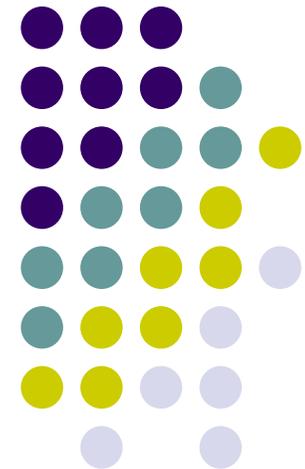


共分散構造分析

心理データ解析演習 07.11.28
工学研究科 M1 森安洋平





はじめに

- 共分散構造分析とは
数理的な定義: 観測変数間の分散・共分散の構造を分析する手法
- これではわかりにくいので・・・
定義: 直接観測できない**潜在変数**を導入し、その潜在変数と観測変数との間の**因果関係**を同定することにより社会現象や自然現象を理解するための統計的アプローチ。
因子分析と多重回帰分析(パス解析)の拡張。
- 潜在変数とは、心理学や社会学でいう「**構成概念**」を数学的に表現したもの。



コンテンツ

- 因果関係とは
- 2変数の関係
- 回帰モデル
- 共分散構造モデルの導入
- モデルの評価
- Amosを使った共分散構造分析

因果関係とは 潜在した共通の原因



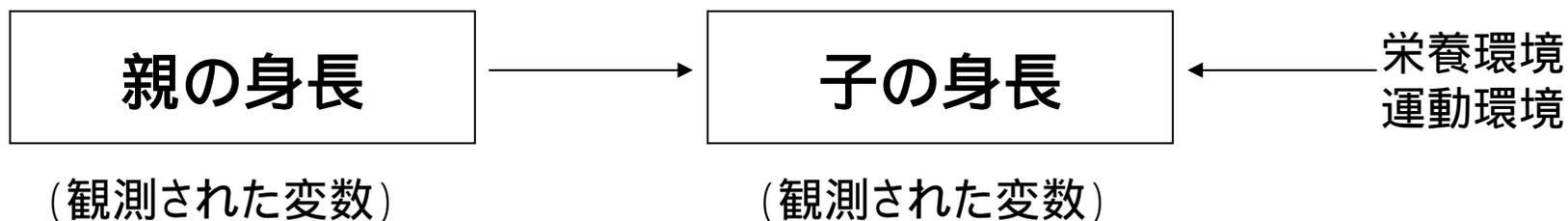
現象の背後に共通して存在する原因

構成概念

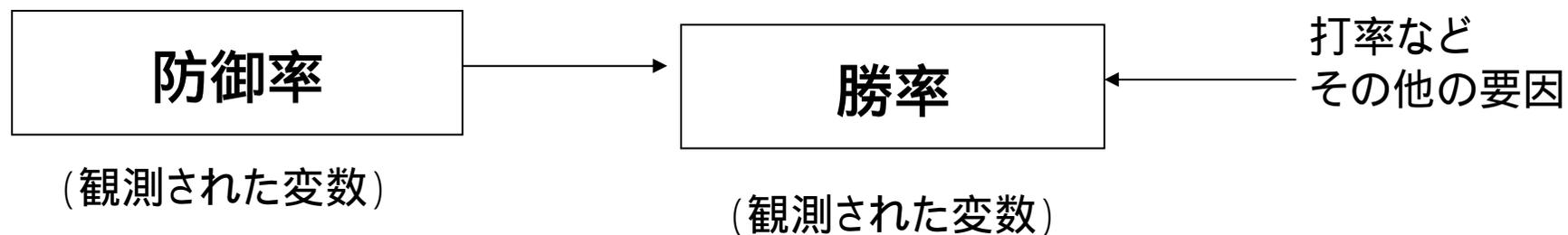
観測された現象

想定した原因以外の個別の要因(誤差)

因果関係とは 一方が原因で他方が結果

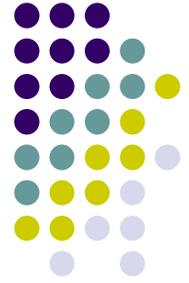


時間的な前後関係



意味的な前後関係

因果関係とは まとめ

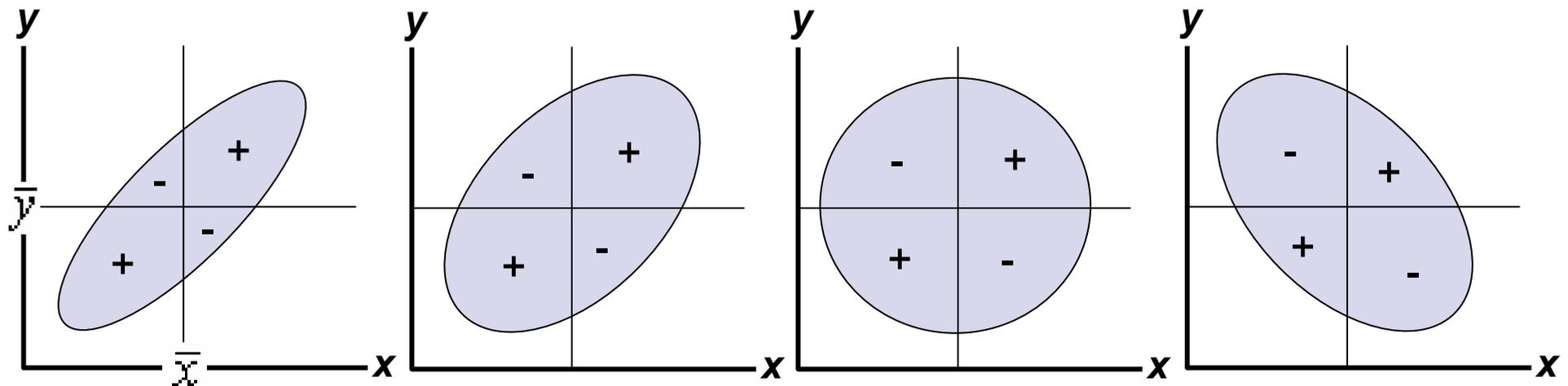


- 統計学において因果関係を扱う場合には大きく分けて2つのタイプの考え方がある
- 「2つの現象の背後に、共通の原因となる潜在的な変数を想定することができる」場合
- 「2つの現象の間に、一方が原因となって他方を規定しているという形で因果関係が認められる」場合

2変数の関係 関係の強さを示す指標



- 共分散 $COV = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$



強い正の共変関係

$COV > 0$

弱い正の共変関係

$COV > 0$

共変関係なし

$COV = 0$

弱い負の共変関係

$COV < 0$

2変数の関係

共分散と相関係数



- 共分散は変数の偏差の大きさ(単位)に影響を受ける
- 共分散が大きくても、それが共変動が大きいいためなのか、それとも一方(あるいは両方)の変数の標準偏差が大きいためなのか、区別がつかない
- そこで相関係数の導入
- 共分散を2変数の標準偏差で割る

$$r = \frac{\sum(x_1 - \bar{x})(y_1 - \bar{y})}{\sqrt{\sum(x_1 - \bar{x})^2 \sum(y_1 - \bar{y})^2}}$$

2変数の関係 相関係数



相関係数の 絶対値の範囲	相関のある無し
0 r 0.2	ほとんど相関がない
0.2 r 0.4	やや相関がある
0.4 r 0.7	かなり相関がある
0.7 r 1	強い相関がある

- 2つの変数に相関があるのは、
 1. 2つの変数の背後に共通の原因がある
 2. 2つの変数の間に因果関係がある

2変数の関係

相関関係から因果関係へ



- 因果関係が成立する基本条件
- 変数 x が変数 y の原因となるために必要とされる条件
 - 時間的先行性
 - x と y の関連の強さ
 - 関連の普遍性
 - 関連の意味的な整合性

回帰モデル モデルの条件

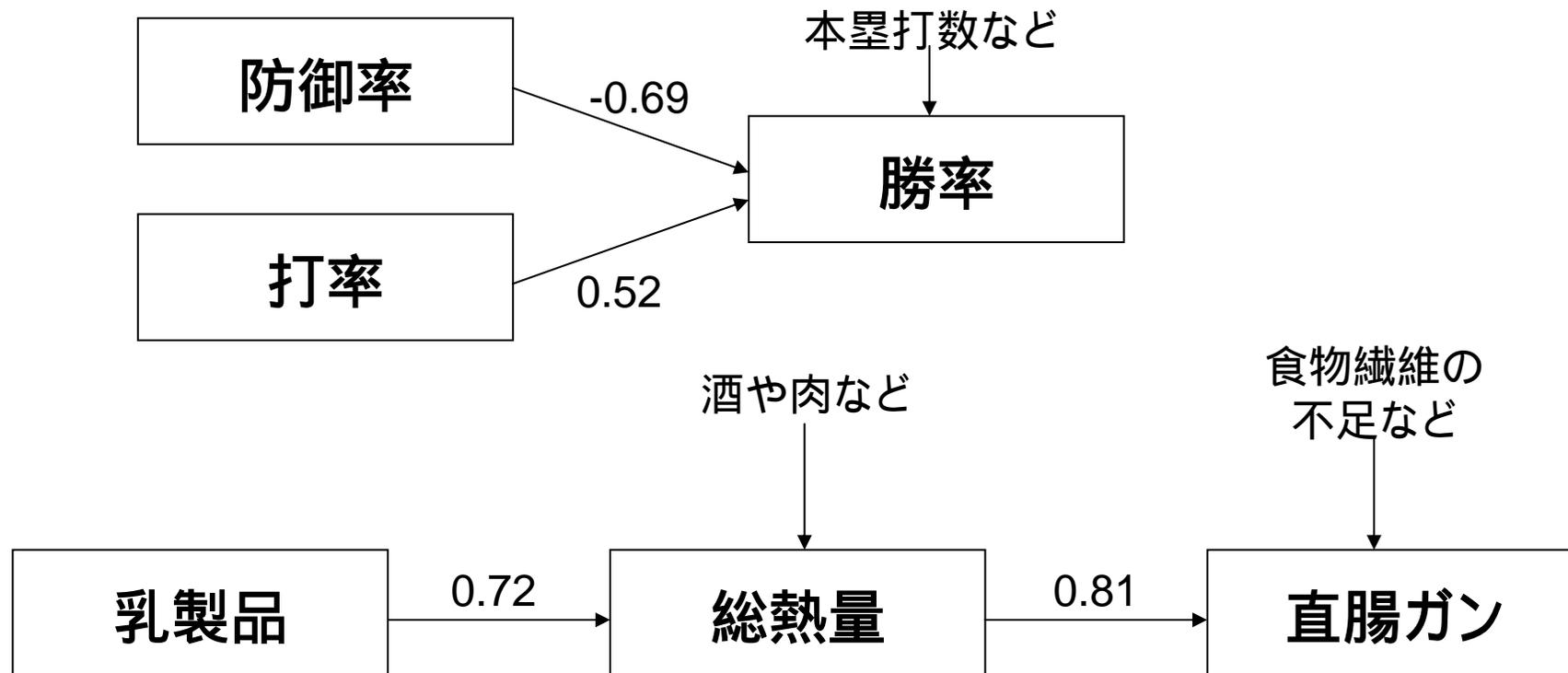


- 因果関係を明らかにするためには、まず分析者の側に、現象に対する**仮説**があり、その仮説を矛盾なく説明できる**理論**があることが前提となる。
- その上で統計的な分析を行った結果が、その理論と合っていることを確認する。
- これは、回帰モデルにも、共分散構造モデルにもあてはまる。

回帰モデル 例



- 回帰分析に関する説明は省略



(豊田、1992)

共分散構造モデルの導入 回帰モデルの限界



- 回帰モデルでは、原因となるいくつかの変数と結果となる1つの変数の間の因果関係を扱う。
- しかし、多数の原因が複雑にからみあって結果を規定している現象は多い。
- また、結果の側に複数の変数が考えられる場合もある。
 - 結果となる変数が直接観測できない「構成概念」であり、構成概念を定義するために利用する変数の候補がいくつもある場合など。

共分散構造モデルの導入 潜在変数を導入する



- 構成概念を扱う場合には、「潜在変数」と呼ばれる変数を用いる。
- 潜在変数とは、観測される多くの現象(変数)の背後に潜み、それらの現象に影響を与えている要因で、目に見えない仮説的な変数のこと。
- それに対し、実際に値が観測される変数のことを「観測変数」と呼ぶ。

潜在変数を導入すると、内容の似通った観測変数をまとめて1つのものとして扱える。

観測変数をまとめあげ、潜在変数の間の因果関係を考えると、多数の変数間の因果関係も検討しやすくなる。

共分散構造モデルの導入

共分散構造モデルとは



1. 潜在変数を導入することによって、類似した傾向を示す観測変数をまとめることができる。
 2. 潜在変数の間で因果関係を検討すれば、多くの変数の間を直接扱うより効率が良い。
という考えをもとにして、潜在変数間の因果関係を検討するための統計的手法である。
- つまり、因果関係に関する2つの考え方を同時に適応した手法。

共分散構造モデルの導入 下位モデル



- 共分散構造モデルとは多変量解析の色々な手法を統合したモデルといえる。

測定方程式だけを用いたモデル

- ・(確認的)因子分析
- ・分散成分の推定モデル
- ・主成分分析
- ・他方法他特性行列の分析
- ・古典的テストモデル
- ・一般化可能性係数の推定モデル
- ・ワイナー・シンプレックス・モデル

測定方程式と構造方程式を用いたモデル

- ・MIMICモデル
- ・**多重指標モデル**
- ・PLSモデル
- ・高次因子分析
- ・シンプレックス構造モデル
- ・重判別分析
- ・正準相関分析
- ・数量化 類
- ・サーカムプレックス構造モデル
- ・パネル・データの分析

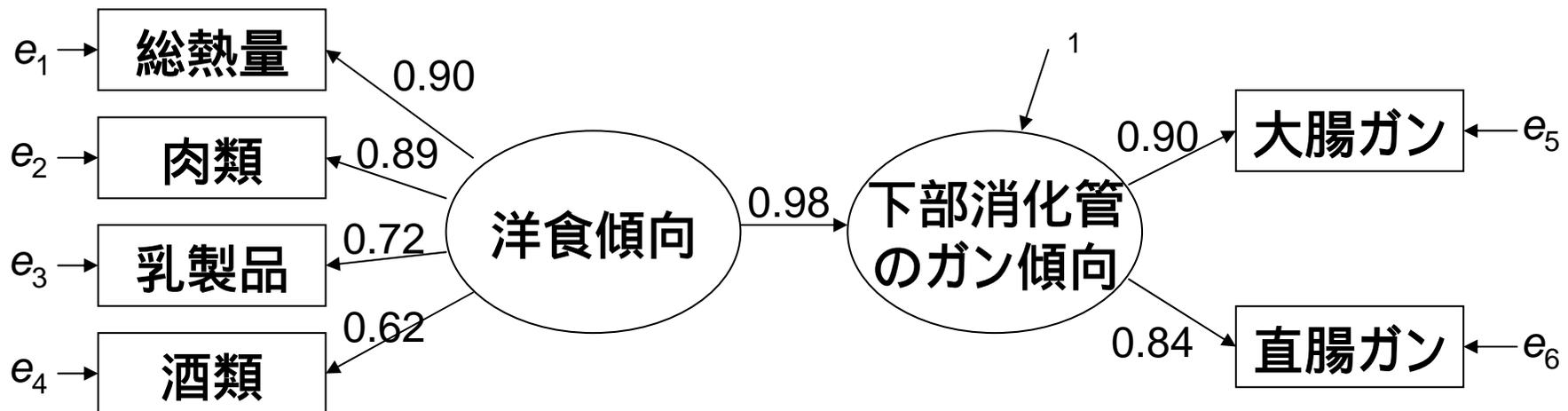
構造方程式だけを用いたモデル

- ・パス解析
- ・**単回帰分析**
- ・**重回帰分析**
- ・同時方程式モデル
- ・多変量回帰分析
- ・**分散分析**
- ・共分散分析
- ・多変量分散分析
- ・多変量共分散分析
- ・判別分析
- ・数量化 類
- ・数量化 類

共分散構造モデルの導入 多重指標モデル



- 「構成概念」を潜在変数の名前として定義。
- 仮説と理論が必要
- 食物摂取とガンの例



(豊田、1992)

共分散構造モデルの導入 パス図の書き方



観測変数は**四角形**で囲む

潜在変数は円または**楕円**で囲む

誤差変数は囲まない

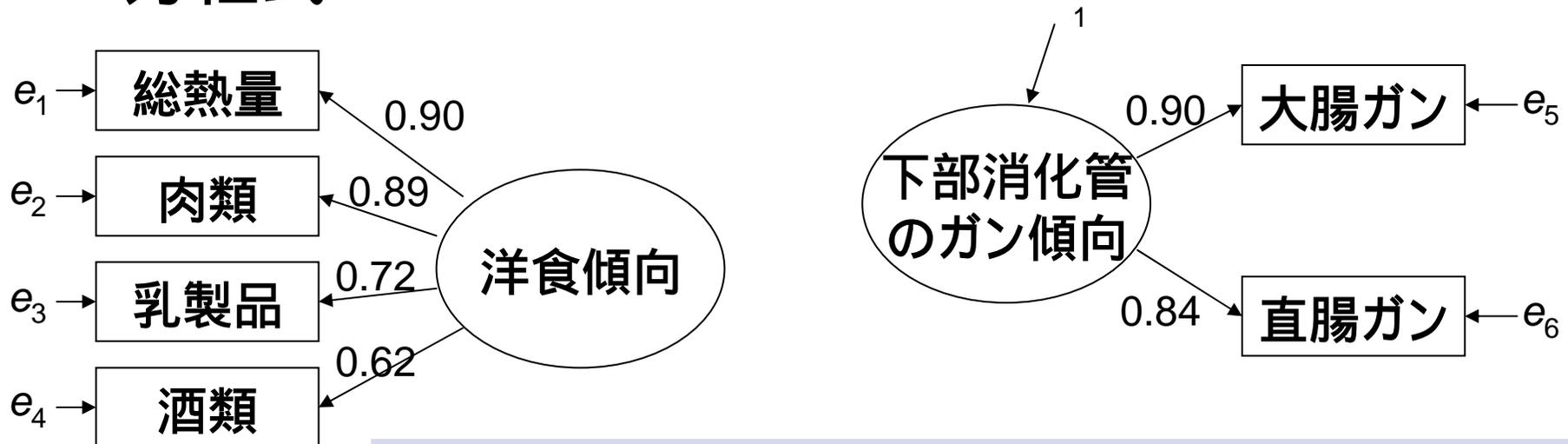
因果的な影響を与える変数から与えられる変数に**単方向の矢印**を書き、矢印に因果関係の**影響力を示す数値**を付与する

共変動を示す2つの変数に因果関係を仮定しないときには**双方向の矢印**を書き、矢印に**共分散(または相関)**を示す数値を付与する

共分散構造モデルの導入 測定方程式



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

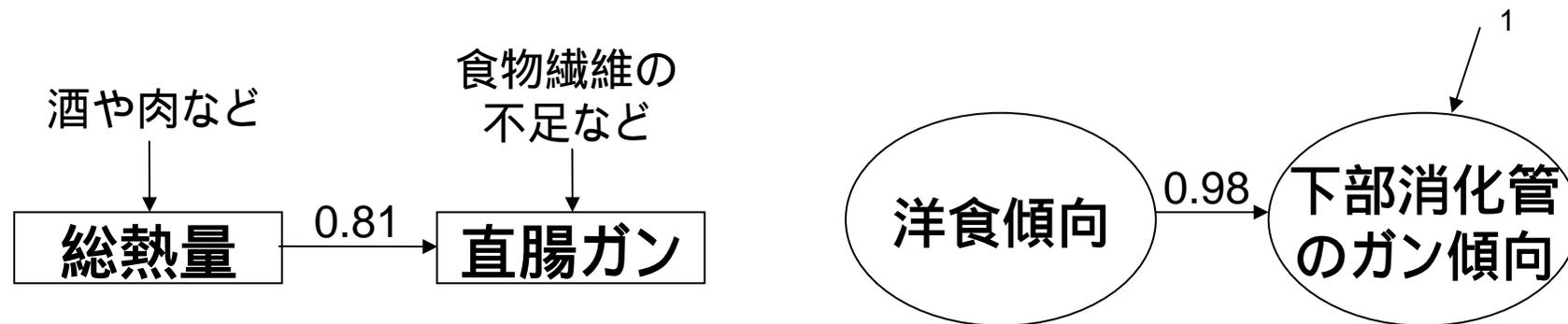


観測変数 = 因果係数 × 潜在変数1 [+ 因果係数 × 潜在変数2 ···] + 誤差変数

共分散構造モデルの導入 構造方程式



- 変数間の因果関係を表現するための方程式
「潜在変数が別の潜在変数の原因になる」
「観測変数が別の観測変数の原因になる」
「観測変数が潜在変数の原因になる」



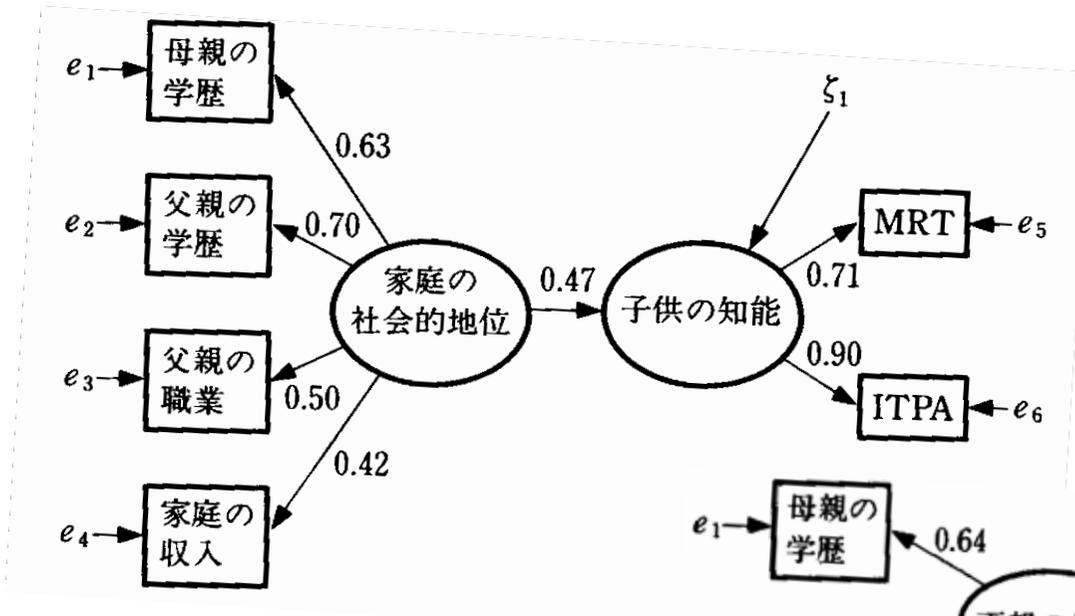
(豊田、1992)

共分散構造モデルの導入

共分散構造分析の例

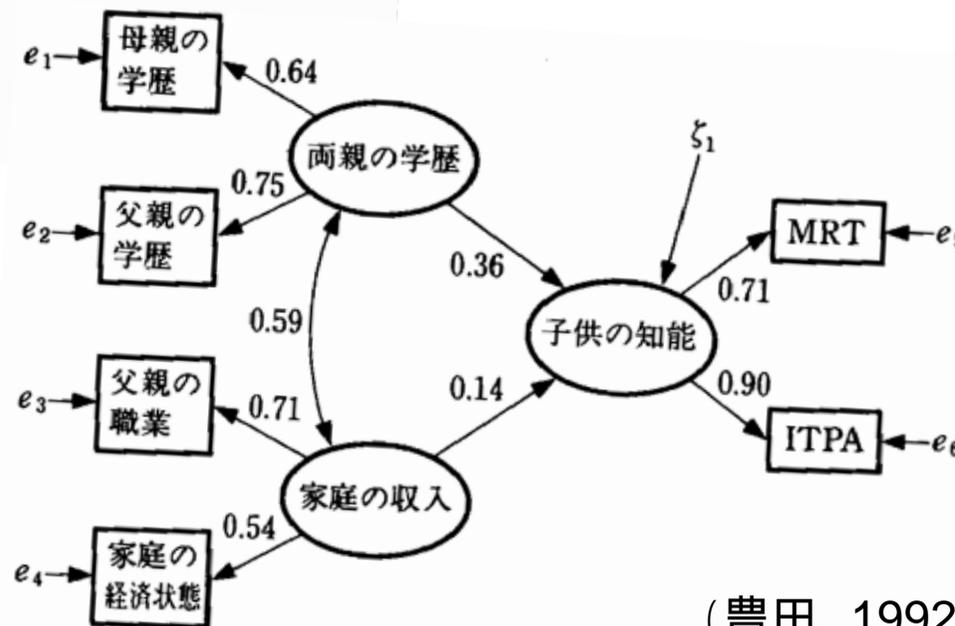


- 家庭環境と子供の知的能力の関係



分析者の仮説によっては
同一の観測データから
複数のモデルが考えられる

モデルの比較評価により
最適なモデルを探索できる

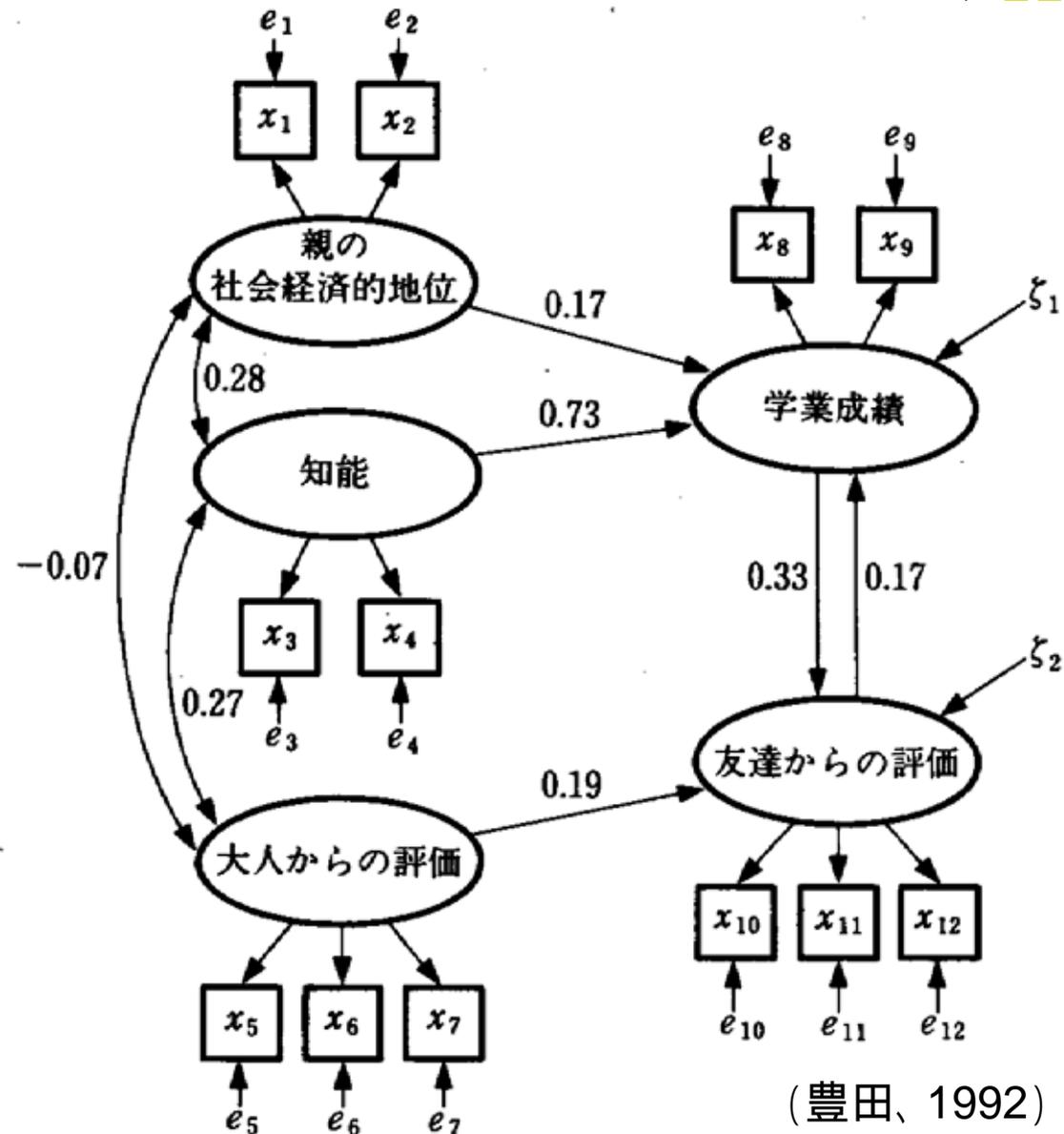


(豊田、1992)

共分散構造モデルの導入 共分散構造分析の例



- 双方向の因果関係を含んだモデル



(豊田、1992)

モデルの評価

モデルの全体の適合性の点検



- χ^2 検定
- 「モデルが正しい」ことを帰無仮説とする
- 5%の有意さではモデルと観測データが等しいという仮説は棄却されない
モデルは観測データに適合
- モデルが正しい可能性を示してはいるが、積極的にモデルを支持しているわけではない

モデルの評価

モデル全体の適合性の点検



- 適合度指標GFI,AGFI
 - GFI (Goodness of Fit Index) : 適合度指標
 - 0から1までの値をとり、分析者が構成した因果モデルが、データを何パーセント説明したかという、モデルの「説明力」の目安。1に近いほど説明力があるモデルと判断。
 - しかし、統計モデルは、母数の数を増やして複雑なモデルを構成すると、モデルの見かけ上の説明力が上昇するという一般的な傾向がある。
 - AGFI (Adjusted GFI) : 調整済み適合度指標
 - GFI AGFIという関係
 - GFIに比べてAGFIが著しく低下する場合は、あまり好ましいモデルではない

モデルの評価

モデルの部分ごとの点検



- 個々の母数の有意性 (t検定)
 - 推定された母数のうち、測定方程式の影響指標や構造方程式の因果係数は、0に比べて十分に大きい値でなければ意味がない
 - ここで「t検定」を用いる
 - 「母数の推定値が0である」という帰無仮説を棄却することで有意性を確認
 - 構造方程式の因果係数に関する部分の有意性に要注意

モデルの評価

因果の規定力と有意性



- 決定係数
 - t検定は因果係数が0でないことを保証するだけで、その因果係数が十分に大きな因果的規定力をもつことを保証するわけではない
 - 結果となる変数の決定係数があまりにも小さいモデルは、原因としてモデルに導入した構造変数(原因となる変数のうち誤差変数以外)では結果変数の変動が予測できないことになる

モデルの評価

複数のモデルの比較



- 情報量基準AIC
 - AIC (Akaike's Information Criterion)
 - 「説明力」と「安定性」を統合的な立場から評価する
 - 最尤推定法によって母数を推定した「統計モデルの良さ」を図る指標
 - 小さいほどデータに対するあてはまりが良い
 - 複数のモデルを比較する際の、モデルの相対的な良さを評価するための指標

Amosを使った共分散構造分析

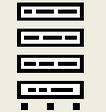
Amosとは



- spss社の提供する、共分散構造分析を実行できる統計ソフト
- 仮説のモデルを数式ではなく、「パス図」で視覚的に表し、同時に分析することができる
- 食物摂取とガンの例を実際にAmosを使って分析することで使い方を学んでみましょう

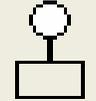
Amosを使った共分散構造分析 演習



- 「発ガンデータ.sav」をデスクトップに保存
- 「スタート」から「Amos Graphics」を起動
- データファイルの指定
で「発ガンデータ.sav」を選択
- でデータファイル内の変数を表示
- 表示された変数のリストからドラッグ&ドロップで変数(観測変数)を枠内に描画
- スライド17の図を目標にパス図を作成

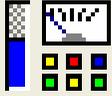
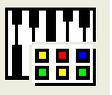
Amosを使った共分散構造分析 演習



- パス図作成ツールの紹介
-  潜在変数を描く
-  誤差変数を描く (変数の上でクリック)
-  因果関係の方向を描く
-  オブジェクトを移動する (ドラッグ&ドロップ)
-  オブジェクトを削除する (変数の上でクリック)

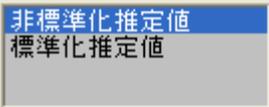
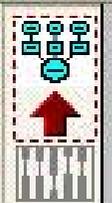
Amosを使った共分散構造分析 演習



-  で、潜在変数 (洋食傾向、下部消化管のガン傾向) と誤差変数 (e1~e6,e7) に名前をつける
-  で、「洋食傾向」から「総熱量」の に係数「1」を設定する
-  で、「下部消化管のガン傾向」から「大腸ガン」の に係数「1」を設定する
-  から、「出力」の「標準化推定値」と「重相関係数の平方」にチェックを入れる
-  で分析開始！

Amosを使った共分散構造分析 演習



- 「XX:モデル番号1」が「OK:モデル番号1」になれば分析完了
-  で、「標準化推定値」を選択し  を選択すると・・・スライド17と同じ係数が現れるはず！
-  で分析結果のテキスト出力を表示
- 「モデルについての注釈」 「 χ^2 値」「有意確率」
「モデル適合」 「GFI」「AGFI」「AIC」
「推定値」 「検定統計量」「確率」(t検定結果)
をそれぞれ確認



参考文献

- 原因をさぐる統計学 共分散構造分析入門
講談社: 豊田秀樹、前田忠彦、柳井晴夫 (1992)
- SPSS完全活用法 共分散構造分析 (Amos)
によるアンケート処理
東京図書: 田部井明美 (2001)
- 日本統計学会チュートリアルセミナー
共分散構造分析
狩野裕、市川雅教 (1999)
- 過去の「心理データ解析演習」発表資料