

## 書 評

## 目 次

Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons (1995) John Wiley & Sons	R. E. Bechhofer, T. J. Santner and D. M. Goldman 著	百武弘登
Neural Networks for Pattern Recognition (1996) Oxford University Press	C. M. Bishop 著	辻谷将明
Stochastic Geometry and its Application Second Edition (1995) John Wiley & Sons	D. Stoyan, W. S. Kendall, and J. Mecke 著	間瀬 茂
Growth Curves (1995) Marcel Dekker	A. M. Kshirsager and W. B. Smith 著	神田隆至
為替レート変動の時系列分析 (1995) 牧野書店	小島平夫著	山本 拓
Excel による統計入門 (1996) 朝倉書店	縄田和満著	松原 望
貯蓄と資産形成—家計資産のマイクロデータ分析— (1996) 岩波書店	高山 憲之・有田富美子著	寺崎康博

**Design and Analysis of Experiments for  
Statistical Selection, Screening, and  
Multiple Comparisons**

R.E. Bechhofer, T.J. Santner  
and D.M. Goldsman

John Wiley and Sons, Inc. (1995, xii + 325 頁)

九州大学数理学研究科 百武弘登

本書は選択問題における実験計画法に関するガイドブックであり、ある程度の統計学の知識を前提に書かれている。実験計画と書いたが、それほど知らなくても読めるであろう。また、ある程度の知識とは正規分布や2項分布などの簡単な分布論くらいであろうか。ここでいう選択問題とは、数個の母集団があり、たとえば最大平均をもつ母集団を最良として、それをある確率以上で選択することである。選択問題については、本書の第一著者である Bechhofer の Indifference Zone Approach (IZ) と Gupta の Subset Selection Approach (SS) による解決法がある。実際、選択問題に関する著書において両者の文献が多く引用されている。IZ は最良母集団とその他の母集団に少なくとも一定の差があることを仮定して最良を選択するものであり、理論的には各母集団からの標本数に関する議論が中心となる。したが

って、2段階法や逐次法と関連した論文が少なくない。一方、SS はその差を仮定せずに、最良母集団を含むような部分集合を選択する方法であり、部分集合の大きさに関する議論となることが多い。本書は第一著者が Bechhofer という点もあって、IZ について詳しく述べてあるようだ。選択問題に関する分野において、本書と類似した有名な著書としては、Gibbons, Olkin and Sobel (1977) と Gupta and Panchapakesan (1979) が挙げられる。この2冊の印象が強いため、本書を手にしたとき、選択問題に関する著書の割に厚みがないことに驚いた。最近では、95年の本誌で青嶋により書評がなされた Mukhopadhyay and Solanky (1994) があるが、これは応用を意識して書かれたものではなく、選択問題というより逐次解析の著書という印象である。前の2冊はともに絶版になっているから、現在のところ選択問題に関して手にできるのは本書だけではないだろうか。本書は Preface にあるが、過度に長くなることを避けるため理論面は最小限にして、統計的方法を記述するに留めている。また、扱う分布(母集団)も限られており、8章のうち第2章から第6章までが正規分布であり、分散に関する選択については省いている。特に、第2章は54ページにわたって IZ による方法が広く述べられている。大部分が数表である Appendix を除けば246ページであるから、第2章だけで2割以上を割いている。あとは第7章

で2項分布を、第8章で多項分布を扱っている。また、Appendixには数表の他にプログラムも掲載してある。ノンパラメトリックな選択については、章末のChapter Noteに参考文献が紹介されている。

本書の構成は次のようになっている。

1. The Rationale of Selection, Screening and Multiple Comparisons
2. Selecting the Best Treatment in a Single-Factor Normal Response Experiment Using the Indifference-Zone Approach
3. Selecting a Subset Containing the Best Treatment in a Normal Response Experiment
4. Multiple Comparison Approaches for Normal Response Experiments
5. Problems Involving a Standard or Control Treatment in Normal Response Experiments
6. Selection Problems in Two-Factor Normal Response Experiments
7. Selecting Best Treatments in Single-Factor Bernoulli Response Experiments
8. Selection Problems for Categorical Response Experiments

第2章は、分散が共通で既知の場合から分散不均一で未知の場合までのIZによる方法が述べられている。最初は、分散が共通で既知の場合において、 $t$ 個の母集団から最良である1個の母集団を選択するための各母集団からの標本数の決定法、頑健性、多段階法についてである。ここでの多段階法は標本の取りすぎを避けるための方法であり、主にEliminationを目的としている。さらに、最良から $s$ 個を選択する方法まで紹介しているが、これについては第3章とも関連しており、そこにも述べられている。分散既知の場合については、わりと詳しく書いてあるから選択問題について理解しやすいであろう。それから分散が未知の場合についてであるが、問題設定によっては1段階では選択法が存在しないため、その解決法として2段階法を用いている。ただ、1段階では不可能なことが参考文献を挙げているだけである。後の章にも、この事実が引用されることから、せめて証明の概略だけでも書いてあげれば残念である。ここでもEliminationを伴う多段階法も述べてあり、2段階法と多段階法の期待標本数の比較もなされている。最後は、等分散を仮定しない場合についてである。この問題については、まだ改良の可能性が残されていると思われ、あまり詳しく書か

れていない。

第3章は、SSについてである。釣合型、不釣合型、ブロック計画における選択法が述べられている。SSの大きな目的はScreeningにあると考えられる。また、SSは具体的なデータ解析より理論的な研究が多いようである。たとえば、選択される母集団の個数の期待値の評価などがあり、その紹介がなされている。この分野は、ベイズアプローチもなされているが、本書は参考文献を挙げているだけである。

第4章は、同時信頼区間について述べられている。ここでは、Contrast、すべてのペアの差、および最大平均との差の信頼区間を構成している。ペアの差については、長さ一定の信頼区間についても触れており、ここでも2段階法が紹介されている。Contrastの節は割と親切に書かれているが、ペアの差と最大との差があまりにあっさり書いてあり、例もなくものたりない。Contrastの節の一部分を削ってでも、もっと詳しく書いて欲しいものである。

第5章は、 $t$ 個の母集団にコントロール母集団を加えた $t+1$ 個の母集団から最良を選択する方法がIZとSSについて述べられている。また、コントロールとの差に対する同時信頼区間の構成をしている。ここでは片側信頼区間も与えており、コントロールより良い母集団の選択へも応用できる。

第6章は、二元配置分散分析モデルにおいて、効果が最大となる組み合わせを選択することが目的となる。ただし、交互作用は仮定していない。この章も第2章のように、分散既知の場合から未知の場合へとIZとSSが述べられているが、IZの既知の場合については長すぎるように思う。その割に未知の場合には2段階法を紹介しているだけで不親切である。特に、この章のIZについてはガイドブックとしての役割を果たしているか疑問がある。この章は思い切って削除してもいいのではないだろうか。

第7章は、いくつかのBernoulli試行の中から最大の成功確率を持つ試行の選択について述べられている。通常、平均に関するIZは差で与えられるが、そのみではなくオッズ比によるIZも与えている。数表により標本数を決定する1段階法と逐次法が与えられ、その比較がなされている。SSについては、やはり数表を用いる1段階法が述べられているだけであるが、正規近似による方法も紹介されている。同時信頼区間については述べられていない。また、この章の問題は具体例が多いと思われるが、例が少ないのも残念である。

第8章は、多項分布において最大の確率を持つカテゴリーの選択について述べられている。ここでは、

確率の比による IZ を与えている。IZ と SS の両方に対して、1 段階法と逐次法が紹介してあり、IZ ではこれらの方法の標本数の違いが述べられている。数表は欠くことのできないものであるが、ここでは本文中に 1 段階法で用いる数表が数ページを割いて掲載されているので読みづらくなっている。Appendix にまとめて欲しい。

正規分布を扱った章では、だいたい 1 段階法から 2 段階法、逐次法といった同じような流れであるから、少々退屈してくる。分散が共通でない場合については、第 2 章で簡単に紹介しているだけであり、物足りない。これに関しては、まだ応用できる段階まで達していないのかもしれないが、コメントくらいは述べてもいいように思う。ガイドブックという性質から、理論面は少ないが章や節によってはくどのような部分もある。著者との興味の違いかも知れないが、1 段階法では解決できず 2 段階法を用いる理由をもっと強調して欲しかった。また、データ解析への応用を考えた場合、分散が既知であることは稀であると思われる。そこで、分散既知に対しては、選択法を解説するために方法の記述は必要であろうが、数表をかなり削ってもよかった。そのぶん、具体例をもっと詳細に記述すれば、選択問題に興味をもつ人が増えるのではないだろうか。タイトルに多重比較と書かれているが、簡単に触れているだけであるから、そのつもりで本書を手にはすれば戸惑うであろう。多重比較と選択問題の関連をもう少し詳しく述べられていれば、おもしろくなったかもしれない。最初に、統計の知識がある人が対象と書いたが、読み終えてみると、どのような読者を想定しているかがよくわからない。統計を学んだ人にとっては物足りない気がするし、初学者にとっては取っつきにくいように思える。データ解析への応用に関心がある人が対象とするには、応用例が少ないように感じる。しかし、扱う分布を正規分布に絞ったことで、選択問題の問題設定とその解決法は一応理解できるであろうし、短時間で読めると思う。

#### 参考文献

1. Gibbons, J. D., Olkin, I., and Sobel, M. (1977). *Selecting and Ordering Populations: A New Statistical Methodology*. John Wiley & Sons.
2. Gupta, S. S. and Panchapakesan, S. (1979). *Multiple Decision Procedures*. John Wiley & Sons.
3. Mukhopadhyay, N. and Solanky, T. K. S. (1994). *Multistage Selection and Ranking Procedure*. Marcel Dekker.

## Neural Networks for Pattern Recognition

Christopher M. Bishop

Oxford University Press 1996, xiv+482 頁

大阪電気通信大学 辻谷 将明

近年、人間の脳・神経系の情報処理を模倣したニューロコンピューティングが新しい展開をみせ、パターン認識（例えば、文字・形状認識、音声認識）、画像処理、ロボット制御などの分野で広範に活用されている。1985 年以降、階層型ニューラルネットワークは、バックプロパゲーション学習則の開発を契機とし、分類問題の視座から脚光を浴びている。特に、統計的非線形判別分析やパターン認識との融合という観点から、現象の解明や情報圧縮を通じ、新たな生産的知見の獲得が期待される。統計的多変量データ解析の分野に眼を転ずれば、射影追跡、平滑化法あるいは一般化加法モデルなどの有望な方法で、非線形性に対する挑戦が続いている。

本書では、統計的パターン認識の観点から、階層型ニューラルネットワークについて概説している。パターン認識の基礎的知識を与え、確率密度関数のモデル構築法について解説し、階層型ニューラルネットワークの代表である多層パーセプトロンと RBF (Radial Basis Function) ネットワークの有効性について論じている。更に、種々様々な誤差関数の利用を推奨し、誤差関数を最小化する主要なアルゴリズムについてもふれている。また、ニューラルネットワークの学習と汎化性についても詳述している。データ加工、特徴抽出および事前情報に関する話題も取上げている。最後に、ベイズ法の広範な取扱いおよびニューラルネットワークへの応用について議論している。

本書は、次のような 10 章から構成されている。

- 第 1 章 統計的パターン認識
- 第 2 章 確率密度推定
- 第 3 章 単純ネットワーク
- 第 4 章 多層ネットワーク
- 第 5 章 RBF (Radial Basis Function)
- 第 6 章 誤差関数
- 第 7 章 パラメータ最適化アルゴリズム
- 第 8 章 事前加工と特徴抽出
- 第 9 章 学習と汎化性
- 第 10 章 ベイズ法

以下、本書の内容を簡単に紹介する。まず第 1 章

は、パターン認識に関する基礎的な考え方を述べている。多項関数の当てはめを例に取上げ、線形回帰分析と対比させて説明している。また、二変量の二群判別をグラフ表示し、線形判別、二次判別そしてニューラルネットワークによる非線形判別への流れが明瞭である。2章以降の話題であるパラメータ最適化、汎化性についてふれている。特に、分類(classification)問題と回帰(regression)問題を取上げ、統計的パターン認識、決定規準およびベイズ法についても簡単に解説している。

第2章では、データ集合の確率分布をモデル化する問題を取扱っている。多次元正規分布を基礎とした伝統的なパラメトリック法、およびノンパラメトリック法について概説するとともに、混合分布に基づくセミパラメトリック法も取上げている。パラメトリック法では、多次元正規分布が基礎となり、最尤法が利用される。そして、ベイズ推測へと進展する。ノンパラメトリック法では、データのみ依存し、関数型があらかじめ規定されていない確率密度関数を考える。そのため、ヒストグラムによる密度推定を扱う。他のノンパラメトリック推定法としては、最短距離法について述べている。混合分布に基づくセミパラメトリック法は、モデルの訓練のためにのみデータ集合を利用するため、パラメトリックやノンパラメトリック法に比べて計算量が膨大になる。特に、最尤解を得るためのEMアルゴリズムの有効性について論じている。

第3章では、単純ニューラルネットワーク(すなわち、入力層と出力層のみをもつ)を紹介している。単純ニューラルネットワークは、多層ニューラルネットワークに比べて柔軟性に欠けるものの、実践の場では重要である。特に単純パーセプトロンは、多変量データ解析と密接な関係がある。すなわち、出力層のユニットの出力関数が線形なら、単純パーセプトロンは線形重回帰モデルになる。出力関数が、ロジスティック関数なら、単純パーセプトロンはロジスティック回帰モデルになる。第2、3章では、確率密度の推定法を取上げたが、密度推定を回避する方法として、判別関数の援用がある。そこでは、パラメータ化された特定の関数が選ばれ、適切な学習則を用いて訓練データ集合からパラメータ値を決定する。入力変数の線形結合から、パターン認識で用いられる最も簡単な判別関数を構成する。線形結合の係数がモデルのパラメータで、ニューラルネットワークでは、結合荷重に対応する。非線形関数(活性化関数と呼ばれる)を用いて変換すれば、この簡単な判別関数が一般化される。そして、ロジスティ

ック回帰やパーセプトロンへと展開できる。この章では、統計的な二群判別や多群判別分析と単純ニューラルネットワークとの関連性について分りやすく解説している。また、ロジスティック判別の観点からニューラルネットワークの活性化関数を取上げている。平方和誤差関数を定義し、多層ニューラルネットワークにおけるバックプロパゲーション学習則の予備段階として、単純ニューラルネットワークに関する勾配法による誤差最小化アルゴリズムについて述べている。最後に、次元縮約に注視した相関比に基づくFisherの線形判別分析について説明している。

第3章の単純ニューラルネットワークは、第4章で多層パーセプトロンに拡張される。入力層と出力層のみからなる単純ニューラルネットワークに、隠れ層(中間層)を含めた多層パーセプトロンは、多くの問題解決に威力を発揮してきた。各層間の結合荷重は、第3章の最小二乗法を基盤にしたバックプロパゲーション学習則を用いることにより容易に推定される。ベイズ法や多パラメータ最適化アルゴリズムで中心的役割を果たすヘシアン行列についても論じている。バックプロパゲーション学習則により結合荷重を推定する際、シグモイド関数(ロジスティック関数)の有効性についても議論している。また、ニューラルネットワークと分類や回帰問題で広範に活用される射影追跡法との関連性は興味深い。

第5章は、多層パーセプトロンとともに階層型ニューラルネットワークの一つであるRBF(Radial Basis Function)ネットワークを取上げている。それは、隠れ層の基底関数(Basis Function)の出力を線形結合して出力する。基底関数としては、ガウス関数を用い、入力为中心付近にある場合のみ出力を算出する。出力層の各ユニットの出力は、隠れ層の出力の線形結合によって算出される。

第6章では、ニューラルネットワークを訓練するために用いられてきた誤差関数について考察している。5章までは、解析的容易さから平方和誤差を利用してきた。しかし、特定の応用の場面では、他の誤差関数のほうが有効なこともある。本章では、種々の誤差関数を取上げ、その有効性を示している。回帰問題において、平方和誤差関数が重要な性質を持っており、ニューラルネットワークでもその誤差関数が用いられる。分類問題では、事後確率に注目し、平方和誤差関数が採用されるが、更に適切な誤差関数も考えられる。本章では、ニューラルネットワークにおける目標値が正規分布に従うという仮定のも

とで、最尤原理から平方和誤差関数が導出される。また、分類問題のように、教師値が2値の場合、二項分布に従う。そのため二項分布の積が尤度になり、誤差関数として Kullback-Leibler 距離を導いている。

前章までは、誤差関数の最小化という視点から、ニューラルネットワークの学習の問題を定式化してきた。この誤差関数は、未知パラメータ（結合荷重とバイアス）の関数である。特に、第4章では、バックプロパゲーション学習則を用いれば、ネットワークの未知パラメータについて誤差関数の偏導関数が容易に得られた。第7章では、ニューラルネットワークの結合荷重（未知パラメータ）が最適化される種々の実用的アルゴリズムを提示している。その中で、最も単純なアルゴリズムは最急降下法である。しかし、慣性項をもつ最急降下法は、厳しい制約を受けるため、もっと強力なアルゴリズム、例えば共役勾配法やガウス-ニュートン法を推奨している。最後に、平方和誤差関数に有力な Levenberg-Marquardt アルゴリズムについても論じている。

実践の場でのニューラルネットワーク適用の成否は、データの事前加工にある。第8章では、データの事前加工に関する話題を取上げ、次元縮約や事前情報の活用に関する実用的な技法を紹介している。

第9章は、汎化性の問題に注視している。ニューラルネットワークを訓練する目的は、訓練データそのものを厳密に表現するのではなく、その生成機構を統計的にモデル化することにある。すなわち、隠れ層のユニット数を多くすれば訓練データに対する良い関数近似が得られるが、必ずしも未知データに対して良い予測値が得られる保証はない。実際の場合へ多層パーセプトロンを応用するには、未知データに対しても有効性を発揮する（汎化能力の高い）ネットワークが望まれる。バックプロパゲーション学習における未解決な問題の一つとして、ネットワーク構造に関する事前情報が得られない点が残されている。そのため、階層型ネットワークのコンパクト構造化について論じている。具体的には、不要な結合荷重を枝刈り (pruning) することにより汎化能力の向上が期待される。

第10章では、ベイズ流の観点からニューラルネットワークの取扱いについて論じている。訓練データの集合に基づいてニューラルネットワークの結合荷重の学習にベイズ法を適用できる。前章までは、最尤法（誤差関数の最小化と同等）を用い、ネットワークの結合荷重の推定値を求めてきた。一方、ベイズ法では、確信の度合を示す結合荷重の確率分布関

数を考える。データが観測されるとベイズの定理を媒介にして事後分布に変換できる。そして、新しい入力データを獲得したとき、事後分布を用い、訓練されたネットワークの予測値を推定し、結合荷重の分布からネットワークの出力の分布も得られる。ニューラルネットワークに対する種々のモデル（例えば、隠れ層に種々のユニットをもつ多層パーセプトロン、RBF ネットワーク、線形モデルなど）が存在するとき、モデル選択にもベイズ法は有力な武器になる。付録には、有効な数学的結果が載っている。特に、対称行列の性質、ラグランジュ乗数法および主成分分析などについて説明している。

以上、本書は簡単な例示による入門から、かなり高度な理論まで網羅しており、ニューラルネットワークと統計的パターン認識を学ぶうえでは、非常に読みごたえのある好適書である。ニューラルネットワークと非線形多変量解析との関連性については、豊田 (1996) が分かり易いテキストであろう。判別分析と統計的パターン認識について深く学びたい向きには McLachlan (1992) を薦めたい。また、バックプロパゲーション学習則については、安居院 (1993) が丁寧に説明しており、C言語による使い勝手の良いコンピュータプログラムも掲載している。

#### 参 考 文 献

- 安居院猛, 長橋 宏, 高橋裕樹 (1993): ニューラルプログラム, 昭晃堂  
 豊田秀樹 (1996): 非線形多変量解析—ニューラルネットワークによるアプローチ, 朝倉書店  
 McLachlan, G. J. (1992): Discriminant Analysis and Statistical Pattern Recognition, John Wiley.

### Stochastic Geometry and its Applications 第二版

D. Stoyan, W.S. Kendall,

and J. Mecke

John Wiley & Sons (1995, xix+436)

広島大学 間 瀬 茂

1987年にこの本の初版が出版されて後、著者の Stoyan と Mecke は祖国東ドイツの崩壊という混乱に巻き込まれた。特に Stoyan は大学 (TU Bergakademie Freiberg, 鉱山学校を前身とする技術系大学) の政治ボスが追放された後を受け学長に担ぎ出され、統一後の大学の再建という重責を荷なうことになった。その激務の中で第二版が準備されたわけ

で、序文に書かれてあるように、この間著者同士直接顔を合わせる暇も無く、電子メールと製版ソフト LaTeX という道具無しには到底完成はおぼつかなく、第二版では頁数が345頁から436頁へと大幅に増えたが、章立て等基本的な構成は初版と同じで、頁数の増加はここ10年程の間に得られた数多くの新しい結果の紹介・引用によるものである。

この世のすべての現象は時間と空間の中で観測されるわけで、最初からデータが空間的構造を持つ場合が多い。Buffonの針問題に代表される古典的な幾何確率 (geometric probability) 論は、積分幾何学により大成されたが、一方で現実から遊離し、その後空間的対象を扱う確率・統計理論は長い冬眠状態に入る。唯一の例外は対象の三次元的性質をその断面の観測から推定するステレオロジー (stereology) と呼ばれる理論で、生理学・材料科学等の研究者により、統計学と没交渉に研究が行われてきた。こうした閉塞状態はしかしながら1970年代後半、フランスのMatheronによるランダム閉集合理論と、東ドイツの研究者を中心とする空間点過程理論の誕生で鮮やかに打破されることになった。こうした理論の誕生の背景として観測機器や計算機の発達により複雑な空間データの観測・解析が容易になったことと、環境問題を典型とする空間的広がりがある本質的な問題の増加があげられる。Matheron自身もStoyanと同じくその名も鉱山学校 (Ecoles des Mines) という名のフランスの大学に所属していることは決して偶然では無い。こうした空間的現象の確率論と統計解析の新しい波を指す言葉として確率幾何 (stochastic geometry) と空間統計 (spatial statistics) という新しい言葉が作られたのは1980年代に入ってからであった。この新興の理論を始めて体系的に紹介したのがこの本の初版であり、たちまち基本的文献の位置を占めることになった。ただしこの本に於ける空間統計の紹介は表面的であり、数学的基礎と確率モデルに関する体系的であるが過度の数学的詳細には立ち入らぬ内容になっている。そういう意味で空間データの解析に関する基礎を多少なりとも本気で学びたい人には格好の入門書といえるであろう。空間統計の理論そのものはN. Cressieによる900頁に及ぶ大著があり (Statistics for Spatial Data, Wiley Sons, 1991)、かつての評者の様に余り文献を読まずに済んだ (読むべき文献がなかった) 良き時代はすでに過去のことである。

以下内容を順不同に紹介しよう。第2, 4, 5章では空間点過程 (spatial point process) の理論を紹

介している。空間点過程は一次元点過程の多次元径数への拡張であり、平面・空間中に位置も個数もランダムに生起する点群をモデル化する。Stoyan自身は最初一次元点過程理論を重要な道具とするORの研究者であった。但し時間径数と空間径数の違いは本質的であり、かつて安易に試みられた一次元点過程の議論の単純な拡張は不毛であることがわかっている。ポアソン点過程は場所が違えば起こる現象が独立になるモデルで単純すぎるモデルであるが、他のモデルの基礎として重要である。各点に更に付随する特性量 (例えば震央と地震規模) を合わせて考えたものをマーク付き空間点過程 (marked point process) と呼び、空間現象を定式化する際に特に重要なモデルである。第4章では空間点過程の基本的統計量を導き出す基礎となるモーメント測度と、特定の場所に点があるという条件の下での条件付分布を与えるパーム分布 (これも元来ORの分野から生まれた概念である) の紹介がされている。簡単にいえば点過程の少なくとも二次までのモーメント特性量はパーム分布の言葉で書き下せることがわかる。第6章では実用上重要な点過程モデルの例がいくつか紹介されている。特に統計力学由来のギブス点過程は現在も活発な研究の対象であり、特にマルコフ確率場と呼ばれるサブクラスは、画像処理への応用が注目を浴びている。但し点過程モデルは時系列解析のように正規モデルという便利な逃げ道が無く、ポアソン点過程を除けばその確率分布はほとんどの場合直接的な表現、解析は不可能である。

第3, 6章はMatheronによるランダム閉集合の紹介である。ランダム閉集合とは閉集合の作る空間上の確率測度であり、古典的幾何確率論では扱えぬ、もはや幾何図形とは呼べぬような複雑な空間パターンもモデル化できる。特にプーリアンモデルと呼ばれるクラスは数学的な扱いがかなり容易であり、詳しく研究されている。プーリアンモデルとは、それ自身ランダムな基本パターン集合を空間内に無作為に置いて出来る結果 (和集合) の集合である。見方を変えれば図形を点と考えた点過程であり、点過程の一般理論で定式化出来る。

第8, 9, 10章の内容は古典的幾何確率論由来の問題の空間点過程理論による定式化である。但しここでいう点は、文字通りの点では無く、直線や平面等を表す (線過程, 面過程)。こうしたモデルは自然現象そのもののモデルというよりは、ステレオロジー等で用いられるランダムなプローブ (観測用の試験線・面) として登場する。またポロノイ分割を代表とする空間分割の理論が紹介されている。

最後の第11章はステレオロジーの理論の紹介に当てられている。ステレオロジー(計量形態学)は多くの三次元試料の内部構造が直接観測不可能であり、断面の二次元的観測から推定せざるを得ないという、観測上の制限に由来する統計技術を指す。解剖学、生理学、金属学などの材料科学等で必須の道具となる。関心のある構造がしばしば顕微鏡サイズであるため、計量顕微鏡学と呼ばれることもある。理論の前提として、二次元特性量から三次元特性量(例えば対象物の断面上の面積比率から空間中の体積比率)が復元可能である必要があり、何でも推定出来るわけでは無いが、紹介されているように多くの興味ある幾何学的特性量が(少なくとも理論的には)推定可能なことがわかる。ステレオロジーにおける古典的なアプローチはランダムな断面(つまり面過程)を幾つも作って観測することにある。但し同じ試料を無数に何度も切断出来るわけでは無く、現実にはこの定式化には無理がある。二つめのアプローチは試料の三次元構造そのものがランダムと仮定する(ランダム閉集合)ことで、原理的には十分広ければたった一つの断面の観測から必要な三次元特性量を推定することが正当化される。このアプローチは更に空間点過程の二次モーメント特性量を用いた推論を可能にする。この章で特に詳しく解説されている話題に Wicksell の小球問題がある。これは空間中のランダムなサイズの排反な球群の半径分布を、断面上に現れるランダムな円盤の半径分布から推定する問題で、典型的な不適切数値問題の例として知られている。Wicksell (1925) による有名な積分等式に基づくさまざまな推定法(評者によるものも一つ)が紹介されている。

終りになるが第7章はランダム測度の理論を紹介している。これまでの紹介からわかるように、確率幾何に登場するさまざまなモデルは、見方によればある抽象的な空間の抽象的な点のランダムな集まりになっており、実際これらを抽象的に点過程として一般的にとらえることが出来る。点過程  $X$  に対し、集合  $D$  中に入る点の数を  $M(D)$  とすると、これは整数値をとる( $X$ の実現値毎に決まる、従ってランダムな)測度になる。つまり点過程は測度を値を持つ確率過程、ランダム測度、と取らえることが出来る。またランダム閉集合  $X$  に対し、例えば  $M(D)$  を集合  $D$  中に含まれる  $X$  の部分の面積とすれば、これもまたランダムな測度となる。このように確率幾何ではランダムな測度のごく自然に登場し、またそれを考えることが有益な特徴付けを与えることが多い。この章ではランダム測度に関する二次モーメント特性量

の理論と、条件付分布であるパーム分布の理論が紹介されている。基本はキャンベル測度と呼ばれるある測度を構成することにある。

なお Stoyan には細君との共著になるこの本の姉妹編とも呼べる近著 (D. Stoyan & H. Stoyan, *Fractals, Random Shapes and Point Fields*, Wiley & Sons, 1994) があることを注意しておこう。

## Growth Curves

A. M. Kshirsagar and W. B. Smith  
Marcel Dekker (1995, xv+359)

広島工業大学 神田隆至

本書は Potthoff and Roy によって提唱された growth curve model (成長曲線モデル, PR モデルと略す) 及びその関連モデル(以下これらを単に成長曲線モデルという)について解説したものである。PR モデルについては、これ以前にもいくつかの多変量解析の本の中で扱われているけれども、それらはおおむね本の一部の章または一部の節としての扱いであって、この本のように一冊全部成長曲線モデルについて述べてある本は少ない。承知のとおり経時データに関する書籍は最近ほかにも出版されており、この本とは丁度時期が同じ頃なので引用はされていないが、例えば Lindsey [1], Diggle, Liang and Zeger [2], Crowder and Hand [3] などがある。また、本書は主として PR モデル及びその関連モデルを扱っているのが特徴であり、取りあえず成長曲線モデルを知りたい者にとっては、以下で分かるようにならりの範囲の題材を取り上げており、総括的な把握をするのには便利かと思う。また、本文の他に3つの事例の説明及び多変量検定のための数値表に関する Appendix と文献もおおよそ200編があり、これらが全体の1/4を占めている。

本書は以下のような章からなり、各章は代表的な論文を主体にその他の論文等をちりばめて構成されている。さらに、各章はいくつかの節をもち、最後の節はまとめになっているので便利である。

1. Introduction
2. The Growth Curve Model
3. A Multidimensional Growth Curve Model
4. The Sum of Profiles and Time Moving Covariates Models
5. A Growth Curve Model with Exchangeably Distributed Errors
6. Structured Covariance Matrices, Model Selection Prediction from Growth Curves and Other

## Topics

7. Growth Curves with Incomplete or Unbalanced Data
8. Potthoff-Roy Growth Curve Model: Derivations of Main Results
9. Bayesian Analysis of the Potthoff-Roy Model
10. Nonparametric Methods in Growth Curve Analysis

第1章の前半では分割実験計画に代表される繰り返し測定モデル (repeated measures models) と成長曲線モデルについて、それらの類似性と違いについて述べている。成長曲線モデルは繰り返し測定モデルよりもより一般的であることを指摘している。成長曲線モデルにおける基本的な前提は処理効果とそれが適用される時間との間の関数的な関係があることであって、この関係をモデル化することに強い関心がよせられる。実際の関数関係はより一般的であるかも知れないが、特に特化されないときは、多くの場合多項式近似として扱われる。この場合、多項式の係数は成長曲線モデルにおいては興味があり、分散共分散行列に加えてこれらの母数の推定、有意性の検定、未来の観測値の予測、異なるグループに対する成長曲線の比較などはすべて重要である。繰り返し測定モデルと成長曲線モデルの差は、分散分析と回帰モデルとの間の感覚差に類似している。前者は級効果と交互作用の言葉で表され、後者は既知観測値をもつ回帰説明変数と未知係数をもつ線型または多項式関数の言葉で述べられる。線形回帰モデルや多変量分散分析モデルについての基本的な結果を概観した後、第1章の後半で3つの事例研究について説明している。事例研究1は「犬の両眼の瞳孔の大きさと眼内圧についての血管拡張剤とロイコトリエンの効果」で、30匹の犬について5つの処理を6時点で測定したデータである。これはこの章と2、6章でも扱われる。事例研究2は「1/4才馬のナトリウム・ゼオライト A (SZA) の食養生処理から生じるプラズマ・シリコンの濃度」に関するものである。60頭の離乳したての1/4才馬について、4種類の SZA を含む飼料を与え、毎日5時点で測定されるが、実験の性質上欠測値をもつデータである。事例研究3は「集中看護における気管吸引」に関するもので、25人の入院患者について、5時点で3つの気管内チューブの吸引による酸素の浸透効果が測定されたデータである。このデータは付録Cで示されており、その解析は5章で紹介される。これら3つの事例研究は統計的な方法論を例示するのに、このモノグラフを通じて用いられている。

第2章では通常の成長曲線モデル、すなわち PR モデルが紹介され、統計的解析の主な目標として (i) モデルの多項式の次数の検定 (ii) 成長曲線の係数の推定 (iii) 推定された係数の分散共分散を求める (iv) 推定された係数の分散共分散の推定 (v) 一般化線型仮説の検定 (vi) 係数のベクトルの同等性の仮説の検定 (vii) 母数の線型関数に対する同時信頼区間の構成等がある。これらについて、具体的な解析の方法を数値計算を頭の中に入れた上で、その手順を示しながら解説を試みている。初心者にとっては事例研究1や「ネズミのデータ」等具体的に引きながら進めているので理解し易いであろう。

係数パラメータ行列の次元が縮小した場合、誤差が多変量正規でなくエリプティカルが球形分布、Wilks の基準等についてふれている。

第3章は PR モデルの1つの一般化として、繰り返し測定される変数が複数個ある場合、すなわち多変量成長曲線モデルを扱っている。多変量の場合の PR モデルは、共分散行列に含まれる未知パラメータの数が多くなり、標本の数が十分大きくないときには共分散行列に、ある構造を仮定する必要がある。ここでは固体の成長曲線係数のベクトルが、時間には依存しないが、固体または他の共変量の影響を受ける固定効果とランダム効果の両方から構成されていると仮定して導入されるランダム効果共分散構造の場合を扱っている。この章ではランダム効果モデルについて、推定、一般化線型仮説検定、同時信頼区間の構成等について記述している。ここではいきなり多変量のランダム効果共分散構造が取り上げられているが、1変量の場合から引き起こせば、より良い理解が出来るかも知れない。

第4章では種々の縦断的データへの適用を目的とした PR モデルの拡張を取り上げている。PR モデルでは固体内 (プロフィール) 計画行列がすべての固体について同一であったが、いくつかの異なる固体内計画行列が現れる場合の成長曲線モデルを扱っている。ここでは考え方の流れのみに止め、後は「子羊」の例を紹介している。また、このモデルによく似ているものとして時間依存する共変量がある場合のモデルについて、最尤推定量を求める手順を与え、その母数の線型仮説の検定の手順も述べている。これらのモデルは多変量の場合の “Seemingly Unrelated Regression” モデルに帰着されるものである。

応用としてはおもしろいが、全体としてはまだ完結できていない部分を含んでおり、推測理論の発展が望まれる。



第5章は観測値が時間間で依存するのに加えて、グループ又は処理で相互に関係している成長曲線モデルを扱っている。これはデータ行列の共分散行列に対して、級内相関タイプの構造を追加することになる。第2章の標準タイプのモデルへの変換を行うと、共分散行列が異なる2つのPRモデルに帰着される。この結果を利用して、母数に対して2章における手順に従って推定量を求めている。

第6章では共分散行列が具体的な構造、例えば級内相関構造、系列共分散構造をもつ場合を扱っている。これらの分散構造に関する仮説検定、さらに Miscellaneous topics として、モデル選択、要因分析的モデル、混合モデルおよび多重比較にまで及んでいる。モデル選択については、一般的言及ではなく Geisser and Eddy (1979) による予測尤度法の紹介である。

繰り返し測定データでは不完全または不釣り合いのデータになることが多い。第7章では欠測値データを扱っている。これらの解析にあたって、多くの場合これらの情報を削除しての母数の推定、検定を行っているが効果的ではない。ここでは、欠測値がある場合の推測技法としてEMアルゴリズムを中心に述べている。しかし、この方法は数値的に計算できても、推定量等の性質などを調べるのには適していない。正規性のもとでのBAN推定量を求め方法や、不完全データに対する成長曲線モデルとSURモデルとの関連にも及んでいる。欠測値のある成長曲線モデルを初めて考える者にとっては種々な方法が紹介されていて取りあえずは便利かと思う。

第8章では第2章で述べたPRモデルについての主な結果の導出を行っている。1つの処理の場合には Rao (1959) によって導出されたが、2つ以上の処理の場合には Khatri (1966) 等による。ここでは教育的見地から Khatri の方法、 $p$ 個の時点での測定値を主変量と共変量に変換し、共変量を与えたときの主変量の条件付き分布を利用した母数の推定、検定、信頼領域等の結果の導出を紹介している。また、共変量の選択による、より有効な推定法についても議論している。

第9章では Bayesian の観点からの成長曲線モデルの Bayesian 接近を与えている。母数の事前密度を仮定して事後密度を得、Bayesian 推定及び信頼区間を構成している。これらの区間は第2章で構成された信頼区間の内部にあることを指摘している。また、ただ1つの処理の場合について、第2章の結果との比較を行っている。未来の観測値の予測では予測値の予測密度を導きその点推定を与えている。さ

らにこれらの技法の応用も与えている。

第10章では繰り返し測定データへのノンパラメトリックな接近を説明している。ここでは、核型密度推定法、ランクに基づく検定法を紹介している。これらの方法を例示するために事例研究1を用いて分析している。最後に、測定変量が離散型変量を含み、かつ、共変量が時点で依存する一般的な状況のもとでの解析法として、いわゆる GE (General Equation) 法を説明している。この部分については詳しい展開はなされていない。

本書は3つの事例研究を引いて解説をしているので、理論を理解する上で役立つであろう。なお、本書よりも簡単かつ俯瞰的にPRモデル及びその関連成長曲線モデル(多くの場合は連続型で多変量正規性を仮定している)についての最近の発展までを含む解説を知りたい者は、参考文献の中の von Rosen (1991) は見落とせない。一般的な経時データの解析法として最近注目されている GE 法の理論と応用については、本書はあまり言及していないので冒頭にあげたような他のテキストが参考になるう。

#### 参 考 文 献

- [1] Lindsey (1993). *Models for Repeated Measurements*. Oxford University Press.
- [2] Diggle, Liang and Zeger (1994). *Analysis of Longitudinal Data*. Oxford University Press.
- [3] Crowder and Hand (1995). *Analysis of Repeated Measures*. Chapman and Hall.

#### 「為替レート変動の時系列分析」

小島平夫著  
(牧野書店, 1995)

一橋大学 山本 拓

本書は、1変量時系列分析を用いて為替レートの変動についての統計的分析を試みたものである。実質為替レートに関して、種々の構造変化を含むモデルを考え、その構造変化を検出することを主要なテーマとしている。用いられている統計的手法についてもかなり詳しい理論的説明が与えられており、その意味では構造変化を取り入れた1変量時系列分析の手法の解説書としての役割も持っている。以下ではまず本書の章立てを示し、その後それぞれの章の内容について簡単な説明を行う。最後に本書についてのコメントならびに評価を述べることにする。

本書は8章からなっており、そのタイトルは以下の通りである。

1. 序：為替変動分析と1変量時系列モデリング
2. 実質為替レートの1変量時系列分析
3. 異常値モデルの統計的数理
4. 時系列構造変化の一般モデル
5. 為替レートの構造変化と為替レートボラティリティー
6. 公的為替介入の為替レート効果
7. 時間序列関係としての因果性の分析
8. むすび：まとめと展望

さらに、A. 1からA. 12まで主に数理的な解説をする12の付録がついている。

次に各章毎の内容について若干の説明を加える。第1章ではまず本書全体の概観が与えられ、続いてBox-Jenkins流の標準的な1変量のARIMAモデルに基づく時系列分析の考え方が簡単に説明されている。

第2章では、対ドル円レートのWPI(卸売り物価指数)に基づく実質為替レートと、CPI(消費者物価指数)に基づく実質為替レートの2種類のデータを用いてのARIMAモデルの構築が実際に行われる(それぞれWPI実質レート、CPI実質レートと呼ばれている)。データはまず対数を取りそのあと1次の階差を取ってある。(従って、レベルの対数を取ったデータについての単位根の存在が前提とされていることになる。)以降の分析の殆どは、このように変換されたデータが定常であるという前提のもとに行われる。データは月次で、標本期間は1973年3月から1987年12月である。

以降の章での研究の基礎として、全期間についてのBox-Jenkins流の1変量ARMAモデルが同定、推定されている。ここでの特徴は極めて長いラグのモデルが推定されていることである。たとえば表2.5では、1次と32次の遅れもつARMAモデルが推定されている。

第3章から第5章にかけては、本書の主たる貢献である、構造変化を持つ時系列モデルの分析についての理論的解説と実証分析が与えられている。ここでいう構造変化とは、Fox (1972), Tiao (1985), Tsay (1988) 等の一連の研究で異常値 (outlier) の問題あるいは干渉 (intervention) モデルとして扱われた現象のことである。第3章と第4章はいわば理論編であるが、それは上記の論文等からの解説ならびにその拡張である。

第3章ではまず基本的な2つの異常値モデル、すなわち加法的異常値 (AO: additive outlier) モデルとイノベーション異常値 (IO: innovational outlier) モデルが紹介される。AOモデルとは基準のARMA

モデルに、ダミー変数のように増分として外生的変化が加わるものである。IOモデルは変化がinnovationに起こる場合である。従ってこの変化の影響はARMAモデルに引き継がれていく。まず異常値が存在しない時のモデルを帰無仮説とし、異常値が存在する時を対立仮説とする尤度比統計量の分布がそれぞれのモデルについて導出されている。厳密な尤度比統計量の分布は複雑なので、より簡単な実用的な統計量が提示されている。そしてこの統計量の検出力関数が解析的に検討されている。ここでは、AR(1)モデルを想定し、モデルの1次の係数が1に近づくほどモデルがAOモデルかIOモデルかを正しく判定することが重要であるとの結果を得ている。さらにAOとIOが混在するMO (mixed outlier) モデルについて、その存在の検定のための仮説検定の統計量、ならびにAOあるいはIOに特定化するためのベイズルールが提案されている。

第4章では、第3章の結果を拡張し、異常値が未知の時点で生じたり、複数の異常値が生じたりする、より一般的あるいは現実的な状況における分析を取り上げている。さまざまな状況におけるベイズ分析の方法が示されている。IOモデルとAOモデルにおいて異常の発生が1回の場合のパラメータの事後分布をt分布として求めている。異常の発生的一般の場合として、干渉モデルの枠組みで整理が行われ、平均のシフトとしてシフトが永遠に持続する恒常的水準シフト (PS: permanent shift) モデルと、単調に減衰する一次的シフト (TS: transient shift) モデルを考察の対象に加えた。そしてこれらのシフトとAO、IO間の差異、類似点を考察している。そしてPSとTSはIOの特殊ケースであるという指摘が行われている。

第5章では、前章で示された干渉モデルの方法による構造変化の検出を2種のデータに適用する。第1種のデータは実質為替レートであり、第2種のデータは名目実効為替レートのボラティリティーである。

まず実質為替レートの構造変化の分析については、前章までに解説されてきた平均あるいはレベルのシフトの分析を行う前に、Tsay (1988) の反復手続きによって分散のシフトの検出を行っている。分散のシフトの影響を除去したあとの系列について、平均のシフトの検出が反復の手続きによって行われている。これはTiao (1985) やTsay (1988) の手続きに従っている。この分析の標本期間は1957年1月-1988年12月である。WPI実質為替レートとCPI実質為替レートでは大きな構造変化の時期が大幅に

異なるという結果を得ている。従って、種々の計量分析においていずれの実質レートをを用いるかで分析結果が大きく異なりうるという重要な指摘がなされている。

第5章の第4節では、これまでの分析対象とは異なり、第2種のデータとして名目実効為替レートのボラティリティーがシフト（ここでは変則的変動とよばれている）検出の対象とされている。名目実効レートとはIMFがInternational Financial Statisticsの中で発表している、当該通貨の他の全ての通貨に対する変動を測る加重に基づいた指標である。ここで対象としているのは円他に、ドイツマルクとスイスフランである。標本期間は1973年3月から1988年12月である。ボラティリティーのデータは名目実効レートの移動標本標準偏差として作成されている。移動標本標準偏差を計算する範囲としては、6カ月の場合と24カ月の2種が採用されている。さらにデータは対数変換され、そのデータに対して、上記と同様にシフトの反復的検出法が適用されている。やはり通貨によって大きな変則の時期が異なっていることが明らかにされている。

第6章と第7章は、それまでの構造変化の問題から離れ、異なった角度からの為替レートの分析を行っている。第6章は、為替のボラティリティーに対する公的介入の効果を上げている。まずデータ分析として、名目為替レートの変化率と公的介入をプロットしたグラフにより、介入が多くの国で向かい風(leaning against the wind)であることを確認している。そして向かい風政策を前提として、公的介入がボラティリティーを減少させることができるのか、あるいは増加させることになるかをさまざまな状況(仮定)の場合について、理論的に分析している。本章は純経済理論に基づく分析に特化した、本書の中でユニークな章である。

第7章は、前章まで結果についての異なった角度からのアプローチを試みている。2変量の世界を取り上げ、グレンジャーの因果関係の検定を2種のモデルに適用している。紙幅の関係でその一つだけについて述べると、名目為替レートと相対価格の間の関係を調べており、為替レートからWPI物価比への因果関係はあるが、為替レートとCPI物価比には因果関係が無いという結果を導いている。これは第5章で示されたWPI実質為替レートとCPI実質為替レートの構造変化時期の大幅な違いの原因であろうとの興味ある指摘がなされている。

最後に第8章では、かなり丁寧な各章毎のまとめと、今後の研究への展望が述べられている。

以下では、本書についてのいくつかのコメントを述べよう。なお本書では時系列データの統計的分析という側面と、一部ではあるが経済モデルの理論的分析という側面がある。私自身の専門の関係から、以下のコメントは前者を主たる対象とすることをお断りしておく。

まず全体の読後感としては、ユニークな著作だという印象を得た。それはまず本書が取り上げている統計的アプローチは近年大流行の手法を全く用いていないという点である。例えば最近の金融時系列データの分析で極めてよく取り上げられるアプローチである単位根の検定、共和分の検定、エラーコレクションモデル、ARCHモデル等については、言及はされているが、実際には用いられていない。単位根はその存在が前提され、階差データが分析の対象とされており、1変量モデルに特化しているために、共和分やエラーコレクションモデルは対象外となっている。

そして本書では、階差によって定常化されたとみなされたデータに対し、ボラティリティー(標準偏差)の変化を考慮に入れた複数の構造変化を含むモデルを実証的に追求している。私の知る限り、このようなFox, Tiao, Tsay等の方法が日本のデータに適用されたことは初めてだと思われるので、実証分析における貴重な貢献だと思われる。またその構造変化検出の手法やその考え方も丁寧に解説しており、当該手法の解説書としての価値も高い。以上が、私が本書をユニークと考える根拠である。

しかし、無い物ねだりを含めていくつかの気になる点があり、以下ではそれらを指摘する。まず実証的な中心である第5章では、構造変化の特定化に関して反復法によりシフト時点を探し出している。これ自体は極めて面白く興味深い。しかし分析自体は、反復手法による変化の特定ならびに初期的推定で止まっており、最終的に全てのシフトを考慮に入れたモデルの再推定は行われていない。これは分析としてやや中途半端であり、残念な点である。

構造変化を記述するモデルの分類に関して、第3章では異常値モデル(AOモデルとIOモデル)を取り上げ、第4章2節では干渉モデル(PSモデルとTSモデル)を取り上げている。そしてPSモデルとTSモデルはIOの特殊ケースであることが示されている。モデルの分類の観点からいえば、2段構えでやや複雑であり、最初から統一的な枠組みで整理した方が分かり良いのではないかと思われる。

本書で用いられたFox, Tiao, Tsay流のアプローチを用いたモデルによる予測が優れているとの過去

の論文が言及されている。その正当性を実際に本書の分析対象についても調べてみると本書の分析結果の重みもより増したのではないと思われる。

第7章ではグレンジャーの因果関係の検定が1フィルター法に基づく標本相互相関を調べる方法で行われている。一般に、グレンジャーの因果関係の検定には、VARモデルに基づく検定が簡便で優れていると言われている。なぜこの方法が用いられたのか、分かりにくい。

以上2, 3の不満・希望を述べたが、コメントの冒頭で述べたように、日本のデータについて、Fox, Tiao, Tsay 流のアプローチを用いた文献として、初めての著作であり、興味深い。従って、為替レートの実証分析に関心のある読者、ならびに複雑な構造変化を含む時系列分析に興味のある読者にとって一読の価値がある著作であると評価する。

#### 参 考 文 献

- Fox, A. J. (1972). "Outliers in Time Series," *Journal of the Royal Statistical Society, Ser. B*, 34, 350-363.
- Tiao, G. C. (1985). "Autoregressive Moving Average Models, Intervention Problems and Outlier Detection in Time Series," in *Handbook of Statistics*, vol. 5, ed. by E. J. Hannan, P. R. Krishnaiah and M. M. Rao, North-Holland.
- Tsay, R. S. (1988). "Outliers, Level Shifts, and Variance Changes in Time Series," *Journal of Forecasting*, 7, 1-20.

### Excel による統計入門

縄田和満著

朝倉書店 (1996、v + 198 頁)

東京大学教養学部 松原 望

大学設置基準が廃止されて「要綱」化されて以来、統計学の研究・教育環境は大きく変わった。ことに、教養課程における基礎教育の中の統計学の位置づけは、おそらく(大学により事情はさまざまであろうが)自由選択に近い形に近づいた場合が多いであろう。統計学の科目は他の人文、社会、自然の各科目と競争状態に入ったことになる。これが「自由化」といわれる措置だが、良いことか否かは歴史に判断が任されることとなるであろう。とはいえ、統計学教育の現場での真価が問われることに変わりはない。

統計学を科学的真理の追究、いわば「サイエンス」と考える人は多いし、それはそれでよいが、コミュ

ニケーション(統計データによる伝達)の手段、説得の技法と考えることもできよう。統計学を教えることは、自ら実演でそれを証明せねばならず、まさに「紺屋の白袴」といわれかねない場におかれてる。学生の聴講数が回とともに指数曲線のように減衰してゆくのを前にすれば、負担が減ったと喜ぶ反面、人への伝達技術としての統計学の教育において、いささか内心おだやかならぬ気持ちを抱くのも私だけではあるまい。実際、統計学を専攻とする学生は100人に1人もいない。99人までがユーザーである。これは統計学の宿命であると同時に、いつも絶えることのない社会的ニーズを表している。世に絶えた学問は多い。

本書はその意味で(ユーザーのためのわかりやすい統計学)のテキストである。(わかりやすい本)は学者にとって難しい。第1にわかり易く書くことは著者が本当にわかっていないと書けない。わかっていることと書けることの比は5:1位であろう。つまり、本の著者が達人でなければならない。第2に、わかりやすい本は学者が書いてはならないとする掟がある。若手である著者はこの2点に挑んでいる。なぜそれが可能だといえ、若者独得のコンピュータに対する勘の良さであろう。今までも、多くのコンピュータ統計学の指導書があるが、一般的には難しく高レベルである。実際、通常の学問書より却って範囲も広くかつ難しい。これは何とかならないであろうか。

本書の特色は、Excelにできることをすべては扱っていない点である。できることから演習を通じて進み、理論へと向かう。これは統計学の学問史や精神にかなっている。煩雑さは概ね省かれている(完全ではない)。初心ユーザーが陥りやすい誤りをもカバーしている心配りよりもよい、データは本当の実際から取られており、架空例、人工例などはない。全体として、今後の統計学の教育の進むべき(さらにいえば、そう進まざるをえない)方向を示したものと見えよう。ただし、統計学のテキストかコンピュータの指導書かの中間の性格はやはり否めない。もっとも、これは統計学の本質的な二重性、つまり一方に理論的厳密性、他の一方に実用性(あるいは実利性)を目指すことの表れである。実際、これをはっきりと分離すると、クックブックないしはケースブックになる可能性もあり、難しい所である。読み返して見て結局これ以外にはないという感もある。

内容は、データ・プレゼンテーション、記述統計、正規分布の検定、推定をカバーする。データ・プレ

ゼンテーションは「統計」とはいえない統計と考えられてきたが、今後を考えると入れるのが良い。ユーザーの関心の深まり方の経路もこれが多いであろう。記述統計は可も不可もない内容であるが、順位、連、最大、最小、記録値、Trimmingを入れてノンパラメトリックにした方がよかった。これこそ統計分析の華だからである。検定、推定は妥当な所である。分布論はマクロ命令でさらっと済ませたのがよい。一般的には分布論(総論)をどう扱ってゆくかが、今後の大きな課題の一つである。もっとも、分布論の発想法自体はたいへん大切なものでおろそかにしてはならないが、

統計学とは面白い分野で、全く畑違いの分野の人でも本が書けると思われているようである。実際、書店で数学全分野と並ぶ程の点数が出版されているが、うち半数程が素人の著者による。真の統計学書は少ない。問題はこれらが難しくユーザーには歯が立たないという矛盾である。「やさしい物理学」「わかりやすい化学」「経済学入門」などは物理学者、化学者、経済学者が多く著しているが、統計学ではそのようなジャンルそのものが専門家によって成り立っているとは思えない。一つの問題である。本書がそういう今後の一つの展開になれば幸いである。

以下内容を簡略に示しておこう。

- 1～4章: Excelの基本操作の学習
  - 1章: Excelの使い方入門
  - 2章: 表計算機能の説明
  - 3章: 棒グラフ、円グラフ、線グラフ、X-Yグラフなどのグラフ作成
  - 4章: 78カ国の国別人口データの入力(5～8章、11～13章の演習は、このデータを使う)
- 5～8章: 記述統計
  - 5章: データの並べ替え、条件にあったデータの抽出
  - 6章: 度数分布表の作成(6～7章は1次元データの整理・分析)
  - 7章: 代表値、散らばりの尺度の計算
  - 8章: 分割表の作成と相関係数の計算による2次元データの整理・分析)
- 9～10章: マクロ命令と確率分布の学習
  - 9章: マクロ命令と Visual Basicの説明
  - 10章: マクロ命令と Visual Basicを使った確率分布と乱数の発生学習
- 11～13章: 推測統計
  - 11章: 正規母集団に関する推定と検定(点推定と区間推定、仮説検定、2つの母集団の同一性の検定)

12章: 2次元データにおける2変数間の関係の検定(分割表による独立性の検定、一元配置分散分析、相関係数のフィッシャーのZ変換を使った検定)

13章: 回帰分析

付章A: Windows 95/Excel 7.0に関する説明。

## 貯蓄と資産形成

高山憲之・有田富美子著

岩波書店(1996, xvi+238頁)

東京理科大学経営学部 寺崎 康博

「資産形成」というとバブル経済が崩壊した現在ではもの悲しい響きを持つ言葉かもしれない。土地評価額の高騰とその後の急落は人々の生活の尺度に狂いをもたらした。将来に向けた様々な構造改革が議論されているが、必ずしも明るい見通しが定まっているわけではない。しかし、家計資産が着実に増加してきたのも事実である。金融資産残高だけを見ても、1975年では年収のほぼ106%であったのが、1985年では153%、1995年では210%にのぼる(農家を除く全世帯)。住居についても1住居あたりの広さが1973年の77平方メートルから1983年では86、1993年では92に増加している(住宅統計調査)。また、1992年では土地を持つ家計は57%に達している(土地基本調査)。少なくとも平均水準では蓄えが蓄積され、豊かな生活が実現しているといえよう。しかし、家計資産といっても、金融資産もあれば、土地、住宅等の実物資産も含まれる。年金請求権等を含んだ考え方も存在する。一般に、金融資産がある程度蓄積された段階で実物資産の購入が進むといわれているので金融資産、あるいは実物資産のどちらか片方のみを見ていたのでは家計資産の全貌がわからないことになる。我が国においては1970年の国富調査を最後として、ストックに関する総合的な調査が実施されておらず、資産の保有状況についての分析は不十分な段階でとどまっていた。1970年の国富調査ですら土地保有に関しては調査していないので完全なものとは言えない。このような状況の中で、既存の統計調査を駆使して家計資産の全体像をとらえるプロジェクトが著者達を中心に経済企画庁で始まり、1986年から1990年まで続いたとされる。このときの共同研究の成果は「経済分析」に掲載されるときに、『ストックエコノミー』として公開されている。また、それらの成果を反映する形で『1989年全国消費実態調査報告 資料編(家計資産)』が編纂された。本書はこれらの研究の続報であり、ここ4、

5年ほど精力的に発表してきた研究成果を一書として整理したものである。その構成は以下の通りになっている。

### 第I部 資産保有の実態と資産増加の諸要因

#### 第1章 家計資産の分配とその変遷

#### 第2章 家計資産の増加とその要因

#### 第3章 家計資産の蓄積と遺産・相続の実態

#### 第4章 共稼ぎ世帯の家計実態と妻の就業選択

### 第II部 高齢者の生活実態資産保有

#### 第5章 高齢夫婦世帯の所得・消費・資産

#### 第6章 高齢単身世帯の所得・消費・資産

#### 第7章 同居高齢者の経済状況

#### 第7章補論 同居・非同居の決定要因

### 第三部 世代間の所得再分配

#### 第8章 可処分所得の世代間分配

#### 第9章 1994年の年金改正とその評価

本書は初出の論文をほぼ全般にわたって書き直し、数多く得られた推計結果を整理し、それに関する主要な分析が平易な文章でまとめられているので読みやすい。しかも、立ち入った分析が行われている第3章、第5章および第7章では章末に要約が掲載されているので、著者が何を見だし、どのように解釈しているのかがよくわかるように工夫されている。また、本書の特長も著者達自身のはしがきに述べられており、周到である。

しかし、本書の価値を比類のないものにしてるのは主に『全国消費実態調査』のマイクロデータを利用して再集計を行って新しい事実を見だしていること、およびマイクロデータから得られる情報を利用して推計を行い新たな知見を得ているところにある。『全国消費実態調査』は政府が実施する統計調査の中でも「指定統計」であり、マイクロデータの形で研究者自身が分析できる機会は「目的外使用申請」を行って承認された場合に限られている。ちなみに、ここ10年ほどの学問的研究(医学研究を除く)を目的とした「目的外使用申請」の承認件数は年間10件を下回る数で推移しており、16件承認された1987年を最後に停滞している(舟岡・寺崎(1995))。このような現状において、行政機関である経済企画庁で約5年間にわたってマイクロデータの分析を行い、十分な経験を積んだ著者達は『全国消費実態調査』を実にさまざまな角度から再吟味している。マイクロデータの形式でこのような長期間にわたって分析できる機会を得た研究者は他にあまり存在しないのではないか。しかし、著者達は基礎的作業としてできるだけデータから読みとることのできる事実を丹念に分析することに専心しており、恵まれた機会を得た

研究者としての努めを十分に果たしていると言える。さらに、分析の視点や問題意識も豊富であり、教えられるところが多い。

以下、各章についてコメントを交えながら概観する。まず、第1章は前書『ストックエコノミー』との接続的な役割を果たし、1984年から1989年にかけての家計の総資産(正味資産)の推計とその分配状況、及びその変化に関する記述にあてられている。第2章は家計資産の増加の要因を分析している。40歳未満では相続贈与による資産増加が一番大きく、中高年層の資産増加はキャピタルゲインによるところが大きいことを発見している。前者は予想通りであるが、後者については1989年がバブル最盛期でもあり、新しいデータで再確認する必要があるだろう。

第3章は「家計における金融資産選択に関する調査」(郵政研究所)のマイクロデータを使用して、家計資産の蓄積と遺産・相続の実態を明らかにしている。遺産動機には職業差、地域差が見られるほか、相続経験者の方が強い遺産動機を持っていること等を見だしている。これらの差をもたらす背景については深く分析が行われているわけではないが、今後この分野で取り組むべき課題を提起したところに意義があるだろう。経済学者のみならず幅広い分野からの立ち入った分析が望まれる。

第4章は就業している妻をもつ世帯の分析に当てられているが、貯蓄あるいは資産形成に関しては現状の記述があるだけで、分析はほとんど行われていない。分析の中心はどちらかと言えば、両者のフローの所得の関係、妻の就業選択、夫婦の職業の関係等におかれている。就業選択モデルとして比較的新しい方法である多項ロジットモデルを試みるなど興味深い点が多い。なお、夫婦の職業面から見て結婚のペアリングが必ずしもランダムでないことを見出したとしている点については、共稼ぎ世帯に限った分析であることにも配慮すべきであり、また社会学の分野での研究成果等を参考にして幅広く吟味する必要があるように思われる。

第5章から第7章までは高齢者世帯の経済状況の記述と資産形成に関する分析に当てられている。ここでの主要な論点は著者達が従来から主張している「高齢者の多数派はもはや貧困ではない」ということである。数多くの反論を受けそうな主張であるが、これだけ高齢化が進展して深刻な社会不安が発生していないことを考慮すると著者達の主張は単に数字のみならず、実感をも反映したものであると言える。丹念なデータの吟味が行われているが、特に興味深いのは同居高齢者の所得や消費等の経済状況を切り

離して推計した点にある。一般に、世帯所得や世帯消費に関するデータは得られるが、それが世帯内でのように獲得され、あるいは消費されているかを知ることは公表されたデータを使用する限りでは難しかった。しかし、マイクロデータの使用はそれを可能にしたのである。著者達の推計によれば同居高齢者の経済状況は比較的高い水準にあるとしている。「日本における同居高齢者が経済的に恵まれない高齢者 (the dependent old) であるという従来の通念は事実の半分にすぎない」(180頁)と言うのか、「半分も存在する」と見るのかは評価が分かれるところかもしれないが、今後さらに検討が重ねられるべき課題である。同居高齢者の経済状況の分離にあたってはいくつかの仮定がおかれており、それらの妥当性を含めた検討は必要になる。

第8章と第9章は世代間の所得分配を扱っている。高齢化が進展し、年金の負担が増加しつつある近年では世代間の公平さが意識されるようになってきた。「世代会計」に関する議論はその一つである。新たに世代間で再分配された可処分所得を推計した第8章では、高齢者は年金受給によって相対的には平均水準にまで回復するが、30歳代は相対的に最も所得水準が低いことを明らかにしている。興味深い指摘である。なお、この章の図8.2から8.5の縦軸は相対度数と表示されているが、相対比率とか相対所得とすべきであろう。第9章は1994年の年金法改正についての評価が行われているが、本書のテーマであるマイクロデータの分析は表だっでは使用されていない。この年金法改正に当たっては著者の1人である高山氏は大きな役割を果たしたといわれ、「年金改革の構想」をはじめさまざまな機会に発言をしてきているので本章はその要約としても便利なものであろう。

わが国の現状では政府統計のマイクロデータを扱える研究者は限られており、本書で示された実証的な結果には新しい事実として発見されたものが多い。また、著者達もその事実発見を中心に作業を行っており、複雑な現象を整理して新しいモデルを提案するということは将来の課題とされている。プロビット分析とか多項ロジット分析等が一部試みられているが、適用例としてとどまっており、本格的な取り組みは今後の課題と言える。また、このような分析が深まるためには多くの研究者がマイクロデータを使用できるような環境が必要である。実際、欧米における傾向を見ると今後の実証研究の中心はマイクロデータの分析に移っていくと思われる。現在、「統計行政の中・長期構想—統計審議会答申」を受けて「標

本データ」(マイクロデータ)の提供の可能性等が検討されており、前向きの結果が出ることを期待したい。

最後に、本書は高齢化社会の生活状況に関心ある一般の読者、ならびに世帯に関して社会、あるいは経済的側面から実証分析を志す研究者にとって得るところの多い一書であり、推薦したい。なお、著者の一人である高山氏はマイクロ統計データの取り扱いでは最も経験豊富な研究者の一人になられたわけであるので、わが国における実証研究水準の向上のためにもそのノウハウ等を今後の若手の研究者に何らかの形で伝えていただくことを期待したい。

#### 参 考 文 献

- 高山憲之他(編著)(1992)『ストックエコノミー』東洋経済新報社。  
高山憲之(1992)『年金改革の構想』日本経済新聞社。  
舟岡史雄・寺崎康博(1995)「最近の日本の micro-data 分析の状況とその背景」『新重点領域「統計情報活用」のフロンティアの拡大』に関するセミナー—マイクロデータによる社会構造分析—』35-42頁所収。