

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 (大阪大学人間科学部)

市川雅教 (東京外国語大学)

はじめに

本セミナーでは、「共分散構造分析」についてその入門とやや高度な分析事例を解説します。セミナーを通じて、「共分散構造分析」とは何か、どんな分析ができるのか、その限界は何かなどを理解して頂ければ、われわれの目的は達せられることとなります。

一般に、チュートリアルセミナーでは、受講者のレベルとニーズを正確につかむことが重要です。私たちの今までのコンサルティングや講習会の経験を総合し、次の三つのグループを念頭においています：

- ◇ 新しく共分散構造分析を勉強してみたいという初級者層
- ◇ 易しいテキストに出ている事例は理解し共分散構造分析のソフトも使ったことがあるが、実際論文で使うような分析となると不安があるという中級者層
- ◇ 共分散構造分析を用いて論文を執筆したことがあるが、自信をもって分析するにはその理論背景を知りたい、もしくは、当面分析する機会はないが、(方法論者として) その理論を知りたいという理論派層。

今回のセミナーでは、受講対象として初級者層と中級者層を念頭におき、理論背景については、統計ユーザーが直接関連する内容、もしくは疑問をもつことが多い点を中心に解説します。

第2章は初級者向けの内容で、共分散構造分析で最も基礎的かつ重要なモデルである「パス解析」「検証的因子分析」「多重指標分析」を紹介します。この3つのモデルを習得するだけで共分散構造分析が相当使えるようになります。この章は適合度の解説などもあり、この章だけで self-contained になるよう工夫しました。第3章は中級者向けです。ここで紹介するモデルを使って統計分析し論文を執筆すると、査読者にインパクトを与えることは間違いないでしょう。第4章では、共分散構造分析による因果推論の可能性と限界について解説してあります。

今回は、少し新しい試みとして、実験データの共分散構造分析とブートストラップ法をやや詳しく紹介しています。多母集団の同時分析は代表的な実験データの分析方法ですが、比較する母集団が多くなるとその適用は難しくなります。そこで、簡便法が注目されます。簡便法とは、例えば、男女を 0,1 という変数 X で表したとき、 X を連続変数と思って他の変数と一緒に共分散構造分析をする方法です。この方法は、 X が独立変数であるときには「罪は重くない」のです。第5章では、「罪は重くない」という主張の理由も理論的に解説しています。

計算機の発達に伴い、共分散構造分析でもブートストラップ法が手軽に適用できる環境になってきました。推定量の標準誤差や適合度検定は、正規性などの母集団分布に関する仮定の下で導かれています。ブートストラップ法はそのような仮定をおかず、計算機のパワーで標準誤差や適合度検定統計量を計算します。正規性の仮定が疑わしい社会科学では、多くのソフトウェアがブートストラップ法をオプション化していることは朗報です。

「共分散構造分析」を統計ユーザーと、これから理論的なことも勉強してみたい方へお贈りする日本統計学会チュートリアルセミナー、ごゆっくりとお楽しみください。

目次

はじめに	i
第1章 共分散構造分析とは	1
1.1 定義	1
1.2 潜在変数とパス図	2
1.3 歴史とソフトウェア	4
第2章 共分散構造分析で何ができる：基礎編	9
2.1 基本的なモデル... 3つの御利益	9
2.2 モデルファイルと共分散構造	15
2.2.1 モデルファイル作成のコツ	15
2.2.2 共分散構造	16
2.3 モデルの修正	17
第3章 共分散構造分析で何ができる：応用編	19
3.1 平均構造モデルと多母集団の同時分析	19
3.1.1 データと標準解・非標準解	19
3.1.2 平均構造モデル	21
3.1.3 多母集団の同時分析	23
3.1.4 簡便法	26
3.2 実験データの共分散構造分析	29
3.3 潜在曲線モデル	32
3.3.1 例題と分析結果	32
3.3.2 考え方	34
第4章 因果分析について	37
4.1 因果の方向を探る	37
4.2 因果の大きさを正確に測る	41
第5章 統計的推測の基礎	45
5.1 最尤推定とカイ2乗値	45
5.2 適合度指標	49
5.3 二値データなどの扱い	54
5.3.1 独立観測変数の場合：Muthénのフレームワーク	54
5.3.2 3.1.4節の簡便法の吟味	57

第 6 章 共分散構造分析におけるブートストラップ法	61
6.1 ブートストラップ法とは	61
6.2 ブートストラップ法の考え方	61
6.3 共分散構造分析におけるブートストラップ法の利用	64
6.4 ソフトウェアの対応	66
付録：モデルファイル例	68
参考文献	77

第1章

共分散構造分析とは

1.1 定義

共分散構造分析 (covariance structure analysis) とは、その名の通り、観測変数間の分散・共分散の構造を分析する方法である。この定義は数理的側面からのもので、共分散構造分析で何が出来るかは伝わってこない。より分かり易く表現するため、私たちは、共分散構造分析を次のように考えている：

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。

潜在変数と観測変数という概念は、共分散構造分析を理解する上できわめて重要である。これらの解説は次節にゆずる。

因果関係を表現する線形の方程式を構造方程式 (structural equation) という。このことから、共分散構造分析は構造方程式モデリング (structural equation modeling; SEM) と呼ばれることも多い。

因子分析では平均の情報は用いない。観測変数や因子の平均は0に固定されている。一方、共分散構造分析では平均に関する推測を行うことができ、ときにはきわめて有用な道具になる (3.1.2 節参照)。このモデルは平均・共分散構造モデルとよばれることがある。共分散構造分析という呼称は「平均・共分散構造モデル」を含まない印象を与えるのであまり良い名称とはいえない。そこに、構造方程式モデリング (SEM) という別称の価値があり、最近では SEM の方がよく用いられている。

以前は、上記の定義に「基本的に非実験多変量データの分析方法」という説明も加えていた。しかし、最近、実験データへの適用可能性が広がり、分析事例も多くなってきたので、今回はこの説明を外した。

非実験多変量データとは、調査データ・観察データともいわれるもので、アンケート調査や試験の成績のデータと考えればよい。一方、実験データとは分散分析できるデータである。つまり、特性に影響を及ぼすであろう要因に水準を定め、各水準ごとに収集されたデータである。3.1 節で紹介する語彙能力の分析では、コントロール群とトレーニング群の比較を行う。2群の比較であるから、これは小さな実験データの分析である。3.2 節で紹介するのはより本格的で、2要因で各3水準からなる実験データの分析事例である。

要因や水準数が多いと比較すべき群が多くなる。そのような場合は、各群の分散共分散行列が等質であるという条件のもとで分析する簡便法しか適用できないことが多い。3.2節での分析も簡便法である。現在のところ簡便法の評価は定まっていないが、将来的には豊かな可能性を秘めた分析方法であると考え、今回ここで紹介する方法に学んで、色々な分野で試していただきたい。

1.2 潜在変数とパス図

直接観測できない変数を潜在変数 (latent variable) という。因子分析でいうところの共通因子 (common factor) もしくは因子 (factor) が潜在変数にあたる。これに対して直接観測できる変数を観測変数 (observable variable), または顕在変数 (manifest variable) とよんで潜在変数と区別する。具体的には、観測できる変数とはアンケート項目であり、試験の成績である。つまり、ソフトウェアに入力するデータ行列の変数である。

因果関係を記述した図をパス図 (path diagram) という。パス図では、観測変数を四角で、潜在変数を楕円で囲んで区別する。因果関係は片方向の矢印で表し、単なる相関は双方向の矢印で描く。

共分散構造分析での潜在変数の使われ方は大きく分けて次の二通りがある。

- 心理学や社会科学でいう「構成概念 (construct)」を数学的に表現したもの
- 大きな誤差を伴ってしか観測できないもの

構成概念とは、例えば、「知能」「有能感」「満足度」など、日常会話にはよく出てくるが、物理量のように正確に定義しにくいものをいう。日常会話によく出てくるということは、その概念は便利であり、また、それなりに普遍的な意味をもっていることを意味する。

「学力」という構成概念を測定するために入学試験が実施されているものとしよう。大学入試では、多くの大学で、「X1: センター試験」と「X2: 個別試験」の結果に、「X3: 高校の成績」が加味されて合否が決まる。「学力」の高低は観測変数 X1, X2, X3 の結果を相当左右すると考えられるので、図 1.1 のようなパス図が描ける。このパス図は、X1, X2, X3 間の相関が潜在変数である学力で説明できるということを意味している。X1 が高い受験生は、X2 も X3 もそれなりに高い傾向にあり、X1 が低い受験生は X2 も X3 も低いことが予想される。つまり、これらの3つの変数は互いに連動している。これを、X1, X2, X3 は共変する (covary) という。図 1.1 のパス図は、3変数が共変する要因が (第三変数の) 潜在変数である「学力」であることを表している。潜在変数「学力」で説明できない変動を表すのが誤差変数 e_1, e_2, e_3 である。直接観測できない「学力」を計測するために観測される変数 X1, X2, X3 を「学力」の指標 (indicator) という。

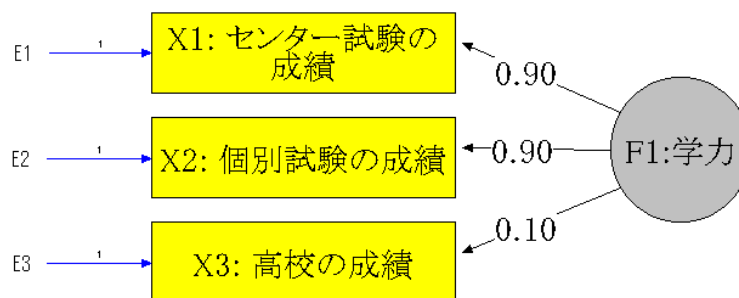


図 1.1: 「学力」を測定するモデル

因果を表す片方向の矢印の上には因果の影響の大きさを表すパス係数¹の推定値が付される。X3へのパス係数が小さい値になっているのは、高校間格差によって、高校での成績が学力と共変する部分があまり大きくないと考えられるからである。

図 1.1 のモデルを考えることで、3つの観測変数が「学力」という1つの変数にまとめられ、次元縮小が行われたことになる。次元縮小が目的ならば、単純和 $X1+X2+X3$ や重み付きの和 $X1+X2+0.1X3$ を考えればよいではないか、という意見が聞こえてきそうである。もちろんそれも有効な方法である。図 1.1 のモデルを考えることの利点は、(i) パス係数がデータから決められること (ii) 図 1.1 の「学力」には誤差が含まれていないこと、である。

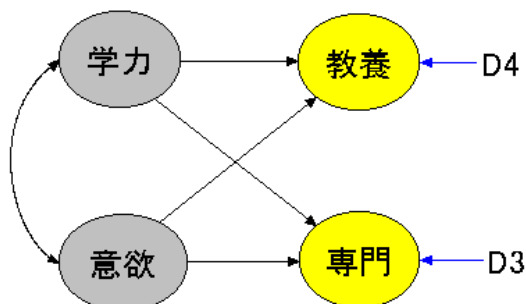


図 1.2: 「専門知識と一般教養」の因果モデル 1

大学教育の大きな目的は、高度な専門教育と幅広い教養を身につけ、社会に貢献できる人材を育成することである。専門知識と一般教養に関して図 1.2 の因果モデルを考えよう。入学時の学力と学習意欲が一般教養科目と専門科目の成績を規定するというモデルである。このモデルが不十分ならば、例えば、一般教養から専門へパスを引いてもよいし、その代わりに、学力から専門へのパスを消してもよい (図 1.3)。このモデルは、入学時の学力は、必ず一般教養の成績を経由して専門に影響を及ぼすことを示している。別の観点として、学習意欲を1・2年次のものと3・4年次のものとに分けて考えるモデルも考えられる。

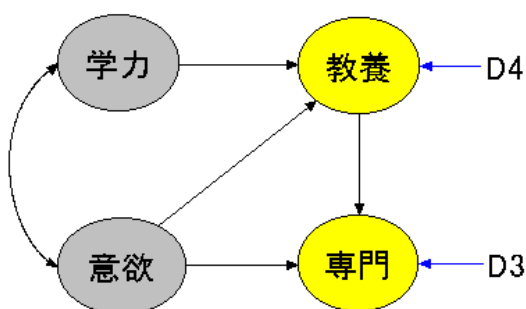


図 1.3: 「専門知識と一般教養」の因果モデル 2

共分散構造分析は、図 1.2 や 図 1.3 のように、検討したい因果に関する仮説を設定することから始まる。次に、潜在変数の指標としてどのような観測変数を観測するのが望ましいかを検討する。この手順は、まず観測変数ありきで始まる古典的な探索的多変量解析とはまったく逆である。図 1.4 に観測変数を含めた全体モデルの例を示してある。

¹偏回帰係数や影響指数とよばれることもある。

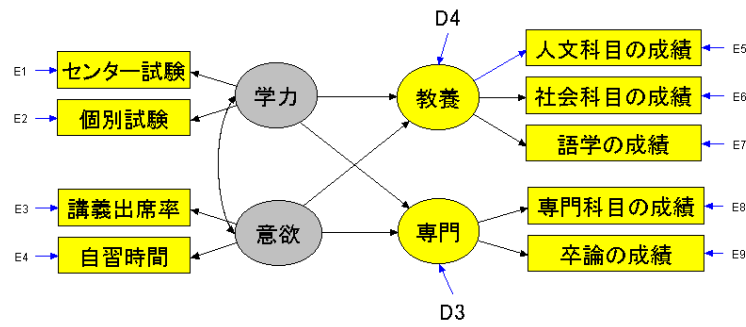


図 1.4: 因果モデル1での指標（観測変数）の例（仮想モデル）

共分散構造分析の最大の特徴は、研究者自身が構成した因果に関する仮説を直接モデル化し分析することができる点にある。古典的な因子分析・主成分分析や回帰分析などでは、このような自由なモデル構成ができないのである。この意味で、Fornell(1982)は「共分散構造分析は第二世代の多変量解析」、豊田²は「共分散構造分析は多変量解析のための簡易高級言語」としている。

以下にパス図を描くときの基本原則をまとめておく。

観測変数は四角形で囲む
 潜在変数は（楕）円で囲む
 誤差変数は記号のみか（楕）円で囲む
 片方向矢印は因果を表す
 双方向矢印は（単なる）相関関係を表す
 片方向矢印を受けた変数（従属変数）には必ず誤差変数が付属する
 片方向矢印の上には、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
 双方向矢印の上には、相関係数または共分散の推定値が付される

1.3 歴史とソフトウェア

本節では、共分散構造分析の歴史とソフトウェアを簡単に紹介する。

構成概念を導入した統計モデルの端緒は英国の心理学者 Spearman (1904) の1因子（分析）モデルと考えられる。彼は、古典、仏語、英語、数学、音の弁別、音楽の6つのテストを子供に実施し、結果は一般知能「g」を持つ1因子モデルで説明できるとした。この1因子モデルを多因子に拡張したのが Thurstone(1935)であり、現在、いわゆる因子（分析）モデルもしくは探索的（因子）分析モデル（exploratory factor analysis model）とよばれているものである。多因子モデルが開発されると因子回転が問題となった。現在最もしばしば用いられている varimax 回転法は、Kaiser によって 1958 年に開発された（Kaiser 1958）。

探索的因子分析とやや趣を異にする検証的因子分析（confirmatory factor analysis）は、共分散構造分析の最も重要なコンポーネントである。

パス図 1.5 で表されるようなモデルを検証的因子分析モデルという。検証的という術語は、背後に潜む因子を事前に想定し、因子 F_1 は観測変数 X_1, X_2, X_3 だけに影響を及ぼし因子 F_2 は

²メイリングリスト f p r の記事より。次の URL 参照 <http://www.nuis.ac.jp/~mat/fpr/fpr1997/0130.html>

観測変数 X_4, X_5, X_6 だけに影響を及ぼすというような事前の仮説を導入し、それを検証しようとするに由来する。言い換えると、 F_1 は X_4, X_5, X_6 に影響を及ぼさず、 F_2 は X_1, X_2, X_3 に影響を及ぼさないという仮説を explicit に記述したモデルといえる。

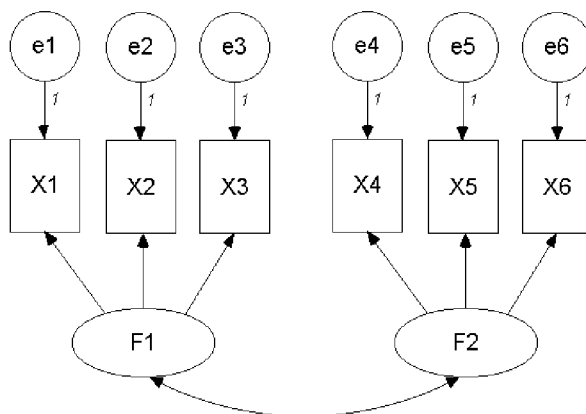


図 1.5: 検証的因子分析モデル

探索的因子分析モデルとは、因子がどの観測変数に影響を及ぼすかを仮定せず、どのような因子があるか探索するモデルをいう。

まとめると、検証的モデルは、因子の数、名前、因子がどの観測変数に影響を及ぼすかなどが既知である状況に適用するのに対して、探索的モデルでは、これらがすべて未知のときに使うことになる。

1960年代半ばには、探索的モデルは未知の部分が多すぎ、同モデルで適切に分析できないデータがかなりあることが分かってきた。それが、計算機の発達や Newton-Raphson 法などの優れた計算アルゴリズムが因子分析の計算に導入されたことにより顕在化したことは皮肉なことである(丘本 1987, 17 ページ)。そこで、1960年代の後半、スウェーデンの統計学者 Jöreskog は検証的因子分析モデルを導入した。検証的モデルでは分析できるが、探索的モデルでは分析できないモデルがかなりあることが、数学的にも分かっている。

検証的因子分析モデルをさらに発展させたのが共分散構造モデルである。Jöreskog は、「いくつかの因子が特定の観測変数に影響を及ぼす」という因子分析の構造にとらわれず、もっと自由に因果の関係を分析できるようモデルを拡張した。その際、観測変数間の因果関係を記述するモデルであったパス解析の考え方が採用されている。パス解析は生物学者 Write が中心になって 1920-30 年代に発展させられた因果分析法である。なお、初めて共分散構造分析という言葉を使ったのは、Bock-Bergmann (1966) と言われている。

共分散構造分析の発展に重要な貢献をしたのが、Jöreskog-Sörbom の LISREL モデルであり、同名のソフトウェアであった。LISREL とは LInear Structural RELationships の頭文字をとったものである。LISREL は 最初 1973 年に発売され、現在のバージョンは version 8.30 である。LISREL の歴史は清水 (1994) に詳しい。

共分散構造分析のもう一つの流れは、経済学の同時方程式モデルと考えられている。Klein (1950) は、大恐慌期を含む 1921 年から 1941 年までの米国の経済変動をある線形モデルで記述することに成功した。この研究を端緒にして同時方程式モデルといわれる分析方法が発展した。ここではこれ以上述べるスペースがないので、興味のある読者は森棟 (1985) を参照されたい。

1970 年前後は共分散構造分析の黎明期であったが、それに続く 1970 年代は、多母集団の同時分

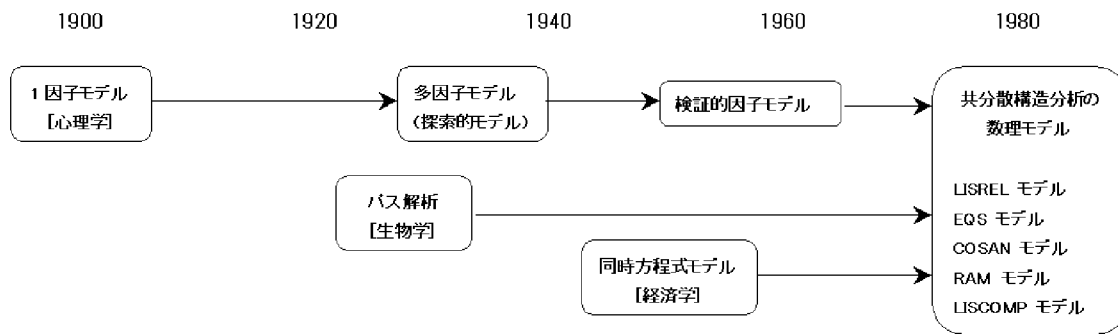


図 1.6: 共分散構造分析の生い立ち

析, 平均構造の分析などが導入され, 共分散構造分析の拡張期といえる. そして, 1970 年代後半から 80 年代前半にかけて, これらの拡張モデルを含めた各種モデルの数学的記述方法が開発された. Jöreskog-Sörbom の LISREL モデルの改訂に加えて, McDonald (1978, 1980) の COSAN, Bentler-Weeks (1980) の EQS, McArdle-McDonald (1984) の RAM などである. 詳しく解説するスペースはないが, 強調しておきたいのは, Muthén (1984) の LISCOMP である. 同モデルは, 共分散構造分析でカテゴリカルデータの扱いを可能にした. 特に, 因果モデルにおいて, カテゴリカルである項目が従属変数であるか独立変数であるかで扱いが異なることを指摘し, 独立変数の場合は連続変数とみなしても大きな損失はないこと, 従属変数の場合はプロビット変換やロジット変換などの考え方に基づく特殊な解法が必要になることを指摘し, プログラム LISCOMP を開発した. 最近, LISCOMP を発展させた Mplus (Muthén-Muthén 1998) が発売され, 同方法がエンドユーザーにも手軽に使えるようになった. 以上の歴史を図 1.6 にまとめている.

1980 年代は, 理論的に興味ある結果が導かれた時代であった. 共分散構造モデルは多変量正規分布の仮定の下で発展させられてきた手法である. しかし, 社会科学のデータの多くは多変量正規分布を仮定できない. そこで, 非正規分布の下での統計推測理論が発展させられることとなった.

1990 年代はソフトウェアの時代といえよう. Windows の GUI を利用したユーザーに易しいソ

表 1.1: ソフトウェアの特徴

	AMOS 3.60	EQS 5.7	LISREL 8.30	Mplus 1.0	SAS CALIS 6.12	STATISTICA SEPATH 5.0J
パス図	○	○	○			
多母集団の同時分析	○	○	○	○		○
カテゴリカルデータの分析		△	○	○		
LM検定・修正指標	△	○	○	○	○	
潜在曲線分析	○	○	○	○	○	○
ミッシングデータの最尤分析	○			○		
相関行列の正確な分析						○
ブートストラップ法	○	△				○
シミュレーション	△	○	○	○		○

ソフトウェアが続々と発売された。共分散構造分析は、多くのエンドユーザーにとって手軽に適用できる分析方法になったのである。

CALIS は汎用ソフトである SAS に搭載され、共分散構造分析の普及に貢献している。総合的な統計プログラムである STATISTICA には Steiger による SEPATH が付属している。独立系の共分散構造分析ソフトウェアである、Arbuckle の AMOS, Bentler の EQS, Jöreskog-Sörbom の LISREL は、Windows の GUI を活かして、パス図を入力ファイルにしたり、パス図の自動描画のオプションを備えている。表 1.1 で、これらのソフトウェアの特徴比較をしているので参考にされたい。

日本においても共分散構造分析の解説書が充実してきた。ソフトウェアに準拠したものとして、豊田 (1992, CALIS), 狩野 (1997, AMOS, EQS, LISREL), 山本-小野寺 (1999, AMOS) がある。総合的な解説書としては、豊田-前田-柳井 (1992) と豊田 (1998a) が出版されている。³ 豊田 (1998b) は共分散構造分析の貴重な事例集である。

柳井他 (1990) や服部-海保 (1996) は共分散構造分析に一章を割いている。

³豊田 (1992, 1998a) はやや高級で中級者向けである。その他は初級者向けの解説書。

第2章

共分散構造分析で何ができる：基礎編

2.1 基本的なモデル... 3つの御利益

共分散構造分析によって、こういったデータに対してどのような分析ができるのか、ということ解説する。分析対象となるデータのタイプは、いわゆる観察によって得られる多変量データである。従来の因子分析や主成分分析、連続量の回帰分析などが適用できるデータと考えればよい。基本的には連続データが望ましいが、5件法や7件法による順序尺度データも連続データとみなして共分散構造分析されることも多い。なお、グループ変数や二値の変数が含まれるデータの分析は応用編に属する。

このようなデータに対して、まず、以下の3通りの分析を理解したい：

1. 観測変数（調査項目）間の因果関係を調べる（多重回帰分析，パス解析）
2. 観測変数（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. 観測変数（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）

表 2.1: 中古車価格のデータ

No	価格 (万円)	走行距離 (KM)	乗車年数 (年)	車検の残 (月)
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

パス解析の例として、中古車価格の分析例を紹介する。¹

¹本節の例は狩野(1997)からのものである。

表 2.1 に示したデータは、首都圏の物件を中心にした中古車情報誌からのもので、ある車種 12 台の中古車価格とその関連要因を調べたものである。このデータをパス解析モデルで分析した結果を図 2.1 に示してある。中古車価格を乗車年数と走行距離、車検（の残り）によって説明しようというモデルである。さらに、走行距離は乗車年数から影響をうけることを想定している。また、乗車年数と車検とは本来関係がないので“相関関係にない”という情報もモデルに組み込まれている。² このように、パス解析では、観測変数間に実質科学的観点から因果関係を想定すること

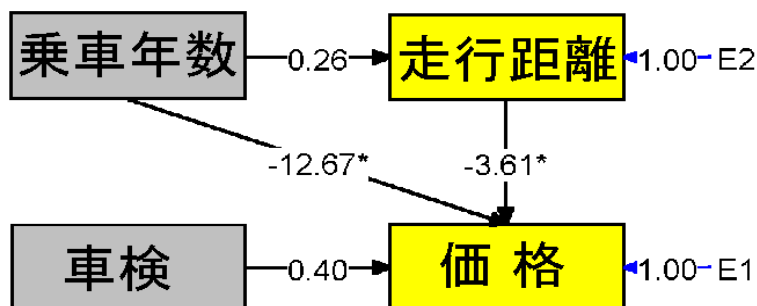


図 2.1: 中古車価格のモデルと推定結果

で、因果の大きさを推定することができる。

この分析の一つの成果は効果の分解である。乗車年数から価格へのパスは 2 種類ある。直接のパス（パス係数：-12.67）と走行距離を経由する間接のパスである。直接のパスは、走行距離を経由しない、つまり、走行距離が一定であるとき、乗車年数が 1 年高い（長い？）中古車の価格は 12.67 万円安いことを意味している。これを**直接効果（direct effect）**という。走行距離から価格へのパス係数は、乗車年数が一定であるならば、走行距離が 1 万キロ長い車は、3.61 万円安くなることを示している。一方、乗車年数から走行距離へのパス係数は 0.26 であり、これは、1 年当たり平均的には 0.26 万キロ乗車することを示すから、乗車年数が 1 年増加したときの走行距離を経由する価格への効果は

$$0.26 \times (-3.61) = -0.94$$

と推定できる。これを乗車年数から価格への**間接効果（indirect effect）**という。直接効果と間接効果を合わせて**総合効果（total effect）**という。

この分析から、乗車年数の価格への（総合）効果は

$$-12.67 - 0.26 \times (-3.61) = -13.61$$

と計算できる。つまり、1 年乗車年数が高くなると平均的に見て 13.61 万円価格が下がることが分かる。さらに、この価格下落は、走行距離に無関係な部分（直接効果：所有しているだけで下がる部分）と走行距離が増えるから価格が下がるという部分（間接効果：使ったから価値が下がる部分）に分解することができるのである。

さて、少し基本的なことを補足する。極論すれば、パス解析では、矢印（パス）をどのように引いてもよいのである。ソフトウェアはどのようなモデルに対しても、それなりの推定値を出力してくれるであろう。パス解析が科学的な方法論とはなるためには、パスを引くことの妥当性を科学的に評価する必要がある。そこで登場するのがモデルの**適合度（goodness-of-fit）**の評価である。適合度とは、モデルとデータの距離である。距離が遠ければ悪いモデル、近ければ良いモデルと判断される。

²乗車年数と車検の間に双方向の矢印が描かれていない。

具体的に説明すると、パス図に基づいて計算される観測変数間の分散・共分散 (implied variance-covariances, しばしば $\Sigma(\hat{\theta})$ 表される)³とデータから計算される分散・共分散 (sample variance-covariances, しばしば S で表す) の食い違いの程度を測ろうとするものである。

適合度指標はソフトウェアによっては30以上も出力されるので、ユーザーとしては迷うばかりである。しかし、一般には、カイ2乗値、GFI、CFI、RMSEAの4つをおさえておけばよいだろう。カイ2乗値は上記の差 $S - \Sigma(\hat{\theta})$ をモデルが正しいという仮定の下で確率評価する。P-値が出力されるので、0.05以上であればモデルを受容する。GFIは、 S と $\Sigma(\hat{\theta})$ を回帰分析における基準変数とその予測値に見立て、回帰分析の決定係数 R^2 として定義される。CFIは、最も制約的なモデルである観測変数間に相関がないというモデル (独立モデル) を導入し、独立モデルのカイ2乗値を当該モデルのカイ2乗値がどの程度改善したかを評価する。GFIとCFIともに、0.90または0.95以上ならば当てはまりのよいモデルと判断する。RMSEAは、モデルを真のデータ発生機構と少しずれているということを認め、そのずれをカイ2乗分布の非心母数で評価したものにに基づいている。0.05以下であればよい当てはまり、0.10以上は悪い当てはまりであると判断する。0.05 < RMSEA < 0.10 はグレーゾーンである。

図2.1のモデルの適合度は次のようになり、よい当てはまりであることが分かる：

カイ2乗値=1.031(P-値=0.597,df=2,n=12)

GFI=0.957, CFI=1.000, RMSEA=0.000

2番目の御利益は因子分析である。例題として、6科目のテストの成績を取り上げる。表2.2に生データと相関行列を示す。

表 2.2: 6科目のテストの生データと相関行列

被験者	ゲール語	英語	歴史	計算	代数	幾何
1	77	82	67	67	81	74
2	63	78	80	70	81	79
3	75	73	71	66	81	52
4	55	72	63	70	68	67
5	63	63	65	70	63	89
6						

	ゲール語	英語	歴史	計算	代数	幾何
X1	1.000					
X2	0.439	1.000				
X3	0.410	0.351	1.000			
X4	0.288	0.354	0.164	1.000		
X5	0.329	0.320	0.190	0.595	1.000	
X6	0.248	0.329	0.181	0.470	0.464	1.000

6科目のテストデータを2因子の検証的因子分析した結果をパス図で示す(図2.2)。パス図から、本モデルは、文科的能力を表す潜在変数 F1 は数学系科目に影響を及ぼさず、また、数学的能力を表す潜在変数 F2 は文科系科目に影響しないことを仮定していることが分かる。このような仮説がデータによって指示されているかどうかは、先ほどの例と同様、適合度を調べることによって確認できる。本モデルの適合度は以下のようになり、適合度は高いことが分かる：

カイ2乗値=7.953(P-値=0.438,df=8,n=220)

GFI=0.988, CFI=1.000, RMSEA=0.000

³モデルによって「誘導された」という意味で implied とよぶ。

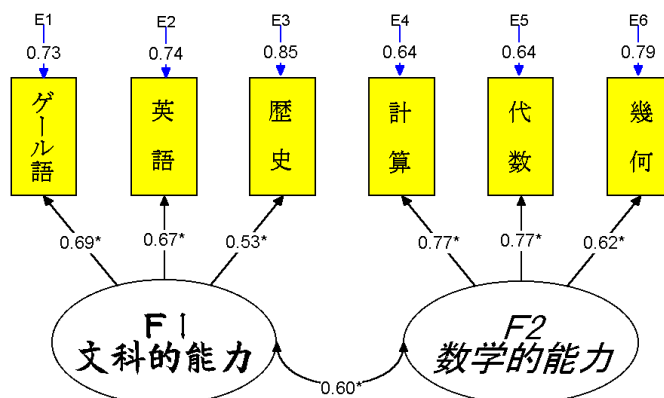


図 2.2: 因子分析モデルと推定結果

この分析結果から導かれる解釈を述べる。まず、パス係数（因子負荷量）を解釈することで、テスト項目についての知見が得られる。例えば、数学系科目の中では、「幾何」が少し外れており、文科系科目の中では「歴史」が外れている。

共分散構造分析のコンテキストでより重要なのは、潜在変数間の相関である。ここで推定された 0.60 は、文科系群と数学系群の間のどの「観測変数間の相関」よりもずっと高い値を示している。これを希薄化が修正されているという。観測変数には誤差 E1 ~ E6 が含まれており、これらの誤差のおかげで、観測変数間の相関は不当に薄められているのである。

因子分析の目的が次元縮小であることもしばしばである。次元縮小して、次のステップの分析の助けにするのである。例えば、文科系の 3 科目を一つのもの、数学系 3 科目を一つのものと考え、これらの 2 変数にいくつかの変数を加えて相関分析や回帰分析をするのである。検証的因子分析結果は、次元縮小を行い相関分析や回帰分析へ移るためのお墨付きを与えてくれる。

共分散構造分析が身近になる前は、「一つのもの」にする手段は、これら 3 科目の合計点の算出であった。これを尺度化という。しかし、先に述べたように、観測変数の合計点は誤差を含んだままであり、たとえ回帰分析を実行しても偏回帰係数は希薄化され絶対値が小さな値となる。つまり、バイアスをもってしまうのである。では、因子得点を用いて回帰分析を行うのはどうだろうか。実は、これもあまり有効でない。理由は 4 つある。(i) 因子得点の計算式にはたくさんの流儀がありベストがない、(ii) 因子得点の計算式には因子負荷量や独自性などの未知パラメータが含まれており、それらは推定値で置き換えることになる。その際の推定誤差が無視できない、(iii) 因子不確定性という「観測変数から因子得点を一意に決められないという問題」がある、⁴ (iv) 因子得点の推定式は観測変数の線形結合であり、依然として誤差項を含んでいる。従って、希薄化の問題を解決しない。さらに、探索的因子分析結果に基づいて因子得点を計算する場合は、(i)–(iv) に加えて、どの因子回帰法を選ぶかによる結果の不安定性の問題がある。

尺度化や因子得点を経由せずに、潜在因子を用いた回帰分析を実行できれば、上記のような問題は回避されるかもしれない。このための方法論が、第 3 番目の御利益「観測変数をまとめて潜在変数化してから因果関係を調べる」である。

表 2.3 のデータは、自然食品の購買と食物に関する質問紙調査の結果である。この調査に基づいて図 2.3 のモデルを考える。本分析の主要な興味は、食物に関する意識が自然食品の購買意欲をどの程度規定するかを調べることである。潜在変数である「食物意識」の指標として X1 と X2 を考え、「購買意欲」の指標として X3 と X4 を考えている。モデルの適合度は以下のような

⁴ 因子不確定性についての詳細は、丘本 (1986) を参照のこと

表 2.3: 「自然食品の購買行動」の生データと相関行列

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
6	4	4	115	5
7	1	2	85	3
8	5			

	X1	X2	X3	X4
X1: 食品添加物に気を使う	1.000			
X2: 栄養のバランスに気を使う	0.301	1.000		
X3: 自然食料品店での購買額	0.168	0.188	1.000	
X4: 自然食料品店での購買回数	0.257	0.328	0.530	1.000

ており、本モデルの適合度の高さが理解されよう：

カイ 2 乗値=0.430(P-値=0.5120,df=1,n=831)

GFI=1.000, CFI=1.000, RMSEA=0.000

分析結果から、食物意識から購買意欲へのパス係数が 0.56 と推定されていることが分かる。この値は、これら 2つの潜在変数の相関係数と考えるとよい。というのは、潜在変数の分散は 1 に標準化してあるからである。再び、この分析によるパス係数（相関係数）は、観測変数間の相関係数に生じている希薄化を修正していることが分かる。実際、X1, X2 と X3, X4 間の相関係数は、0.168, 0.188, 0.257, 0.328 であり、0.56 よりずいぶん小さい。

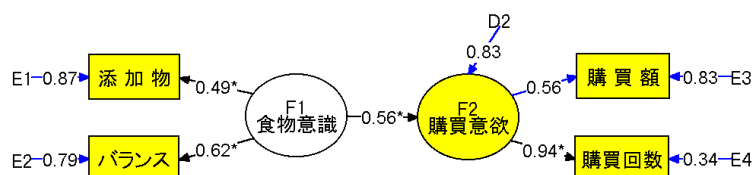


図 2.3: 多重指標モデルと推定結果

さて、ここで、従来からよく行われている「尺度化してからの回帰分析」を実行してみよう。X1+X2 を説明変数、X3+X4 を基準変数として回帰分析するのである。結果を図 2.4 (上) に示す。パス係数の推定値は 0.33 であり、0.56 と比べてずいぶん低く推定されていることが分かる。これが、誤差を分離しなかったことによる希薄化である。

尺度化するときは、できるだけ多くの項目を加算し、ここでのように 2 項目の和得点を分析することはあまりない。本例題の原典 (Homer-Kahle, 1988) では、食物意識の指標として「添加物」と「バランス」に加えて「味」と「認識」が採用されている。⁵ この 4 変数の和得点として尺度を構成し、回帰分析を実行した結果を図 2.4 (下) に示す。パス係数は 0.40 となり、図 2.4 (上) の結果である 0.33 を少し改善している。尺度構成するときは、なるべく多くの下位尺度をもってくる

⁵ 「味」は、「自然食品の味は良くない」という意見にどの程度賛成するかという項目 (逆転項目)。「認識 (perception)」は、自然食品店のイメージ (好き嫌い) という項目。心理学での定訳は「知覚」。

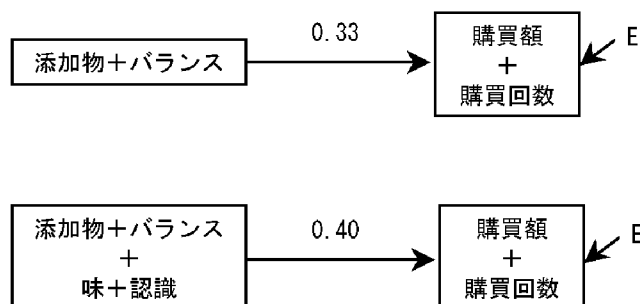


図 2.4: 尺度化したモデルの推定結果

べしという rule of thumb の意味はここにある。しかしながら、多重指標モデル（図 2.3）での推定値である 0.56 には及ばず、希薄化の修正は十分ではない。

尺度化してからの回帰分析には、希薄化の問題のほかに、モデルを吟味できないという問題もある。一般に、(重)回帰分析には適合度という概念が定義できないのである。共分散構造分析の言葉で言えば、回帰分析はいつも飽和モデルなのであり、モデルとデータのずれは、いつも 0 である。

この節を終わるにあたって、**標準解 (standardized solution)** について解説する。すべての観測変数と潜在変数の分散を 1 に標準化したときの解を標準解という。そうでない解を非標準解 (unstandardized solution) という。次節で述べる多母集団の同時分析の場合を除いて、分析結果として標準解を見ることが多い。6 科目の分析 (図 2.2) と自然食品の購買行動の分析 (図 2.3) の例は標準解を示している。標準解を見る理由は、(i) 潜在変数の尺度の不定性と (ii) 一般に観測変数間の尺度が比較可能でないことである。潜在変数の尺度は通常、その分散を 1 に固定する。探索的因子分析では、「共通因子の分散 = 1」は定義の一つである。尺度が比較可能でないときは、分散を (1 に) そろえて分析するのが常道である。

中古車価格の分析 (図 2.1) では非標準解を報告している。その理由は、「車が 1 年古くなると 13.61 万円下がる」というような、パス係数の大きさに関して解釈ができるからである。つまり、**因果の大きさ**に実質的な意味があるのである。しかし、**因果の強さ**を見るには標準解が必要になる。価格への関連が一番強い要因がどれであろうか。パス係数の大きさが一番 (絶対値) 大きい、乗車年数 (-12.67) であろうか。もし、乗車期間のデータが年数ではなく月数であったとしたら、パス係数は $-12.67/12=1.06$ となる。このようにパス係数の大きさは単位依存であり、直接の比較に意味がない。つまり、非標準解から影響の強さの情報を得ることはできない。そこで、すべて

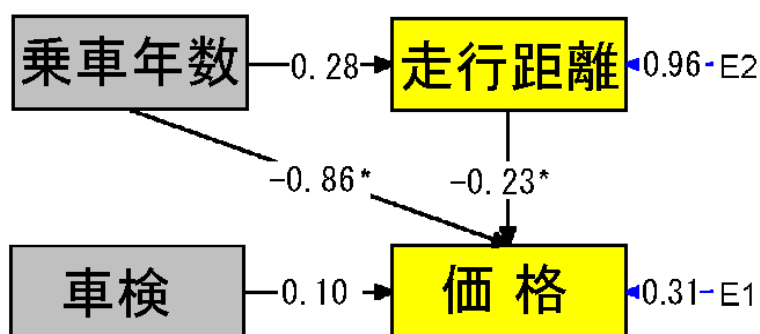


図 2.5: 中古車価格の推定結果：標準解

の分散を1に標準化した解の登場となる。中古車価格の分析の標準解を図 2.5 に示す。標準解から、価格に最も強い影響を与える要因が乗車年数であることが分かる。

標準解ではパス係数は標準（偏）回帰係数となる。従って、その値が1に近ければ、誤差変数の分散は小さく、原因変数は結果変数の変動をほとんど説明する。この意味で、因果関係は強いといえる。

影響の大きさと強さを区別することは重要である。いくら大きな影響を与えていても誤差がさらに大きければ影響の強さは小さい（相関係数は小さくなる）。標準解は、原因変数の影響の大きさを誤差の大きさと比較してその強さを評価してくれる。

2.2 モデルファイルと共分散構造

2.2.1 モデルファイル作成のコツ

共分散構造モデルを実行するには、解析者がどのようなモデルで分析したいかをソフトウェアに伝える必要がある。探索的因子分析では、分析する変数、推定方法、因子数、因子回転などを指定するだけである。しかし、共分散構造モデルは、種々多様な因果モデルが扱えるだけに、因果モデルの情報をすべて記述することになる。

パス図において、因果を表す片方矢印が一本も向けられていない変数を **独立変数 (independent variable)** という。片方矢印が一本以上向けられている変数を **従属変数 (dependent variable)** という。この定義は、観測変数・潜在変数・誤差変数のすべてに共通である。通常、誤差変数は独立変数になる。

表 2.4 に、モデルファイルの作成の要点をまとめてある。

表 2.4: モデルファイル作成の要点

- データファイルの記述
- 従属変数には方程式を作成
- 独立変数には分散・共分散（相関）を設定
- 出力の設定
- 潜在変数の尺度を定める
 - － 潜在変数の数だけ1とおくものがある
 - － 独立潜在変数の場合は「分散=1」または、「当該潜在変数から出るパス係数を一つ1に固定」
 - － 従属潜在変数の場合は「当該潜在変数から出るパス係数を一つ1に固定」

図 2.3 にある自然食品の購買行動の分析におけるモデルファイルを例にとり、モデルファイルの作成方法を解説しよう。表 2.5 に SAS CALIS でのモデル記述部分を抜き出してある。本モデルでは、観測変数はすべて従属変数である。潜在変数「食物意識 (F1SHIKI)」は独立変数、潜在変数「購買意欲 (F1YOKU)」は従属変数である。誤差変数は5つあり、すべて独立変数である。LINEQS のセクションには、先に述べた5つの従属変数に方程式を、回帰モデルの要領で作成

表 2.5: 「自然食品の購買行動」のモデルファイル (図 2.3)

```

LINEQS
TENKA   = L_11 F_ISHIKI + E1,
BARANSU = L_21 F_ISHIKI + E2,
GAKU    = 1.00 F_IYOKU  + E3,
KAISU   = L_42 F_IYOKU  + E4,
F_IYOKU = B_21 F_ISHIKI + D2;
STD
E1-E4   = DEL1-DEL4,
D2       = PSI2,
F_ISHIKI = 1.00;

```

する。L_11 などとあるのは、推定すべきパス係数（回帰係数）である。1.00 F_ISHIKI では、従属潜在変数の尺度を定めるためにパス係数を1に固定してある。STD では、独立変数の分散を指定している。F_ISHIKI=1.00 は独立潜在変数の尺度を定めるため分散を1に固定することを宣言している。本モデルには共分散はない。本モデルのモデルファイル全体とその他のモデルファイルは、本原稿「付録」にまとめて掲載してある。

独立変数における分散・共分散の設定方法について、表 2.6 で補足する。

表 2.6: 独立変数の分散・共分散

- 独立変数のすべてに、分散を未知パラメータとして設定する。ただし、潜在変数の場合はその尺度を定めるため分散を1に固定することがある。
- 誤差変数間の共分散（相関）は原則として設定せず0とする。共分散を設定する場合は、しっかりとした理由づけが必要。
- 誤差変数や攪乱変数以外の独立変数間には、原則として共分散（相関）を設定する。設定しない場合は理由が必要。

2.2.2 共分散構造

ソフトウェアが推定を行うとき必要とするのは、観測変数の分散共分散行列をパス係数などの未知パラメータで表したものである。われわれはそれを共分散構造（covariance structure）といい、しばしば $\Sigma(\theta)$ で表す。つまり、

$$\text{Var}(X) = \Sigma(\theta)$$

である。ここで、 θ は、未知パラメータをまとめてベクトルで表したものである。ソフトウェアは、モデルファイルから計算された $\Sigma(\theta)$ とデータから計算された相関行列 S を比べて、モデルがデータを再現するように、つまり、 $\Sigma(\theta)$ が S に近くなるようにパラメータ θ の値を決めてくれる。これが推定値の計算メカニズムである。

さて、未知パラメータとは具体的に何を指すのであろうか。それは、パス係数と独立変数の分散・共分散である。それゆえ、モデルでこれらを指定したのである。従属変数に係わる分散・共分散は、従属変数に対して構成される構造方程式を用いて導くことができる。つまり、従属変数の分散・共分散は、パス係数と独立変数の分散・共分散の関数となっており、新たな未知パラメータではない。

例えば、表 2.4 で従属変数である TENKA の分散は

$$\begin{aligned}\text{Var}(\text{TENKA}) &= \text{Var}(\text{L}_{11}\text{F_ISHIKI}+\text{E1}) = (\text{L}_{11})^2\text{Var}(\text{F_ISHIKI})+\text{Var}(\text{E1}) \\ &= (\text{L}_{11})^2+\text{DEL1}\end{aligned}$$

のように、パス係数 L_{11} と独立変数 $E1$ の分散である $DEL1$ の関数になっている。TENKA と BARANSU との共分散も同様に

$$\text{Cov}(\text{TENKA},\text{BARANSU}) = L_{11}\times L_{21}$$

となる。このように、従属変数にともなう方程式と独立変数の分散・共分散が分かれば、観測変数全体の分散・共分散、つまり共分散構造を導くことができる。それゆえ、これらをモデルファイルとしてソフトウェアに与えてやるのである。

2.3 モデルの修正

固有技術的観点から十分に検討され構築されたモデルも、データに合わず棄却されることがある。理屈としては、データを採る前に最大限の吟味・検討をしてあれば、モデルはデータに適合するはずである。しかし、現実問題として、データがモデルに十分適合せず、モデルを修正する必要に迫られることがある。モデルを修正して新たな知見が得られれば、データ解析をした価値があったといってもよいだろう。

考えているモデルの適合度が低いとき、常套手段は新たなパスを引いて適合度を改善することである。そのプロセスで有益な情報を与えてくれるのが、LM検定や修正指標 (modification index) である。これらの統計量は、統計的観点からどこにパスを入れるべきかを教えてくれる。パスは同時に複数個入れず一本ずつ入れていく。LM検定や修正指標は多くのソフトウェアに装備されている。一方、不必要なパスを取り除いてシンプルなモデルへ修正したいこともある。そのようなときは、パスの t -値 (ワルド検定; CR, critical ratio) を用い、有意でないパスを一つずつ消していく。

因果関係の修正ではなく、より根本的な変更として、観測変数を落とすこともある。潜在変数を測定している指標が適切でない場合は、因果モデルのアイデアは正しくとも、適合度は上がらない。この意味で、指標が適切であるかどうかの検討は重要である。探索的因子分析という限られた状況ではあるが、指標を選択するための方法論が提供されている (Kano-Harada 1999).⁶

⁶原田-狩野 (1998) は、ULR 上で実行できるプログラムを開発している。アドレスは以下のとおり：
<http://koko15.hus.osaka-u.ac.jp/~harada/factor/>

第3章

共分散構造分析で何ができる：応用編

本章では，共分散構造分析で扱えるやや高度なモデルを紹介する．

3.1 平均構造モデルと多母集団の同時分析

3.1.1 データと標準解・非標準解

共分散構造分析で拡張モデルといわれるものの代表が，平均構造モデルと多母集団の同時分析である．本節では，それらを表 3.1 の「語彙の学習データ」の分析を通して学ぶ。¹

表 3.1: 語彙の学習データ

コントロール群(n=105)				トレーニング群(n=108)			
一回目のテスト		二回目のテスト		一回目のテスト		二回目のテスト	
類似語	反対語	類似語	反対語	類似語	反対語	類似語	反対語
分散・共分散				分散・共分散			
37.626				50.084			
24.933	34.680			42.373	49.872		
26.639	24.236	32.013		40.760	36.094	51.237	
23.649	27.760	23.565	33.443	37.343	40.396	39.890	53.641
平均				平均			
18.381	20.229	20.400	21.343	20.556	21.241	25.667	25.870

データの構造は以下の通りである．11歳の被験者 213 人をランダムに 2つのグループに分け，それぞれ，コントロール群 (n=105)，トレーニング群 (n=108) とする．それぞれの群に，類似語 (synonyms) と反意語 (opposites) の試験を 1ヶ月間をおいて 2回繰り返す．従って，各群とも 4変数の観察データが得られる．ただし，トレーニング群の被験者には，間の一ヶ月に 3回，類似語と反意語のトレーニングを行う．コントロール群にはトレーニングは行わない．このようにして収集されたデータを平均ベクトルと分散共分散行列とにまとめたものが表 3.1 である．研究目的は，トレーニング効果を調べることである．

¹データの出展は Sörbom (1978)．本データの分析は多くのテキストに紹介されている．例えば，Bentler (1995, Chapter 9), Arbuckle (1995, Examples 9 and 16), 狩野 (1997)．いずれも細部で分析モデルが異なり，著者の特徴が出ていて面白い．

表 3.1 をやや詳しく見ると、1 回目のテストにおいても、トレーニング群の平均が高いようである。また、全変数を通して、トレーニング群の分散が大きいようである。理由は分からないが、被験者のランダム割り付けに失敗しているのではないかと疑われる。

このデータの構造はいわゆる二標本問題であり、実験データといってよい。2.1 節で紹介した、普通の多変量（観察）データとは状況が異なる。

このデータを分析する際、最初に頭に浮かぶのは多変量分散分析 (MANOVA) であろう。ただ、群間で分散共分散行列が異なるようで、MANOVA では分析しづらい。また、平均ベクトルに違いがあるという結論になっても、各平均に関する細かい分析がしにくい。特に、1 回目のテストの練習効果の評価をどうするか、一直線の分析では済まない。

ここでは共分散構造分析によるアプローチを紹介する。解析の目的が、「類似度テスト」と「反意語テスト」個々のトレーニング効果を見たいということであれば、一変量の分析を行えばよい。「類似度テスト」と「反意語テスト」を語彙能力や言語能力の指標として観測しているということであれば、これら2つのテストから、潜在変数である語彙能力や言語能力を導入し分析することになる。共分散構造分析が活躍するのは後者の場合である。

1 回目のテストの背後に「pre_verbal」、2 回目のテストの背後に「post_verbal」という潜在変数を想定し、「pre_verbal」から「post_verbal」へ因果のパスを引くことにする。以上がモデルの骨格である。本モデルをパス図で表すと、図 3.1 のようになる。分析結果はコントロール群を分析したものである。パス係数で 1.00 なる表記があるのは、潜在変数の尺度を定めるために 1.00 に固定されたパスを表している。²

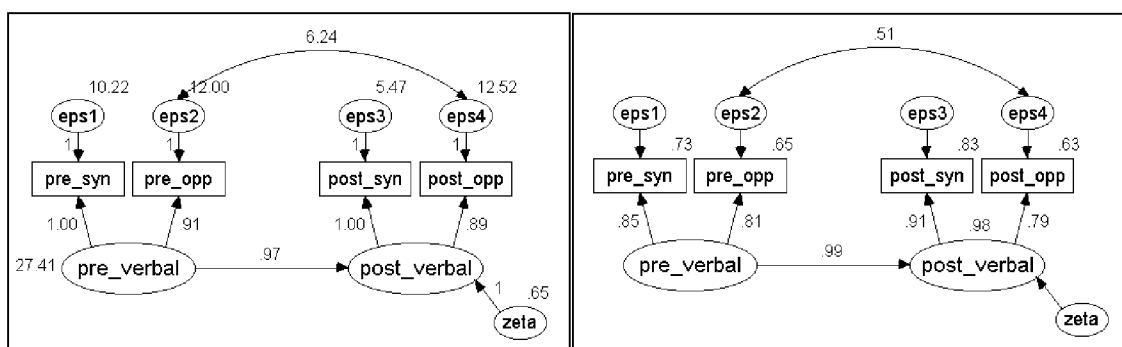


図 3.1: モデルと分析結果 (左：非標準解, 右：標準解)

パス図のモデルには誤差相関を入れてある。誤差相関を入れないモデルは適合度が悪く棄却されるのである。といっても、誤差相関は安易に許すべきものではない。誤差相関を許す根拠は、eps2 と eps4 はどちらも「反意語」に対する誤差であり、対応があるということである。ここでは、誤差相関については深入りしない。

標準解ではすべての分散が 1 に標準化されている。一方、非標準解では、2 つの潜在変数の分散も異なっている。実際、分散の推定値を計算すると以下ようになる：

$$\begin{aligned}\text{Var}(\widehat{\text{pre_verbal}}) &= 27.41 \\ \text{Var}(\widehat{\text{post_verbal}}) &= 0.97^2 \times 27.41 + 0.64 = 26.44\end{aligned}$$

post_verbal の分散がやや小さいのは、観測変数の分散が、1 回目が 37.626, 34.680, 2 回目が

²本節の分析は AMOS で行っている。他のソフトウェアとの整合性をとるため、\$useunbiased を \$ コマンドファイルに設定している。

32.013, 33.443 であることから容易に想像できる (表 3.1 を参照). 同じテストを受け同じ基準で採点しているから, 対応する潜在変数の分散を比較することに意味があるのである. 2.1 節で紹介した「文物的能力」と「数学的能力」とは比較できない. 問題が違えば採点基準も異なるからである. 分散を小さく採点したり, 平均を任意の点になるように採点することはたやすく, それゆえ, 潜在変数の比較に意味がないのである.

潜在変数から「類似語」へのパスを等置³ (共に 1 と固定) するという仮定をおいていることに注意すべきである. 語彙能力 (verbal) が「類似語」へ及ぼす影響が同一であるという仮定をおいているのである.⁴

一方, 標準解から「pre_verbal」から「post_verbal」への因果の強さは相関係数でみて 0.99 であることが分かる. この結果は, 「語彙能力 (verbal)」に関して, 一ヶ月程度の間隔をおいても成績の良し悪しは変化しないことを示している.

このように, 標準解と非標準解では得られる情報が異なる. この分析の最終目的は「pre_verbal」と「post_verbal」の群間の比較である. この目的に適しているのは非標準解である. 以後の解析は非標準解を用いて行う.

3.1.2 平均構造モデル

前節では, 非標準解 (図 3.1) を用いれば, 潜在変数の分散が比較できることを紹介した. 同じ意味で, 潜在変数の平均も比較可能である.

図 3.1 は, 平均を分析しないモデル, つまり, 平均から情報は取らないモデルであった. このモデルをより分かりやすく, 平均についての情報を加えると図 3.2 が得られる. この図での推定値等の見方を少し説明する. 例えば, 独立潜在変数「pre_verbal」には 2 つの数値 0, 27.41 が添えられている. これらは $E(\text{pre_verbal})=0$, $\widehat{\text{Var}}(\text{pre_verbal})=27.41$ を示している. 本モデ

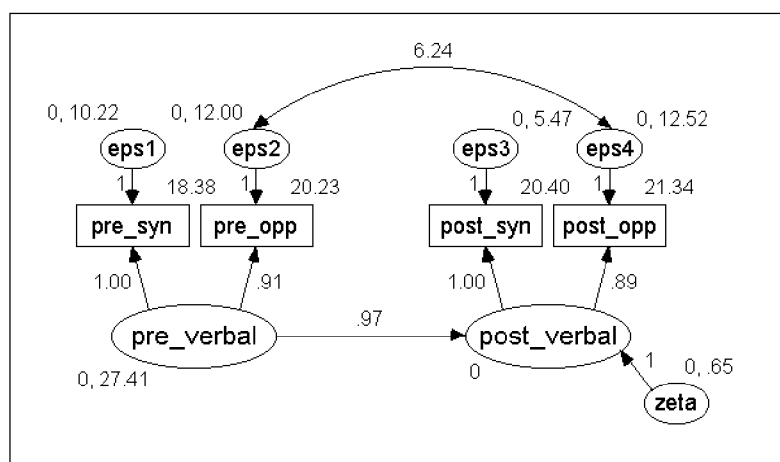


図 3.2: 潜在変数の平均 = 0

ルでは, 潜在変数と誤差変数の平均が 0 に固定されていることが分かる. 一方, 観測変数の右肩には, 入力データの平均値をそのまま表示してある. 潜在変数の平均が 0 に固定されているので,

³等しいという制約を置くこと. equate. 数学用語.

⁴パス係数を 1 と固定された観測変数を marker variable という. 「反意語」へのパスもテスト間で等置しておいた方がよいかもしれない. その方が議論が consistent であり, 潜在変数間の比較もよりはっきりするであろう. 今回はそのようなモデルは採用しないが, 受講生には一度試していただきたい.

観測変数は平均に関して、潜在変数から影響を受けていないと解釈できる。従って、観測変数の肩に付与されている平均値は、潜在変数の影響を受ける前の平均効果と考えることができる。

平均値を眺めてみると、2回目のテストの方がやや成績が良いことが分かる。同じテストを受けたことによる練習効果である。このような情報を得ることができるのは、観測変数が比較可能であるからである。

さてここでは、成績の上昇分が、語彙能力 (verbal) が向上したためであると考えてみよう。潜在変数から影響を受ける前の平均は、テスト間で等しいのであるが、2回目のテストの成績が良いのは潜在変数からの影響があるからだとするのである。そこで、post_verbal に平均の項を加えて推定する。pre_verbal の平均は0のままとする。このとき、post_verbal の平均は、pre_verbal との差を表すことになる。推定結果を図 3.3 に示す。post_verbal の平均は 1.68 と推定されてい

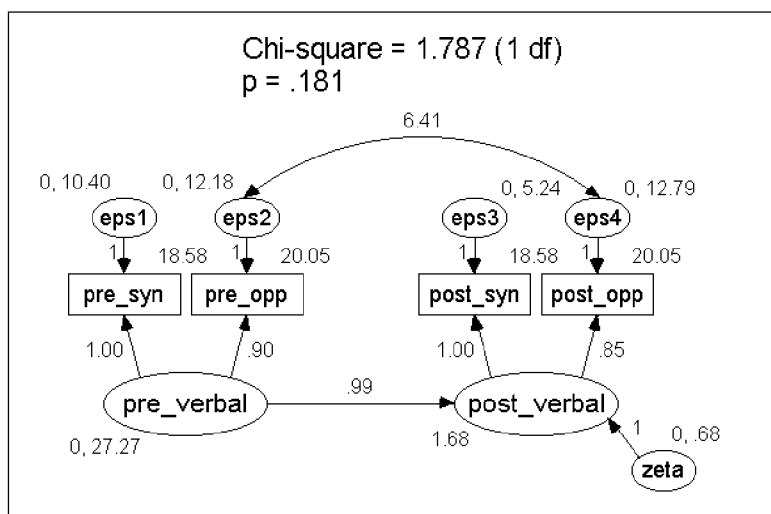


図 3.3: 潜在変数に平均を導入する

る。これは、pre_syn と post_syn を marker variable とする「語彙能力」には、1回目のテストを受けると練習効果が生じ、平均で 1.68 の上昇がみられることを示す。これを観測変数レベルで見ると

$$\text{同意語 (post_syn)} : 1.00 \times 1.68 = 1.68$$

$$\text{反意語 (post_opp)} : 0.85 \times 1.68 = 1.43$$

の上昇になる。

この効果を統計的に検定することができる。post_verbal に設定した平均を effect_c と表すと、検定すべき仮説は

$$H_0 : \text{effect_c} = 0$$

である。そのための一つの方法はワルド検定である。 $\widehat{\text{effect_c}} = 1.68$ を、その $SE = 0.30$ で除して $1.68/0.30 = 5.57$ を得る。これを標準正規分布の上側 2.5 % 点である 1.96 と比較することにより有意性が示される。もう一つの方法は、 $\text{effect_c} = 0$ とおいたモデルの適合度を調べることである。表 3.2 に $\text{effect_c} \neq 0$ としたモデル (図 3.3 のモデル) と $\text{effect_c} = 0$ とおいたモデルのカイ 2 乗値を報告している。表 3.2 で、「母数の数」とは、モデルにおいて推定すべき母数 (パラメータ) の数を表す。この検定では、 $\text{effect_c} = 0$ としたモデルが棄却され、「練習効果あり」という結論になる。

表 3.2: 練習効果の検定

モデル	母数の数	カイ 2 乗値	df	P-値
effect_c \neq 0	13	1.787	1	0.181
effect_c = 0	12	30.038	2	0.000

共分散構造分析を適用したい状況では、多くの場合、興味の対象は観測変数ではなく構成概念を実現する潜在変数である。平均構造モデルは、成長、変化、違いなどを構成概念のレベルで記述してくれる。注意しておきたいのは、ここでのように潜在変数が比較可能でないといけないことである。同一の内容を測定している潜在変数は、多くの場合比較可能である。1つの潜在変数の平均を0とおく。典型例が、ここで扱ったような反復測定データや縦断的データである。

比較可能な潜在変数のもう一つの典型例は、男女や国籍など属性間での比較である。例えば、男女の潜在変数の平均差を評価することができる。このような分析は、多母集団の同時分析で実行できる。

3.1.3 多母集団の同時分析

多母集団の同時分析とは、いわゆる、多変量解析における多標本問題である。いくつかの母集団間で分散共分散行列や平均ベクトルの比較を行うわけであるが、一般的な相等性の検定だけではなく、各母集団における（パス図で記述された）共分散構造モデルの全体もしくは一部の相等性が検定できる。

本節では、表 3.1 のデータを用いて、コントロール群とトレーニング群の2群比較を行う。特に、練習効果を除いた（調整した）後でのトレーニング効果を調べることを最終目的とする。

図 3.1 のモデルで両群を個別に分析した結果を図 3.4 に示す。ここでは平均構造は考えていない。両群では、トレーニングを施したかどうかという点が違うだけである。従って、いくつかの

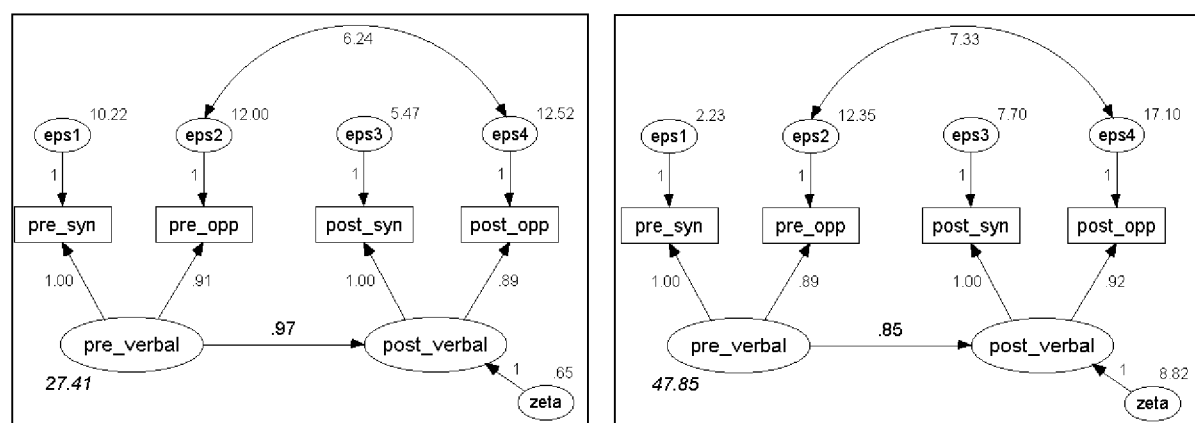


図 3.4: 同一モデルによる個別分析 (左: コントロール群, 右: トレーニング群)

パス係数や分散・共分散が、群間で等置しても大きな損失はないかもしれない。多母集団の同時分析では、どの母数に等置制約を課してもよいかを統計的に調べるのが主要なプロセスになる。

等置可能性を調べる第一歩は、測定モデルの同等性、つまり、潜在変数から観測変数へのパス

表 3.3: 同時分析結果（上：平均構造がないモデル，下：平均構造があるモデル）

検討すべきモデル	推定する母数の数	df	モデルのカイ2乗検定		RMSEA	AIC	付加された等置制約の直接検定		
			χ^2 -値	P-値			df	χ^2 -値	P-値
等置制約なし	20	0	0.00	1.00	0.00	00.00	—	—	—
因子負荷が等しい	18	2	0.15	0.93	0.00	-3.85	2	0.15	0.93
因子間パス係数が等しい	17	3	1.31	0.73	0.00	-4.69	1	1.16	0.28
Var(pre_verbal) が等しい	16	4	6.05	0.20	0.05	-1.95	1	4.74	0.03
Var(zeta) が等しい	15	5	15.92	0.01	0.10	5.92	1	9.88	0.00

検討すべきモデル	推定する母数の数	df	モデルのカイ2乗検定		RMSEA	AIC	付加された等置制約の直接検定		
			χ^2 -値	P-値			df	χ^2 -値	P-値
等置制約なし	25	3	3.04	0.39	0.01	-2.96	—	—	—
因子負荷が等しい	23	5	3.28	0.66	0.00	-6.72	2	0.24	0.89
因子間パス係数が等しい	22	6	4.56	0.60	0.00	-7.44	1	1.28	0.26
Var(pre_verbal) が等しい	21	7	9.27	0.23	0.04	-4.73	1	4.71	0.03
Var(zeta) が等しい	20	8	19.35	0.01	0.08	2.35	1	10.08	0.00

係数（因子負荷）が群間で等しいかどうかの検定である。次いで、潜在変数に係わる母数の相等性を調べていく。一連の検定結果を表 3.3 に示す。この表において、Var(zeta) とは、post_verbal に付随する攪乱変数（誤差変数）の分散である。

表 3.3 から、因子負荷と因子間パス係数が群間で等しいというモデルが妥当であろうと結論できる。このことは、Var(pre_verbal) が群間で異なると判断することであり、被験者のランダム配置がうまく実行できなかったことを統計的に示している。因子負荷の相等性が成り立つとき、モデルは測定不変（measurement invariant）であるという。

採用されたモデル（平均構造あり）での推定結果を図 3.5 に報告してある。

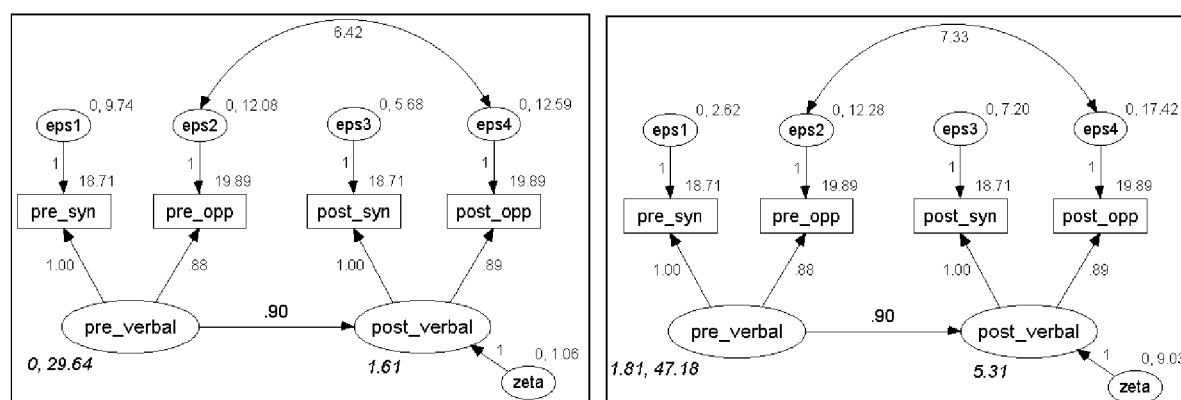


図 3.5: 同時分析最終結果（左：コントロール群，右：トレーニング群）

まず、図 3.5 でのモデルにおいて、平均構造がどのように入れられているかをみておく。3.1.2 節での個別分析では、潜在変数の影響を受ける前の観測変数の平均を 1 回目と 2 回目とで等置することで、潜在変数の平均差を評価した。そうすることで、練習効果が $E(\text{post_verbal})$ に現れたのである。

同様のテクニックを群間にも使う。潜在変数の影響を受ける前の観測変数の平均を、コントロール群とトレーニング群とで等置する。観測変数の平均の違いを、潜在変数の平均の違いに押し込むのである。

トレーニング群の $E(\text{pre_verbal})$ は 1.81 と推定されている。これは、被験者を 2 つの群に割り付けたときの語彙能力差である。先に pre_verbal の分散に違いがあることを指摘したが、平均にも違いが見られるようである。

目的のトレーニング効果は、 post_verbal に係わる平均から検討することができる。図 3.5 での潜在変数の平均についての結果を表 3.4 にまとめてあるのでご覧きたい。

表 3.4: 平均構造とトレーニング効果

群	$E(\text{post_v})$		切片項		パス係数		$E(\text{pre_v})$
トレーニング群	6.94	=	5.31	+	0.90	×	1.81
コントロール群	1.61	=	1.61	+	0.90	×	0
差	5.33		3.70				1.81

平均構造の入れ方は、独立変数と従属変数で異なる。独立変数にはその平均を設定するが、従属変数には切片項を入れる。つまり、因子間の関係は、構造方程式で表せば

$$\text{post_verbal} = \alpha + b \times \text{pre_verbal} + \text{zeta} \quad (3.1)$$

となる。ここで b はパス係数、 zeta は誤差項である。独立変数である「 pre_verbal 」に対しては、その平均 $E(\text{pre_verbal})$ が推定され出力される。一方、従属変数に対しては、構造方程式 (3.1) における切片項 α が推定・出力される。パス図に出力されている 5.31, 1.61 は、潜在変数の平均の推定値ではなく、切片項の推定値である。切片項は、独立変数の影響を受ける前の平均効果と解釈できる。従属変数の平均 $E(\text{post_verbal})$ は、構造方程式 (3.1) を用いて表 3.4 のように計算できる。

観測変数の平均で、潜在変数の影響を受ける前の平均という術語を使った。これは、切片項のことである。このモデルでの観測変数は、すべて従属変数である。

さて、最終目的であるトレーニング効果の検討に移ろう。再び、表 3.4 をご覧いただきたい。2 回目の語彙能力の平均 ($E(\text{post_verbal})$) が、両群で

$$5.33 = 6.94 - 1.61$$

だけ異なることが分かる。トレーニング群が 5.33 だけ高いのである。しかし、この差には、被験者の割り付けによる違いも含まれている。つまり、トレーニング群の方は、トレーニングをする前から、1.81 だけ能力が高かったのである。この効果を取り除いた純粹のトレーニング効果が、切片項の違い

$$3.70 = 5.31 - 1.61$$

表 3.5: 平均効果の検定

モデル	推定する母数の数	df	モデルのカイ2乗検定		RMSEA	AIC	付加された等置制約の直接検定		
			χ^2 -値	P-値			df	χ^2 -値	P-値
最終モデル (図 3.5)	22	6	4.56	0.60	0.00	-8.44	-	-	-
post_verbal の切片項が群間で等しい	21	7	56.58	0.00	0.18	42.58	1	52.02	0.000
E(pre_verbal) が群間で等しい	21	7	8.58	0.28	0.03	-5.42	1	4.02	0.045

に現れている。

群間で、「E(pre_verbal) が異なる」「post_verbal の切片項が異なる」ということが推定値上で示された。これらは統計的検定により確認することができる。検定結果を表 3.5 に示す。トレーニング効果がない、つまり、post_verbal の切片項が群間で等しいモデルは、明らかに棄却される。一方、被験者がランダムに両群へ割り付けられているという仮説である「E(pre_verbal) が群間で等しい」というモデルは受容されている。が、付加された等置制約の直接検定では棄却される。

3.1.4 簡便法

前節で紹介した「平均構造のある多母集団の分析」はやや高度であり、共分散構造分析をかなり習熟している必要がある。本節では、トレーニング効果を検討するための簡便法を紹介する。簡便であるがゆえの落とし穴もある。簡便法を適用するための注意点も併せて習得されたい。

トレーニングの有無を二値変数 training で表し、値を 1, 0 とする。training を加え 5 変数の分散共分散行列を全被験者 (n=213) を用いて作成する (表 3.6)。

表 3.6: 両群を合併したデータ

subject	pre-syn	pre-opp	post-syn	post-opp	training
1	15	18	20	25	0
2	18	17	16	27	0
3	25	27	30	32	0
...
105	15	20	24	20	0
106	14	23	28	16	1
...
213	19	20	23	24	1

pre-syn	pre-syn	pre-opp	post-syn	post-opp	training
	1				
pre-opp	0.783	1			
post-syn	0.782	0.693	1		
post-opp	0.704	0.774	0.776	1	
training	0.163	0.078	0.379	0.325	1
standard deviation	6.687	6.499	6.950	6.957	0.500

前節まで検討してきたモデルに、training が post_verbal に効果を及ぼすというモデルを付け加える。分析結果を図 3.6 に示す。適合度は良いようである。training から post_verbal へのパス係

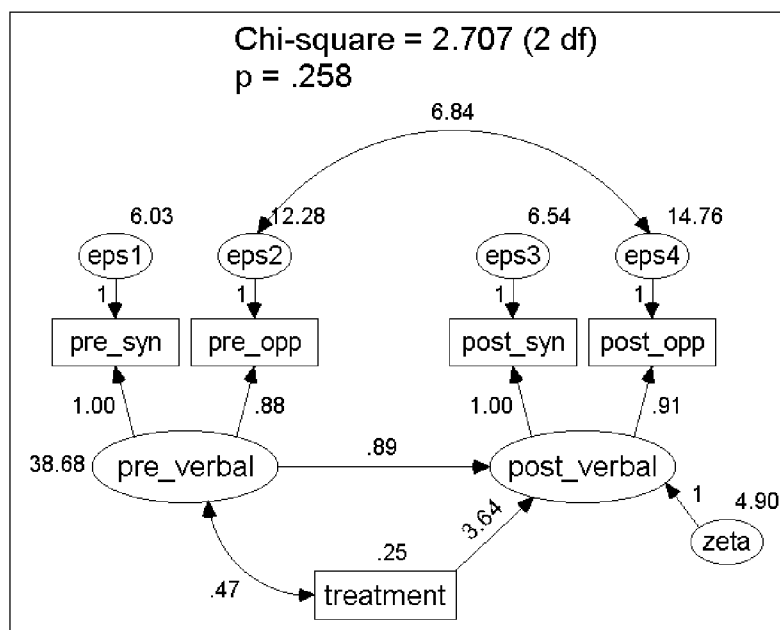


図 3.6: 簡便法による分析結果

数が 3.64 と推定されている。training が 0, 1 の値を取る変数であるから、training の効果として 3.64 が期待できることを示す。前節の複雑な分析の結果は 3.70 であったから、遜色ない値が得られている。

では、トレーニング群の方へやや語彙能力が高い被験者を割り付けたという、割り付けの偏りの影響はどこへいったのだろうか。割り付けに偏りがあることは、pre_verbal と training との間に相関が認められることから容易に想像できる。training=0 のときの pre_verbal の条件付き平均は 0 であるが、training=1 のときは

$$E(\widehat{\text{pre_verbal}} | \text{training}=1) = 0.47/0.25 = 1.88$$

と推定される。多母集団の分析からは 1.81 が得られており、十分近い値といえよう。

簡便法の分析において、割り付けの偏りがトレーニング効果の推定に偏りを生じさせていないのだろうか。この問題も OK である。というのは、training から post_verbal へのパス係数は、回帰分析の偏回帰係数であり、pre_verbal の影響を取り除いた下での推定値だからである。

以上のことから、この簡便法は、複雑な平均構造のある多母集団の同時分析に取って代わりうる、極めて巧妙な分析方法であると考えられよう。

しかし、実は、この簡便法には落とし穴がある。中級レベルの共分散構造分析ユーザーが気になるのは、training のような実験条件を表す変数（グループ変数）を、何の注意もなく分析に加えてもよいのかという問題であろう。

この問題については明快な回答が得られている。考えている共分散構造モデルにおいて、二値変数やグループ変数が**独立変数**であるならば、ある条件の下で、連続変数と考えて分析してよいのである。

問題は「ある条件」とは何かである。実は、このような分析が正当化されるためには、連続変数の（条件付）分散共分散行列が、二値変数やグループ変数の値によらず一定であることが要求される。これが「ある条件」である。今回の例に戻れば、トレーニングの有無で、テストの結果である 4 変数間の分散共分散行列が変化しないことが要求される。そしてこの仮定は偽であること

が、多母集団の分析のときに明らかになっているのである。図 3.6 に報告されている「カイ 2 乗値=2.707」は、4 変数間の分散共分散行列が変化しないということを前提にした値であって、そうでないときには信用できない。

以上をまとめる。この節で紹介した簡便法はきわめて巧妙な分析法である。独立変数である限りグループ変数をモデルの中に組み込んでもよいが、母集団間で分散共分散行列が均質であることが要求される。本例題では残念ながらこの条件が満たされない。よって、このデータに対しては簡便法は不適切な分析方法と言わざるを得ない。

多母集団の同時分析は、母集団間で部分的に構造が違っていても適用することができるのが大きなメリットであり、簡便法より複雑であったとしても、その存在価値は大である。

5.3 節に、本節の内容に関する数理的解説があるので参考にされたい。

3.2 実験データの共分散構造分析

3.1.3 節でも指摘したように、多母集団の同時分析は簡単な実験データの分析である。“簡単な”という意味は、比較する母集団の数が少ないということである。実際、この例では比較する母集団は二つである。一般の実験データは比較する母集団の数がもっと多い。例えば、要因が二つで、各要因が3水準であるとすれば、計9つの母集団を比較することになる。

たかが二つの母集団比較でこんなに複雑な分析をしないといけないのであれば、一般の実験計画によって採られたデータの分析はかなり厄介だと予想されるであろう。さらに、共分散構造分析は相当の標本サイズを各母集団に要求するから、そんなにデータが採取できるのかという問題も生ずる。

そこで注目されるのが、3.1.4 節で紹介した簡便法である。処理の効果を 0, 1 で表し共分散構造分析へ持ち込むというアプローチは非常に簡明である。群間で分散共分散行列が等質であるという仮定が必要であるが、このような実験データの共分散構造分析はもっと注目されてよい。

あまりスペースがないので、ここでは、やや本格的な実験データの解析事例を、簡単に紹介する。図 3.7 に示したパス図は、Churchill-Surprenant (1982) による顧客満足度の決定要因を探るモデルとその分析結果である。⁵ 実験では、ビデオディスクプレーヤーと菊の花（鉢植え？）が提

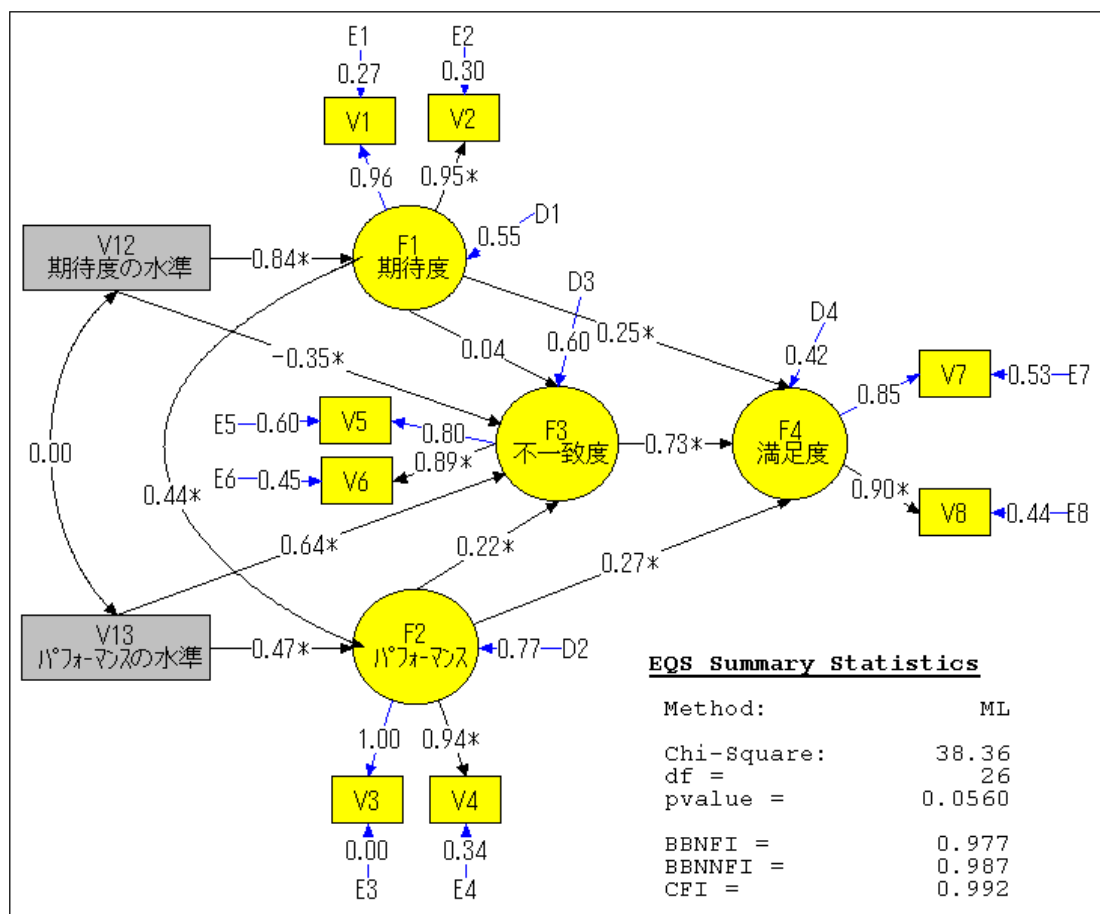


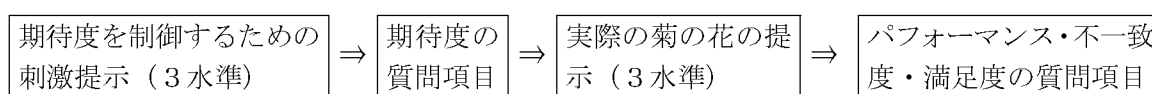
図 3.7: 実験データの共分散構造分析：顧客満足度の分析

示され、その満足度などが調査された。ここでは、菊の花のデータ分析のみを紹介する。

⁵Churchill-Surprenant (1982) の分析モデルとはやや異なっている。

満足度 (F4) を規定する要因として、商品への期待度 (F1)、商品のパフォーマンス (性能 F2)、そして、期待度とパフォーマンスの不一致度 (F3) を取り上げている。不一致度を取り上げているのは、商品が期待以上のパフォーマンスであれば満足度が高まるという仮説に基づくものである。概念的には、「不一致度=パフォーマンス-期待度」であるが、実際には差で定義するのではなく、「被験者自身の期待度は、実際の商品と比べて高かったですか」という質問項目で評価している。

このデータの最大の特徴は、期待度とパフォーマンスを実験条件でコントロールしていることにある。商品の期待度をコントロールするため、「菊の花というものはこんなものだ」という3種のレベルの事前情報を用意し、各レベルの情報に被験者をランダムに割り付けている。一方、パフォーマンスについては、3種のレベルの菊の花を実際に用意し、各レベルの菊の花に被験者をランダムに割り付けている。従って、このデータは3×3要因計画による実験データとなる。被験者は各セルに20名が割り付けられ、トータルの被験者数は180である。以下に実験の流れを示す。



Churchill-Surprenant (1982) では、刺激提示が期待どおりの成果を上げているかをチェックするために、分散分析を行っている。分散分析の結果、ならびに各水準における点推定値と信頼区

表 3.7: 分散分析 (Churchill-Surprenant, 1982, Table 2 より再掲)

要因	分散分析					各水準における推定		
	平方和	自由度	SS	MS	F	水準	推定値	95%信頼区間
期待度	級内	2	288.24	144.12	161.28**	低	3.25	2.94-3.55
	級間	177	158.17	0.89		並	5.03	4.81-5.25
	合計	179	446.41			高	6.34	6.14-6.53
パフォーマンス	級内	2	249.03	124.52	24.82**	低	8.13	7.57-8.68
	級間	177	887.92	5.02		並	9.80	9.20-10.40
	合計	179	1136.96			高	10.99	10.41-11.57

** は 0.1%有意を示す。

間を表 3.7 に示す。要因「期待度」と「パフォーマンス」はもちろん有意であり、母平均の信頼区間が重なっていない。これらのことから「期待度」と「パフォーマンス」は、実験者の意図どおりに制御されていることが分かる。また、推定値は、水準ごとに等間隔に並んでいるとみなして大きな損失はないであろう。

共分散構造分析は以下のように実行される。実験条件を表す変数「V12: 期待度の水準」「V13: パフォーマンスの水準」には、それぞれ、1, 2, 3 のスコアを与える、これらの変数と質問紙の項目とを合わせて相関行列を作成する。そして、図 3.7 のパス図のモデルで分析する。

V12 と V13 は固定変数であり、他の調査項目と一緒に連続変数と考えて分析するのは、一見抵抗がある。しかし、前節で述べたように、V12 と V13 は独立変数であるから、そのような分析をする罪は重くない。

図 3.7 に示しているのは、推定値 (標準解) と適合度指標の結果である。本モデルの適合度は悪くない (カイ 2 乗値=38.86, df=26, P-値=0.0560)。

この分析から以下のことが理解される：

1. 本来は、9 個の母集団の分散共分散行列の相等性をチェックしておかなければならない。しかしながら、各母集団の標本サイズが 20 であることを考えると相等性のチェックは難しいであろう。状況は、観察データの回帰分析で、誤差分散の相等性がチェックし難いのと似ている。
2. 満足度を規定する要因としては、「不一致度」が一番高い。全変動の約半分 ($0.73^2 = 0.53$) が不一致度によって説明できる。
3. 期待度とパフォーマンスをコントロールするために導入した固定変数 V12 と V13 は、期待度とパフォーマンスだけでなく、不一致度にも影響する。
4. 固定変数からの影響とは、平均に影響を及ぼすことを示す。例えば、V12 から F1 へのパス係数は 0.84 と推定されている。この意味は、V12 が 1 単位動いたとき、期待値 $E(F1)$ が 0.84 だけ動くことを示す。
5. 「不一致度」は「パフォーマンスー期待度」に近いはずである。従って、V12 からのパス係数はマイナス、V13 からはプラスに評価されている。
6. E3 の分散が 0 として推定されている。不適解 (improper solution) である。不適解は報告しづらいが、最終解として解釈できるときもある。3.3 節の「注意」をみよ。
7. 本データを、実験条件を表す固定変数 V12 と V13 を入れずに分析するのは誤りである。表 3.7 に示してあるとおり実験条件ごとに平均が異なるので、それらを調整せずに分析してはいけない。平均が異なる異質なデータを混在させると相関係数が適切に推定できない。
8. 本モデルは、9 個の実験条件において、共分散構造は同一で平均が異なるモデルを当てはめていることになる。

多変量分散分析 MANOVA では、各セルにおける母平均ベクトルの違いに興味がある。通常、各セルでの母分散共分散行列は等質であることが仮定されている。母分散共分散行列に構造は導入しない。一変量の分散分析でも事情は同じである。ここで紹介した実験データの共分散構造分析では、母分散共分散行列は等質であるが、因果モデルを構築することで母分散共分散行列に構造を導入する。構築された因果モデルは、平均ベクトルの違いをも記述している (Bagizzi and Yi 1989)。

5.3 節に、本節の内容に関する数理解説があるので参考にされたい。

3.3 潜在曲線モデル

最近注目を浴びている応用モデルが潜在曲線モデル (latent curve model) である。いくつかのソフトウェアは、潜在曲線モデルを実行しやすくするオプションを備えている (e.g., EQS)。今後、適用例が急速に増加する方法論と考えられる。

3.3.1 例題と分析結果

表 3.8 のデータを見てみよう。⁶ 10 歳から 15 歳までの青少年のアルコール消費回数が 3 時点 (1 年間隔: T1, T2, T3) で調査されている。同一被験者に 3 回の調査を行う縦断的データである。この変数には、過去 1 年間での、ビール・ワインやウィスキーを飲んだ回数、泥酔した回数などが下位尺度に含まれている。それらの平均は、1.36, 2.12, 3.18 と (線形に?) 増加することが見て

表 3.8: 青少年のアルコールのデータ n=363

	X1	X2	X3	X4	X5	X6
X1: T1 におけるアルコール消費回数	1.00					
X2: T2 におけるアルコール消費回数	.68	1.00				
X3: T3 におけるアルコール消費回数	.50	.68	1.00			
X4: 年齢	.31	.31	.23	1.00		
X5: 性別	.01	.06	.10	.14	1.00	
X6: 親のアルコール依存	.12	.17	.20	.12	-.15	1.00
平均	1.36	2.12	3.18	12.91	0.52	0.56
標準偏差	2.81	3.98	4.79	1.41	0.50	0.50

取れる。残りの 3 つの変数は被験者の属性を表す。「親のアルコール依存」とは、二値変数で、両親もしくは片親がアルコール依存症と診断されているならば 1、そうでなければ 0 の値をとる。

分析の目的は、「増加するアルコール消費回数」が、「X4: 年齢」「X5: 性別」「X6: 親のアルコール依存」とどのような関係があるかを調べることである。「増加するアルコール消費回数」は、直線を当てはめるならば、切片と傾きで記述できる。切片や傾きの違いが、変数 X4 ~ X6 によってどのように説明できるかがポイントになる。このとき、切片と傾きは、回帰分析の基準変数であり確率変数とみなすことになる。

この分析には、時間にアルコール消費回数を回帰する回帰分析と、変数 X4 ~ X6 に切片項と傾きを回帰する回帰分析があり、回帰分析を二回行うことになる。このような分析を実行する他の方法として MLM(Multilevel model) や HLM(hierarchical linear model) がある。⁷

まず、最初に、Curran (1999) による分析結果を見てみよう。彼は図 3.8 のモデルにて分析を行った。適合度は以下の通りで、本モデルはよい適合を示している：

$$\begin{aligned} \text{カイ 2 乗値} &= 5.530 (\text{P-値} = 0.237, \text{df} = 4, \text{n} = 363) \\ \text{GFI} &= 0.996, \text{CFI} = 0.997, \text{RMSEA} = 0.033 \end{aligned}$$

図 3.8 の推定結果から以下のことが分かる：

⁶出展は Curran (1999).

⁷例えば、Bryk-Raudenbush (1992) や Goldstein (1995) を見よ。

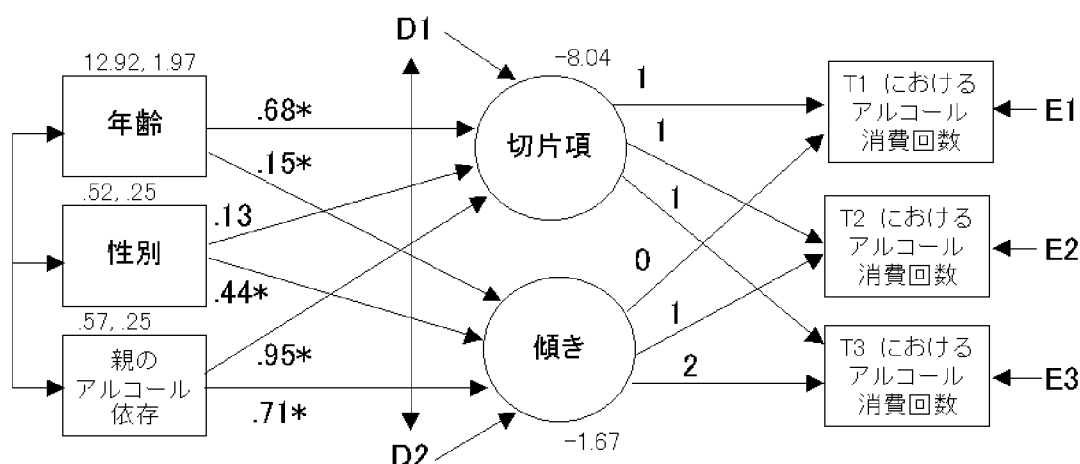


図 3.8: 潜在曲線モデル

共分散と誤差分散の推定値は略してある。有意性を表す*はパス係数にのみ付与してある。

年齢や性別、環境（親のアルコール依存）によって、T1 時点のアルコールの消費回数（切片項）・増加の割合（傾き）は変化する。特に

1. T1 時点で年齢が 1 つ上だと消費回数が 0.68 増加し、経年による増加量は傾きで 0.15 増す。
2. 性別は T1 時点での消費回数には有意に影響しないが、経年による増加量には影響し、傾きで 0.44 だけ男子が高い。
3. 親がアルコール依存であるかどうかは、T1 時点での消費回数 (0.95) にも経年による増加量 (0.71) にも影響をもつ。

注意: 実は、図 3.8 の解は不適解である。パス図には書いていないが、E1 の分散が 0 と推定されている。不適解の扱いは難しい。ここでは以下の理由により、この不適解の原因は標本変動であろうと判断し、境界解 (i.e., $\text{Var}(E1) = 0$) を報告している。

1. $\text{Var}(E1) \geq 0$ の条件を外して推定したところ、 $\widehat{\text{Var}}(E1) = -.127 (SE = .698)$ となった。この推定値は、 $\widehat{\text{Var}}(E2) = 5.264 (SE = .562)$, $\widehat{\text{Var}}(E3) = 4.648 (SE = 1.221)$ と比較してかなり小さな値である (括弧内は推定値の標準誤差 SE を表す)。
2. $H_0: \text{Var}(E1) = 0$ を SE を用いてワルド検定すると $z = -.127/.698 = -.182$ となり、仮説は棄却されない。
3. 適合度が良いので、モデルがデータに合わないことによる不適解ではない。
4. 識別性が達成されていないなどモデルに内在する問題による不適解ではない。その理由は以下のとおり：
 - (a) この推定値以外に異常な値として推定されている母数はない。
 - (b) $\widehat{\text{Var}}(E1)$ の SE は 0.698 であり、他の SE と比較しても異常に大きな値ではない。

不適解を報告するには勇気がある。どうしても不適解を最終的な解としたいときは、最低上記程度の考察が必要である。探索的因子分析の枠組みにおける不適解の原因とその処理については、狩野 (1997) を参照されたい。

3.3.2 考え方

本節では、潜在曲線モデルの基本的な考え方を解説する。時点 t におけるアルコール消費回数を x_t で表す。前節の例では、 $t = 1, 2, 3$ である。⁸ 調査時点に対して x_t をプロットし、直線を当てはめる。切片項と回帰係数を β_0, β_1 とすると、

$$x_t = \beta_0 + (t - 1)\beta_1 + e_t \quad (t = 1, 2, 3) \quad (3.2)$$

となる。⁹ここで e_t は直線からのずれを表す誤差項である。被験者を表す添え字を付加すれば、

$$x_t^{(i)} = \beta_0^{(i)} + (t - 1)\beta_1^{(i)} + e_t^{(i)} \quad (t = 1, 2, 3; i = 1, \dots, 363)$$

となる。このように書くと、切片項や傾きが被験者ごとに変化するということがよく分かる。しかし、この表現はやや複雑なので、以後は、式 (3.2) で議論を進める。

式 (3.2) は β_0, β_1 を共通因子とする因子分析モデルの形をしている。行列表記すれば

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} \quad (3.3)$$

となり、 $[\beta_0, \beta_1]'$ を共通因子、その前にある $0, 1, 2$ からなる定数行列が因子負荷量であることが分かる。図 3.8 において、切片項と傾きから消費回数へのパス係数が $1, 1, 1$ と $0, 1, 2$ の固定された値が割り振られているのは、(3.3) の因子負荷量からきている。

因子分析モデルでは、共通因子は確率変数とすることが多い。共分散構造モデルの潜在変数は確率変数である。ここでの β_0, β_1 も確率変数である。切片項や傾きを確率変数と考えるのには少々抵抗があるかもしれない。アルコール消費回数は個人によって異なる。それゆえ、切片項・傾きは被験者個人に貼りついている量と考えられる。この意味で知能や学力と同じであり、切片項・傾きを確率変数と考えるのに抵抗感を持つ必要はない。

参考のため、観測時点が5つで等間隔の場合に2次式を当てはめるときのモデル式を以下に紹介しておく：

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

式 (3.3) のモデルをパス図で表すと図 3.9 になる。このパス図で β_0 と β_1 とに相関を許していることに注意する。(誤差変数を除く) 独立変数間には通常、相関を設定するという基本ルールに則ったものである。実際、調査開始時点 (T1) でアルコール消費回数が多い被験者は、その後の消費回数の増加率も高いとか、逆に、十分消費しているからあまり伸びないなどの可能性があり、切片項と傾きは何らかの相関が認められてもおかしくはない。通常因子分析との相違点は、確率変数 β_0, β_1 の平均が 0 ではないことである。平均 $E(\beta_0)$ は調査開始時点 (T1) での被験者の平均アルコール消費回数を表し、平均 $E(\beta_1)$ は、平均伸び率 (1年に伸びる消費回数) を表している。パス図にはその意味で $E(\beta_0)$ と $E(\beta_1)$ を入れてある。図 3.9 のモデルで分析するには、平均構造のある共分散構造モデルを用いることになる。

⁸ 調査地点が等間隔でない場合は、 $t = 1, 4, 5$ などとする。

⁹ $x_t = \beta_0 + t\beta_1 + e_t \quad (t = 0, 1, 2)$ と書く方が分かりやすいかもしれない。共分散構造分析では、変数は $X_1, X_2 \dots$ と (0 からではなく) 1 から始めるのが普通なので本文のようにした。

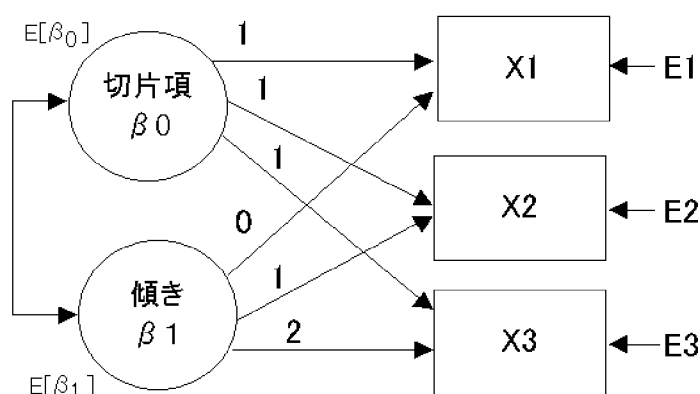


図 3.9: 式 (3.3) のモデル

より興味ある分析は、被験者ごとに異なる切片項や傾きが、どのような要因によって左右されているのかを調べることである。年齢によって異なるのか、男女差があるのか、また、環境（両親のアルコール依存）が関係してくるのか。このような変数を説明変数に配し、切片項と傾きを基準変数にしたパス解析を実行すると、有益な情報が得られる可能性がある。これを実行したのが、前節で紹介した図 3.8 での解析である。

Curran (1999) では、図 3.8 でのモデルに、「友人の影響」を加えたモデルで分析している。最終結果のパス図だけを図 3.10 に紹介する。

本節で紹介した分析は、観測変数に当てはめる直線や曲線の係数を潜在変数に見たてるところから、潜在曲線分析 (latent curve analysis) とよばれている。潜在成長曲線分析 (latent growth curve analysis) という用語が用いられることもある。共分散構造分析の枠組みで、このような分析ができることを指摘したのは Meredith-Tisak (1990) である。Chou-Bentler-Pentz (1995) には、二次関数を用いた潜在曲線モデルの例が紹介されている。また、歴史的によく用いられてきた MLM, HLM との興味ある比較も行われている。参考にされたい。

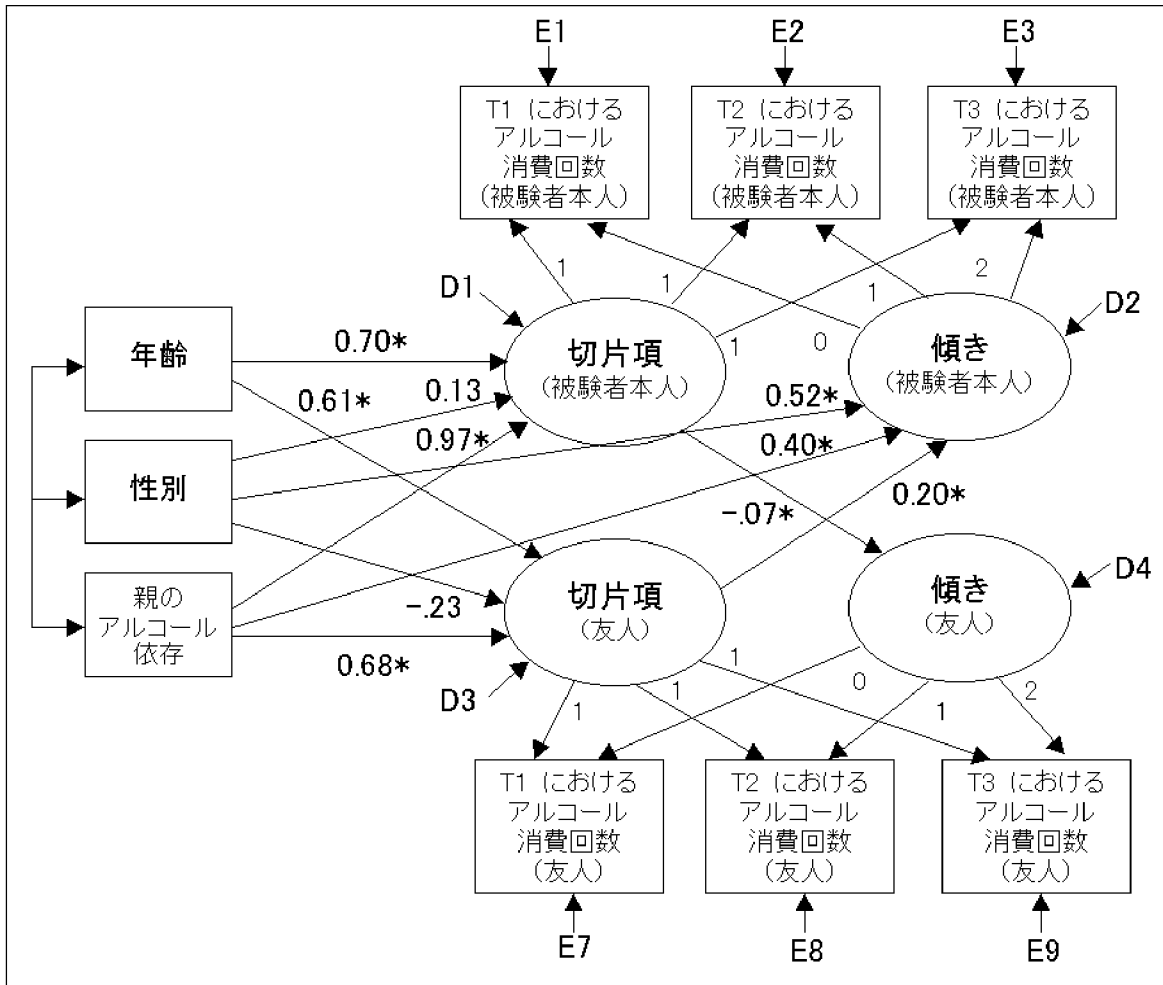


図 3.10: Curran (1999) の解析 : 「友人のアルコール消費回数」を加えた分析
 攪乱変数間と誤差変数間の共分散の推定値は以下のとおりである：
 $Cov(D1, D2) = -.95^*$, $Cov(D2, D4) = 0.55^*$, $Cov(D3, D1) = 2.30^*$, $Cov(D4, D3) = -.09$
 $Cov(E1, E7) = 0.04$, $Cov(E2, E8) = 0.73^*$, $Cov(E3, E9) = 0.71^*$

第4章 因果分析について

共分散構造分析は、ときに因果構造分析とも言われる。複雑に絡んだ因果関係をズバリ決定するエースというプロパガンダ的宣伝文句もあるようである。

結論を先に申し上げると、「AならばBが証明された」というような強い結論を、共分散構造分析だけから求めるのは難しい。データを共分散構造分析する際、因果推論に関して何が仮定され、どのような結論が引き出せるのかを事前に検討し、モデルのデザインを組む必要がある。本章では、そのようなことに資する情報を紹介する。

因果推論には、大きく分けて二つの側面がある。一つは因果の方向を如何にして探るかという問題であり、他の一つは、因果の方向は既知であるという状況で、因果の大きさを如何に正確に（バイアスなく）測定するかという問題である。「やる気」と「成績」ではどちらが原因でどちらが結果か、という状況は前者の問題であり、喫煙が肺がんにどのような影響を及ぼすか、という状況は後者である。前者の問題について統計学辞典（竹内啓 1989）をひもといてみると、因果関係の決定には (i) 時間的先行性 (ii) 関連性の強さ (iii) 直線的関係 (iv) 関連の一致性・普遍性 (v) 関連の整合性、を十分に検討すべしとある。後者の問題で重要なのは、いわゆる第三変数の影響である。

4.1 因果の方向を探る

共分散構造分析によって因果を同定するための道具は、モデルの適合度である。因果モデルAが受容され、対抗馬である因果モデルBが棄却されれば、モデルAの因果関係が立証されたと考えようというロジックである。この論理に立ちはだかる問題の一つは同値モデルの存在である。二

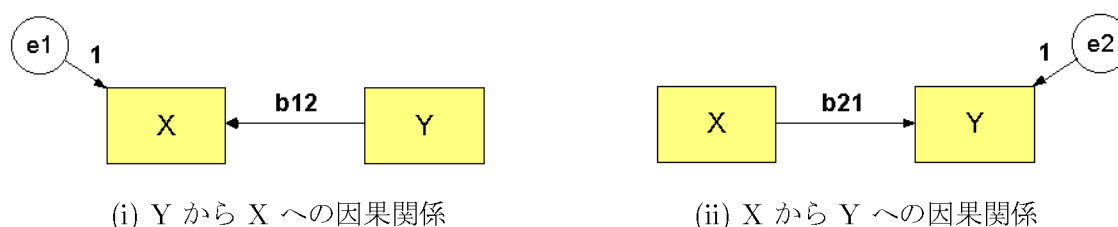


図 4.1: 同値モデル

つの観測変数 X , Y があつたとき、それらの相関係数から因果の方向を決めることはできない。相

関係数は X と Y に関して対称だからである．因果関係が， X から Y への場合とその逆の場合をパス図で表すと図 4.1 となる．方程式で表すと

$$Y = b_{21}X + e_2$$

$$X = b_{12}Y + e_1$$

となる．これらのモデルの適合度はまったく同じである．

上記の二つのモデルの共分散構造は，パラメトリゼーションは異なるが集合としては一致する．二つのモデルがこのような関係にあるとき，(互いに) 同値モデルであるという．同値モデルの適合度は完全に一致し，データからどちらのモデルが良いかは判断できない．

潜在変数を含むモデルでも同様のことが起こる．例えば図 4.2 の二つのモデルも同値であり， $F1$ が原因か $F2$ が原因かはこれらのモデルからは同定できない．

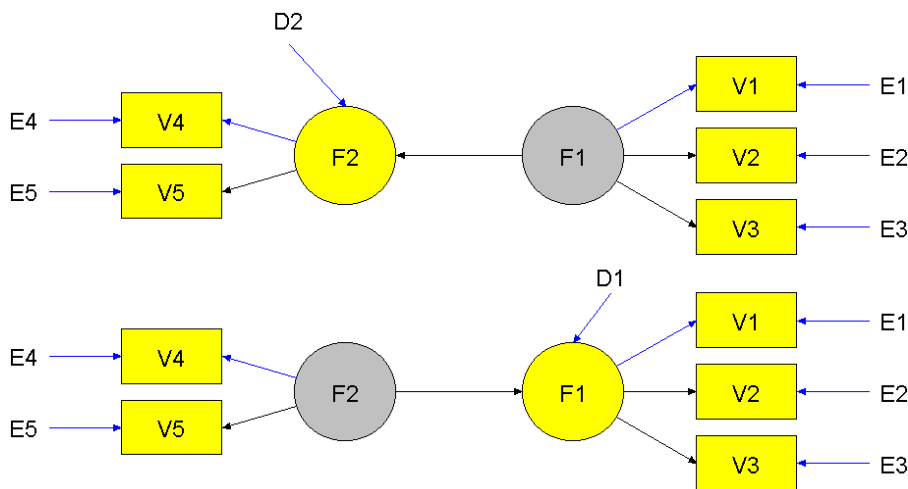
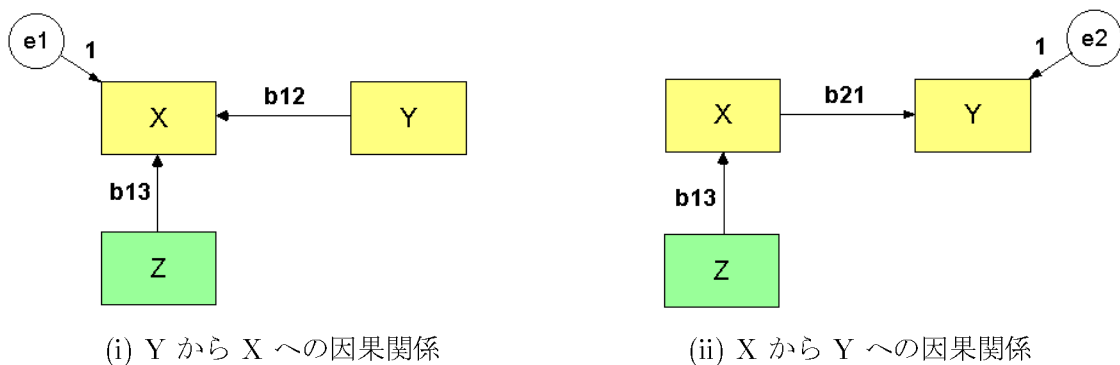


図 4.2: 同値な二つのモデル (上・下)

因果関係をデータから同定するには，互いに同値モデルにならないようなモデルを作ればよい．そのための方法論が操作変数法 (instrumental variables method) である．図 4.1 の因果の方向を操作変数法で決定するには，変数 X には直接効果をもつが，変数 Y には直接の影響を及ぼさない新しい変数 Z を観測する．「 Y には直接の影響を及ぼさない」という条件から，これらの変数



(i) Y から X への因果関係

(ii) X から Y への因果関係

図 4.3: 操作変数 Z を導入

間の相関行列は表 4.1 のようになる。これらの相関構造を比較すると、 $\text{Cor}(Y, Z)$ に顕著な違いが

表 4.1: 図 4.3 のモデルの相関構造

	X	Y	Z
X	1		
Y	b_{12}	1	
Z	b_{13}	0	1

	X	Y	Z
X	1		
Y	b_{21}	1	
Z	b_{13}	$b_{21}b_{13}$	1

(i) Y から X への因果関係

(ii) X から Y への因果関係

見られる。従って、Y と Z 間の標本相関係数 r_{YZ} が 0 に近ければ (i) のモデルと判断し、 r_{YZ} が大きく、かつ、 $r_{YZ} \approx r_{YX}r_{XZ}$ であれば (ii) のモデルを選ぶことになる。実際の判断は目の子で行わず、モデルの適合度が用いられる。

操作変数法によれば、双方向因果モデルの同定も可能である。ここでは、その実際例として、Felson-Bohrnstedt(1979) のデータを用いた分析を紹介する。¹ 彼らの研究は、12 歳から 14 歳までの女子学生 209 人に対する質問紙調査に基づいている。「スマートさ (smart)」と「美しさ (good-looking)」との間の因果関係において、どちらからの因果が強いかを調べるのが目的である。スマート (頭が良い) だから美しく見えるのか、美しいからスマートに感じられるのか、を同定しようというのである。操作変数法のメカニズムを図 4.4 を見ながら解説しよう。「スマートさ」と

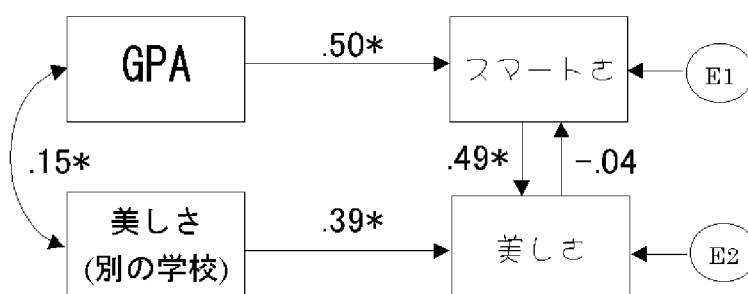


図 4.4: 双方向因果モデル:女子生徒

「美しさ」は、ある学校 A の生徒間の相互評価である。「美しさ (別の学校)」とは、学校 A とは別の学校の生徒が、写真によって「美しさ」を評価したものである。この場合、評価者は学校 A の生徒を知らない。従って、この評価は「スマートさ」に影響されないし、「スマートさ」へ直接の影響を及ぼすこともない。「美しさ (別の学校)」から「スマートさ」へパスが引かれていない所以である。一方、GPA (Grade Point Average) は教師が評価する科目成績の平均点であり、学校 A の生徒のお互いが他人の具体的な点数を知っているわけではない。従って、「GPA」から「美しさ」への直接のパスは引かれていない。このモデルは、「スマートさ」と「美しさ」との間の因果関係を調べるために、操作変数として「GPA」と「美しさ (別の学校)」を導入したものといえる。双方向因果モデルの場合は、誤差変数 E1, E2 間に相関を許すことが多いのであるが、このデータについては、その相関が有意にならなかったので外してある。

結果を見てみよう。モデルの適合度は

¹本データの分析例は、AMOS(Arbuckle 1995) にも紹介されている。

カイ 2 乗値=0.508(P-値=0.476, df=1)

となっている。「美しさ」から「スマートさ」へのパス係数は、-.04 と小さな値として推定され、唯一有意でないパスであることが分かる。従って、「スマートさ」から「美しさ」への因果関係は認められるが、逆の効果は無いと判断できることになる。

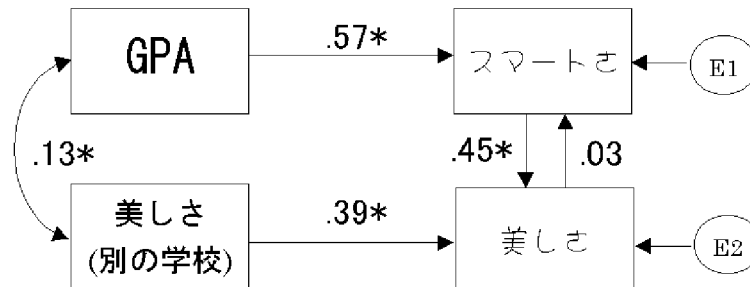


図 4.5: 双方向因果モデル:男子生徒

男子生徒の場合の結果を図 4.5 に示す。結論は女子生徒の場合と同じである。男女のデータを用いて、3.1.3 節で紹介した多母集団の同時分析を行うこともできるが、ここでは述べない。

以下に、操作変数法に関する注意点を述べる。

- (i) **同値モデルをにらんだモデル構成** 因果の方向に興味がある、つまり、「 $X \rightarrow Y$ 」を主張したい場合は、このモデルの適合度が良いということを示すだけでは不十分である。対抗馬である「 $X \leftarrow Y$ 」なるモデルが棄却されるということが必要である。そのためには、まず、「 $X \rightarrow Y$ 」のモデルと「 $X \leftarrow Y$ 」のモデルが同値モデル（適合度がまったく同じ）にならないようなモデル構成が必要である。
- (ii) **適切な操作変数の探索** 上記 (i) のための有効な方法が操作変数法である。同方法では、適切な操作変数を見つけることが最大の問題になる。本節での例のような適切な操作変数がいつも見つかるとはかぎらない。
- (iii) **操作変数に対する安定性** パス係数の推定値が操作変数の選択にどの程度依存するか検討しておく必要がある。しかし、この種の研究は未だないようである。
- (iv) **関係の安定** 双方向の因果関係を分析する際は、関係が安定している必要がある。例えば、クラス替えがあった直後のデータだと、スマートさや美しさなどの評価が安定しないので、上記のような分析は不適である。

以上、共分散構造分析による因果の方向の決定について紹介してきた。ここまで読んできた読者は、「共分散構造分析で因果が完全に決定できるではないか」という感想をもったかもしれない。しかし、過大な期待は禁物である。

図 4.3 での因果決定について考えてみる。ここでの選択は、「 $X \rightarrow Y$ 」 or 「 $X \leftarrow Y$ 」の何れかであった。「 $X \rightarrow Y$ 」 or 「 $X \leftarrow Y$ 」の何れかであることが事前に分かっておれば、操作変数法は有効な手段である。問題は、「 X と Y は因果の関係にない」という選択肢がないことである。例えば、 X と Y の相関が、解析者が気づかない第三変数 W によって引き起こされているとしよう (図 4.6)。 W と Z の間に相関がなければ、 Z と Y の間に相関は生じない。すると、図 4.3 (i) の因果関係が支持されることになってしまう。非実験 (観察) データの分析には、このような第三変数の影響が出ることがあり、共分散構造分析もこの効果に無縁ではない。

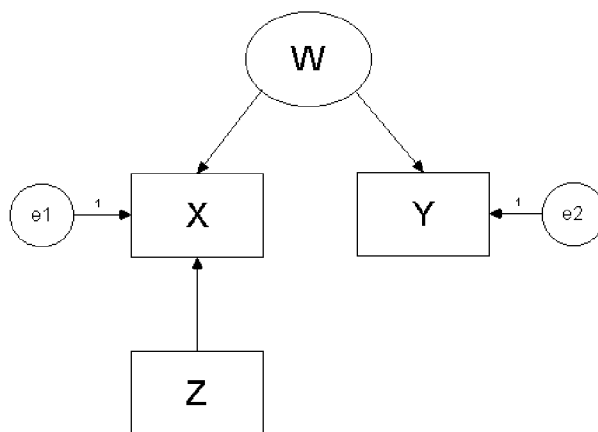


図 4.6: 第三変数の影響

例えば、図 4.4 や図 4.5 で、「GPA」と「美しさ（別の学校）」の間に有意な相関が生じている。本来これらの変数は無相関であるべきである。これら二つの項目には、分析者の関知しない何らかの影響があったと考えるのが自然である。

もう一つは、因果の大きさの問題である。「 $X \rightarrow Y$ 」という因果を含むモデルが受容され、この方向の因果関係が統計的に認められても、「 $X \rightarrow Y$ 」の決定係数が大きくないことがしばしばある。このようなときは、 X によって説明される Y の変動が大きいということであり、「 $X \rightarrow Y$ 」であっても X が Y の主要な要因ではないことになる。このことを認めた上で分析結果を解釈する必要がある。

最後に注意しておきたいのは、共分散構造モデルは線形の因果構造を扱っているということである。交互作用がある場合など非線形の関係については何も言及していない。非線形の関係に対して線形モデルを当てはめたとき、モデルが棄却されれば問題はないが、ある線形モデルと同値になっていると深刻なことになる。間違った線形モデルが採用され間違った解釈が行われるからである。

非線形共分散構造モデルへの接近については、Bollen-Paxton (1998) や Schumacker-Marcoulides (1998) などがあるが、まだ全貌はつかめていない。

4.2 因果の大きさを正確に測る

喫煙の量と肺がんの発症率の関係を調べたいとき、因果の方向は明らかである。喫煙が原因で肺がんが結果である。問題は、因果の大きさをどのようにして正確に求めるかである。ここでの主要なテーマは、前節でも登場したが、第三変数の影響である。図 4.7 を見ていただきたい。われわれが通常興味あるのは、「喫煙量」から「肺がん発症率」への効果 b_{21} である。ところがこのパス図のように、両者への共通要因である F が存在するならば、「喫煙量」から「肺がん発症率」への回帰係数は

$$b_{21} + \lambda_1 \lambda_2$$

となり、 $\lambda_1 \lambda_2$ だけバイアスをもつことになる。²

F から「喫煙量」へのパスをぶった切ることができれば、問題は減少する。 F から「肺がん発症

²ただし、全変数の分散は1に標準化してある。

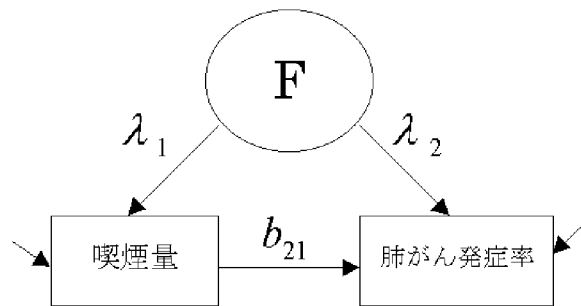


図 4.7: 推定値のバイアス：第三変数の影響

率」への影響は誤差として処理される。「喫煙量」へのパスを切る一つの方法は、喫煙量を解析者がコントロールすることである。例えば、「無喫煙」「1箱未満」「1箱から2箱」「2箱以上」³という群を作り被験者をランダムに割り当てる。そして、数年後に肺がんの発症率を群ごとに比較する。これはいわゆる実験データであり、実験データを収集することのメリットの一つはここにある。第三変数の影響をシャットアウトできるのである。

しかしながら、倫理上、このような実験を組むことはできない。従って、観察研究的方法ではあるが、第三変数の影響をできる限り減らせるような分析方法を採る。例えば、ケース⁴として、肺がん発症者の喫煙率を調べる。コントロールとして、健常者から、肺がん発症者と性別や年齢、その他肺がん発症に関係すると思われる属性をできるだけマッチングさせたサンプルを選び、喫煙率を調べるのである。属性をマッチングさせることで、第三変数の影響が減少することが期待される。

以上述べてきたように、実験データに基づく因果推論よりも、観察データによる因果推論の結論が強くないのは、第三変数の影響を拭い去れないからであり、観察データに基づく共分散構造分析もこの弱点を克服しているわけではない。そして、喫煙と肺がんの研究のような「第三変数の影響をできる限り減らせる」ような工夫もあまりしていないと思われる。共分散構造分析で因果推論を実行するときには、以上の点をわきまえておく必要がある。

共分散構造分析は、第三変数の影響にまったく無防備かといえ、そうでないこともある。図 4.8

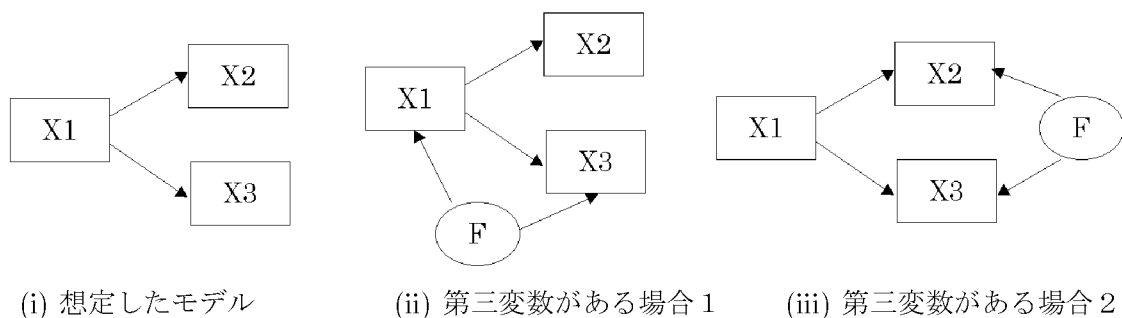


図 4.8: 共分散構造分析における第三変数の影響
(誤差変数は略してある)

には三つのパス図がある。パス図 (i) が想定したモデルである。パス図 (ii), (iii) には、解析者が

³もちろん一日の喫煙量のことである。

⁴ケースとは、病気を発症している患者から選ばれたサンプルをいう。

想定していない第三変数 F の影響が存在する。真の状況が (ii) または (iii) のとき、モデル (i) で分析を行うと適合度が悪くなる。解析者が想定していない変数が因果モデルに影響しているときは、モデルがデータに合わないという警告が発せられ、モデル修正する機会を与えてくれる場合がある。

しかし、第三変数の検出をいつも適合度に頼るのは勧められない。例えば、想定したモデルが、(i) において X2 から X3 へパスを引いていたならば、適合度は (ii) や (iii) の第三変数の検出をしてくれない。推定値にバイアスが生じたモデルをそのまま受容してしまう。適合度によって第三変数の存在が示唆されたならば、それは非常に幸運であったと考えるべきであろう。

観察データに基づく因果分析の弱点を、実験データの分析と比して述べてきたわけであるが、実験データによる因果分析がパーフェクトであるというわけではない。まず、共分散構造分析を適用するような複雑な因果関係が研究対象であるとき、適切な実験デザインが組めるかという問題がある。次に、水準・制御という実験を行うための非常に基本的な事柄に関する問題がある。実験をするためには水準が必要であり、事前に（恣意的に）設定された水準で因果効果が認められたとしても、それ以外の水準での因果効果は想像するしかない。変数を固定（制御）したり、喫煙効果の分析のようにマッチングをしたコントロール群を作成すると、因果効果が有意になりやすいのは事実である。しかし、固定された値に関する一般化可能性や、マッチングで登場する属性の組合せ以外の状況での因果効果はどうなるか、という問題が生じるのである。一方、観察データは、水準などは分析者が定めず、ランダムネスへ押し込んでしまう。⁵

因果分析で共分散構造分析を有効に利用するためには、できるだけ正確なパス図を作成することがポイントである。パス図ができれば、データが採取できるかどうか、そして、そのモデルがデータから推定できるかどうかを吟味する。重要な変数のデータが取れないときには、その変数を制御した下で、つまり、その変数をある一定の値に固定した下でのデータ収集の可能性を探ることになる。

もう一つの可能性は、実験データの共分散構造分析を適用してみることであろう。多母集団の同時分析や 3.1.4 節で紹介した簡便法が有効な手法になると考えられる。

共分散構造分析による因果推論に興味のある受講生は、Bullock-Harlow-Mulaik (1994) も参照されたい。

⁵この議論については豊田 (1998, 9 章) が詳しい。

第5章

統計的推測の基礎

前章までで、種々の共分散構造分析事例を紹介してきた。本章では、それらの統計的基礎について簡単に紹介する。

5.1 最尤推定とカイ2乗値

共分散構造モデルは、その名のとおりに、観測変数ベクトル \mathbf{X} の (分散) 共分散行列 Σ に構造があるモデルのことである。構造があるとは、 Σ がより基本的な母数 $\boldsymbol{\theta}$ の関数になっていることをいう。式で書くと

$$\text{Var}(\mathbf{X}) = \Sigma(\boldsymbol{\theta}) \quad (5.1)$$

となる。パラメータ $\boldsymbol{\theta}$ の具体的な中身は、パス係数や独立変数の分散・共分散である。

p 次元のデータ $\mathbf{X}_1, \dots, \mathbf{X}_N$ を多変量正規分布 $N_p(\boldsymbol{\mu}, \Sigma)$ からの独立標本とし、その分散共分散行列が式 (5.1) の構造を持っているとしよう。このとき、データ $(\mathbf{X}_1, \dots, \mathbf{X}_N)$ の同時確率密度関数は

$$\prod_{i=1}^N N_p(\mathbf{x}_i | \boldsymbol{\mu}, \Sigma(\boldsymbol{\theta})) \quad (5.2)$$

となる。ここで、 $N_p(\mathbf{x} | \boldsymbol{\mu}, \Sigma)$ は (p -次元) 多変量正規分布の確率密度関数で、具体的な形は以下で与えられる。

$$N_p(\mathbf{x} | \boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right)$$

平均構造モデルを考える場合は、式 (5.1) に加えて

$$E(\mathbf{X}) = \boldsymbol{\mu}(\boldsymbol{\theta}) \quad (5.3)$$

が導入される。さらに、多母集団の同時分析を考えるときは、 g 番目の母集団からのデータを $(\mathbf{X}_1^{(g)}, \dots, \mathbf{X}_{N_g}^{(g)})$ ($g = 1, \dots, G$) と書き、その母集団分布を $N_p(\boldsymbol{\mu}^{(g)}(\boldsymbol{\theta}), \Sigma^{(g)}(\boldsymbol{\theta}))$ とするならば、同時確率密度関数は

$$\prod_{g=1}^G \prod_{i=1}^{N_g} N_p\left(\mathbf{x}_i^{(g)} \middle| \boldsymbol{\mu}^{(g)}(\boldsymbol{\theta}), \Sigma^{(g)}(\boldsymbol{\theta})\right) \quad (5.4)$$

と表すことができる。

本稿では、議論を簡単にするため、一般的な (5.4) は扱わず、(5.2) の場合 (単一母集団で平均構造を考えない) について解説していく。

確率密度関数の意味を直感的に述べれば、その値が高いところは生起確率が高い (起こり易い) ことを表すといえる。ということは、逆に、 $\mathbf{X}_1, \dots, \mathbf{X}_N$ が起これば、その点での確率密度関数の値は小さくはないと考えるのは自然である。従って、未知母数 $\boldsymbol{\theta}$ の値は、 $\mathbf{X}_1, \dots, \mathbf{X}_N$ での確率密度関数の値が最大になる点として定めるのは無理がない。この考え方によって推定量を構成する方法を最尤法 (method of maximum likelihood) という。最尤法によって定められる推定量を最尤推定量 (maximum likelihood estimator; MLE) という。確率密度関数において $(\mathbf{x}_1, \dots, \mathbf{x}_N)$ にデータ $(\mathbf{X}_1, \dots, \mathbf{X}_N)$ を代入し、パラメータの関数と考えたものを尤度 (likelihood) といい、ここでは、 $L(\boldsymbol{\mu}, \Sigma(\boldsymbol{\theta}))$ と書くことにする。つまり、

$$L(\boldsymbol{\mu}, \Sigma(\boldsymbol{\theta})) = \prod_{i=1}^N N_p(\mathbf{X}_i | \boldsymbol{\mu}, \Sigma(\boldsymbol{\theta}))$$

言い換えると、最尤推定量 $(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\theta}})$ は最大化問題 (5.5) の解である。

$$\max_{(\boldsymbol{\mu}, \boldsymbol{\theta})} L(\boldsymbol{\mu}, \Sigma(\boldsymbol{\theta})) \quad (5.5)$$

簡単な計算で、 $L(\boldsymbol{\mu}, \Sigma(\boldsymbol{\theta}))$ は次のように変形することができる。

$$L(\boldsymbol{\mu}, \Sigma(\boldsymbol{\theta})) = \frac{1}{(2\pi)^{pN/2} |\Sigma(\boldsymbol{\theta})|^{N/2}} \exp \left[-\frac{N}{2} (\bar{\mathbf{X}} - \boldsymbol{\mu})' \Sigma(\boldsymbol{\theta})^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) - \frac{1}{2} \text{tr} \left\{ \Sigma(\boldsymbol{\theta})^{-1} \sum_{i=1}^N (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})' \right\} \right] \quad (5.6)$$

ここで、 $\bar{\mathbf{X}} = \frac{1}{N} \sum_{i=1}^N \mathbf{X}_i$ である。 $\Sigma(\boldsymbol{\theta})$ は正定値行列であるから、 $L(\boldsymbol{\mu}, \Sigma(\boldsymbol{\theta}))$ は $\boldsymbol{\mu} = \bar{\mathbf{X}}$ (= $\hat{\boldsymbol{\mu}}$, と書く) で最大になる。このとき、 $L(\boldsymbol{\mu}, \Sigma(\boldsymbol{\theta}))$ は (5.7) のようになる。

$$L(\hat{\boldsymbol{\mu}}, \Sigma(\boldsymbol{\theta})) = \frac{1}{(2\pi)^{pN/2} |\Sigma(\boldsymbol{\theta})|^{N/2}} \exp \left[-\frac{N}{2} \text{tr} \left\{ \Sigma(\boldsymbol{\theta})^{-1} S_N \right\} \right] \quad (5.7)$$

ここで S_N は \mathbf{X}_i の標本分散共分散行列であり、具体的に書けば $S_N = \frac{1}{N} \sum_{i=1}^N (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})'$ である。

最尤推定量を計算するには、(5.7) で $\boldsymbol{\theta}$ に関する最大化を実行すればよいのであるが、もう少し式変形した方が分かりやすくなる。(5.7) において、もし Σ に構造がなければ、最大化問題は $\boldsymbol{\theta}$ ではなく、正定値行列全体 $\{\Sigma > 0\}$ にわたって最大化することになる。その結果は次のようになることが分かっている：

$$\max_{\Sigma > 0} L(\hat{\boldsymbol{\mu}}, \Sigma) = L(\hat{\boldsymbol{\mu}}, S_N) = \frac{1}{(2\pi)^{pN/2} |S_N|^{N/2}} \exp \left[-\frac{Np}{2} \right]$$

ここで、次の尤度比 (likelihood ratio) を作る：

$$\lambda = \frac{\max_{\boldsymbol{\theta}} L(\hat{\boldsymbol{\mu}}, \Sigma(\boldsymbol{\theta}))}{\max_{\Sigma > 0} L(\hat{\boldsymbol{\mu}}, \Sigma)} = \max_{\boldsymbol{\theta}} \frac{|\Sigma(\boldsymbol{\theta})|^{-N/2}}{|S_N|^{-N/2}} \exp \left[-\frac{N}{2} \text{tr} \left\{ \Sigma(\boldsymbol{\theta})^{-1} S_N \right\} + \frac{Np}{2} \right] \quad (5.8)$$

$-2 \log \lambda$ を計算すれば

$$-2 \log \lambda = N \cdot \min_{\boldsymbol{\theta}} \left[\log |\Sigma(\boldsymbol{\theta})| - \log |S_N| + \text{tr} \left\{ \Sigma(\boldsymbol{\theta})^{-1} S_N \right\} - p \right] \quad (5.9)$$

となる。もちろん、(5.7) を θ について最大化することと、(5.9) は同等であり、(5.10) の右辺の係数である標本サイズ N も最小化には関係しない。

共分散構造分析のコンベンションとして、 N の代わりに $n = N - 1$ を用いることに注意する。つまり、(5.9) で、 S_N の代わりに不偏分散共分散行列

$$S = \frac{1}{n} \sum_{i=1}^N (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})'$$

を用い、 N の代わりに n を使うのである。

以上まとめると、 θ の最尤推定量は、最小化問題

$$\min_{\theta} F_{ML}(S, \Sigma(\theta))$$

の解として得られる。ここで

$$F_{ML}(S, \Sigma(\theta)) = \log |\Sigma(\theta)| - \log |S| + \text{tr} \{ \Sigma(\theta)^{-1} S \} - p \quad (5.10)$$

である。

最尤推定量を求める際、(5.7) の最大化から $F_{ML}(S, \Sigma(\theta))$ の最小化まで変形してきた。その理由を二つ述べる。

一つ目の理由は、関数 $F_{ML}(S, \Sigma)$ は次の性質をもっていることである：

- (i) $F_{ML}(S, \Sigma)$ は S と Σ に関して連続
 - (ii) $F_{ML}(S, \Sigma) \geq 0$
 - (iii) $F_{ML}(S, \Sigma) = 0 \iff S = \Sigma$
- (5.11)

これらの性質から、 $F_{ML}(S, \Sigma)$ は S と Σ との距離のようなものを表していることが分かる。従って、多変量正規性を仮定した最尤推定量は、『モデル $\Sigma(\theta)$ をデータ S に近づけ、最も近くなる θ の値として定められている』ことが分かる。この事実から、データ S がモデル $\Sigma(\theta)$ に十分近いとき、MLE $\hat{\theta}$ は、真の母数の値 θ_0 の近似になっていることが理解されよう。

上記でカギ括弧『...』で囲まれた部分は、多変量正規分布の仮定とは関係のない部分であり、データが多変量正規分布に従っていないとも成立する。従って、 $F_{ML}(S, \Sigma(\theta))$ の最小化で定められる $\hat{\theta}$ は、たとえ母集団分布が多変量正規分布でなくとも、真の母数の値 θ_0 の近似になっていることが保証される。これが、母集団分布が多変量正規でないときでも推定値は正確である、ということの理由である。統計用語を用いれば、正規であろうとなかろうと一貫性 (consistency) が成立する。

われわれは、性質 (5.11) を有する関数を不一致度関数 (discrepancy function) とよんでいる。

二つ目の理由は、 $\min_{\theta} F_{ML}(S, \Sigma)$ がデータとモデルの距離ならば、その値に意味付けができることである。値が0に近ければ良いモデル、そうでなければ悪いモデルと判断できよう。さらに、この値はカイ2乗値と密接な関係がある。尤度比 (5.8) は、仮説

$$H_0 : \text{Var}(\mathbf{X}) = \Sigma(\theta) \quad \text{versus} \quad H_1 : \text{Var}(\mathbf{X}) > 0$$

を検定するための尤度比検定統計量になっている。この検定は共分散構造モデルの適合度検定である。つまり、統計量

$$-2 \log \lambda = n \cdot \min_{\theta} F_{ML}(S, \Sigma(\theta))$$

が、ソフトウェアからカイ2乗値として出力されるものになっている。

統計学の一般理論から、 $-2 \log \lambda$ の帰無分布は、多変量正規性の下でカイ2乗分布で近似できることが知られている。自由度は、帰無仮説と対立仮説における推定するパラメータ数の差である。

最後に、推定量の標準誤差の公式を紹介する。行列 A に対して、 A の縦ベクトルを重ねて積み上げた長い縦ベクトルを $\text{vec}(A)$ で表す。¹ 共分散構造の一階微分を

$$\Delta = \frac{\partial \text{vec}(\Sigma(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}'}$$

とおく。このとき、多変量正規分布の仮定の下で、 $\hat{\boldsymbol{\theta}}$ の漸近分散行列は以下で与えられる。

$$\frac{2}{n} \left(\Delta' \{ \Sigma(\boldsymbol{\theta}) \otimes \Sigma(\boldsymbol{\theta}) \}^{-1} \Delta \right)^{-1} \quad (5.12)$$

ここで \otimes は行列のクロネッカー積である。推定量の標準誤差は、(5.12) の対角要素の平方根で与えられる。

先に述べたカイ2乗分布による近似や推定量の標準誤差は、多変量正規分布の仮定の下で導かれている。しかし、社会科学でのデータは多変量正規分布と見なせないこともしばしばある。非正規分布の下では、カイ2乗分布の近似や標準誤差は正しいのであろうか。自然なこの疑問に対して、以下のことが分かっている。

i) 大雑把に言って、潜在変数（誤差変数を含む）が独立に分布し、独立潜在変数の分散・共分散が自由なパラメータであれば、パス係数の標準誤差は、非正規連続分布に対しても正しい。カイ2乗分布による近似も正しい (Browne-Shapiro 1988; 狩野 1990; Kano 1993)。

ii) 観測変数の4次モーメントが次の構造をしているとしよう：

$$\begin{aligned} E \left((X_i - \mu_i)(X_j - \mu_j)(X_k - \mu_k)(X_\ell - \mu_\ell) \right) \\ = (1 + \kappa)(\sigma_{ij}\sigma_{kl} + \sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk}) \end{aligned} \quad (5.13)$$

このとき、推定量の標準誤差とカイ2乗値は、それぞれ、ソフトウェアの出力を $\sqrt{1+\kappa}$ と $1+\kappa$ で除したものになる (Browne 1982, 1984; Shapiro-Browne 1987; 狩野 1990; Kano 1993)。

状況 i) は漸近ロバストの場合 (asymptotic robustness case) といわれる。理論的に大変興味深い結果であり、美しい定理である。しかしながら、潜在変数の（無相関性だけでなく）独立性の仮定は、確認のしようがないという欠点がある。状況 ii) は重要である。 κ は観測変数の尖度を表しており、多変量正規分布の場合は $\kappa = 0$ となる。 $\hat{\kappa} \approx 0$ であれば問題は少ないが、0 から大きく離れていると、標準誤差やカイ2乗値が信頼できなくなる。なお、観測変数が楕円分布にしたがっている場合に、(5.13) が満たされる。

ここでの議論は連続型分布に限定した。離散型分布については、5.3 節を参照されたい。

¹ $\text{vec}(A)$ や以下に出てくるクロネッカー積をより詳しく知りたい受講者は、Magnus-Neudecker (1988) を参照されたい。

5.2 適合度指標

²考えたモデルがデータと矛盾しないかどうかを検査することは極めて重要である。共分散構造分析では、因果に関する(恣意的な)仮説を出発点とするので、唯我独尊に陥らないためにも、モデルの良さを客観評価することが必要である。

回帰分析では、データと推定されたモデルの差

$$\mathbf{y} - X\hat{\boldsymbol{\beta}}$$

に基づいて、モデルの適合性や異常値の有無などを検討する。データと推定されたモデルの差を残差といい、このような作業を残差分析という。残差分析は回帰分析の重要なプロセスである。

共分散構造分析でも同様に、データと推定された構造の差(残差)

$$S - \Sigma(\hat{\boldsymbol{\theta}})$$

や、何らかの意味での S と $\Sigma(\hat{\boldsymbol{\theta}})$ の食い違いの程度に基づいて、モデルの妥当性の検討を行う。

残差が全体的に大きければ考えたモデルが不十分であり、因果モデルを全体的に見直す必要がある。一方、一部分の残差が大きいつきは、対応する共分散が適切に推定されているかどうか、つまり、データに異常値などが含まれていないかどうかを散布図などで吟味する。サンプルに異常がなければ、モデルを部分修正する。この場合、どこにパスを引けばよいか問題になるが、LM 検定や修正指標が有用な情報を与えてくれる。今回は、モデル修正にはふれない。

多重指標モデル(図 2.3)での分析では、残差行列は次のようになる。

$$S - \Sigma(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} 0.000 & & & & \\ 0.000 & 0.000 & & & \\ 0.013 & -0.008 & 0.000 & & \\ -0.001 & 0.001 & 0.000 & 0.000 & \end{bmatrix}$$

最大の絶対残差は 0.013 であり、これは十分小さいと考えてよい。

分析を、(標本)相関行列ではなく標本分散共分散行列からスタートした場合は、標準偏差 $\sqrt{s_{ii}}$ で基準化した残差を用いる。多くのソフトウェアで(基準化)残差のヒストグラムが出力される。このことは、共分散構造分析する際、残差の吟味が如何に重要かを物語っている。

残差行列は多次元なので、その大きさについて判断しにくい。そこで 1 次元に縮約したのが適合度指標である。この縮約の方法にさまざまな工夫があり、極めてたくさんの適合度指標が提案されている。

まず、適合度指標として最も考え易いのは、(基準化された)残差の 2 乗和を平均し平方根を取ったもの RMR (Root Mean square Residual)³、そして、GFI (Goodness-of-Fit Index)、AGFI (Adjusted GFI) であろう。

$$\begin{aligned} \text{RMR} &= \sqrt{\frac{2}{p(p+1)} \sum_{i \leq j} \{s_{ij} - \sigma_{ij}(\hat{\boldsymbol{\theta}})\}^2 / s_{ii}s_{jj}} \\ \text{GFI} &= 1 - \frac{\text{tr}[\{\Sigma(\hat{\boldsymbol{\theta}})^{-1}(S - \Sigma(\hat{\boldsymbol{\theta}}))\}^2]}{\text{tr}[\{\Sigma(\hat{\boldsymbol{\theta}})^{-1}S\}^2]} \\ \text{AGFI} &= 1 - \frac{p(p+1)(1 - \text{GFI})}{p(p+1) - 2q} \end{aligned}$$

²本節は、狩野(1997)の 4.7 章と同等である。

³Jöreskog-Sörbom(1981)では、 $s_{ii}s_{jj}$ による基準化はなされていない。

ここで、 q は $\boldsymbol{\theta}$ の次元で推定すべきパラメータの数、 $\sigma_{ij}(\hat{\boldsymbol{\theta}})$ は $\Sigma(\hat{\boldsymbol{\theta}})$ の (i, j) 成分である。以上は、LISREL 5 (Jöreskog-Sörbom 1981) に初めて搭載された。RMR は小さいとき、GFI と AGFI は 1 に近いとき、良い当てはまりを示す。回帰分析の言葉でいえば、GFI は重相関係数に、AGFI は自由度調整済みの重相関係数に対応する。

回帰分析では、説明変数の増加にともなって残差は小さくなる。これと同じことが共分散構造分析においても起こる。つまり、パスを引けば引くほど、推定すべき自由なパラメータが増加し残差は小さくなる。残差が小さければ良いというだけならば残差が 0 の飽和モデル⁴が一番良いということになる。如何に少数個のパラメータで良い適合を達成するかがポイントである。このような観点で、GFI にパラメータ数 q に関するペナルティを付加したものが AGFI である。数学的に $GFI \geq AGFI$ なる関係がある。

最近、次の指標が提案され注目されている。

$$RMSEA = \sqrt{\max \left\{ \frac{\hat{F}}{d} - \frac{1}{n}, 0 \right\}}$$

ここで、 $d = p(p+1)/2 - q$ 、 \hat{F} はモデルとデータの最小距離であり、次で定義される：

$$\hat{F} = \min_{\boldsymbol{\theta} \in \Omega} F_{ML}(S, \Sigma(\boldsymbol{\theta}))$$

RMSEA (Root Mean Square Error of Approximation) は、もともとは Steiger-Lind が 1980 年に米国アイオワ州で行われた計量心理学会 (Psychometric Society) で発表したもので、Browne-Cudeck (1993) がその理論を整備した。その考え方の基礎は、真の母分散共分散行列 Σ_0 は、厳密にはモデル $\Sigma(\boldsymbol{\theta})$ とは一致せず、モデルと Σ_0 との距離

$$F_0 = \min_{\boldsymbol{\theta} \in \Omega} F_{ML}(\Sigma_0, \Sigma(\boldsymbol{\theta}))$$

は 0 ではないとする。RMSEA は、この最小距離 F_0 を自由度で割った量 F_0/d を推定しようとするものである。また、 F_0/d の区間推定も提案されている。

これらの指標を使ってどの程度の値であればモデルを受容するか、という具体的な基準を与えることは困難である。それは、回帰分析での決定係数 R^2 に基準がないのと同じである。10 年ぐらい前までは、GFI や以下で紹介する CFI は 0.9 以上であればよい適合とする rule of thumb が流布していたが、最近では、0.95 以上を要求される場合がある (e.g., Hu-Bentler 1999)。

Browne-Cudeck (1993) は、RMSEA の値が 0.05 以下であれば当てはまりが良く、0.1 以上であれば当てはまりは悪いと判断してよいと述べている。

「自然食品の購買行動」の分析では、これらの統計量の値は次のようになる。

$$\left. \begin{array}{l} RMSEA = 0.000 \\ RMR = 0.005 \end{array} \right\} \text{小さい値が良いモデルを表す}$$

$$\left. \begin{array}{l} GFI = 1.000 \\ AGFI = 0.997 \end{array} \right\} \text{大きい値が良いモデルを表す}$$

どの基準でみてもこのモデルはよい当てはまりであると判断できる。

次に、前節でも少しふれたカイ 2 乗値を紹介する。カイ 2 乗値とは、以下の仮説を統計的に検定する統計量を指す：

$$H : \text{Var}(\boldsymbol{x}) = \Sigma(\boldsymbol{\theta}) \text{ versus } A : \text{Var}(\boldsymbol{x}) \text{ は構造をもたない。} \quad (5.14)$$

⁴パラメータ数最大のモデル。観測変数の分散共分散行列に構造がないモデル。

「 $\text{Var}(\mathbf{x})$ が構造をもたない」とは、 $\text{Var}(\mathbf{x})$ の要素はすべて自由なパラメータであることを意味し、飽和モデル (saturated model) とよばれる。飽和モデルのパラメータの数は $d = p(p+1)/2$ であるから、飽和モデルの自由度は $d = p(p+1)/2 - p(p+1)/2 = 0$ になる。⁵

飽和モデルでは、 $\widehat{\text{Var}}(\mathbf{x}) = S$ と推定する。一方、 $\text{Var}(\mathbf{x}) = \Sigma(\boldsymbol{\theta})$ のときは、 $\text{Var}(\mathbf{x})$ の要素は自由なパラメータではなく、一般に、 $\Sigma(\boldsymbol{\theta}) = S$ となるような $\boldsymbol{\theta}$ は存在しない。

この仮説に対する尤度比検定統計量は、前節で述べたとおり、

$$T_{ML} = n \cdot F_{ML}(S, \Sigma(\hat{\boldsymbol{\theta}}))$$

となり、数理統計学の一般論から、標本数 n が十分大きいとき、帰無仮説のもとで、 T_{ML} の分布は自由度

$$\frac{1}{2}p(p+1) - q \quad (= d \text{ と書く}) \quad (5.15)$$

のカイ 2 乗分布 χ_d^2 で近似できることが分かっている。従って、検定方式は

$$\begin{aligned} T_{ML} \leq \chi_d^2(\alpha) &\implies \text{モデルを受容} \\ T_{ML} > \chi_d^2(\alpha) &\implies \text{モデルを棄却} \end{aligned}$$

となる。ここで、 $\chi_d^2(\alpha)$ は自由度 d のカイ 2 乗分布の上側 $100\alpha\%$ 点である。有意水準 α の値としては 0.05 や 0.10 がよく用いられる。

モデルの比較には、AIC がよく用いられる (e.g., Akaike 1987)。AIC は次の式で定義される：

$$AIC = T_{ML} - 2d \quad \text{または} \quad AIC = T_{ML} + 2q$$

d と q の関係は (5.15) にある。AIC は絶対的な値には意味がなく、モデル間の差にのみ意味がある。従って、AIC は上記どちらで定義しても同等なのである。AIC の値が小さい方が良いモデルと判断される。

AIC は、大きくなりすぎるモデルを抑制する意味をもつ。⁶ つまり、パスをたくさん引いてモデルを大きくすると、推定すべきパラメータ数 q が大きくなる。その数に見合うだけカイ 2 乗値が小さくならないと、そのモデルは採択されない。具体的に述べると、パスを 1 本引いてカイ 2 乗値が 2 以上下がらなければ、そのパスを引く価値はないと判断される。

今まで紹介した適合度指標はすべて、推定された共分散構造 $\Sigma(\hat{\boldsymbol{\theta}})$ と標本分散共分散行列 S との食い違いの程度を測定しようとするものであった。Tucker-Lewis (1973) や Bentler-Bonett (1980) は少し違った指標を提案している。この指標を紹介する前に、回帰モデル

$$\begin{aligned} y_i &= \beta_0 + \sum_{j=1}^p x_{ij}\beta_j + \epsilon_i \quad (i = 1, \dots, n) \\ \mathbf{y} &= X\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (\text{ベクトル表示}) \end{aligned}$$

における、⁷ 偏回帰係数の零仮説の検定を考える：

$$H : \beta_1 = \dots = \beta_p = 0, \text{ versus } A : \text{仮説 } H \text{ は正しくない}$$

⁵逆に、自由度が 0 であるモデルを飽和モデルということもある。

⁶小さいモデルとは、パラメータ数 q が小さく、モデルの自由度 d が大きいモデルのことをいう。大きいモデルはその反対である。

⁷ $X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix}$ である。

注意したいのは H において $\beta_0 = 0$ を要求していないことである。この検定は、一般平均 β_0 を除く説明変数 X が、 \mathbf{y} の変動に影響を及ぼしているかどうかを検討するもので、分析の最初に行われる説明変数の全体的評価である。

上記の仮説 H の下で $E(y_i) = \beta_0$ 、仮説 A の下で $E(y_i) = \beta_0 + \sum_{j=1}^p x_{ij}\beta_j$ となるので、これらをベクトルで表せば

$$H : E(\mathbf{y}) = \mathbf{1}_n\beta_0 \text{ versus } A : E(\mathbf{y}) = X\boldsymbol{\beta} \quad (5.16)$$

となる。ここで、 $\mathbf{1}_n = [1, \dots, 1]' (\in \mathbf{R}^n)$ である。

先ほどの共分散構造分析における適合度検定との違いは、適合度検定は興味あるモデル $\Sigma(\boldsymbol{\theta})$ を最大モデルである飽和モデルと比較しているのに対して、回帰係数の零仮説の検定は、興味ある回帰モデルを $E(\mathbf{y}) = \mathbf{1}_n\beta_0$ という最小モデルと比較しているところにある。

Bentler-Bonett (1980) は、共分散構造分析における最小モデルとして、観測変数間に相関がないというモデルを導入した：

$$\text{Var}(\mathbf{x}) = \begin{bmatrix} \sigma_{11} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sigma_{pp} \end{bmatrix} \quad (= \text{diag}(\sigma_{11}, \dots, \sigma_{pp}), \text{ と書く})$$

このモデルは伝統的に、独立モデルとよばれているものである。(5.16) に対応する仮説は

$$H : \text{Var}(\mathbf{x}) = \text{diag}(\sigma_{11}, \dots, \sigma_{pp}) \text{ versus } A : \text{Var}(\mathbf{x}) = \Sigma(\boldsymbol{\theta}) \quad (5.17)$$

となる。回帰分析での検定 (5.16) では説明変数が観測変数に影響を及ぼしているかどうかを検討したのに対して、(5.17) は観測変数間の相関を説明するパスや相関が存在するかどうかを検討している。Bentler-Bonett (1980) は、この検定だけによってモデルの妥当性を検討するのではなく、現在考慮しているモデル $\Sigma(\boldsymbol{\theta})$ が、独立モデル (最小モデル) と飽和モデル (最大モデル) とを結ぶ直線上のどのあたりに位置するのかをみることを提案した。

飽和モデルのカイ 2 乗値 T_s は、 $\widehat{\text{Var}}(\mathbf{x}) = S$ に注意すると $T_s = n \cdot F(S, S) = 0$ となり最小値をとる。独立モデルのカイ 2 乗値を T_i とすると、このモデルが最も制約された小さいモデルであることから、 T_i の値は最も大きくなる。⁸ 現在考慮しているモデル $\Sigma(\boldsymbol{\theta})$ のカイ 2 乗値を

$$T = n \cdot \min_{\boldsymbol{\theta} \in \Omega} F(S, \Sigma(\boldsymbol{\theta}))$$

とすると、 S と $\Sigma(\hat{\boldsymbol{\theta}})$ が近いほど良いモデルである、つまり、 T が $T_s = 0$ に近ければ近いほど良いモデルであり、 T が T_i に近ければ悪いモデルであると考えられる (図 5.1)。そこで、Bentler-Bonett (1980) は T の相対的位置

$$\text{NFI} = \frac{T_i - T}{T_i - T_s} = \frac{T_i - T}{T_i} = 1 - \frac{T}{T_i}$$

を適合度の指標にした。 $0 \leq \text{NFI} \leq 1$ であり、⁹ もちろん、1 に近ければ良いモデルということになる。

NFI は標本サイズ N が大きくないとき、モデルが正しくても 1 に近くないという欠点がある。そこで、モデルの自由度 d を導入することによってその欠点を改善したものが、次の NNFI と考えられる。

$$\text{NNFI} = 1 - \frac{T/d - 1}{T_i/d_i - 1}$$

⁸添字 i は independent, s は saturated の意味。

⁹それゆえ、Normed Fit Index の名がある。

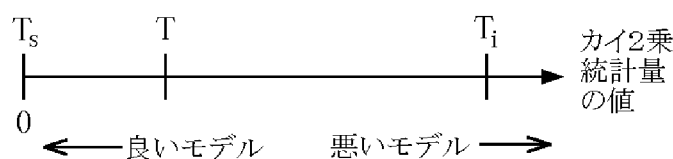


図 5.1: Bentler-Bonett (1980) の考え方

ここで、 d_i は独立モデルの自由度で、 $d_i = p(p+1)/2 - p = p(p-1)/2$ である。モデルが正しいとき、 T は近似的に自由度 d のカイ 2 乗分布 χ_d^2 にしたがう。 T が $E(\chi_d^2) = d$ なる値をとるときは十分良い適合だと考えられる。このとき ($T = d$ のとき)、1 なる値をとるように NFI を改良したものが NNFI である。この指標は、Tucker-Lewis (1973) が探索的因子分析モデルに対して提案した信頼性係数 (reliability coefficient) の共分散構造モデルへの拡張になっている。

NNFI は $0 \leq \text{NNFI} \leq 1$ なる性質をもたない。¹⁰ そこで、Bentler (1990) は

$$\text{CFI} = 1 - \frac{\tau}{\tau_i}$$

を提案した。¹¹ ここで

$$\begin{aligned} \tau &= \max\{T - d, 0\} \\ \tau_i &= \max\{T - d, T_i - d_i, 0\} \end{aligned}$$

である。

多重指標モデルでの分析では、

$$\text{NFI}=0.999, \text{NNFI}=1.007, \text{CFI}=1.000$$

となっており、このモデルは十分良い適合であることが分かる。

重要な適合度指標の紹介が終わったところで、「自然食品の購買行動」のデータを図 5.2 のパス解析モデルで解析したときの適合度をみてる。表 5.1 の結果から、このデータに対してパス解析モデルは良い当てはまりであるとはいえない。最大絶対残差は

$$r_{34} - \sigma_{34}(\hat{\theta}) = 0.449$$

と出力され、このモデルでは X_3 (購買額) と X_4 (購買回数) の相関が十分説明できていないことが分かる。このように、適切でないモデルに対して適合度指標は、クリアに「No」と言ってく

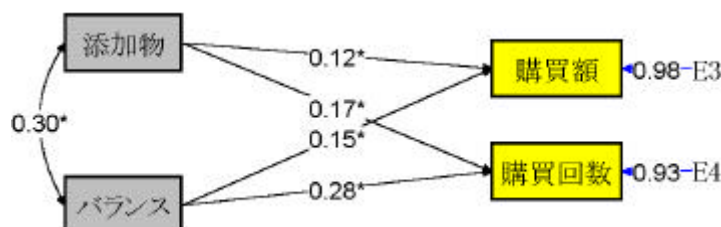


図 5.2: パス解析結果

¹⁰それゆえ、NonNormed Fit Index の名がある。

¹¹CFI は Comparative Fit Index の略。

表 5.1: パス解析モデルの適合度

カイ 2 乗値=233.32(P-値=0.0000, df=1, n=831) RMSEA=0.529, RMR=0.142, GFI=0.891, AGFI =-.092 NFI=0.508, NNFI=-1.976, CFI=0.504
--

れる。

回帰分析や共分散構造分析で、以下の2つの概念を混同してはいけない:

適合度指標: モデルがデータに近いか遠いか

決定係数: 説明変数が従属変数をどの程度説明しているか

重回帰モデルの良し悪しは、決定係数 R^2 で測ることが多い。しかし、 R^2 はモデルがデータに矛盾しないかどうかを見る尺度でない。 R^2 は、モデルは正しいという仮定の下で、説明変数の変動が従属変数の変動をどの程度説明しているかを表す尺度である。回帰分析は予測に使われることが多く、 R^2 が大きくあって欲しいという希望がある。それゆえ、 R^2 の値が小さければ悪いモデルであるというレッテルが貼られる。しかしながら、それでも、誤差分散が大きいモデルが当てはまっている。

一方、4.1 節で注意したように、共分散構造モデルの適合度がすばらしく良くても、決定係数が非常に小さいということも起こり得る。

5.3 二値データなどの扱い

5.1 節では、(i) 共分散構造分析の理論枠組みは、多変量正規性の仮定の下で構築されてきたこと; (ii) 非正規連続分布についてもおおむね適用可能であること; そして、(iii) より詳しくは、多変量正規性の下での理論を適用できるための条件、を述べた。

本節では、離散型の分布について取り上げる。まず、観測変数が5件法や7件法などの順序カテゴリカル変数の場合は、連続変数とみなして大きな損失がないことが分かっている。歪度・尖度・ヒストグラムを吟味して、きつい天井効果や床効果を受けていないこと、極端に、あるカテゴリ（例えば、「どちらでもない」）に回答が集中していないことなどを確認するとよい。

二値変数が含まれている場合を考える。二値変数とは、例えば、被験者の属性（性別、関東と関西など）、賛成と反対、処理群とコントロール群などを0, 1の値で区別するものである。二値変数を分析する際、それが従属変数であるか独立変数であるかで、扱いが大きく異なる。独立変数の場合は、次節で述べるように大きな問題にならないことが多い。従属変数の場合は、特殊な分析方法が必要である。最近のソフトウェアには、そのためのオプションがつけられていることが多い。

5.3.1 独立観測変数の場合: Muthén のフレームワーク

Muthén (1984) や Muthén-Muthén (1998) による「観測変数を独立変数 X と従属変数 Y に分ける」というアイデアは秀逸である。このアイデアの下で、以下の構造式を考える:

$$\begin{cases} Y &= \nu + \Lambda f + KX + \epsilon \\ f &= \alpha + Bf + \Gamma X + \zeta \end{cases} \quad (5.18)$$

ここで、 \mathbf{f} は潜在変数、 $\boldsymbol{\epsilon}$ は従属観測変数に付随する誤差変数、 $\boldsymbol{\zeta}$ は潜在変数 \mathbf{f} に付随する攪乱変数 (disturbance; 誤差項) を表す. 係数行列 B の対角要素は 0 とし、 $I - B$ が正則であると仮定する. この下で、

$$\mathbf{Y} = \boldsymbol{\nu} + \Lambda(I - B)^{-1}(\boldsymbol{\alpha} + \Gamma\mathbf{X} + \boldsymbol{\zeta}) + K\mathbf{X} + \boldsymbol{\epsilon} \quad (5.19)$$

独立変数に、次の分散・共分散行列を設定する：

$$E(\mathbf{X}) = \boldsymbol{\tau}, \quad \text{Var}(\mathbf{X}) = \Phi, \quad \text{Var}(\boldsymbol{\epsilon}) = \Theta, \quad \text{Var}(\boldsymbol{\zeta}) = \Psi$$

なお、 $E(\boldsymbol{\epsilon}) = 0$, $E(\boldsymbol{\zeta}) = 0$ である. さらに、これらの変数は互いに独立に多変量正規分布すると仮定する. ここでのポイントは、 \mathbf{X} の分布については何の仮定もおかないことである.

以上の仮定の下で、 \mathbf{X} を与えた下での \mathbf{Y} の条件付分布は多変量正規分布となり、条件付平均と分散は以下のとおり与えられる：

$$\begin{cases} E(\mathbf{Y}|\mathbf{X}) &= \boldsymbol{\nu} + \Lambda(I - B)^{-1}(\boldsymbol{\alpha} + \Gamma\mathbf{X}) + K\mathbf{X} \\ \text{Var}(\mathbf{Y}|\mathbf{X}) &= \Lambda(I - B)^{-1}\Psi(I - B)^{-1}\Lambda' + \Theta \end{cases} \quad (5.20)$$

数学的には、条件付分布 $\mathbf{Y}|\mathbf{X}$ が多変量正規で (5.20) の平均と分散をもつことを仮定する方が良い (弱い) 条件になる.

観測変数 $(\mathbf{Y}_i, \mathbf{X}_i)$ に基づく尤度は

$$\prod_{i=1}^N N\left(\mathbf{Y}_i \mid E(\mathbf{Y}_i|\mathbf{X}_i), \text{Var}(\mathbf{Y}_i|\mathbf{X}_i)\right) f_x(\mathbf{X}_i|\boldsymbol{\tau}, \Phi) \quad (5.21)$$

となる. ここで、 $f_x(\mathbf{X}_i|\boldsymbol{\tau}, \Phi)$ は \mathbf{X}_i の確率密度関数、または、確率関数である. 特に、 \mathbf{X}_i が固定変数の場合は、 $f_x(\mathbf{X}_i|\Phi) = 1$ である. もし $\mathbf{X}_i \sim N(\boldsymbol{\tau}, \Phi)$ であれば、観測変数を分割しない通常の方法 (5.1 節で述べたもの) に帰着する.

ここで、 $\boldsymbol{\tau}$ と Φ がその他のパラメータと無関係であることを仮定する. この仮定は多くのモデルで成立しており、無理のないものである. この仮定の下で、(5.21) の最大化は、 $\mathbf{Y}_i|\mathbf{X}_i$ を含む部分と \mathbf{X}_i を含む部分とを独立に行うことができる. つまり

$$\begin{aligned} & \max_{\text{すべてのパラメータ}} \prod_{i=1}^N N\left(\mathbf{Y}_i \mid E(\mathbf{Y}_i|\mathbf{X}_i), \text{Var}(\mathbf{Y}_i|\mathbf{X}_i)\right) f_x(\mathbf{X}_i|\boldsymbol{\tau}, \Phi) \\ &= \max_{\nu, \alpha, \Lambda, B, \Gamma, \Theta, \Psi} \prod_{i=1}^N N\left(\mathbf{Y}_i \mid E(\mathbf{Y}_i|\mathbf{X}_i), \text{Var}(\mathbf{Y}_i|\mathbf{X}_i)\right) \times \max_{\boldsymbol{\tau}, \Phi} \prod_{i=1}^N f_x(\mathbf{X}_i|\boldsymbol{\tau}, \Phi) \end{aligned}$$

従って、 $f_x(\mathbf{X}_i|\boldsymbol{\tau}, \Phi)$ を $N(\boldsymbol{\tau}, \Phi)$ で置き換えても、 $\mathbf{Y}_i|\mathbf{X}_i$ を含む項の最大化に関係しないことが分かる. このことから、(5.21) を以下のように変更する:

$$\prod_{i=1}^N N\left(\mathbf{Y}_i \mid E(\mathbf{Y}_i|\mathbf{X}_i), \text{Var}(\mathbf{Y}_i|\mathbf{X}_i)\right) N(\mathbf{X}_i|\boldsymbol{\tau}, \Phi) \quad (5.22)$$

(5.22) は、先に述べたように、観測変数全体に多変量正規分布を仮定したもとの尤度に他ならない.

パラメータ $\boldsymbol{\tau}$ と Φ は \mathbf{X}_i の標本平均と分散で推定される. 独立観測変数の分布について正確な情報がある場合は、(5.21) で推定する方が推定効率が良いかもしれないが、標本平均と分散で推定しても一貫性は保たれ、大きな情報損失にはならない.

(5.22) の下で計算された推定値の標準誤差は、一般に信用できないが、 \mathbf{X}_i が固定変数の場合は正しいときもある。

適合度の尤度比検定（カイ2乗値）はどうだろうか。飽和モデルの確率密度関数を次のように定義してみよう：

$$N\left(\mathbf{y} \mid \boldsymbol{\nu} + K\mathbf{x}, \Theta\right) f_x(\mathbf{x} \mid \boldsymbol{\tau}, \Phi)$$

ここで $\boldsymbol{\nu}$, K の要素はすべて自由なパラメータであり、 Θ は任意の正定値行列である。このとき、尤度比検定は以下ようになる：

$$\begin{aligned} & \frac{\max \prod_{i=1}^N N\left(\mathbf{Y}_i \mid E(\mathbf{Y}_i \mid \mathbf{X}_i), \text{Var}(\mathbf{Y}_i \mid \mathbf{X}_i)\right) f_x(\mathbf{X}_i \mid \boldsymbol{\tau}, \Phi)}{\max \prod_{i=1}^N N\left(\mathbf{Y}_i \mid \boldsymbol{\nu} + K\mathbf{X}_i, \Theta\right) f_x(\mathbf{X}_i \mid \boldsymbol{\tau}, \Phi)} \\ &= \frac{\max \prod_{i=1}^N N\left(\mathbf{Y}_i \mid E(\mathbf{Y}_i \mid \mathbf{X}_i), \text{Var}(\mathbf{Y}_i \mid \mathbf{X}_i)\right) N(\mathbf{X}_i \mid \boldsymbol{\tau}, \Phi)}{\max \prod_{i=1}^N N\left(\mathbf{Y}_i \mid \boldsymbol{\nu} + K\mathbf{X}_i, \Theta\right) N(\mathbf{X}_i \mid \boldsymbol{\tau}, \Phi)} \quad (5.23) \end{aligned}$$

(5.23) の成立条件は、 $\boldsymbol{\tau}$ と Φ が他のパラメータと関連していないという条件に加えて、 $\boldsymbol{\tau}$ と Φ はすべて自由なパラメータである（構造がない）ことである。

(5.23) は \mathbf{X}_i の正規性の下での尤度比検定になっている。従って、カイ2乗値は信頼してもよいことが分かる。

ただし、注意しておかなければならないのは、対立仮説では、 \mathbf{X}_i の値によらず共通の分散・共分散が設定されていることである（もちろんすべてが自由なパラメータである）。(5.23) で見れば、分母の \mathbf{Y}_i の分散共分散行列 Θ は \mathbf{X}_i の値によらず一定になっている。例えば、男性被験者の場合 $\mathbf{X}_i = 1$ 、女性被験者の場合 $\mathbf{X}_i = 0$ という変数の場合では、男女間で分散・共分散が共通であるモデルが対立仮説のモデル（飽和モデル）になっている。従って、もし、この前提が崩れているとカイ2乗値自体に意味がなくなる。

以上で分かったことをまとめると以下ようになる：

- ◇ **[仮定1]** 独立観測変数 \mathbf{X} の分布形は指定しないが、条件付分布 $\mathbf{Y} \mid \mathbf{X}$ は多変量正規とする。独立変数である \mathbf{X} , $\boldsymbol{\epsilon}$, $\boldsymbol{\zeta}$ は互いに独立に分布するとする。
- ◇ **[仮定2]** $E(\mathbf{X}) = \boldsymbol{\tau}$, $\text{Var}(\mathbf{X}) = \Phi$ とする。 $\boldsymbol{\tau}$ と Φ は構造を持たず、これら以外のパラメータ Λ , K , B , Γ , Θ , Ψ と無関係であるとする。
- ◇ パラメータ Λ , K , B , Γ , Θ , Ψ の最尤推定量は、(間違つて) $(\mathbf{Y}_i, \mathbf{X}_i)$ に多変量正規分布を仮定したときの最尤推定量に一致する。この意味で、推定値はソフトウェアからの通常の実出力を信じてよい。
- ◇ カイ2乗値も信用してよいが、適合度検定の対立仮説に注意。
- ◇ 推定値の標準誤差は参考にとどめる方がよい。

5.3.2 3.1.4 節の簡便法の吟味

本節は、前節での理論的な結果をふまえて、3.1.4 節で紹介した（実験データの）簡便法分析の例を理論的に吟味してみよう．図 5.3 にモデルと記号を載せてある．このモデルの方程式は以下のようになる：

$$\begin{aligned} Y_1 &= \nu_1 + 1.0f_1 && + \epsilon_1 \\ Y_2 &= \nu_2 + \lambda_{11}f_1 && + \epsilon_2 \\ Y_3 &= \nu_3 && + 1.0f_2 + \epsilon_3 \\ Y_4 &= \nu_4 && + \lambda_{11}f_2 + \epsilon_4 \\ f_2 &= &b_{21}f_1 + \gamma_{21}x_1 + \zeta_2 \end{aligned}$$

これを、Muthén の枠組みに合わせるため、式 (5.18) のように書きかえると、(5.24) が得られる．

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \end{bmatrix} + \begin{bmatrix} 1.0 & 0 \\ \lambda_{21} & 0 \\ 0 & 1.0 \\ 0 & \lambda_{42} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{bmatrix} \quad (5.24)$$

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ b_{21} & 0 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \gamma_{21} \end{bmatrix} X + \begin{bmatrix} f_1 \\ d_2 \end{bmatrix}$$

(5.18) との対応は詳しくは述べないが、 $K = O$ 、 $\alpha = 0$ に注意しておく．一方、独立変数の平均や分散・共分散は以下のようになる：

$$E \left(\begin{bmatrix} X \\ f_1 \\ d_2 \end{bmatrix} \right) = \begin{bmatrix} \tau_x \\ 0 \\ 0 \end{bmatrix}, \quad \text{Var} \left(\begin{bmatrix} X \\ f_1 \\ d_2 \end{bmatrix} \right) = \begin{bmatrix} \sigma_{xx} & \sigma_{xf} & 0 \\ \sigma_{fx} & \sigma_{ff} & 0 \\ 0 & 0 & \sigma_{dd} \end{bmatrix} \quad (5.25)$$

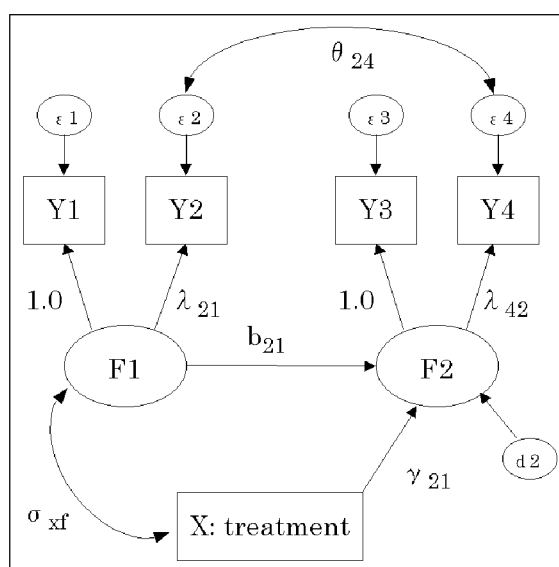


図 5.3: トレーニング効果の分析：簡便法のモデル

変数 X と f_1 の間に共分散が設定されている．このおかげで前節で述べた理論結果は使えないことに注意する．

変数 X は実験条件 ($X = 0$: トレーニングをしない; $X = 1$: トレーニングをする) を表しているから固定変数である．従って, (5.25) での意味の分散を計算するのは適切ではない．しかし, ソフトウェアの内部では, (5.24) と (5.25) のモデルで計算が実行される．

(5.19) と同様にして, (5.24) から,

$$\mathbf{Y} = \boldsymbol{\nu} + \Lambda(I - B)^{-1}(\Gamma X + \boldsymbol{\zeta}) + \boldsymbol{\epsilon}$$

が導かれる．従って, 多変量正規性の下での尤度は次のように書ける:

$$L = \prod_{i=1}^{N_0+N_1} N\left(\mathbf{Y}_i \mid \mathbf{E}(\mathbf{Y}_i|X_i), \text{Var}(\mathbf{Y}_i|X_i)\right) N(X_i \mid \tau_x, \sigma_{xx}) \quad (5.26)$$

$$\begin{cases} \mathbf{E}(\mathbf{Y}_i|X_i) &= \boldsymbol{\nu} + \Lambda(I - B)^{-1}(\Gamma X_i + \mathbf{E}(\boldsymbol{\zeta}_i|X_i)) \\ \text{Var}(\mathbf{Y}_i|X_i) &= \Lambda(I - B)^{-1}\text{Var}(\boldsymbol{\zeta}_i|X_i)(I - B)^{-1}\Lambda' + \Theta \end{cases}$$

ここで, $N_0 = 105$, $N_1 = 108$ であり,

$$\mathbf{E}(\boldsymbol{\zeta}_i|X_i) = \mathbf{E}\left(\begin{bmatrix} f_1 \\ d_2 \end{bmatrix} \mid X_i\right) = \begin{bmatrix} \sigma_{fx}\sigma_{xx}^{-1}(X_i - \tau_x) \\ 0 \end{bmatrix}$$

$$\text{Var}(\boldsymbol{\zeta}_i|X_i) = \text{Var}\left(\begin{bmatrix} f_1 \\ d_2 \end{bmatrix} \mid X_i\right) = \begin{bmatrix} \sigma_{ff} - \sigma_{fx}\sigma_{xx}^{-1}\sigma_{xf} & 0 \\ 0 & \sigma_{dd} \end{bmatrix}$$

である．

さて, X_i は次のような二値変数であるから

$$X_i = \begin{cases} 0 & \text{if } 1 \leq i \leq N_0 \\ 1 & \text{if } N_0 + 1 \leq i \leq N_0 + N_1 \end{cases}$$

これを (5.26) へ代入すると

$$L = \prod_{i=1}^{N_0} N\left(\mathbf{Y}_i \mid \boldsymbol{\mu}_0, \Xi\right) N(X_i = 0 \mid \tau_x, \sigma_{xx}) \times \prod_{i=N_0+1}^{N_0+N_1} N\left(\mathbf{Y}_i \mid \boldsymbol{\mu}_1, \Xi\right) N(X_i = 1 \mid \tau_x, \sigma_{xx}) \quad (5.27)$$

となる．ここで, 条件付分散共分散行列 Ξ は群間で共通であり以下で与えられる:

$$\Xi = \text{Var}(\mathbf{Y}_i|X_i) = \Lambda(I - B)^{-1} \begin{bmatrix} \sigma_{ff} - \sigma_{fx}\sigma_{xx}^{-1}\sigma_{xf} & 0 \\ 0 & \sigma_{dd} \end{bmatrix} (I - B)^{-1}\Lambda' + \Theta \quad (5.28)$$

一方, 条件付き平均の方は群間で異なり

$$\boldsymbol{\mu}_0 = \mathbf{E}(\mathbf{Y}_i|X_i = 0) = \boldsymbol{\nu} + \Lambda(I - B)^{-1} \begin{bmatrix} \sigma_{fx}\sigma_{xx}^{-1}(0 - \tau_x) \\ 0 \end{bmatrix}$$

$$\boldsymbol{\mu}_1 = \mathbf{E}(\mathbf{Y}_i|X_i = 1) = \boldsymbol{\nu} + \Lambda(I - B)^{-1} \begin{bmatrix} \sigma_{fx}\sigma_{xx}^{-1}(1 - \tau_x) \\ \gamma_{21} \end{bmatrix}$$

となる． $\boldsymbol{\mu}_0$ での右辺全体を改めて $\boldsymbol{\nu}$ と書きなおせば, 最終的に

$$\boldsymbol{\mu}_0 = \mathbf{E}(\mathbf{Y}_i|X_i = 0) = \boldsymbol{\nu} \quad (5.29)$$

$$\boldsymbol{\mu}_1 = \mathbf{E}(\mathbf{Y}_i|X_i = 1) = \boldsymbol{\nu} + \Lambda(I - B)^{-1} \begin{bmatrix} \sigma_{fx}\sigma_{xx}^{-1} \\ \gamma_{21} \end{bmatrix} \quad (5.30)$$

が得られる。

以上から、図 5.3 のパス図による簡便法は、次のモデルによる多母集団の同時分析と同等であることが分かる：

- ◇ 共分散構造に関して (5.28) から、群間ですべてのパラメータを等置した共分散構造モデルを適用
- ◇ 平均構造に関して (5.29) と (5.30) から
 - 観測変数の切片項を群間で等置
 - コントロール群の潜在変数の平均を 0 に固定
 - トレーニング群の潜在変数 F1 に平均 ($\sigma_{fx}\sigma_{xx}^{-1}$) を、F2 に切片項 (γ_{21}) を設定

第6章

共分散構造分析におけるブートストラップ法

6.1 ブートストラップ法とは

ブートストラップ法は、コンピュータ指向型の統計的推測の方法で、複雑な理論や数式を用いた解析的導出を、コンピュータによる大量の反復計算で置き換えるものである。つまり、コンピュータのパワーを生かしたシミュレーションにより、推定量の標準誤差やバイアス（偏り）を推定したり、信頼区間の構成や仮説検定を行うのである。その根底にある考え方は、「未知の母集団分布を経験分布で推定する」という単純かつ洗練されたものであることから、極めて柔軟性に富んでおり、広い範囲の問題に適用できる。解析的に答えを得るのが困難な多くの現実の問題に対しても、数値的に答えを与えてくれる。

ブートストラップ法は、Efron (1979) により提案されて以来、多くの人の関心を集め、理論と応用の両面から膨大な量の研究がなされている。こうしたことから、ブートストラップ法を提案した Efron の論文は、Kotz-Johnson (1992) による *Breakthroughs in Statistics* にも選定されている。ブートストラップ法の利用は、日本でも多くの人に注目されており、昨年 (1998) 度の日本統計学会のチュートリアル・セミナーがブートストラップ法を取り上げていることからその関心のほどがうかがえる。

ブートストラップ法全般について解説した本としては、Efron (1982), Efron-Tibshirani (1993), Davison-Hinkley (1997) などがある。残念ながら、これらの本の翻訳はいまのところない。邦文で書かれたものとしては、小西 (1988) がブートストラップ法の考え方を丁寧に説明している。

6.2 ブートストラップ法の考え方

ここでは、標準誤差の推定を例として、ブートストラップ法の考え方について説明する。簡単のために、観測変量が 1 変量の場合について説明する。

未知の確率分布 G からの大きさ N の無作為標本を

$$\mathcal{X} = \{x_1, x_2, \dots, x_N\}$$

とする。 \mathcal{X} にもとづく、母数 θ の推定量を $\hat{\theta}$ とする。いま、 $\hat{\theta}$ の標準誤差

$$SE = [E_G\{\hat{\theta} - E_G(\hat{\theta})\}^2]^{1/2} \quad (6.1)$$

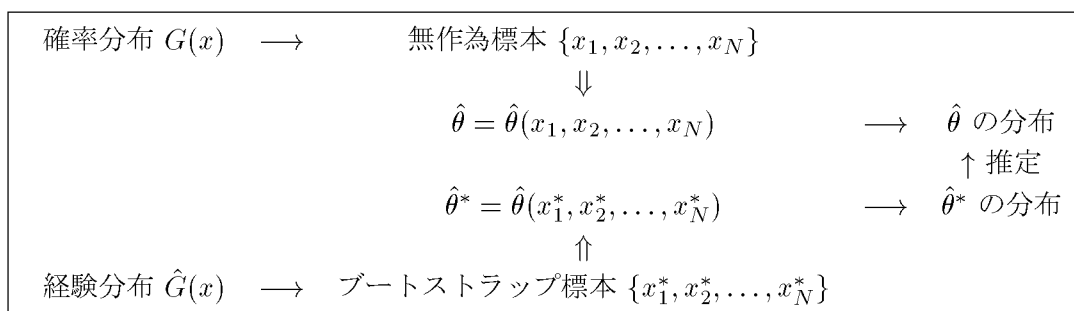


図 6.1: ブートストラップ法による分布推定

の推定を考える。例として、 G を平均 μ 、分散 σ^2 の正規分布 $N(\mu, \sigma^2)$ とし、 $\theta = \mu, \hat{\theta} = \bar{x}$ とする。よく知られているように、 \bar{x} は $N(\mu, \sigma^2/N)$ にしたがうので、この場合、 \bar{x} の標準誤差は σ/\sqrt{N} となる。未知母数 σ をその推定値 s で置き換え、 $\widehat{SE} = s/\sqrt{N}$ により SE を推定する。この例で分かるように、通常の方法は、 $\hat{\theta}$ の標準誤差を母数の関数として表し、母数をその推定値で置き換えることにより標準誤差を推定する。一般には、 $\hat{\theta}$ の標準誤差の導出には、複雑な理論や数式が必要とされる。

ところで、標準誤差の定義 (6.1) 式を見ると分かるように、標準誤差は、 $\hat{\theta}$ と $\hat{\theta}$ の平均との差の 2 乗の (確率分布 G についての) 平均の平方根である。従って、もしも G が既知ならば、数式を用いなくても、モンテカルロ法により、

1. 母集団 G から無作為標本 \mathcal{X} を抽出する
2. 推定値 $\hat{\theta}$ を計算する

という操作を多数回 (B 回) 繰り返せば、得られた B 個の推定値 $\hat{\theta}_{(1)}, \hat{\theta}_{(2)}, \dots, \hat{\theta}_{(B)}$ の標準偏差を計算することにより、標準誤差を数値的に求めることができる。(もちろん、こうして計算された標準誤差には、シミュレーションによる誤差が含まれるが、必要な精度に応じてシミュレーションの回数を増やしてやればよい。) このことは、未知の分布 G をなんらかの方法で推定することができれば、標準誤差を推定できることを示唆している。

無作為標本 \mathcal{X} というのは、母集団の「縮図」であるから、 \mathcal{X} には、母集団、すなわちの確率分布 G についての情報が含まれている。ブートストラップ法の基本的な考え方は、未知の確率分布 G を、**経験分布** (empirical distribution)

$$\hat{G}(x) = \#\{x_i \leq x\}/N$$

により推定するというものである¹。ここで、 $\#\{x_i \leq x\}$ は、 x 以下の観測値の個数を表すものとする。すなわち、経験分布とは、(観測値に同一のものがない場合には) 観測値 x_i で値が $1/N$ 増加する階段状の関数を分布関数とするような分布である。未知の確率分布 G を経験分布 \hat{G} で推定したので、標準誤差の定義 (6.1) 式において、 G を \hat{G} で置き換える。そして、 G からの無作為標本 \mathcal{X} にもとづく推定値 $\hat{\theta}$ を、 \hat{G} からの無作為標本

$$\mathcal{X}^* = \{x_1^*, x_2^*, \dots, x_N^*\}$$

にもとづく推定値 $\hat{\theta}^*$ で置き換えると、標準誤差のブートストラップ推定値

$$\widehat{SE}_{\text{BOOT}} = \left[E_{\hat{G}} \{ \hat{\theta}^* - E_{\hat{G}}(\hat{\theta}^*) \}^2 \right]^{1/2} \quad (6.2)$$

¹ここで説明しているのは、ノンパラメトリックなブートストラップ法である。 G の分布形を既知とするパラメトリックなブートストラップ法もあるが、ここでは取り上げない。

が得られる。ここで経験分布からの無作為標本とは、観測データ $\{x_1, x_2, \dots, x_N\}$ からの復元抽出に他ならない。経験分布からの無作為標本を**ブートストラップ標本** (bootstrap sample) とよび、ブートストラップ標本の抽出を**リサンプリング** (resampling) とよぶ。また、ブートストラップ標本から計算された推定量 $\hat{\theta}^*$ の分布を**ブートストラップ分布** (bootstrap distribution) とよぶ。すなわち、ブートストラップ法では、推定量 $\hat{\theta}$ の標本分布を、ブートストラップ分布で推定するのである。

ところで、標準誤差のブートストラップ推定値 $\widehat{SE}_{\text{BOOT}}$ をどのように計算したらよいのであろうか。 $\widehat{SE}_{\text{BOOT}}$ が具体的にどのような式で表されるかを、個々の問題についてあらかじめ導出しなければならないのがある。ここで、先ほど標準誤差について、もしも確率分布 G が既知ならば、モンテカルロ法により標準誤差を求めることができることを指摘したことを思い起こしてほしい。経験分布 \hat{G} は、データが得られた段階で既知となるので、標準誤差のブートストラップ推定値も、モンテカルロ法により求めることができるのである。すなわち、

1. \hat{G} からブートストラップ標本 \mathcal{X}^* を抽出する
2. \mathcal{X}^* にもとづく推定値 $\hat{\theta}^*$ を計算する

という操作を B 回繰り返す。そして、得られた B 個の推定値 $\hat{\theta}_{(1)}^*, \hat{\theta}_{(2)}^*, \dots, \hat{\theta}_{(B)}^*$ の標準偏差

$$\left[\frac{1}{B-1} \sum_{b=1}^B \{\hat{\theta}_{(b)}^* - \hat{\theta}_{(\cdot)}^*\}^2 \right]^{1/2} \quad (\text{ただし, } \hat{\theta}_{(\cdot)}^* = \frac{1}{B} \sum_{b=1}^B \hat{\theta}_{(b)}^*) \quad (6.3)$$

により $\widehat{SE}_{\text{BOOT}}$ を数値的に近似する。

バイアスのブートストラップ推定も同様である。バイアスのブートストラップ推定値を $\widehat{BIAS}_{\text{BOOT}}$ とすると、

$$\widehat{BIAS}_{\text{BOOT}} = E_{\hat{G}} \hat{\theta}^* - \hat{\theta} \quad (6.4)$$

$$\approx \hat{\theta}_{(\cdot)}^* - \hat{\theta} \quad (6.5)$$

ここで、(6.4) が $\widehat{BIAS}_{\text{BOOT}}$ の定義式であり、(6.5) は、そのモンテカルロ近似である。

ブートストラップ法を利用した信頼区間の構成には、さまざまな方法がある。よく知られているのは、ブートストラップ分布の分位点を用いたパーセンタイル法である。すなわち、 B 個のブートストラップ推定値 $\hat{\theta}_{(1)}^*, \hat{\theta}_{(2)}^*, \dots, \hat{\theta}_{(B)}^*$ に対して、 $\hat{\theta}_\alpha^*$ を

$$\#\{\hat{\theta}_{(b)}^* < \hat{\theta}_\alpha^*\} / B = \alpha$$

とすると、パーセンタイル法による信頼係数 $1 - 2\alpha$ の信頼区間は、

$$(\hat{\theta}_\alpha^*, \hat{\theta}_{1-\alpha}^*) \quad (6.6)$$

で与えられる。パーセンタイル法による信頼区間は、Efron (1979) により提案されたもので、ブートストラップ法を利用した信頼区間の中では最も簡単なものである。その後、いろいろな改良が提案されている。しかし、共分散構造分析の場合、既存のソフトウェアの利用を前提にすると、利用可能なブートストラップ信頼区間は、パーセンタイル法に限られる。

共分散構造分析に限らず、しばしば用いられるのは、

$$\frac{\hat{\theta} - \theta}{\widehat{SE}}$$

の分布を標準正規分布 $N(0, 1)$ で近似した信頼区間

$$(\hat{\theta} - z_{1-\alpha}\widehat{SE}, \hat{\theta} - z_{\alpha}\widehat{SE})$$

である。(ここで、 z_{α} は、標準正規分布の $100\alpha\%$ 点である。)パーセンタイル法による信頼区間の精度は、こうした信頼区間と同程度である。ただし、個々の問題において、どちらの信頼区間が優れているかは、シミュレーションにより確認する必要がある。

なお、必要とされるブートストラップ標本の数であるが、標準誤差やバイアスの推定の場合には $50 \sim 200$ 、信頼区間の構成の場合には $1,000 \sim 2,000$ とされている。

6.3 共分散構造分析におけるブートストラップ法の利用

共分散構造分析におけるブートストラップ法の利用は、次の4つに分類することができるであろう。

1. 標準誤差の推定
2. バイアスの推定
3. 信頼区間の構成
4. モデルの評価

このうち、1,2,3については、一般の統計モデルの場合と特に変わるところはない。ブートストラップ法が正しい結果を与えることは、漸近理論により、一般的に示されている。従って、共分散構造分析の場合にも、実際にわれわれが経験するような状況でブートストラップ法の「よさ」が確認されれば、あとは「使ってみればよい……」ということになる。ただし、共分散構造分析の場合には、モデルのあてはめに数値的最適化の方法を用い、繰り返し計算を行うために、計算時間がかかる、という問題がある。

また、探索的因子分析モデルの場合のように、係数の符号や係数行列(因子負荷量行列など)の列の順番の任意性がある場合には、リサンプリングにともなう変動を考慮し、係数の正負や係数行列の列の順番をもとの標本と一致させる必要がある。さもないと、標準誤差のブートストラップ推定値が、異様に大きくなったりする。

モデルの評価への利用には、やや共分散構造分析に特有な点がある。最もよく知られているのは、ブートストラップ適合度検定であろう。ブートストラップ適合度検定は、帰無仮説 H_0 と対立仮説 H_1 をそれぞれ

$$H_0 : \Sigma = \Sigma(\theta)$$

$$H_1 : \Sigma > 0$$

とした共分散構造モデルに関する仮説検定にブートストラップ法を利用するというものである。正規分布の仮定の下では、 H_0 が正しいければ、 $n \cdot F_{ML}(S, \Sigma(\hat{\theta}))$ の分布が、自由度 $d = p(p+1)/2 - q$ (ただし、 q はモデルに含まれる自由パラメータ数)のカイ2乗分布で近似できることが知られている。そこで、検定の有意水準を α とすると、データから計算された $n \cdot F_{ML}(S, \Sigma(\hat{\theta}))$ の値が自由度 d のカイ2乗分布の上側 $100\alpha\%$ 点より大きければ H_0 を棄却する。ブートストラップ適合度検定の場合には、検定に用いる統計量の分布を、そのブートストラップ分布で近似する。ただし、ブートストラップ適合度検定では、データ $\{X_1, X_2, \dots, X_N\}$ からのリサンプリングを行わず、各観測値を

$$Y_i = \Sigma(\hat{\theta})^{1/2} S^{-1/2} X_i \quad (i = 1, \dots, N)$$

により変換したデータ $\{Y_1, Y_2, \dots, Y_N\}$ からのリサンプリングを行う。(ここで $A^{1/2}$ は、正定値行列 A に対して、 $A^{1/2}(A^{1/2})' = A$ を満たす行列である。また、 $A^{-1/2}$ は、正定値行列 A の逆行列 A^{-1} (これも正定値行列である) に対して、 $A^{-1/2}(A^{-1/2})' = A^{-1}$ を満たす行列である。) こうすることにより、ブートストラップ分布の母分散共分散行列が $\Sigma(\hat{\theta})$ となり、帰無仮説で指定した構造を持つ。この変換したデータからのリサンプリングの方法は、Bollen-Stine (1993) により提案された。

検定の実際の手順であるが、 b 番目のブートストラップ標本 (変換したデータからリサンプリングしたものである) から計算された分散共分散行列を $S_{(b)}^*$ 、推定値を $\hat{\theta}_{(b)}^*$ とし、 $n \cdot \hat{F}_{(b)}^* = n \cdot F(S_{(b)}^*, \Sigma(\hat{\theta}_{(b)}^*))$ とする。そして B 個の $n \cdot \hat{F}_{(b)}^*$ の上側 $100\alpha\%$ 点を $n \cdot \hat{F}_\alpha^*$ とすると、

$$n \cdot F(S, \Sigma(\hat{\theta})) > n \cdot \hat{F}_\alpha^*$$

の場合に H_0 を棄却する、というものである。探索的因子分析モデルのブートストラップ適合度検定については、Ichikawa-Konishi (1997) を参照されたい。

その外に、共分散構造分析におけるモデル評価へのブートストラップ法の利用としては、Yung-Bentler (1994) による、適合度検定統計量の分布のカイ2乗近似の精度を改善する試みがある。また、ブートストラップ法を利用して GFI や AGFI などの適合度指標の信頼区間を構成する方法も提案されている (Bone-Sharma-Shimp 1989)。さらに、ブートストラップ法を用いることにより AIC を拡張し、より柔軟なモデル評価を可能にした情報量規準 EIC (Ishiguro-Sakamoto-Kitagawa 1997) の共分散構造分析への利用については、Ichikawa-Konishi (1999) を参照されたい。

ところで、どのような場合にブートストラップ法を使うべきか、ということであるが、共分散構造分析における統計的推測で最も多く用いられているのは、正規分布を仮定した最尤法であろう。正規分布の仮定のもとでは、ブートストラップ法は正しい結果を与えるが、ブートストラップ法を用いる積極的な理由はない。しかし、実際には正規分布の仮定は、程度の差はあっても疑わしい。正規分布を仮定した最尤法を、非正規分布のもとで用いても、母数の推定量は一致性をもつが、標準誤差の推定や適合度検定の結果などは正しくない。もちろん、真の分布が楕円分布の場合には、正規分布を仮定した最尤法にもとづく統計的推測の修正が知られているが、それは楕円分布というかなり限定された分布に対してしか正しくない。ところが、非正規分布のもとで最尤法を用いた場合でも、標準誤差の推定や適合度検定をブートストラップ法により行えば、正しい結果を与える。これは、ブートストラップ法が非正規性の影響の補正を自動的に行うことによる。実は、一般に正規分布・非正規分布を問わず、ブートストラップ法による統計的推測は、推定方法 (不一致度関数の選択に対応する) にかかわらず正しい結果を与える。大量の計算は、ブートストラップ法のこうした好ましい性質の代償といえるであろう。

共分散構造分析におけるブートストラップ法の利用については、Yung-Bentler (1995) による総合報告があるので、興味や関心のある方は参照していただきたい。ところで、この総合報告を見て分かることは、ブートストラップ法の利用についての研究はあまり多くはない、ということである。なぜ、共分散構造分析におけるブートストラップ法のよさ (悪さ) を検証した論文が少ないのであろうか。あるデータに適用してみて、ブートストラップ法がもっともらしい結果を与えることを確認しただけでは不十分なのであろうか。実は、ブートストラップ法が単に理論的だけでなく、実際に「よい」ことを確かめるためには、ある母集団から無作為標本を抽出し、ブートストラップ法を適用する、という操作を多数回繰り返すシミュレーションを行わなければならない。そして、その多数回の繰り返しを通して、ブートストラップ法がよい成績を示すことを確認しなければならない (Ichikawa-Konishi 1995)。ブートストラップ法そのものが、リサンプリングとモデルのあてはめの繰り返しというシミュレーションを含んでいるので、入れ子になったシミュレー

ションを行う必要がある。共分散構造分析では、1回のモデルのあてはめにも多くの計算が必要なので、ブートストラップ法の「よさ」をシミュレーションにより検証するのは容易ではない。しかも、シミュレーションの結果は、用いた特定の共分散構造モデルやパラメータの値に依存するものであり、ただちに共分散構造分析モデル全般に一般化できるわけではない。

最後に、ブートストラップ法を利用する場合の注意であるが、ブートストラップ標本には、同じ観測値が重複して選ばれるので、数値的には条件が悪くなり、不適解や非収束が発生しやすくなる。ブートストラップ標本での不適解・非収束の頻度が高い場合には、ブートストラップ法の結果の利用は避けた方がよいであろう。

6.4 ソフトウェアの対応

共分散構造分析のソフトウェアのいくつかは、ブートストラップ法に対応していることをうたっている。ソフトウェアの製作者にとって、とりあえずブートストラップ法に対応するのは、それほど難しいことではないであろう。特に、シミュレーションの機能を持っている場合には、ブートストラップ法への対応は、「あと一歩」である。母集団からの標本抽出(サンプリング)に、観測データからのブートストラップ標本の抽出、すなわち、リサンプリングの機能を加えればよい。ここでは、共分散構造分析のソフトウェアのうち、AMOS (3.6), EQS (5.7A), SEPATH (STATISTICA 5.0J) について見ていく。それによって、製作者が、ブートストラップ法をどのように評価しているかを窺い知ることでもできるであろう。

EQS と SEPATH では、ブートストラップのメニューの中から、標準誤差の推定、バイアス推定、信頼区間の構成、ブートストラップ適合度検定を選ぶ、という形式にはなっていない。では、どのようになっているのかというと、ブートストラップをシミュレーション機能のひとつとして位置づけ、ブートストラップ標本 $\mathcal{X}_{(1)}^*, \mathcal{X}_{(2)}^*, \dots, \mathcal{X}_{(B)}^*$ に共分散構造モデルをあてはめた結果を保存しておく。そして、その結果を記述統計のプログラムで用いるのである。つまり、 $\hat{\theta}_{(1)}^*, \hat{\theta}_{(2)}^*, \dots, \hat{\theta}_{(B)}^*$ の標準偏差を計算することにより、標準誤差のブートストラップ推定値を得る ((6.3) 式参照)。同様に、 $\hat{\theta}_{(1)}^*, \hat{\theta}_{(2)}^*, \dots, \hat{\theta}_{(B)}^*$ の平均 $\hat{\theta}_{(\cdot)}^*$ から $\hat{\theta}$ を引くことにより、バイアスのブートストラップ推定値を得る ((6.5) 式参照)。さらに、 $\hat{\theta}_{(1)}^*, \hat{\theta}_{(2)}^*, \dots, \hat{\theta}_{(B)}^*$ をソートして大きさの順に並べ替えるかヒストグラムを描くことにより、 B 個の値の分位点 $\hat{\theta}_{\alpha}^*$ と $\hat{\theta}_{1-\alpha}^*$ を求めれば、パーセンタイル法による信頼区間を構成することができる。(同様のことを、各種の適合度指標に対しても行うことができる)。このような形でブートストラップに対応すれば、各ブートストラップ標本の分析結果で保存可能なものは、そのブートストラップ分布を調べることができる。ただし、Bollen-Stine (1993) の修正を用いたブートストラップ適合度検定を行うことはできない。

これに対して、AMOS はブートストラップに関連したコマンドをいくつか用意している。すべての母数の推定値について、コマンドひとつでその標準誤差のブートストラップ推定値を計算させることができる。また、不一致度関数の値 $F(S_{(b)}^*, \Sigma(\hat{\theta}_{(b)}^*))$ のヒストグラムもコマンドひとつで作成する。さらに、EQS や SEPATH では不可能だった、Bollen-Stine (1993) の方法によるブートストラップ適合度検定もできる。

以上のように、ソフトウェアのブートストラップへの対応は、現在のところ、あまり積極的とはいえない。特に EQS の場合には、マニュアルの説明もわずかであることから、ライバルに遅れをとらないために、シミュレーションの中にリサンプリングを加えることにより、とりあえずブートストラップに対応した、という印象を受ける。ブートストラップ法を利用したいという場合には、EQS や SEPATH よりも、AMOS の方がよいであろう。まとめると、ブートストラップへの

対応は、

$$\text{EQS} < \text{SEPATH} < \text{AMOS}$$

ということになる。ところで、これは、それぞれのソフトウェアの開発者の共分散構造分析における統計的推測理論への貢献が

$$\text{Arbuckle(AMOS)} < \text{Steiger(SEPATH)} < \text{Bentler(EQS)}$$

であることとちょうど逆の順になっている、というのうがちすぎた見方であろうか。

ソフトウェアのブートストラップ法への対応が積極的でない理由としては、次のようなものが考えられる。

1. 計算量の問題
2. 開発コストの問題
3. 共分散構造分析におけるブートストラップ法の「よさ」について、十分な知見がないこと

1. の計算量の問題についてはすでにのべた。

2. については、ブートストラップ法にきちんと対応するのであれば、計算時間の短縮のための効率的なリサンプリングのアルゴリズムや精度の高い信頼区間をサポートする必要がある。しかしこれらをプログラムにより実現しようと思えば、その分のコストがかかるのである。

3. については、すでに見たように、ブートストラップ法が単に理論的によいばかりでなく、実際によりよいことを検証するのが、共分散構造分析の場合には大変であるために、十分な研究がすすんでいない。例えば、共分散構造分析におけるさまざまなブートストラップ信頼区間の比較は、殆ど行われていない。ソフトウェアのブートストラップ法への対応は、まだまだこれからである。

付録：モデルファイル例

ここでは、本文で紹介した分析を実行するためのモデルファイル例をいくつか紹介する。主に EQS のモデルファイルを扱うが、基本モデルについては SAS CALIS のモデルファイルも紹介する。

図 2.1: 中古車データのパス解析 [CALIS]

```
DATA USED CAR;
  INPUT YEN KM YEAR SHAKEN;
CARDS;
  89    4.3    5     24
  99    1.9    4     18
 128    5.2    2     13
  98    5.1    3     4
  52    4.0    6    15
  47    4.8    8    24
  40    8.7    7     3
  39    8.2    7     6
  38    3.3   10    14
  48    3.9    6     0
  27    8.2    8    24
  23    7.2    8    24
;
PROC CALIS DATA=USED CAR M=ML ALL NOMOD;
LINEQS
  YEN = K:YEAR + K:KM + K:SHAKEN + E1,
  KM = K:YEAR + E2;
STD
  SHAKEN = PHI1,
  YEAR = PHI2,
  E1 - E2 = DEL1-DEL2;
RUN;
```

注 1 パス係数は、K: という省略形で指定することが可能。

注 2 ALL NOMOD で、修正指標 (modification index) 以外の出力をすべて要求する。修正指標は計算に時間がかかるため、必要なときのみ出力する。

図 2.1: 中古車データのパス解析 [EQS]

```
/TITLE
  chukosha data no kaiseki
/SPECIFICATIONS
  DATA='CHUKOSHA.DAT';
  VARIABLES= 4; CASES= 12;
  METHODS=ML; MATRIX=RAW;
/LABELS
  V1=PRICE; V2=KM; V3=YEAR; V4=SHAKEN;
/EQUATIONS
  V1 = *V2 + *V3 + *V4 + E1;
  V2 = *V3 + E2;
/VARIANCES
  V3 = *; V4= *;
  E1 = *; E2= *;
/COVARIANCES
/OUTPUT
  PARAMETERS; STANDARD ERRORS; LISTING;
/END
```

注 1 EQS では、生データは外部ファイルとして保存し、モデルファイルでパスとファイル名を指定する。モデルファイル内に記述できるデータは、「相関行列+標準偏差」または「分散共分散行列」である。上記では、CALIS プログラムのデータの部分を CHUKOSHA.DAT ファイルに格納し、上記モデルファイルと同一ディレクトリ (フォルダ) に保存してある。

注 2 EQS では、パス係数や分散・共分散の未知パラメータを * で表す。

注 3 /OUTPUT は出力のコントロールを行うが、ここでは、「おまじない」と考えておけばよい。

図 2.2: 6 科目のデータの因子分析 [CALIS]

```

DATA SCHOOL(TYPE=CORR);
  _TYPE_ = 'CORR'; INPUT _NAME_ $ X1-X6;
LABEL
  X1='Gaelic' X2='English'
  X3='History' X4='Arithmet'
  X5='Algebra' X6='Geometry';
CARDS;
  X1 1      .      .      .      .      .
  X2 0.439 1      .      .      .      .
  X3 0.410 0.351 1      .      .      .
  X4 0.288 0.354 0.164 1      .      .
  X5 0.329 0.320 0.190 0.595 1      .
  X6 0.248 0.329 0.181 0.470 0.464 1
;
PROC CALIS DATA=SCHOOL M=ML DF=219
  ALL NOMOD;
  LINEQS
    X1= L_11 F1      + E1,
    X2= L_21 F1      + E2,
    X3= L_31 F1      + E3,
    X4=      L_42 F2 + E4,
    X5=      L_52 F2 + E5,
    X6=      L_62 F2 + E6;
  STD
    E1-E6 = DEL1-DEL6,
    F1-F2 = 2*1.00;
  COV
    F1 F2 = PHI12;
RUN;

```

注 1 DATA ステップで、相関行列の上三角部分を省略するときは、“.”を入れる。

注 2 DATA ステップで標本サイズを指定していない場合は、PROC ステップで指定できる。DF=219 における 219 は、標本サイズ-1 である。

注 3 F1, F2 は独立潜在変数である。独立潜在変数の尺度は、「分散=1」で指定することが多い。2*1.00 は 1.00 が 2 つ並ぶことを意味する。

図 2.2: 6 科目のデータの因子分析 [EQS]

```

/TITLE
  cfa with two factors
/SPECIFICATIONS
  VARIABLES=6; CASES=220;
  METHODS=ML; MATRIX=COVARIANCE;
/LABELS
  V1=Gaelic;V2=English;V3=History;
  V4=Arithmet;V5=Algebra;V6=Geometry;
/EQUATIONS
  V1 = *F1      + E1;
  V2 = *F1      + E2;
  V3 = *F1      + E3;
  V4 =      *F2 + E4;
  V5 =      *F2 + E5;
  V6 =      *F2 + E6;
/VARIANCES
  F1 = 1.00; F2 = 1.00;
  E1 TO E6 = *;
/COVARIANCES
  F2 , F1 = *;
/PRINT
  FIT=ALL;
/MATRIX
  1.000
  0.439 1.000
  0.410 0.351 1.000
  0.288 0.354 0.164 1.000
  0.329 0.320 0.190 0.595 1.000
  0.248 0.329 0.181 0.470 0.464 1.000
/OUTPUT
  PARAMETERS;STANDARD ERRORS;LISTING;
/END

```

注 1 MATRIX=COVARIANCE は入力データが分散共分散行列であることを示す。しかし、実際の入力相関行列である。厳密に言えば、相関行列を分散共分散行列とみなして分析するのは間違いである。特に、推定量の標準誤差が正しくない。しかし、現在のところ、このような分析が一般的になっている。

注 2 FIT=ALL; によって、EQS のデフォルトに加え各種適合度指標の出力を要求する。

図 2.3: 多重指標モデルと推定結果 [CALIS]

```

DATA FOOD(TYPE=CORR);
  INPUT _TYPE_ $ _NAME_ $
  TENKA BARANSU GAKU KAISU;
CARDS;
  N      .      831   831   831   831
CORR TENKA   1.000 0.301 0.168 0.257
CORR BARANSU 0.301 1.000 0.188 0.328
CORR GAKU    0.168 0.188 1.000 0.530
CORR KAISU   0.257 0.328 0.530 1.000
;
PROC CALIS DATA=FOOD M=ML ALL NOMOD;
LINEQS
  TENKA   = L_11 F_ISHIKI + E1,
  BARANSU = L_21 F_ISHIKI + E2,
  GAKU    = 1.00 F_IYOKU  + E3,
  KAISU   = L_42 F_IYOKU  + E4,
  F_IYOKU = B_21 F_ISHIKI + D2;
STD
  E1-E4    = DEL1-DEL4,
  D2       = PSI2,
  F_ISHIKI = 1.00;
RUN;

```

注1 DATA ステップで標本サイズを指定している。

注2 従属潜在変数である F_IYOKU の尺度は、GAKU へのパス係数を 1 と固定することで定める。独立潜在変数である F_ISHIKI の尺度はその分散を 1 とおいている。

図 2.3: 多重指標モデルと推定結果 [EQS]

```

/TITLE
  Multiple indicator model
/SPECIFICATIONS
  VARIABLES= 4; CASES= 831;
  METHODS=ML; MATRIX=COVARIANCE;
/LABELS
  V1=TENKA; V2=BARANSU;
  V3=GAKU; V4=KAISU;
/EQUATIONS
  V1 = *F1 + E1;
  V2 = *F1 + E2;
  V3 = 1.0F2 + E3;
  V4 = *F2 + E4;
  F2 = *F1 + D2;
/VARIANCES
  F1= 1.00;
  E1 TO E4= *; D2= *;
/COVARIANCES
/OUTPUT
  PARAMETERS;STANDARD ERRORS;LISTING;
/MATRIX
  1.000
  0.301 1.000
  0.168 0.188 1.000
  0.257 0.328 0.530 1.000
/END

```

注1 従属潜在変数である F2 の尺度は、V3 へのパス係数を 1 と固定することで定める。独立潜在変数である F1 の尺度はその分散を 1 とおいている。

図 3.1: 語彙学習データの分析 [EQS]:
コントロール群, 平均構造なし

```

/TITLE
  analysis of control group
  without mean
/SPECIFICATIONS
  VARIABLES= 4; CASES= 105;
  METHODS=ML; MATRIX=COVARIANCE;
/LABELS
  V1=pre_syn; V2=pre_opp;
  V3=post_syn; V4=post_opp;
/EQUATIONS
  V1 = 1.0F1 + E1;
  V2 = *F1 + E2;
  V3 = 1.0F2 + E3;
  V4 = *F2 + E4;
  F2 = *F1 + D2;
/VARIANCES
  F1 = *;
  E1 TO E4 = *;
  D2 = *;
/COVARIANCES
  E4,E2 = *;
/OUTPUT
  PARAMETERS;STANDARD ERRORS;LISTING;
/MATRIX
  37.626
  24.933 34.680
  26.639 24.236 32.013
  23.649 27.760 23.565 33.443
/END

```

注 1 平均構造があるモデルや多母集団の分析を考
えるときは、潜在変数の尺度は独立変数であつ
ても、パス係数=1 とする方がよい。本解析
は平均構造があるモデルや多母集団の分析へ
の第一歩であるので、パス係数=1 のように
して尺度を定めた。

注 2 誤差変数間の相関を許している。

図 3.3: 語彙学習データの分析 [EQS]:
コントロール群, 平均構造あり

```

/TITLE
  analysis of control group
  with mean structure
/SPECIFICATIONS
  VARIABLES= 4; CASES= 105;
  METHODS=ML; MATRIX=COVARIANCE;
  ANALYSIS=MOMENT;
/LABELS
  V1=pre_syn; V2=pre_opp;
  V3=post_syn; V4=post_opp;
/EQUATIONS
  V1 = *V999 + 1.0F1 + E1;
  V2 = *V999 + *F1 + E2;
  V3 = *V999 + 1.0F2 + E3;
  V4 = *V999 + *F2 + E4;
  F2 = *V999 + *F1 + D2;
/VARIANCES
  F1 = *;
  E1 TO E4 = *;
  D2 = *;
/COVARIANCES
  E4,E2 = *;
/OUTPUT
  PARAMETERS;STANDARD ERRORS;LISTING;
/CONSTRAINTS
  (V1,V999)=(V3,V999);
  (V2,V999)=(V4,V999);
/MATRIX
  37.626
  24.933 34.680
  26.639 24.236 32.013
  23.649 27.760 23.565 33.443
/MEAN
  18.381 20.229 20.400 21.343
/LMT
/END

```

注 1 ANALYSIS=MOMENT; を記述することで、平均
構造のあるモデルで分析することを宣言する。
/MEAN セクションに標本平均を入力する。

注 2 変数に (切片としての) 平均を設定するには、
*V999 を方程式に付加する。

注 3 (V1,V999)=(V3,V999); は、変数 V1 と V3
の平均 (切片項) を等置することを示す。

注 4 等式制約を課したときは、/LMT を設定して、
等式制約が妥当であったかどうかを検討して
おく。LMT は Lagrange Multiplier Test の
略。

図 3.5: 語彙学習データの分析 (最終モデル)
[EQS]: 同時分析, 平均構造あり

```

/TITLE
  analysis of control and training
  group with mean structure
  CONTROLE GROUP
/SPECIFICATIONS
  VARIABLES= 4; CASES= 105;
  METHODS=ML; MATRIX=COVARIANCE;
  ANALYSIS=MOMENT; GROUPS=2;
/LABELS
  V1=pre_syn; V2=pre_opp;
  V3=post_syn; V4=post_opp;
/EQUATIONS
  V1 = *V999 + 1.0F1 + E1;
  V2 = *V999 + *F1 + E2;
  V3 = *V999 + 1.0F2 + E3;
  V4 = *V999 + *F2 + E4;
  F2 = *V999 + *F1 + D2;

/VARIANCES
  E1 TO E4 = *;
  F1 = *; D2 = *;
/COVARIANCES
  E4,E2 = *;
/OUTPUT
  PARAMETERS;STANDARD ERRORS;LISTING;
/MATRIX
  37.626
  24.933 34.680
  26.639 24.236 32.013
  23.649 27.760 23.565 33.443
/MEAN
  18.381 20.229 20.400 21.343
/CONSTRAINTS
  (V1,V999)=(V3,V999);
  (V2,V999)=(V4,V999);
/END

```

右のカラムへ続く

- 注 1** 多母集団の同時分析をするには, 比較する群の数を `GROUP=2`; によって宣言する.
- 注 2** 各群においてモデルファイルを作成する. 上記で, 左カラムがコントロール群の, 右カラムがトレーニング群のモデルファイルである. これらを比較して違いを確認されたい.
- 注 3** 群間の制約は最終群に設定する.
- 注 4** 従属変数に設定する切片項は, 方程式に `*V999` を付加する.

左のカラムからの続き

```

/TITLE

  TRAINING GROUP
/SPECIFICATIONS
  VARIABLES= 4; CASES= 108;
  METHODS=ML; MATRIX=COVARIANCE;
  ANALYSIS=MOMENT;
/LABELS
  V1=pre_syn; V2=pre_opp;
  V3=post_syn; V4=post_opp;
/EQUATIONS
  V1 = *V999 + 1.0F1 + E1;
  V2 = *V999 + *F1 + E2;
  V3 = *V999 + 1.0F2 + E3;
  V4 = *V999 + *F2 + E4;
  F2 = *V999 + *F1 + D2;
  F1 = *V999 + D1;
/VARIANCES
  E1 TO E4 = *;
  D1 TO D2 = *;
/COVARIANCES
  E4,E2 = *;
/OUTPUT
  PARAMETERS;STANDARD ERRORS;LISTING;
/MATRIX
  50.084
  42.373 49.872
  40.760 36.094 51.237
  37.343 40.396 39.890 53.641
/MEAN
  20.556 21.241 25.667 25.870
/CONSTRAINTS
  (1,V1,V999)=(2,V1,V999);
  (1,V2,V999)=(2,V2,V999);
  (1,V3,V999)=(2,V3,V999);
  (1,V4,V999)=(2,V4,V999);
  (1,V2,F1)=(2,V2,F1);
  (1,V4,F2)=(2,V4,F2);
  (1,F2,F1)=(2,F2,F1);
/LMT
/END

```

- 注 5** 独立変数の平均の設定は, 現在のところ, やや複雑である. 独立変数 F1 に平均を設定するには

$$F1 = *V999 + D1;$$

とする. F1 を従属変数として扱い, 新たに独立変数 D1 を導入し, D1 に関して分散・共分散を設定する.

図 3.6: 語彙学習データの同時分析 [EQS]:
簡便法

```

/TITLE
  analysis of experimental data:
  simple method
/SPECIFICATIONS
  VARIABLES = 5; CASES = 213;
  METHODS=ML; MATRIX=CORRELATION;
/LABELS
  V1=pre_syn; V2=pre_opp;
  V3=post_syn; V4=post_opp;
  V5=treatment;
/EQUATIONS
  V1 = 1.0F1 + E1;
  V2 = *F1 + E2;
  V3 = 1.0F2 + E3;
  V4 = *F2 + E4;
  F2 = *F1 + *V5 + D2;
/VARIANCES
  V5 = .25*;
  F1 = 40*;
  E1 TO E4 = 10*;
  D2 = 5*;
/COVARIANCES
  V5,F1 = .5*;
  E4,E2 = 7*;
/OUTPUT
  PARAMETERS;STANDARD ERRORS;LISTING;
/MATRIX
  1.000
  0.783 1.000
  0.782 0.693 1.000
  0.704 0.774 0.776 1.000
  0.163 0.078 0.379 0.325 1.000
/STANDARD DEVIATIONS
  6.687 6.499 6.950 6.957 0.500
/END

```

注 1 未知パラメータを表す“*”の前に、初期値を設定することができる。推定値は、適当に選ばれた初期値からスタートする反復法で計算される。平均構造のあるモデルでは初期値の設定が難しく、時に、反復が収束しないことがある。そのような場合には、解析者が適切な初期値を与えてやることになる。

図 3.7: 顧客満足度の分析 [EQS]

```

/TITLE
  CS(kokyaku manzokudo) no bunseki
  Jikken deta no bunseki
/SPECIFICATIONS
  VARIABLES= 13; CASES= 180;
  METHODS=ML; FO='(13F3.2)';
  MATRIX=COVARIANCE;
/LABELS
  V1=EXPECT1;V2=EXPECT2;V3=PERFORM1;
  V4=PERFORM2;V5=DISCONF1;V6=DISCONF2;
  V7=SATISF1;V8=SATISF2;V9=SATISF3;
  V10=SATISF4;V11=SATISF5;
  V12=M-EXP; V13=M-PERF;
/EQUATIONS
  V1= 1.0F1 + E1;
  V2= *F1 + E2;
  V3= 1.0F2 + E3;
  V4= *F2 + E4;
  V5= 1.0F3 + E5;
  V6= *F3 + E6;
  V7= 1.0F4 + E7;
  V8= *F4 + E8;
  F1= *V12 + D1;
  F2= *F1 + *V13 + D2;
  F3= *F1 + *F2 + *V12 + *V13 + D3;
  F4= *F1 + *F2 + *F3 + D4;
/VARIANCES
  V12 TO V13 = *;
  E1 TO E8 = *; D1 TO D4 = *;
/COVARIANCES
  V13,V12 = *;
/PRINT
  fit=all;
/OUTPUT
  PARAMETERS;STANDARD ERRORS;LISTING;
/MATRIX
  100
  92100
  42 40100
  37 37 94100
  -18-18 26 27100
  -10-13 36 32 71100
  24 22 58 57 51 59100
  20 16 58 54 55 66 76100
  21 19 57 56 49 61 69 71100
  16 16 57 53 53 64 74 86 75100
  10 12 34 34 39 37 57 57 57 59100
  80 80 43 40-26-19 22 15 18 11 03100
  00 00 47 41 57 66 57 63 45 57 33 00100
/END

```

注 1 データのフォーマットを FO='(13F3.2)'; のように Fortran 形式で記述可能。

図 3.8: 潜在曲線モデル [EQS]

```

/TITLE
  Latent Curve Model
/SPECIFICATION
  cases=363; variables=6; method=ml;
  matrix=covariance; anaysis=moment;
/LABELS
  V1=alc1; V2=alc2; V3=alc3;
  V4=age; V5=gender; V6=coa;
  F1=intercpt; F2=slope;
/EQUATIONS
  V1= 1.0F1 + 0.0F2 + E1;
  V2= 1.0F1 + 1.0F2 + E2;
  V3= 1.0F1 + 2.0F2 + E3;
  V4= *V999 + E4;
  V5= *V999 + E5;
  V6= *V999 + E6;
  F1= *V999 + *V4 + *V5 + *V6 + D1;
  F2= *V999 + *V4 + *V5 + *V6 + D2;
/VARIANCES
  E1 TO E6 = *; D1 TO D2 = *;
/COVARIANCES
  E4,E5 = *; E5,E6 = *; E6,E4 = *;
  D1,D2 = *;
/PRINT
  fit=all;
/INEQUALITIES
  !(E1,E1) > -999999;
/MATRIX
  7.906
  7.634 15.828
  6.726 12.917 23.013
  1.238 1.750 1.532 1.968
  0.019 0.110 0.239 -0.022 0.250
  0.164 0.331 0.478 -0.104 0.001
  0.246
/MEANS
  1.3636 2.1185 3.1846 12.9146 0.5152
  0.5647
/OUTPUT
  PARAMETERS;STANDARD ERRORS;LISTING;
/END

```

注1 図 3.5 のモデルファイル注 3 でも述べたが、独立変数の平均の設定は面倒である。独立変数 V4 に平均を設定するには

$$V4 = *V999 + E4;$$

とし、E4 に分散・共分散を設定する。

注2 デフォルトでは、分散の推定値は非負になるよう設定されている。E1 の分散に負の値を許すには、/INEQUALITIES セクションに

$$(E1, E1) > -999999;$$

を記述する。なお、上記モデルファイルではコメント文にしてある。

図 3.10: Curran (1999) の解析 [EQS]

```

/TITLE
  growth curve model between adolescent
  and peer alcohol use regressed on age
  gender and parent alcoholism diagnosis
  Curran (1999) Model D
/SPECIFICATIONS
  CASES=363; VARIABLES=9; METHOD=ML;
  MATRIX=COVARIANCE; ANALYSIS=MOMENT;
/LABELS
  V1=alc1; V2=alc2; V3=alc3;
  V4=age; V5=gen; V6=coa;
  V7=peer1; V8=peer2; V9=peer3;
  F1=alcint; F2=alclsp;
  F3=peerint; F4=peerslp;
/EQUATIONS
  V1= 1.0F1 + 0F2 + E1;
  V2= 1.0F1 + 1F2 + E2;
  V3= 1.0F1 + 2F2 + E3;
  V4= 12.9*V999 + E4;
  V5= .515*V999 + E5;
  V6= .554*V999 + E6;
  V7= 1.0F3 + 0F4 + E7;
  V8= 1.0F3 + 1F4 + E8;
  V9= 1.0F3 + 2F4 + E9;
  F1= *V999 + *V4 + *V5 + *V6 + D1;
  F2= *V999 + *V5 + *V6 + *F3 + D2;
  F3= *V999 + *V4 + *V5 + *V6 + D3;
  F4= *V999 + *F1 + D4;
/VARIANCES
  E1 TO E9=*; D1 TO D4=*;
/COVARIANCES
  E6,E5=0.0*; E6,E4=-0.1*; E5,E4=0.0*;
  E7,E1=*; E8,E2=*; E9,E3=*;
  D2,D1=*; D4,D3=*; D4,D2=*; D3,D1=*;
/MEANS
  1.3636 2.1185 3.1846 12.9146 0.5152
  0.5647 1.3289 1.7543 2.4749
/MATRIX
  7.906
  7.634 15.828
  6.726 12.917 23.013
  1.238 1.750 1.532 1.968
  0.019 0.110 0.239 -0.022 0.250
  0.164 0.331 0.478 -0.104 0.001
  0.246
  3.246 4.021 3.853 1.179 -0.081
  0.078 3.047
  2.691 4.749 4.555 1.118 -0.022
  0.107 2.077 3.321
  2.147 4.055 5.849 0.822 -0.132
  0.159 1.691 2.126 4.039
/OUTPUT
  PARAMETERS;STANDARD ERRORS;LISTING;
/END

```

図 4.4: 双方向因果モデル：女子生徒 [EQS]

```

/TITLE
  good-looking vs smartness
/SPECIFICATIONS
  VARIABLES=4; CASES=209;
  METHODS=ML; MATRIX=CORRELATION;
/LABELS
  V1=smart;V2=looking;
  V3=GAP;V4=rating;
/TEC
  !ITER=300;
/EQUATIONS
  V1 = 0.0*V2 + *V3 + E1;
  V2 = 1.5*V1 + *V4 + E2;
/VARIANCES
  V3 TO V4 = *;;
  E1 TO E2 = *;
/COVARIANCES
  V4, V3 = *;
/PRINT
  FIT=ALL;
/MATRIX
  1.00
  .50 1.00
  .49 .32 1.00
  .09 .43 .15 1.00
/STANDARD DEVIATIONS
  .16 .49 3.49 1.01
/OUTPUT
  PARAMETERS;STANDARD ERRORS;LISTING;
/END

```

注1 MATRIX=CORRELATION; を設定すると、必ず /STANDARD DEVIATIONS を記述する必要がある。

注2 双方向因果モデルは初期値に敏感である。解析者が適切な初期値を設定しないと収束しないことがある。

注3 EQS の最大反復回数は 30 である。これをコントロールするには、/TEC セクションに ITER=300; などと記述する。なお、上記モデルファイルではコメント文にしてある。

参考文献

邦文

- 丘本 正 (1986). 因子分析の基礎. 日科技連
- 丘本 正 (1987). 因子分析の最近の発展. 「社会科学の計量分析 (鈴木雪夫・竹内 啓 編)」 pp. 7-26. 東京大学出版会.
- 狩野 裕 (1990). 因子分析における統計的推測: 最近の発展. 行動計量学, **18**, 3-12.
- 狩野 裕 (1996). 因果構造分析の活用法 — EQS による共分散構造分析入門. マーケティング・リサーチャー, **74**, 40-48. 日本マーケティング・リサーチ協会.
- 狩野 裕 (1996.2-1997.3). 共分散構造分析とソフトウェア. BASIC 数学 連載記事. 現代数学社.
- 狩野 裕 (1997). グラフィカル多変量解析: 目で見る共分散構造分析. 現代数学社.
- 狩野 裕 (1998). 不適解の原因と処理: 探索的因子分析. 大阪大学人間科学部紀要, **24**, 303-327.
- 小西貞則 (1988). ブートストラップ法による推定量の誤差評価. 「村上征勝・田村義保 (編) パソコンによるデータ解析」 pp. 123-142. 朝倉書店.
- 清水和秋 (1994). Jöreskog と Sörbom によるコンピュータプログラムと構造方程式モデル. 関西大学社会学部紀要, **25**, 1-41.
- 竹内 啓編 (1989). 統計学辞典. 東洋経済.
- 豊田秀樹 (1992). SAS による共分散構造分析. 東京大学出版会.
- 豊田秀樹 (1998a). 共分散構造分析 [入門編]. 朝倉書店.
- 豊田秀樹 (1998b). 共分散構造分析 [事例編]. 北大路書房.
- 豊田秀樹・前田忠彦・柳井晴夫 (1992). 原因をさぐる統計学. BLUE BACKS 講談社.
- 服部 環・海保博之 (1996). 心理データ解析. 福村出版.
- 原田 章・狩野 裕 (1998). 探索的因子分析における変数選択プログラム SEFA: オンラインマニュアル. <http://koko15.hus.osaka-u.ac.jp/~harada/factor/>
- 森棟公夫 (1985). 経済モデルの推定と検定. 共立出版.
- 柳井晴夫・繁樹算男・前川眞一・市川雅教 (1990). 因子分析: その理論と応用. 朝倉書店.
- 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.

英文

- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, **52**, 317-332.
- Arbuckle, J. L. (1995). AMOS 3.5. Smallwaters: Chicago.
- Bagozzi, R. P. and Yi, Y. (1989). On the use of structural equation models in experimental design. *Journal of Marketing Research*, **6**, 271-284.
- Bekker, P. A., Merckens, A. and Wansbeek, T. J. (1994). *Identification, Equivalent Models and Computer Algebra*. Academic Press, Inc.: San Diego.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, **107**, 238-246.
- Bentler, P. M. (1995). *EQS Structural Equations Program Manual*. Multivariate Software, Inc.: CA.
- Bentler, P. M. and Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, **88**, 588-606.
- Bentler, P. M. and Weeks, D. G. (1980). Linear structural equations with latent variables. *Psychometrika*, **45**, 289-308.

- Bentler, P. M. and Wu, E. J. C. (1995). *EQS for Windows User's Guide*. Multivariate Software, Inc.: CA.
- Bock, R. D. and Bergmann, R. E. (1966). Analysis of covariance structures. *Psychometrika*, **31**, 507-534.
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. Wiley: New York.
- Bollen, K. A. and Long, J. S. (1993). *Testing Structural Equation Models*. Sage Publications: Newbury Park.
- Bollen, K. N. and Paxton, P. (1998). Interactions of latent variables in structural equation models. *Structural Equation Modeling*, **5**, 267-293.
- Bollen, K. A. and Stine, R. A. (1993). Bootstrapping goodness-of-fit measures in structural equation models. In *Testing Structural Equation Models* (Bollen, K. A. and Long, J. S., Eds.), pp. 111-135. Sage Publications: CA.
- Bone, P. F., Sharma, S. and Shimp, T. A. (1989). A bootstrap procedure for evaluating goodness-of-fit indices of structural equation and confirmatory factor models. *Journal of Marketing Research*, **26**, 105-111.
- Browne, M. W. (1982). Covariance structures. In *Topics in Applied Multivariate Analysis* (Hawkins, D.M., Ed.), pp.72-141. Cambridge University Press: Cambridge, England.
- Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, **37**, 62-83.
- Browne, M. W. and Cudeck, R. (1993). Alternative ways of assessing model fit. In *Testing Structural Equation Models* (Bollen, K. and J. S. Long, Eds.), pp. 137-162. Sage Publications: Newbury Park.
- Browne, M. W. and Shapiro, A. (1988). Robustness of normal theory methods in the analysis of linear latent variate models. *British Journal of Mathematical and Statistical Psychology*, **41**, 193-208.
- Bryk, A. S. and Raudenbush, S. W. (1992). *Hierarchical Linear Model: Applications and Data Analysis Methods*. Sage: Newbury Park, CA.
- Bullock, H. E., Harlow, L. L. and Mulaik, S. A. (1998). Causal issues in structural equation modeling research. *Structural Equation Modeling*, **1**, 253-267.
- Chou, Chih-Ping, Bentler, P. M. and Pentz, M. A. (1995). Comparisons of two statistical approaches to study growth curves: The multilevel model and latent curve analysis. *Structural Equation Modeling*, **5**, 247-266.
- Churchill, G. A. and Surprenant, C. (1982). An investigation into the determinants of customer satisfaction. *Journal of Marketing Research*, **19**, 491-504.
- Curran, P. J. (1999). A latent curve framework for the study of development trajectories in adolescent substance use. In *Multivariate Applications in Substance Use Research* (Rose, J. et. al., Eds.), in press.
- Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap Methods and their Application*. Cambridge University Press: Cambridge.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, **7**, 1-26.
- Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. SIAM: Philadelphia.
- Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall: New York.
- Felson, R. B. and Bohrnstedt, G. W. (1979). Are the good beautiful or the beautiful good? The relationship between children's perceptions of ability and perceptions of physical attractiveness. *Social Psychology Quarterly*, **42**, 386-392.
- Fornell, C. (1982). *A Second Generation of Multivariate Analysis. Vol. 1 & 2*. Praeger: New York.
- Goldstein, H. (1995). *Multilevel Statistical Models* (2nd ed.). Edward Arnold and New York: Halsted, London.
- Homer, P. M. and Kahle, L. R. (1988). A structural equation test of the value-attitude-behavior hierarchy. *Journal of Personality and Social Psychology*, **54**, 638-646.
- Hu, L. and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, **6**, 1-55.
- Hu, L., Bentler, P. M. and Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, **112**, 351-362.
- Ichikawa, M. and Konishi, S. (1995). Application of the bootstrap methods in factor analysis. *Psychometrika*, **60**, 77-93.

- Ichikawa, M. and Konishi, S. (1997). Bootstrap tests for the goodness of fit in factor analysis. *Behaviormetrika*, **24**, 27-38.
- Ichikawa, M. and Konishi, S. (1999). Model evaluation and information criteria in covariance structure analysis. to appear in *British Journal of Mathematical and Statistical Psychology*.
- Ishiguro, M., Sakamoto, Y. and Kitagawa, G. (1997). Bootstrapping log likelihood and EIC, and extension of AIC. *Annals of Institute of Statistical Mathematics*, **49**, 411-434.
- Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, **34**, 183-202.
- Jöreskog, K. G. and Sörbom, D. (1981). *LISREL 5: User's Guide*. Chicago, IL.
- Jöreskog, K. G. and Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Scientific Software International, Inc.: Chicago, IL.
- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187-200.
- Kano, Y. (1993). Asymptotic properties of statistical inference based on Fisher consistent estimators in the analysis of covariance structures. In *Statistical Modelling and Latent Variables* (K. Haagen, D. J. Bartholomew and M. Deistler, Eds.), pp.173-190. Elsevier Science Publisher: Amsterdam.
- Kano, Y. (1997). Software. *Behaviormetrika*, **22**, 85-125.
- Kano, Y. and Harada, A. (1999). Stepwise variable selection in factor analysis. *Psychometrika*. in press.
- Kano, Y. and Ihara, M. (1994). Identification of inconsistent variates in factor analysis. *Psychometrika*, **59**, 5-20.
- Klein, L. R. (1950). *Economic Fluctuations in the United States 1921-1941*. Wiley: New York.
- Kotz, S. and Johnson, N. L. (1992). *Breakthroughs in Statistics Volume II Methodology and Distribution*. Springer: New York.
- Lawley, D. N. and Maxwell, A. E. (1963). *Factor Analysis as a Statistical Method*. Butterworth. [丘本正監訳 (1970). 因子分析法 日科技連]
- Magnus, J. R. and Neudecker, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley: New York.
- McArdle, J. J. and McDonald, R. P. (1984). Some algebraic properties of the Reticular Action Model for moment structures. *British Journal of Mathematical and Statistical Psychology*, **37**, 234-251.
- McDonald, R. P. (1978). A simple comprehensive model for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, **31**, 59-72.
- McDonald, R. P. (1980). A simple comprehensive model for the analysis of covariance structures: Some remarks and applications. *British Journal of Mathematical and Statistical Psychology*, **33**, 161-183.
- Meredith, W. and Tisak, J. (1990). Latent curve analysis. *Psychometrika*, **55**, 107-122.
- Muthén, B. O. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, **49**, 115-132.
- Muthén, L. K. and Muthén, B. O. (1998). *Mplus: User's Guide*. Muthén & Muthén: LA, CA.
- Schumacker, R. E. and Marcoulides, G. A. (1998). *Interactions and Nonlinear Effects in Structural Equation Modeling*. LEA: New Jersey.
- Shapiro, A. and Browne, M. W. (1987). Analysis of covariance structures under elliptical distributions. *Journal of American Statistical Association*, **82**, 1092-1097.
- Sörbom, D. G. (1978). An alternative to the methodology for the analysis of covariance. *Psychometrika*, **43**, 381-396.
- Spearman, C. (1904) "General intelligence," objectively determined and measured. *American Journal of Psychology*, **15**, 201-293.
- Steiger, J. H. (1994). *Structural Equation Modeling: Technical Documentation*. StatSoft, STATISTICA.
- Steiger, J. H. and Lind, J. C. (1980). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa.
- Thurstone, L. L. (1935). *Vector of Mind*. The University of Chicago Press: Chicago.
- Tucker, L. R. and Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, **38**, 1-10.
- Yung, Y. -F. and Bentler, P. M. (1994). Bootstrap-corrected ADF test statistics in covariance structure analysis. *British Journal of Mathematical and Statistical Psychology*, **47**, 63-84.

- Yung, Y. -F. and Bentler, P. M. (1996). Bootstrapping techniques in analysis of mean and covariance structures. In *Advanced Structural Equation Modeling: Issues and Techniques*. (Marcoulides, G. A. and Schumacker, R. E., Eds.), pp. 195-226. Lawrence Erlbaum Associates: Hillsdale, NJ.