で扱う内容

- Ver 4.1 (現行バージョン) 準拠
- 特に順序カテゴリカルデータの分析を念頭に置いて、基本的な分析を方法を紹介する
  - 部分点のある項目を含むテスト
  - リッカート形式の項目を含むテスト
- 発表の流れ
  1. データの準備と読み込み
  2. 分析モデルの指定
  3. 推定の実行と結果の確認

# データの準備と読み込み指定

# データの準備

```
1 20081001 432343454454134232254421242243
2 20081002 433121443433134213244421522133
3 20081003 413233344322233324444431223313
4 20081004 42435241445412211345253151345
5 20081005 312351355223243311254521112229
6 20081006 541451444444115113242421552243
7 20081007 54135154311111521115151121111
8 20081008 2314144351142452342552322221
```

ID

回答

- テキストファイルで用意する
  - ID+回答というのが想定される基本的な形式
- IDと回答は, それぞれ桁数を揃えて入力する
  - IDの桁数と, 1項目への回答の桁数
- 回答は順序を持った整数値で入力しておくこと楽
  - 2値データの場合, 「1」「2」
  - 順序カテゴリカルデータの場合, 「1」「2」「3」...

# 欠損値の入力

データ

```
1 20081001 432343454454134232254421242243
2 20081002 433121443433134213244421522133
3 20081003 413233344322233324444431223313
4 20081004 42435241445412211345253151345
5 20081005 312351355223243311254521112229
6 20081006 541451444444115113242421552243
7 20081007 54135154311111521115151121111
8 20081008 2314144351142452342552322221
```

欠損コード

```
mislode.txt
1 dummysig 99999999999999999999999999999999
2 ↓
3 [EOF]
```

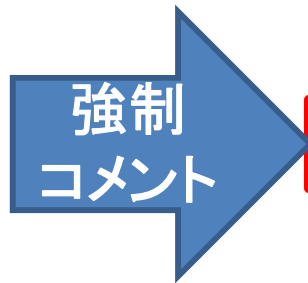
- 回答とは被らない欠損を表すコードを用意し、データファイルにはその値を入力する
  - 回答が「1」～「5」のときに、欠損は「9」など
- 回答データとは別に、欠損コードを指定するファイルを用意する
  - 桁数は、IDも含めてデータファイルと同様にする
  - 回答に相当する部分に、各項目の欠損コードを入力する

# 複数のテストをまとめる場合

	項目1～50	項目51～100
集団A	提示	未提示
集団B	未提示	提示

- グループごとに異なる構成のテストを用いた場合、提示していない項目への回答は「未提示(Not Presented)」として扱う
  - 集団Aは前半の50個分に回答を入力
  - 集団Bは後半の50個分に回答を入力
- 使わなかった項目への回答は、未提示を表すコードを設定して入力する
  - 全体としては集団A, Bともに100項目への反応があるという形式のデータになる
  - 欠損ファイルと同様に、未提示コードを指定するファイルを用意する
    - ただし現バージョンでは、推定時に未提示と欠損を区別していない

# PARSCALEの実行



```
1 | CO1.PSL:↓  
2 | Samejima's Graded Response Model for CO1 items.↓  
3 | >FILE  DFNAME='TCIRid.txt', OFNAME='misCode.txt', SAVE;↓  
4 | >SAVE  PARM='CO1item.txt', SCORE='CO1trait.txt';↓  
5 | >INPUT NIDCH=8, NTOTAL=240, NTEST=1, LENGTH=8, NFMT=1;↓  
6 | (8A1.1X.240A1)↓
```

- コマンドを記述したスクリプトを読み込ませて実行する
  - 拡張子は「.psl」だが、実体はテキストファイル
- 冒頭2行は、強制的に「タイトル」として認識される
  - この部分に命令を書いてもコマンドとして認識されないの  
で注意
- 3行目以降にコマンドを記述する
  - 「>コマンド オプション;」という形式
  - オプションは「,」で区切って記述する



# コマンド一覧

>COMMENT: コメント  
>FILES: 読み込むファイルの設定  
>SAVE: 書き出すファイルの設定  
>INPUT: 変数読み込みの設定  
(細かい読み込み指定)  
>TEST: テスト構成の設定  
>BLOCK: テストに含まれるブロックの設定  
>MGROUP: 多母集団に関する設定  
>MRATER: 評定者に関する設定  
>CALIB: 項目母数推定に関する設定  
>QUADP: 項目母数推定時の求積点の設定  
>PRIORS: 推定時の事前分布の設定  
>SCORE: 被験者母数推定に関する設定  
>QUADS: 被験者母数推定時の求積点の設定  
>COMBINE: 部分得点を合算するための重みの設定

- 上からこの順番で指定する必要がある
  - 下線が引かれているのは、必須コマンド

# ファイル読み込みの指定

```
>FILE DFNAME='...',  
      OFNAME='...', NFNAME='...',  
      SAVE, ... ;
```

- 「>FILE」以下に、分析で読み込む各種ファイルの名前を記述する
  - DFNAME: データファイル
  - OFNAME: 欠損コードファイル
  - NFNAME: 未提示コードファイル
  - SAVE: 結果を別ファイルとして出力する場合に指定

# ファイル保存の設定

```
>SAVE PARM='...', SCORE='...',  
      INFORMATION='...', ...;
```

- 項目母数や被験者母数を別ファイルとして保存したい場合の設定
  - 推定結果そのものは、ここで指定しなくても表示される
  - さらに別の分析などで利用する場合に必要
- 事前に「>FILE」コマンドでSAVEオプションを指定しておく必要がある
  - PARM: 項目母数を保存するファイルの名前
  - SCORE: 被験者母数を保存するファイルの名前
  - INFORMATION: 項目情報量を保存するファイルの名前

# 実際の記述例

```
2 | samejima s graded response model for LUI items.↓  
3 | >FILE DFNAME='TCIRid.txt', OFNAME='misCode.txt', SAVE;↓  
4 | >SAVE PARM='CO1item.txt', SCORE='CO1trait.txt';↓  
5 | >INPUT NIDCH=8, NTOTAL=240, NTEST=1, LENGTH=8, NFMT=1;↓  
6 | (8A1,1X,240A1)↓  
7 | \TEST1 TRAIT NO UNBLOCK 0 1
```

- >FILE
  - データは”TCIRid.txt”というファイルから読む
  - 欠損パターンは”misCode.txt”というファイルから読む
  - 結果を別ファイルに出力する
- >SAVE
  - 項目母数を”CO1item.txt”というファイルに出力
  - 被験者母数を”CO1trait.txt”というファイルに出力

# データ読み取りの指定(1)

```
>INPUT NTEST=..., LENGTH=(...),  
      NTOTAL=..., NFMT=...,  
      NIDCH=..., ... ;
```

- 「>INPUT」以下に, データ読み取りの指定を行う
  - NTEST: 提示したテストの種類
  - LENGTH: テストごとの項目数
  - NTOTAL: データファイル全体での項目数
  - NFMT: 1問への回答の桁数
  - NIDCH: ID部の桁数

# データ読み取りの指定(2)

(○A1,△X,□A◇)

- 「>INPUT」直下の行に, フォートラン形式で細かい入力指定を行う
  - ○: ID部の桁数
  - △: ID部と回答部の間に関隔を開けている場合, その桁数
  - □: データ全体での項目数
  - ◇: 回答1問辺りの桁数

# 実際の記述例

```
2 | samejima s graded response model for GUI items.↓
3 | >FILE DFNAME='TCIRid.txt', OFNAME='misCode.txt', SAVE;↓
4 | >SAVE PARM='CO1item.txt', SCORE='CO1trait.txt';↓
5 | >INPUT NIDCH=8, NTOTAL=240, NTEST=1, LENGTH=8, NFMT=1;↓
6 | (8A1,1X,240A1)↓
7 | >TEST1 TRAIT=NO UNBLOCK=0
```

- >INPUT

- IDは8桁で入力
- 項目は全部で240個
- テストは1種類
- テストは8問
- 1問への回答は1桁で入力

- (8A1, 1X, 240A1)

- IDは8桁
- IDと回答部の間に1桁空いている
- 回答部は, 240項目を各1桁ずつで入力

	0	10	20	30
1	20081001	432343454454134232254421242243		
2	20081002	433121443433134213244421522133		
3	20081003	413233344322233324444431223313		
4	20081004	42435241445412211345253151345.		
5	20081005	312351355223243311254521112229		
6	20081006	541451444444115113242421552243		
7	20081007	54135154311111521115151121111		
8	20081008	23141443511424523425523222211		

# 分析に関する指定



# テスト構成の指定

```
>TEST TNAME=..., NBLOCK=...,  
      ITEM=(...), INAME=(...);
```

- 「>TEST」以下に，分析するテストの項目構成を指定する
  - TNAME: テスト名
  - NBLOCK: テストを構成するブロックの数(後述)
  - ITEM: テストに含まれる項目の通し番号
  - INAME: テストに含まれる項目の名前

# 実際の記述例

```
5 | (8AI,IX,Z4UAI)↓  
7 | >TEST1  TNAME=NS, NBLOCK=8,↓  
8 |         ITEM=(4,28,124,133,142,198,203,224),↓  
9 |         INAME=(X004,X028,X124,X133,X142,X198,X203,X224);↓  
7 | \NBLOCK  RNAME=(V004 V028 V124 V133 V142 V198 V203 V224) |
```

- >TEST

- TNAME=NS: テスト名は「NS」
- NBLOCK=8: ブロック数は8
- ITEM=(...): データ全体のうち何番目の項目が含まれるかを, 全体での通し番号で指定
- INAME=(...): 項目名のラベル

# ブロックとは？

- ブロックとは1つのテストの中で、「選択肢間の困難度の差が等しい」と仮定できる項目のまとまりのこと
  - 同じブロックに属する項目は、困難度によって相対的な位置関係が変わっても、選択肢間の距離は等しくなる
  - ブロックに含まれる項目は、回答カテゴリの数が同じでなければならない
- 項目ごとに選択肢間の距離が異なるように推定を行うためには、1問1ブロックとして指定する必要がある
  - この場合、ブロック数＝項目数
  - GPCM, GRMにおいて一般的な設定

# ブロックの指定

```
>BLOCK BNAME=(...), NITEMS=...,  
NCAT=..., REPEAT=...,  
GUESSING=..., ;
```

- 「>BLOCK」以下に, テストを構成するブロックの指定を行う
  - BNAME: ブロック名
  - NITEMS: ブロックに含まれる項目数
  - NCAT: ブロックに含まれる項目の回答カテゴリ数
  - REPEAT: 「NITEMS」「NCAT」の設定が同じブロックを複数設定する場合, その繰り返し回数
  - GUESSING: 当て推量母数に関する設定

# 実際の記述例

```
5 | >BLOCKS BNAME=(X004,X028,X124,X133,X142,X198,X203,X224),↓  
1 |         NITEMS=1, NCAT=5, REPEAT=8;↓  
2 | ▶CAT IR GRADFD LOGISTIC CYCLES=20 NOPTS=31 NEWTON=20 T
```

- 8問からなるテストに対して, 1ブロック1問で全体で8ブロックを設定
  - NITEMS=1: ブロックに含まれるのは1問
  - NCAT=5: 項目への回答は5件法
  - REPEAT=8: 上記設定のブロックを8個作る
  - BNAME: ブロック名 = 項目名としている

# 複数のテスト・ブロックがある場合

- 「>TEST」や「>BLOCK」が複数ある場合, 「>TEST1」「>TEST2」のように順番に番号を振って並べていく
  - テストごとにブロックを指定して1セット

スクリプト全体
>TEST1
>BLOCK (TEST1のブロック指定)
>TEST2
>BLOCK1 (TEST2のブロック1の指定)
>BLOCK2 (TEST2のブロック2の指定)
>TEST3
>BLOCK1 (TEST3のブロック1の指定)
>BLOCK2 (TEST3のブロック2の指定)

# その他>BLOCKで可能なこと

```
1 >BLOCKS BNAME=(MC01,MC02,MC03,MC04,MC05,MC06,MC07,MC08,MC09,MC10,MC11,MC12), ↓
2         NITEMS=1, NCAT=2, ORIGINAL=(0,1), MODIFIED=(1,2), ↓
3         REPEAT=12, GUESSING=(2,ESTIMATE) ; ↓
4 \BLOCK BNAME-OF NITEMS-1 NCAT-2 COORDINATE-(1 2 2) . .
```

- 当て推量母数の推定を行うかどうかの指定
  - 推定を行いたい場合, 以下のオプションを追加(デフォルトでは推定されない)  
「GUESSING=(N, ESTIMATE)」
    - Nには回答における選択肢数を入力する
- データにおいて入力された値を, 分析用に変換する
  - MODIFIEDオプションなどを利用する

# 項目母数推定の指定

>CALIB GRADED/PARTIAL,  
LOGISTIC/NORMAL,  
SPRIOR, TPRIOR,  
NOCALIB, ... ;

- 「>CALIB」以下に、項目母数の推定に関する指定を行う
  - GRADED/PARTIAL: 段階反応モデルか部分採点モデルか
  - LOGISTIC/NORMAL: ロジスティックモデルか正規累積モデルか (GRADEDのときのみ有効)
  - SPRIOR: 識別力母数に対して、対数正規分布を事前分布として利用する
  - TPRIOR: 困難度／ステップ母数に対して、正規分布を事前分布として利用する
  - NOCALIB: 推定を行わず、既存の分析結果を利用する



# 実際の記述例

```
1 |          NITEMS=1, NCH=0, NPLAT=0, ↓  
2 | >CALIB  GRADED, LOGISTIC, CYCLES=20, NQPTS=31, NEWTON=20, ITEMFIT=31; ↓  
3 | >SCORE  F&P  NAME=C02. ↓
```

- GRADED, LOGISTIC
  - ロジスティック関数を用いた段階反応モデルによる推定を行う
- 推定が上手くいかない場合、細かい設定をいじる必要があるが、こうすべきという指針は無い
  - CYCLES: EMアルゴリズムにおける繰り返し回数
  - NQPTS: EMアルゴリズムおよびガウス・ニュートン法で用いる求積点の個数
  - NEWTON: ガウス・ニュートン法の最大繰り返し回数
  - ITEMFIT: 適合度計算の際に、回答者を大雑把にいくつかのグループに分けるか

# 被験者母数の推定に関する指定

```
>SCORE EAP/MLE/WML,  
SMEAN=..., SSD=...,  
NOSCORE, ... ;
```

- 「>SCORE」以下に、被験者母数の推定に関する指定を行う
  - EAP/MLE/WML: ベイズ推定 / 最尤推定 / 重み付き最尤推定
  - SMEAN, SSD: 推定された被験者母数を、特定の平均(SMEAN)、分散(SSD)を持つように変換する
  - NOSCORE: 被験者母数の推定を行わない

# 実際の記述例

```
2 | >CALIB GRADED, LOGISTIC  
3 | >SCORE EAP, NAME=CO2;↓  
4 | [EOF]
```

- EAP
  - ベイズ推定を行う
- NAME
  - 「CO2」という名前を付ける

# テスト構成が複雑な場合の注意

## スクリプト全体

>TEST1

>BLOCK (TEST1のブロック指定)

>CALIB (TEST1の項目母数推定)

>SCORE (TEST1の被験者母数推定)

>TEST2

>BLOCK1 (TEST2のブロック1の指定)

>BLOCK2 (TEST2のブロック2の指定)

>CALIB (TEST2の項目母数推定)

>SCORE (TEST2の被験者母数推定)

>TEST3

>BLOCK1 (TEST3のブロック1の指定)

>BLOCK2 (TEST3のブロック2の指定)

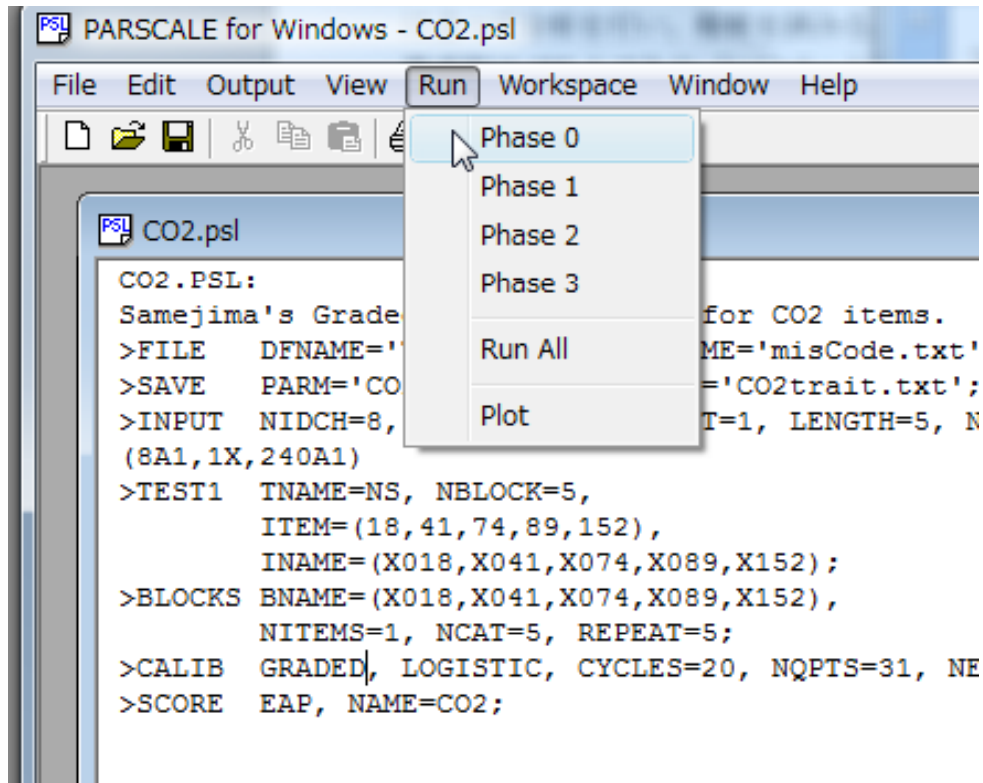
>CALIB (TEST3の項目母数推定)

>SCORE (TEST3の被験者母数推定)

- 「CALIB」と「SCORE」は、1つのテストごとに1セットずつ必要になる
  - 母数推定はテスト単位で行われる

# 推定の実行と結果の解釈

# 推定の実行



- PARSCALEを起動して、スクリプトを読み込む
  - データファイル等はスクリプトと同じディレクトリに用意する
- メニューの「Run」から推定を行う
  - Phase0～3の4段階があるので、0から順に実行していく
  - 各結果は「.PH0」～「.PH3」の拡張子で保存・出力される

# Phase 0

## スクリプトとデータの読み込み

```
OBSERVATION #      1
GROUP: 1
ID: 20081001
INPUT RESPONSES:  4  3  2  3  4  3  4  5  4  4  5  4  1  3  4  2  3  2  2  5
                  4  4  2  1  2  4  2  2  4  3  5  1  2  2  2  4  3  2  2  3
                  1  4  4  2  4  3  4  3  3  5  3  3  3  4  4  3  3  4  3  1
                  2  3  3  3  1  2  1  5  2  3  4  3  4  4  1  4  3  2  2  4
                  3  3  4  1  1  1  3  1  2  3  3  3  2  4  1  4  5  3  3  3
                  1  4  1  2  4  3  3  4  3  3  1  4  4  4  2  3  4  1  4  5
                  2  2  3  1  4  3  3  1  2  3  2  2  2  1  3  3  3  3  4  3
                  4  5  4  4  3  3  3  3  3  3  1  4  3  1  3  3  2  4  3  3
                  2  3  4  4  3  3  3  4  3  2  3  3  3  2  4  1  5  5  1  3
                  4  4  3  2  3  3  4  3  3  2  3  4  3  3  3  3  4  3  5  2
                  3  3  3  3  2  2  3  2  3  4  2  3  3  3  3  2  4  3  2
                  3  5  3  3  2  3  5  4  3  2  4  2  3  2  3  4  2  3  3  3
```

- 特にエラーが出ない場合，結果はざっと見る程度で構わない
  - 文法エラーなどはここで指摘されるので，問題があればスクリプトを書き直す
  - 出力の一番下に，最初の方のオブザベーションについてデータをどう読み取ったかが表示される

# Phase 1

## 古典的な項目分析

```
BLOCK NO.: 1      NAME: X018
```

---

ITEM	TOTAL	NOT PRESENT	OMIT	CATEGORIES				
				1	2	3	4	5
X018								
FREQ.	671	0	5	74	217	178	149	48
PERC.		0.0	0.7	11.0	32.3	26.5	22.2	7.2

---

- 項目ごとに、欠損および各選択肢の回答頻度が表示される
  - 回答頻度が0の選択肢がある場合、これ以上の分析ができないので、選択肢の併合を行う



# Phase 1

## 古典的な項目分析

BLOCK ITEM	RESPONSE MEAN S.D.*	TOTAL SCORE MEAN S.D.*	PEARSON & POLYSERIAL CORRELATION	INITIAL SLOPE	INITIAL LOCATION I
X018 1 X018	2.820 1.118*	15.009 2.680*	0.028 0.029	0.029	-0.002
CATEGORY		MEAN	S.D.	PARAMETER	
1		14.622	3.199	47.505	
2		15.171	2.439	10.807	
3		14.753	2.604	-13.185	
4		15.315	2.786	-45.127	
5		14.875	2.611		

- 項目得点の平均とSD
  - 通常の集計作業も行うことが可能
- 総合得点と項目得点との相関係数
  - 相関が低い項目は, 1次元性が高くない可能性がある
- 項目母数推定の初期値
  - Phase2の推定が上手くいかない場合, ここがおかしくないかをチェック

## Phase 2

# 項目母数の推定

- 上半分は推定過程の履歴なので、細かく見なくても良い
  - この段階でエラーが出る場合、「>CALIB」コマンドのオプションを調整する必要がある
  - または各種母数の推定の初期値を指定する
- 下半分に、項目母数の推定値などが出力される
  - 推定モデルによって微妙に出力が異なる

# 一般化部分採点モデル

- 特性値  $\theta$  を持つ回答者  $j$  が項目  $i$  のカテゴリ  $k$  ( $=1, \dots, K$ ) を選択する確率を

$$P_{ik}(\theta_j) = \frac{\exp \sum_{k'=0}^k [a_i (\theta_j - \beta_{ik'})]}{\sum_{k=0}^K \exp \sum_{k'=0}^k [a_i (\theta_j - \beta_{ik'})]}$$

- $\theta$  = 被験者母数: 測定対象である潜在特性
- $a$  = 識別力母数: 選択率が  $\theta$  の変化にどの程度敏感に反応するか
- $\beta$  = ステップ母数: カテゴリ  $k-1$  よりもカテゴリ  $k$  を選ぶことの難しさ

# 一般化部分採点モデルの出力

ブロックごとの  
ステップ(閾値)

ITEM BLOCK 5 X152

SCORING FUNCTION :	1.000	2.000	3.000	4.000	5.000
STEP PARAMTER :	0.000	1.675	0.228	-0.038	-1.866
S.E. :	0.000	0.186	0.133	0.119	0.141

ITEM	BLOCK	SLOPE	S.E.	LOCATION	S.E.	GUESSING	S.E.
X018	1	0.007	0.001	8.432	2.986	0.000	0.000
X041	2	0.228	0.016	0.439	0.119	0.000	0.000
X074	3	0.766	0.072	-0.184	0.064	0.000	0.000
X089	4	0.236	0.017	-0.078	0.109	0.000	0.000
X152	5	0.543	0.060	-0.311	0.063	0.000	0.000

識別力

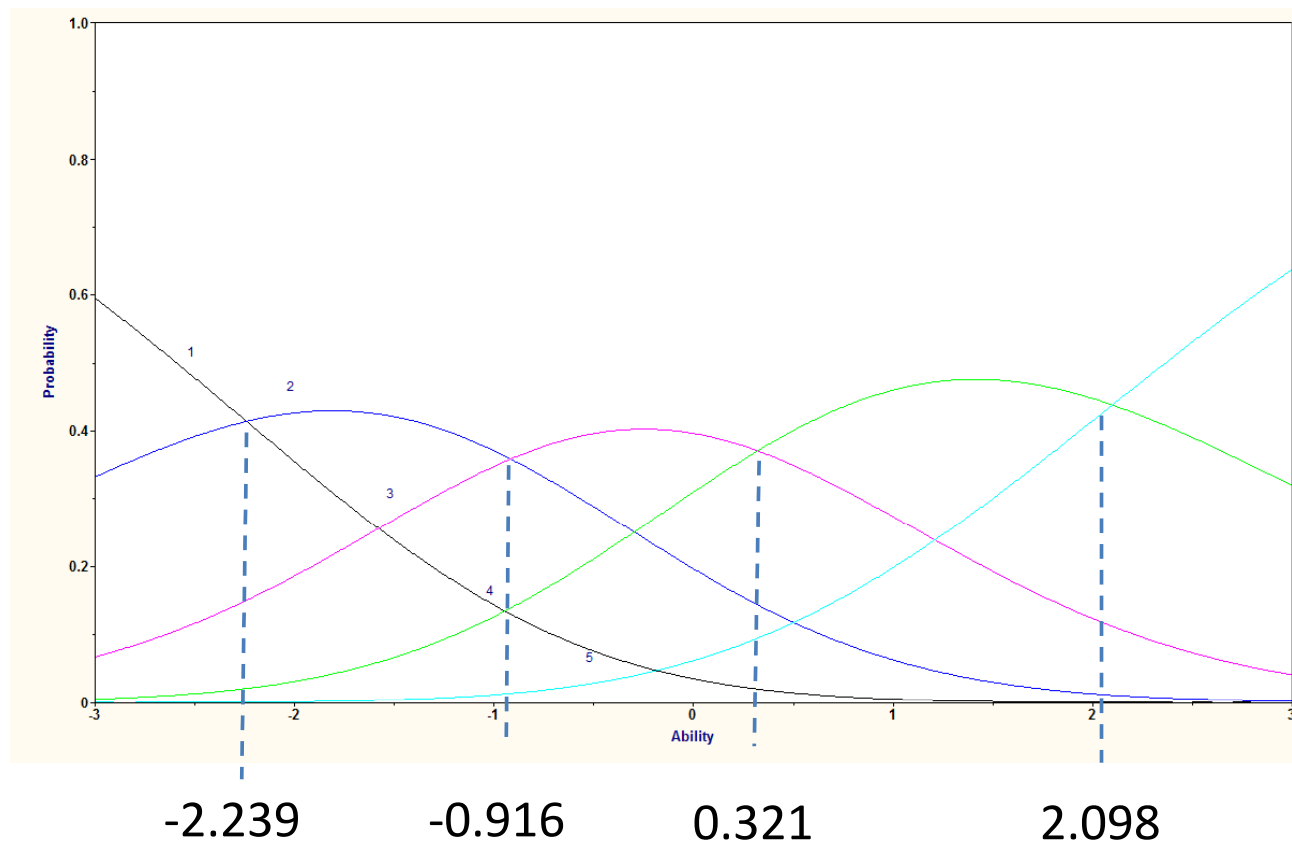
位置

当て推量



# ステップ母数の解釈

- ステップ母数 $\beta_x$ は, 1つ下の選択肢よりも選択肢 $x$ を選びやすくなる境界を表す
  - 項目特性曲線において, 1つ下のカテゴリの曲線と交差する点の $\theta$ の値に相当する



# 段階反応モデル(1)

- 特性値  $\theta$ を持つ回答者 $j$ が項目 $i$ のカテゴリ $k(=1, \dots, K)$ 以上を選択する確率を

$$P_{ik}^*(\theta_j) = \frac{1}{1 + \exp[-Da_i(\theta_j - b_{ik}^*)]}$$

- $\theta$  = 被験者母数: 測定対象である潜在特性
- $a$  = 識別力母数: 選択率が  $\theta$  の変化にどの程度敏感に反応するか
- $b^*$  = 困難度母数: カテゴリ $k$ 以上を選ぶことの難しさ

# 段階反応モデル(2)

- ただし, 以下は特別とする

$$P_{i1}^*(\theta_j) = 1$$

$$P_{i(K+1)}^*(\theta_j) = 0$$

- 以上を利用して, カテゴリ $k$ を選択する確率を  
下のように求める

$$P_{ik}(\theta_j) = P_{ik}^*(\theta_j) - P_{i(k+1)}^*(\theta_j)$$



# 段階反応モデルの出力

ITEM	BLOCK	SLOPE	S.E.	LOCATION	S.E.	GUESSING	S.E.
X018	1	0.028	0.001	5.181	1.340	0.000	0.000
X041	2	0.441	0.017	0.257	0.121	0.000	0.000
X074	3	1.126	0.055	-0.222	0.069	0.000	0.000
X089	4	0.486	0.019	-0.062	0.104	0.000	0.000
X152	5	0.931	0.058	-0.324	0.059	0.000	0.000

ブロックごとの  
カテゴリ困難度

識別力

位置

当て推量

# 項目ごとの困難度の算出(1)

- 項目ごとに「位置-カテゴリ困難度」で, BCC(境界特性曲線)の困難度母数が算出される
  - 一般的に $b^*$ という記号で表されることが多い
- 例えば以下のような項目の場合,
  - $b^*2 = -0.222 - 2.194 = -2.416$
  - $b^*3 = -0.222 - 0.735 = -0.957$
  - $b^*4 = -0.222 - (-0.62) = 0.398$
  - $b^*5 = -0.222 - (-2.309) = 2.087$

ITEM BLOCK 3 X074

CATEGORY PARAMETER	:	2.194	0.735	-0.620	-2.309
S.E.	:	0.095	0.054	0.049	0.081

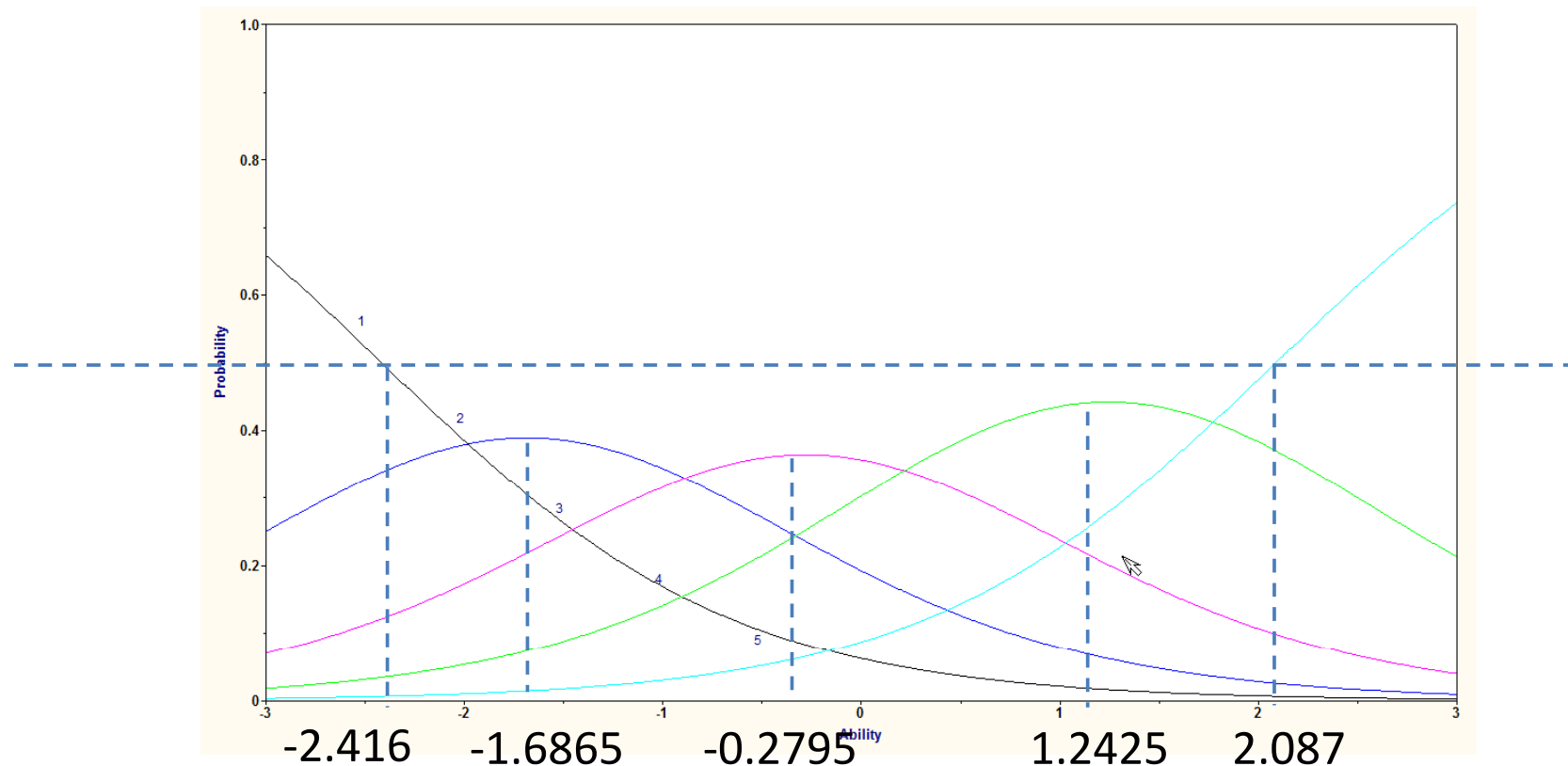
X074	3	1.126	0.055	-0.222	0.069	0.000	0.000
------	---	-------	-------	--------	-------	-------	-------

# 項目ごとの困難度の算出(2)

- 以下のような計算をすることで, IRCCC(項目反応カテゴリ特性曲線)における各カテゴリの曲線と対応する困難度母数( $b$ )を求められる
  - $b_1 = b^*2 = -2.416$
  - $b_2 = (b^*3 - b^*2) / 2 = -1.6865$
  - $b_3 = (b^*4 - b^*3) / 2 = -0.2795$
  - $b_4 = (b^*5 - b^*4) / 2 = 1.2425$
  - $b_5 = b^*5 = 2.087$

# 困難度母数の解釈

- 両端のカテゴリの困難度母数は、各選択肢への回答率が50%になる $\theta$ の値を示す
- それ以外のカテゴリの困難度母数は、その選択肢への回答率が最も高くなる $\theta$ の値を示す



# どちらを選ぶべきか？

- 一般化部分採点モデルと段階反応モデルは、どちらも利用されている
  - 基本的には好みで選んで構わない
- 段階反応モデルは、 $K$ 個の選択肢に対して $K$ 個の困難度母数を推定する
  - モデルとして自然で無理がない
- 一般化部分採点モデルは、回答に $K$ 個の選択肢があった場合、困難度として $(K-1)$ 個のステップ母数を推定する
  - 段階反応モデルよりも母数が少ないので、標本数が少ない場合でも推定が安定しやすい

# 項目・テスト全体の適合度

ITEM FIT STATISTICS

BLOCK	ITEM	CHI-SQUARE	D.F.	PROB.
X018	X018	144.31169	62.	0.000
X041	X041	96.51670	54.	0.000
X074	X074	168.22560	46.	0.000
X089	X089	110.11793	55.	0.000
X152	X152	132.41061	49.	0.000
TOTAL		651.58252	266.	0.000

- 帰無仮説は「IRTモデルがデータに適合している」
  - 有意確率「PROB.」が有意水準よりも低い場合, IRTモデルの適合が悪いことになる
  - あまり利用されていないので, 気にしない方向で

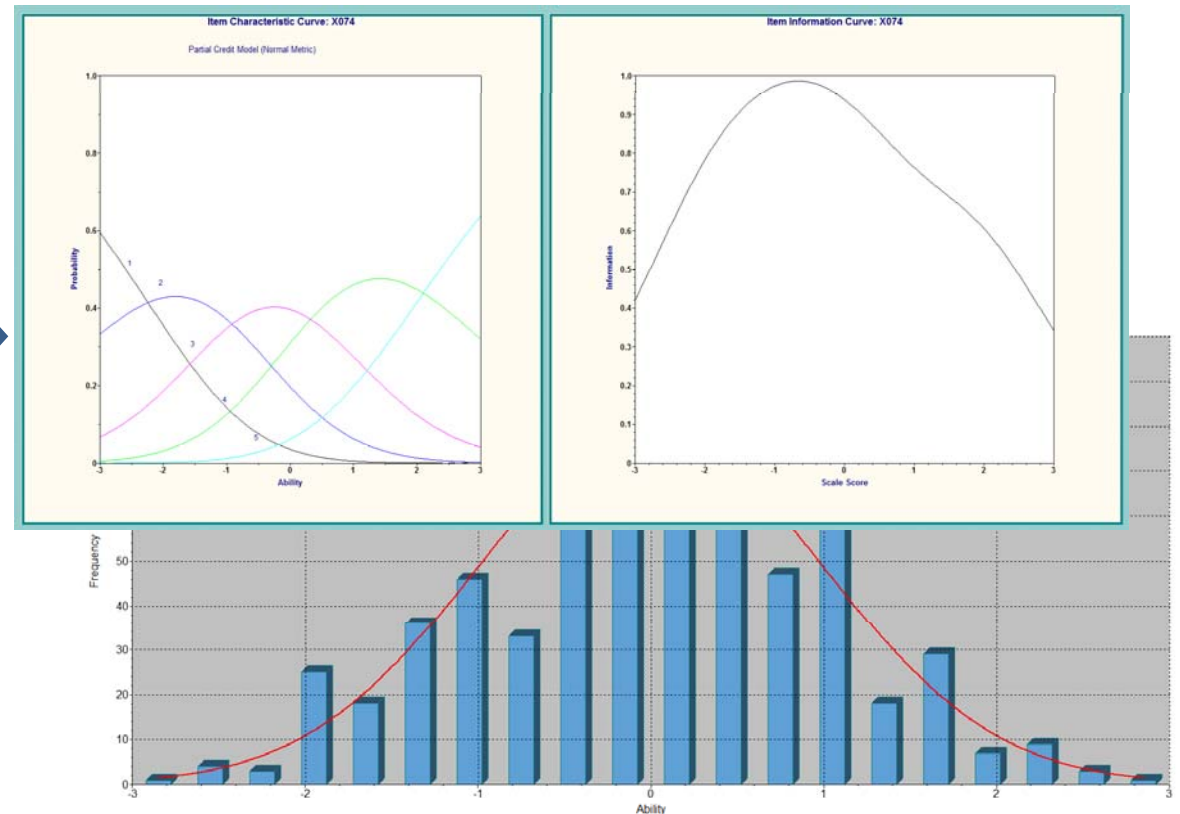
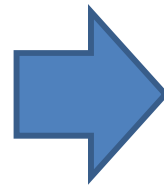
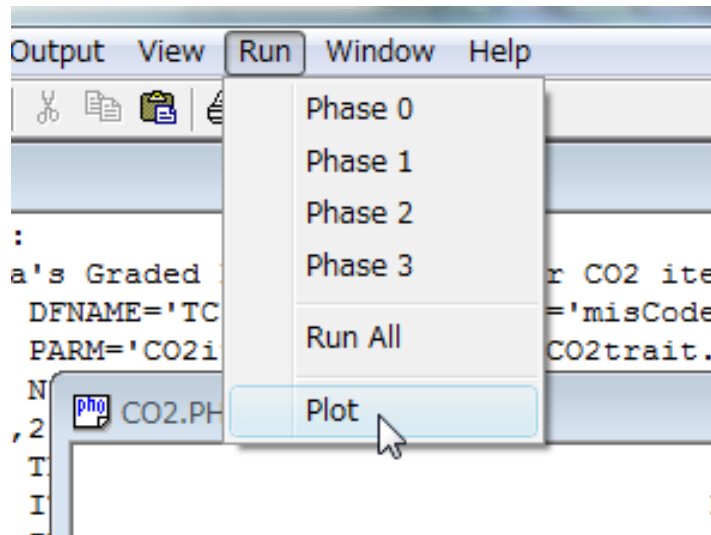
# Phase 3

## 被験者母数の推定

SUBJECT IDENTIFICATION			WEIGHT/FREQUENCY					
SCORE NAME	GROUP	WEIGHT	MEAN	CATEGORY	ATTEMPTS	ABILITY	S.E.	
20081001		1		GROUP 01	1.00			
1 CO2	1	1.00	2.60		1.00	0.1887	0.5869	
20081002		2		GROUP 01	1.00			
1 CO2	1	1.00	2.60		1.00	-0.5970	0.5750	
20081003		3		GROUP 01	1.00			
1 CO2	1	1.00	2.80		1.00	-0.7184	0.5762	
20081004		4		GROUP 01	1.00			
1 CO2	1	1.00	3.40		1.00	0.8081	0.6136	

- 左端にデータとして入力したID
- 右端の「ABILITY」が被験者母数  
– 隣の「S.E.」が推定の標準誤差

# 基本的なグラフの描画



- 全推定が完了した後で上部メニューの「Run」→「Plot」を選ぶと、いくつかの基本的なグラフを描くことができる
  - － 項目特性曲線
  - － 項目／テスト情報量曲線
  - － 平均項目得点と推定された被験者母数との対応
  - － 推定された被験者母数の分布