

書 評

目 次

Introduction to Statistical Pattern Recognition, Second Edition (1990) Academic Press Keinosuke Fukunaga 著	西井龍映
多変量データ解析法—理論と応用— (1994) 朝倉書店 柳井晴夫著	岩崎 学
An Introduction to Regression Graphics (1994) John Wiley & Sons R. D. Cook and S. Weisbrg 著	水田正弘
Image Analysis for the Biological Sciences (1995) John Wiley & Sons C. A. Glasbey and G. W. Horgan 著	鎌倉稔成
消費者行動と日本の資産市場 (1996) 東洋経済新報社 羽森茂之著	下野恵子
計量経済学の理論と応用 (1996) 日本評論社 藁谷千風彦著	西山 茂
Behind State Company Nexus: One and half century experience of Japanese economic development in a statistical mirror (1996) MARUZEN Yoshiro Matsuda and Fumiko Arita 著	大平純彦
非線形経済時系列分析法とその応用 (1997) 岩波書店 刈屋武昭・照井伸彦著	吉田あつし
Analysis of Longitudinal Data (1994) Oxford University Press Peter J. Diggle., Kung-Yee Liang., and Scott L. Zeger 著	松山 裕
Health Measurement Scales: A Practical Guide to Their Development and Use, Second Edition (1995) Oxford University Press David L. Steiner & Geoffrey R. Norman 著	橋本修二
統計的多重比較の基礎 (1997) サイエンティスト社 永田 靖・吉田道弘著	浜田知久馬

**Introduction to Statistical Pattern
Recognition
Second Edition**

Keinosuke Fukunaga
Academic Press (1990, 591 頁)

広島大学 西井龍映

リモートセンシングの分野では工学センサーの発達によって多次元で大量のデータが同時に取得されるようになってきた。100 万以上の画素を持つ衛星画像は標準的である。また航空機に搭載するセンサーのなかには 200 次元以上のスペクトルの帯を観測

できる場合もある。このような多次元データを解析するために、各種の統計手法・多変量解析が広く用いられている。中でも判別分析は衛星データから土地利用を推定するための必須の手法である。日本リモートセンシング学会の講演会でも判別手法に関するセッションが常設されるくらいである。そこでこの分野の文献に多く引用される本書を書評の対象とした。統計専門家を対象としていない本書を書評することをお許し頂きたい。

本書は“統計的パターン認識入門”というタイトルからもわかるとおり、工学的な見地からの判別分析及び関連する統計手法を取り扱っている。パターン認識の勉強を始める人や関連分野の実務家を読者として想定している。第 1, 2 章の準備に続く各章

のタイトルと概要は以下の通りである。なお、本書では判別基準は判別子 (Classifier) と呼ばれる。

第3章 Hypothesis Testing

タイトルの“仮説検定”は観測値がどの群に属するかを決定するという意味である。母集団におけるベイズ判別法、その誤判別確率の上限を与える Chernoff や Bhattacharyya 距離を用いて分散共分散行列が異なる場合の誤判別確率を評価している。また逐次検定を用いた判別も紹介している。

第4章 Parametric Classifiers

分散共分散行列が異なる場合でも線形判別を行なう時の決定法が示されている。Mahalanobis 距離の2乗の散布図に対して、群を分類する一次関数を判別関数とする適応性のある手法を紹介している。また時系列モデルで、周波数領域での判別を提示した。

第5章 Parameter Estimation

母数を推定し (Parameter Estimation)、プラグインルールで判別基準を導出したときの判別効率への影響を考察している。2つの正規母集団の標本 Bhattacharyya 距離のバイアスを評価し、数表を作成した。また、見かけの誤判別確率の上限、下限を与え、それらを推定する方法としてリサブステューション、クロスバリデーション法、ブートストラップ法を考察した。

第6章 Nonparametric Density Estimation

分布形が未知の場合の核関数による密度関数の推定法を紹介する章である。平均自乗誤差基準を用いて、最適な一様核関数の半径を決定した。また k 個のデータを含む近傍で構成する密度推定の手法 (k NN) についても、平均自乗誤差基準により最適な k の決定を行っている。一次元の直交関数系による展開、主に Gram-Charlier 展開についてもふれている。

第7章 Nonparametric Classification and Error Estimation

前章で述べたノンパラメトリックに推定された密度によるベイズ判別手法が述べられている。また、すべてのテストデータが混ざっている空間の中に分類したいデータを置き、 k 個のテストデータが含まれる近傍を考える。 k 個のテストデータのうち最も数の多い群に分類したいデータを割り当てる方法 (voting k NN) について、2群の場合、誤判別確率を事後確率により表現した。テストデータが無限にたくさんあれば (分類したいデータの事後確率が多くのテストデータのそれで近似できれば)、誤判別確率は簡単に表現が可能であり、誤判別確率がベイズ誤判別確率の半分以上2倍以下であること、誤判別

確率が k についての単調減少性が議論されている。多群で k が1の場合の誤判別確率も求められ、同様の関係が成立することを示した。これらのノンパラメトリック手法の性能が正規分布の種々の設定の下で、またレーダーデータで比較されている。

第8章 Successive Parameter Estimation

初期のテストデータで判別基準を構成したのち、新たなテストデータにより逐次的に判別ルールを更新していく手法を扱っている。いくつかの超平面で分離可能な多群の場合は、線形判別関数の逐次更新により、漸的に分離可能であることが示されている。また非線形回帰分析における回帰曲線の零点や最大値、最小値を与える点を求めるための Robbins-Monro 法、Kiefer-Wolfowitz 法の紹介とその収束性を扱っている。

第9章 Feature Extraction and Linear Mapping for Signal Representation

全データの変数のうち、あるいは線形変換した変数の中からデータの特徴を良く表している変数の組を選ぶことを特徴抽出と呼ぶことにする。データの次元縮小のために平均自乗誤差をもちいて主成分分析を紹介している。また正規モデルのエントロピー最大化という基準からも主成分が得られることに注意している。さらに固有値、固有ベクトルのモーメントの推定量を導出している。確率過程に対しては Karhunen-Loeve 展開による主成分を構成している。

第10章 Feature Extraction and Linear Mapping for Classification

特徴抽出をすることは必ずしも判別効率を上げることに限らない。誤判別確率そのものが判別基準の良さの尺度となりうるが、特別な場合をのぞいては閉じた形での表現が得られない。そこで、多群間の分離度の尺度として、群間変動行列 S_b と群内変動行列 S_w の逆行列との積のトレースが考えられる。その関数を最大にする全変数の線形結合として、固有値が大きいものに対応する正準判別変数が得られる。2群の場合の分離度として Bhattacharyya 距離を考えると、分散共分散行列が同じなら、線形判別関数が得られる。平均が同じなら、正準判別変数のうち固有値の関数 $\log(r+1/r+2)$ の大きいものから順に選ばれた変数が特徴抽出に役立つ。群間の分離度を線形変換で不変な関数として考えた場合、一般化された群間変動行列と重み付き分散共分散の逆行列との積の固有値、固有ベクトルにより、特徴抽出が可能であることを示している。さらに、 S_b や S_w をノンパラメトリックに推定することにより、前

述の方法で特徴抽出が可能であることを示している。第2主成分までの散布図では2群に分離出来ないが、ノンパラメトリックアプローチでは分離可能となっている数値例が示されている。また全変数から分離度を大きくするよう変数の組を選び出すための変数増加法、減少法が紹介されているが、変数の数が異なるときの対処法については記述がない。

第11章 Clustering

クラスター分析と判別分析は、それぞれ教師あり分類 (supervised classification) と教師なし分類 (unsupervised classification) と呼ばれることもある。パラメトリックアプローチでは、クラスター数を固定して分離度を最大にするように変数組を選ぶための逐次的アルゴリズムを提唱している。10章で扱った分離度はクラスター数が増えると単調に増加するが、クラスター数決定のための理論はないと書いている。ノンパラメトリックアプローチではクラスター間にデータが少なくなっている谷の部分で密度関数のグラディエントを用いて推定し、グラディエントの方向にデータを移動する。最終的に全てのサンプルをいくつかのベクトルに収束させて、クラスターを形成する手法が述べられている。

本書は1972年の初版本に、主としてノンパラメトリックなアプローチに関する記述を強化することにより、200ページ以上の内容を追加したものである。各章末には計算機を用いて解答する問題、理論の演習問題、及び参考文献表が付加されるよう変更された。各章をある程度独立させたため、実務家が興味のある特定の章だけを読むことができる配慮がなされている。また多くの散布図、流れ図、グラフィック、表を用いて、直感的理解が得られるよう心がけていることも本書の特徴であろう。3種類の2群正規母集団、および2台の自動車をレーダーで観測したデータセット (観測次元66) が本書を通じて用いられている。

実は本書では Mahalanobis 距離、主成分分析、正準判別、Gram-Charlier 展開等の専門用語自身は用いられてはいない。もし、本書を勉強した人が統計家の書いた判別分析の本を読もうとすると、あるいはその逆の場合もキーワードの違いで違和感を感じるのではなかろうか。また特徴抽出のための変数選択やクラスター数の決定のために AIC を代表とする情報量基準の応用の記述が含まれるべきである。変数が多くなることが必ずしも誤判別確率を減らしはしないという著者の強い信念が読みとれるだけに、残念である。(これを誤解している実務家は少なくない)

判別分析に関する成書として McLachlan (1992) があり、すでに本学会誌 1993 年 23 巻 2 号で若木氏による書評もなされている。この本には正規理論に基づいた判別理論からリモートセンシングへの応用に渡り広く判別分析が紹介されている。(評者は判別分析を勉強したい人にはためらいなくこの本を奨めたい) McLachlan はこのようにリモートセンシングも視野に入れてはいるが、それでも “the barrier imposed by a different jargon” のためにリモートセンシング分野に対する統計家の貢献が限られたものであったと指摘している。たしかに本書にも当てはまる。しかし、用語の違いだけではない障壁、たとえばデータに対する考え方の違い、も存在するよう思う。

リモートセンシングでの判別の特徴として 1) 空間データであること、2) 複数のカテゴリーが混合しているサンプルが存在すること、3) 分類したいデータ数が巨大であること、4) 観測次元が大きい場合があることが挙げられる。2) や 4) は統計学に新しい問題を提供している。統計家が各応用分野との用語の統一に始まり、当該分野にふさわしい状況下での議論展開をすることにより、障壁を乗り越え、他の分野に貢献することができるのではないかと考える。

参考文献

McLachlan, G. J. (1992). Discriminant Analysis and Statistical Pattern Recognition. John Wiley and Sons.

多変量データ解析法

—理論と応用—

柳井晴夫著

朝倉書店 1994 年 12 月発行 x+198 頁

成蹊大学工学部経営工学科 岩崎 学

本書は、日本行動計量学会の編集による行動計量シリーズの中の一冊であり、目次は次のとおりである。

1. 多変量解析概論
2. 基本的数理
3. 多変量データの構造分析
4. 予測と判別
5. 多群の変量間の関連分析—正準相関分析
6. 質的データの数量化—数量化理論と関連手法
7. 潜在構造分析—因子分析と共分散構造分析

8. その他の手法

付録：数学的性質

参考文献

多変量解析も、コンピュータの小型化、高速化により急速に身近なものになってきた。実際の使用者は、20年前に比べると格段にその数を増しているであろう。それに伴い、多変量解析の書物の刊行も急ピッチで進んでいる。その中で本書の位置付けを探ってみたい。また、本書の副題は「理論と応用」であり、その両者がうまく融合しているかどうかも本書を評価するポイントである。

本書では、主成分分析、重回帰分析、判別分析、正準相関分析、数量化理論、因子分析といった標準的な多変量解析の手法に加え、冗長性モデル、成長モデル、多変量家族モデル、制約付き正準相関分析、共分散構造モデル、対応分析と偏対応分析、Tuckerモデル、PARAFACモデルなどが扱われている。これらの多くの手法は、詳細に議論するとなれば優に1冊の本を必要とするものばかりであり、本文200ページ足らずの中にすべてを含めることは当然のことながら不可能である。もちろん著者もそれを意図している訳ではなく、手法の紹介と応用例の提示にとどめている。

「まえがき」で述べられているように、多変量データ解析の手法の多様さおよび応用範囲の広範さを具体的に示すことが本書の目的の一つである。扱われている例が著者が実際にかかわったものであることも手伝って、その目的はある程度達せられていると評価できよう。具体的な問題を抱え、多変量解析を適用してみたいとする読者、やってみたらこれでいいのか不安であるという読者には有益であると思われる。また、巻末には287にもものぼる詳細な文献表が付けられていて、これらは「次のステップ」に進みたいと望む読者にとって大変有難い道標となる。

次に「理論」についてみてみる。多変量解析の理論とすぐに多変量正規分布に基礎を置く多変量分布論を思い浮かべるが、本書の「理論」はそうではない。著者の柳井教授は、長年「線型空間における射影」の観点から多変量解析を見つめ、種々の解析手法に対する統一的な視点を提供すると共に新しい結果を導く研究において顕著な業績を挙げられてきた人である。本書の「理論」はその流れに沿ったいわば線型代数（とその応用）である。母集団分布を想定し、そのパラメータに関する統計的な推測を行なうという古典的な推測理論でなく、本書で展開されている理論は、多変量データそのものを線型空間という枠組みのなかで捉えたいわば記述統計的

なものである。

私自身、二十年程前に本書の理論と同じアプローチの仕方によって書かれた竹内・柳井（1972）の本によって多変量解析に対する理解が深まったという経験をもつ人間であるので、本書の理論的アプローチに反対するものではない。しかし、線型代数に不慣れな読者にとっては、本書の記述の理解は並大抵のものではないとも思う。付録として16ページにわたり線型代数の解説が付けられているが、その内容は、高度である。

多変量データ解析の手法は多岐にわたり、その応用範囲は広大である。応用面についていえば、本書はその役割をある程度果たしている。しかし、理論面については、そのアプローチは魅力的であり、きちんとトレースすることによって解析手法に対する理解は深まることは事実であるが、初学者にとっては負担がかなり大きいといわざるを得ない。

本書は、理論および応用の両方に造詣が深い柳井教授であればこそ誕生した書物であるといえ、その意味で類書はない。しかし、ここに盛り込まれたトピックスをすべて含むにはページ数が足りないように思われる。

参 考 文 献

竹内 啓・柳井晴夫（1972）多変量解析の基礎—線型空間への射影による方法—。東洋経済。

An Introduction to Regression Graphics

R. D. Cook and S. Weisberg

John Wiley & Sons, 1994, xx+253頁

北海道大学 水田正弘

本書は、グラフィックスを利用した回帰分析について扱っている。回帰分析はデータ解析において基礎的であり、最も良く利用されている手法である。従って、これまで和書も含め多くの教科書が発行されてきた。その中で、本書の大きな特徴は、R-codeと名付けたソフトウェアをフロッピーにより付録として添付していることである。R-codeはXlisp-Stat上で動作するインタフェースである。

すでにご存知の方も多いと思うが、Xlisp-StatはTierney（1990）によって、Xlispをもとに開発されたデータ解析用システムである。Windows, MacOS, UNIXなど多くのOS（コンピュータ）で実行が可能であり、かつ無料で配布されている。対話的なグラフィック機能、行列演算などがすぐれており、利用価値の高いシステムである。例えば、Xlisp-Stat

の利用を前提とした統計学の教科書「統計」が竹村(1997)によって書かれている。さらに、最近、パソコン等のOSとして注目されているlinuxではXlisp-Statを含めて配布する場合が多く、統計に関心のない人にとって目に振れる機会が最も多いデータ解析用システムとなりつつある。

Xlisp-Statの欠点として、lispをベースにしているため馴染みにくい点がしばしば指摘される。R-codeを利用すれば、マウスの操作だけで、回帰分析および、その前処理・後処理を容易に実行することができる。また、「コンピュータパッケージを利用したデータ解析」に関して古くから議論されてきた事項として、「データ自身やデータ解析の知識がなくても自動的に結果が出てしまう」点があげられる。それに対して本書では、グラフィックを用いた直線(回帰直線)の当てはめ(Fitting by Eye)や次元の検出などが扱われている。

本書では、人工データや実際のデータなど多くの例題を使っているが、これらのデータは全て、R-codeで扱うことのできるR-dataとして付録のフロッピーに含まれている。コンピュータを使いながら、本書の指示に従った操作をすることにより、回帰分析の実際的な知識を修得することができる。また、練習問題は、例題の捕足や、別のデータの解析の要求であり、実際的で丁寧な問題が揃っていると思われる。

R-codeをWindows 95およびUNIX (sun OS), RedHat linuxで実行させてみたが、どれも同様な操作環境で利用することができた。ただ、私のWindowsのコンピュータ環境においては、表示に小さな問題が生じた。この原因は、ソフトとハードのどちらに原因があるのか不明である。

R-dataを含むR-codeは、本書の付録以外にも、stat.umn.eduなどからftpにより入手できる。しかし、本書のAppendix BのCopyrightsにあるように、興味深い著作権の設定がなされている。

XlispとXlisp-Statはそれぞれ、David BetzとLuke TierneyがCopyrightを保有しているが、「目的にかかわらずコピー、改変、配布、販売」(copy, modify, distribute, and sell this software and its documentation for any purpose)を許可している。これに対してR-codeは、「この本を個人で購入したのなら、購入者は任意の1台のコンピュータシステムにおいてR-codeを利用してもよい、この本を図書館で購入したのなら、その本を借りた人は借用期間中のみR-codeを利用してもよい。それ以外の団体などで購入したのなら、同時に1人だけがR-code

を利用してもよい。もし、同時に複数の人がR-codeを利用するのなら、その分だけ本書を購入しなくてはならない。」との条件がついている。

本書は、コンピュータの利用を前提とした回帰分析の教科書として評価できると思われる。

最後に、本書の各章のタイトルを列挙する。

1. Getting Started
 2. Simple Regression Plots
 3. Two-Dimensional Plots
 4. Scatterplot Matrices
 5. Three-Dimensional Plots
 6. Visualizing Linear Regression with Two Predictors
 7. Visualizing Regression without Linearity
 8. Finding Dimension
 9. Predictor Transformations
 10. Response Transformations
 11. Checking Models
 12. Assessing Predictors
 13. Influence and Outliers
 14. Confidence Regions
- Appendix A The R-code
Appendix B Copyrights

参 考 文 献

- [1] 竹村彰通(1997)「統計」, 共立講座21世紀の数学, 共立出版.
- [2] Tierney, L. (1990). *LISP-STAT: an object-oriented environment for statistical computing and dynamic graphics*, John Wiley & Sons [垂水共之, 鎌倉稔成, 林 篤裕, 奥村晴彦, 水田正弘 訳 (1996). 「LISP-STAT」, 共立出版, 東京.]

Image Analysis for the Biological Sciences

C. A. Glasbey and G. W. Horgan
John Wiley & Sons 1995 xi+218 頁

中央大学 鎌倉稔成

本書は、生物関係の画像解析を行うための入門書として、わかりやすい例題を豊富に取り入れた書として群を抜いている。例題として扱う画像データは、一部のデータ(ランドサットのデータ)を除いてはインターネットで著者のサーバーからフリーウェアとして利用することができる。したがって、プログラミングの得意な人は、本書の解説どおりに画像解

析がいくかどうかを試してみることができる。残念ながら解析するためのプログラムは公開されていない。

各章の構成は以下の通りである。

- 第1章 序章
 - 第2章 画像表示法
 - 第3章 フィルター
 - 第4章 セグメンテーション
 - 第5章 数学的形態学
 - 第6章 測定
 - 第7章 結果の要約
- 付録

第1章では、画像解析を行うのになぜコンピュータを使用するのか、画像解析の対象となるデータは何か、画像解析ではどんなことをするのかといった内容がわかりやすく、入門的に解説されている。特に、シュルツ作のピーナッツの漫画を引用している件は他書に例をみない導入である。スヌーピーの仲間として登場する、大の勉強嫌いである、サリー・ブラウンが2000語の作文という宿題を課され、“One picture is worth a thousand words”ということわざを思いだし、2つの簡単なイラストを描いて宿題を済ませているのである。また、この本で例題として扱うデータを詳しく解説しているので、興味の持てるものかどうか判断することができる。

第2章では、画像の表示法について解説している。2値表示、グレースケール表示、カラー表示、画像の拡大と縮小、レジストレーションについて述べている。統計との関わり合いでは、レジストレーションは同一の対象に対する複数の画像の位置合わせの問題であり、画像のアフィン変換係数を回帰係数の推定で行うことが述べられている。

第3章では、各種フィルターについて解説している。フィルターは最も統計の分野に関係している部分でもある。空間領域における線形フィルターである移動平均フィルター、ガウシアンフィルターはウィンドウサイズが 3×3 , 5×5 , 9×9 の場合についてX線画像について解析例が示されている。また、エッジを強調するフィルターとして1次微分フィルターと2次微分であるラプラシアンが紹介されている。フーリエ変換を利用した周波数領域の線形フィルターについても例証している。また、画像にぼけ(blur)が入った場合の画像復元の問題についてもウィナーフィルターを紹介し、SAR画像、DNA画像についてこのフィルターを適用した例を示している。さらに、エッジ検出のために7つのフィルターを紹介している。

第4章では画像をカテゴリーや領域に分割するためのセグメンテーションの技法が扱われている。セグメンテーションは画像解析の1つの解析目標の1つでもあり、統計的にも重要な問題である。最も簡単な方法はピクセル値に閾値を設定してこの値を越えているかどうかで分割する閾値法である。自動化の閾値の設定法として適当な閾値 t をとり、その上下に分割されたピクセルの平均を2つ計算し、その2つの平均の平均を新たに閾値として収束するまで計算を繰り返すアルゴリズムが紹介されている。さらに、ピクセル値に2つの正規分布を仮定して議論を行っている。これは、混合モデルの判別の問題であり統計的にも興味深いところである。さらに、多変量の画像データについてもマハラノビスの距離を利用したセグメンテーションおよびクラスタリングの技法が述べられている。また、エッジによるセグメンテーション、領域によるセグメンテーションのアルゴリズムも各種紹介されているが、本書のみですべてを理解するには解説が十分とはいえない。

第5章では数学的形態学について集合論の立場で述べられている。まず、収縮(erosion)、膨張(dilation)、オープニング、クロージングといった形態学上の基本的概念が導入され、画像処理に適用した例が紹介されている。また、接続性に関しての形態学的な技法、グレースケールのデータの形態学的扱いについても記述されている。

第6章では画像から量的な情報をどのようにに引き出すかという、画像解析の量的な目標の1つである測定の問題を扱っている。ここでいう測定は実にさまざまであり、対象物の数、特定領域の面積、対象物間の長さ、フェレ径、形態学的なコンパクトネス、接続性、丸み、伸び、また境界を記述する各種量である。詳しくは本書を参照されたい。

第7章はこれまでのまとめ章であり、これまでの章で説明された方法によって解析された例題のデータの結果がまとめられている。それぞれの例題の問題意識を再確認することもできる。

画像解析の本は工学系を中心に数多く見受けられるが、画像解析の用語から始まり実際の画像を扱いながら読んでいける本はそうはないだろう。フィルタリング等の解説の数学的扱いは必ずしも読者の期待にはそえないかもしれないが、引用はきちんとしているので詳しくはそれぞれの論文にあたることとして方法論の紹介程度に考えて一読を進めたい。

消費者行動と日本の資産市場

羽森茂之著

東洋経済新報社, 1996年

名古屋市立大学経済研究所 下野恵子

この本の主要な目的は、「消費に基づく資本資産価格モデル (Consumption based Capital Asset Pricing Model; 以下では C-CAPM を用いる)」の考え方を説明し、1970年、1980年代の日本の資産市場が C-CAPM が当てはまるような合理的な市場であったか否かを検証することにある。さらに、このモデルから導かれるオイラー方程式を推定する方法として Hansen が開発した一般化積率法 (Generalised Method of Moments; 以下では GMM を用いる) についても第3章の補論で詳しく紹介されており、実証研究者には有用な部分である。その際、GMM の訳は一般化積率法でよいが、計量経済学の方法論からは、Hansen and Singleton [1982] の論文の題名 (Generalized Instrumental Variables Estimation) のように、操作変数法の一般化として紹介した方が理解しやすいかもしれない。GMM は複雑な尤度関数を導かなくてはならない最尤法に比べると、格段に簡単な推定法であり、非線形の2段階最小二乗法や3段階最小二乗法に代わる推定法として広く用いられ始めている。

本書は以下の7章からなっている。

- 第1章 資産市場
- 第2章 消費に基づく資産資本価格モデル
- 第3章 C-CAPM と日本の資産市場
- 第4章 資産市場の変化と選好パラメーターの安定性
- 第5章 Kreps-Porteus 型選好と資産市場
- 第6章 Kreps-Porteus 型選好モデルと資本の資産市場
- 第7章 日米間における選好パラメーターの異質性について

第1章では、資産市場の分析に必要な危険回避度、リスク、リスクプレミアムの説明がなされ、次に資産価格の決定モデルが示され、最後に日本の資産市場に関するバブル論争が紹介されている。その際、もし日本の資産市場に合理的な消費者行動を前提とした C-CAPM が当てはまるならば、1980年代後半の資産価格の高騰は合理的バブルによるものと考えられる。第2章は C-CAPM の詳細な紹介である。第3章では、1970年、1980年代のデータを用いて、日本の資産市場に C-CAPM が当てはまるか否かの計

測を行っている。そして、(1)日本のデータでは C-CAPM を否定できない、(2)推定パラメータは符号条件を満たし、統計的にも有意である、という重要な結果を得ている。第4章は、GMM によって得られるパラメータの安定性、および、モデルの妥当性の検定を行う J 統計量とプレディクティブテストの性質をモンテカルロ実験により明らかにしている。第5章では、通常の消費関数では同じものとなる異時点間の代替の弾力性と相対的危険回避度の逆数を別のパラメータとして推定する方法 (Kreps-Porteus 型効用関数) を説明し、第6章でその推定を行っている。推定結果は良好である。第7章では、日米の資産市場を同時に推定して日米間の選好パラメータの比較を行い、通常考えられるのとは逆に、アメリカの方が相対的危険回避度が高いという結果を得ている。なお、すべての推定が GMM を用いて行われている。

本書の記述は平明で読みやすい。また、C-CAPM の妥当性を日本の資産市場でチェックするという意欲的な試みをまとめた形で出版されたことは、ファイナンス、マクロ経済学分野への大きな貢献である。私も本書を読むことによって、C-CAPM を正確に理解し、さらに GMM についても新たな知識を得る機会を持てた。また、以下に述べるような疑問点があるものの、著者の立場がはっきりしており、論理の流れが明せきなので、専門書であるものの面白く読むことができた。個人的には、第3章の推計、および、第4章のモンテカルロ実験の結果をもっと興味深く読んだ。

ここで、本書を読んだ後の疑問点を3点述べたい。

第1に、「合理的な消費者行動から導かれる C-CAPM が日本のデータを用いた推定結果によって否定できないので、日本の資産市場は合理的である」という著者の含意は強すぎるように思われる。その理由は、本書のメインとなる第3章の推定結果についての疑問のためである。

第3章は、時間に関して分離可能な効用関数を用いて、消費者の効用最大化行動から導かれるオイラー方程式を推定している。アメリカのデータを用いた場合、Hansen and Singleton [1982] 以来数多くの実証研究が行われたにもかかわらず、過剰識別条件によりモデルが否定されるか、あるいは、パラメータの符号条件が正しくないため、ほとんどのケースで C-CAPM は否定されている。これに対し、著者は、日本のデータを用いた場合、時間に関して分離可能な効用関数を用いた場合でも、C-CAPM を否定できないばかりか、相対的危険回避度、時間割引率

という消費者の選好パラメータも符号条件を満たし、統計的にも有意であるという重要な結果を提示している。

しかしながら、ここで著者が用いている消費の概念は1世帯当たりの消費支出である。時間に関して分離可能な効用関数を前提とすれば、C-CAPMの理論において、資産選択と同時決定される消費には耐久消費財は含まれない。それゆえ、アメリカの一連の実証研究は理論に沿って耐久財を除いた消費を用いて推定されている。著者もその点を認識しており(第3章の注10)、そして、第6章、第7章では耐久財を除いた消費を用いた推定を行っているにもかかわらず、第3章での推定には耐久財を含んだ消費を用いている。これまでに蓄積された多くの実証結果は消費の概念によって推定結果が有意に異なることを示しており、第3章での結果のみから、日本の資産市場でC-CAPMが当てはまり、ゆえに合理的な市場であるとは結論づけられない。もし、時間に関して分離可能な効用関数を用いるのであれば、耐久消費財をのぞいた消費で推定した方がよしい、耐久消費財を含むのであれば、時間に関して分離不可能な効用関数を用いるべきであろう。

さらにつけ加えれば、上記の理由から、従来の研究との比較を行った表3-6に関しても、まず、消費概念の違いを明確に示しておくべきであろう。また、表3-6では、Hansen and Singletonの1982年の論文からGMMによる推定結果を引用しているが、この結果は消費系列の取り扱いに誤りがあったとして、1984年のEconometrica (p. 267-268)に訂正した推定結果が報告されている。この結果を引用すべきである。

第2の疑問点は、モデルの評価と推定法に関連する。例えば、推定された相対的危険回避度の大きさは、第3章では0.15前後であり、第6章では、2.5前後となっている。どちらが妥当なのであろうか。第3章と第6章では異なる消費概念、異なる効用関数を用いているが、どちらのモデルも統計的に否定できない。つまり、モデルの妥当性を検定するJ統計量は弱い検定量であり、どちらのモデルがより妥当性を持つかをいうことができない。さらに、推定法の問題点として、GMMは操作変数法の拡張であり、一般に操作変数法では操作変数の選択によって結果が大きく異なる可能性がある。実際、第7章の推定では、操作変数の選択によって推定結果は大きく異なっている。それゆえ、GMMの推定結果のみから強い結論、例えば、第7章のように、アメリカの消費者の相対的危険回避度が日本の消費者より大きく、危

険回避的であるという結論を提示するのは妥当ではない。非常に複雑になるにしても、最尤法を用いた推定も考えた方がよいのかもしれない。

第3点は、C-CAPMの評価である。C-CAPMは、消費とポートフォリオの決定を同時に扱っており、経済学を学ぶ者にとっても理解しやすい理論である。著者は「C-CAPMが欧米において標準的な(ファイナンス)理論としての地位を確立しつつある(p. 3)」と述べているが、私の印象では必ずしもそうではない。というのは、理論的な理解のしやすさとは別に、本書でも述べられているが、Hansen and Singleton [1982]以降のアメリカのデータを用いた多くの実証研究ではいい結果が得られていない。最近でも、本書で紹介されているKreps-Porteus型の効用関数の導入や、消費における習慣形成、耐久財を考慮した局所的代替性を考えた時間に関して分離不可能な効用関数を想定した分析が行われているが、モデルのあてはまりは良くなっているものの、相対的危険回避度の推定に関してはさほど改善されていない。国際比較をした最近のいくつかの論文でもC-CAPMの妥当性に関して否定的である。さらに、Mehra and Prescott [1985]が指摘した、C-CAPMでは株式収益率と財務省証券との間の大きなリスクプレミアムの格差を説明できないという問題も解決されていない。その結果、資産価格の変動を説明しようとするファイナンス研究の分野において、C-CAPMは必ずしもメインストリームにあるとはいえない。むしろ、長い歴史をもつCAPMやAPT (Arbitrage Pricing Theory)の方が相対的にうまく資産価格の動きを説明できているように思われる(Kariya [1993]を参照)。

C-CAPMに積極的な意味を見いだすのは、ファイナンス研究者以上に、著者も含めたマクロ経済学者かもしれない。ルーカスに始まる合理的期待形成の考え方は、マクロ経済学にミクロの基礎を持ち込もうとしたものであり、C-CAPMもその流れに沿った魅力的なモデルである。

最後に、著者に対しては、本当に日本の資本市場は合理的なのか、という点についてより研究を深められることを希望する。私を含めて、アメリカで成立しないC-CAPMが日本で成立するはずがない(日本の資本市場がアメリカ以上に効率的とは考えにくい)、と思いこんでいる多くの研究者を説得するためには、本書で示された以上の実証結果が必要である。GMMによって推定されたパラメータの性質を含め、より説得的な実証結果を提出されるのを楽しみにしている。

参考文献

- Hansen, L. P. and K. J. Singleton, "Generalized Instrumental Variables Estimation of Nonlinear Expectations Models", *Econometrica*, Vol. 50, No. 5, 1269-1286, 1982.
(errata: *Econometrica*, Vol. 52, No. 1, 267-268, 1984.)
- Kariya, T., *Quantitative Methods for Portfolio Analysis*, Kluwer Academic Publishers, 1993.
- Mehra, R. and E. C. Prescott, "The Equity Premium: A Puzzle", *Journal of Monetary Economics*, Vol. 15, 145-161, 1985.

計量経済学の理論と応用

養谷千凰彦著

日本評論社, 1996年

小樽商科大学 西山 茂

実際の経済現象から様々な問題を抽出し、その発生メカニズム、解決への道筋を探ることが計量経済学の目的の一つであるとすれば、その取り組みが様々な疾患と格闘する医学と酷似してくるのは、ある意味で当然のように思える。本書は、計量経済学の最前線で広く利用されている分析手法や概念を懇切丁寧に解説した研究書であり、計量経済学を専門領域とする人達にとって座右の書となりうるとともに、経済分析の現場に立つ専門家がデータと取り組む際の便覧としても価値ある好著である。

先ず通例に則して、本書の構成を概括しておこう。本書は全10章から成り立っているが、全体を通読する際の折り返し点近くに「計量経済学の伝統的方法と混迷の70年代」という大変ユニークな一章が設けられ、科学哲学的叙述が展開されている。この章をはさみ前半では、古典的回帰モデル(第1章)、自己相関(第2章)、不均一分散(第3章)、ARCHモデル(第4章)が採りあげられ、最後に非線形モデル(第5章)のいくつかの推定技術がまとめられている。その次の第6章が上述の計量経済学の混迷を説いた箇所であり、伝統的計量経済学に対する方法論上の疑惑がなぜ生じたことになったかが語られている。この後の各章では、「適切なモデルの選択」を話題の軸としながら、LSE(London School of Economics)アプローチ(第7章)、パラメータの安定性テスト(第8章)、関数形の選択(第9章)を経て、ノンネステッド・モデル(Non-nested Model)に関する包括テスト(第10章)が解説されている。巻末には線型代数、数理統計に関連する主要な結果と証

明が数学付録として含まれている。

第1章は古典的回帰モデルの説明である。最小二乗推定量から説き起こすやり方は定石と言ってもよい。残差分散推定量の最良二次不偏推定量としての特性が、比較的細かな計算とともに説明されているのは、中々面白いと感じた部分の一つである。この箇所の他、本書を通じて、多くの演算過程が本文ないし数学注に丁寧に示されている。この辺の処理にはバランス感覚が要求されるのだが、本書の場合、適度に懇切であり著者の感覚を感じさせる。線型代数の初歩的知識が前提されるが、一夜漬けで主要な公式を頭に入れておきさえすれば、この程度は「習うより慣れよ」の格言を身をもって実践したほうがよい。最尤推定量、線型制約検定、一般化最小二乗法と続くあたりは、馴染みの食材が調理されるプロセスを見るのと似た興味がある。

第2章と第3章は古典的回帰モデルの前提が崩れるケースである。「崩れる」とはいえ前者は自己相関、後者は不均一分散を採りあげているので、例外とはいえ古典的な場合である。自己相関といえば、どうしてもDurbin-Watson比からということになるが、そのあと、高階の自己相関の検定量としてBox-Pierce, Ljung-Box統計量がラグランジュ乗数(LM)検定の観点から詳しく説明されている。推定についてはPrais-Winsten変換を軸に最小二乗原理、最尤原理の双方から説明が与えられる。自己相関の場合もそうだが不均一分散においても、対立仮説下の最小二乗推定量の分散を知ることが重要である。第3章ではまずWhiteの不均一分散一致推定量が紹介される。検定統計量としてはKoenker, Breusch-Pagan-Godfrey統計量が解説されている。検定に進む予備段階として本書のように $e \cdot \hat{Y}$ プロット、L-Rプロットを用いながら回帰診断を行うておくことは非常に有効である。

第4章では自己相関も不均一分散もないが、攪乱項の既往値を所与とした条件付き分散は均一ではなくなる場合をとりあげる。即ちARCHモデルである。攪乱項が独立かつ不均一分散と言っても最小二乗推定量は有効性を失い、非線形の最尤推定量が有効推定量となる。本書ではARCH(1), ARCH(q), GARCH(p, q)モデルの最尤推定とLM検定が詳説され、最近の展開としてARCH-Mモデル, EGARCHモデルが紹介されている。非線形モデルの推定については第5章で整理されている。本章は極めて短いですが、非線形最大化の数値計算法のいくつかが要領良くまとめられており研究者にとっては結構役に立つのではないかと感じられる。

第6章では、上述したように、1970年代にかけて著しく高まった伝統的計量経済学に対する不信と方法論上の混迷が回顧されている。先験的理論仮説を観測されたデータに基づいて検証するというコールズ委員会以来の伝統的立場がいかにして確認主義に墮し、批判を浴びることになったかが振り返られ、経済予測上の失策、ルーカス批判を契機に一層進んだ計量経済学の地盤沈下、新たな分析原理として登場したLSEアプローチまでが語られる。

第7章はLSEアプローチとは何かである。著者は、(1)一般形から特定の関数形へ、(2)エラー修正モデル、(3)代替モデルをモデルに包括する、という三点を挙げている。いずれをとっても特殊な技術ではなく先端的な理論に立脚するものでもない。寧ろ研究者として反対する者としてないような心構えに属することは、その後の計量経済学史の進展に占めているLSE接近法の重要性を思うと、驚きに値すると言ってよいかもしれない。LSEアプローチが受容された前提として、やはり70年代の混迷状況を指摘しなければならないのであろう。

第8章ではパラメータの安定性テストのいくつかを紹介される。チョウ検定、個々の係数の安定性に関するT検定、不均一分散下のウェルチ検定等、いずれも頻繁に参照される箇所であろう。チョウの予測誤差検定を含む逐次チョウテストなど、実際の場面では広く使用されているものの、それをきちんと説明している専門書は案外少ない。計量経済分析ソフトウェアで利用可能なことから普及してきているCUSUMテスト、CUSUM二乗テストも紹介されているが、CUSUM二乗テストとチョウの予測誤差検定との関係など、本書のように明示的に説明してくれたほうが親切であるし、その検出力が高いものではないことの指摘、Nyblom-Hansenテストを紹介している点など、本書の価値を高くしている所の一つだろう。本章だけではなく全体を通じて言えることだが、本文中に重要な参考文献が示されており、巻末の一覧とあいまって、本書を土台に個別の研究を深める際の便宜を高めている。おおよそ1990年代前半までの文献がカバーされているようである。

第9章、第10章はモデル選択について詳説している。第9章ではRESETテスト、ホワイトの関数形テスト、ボックス・コックス変換に関するLM検定が述べられ、第10章ではNon-nested ModelについてCox-Pesaranテスト、小標本を対象とする際のGodfrey=Pesaranの \tilde{N} テストとWテスト、Davidson=MacKinnonのJテストが詳しく説明されている。Fisher=McAleerのJAテスト、GodfreyのTテス

ト、Bera=McAleerのBMテストも簡単に紹介される。複数の検定統計量が短期間の内に相次いで提案されているということは、この分野が現在の計量経済学において豊かな研究成果を約束する領域であるという事実を反映するものである。なお、第10章はモデル選択基準としての R^2 の不十分性から議論が始められている。これは評者の感想に過ぎないが、第1章で紹介されているAICとまとめて、モデル選択基準の全般的解説を、ここで行ったほうが適切ではなかったかと思う。

さて、「混迷の70年代」のさ中に計量経済学の基本的訓練を受け、現在なお同じ学問領域で仕事をしている評者にとっては、何よりも上のような本書の基本構成そのものが大いに刺激的である。そこに著者の最大のメッセージが巧妙に隠されているのではなからうかと感じられる。

本書の前半部分からも明らかのように伝統的な計量経済学とは何かという基本的な点検を行えば、数理統計的には完璧に見えたその伝統的立場が、1970年代に入って以降、なぜ不信の対象となったのか、そのこと自体が不思議であろう。著者も振り返るように、石油危機を伴うインフレの70年代、予測面で示された様々な失策、政策立案段階における当事者能力の欠如(?)、科学哲学から寄せられた諸批判の中で、計量経済学は大きく地盤沈下した。更に、1976年の「ルーカス批判」が登場するに及んで、それまでの計量経済学信仰は雲散霧消し、どのような結論をも引き出しうる統計的黒魔術と形容されるまで状況は悪化したものである。

その頃、評者はマクロ計量モデル構築と隣接する部所で仕事をしていたが、同時代の本質的批判とは裏腹にマクロモデルの細分化と巨大化に歯止めがかかることはなかったと言ってよい。為替相場決定がフロート制に移行し、かつまた日本経済の拡大の中でそれが外生変数とは考えられないならば、世界経済モデルを構築して内生変数とすればよい... 振り返ると、確かにそれは計量経済学の失策をより大規模に反復する契機となったのだろうが、仮に、そこに「失敗」とでも形容するべき一連の結末が何がしか在ったとすれば、失敗の原因は当時の計量経済学にも経済理論にもなかったように思われる。むしろ、大艦巨砲主義にも似た巨大プロジェクト信仰や、科学技術の信頼性を等閑視した資源の投入に直接的関連性を求めるべきだろう。

失策には失策の原因がある。とはいえ、その失策の中で1970年代の計量経済学の混迷があったとすれば、計量経済学のどこかに科学として不十分な点

があったのではなからうかと自問することは決して無用のことではない。

正直なところ、評者は本書のような計量経済学のテキストとしても使用可能な高水準の理論書の真ん中に、『計量経済学の伝統的方法と混迷の70年代』という科学哲学的思弁に属する叙述がなぜ挿入されているのか、最初は不審に感じたものである。その方法論的展望をLSEアプローチの登場で締めくくる著者の立場にも深い理解を感得することができなかったことを告白したい。1970年代末から現在にかけて急展開した計量経済学全般を包括しようとするれば、より一層広範に時系列分析の成果に触れるべきであるし、また質的変数モデルとマイクロ・エコノメトリックスの急速な発展を概説する方が一層バランスのとれた選択であると思われるからである。更には言えば、同時方程式や識別問題に全く触れないというのも大胆に過ぎる構成であるように感じた。

しかしながら、何度も反復して本書を通読するにともない、本書の後半部分は著者自身の回答ではなからうかと思いついた。計量経済学を科学の一領域と考えた場合、どのような点に不信と混迷を招く一因があったか？社会科学として一層の成熟を図ろうとすれば計量経済分析はいかなるディシプリンの下で行われるべきか？このような問いかけに対する回答である。

第6章にも言及されていることだが、伝統的計量経済学では、仮説は理論的仮説のことであり、データに先立って与えられる性格を有している。その仮説が観測されたデータによって反証されるか否かがポイントとなる。ところが、実際に応用するとすると、理論に対する十分な反証を提示するのが相当に困難であることが判明する。実は、この点は当初から認識されていた。たとえば、計量経済学の古典的教科書であるKleinの“A Textbook of Econometrics”には、観測されたデータから正当な理論を取捨選択するのが実際には難しいことが述べられている。即ち、

経験によると、計量経済学において通常使用される標本データは、各種の仮説と斉合的であることを示している。受容可能な組は、推定された方程式を標本点の外にある観測値に外挿することによって、またデータの独立標本を繰り返し検討することによって、絶えずせめられていく。標本の外にうまく外挿できない方程式、あるいは独立標本によって繰返し行う推定の中でぐらつく方程式は、漸次棄却されていく。

—L. R. クライン『計量経済学』（岩波書店）宮沢光一、中村貢訳、pp. 27

外挿テストの反復と標本期間の拡大に希望を求めてはいるが、計量経済分析の現場ではクラインの述べたような方向には事態は進んでいかなかった。データは誤った理論に対する反証となるのではなく、複数の相対立する理論の確証として使われた。先験的公理から演繹される仮説的結論に対して、それ自らを支持する確証をのみ探し求めることが計量経済学であるとすれば、どのような「実証的結論」も本質的には単なる「思い込み」とさして違いがなくなるであろう。それは経験科学の名には値しない。

確証主義に墮するかどうか？端的に言えば、それは分析者の側の研究姿勢に依存すると考えられるかもしれない。ただ、著者が陰鬱な第6章をLSEアプローチの登場をもって明るい調子で閉じているのは、実は経済学の純粋理論からどれだけ有意な仮説が提供されるか、この辺りに関する疑念も反映されているような気がする。統計的仮説は確かに標本以前に確定していなければならない。これが伝統的立場である。そうであるにしても、標本情報によって検証すべき仮説が事前にどれほど在りうるかといえ、それは個々の学問領域の精緻化の度合にも依存する。経済学のようなソフトサイエンスで、データと対決させるべき純粋の仮説がどれだけ提供されるのか？LSEグループの唱える「データとの対話」に著者が今後の計量経済分析の行くべき方向を認めるのは評者も納得の行くところである。

この点に立てば、本書の後半部分がLSEアプローチの発想に沿ったモデル選択の考え方を中心に展開されていることは、全体の構想として当然であり、余分な内容は混在しないほうがいい。“General to Specific”というモデル構築原理、“Non-nested Model”に対する包括テスト、著者が紹介している様々な分析用具のいずれもが確証主義に墮した不毛の論争を回避し客観的議論を可能にする手法として有効なものである。

本書の全体を通じて豊富に提供されている具体的な例題で、著者は慎重な回帰診断とともに複数の検定統計量を駆使し、その結果のいかに応じて関数形を修正し、推定量を吟味しつつ再推計を繰り返している。その様子は実施可能な複数の検査結果を手にしなが患者から症状を聴取し正確な病名を結論として導こうとしている冷徹な医師の技を髣髴とさせる。それは、 R^2 , T , Durbin-Watsonの三種の神器と最小二乗法によって、マクロ計量経済モデルを構

築していた頃の現場の熱狂とはほど遠いが、情熱では実現しきれなかった科学者としての条件をより確実に充たしていると、評者には感じとれる。

以上のように本書は計量経済学研究の第一線に立ってきた著者の理念が色濃く反映された独自の専門書であり、文芸書ではない本書を形容する言葉としては適切ではないかもしれないが、一個の「作品」であると言うのも可能なような気がする。彫刻を学ぶのに作品が教科書とはならないように、本書を読むことによって現在の計量経済学の全般を習得することは難しいと思われる。しかしながら、経済問題に対して計量経済分析を行う意図を有している専門家、研究を志す大学院生にとって、本書がマスターピースとしての役割を演じうることは確実である。

**Behind State Company Nexus :
One and half century experience of
Japanese economic development in
a statistical mirror**

Yoshiro Matsuda and Fumiko Arita
MARUZEN (1996, xiv+352 頁)

静岡県立大学 大平純彦

この魅力的な表題を持つ本は、現在、マイクロ統計データに関する重点領域研究の研究代表者として研究を主導するとともに、統計審議会委員として現実の統計行政にも大きな役割を果たしておられる松田芳郎教授の長期統計・歴史統計の分野における研究成果を集大成したものである。その最も核心的な部分は共著者の有田富美子教授との共同研究の成果である。これまで松田教授の長期統計・歴史統計の仕事は、専門性が高く、その成果が入手しやすい形ではまとめられていなかったことなどから、一般の統計家や経済史家の間には広く知られておらず、その意義は正しく認識されていないように思われる。本書は松田教授が、これまでの仕事を単なる論文集という形ではなく、明治以降の統計制度の発達と日本経済の発展の研究に関する流れの中に位置づけ体系化するという形で再構成されたものである。本書の刊行が契機となって松田・有田両教授によって開発された「復元調査」の手法が普及し、長期統計や数量経済史の研究者によって活用されるようになることが期待される。

松田教授の長期統計・歴史統計の分野での業績の中で一般に知られているのは、統計の乏しい明治期について、それまで着目されていなかった「勸業年

報」や『徴発物件一覧表』などを整備し、それを統計として用いる道を開いたことである。特に府県統計書に対する勸業年報の優位性を主張したことは一般の経済史家にも広く知られており、現在では地方史家などが勸業年報と府県統計書と併用して分析を行う例が増えている。

こうした業績と比較すると余り知られていないが、最大の業績はデータ・リンケージの手法を初めて長期統計・歴史統計の分野に導入し、それによって新たなデータベースを開発しうることを実証したことにある。この新たな手法は「復元調査」(Reconstruction survey)と命名されている。松田教授は長年にわたって統計審議会の場などを通じてデータ・リンケージ手法の開発に取り組んでこられた(参考文献 [1])。その成果は現実の統計行政においては「統計行政の新中・長期構想」にいう「企業・事業所フレーム」の導入などによって定着しつつある。このデータ・リンケージの手法を長期統計・歴史統計の分野に導入しようというのは、あるいは松田教授にとっては自然な発想であったかもしれないが、極めて斬新な発想であると評価することができよう。山口和雄、古島敏雄などの研究に萌芽的な形態がみられたとしても、この新たな手法の開発の功績は松田教授に帰属するものである。

著者は今までに得られた結果はtentativeなものであると強調している。たしかに今後復元調査の手法を用いて解明することができる問題領域は広く、その手法の潜在的な可能性との対比でみると現在までに完成した部分は少ないかもしれない。しかし方法的な革新性という観点からすると、速水融教授による宗門人別帳を用いた徳川期の人口分析の仕事に匹敵する大きな意義を有していると考えられるにもかかわらず、速水教授の仕事が生み出したような新たな研究の潮流を生み出すには至っていない。これにはいくつかの要因が考えられる。第一は、多くの人にとってデータ・リンケージの手法自体になじみが薄いということである。第二は、わが国ではデータの利用上の制約が大きいので、マイクロ・データやロンジチュージュナル・データによる分析を手がける研究者の数が極めて限られていることである。今後マイクロ・データやロンジチュージュナル・データの分析が普及するならば、歴史的な実証研究において復元調査の手法は不可欠のものとして使われることになる。

本書は次のように序文と3つのパートから構成されている。以下ではその構成にしたがって、内容をみていくことにしよう。

Introduction: STATISTICAL SURVEY AND SOCIETY

Part I Statistical Survey System

Part II Formation of State Company 'Nexus'

Part III Database System for Quantitative Historical Economic Analysis

パートIは、統計調査の発展という統一的な観点から、明治維新後に行われた各種の調査から今日の統計システムが抱える問題点まで、わが国の統計制度の発展を限られたスペースで簡潔にまとめている。日本帝国の海外植民地における統計調査にかなりのスペースを割き、内地における統計調査との並行関係を明らかにしたことが、類書にみられない特色となっている。また戦時期の混乱による統計調査の重複の実態なども貴重な情報である。

パートIIは、本書の最も重要な部分であり、明治維新以降の日本経済の発展を統計の視点からまとめたものである。その内容は著者の研究を中心として、それ以外の部分については他の研究者の研究を紹介する形で歴史的な順番にしたがって記述されている。しかしここでの記述は教科書的な意味での包括性を意図しているようには思われない。むしろ著者自身の研究も含めて、教科書などにあまり取り上げられていない研究を重点的に取り上げているので、この部分は日本経済史、あるいは日本経済論の副読本として用いるならば視野を広げるという意味で大変有用であろう。

第1節では、まず徳川期の遺産として近代化の初期条件をなす人口や生産の状態についてまとめられている。第2、3節が著者による重要な貢献の部分である。第2節では、工場生産と会社制度の形成を扱っており、近代化の初期段階で工場制度、会社制度がどのように普及していったかについてデータ面から詳細に明らかにしている。特に、表II.2として提示されている1889、1896、1902、1908、1920年における会社数、資本金、払込資本を県別に時系列比較が可能な形でとらえた結果は、工業化の空間的な広がりを示すデータとして画期的な意義を持つもので、地方史研究などに大きなインパクトを与えらると思われる。ただデータの持つ情報量が多いこともあって、結果が直ちに読みとれるような形で提示されていないことが残念である。

第3節では、日本経済が工業化の道を疾走するさまを復元調査の手法によって開発された工場と企業とを統合した形態をとる1902、1909、1920年のデータベース、1902-1920年のロンジチュージナル・データベースを活用して分析している。これによって工

場数の変化、工場の地域的な拡大、企業の形態の変化、企業規模の変化など様々な観点から工業化の実態を明らかにしている。現代の統計データを用いて官庁の白書などが行っているのと同様の分析を工業化の時期に適用することができるのであるから、このデータベースが持つ意義は大きい。しかしたとえば経済史家などに対してその意義を示すためには、これを用いることによってこれまでの見解に何を付け加えることができるかを提示していかなければならないであろう。このデータベースのもつ情報量の多さがかえって理解を妨げる要因になるおそれがあると思われる。

第4節では、日本帝国の形成と崩壊を扱っている。表II.20は植民地をも含めた日本帝国の人口構造について「人種」と「地域」をクロスさせて人口の変化をとらえたものである。これは日本人の海外進出の実態を数量的に示す表として最も優れたものであり、今後は標準的な教科書などで紹介されることになるだろう。そのほか家族構造、職業構造、関東大地震の被害と回復の状況、第二次世界大戦下の労働移動、戦争被害の実態などをあつかった研究を取り上げ、興味深い結果が紹介されている。最後に、戦後の経済成長の要因として、戦時中に得られた技術的な蓄積、戦時中の経験が企業経営に与えたインパクト、所得・富・教育の平等化という3つの要因を指摘している。

パートIIIでは、松田一有田による復元調査の手法とそれによるデータベースの編成法がかなりのスペースを用いて丁寧に解説されている。しかしこの記述から得られるのは、未使用ソフトのマニュアルを読んでいるような感じで、全体像をつかめたという充足感は得られなかった。これは記述の問題というよりは、このデータベースに実際にアクセスしたことがないという経験の欠如に起因するものと考えられる。本書の最後の展望の部分において開発されたデータベースをオンラインで公表する予定があることを述べている。もしそれが実現するならば、本書の刊行と並んで復元調査の手法の普及に大きなインパクトをもたらすものと期待される。

著者は、戦前の統計制度を統計調査の発展の観点から、発生期、成長期、安定期、戦争による混乱期の4つの時期に分けている。マクロデータを用いる際に生じる統計制度の拡大・充実に伴ってマクロの値が増大するというバイアスをただすという観点から著者は分析の対象として成長期を選択したと述べている。では、それ以外の時期に対しては復元調査を適用する必要はないのであろうか。著者は昭和期

についてデータが比較的信頼できるものと評価し、安定期と呼んでいる。たしかに昭和期には制度面での整備は進んだが、軍機保護法の強化とともに陸海軍から軍機工場に対して統計調査に対する非協力の指示が出され、さらに軍資源秘密保護法によって工業統計の一部が「別冊」として秘匿されるなど、統計調査に対する軍部からの統制が強化されたという一面もある。私は工業統計にみられる不自然な変動などからみて統計データの信頼性はむしろ低下している可能性があると考えている。そうして復元調査の手法を昭和期におけるデータに適用するならば、その精度を解明するのに有効性を発揮するのではないかと考えている。

最後に、本書の刊行を機に松田、有田両教授が研究者の少ない長期統計・歴史統計の分野においても、研究を主導されることを希望しておきたい。

参考文献

- [1] 工藤弘安・大屋祐雪・山田 茂・森 博美 (1993). 官庁統計制度と統計調査の現状, 日本統計学会誌, 22-3, 613~654.

非線形経済時系列分析法とその応用

著者：刈屋武昭・照井伸彦
岩波書店, 1997年

大阪府立大学 吉田あつし

経済データの時系列分析は、70年代に Box-Jenkins 法が確立され簡易な方法でモデルの同定、推定、診断が可能になったことによって急速に広まってきた。一変量の経済時系列に自己回帰移動平均 (ARMA) モデルを当てはめることから始まり、多変量経済時系列への拡張 (VARMA) が行われ、従来の回帰分析を基礎にした同時方程式モデルにかわる、経済変量間の関係を分析し予測をする統計的な方法として受け入れられてきた。しかしながら、これらの方法はモデルの線形性、プロセスのガウス性 (それらの結果としてプロセスの定常性) を前提にしており、現実の経済時系列がこれらの前提をみたしているかが80年代に入って以下の二つの経済時系列の分野から問題にされてきた。

長期マクロ経済時系列を分析した Nelson と Plosser は、彼らの検討した経済時系列のうちの多くがタイムトレンド回りで定常というよりも、階差をとった系列が定常である (単位根が存在する) という結論を得た。もう一つは、ファイナンスの分野で扱う証券の収益率のデータについてである。

Engle は、収益率の自己相関はあまり高くないが収益率の二乗の自己相関が高いことに着目して、収益率の条件付き分散に自己回帰構造を導入したモデル (ARCH モデル) を提案した。彼らの研究以降、経済時系列の非線形性や定常性に対する関心が高まり、この分野に対する多くの研究がなされてきた。

本書で扱っているのは、主に以下の二つの点である。第一の点は、ガウス性検定である。経済データのガウス性は経済理論モデルの前提にされていたり、パラメータの推定や検定の前提になっていたりする。Markowitz の平均一分散モデルや Mossin の CAPM は、効用関数が二つのパラメータで記述できると仮定している点で、明示的ではないが収益率のガウス性を仮定している。また、上記の ARCH モデルは、収益率の条件付き分布のガウス性を仮定し、そこから最尤法を用いてパラメータの推定値を得ている。また、マクロ経済時系列を扱う場合でも、Johansen の共和分の検定は誤差項のガウス性を仮定している。本書では、Kariya, Tsay, Terui and Li によって提案された新しい検定統計量を紹介し (KTTL 検定)、それとこれまでに提案されたいくつかのガウス性の検定とをシミュレーションによって比較検討している。

第二の点は、統計モデルの線形性を検定する方法を紹介し、代表的な非線形モデルについてその統計モデル上の特性やパラメータの推定方法を明らかにすることである。モデルの線形性を検定するためには、対立仮説モデルの非線形性のクラスをある程度制約しておかないと検定の検出力が小さくなってしまふ。本書では、対立仮説の非線形モデルを多項式型の非線形モデルと閾値型の非線形モデルに分けて、それぞれについて検定方法を紹介している。ここで帰無仮説が棄却されれば、データに照らし合わせた上で、対立仮説の非線形モデルの妥当性が問題になってくる。本書で現実のデータにあてはめられて検討される非線形モデルは、主に閾値型の非線形モデルおよび ARCH タイプの非線形モデルである。

上記のガウス性の検定や非線形性の検定は、日経225の時系列データやマクロ経済時系列データに応用されて、それらの時系列にいかなるモデルがあてはめられるかについて詳細な議論がなされている。

ここで、本書の構成を概観してみよう。第2章「ガウス性の検定と線形モデル」では、時系列分析の基礎的な概念である (1) 定常性 (2) 線形性 (3) ガウス性の定義が明らかにされ、それらの概念の相互関係が示される。また、ガウス性が統計的推測に

において重要なばかりでなく、ポートフォリオ理論においても重要であることが指摘されている。第3章「非線形時系列モデル」では、ARCHタイプのモデル、双線形モデル、閾値自己回帰モデル、ExpARモデル、ランダム係数モデル、状態依存モデルが紹介されている。第4章「ガウス性検定と非線形性検定」が、本書の理論的中心部分である。ここでまずKTTL検定が紹介される。KTTL検定の基本的なアイデアは次のとおりである。

y を平均 μ 、分散 Σ の $n \times 1$ 多変量正規確率ベクトルとする。ここで、 y の標準化変量を

$$z_i = (y_i - \mu_i) / \sqrt{\sigma_{ii}} \quad (i=1, \dots, n)$$

とする。この変量 z_i の p 次エルミート多項式変換を $w_i^{(p)}$ とする。たとえば4次までのエルミート変換を考えると、

$$\begin{aligned} w_i^{(1)} &= z_i \\ w_i^{(2)} &= (z_i^2 - 1) / \sqrt{2} \\ w_i^{(3)} &= (z_i^3 - 3z_i) / \sqrt{6} \\ w_i^{(4)} &= (z_i^4 - 6z_i^2 + 3) / \sqrt{24} \end{aligned}$$

となる。この変換された確率変量 $w_i^{(p)}$ について次の様な関係が成り立つ。

$$\begin{aligned} E(w_i^{(p)}) &= 0, \\ \phi_{ij}^{(p,q)} &\equiv \text{Cov}(w_i^{(p)}, w_j^{(q)}) = \begin{cases} \phi_{ij}^{(p)} & \text{if } p=q \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

ここで、 $\phi_{ij} = \text{Cov}(w_i^{(1)}, w_j^{(1)})$ である。この関係の逆も言えるので、そのことを用いて適当な大きさの p について、 y が多変量正規確率ベクトルであるという帰無仮説の下でのみ一致性を持つ $\phi_{ij}^{(p,q)}$ の推定量、すなわち $\hat{\phi}_{ij}^{(p,q)}$ または 0 ($p=q$)と、対立仮説の下でも一致性を持つ $\hat{\phi}_{ij}^{(p,q)}$ の推定量とを比較すれば、検定統計量を作ることが可能である。対立仮説の下でも一致性を持つ推定量は、エルミート変換後の変量の標本共分散を用いればよい。これまでの話は多変量確率ベクトルの正規性の検定の話であったが、これを一変量の時系列モデルに拡張することも可能である。

モンテカルロシミュレーションにより、多変量正規性の検定についてはマーディアの検定法と比較し、一変量時系列の正規性の検定については、マーディアの検定、パイスペクトラム検定と比較される。いずれの場合についても、KTTL検定がサイズの点からも検出力の点からも他の方法よりも優れていることが示されている。さらにこの章では、いくつかの非線形性検定も説明されている。対立仮説が、多

項式型、双線形型、閾値型非線形モデルかによって検定方法が変わってくる。

第5章「日本の株価収益率」では日経225を構成する銘柄のうち214銘柄の日次および週次データを取り上げ、KTTL検定を適用して収益率のガウス性を議論している。同時に、前章で紹介した非線形性の検定を同じデータに適用して、収益率プロセスの非線形性を考察している。その結果、KTTL検定ではほとんどの銘柄に対してガウス性が棄却されている。また、非線形性の検定でも同様に線形性が棄却されている。さらにいくつかの銘柄については、どのような非線形型の可能性があるかを考察している。

第6章「マクロ経済データ」では、第5章と同様のことを(1) Nelson and Plosserの長期マクロ時系列、(2) OECD諸国のGNP、失業率、消費、マネーサプライ、利子率、(3)為替レート、について行っている。(1)についてはガウス性が否定され、(2)についてはガウス性と矛盾しない結果が得られ、(3)については観測間隔の取り方によって結論は異なってくる。

第7章「非線形モデルの実証」では、株価収益率および為替レート変化率の日次系列についてARCH(GARCH)モデルのあてはめを行い、またOECD諸国のマクロ時系列および日本の株価収益率に対して閾値自己回帰モデルのあてはめ可能性を検討している。

以上概観してきたように、本書はガウス性と非線形性についての検定方法を提供して、それを現実のマクロ時系列、金融時系列に適用してそれらがガウス性をみたまいかいなか、線形性をみたまいかについて詳細な検討を行ったという点は特筆されるべきである。類書としてはGranger and Teräsvirta(1993)がある。こちらはLong Memory Modelについても言及しており、非線形性の検定についてもやや詳しく説明はされているものの、実際の経済データへの応用という点から見ると、例題という位置づけしかなくおらず、本書に比べて見劣りがする。統計理論的にも、ガウス性に対する問題意識は小さい。本書は、経済時系列のガウス性と非線形性を理論的・実証的に包括的に取り扱ったベンチマーク的な研究といえよう。

最後に、蛇足になるかもしれないが、経済時系列データを扱う経済学者・計量経済学者の多くがARCHタイプのモデルを除いて、非線形時系列モデルにあまり関心を示してこなかったのは何故かについて少し考えて見る必要がある。例えば、経済学者

向けに書かれた比較的新しい時系列分析の教科書 (Hamilton (1994)) では、ARCH モデルには1章が割かれてはいるが、その他の非線形モデルはわずかに最後の章で紹介されているだけである。一方、単位根や共和分に関する記述は、本文の3分の1程度になる。もちろん、非線形モデルの多くが70年代の後半に提案されてきたモデルであるということも理由の一つではあるが、同じ時期かそれより後に現れた単位根と共和分は、80年代後半から今にいたるまで計量経済学を席卷してきたことも事実である。マクロ経済時系列を用いた実証研究では、必ず単位根の存在について言及しなければならない、というのが現実である。私見になるが、従来の経済学は均衡の存在とその安定性を中心概念にしており、たとえ複数均衡の存在を認めたととしても、安定的な均衡は唯一であり、必ず定常状態に収束するという信念を多くの経済学者・計量経済学者は持っている。従って、長期的な均衡関係をあらわしている誤差修正モデルとしての表現が可能な共和分モデルは経済学者に受け入れられるが、非線形時系列モデルは経済学者の間にそれほど普及しなかったと思われる。

しかしながら、近年の「複雑系」経済学が経済学者一般に受け入れられるようになれば、上記のような信念は変わるかもしれない。また、非線形時系列モデルによる実証研究の積み重ねが、この動きを強めることになるかもしれない。本書はそこまで意図されていないが、非線形時系列モデルに対する経済学者・計量経済学者の理解を深める助けとなることは間違いない。

参 考 文 献

- C. W. J. Granger and T. Teräsvirta (1993) : *Modeling Nonlinear Economic Relationships*, Oxford University Press, Oxford.
 J. D. Hamilton (1994) : *Time Series Analysis*, Princeton University Press, New Jersey.

Analysis of Longitudinal Data

Peter J. Diggle, Kung-Yee Liang,
and Scott L. Zeger.

Oxford Statistical Science Series 13,
Oxford University Press, 1994, xi+253 頁

東京大学大学院医学系研究科
健康科学・看護学専攻 生物統計学 松山 裕

健康科学 (health science) の分野に限らず、多くの研究分野において、研究対象者毎に時間を追って

繰り返し反応変数 (結果変数) を測定することがある。このような研究は経時観察研究 (longitudinal study) とか反復測定研究 (repeated measures study) などと呼ばれる。同一対象者に対して繰り返し測定される反応変数には相関があるので、統計的に妥当な推論を行うためにはこの測定単位に起因する相関を考慮しなければならない。近年、そのような反応変数間の相関を考慮したモデルに基づいた解析手法が理論的にも、また現実のデータへの応用に関しても、1986年に本書の著者でもある Liang と Zeger が Generalized Estimating Equations (GEE) を提案して以来 (Biometrika, 73: 13-22, 1986)、注目を浴びてきている。連続的な反応変数の場合には分散分析流の解析手法がこれまで利用されてきたが、医学研究でよくみられる二値データ、カウントデータのような離散的な反応変数の場合には、話が複雑になる。さらに、反応変数間の相関の問題だけでなく、対象者の観察途中での脱落 (drop-out) に伴う欠損データ (missing data) の取り扱いという問題も経時観察データ解析には存在する。

本書はこのような経時観察データに対する解析手法として、1. 周辺モデル (marginal models) 2. 変量効果モデル (random effects models) 3. 推移モデル (transition models) の3つのアプローチを紹介している。その内容は、理論的側面よりも生物学、あるいは健康科学の分野での現実の応用例を中心に議論されている。これは、著者たちも述べているように、現実のデータ解析に携わる者で数理的側面を十分に理解していない読者をも対象にしているためと思われる。また、特定の統計パッケージに関する話題は議論されていないが、生データがテキストに記載されている。

本書は11章と理論的側面を記述した Appendix からなっており、構成は以下の通りである。

1. Introduction
 2. Design considerations
 3. Exploring longitudinal data
 4. General linear models for longitudinal data
 5. Parametric models for covariance structure
 6. Analysis of variance models
 7. Generalized linear models for longitudinal data
 8. Marginal models
 9. Random effects models
 10. Transition models
 11. Missing values in longitudinal data
- Appendix (Statistical background)

著者たちが序文で述べているように、上記の11章を大きく分けると次の3つの部分に分かれる。

1. 基本的内容の概説 (1-3章)
2. 連続的な反応変数に対する一般線型モデル (general linear model) に基づく推論 (4-6, 11章)
3. 一般化線型モデル (generalized linear model) を拡張することによる離散型反応変数に対する解析手法 (7-10章)

基本的内容の概説 (1-3章)

1章は本書で例として用いられる現実のデータ (6例) の内容についての紹介である。1つ目の例は、HIV (Human Immune Deficiency Virus) に感染した患者のCD4+セル数の感染時からの経時変化のデータである。このデータは実際に行われた多施設AIDSコホート研究からのデータである。2つ目の例はインドネシアの子供の健康状態を経時的に調べた疫学研究のデータである。この研究では約3000人の子供を対象に呼吸器疾患、感染症、ビタミンA不足から生じる眼球乾燥症の有無を年4回、7年間調べている。3つ目の例は、樹木の生長に対するオゾン汚染の影響を調べる実験研究である。オゾン曝露を与えた樹木と与えない樹木の間で成長率が異なるかどうかに関心事である。4つ目の例は異なる食事を与えて育てた79頭の牛のミルクに含まれるタンパク質の量を経時的に測定したデータである。出産のため観察途中で打ち切り (censoring) を受ける牛をどのように扱うかがこの研究では問題である。5つ目の例は薬剤の臨床試験におけるクロス・オーバー試験 (cross-over trial) の解析である。最後の6つ目の例はてんかん発作に対する薬剤の臨床試験の例である。この例における反応変数は発作回数というカウントデータであり、個体差を考慮した解析が望まれる。

この章の最後に、本書で紹介する経時観察データに対する3つの解析アプローチ (1. marginal models 2. random effects models 3. transition models) を概説している。これらのモデルの詳細はそれぞれ8, 9, 10章で紹介されている。

2章は研究デザインに関する内容を扱っている。断面研究 (cross-sectional study) に対する経時観察研究の長所を議論し、経時観察研究におけるサンプルサイズ設計について議論している。

3章はデータのグラフ表示、平滑化法 (smoothing method)、自己相関関数の推定、バリオグラム (variograms) といった経時観察研究における探索的データ解析 (exploratory data analysis) を扱っている。

連続的な反応変数に対する一般線型モデルに基づく推論 (4-6章)

4章と5章では、連続的な反応変数に対する一般線型モデルに基づく推論を扱っている。反応変数のパラッキに対する平均的な挙動を「回帰 (regression)」により定式化することに加えて、3つの質的に異なるランダム誤差の原因を特定化する (1. 変量効果, 2. 経時相関, 3. 測定誤差)。このような相関構造に対しては、測定間隔が不規則で観察時間が長いデータと測定間隔は規則的で観察時間は短いデータを比較しながら、パラメトリック、およびノンパラメトリックな構造の双方について議論されている。

回帰パラメータに対する未知パラメータ推定は、重み付き最小二乗法と正規性を前提とした最尤法が議論されており、相関構造に対するパラメータ推定は、最尤法、および制限付き最尤法 (restricted maximum likelihood: REML) が議論されている。

個人内の反応変数間の相関構造が正しく特定されていれば、推定された回帰パラメータの漸近分散はフィッシャー情報量の逆数で推定できる。しかしながら、実際には相関構造を正しく特定することは難しく、その誤特定 (misspecification) が起きると考えられる。この場合には通常の一貫分散の一致性は保証されず、回帰パラメータに対する妥当な推論が行えない。そこで、ロバスト分散 (サンドイッチ分散と呼ばれることもある) を用いることが提案されている。回帰パラメータの推定に第一義的な目標があり、かつ相関構造についての情報が相対的に少ない状況では、相関構造のモデル化を誤った場合でも回帰パラメータの推定にバイアスは入らず、正しい信頼区間 (分散) を与えてくれる方法が望ましい。ロバスト分散はまさにこの要望に応えてくれるものである。

5章の後半は、以下の3つの話題を扱っている。

1. 当てはめたモデルの当てはまりの良さ、回帰診断、2. 平均的な反応に対するノンパラメトリックなモデリング、3. 個人の反応プロファイルに対する経路ベイズ予測

6章は古典的な分散分析モデルの話題を扱っている。

一般化線型モデルを拡張することによる離散型反応変数に対する解析手法 (7-10章)

7章から10章では、二値データ、カウントデータなどの離散型経時観察データに対する解析手法を扱っている。これらは全て反応変数が独立である場合の一般化線型モデルを拡張する形で議論されている。この分野はここ10年間で急速に発展してきた分

野である。

離散型データに対する非線型モデルが、4章、5章で述べられた連続型データに対する線型モデルと異なる点は、相関構造に対する仮定が異なれば、解釈の異なる回帰パラメータを導くという点である。本書では、*marginal models*, *random effects models*, *transition (Markov) models* の3つのアプローチが、二値データ、カウントデータのそれぞれに対して議論されている。

各モデルに対する未知パラメータ推定方法も異なる。周辺モデルに対しては、一般化推定方程式 (*generalized estimating equations*) により周辺平均に関する回帰パラメータを推定する。このモデルは、*Generalized Estimating Equations (GEE)* が経時観察データ解析の分野で提案されて以来、理論的にも、また現実のデータへの応用に関しても注目を浴びてきている。統計ソフト SAS (Release. 6.12) の GENMOD プロシジャではこのモデルを扱うことができる。変量効果モデルに対しては、条件付き尤度、あるいは周辺尤度に対して、EM アルゴリズム、あるいはマルコフ連鎖モンテカルロ法 (*Markov chain monte carlo method: MCMC*) を用いるアプローチが議論されている。変量効果モデルの周辺尤度に対するパラメータ推定は統計ソフト SAS のマクロ GLIMMIX により実行でき、MCMC 法による事後分布の計算は BUGS (*Bayesian inference using Gibbs sampling*) により行うことが可能である。推移モデルに対しては、近似条件付き尤度を考えるが、これは反復重み付き最小二乗法に帰着することが議論されている。

最後の11章は、対象者の観察途中での脱落 (*dropout*) に伴う欠損データ (*missing data*) の問題を扱っている。完全ランダム欠損 (*completely random missing*)、ランダム欠損 (*random missing*)、情報のある欠損 (*informative missing*) という3つの欠損メカニズムについて議論されているが、さらなる研究が、特に、離散型反応変数に対する非線型モデルにおいては必要であると結論づけられている。

以上が各章の概略である。1986年に本書の著者でもある Liang と Zeger が GEE を提案して以来、海外の統計関連雑誌の論文発表や学会での発表演題のタイトルを見ても、経時観察データ解析に関する話題が多く見受けられる。頭の痛い問題である経時観察データ解析のために提案されたこともあり、医学領域の統計解析に携わる者にとってこのような手法は大変興味深いものであることも注目を浴びている理由であろう。また、本書で紹介されている解析手

法は、繰り返し測定に限らず、施設を単位としてクラスター化されたデータ、毒性実験での同腹仔データなど、相関のあるデータ解析一般に応用できる方法論でもある。本書はこのような広い応用分野に対する解析手法を見通しよく数理的側面はなるべく避け、応用例を中心にコンパクトにまとめており、理論的側面に興味をもっている者のみならず、現実のデータ解析に携わる者で数理的側面を十分に理解していない者にとっても非常に貴重なものといえよう。

Health Measurement Scales: A Practical Guide to Their Development and Use Second Edition

David L. Steiner & Geoffrey R. Norman
Oxford University Press
(1995; viii + 231 頁)

東京大学大学院医学系研究科
健康科学・看護学専攻 橋本修二

人の健康問題を考えるとき、多種多様な項目が必要となり、それらの測定や調査が行われる。たとえば、生死、血清中の総コレステロール濃度、喫煙状況などである。近年、精神・心理的な側面が注目されつつあり、とくに、生活の質 (*quality of life; QOL*) は、疫学研究、臨床研究などの広い分野で重要視されている。QOL の概念を厳格に規定することは難しいが、生活の満足度などの精神・心理的な内容を含むものである。

QOL などの精神・心理的な概念について、疾患や研究目的などによって様々な測定方法 (QOL 尺度など) が提案されている。また、それら測定方法を用いて、データが収集され、多くの研究に利用されている。しかし、その測定方法自体はそれほど十分に検討されているわけではなく、現在、その点が大きな問題といえる。その主な原因として、QOL などの精神・心理的な概念の測定にはきわめて大きな困難を伴うことが挙げられるが、測定方法論に関する成書がきわめて少なかったことも関係が大きいと思われる。

本書は、健康に関連する概念 (QOL などの精神・心理的なものを含む) の測定方法について、精神・心理学的な側面、数理統計学的側面なども含め、総合的に議論したものである。本書は14の章と2つの Appendix からなっており、構成は以下の通りである。

1. Introduction
2. Basic concepts
3. Devising the items
4. Scaling responses
5. Selecting the items
6. Biases in responding
7. From items to scales
8. Reliability
9. Generalizability theory
10. Validity
11. Measuring changes
12. Item response theory
13. Methods of administration
14. Ethical considerations

Appendix A. Further reading

Appendix B. Where to find tests

第1章では、様々な研究分野における測定的重要性を強調するとともに、測定方法に関する過去の研究の流れを概説している。とくに、QOLなどでは主観的評価が本質的であることから、教育学や心理学の研究成果の健康関連分野への応用の必要性にも言及している。本書の対象は主として健康科学などの研究者であり、そのねらいが測定に関わる概念や方法の提示にあるとされている。また、本書の特徴として、最新性と包括性が挙げられている。

第2章では、研究に適用するスケール (scale) の検索から、その評価までを扱っている。文献調査、批判的評価 (critical review)、内容的妥当性 (content validity)、信頼性 (reliability) などを簡単にまとめている。しかし、一方では、研究目的に合致し、かつ、評価基準を満たすスケールが現実にはほとんどないことを指摘し、新たなスケールの開発の必要性を挙げている。

第3章から第7章までは、新たなスケールの開発方法を扱っている。一般に、スケール開発の流れとしては、項目 (item) の収集、回答肢 (response) の選定、項目の選定、スケールの作成となる。本書もその流れに沿って、章の構成がなされている。

第3章では項目の収集と内容的妥当性を扱っている。項目の収集方法については、項目のソースとその特徴を整理している。内容的妥当性の定義を示すとともに、具体的な事例を用いて議論を進めている。第4章では様々な回答肢を挙げ、その特徴を示している。とくに、visual analogue scaleのような標準的な方法から、Thurstone や Guttman などの特殊な方法までも網羅的に取り上げている。第5章では項目の選定を扱っている。解釈面からプリテストによ

る頻度面までも含めて、項目選定の基準を提示するとともに、項目の均質性の評価方法を解説している。第6章では回答の偏り (bias) を扱っている。回答の偏りの原因について、心理学的側面から整理し、その1つ1つについて対応までも言及している。第7章ではスケールの作成を扱っている。とくに、項目の重みの選定、スコアの変換、カットポイントの選定について、統計的側面を含めて詳細に論じている。いうまでもないことながら、スケール開発は形式的に行っても適切なものができるわけではなく、また、それほど安易に行うものでもない。本書の各章でも、様々な留意点が挙げられており、参考にすべきものが多い。

第8章から第12章までは、スケールの評価方法とその基礎理論を扱っている。スケール評価の主な視点は、信頼性と妥当性であり、それ以外に変化に対する感度などがある。簡単にいえば、信頼性とは測定誤差が小さいことを意味し、妥当性とはスケールが測定したい概念を表していることをいう。感度とは、測定したい概念の状況変化にスケールが鋭敏に対応して変化することである。

第8章では信頼性を扱っている。信頼性係数の定義と哲学的解釈を与えるとともに、分散分析を用いた具体的な信頼性係数の推定方法を提示している。また、信頼性係数値の解釈に関わる問題を考察し、信頼性の向上を図る方策を議論するとともに、信頼性研究のためのサンプルサイズ設計方法を提示している。第9章の Generalizability theory は、Cronbach らにより提案されたものである。以前の理論は測定値が真値と誤差の和という枠組みの中で議論していたのに対し、この理論では測定上の変動要因を実際に探っていくところに特徴がある。ここでは、理論の詳細というよりも、いくつかの事例を挙げて具体的な適用方法を示している。評価者内、評価者間、測定日や場所などの影響を評価するための調査デザインや推定方法を挙げている。

第10章では妥当性を扱っている。妥当性は内容的妥当性、基準関連妥当性 (criterion validity) と構成概念妥当性 (construct validity) の3つに大別される。まず、内容的妥当性を定義 (項目の内容からみた妥当性) し、その評価を考察している。次に、基準関連妥当性を定義し、その具体的な評価方法を与えている。基準関連妥当性とは標準的スケール (gold standard) がある場合に、それとの一致性を指すものである。また、構成概念妥当性とは構成概念から見た妥当性を指すが、その理解と評価が難しいものである。ここでは、その点を様々な事例を挙げるこ

によって詳細な説明を行っている。最後に、妥当性評価における偏りに言及し、その補正方法を議論している。

第11章では変化の測定を扱っている。スケールの適用目的として、個人のスコアの変化を観察することは重要である。とくに、介入研究では、通常、介入前後のスコアの差を介入群と非介入群の間で比較される。ここでは、変化の測定方法上の問題を考察するとともに、スケールの信頼性と変化に対する感度についての評価方法、成長曲線 (growth curve) を用いた評価などを提示している。第12章では項目反応理論 (item response theory) を紹介している。これは、スケールの項目に対する反応と潜在的な特性 (trait) との関係を数学モデルで記述し (item characteristic curve)、データから推論するものである。ただ、スケールを構成する各項目が同じ1つの特性を表し、それらが互いに独立に回答されるといような仮定が必要であり、その是非などについても議論している。

第13章では調査法を扱っている。面接法、電話法、郵送法などを取り上げるとともに、コンピュータを利用した調査法にも触れ、その長所・短所を議論している。第14章では倫理面の議論をしている。スケールの開発上、倫理はきわめて重要な問題であることを強調している。Appendix A では、各章ごとに参考文献を挙げている。Appendix B では、過去に提案されているスケール一覧を、一般的なもの、社会・態度、パーソナリティなど別に示している。これは、大変参考になるものと思われる。

以上が各章の概略である。本書は、信頼性、妥当性などの主要な概念については詳細に記述されているが、統計学的な側面は必ずしも詳細に解説されていない。これは、統計学を十分に理解していない読者をも対象としているためと思われる。また、本書の特徴として、著者も述べているように、スケールの開発・評価に関する知見を、最新のものを含め包括的に取り上げ、それをコンパクトに整理されていることも関係している。一般に、データが重要であることはいうまでもなく、また、この点は広く理解されていると思われる。一方、データを生み出す測定方法については、たとえばデータ解析を行う際などでは、必ずしも十分に注意していない面があるように感じられる。スケールを開発する者だけでなく、スケールにより測定されたデータを解析する者にとっても、本書は貴重なものといえよう。

統計的多重比較の基礎

永田 靖・吉田道弘著

サイエンティスト社、1997年

東京大学医学部薬剤疫学教室 浜田知久馬

一昔前、医学・生物学系の論文では、統計解析といえば、ほとんどが *t* 検定ですまされていた。その後、欧米の一流のジャーナルに統計学の専門家がレフリーとして参加するようになり、論文の実験計画、統計解析について適切であるか評価するようになった。現在では、統計の専門家がチェックしたことを認めるサインをしなければ、投稿自体が認められないジャーナルすら存在し、かなり厳しい状況になっている。以前と比べて最も大きく変わった点は、検定の多重性の問題を避けるため、多重比較の手法が多く用いられるようになった点である (多時点の多重性の問題を避けるため経時的な分散分析も多く用いられるようになった)。ある人が1996年1月～3月までの3ヶ月間に発行された欧米の代表的な薬理学雑誌である European Journal of Pharmacology (196報) と The Journal of Pharmacology and Experimental Therapeutics (165報) について、どのような多重比較法が用いられているか文献調査を行った。その結果を次に示す。

	論文数
Dunnett	39報
Williams	1報
Tukey (Tukey-Kramer)	18報
Bonferroni	28報
Sidak	1報
Scheffe	14報
Fisher LSD	22報
Student-Newman-Kuels	40報
Duncan	11報

多様な多重比較法が薬理試験の現場で用いられていることがわかる。これに対して我が国では、前臨床試験でよく用いられる対照群+3ないし4用量群の実験において、対照群と各用量群の比較で、例数のバランスがとれていれば Dunnett、そうでなければ Scheffe 法を使い分けるアルゴリズムが長年用いられており、現在までのところ毒性試験用の標準的な GLP システムでも、この方法が組みこまれたままである。GLP システムが作られた当時、例数がアンバランスなときでも問題なく使える多重比較法として、Scheffe 法以外で、一般に広く知られている手法は少なかったためである。いうまでもないことで

あるが、Scheffe 法は全ての可能な対比を対象とした多重比較法であり、Dunnett 型の比較に用いると過度に保守的になる。特に毒性試験でアンバランスになるケースは、高用量で毒性による死亡が起きている場合が多く、このように毒性の発現が予想されるケースで、必要以上に保守的な方法を用いるのは明らかに好ましくなく、多重比較の誤用といえよう。別な例では薬効薬理試験の分野では、一頃 Duncan の方法が頻用されたが、この方法は第一種の過誤を名義水準以下に抑えることはできない。このように生物系の論文では、多重比較の方法が適切に用いられてない例が非常に多い。

非臨床試験の現場の実験者の間でも、適切な統計手法が用いられていないという認識が、数年前から広まっていった。にもかかわらず、前述の明らかにおかしな統計手法が長年に渡って用いられてきたのは、現場の実験者にわかるような形式で統計手法、特に多重比較法を紹介したテキストがこれまで存在しなかったのが一因である。

このような背景をふまえて、今回 永田靖・吉田道弘先生共著によって書かれた「統計的多重比較の基礎」は、現場の実験者が長年待ち焦がれていた、日本語の多重比較のテキストといえる。以下にこの本の構成を示す。

統計的多重比較の基礎

- 第1章 イントロダクション
- 第2章 多重比較法はなぜ必要か
- 第3章 多重比較法の考え方
- 第4章 基本的な手法
- 第5章 ノンパラメトリック法
- 第6章 ボンフェローニ法と関連する方法
- 第7章 ステップダウン法
- 第8章 多重比較法における検出力
- 第9章 多重比較法で用いる確率分布

この本では、前述の非臨床試験の現場でよく用いられる統計手法である Dunnett, Williams, Tukey, Bonferroni, Sidak, Scheffe, Fisher LSD, Student-Newman-Kuels などのポピュラーな多重比較法をすべてカバーしている。それぞれの手法ごとに、想

定する帰無仮説と対立仮説、計算手順、適用例を示したそのスタイルは吉村功編著「毒性・薬効試験データの統計解析」(通称ピンク本)を踏襲したものであり、いわばピンク本の多重比較版と位置づけられる。ただしこの本は、単なる多重比較のクック本ではない。著者達の数理統計学者としての最新の知識と豊富な経験に基づき、なおかつ平易な文章で、閉手順などの多重比較を理解するのに必要な基礎的概念、Peritz 法などのより性能の高い高度な多重比較法についても解説されている。

第2章では t 検定を繰り返すことの問題点を丁寧に説明し、第3章では、多重比較法の基本的な用語についての解説がなされている。特に「多重比較における帰無仮説と第1種の過誤」については、重要な概念でありながら、邦文の成書できちんと解説されたものが少なかっただけに是非一読をお薦めする。

第4章では一元配置分散分析手法についてのパラメトリック手法を、第5章では対応するノンパラメトリック手法を、読者が実際に計算できるような形式で説明している。

第6章ではボンフェローニの不等式に基づく多重比較を紹介し、第7章では逐次的な多重比較法として、高い検出力を持ちながら、一般にはあまり知られていなかった Tukey-Welch 法、Peritz 法が紹介されており、本書ならではのユニークな特徴といえる。

第8章では、多重比較法における検出力について紹介し、議論している。多重比較を行う際には第1種の過誤の制御に注意が集中しがちだが、多重比較法を選択する際には、なるべく検出力が高くなるような方法を選択する必要がある。検出力を実験前に評価しておくことは重要である。第9章では多重比較法で用いる確率分布を紹介し、正確な棄却限界値の求め方を述べている。

以上紹介したように、本書は多重比較法の基礎的事項を解説することを意図して書かれており、応用家、実務家、のみならず統計の専門家を志す学生の方にも、多重比較の教科書としての利用をお勧めする。