

SEMの理論と多母集団分析

7/3 教育学研究科

M1

藤田 弥世

SEMとは

- structural equation modelの略

；構造方程式モデル(別名.共分散構造分析)

多変量解析の色々な手法を統合したモデル

⇒相関行列や共分散行列を利用して、多くの
変数間の関係を総合的に分析する手法



共分散(≡相関係数)の観点から、相関係数で関連の大小を評価することができるデータすべてに適用可能

パス解析との違い

前回の授業の修正点...

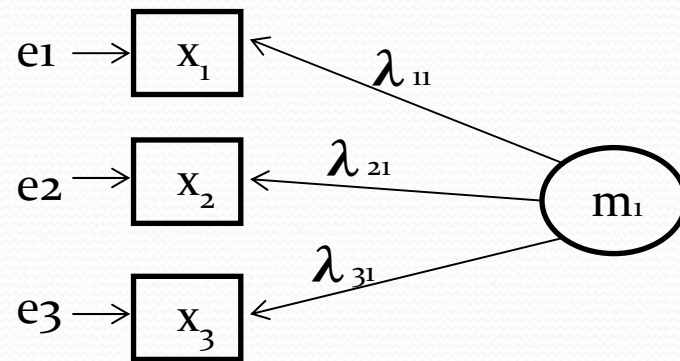
- パス解析は、複数の外生&内生変数を扱うと述べたが、それらはすべて観測変数でなければならず、したがって観測変数のみを取り扱う分析
- SEMは、観測変数だけでなく潜在変数も取り入れた分析
⇒構成概念(潜在変数)間のパス解析モデル
因子分析と回帰分析を同時に行うモデル

SEMの下位モデル

- 測定方程式だけを用いたモデル
 - ：ex) 因子分析, 主成分分析
 - 一般化可能性係数の推定モデル
- 測定方程式&構造方程式を用いたモデル
 - ：ex) MIMICモデル, 多重指標モデル,
 - 正準相関分析
- 構造方程式だけを用いたモデル
 - ：ex) パス解析, 単/重/多変量回帰分析,
 - 分散/共分散分析 数量化Ⅱ類

付録1. 測定方程式？構造方程式？

● 測定方程式



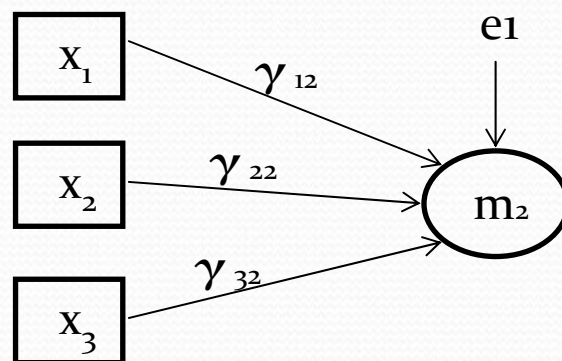
$$x_1 = \lambda_{11} \times m_1 + e_1$$

$$x_2 = \lambda_{21} \times m_1 + e_2$$

$$x_3 = \lambda_{31} \times m_1 + e_3$$

⇒ 共通の原因としての潜在変数 (m_1) が複数個の観測変数 (x) に影響を与えている様子を記述するための方程式

● 構造方程式



$$m_2 = \gamma_{12} \times x_1 + \gamma_{22} \times x_2 + \gamma_{32} \times x_3 + e_1$$

⇒ 変数間の因果関係を表現するための方程式

なぜSEMを使うのか？

下位モデルを見れば分かるように、SEMでなくとも分析可能な手法は盛りだくさんですが・・・

SEMで行う主なメリット3点

①適合度の評価が可能

；適合度指標

②モデルの改善が容易

；LM検定 (Lagrange Multiplier Test) —引いていないパス

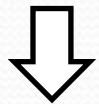
ワルド検定 —引いているパス

③推定値の標準誤差を評価できる

SEMに適さないデータはあるのか？

相関係数の定義をおさらい・・・

⇒ 間隔or比率尺度について、直線的な関係の度合いを評価するための指標



つまり、SEMは本来、量的変数の為の分析法



- 名義尺度
- 順序尺度は5水準以上なら○
※2は×, 3~4はグレーゾーン
- 非線形データ
- セレクションバイアスデータ

内生変数

標本数は？

- 標本数が少ないと、一般化の妥当性が得られなかったり、検定力が低くなったりと色々支障が出ますが、必要な標本数の基準は？
- 明確な一般的基準は存在しない
 - ⇒モデルの複雑さが必要となる観測変数に影響を与える
 - ex) 先行研究の結果(最尤法を用いた検討)
 - ；5つの観測変数が1つの因子を測定するモデル
 - 標本数10で○
 - 5つの観測変数が2つの因子を測定するモデル
 - 標本数20～30で○
- 観測変数より標本数の方が多いこと





SEMについてざっくり説明したところで・・・

今回はSEMならではの分析の一つ、

多母集団分析

について説明致します。

単一母集団分析の罠

例えば・・・

ある集団AとBそれぞれに「睡眠時間と課題への集中時間」の関係について実験を行い、別々/併合して分析

↓

別々ver. 集団A $r_1=.65$ / 集団B $r_2=.87$

併合ver. $r_3=.61$

！安易に多集団同士を併合すると測定結果が変わってしまう！

⇒単一の母集団から抽出したことを述べるには、集団間に上記のような回答傾向の差がないことを証明しなければならない。

⇒今回の場合、母集団がAかBかによって測定値の値が大きく変化しうるため、測定値の独立性が失われている。

値の独立性

①標本の独立性

;表の行間において互いに独立

特に多母集団解析の際に注意が必要

②観測変数の独立性

;表の列間において互いに独立(=無相関)

変数同士が無相関であるデータを分析する意義なし

=値の変動に共通する部分のある多変数間の関係を分析することに意義がある

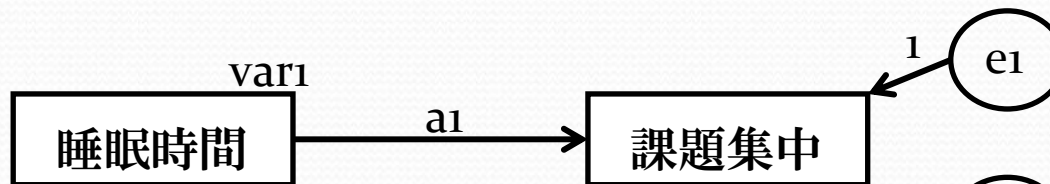
	学習時間(分)	テストの点数(点)	所持スカート数
A	180	92	5
B	35	47	0
C	70	68	2

集団ごとの分析の弊害

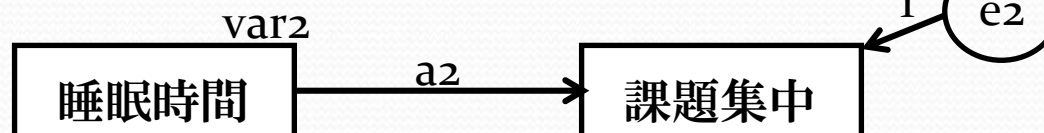
異質な集団であるならば、集団ごとに分析をすれば？

↓

集団A;



集団B;



⇒ 集団間で回帰モデルの解が正しいか否かを検討するためには、同種類の推定値同士の有意差検討

: a_1 と a_2 , var_1 と var_2 , e_1 と e_2

モデルの各部位における集団間の母数差異は評価○
モデル全体としての差異についての言及は難しい

SEMで用いる母数—1. 自由母数

- 自由に値を取りうる未知母数
- 回帰分析、探索的因子分析など、一般的な多変量解析において想定される母数

ex) $y = bx + a$ ・・・単回帰式

⇒ 母数 = a, b (母集団における値)

※この式が単なるサンプルとして取られたデータのための記述であれば、

$a, b =$ (標本) 統計量

SEMで用いる母数—2. 固定母数

- ある値に特定されてモデルに組み込まれる母数

- 利用の仕方

- ①モデル識別のための利用

- ;任意の観測変数への影響指標 1 固定

- 因子の分散を 1 に固定

- 因子平均を比較時に一方の因子平均 0 固定

- ②モデルに実質科学的知見を反映させるための利用

- ;確認的因子分析時に影響指標 0 固定

測定方程式

SEMで用いる母数—3. 制約母数

- 他の母数の関数に等しいという制約条件が課せられる母数(等値制約)
- 多母集団の同時解析や縦断的データ解析において良く利用
 - ⇒2群/時点間で同一の構成概念の違いや変化を検討するためには、因子の意味が同じでなければならないから

多母集団分析の登場

- データが複数の母集団から抽出されたことを認める分析
- 複数の母集団から抽出された標本を同時分析
- 母集団間の等質/異質性を検討することが目的

多母集団分析

- 標準的な手順

多母集団パス解析

- ①モデルの構成
- ②母集団ごとの分析
- ③配置不変性の検討
- ④等値制約

多母集団因子分析

- ①モデルの構成
- ②母集団ごとの分析
- ③配置不変性の検討
- ④測定不変性の検討
- ⑤外生変数の分散・共分散の制約

⇒今回の例では、

男女別、他者評価と自身のなさの関係

について分析を行います(豊田, 2006)

多母集団分析—①モデルの構成

- 複数の集団で共通して適用されるパス図の作成

今回の場合・・・

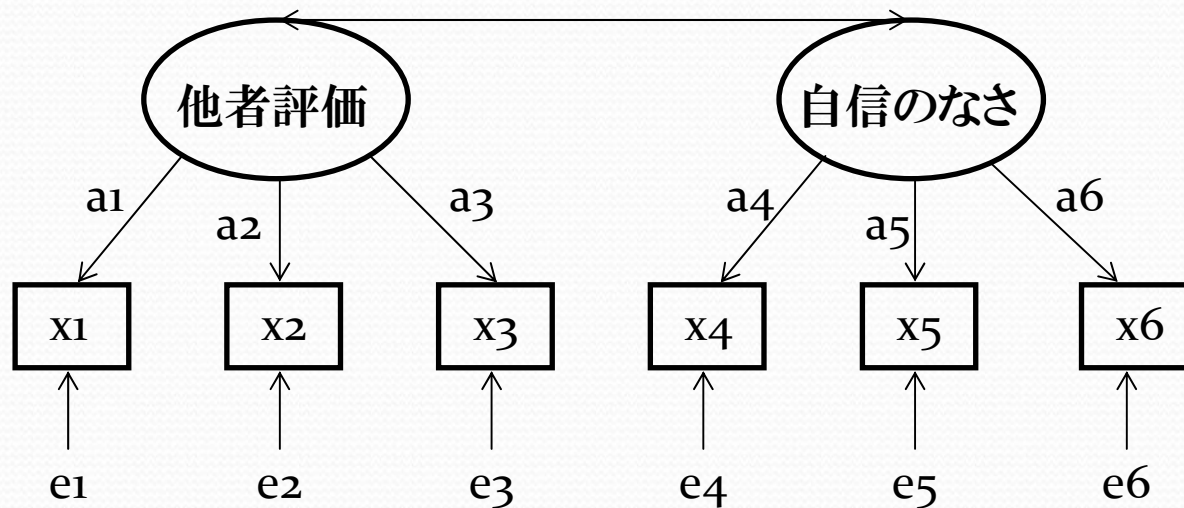
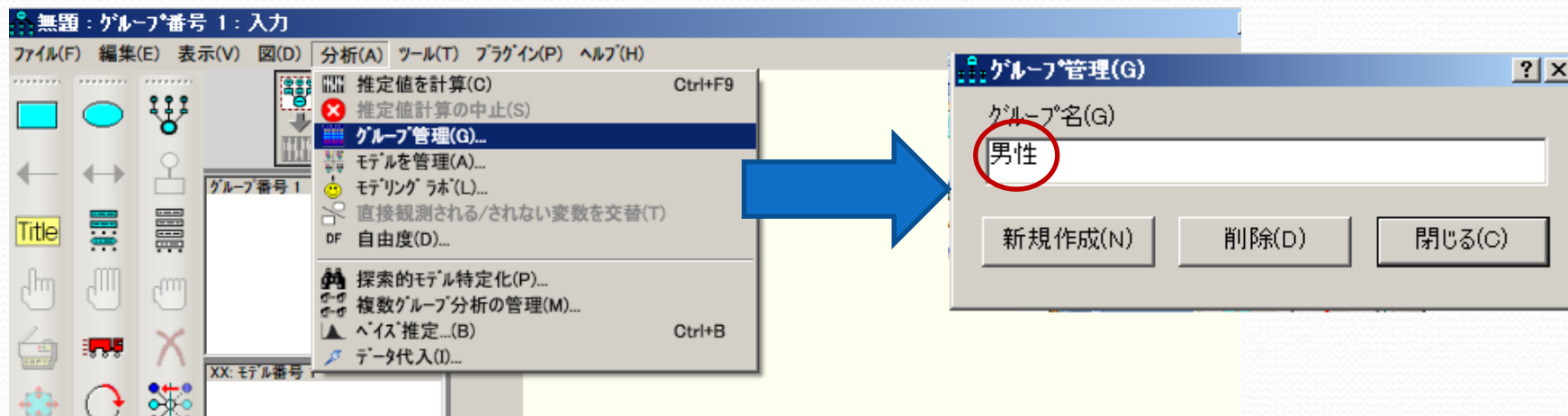


Figure. 1 モデル例

多母集団分析—②母集団ごとの分析

- Figure.1について、集団ごとに(この場合は男/女)モデル分析を行い、適合度が高いことを確認する
- 手順・・・Amosを立ち上げ、「分析」から「グループ管理」を選択し、グループ名を「男性」にする



多母集団分析—②母集団ごとの分析

「ファイル名」から→「発表用データ」選択
「グループ化変数」から→「性別」選択
「グループ値」から→「1」選択

The screenshot shows the SPSS 'Data File' dialog box. The 'File Name' field is highlighted with a red circle. The 'Grouping Variable' dropdown menu is also highlighted with a red circle, showing '性別' (Gender) selected. The 'Group Value' dropdown menu is highlighted with a red circle, showing '1' selected. The 'File Name' field contains '発表用データ.sav'. The 'Grouping Variable' dropdown shows '性別' (Gender) selected. The 'Group Value' dropdown shows '1' selected. The 'Data File' dialog box is open, showing the 'File Name' field, the 'Grouping Variable' dropdown, and the 'Group Value' dropdown. The 'File Name' field contains '発表用データ.sav'. The 'Grouping Variable' dropdown shows '性別' (Gender) selected. The 'Group Value' dropdown shows '1' selected. The 'Data File' dialog box is open, showing the 'File Name' field, the 'Grouping Variable' dropdown, and the 'Group Value' dropdown. The 'File Name' field contains '発表用データ.sav'. The 'Grouping Variable' dropdown shows '性別' (Gender) selected. The 'Group Value' dropdown shows '1' selected.

グループ名	ファイル	変数	数値	標本数
男性	<ワーキング>			
グループ番号 2	<ワーキング>			

グループ名	ファイル	変数	数値	標本数
男性	発表用データ.sav	性別	1	206/320
グループ番号 2	<ワーキング>			

多母集団分析—②母集団ごとの分析

- Figure.1を描き描きし、データ分析実行
- この手順を女性のグループに対しても行い、結果をみる。

多母集団分析—③配置不変性の検討

- ここから同時分析の手続き
- ここでは、母数に関して集団間で等値制約を行わない
;ex) 男 a_1 =女 a_1
- 等値制約を行わないことで、集団間でパス図は一緒でも推定値はそれぞれ異なってもよいという仮説を表現
= 配置不変性

多母集団分析—③配置不変性の検討

データ ファイル(D)

グループ名	ファイル	変数	数値	標本数
男性	発表用データ.sav	性別	1	206/320
女性	発表用データ.sav	性別	2	114/320

ファイル名(N) 実行中のファイル(W) ヘルプ(H)

データの表示(D) グループ化変数(G) グループ値(V)

OK キャンセル

数値でないデータを許可する(A) ケースをグループに割当てる(C)

分析(A) ツール(T) プラグイン(P) ヘルプ(H)

- 推定値を計算(C) Ctrl+F9
- 推定値計算の中止(S)
- グループ管理(G)...
- モデルを管理(A)...**
- モデルグラフ(G)...
- 直接観測される/されない変数を管理(M)...
- 自由度(D)...

探索的モデル特定化(P)...

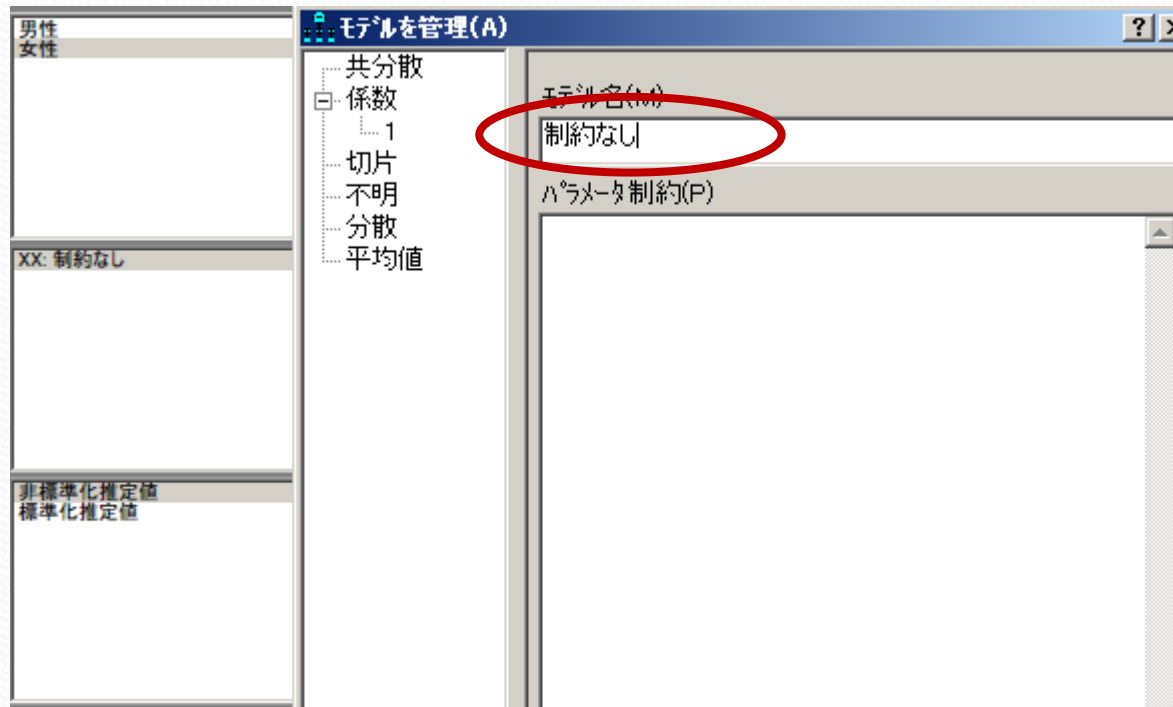
複数グループ分析の管理(M)...

ベイズ推定...(B) Ctrl+B

データ代入(I)...

モデル管理

多母集団分析—③配置不変性の検討



分析を実行

適合度指標を確認し、当てはまりが良ければ作成したモデルは両集団に共通して適合が良く、配置不変が成り立つ可能性が高いと言える

多母集団パス解析の場合

- 手順①～③の後・・・

- 手順④として

⇒ 集団間でモデルの各推定値の差異を検討

— 「差に対する検定統計量」 @ 「出力」タブ

= モデル全体として集団間の差異を述べてはいない

- ここから、モデル全体における適合を見るためには、異質性が疑われる (= 母数間の比較で差が大きい) 母数に対して集団間で等値の制約を置いて同時分析を行い、その適合の変化を検討する

⇒ モデル名を制約ありにし、パラメータ制約作成

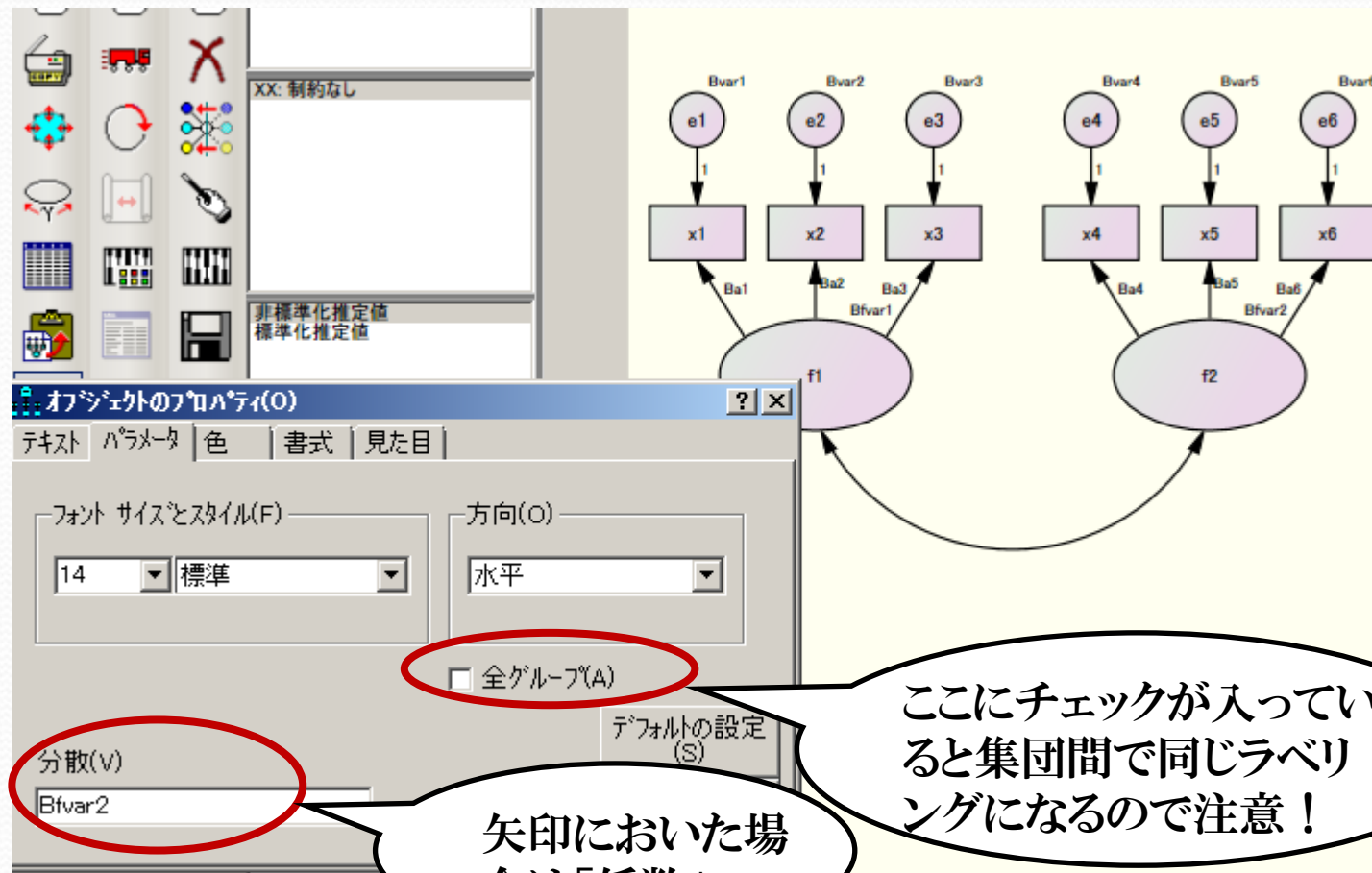
※ 先に推定値すべてにラベリングの必要あり

モデル名(M)
制約なし
パラメータ制約(P)

多母集団因子分析—④測定不変性の検討

- 母集団間で因子パターンが等しいと仮定したモデル作成
＝測定不変性
- 集団間で因子平均や因子分散を比較したい場合は、これが満たされている必要がある

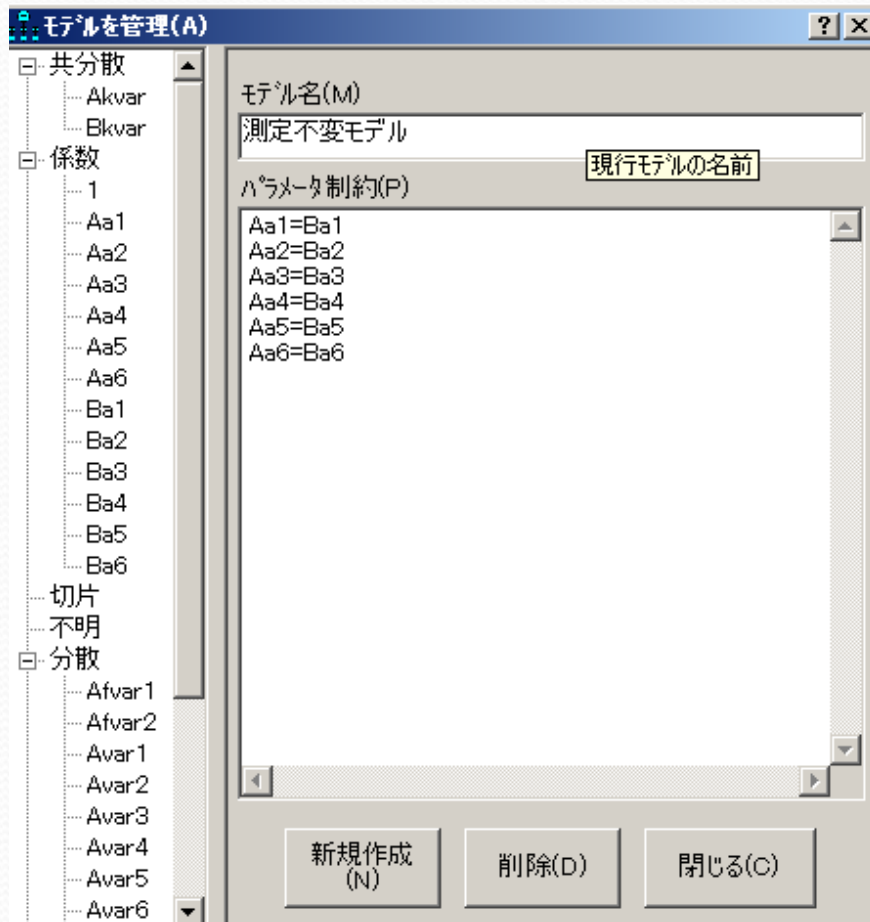
多母集団因子分析—④測定不変性の検討



ここにチェックが入っていると集団間で同じラベリングになるので注意！

矢印においた場合は「係数」

多母集団因子分析—④測定不変性の検討



多母集団因子分析

—⑤外生変数の分散・共分散の制約

- 母集団間で誤差分散や因子間の分散共分散に等値制約を置く。
- 集団間で、母集団の分散共分散行列が完全に一致するといふかなり強い制約
- Figure.1の例で述べるなら
 - ；・各因子 (f_1/f_2) での制約→2通り
 - ・因子 ($f_1 \& f_2$)間の共分散の制約→1通り
 - ・観測変数の誤差分散 ($e_1 \sim e_6$)の制約→6通り

多母集団因子分析

—⑤外生変数の分散・共分散の制約

- 測定不変性と合わせ、全部で16通りのモデル比較可能

モデル	測定	因子1	因子2	共分散	誤差	モデル	測定	因子1	因子2	共分散	誤差
1	○					9	○		○	○	
2	○	○				10	○		○		○
3	○		○			11	○			○	○
4	○			○		12	○	○	○	○	
5	○				○	13	○	○	○		○
6	○	○	○			14	○	○		○	○
7	○	○		○		15	○		○	○	○
8	○	○			○	16	○	○	○	○	○

(豊田, 2006)

多母集団分析—結果が出力されたら

テキスト出力の表示を見て、適合度をそれぞれ比べる

望ましい方向	指標名	とりうる値	「非常に良い」範囲	「悪い」範囲
小さい方が 良い	RMSEA	$RMSEA \geq 0$.05未満	.1以上
	AIC	制限なし	相対的比較	相対的比較
大きい方が 良い	GFI	$GFI \leq 1$.95以上	.9未満
	AGFI	$AGFI \leq GFI$.95以上	.9未満
	NFI	$0 \leq NFI \leq 1$.95以上	.9未満
	CFI	$0 \leq CFI \leq 1$.95以上	.9未満

(朝野・鈴木・小島, 2005)

多母集団分析の結果を受け

- 各分析に関し、すべての手順をこなした上で尚モデル適合が良好な場合には、複数の集団は同一の分散・共分散をもつ母集団に属すると判断できる

！ただし、平均の等質性について検討していないため、これまでの2つの分析のみで複数の集団の等質性までは判断できない。

- 反対に、異質な結果が出た場合、どちらの集団がどうなのかという比較も可能

集団の等質性を言及するために

- 平均の情報を用いて分析する、
多母集団の平均構造分析も行うと○
- これも測定不変が成立しないと適用×
- もっとも代表的なものは、因子分析における母集団間での因子平均の比較
⇒これにより、複数の集団や時点における構成概念の値の違いor推移を、直感的に分かりやすく解釈することが可能

欠損データへの対処

- 多母集団分析は欠損値への対応もO.K.
 - ；原因が分かっている欠損値
- この場合、欠損値のある群とない群に分けて分析を行う
- その他の欠損の種類に関しては、別の方法を採用
 - ；完全にランダムな欠損 (Missing Completely At Random)
 - ランダムな欠損 (Missing At Random)
 - ⇒リストワイズ削除・・・全消去
 - ペアワイズ削除・・・部分削除
 - 欠損値を補完する方法もあるがここでは割愛

引用文献

- 朝野熙彦, 鈴木督久, 小島隆矢(2005). 入門 共分散構造分析の実際. 講談社
- 狩野裕(2002). 構造方程式モデリングは, 因子分析, 分散分析, パス解析のすべてにとって代わるのか?. 行動計量学, 29, 138-159
- 豊田秀樹, 前田忠彦, 柳井晴夫(1993). 原因を探る統計学 共分散構造分析入門. 講談社
- 豊田秀樹(2006). 共分散構造分析[技術編]. 朝倉書店
- 豊田秀樹(2009). 共分散構造分析[疑問編]. 朝倉書店
- 使用データは豊田研究室HPより

<http://www.waseda.jp/sem-toyoda-lab/data/2boinnsi.txt>