高次因子分析モデルと階層因子分析モデルについて

○松田淑美* ・ 狩野裕**

* 大阪大学大学院人間科学研究科
** 大阪大学大学院基礎工学研究科

内容

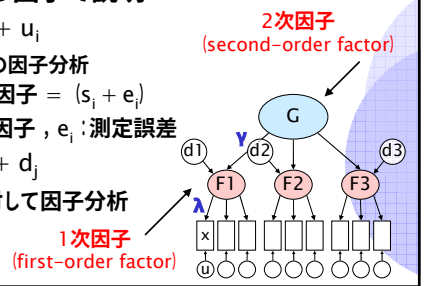
- 高次因子分析モデルおよび階層因子分析モデルについての説明
- 2 つのモデルの関係
- 2 つのモデルの比較
  - 因子間相関，γ へのパス
- 階層因子分析モデルの利用
  - 他の変数 γ との関連
- まとめ

高次因子分析モデルと階層因子分析モデル

2つのモデルの説明

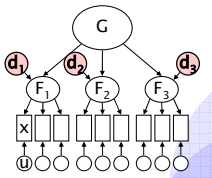
高次因子分析モデル
(higher-order factor analysis model)

- 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
  - $x_i = \lambda_i F_j + u_i$
    - 通常の因子分析
    - u_i :独自因子 = $(s_i + e_i)$
    - s_i :特殊因子，e_i :測定誤差
  - $F_j = \gamma_j G + d_j$
    - 因子に対して因子分析



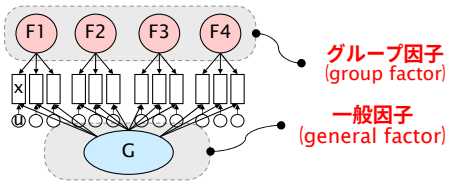
1次因子の独自因子 (d)

- $F_j = \gamma_j G + d_j$
  - G : 共通因子
  - d : 独自因子
    - ≠ *特殊因子 + 誤差*
- 測定誤差 (e) は，1次因子 (F) の測定時に除かれている
  - d = 因子の特殊因子 + ~~誤差~~
  - 独自因子 d は各1次因子の独自の特徴(個性) として解釈可能



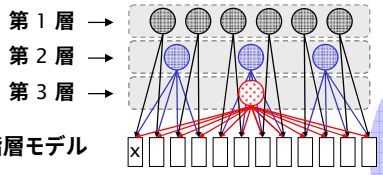
階層因子分析モデル
(hierarchical factor analysis model)

- 一般因子 (general factor)
  - 全ての観測変数を説明する因子
- グループ因子 (group factor)
  - 一部の観測変数を説明する因子



階層因子分析モデルの構造

- いくつかの "*layer (s)*" (「層」) がある
  - 最も因子が多い層を第1層，その次を第2層…



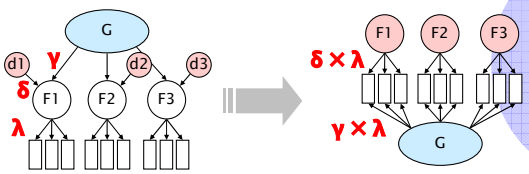
- e.g.)
  3層の階層モデル
  - 全ての因子は同じ次元 (order)，互いに直交
- 知能研究の分野でよく使われる

2つのモデルの関係

高次モデル ⇒ 階層モデル

- Schmid & Leiman (1957)
  - 高次因子分析モデルから階層因子分析モデルを導く方法を提案

e.g.) 2次因子分析モデル
⇒ 2層の階層因子分析モデル



高次モデル ⇒ 階層モデル

- 同じような内容を示す因子からなる
  - 2次因子 ＝ 一般因子
  - 独自因子d ＝ グループ因子 (因子の特殊因子)
- しかし，こうして導かれる階層モデルには，一般因子とグループ因子の因子負荷間に制約

$$\frac{\text{グループ因子の因子負荷}}{\text{一般因子の因子負荷}} = \frac{\delta_j \lambda_i}{\gamma_j \lambda_i} = \frac{\delta_j}{\gamma_j} = \text{定数}$$

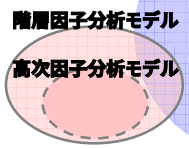
モデルの包含関係

- 高次因子分析モデルから，階層因子分析モデルを導くことができる
- ただし，その階層因子分析モデルには制約

⇒ 2つのモデルは同じではない (等しくない)

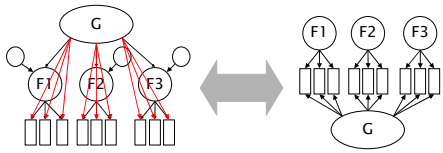
- 高次因子分析モデル ⊂ 階層因子分析モデル

- 制約が満たされている場合，2つのモデルは等しい
  - どちらのモデルも適合

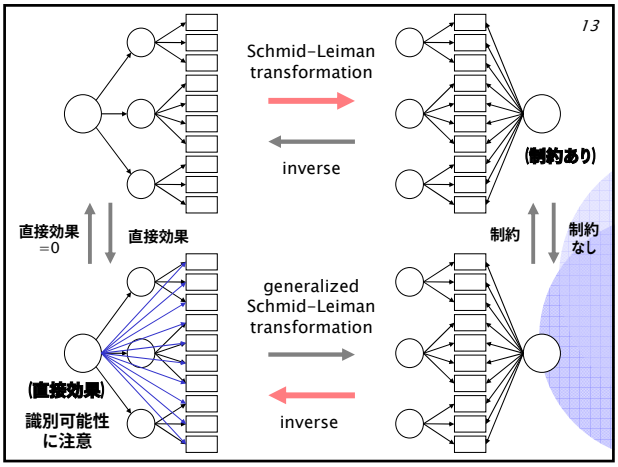


等しいモデル

- 「高次因子から観測変数への直接のパス」を加えると，2つのモデルは (数学的に) 等しくなる



- 識別のためには，制約が必要
  - 因子負荷の一部を 0
- 識別可能な場合に，自由に推定した階層因子分析モデルと (数学的に) 等しい

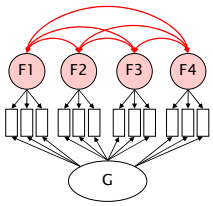


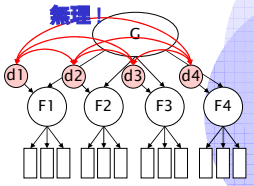
2つのモデルの比較

因子間相関
他の変数 y へのパス

因子間相関

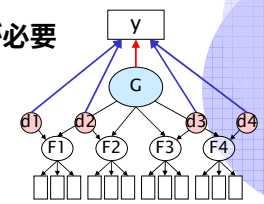
- 階層因子分析モデル
  - グループ因子間の相関を全て推定できる
- 高次因子分析モデル
  - d 間の全ての相関を推定することはできない





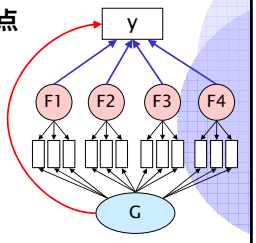
高次因子分析モデルの場合

- 独自因子d は因子の個性（誤差ではない）
  - y に対して何らかの影響がある可能性
- 「高次因子 & 全ての独自因子 d」によって y を予測するモデルは識別可能ではない
- 識別のためには工夫が必要
  - パスをひとつ減らす
  - 一部の推定値を事前の推定値で置き換える
    - Kano (2003)



階層因子分析モデルの場合

- 「一般因子 & 全てのグループ因子」によって y を予測するモデルは識別可能
  - 高次因子分析モデルより使いやすい
  - 階層因子分析モデルの利点
- 注意点
  - 制約を満たす場合には識別可能ではない
    - 高次因子分析モデルが適合する場合

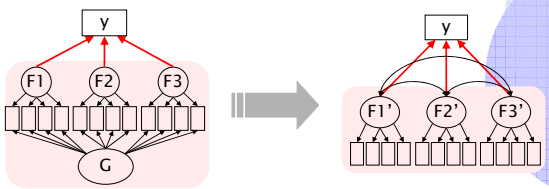


階層因子分析モデルの利用

他の変数との関連を検討する際の注意点

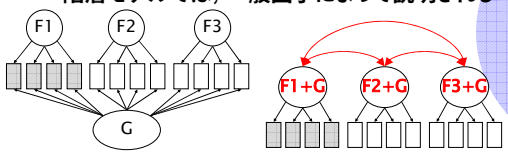
モデルの misspecification

- 階層モデルを，斜交モデルとして分析すると，どのような問題が生じるか？
  - y へのパス係数に関して
  - 制約を満たす場合，2つのモデルの適合は同じ



モデルの違い

- 因子の意味
  - 斜交モデルは，グループ因子による相関と一般因子による相関とをひとつの因子で説明
- 因子間相関
  - 異なる因子に属する項目間の相関を説明
    - 階層モデルでは，一般因子によって説明される

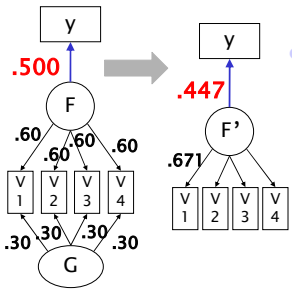


比較
(y へのパス係数)

- 2つの違いがどのような影響を与えるかについて検討
  - ① 1因子のモデル ・・・ 因子の意味の違い
  - ② 3因子のモデル ・・・ ＋ 因子間相関
  - 加えて，尺度得点の場合についても検討
- 制約を満たす階層モデルの場合を考える
  - 適合の変化による違いはない
    - 制約を満たす場合，斜交モデルでも適合
      - 階層モデル ⇒ 高次モデル ⇒ 斜交モデル
- 階層モデルの相関行列を用いて分析

1因子の場合
(因子の意味の違いによる影響)

[真のモデル]

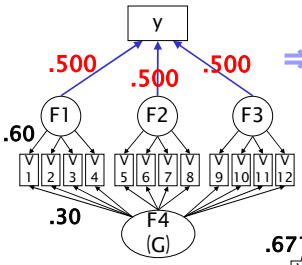


- y へのパス係数は小さくなっている
- 1因子モデルの因子は，G も含む
  - F' = F + G

⇒ G→y=0 のため，F' → y が小さくなる
- G→y=0 の影響
- 「一般因子による希薄化」

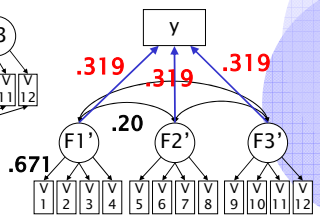
3因子の場合
(＋ 因子間相関の影響)

[真のモデル]

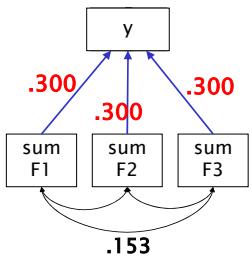


- 1因子の場合よりもさらに小さな値
  - (1因子の場合=.447)

⇒因子間相関があるため



尺度得点の場合
(＋ 測定誤差による希薄化)



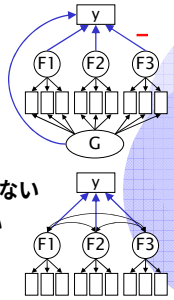
- 尺度得点の場合は，さらに小さな値
  - (斜交モデル=.319)
  - 測定誤差による希薄化の影響

まとめ
(y へのパス係数)

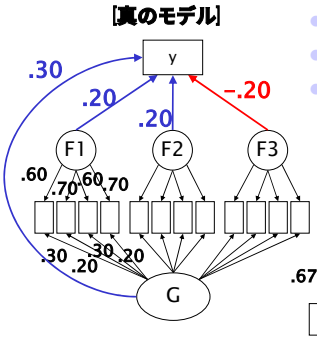
- 制約が満たされている場合，階層モデルと斜交モデルの適合は同じ
  - どちらかを選択する必要
- 他の変数 y へのパス係数は大きく異なる
  - 斜交モデルの場合
    - 「Gによる希薄化」と 因子間相関
  - 尺度得点の場合，さらに「誤差による希薄化」
- 注意しなければならない

より複雑なモデル

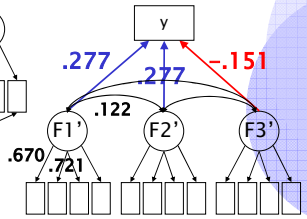
- G→y≠0かつ，さらに，y へのパス係数がグループ因子に異なる場合
  - さらに複雑な問題が生じる
    - G→y の影響
- 制約は満たさない
  階層因子分析モデルで考える
    - 満たしている場合には推定できない
  - 斜交モデルでは少し適合が悪い



3因子の場合



- G →y = .30
- F1,F2 →y = .20
- F3→y = −.20



3因子の場合

- 斜交モデルによる分析の場合，正の方向に大きくなっている
  - G が加わることの影響
    - F'=F+G
  - G→y =.30 であり，パス係数を大きくする方向に働く

	F1	F2	F3	G
階層モデル	.200	.200	−.200	.300
斜交モデル	.277	.277	−.151	---

6因子の場合

- グループ因子が6因子の場合について検討
  - 真のモデル (階層モデル)
    - G →y = .30
    - F1~F5→y = .20, F6→y = −.20

	F1	F2	F3	F4	F5	F6	G
階層モデル	.200	.200	.200	.200	.200	−.200	.300
斜交モデル	.247	.247	.247	.247	.247	−.234	---

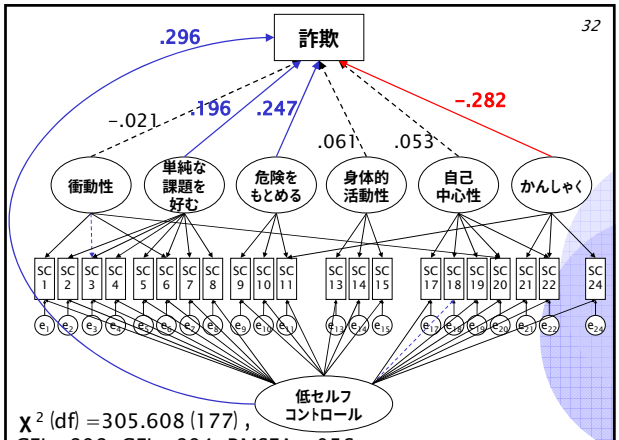
  - F1~F5→y は大きく，F6→y は小さくなっている

6因子の場合

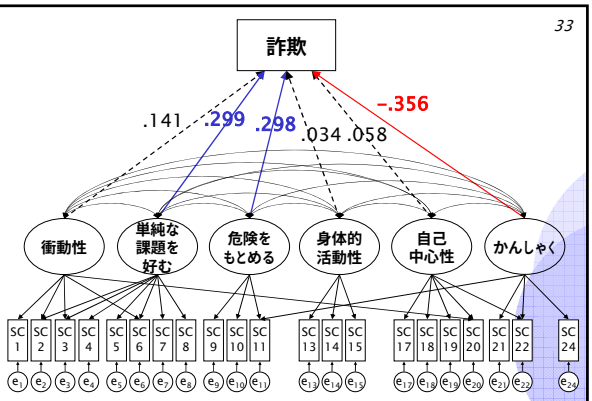
- F1'~F5' → y ： 大きくなっている
  - G→y = .30 の影響 (3因子の場合と同じ)
  - + 他の因子を経由する疑相関の影響
    - 因子数が6因子と多いため疑相関が大きい
    - そのため，3因子の場合に比べると小さい
      - 3因子･斜交= .277，6因子･斜交= .247
- F6' → y ：負の小さな値になっている
    - 3因子の場合= .151，正の方向に大
  - G→y =.30 により，少し正の方向に大きくなる
  - + F1 ~ F5 を経由する正の疑相関が大きい
    - 3因子の場合よりも疑相関が大きい

実データの分析

- セルフコントロール理論
  - 犯罪の原因は個人のセルフコントロールの低さにある
- セルフコントロール尺度（6因子）
  - 衝動性，単純な課題を好む，危険を求める，身体的活動性を好む，自己中心性，かんしゃく
- 詐欺犯罪
  - カンニング，他人のレポートや宿題を写す，授業をさぼる，万引き，自転車泥棒
- 「かんしゃく」→「詐欺」のみ負



χ^2 (df) =305.608 (177)，
CFI=.898，GFI=.894，RMSEA=.056



χ^2 (df) =298.833 (183)，
CFI=.909，GFI=.896，RMSEA=.052

実データの分析

- 因子分析部分の，モデルは少し違う
  - 階層因子分析モデル：衝動性→x2 なし
  - 斜交モデル：衝動性→x2 あり
  - 階層モデルの場合，少し不安定
- yへのパス係数
  - 斜交モデルで分析すると…
    - 「かんしゃく→詐欺」負の傾向が強まる
    - 「単純な課題を好む」と「危険をもとめる」は，より大きな値
      - 6因子の場合の例と同じ傾向

さまざまな状況

- 同じ6因子のモデルであっても，y へのパス係数の値によって，さまざまなパターン

Model ①	F1	F2	F3	F4	F5	F6	G
階層モデル	.200	.200	.200	.200	.200	-.200	.300
斜交モデル	.247	.247	.247	.247	.247	-.234	---

Model ②	F1	F2	F3	F4	F5	F6	G
階層モデル	.300	.300	.300	.300	.300	-.300	.300
斜交モデル	.289	.289	.289	.289	.289	-.353	---

さまざまな状況

- 斜交モデルの場合の，F1~F5のパス係数
  - Model① では大きく，Model② では小さくなる
  - ⇒Model② は，グループ因子のパス係数が大きく，そのため疑相関も大きい
- y へのパス係数の値
  - 因子数やグループ因子と一般因子のyへのパス係数によって，さまざまな変化
  - 注意が必要
  - 尺度得点の場合，誤差による希薄化が加わる

終わりに

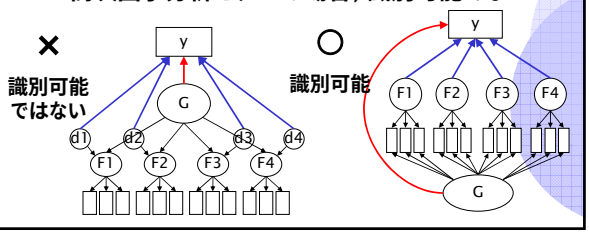
まとめなど

まとめ
(モデル選択)

- 高次因子分析モデルと階層因子分析モデルを紹介し，モデル間の関係を整理
  - 2つのモデルはよく似たモデル
    - 高次因子分析モデル ⊂ 階層因子分析モデル
      - 高次因子分析モデルがfit する場合，階層因子分析モデルも必ず fit する
  - 両モデルとも心理学における尺度構成に有用
- どちらのモデルを用いるか
  - モデルの適合度が判断材料のひとつ
  - 基本的には，研究者の仮説が大事

まとめ
(yの因子への回帰)

- 他の変数 y との関係を検討する場合，階層因子分析モデルの方が有用
  - 因子の特殊因子 と一般因子から y へのパス
  - 高次因子分析モデルの場合，識別可能でない



まとめ
(yの因子への回帰)

- 階層因子分析モデルが正しいとき
  - 斜交モデルや尺度得点を利用すると，誤った解釈をする可能性
    - 一般因子とグループ因子の両方を含む
    - 因子間相関の違い
    - どのように値が変化するかは，因子数や y へのパス係数による
    - ⇒貴重な情報を見落とす可能性がある
  - 基本的には，そのまま SEM を利用すべき

ありがとうございました