

Σ 

SSJデータ・アーカイブ 第6回公開セミナー
社会調査の2次分析シリーズ 第1回
2002年11月11日(月)13:00-15:00

共分散構造分析の基礎と実際 -----基礎編-----

狩野 裕 (大阪大学 大学院人間学研究科)



共分散構造分析とは

直接観測できない潜在変数を導入し、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。

やや具体的にかくと

研究者が想定した因果に関する仮説をモデル化する。以下の情報が得られる

- (i) モデル(仮説)の妥当性の検討
- (ii) モデル(仮説)修正へのsuggestion
- (iii) 因果の大きさ・強さの推定・検定

呼称「SEM」について

- Structural Equation Modeling
構造方程式モデル(モデリング)
 - 近年は共分散構造分析よりもメジャーな名称
 - 平均構造も分析できるので共分散構造分析では誤解を招く
 - “構造方程式”は真の因果関係を表すという意味がある

プログラム

プログラム	パス図	多母集団 の分析	離散順 序尺度 の分析	非線型 制約で の分析	LM検定・ 修正 指標	相関行列 の分析	潜在曲線 モデルへ の対応	二段抽出 モデル	交互作用モ デル	ミッシング データの 処理
AMOS4										
EQS6										
LISREL8.50										
SAS Calis									SAS syslin	
STATISTICA Sepath										
Mplus2										

基本的な3つの利用法

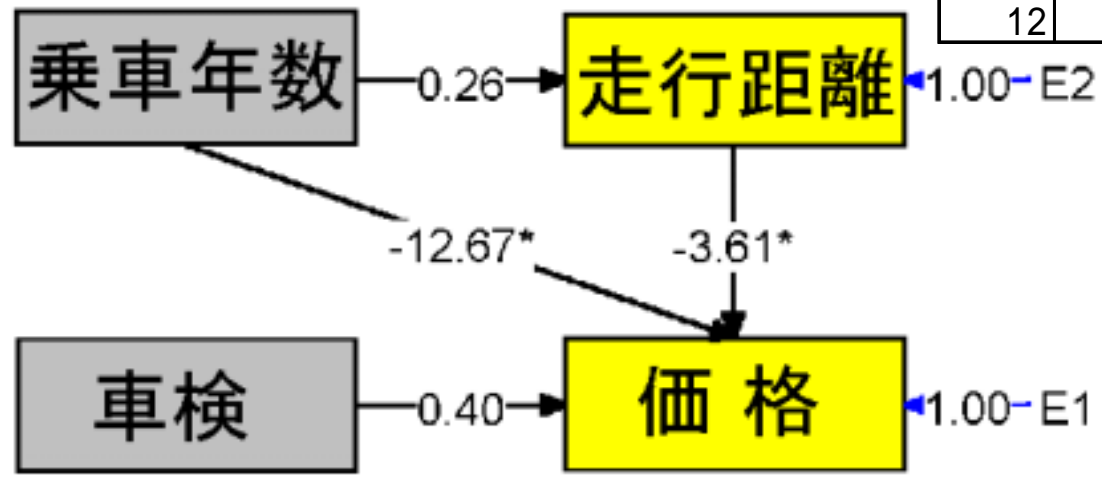
1. 調査項目間の因果関係を調べる
 - (多)重回帰分析(パス解析)モデル
2. 調査項目をまとめて単純化(潜在変数化)する
 - 因子分析モデル
3. 調査項目をまとめて単純化(潜在変数化)してから因果関係を調べる
 - 多重指標モデル(典型的な共分散構造モデル)

1. 調査項目間の因果関係を調べる

(多)重回帰分析(パス解析)モデル

中古車価格の要因分析

番号	価格	走行距離	乗車年数	車検
1	89	4.3	5	24
2	99	1.9	4	18
3	128	5.2	2	13
4	98	5.1	3	4
5	52	4.0	6	15
6	47	4.8	8	24
7	40	8.7	7	3
8	39	8.2	7	6
9	38	3.3	10	14
10	48	3.9	6	0
11	27	8.2	8	24
12	23	7.2	8	24



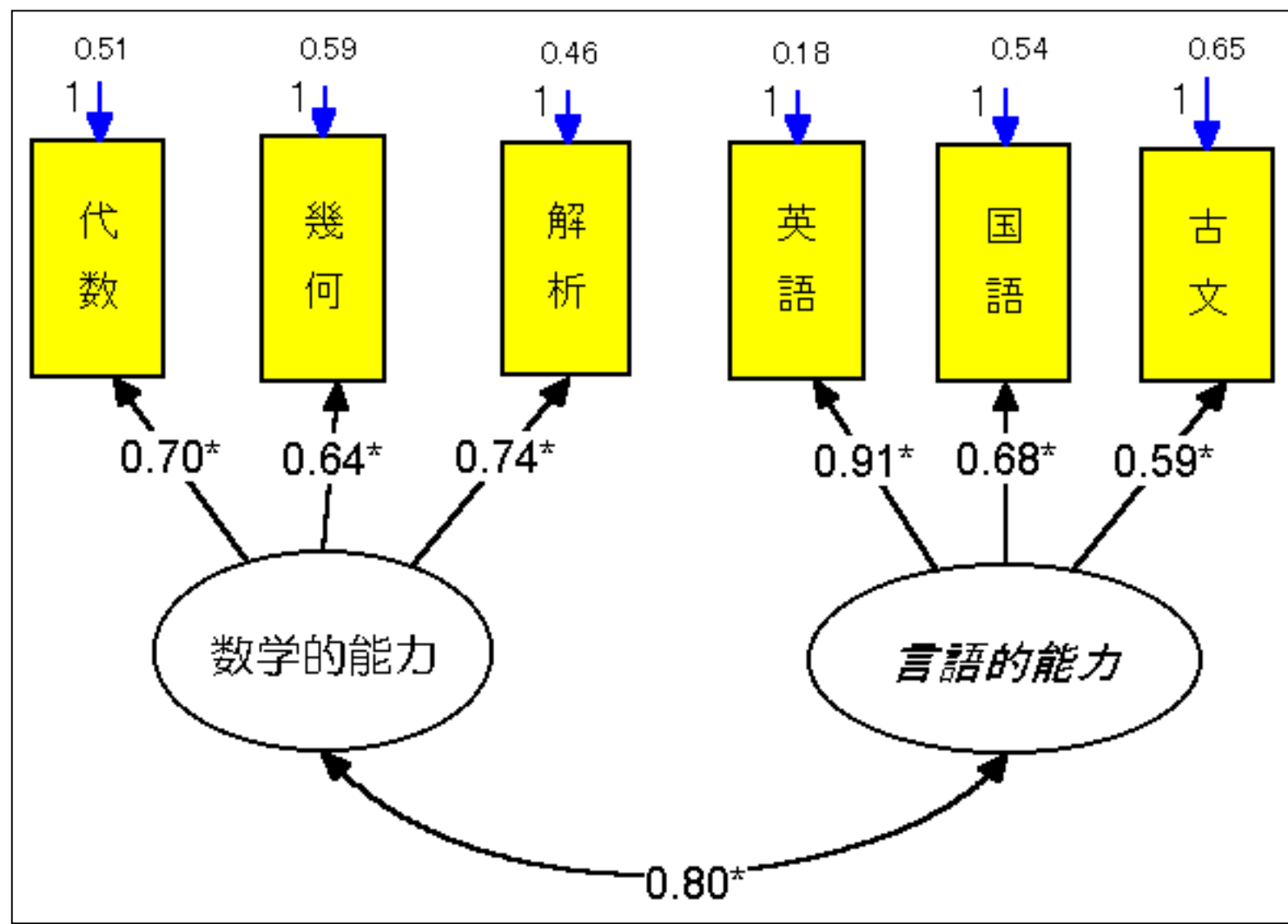
2. 調査項目をまとめて単純化 (潜在変数化)する ---因子分析モデル---

被験者	代数	幾何	解析	英語	国語	古文
1	77	82	67	67	81	74
2	63	78	80	70	81	79
3	75	72	71	66	81	52



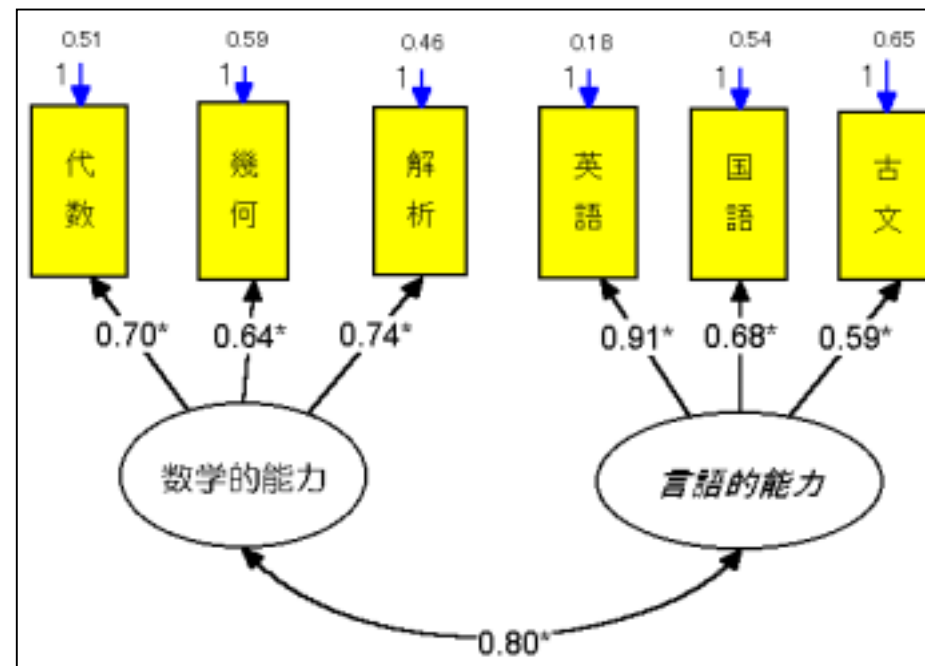
	代数	幾何	解析	英語	国語	古文
代数	1					
幾何	0.412	1				
解析	0.521	0.495	1			
英語	0.538	0.499	0.525	1		
国語	0.334	0.293	0.364	0.607	1	
古文	0.346	0.248	0.323	0.517	0.506	1

2. 調査項目をまとめて単純化 (潜在変数化)する ---因子分析モデル---



2. 調査項目をまとめて単純化 (潜在変数化)する --- 因子分析モデル ---

- 探索的因子分析との違い
 - 影響を及ぼさないと考えられるパス(因子負荷)は積極的に0に固定して推定する



潜在変数の導入(単純化)の意義

- 単純化したものは理解しやすい
 - 次元縮小
 - サイエンスは, シンプルな基本原理を見つけその原理で多くのことを説明しようとする
- (心理学などでの) 構成概念の数理モデル
- 誤差を伴ってしか測定できない状況
 - 測定道具(コスト)の問題
 - 低い相関の補正

3. 調査項目をまとめて単純化(潜在¹²変数化)してから因果関係を調べる

---多重指標モデル---



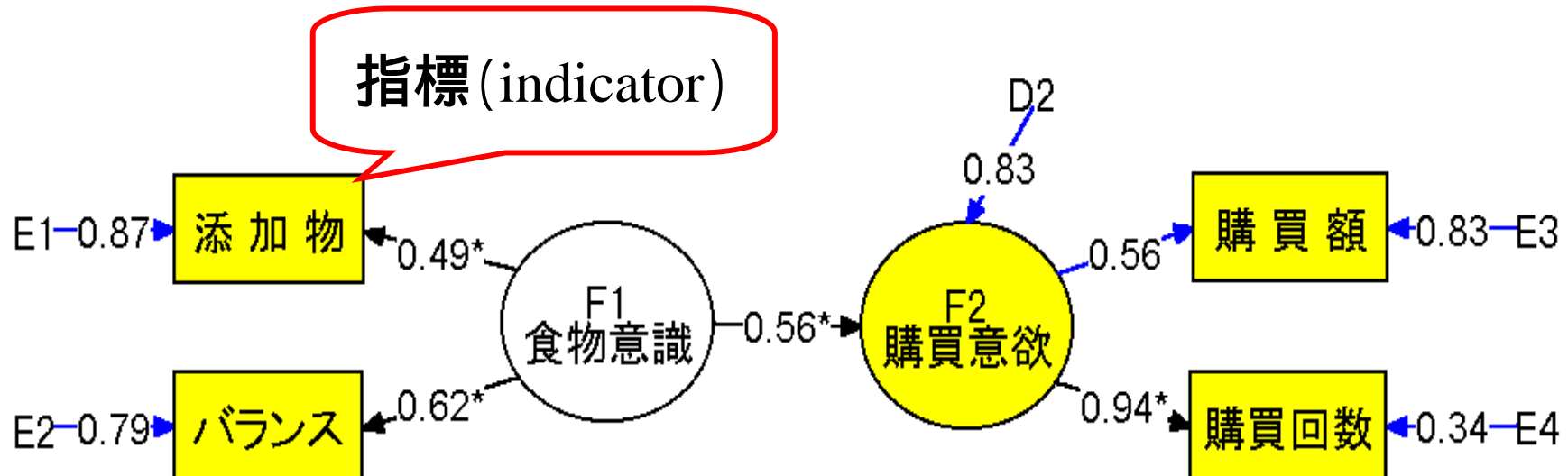
被験者	X1	X2	X3	X4
1	3	5	20	1
2	3	4	112	4
3	5	2	45	2
4	3	4	130	4
5	2	2	120	4

	X1	X2	X3	X4
X1: 食品添加物に気を使う	1			
X2: 栄養のバランスに気を使う	0.301	1		
X3: 自然食品店での購買額	0.168	0.188	1	
X4: 自然食品店での購買回数	0.257	0.328	0.530	1

3. 調査項目をまとめて単純化(潜在¹³変数化)してから因果関係を調べる

---多重指標モデル---

	X1	X2	X3	X4
X1: 食品添加物に気を使う	1			
X2: 栄養のバランスに気を使う	0.301	1		
X3: 自然食料品店での購買額	0.168	0.188	1	
X4: 自然食料品店での購買回数	0.257	0.328	0.53	1



パス図の決まり

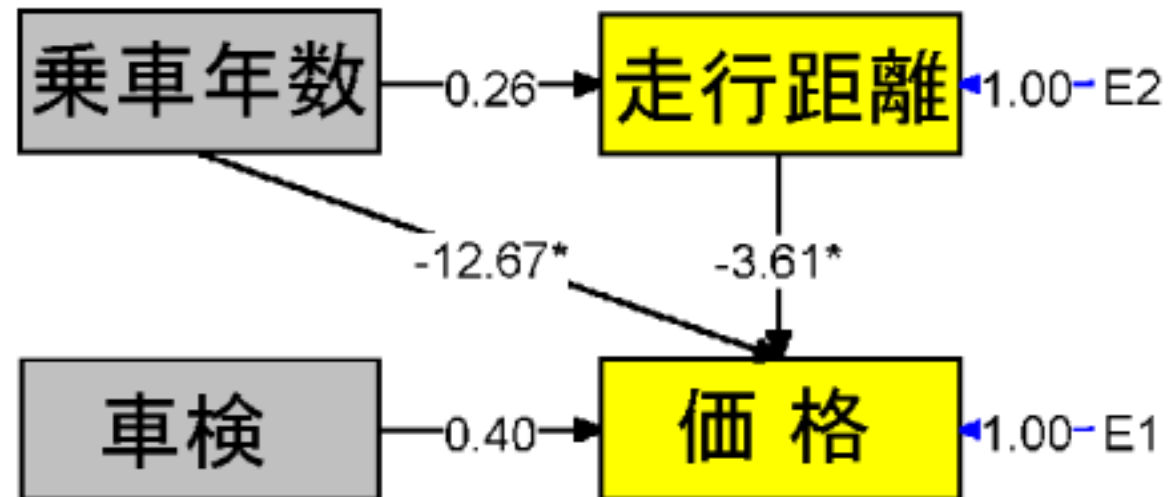
- 観測変数は四角形で囲む
- 潜在変数は円または楕円で囲む
- 誤差変数は記号のみか(楕)円で囲む
- 片方矢印は因果を表す
- 双方矢印は(単なる)相関関係を表す
- 片方矢印を受けた変数(従属変数)には必ず誤差変数が付属する
- 片方矢印の上には, パス係数(影響指標, 因果の大きさ, 強さ)の推定値が付される
- 双方矢印の上には, 相関係数または共分散の推定値が付される

3つのモデルのご利益

- 効果の分解
 - 直接効果, 間接効果, 総合効果
- 次元縮小
 - 潜在変数化
- 希薄化の修正
 - 誤差を取り除く

直接効果・間接効果・総合効果

- X1からX2へ直接のパスがあるとき,それを**直接効果 (direct effect)**という
- X1からいくつかの変数を経由してX2へつながるとき,それを**間接効果 (indirect effect)**という
- 直接効果と間接効果を併せて**総合効果 (total effect)**という



- 注:片方向の矢印だけを辿る

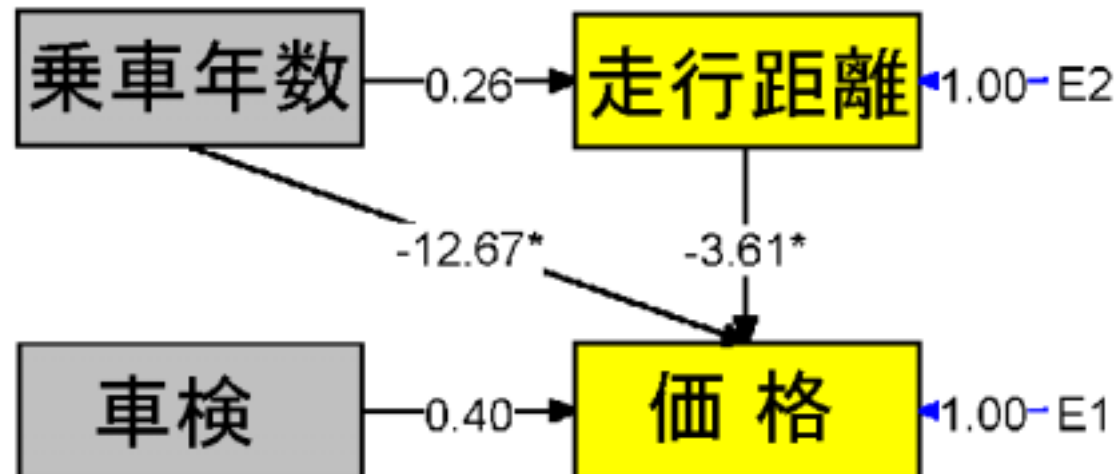
計算方法

乗車年数から価格への

直接効果 -12.67

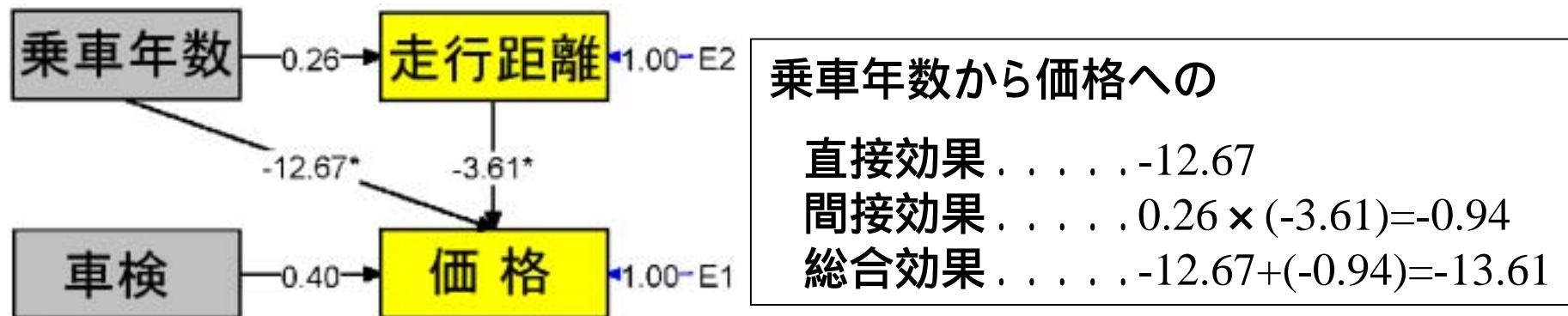
間接効果 $0.26 \times (-3.61) = -0.94$

総合効果 $-12.67 + (-0.94) = -13.61$

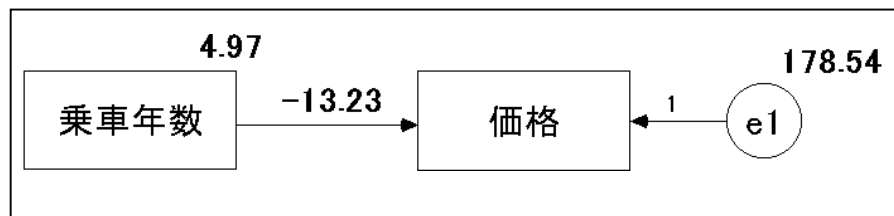
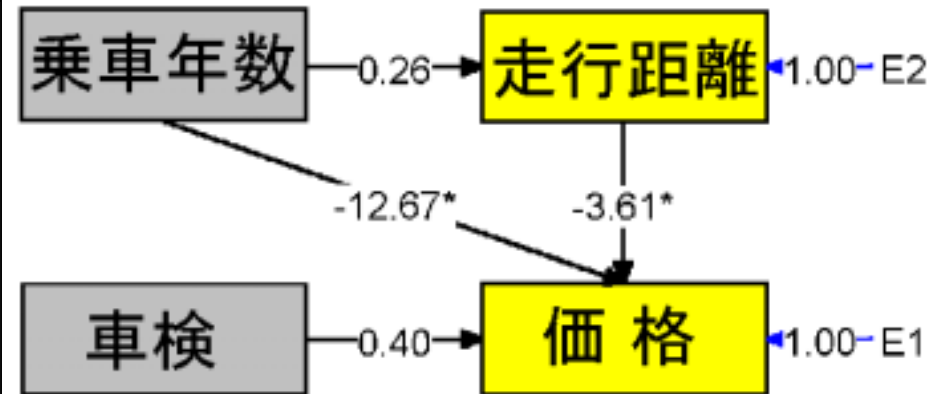
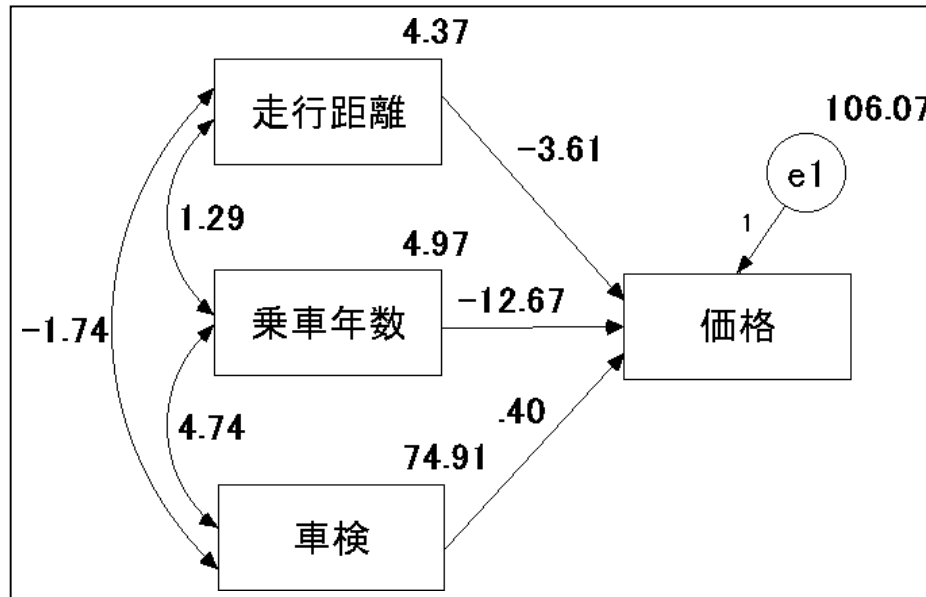


効果の分解の解釈

- 直接効果 . . . 所有することによる価値の減少
 - 走行距離が一定の下での乗車年数の効果
- 間接効果 . . . 使うことによる価値の減少



回帰分析との比較



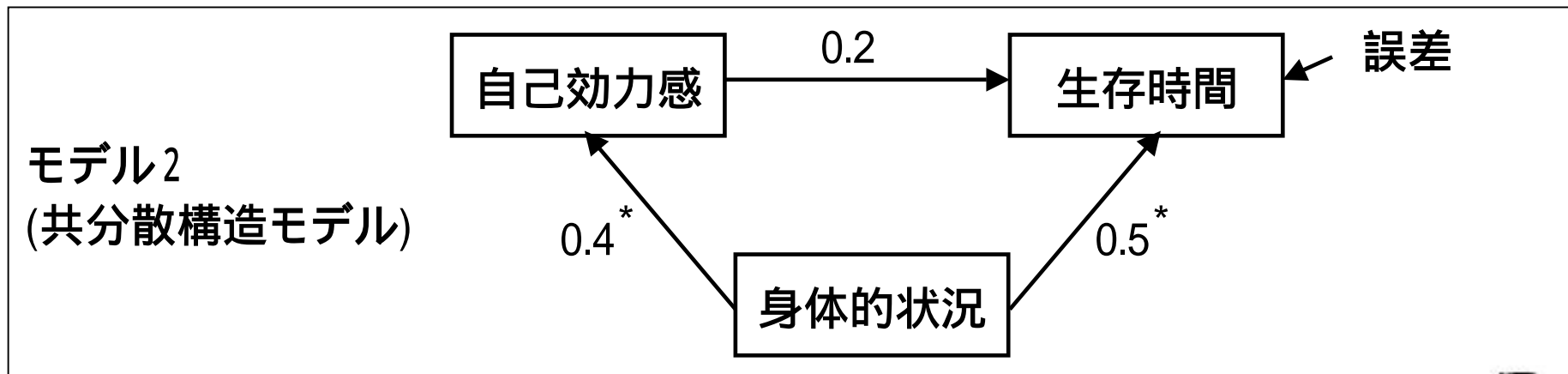
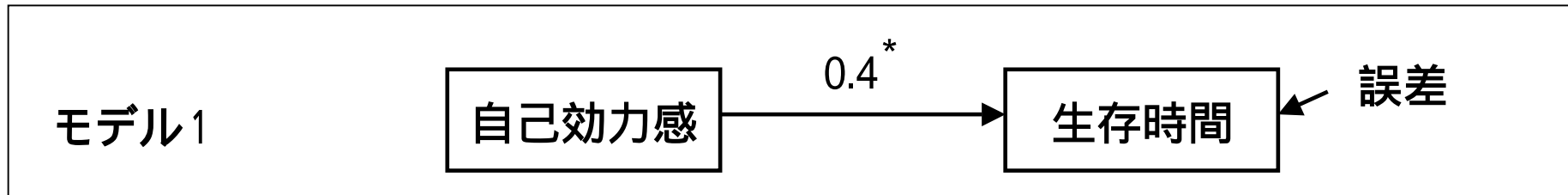
パス解析の効用

末期がん患者とSE:平井氏の研究から

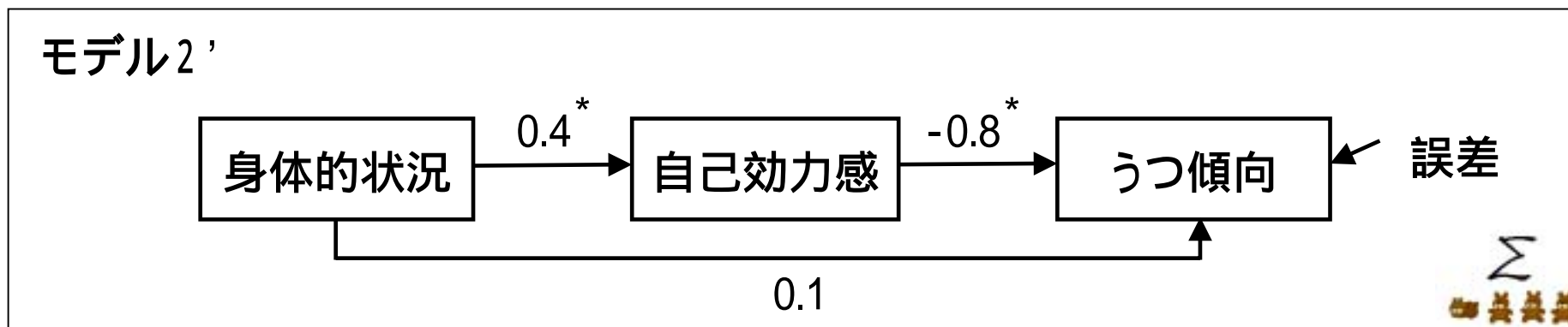
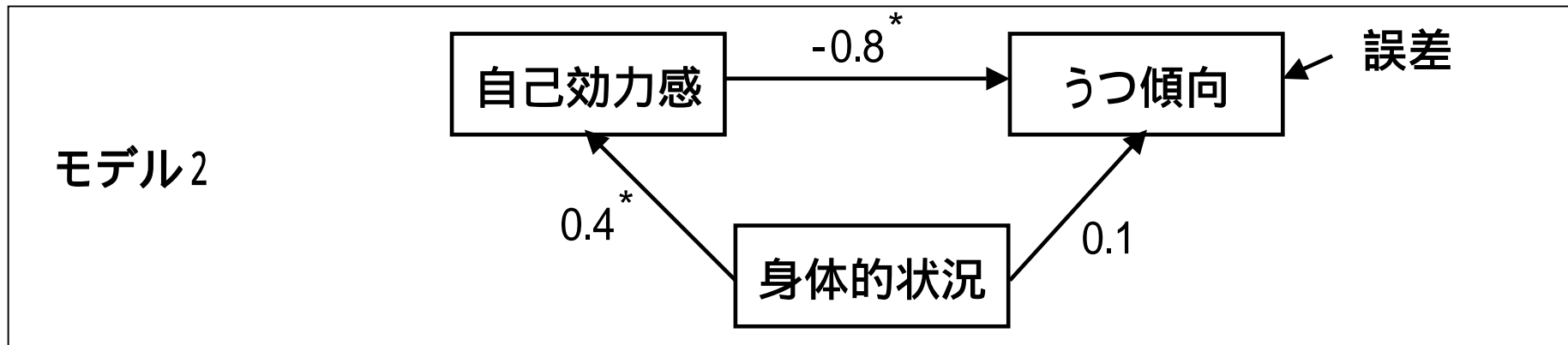
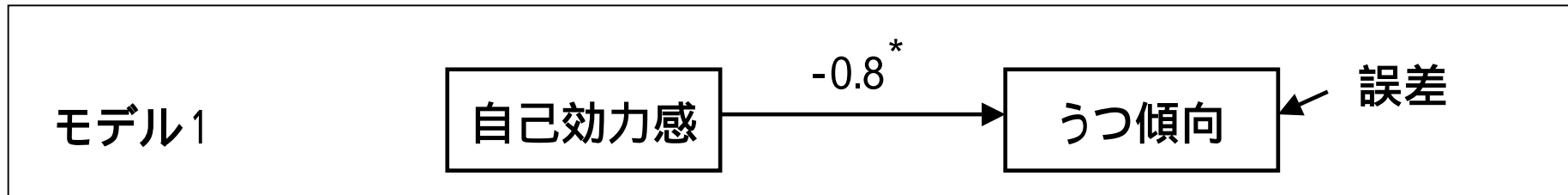
- データ
 - 末期がん患者50名

respondant	自己効力感	身体的状態	生存時間	うつ傾向
1				
2				
49				
50				

パス解析の効用：例1

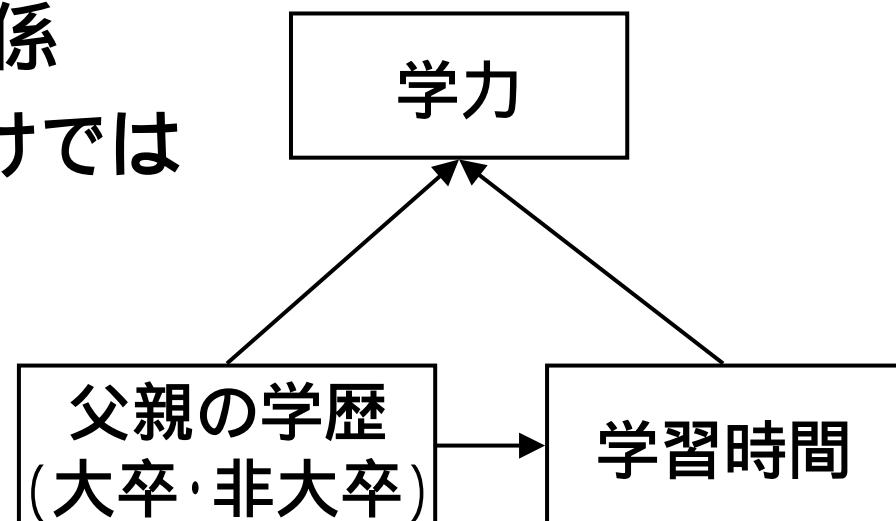


パス解析の効用：例2

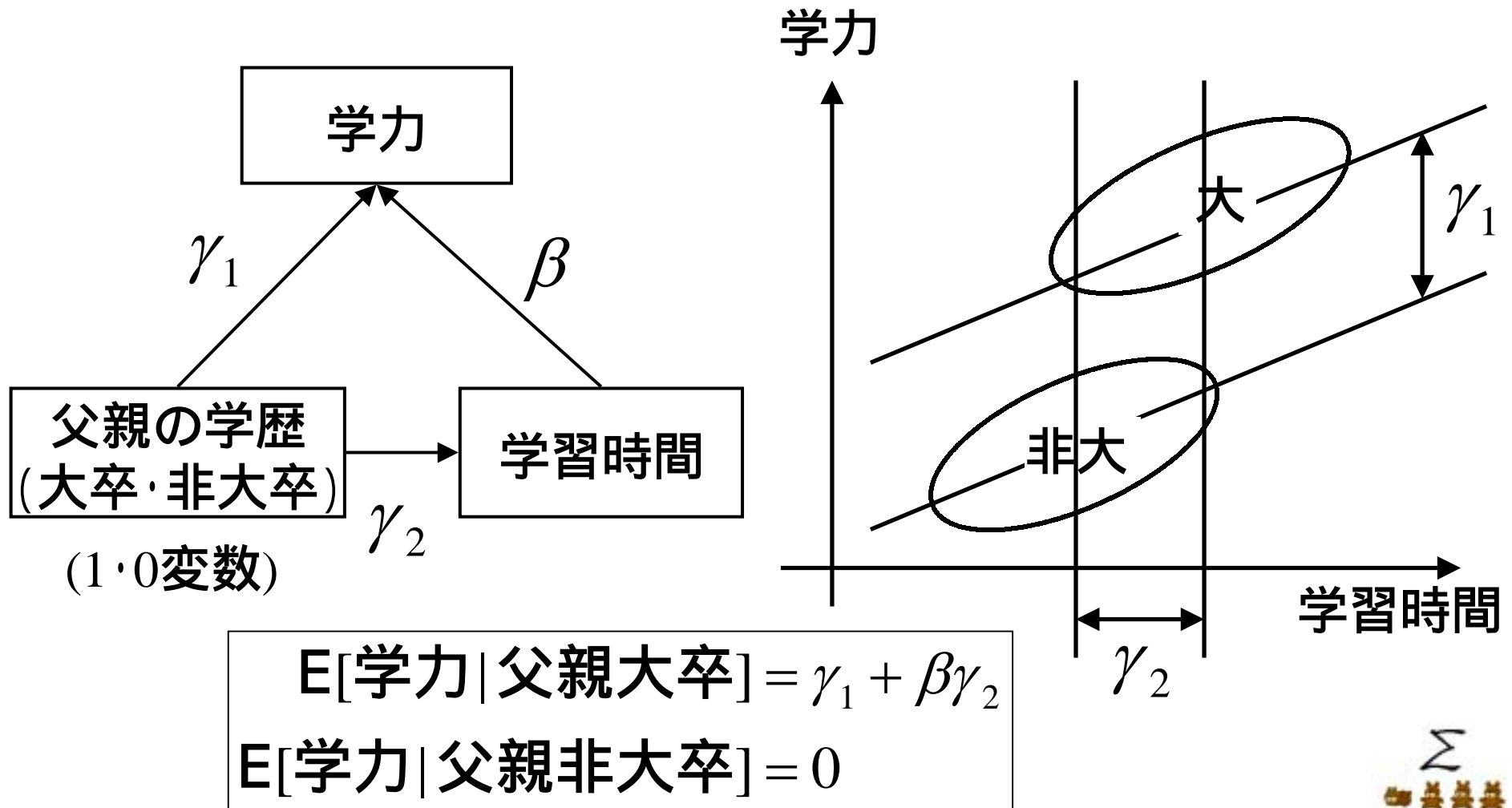


平均における効果の分解

- 出典: 2002/11/2 日経朝刊(教育)
 - 小学生の学力
 - 「学歴階層化社会」が進展
 - 学力低下の陰に格差拡大
 - 父親の学歴と密接関係
 - 機会を均等にするだけでは不平等は解決しない



学歴階層化社会



平均効果の分解

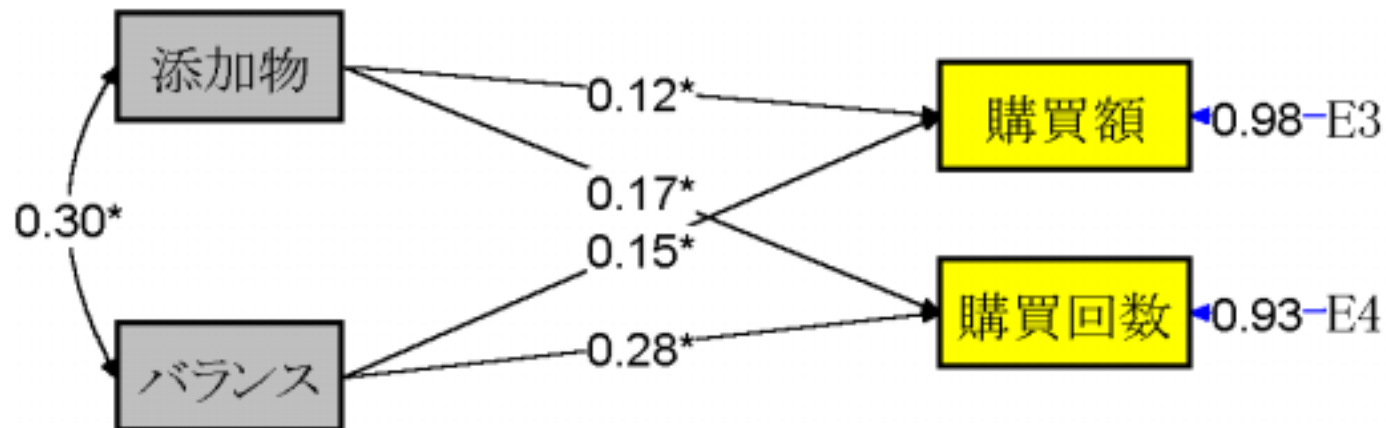
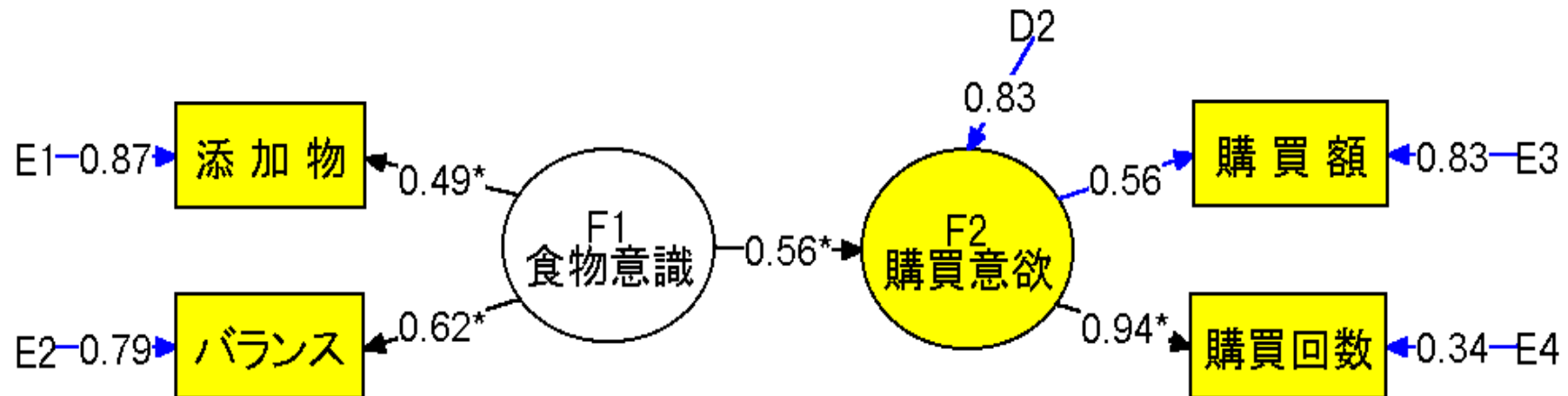
- 二値独立変数からの矢印は平均効果を表す
 - γ_1 : 直接効果
 - 学習時間が同じであっても生じる学力への影響
 - 生得的なもの, 学習時間以外の環境
 - $\beta\gamma_2$: 間接効果
 - 学習時間の違いによる学力への影響

$$E[\text{学力} | \text{父親大卒}] = \gamma_1 + \beta\gamma_2$$

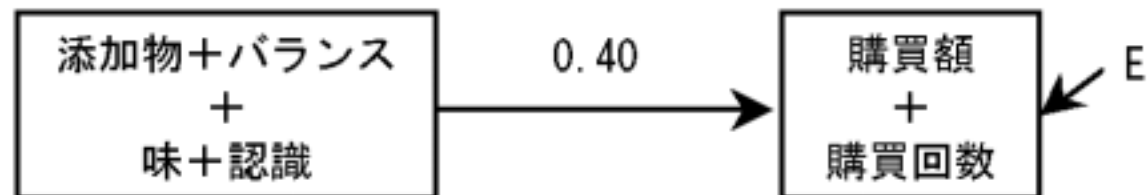
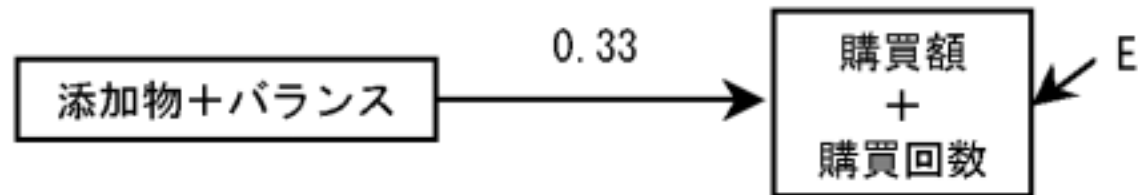
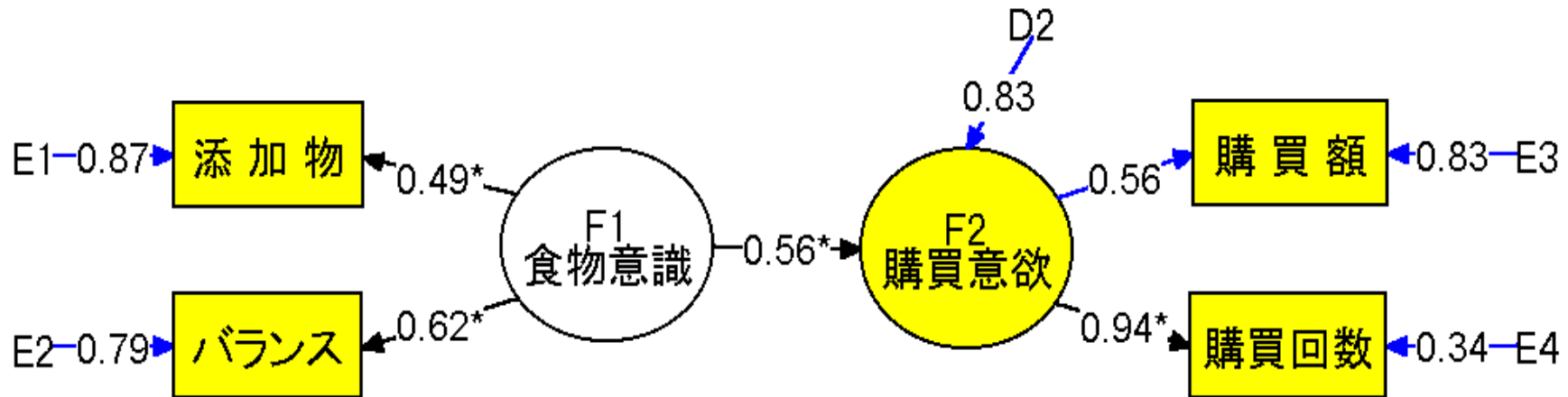
$$E[\text{学力} | \text{父親非大卒}] = 0$$

希薄化の修正1：パス解析との比較

	X1	X2	X3	X4
X1: 食品添加物に気を使う	1			
X2: 栄養のバランスに気を使う	0.301	1		
X3: 自然食料品店での購買額	0.168	0.188	1	
X4: 自然食料品店での購買回数	0.257	0.328	0.530	1



希薄化の修正2：尺度化との比較

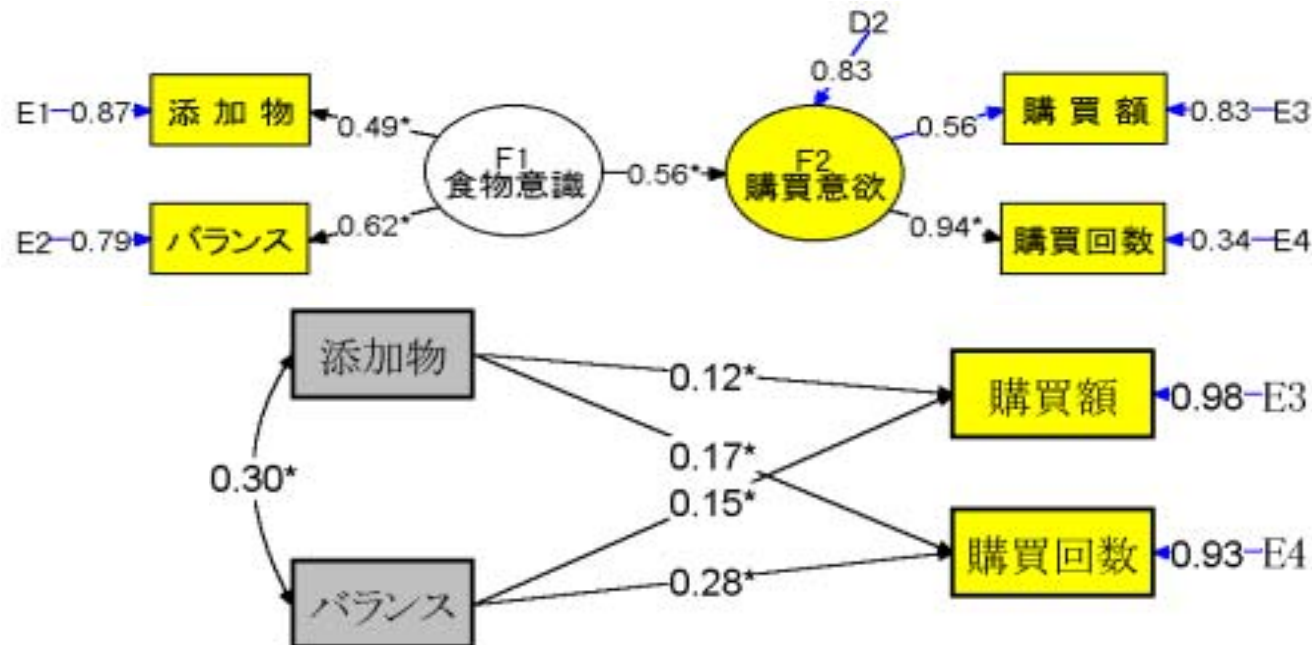


パス解析と多重指標分析 - 1

- 係数の大きさ

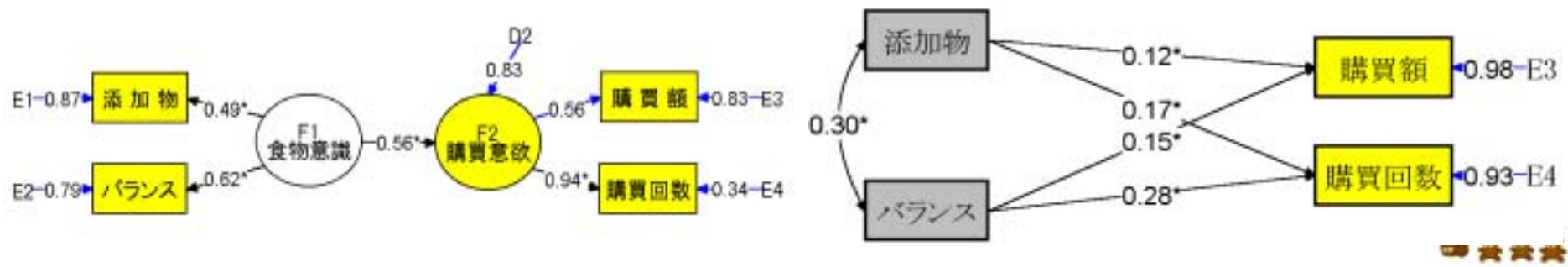
多重指標分析 > 単相関 > パス解析

	X1	X2	X3	X4
X1: 食品添加物に気を使う	1			
X2: 栄養のバランスに気を使う	0.301	1		
X3: 自然食料品店での購買額	0.168	0.188	1	
X4: 自然食料品店での購買回数	0.257	0.328	0.530	1



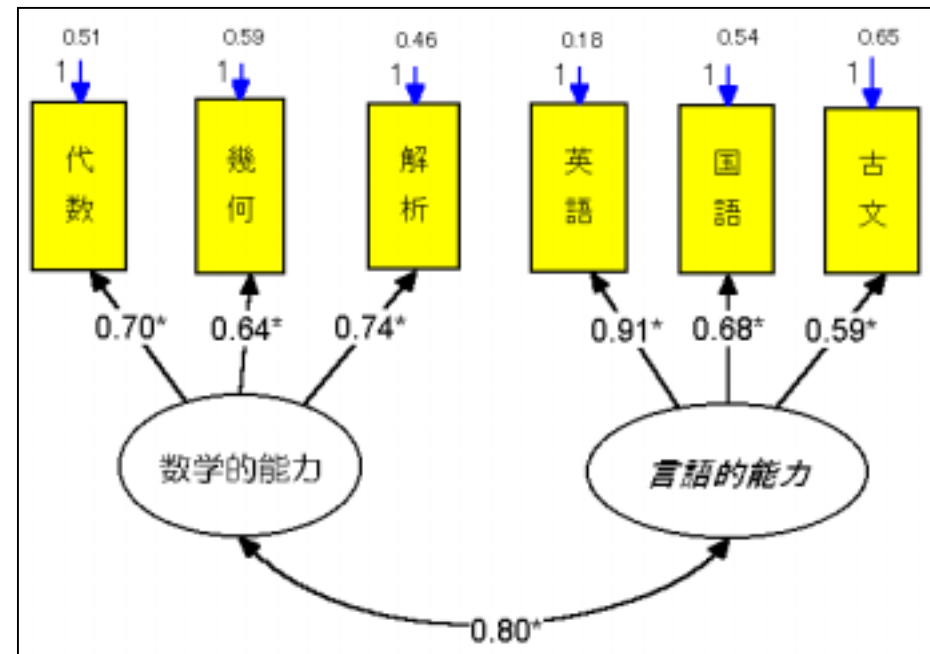
パス解析と多重指標分析 - 2

- 「バランス」が一定の下での「添加物」の影響とは何かよく分からない
- 共変する似た項目は潜在変数でまとめる方がよい
- 違った側面をもつ項目は潜在変数化しない方がよい(中古車価格の例)



適合度の吟味

- モデルがデータをどの程度反映しているかを定量的にみる
 - モデルとデータの距離を測る
- 因子分析の例
 - どうもオカシイ
 - EFAの結果と合わない
 - 因子相関が高すぎる



補遺：探索的因子分析との比較

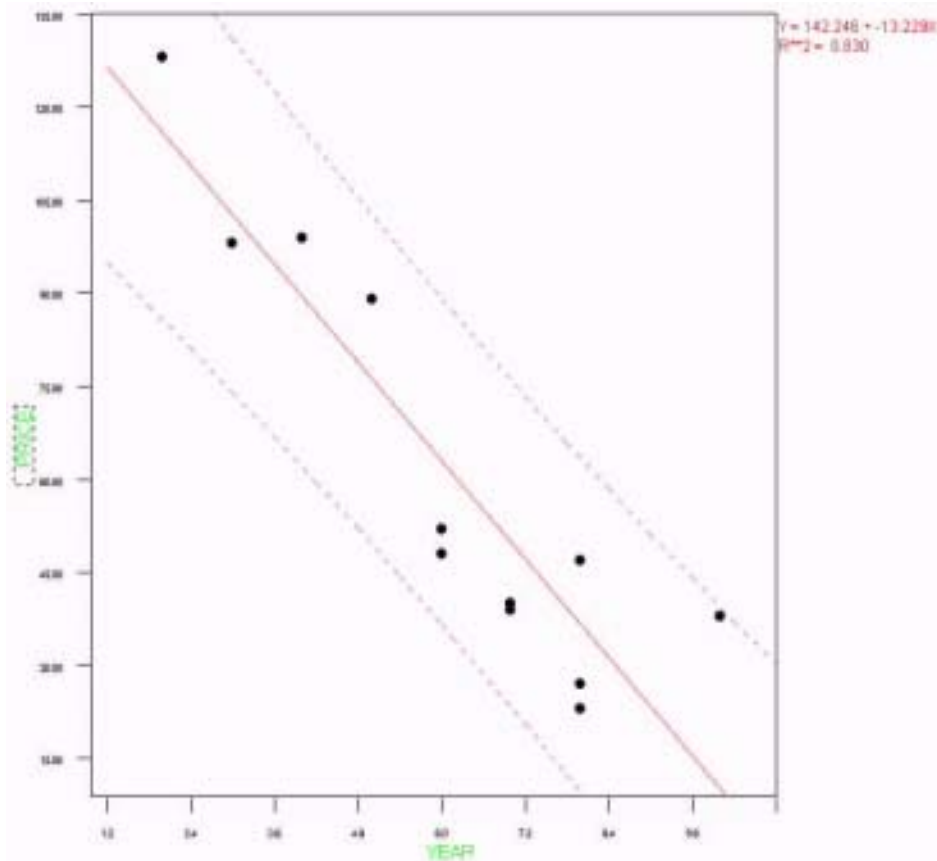
	探索的因子分析 (プロマックス斜交回転)		検証的因子分析	
	F1	F2	F1	F2
	代数	0.63	0.08	0.70
幾何	0.67	-0.02	0.64	0
解析	0.73	0.01	0.74	0
英語	0.39	0.54	0	0.91
国語	-0.06	0.82	0	0.68
古文	0.02	0.63	0	0.59
因子相関	1	0.64	1	0.80
	0.64	1	0.80	1

推定と適合度

回帰分析--- データとモデルの距離 ---

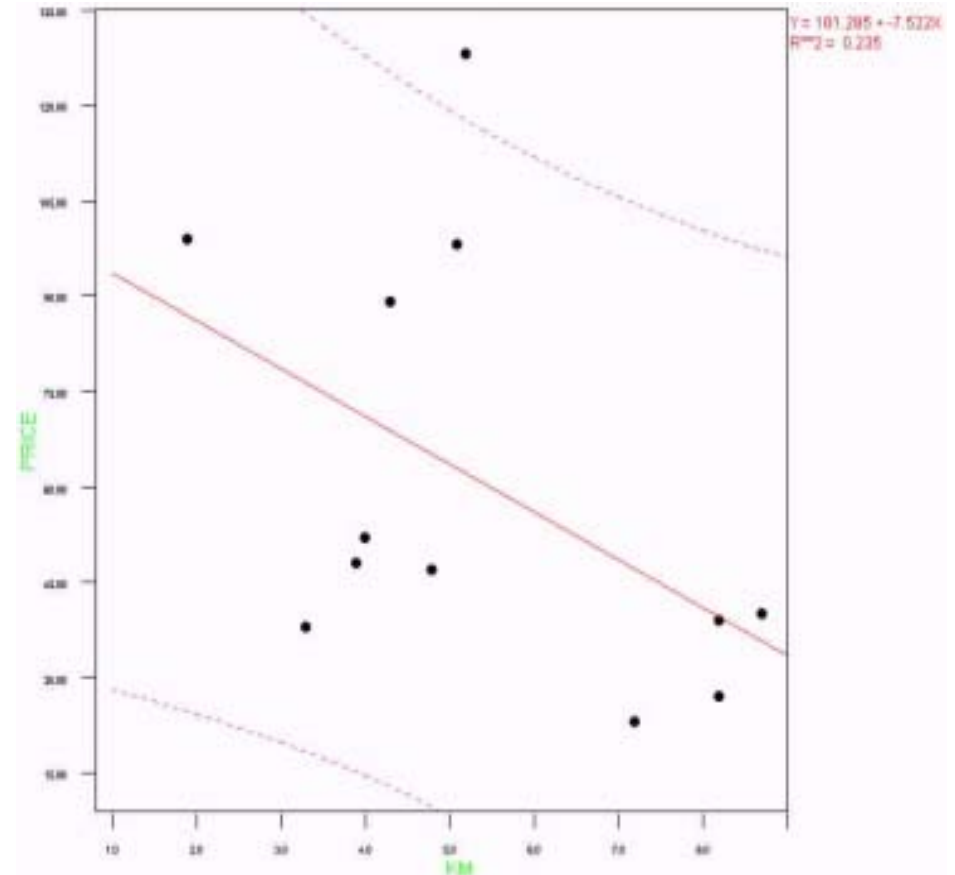
良い当てはまり

乗車年数と価格($r=-.91$)



悪い当てはまり

走行距離と価格($r=-.49$)

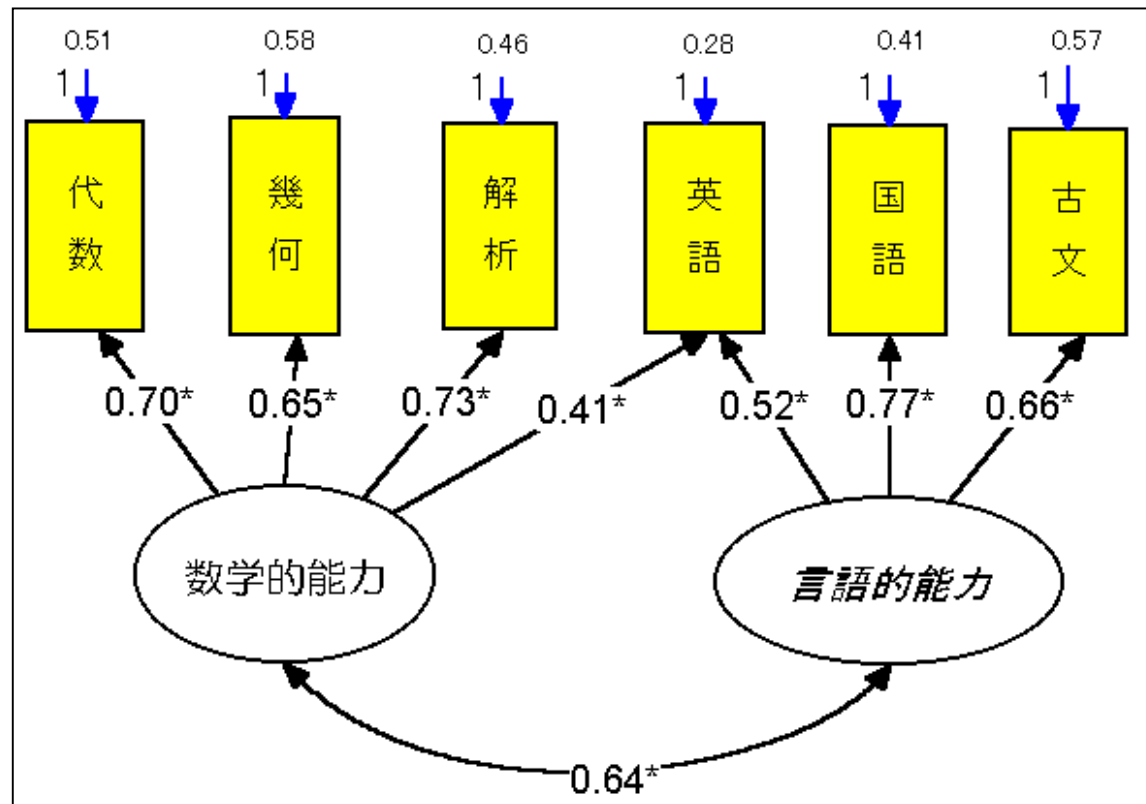


適合度検定: カイ2乗値

- H0: モデルが正しい
versus
H1: モデルが正しくない
- 上記仮説をカイ2乗検定する
結果 カイ2乗値=18.244 (df=8)
有意確率(p値) = 0.019
- モデルは有意水準5%で棄却
- モデル修正の必要性

修正されたモデル

- カイ2乗値=6.307 (df=7)
有意確率(p値) = 0.504



まとめ

- 分析対象データ: 多変量観察データ
- 基本モデル: パス解析, 因子分析, 多重指標分析
 - 調査項目間の因果関係の記述
 - 調査項目をまとめて単純化(潜在変数化)する
 - 調査項目をまとめてから因果関係を調べる
- 効果の分解
 - 直接効果, 間接効果, 総合効果
- 次元縮小
- 希薄化修正
 - 誤差を分離し, 本来検討したいものののみを取り出してから分析
- 仮説モデルの修正
 - 適合度の吟味

参考書籍：初級者向け

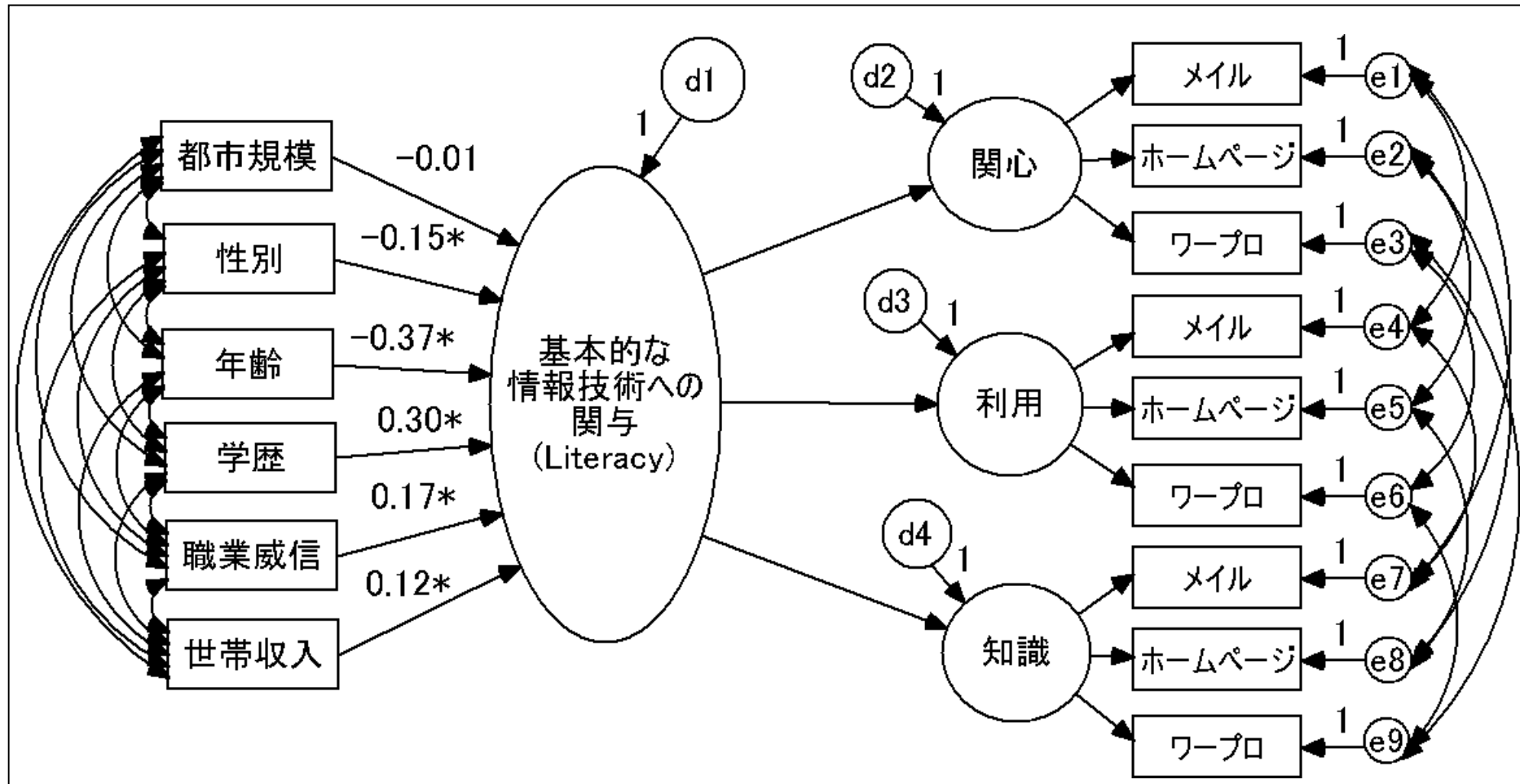
- 豊田秀樹・前田忠彦・柳井晴夫(1992). 原因を探る統計学
ブルーバックス. 講談社
- 狩野 裕(1996.2 ~ 1997.3). 「共分散構造分析とソフトウェア」
BASIC数学連載. 現代数学社
- 豊田秀樹(1998). 共分散構造分析(1998)事例編. 北大路書房
山本嘉一郎・小野寺孝義(1999). Amosによる共分散構造分析と解
析事例. ナカニシヤ出版
- 田部井明美(2001). S P S S 完全活用法 共分散構造分析
(Amos)によるアンケート処理. 東京図書
- 狩野裕・三浦麻子(2002). グラフィカル多変量解析:目で見える共
分散構造分析(増補版). 現代数学社.

参考書籍：中級者向け

- 豊田秀樹(1992). SASによる共分散構造分析. 東大出版
- 豊田秀樹(1998). 共分散構造分析[入門編]. 朝倉書店
- 豊田秀樹(2000). 共分散構造分析[応用編]. 朝倉書店
- Bollen, K. (1989). Structural Equation with Latent Variables. Wiley.

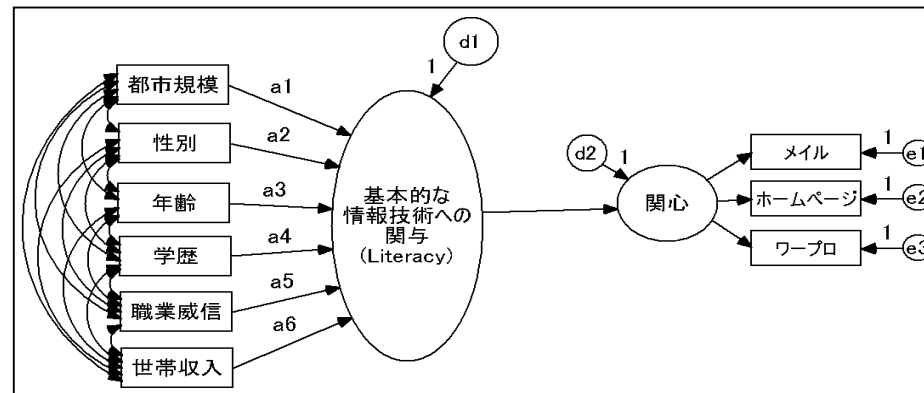
付録：仮想例と様々な分析

社会的属性とデジタルデバイド？



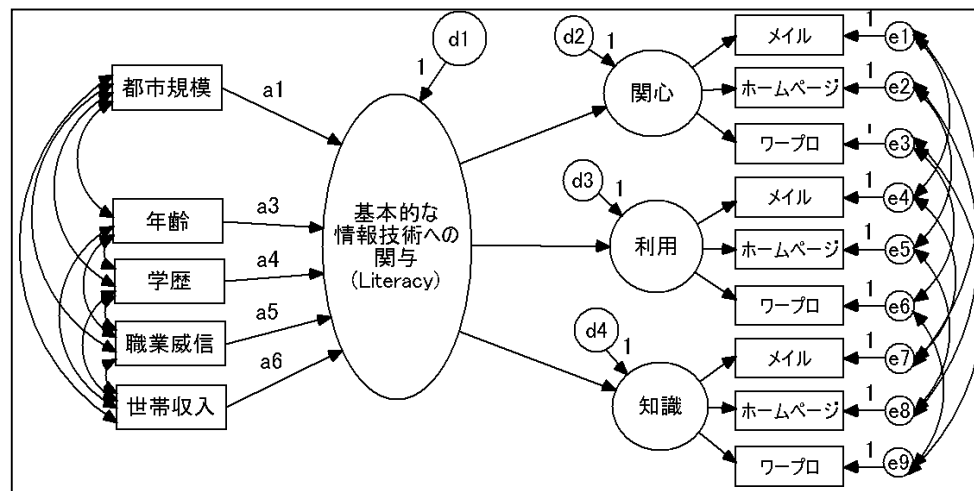
従属潜在変数ごとの分析

- 関心・利用・知識ごとに分析
 - 「Literacy」へのパス係数が分析ごとに異なってよい
 - 関心・利用・知識が3因子構造でなくてもよい



男女別に分析

- すべてのパス係数が男女ごとに異なりえる
 - 性別変数を入れた分析は, パス係数は共通
 - 交互作用がないという分析



「職業威信」「世帯収入」に欠測

- 欠測のパターンによって4つの集団を作成
 - すべてのパス係数を等値
 - 平均も等値
 - AMOSは自動的にこの分析を行う
 - MARの条件

