

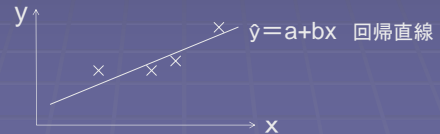
## 重回帰分析からパス解析へ

心理データ解析演習 2005/06/22

発表者 教育学研究科 教育認知心理学講座  
M1 常深 浩平

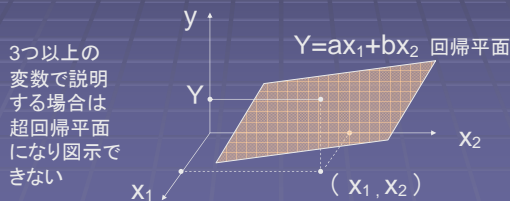
### 復習: 回帰分析 ①

- 栗本君の発表を思い出しながらお聞き下さい
- 回帰分析とは→ある変数の変動をいくつかの変数の変動によって、説明・予測するための統計手法
- 単回帰分析: ある変数をひとつの変数で説明・予測
  - $\hat{y} = a + bx$  ( $\hat{y}$ : 予測値,  $a$ : 切片,  $b$ : 回帰係数)



### 復習: 回帰分析 ②

- 重回帰分析
  - ある変数を複数の変数  $x_1, x_2, \dots, x_p$  で説明・予測 (下図は2つの変数で説明・予測する例)



### 重回帰分析の使いどころと利点

- どんな時に重回帰分析は力を発揮するのか
  - 沢山の 변수がある一つの変数を説明・予測すると仮定されるとき
  - 他の変数の影響を取り除いた上で、個々の目的変数と説明変数の間の関係を見ることが出来る → **偏回帰係数**
  - 加えて、変数間の寄与率の差異を **標準化偏回帰係数 ( $\beta$ )** という形で比較することが出来る

### 傾向と対策

- 一般に、説明変数の数を増やすと、分析に用いたデータへの当てはまり ( $R$ ,  $R^2$ ) は良くなるが、将来の観測値に対する予測力は必ずしも向上しない
- 多重共線性の問題が起こる可能性も(単純に)上がる
- そこで、**説明変数の取捨選択**を行う方法がいくつかある

### 説明変数の取捨選択法 ①

- SPSSでは以下が実装されている
  - 強制投入法
  - (強制除去法) 割愛
  - 変数減少法
  - 変数増加法
  - ステップワイズ法
- 以下、それぞれについて簡単に説明

### 説明変数の取捨選択法 ②

- 強制投入法
  - 全ての変数を投入する(変数の取捨選択なし)
- 変数減少法
  - 最初に全ての変数を投入してモデルを作った後、最も予測に有効ではない(目的変数との偏相関が低い)変数をモデルから除去、予測精度(R・R<sup>2</sup>)の低下が一定程度以内である間、その作業を続ける

### 説明変数の取捨選択法 ③

- 変数増加法
  - まず最も予測に有効な説明変数一つモデルに取り入れ、次に残りの変数の中で最も予測に有効な変数一つ取り入れる、予測精度が一定程度上昇する間、その作業を続ける
- ステップワイズ法
  - 変数増加法に似るが、先に採用された説明変数も、後で採用された変数との関係で不要になる場合があるので、新たな採用の際に、すでに採用された変数を取除くかどうかをチェックする

### 説明変数の取捨選択法 ④

- SPSSでは、変数を投入する際に幾つかのブロックを設定することも可能なので、ブロックごとに前述の変数取捨の方法を変えることも出来る
  - 例えば、取捨選択すると落ちてしまう変数をどうしてもモデルに入れたい→その変数だけブロックを変えて、強制投入法にする等
- ただし、そのような選択がどのような根拠の下で行われるのかは考えるべきである

### 更なる問題

- 重回帰分析では各説明変数と一個の目的変数の間の一方の因果のみを仮定している(スライド16参照のこと)
- しかし、変数間に因果関係が仮定される場合や目的変数が複数仮定される場合もある
- その場合、重回帰モデルでは説明しきれない
- そこで・・・パス解析

### パス解析とは？

- 変数間に幾つかの因果を仮定し、共分散行列や相関行列をもとに因果推論を行う統計的な分析手法
- 生物学者ライト(Wright, S.)が考案
- 利点: 重回帰分析では各説明変数から従属変数への一方の因果のみを仮定するのに対し、パス解析では変数間に自由に因果関係を仮定してモデルをつくる事が可能

### パス解析のモデル式

- $y = By + \Gamma x + e$
  - $y$ : 内生変数
  - $x$ : 外生変数
  - $e$ : 誤差変数
- 数式のごとはよく分からないが、線形モデルだということはおさえておこう
- 内生? 外生? 誤差変数?
  - 内生変数: 一度以上、因果の「果」になる
    - 必ず誤差変数を伴う
  - 外生変数: 一度も「果」にならない、完全な「因」
  - 誤差変数: 実際に測定された変数以外で結果に影響を与えていると想定される変数

### 実際に描くパス解析 [Amos] ①

- 今回はAmos (Analysis of Moment structures: モーメント構造分析) というソフトを利用する (院生室の共用パソコン・サテライトの一部には入っている)
- 観測変数
  - 実際に測定した数値
    - つまり、SPSSなどに入力されている変数
  - 内生変数や外生変数

[Amosではこう描く]



13

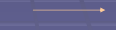
### 実際に描くパス解析 ②

- 潜在変数
  - 測定されていない変数
  - パス解析では誤差変数など
- 因果関係
- 相関関係

[Amosではこう描く]



[Amosではこう描く]

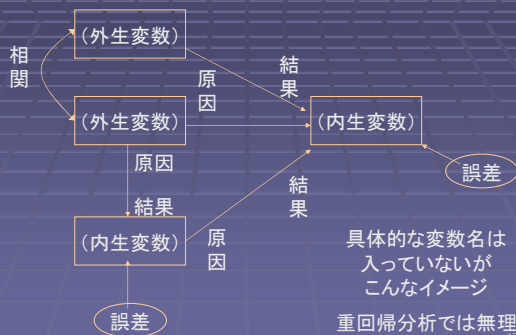


[Amosではこう描く]



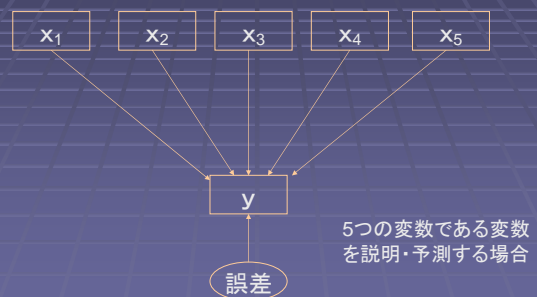
14

### ここまでの簡単なまとめとしてのパス図



15

### 参考に 重回帰モデルを表すパス図



16

### ところで

- よく似てるけれど、パス解析と**共分散構造分析** (木戸さんの発表参照) の違いは何如?
- 過去には、パス解析は重回帰分析を内生的観測変数の数だけ繰り返してパス図を描いたが、正確な数値が出ない、モデルの適合度を計算できない、モデルの自由度が少ないなどの問題があった
- 現在では技術の進歩によってパス解析の方法は変化し、**共分散構造分析の特殊なケース**として考えられている
- その上で、パス解析は観測変数間の因果モデル、共分散構造分析は潜在変数間の因果モデルという点に違いがある

17

### モデルの評価 ①

- 重回帰分析と同様に、モデルの評価のための基準が存在する
- AMOSでは以下のような基準が出力される
  - $\chi^2$ 値、GFI、AGFI、AIC、CFI、RMR、RMSEA等 (添付資料参照)
- なお、これらは共分散構造分析モデルの評価基準であるが、前述の理由からAmosではパス解析と共分散構造分析を区別しないのでパス解析の基準としても用いることが可能

18

## モデルの評価 ②

- $\chi^2$ 検定
- $\chi^2$ 値
  - 帰無仮説:「構成されたモデルは正しい」を検定
  - $\chi^2$ 値が有意であれば採択されない
  - モデルが正しい可能性を示しているだけで、積極的にモデルを支持しているわけではない
  - データ数が大きいときは殆ど確実にモデルを棄却してしまう

19

## モデルの評価 ③

- GFI (goodness of fit index) :モデル内のデータの何%説明できるのか
  - .90以上か .95以上なら当てはまりが良いモデルと判断
  - 複雑なモデルだと値は上がるが安定性が下がる(説明はできるが予測には向かない)
- AGFI (Adjusted GFI) :GFIから不安定度を割引して示す指標
  - 値が1に近いほどデータへのあてはまりが良い
  - GFIに比べAGFIが極端に低下する場合は注意

20

## モデルの評価 ④

- 赤池の情報量基準: AIC (Akaike's information criterion)
  - どの程度予測力(≠説明力)があるかの指標
  - 値が小さいほど良いモデル
- 比較適合度指標: CFI (comparative fit index)
  - 独立モデル(観測変数間に相関が無いと仮定するモデル)と比較
  - .90以上か .95以上なら当てはまりが良いと判断

21

## モデルの評価 ⑤

- 残差平方平均平方根(Root Mean square Residual: RMR) : 値が0に近いほど、モデルがデータにうまく適合している
- 平均二乗誤差平方根 (root mean square error of approximation : RMSEA) :モデルの分布と真の分布との乖離を1自由度あたりの量として表現した指標
  - 一般に0.05以下なら当てはまりが良く、0.10以上なら当てはまりが悪いと判断

22

## モデルの評価 ⑥

- モデルの部分評価
- t 検定
- パス係数の数値が0にくらべて(絶対値で)十分に大きい値である必要がある
- 0に近ければ、2つの変数間の関係が「ない」ということになる
- その係数が有意であるかどうかを検定する際に、t 検定を用いる

23

## 留意点

- パス解析では、モデルによって表現された因果関係が存在したと仮定できる場合に、その規定力を推定できるのであって、パス解析によって直接的に因果関係の存在が確認されるものではない
- よって、やはり仮説としてのモデル構築とそれに基づいた変数の測定が重要になってくる
- また、パス解析ではモデル同士を比べるための指標が豊かであるため、仮説に基づく複数のモデルを比較検討できる利点も活用したい
- 重回帰分析同様、交互作用項を作るなどしないと、基本的には交互作用を検討できない



24

## 実際にやってみました

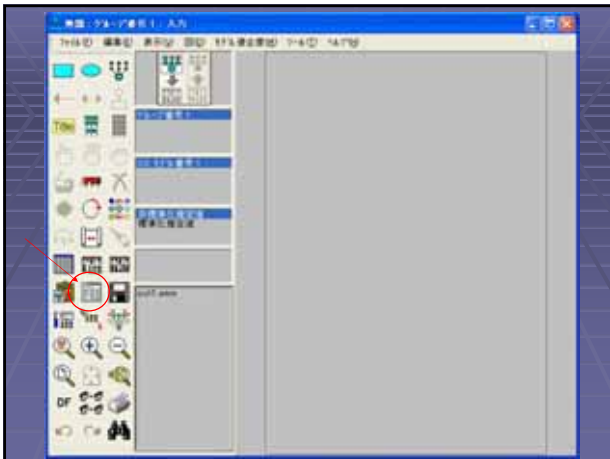
- 文章の読解時間を説明するモデル
- 音節(モーラ)数や文番号(何文目か)、名詞の新出や重複など沢山の説明変数が仮定される
- また変数が互いに影響しあっている可能性が高いが、それぞれの寄与率を独立に見たい
- AmosがすべてのPCに入っているわけではないので、今回はデモンストレーション中心

25





## 簡略に手順を記述 ①

- まず、パス解析を行いたいデータをSPSSに打ち込む
- データを保存してSPSSを閉じる
- Amos Graphicsを起動する
-  こんなアイコン
-  を押して、SPSSのデータファイルを読み込む

26





## 簡略に手順を記述 ②

- 次に、 を押して読み込んだデータ内の変数一覧を表示する
- 一覧表からドラッグ&ドロップで観測変数を配置できる( を押して四角を描き、左クリックでプロパティを開いて、自分で変数名を入力してもOK.)
-  を押して、ワードなどのオートシェイプで描く感じで因果関係を表す矢印を変数間に結ぶ(相関関係を記述したいときは  で同様に描く)

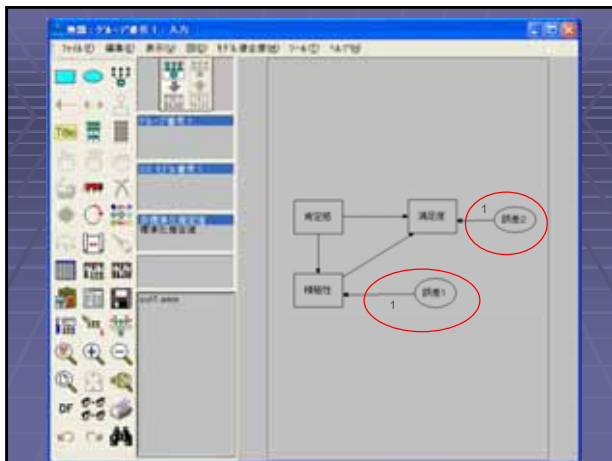
28



## 簡略に手順を記述 ③

- 内生変数には誤差変数を付けないといけないので  を押して、円を描き、プロパティで誤差1、誤差2などの名称をつける
-  で誤差変数から内生変数へ矢印を描く
- その矢印のプロパティを開き、係数を1にする

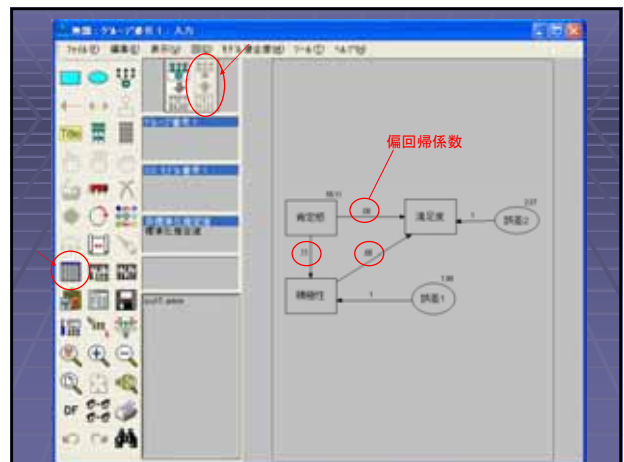
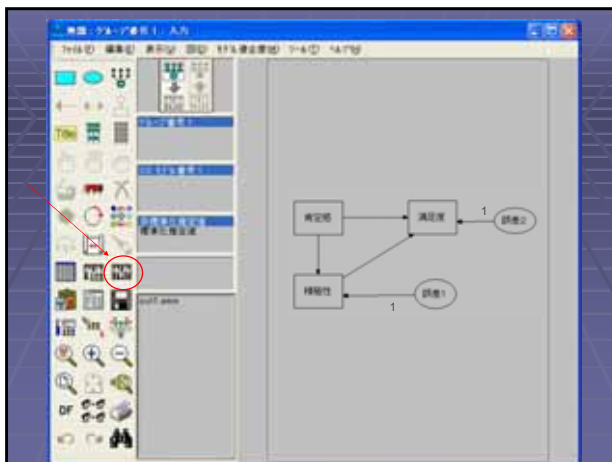
30



### 簡略に手順を記述 ④

- いよいよ計算する
- を押して、計算を開始
- を押すと、偏回帰係数がパス図に書き込まれる
- を押すとテキスト出力がされる
  - この中にモデルの評価基準などが記述されている

32



### 簡略に手順を記述 ⑤

- 後は、テキスト出力を見て適宜モデルを改変していく
  - 例えば
    - 係数が有意ではないパス(矢印)を消す
    - それに伴い必要の無くなった変数を消す
    - 再び計算をしてみて、前のモデルとモデルの評価基準を比べる、など

35

### 引用・参考文献 ①

- 楠見先生のホームページより
- 心理データ解析演習 発表資料
  - 木戸口さん(2005)
  - 栗本くん(2005)
  - 前原さん(2004)
  - 森本さん(2004)

36

### 引用・参考文献 ②

- 石村貞夫 SPSSによる多変量データ解析の手順 東京図書 1998
- 小塩真司 心理データ解析  
[http://psy.isc.chubu.ac.jp/~oshiolab/teaching\\_folder/datakaiseki\\_folder/top\\_kaiseki.html](http://psy.isc.chubu.ac.jp/~oshiolab/teaching_folder/datakaiseki_folder/top_kaiseki.html)
- 田部井明美 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 2001

37

### 引用・参考文献 ③

- 豊田秀樹 FPR 544  
<http://www.nuis.ac.jp/~mat/fpr/fpr1997/0019.html> (注: Googleのキャッシュにしか残ってませんでした2005/06/08)
- 豊田秀樹 共分散構造分析[疑問編] 朝倉書店 2003 【配布資料】 P122-125
- 中島義明・他(編) 心理学辞典 有斐閣 1999

38

### 引用・参考文献 ④

- 室淳子・石村貞夫 SPSSでやさしく学ぶ多変量解析 東京図書 1999
- 森敏昭・吉田寿夫 心理学のためのデータ解析テクニカルブック 北大路書房 1990

39