今日からはじめる！ Mplus と R による 構造方程式モデリング

小杉考司（山口大学）
清水裕士（広島大学）

おしながき

- M*plus*導入編　11：35～12：05(30min)
- M*plus*活用編　12：05～12：25(20min)
- Rでやってみる　12：25～12：40(15min)
- 質疑など　12：40～13：00(20min)

講師自己紹介

- 小杉考司
  - 所属;山口大学教育学部
  - 専門;社会心理学
  - 経歴;M*plus*歴8年，R歴7年
- 清水裕士
  - 所属;広島大学大学院総合科学研究科
  - 専門;社会心理学
  - 経歴;M*plus*歴7年，R歴1年

受講者の皆様は・・・

- 構造方程式モデリング（SEM）を始めたい
  - 敷居は低くあってほしい
  - 機能は豊富にほしい（後で学び直したくない）
- SEMのソフトは触ったことあるけど思うようにモデリングできない
  - もっと自由に，もっといろいろなモデルを

どちらのニーズにも，M*plus*(とR)は応えます！

What is M*plus*?

- Muthén＆Muthénが作った統計ソフトウェア
  - 構造方程式モデリングのためのソフト
  - http://www.statmodel.com/
- 非常に高度な分析が可能
  - 他のSEMソフトができることは全部可能
  - カテゴリカルデータ分析、マルチレベル分析、潜在クラス分析、それらの組み合わせ、など
- CUI（Character User Interface)を採用
  - 最初はやや慣れが必要
  - 最新版（M*plus*7以降）は、一部GUIに対応

What is M*plus*?

- 機能の割にかなり安い
  - 学生の場合

Mplus Version 7.11 Products	Version 7.11 Student Pricing
Mplus Base Program	$195
Mplus Base Program and Mixture Add-On	$240
Mplus Base Program and Multilevel Add-On	$240
Mplus Base Program and Combination Add-On	$350

  - 教員の場合

Mplus Version 7.11 Products	Version 7.11 Pricing For New Licencees
Mplus Base Program	$595
Mplus Base Program and Mixture Add-On	$745
Mplus Base Program and Multilevel Add-On	$745
Mplus Base Program and Combination Add-On	$895

  - しかも、同じバージョンなら永久に使える
    - アップデートする場合は、1年間は無料。それ以降は有料
    - でも更新料は175ドルほど！

他のソフトとの違い

- GUI（Amosなど）かCUI（M*plus*）か
  - これには一長一短がある
    - 簡単なパスモデルはGUIのほうが簡単
    - M*plus*だと、”F1 by v1-v40;” というコードだけで40項目の因子分析ができるが、GUIだと大変。
- ライセンスの扱い
  - M*plus*は一度購入すれば同じバージョンならずっと使える
  - Win/Mac/Linux版がひとつのアカウントで入手可能
- 日本語への対応
  - M*plus*は完全非対応。Rは日本語でできなくもないが，まだ不便＆不完全なところも…

What is R?

- フリーソフトウェアの統計プラットフォーム
- できる分析モデルの豊富さは抜群
  - あらゆる統計分析ができる言語/環境
- フリーソフトウェアだから金銭的コストはゼロ
  - 自由かつオープンであることの強み
- 4,000を超えるパッケージが公開されている
  - 勿論SEMに関するパッケージもね

M*plus*導入編

まずは手に入れましょう

“Mplus　Muthen”などで検索するとstatmodel.comにすぐ着きます

アカウントを作って購入

Muthen & Muthen

Home | Mplus Home Page | Terms & Conditions

Mplus

Categories
- Mplus Software
  - University Pricing
  - Commercial/Non-Profit/ Government Pricing
  - Student Pricing
- Mplus Version 7 User's Guide
- Add-on Purchases
- Discount Information
- Payment Information

Home > ... Previous Page > Sign in

Create a New Account

Create an account with us and you'll be able to:

- Check out faster
- Save multiple shipping addresses
- Access your order history
- Track new orders
- Save items to your wish list

Click here to create a new account.

Sign in to Your Account

Email Address:

Password:

Sign In

Forgot your password?

Your Account Has Been Created

Thank you for creating your account at Muthen & Muthen. Your account details have been emailed to [redacted]

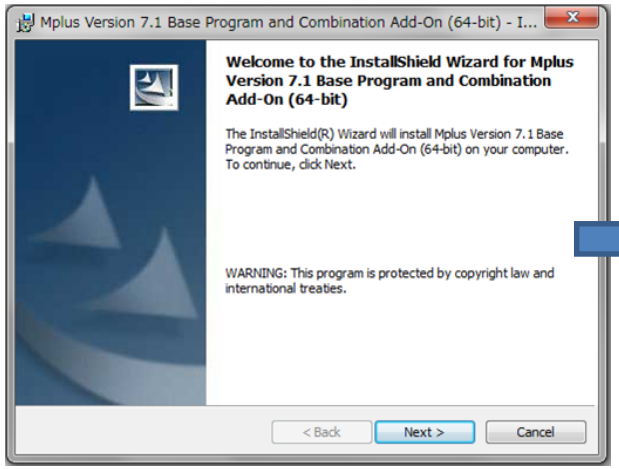
Click here to continue shopping

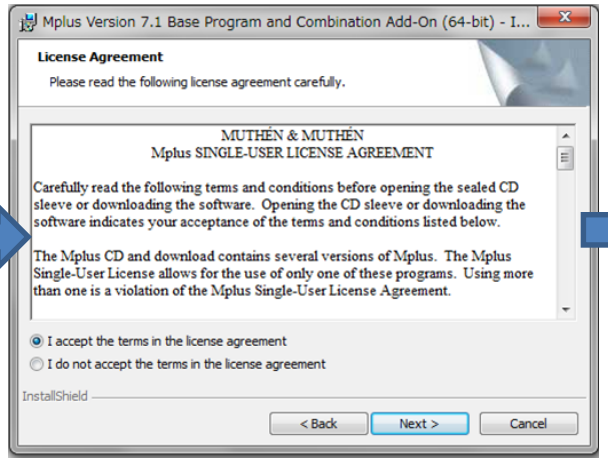
Mplus Version 7.1 Base Program and the Combination Add-On Single-User License with pdf User's Guide (Students). Download contains both 32-bit and 64-bit programs for Microsoft Windows, Mac OS X, and Linux. **$350.00** 1

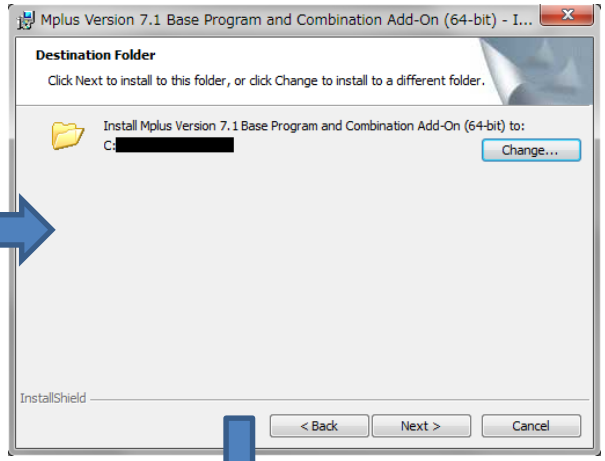
The Mplus Base Program and Combination Add-On contains all of the features of the Mplus Base Program and the Mixture and Multilevel Add-Ons. In addition, it includes models that handle both clustered data and latent classes...

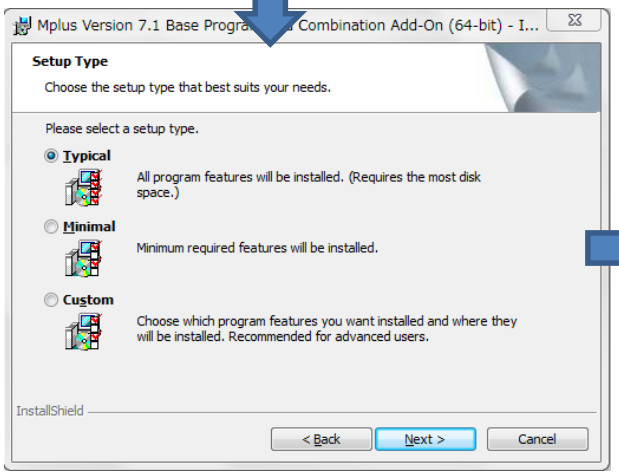
add to cart

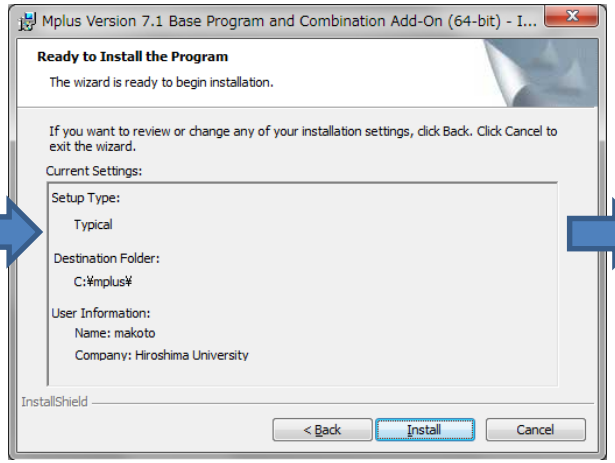
インストールはNext,Next....Finish!

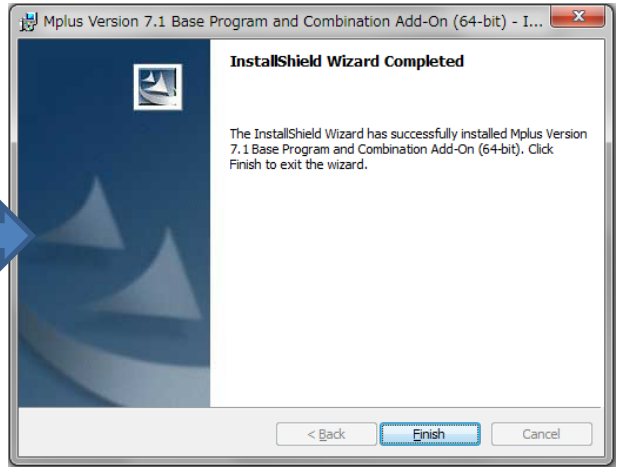




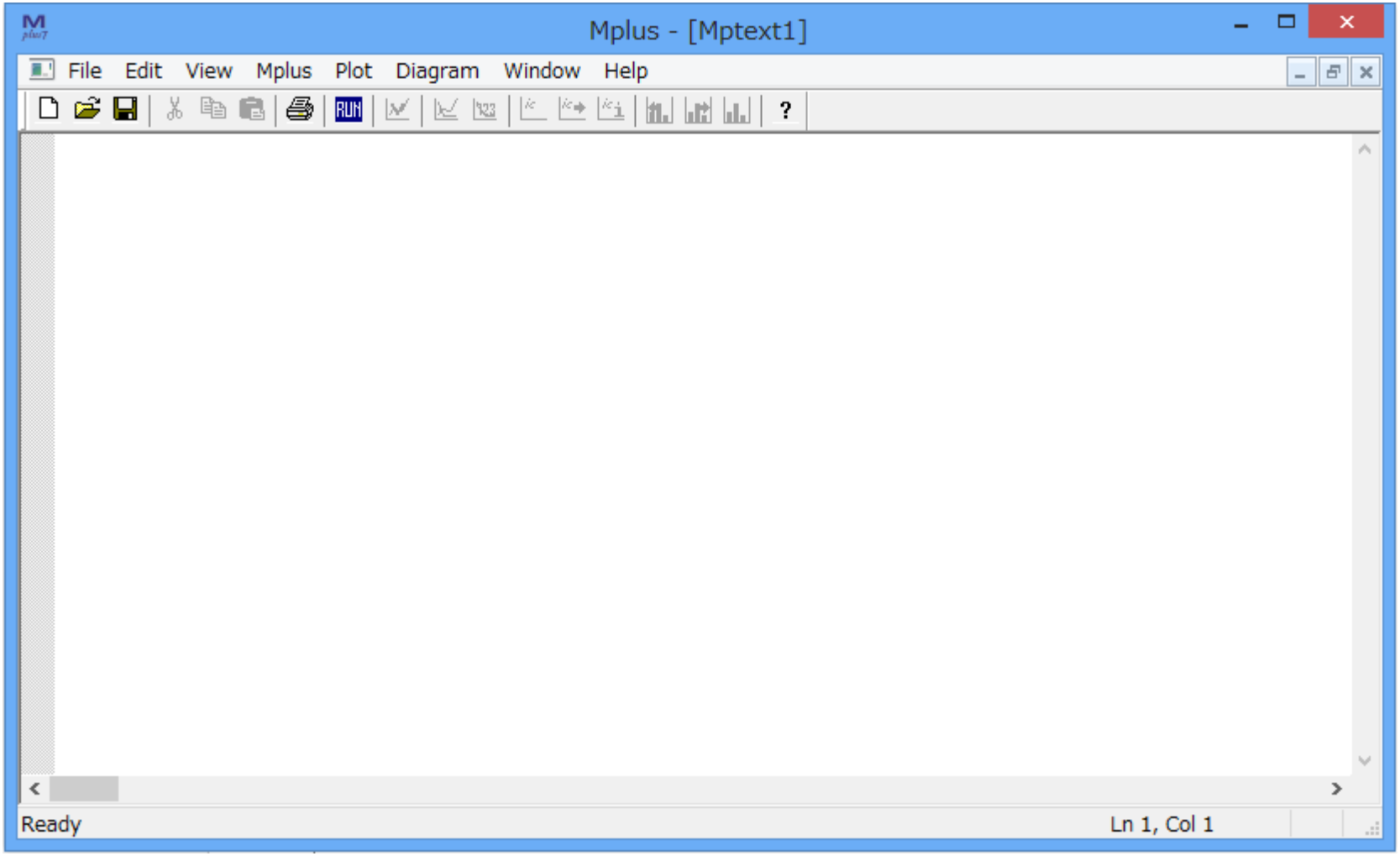








起動した画面がこちら


Mplus - [Mptext1]
File Edit View Mplus Plot Diagram Window Help
RUN
?
Ready
Ln 1, Col 1

この後の流れ

- データの読み込み
  - データセットの準備
  - 基本的なコードの書き方
- とりあえず回帰分析
  - 重回帰分析
- とりあえず因子分析
  - 確認的因子分析

Mplusの構成

- データファイル
  - プレーンテキストのデータ形式。Mplusのデフォルトでは拡張子「.dat」だが，「.txt」でも可能。
- 入力ファイル
  - inpファイルと呼ばれる。これにすべてのコードを書く。
  - 拡張子は「.inp」，中身はプレーンテキスト。
- 出力ファイル
  - outファイルと呼ばれる。これに分析結果が出力される。
  - 拡張子は「.out」，中身はプレーンテキスト。

データファイルの読み込み

- 読み込みファイル
  - datファイル(.dat)
  - テキストファイル(.txt)
  - csvファイル(.csv)
- データ区切り
  - スペース
  - タブ
  - カンマ
- オススメは、タブ区切りのテキストファイル

使いはじめる前に

- M*plus*用のフォルダを作ろう
  - M*plus*は日本語のパスを読めない！
    - ユーザー名が日本語の人は、マイドキュメントにおいても認識されない
  - オススメの使い方は・・・
    - C:¥mplusfile¥～など、Cドライブに直接フォルダを作るとよい。

例：エクセルのデータからMplusへ　1

- エクセルのデータを選択してコピー
  - 変数名は含めないのがミソ

	A	B	C	D	E	F	G	H	I	J	K
1	ID	V1	V2	V3	V4	V5	V6	x1	x2	x3	y
2	1	5	5	6	5	5	4	5	72	4	26
3	2	5	6	6	6	7	6	6	31	7	18
4	3	4	4	5	5	4	4	4	83	0	36
5	4	6	4	4	7	6	5	6	90	8	18
6	5	7	6	6	5	6	4	7	102	15	44
7	6	5	6	5	5	6	4	2	86	15	35
8	7	5	5	5	4	4	6	6	6	36	11
9	8	4	4	5	6	5	6	5	71	18	36
10	9	7	6	6	5	5	5	7	35	5	25
11	10	4	3	3	5	5	4	4	100	19	46
12	11	4	5	4	6	5	6	5	79	23	0
13	12	5	4	4	5	6	4	4	39	22	24

例：エクセルのデータからMplusへ　2

- メモ帳にそのまま張り付ける

無題 - メモ帳

ファイル(F)　編集(E)　書式(O)　表示(V)　ヘルプ(H)

1	5	5	6	5	5	4	5	72	4	26
2	5	6	6	6	7	6	6	31	7	18
3	4	4	5	5	4	4	4	83	0	36
4	6	4	4	7	6	5	6	90	8	18
5	7	6	6	5	6	4	7	102	15	44
6	5	6	5	5	6	4	2	86	15	35
7	5	5	5	4	4	6	6	6	36	11
8	4	4	5	6	5	6	5	71	18	36
9	7	6	6	5	5	5	7	35	5	25
10	4	3	3	5	5	4	4	100	19	46
11	4	5	4	6	5	6	5	79	23	0
12	5	4	4	5	6	4	4	39	22	24
13	5	4	5	6	4	5	4	81	7	35
14	5	6	5	5	3	6	4	51	15	27
15	5	6	5	4	4	6	6	33	19	18
16	5	7	6	5	5	5	5	86	16	56
17	5	4	4	5	4	5	2	18	10	-3
18	4	4	4	6	6	6	4	48	-1	21
19	4	6	5	7	7	7	3	46	24	12
20	5	5	5	5	4	5	4	20	19	30
21	5	6	6	6	6	7	5	43	20	26
22	6	6	6	7	7	6	5	66	0	36
23	5	4	3	5	4	4	5	30	-2	14
24	5	5	6	6	6	5	5	75	14	27
25	5	7	5	5	5	5	4	102	21	42
26	5	5	5	6	6	6	5	39	35	7
27	4	4	4	5	5	6	4	42	3	3

例：エクセルのデータからMplusへ　3

- ファイルの保存場所
  - Mplusの入力ファイル(.inp)と同じフォルダに入れること。詳細は後述。
  - さっき作った、Mplus用のフォルダの中に入れておけばよろしいかと。

- 保存方法
  - 保存するとき、拡張子は.datのほうが便利かもしれないけど、そこは好みで。
    - .datか、.txt、.csvあたりで。
  - 日本語は不対応なので、半角英数字で。

Mplus言語のルール

- 文章はすべて半角英数字
  - 日本語不可
  - 大文字・小文字の区別はない
- 文末にはセミコロン”;”を付ける
  - 改行だけでは、同じ文章だと判断される
- “IS”や”ARE”は”=“で代用できる
  - マニュアルには”IS”とか”ARE”が使われているが、”=“で大丈夫。
- その他
  - コメントアウトは”!”

Mplusの基本コマンド

- DATA:
  - データファイルのある場所を指定
- VARIABLE:
  - 分析に使う変数の指定
- ANALYSIS:
  - 分析方法の指定
- MODEL:
  - モデルの指定
- OUTPUT:
  - 出力の設定
- その他
  - PLOT、SAVEDATA、DEFINEなどなど

こんな感じ

Mplus - [Mptext1]

File Edit View Mplus Plot Diagram Window Help

```
TITLE:  Mplus sample

DATA:
  FILE IS "Y:¥Dropbox¥Mplus¥sample1.txt";

VARIABLE:
  NAMES ARE ID v1-v6 x1-x3 y;
  USEVARIABLES ARE x1-x3 y;

ANALYSIS:
  TYPE IS GENERAL;
  ESTIMATOR IS ML;
  ITERATIONS = 1000;
  CONVERGENCE = 0.00005;

MODEL:


OUTPUT:
  SAMPSTAT STDYX MODINDICES(ALL);
```

Ready Ln 16, Col 1 CAP

DATA: コマンド

- 使うデータファイルを指定する
  - ファイルのパスを指定する
    - FILE = “D:¥Dropbox¥・・・¥sample1.txt”;
    - パスの最後にセミコロンを忘れずに！
  - inpファイルと同じフォルダに入れる場合・・・
    - FILE = “sample1.txt”;
    - ファイル名だけでよい。こちらが簡単なのでオススメ

VARIABLE: コマンド

- 読み込んだ変数を指定する
  - データファイルにある変数名をすべて書く
  - NAMES = v1 v2 v3;
    - 上の場合、v1-v3; と書くこともできる
- 分析に使用する変数を指定する
  - USEVARIABLES = v1 v2;
    - 書き方は上と同じ
- 欠損値を指定する
  - MISSING = .;
    - ピリオドの場合は上のように書く
    - ダブルクォーテーションはつけない

ANALYSIS: コマンド

- 分析方法を指定する
  - TYPE = GENERAL;
    - 基本的なSEMをする場合はこれだけでOK。
    - 他に、TWOLEVEL（マルチレベルモデル）、MIXTURE（潜在クラス分析）、EFA（探索的因子分析）などがある。
  - ESTIMATOR = MLR;
    - デフォルトはMLR。ロバスト最尤法のこと。
    - 他に、ML（最尤法）、WLSMV（重みつき最小二乗法）、BAYES（ベイズ法）、GLS（最小二乗法）などがある。
      - Amosと結果を一致させたいなら、MLを使用するといい。

MODEL: コマンド

- モデルを指定する
  - モデルの書き方の基本は以下の5つ。
    - 回帰分析（構造方程式）は ON
    - 因子分析（測定方程式）は BY
    - 共分散は WITH
    - 変数名だけを書くと、分散を推定する指定
    - [変数名]と書くと、平均値を推定する指定
- パラメータの制約をする
  - MODEL CONSTRAINT:の後に書く
    - これについては、後述

OUTPUT: コマンド

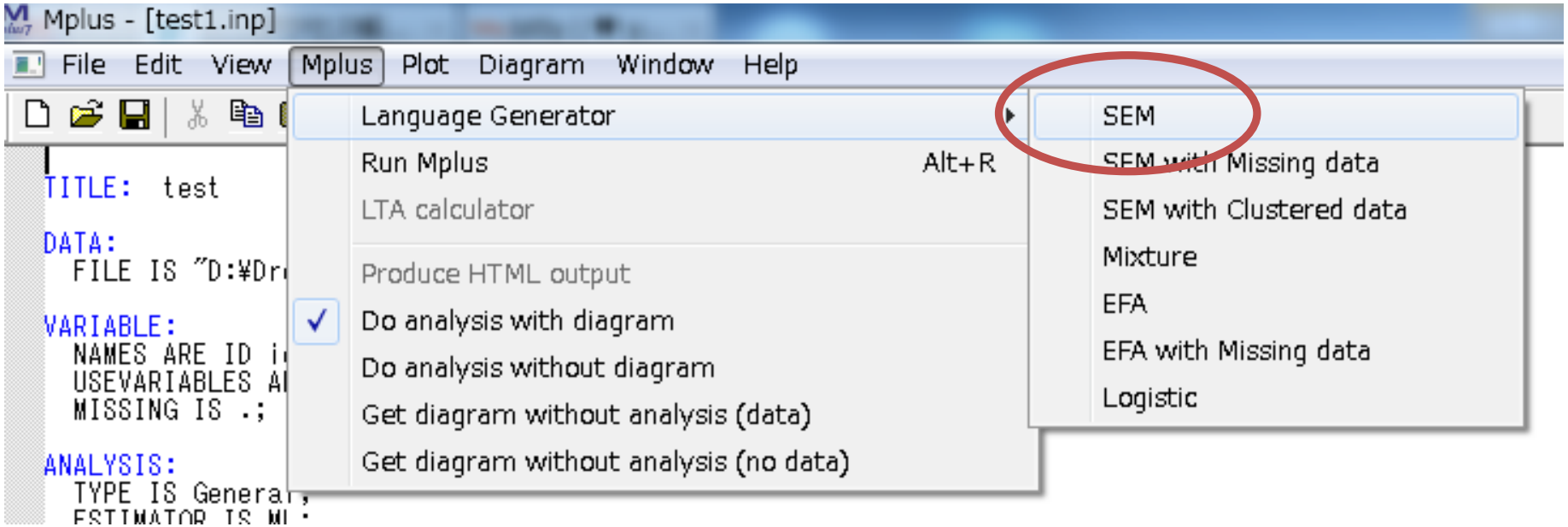
- 出力の設定をする
  - 記述統計量を出力する
    - SAMPSTAT
  - 標準化係数を出力する
    - STANDARIZED
    - しかし、STDYXと書いた方が余計なのが出なくて便利。
  - 信頼区間を出力する
    - CINTERVAL
    - ブートストラップ信頼区間を出力する場合は、CINTERVAL（BOOTSTRAP)と書く。

その他のコマンド

- SAVEDATA: コマンド
  - 因子得点などを保存する
  - SAVE ＝ FSOCRES;　で因子得点を保存
    - File = “###.txt”;　でファイル名を指定。日本語不可。
- PLOT:コマンド
  - グラフなどを出力する
    - IRTやベイズ推定の時に使う
- DEFINE:コマンド
  - 新しい変数を作ったり、変換したりするのに使う

Mplusの基本コードの書き方

- 最初は、Language Generatorで。



  - 基本的なモデリングはこれで解決する。

これだけは覚えておこう

ここにファイル名を書く。日本語不可。

```
DATA:
  FILE = "sample1.txt";
VARIABLE:
  NAMES = ID v1-v6 x1-x3 Y ;
  USEVARIABLES = x1 x2 Y;
  MISSING = .;
ANALYSIS:
  TYPE = GENERAL;
MODEL:
OUTPUT:
  SAMPSTAT   STDYX   MODINDICES(ALL);
```

ファイルに入っている変数名すべて書く。

分析に使う変数名を書く。

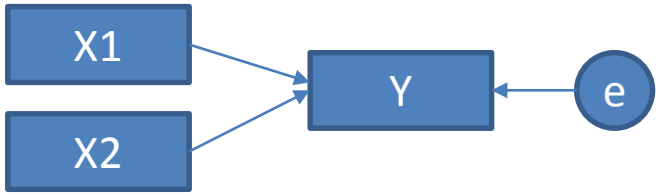
欠損値記号の指定

とりあえずこう書く!

モデルを記述する一番大事な箇所

MODEL:コマンドの書き方

- 重回帰分析をやってみよう･･･ONを使う。
  - 独立変数X1,X2でYを予測する




```
Y on X1 X2;
```

  - これで終わり！
  - Regressed on のON!

全体見てもこれだけ

```
DATA:
  FILE = "sample1.dat";
VARIABLE:
  NAMES = ID v1-v6 x1-x3 Y;
  USEVARIABLES = x1 x2 Y;
  MISSING = .;
ANALYSIS:
  TYPE = GENERAL;
MODEL:
   Y on X1 X2;
OUTPUT:
  SAMPSTAT   STDYX   MODINDICES(ALL);
```

入力ファイルの保存と実行

- 分析を実行する前に、inpファイルを保存
  - 場所は、さっきつくったフォルダの中に
  - データファイルがある場所が便利
    - フルパスを書かなくてよいから
  - 拡張子は.inp

- inpファイルが保存できたら、RUNボタンを押す
  - 最初は、何かしらエラーが出る。
  - こころが折れそうになることもあるが、がんばる。

出力の見方

- 結果を見るときの心得
  - できるだけ詳細をちゃんと見る
    - 分析がうまくいってない場合、走っても警告が出ることがある。
    - したかった分析がちゃんとできているか、チェック。
      - 手違いで分析法などが間違えているかもしれない
  - 適合度は必ず見る
    - 適合度がかなり悪いなら、データ入力のミスの可能性
    - 慣れないCUIでのモデリングは、ミスも多い

分析の要約

```
Mplus sample

SUMMARY OF ANALYSIS

Number of groups                                                 1
Number of observations                                         200

Number of dependent variables                                    1
Number of independent variables                                  2
Number of continuous latent variables                            0

Observed dependent variables

  Continuous
   Y

Observed independent variables
   X1          X2


Estimator                                                       ML
Information matrix                                        OBSERVED
Maximum number of iterations                                  1000
Convergence criterion                                    0.500D-04
Maximum number of steepest descent iterations                   20

Input data file(s)
  sample1.txt

Input data format  FREE
```

適合度指標

```
MODEL FIT INFORMATION

Number of Free Parameters                        4

Loglikelihood

          H0 Value                        -755.271
          H1 Value                        -755.271

Information Criteria

          Akaike (AIC)                    1518.542
          Bayesian (BIC)                  1531.735
          Sample-Size Adjusted BIC        1519.063
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                              0.000
          Degrees of Freedom                     0
          P-Value                           0.0000

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                           0.000
          90 Percent C.I.                    0.000  0.000
          Probability RMSEA <= .05           0.000

CFI/TLI

          CFI                                1.000
          TLI                                1.000

Chi-Square Test of Model Fit for the Baseline Model

          Value                            124.390
          Degrees of Freedom                     2
          P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                              0.000
```

- 情報量基準
  - モデル比較に使う
- X2乗値
  - モデルがデータと等しいという帰無仮説の確率
- RMSEA
  - 0.05（0.10）以下がよい
  - 倹約性を考慮した指標
- CFI
  - 0.95以上がよい
  - 倹約性をやや考慮した指標
- SRMR
  - 0.05（0.10）以下がよい
  - モデルとデータの距離

推定結果

```
MODEL RESULTS

                    Estimate       S.E.  Est./S.E.    Two-Tailed
                                                        P-Value

 Y        ON
    X1                 0.506      0.701      0.722      0.470
    X2                 0.346      0.027     13.033      0.000

 Intercepts
    Y                  2.120      4.019      0.528      0.598

 Residual Variances
    Y                111.593     11.160     10.000      0.000

STANDARDIZED MODEL RESULTS

STDYX Standardization

                    Estimate       S.E.  Est./S.E.    Two-Tailed
                                                        P-Value

 Y        ON
    X1                 0.038      0.053      0.722      0.470
    X2                 0.686      0.039     17.652      0.000

 Intercepts
    Y                  0.147      0.281      0.524      0.601

 Residual Variances
    Y                  0.537      0.052     10.391      0.000
```

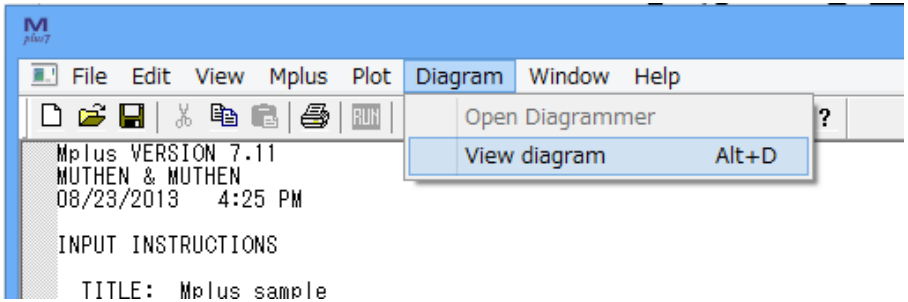
- Estimate：推定値
- S.E.：標準誤差
- EST./S.E.　z値

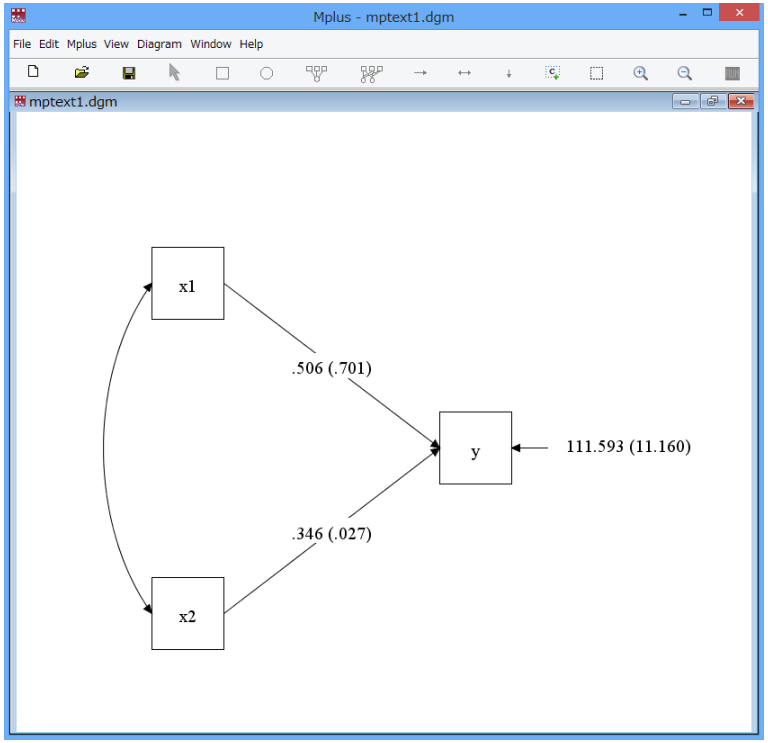
  1.96以上で5%有意
- STDYXオプションで

標準化係数も出力される

ダイアグラムの出力

- メニューバーのDiagramを選択
  - View diagramをクリックすると、パス図が見られる
  - ()内は標準誤差　出力設定で変えられる





分散・共分散の推定

- 分散と共分散を推定してみよう・・・WITHを使う。
  - さっきと同じモデルで
  - ただし、独立変数の分散も同時に推定する

```
Y on X1 X2;
X1 X2;
X1 with X2;
```

  - Mplusでは、デフォルトでは内生変数の分散しか推定しない。
  - 外生変数の分散を指定すると、それらの変数間の共分散も自動的に推定する。
    - すべての欠損値を推定するためには、独立変数も分散を推定する必要がある。ただ、データ規模が大きくなるため、適合度が悪くなることもある

結果は･･･

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 Y        ON
    X1                 0.506      0.701      0.722      0.470
    X2                 0.346      0.027     13.034      0.000

 X1       WITH
    X2                -5.519      2.227     -2.478      0.013

 Means
    X1                 4.920      0.077     64.226      0.000
    X2                53.075      2.024     26.223      0.000

 Intercepts
    Y                  2.120      4.019      0.528      0.598

 Variances
    X1                 1.174      0.117      9.999      0.000
    X2               819.326     81.933     10.000      0.000

 Residual Variances
    Y                111.592     11.159     10.000      0.000
```

共分散を推定したくない場合

- 固定母数の制約をする･･･@を使う。
  - 共分散を0に固定する

```
X1 with X2@0;
```

  - これでX1とX2の共分散が0に固定される。

```
Y with X2@0;
```

  - とすれば、パス係数を0に固定することもできる。

共分散の指定あれこれ

- 一度に全部の変数の共分散を指定

```
X1 X2 Y WITH X1 X2 Y;
```

  - これで3変数間の共分散を推定できる
- ペアごとの共分散を指定

```
v1 v2 v3 PWITH v4 v5 v6;
```

  - v1とv4、v2とv5、v3とv6の共分散を推定
- 共分散の自動推定をオフにする
  - ANALYSISに**MODEL = NOCOVARIANCES;** と書く。
    - これで、因子間や分散を推定した外生変数間の共分散が自動で推定されなくなる

平均値の指定

- 平均構造の推定
  - TYPE = GENERALにしておけば、自動的に平均構造も推定される
  - 自分で指定しておきたい場合は、[]を使い，特に平均値に固定制約を与えたい場合は、@で。

```
[X1];
[X2]@0;
```

  - 平均値を推定したくない場合は、ANALYSISに、**MODEL = NOMEANSTRUCTURE;** と書く。

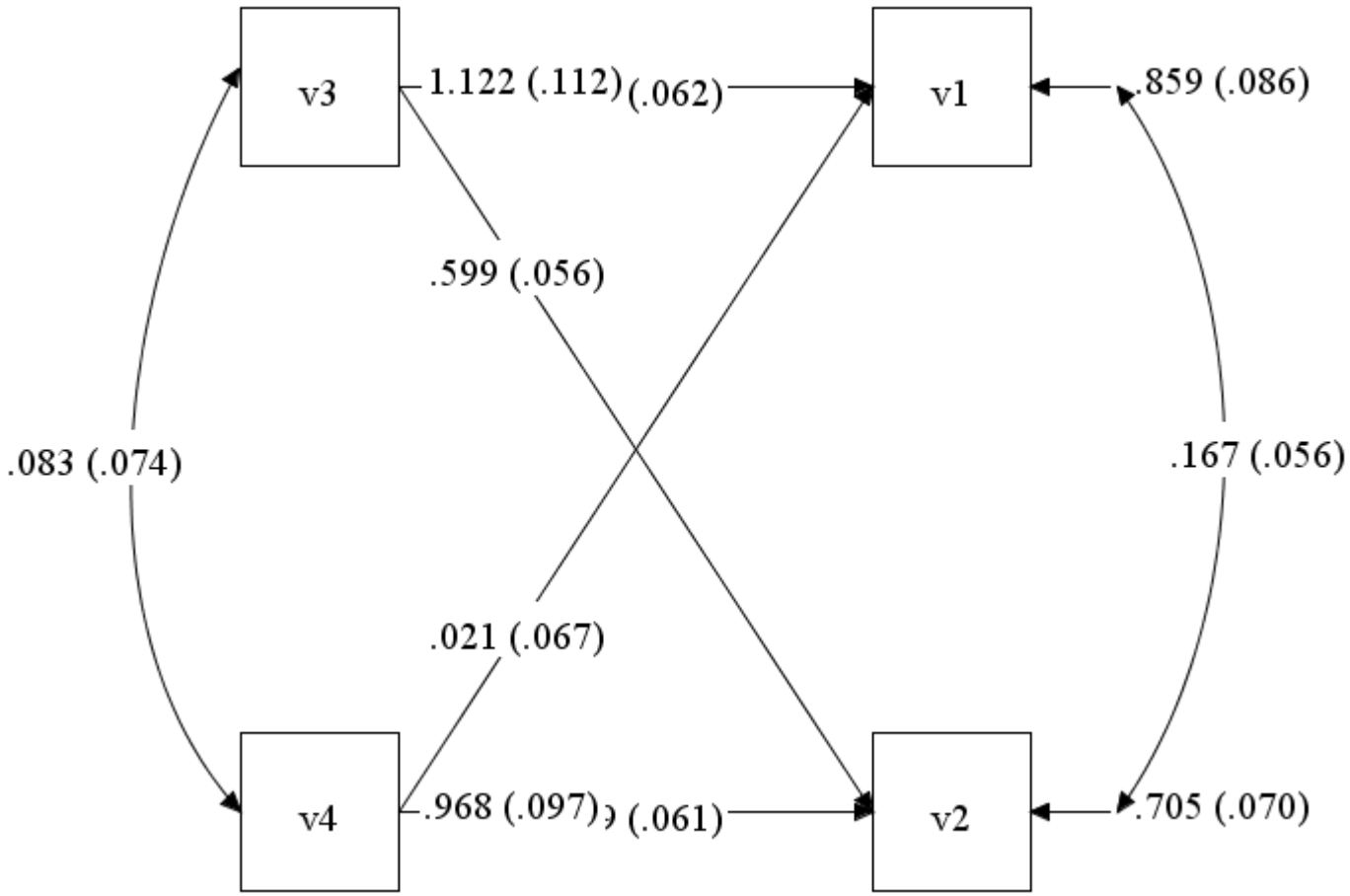
パス解析

- ONとWITHでパス解析が可能
  - v1 と v2 を v3 と v4 で予測するモデル
    - Usevariablesの変更をお忘れなく！

```
V1 V2 ON V3 V4;
V3 WITH V4;
```

  - 複数の目的変数に同じ説明変数が予測する場合は、一行で書ける。

パス解析のダイアグラム



確認的因子分析

- 因子分析をする･･･BYを使う。
  - F1とF2という潜在変数がv1-v6で測定されている

```
F1 BY v1-v3;
F2 BY v4-v6;
```

  - F1とF2の因子間相関は自動的に推定される
    - もちろん、F1 with F2@0; と書けば、直交回転となる。

確認的因子分析

- いくつかの注意点
  - 単にF1 by v1-v3;とかくと、v1への因子負荷量は、自動的に1に固定される
    - これは、モデルの識別のために必要な制約
  - ただ、分散を1に固定しても因子分析は可能
    - ただし、因子が外生変数の場合に限る

```
F1@1;
```

    - と書けば、F1の分散を1に固定することができる。
    - また、v1を自由推定するように指定することができる。

```
F1 by v1* v2-v3;
```

    - あるいは F1 by v1-v5*;　*は自由推定の指定。

コードも簡単でしょ

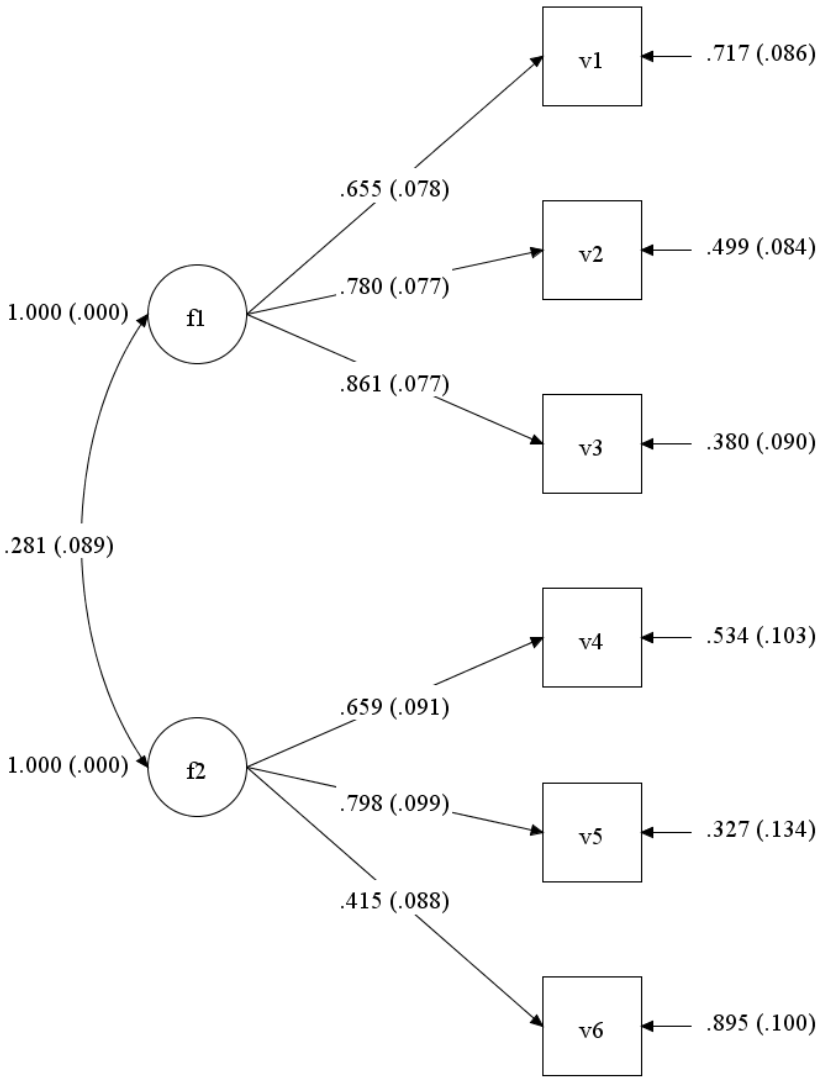
```
DATA:
  FILE = "sample1.txt";
VARIABLE:
  NAMES = ID v1-v6 x1-x3 Y;
  USEVARIABLES = v1-v6;
  MISSING = .;
ANALYSIS:
  TYPE = GENERAL;
MODEL:
   F1 by v1-v3*;
   F2 by v4-v6*;
   F1-F2@1;
OUTPUT:
  SAMPSTAT   STDYX   MODINDICES(ALL);
```

確認的因子分析の結果

STANDARDIZED MODEL RESULTS

STDYX Standardization

		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1	BY				
V1		0.612	0.056	10.908	0.000
V2		0.741	0.052	14.305	0.000
V3		0.813	0.051	16.083	0.000
F2	BY				
V4		0.669	0.078	8.549	0.000
V5		0.812	0.086	9.425	0.000
V6		0.402	0.078	5.138	0.000
F2	WITH				
F1		0.281	0.089	3.166	0.002
Intercepts					



- 因子分析の場合は標準化のほうが解釈しやすい

確認的因子分析の誤差相関

- 誤差相関とは
  - 測定項目の誤差項の間に共分散を仮定するモデルも書ける

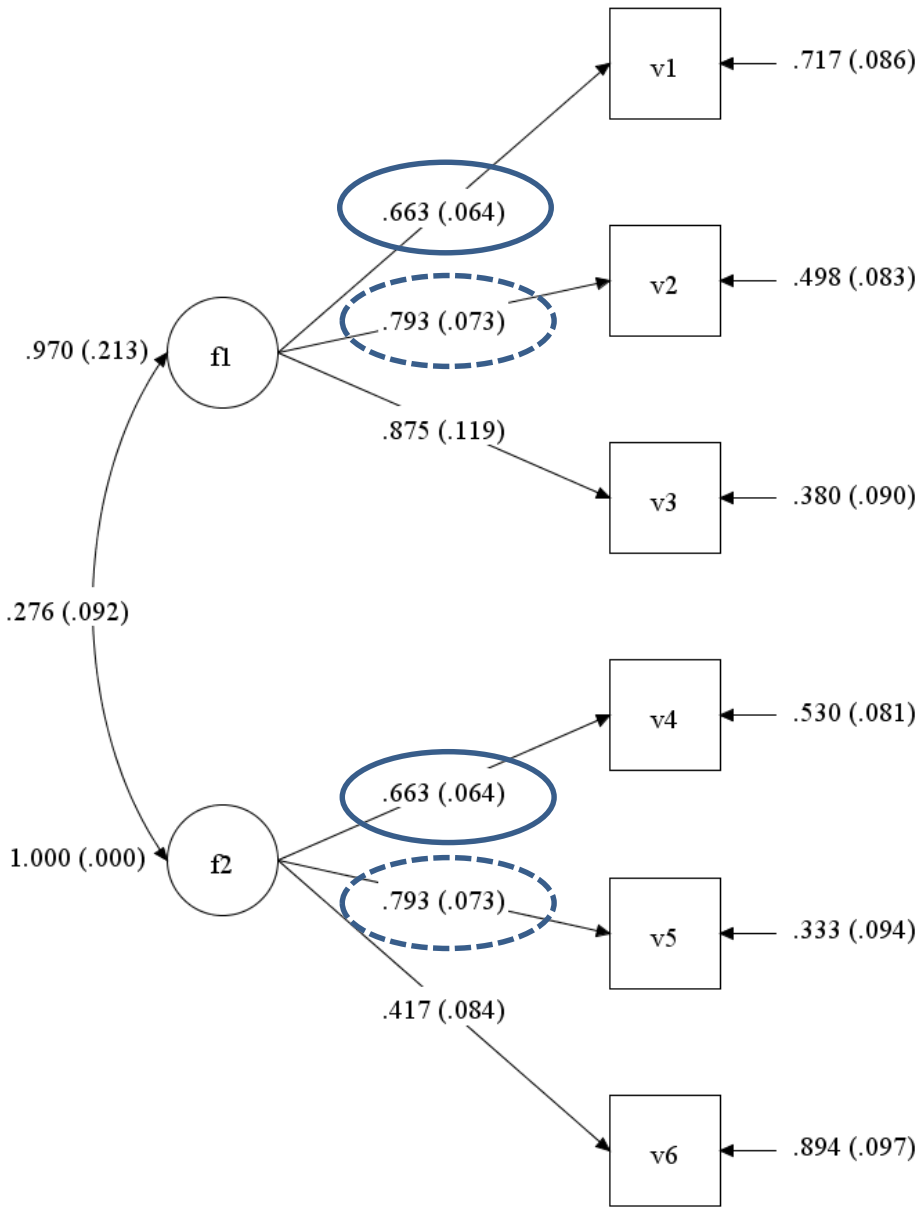
```
F1 by v1-v3;
v1 with v2;
```

- どういう時に誤差相関を仮定するか
  - 測定方法論上、測定誤差に相関が仮定される場合
    - 質問項目が似ている（同じ言葉が含まれている、など）
    - 同じことについて、違う側面から質問している
  - 適合度を上げるためだけにむやみに仮定するべきではない
    - しかし、相関を仮定することが妥当なら、仮定する方がよい。

因子負荷量の等値制約

- 各項目の因子負荷量が等しいという仮定
  - モデルの倹約性が向上する可能性がある
  - パラメータ名を設定し、等値に指定する

```
F1 by v1* (p1)
      v2  (p2)
      v3;
F2 by v4* (p1)
      v5  (p2)
      v6;
F1 F2 @1;
```




```
MODEL FIT INFORMATION

Number of Free Parameters                       18

Loglikelihood

          H0 Value                       -1606.555
          H1 Value                       -1602.844

Information Criteria

          Akaike (AIC)                    3249.109
          Bayesian (BIC)                  3308.479
          Sample-Size Adjusted BIC        3251.453
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

          Value                              7.421
          Degrees of Freedom                     9
          P-Value                           0.5934

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                           0.000
          90 Percent C.I.                    0.000  0.069
          Probability RMSEA <= .05           0.855

CFI/TLI

          CFI                                1.000
          TLI                                1.010

Chi-Square Test of Model Fit for the Baseline Model

          Value                            269.603
          Degrees of Freedom                    15
          P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)

          Value                              0.031
```

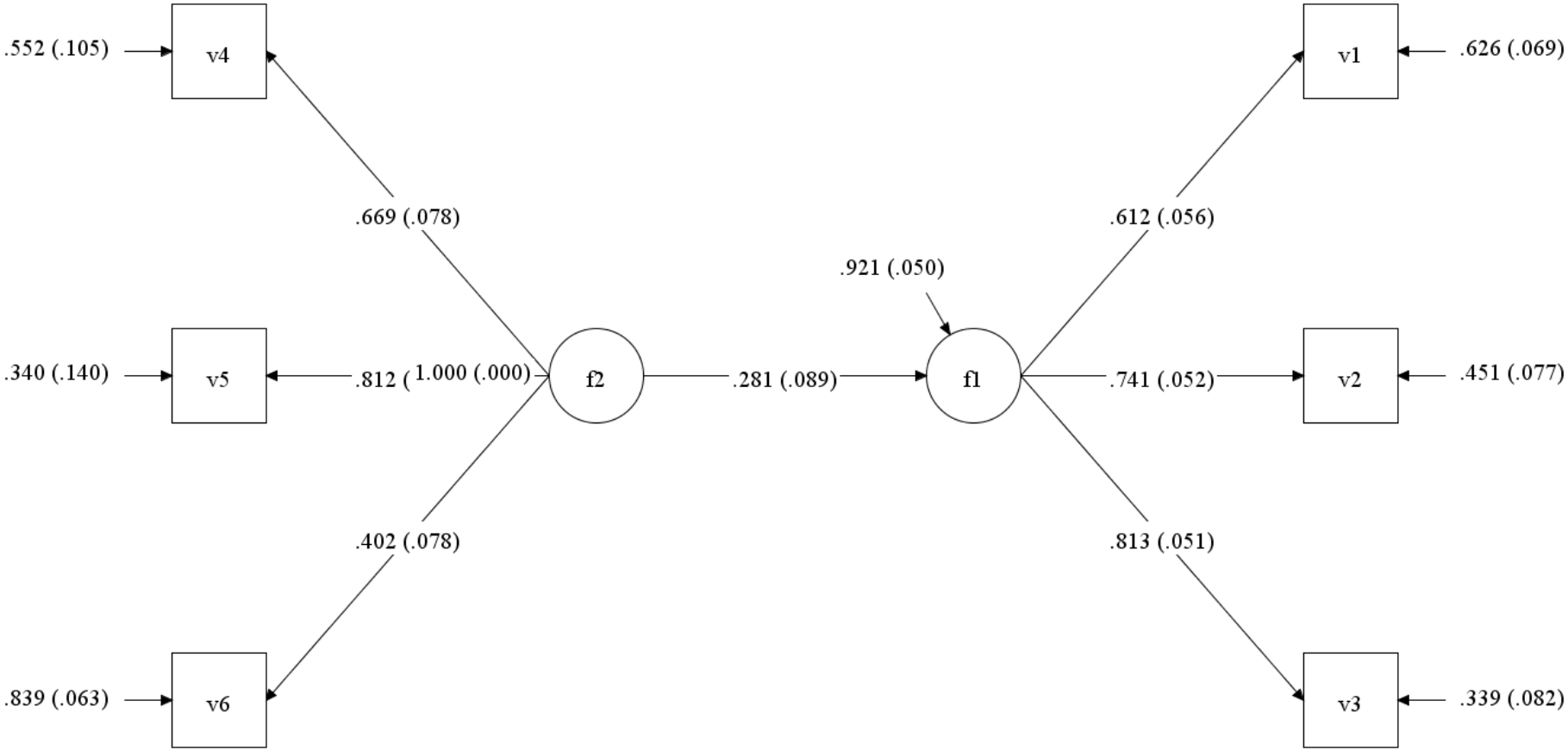
潜在変数を含むパス解析

- 因子間に因果関係を仮定する
  - F1がF2を回帰するような場合。

```
F1 by v1-v3;
F2 b7 v4-v6;
F1 on F2;
```

  - 今回は、因子負荷量を1に固定して推定

分析結果（標準化係数）


.552 (.105)
v4
.669 (.078)
.340 (.140)
v5
.812 (
1.000 (.000)
f2
.402 (.078)
.839 (.063)
v6
.281 (.089)
.921 (.050)
f1
.612 (.056)
v1
.626 (.069)
.741 (.052)
v2
.451 (.077)
.813 (.051)
v3
.339 (.082)

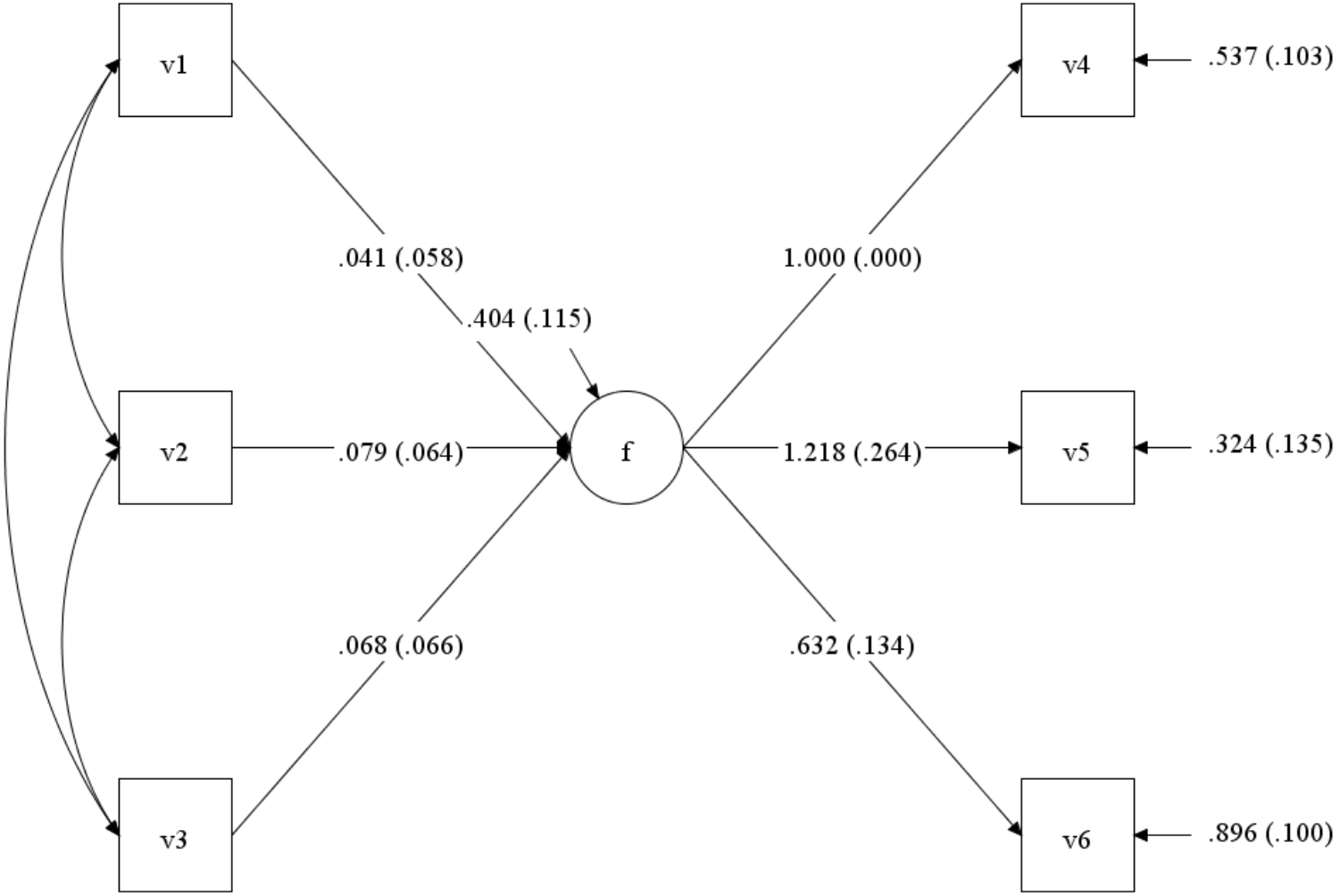
MIMICモデル

- Multiple indicator Multiple Cause Model
  - 複数の説明変数が、潜在変数を経て複数の目的変数に影響するモデル。

```
F by v4-v6;
F on v1-v3;
```

  - 特に難しい制約は必要ない。

MIMICモデル　ダイアグラム



導入編まとめ

- 英語&CUIがハードルだけど，すごく簡単にすごくたくさんのことができます！
- データもコードも出力も，みんなプレーンテキスト！ 日本語のパスはダメ，絶対。
- モデルの指定はonとbyでだいたいOK
  - **回帰分析（構造方程式）は ON**
  - **因子分析（測定方程式）は BY**
  - **共分散は WITH**
- これであなたも今日からM*plus*ユーザー！

Mplus活用編

活用編のメニュー

- カテゴリカル・非正規データの分析
  - 2値データ、順序データ、様々な分布データ
  - 項目反応理論(Item Response Theory)
- マルチレベルSEM(Multilevel Modeling)
  - 階層構造にあるデータを分析
- 潜在混合分布モデル(Mixture Modeling)
  - モデルからサンプルをグループ分けする分析

多変量正規性の仮定と心理尺度

- 一般に、構造方程式モデルは最尤法を使う
  - よって、多変量正規性の仮定が必要
  - しかし、心理尺度の多くは正規分布にならない
    - 臨床尺度は、「極端な人」を識別することを目的とする
    - 回答の負担を減らすため、２～4件法もよく使われる
    - 天井・床効果もよく起こる
- 連続・正規性から、順序・非正規性へ
  - 無理に連続・正規性にこだわる必要はない
  - データにあったモデリングを柔軟に選択できる

カテゴリカルデータの分析

- ここからが、M*plus*の真骨頂
  - 変数の尺度や分布に様々な仮定を当てはめられる
- 尺度の水準
  - 連続尺度・・・今までのモデル
  - 順序尺度・・・順序性があるカテゴリカルデータ
  - 名義尺度・・・順序性がないカテゴリカルデータ
- 目的変数の分布（後述）
  - 正規分布・・・今までのモデル
  - ポワソン分布・・・イベント生起度数の分布

順序回帰分析

- 従属変数がカテゴリカル変数の回帰分析
  - v1が1項目の尺度で測定されている
    - 連続変量とは言い難い。

- 順序カテゴリカルの指定
  - VARIABLES:コマンドにCATEGORICALオプション

CATEGORICAL = v1;

    - ※順序、名義尺度は値がすべて整数である必要

  - ANALYSIS:コマンドで推定法を選択
    - WLS:重みつき最小二乗法・・・順序プロビット回帰分析
    - ML:最尤法・・・順序ロジスティック回帰分析

以下のコードを書く（ロジスティック回帰）

```
VARIABLE:
  NAMES = ID v1-v3;
  USEVARIABLES = v1-v3;
  MISSING = .;
  CATEGORICAL = v1;

ANALYSIS:
  TYPE = GENERAL;
  ESTIMATOR = MLR;

MODEL:
 v1 ON v2-v3;
```

推定結果

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 V1         ON
    V2                 0.624      0.128      4.885      0.000
    V3                 0.342      0.075      4.542      0.000

 Thresholds
    V1$1              -0.093      0.522     -0.177      0.859
    V1$2               1.429      0.458      3.118      0.002
    V1$3               3.651      0.500      7.300      0.000
    V1$4               5.465      0.559      9.768      0.000


LOGISTIC REGRESSION ODDS RATIO RESULTS

 V1         ON
    V2                 1.866
    V3                 1.408
```

プロビット回帰分析

```
VARIABLE:
  NAMES = ID v1-v3;
  USEVARIABLES = v1-v3;
  MISSING = .;
  CATEGORICAL = v1;

ANALYSIS:
  TYPE = GENERAL;
  ESTIMATOR = WLSMV;

MODEL:
 v1 ON v2-v3;
```

推定結果

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
V1 ON				
V2	0.345	0.057	6.052	0.000
V3	0.195	0.032	6.092	0.000
Thresholds				
V1$1	-0.012	0.255	-0.046	0.963
V1$2	0.756	0.241	3.130	0.002
V1$3	2.058	0.249	8.281	0.000
V1$4	3.111	0.266	11.714	0.000

順序因子分析

- 順序カテゴリカルデータで因子分析
  - 5件法で測定された尺度で因子分析
    - 基本的には5件であれば連続変量とみなしても構わないが、正規性が満たされない場合は順序尺度としてみなすのも一つの方法

- 指定方法は同じ

  CATEGORICAL = v1-v6;

  - 因子分析ではWLS（あるいはWLSMV）がオススメ
    - 項目数と水準数が多い場合は、MLではメモリが足りなくなる場合もある
    - また、順序のMLは適合度指標が限定される

以下のコードを書く

```
VARIABLE:
  NAMES = v1-v6 sex;
  USEVARIABLES = v1-v6;
  MISSING = .;
  CATEGORICAL = v1-v6;

ANALYSIS:
  TYPE = GENERAL;
  ESTIMATOR = WLSMV;

MODEL:
  F1 by v1-v3*;
  F2 by v4-v6*;
  F1-F2@1;  [F1-F2@0];
```

推定結果

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1 BY				
V1	0.839	0.041	20.596	0.000
V2	0.740	0.047	15.794	0.000
V3	0.755	0.042	17.781	0.000
F2 BY				
V4	0.835	0.034	24.598	0.000
V5	0.830	0.031	26.966	0.000
V6	0.897	0.032	27.792	0.000
F2 WITH				
F1	0.730	0.054	13.533	0.000
Means				
F1	0.000	0.000	999.000	999.000
F2	0.000	0.000	999.000	999.000

名義回帰分析

- 名義カテゴリカルデータの分析
  - 順序性がない回帰分析
  - 男性と女性（順序性はない）
    - sexを名義尺度として分析
    - ただし、2値の場合は順序も名義も結果は同じ

- 名義カテゴリカルの指定
  - VARIABLES:コマンドにNOMINALオプション

NOMINAL = sex;

  - 名義回帰は、ML（あるいはMLR）でしか推定できない

以下のコードを書く

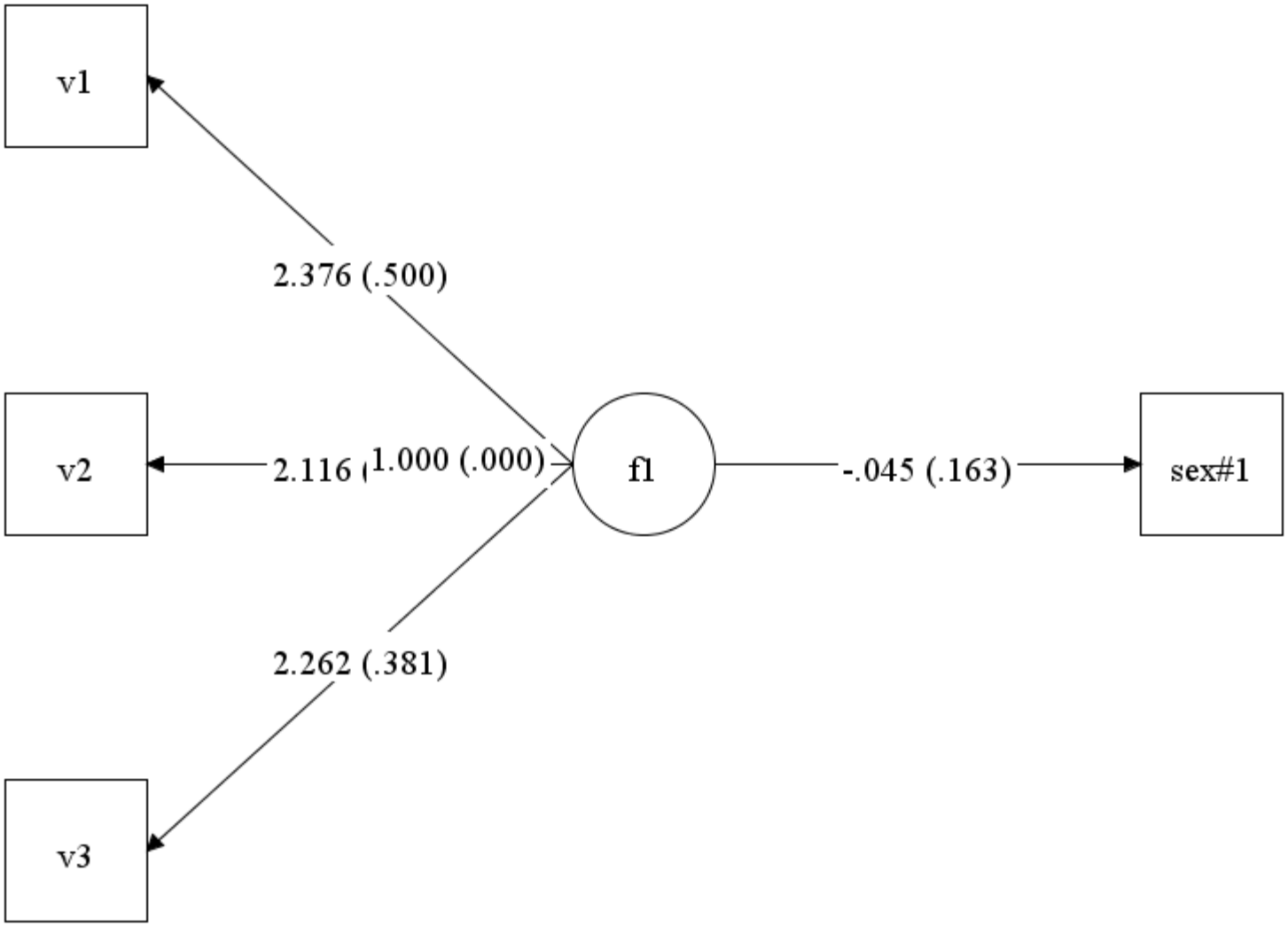
```
VARIABLE:
  NAMES = v1-v6 sex;
  USEVARIABLES = v1-v3 sex;
  MISSING = .;
  CATEGORICAL = v1-v3;
  NOMINAL = sex;

ANALYSIS:
  TYPE = GENERAL;
  ESTIMATOR = MLR;

MODEL:
  F1 by v1-v3*;
  F1@1;
  [F1@0];

  sex on F1;
```

推定結果

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1 BY				
V1	2.376	0.500	4.755	0.000
V2	2.116	0.391	5.408	0.000
V3	2.262	0.381	5.938	0.000
SEX#1 ON				
F1	-0.045	0.163	-0.274	0.784
Means				
F1	0.000	0.000	999.000	999.000
Intercepts				
SEX#1	0.000	0.141	0.000	1.000

ダイアグラム



カウントデータの分析

- カウントデータ
  - 一定時間内の特定のイベントの生起度数
    - ポワソン分布に従うことが知られている
      - あるいは負の二項分布

- M*plus*でのカウントデータの分析
  - USEVARIABLESコマンドに

COUNT = v1(p);

  - と書く。(p)はポワソン分布のこと
  - 負の二項分布は(nb)と書く。

ポワソンを仮定した分析結果

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 V1         ON
    V2                 0.083      0.017      4.757      0.000
    V3                 0.047      0.010      4.635      0.000

 Intercepts
    V1                 0.755      0.068     11.041      0.000
```

ポアソン分布か正規分布か

ポアソン分布

```
Information Criteria

          Akaike (AIC)                          992.198
          Bayesian (BIC)                       1003.269
          Sample-Size Adjusted BIC              993.755
            (n* = (n + 2) / 24)
```

正規分布（通常の回帰分析）

```
Information Criteria

          Akaike (AIC)                          784.900
          Bayesian (BIC)                        799.661
          Sample-Size Adjusted BIC              786.976
            (n* = (n + 2) / 24)
```

- 正規分布のほうが適合している

項目反応理論

項目反応理論（段階反応モデル）

- M*plus*でも項目反応理論が可能
  - 他のソフトウェアと一致させるためには要工夫
    - 1因子モデル
    - 各項目をCATEGORICALで指定
    - 推定法はML（あるいはMLR）で指定
    - 因子の分散を1、平均を0に固定
    - 推定値を1.702で割る
  - PLOTコマンドで、情報量関数や項目特性を出力
    - TYPE = PLOT3;

以下のコードを書く

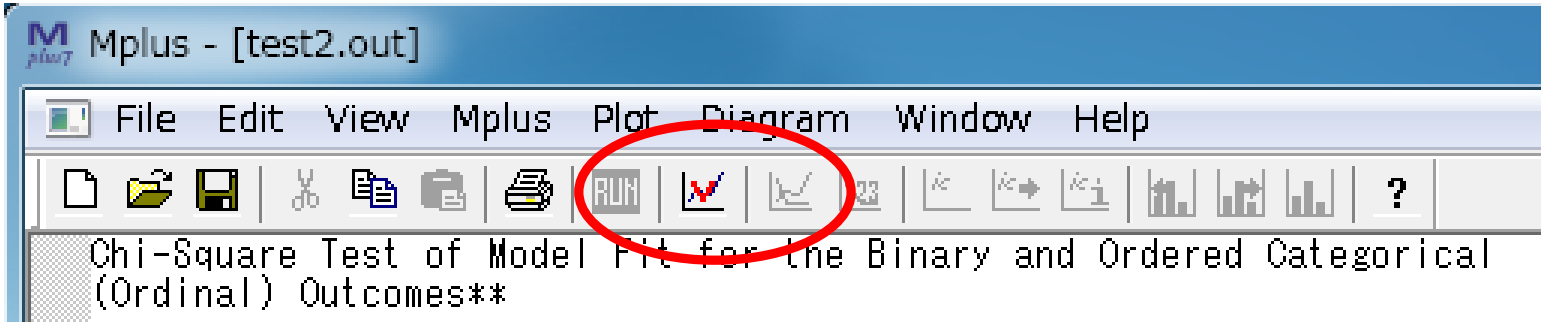
```
VARIABLE:
  NAMES = v1-v6 sex;
  USEVARIABLES = v1-v6;
  MISSING = .;
  CATEGORICAL = v1-v6;

ANALYSIS:
  ESTIMATOR = ML;

MODEL:
  F1 by v1-v6*(p1-p6);
  F1@1;    [F1@0];
 MODEL Constraint:
  NEW(a1-a6);
  DO(1,6)p# = 1.702*a#;

PLOT:
  TYPE = PLOT3;
```

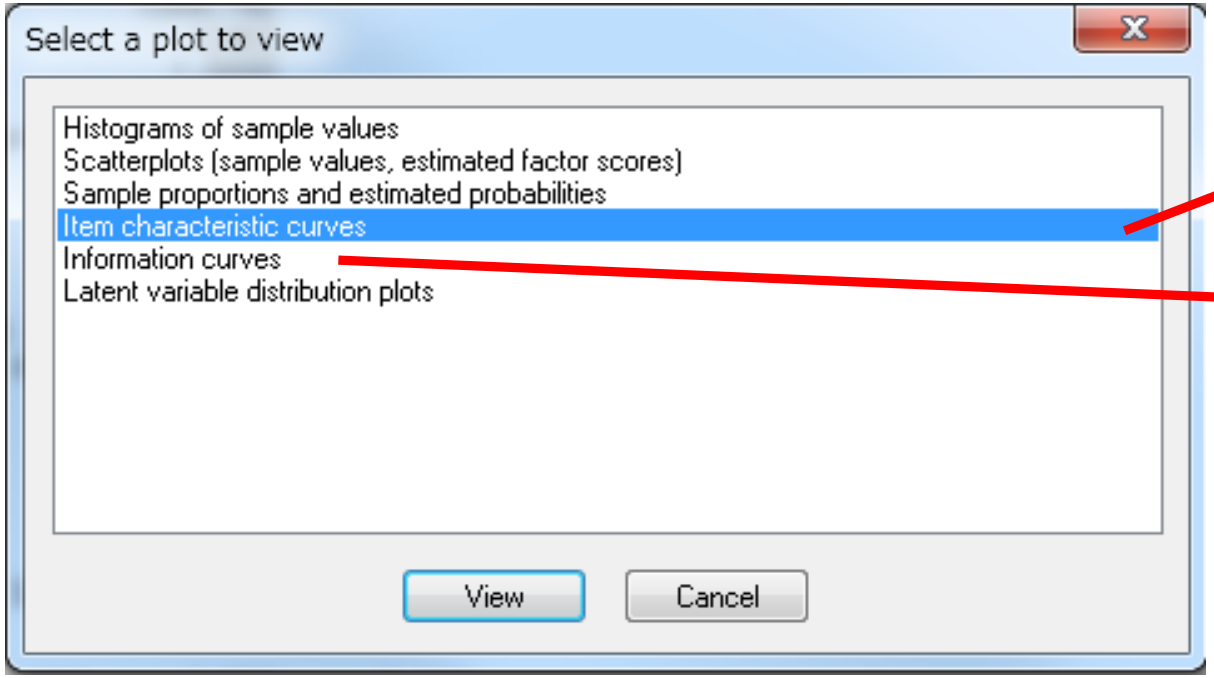
推定結果

- 項目の係数ではなく、a1-a6を確認
  - これはIRTの識別力を表している。

```
New/Additional Parameters
    A1                    1.012      0.154      6.573      0.000
    A2                    0.837      0.130      6.450      0.000
    A3                    0.825      0.128      6.454      0.000
    A4                    1.427      0.201      7.115      0.000
    A5                    1.419      0.199      7.131      0.000
    A6                    1.905      0.304      6.259      0.000
```

項目特性曲線(ICC)の見方

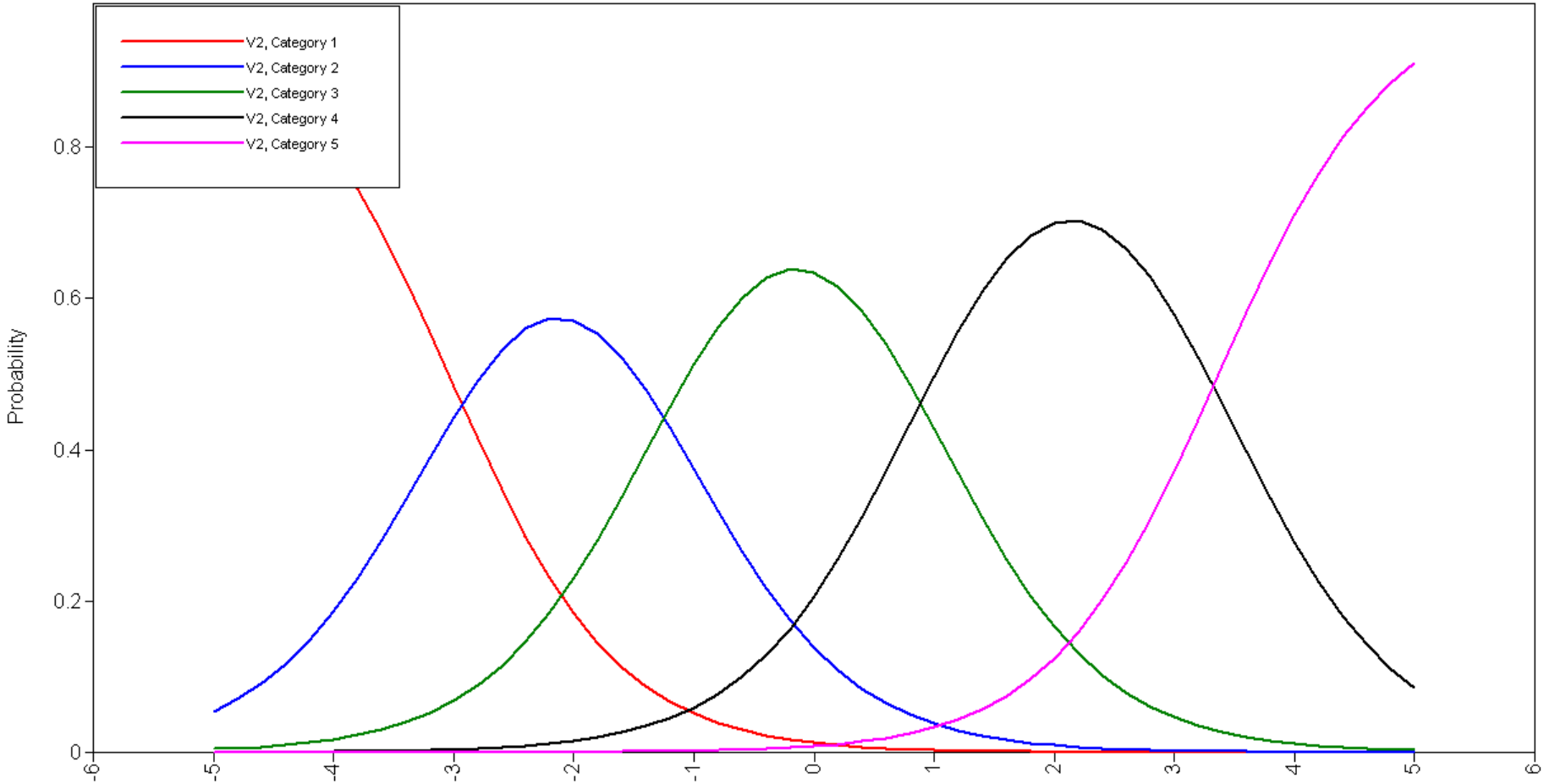




項目特性曲線

テスト情報関数

項目特性曲線


V2, Category 1
V2, Category 2
V2, Category 3
V2, Category 4
V2, Category 5
Probability
0
0.2
0.4
0.6
0.8
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6

マルチレベルモデル

マルチレベルモデル

- データが階層性を持った場合の回帰モデル
  - 学校をサンプリングした後，生徒をサンプリング
  - 学校レベルと生徒レベルの2段階の階層を持つ
- 各レベルでモデリングを行う
  - 学校単位のモデルと，生徒単位のモデルをそれぞれ推定
  - 集団レベルをBetween，個人レベルをWithinと呼ぶ
- マルチレベルアドオンが必要
  - ベースに追加料金を払えば，手に入る

マルチレベルモデルのコード

- 集団を識別する変数を指定
  - CLUSTER = group;　など，変数を指定
- マルチレベルアドオンの使用を指定
  - TYPE = TWOLEVEL;
  - GENERALでは実行できない
- 各レベルのモデル
  - %Between%と書いて，Betweenモデルを記述
  - %Within%と書いて，Withinモデルを記述

データセット

	A	B	C	D	E
1	ID	v1	v2	v3	
2	1	3	3	3	
3	1	3	2	3	
4	1	3	3	3	
5	2	3	3	3	
6	2	2	1	3	
7	2	2	1	3	
8	3	1	3	2	
9	3	3	3	2	
10	3	3	2	2	
11	4	1	4	1	
12	4	1	4	1	
13	4	1	4	1	
14	5	2	2	5	

以下のコードを書く

```
VARIABLE:
  NAMES = ID v1-v3;
  USEVARIABLES = v1-v3;
  MISSING = .;
  CLUSTER = ID;
  BETWEEN = v3;

ANALYSIS:
  TYPE = TWOLEVEL;
  ESTIMATOR = MLR;

MODEL:
 %WITHIN%
    v1 ON v2;
 %BETWEEN%
    v1 ON v2 v3;
```

推定結果

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Within Level

 V1         ON
    V2                 0.221      0.082      2.689      0.007

 Residual Variances
    V1                 0.603      0.088      6.891      0.000

Between Level

 V1         ON
    V2                 0.435      0.198      2.196      0.028
    V3                 0.158      0.048      3.327      0.001
```

階層線形モデル

```
VARIABLE:
  NAMES = ID v1-v3;
  USEVARIABLES = v1-v3 v2m;
  MISSING = .;
  CLUSTER = ID;
  WITHIN = v2;
  BETWEEN = v2m v3;

DEFINE:
  v2m = CLUSTER_MEAN(v2);
  CENTER v2 (Groupmean);
  CENTER v2m (Grandmean);
```

集団平均と全体平均で中心化

```
ANALYSIS:
  TYPE = TWOLEVEL;
  ESTIMATOR = MLR;

MODEL:
  %BETWEEN%
    v1 ON v2m v3;
 %WITHIN%
    v1 ON v2;
```

集団平均値を変数として利用

推定結果

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Within Level

 V1         ON
    V2                 0.218      0.082      2.657      0.008

 Residual Variances
    V1                 0.603      0.087      6.890      0.000

Between Level

 V1         ON
    V2M                0.347      0.113      3.072      0.002
    V3                 0.162      0.046      3.517      0.000
```

潜在混合分布モデル

- データから潜在的なカテゴリを見つける
  - カテゴリカルデータ：潜在クラスモデル
  - 連続データ：潜在プロフィルモデル

- 混合分布モデルアドオンを使用
  - TYPE = MIXTURE;
- 潜在カテゴリ数を指定
  - CLASS = c(2); ・・・cという変数で2カテゴリを仮定
  - %OVERALL%で全カテゴリについてのモデル
  - %c#1%で，カテゴリ1に限定したモデルを指定

以下のコードを書く

```
VARIABLE:
  NAMES ARE v1-v6 sex;
  USEVARIABLES ARE v1-v6 sex;
  MISSING IS .;
  CATEGORICAL = v1-v6;
  CLASS = c(2);

ANALYSIS:
  TYPE IS MIXTURE;
  ESTIMATOR IS MLR;

MODEL:
%OVERALL%
 c#1 on sex;
```

潜在カテゴリを性別で予測

推定結果

```
CLASSIFICATION OF INDIVIDUALS BASED ON THEIR MOST LIKELY LATENT CLASS MEMBERSHIP

Class Counts and Proportions

    Latent
   Classes

       1                85          0.42500
       2               115          0.57500


Average Latent Class Probabilities for Most Likely Latent Class Membership (Row)
by Latent Class (Column)

            1        2

    1   0.929    0.071
    2   0.037    0.963
```

推定結果

```
Categorical Latent Variables

 C#1          ON
    SEX                -0.288       0.330      -0.873       0.383

 Intercepts
    C#1                 0.092       0.602       0.152       0.879
```

その他Mplusの特徴

Mplusで選択できる推定法

- 最尤法
  - ML　MLRなど
- 重みつき最小二乗法
  - WLS　WLSMVなど
- ベイズ推定法
  - Bayes

標準誤差の推定法

- 普通の推定方法
  - MLやWLS
  - モデルの仮定に基づいて標準誤差を推定
    - 多変量正規性の仮定

- ロバスト標準誤差の推定
  - MLRやWLSMV
  - データの分布に合わせて標準誤差を補正
    - 多少多変量正規性から逸脱しても、妥当な結果
      - 推定値そのものは同じ
    - M*plus*のデフォルトはこちらを使う。オススメはこちら。
      - ただ、ロバスト標準誤差の推定では、χ2乗検定がやや複雑になる。

欠損値の推定法

- 完全情報最尤法(Full information ML)
  - 特に指定はいらない。基本はこれを使ってくれる。
  - サブジェクト全体を消すのではなく、欠損していない部分をすべて活用して推定

- 多重代入法
  - ベイズ推定を用いて，欠損値を推定する

様々なモデルの組み合わせ

- M*plus*の最大の特徴は，組合せの多様さ
  - 項目反応理論　＋　多母集団分析
    - 群ごとに項目特性が異なるか，などを検討可能
  - カテゴリカル（ポアソン）　＋　マルチレベル
    - カテゴリカルデータのマルチレベルモデル
      - 一般化線形混合モデルになる
  - マルチレベル　＋　潜在混合分布モデル
    - 集団を潜在カテゴリに分ける

Rでやってみる

Rで本気出してみる

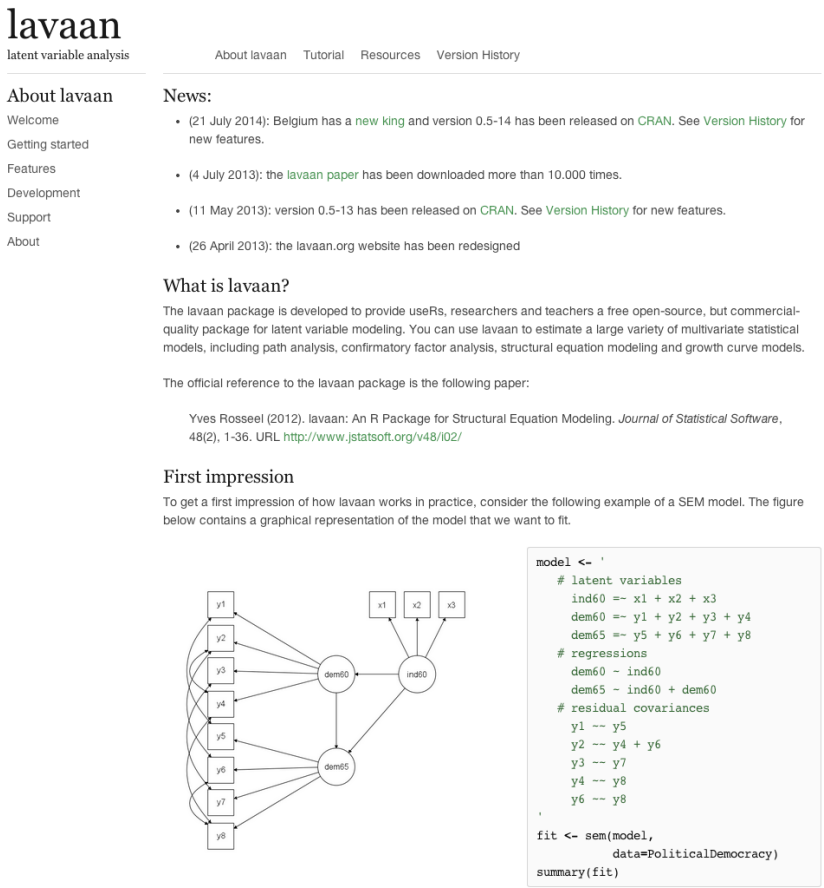
Rの導入について

- Rの導入については本TSでは射程外。
  - Tsukuba.R, Tokyo.R, Nagoya.R, Sapporo.R, Osaka.R, Yamadai.Rなど全国のRユーザーによるセミナーがあります。
  - 中国四国心理学会第69回大会Rチュートリアルセミナー（於：山口大学）もね。
  - インターネット上の解説サイトも多数。
    - 動画による解説，Rオンラインガイドもね。
    - http://psycho.edu.yamaguchi-u.ac.jp/?page_id=242

Rの特徴としてのパッケージ

- R本体だけで基本的な統計処理は全部できますが，パッケージを追加することで様々な分析モデルが手に入ることがRの真骨頂
  - 4000を超えるパッケージ群！
  - 心理学関係のパッケージも多数！
- SEMをするためのパッケージも
  - semパッケージ，OpenMXパッケージなど
  - なかでもlavaanパッケージはM*plus*クローンか

lavaan pacakage 0.5-14



lavaan
latent variable analysis

About lavaan Tutorial Resources Version History

About lavaan
- Welcome
- Getting started
- Features
- Development
- Support
- About

News:
- (21 July 2014): Belgium has a new king and version 0.5-14 has been released on CRAN. See Version History for new features.
- (4 July 2013): the lavaan paper has been downloaded more than 10.000 times.
- (11 May 2013): version 0.5-13 has been released on CRAN. See Version History for new features.
- (26 April 2013): the lavaan.org website has been redesigned

What is lavaan?

The lavaan package is developed to provide useRs, researchers and teachers a free open-source, but commercial-quality package for latent variable modeling. You can use lavaan to estimate a large variety of multivariate statistical models, including path analysis, confirmatory factor analysis, structural equation modeling and growth curve models.

The official reference to the lavaan package is the following paper:

Yves Rosseel (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36. URL http://www.jstatsoft.org/v48/i02/

First impression

To get a first impression of how lavaan works in practice, consider the following example of a SEM model. The figure below contains a graphical representation of the model that we want to fit.


```
model <- '
  # latent variables
    ind60 =~ x1 + x2 + x3
    dem60 =~ y1 + y2 + y3 + y4
    dem65 =~ y5 + y6 + y7 + y8
  # regressions
    dem60 ~ ind60
    dem65 ~ ind60 + dem60
  # residual covariances
    y1 ~~ y5
    y2 ~~ y4 + y6
    y3 ~~ y7
    y4 ~~ y8
    y6 ~~ y8
'
fit <- sem(model,
           data=PoliticalDemocracy)
summary(fit)
```

- LAtentVAriableANalysis の略で，RにおけるSEM用パッケージとしては後発。
- まだベータ版でバグがあるかも，だけどオープンにやってます！というところがウリ。

lavaanの書式

- モデルの指定の基本
  - 潜在変数=~観測変数1+2+3+...で測定方程式
  - 変数X~変数Y で構造方程式
  - 変数X~~変数Yで相関関係（共分散）
  - 他にも~1で切片，f1 ~ 0*f2で直交，など。
- 結果はsummary関数で出す
  - オプションとして
    - 適合度指標　fit.measure=TRUE
    - 標準化係数　standardrized=TRUE

サンプルコード・確証的因子分析

```
library(lavaan)
model <- '
F1 =~ V1+V2+V3
F2 =~ V4+V5+V6
F1~~F2
'
fit <- sem(model,data=sample1)
summary(fit,standardraized=TRUE,fit.measure=TRUE)
```

パッケージの読み込み

モデル式の記述

Fitting

出力

サンプル結果・確証的因子分析

```
lavaan (0.5-14) converged normally after  26 iterations

  Number of observations                           200

  Estimator                                         ML
  Minimum Function Test Statistic                7.417
  Degrees of freedom                                 8
  P-value (Chi-square)                           0.492

Model test baseline model:

  Minimum Function Test Statistic              269.603
  Degrees of freedom                                15
  P-value                                        0.000

Full model versus baseline model:

  Comparative Fit Index (CFI)                    1.000
  Tucker-Lewis Index (TLI)                       1.004
```

サンプル結果・確証的因子分析

```
Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)                  -1606.553
  Loglikelihood unrestricted model (H1)          -1602.844

  Number of free parameters                             13
  Akaike (AIC)                                    3239.105
  Bayesian (BIC)                                  3281.983
  Sample-size adjusted Bayesian (BIC)             3240.798

Root Mean Square Error of Approximation:

  RMSEA                                              0.000
  90 Percent Confidence Interval                0.000  0.079
  P-value RMSEA <= 0.05                              0.781

Standardized Root Mean Square Residual:

  SRMR                                               0.035
```

サンプル結果・確証的因子分析

	Estimate	Std.err	Z-value	P(>|z|)	Std.lv	Std.all
Latent variables:						
F1 =~						
V1	1.000				0.655	0.612
V2	1.192	0.163	7.333	0.000	0.780	0.741
V3	1.316	0.183	7.208	0.000	0.861	0.813
F2 =~						
V4	1.000				0.659	0.669
V5	1.211	0.248	4.881	0.000	0.798	0.812
V6	0.630	0.137	4.589	0.000	0.415	0.402
Covariances:						
F1 ~~						
F2	0.121					
Variances:						
V1	0.717					
V2	0.499					
V3	0.380					
V4	0.534					
V5	0.327					
V6	0.896					
F1	0.428					
F2	0.434					

c.f./M*plus* Output

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
F1 BY				
V1	0.612	0.056	10.908	0.000
V2	0.741	0.052	14.305	0.000
V3	0.813	0.051	16.083	0.000
F2 BY				
V4	0.669	0.078	8.549	0.000
V5	0.812	0.086	9.425	0.000
V6	0.402	0.078	5.138	0.000
F2 WITH				
F1	0.281	0.089	3.166	0.002
Intercepts				

オプションなど

- 推定方法
  - estimatorオプションで指定
    - 一般的な推定法；ML,GLS,WLS
    - ロバストな推定；MLM,MLMVS,MLMV,MLR
- 欠損値の扱い
  - missingオプションで
    - liswiseを指定すると欠損のあるデータを除外
    - ml,fimlを指定し，推定がML系なら完全情報最尤推定をしてくれる

オプションなど

- 平均構造
  - Fitting関数のオプションでmeanstructure=TRUEとするだけ

```
fit <- cfa(HS.model,
           data = HolzingerSwineford1939,
           meanstructure = TRUE)
summary(fit)
```

- 多母集団同時分析
  - 群分け変数をgroup変数で指定
  - 等値制約をかけるときは，制約する係数をgroup.eauql関数で“loadings”, “intercepts”など指定

```
HS.model <- '  visual =~ x1 + x2 + x3
              textual =~ x4 + x5 + x6
              speed   =~ x7 + x8 + x9 '
fit <- cfa(HS.model,
           data = HolzingerSwineford1939,
           group = "school")
summary(fit)
```

オプションなど

- データをcov行列で渡すことも可能

```
lower <- '
 11.834
  6.947   9.364
  6.819   5.091  12.532
  4.783   5.028   7.495   9.986
 -3.839  -3.889  -3.841  -3.625   9.610
-21.899 -18.831 -21.748 -18.775 35.522 450.288 '

wheaton.cov <-
    getCov(lower, names = c("anomia67", "powerless67",
                            "anomia71", "powerless71",
                            "education", "sei"))
```

```
# classic wheaton et al model
wheaton.model <- '
  # latent variables
    ses     =~ education + sei
    alien67 =~ anomia67 + powerless67
    alien71 =~ anomia71 + powerless71
  # regressions
    alien71 ~ alien67 + ses
    alien67 ~ ses
  # correlated residuals
    anomia67 ~~ anomia71
    powerless67 ~~ powerless71
'
fit <- sem(wheaton.model,
           sample.cov = wheaton.cov,
           sample.nobs = 932)
summary(fit, standardized = TRUE)
```

その他できること

- 適合度指標
  - fitMeasures関数を使うと50種類近い指標
  - もちろん選択して出すことも可能

```
> fitMeasures(fit, c("chisq", "df", "pvalue", "cfi", "rmsea"))
 chisq     df pvalue    cfi  rmsea
85.306 24.000  0.000  0.931  0.092
```

- 修正指標
  - modindices(フィットさせたモデル)で出力される

```
> modindices(fit)
       lhs op     rhs group     mi mi.scaled    epc sepc.lv sepc.all sepc.nox
1   visual =~      x1     1     NA        NA     NA      NA       NA       NA
2   visual =~      x2     1  0.000     0.000  0.000   0.000    0.000    0.000
3   visual =~      x3     1  0.000     0.000  0.000   0.000    0.000    0.000
4   visual =~      x4     1  1.498     1.026  0.099   0.104    0.090    0.090
5   visual =~      x5     1  9.593     6.565 -0.281  -0.294   -0.225   -0.225
6   visual =~      x6     1  3.972     2.718  0.137   0.144    0.146    0.146
```

その他できること

- 潜在曲線モデルを推定する関数growthも

```
model <- ' i =~ 1*t1 + 1*t2 + 1*t3 + 1*t4
           s =~ 0*t1 + 1*t2 + 2*t3 + 3*t4 '
fit <- growth(model, data=Demo.growth)
summary(fit)
```

- 順序尺度変数
  - カテゴリカル・データであることを指定する方法は二通り
    - 与えるデータセット変数が順序尺度水準であることを指定する；data.frameのordered型を設定
    - Fitting関数内で orderd=変数名で指定

```
fit <- cfa(myModel, data = myData,
           ordered=c("item1","item2",
                     "item3","item4"))
```

M*plus*との連携

- lavExport
  - lavaanのモデルをM*plus*書式で出力する
- Mplus2lavaan
  - M*plus*のinpファイルを読み込む

M*plus*とR with lavaanとの比較

- M*plus*が勝るところ
  - ベイズ推定
  - マルチレベルモデル
  - 混合分布モデル
- lavaanが勝るところ
  - モデル的には特にない
  - Rという分析環境のメリット（環境を統一できる，他の分析モデルと合わせる，フリーソフトウェアであることなど）

Enjoy!
Mplus and R Life

提供
小杉考司（山口大学）
清水裕士（広島大学）