構造方程式モデリング

聖路加国際大学大学院看護学研究科
看護情報学分野
米倉佑貴

今回の内容

- 構造方程式モデリング(Structural Equation Modeling; SEM)
  - SEMの利点
  - 潜在変数と観測変数
  - 適合度
  - SEMで表現できる色々なモデル
  - 確認的因子分析
  - パス解析

構造方程式モデリング(SEM)とは

- 直接観測できる変数(観測変数)や直接観測できない変数(潜在変数)の関連性を分析するモデル
- いろいろな呼び方
  - 構造方程式モデル(Structural Equation Modeling: SEM)
  - 共分散構造分析(Covariance Structure Analysis: CSA)
- 多変量解析のいろいろなモデルを表現することができる
  - 因子分析モデル
  - 主成分分析モデル
  - 一般線形モデル

  など
- **主に仮説の検証に用いられる**
  - **仮説を設定しモデルとして表現する必要がある**
  - 探索的な分析をしたいなら重回帰分析や探索的因子分析を使う

用語の整理

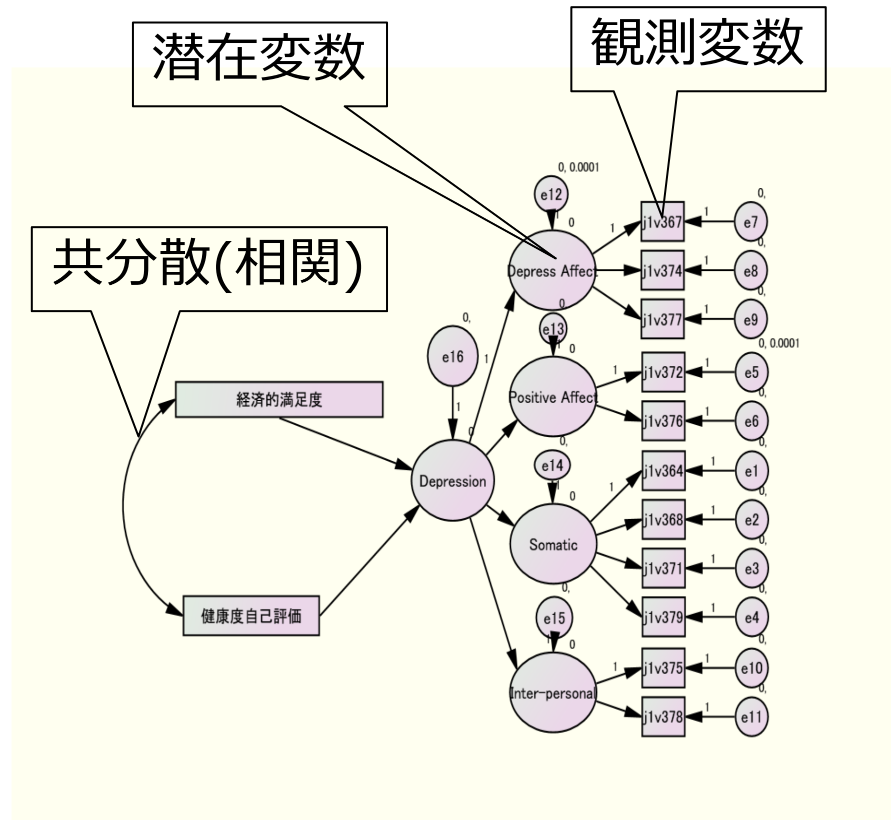
- 観測変数(observed variable)
  - 実際にデータが得られている変数のこと
- 潜在変数(latent variable)
  - 実際にデータが得られない変数のこと
- 内生変数(endogenous variable)
  - 少なくとも一回は他の変数の結果になる変数
- 外生変数(exogenous variable)
  - 一度も他の変数の結果にならない変数
- 測定方程式(measurement equation/model)
  - 因子分析の方程式のこと
- 構造方程式(structural equation/model)
  - 変数と変数の関連性を表す方程式のこと
- パス係数
  - 変数から変数への影響力の大きさ
  - 標準化したものと標準化していないものがある
  - 重回帰分析の偏回帰係数，因子分析の因子負荷量に相当するもの

構造方程式モデルの考え方

1. 変数間の関連の方程式（測定方程式 and/or 構造方程式）をたてる
2. 方程式から観測変数の分散、共分散を回帰係数と誤差分散等のパラメータで表現する
   - 観測変数の分散，共分散は(観測変数の数)×(観測変数の数+1)/2個あるので，立てられる方程式はその数が上限
   - 上限を超える場合は推定できない
   - 上限を超えていなくても推定できない場合がある
3. 目的関数に2の関係式を代入して，最適化する
   - 目的関数は推定方法によって変わる．最小二乗法なら誤差関数，最尤法なら尤度関数など
   - SEMでは観測変数が多変量正規分布に従うと仮定して最尤法を使うことが多い

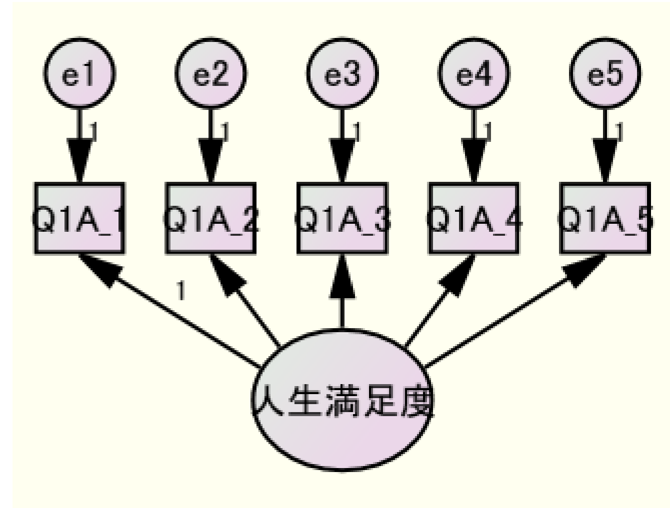
パス図

- 構造方程式モデルでは複雑な変数間の関係を扱うことができる
- 変数と変数の関係を表した図が**パス図**
- パス図のルール
  - 潜在変数(因子分析の因子など)は楕円で表す
  - 観測変数は四角で表す
  - 原因(説明変数)から結果(目的変数)に向けて矢印(→)を引く
  - 矢印が刺さる変数(=結果になる変数→内生変数)には誤差をつける
  - 共変動（共分散、相関）は双方向の矢印(↔)で表す



パス図の例

測定方程式とパス図



$$v_{1i} = \alpha_1 f_{1i} + e_1$$
$$v_{2i} = \alpha_2 f_{1i} + e_2$$
$$v_{3i} = \alpha_3 f_{1i} + e_3$$
$$v_{4i} = \alpha_4 f_{1i} + e_2$$
$$v_{5i} = \alpha_5 f_{1i} + e_n$$

v_i : 観測変数, α_i : パス係数(因子f_1からの因子負荷量)

f_1 : 因子得点, e_i : 誤差・独自因子

- 測定方程式=因子分析モデル
- 原則として観測変数は量的変数である必要がある

構造方程式とパス図

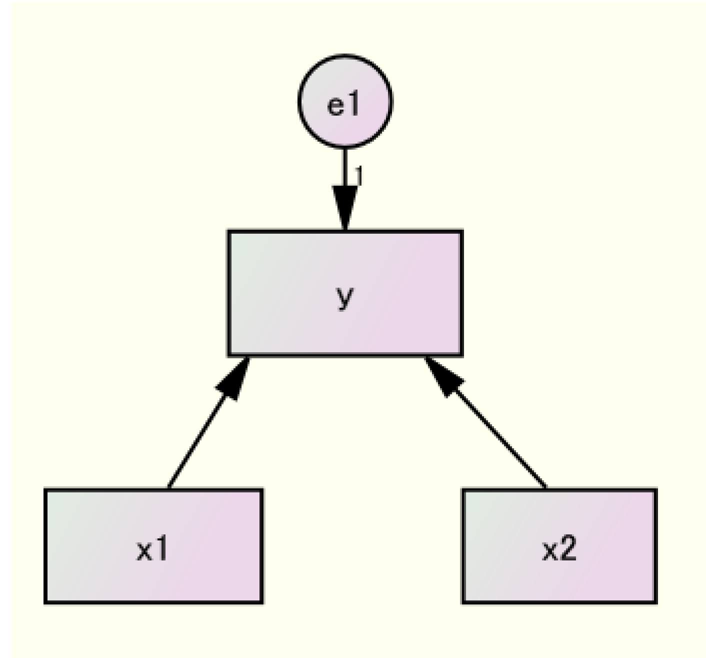
$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + e_1$$

y_i :従属変数

x_i :説明変数

β_i :パス係数(偏回帰係数)

e_i :誤差



- 構造方程式=測定方程式以外の変数の関連性を表す方程式
  - 重回帰式など
  - 目的変数，説明変数は観測変数でも潜在変数でもOK
  - 原則として内生変数は量的変数である必要がある
  - 外生変数が質的変数の場合はダミー変数にすればOK

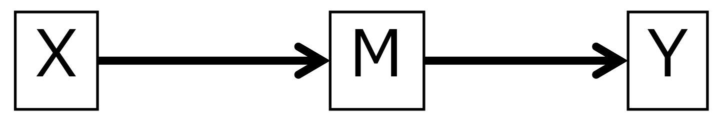
SEMの利点

- 柔軟・自由なモデル設定
  - 複雑な回帰モデル
    - パス解析
    - 直接効果・間接効果・総合効果の検討
    - 複数の目的変数→媒介効果などの検討
  - 因子分析と重回帰分析の組み合わせ
- 希薄化の修正
  - 多項目尺度の合計点を使った分析よりも関連性が強く出ることが多い
  - 因子分析モデルと一般線形モデルの組み合わせ
- モデル適合度の算出とモデル比較
  - 豊富な適合度指標でモデル全体の当てはまりを評価，比較できる

効果の分解と統合
直接効果・間接効果・総合効果

- 直接効果
  - 説明変数から目的変数に直接ささる場合
- 間接効果
  - 他の説明変数を経由して目的変数にささる場合
  - 経路のパス係数を全て掛けあわせることで算出できる
- 総合効果
  - 直接効果+間接効果
- 重回帰分析では間接効果の推定にX→M，M→Yの2つのモデルを推定(パス解析)しないと算出できないが，SEMでは同時に推定して1回で出せる
- 重回帰分析の繰り返しによるパス解析よりも，SEMによるパス解析の方が望ましい

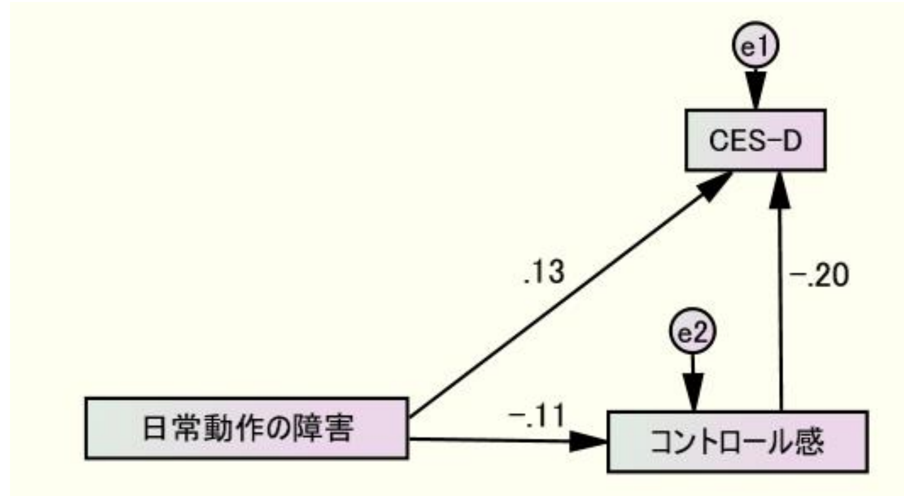
XからYへの直接効果



XからYへのMを介した間接効果

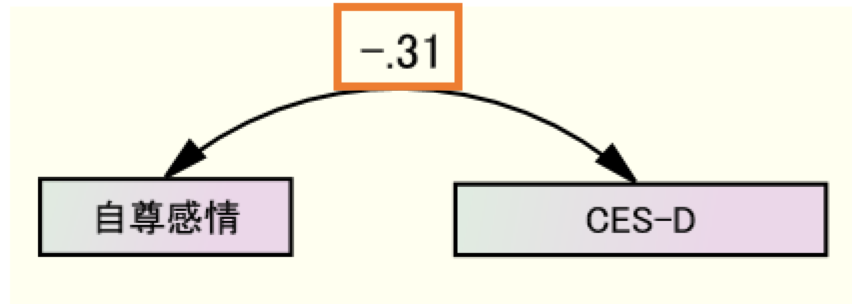
直接効果・間接効果・総合効果の例

- 「日常動作の障害」のCES-Dへの直接効果=0.13
- 「日常動作の障害」のコントロール感を介したCES-Dへの間接効果 =-0.11×-0.20=0.022
- 「日常動作の障害」のCES-Dへの総合効果 =0.13+0.022=0.152

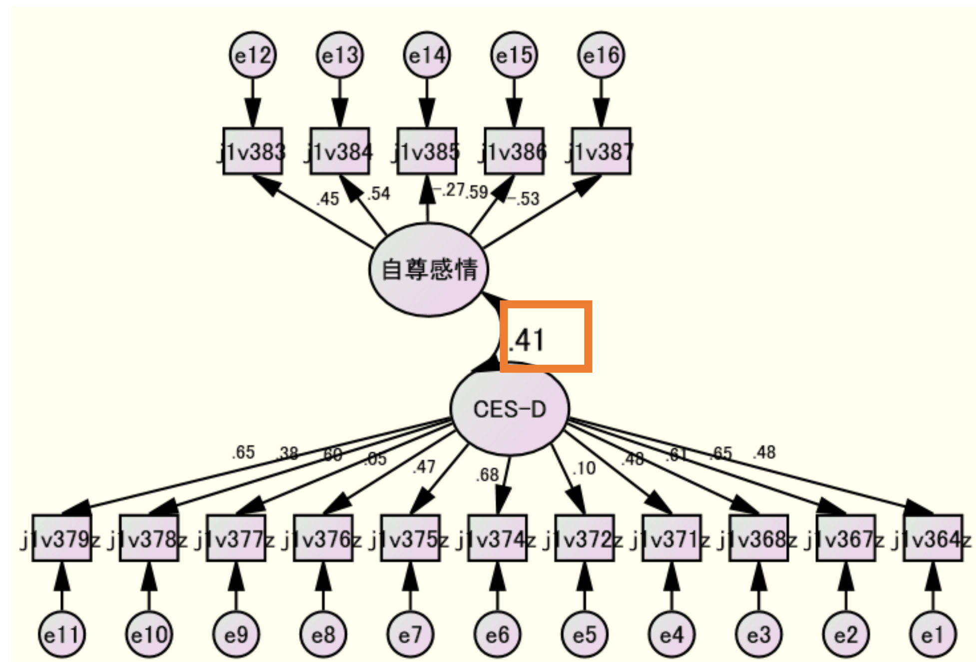


希薄化の修正

- 尺度で測定する構成概念を潜在変数化することで，構成概念との関連性(パス係数)が強く出る
- 例:自尊感情とCES-Dの関連(右図)
  - 合計点を観測変数として使った場合(右図上)
    - 相関係数は-0.31
  - 潜在変数化した場合(右図下)
    - 相関係数は0.41



尺度の合計点を観測変数として使った場合



潜在変数化した場合

適合度指標(1) 乖離度

- カイ二乗値
  - 観測値とモデルのズレ
  - ズレが小さいほうがよい
  - カイ二乗値を使って検定ができる
    - 帰無仮説は「モデルがデータに適合している」
      →帰無仮説が棄却されなければモデルがデータに適合している
    - サンプルサイズが大きいと帰無仮説が棄却されやすくなるのが欠点
- RMSEA(Root Mean Square Error of Approximation)
  - 1自由度あたりのモデルの分布と真の分布のズレ
  - 0.05以下で当てはまりが良い，0.08以下は許容範囲，0.1以上は当てはまりが悪い
  - 値が0.05未満かの検定を用いることもできる

適合度指標(2)
独立モデルとの比較によるもの

- NFI(Normed Fit Index)
  - 独立モデル(観測変数間の相関がないモデル)との距離
  - 値が大きいほどよい
- CFI(Comparative Fit Index)
  - NFIを0から1の間の値を取るように変換したもの
  - 1に近いほど当てはまりがよい
  - 0.9以上は許容範囲，0.95以上なら当てはまりがよいと評価
- PCFI(Parsimonious Comparative Fit Index)
  - CFIにモデルの複雑さでペナルティをかけたもの
  - シンプルなモデルの方が評価される

適合度指標(3) 情報量基準

- AIC (Akaike’s Information Criterion)
  - $AIC = \chi^2 - 2 \times (自由度)$
  - AICの値が小さいほどよいモデルであると評価される
  - 絶対基準はなく複数のモデルを比較する時に使用される
  - SEMに限らず尤度に基づく統計モデル全般に使用できる
- CAIC (Consistent Akaike’s Information Criterion)
  - $CAIC = \chi^2 - (log(N) + 1) \times (自由度)$
  - AICよりも自由度の小ささ(=モデルの複雑さ)に対するペナルティを強めたもの
  - AICよりもシンプルなモデルの方が評価が高くなる
- BIC/SBC (Bayesian Information Criterion/ Schwarz’s Bayesian Criterion)
  - $BIC = \chi^2 - log(N) \times (自由度)$
  - CAICよりも自由度の小ささ(=モデルの複雑さ)に対するペナルティを強めたもの
  - CAICよりもシンプルなモデルの評価が高くなる

適合度指標(4) GFI関連

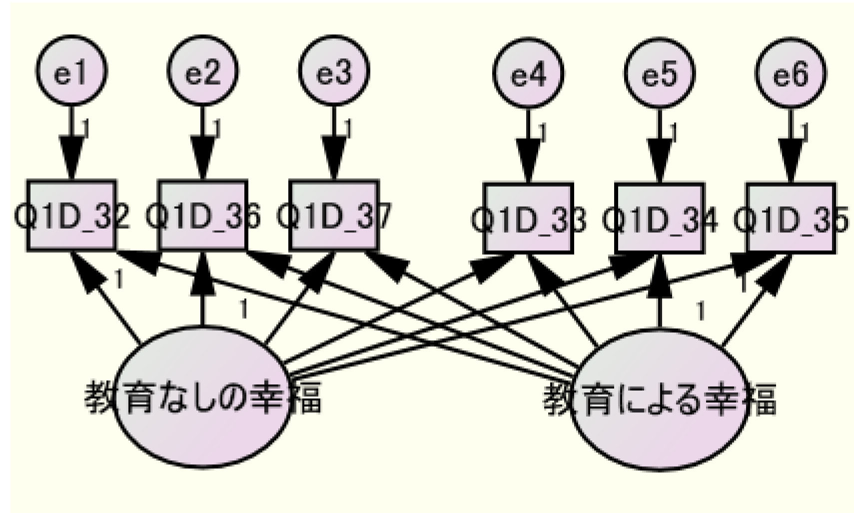
- GFI (Goodness of Fit Index)
  - 標本分散共分散行列とモデルを当てはめて推定した分散共分散行列を比較して評価
  - 重回帰分析の決定係数と似た指標
  - 推定するパラメータが多くなるほど高くなる
  - 1以下の値を取り，1に近いほど適合度がよい
  - 0.9以上が経験的な基準
- AGFI (Adjusted Goodness of Fit Index)
  - GFIを推定するパラメータの数で調整したもの
  - 重回帰分析における自由度調整済み決定係数と同じようなもの
  - 同じGFIを取るモデルが2つあれば，パラメータが少ないモデルのほうがよいと評価される

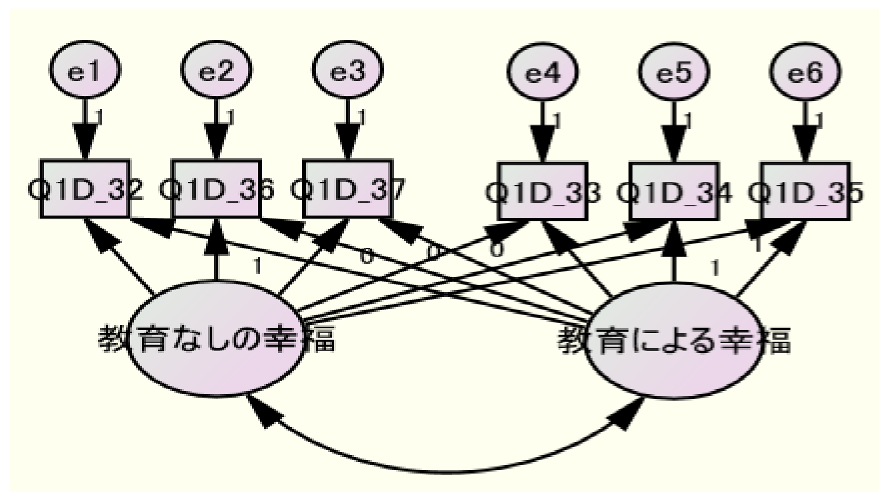
モデル評価・比較の方法

- 1つのモデルを評価する場合
  - 基準がある指標を用いて評価する
    - RMSEA，CFI，GFIなど
  - 適合度指標は複数の指標の評価から総合的に判断する
- 複数のモデルを比較する場合
  - 相対的な評価が可能な基準で比較する
    - 情報量基準(AIC，BIC，CAICなど)による比較，自由度調整をした指標(PCFI，AGFI，RMSEAなど)による比較
  - 入れ子(nested)関係にあるモデルの比較の場合は尤度比検定も利用できる
    - 入れ子モデル: あるモデルに何らかの制約を加えた形で表現できるモデルのこと

SEMで実行できるモデル(1)
因子分析モデル-探索的因子分析

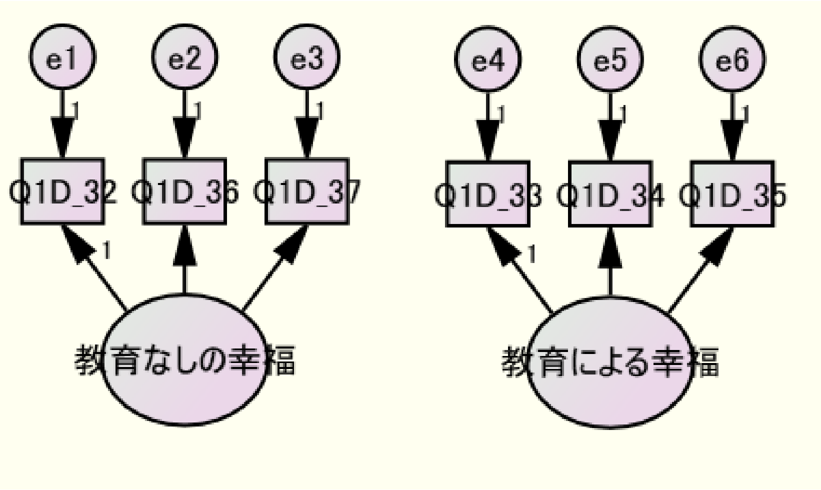
- それぞれの因子(潜在変数)からすべての項目にパスを引いたモデル
- 直交回転は潜在変数間の相関のパスを引かない(右上図)
- 斜交回転は潜在変数間に相関のパスを引く(右下図)
- SEMでやる意味はあまりない





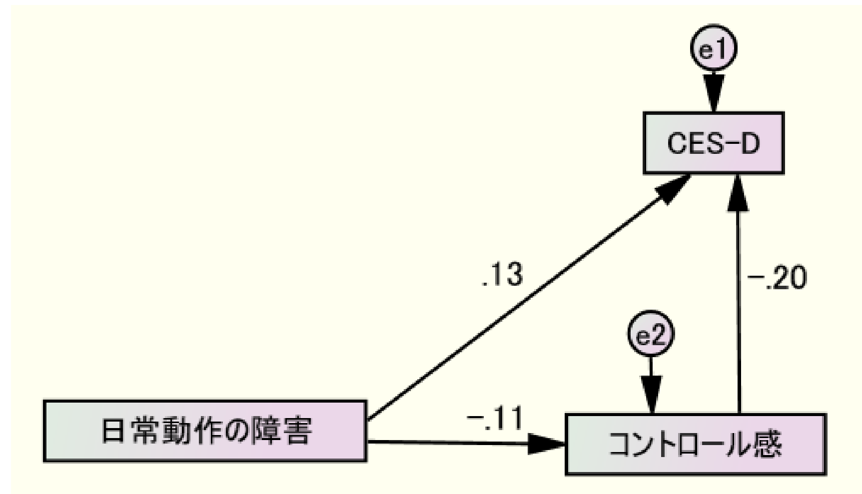
SEMで実行できるモデル(2) 因子分析モデル-確認的因子分析

- 仮説に基づいて因子負荷を設定
  - 因子帰属がない項目の因子負荷は積極的にゼロに固定(=パスを引かない)
- 全てのパラメータ(分散，共分散，因子負荷量)が自由パラメータ(推定するパラメータ)のモデルは推定できないのでパラメータに以下のどちらかの制約をかける
  - すべての因子の分散を1に固定
  - それぞれの因子から項目への因子負荷量のうち一つを1に固定
- 観測変数が2つ以下の場合はさらに制約が必要
  - 観測変数が1つ
    - 誤差分散を固定する
  - 観測変数が2つ
    - 因子負荷が等しいという制約か誤差分散が等しいという制約を置く
    - 他の因子と相関があるモデルにする



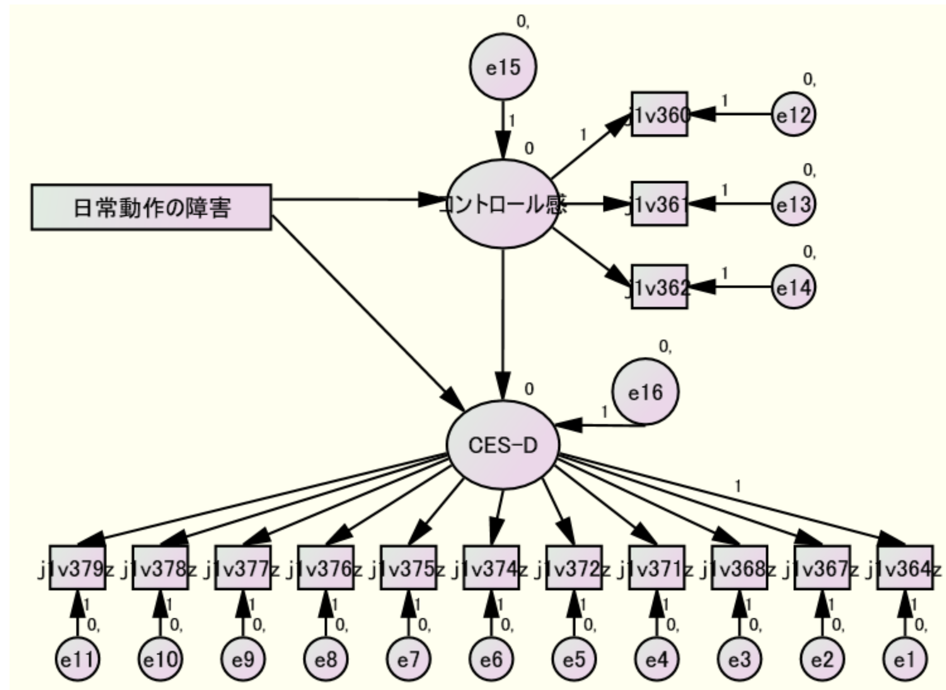
SEMで実行できるモデル(3)
重回帰モデル(パス解析モデル)

- 説明変数間の因果関係の仮説がある場合に使う
  - 特にない場合は普通の重回帰をすればOK
- 質的変数はダミー変数にすればモデルに組み込める
  - ただし質的変数が目的変数になる場合は特別な処理が必要



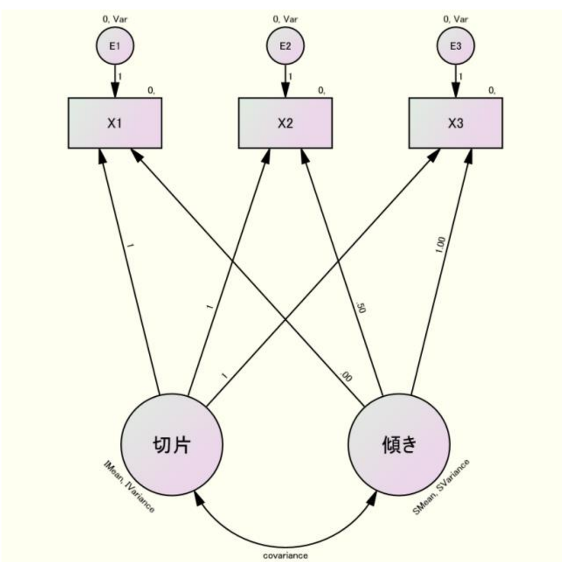
SEMで実行できるモデル(4)
因子分析モデルと重回帰モデルの組み合わせ

- 多項目尺度で測定する概念を潜在変数にして回帰モデルに組み込むことも可能



SEMで実行できるモデル(5)

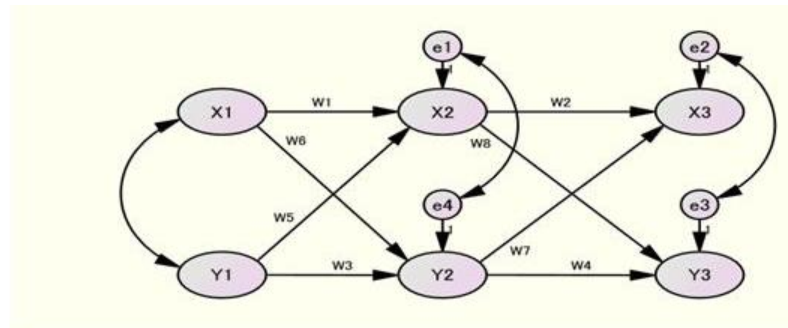
- 多母集団モデル
  - いくつかのグループごとにモデルを当てはめるモデル
  - グループ間で因子構造が違うかどうかや変数同士の関連性が違うかどうかなどを検討できる
- 潜在曲線モデル
  - 時系列データを切片と傾きを潜在変数化して記述するモデル
  - 切片や傾きを目的変数，説明変数にしてさらに複雑なモデルにすることも可能



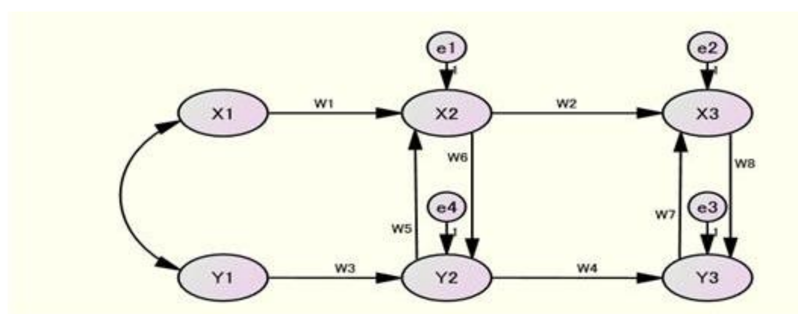
潜在曲線モデル

SEMで実行できるいろいろなモデル(6)

- ラグ付き変数モデル(交差遅延モデル，Cross-Lagged Model)
  - 縦断データで時点間の因果関係を検討するモデル



- 道具的変数モデル(同時効果モデル， synchronous effects model)
  - ある時点における2変数の双方向の因果関係を検証するモデル



分析例(1) 確認的因子分析

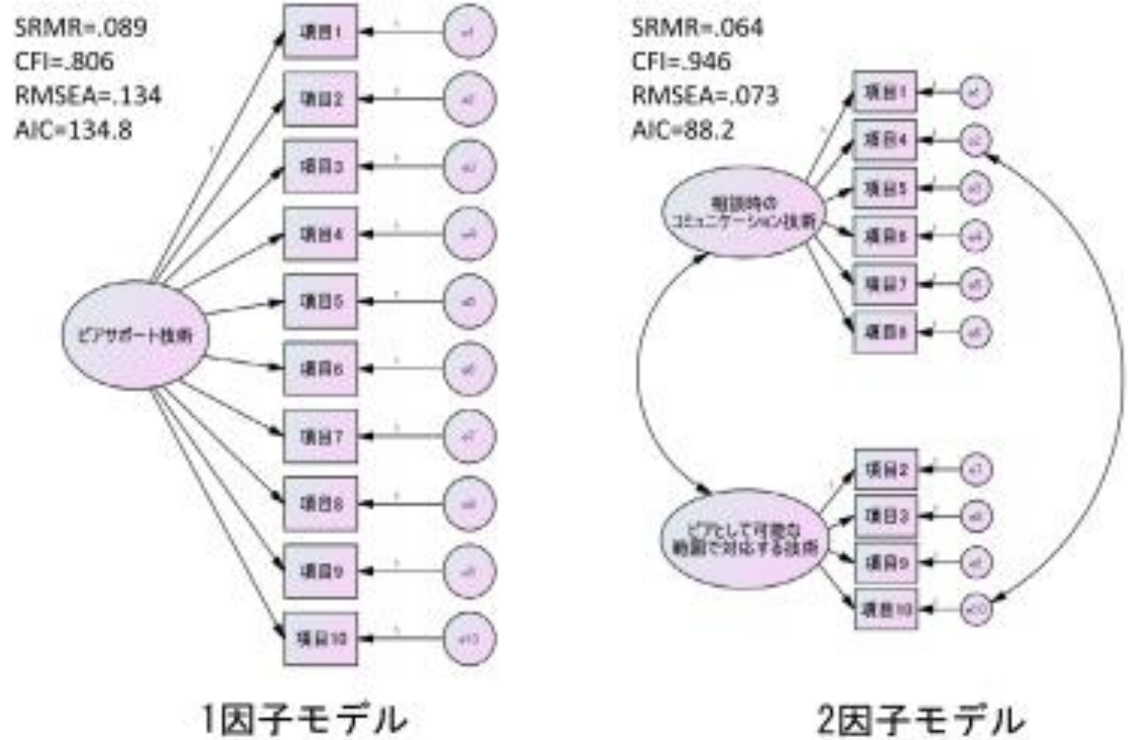


図１　各尺度の因子構造

上段：活動に対する負担感尺度，中段：活動に対する満足度尺度，下段：ピアサポート技術尺度

負担感は項目の一部に欠損があるため SRMR は算出していない

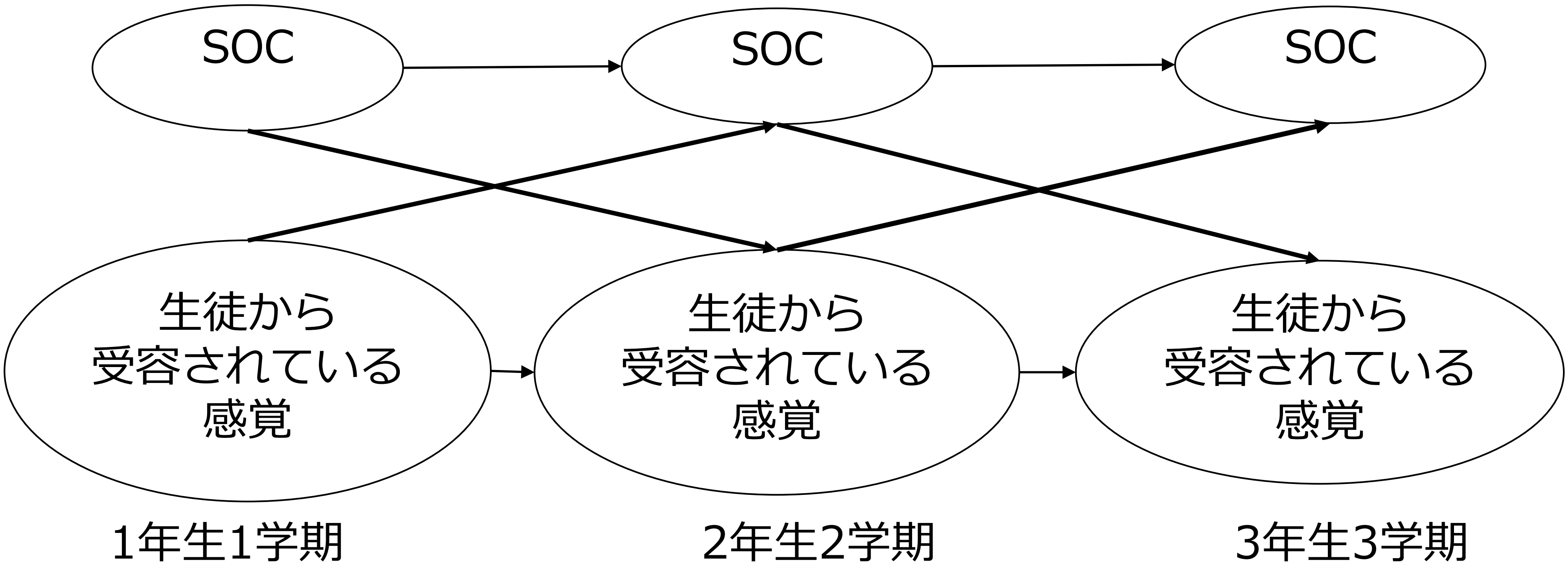
SRMR: Standardized Root Mean-Square Residual, CFI: Comparative Fit Index, RMSEA: Root Mean Square Error of Approximation,

AIC: Akaike's Information Criterion

- 慢性疾患患者を対象としたピアサポートの技術を測定する尺度の因子構造について，1因子モデルと2因子モデルを比較
  - 2因子モデルの方が適合度が良好かつ概ね基準をみたしているため，2因子モデルを採択

米倉 佑貴, 慢性疾患患者を対象としたピアサポートの提供者の負担感，満足感，サポート技術を測定する尺度の信頼性・妥当性の検討, 日本看護科学会誌, 2021, 41 巻, p. 556-566

分析例2: 交差遅延モデルとモデル比較



- 高校生を対象とした追跡調査のデータを用いた，ストレス対処能力SOCと生徒から受容されている感覚の因果関係の検討
  - 交差遅延モデルでSOCと学校帰属感覚のどちらが原因か検討する

出典: 山崎喜比古，戸ヶ里泰典編「思春期のストレス対処能力SOC」

適合度の比較

基準比較

モデル	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
both	.875	.816	.941	.911	.939
soc to pssm	.874	.817	.941	.911	.939
pssm to soc	.867	.807	.934	.901	.932
no relationship	.866	.807	.933	.901	.931
飽和モデル	1.000		1.000		1.000
独立モデル	.000	.000	.000	.000	.000

RMSEA

モデル	RMSEA	LO 90	HI 90	PCLOSE
both	.057	.046	.068	.138
soc to pssm	.057	.046	.068	.143
pssm to soc	.060	.049	.071	.058
no relationship	.060	.050	.071	.057
独立モデル	.191	.184	.199	.000

AIC

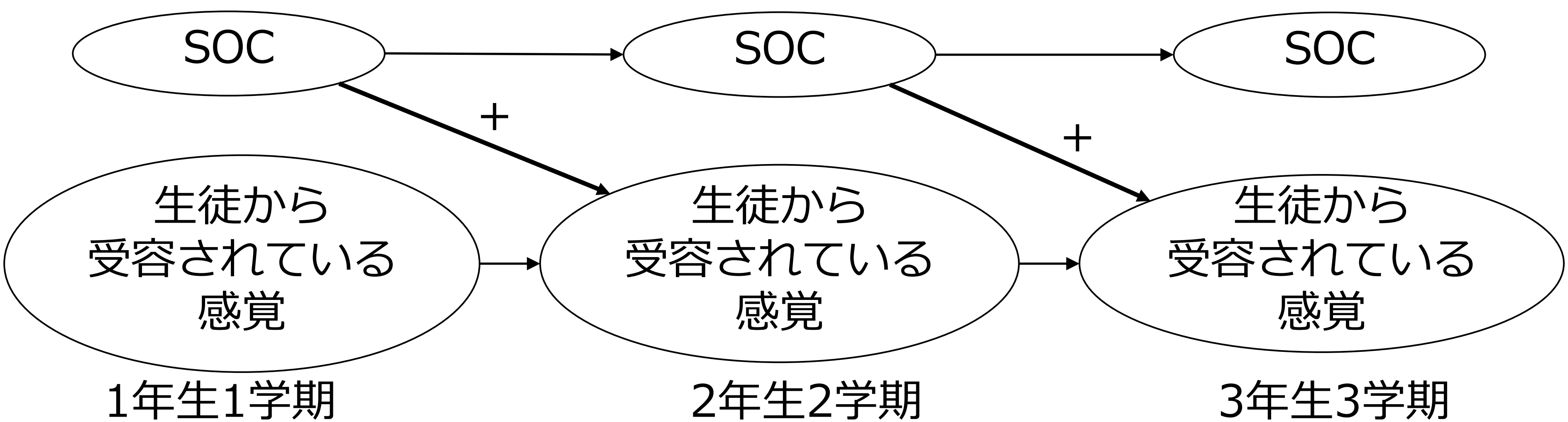
モデル	AIC	BCC	BIC	CAIC
both	468.202	487.735		
soc to pssm	467.042	486.164		
pssm to soc	481.308	500.430		
no relationship	481.115	499.825		
飽和モデル	504.000	555.813		
独立モデル	2269.820	2274.137		

- CFI，RMSEA
  - 双方向の因果(both)とSOCが原因になっているモデルがほぼ互角
- AIC，BCC
  - SOCが受容されている感覚の原因になっているモデル(soc to pssm)が低い
    → SOCが受容されている感覚の原因になっているモデルを採択

パス係数の確認

			推定値	標準誤差	検定統計量	確率	ラベル
SOCT2	<---	SOCT1	.705	.093	7.587	***	
AST2	<---	AST1	.426	.066	6.426	***	
AST2	<---	SOCT1	-.050	.019	-2.635	.008	W7
SOCT2	<---	AST1	-.317	.186	-1.704	.088	W8
SOCT3	<---	SOCT2	.681	.103	6.619	***	
AST3	<---	AST2	.557	.131	4.243	***	
AST3	<---	SOCT2	-.079	.029	-2.707	.007	W9
SOCT3	<---	AST2	-.054	.346	-.155	.877	W10

- 1年時のSOC(SOCT1)から2年時の受容されている感覚(AST2)が有意な関連
- 2年時のSOC(SOCT2)から3年時の受容されている感覚(AST3)が有意な関連
  →SOCが高い人ほど受容されている感覚を得られている



論文等に書くべきこと

- 仮説モデルの設定根拠→緒言か方法に記述する
  - 仮説モデルは事前に示しておく
  - 測定モデル(因子構造)，構造モデル(変数間の関連)についての理論的な背景を踏まえてパスを引く根拠を示す
- 適合度指標→結果に記述する
  - モデルの当てはまりの評価に使用した適合度指標の値を示す
  - 適合度指標は1つだけではなく，性質の違う複数の指標を示すのが望ましい
  - 複数のモデルを比較した場合は表にすると見やすい
- パス係数・負荷量→結果に記述する
  - パス図に示すか表にして示す

参考文献

- 豊田秀樹. 共分散構造分析 入門編-構造方程式モデリング. 朝倉書店.
- 豊田秀樹. 共分散構造分析 応用編-構造方程式モデリング. 朝倉書店.
- 豊田秀樹. 共分散構造分析 疑問編-構造方程式モデリング. 朝倉書店.