

## 日本心理学会第82回大会 チュートリアルワークショップ

### Mplusによる多変量解析： 心理学研究における実践

中部大学 伊藤大幸  
愛知淑徳大学 谷 伊織

## 企画趣旨

### • 構造方程式モデリング (SEM)

- 社会科学領域で用いられる解析手法の大部分を下位モデルとして扱える包括的な解析の枠組み
  - 一般線形モデル (t検定、分散分析、重回帰分析など)
  - 一般化線形モデル (ロジスティック回帰分析など)
  - 因子分析
  - マルチレベルモデル
  - 項目反応理論 など
- SEMを理解することは、現在の心理統計の技術で何ができる、何ができないかのラインを知ること、ほぼイコール
- 基本原理を理解していないと誤用につながりやすい
  - 心理学研究における実践例と注意点を示す

## 企画趣旨

### • Mplus

- 数あるSEMのソフトウェアの中でも、最も柔軟性が高く、広範囲のモデルや推定法を扱える
  - ロバスト推定法
  - 特殊な変数 (カテゴリカル変数、打ち切り変数、回数データなど) を含むモデル
  - マルチレベルSEM
  - 混合分布モデル など
- Mplusの使い方そのものではなく、Mplusを利用して、心理学研究上の様々なリサーチクエストをどのように解決していくかを議論する

## 企画趣旨

### • 心理学・社会科学研究のための構造方程式モデリング：Mplusによる実践

- 村上 隆・行廣隆次 (監修) 伊藤大幸 (編著) 谷 伊織・平島太郎 (著) (ナカニシヤ出版)
- 基礎編 2018年9月 刊行
- 発展編 2019年夏 刊行予定

#### 【基礎編】

SEMの基本原理・パス解析・探索的因子分析・確認的因子分析・フルSEM (潜在変数間のパス解析)・カテゴリカルデータの分析・トラブルシューティング等

#### 【発展編】

多母集団同時分析・発展的な因子分析 (高次、階層)・縦断データの分析・マルチレベルモデル・混合分布モデル等

## 今日の内容

### • Mplusの基本 (谷) 15分

### • 解析の実践

- カテゴリカルデータの分析 (伊藤) 15分
- 多母集団同時分析 (伊藤) 15分
- 発展的な因子分析 (谷) 20分
  - 高次因子分析
  - 階層的因子分析
- 縦断データの分析 (伊藤) 50分
  - 潜在成長モデル
  - 交差遅延モデル

日本心理学会第82回 TWS

## Mplusによる多変量解析： 心理学研究における実践 —Mplusの基本— (10min)

谷伊織 (愛知淑徳大学)

Mplus

6

## Mplusとは

- Muthénらによって開発された構造方程式モデリングのソフトウェア
  - 広範囲のモデルを扱うことができる
  - SPSSやAMOSなどのGUIではない R、SAS、CALISなどと同じく、CUI
- 初学者にはGUIが便利のように思えるが、多くの指標をともなう測定モデルでは大変長期的にはCUIに慣れる方が便利

7

## Mplusの特徴

- よく使用する解析が最低限の命令で実行可能
- 基本的な解析のために覚えるルールは少ない
- モデルをパス図として確認することも可能
- 初学者が犯しやすいモデル指定上のミス (例えば識別の問題) が極力起こらないよう、仕様上の工夫がなされている
- サポートが充実、比較的安価 (学生版125\$~175\$、教員595~895\$)

8

## ソフトウェアの準備



- Mplusで検索すればwebサイトから購入可能
- 日本語は使用不可  
→ファイル名や場所に注意

9

## データセットの準備

	A	B	C	D	E	F	G
1	ID	gender	grade	a1	a2	a3	a4
2	40001	男	1	3	1	4	2
3	40002	男	1	2	2	2	3
4	40003	男	1	3	1	2	2
5	40004	男	1		2	3	1
6	40005	男	1	2	2	2	2
7	40006	男	1	3	1	2	3
8	40007	男	1	2	2	3	3
9	40008	男	1	3	2	3	2
10	40009	女	1	3	1	3	2
11	40010	女	1	3	2		2
12	40011	女	1	1	3	3	3

- 元のデータはExcel上での入力OK

10

## データセット 注意事項

- データは全て数字に置換  
(例:「男」=1、「女」=2)
- 桁数が8桁を超えないように
- 変数のラベルを示す最初の行は削除  
→シンタックスにて後ほど指定
- 欠損値にはコードを割り当てる(例えば999)  
→実データには含まれない値を使用

11

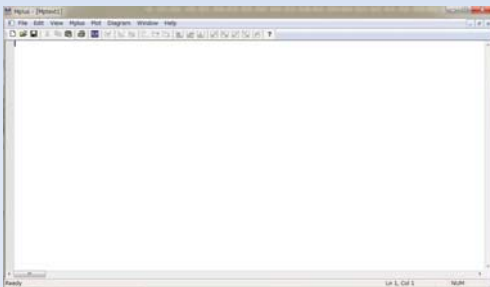
## データセット できあがり

	A	B	C	D	E	F	G
1	40001	1	1	3	1	4	2
2	40002	1	1	2	2	2	3
3	40003	1	1	3	1	2	2
4	40004	1	1	999	2	3	1
5	40005	1	1	2	2	2	2
6	40006	1	1	3	1	2	3
7	40007	1	1	2	2	3	3
8	40008	1	1	3	2	3	2
9	40009	2	1	3	1	3	2
10	40010	2	1	3	2	999	2
11	40011	2	1	1	3	3	3

- ファイルは「テキスト(タブ区切り)」で保存
- ファイル名に日本語や全角文字を含まない  
(補足) 共分散行列等の要約データセットにも対応

12

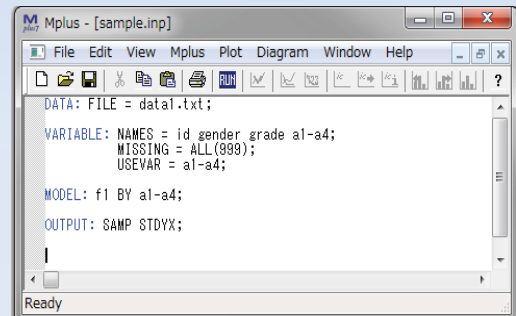
## 起動



- Mplus Editorを起動すると上のような画面が見れる

13

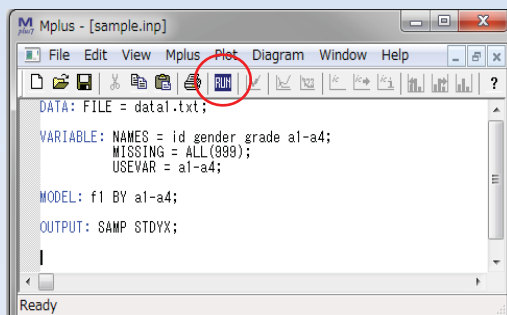
## シンタックスの入力



- 上のように、シンタックスを入力する

14

## 実行



15

## 出力

SUMMARY OF ANALYSIS	
Number of groups	1
Number of observations	1197
Number of dependent variables	1
Number of independent variables	1
Number of continuous latent variables	0
Observed dependent variables	
Continuous	DEPT
Observed independent variables	
Estimator	MLF1
Information matrix	OBSERVED
Maximum number of iterations	1000
Convergence criterion	0.5000-04
Maximum number of steepest descent iterations	30
Maximum number of iterations for H1	2000
Convergence criterion for H1	0.1000-03

- テキスト、図による出力が得られる

16

# Mplusによる 基本的な解析

## 潜在変数を含むパス解析

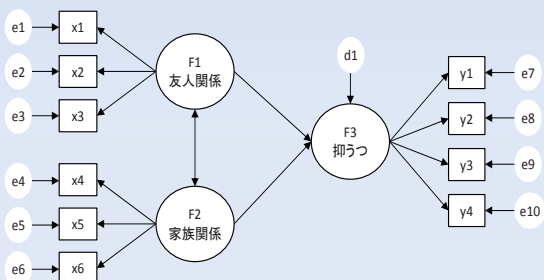
17

## データについて

- Mplusを使いたいシンプルな解析例
  - 小中学生を対象とした仮想データ
  - メンタルヘルスや問題行動に関する調査
- 今回のデータは非公開ですが・・・
  - 簡単なシンタックスは本資料で十分に活用可能
  - 書籍に対応した学習用のデータを公開中

18

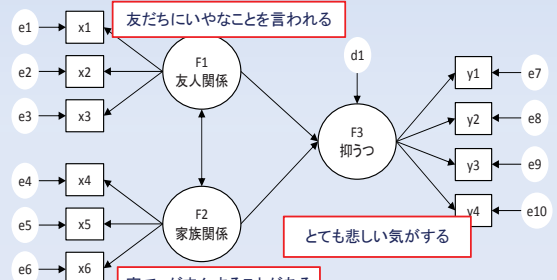
### モデル1: 対人関係が抑うつに与える影響



- シンタックスの記述方法の説明をかねて、横断データの解析例を挙げる

19

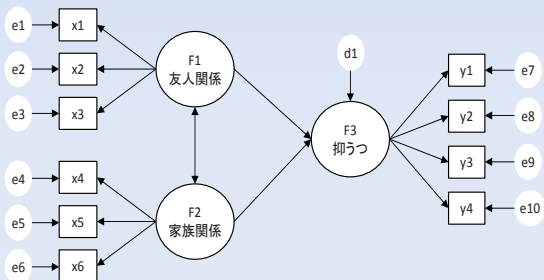
### モデル1: 対人関係が抑うつに与える影響



- 家族と友人関係での問題が抑うつに影響するというモデルを検討

20

### モデル1: 対人関係が抑うつに与える影響



- いずれも質問紙による本人評定(1230名)
- 潜在変数を含むパス解析によって検討

21

## シンタックスのルール

- Mplusのシンタックスは・・・  
コマンドとオプションによって構成
- コマンドは命令の種類を表す
- オプションは命令の内容を表す

22

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVARIABLES = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
        f2 BY x4-x6;
        f3 BY y1-y4;
        f3 ON f1 f2;
        f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
    
```

- TITLE、DATA、VARIABLE、ANALYSIS、MODEL、OUTPUT コマンドが使用

23

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVARIABLES = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
        f2 BY x4-x6;
        f3 BY y1-y4;
        f3 ON f1 f2;
        !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
    
```

- DATAコマンドの下位にFILEオプション、
- VARIABLEコマンドの下位にNAMES、MISSING等のオプションが使用

24

## 覚えておくこと

- コマンドはコロン(:)で終わり、オプションはセミコロン(; )で終わる (※オプションによってはまとめることも可)
- 各行の命令は90字まで 超えるときには改行する
- 個々の変数に与える名称は8字まで
- Mplusは日本語を処理できない
- Mplusは大文字と小文字を区別しない

25

## 覚えておくと便利

- 一部のコマンドやオプションは短く省略可能
- 一連の変数を表すためにハイフン(-)を使用できる 例「a1-a4」は、「a1, a2, a3, a4」
- 特定のオプションで、全変数を表す「ALL」を使用可能
- コマンドの順序は、ほとんどの場合任意、どのコマンドから書き始めてもOK
- 任意の位置にコメントを挿入できる →「!»以降の記述は解析に使用されない

26

## 基本的なコマンドとオプション

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVAR = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
        f2 BY x4-x6;
        f3 BY y1-y4;
        f3 ON f1 f2;
        !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
Ready
    
```

27

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVAR = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
        f2 BY x4-x6;
        f3 BY y1-y4;
        f3 ON f1 f2;
        !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
Ready
    
```

- 「TITLE」は、タイトルを指定するコマンド
- 解析には使用されない

28

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVAR = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
        f2 BY x4-x6;
        f3 BY y1-y4;
        f3 ON f1 f2;
        !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
Ready
    
```

- 「DATA」は、データに関する指定を行うコマンド
  - 「FILE」はファイルを指定するオプション
- ※シンタックスとデータファイルを同じフォルダに保存すれば、ファイルのパスを指定する必要がない

29

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVAR = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
        f2 BY x4-x6;
        f3 BY y1-y4;
        f3 ON f1 f2;
        !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
Ready
    
```

- 「VARIABLE」コマンドは、変数に関する指定を行う
- 「NAMES」オプションは、観測変数の名称を指定、

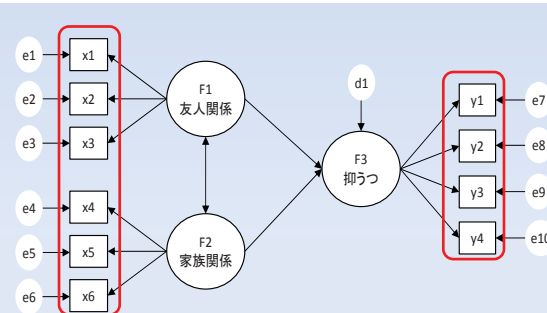
30

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVAR = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
        f2 BY x4-x6;
        f3 BY y1-y4;
        f3 ON f1 f2;
        !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
Ready
    
```

- 「USEVAR」は、モデルに含める観測変数を指定するオプション
- 分析に使わないものは指定しない

31



- 観測変数として、x1-x6, y1-y4が指定され、分析に用いられている
- 他にも、id(番号)、gender(性別)、grade(学年)が指定

32

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVARIABLES = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
       f2 BY x4-x6;
       f3 BY y1-y4;
       f3 ON f1 f2;
       !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
Ready

```

- 「MISSING」オプションは、欠損値コードを指定
- Mplusでは特に指定しなければ、デフォルトで完全情報最尤法によって欠損値を処理

33

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVARIABLES = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
       f2 BY x4-x6;
       f3 BY y1-y4;
       f3 ON f1 f2;
       !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
Ready

```

- 「ANALYSIS」コマンドは、分析の詳細を指定
- 「TYPE」オプションは、分析のタイプを指定

34

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVARIABLES = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
       f2 BY x4-x6;
       f3 BY y1-y4;
       f3 ON f1 f2;
       !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
Ready

```

- 通常のパス解析や確認的因子分析は「GENERAL」というタイプに含まれる
- 特に指定しなければデフォルトで「GENERAL」

35

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVARIABLES = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
       f2 BY x4-x6;
       f3 BY y1-y4;
       f3 ON f1 f2;
       !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
Ready

```

- 「ESTIMATOR」オプションは推定法を指定
- 種類が多い
- ここではMLR(ロバスト最尤法)を選択

36

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVARIABLES = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
       f2 BY x4-x6;
       f3 BY y1-y4;
       f3 ON f1 f2;
       !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
Ready

```

- 「MODEL」コマンドは、モデルに関する指定を行う
- シンタックスの中でも特に重要

37

```

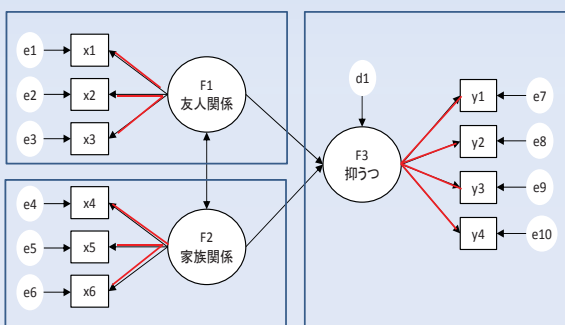
Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVARIABLES = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
       f2 BY x4-x6;
       f3 BY y1-y4;
       f3 ON f1 f2;
       !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
Ready

```

- 「BY」オプションは、「measured by」を意味
- 測定方程式(因子分析)の指定を行う
- 「因子 BY 指標」

38

## 測定モデル「BY」



39

```

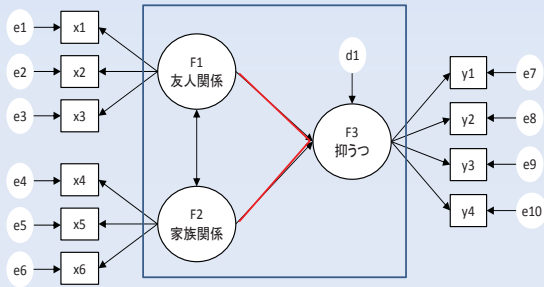
Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVARIABLES = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
       f2 BY x4-x6;
       f3 BY y1-y4;
       f3 ON f1 f2;
       !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
Ready

```

- 「ON」オプションは「regressed on」を意味
- 構造方程式(回帰分析、パス解析)の指定を行う
- 「結果変数 ON 原因変数」

40

## 構造モデル「ON」



41

## 注意

- 「BY」オプションや「ON」オプションでは、明示しないが、全ての従属変数(指標含む)に誤差変数が自動的に仮定される
- 「BY」オプションでは原因変数(因子)が先で結果変数(指標)が後という形式
- 「ON」オプションでは結果変数が先で原因変数が後という形式
- 慣れないうちは、この順序をよく間違える
- 「BYは因子が先、ONは結果変数が先」

42

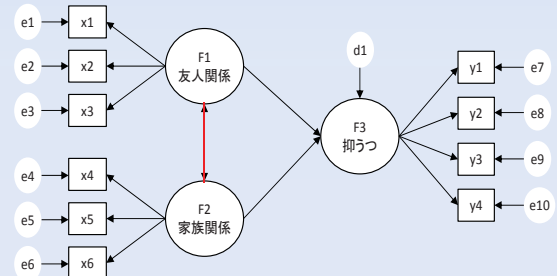
```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVARIABLES = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
       f2 BY x4-x6;
       f3 BY y1-y4;
       f3 ON f1 f2;
       !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
    
```

- 「WITH」オプションは「correlated with」を意味
- 変数間の相関(共分散)を指定
- ここでは独立変数であるf1とf2の相関を仮定

43

## 変数間の共分散「WITH」



44

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVARIABLES = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
       f2 BY x4-x6;
       f3 BY y1-y4;
       f3 ON f1 f2;
       !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
    
```

- 「OUTPUT」コマンドは、出力に関する指定
- 「SAMP」は、「SAMPSTAT」オプションの略
- 観測変数の記述統計を出力

45

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVARIABLES = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
       f2 BY x4-x6;
       f3 BY y1-y4;
       f3 ON f1 f2;
       !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
    
```

- 「STDYX」オプションは、パラメータの標準化推定値を出力
- 「RESIDUAL」オプションは、モデルの部分的適合を評価するための残差行列を出力

46

```

Mplus - [sample2.inp]
File Edit View Mplus Plot Diagram Window Help
TITLE: sample syntax
DATA: FILE = data2.txt;
VARIABLE: NAMES = id gender grade x1-x6 y1-y4;
          MISSING = ALL(999);
          USEVARIABLES = x1-x6 y1-y4;
ANALYSIS: !TYPE = GENERAL; default
          ESTIMATOR = MLR;
MODEL: f1 BY x1-x3;
       f2 BY x4-x6;
       f3 BY y1-y4;
       f3 ON f1 f2;
       !f1 WITH f2; default
OUTPUT: SAMP STDYX RESIDUAL MOD;
    
```

- 「MOD」は修正指標を出力

47

## 注意

- モデルの全体的適合に関する指標(CFI, RMSEAなど)はデフォルトで出力
- 部分的適合を評価するための残差行列や修正指標は、指定する必要がある

48

## シンタックスについて

- 基本的な構造モデルや測定モデルの解析を行うには、ここまでで解説したコマンドやオプションの範囲で対応可能
- コマンドを5つ、それぞれのオプションを3つほど覚えておけば十分に使用できる

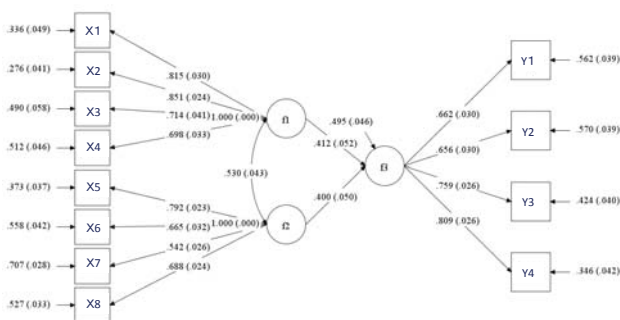
49

## 出力の確認

- 実行すると、入力ファイル.outという出力ファイルが画面に表示
- 正しくモデル指定ができたか確認するため、メニューバーの「Diagram→View diagram」を選択、パス図を確認
- 「View→STDYX estimates」を選択すると標準化係数が表示

50

## パス図



- 視覚的にモデルの確認が可能

51

## 出力

```
SUMMARY OF ANALYSIS
Number of groups                1
Number of observations          1230
Number of dependent variables   12
Number of independent variables  0
Number of continuous latent variables  3

Observed dependent variables
Continuous
  x1    x2    x3    x4    x5    x6
  x7    x8    y1    y2    y3    y4

Continuous latent variables
  F1    F2    F3

Estimator                       MLR
Information matrix              OBSERVED
Maximum number of iterations    1000
Convergence criterion           0.500E-04
Maximum number of steepest descent iterations  20
Maximum number of iterations for HI  2000
Convergence criterion for HI    0.100E-03

Input data file(s)
data2.txt

Input data format  FREE

Number of missing data patterns  12
```

52

## SUMMARY OF ANALYSIS

- Number of groups 集団数 1
- Number of observations 1,230名
- Number of dependent variables 従属変数の数 12
- Number of independent variables 独立変数の数 0
- Number of continuous latent variables 連続的な潜在変数の数 3

53

## SUMMARY OF ANALYSIS

- Observed dependent variables 従属変数の一覧
- Observed independent variables 独立変数の一覧
- Estimator 推定法 (MLR)
- Information matrix以降・・・推定の詳細な設定 行列の種類、反復の最大数、収束の基準など  
- 設定は変更可能だが、今回は指定しないのでデフォルト設定が適用

54

## SUMMARY OF ANALYSIS

- Input data file(s) 読み込んだファイル名
- Input data format FREE フォーマット
- Number of missing data patterns 欠損データ数、欠損パターン

55

## SAMPLE STATISTICS

```
SAMPLE STATISTICS
ESTIMATED SAMPLE STATISTICS
Means
  x1    x2    x3    x4    x5
  1  0.138  0.110  0.088  0.185  0.256
Means
  x6    x7    x8    y1    y2
  1  0.170  0.859  0.491  0.314  0.244
Means
  x3    x4
  1  0.254  0.195
```

- SAMPLE STATISTICSが表示 (OUTPUTコマンドのSAMPオプション)
- Means 観測変数の平均値

56

## SAMPLE STATISTICS

	Covariances x1	x2	x3	x4	x5
x1	0.150				
x2	0.095	0.122			
x3	0.069	0.069	0.103		
x4	0.093	0.090	0.083	0.198	
x5	0.075	0.061	0.038	0.066	0.279
x6	0.057	0.055	0.033	0.050	0.123
x7	0.057	0.056	0.032	0.071	0.154
x8	0.091	0.079	0.042	0.086	0.204
y1	0.070	0.053	0.037	0.068	0.098
y2	0.085	0.060	0.037	0.057	0.103
y3	0.096	0.078	0.053	0.091	0.087
y4	0.077	0.061	0.047	0.065	0.091

- Covariances 分散共分散行列
- Mplusは標本分散を出力

57

## SAMPLE STATISTICS

	Correlations x1	x2	x3	x4	x5
x1	1.000				
x2	0.703	1.000			
x3	0.552	0.616	1.000		
x4	0.536	0.580	0.580	1.000	
x5	0.366	0.332	0.226	0.282	1.000
x6	0.338	0.360	0.234	0.258	0.539
x7	0.206	0.225	0.140	0.222	0.410
x8	0.335	0.323	0.187	0.278	0.553
y1	0.327	0.275	0.211	0.278	0.339
y2	0.434	0.343	0.233	0.256	0.390
y3	0.489	0.444	0.326	0.406	0.329
y4	0.432	0.380	0.320	0.317	0.373

- Correlations 相関行列

58

## MODEL FIT INFORMATION

### MODEL FIT INFORMATION

Number of Free Parameters	39
Loglikelihood	
HO Value	-7125.524
HO Scaling Correction Factor for MLR	2.4974
HI Value	-6972.236
HI Scaling Correction Factor for MLR	2.0551
Information Criteria	
Akaike (AIC)	14329.048
Bayesian (BIC)	14528.524
Sample-Size Adjusted BIC (n* = (n + 2) / 24)	14404.643

59

## MODEL FIT INFORMATION

- Number of Free Parameters 自由推定されたパラメータの数 39  
パス係数、誤差分散、切片が推定
- Loglikelihood 対数尤度  
– 最尤法は、これを最大化するパラメータの組み合わせを探索  
– 対数尤度比検定に使用、単独では解釈しない
- Information Criteria 情報量基準  
– AIC、BIC、サンプルサイズ調整BIC(aBIC)が表示  
– 単独で解釈しない、複数のモデルの比較に使用

60

### Chi-Square Test of Model Fit

Value	178.555*
Degrees of Freedom	51
P-Value	0.0000
Scaling Correction Factor for MLR	1.7170

\* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.

- Chi-Square Test of Model Fit  
カイ二乗検定の結果
- 4行のメッセージ・・・MLR等のロバスト推定法を使用した場合に表示
- 別のモデルとの差異を検定する際に、通常の方法では行えないことを通知

61

### RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.045
90 Percent C.I.	0.038 0.052
Probability RMSEA <= .05	0.863

### CFI/TLI

CFI	0.957
TLI	0.945

### Chi-Square Test of Model Fit for the Baseline Model

Value	3048.538
Degrees of Freedom	66
P-Value	0.0000

### SRMR (Standardized Root Mean Square Residual)

Value	0.037
-------	-------

62

## MODEL FIT INFORMATION

- RMSEA  
– Estimate 0.045 経験的な基準として・・・  
.05以下→良い、.05～.10→中程度、.10以上→悪い適合  
– 90 Percent C.I. RMSEAの90%信頼区間 0.038 0.052  
サンプルサイズが比較的大きいので幅が狭い  
– Probability RMSEA <= .05 0.863  
RMSEAが.05を下回る確率 クロスフィット検定
- CFI 0.957  
経験的な基準として .90以上や.95以上が望ましい
- TLI 0.945  
経験的な基準として .90以上や.95以上が望ましい

63

## MODEL FIT INFORMATION

- Chi-Square Test of Model Fit for the Baseline Model  
独立モデルを帰無仮説としたカイ二乗検定は有意、ただし独立モデルは根拠のない仮説になりやすく、有効な指標にはなりにくい
- SRMR 0.037  
経験的な基準は.05以下や.08以下が望ましい
- いずれの適合度指標も良好

64



# MODEL RESULT

- パラメータ推定値に関する出力
- まず、非標準化係数が出力

MODEL RESULTS					
		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1	x1 BY	1.000	0.000	999.000	999.000
	x2	0.941	0.085	14.427	0.000
	x3	0.723	0.071	10.166	0.000
	x4	0.984	0.073	13.500	0.000
F2	x5 BY	1.000	0.000	999.000	999.000
	x6	0.689	0.049	14.236	0.000
	x7	0.924	0.059	15.691	0.000
	x8	1.147	0.057	20.179	0.000
F3	x1 BY	1.000	0.000	999.000	999.000
	x2	0.905	0.090	15.144	0.000
	x3	1.052	0.069	15.229	0.000
	x4	1.021	0.057	18.035	0.000
F3	F1 ON	0.474	0.059	8.204	0.000
	F2	0.348	0.051	6.774	0.000
F1	F2 WITH	0.070	0.009	7.546	0.000
Intercepts					
x1		0.138	0.011	12.471	0.000
x5		0.110	0.010	11.015	0.000

# MODEL RESULT

- BYやON、WITHのオプションに対応

MODEL RESULTS					
		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1	x1 BY	1.000	0.000	999.000	999.000
	x2	0.941	0.085	14.427	0.000
	x3	0.723	0.071	10.166	0.000
	x4	0.984	0.073	13.500	0.000
F2	x5 BY	1.000	0.000	999.000	999.000
	x6	0.689	0.049	14.236	0.000
	x7	0.924	0.059	15.691	0.000
	x8	1.147	0.057	20.179	0.000
F3	x1 BY	1.000	0.000	999.000	999.000
	x2	0.905	0.090	15.144	0.000
	x3	1.052	0.069	15.229	0.000
	x4	1.021	0.057	18.035	0.000
F3	F1 ON	0.474	0.059	8.204	0.000
	F2	0.348	0.051	6.774	0.000
F1	F2 WITH	0.070	0.009	7.546	0.000
Intercepts					
x1		0.138	0.011	12.471	0.000
x5		0.110	0.010	11.015	0.000

# MODEL RESULT

- 標準誤差

MODEL RESULTS					
		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1	x1 BY	1.000	0.000	999.000	999.000
	x2	0.941	0.085	14.427	0.000
	x3	0.723	0.071	10.166	0.000
	x4	0.984	0.073	13.500	0.000
F2	x5 BY	1.000	0.000	999.000	999.000
	x6	0.689	0.049	14.236	0.000
	x7	0.924	0.059	15.691	0.000
	x8	1.147	0.057	20.179	0.000
F3	x1 BY	1.000	0.000	999.000	999.000
	x2	0.905	0.090	15.144	0.000
	x3	1.052	0.069	15.229	0.000
	x4	1.021	0.057	18.035	0.000
F3	F1 ON	0.474	0.059	8.204	0.000
	F2	0.348	0.051	6.774	0.000
F1	F2 WITH	0.070	0.009	7.546	0.000
Intercepts					
x1		0.138	0.011	12.471	0.000
x5		0.110	0.010	11.015	0.000

# MODEL RESULT

- zとp

MODEL RESULTS					
		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1	x1 BY	1.000	0.000	999.000	999.000
	x2	0.941	0.085	14.427	0.000
	x3	0.723	0.071	10.166	0.000
	x4	0.984	0.073	13.500	0.000
F2	x5 BY	1.000	0.000	999.000	999.000
	x6	0.689	0.049	14.236	0.000
	x7	0.924	0.059	15.691	0.000
	x8	1.147	0.057	20.179	0.000
F3	x1 BY	1.000	0.000	999.000	999.000
	x2	0.905	0.090	15.144	0.000
	x3	1.052	0.069	15.229	0.000
	x4	1.021	0.057	18.035	0.000
F3	F1 ON	0.474	0.059	8.204	0.000
	F2	0.348	0.051	6.774	0.000
F1	F2 WITH	0.070	0.009	7.546	0.000
Intercepts					
x1		0.138	0.011	12.471	0.000
x5		0.110	0.010	11.015	0.000

# MODEL RESULT

- 切片、分散、残差分散

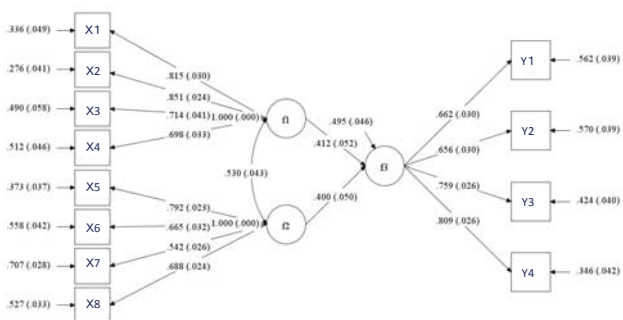
Intercepts					
		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
MAL04		0.138	0.011	12.471	0.000
MAL09		0.110	0.010	11.015	0.000
MAL12		0.088	0.009	9.561	0.000
MAL14		0.185	0.013	14.604	0.000
MAL02		0.256	0.015	17.021	0.000
MAL05		0.170	0.012	13.667	0.000
MAL08		0.659	0.020	32.420	0.000
MAL15		0.491	0.020	24.715	0.000
DEP02		0.314	0.016	19.990	0.000
DEP05		0.244	0.014	17.007	0.000
DEP08		0.254	0.014	17.632	0.000
DEP09		0.195	0.013	14.864	0.000
Variances					
F1		0.100	0.013	7.704	0.000
F2		0.175	0.017	10.372	0.000
Residual Variances					
MAL04		0.051	0.008	6.547	0.000
MAL09		0.034	0.005	6.846	0.000

# STANDARDIZED MODEL RESULTS

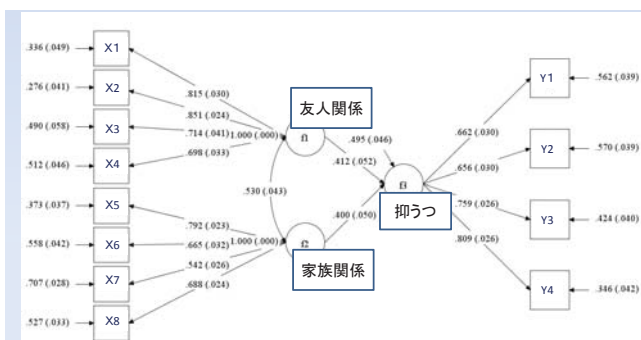
- 標準化係数の出力

STANDARDIZED MODEL RESULTS					
OUTPUTコマンドのSTDYXオプション					
STDYX Standardization					
		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1	x1 BY	0.915	0.030	26.831	0.000
	x2	0.851	0.024	34.951	0.000
	x3	0.714	0.041	17.552	0.000
	x4	0.688	0.033	21.178	0.000
F2	x5 BY	0.792	0.023	34.063	0.000
	x6	0.889	0.032	21.052	0.000
	x7	0.542	0.026	20.858	0.000
	x8	0.688	0.024	28.739	0.000
F3	x1 BY	0.682	0.030	22.357	0.000
	x2	0.856	0.030	21.859	0.000
	x3	0.769	0.026	29.107	0.000
	x4	0.809	0.026	31.387	0.000
F3	F1 ON	0.412	0.052	7.858	0.000
	F2	0.400	0.050	8.038	0.000
F1	F2 WITH	0.530	0.043	12.426	0.000

# パス図



- パス図の標準化係数と同じ
- 有意性については出力を確認



- 友人関係と家族関係 .530 p<.001
- 友人関係から抑うつ .412 p<.001
- 家族関係から抑うつ .400 p<.001

## R-SQUARE QUALITY OF NUMERICAL RESULTS

• R<sup>2</sup>

```
R-SQUARE:
Observed Variable Estimate S.E. Est./S.E. Two-Tailed P-Value
x1 0.664 0.049 13.465 0.000
x2 0.724 0.041 17.445 0.000
x3 0.510 0.059 8.776 0.000
x4 0.488 0.046 10.589 0.000
x5 0.627 0.037 17.031 0.000
x6 0.442 0.042 10.516 0.000
x7 0.293 0.028 10.329 0.000
x8 0.473 0.033 14.369 0.000
y1 0.438 0.039 11.178 0.000
y2 0.430 0.039 10.950 0.000
y3 0.576 0.040 14.654 0.000
y4 0.664 0.042 15.693 0.000

Latent Variable Estimate S.E. Est./S.E. Two-Tailed P-Value
F3 0.505 0.046 10.995 0.000

QUALITY OF NUMERICAL RESULTS:
Condition Number for the Information Matrix (ratio of smallest to largest eigenvalue) 0.745E-03
```

73

## RESIDUAL OUTPUT

```
RESIDUAL OUTPUT:
ESTIMATED MODEL AND RESIDUALS (OBSERVED - ESTIMATED):

Model Estimated Means/Intercepts/Thresholds:
x1 x2 x3 x4 x5
1 0.138 0.110 0.088 0.185 0.256

Model Estimated Means/Intercepts/Thresholds:
x6 x7 x8 y1 y2
1 0.170 0.669 0.491 0.314 0.244

Model Estimated Means/Intercepts/Thresholds:
y3 y4
1 0.254 0.195

Residuals for Means/Intercepts/Thresholds:
x1 x2 x3 x4 x5
1 0.000 0.000 0.000 0.000 0.000
```

74

## RESIDUAL OUTPUT

- OUTPUTコマンドにRESIDUALオプションをとMODオプションが加えることで出力
- サンプルの分散共分散行列と、モデルの制約に基づいて再現される分散共分散行列の間のズレ(残差)を出力
- モデルの部分的な適合を評価する情報として役立つ

75

## MODEL MODIFICATION INDICES

- 修正指標

```
MODEL MODIFICATION INDICES:
NOTE: Modification indices for direct effects of observed dependent variables
regressed on covariates may not be included. To include these, request
MODINDICES (ALL).
Minimum M.I. value for printing the modification index 10.000

M.I. E.P.C. Std E.P.C. StdYX E.P.C.
BY Statements:
F1 BY y3 17.812 0.310 0.098 0.194
F2 BY x3 15.131 -0.121 -0.051 -0.158
F2 BY y2 15.500 0.237 0.099 0.198
F3 BY x1 23.573 0.218 0.079 0.204
F3 BY x3 10.169 -0.125 -0.046 -0.143
WITH Statements:
x4 WITH x3 28.677 0.018 0.018 0.247
x8 WITH x6 15.006 -0.034 -0.034 -0.209
y2 WITH x1 15.715 0.016 0.016 0.182
```

76

## MODEL MODIFICATION INDICES

- 修正指標(modification indices: M.I.)はMODはオプションによって出力
- モデルの部分的な適合の評価に役立つ
- デフォルトでは、修正指標が10以上のみ表示
- 全て表示したいときは「MOD(ALL 0)」

77

## DIAGRAM INFORMATION

### その他の情報

```
DIAGRAM INFORMATION:
Use View Diagram under the Diagram menu in the Mplus Editor to view the diagram.
If running Mplus from the Mplus Diagrammer, the diagram opens automatically.
Diagram output:
c:\Users\koo\Desktop\mplus\sample1.dgm
Beginning Time: 14:51:44
Ending Time: 14:51:45
Elapsed Time: 00:00:01

MUTHEN & MUTHEN:
3463 Stoner Ave.
Los Angeles, CA 90066
Tel: (310) 391-9971
Fax: (310) 391-9971
Web: www.StatModel.com
Support: Support@StatModel.com
Copyright (c) 1998-2014 Muthen & Muthen
```

78

## まとめ

- 基本的なコマンドやオプションを覚えておけば、Mplusによる基本的な構造モデルや測定モデルの解析が可能
- テキスト出力にて得られる情報が充実
- モデルをパス図として確認できる
- 広範囲のモデルを扱うことができるので、必要に応じて覚えていくのが良いのでは

79

## カテゴリカルデータの分析

中部大学 伊藤大幸

## カテゴリカルデータ分析の基本原則

### 変数の種類

従属変数の種類	例	適した解析手法
比尺度	重さ、長さ、時間、絶対温度	線形回帰
間隔尺度	摂氏・華氏、尺度得点	線形回帰
順序尺度	順位、(五件法以下の)項目得点	プロビット回帰
名義尺度	性別、疾患・障害の有無	ロジスティック回帰
回数	事故・災害の件数	ポアソン回帰
打ち切り	天井効果・床効果のある変数	トービット回帰

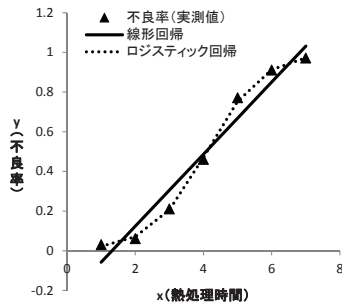
## カテゴリカルデータ分析の基本原則

### なぜ線形回帰ではいけないのか

- 目的の違い
  - 量的変数の分析：「程度」の予測
  - カテゴリ変数の分析：「確率」の予測
- 予測値の範囲
  - 確率は0~1の値しか取らないが、線形回帰の予測値はその範囲を超えることがある
- 非線形の関係
  - 独立変数の「程度」と従属変数の「確率」の関係は非線形

## カテゴリカルデータ分析の基本原則

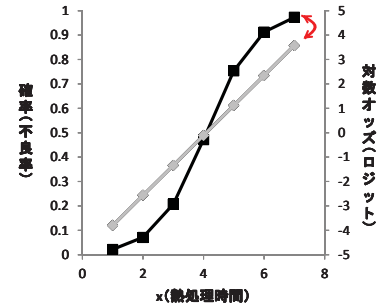
### 確率の予測



## カテゴリカルデータ分析の基本原則

### リンク関数

- ①ロジット変換
- ②回帰分析
- ③逆ロジット変換



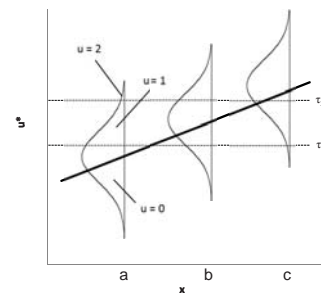
## カテゴリカルデータ分析の基本原則

### リンク関数の種類

	ロジスティック回帰	プロビット回帰
リンク関数	ロジット(対数オッズ)	プロビット(累積正規分布の逆関数)
シグモイド曲線	ロジスティック曲線	累積正規分布曲線
確率密度関数	ロジスティック曲線の導関数	標準正規分布
確率密度関数の平均	0	0
確率密度関数の分散	$\pi^2/3 = 3.29 = (1.81)^2$	1
理論的根拠	質的に異なる2つの群において独立変数が正規分布するとき、独立変数と一方の群に属する確率の関係がロジスティック曲線を成す。	ある反応に関する独立変数の閾値が正規分布するとき、独立変数と反応確率の関係が累積正規分布曲線を成す。
従属変数	(原則的に)名義変数	(原則的に)順序変数
偏回帰係数の解釈	指数変換によってオッズ比として解釈可能(ただし、オッズ比とリスク比は異なる)。	標準正規分布上の偏位の変化量として解釈。標準化係数や限界確率効果などによる解釈も可能。
推定法	(原則的に)最尤法(ML, MLR)	(原則的に)重みづけ最小二乗法(WLS, WLSMV)

## カテゴリカルデータ分析の基本原則

### 順序プロビットモデルと潜在反応変数



## カテゴリカルデータの分析例

### カテゴリカル因子分析

- リッカート法などの心理尺度の評定値は、便宜的に間隔尺度として扱われることが多いが、本来は順序変数
  - 順序プロビットモデルを用いれば、順序変数の性質を考慮して因子分析を行うことが可能
- しかし、順序変数として扱う場合、各評定値の間の閾値を推定する必要が生じるため、(4件法以上の場合)モデルの**儉約性は低下**
  - 適合度と儉約性のバランスを考慮してモデルを選択する必要がある

## カテゴリカルデータの分析例

### データ

- 参加者
  - 大学生521名(男子155名、女子366名)
  - 平均年齢19.4歳(SD=1.9)
- 尺度
  - General Health Questionnaire 12項目版(GHQ-12; Goldberg et al., 1978; 中川・大坊, 1985)
  - 4件法
- モデル
  - ポジティブとネガティブの2因子構造

番号	項目内容	因子
1	集中	ポジティブ
2	心配・不眠	ネガティブ
3	生きがいが	ポジティブ
4	判断	ポジティブ
5	ストレス	ネガティブ
6	困りごと	ネガティブ
7	楽しい	ポジティブ
8	問題解決	ポジティブ
9	ゆううつ	ネガティブ
10	自信喪失	ネガティブ
11	自己否定	ネガティブ
12	幸せ	ポジティブ

## カテゴリカルデータの分析例

### • シンタックス

```
DATA: FILE = data5.2.txt;

VARIABLE: NAMES = id gender age grade data ghq1-ghq12;
MISSING = ALL(999);
USEVAR = ghq1-ghq12;
CATEGORICAL = ghq1-ghq12; !各項目得点を順序変数として指定

MODEL: F1 BY ghq2 ghq5 ghq6 ghq9 ghq10 ghq11;
F2 BY ghq1 ghq3 ghq4 ghq7 ghq8 ghq12;

OUTPUT: SAMP STDYX MOD(0) RESIDUAL;
```

## カテゴリカルデータの分析例

### • 相関係数（一部抜粋）

• 量的変数モデル（Pearsonの積率相関係数）

	GHQ1	GHQ2	GHQ3	GHQ4	GHQ5
GHQ1	1.000				
GHQ2	0.082	1.000			
GHQ3	0.381	0.177	1.000		
GHQ4	0.234	0.103	0.203	1.000	
GHQ5	0.184	0.339	0.169	0.235	1.000

### • 順序変数モデル（ポリコリック相関係数）

	GHQ1	GHQ2	GHQ3	GHQ4	GHQ5
GHQ1					
GHQ2	0.091				
GHQ3	0.469	0.192			
GHQ4	0.277	0.127	0.243		
GHQ5	0.226	0.395	0.206	0.286	

## カテゴリカルデータの分析例

### • 適合度と因子負荷量

	標準化負荷量	
	量的変数 モデル	順序変数 モデル
パラメータ数	37	49
RMSEA	.041	.054
CFI	.970	.979
TLI	.962	.974
SRMR	.047	
WRMR		.722

項目	標準化負荷量	
	量的変数 モデル	順序変数 モデル
F1		
2	.449	.460
5	.692	.761
6	.687	.740
9	.760	.814
10	.717	.788
11	.671	.744
F2		
1	.497	.521
3	.481	.506
4	.503	.597
7	.752	.851
8	.627	.704
12	.587	.652
因子間 相関	.635	.645

## カテゴリカルデータの分析例

### • 3値データの場合

• GHQ12の4件法の評定値のうち、「1」と「2」を統合し、3値データに変換して再分析

	標準化負荷量	
	量的変数 モデル	順序変数 モデル
パラメータ数	37	37
RMSEA	.055	.031
CFI	.946	.992
TLI	.933	.990
SRMR	.054	
WRMR		.691

項目	標準化負荷量	
	量的変数 モデル	順序変数 モデル
F1		
2	.493	.609
5	.672	.733
6	.681	.757
9	.754	.823
10	.710	.806
11	.624	.718
F2		
1	.406	.520
3	.504	.602
4	.353	.517
7	.767	.985
8	.499	.672
12	.665	.770
因子間 相関	.543	.655

## カテゴリカルデータの分析例

### • 2値データの場合

• GHQ12の4件法の評定値のうち、「1」と「2」、「3」と「4」をそれぞれ統合し、2値データに変換して再分析

	標準化負荷量	
	量的変数 モデル	順序変数 モデル
パラメータ数	37	25
RMSEA	.044	.031
CFI	.954	.990
TLI	.943	.987
SRMR	.056	
WRMR		.768

項目	標準化負荷量	
	量的変数 モデル	順序変数 モデル
F1		
2	.467	.671
5	.668	.798
6	.606	.779
9	.754	.897
10	.694	.875
11	.633	.770
F2		
1	.478	.585
3	.493	.620
4	.325	.552
7	.636	.930
8	.518	.631
12	.572	.744
因子間 相関	.536	.681

## カテゴリカルデータ分析のまとめ

### • 従属変数の種類に応じて、適切な分析モデルを用いる必要がある

- 量的変数（正規分布）→線形回帰モデル
- 順序変数→プロビット回帰モデル
- 名義変数→ロジスティック回帰モデル

### • カテゴリカル因子分析には順序プロビットモデルが用いられる

- 順序変数の背後に連続的な潜在反応変数を想定することで、通常の線形回帰モデルと同様の感覚で扱える
  - 標準化係数や適合度指標が得られるため結果の解釈が容易

## カテゴリカルデータ分析のまとめ

### • モデルの使い分け

- 評定値のカテゴリ数
  - 3件法以下の場合、量的変数モデルに比べ、順序変数モデルのパラメータ数が等しいか、少なくなるので、基本的に順序変数モデルの方が望ましい
  - 4件法以上の場合、順序変数モデルのパラメータが増えて、儉約性が低下するので、量的変数モデルと比較して、望ましいモデルを選択する
- 分布の歪み
  - 量的変数モデルは正規分布を仮定しているため、分布に著しい歪みが見られる場合、推定値に偏りが生じる
  - データの性質に応じて、順序変数モデルや打ち切り変数モデルなどを用いる

## 多母集団同時分析

## 多母集団同時分析の基本原則

### • 多母集団同時分析とは

- 複数の母集団のデータを同時に分析し、集団間でパラメータの比較を行う手法
- 構造モデル（パス解析）の文脈
  - 集団間で特定のパス係数や平均値（切片）が異なるか否かを検証する
- 測定モデル（因子分析）の文脈
  - 集団間で因子構造が保たれるか否か（測定不変性）を検証する

## 多母集団同時分析の基本原則

### • マルチレベルモデルとの違い

- 多母集団同時分析
  - 関心：特定の集団間のパラメータの差（集団を**固定因子**と見なす）
  - 集団の数：少数（2～5程度が中心）
- マルチレベルモデル
  - 関心：多数の集団間のパラメータのばらつき（集団を**ランダム因子**と見なす）
  - 集団の数：多数（数十以上）

## 多母集団同時分析の基本原則

### • なぜ測定不変性を検証するのか

- 集団間で平均値やパス係数の比較を行うための前提条件となるから
- 例1：性別によって抑うつ程度は異なるのか
  - 性別間における抑うつ尺度の**強測定不変性**が前提
- 例2：日本とアメリカでは、自閉症特性が対人適応に及ぼす効果が異なるのか
  - 日米間における自閉症特性尺度と対人適応尺度の**弱測定不変性**が前提

## 多母集団同時分析の基本原則

### • 測定不変性の検証手順

1. 個別モデル：同じ因子配置（パス図）で、集団ごとに個別に分析
2. 配置不変モデル：同じ因子配置（パス図）で、パラメータには制約を置かず同時に分析
3. 弱測定不変モデル：2に各項目の因子負荷量の**等値制約**（集団間で値が等しいという制約）を追加
  - 集団間で分散・共分散やパス係数の比較を行うための前提条件
4. 強測定不変モデル：3に各項目の切片の等値制約を追加
  - 集団間で平均値の比較を行うための前提条件

## 多母集団同時分析の分析例

### • データ

- 参加者
  - 保育園年少から中学3年生までの9012名（男子4571名、女子4441名）の母親または父親
- 尺度
  - 肯定的・否定的養育行動尺度（伊藤他, 2014, 印刷中）
  - 24項目（4件法）

### • モデル

- 二次因子モデル
  - 肯定的養育：関与・見守り、肯定的応答性、意思の尊重
  - 否定的養育：過干渉、非一貫性、厳しい叱責・体罰

### 【肯定的養育】

#### 関与・見守り

- 7 学校（または幼稚園・保育園）での出来事や友達のことについて話す
- 13 子どもが遊ぶ友達のことをよく知っている
- 18 学校（幼稚園・保育園）が終わった後、子どもが何をしているのか把握している
- 19 子どもと一緒に遊んだり、楽しいことをしたりする

#### 肯定的応答性

- 8 子どもが何かをしてくれたときに、ありがとうと言う
- 12 子どもが何かうまくできたときには、ほめてあげる
- 17 子どもが喜んでくるときには、一緒になって喜ぶ
- 22 おもしろいことを子どもと一緒に笑う

#### 意思の尊重

- 1 子どもの好きなように遊ばせている
- 2 子どもが問題に直面していても、できるだけ本人に解決させる
- 5 できるだけ子ども自身の意思を尊重する
- 10 失敗することがわかっていても、子どものやりたいようにやらせる

### 【否定的養育】

#### 過干渉

- 16 自分がいないと、子どもは何もできないと感じる
- 20 自分の目を離れている間、子どものことが心配で仕方ない
- 23 子どもは未熟なので、子どもの考えより、自分の考えを優先する
- 24 子どもが立派な大人になるためには、自分たち（保護者）の言うとおりに行動するのが一番だと思う

#### 非一貫性

- 3 子どもを叱ったりほめたりする基準が、その時の気分によって左右される
- 9 個人的なイライラを子どもにぶつけてしまうときがある
- 11 子どもへの叱り方が、自分の気分によって変わる
- 21 子どものペースより、自分のペースを優先する

#### 厳しい叱責・体罰

- 4 子どもが言うことを聞かないとき、頭に血が昇り、冷静さを失う
- 6 子どもに対して、乱暴な言葉遣いになる
- 14 子どもに対して、長時間説教をしたり、文句を言い続ける
- 15 しつけとして、子どもの頭や体を叩くことがある

## 多母集団同時分析の分析例

### • 測定不変性の検証

- 性別
  - 男子
  - 女子
- 回答者の続柄
  - 母親
  - 父親
- 学年段階
  - 保育園
  - 小学校低学年
  - 小学校高学年
  - 中学校

## 多母集団同時分析の分析例

### • 個別モデル (男子) のシNTAXス

```
DATA: FILE = PNPS.csv;

VARIABLE: NAMES = id gender grade cp01-cp24;
USEVAR = cp01-cp24;
MISSING = ALL(999);
USEOBSERVATIONS = gender EQ 0; !男子のデータのみ使用

MODEL: cps01 BY cp07* cp13 cp18 cp19; !負荷量を固定しない
      cps02 BY cp08* cp12 cp17 cp22;
      cps03 BY cp01* cp02 cp05 cp10;
      cps04 BY cp16* cp20 cp23 cp24;
      cps05 BY cp03* cp09 cp11 cp21;
      cps06 BY cp04* cp06 cp14 cp15;
      cps01h BY cps01* cps02 cps03;
      cps02h BY cps04* cps05 cps06;
      cps01-cps02h@1; !モデル識別のため分散・誤差分散を1に固定

ANALYSIS: ESTIMATOR = MLR; !ロバスト最尤法

OUTPUT: SAMP STDYX MOD(0); !記述統計、標準化係数、修正指標を出力
```

## 多母集団同時分析の分析例

### • 個別モデル (女子) のシNTAXス

```
DATA: FILE = PNPS.csv;

VARIABLE: NAMES = id gender grade cp01-cp24;
USEVAR = cp01-cp24;
MISSING = ALL(999);
USEOBSERVATIONS = gender EQ 1; !女子のデータのみ使用

MODEL: cps01 BY cp07* cp13 cp18 cp19; !負荷量を固定しない
      cps02 BY cp08* cp12 cp17 cp22;
      cps03 BY cp01* cp02 cp05 cp10;
      cps04 BY cp16* cp20 cp23 cp24;
      cps05 BY cp03* cp09 cp11 cp21;
      cps06 BY cp04* cp06 cp14 cp15;
      cps01h BY cps01* cps02 cps03;
      cps02h BY cps04* cps05 cps06;
      cps01-cps02h@1; !モデル識別のため分散・誤差分散を1に固定

ANALYSIS: ESTIMATOR = MLR; !ロバスト最尤法

OUTPUT: SAMP STDYX MOD(0); !記述統計、標準化係数、修正指標を出力
```

## 多母集団同時分析の分析例

### • 配置不変モデルのシNTAXス (前半)

```
DATA: FILE = PNPS.csv;

VARIABLE: NAMES = id gender grade cp01-cp24;
USEVAR = cp01-cp24;
MISSING = ALL(999);
GROUPING = gender (0=boy 1=girl); !性別でグループ化

MODEL: cps01 BY cp07* cp13 cp18 cp19;
      cps02 BY cp08* cp12 cp17 cp22;
      cps03 BY cp01* cp02 cp05 cp10;
      cps04 BY cp16* cp20 cp23 cp24;
      cps05 BY cp03* cp09 cp11 cp21;
      cps06 BY cp04* cp06 cp14 cp15;
      cps01h BY cps01* cps02 cps03;
      cps02h BY cps04* cps05 cps06;
      cps01-cps02h@1;
      [cps01-cps02h@0]; !各因子の平均(切片)を0に固定
```

## 多母集団同時分析の分析例

### • 配置不変モデルのシNTAXス (後半)

```
MODEL girl: !女子のモデル
      cps01 BY cp07* cp13 cp18 cp19; !因子負荷量を自由推定
      cps02 BY cp08* cp12 cp17 cp22;
      cps03 BY cp01* cp02 cp05 cp10;
      cps04 BY cp16* cp20 cp23 cp24;
      cps05 BY cp03* cp09 cp11 cp21;
      cps06 BY cp04* cp06 cp14 cp15;
      cps01h BY cps01* cps02 cps03;
      cps02h BY cps04* cps05 cps06;
      [cp01-cp24*]; !各項目の切片を自由推定

ANALYSIS: ESTIMATOR = MLR;

OUTPUT: SAMP STDYX MOD(0);
```

## 多母集団同時分析の分析例

### • 弱測定不変モデルのシNTAXス (抜粋)

```
MODEL: cps01 BY cp07* cp13 cp18 cp19;
      cps02 BY cp08* cp12 cp17 cp22;
      cps03 BY cp01* cp02 cp05 cp10;
      cps04 BY cp16* cp20 cp23 cp24;
      cps05 BY cp03* cp09 cp11 cp21;
      cps06 BY cp04* cp06 cp14 cp15;
      cps01h BY cps01* cps02 cps03;
      cps02h BY cps04* cps05 cps06;
      cps01-cps02h@1;
      [cps01-cps02h@0]; !各因子の平均(切片)を0に固定

MODEL girl: !女子のモデル
      [cp01-cp24*]; !各項目の切片を自由推定
```

## 多母集団同時分析の分析例

### • 強測定不変モデルのシNTAXス (抜粋)

```
MODEL: cps01 BY cp07* cp13 cp18 cp19;
      cps02 BY cp08* cp12 cp17 cp22;
      cps03 BY cp01* cp02 cp05 cp10;
      cps04 BY cp16* cp20 cp23 cp24;
      cps05 BY cp03* cp09 cp11 cp21;
      cps06 BY cp04* cp06 cp14 cp15;
      cps01h BY cps01* cps02 cps03;
      cps02h BY cps04* cps05 cps06;
      cps01-cps02h@1;
      [cps01-cps02h@0]; !各因子の平均(切片)を0に固定

MODEL girl: !女子のモデル (指定なし)
```

## 多母集団同時分析の分析例

### • モデル適合度

	df	$\chi^2$	CFI	TLI	RMSEA	SRMR	BIC
個別モデル							
男子	245	2857.377	.914	.903	.048	.049	
女子	245	2626.782	.916	.906	.047	.051	
配置不変モデル	490	5482.858	.915	.904	.048	.050	389354.6
弱測定不変モデル	514	5508.839	.915	.910	.046	.052	389188.4
強測定不変モデル	538	6001.678	.907	.905	.047	.053	389495.4

## 多母集団同時分析の分析例

### • 修正指標 (抜粋)

	M.I.	E.P.C.	Std E.P.C.	StdYX E.P.C.
Means/Intercepts/Thresholds				
[ CP01 ]	0.950	0.004	0.004	0.006
[ CP02 ]	9.257	0.022	0.022	0.031
[ CP03 ]	16	0.007	0.007	0.011
[ CP04 ]	80	-0.007	-0.007	-0.010
[ CP05 ]	17	0.002	0.002	0.003
[ CP06 ]	1.952	-0.008	-0.008	-0.012
[ CP07 ]	234.486	0.108	0.108	0.141
[ CP08 ]	0.626	0.004	0.004	0.007
[ CP09 ]	14.399	0.020	0.020	0.031
[ CP10 ]	4.867	-0.016	-0.016	-0.022
[ CP11 ]	0.123	-0.002	-0.002	-0.003
[ CP12 ]	30.566	-0.030	-0.030	-0.047
[ CP13 ]	11.221	0.025	0.025	0.032
[ CP14 ]	1.404	-0.007	-0.007	-0.011
[ CP15 ]	45.639	-0.037	-0.037	-0.065
[ CP16 ]	24.262	-0.032	-0.032	-0.049
[ CP17 ]	14.968	-0.027	-0.027	-0.034
[ CP18 ]	7.373	0.024	0.024	0.028
[ CP19 ]	2.415	-0.012	-0.012	-0.014

学校での出来事や  
友達について話す

## 多母集団同時分析の分析例

### • 部分的強測定不変モデルのシNTAXス (抜粋)

```
MODEL: cps01 BY cp07* cp13 cp18 cp19;
      cps02 BY cp08* cp12 cp17 cp22;
      cps03 BY cp01* cp02 cp05 cp10;
      cps04 BY cp16* cp20 cp23 cp24;
      cps05 BY cp03* cp09 cp11 cp21;
      cps06 BY cp04* cp06 cp14 cp15;
      cps01h BY cps01* cps02 cps03;
      cps02h BY cps04* cps05 cps06;
      cps01-cps02h@1;
      [cps01-cps02h@0]; !各因子の平均(切片)を0に固定
```

```
MODEL girl: !女子のモデル
      [cp07]; !項目7の切片を自由推定
```

## 多母集団同時分析の分析例

### • モデル適合度

	df	$\chi^2$	CFI	TLI	RMSEA	SRMR	BIC
個別モデル							
男子	245	2857.377	.914	.903	.048	.049	
女子	245	2626.782	.916	.906	.047	.051	
配置不変モデル	490	5482.858	.915	.904	.048	.050	389354.6
弱測定不変モデル	514	5512.232	.915	.910	.046	.052	389144.8
強測定不変モデル	538	5789.733	.911	.908	.047	.052	389271.2
<b>部分的強測定不変</b>	<b>537</b>	<b>5653.541</b>	<b>.913</b>	<b>.910</b>	<b>.046</b>	<b>.052</b>	<b>389126.7</b>

## 多母集団同時分析の分析例

### • 切片の推定値 (抜粋)

#### • 男子

Intercepts				
CP01	3.227	0.007	481.939	0.000
CP02	2.894	0.007	359.830	0.000
CP03	0.007	0.007	273.635	0.000
CP04	0.008	0.008	227.358	0.000
CP05	0.007	0.007	434.355	0.000
CP06	1.850	0.008	245.563	0.000
CP07	3.022	0.011	268.011	0.000

学校での出来事や  
友達について話す

#### • 女子

Intercepts				
CP01	3.227	0.007	481.939	0.000
CP02	2.894	0.007	359.830	0.000
CP03	1.963	0.007	273.635	0.000
CP04	1.766	0.008	227.358	0.000
CP05	3.005	0.007	434.355	0.000
CP06	1.850	0.008	245.563	0.000
CP07	3.257	0.011	305.869	0.000

## 多母集団同時分析の分析例

### • 続柄の多母集団同時分析

	df	$\chi^2$	CFI	TLI	RMSEA	SRMR	BIC
個別モデル							
母親	245	4960.725	.914	.903	.048	.050	
父親	245	553.314	.916	.905	.047	.051	
配置不変モデル	490	5609.848	.914	.903	.048	.050	386990.8
弱測定不変モデル	514	5619.338	.914	.907	.047	.051	386827.0
強測定不変モデル	538	5996.099	.908	.903	.048	.052	387091.5
<b>部分的強測定不変</b>	<b>534</b>	<b>5715.866</b>	<b>.912</b>	<b>.907</b>	<b>.047</b>	<b>.051</b>	<b>386818.8</b>

#### 4つの切片の等値制約を解除

- 子どもに対して、乱暴な言葉遣いになる 父親>母親
- 学校(または幼稚園・保育園)での出来事や友達のことについて話す 父親>母親
- 個人的なイライラを子どもにぶつけてしまうときがある 母親>父親
- 子どもと一緒に遊んだり、楽しいことをしたりする 父親>母親

## 多母集団同時分析の分析例

### • 学年代階の多母集団同時分析

	df	$\chi^2$	CFI	TLI	RMSEA	SRMR	BIC
個別モデル							
保育園	245	967.279	.896	.883	.050	.054	
小学校低学年	245	1761.126	.912	.901	.048	.049	
小学校高学年	245	1816.055	.911	.900	.049	.052	
中学校	245	1635.562	.917	.907	.047	.049	
配置不変モデル	980	6197.712	.911	.900	.048	.051	389218.4
弱測定不変モデル	1052	6515.256	.907	.905	.047	.066	388814.2
強測定不変モデル	1124	7702.377	.888	.890	.051	.068	389624.4
<b>部分的強測定不変</b>	<b>1108</b>	<b>6824.394</b>	<b>.903</b>	<b>.903</b>	<b>.048</b>	<b>.067</b>	<b>388809.3</b>

#### 6つの切片の等値制約を解除

- 子どもと一緒に遊んだり、楽しいことをしたりする ↓
- 子どもに対して、長時間説教をしたり、文句を言い続けたりする ↑
- 子どもが立派な大人になるためには、自分たちの言うとおりに行動するのが一番 ↑
- 学校(幼稚園・保育園)が終わった後、子どもが何をしているのか把握している ↓
- 子どもが問題に直面していても、できるだけ本人に解決させる ↑
- しつけとして、子どもの頭や体を叩くことがある ↓

## 多母集団同時分析の分析例

### • 考察

- 性別、続柄、学年代階のいずれについても、集団間で因子配置と因子負荷量の不変性が示された
  - 集団間で相関やパス係数の比較が可能
- 切片は一部の項目で不変性が満たされなかった
  - 集団間の切片の差は大きくなかったが、正確を期すならば、属性の異なる対象者間で尺度得点を直接比較することは望ましくない
- 当尺度(伊藤他, 印刷中)では、性別、続柄、学年代階ごとに標準得点のノルム(換算表)を開発しているため、属性が異なっても標準得点上での比較は可能

## 多母集団同時分析のまとめ

- 多母集団同時分析は、特定の集団間のパラメータの差に関心があるときに使用する
  - 不特定の集団間のバラつきに関心がある場合はマルチレベルモデルを使用
- 構造モデルの差を検証する前提条件として、測定モデルの不変性を確かめる必要がある
  - 弱測定不変(因子負荷量の不変性) → 分散・共分散・パス係数の比較
  - 強測定不変(切片の不変性) → 平均値・切片の比較

## Menu

- 因子分析とは
- 因子分析のモデル
- 探索的因子分析 (EFA)
- 確認的因子分析 (CFA)
- 高次因子分析
- 階層的因子分析
- 因子分析の実践

121

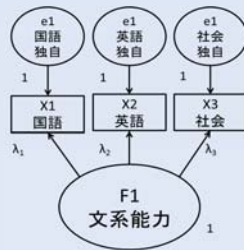
## 因子分析とは

- 多変量解析の中でも広く利用されている分析
- 心理尺度を作成する研究で多用
- 測定可能な変数からその背後にある潜在変数を分析
- 探索的因子分析と確認的因子分析に分けられる

122

## 因子分析のモデル

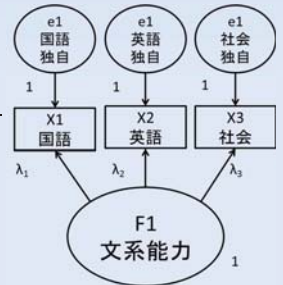
- 因子・・・共通因子 (潜在変数)
- 誤差・・・独自因子 (潜在変数)
- 因子負荷量・・・因子から観測変数へのパス係数
- 独自性・・・誤差の分散



123

## 文系能力の1因子モデル

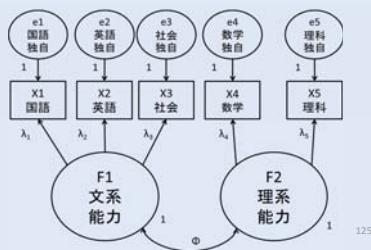
	国語	英語	社会
X1 国語	1.000	.740	.538
X2 英語		1.000	.558
X3 社会			1.000



124

## 学力の2因子モデル

	国語	英語	社会	数学	理科
X1 国語	1.000	.740	.538	.360	.438
X2 英語		1.000	.558	.202	.389
X3 社会			1.000	.163	.269
X4 数学				1.000	.868
X5 理科					1.000



125

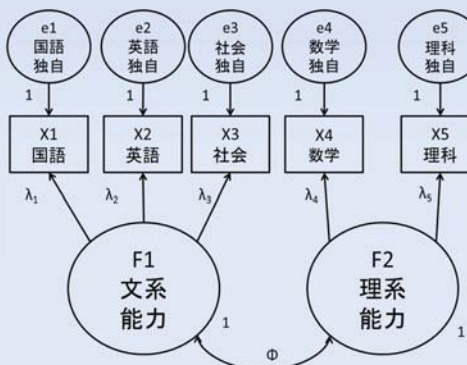
## 確認的因子分析

(CFA: Confirmatory Factor Analysis)

- 事前に何らかの手段によって得られた知見から、因子と観測変数の関係についての仮説を検証的に分析
- 因子数および因子と観測変数の関連に関する仮説モデルを検証する

126

## CFAモデル



127

## 探索的因子分析

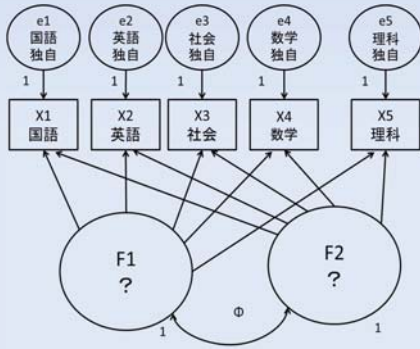
(EFA: Exploratory Factor Analysis)

- 因子に関する明確な仮説がなく、観測変数の背後にある共通の因子構造を探りたいときに使われる
- 多くの観測変数間に見られる相関関係が、いくつかのどのようなまとまりを想定すれば説明できるかを調べる分析となる

128

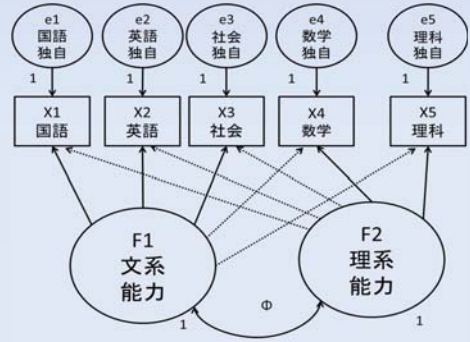


## EFAモデル



129

## EFAの結果



130

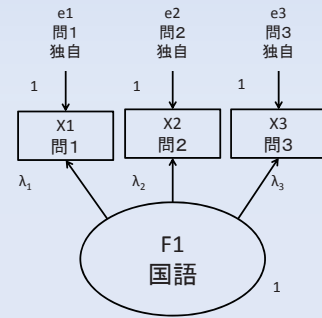
## 高次因子分析

(higher-order factor analysis model)

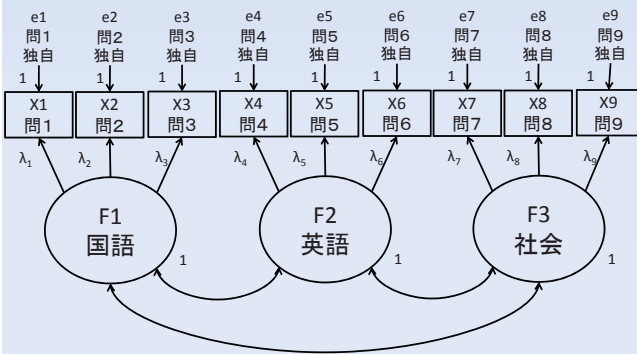
- 通常の因子分析における因子を、さらに少数の因子で説明するモデル
- 二次因子を想定するものが多いが、三次、四次...といった上位の因子を想定することもできる(が、現実的ではないか)。
- 複数の異なる尺度の下位因子をまとめて実施する因子分析はしばしばジョイント因子分析と呼ばれる

131

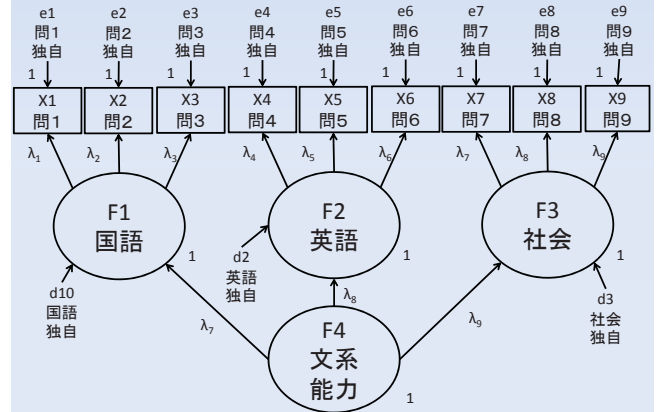
## 項目による因子分析モデル



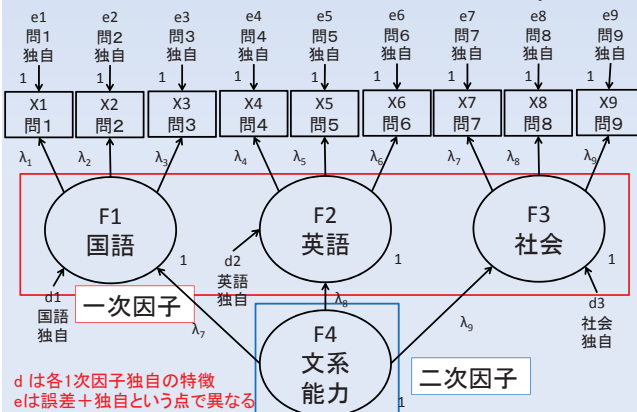
## CFAモデル(3因子)



## 高次因子分析モデル(CFA)

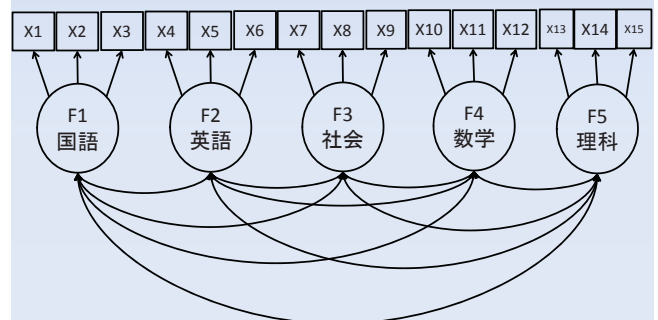


## 高次因子分析モデル(CFA)



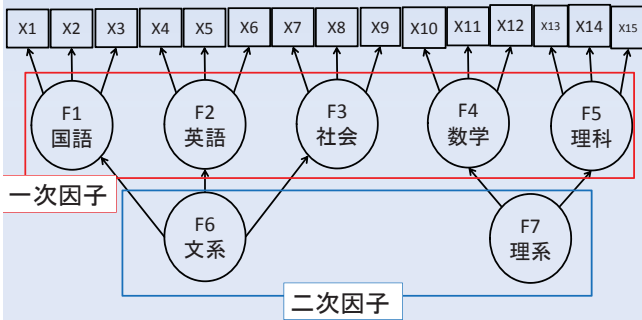
d は各1次因子独自の特徴  
e は誤差+独自という点で異なる

## 通常のCFAモデル(5因子)

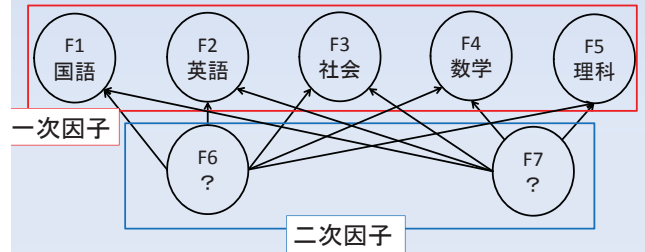


※誤差変数とパス係数、分散は省略

## 高次因子分析モデル(CFA)



## 高次因子分析モデル(EFA)



一次因子に対して通常の因子分析を行う。一次因子の潜在変数(因子得点)や観測変数(尺度得点)よりも、因子間相関を用いるほうが望ましい。

## 高次因子分析の応用

- 通常の因子分析における因子のさらに上位の因子を想定する
- 下位尺度を算出し、それらを加算して総得点を求めるような尺度はこのモデルが適用できる(実際には高次因子分析が行われていないことが多い)。

139

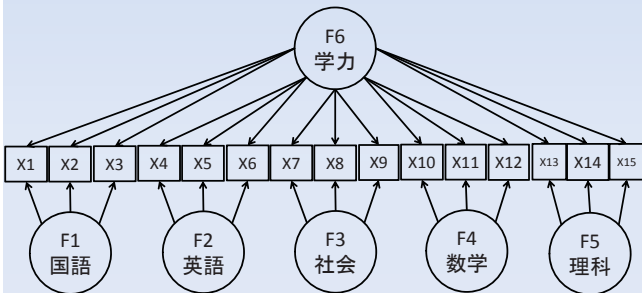
## 階層的因子分析

(hierarchical factor analysis model)

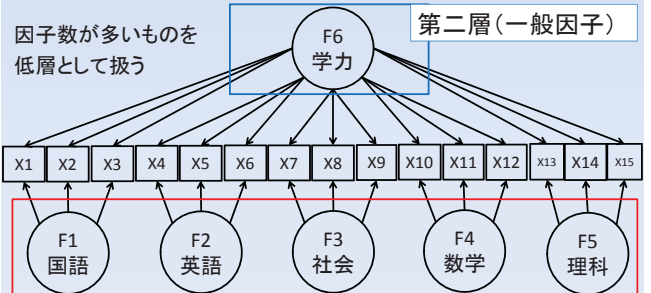
- 因子に複数の層(layer)を想定した因子分析モデルであり、知能研究の分野でよく使用
- 多く使われるモデルは、全ての観測変数を説明する因子として一般因子(general factor)と、一部の観測変数を説明する因子であるグループ因子(group factor)を想定したもの
- 一般因子と二層目の因子によるモデルは双因子モデル(bi-factor model)と呼ばれる。

140

## 階層的因子分析モデル(双因子CFA)



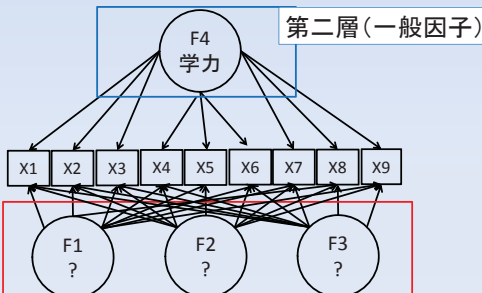
## 階層的因子分析モデル(双因子CFA)



第一層(グループ因子)

一般因子を想定する場合、第一層における因子間相関は指定しない

## 階層的因子分析モデル(双因子EFA)



第一層(グループ因子)

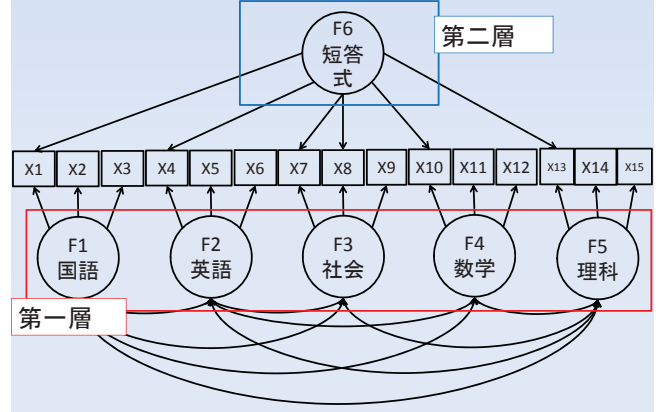
Mplusではbi-factor EFAにて、双因子モデルのみEFAが可能  
一般因子を想定した上でのEFAと考えることができる

## 階層的因子分析の応用

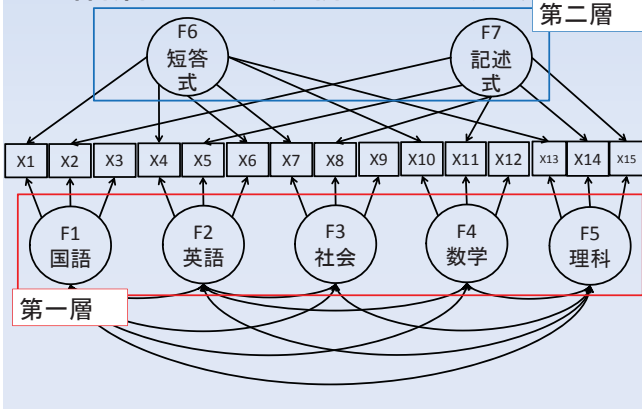
- 全般的な能力を含むモデル  
(一般知能、学力、健康、運動能力、性格のGFP) → 高次因子分析とも対応 (Schmid & Leiman, 1957)
- 方法の因子を含むモデル
  - 多特性多方法行列 (MTMM)
  - 尋ね方の違い (逆転項目など、場面想定)
  - 出題方法の違い (選択式、記述式、短答式など)
  - 評定者の違い (自己、他者、教師など)
 ※誤差相関が解釈可能な因子である可能性も

145

## 階層的因子分析モデル(方法)



## 階層的因子分析モデル(方法)



## 因子分析の実践例

- データ ビッグファイブ尺度 (和田, 1996)
- 5因子 × 12項目 = 60項目
- E 外向性
- N 神経症傾向 (情緒不安定性)
- O 経験への開放性
- C 勤勉性 (誠実性)
- A 協調性 (調和性)

## 因子分析の実践例

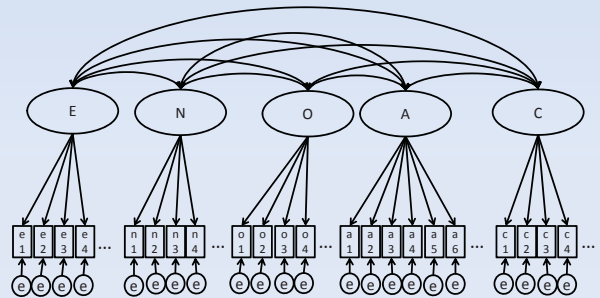
データセット1 (N=2099)

- 5因子モデル
- 6因子モデル (HEXACOモデル)
- 高次因子分析モデル (1因子、2因子)
- 階層的因子分析モデル (逆転項目)

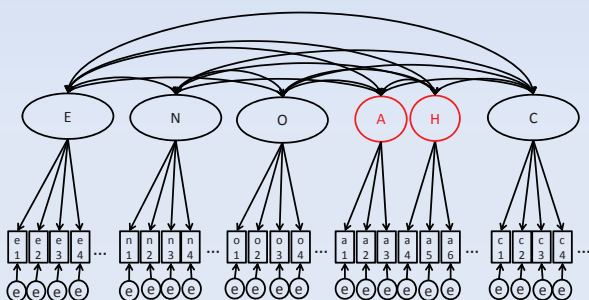
データセット2 (N=250, 60項目 × 2評定)

- 階層的因子分析モデル (自己・他者評価)

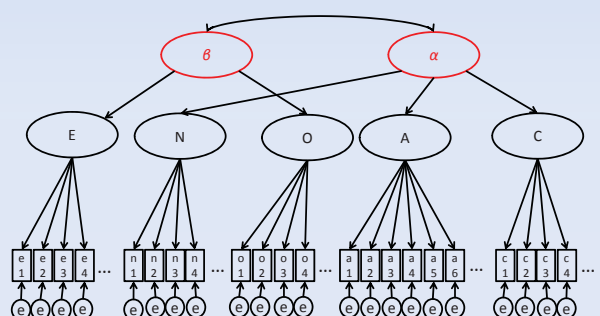
## 一般的なBig Fiveの因子分析モデル



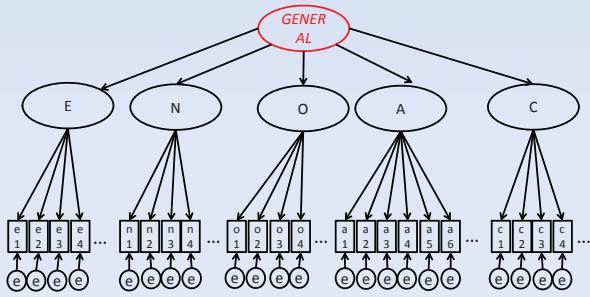
## HEXACOモデル(6因子)



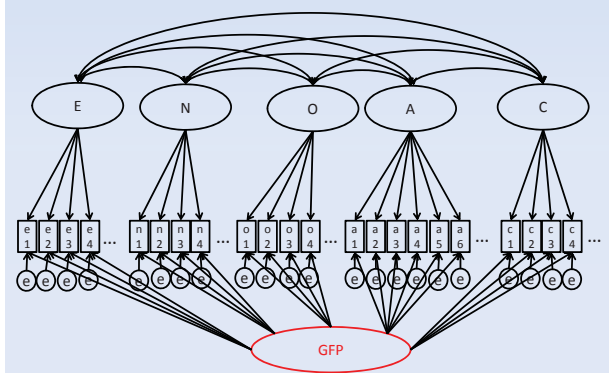
## 高次因子分析モデル(2因子)



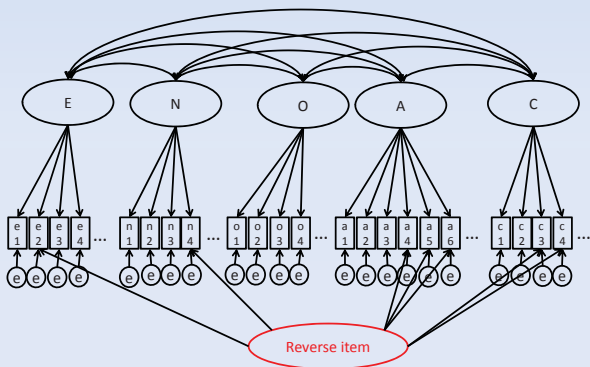
## 高次因子分析モデル(1因子)



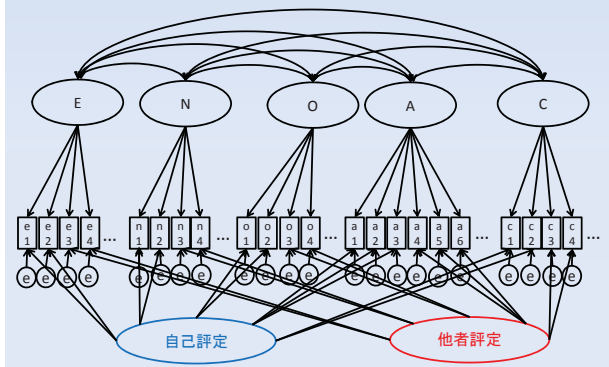
## 階層的因子分析(一般因子, GFP)



## 階層的因子分析(方法の因子:逆転項目)



## 階層的因子分析 (方法の因子:自己評定・他者評定)



## EFAの手順

1. 因子数の候補の選定
2. 因子の抽出・推定
3. 回転
4. 項目の取捨選択

157

## シンタックス例

```
TITLE: Data sample
DATA:
  FILE = "SampleData.csv";
VARIABLE:
  NAMES = ID gend age grade bf01-60;
  USEVARIABLES = bf01- bf60;
  MISSING IS *;
ANALYSIS:
  TYPE = EFA 4 6; ←4~6因子でEFA
  ESTIMATOR = ML; ←最尤推定法
  ROTATION = GEOMIN; ←ジオミン回転
  PARALLEL = 50; ←平行分析
OUTPUT: SAMP STDYX RESIDUAL MOD; ←出力
PLOT:
  TYPE = PLOT3;
```

158

## EFA

```
ANALYSIS:
  TYPE = EFA 4 6; ← 4~6因子でEFA
  ESTIMATOR = ML; ← 最尤推定法
  ROTATION = GEOMIN; ← ジェオミン回転
  PARALLEL = 50; ← 平行分析
OUTPUT: SAMP STDYX RESIDUAL MOD; ←出力
PLOT:
  TYPE = PLOT3;
```

159

## EFA(双因子)

```
ANALYSIS:
  TYPE = EFA 4 6;
  ESTIMATOR = ML;
  ROTATION = BI-GEOMIN; ←双因子ジオミン
  PARALLEL = 50;
OUTPUT: SAMP STDYX RESIDUAL MOD;
```

160

### CFA(通常の5因子)

ANALYSIS: ESTIMATOR = ML; ← 最尤推定法  
 MODEL: f1 BY bf01-bf12; ← E  
 f2 BY bf13-bf24; ← N  
 f3 BY bf25-bf36; ← O  
 f4 BY bf37-bf48; ← C  
 f5 BY bf49-bf60; ← A  
 OUTPUT: SAMP STDYX RESIDUAL MOD(0);

161

### CFA(高次1因子)

ANALYSIS: ESTIMATOR = ML; ← 最尤推定法  
 MODEL: f1 BY bf01-bf12; ← E  
 f2 BY bf13-bf24; ← N  
 f3 BY bf25-bf36; ← O  
 f4 BY bf37-bf48; ← C  
 f5 BY bf49-bf60; ← A  
 f6 BY f1-f5; ← 高次1因子  
 OUTPUT: SAMP STDYX RESIDUAL MOD(0);

162

### CFA(高次2因子)

ANALYSIS: ESTIMATOR = ML; ← 最尤推定法  
 MODEL: f1 BY bf01-bf12; ← E  
 f2 BY bf13-bf24; ← N  
 f3 BY bf25-bf36; ← O  
 f4 BY bf37-bf48; ← C  
 f5 BY bf49-bf60; ← A  
 f6 BY f1 f3; ← 高次1因子  
 f7 BY f2 f4 f5; ← 高次2因子  
 OUTPUT: SAMP STDYX RESIDUAL MOD(0);

163

### CFA(階層的, 方法, 逆転)

ANALYSIS: ESTIMATOR = ML; ← 最尤推定法  
 MODEL: f1 BY bf01-bf12; ← E  
 f2 BY bf13-bf24; ← N  
 f3 BY bf25-bf36; ← O  
 f4 BY bf37-bf48; ← C  
 f5 BY bf49-bf60; ← A  
 f6 BY bf01 bf05...; ← 逆転項目(省略)  
 f6 WITH f1-f5 @0; ← f6との因子相関0  
 OUTPUT: SAMP STDYX RESIDUAL MOD(0);

164

### CFA(階層的, 方法, 評定者)

ANALYSIS: ESTIMATOR = ML; ← 最尤推定法  
 MODEL: f1 BY bf01-bf12 bf61-bf72; ← E  
 f2 BY bf13-bf24 bf73-bf84; ← N  
 f3 BY bf25-bf36 bf85-bf96; ← O  
 f4 BY bf37-bf48 bf97-bf108; ← C  
 f5 BY bf49-bf60 bf109-bf120; ← A  
 f6 BY bf01-bf60; ← 自己評定  
 f7 BY bf61-bf120; ← 他者評定  
 f6 WITH f1-f5 @0; ← f6とf1-f5の因子相関0  
 f7 WITH f1-f5 @0; ← f7とf1-f5の因子相関0  
 OUTPUT: SAMP STDYX RESIDUAL MOD(0);

165

### EFAの出力

SUMMARY OF ANALYSIS、DATA  
 ...変数の数や推定法等の確認、欠測パターン

RESULTS FOR EFA  
 ...固有値の減衰状況、平行分析、適合度、因子負荷量、因子間相関、独自性の順に出力

### CFAの出力

SUMMARY OF ANALYSIS、DATA  
 ...変数の数や推定法等の確認、欠測パターン

RESULTS  
 ...適合度、パラメータの推定値(因子負荷量)、因子間相関、検定統計量(修正指標)の順に出力

### Result: Goodness of fit

	Big 5 (original)	HEXACO	2 High-order	1 High-order	Hierarchical
RESEA	.069	.063	.069	.070	.060
AIC	19135.23	16235.53	19258.36	19391.77	14899.72
BCC	19176.62	16247.27	19269.51	19402.85	14903.07

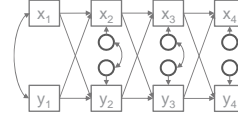
## 縦断データの分析

中部大学 伊藤大幸

## 2つのモデル

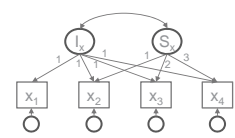
### 交差遅延モデル

- リサーチ・クエスション
  - 各変数の時間的安定性はどの程度か？
  - 変数間の因果関係の方向と強さは？
- 推定の対象
  - 分散・共分散

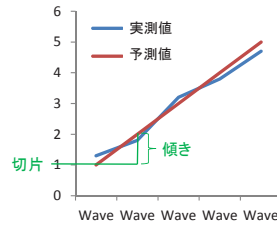
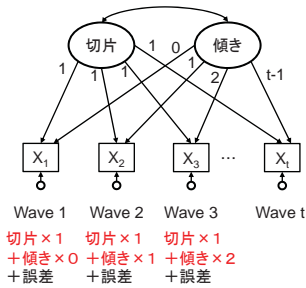


### 潜在成長(曲線)モデル

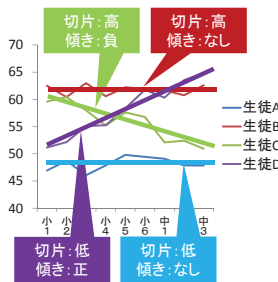
- リサーチ・クエスション
  - 変化の軌跡に個人差はあるか？
  - 変化の軌跡はどのような変数と関連するか？
- 推定の対象
  - 分散・共分散
  - 平均値・切片



## 潜在成長モデルの基本原理



## 潜在成長モデルの基本原理



## 潜在成長モデルの基本原理

### 応用例

- 乳児期の対人的反応の発達の变化を、自閉症スペクトラム障害の早期診断に利用する (軌跡→結果変数)
- うつ病の寛解後の症状の変化を予測する予後規定因子を同定する (原因変数→軌跡)
- 幼児期から思春期にかけての攻撃性の変化の軌跡を成人後の反社会的行動の予測に利用する (軌跡→軌跡)

## 潜在成長モデルの基本原理

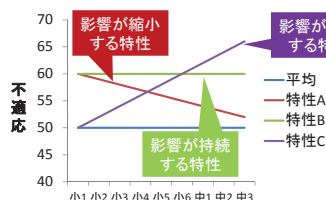
### なぜ「潜在」である必要があるのか？

- 回帰分析を用いて実測値の軌跡から切片と傾きを直接求めることは可能
  - その「顕在」レベルの切片と傾きを用いて、その平均値・分散や他の変数との関連を検討すれば事足りるか？
  - 各回の測定値には偶発的な変動が含まれるため、切片や傾きの値にも誤差が混入してしまう
- 潜在成長モデルでは、切片と傾きを潜在変数としてモデル化し、偶発的な変動を誤差として分離する
  - 切片・傾きの平均値・分散や他の変数との関連について、より正確な推定値が得られる
  - ただし、モデルに適合しない変動は全て誤差として分離されてしまうため、適合度の吟味が重要

## 潜在成長モデルの分析例

### リサーチ・クエスション

- 幼児期に見られる様々な発達障害特性が、就学後の不適応の軌跡をどのように予測するか



## 潜在成長モデルの分析例

### 参加者

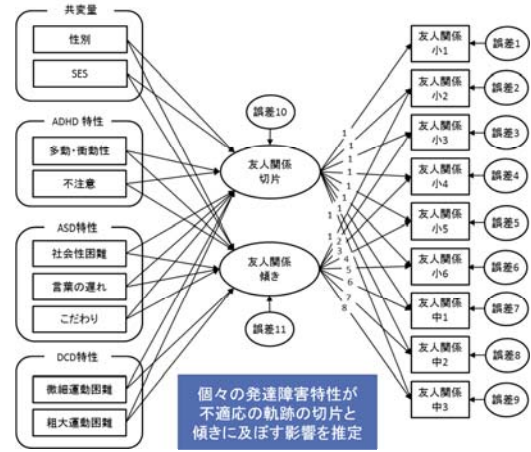
- 2007年度から2016年度にかけて10回の縦断調査
  - 結果変数(就学後の不適応)は2011年度からの6回のみ
- 9つの学年コホートの計3717名が参加
  - 系列コホートデザイン

入学年度	保育園			小学校						中学校		
	男子	女子	合計	1年	2年	3年	4年	5年	6年	1年	2年	3年
2008	198	182	380				380	380	380	380	380	380
2009	208	189	397				397	397	397	397	397	397
2010	156	184	340			340	340	340	340	340	340	340
2011	234	200	434	434	434	434	434	434	434			
2012	238	213	451	451	451	451	451	451	451			
2013	211	216	427	427	427	427	427	427				
2014	221	237	458	458	458	458	458	458				
2015	207	180	387	387	387							
2016	235	208	443	443								
合計	1908	1809	3717	2600	2497	2507	2429	2002	1551	1117	777	380

## 潜在成長モデルの分析例

### 調査内容

- 就学前の発達障害特性（年長2月）
  - TASP (Transitional Assessment Sheet for Preschoolers; 伊藤他, 2017)
    - ADHD特性：多動・衝動性、不注意
    - ASD特性：社会性困難、言葉の遅れ、こだわり
    - DCD特性：微細運動困難、粗大運動困難
- 就学後の不適応（小1～中3の9月）
  - 学業成績：教研式NRT (Norm Referenced Test)
  - 情緒・行動的問題：SDQ (Strength and Difficulties Questionnaire) 教師評定フォーム (Goodman, 1997)
    - 友人関係問題、情緒的症状 (内在化問題)、素行問題 (外在化問題) の3下位尺度を分析に使用



個々の発達障害特性が不適応の軌跡の切片と傾きに及ぼす影響を推定

## 潜在成長モデルの分析例

### 検証の手順

- 測定モデル (因子分析) と構造モデル (パス解析) の両方を含むモデルは2段階で検証を行う
  - 測定モデルの検証
  - 測定モデル+構造モデルの検証
- 理由
  - 測定モデルの適合に問題があれば、構造モデルのパラメータ推定値にも歪みが生じうる
  - モデル全体の適合を評価するだけでは、測定モデルの適合が見逃される場合がある

## 潜在成長モデルの分析例

### シンタックス (測定モデルのみ ; 前半)

```
DATA: FILE = tasp_LGCM2016.t.txt;

VARIABLE: NAMES = id cohort gender SES tasp1-tasp7 int_1-int_9
ext_1-ext_9 peer_1-peer_9 achi_2-achi_9;
MISSING = ALL(999);
USEVAR = peer_1-peer_9;
PATTERN = cohort (
2008 = peer_4 peer_5 peer_6 peer_7 peer_8 peer_9
2009 = peer_3 peer_4 peer_5 peer_6 peer_7 peer_8
2010 = peer_2 peer_3 peer_4 peer_5 peer_6 peer_7
2011 = peer_1 peer_2 peer_3 peer_4 peer_5 peer_6
2012 = peer_1 peer_2 peer_3 peer_4 peer_5
2013 = peer_1 peer_2 peer_3 peer_4
2014 = peer_1 peer_2 peer_3
2015 = peer_1 peer_2
2016 = peer_1
); !系列コホートデザインのデータパターンを指定
```

## 潜在成長モデルの分析例

### シンタックス (測定モデルのみ ; 後半)

ANALYSIS: COVERAGE = 0; !系列コホートデザインによる欠測のため  
ESTIMATOR = MLR; !ロバスト最尤法

MODEL:  
i s | peer\_1@0 peer\_2@1 peer\_3@2 peer\_4@3 peer\_5@4 peer\_6@5  
peer\_7@6 peer\_8@7 peer\_9@8; !切片因子と傾き因子をモデル化

OUTPUT: SAMP STDYX;

PLOT: TYPE = PLOT3; !プロットを出力  
SERIES = peer\_1-peer\_9(\*); !小1～中3の平均値をプロット

## 潜在成長モデルの分析例

### シンタックス (測定モデルのみ ; 後半) (書換)

ANALYSIS: COVERAGE = 0; !系列コホートデザインによる欠測のため  
ESTIMATOR = MLR; !ロバスト最尤法

MODEL:  
i BY peer\_1-peer\_9@1; !切片因子  
s BY peer\_1@0 peer\_2@1 peer\_3@2 peer\_4@3 peer\_5@4 peer\_6@5  
peer\_7@6 peer\_8@7 peer\_9@8; !傾き因子  
[i s]; !切片・傾きの平均値を推定  
[peer\_1-peer\_9@0]; !観測変数の切片を0に固定

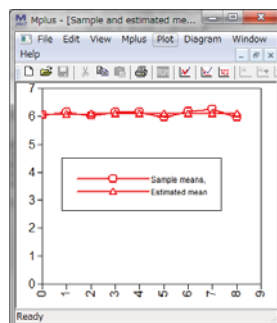
OUTPUT: SAMP STDYX;

PLOT: TYPE = PLOT3; !プロットを出力  
SERIES = peer\_1-peer\_9(\*); !小1～中3の平均値をプロット

## 潜在成長モデルの分析例

### モデル適合度

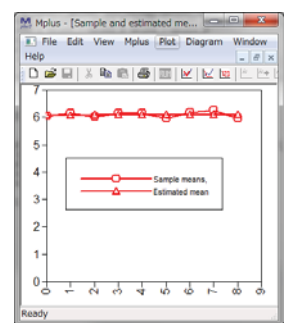
Chi-Square Test of Model Fit	
Value	187.205
Degrees of Freedom	34
P-Value	0.0000
RMSEA (Root Mean Square Error Of Approximation)	
Estimate	0.035
90 Percent C.I.	0.030 0.040
Probability RMSEA <= .05	1.000
CFI/TLI	
CFI	0.944
TLI	0.951
SRMR (Standardized Root Mean Square Residual)	
Value	0.082



## 潜在成長モデルの分析例

### パラメータ推定値 (非標準化)

Means				
I	6.093	0.024	254.138	0.000
S	0.012	0.006	1.984	0.046
(中略)				
Variances				
I	0.805	0.049	16.349	0.000
S	0.013	0.003	3.621	0.000



## 潜在成長モデルの分析例

### ・ シンタックス (フルモデル ; 前半)

DATA: FILE = tasp\_LGCM2016\_t.txt;

```
VARIABLE: NAMES = id cohort gender SES tasp1-tasp7
int_1-int_9 ext_1-ext_9 peer_1-peer_9 achi_2-achi_9;
MISSING = ALL(999);
USEVAR = gender SES tasp1-tasp7 peer_1-peer_9;
PATTERN = cohort (
2008 = peer_4 peer_5 peer_6 peer_7 peer_8 peer_9
2009 = peer_3 peer_4 peer_5 peer_6 peer_7 peer_8
2010 = peer_2 peer_3 peer_4 peer_5 peer_6 peer_7
2011 = peer_1 peer_2 peer_3 peer_4 peer_5 peer_6
2012 = peer_1 peer_2 peer_3 peer_4 peer_5
2013 = peer_1 peer_2 peer_3 peer_4
2014 = peer_1 peer_2 peer_3
2015 = peer_1 peer_2
2016 = peer_1
); !系列コホートデザインのデータパターンを指定
```

## 潜在成長モデルの分析例

### ・ シンタックス (フルモデル ; 後半抜粋)

解釈を容易にするための前処理  
・ 独立変数を標準化  
・ 従属変数を全学年の平均・SDで  
偏差値化

```
DEFINE: STANDARDIZE SES tasp1-tasp7; !標準化
peer_1 = ((peer_1 - 6.1631) / 1.6016) * 10 + 50; !偏差値化
peer_2 = ((peer_2 - 6.1631) / 1.6016) * 10 + 50; !偏差値化
peer_3 = ((peer_3 - 6.1631) / 1.6016) * 10 + 50; !偏差値化
peer_4 = ((peer_4 - 6.1631) / 1.6016) * 10 + 50; !偏差値化
peer_5 = ((peer_5 - 6.1631) / 1.6016) * 10 + 50; !偏差値化
peer_6 = ((peer_6 - 6.1631) / 1.6016) * 10 + 50; !偏差値化
peer_7 = ((peer_7 - 6.1631) / 1.6016) * 10 + 50; !偏差値化
peer_8 = ((peer_8 - 6.1631) / 1.6016) * 10 + 50; !偏差値化
peer_9 = ((peer_9 - 6.1631) / 1.6016) * 10 + 50; !偏差値化

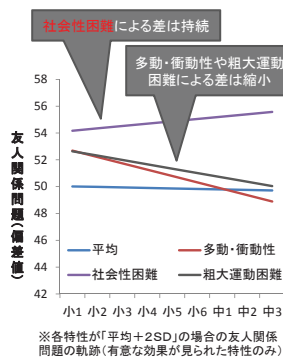
MODEL:
i s | peer_1@0 peer_2@1 peer_3@2 peer_4@3 peer_5@4 peer_6@5
peer_7@6 peer_8@7 peer_9@8; !切片因子と傾き因子をモデル化
i s ON gender SES tasp1-tasp7; !切片・傾き←独立変数
```

## 潜在成長モデルの分析例

### ・ 結果 (友人関係問題)

	切片			傾き		
	B	β	p	B	β	p
性別	.443	.041	.157	.037	.039	.660
SES	-.044	-.008	.762	-.050	-.105	.193
多動・衝動性	1.336	.244	<.001	-.219	-.460	<.001
不注意	-.162	-.029	.521	.100	.206	.132
社会性困難	2.079	.380	<.001	.106	.223	.155
言葉の遅れ	.325	.060	.140	-.025	-.050	.697
こだわり	.256	.046	.250	.006	.010	.933
微細運動困難	-.377	-.069	.059	.056	.120	.293
粗大運動困難	1.311	.240	<.001	-.144	-.296	.006
R <sup>2</sup>	.520 <.001			.141 .009		

$\chi^2(97)=340.5, p<.001$   
CFI=.987, RMSEA=.028, SRMR=.048

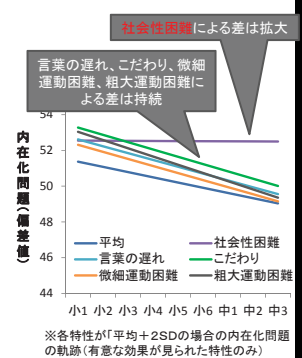


## 潜在成長モデルの分析例

### ・ 結果 (内在化問題)

	切片			傾き		
	B	β	p	B	β	p
性別	1.471	.130	<.001	-.026	-.017	.777
SES	-.311	-.055	.054	-.091	-.127	.025
多動・衝動性	.505	.090	.061	-.110	-.144	.103
不注意	-.343	-.061	.230	.078	.107	.256
社会性困難	.590	.104	.049	.143	.192	.045
言葉の遅れ	.622	.111	.014	-.045	-.059	.496
こだわり	.953	.168	<.001	-.058	-.082	.310
微細運動困難	-.473	-.083	.047	-.052	-.073	.351
粗大運動困難	.836	.148	<.001	-.084	-.110	.144
R <sup>2</sup>	.263 <.001			.041 .076		

$\chi^2(97)=231.9, p<.001$   
CFI=.992, RMSEA=.021, SRMR=.035

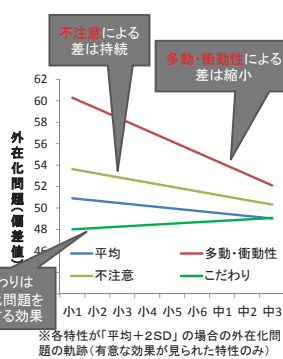


## 潜在成長モデルの分析例

### ・ 結果 (外在化問題)

	切片			傾き		
	B	β	p	B	β	p
性別	-2.457	-.156	<.001	.216	.156	.013
SES	-.787	-.101	<.001	.032	.045	.416
多動・衝動性	4.691	.595	<.001	-.394	-.580	<.001
不注意	1.358	.173	<.001	-.089	-.127	.183
社会性困難	-.133	-.017	.684	.013	.014	.895
言葉の遅れ	-.178	-.023	.473	.006	.008	.928
こだわり	-1.447	-.184	<.001	.184	.269	.002
微細運動困難	-.512	-.066	.054	.095	.135	.091
粗大運動困難	.209	.027	.350	-.006	-.006	.936
R <sup>2</sup>	.474 <.001			.360		

$\chi^2(97)=238.2, p<.001$   
CFI=.993, RMSEA=.021, SRMR=.048

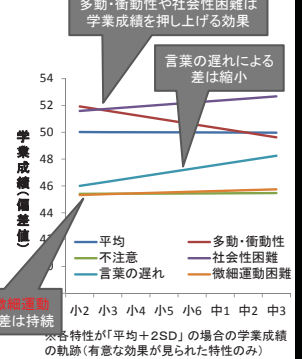


## 潜在成長モデルの分析例

### ・ 結果 (学業成績)

	切片			傾き		
	B	β	p	B	β	p
性別	-1.400	-.074	<.001	-.091	-.050	.200
SES	2.964	.317	<.001	.100	.110	.002
多動・衝動性	.955	.101	<.001	-.161	-.176	.003
不注意	-2.309	-.244	<.001	.008	.008	.890
社会性困難	.788	.083	.017	.081	.089	.220
言葉の遅れ	-2.010	-.212	<.001	.164	.179	.003
こだわり	.409	.043	.105	.012	.013	.806
微細運動困難	-2.350	-.248	<.001	.034	.037	.460
粗大運動困難	-.051	-.005	.824	-.011	-.012	.809
R <sup>2</sup>	.344 <.001			.710.9, p<.001		

$\chi^2(84)=710.9, p<.001$   
CFI=.978, RMSEA=.048, SRMR=.048



## 潜在成長モデルの分析例

### ・ 考察

- ・ 幼児期に評価された発達障害特性により、就学後の不適応が高い精度で予測可能
  - ・ 不注意や社会性困難の影響はとりわけ大きく、児童期から思春期まで持続または拡大する
  - ・ 一方、多動・衝動性、言葉の遅れ、粗大運動の影響は、学年とともに縮小
- ・ ただし、切片に比べると傾きの説明率は比較的低い
- ・ 就学後に生じる様々なライフイベントが軌跡を変化させるためか(クラスメイトや教員の影響、家庭状況の変化、支援・介入など)

## 潜在成長モデルのまとめ

### ・ 潜在成長モデルの強み

- ・ 複数時点にわたる系統的な変化の軌跡をモデル化し、それが他の変数とどのように関連するかを検証できる
  - ・ 一時点の状態よりも、複数時点の軌跡を見ることで、豊かな情報が得られる
  - ・ 例: 影響が持続するのか、拡大するのか、縮小するのか
  - ・ 例: 慢性的ストレスと急性ストレスのどちらが疾患の発症に影響するのか

### ・ 潜在成長モデルの弱み

- ・ 系統的な変化のみを扱うため、時間特異的な変動のメカニズムは明らかにできない
- ・ 変数間の因果関係の方向性を検証できない



## 潜在成長モデルのまとめ

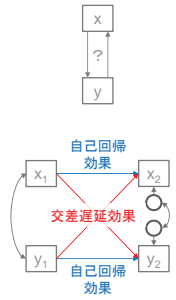
### 研究応用上の注意点

- 少なくとも3時点のデータが必要
  - 適合度の確認には4時点以上のデータが必要
- 測定モデルの適合を確認してから、構造モデルの検証に移る
  - 一次モデルが適合しない場合、二次の傾きを導入するか、時間パラメータ（傾き因子への負荷量）を自由推定するなどして対処する
  - ただし、二次の傾きを含めると、解釈が難解になる
- 理論的に因果関係の方向性が不明確である場合、因果関係の考察には慎重を期す
  - 交差遅延モデルなど、因果関係の検証に適した別のモデルを用いることも検討する

## 交差遅延モデルの基本原則

### 因果関係の方向を知るには

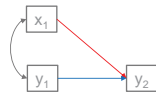
- 1時点のデータ
  - $x \rightarrow y$  と  $y \rightarrow x$  を同時に推定できない
  - 求めるべきパラメータが2つに対し、推定に使用できる共分散の情報1つしかないため（自由度=1-2=-1）
- 2時点のデータ
  - 時間的順序に基づき、1時点目の各変数から2時点目の各変数に対してパスを引く（論理的に逆は必要ない）
    - 同時点の変数間には相関・誤差相関
  - 求めるべきパラメータが6つに対し、共分散の情報4つで6つあるので、推定が可能（自由度=6-6=0）



## 交差遅延モデルの基本原則

### なぜ自己回帰効果の推定が必要なのか

- $x \rightarrow y$  の因果関係にしか関心がないなら、 $x_1$  と  $y_2$  だけで十分なのでは？
- $x_1$  と  $y_2$  の相関は、 $x_1$  と  $y_1$  の相関を介した疑似相関かもしれない
  - つまり、もともと  $x$  と  $y$  に相関があっただけなのか、 $x$  が  $y$  の変化を引き起こしたのか、区別がつかない
- $y_1$  からの自己回帰効果を設定することで、 $y$  の初期値をコントロールした上で  $x_1$  の効果（= $y$  の変化量に対する  $x_1$  の効果）が明らかになる



## 交差遅延モデルの基本原則

### なぜモデルの設定が必要なのか

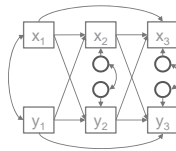
- $y$  の変化量に対する  $x_1$  の効果を知りたいなら、単純に、 $y_2 - y_1$  に対する  $x_1$  の効果を計算すればいいだけでは？
- $y_2 - y_1$  には、 $y$  の真の変化量だけでなく、「平均への回帰」の効果が含まれている
  - 平均への回帰：1回目の測定値が平均から離れているほど、2回目の測定値が平均方向に近づく現象
  - $y_1$  が平均を上回った個人では  $y_2 - y_1$  がマイナス方向に偏り、 $y_1$  が平均を下回った個人では  $y_2 - y_1$  がプラス方向に偏る
  - $y_1$  をモデル上でコントロールすることで、平均への回帰も調整され、 $x_1 \rightarrow y_2$  の適切な推定値が得られる



## 交差遅延モデルの基本原則

### 交差遅延モデルの問題1

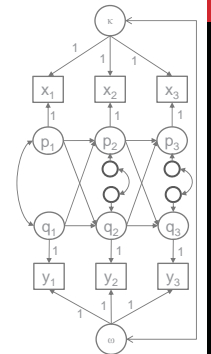
- 3時点以上のデータではモデル適合が悪いケースが多い
  - 2時点前や3時点前からの自己回帰効果を追加してみると改善する なぜ？
  - 多くの構成概念は、安定的な「特性」成分と時間特異的な「状態」成分を併せ持っている
  - 本来、この両方の成分をコントロールした上で交差遅延効果を検証する必要があるが、1時点前からの自己回帰効果を考慮するだけでは不十分
  - 2時点前や3時点前からの自己回帰効果を追加しても十分とは言えない



## 交差遅延モデルの基本原則

### ランダム切片交差遅延モデル (Hamaker et al., 2015)

- 観測変数の変動を、安定的な特性成分（ランダム切片）と時間特異的な状態成分（ランダム切片からの偏差）に分離
- 状態成分で交差遅延モデルを組む
- 特性成分間の相関（共分散）も推定
- 特性成分の分散・共分散を0に固定すれば、通常の交差遅延モデルになる
  - つまり、ネストされた関係にあるため、モデル適合の直接比較が可能



## 交差遅延モデルの基本原則

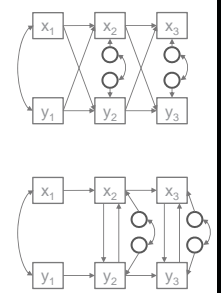
### 交差遅延モデルの問題2

- 測定の時間間隔によって結果が左右される
  - 測定の間隔が実際に因果的影響が生じる間隔と異なっていれば、交差遅延効果は影響の強さを適切に反映しない
  - どの程度の間隔が適切であるかを前もって理論的に判断することは容易でない

## 交差遅延モデルの基本原則

### 同時効果モデル

- 交差遅延効果の代わりに、同時点の測定値間に双方向のパスを設定する
  - 独立変数が従属変数の誤差項と相関しないという前提を逸脱した非逐次モデル
  - 非逐次モデルの識別と推定には、以下の条件を満たす道具的変数の導入が必要
    - 独立変数と相関する
    - 従属変数の誤差項と相関しない
    - 従属変数に直接効果を持たない
  - モデルの仮定のもとで、前時点の観測値がこの条件を満たす
    - 誤差項が自己相関を持っていたり、変数間に同時効果とは独立に交差遅延効果が存在する場合は条件が満たされない



## 交差遅延モデルの分析例

- **リサーチ・クエスチョン**
  - 小中学生の友人関係と抑うつの間にはどのような因果関係があるか
- **調査内容**
  - 抑うつ
    - DSR5-C (Birlerson, 1981; 村田他, 1996)
  - 友人関係
    - SDQ自己評定フォーム (Goodman, 1997)
      - 下位尺度の友人関係問題を分析に使用
- **モデル**
  - 交差遅延モデル
  - ランダム切片交差遅延モデル
  - ランダム切片同時効果モデル

## 交差遅延モデルの分析例

- **参加者**
  - 2011年度から2017年度にかけて7回の縦断調査
  - 13の学年コホートの計9182名が参加
    - 系列コホートデザイン

小学校 入学年度	男子	女子	合計	小学校				中学校		
				3年	4年	5年	6年	1年	2年	3年
2003	392	361	753							753
2004	385	338	723							723
2005	359	393	752						752	752
2006	343	331	674				674		674	674
2007	339	342	681				681	681	681	681
2008	309	349	658			658	658	658	658	658
2009	357	318	675	675	675	675	675	675	675	675
2010	238	303	541	541	541	541	541	541	541	541
2011	318	333	651	651	651	651	651	651	651	651
2012	325	363	688	688	688	688	688	688	688	688
2013	364	386	750	750	750	750				
2014	382	411	793	793	793					
2015	439	404	843	843						
合計	4550	4632	9182	4941	4756	4644	4588	4632	4704	4916

## 交差遅延モデルの分析例

### • ランダム切片モデルのシンタックス (Part 1)

```
DATA: FILE = peer_dep_LCS2.txt;

VARIABLE: NAMES = id gender cohort dep3-dep9 peer3-peer9;
USEVARIABLES = dep3-dep9 peer3-peer9;
MISSING = ALL(999);
PATTERN = cohort;
2003 = dep9 peer9;
2004 = dep8 dep9 peer8 peer9;
2005 = dep7 dep8 dep9 peer7 peer8 peer9;
2006 = dep6 dep7 dep8 dep9 peer6 peer7 peer8 peer9;
2007 = dep5 dep6 dep7 dep8 dep9 peer5 peer6 peer7 peer8 peer9;
2008 = dep4 dep5 dep6 dep7 dep8 dep9 peer4 peer5 peer6 peer7 peer8 peer9;
2009 = dep4 dep5 dep6 dep7 dep8 dep9 peer4 peer5 peer6 peer7 peer8 peer9;
2010 = dep3 dep4 dep5 dep6 dep7 dep8 peer3 peer4 peer5 peer6 peer7 peer8;
2011 = dep3 dep4 dep5 dep6 dep7 peer3 peer4 peer5 peer6 peer7;
2012 = dep3 dep4 dep5 dep6 peer3 peer4 peer5 peer6;
2013 = dep3 dep4 dep5 peer3 peer4 peer5;
2014 = dep3 dep4 peer3 peer4;
2015 = dep3 peer3;
); 系列コホートデザインのデータパターンを指定

ANALYSIS: ESTIMATOR = MLR; !ロバスト最尤法
COVERAGE = 0; !系列コホートデザインによる欠測のため
```

## 交差遅延モデルの分析例

### • ランダム切片モデルのシンタックス (Part 2)

```
MODEL:
!!!抑うつ!!!
!特性成分(ランダム切片)
int_dep BY dep3-dep9@1;

!観測変数の誤差分散を0
dep3-dep9@0;

!状態成分(ランダム切片からの偏差)
dev_dep3 BY dep3@1;
dev_dep4 BY dep4@1;
dev_dep5 BY dep5@1;
dev_dep6 BY dep6@1;
dev_dep7 BY dep7@1;
dev_dep8 BY dep8@1;
dev_dep9 BY dep9@1;

!自己回帰効果
dev_dep4 ON dev_dep3(a);
dev_dep5 ON dev_dep4(a);
dev_dep6 ON dev_dep5(a);
dev_dep7 ON dev_dep6(a);
dev_dep8 ON dev_dep7(a);
dev_dep9 ON dev_dep8(a);
```

## 交差遅延モデルの分析例

### • ランダム切片モデルのシンタックス (Part 3)

```
!!!友人関係!!!
!特性成分(ランダム切片)
int_peer BY peer3-peer9@1;

!観測変数の誤差分散を0
peer3-peer9@0;

!状態成分(ランダム切片からの偏差)
dev_peer3 BY peer3@1;
dev_peer4 BY peer4@1;
dev_peer5 BY peer5@1;
dev_peer6 BY peer6@1;
dev_peer7 BY peer7@1;
dev_peer8 BY peer8@1;
dev_peer9 BY peer9@1;

!自己回帰効果
dev_peer4 ON dev_peer3(b);
dev_peer5 ON dev_peer4(b);
dev_peer6 ON dev_peer5(b);
dev_peer7 ON dev_peer6(b);
dev_peer8 ON dev_peer7(b);
dev_peer9 ON dev_peer8(b);
```

## 交差遅延モデルの分析例

### • ランダム切片モデルのシンタックス (Part 4)

```
!!!Bivariate info!!!
!抑うつ-友人関係の交差遅延効果
dev_dep4 ON dev_peer3(c);
dev_dep5 ON dev_peer4(c);
dev_dep6 ON dev_peer5(c);
dev_dep7 ON dev_peer6(c);
dev_dep8 ON dev_peer7(c);
dev_dep9 ON dev_peer8(c);

!友人関係-抑うつ-友人関係の交差遅延効果
dev_peer4 ON dev_dep3(d);
dev_peer5 ON dev_dep4(d);
dev_peer6 ON dev_dep5(d);
dev_peer7 ON dev_dep6(d);
dev_peer8 ON dev_dep7(d);
dev_peer9 ON dev_dep8(d);

!時点内の相関・誤差相関
dev_dep3 WITH dev_peer3;
dev_dep4 WITH dev_peer4(e);
dev_dep5 WITH dev_peer5(e);
dev_dep6 WITH dev_peer6(e);
dev_dep7 WITH dev_peer7(e);
dev_dep8 WITH dev_peer8(e);
dev_dep9 WITH dev_peer9(e);

!特性成分と状態成分の相関を0に
dev_peer3 dev_dep3 WITH int_dep@0 int_peer@0;

OUTPUT: SAMP STDYX MOD TECH4;
```

## 交差遅延モデルの分析例

### • 交差遅延モデルのシンタックス (追加部分のみ)

```
!!!Bivariate info!!!
(中略)

!特性成分の分散・共分散を0に
int_dep@0 int_peer@0;
int_dep WITH int_peer@0;
```

## 交差遅延モデルの分析例

### • 同時効果モデルのシンタックス (変更部分のみ)

```
!!!Bivariate info!!!
!抑うつ-友人関係の同時効果
dev_dep4 ON dev_peer4(c);
dev_dep5 ON dev_peer5(c);
dev_dep6 ON dev_peer6(c);
dev_dep7 ON dev_peer7(c);
dev_dep8 ON dev_peer8(c);
dev_dep9 ON dev_peer9(c);

!友人関係-抑うつ-友人関係の同時効果
dev_peer4 ON dev_dep4(d);
dev_peer5 ON dev_dep5(d);
dev_peer6 ON dev_dep6(d);
dev_peer7 ON dev_dep7(d);
dev_peer8 ON dev_dep8(d);
dev_peer9 ON dev_dep9(d);
```

## 交差遅延モデルの分析例

### モデル適合度

	交差遅延モデル	ランダム切片 交差遅延モデル	ランダム切片 同時効果モデル
df	81	78	78
$\chi^2$	1382.5	427	423.4
CFI	.938	.983	.983
TLI	.933	.981	.982
RMSEA	.042	.022	.022
SRMR	.110	.063	.064
BIC	261183	259985	259978

## 交差遅延モデルの分析例

	交差遅延モデル		ランダム切片 交差遅延モデル		ランダム切片 同時効果モデル	
	標準化 推定値	$p$	標準化 推定値	$p$	標準化 推定値	$p$
ランダム切片への負荷量						
抑うつ			.643	<.001	.642	<.001
友人関係			.621	<.001	.627	<.001
自己回帰効果						
抑うつ	.580	<.001	.335	<.001	.324	<.001
友人関係	.430	<.001	.197	<.001	.185	<.001
交差遅延効果						
友人関係→抑うつ	.086	<.001	.039	<.001		
抑うつ→友人関係	.188	<.001	.107	<.001		
同時効果						
友人関係→抑うつ					.147	<.001
抑うつ→友人関係					.319	<.001
誤差相関						
抑うつ↔友人関係	.399	<.001	.375	<.001	-.090	.094
ランダム切片間の相関						
抑うつ↔友人関係			.723	<.001	.729	<.001

注: 推定値が複数あるものは中央値を記載

## 交差遅延モデルの分析例

### 考察

- 通常の交差遅延モデルとランダム切片モデルの比較
  - ランダム切片モデルの方が有意に適合が良好
    - 抑うつや友人関係には安定的な個人差が存在する
    - 観測変数の約4割の分散がランダム切片によって説明された (= 特性成分)
  - ランダム切片モデルでは、通常のモデルより自己回帰効果や交差遅延効果の推定値が低かった
    - 本来はランダム切片によって説明されるべき時点間の相関が、通常のモデルでは自己回帰効果に上乗せされている
    - 2変数のランダム切片間に存在する強い相関を、通常のモデルでは交差遅延効果によって穴埋めしようとしている
    - いずれもランダム切片を想定しないことによるパラメータ推定値の偏り (過大推定)

## 交差遅延モデルの分析例

### 考察 (つづき)

- 交差遅延モデルと同時効果モデルの比較
  - 適合度は同時効果モデルの方がわずかに良好
  - 交差遅延効果に比べ、同時効果は3~4倍程度の推定値
    - 1年という時間間隔は、抑うつと友人関係の間の交差遅延効果を評価するには長すぎる可能性がある
    - 年度間でクラス替えがあることも関係するか
  - 交差遅延モデルでは、.375であった誤差相関が、同時効果モデルでは有意でなくなった
    - 交差遅延モデルで誤差相関として表されていた時点内の相関が、同時効果モデルでは同時効果に組み入れられた

## 交差遅延モデルの分析例

### 考察 (つづき)

- 全体の結果から
  - いずれのモデルでも「友人関係→抑うつ」より「抑うつ→友人関係」の効果が強かった
    - 先行研究では友人関係がストレスラーとして抑うつを引き起こすと考えられてきたが、逆の影響の方が強い
    - ストレスラーの蓄積によって抑うつを発症を説明するモデルには限界がある可能性
    - 認知的傾向や生理的変化を考慮に入れる必要がある
  - 抑うつと友人関係のランダム切片間には非常に強い相関
    - 友人関係のランダム切片は自閉症の特性を表す可能性
    - 発達障害と精神疾患は遺伝的リスクを共有しているという見方がある

## 交差遅延モデルのまとめ

### 交差遅延モデルの強み

- 変数間のダイナミックな因果的プロセスを検証できる
- 自己回帰効果とランダム切片によって、一時的安定性 (temporal stability) と時間不変的安定性 (time-invariant stability) の両方を評価できる

### 交差遅延モデルの弱み

- 相関のみを分析の対象とするため、絶対的な水準の変化を検討することができない
- 個々の時点間の変動にのみ着目するため、全ての時点にわたる系統的な変化を扱うことができない

## 交差遅延モデルのまとめ

### 研究応用上の注意点

- ランダム切片を導入する
  - 心理学的構成概念を扱う上で、特性成分を考慮しないのは理論的に不自然である場合がほとんど
  - 3時点以上のデータが必要
- 剰余変数との交絡を回避する
  - 交差遅延モデルは横断データの分析と同じく、剰余変数との交絡には無防備
  - 交絡を生じさせうる変数は統制するかモデルに含める
- 概念の重複に注意する
  - 概念 (項目内容) そのものが重複している場合、交差遅延効果や同時効果が過大推定される
  - あらかじめCFAで重複の有無を確認する

## 縦断データ分析のまとめ

### 潜在成長モデルと交差遅延モデルを統合したモデルもある

- Latent Change Score Model (McArdle, 2009)
- Latent Curve Model with Structural Residuals (Curran et al., 2013)
- 問題
  - 成長モデル部分のわずかな不適合が交差遅延モデル部分のパラメータ推定に歪みをもたらす
  - 変化が「成長モデルに適合した変化」(傾き)と「それ以外の時間特異的な変化」(残差)に分けられるため、結果の解釈が難しい
  - 目的や利用シーンが異なる2つのモデルを同時に推定する必要が高くない

## 縦断データ分析のまとめ

### モデルの使い分け

- 潜在成長モデルに適したシーン
  - 因果関係に関心がない
    - 予測が目的である場合
  - 論理的に因果関係の方向が明らか
    - Time-invariantな要因の影響
    - 時間的順序が明確
- 交差遅延モデルに適したシーン
  - 同じ変数セットを複数回にわたって測定したパネルデータの分析
    - 同時並行的に変化する複数の変数間にどのような因果関係があるかを調べる

## おわりに

### ご清聴ありがとうございました！

### アンケートにご協力ください

- 今後の企画に役立ててまいりますので、率直な御質問、御意見を聞かせください
- いただいた御質問については、後日、可能な限り、メールまたは講師（伊藤）のホームページ上で回答いたします
  - ホームページ
    - [http://www006.upp.so-net.ne.jp/ito\\_h/](http://www006.upp.so-net.ne.jp/ito_h/)
  - メールアドレス
    - ito\_hiroyuki@pd5.so-net.ne.jp

## 交差遅延モデル 補足資料

### 当初のモデルでは時点間で多数の等値制約を置いている

- ランダム切片への負荷量  $6 \times 2 = 12$
- 自己回帰効果  $6 \times 2 = 12$
- 交差遅延効果（同時効果）  $6 \times 2 = 12$
- 誤差相関 6

### 自由度に余裕があれば、これらの制約は解くことができる

- 今回のモデルは自由度が78あるので、全ての制約（42個）を解いても識別可能
- 時点数が多いほど、自由度に余裕が生まれ、自由なモデリングが可能になる

## 交差遅延モデル 補足資料

### 問題

- 全てのパラメータを個別に自由推定するとモデルの検索性が低下し、個々のパラメータの推定精度や検出力も大きく低下する→必要以上に制約を解きたくない

### 解決策

- 方法1：部分ごとに制約を解き、適合度が改善しなければ元に戻す
  - 適合度の悪化は防げるが、制約が解かれた部分の推定精度や検出力は落ちる
- 方法2：等値制約の代わりに、異なる形の制約を課す
  - 例えば、パラメータが時間（学年）とともに直線的に変化するという線形制約を課す
  - 6つのパラメータを推定する代わりに、初期値と傾きを推定するだけでよいので検索性が高く、推定精度も維持できる

## 交差遅延モデル 補足資料

### 線形制約のシンタックス（変更部分のみ1）

```
MODEL:
!!!抑うつ!!!                (中略)
!特性成分(ランダム切片)
int_dep BY dep3@1           !自己回帰効果
dep4(a1)                    dev_dep4 ON dev_dep3(b1);
dep5(a2)                    dev_dep5 ON dev_dep4(b2);
dep6(a3)                    dev_dep6 ON dev_dep5(b3);
dep7(a4)                    dev_dep7 ON dev_dep6(b4);
dep8(a5)                    dev_dep8 ON dev_dep7(b5);
dep9(a6);                   dev_dep9 ON dev_dep8(b6);
```

等値制約を置いていたパラメータに個別にラベルを与える

## 交差遅延モデル 補足資料

### 線形制約のシンタックス（変更部分のみ2）

```
!!!友人関係!!!
!特性成分(ランダム切片)    (中略)
int_peer BY peer3@1
peer4(c1)                   dev_peer4 ON dev_peer3(d1);
peer5(c2)                   dev_peer5 ON dev_peer4(d2);
peer6(c3)                   dev_peer6 ON dev_peer5(d3);
peer7(c4)                   dev_peer7 ON dev_peer6(d4);
peer8(c5)                   dev_peer8 ON dev_peer7(d5);
peer9(c6);                   dev_peer9 ON dev_peer8(d6);
```

## 交差遅延モデル 補足資料

### 線形制約のシンタックス（変更部分のみ3）

```
!!!Bivariate info!!!
!抑うつ←友人関係の同時効果
dev_dep4 ON dev_peer4(e1);
dev_dep5 ON dev_peer5(e2);
dev_dep6 ON dev_peer6(e3);
dev_dep7 ON dev_peer7(e4);
dev_dep8 ON dev_peer8(e5);
dev_dep9 ON dev_peer9(e6);

!友人関係←抑うつ同時効果
dev_peer4 ON dev_dep4(f1);
dev_peer5 ON dev_dep5(f2);
dev_peer6 ON dev_dep6(f3);
dev_peer7 ON dev_dep7(f4);
dev_peer8 ON dev_dep8(f5);
dev_peer9 ON dev_dep9(f6);

!時点内の誤差相関
dev_dep3 WITH dev_peer3;
dev_dep4 WITH dev_peer4(g1);
dev_dep5 WITH dev_peer5(g2);
dev_dep6 WITH dev_peer6(g3);
dev_dep7 WITH dev_peer7(g4);
dev_dep8 WITH dev_peer8(g5);
dev_dep9 WITH dev_peer9(g6);
```

## 交差遅延モデル 補足資料

### 線形制約のシンタックス（変更部分のみ4）

```
MODEL CONSTRAINT:
NEW (as*0 bs*0 cs*0 ds*0 es*0 fs*0 gs*0);

a1 = 1+as*1;
a2 = 1+as*2;
a3 = 1+as*3;
a4 = 1+as*4;
a5 = 1+as*5;
a6 = 1+as*6;

b2 = b1+bs*1;
b3 = b1+bs*2;
b4 = b1+bs*3;
b5 = b1+bs*4;
b6 = b1+bs*5;

c1 = 1+cs*1;
c2 = 1+cs*2;
c3 = 1+cs*3;
c4 = 1+cs*4;
c5 = 1+cs*5;
c6 = 1+cs*6;

d2 = d1+ds*1;
d3 = d1+ds*2;
d4 = d1+ds*3;
d5 = d1+ds*4;
d6 = d1+ds*5;
```

学年が上がること  
に初期値に傾き  
が一つ追加さ  
れていく線形制約  
(因子負荷量は  
1時点目を1に固  
定しているため、  
1が初期値)

初回のパラメータ  
が初期値

パラメータの時間的変化  
の傾きを表すパラメータ  
を導入

## 交差遅延モデル 補足資料

### 線形制約のシンタックス (変更部分のみ5)

$e2 = e1 + es*1;$                        $g2 = g1 + gs*1;$   
 $e3 = e1 + es*2;$                        $g3 = g1 + gs*2;$   
 $e4 = e1 + es*3;$                        $g4 = g1 + gs*3;$   
 $e5 = e1 + es*4;$                        $g5 = g1 + gs*4;$   
 $e6 = e1 + es*5;$                        $g6 = g1 + gs*5;$

$f2 = f1 + fs*1;$   
 $f3 = f1 + fs*2;$   
 $f4 = f1 + fs*3;$   
 $f5 = f1 + fs*4;$   
 $f6 = f1 + fs*5;$

## 交差遅延モデル 補足資料

### モデル適合度

適合が改善

	交差遅延モデル	ランダム切片 交差遅延モデル	ランダム切片 同時効果モデル	ランダム切片 同時効果モデル (線形制約)
df	81	78	78	71
$\chi^2$	1382.5	427	423.4	236.5
CFI	.938	.983	.983	.992
TLI	.933	.981	.982	.990
RMSEA	.042	.022	.022	.016
SRMR	.110	.063	.064	.043
BIC	261183	259985	259978	259799

## 交差遅延モデル 補足資料

### パラメータ推定値

#### MODEL RESULTS

	Estimate	S.E.	Two-Tailed Est./S.E.	P-Value
(中略)				
New/Additional Parameters				
AS	0.081	0.019	4.310	0.000
BS	0.019	0.012	1.550	0.121
CS	0.059	0.017	3.524	0.000
DS	-0.002	0.009	-0.284	0.776
ES	-0.030	0.044	-0.683	0.495
FS	-0.001	0.011	-0.086	0.932
GS	0.059	0.079	0.745	0.456

抑うつ・友人関係のランダム切片の負荷量のみ、傾きパラメータが有意

有意でなかった傾きパラメータは0に固定してもOK (元の等値制約と同じ状態に戻る)

## 交差遅延モデル 補足資料

#### STDYX Standardization

	Estimate	S.E.	Two-Tailed Est./S.E.	P-Value
INT_DEP BY				
DEP3	0.572	0.026	21.814	0.000
DEP4	0.600	0.019	32.081	0.000
DEP5	0.639	0.013	47.430	0.000
DEP6	0.671	0.010	67.806	0.000
DEP7	0.679	0.013	51.769	0.000
DEP8	0.691	0.017	39.701	0.000
DEP9	0.724	0.022	32.834	0.000
(中略)				
INT_PEER BY				
PEER3	0.472	0.023	20.270	0.000
PEER4	0.514	0.019	27.352	0.000
PEER5	0.578	0.014	40.504	0.000
PEER6	0.623	0.010	59.621	0.000
PEER7	0.650	0.011	61.000	0.000
PEER8	0.672	0.016	43.034	0.000
PEER9	0.696	0.018	39.761	0.000

学年が上がるごとに負荷量が増加→安定性が増す

## 付録資料

### SEMの基本原理解

中部大学 伊藤大幸

## 構造方程式モデリング (SEM) とは

### SEM = 因子分析 + パス解析 (≒回帰分析)

#### 因子分析 (測定モデル)

- 概念そのものの構造を検討する
- 探索的因子分析 (Exploratory Factor Analysis: EFA) : 特定の仮説を持たず、データから因子構造を探索する
- 確認的因子分析 (Confirmatory Factor Analysis: CFA) : 因子構造に関する明確な仮説を立て、データによって検証する

#### パス解析 (構造モデル)

- 概念間の因果関係を検討する

## SEMの基本的原理

### SEMの基本的原理

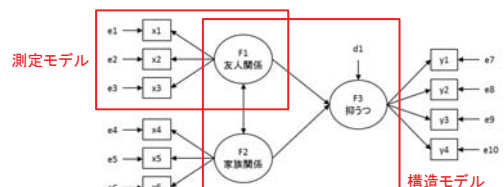
- 複数の変数間の関連性について、特定の定性的な仮説モデルをパス図として設定した上で、
- そのモデルの妥当性を適合度指標によって検証し、
- モデルの仮定のもとでの変数間の定量的な関連性 (パラメータ) を推定する

- 研究者にとっての順序は1→2→3
- 解析上の順序は1→3→2
- 以降、解析上の順序にしたがって解説

## モデルの指定

### パス図

- 四角: 観測変数 (指標)
- 楕円: 潜在変数 (因子)
- 囲まれていない変数: 誤差変数
- 単方向の矢印: 因果関係
- 双方向の矢印: 相関関係



## パラメータの推定

### ・構造モデル (丁度識別)

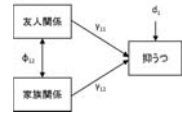
変数間の相関をモデル上のパラメータで表現

相関構造 (モデル)			
友人関係	家族関係	抑うつ	
友人関係	-	-	
家族関係	$\phi_{12}$	-	
抑うつ	$\gamma_{11} + \phi_{12}\gamma_{12}$	$\phi_{12}\gamma_{11} + \gamma_{12}$	-

$$\begin{aligned} \phi_{12} &= 0.5 \\ \gamma_{11} + \phi_{12}\gamma_{12} &= 0.3 \\ \phi_{12}\gamma_{11} + \gamma_{12} &= 0.4 \end{aligned}$$

$$\begin{aligned} \phi_{12} &= 0.5 \\ \gamma_{11} &= 0.13 \\ \gamma_{12} &= 0.33 \end{aligned}$$

※実際は相関行列でなく、分散・共分散行列を使用



相関行列 (データ)			
友人関係	家族関係	抑うつ	
友人関係	-	-	
家族関係	.50	-	
抑うつ	.30	.40	-



複数の独立変数から一つの従属変数にパスが設定されている場合、パス係数は他方の変数を統制したときの当該変数の効果を表す

## パラメータの推定

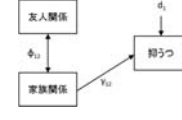
### ・構造モデル (過剰識別)

相関構造 (モデル)			
友人関係	家族関係	抑うつ	
友人関係	-	-	
家族関係	$\phi_{12}$	-	
抑うつ	$\gamma_{11} + \phi_{12}\gamma_{12}$	$\phi_{12}\gamma_{11} + \gamma_{12}$	-

$$.50 \times .40 = .20$$

$$\begin{aligned} \phi_{12} &= 0.5 \\ \phi_{12}\gamma_{12} &= 0.3 \\ \gamma_{12} &= 0.4 \end{aligned}$$

パラメータの数より方程式の数が多く、解が一意に定まらないため、最もデータによくあてはまる解を推定 (最適化)



相関行列 (データ)			
友人関係	家族関係	抑うつ	
友人関係	-	-	
家族関係	.50	-	
抑うつ	.30	.40	-



$\chi^2(1) = 7.30, p = .007, CFI = .971, RMSEA = .118$

適合度指標

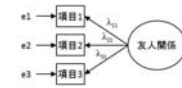
## パラメータの推定

### ・測定モデル (丁度識別)

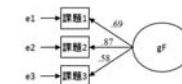
相関構造 (モデル)			
項目1	項目2	項目3	
項目1	-	-	
項目2	$\lambda_{11}\lambda_{21}$	-	
項目3	$\lambda_{11}\lambda_{31}$	$\lambda_{21}\lambda_{31}$	-

$$\begin{aligned} \lambda_{11}\lambda_{21} &= 0.6 \\ \lambda_{11}\lambda_{31} &= 0.4 \\ \lambda_{21}\lambda_{31} &= 0.5 \end{aligned}$$

$$\begin{aligned} \lambda_{11} &= 0.69 \\ \lambda_{21} &= 0.87 \\ \lambda_{31} &= 0.58 \end{aligned}$$



相関行列 (データ)			
項目1	項目2	項目3	
項目1	-	-	
項目2	.60	-	
項目3	.40	.50	-



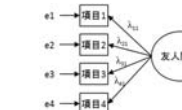
## パラメータの推定

### ・測定モデル (過剰識別)

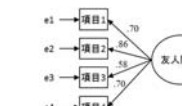
相関構造 (モデル)				
項目1	項目2	項目3	項目4	
項目1	-	-	-	
項目2	$\lambda_{11}\lambda_{21}$	-	-	
項目3	$\lambda_{11}\lambda_{31}$	$\lambda_{21}\lambda_{31}$	-	
項目4	$\lambda_{11}\lambda_{41}$	$\lambda_{21}\lambda_{41}$	$\lambda_{31}\lambda_{41}$	-

$$\begin{aligned} \lambda_{11}\lambda_{21} &= 0.6 \\ \lambda_{11}\lambda_{31} &= 0.4 \\ \lambda_{21}\lambda_{31} &= 0.5 \\ \lambda_{11}\lambda_{41} &= 0.5 \\ \lambda_{21}\lambda_{41} &= 0.6 \\ \lambda_{31}\lambda_{41} &= 0.4 \end{aligned}$$

パラメータの数より方程式の数が多く、解が一意に定まらないため、最もデータによくあてはまる解を推定 (最適化)



相関行列 (データ)				
項目1	項目2	項目3	項目4	
項目1	-	-	-	
項目2	.60	-	-	
項目3	.40	.50	-	
項目4	.50	.60	.40	-



$\chi^2(2) = 0.25, p = .88, CFI = 1.000, RMSEA = .00$

## パラメータの推定

### ・測定モデル (過小識別)

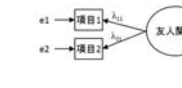
相関構造 (モデル)		
項目1	項目2	
項目1	-	
項目2	$\lambda_{11}\lambda_{21}$	-

$$\lambda_{11}\lambda_{21} = 0.6$$

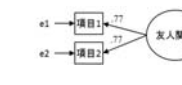
パラメータの数より方程式の数が少ないため、解が一意に定まらず、推定もできない (方程式を満たす解の組み合わせが無数に存在)

何らかの制約を課すことにより、推定すべきパラメータ (自由パラメータ) の数を減らせば、解を求めることが可能になる

$$\lambda_{11} = \lambda_{21}$$



相関行列 (データ)		
項目1	項目2	
項目1	-	
項目2	.60	-



## パラメータの推定

### ・モデルの識別性

- 過剰識別: 自由度 (方程式の数 - 自由パラメータ数) > 0
- 丁度識別: 自由度 = 0
- 過小識別: 自由度 < 0 (※自由度が0以上でもなりうる)
- SEMにおいては、過剰識別が最も望ましい
- なぜ?
  - 丁度識別は誤差を含むデータを「鵜呑み」
  - 過剰識別の状態では、理論的知識に基づく事前情報を利用して、データへの依存度を減らしている
  - 事前情報の確かさは適合度指標によって評価できる

## パラメータの推定

### ・パラメータの制約

- モデルの識別性や推定の精度を高めるために、パラメータに制約を課すことがある
- 等値制約
  - 因子負荷量の等値制約
  - グループ間の等値制約
- 固定
  - 確認的因子分析の不定性の解消
    - 一つの指標の負荷量を1に固定
    - 因子分散を1に固定
  - 変数間にパスを引かない

## パラメータの推定

### ・推定法

- 過剰識別のモデルでは、連立方程式の解が一意に定まらないため、方程式を最大限満たすような最適解を探索する
- どのような解を最適解と見なすかは推定法によって異なる
  - 最小二乗法: モデル上の分散・共分散構造とデータ上の分散・共分散行列の各成分の差の二乗和を最小化するようなパラメータの組み合わせを最適解とする
  - 最尤法: データ上の分散・共分散行列が得られる確率を最大化するような (最も尤もらしい) パラメータの組み合わせを最適解とする

## パラメータの推定

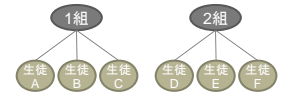
### • 最尤法の利点

- **一貫性**
  - サンプルサイズが大きいほど、真の値に収束
- **効率的**
  - サンプルサイズが十分に大きいとき、他の推定法より推定誤差が小さくなる
- **正規性**
  - サンプルサイズが大きいほど、推定値の分布が正規分布に近似
- **尺度不変性**
  - 変数のスケールによらず同一の解を与える
- 適合度指標としての**カイ二乗値**が得られる
- **欠測値**を含むデータでも偏りのない推定値をもたらす

## パラメータの推定

### • 最尤法の基本的前提

- **観測の独立性**
  - データが局所的な類似性（**級内相関**）を持たないこと
  - 階層構造を成すデータは**多母集団モデル**や**マルチレベルモデル**で処理する必要がある
- **内生変数の多変量正規性**
  1. 全ての内生変数の単変量分布が**正規分布**にしたがうこと
  2. 全ての内生変数のペアの散布図が**線形**であること
  3. 誤差の分布が**等分散** (homoscedastic) であること
- 変数の分布に応じたモデルを用いるか、**ロバスト推定法**によって対処



## パラメータの推定

### • 欠測値

- 最尤法は欠測値を含むデータでも偏りのない推定値を与える (**完全情報最尤法**)
  - リストワイズ法などデータを削除する方法より望ましい
- 最尤法によって対処できる欠測値
  - **MCAR** (Missing Completely at Random)
    - 欠測値の有無が他の変数や当該変数の値と関係しない
  - **MAR** (Missing at Random)
    - 欠測値の有無が当該変数の値と直接関係しない
- 最尤法によって対処できない欠測値
  - **MNAR** (Missing Not at Random)
    - 欠測値の有無が当該変数の値と直接関係する

## 適合度指標

### • 適合度指標とは

- 研究者が指定した仮説モデルが、どの程度、データに適合しているかを定量化した指標
- **なぜ適合度指標が必要なのか?**
  - SEMは、従来の方法に比べ、積極的に研究者の理論や仮説を利用する
    - 構造モデル：特定の因果モデルを仮定することで、本来方向性を持たない相関関係から、方向性のある因果関係を見出そうとする
    - 測定モデル：明確な理論的モデルをあてはめて因子構造を検証する
  - 適合度指標はモデルの**反証**の手段として機能することで、積極的なモデリングの裏づけを提供

## 適合度指標

### • 絶対的な適合の指標

- **カイ二乗検定**：有意でなければ適合が良好
  - ※構造モデルの評価には有用だが、サンプルサイズが中程度以上の場合、測定モデルの評価には厳しくなりすぎる
- **CFI**：.95以上で良好、.90以上でおおむね良好
- **RMSEA**：.05以下で良好、.05~.10でおおむね良好
- **SRMR**：.08以下で良好、.10以下でおおむね良好
  - ※これらの基準は、CFA（測定モデル）におけるシミュレーションに基づいて設定されたもので、**構造モデルの評価には必ずしも適さない**

## 適合度指標

### • 相対的な適合の指標

- **尤度比検定**
  - 基準モデルと**ネストされたモデル**のカイ二乗値の差の検定
  - ネストされたモデルとは
    1. 1つ以上の制約が加わっている
    2. 新しいパラメータが加わっていない
- **AIC・BIC**
  - 値が小さいモデルほど、**儉約性**の割に適合が良好
  - ネストされたモデルでなくても比較が可能だが、観測変数の組み合わせやサンプルは同一でなくてはならない

## 適合度指標

### • 部分的な適合の指標

- **残差行列**
  - モデル上の分散・共分散構造とデータ上の分散・共分散行列の差
- **修正指標**（ラグランジュ乗数検定）
  - 個々のパラメータ制約を解いた場合に、どの程度、モデルのカイ二乗値が低下するか（適合度が改善するか）
    - 3.84以上のとき5%水準、6.63以上のとき1%水準、10.83以上のとき0.1%水準で、モデルが有意に改善することを意味する