

北川 源四郎 著

## 『Rによる時系列モデリング入門』

岩波書店、2020年12月刊、A5判、334ページ、4,180円（税込み）

近年の日本の社会では「データ・サイエンス」という言葉が氾濫している。AI (Artificial Intelligence) 関連の話題がしばしばメディアに登場しているが、我々の身の回りの社会・経済、日々の暮らしの中では消費者物価 (CPI)、賃金・所得 (GDP)、株価、為替レートと云った時間とともに観測されるデータ、時系列データも少なくない。したがって統計データ分析のプロ、「統計家」を名乗るには時系列データの統計的分析方法を理解しておく必要があるだろう。

本書の著者である北川源四郎氏は統計科学、特に統計的時系列解析においては日本を代表する研究者、統計数理研究所の元所長、統計関連の学会では非常によく知られた統計家である。これまで統計学をあまり勉強したことがない読者は恐れをなすかもしれないが、本書は統計的時系列分析の教科書である。

まずは本書の概略を紹介しておこう。第1章「時系列データの解析とその準備」、第2章「共分散関数」、第3章「スペクトルとピリオドグラム」、第4章「モデリング」、第5章「最小二乗法」、第6章「ARMAモデルによる時系列の解析」、第7章「ARモデルの推定」、第8章「局所定常ARモデル」、第9章「状態空間モデルによる時系列解析」、第10章「ARMAモデルの推定」、第11章「トレンドの推定」、第12章「季節調整モデル」、第13章「時変係数ARモデル」、第14章「非ガウス型モデル」、第15章「粒子フィルタ・平滑化」、第16章「シミュレーション」、全16章から本書は構成されている。

説明の便宜上で内容を第1章～第3章+第16章の準備編、第5章～第8章の基礎編、第4章+第9章～第15章の展開編に分けておこう。準備編では準備事項をまとめている。統計的時系列解析は元々は数学の一分野である確率論の展開に基づいてはいるが、その細部に立ち入ることなく、まずは実際の時系列データを分析する際に必要な伝統的枠組み、共分散関数とスペクトル概念を導入している。本書の最大の特徴と思われるが、統計的時系列分析の基本的道具立ての基礎事項が、(評者が学んだ頃に比べて恐縮だが) 近年では自前のPCさえあればかなり容易に活用できることを例証する形で議論が進んでいる。実際、統計的時系列分析を行うのは一昔前まではかなり敷居が高かったが、今日ではフリーソフトウェアRを利用すると多くの道具立てが比較的容易に利用可能となっている。例えば自己共分散やピリオドグラムといったかなり込み入った統計量をデータ入力さえ間違えなければ、北川氏の開発したフリーソフト TSSS (詳しくは <https://jasp.ism.ac.jp/ism/timsac/> 参照)、さらに web 上でだれでも利用可能な RS-Decomp (詳しくは <https://jasp.ism.ac.jp/RS-Decomp/> 参照) を用いるとすぐに実行が可能なのである。次に基礎編では時系列分析を勉強すると必ず出てくる時間領域の統計モデル、AR (自己回帰) モデルと ARMA (自己回帰移動平均) モデルの推定・予測などの説明がある。さらに統計的時系列分析のもう1つの道具立てである周波数領域におけるピリオドグラムの説明

が続く。

本書の大きな特徴は展開編の内容である。実は学問的にもかなり奥が深い内容が多く、他にあまり類を見ない内容にもかかわらず、説明は統計理論にかたよらず具体的な例を多用するなど、北川氏ならではの説明により展開されている。展開編は赤池情報量による統計的モデリング、時系列の状態空間モデルによる統計処理、という観点から首尾一貫した説明が行われ、季節調整法といったかなり実用的な話題から粒子フィルタというかなり先端的話題までカバーされている。こうした統計的モデル分析は先ほど言及したTSSSやRS-Decompによる実装を例示していることには評者もかなり驚かされた。一例としてPCに全く強くない評者も、日本で公表されている公的統計における幾つかの公的データの原時系列の季節性を分析する際には、TSSSとRS-Decompに実際にお世話になっていることを付け加えておこう。

ここで本書の特徴と評者のコメントを幾つか述べておく。本書の読者としては様々な分野においてこれから実際に時系列データを分析したいと考えている学生・院生・研究者である。したがって、なるべく労力・頭脳・時間を使わずに楽しんでデータ・サイエンティストになりたいと希望する学生には本書は不向きだろう。何事にも「投資がなければリターンもない」、というのが経済系の本評者の理解である。しかし、せっかく限られた時間と労力をつぎ込むのならば、優れた導き手を選んだ方が良いだろう。必ずしも数理的な基礎がなくともある程度の努力を厭わなければ、本書とフリーソフトRを利用して比較的短い期間にレベルの高い統計エキスパートになることが可能、とは評者の意見である。例としては、一昔前までならば自分のPCでARIMAモデルを自由に使って将来値を予測したいとすると、これを実現するには多大の労

力が必要であったが、今日ではその必要はかなり軽減されているのである。むしろ状態空間モデルを利用する統計的時系列分析など先端的な時系列分析はいうに及ばず、である。

本書は本誌の一般読者向けというには少し専門的かもしれないと感じつつ、データ・サイエンスという言葉が流行る中で統計分析に関心のある方々に、確固たる内容の書籍を推薦することにも意味がある、と評者は考えている。近年ではHow to …という手軽な書籍が本屋に氾濫、実は書籍ではなくとも少し探せばネット上でも時系列の統計分析の情報が手に入る機会も多くなってきている。こうした動向には歓迎できる側面もあるが、他方、時系列データの生成過程を考慮することなく、「機械学習で日米の株価を予測したほうがよい」、「為替レートをうまく予測できた」などという類の学生レポートも散見される。そうした学生諸君には評者はまず作成した予測モデルを使って訓練データではない将来値を予測するとして、その予測結果を報告しなさいとコメントをすると、沈黙することが多い。過去に観察した時系列データによる統計モデルの構築と将来値の予測は赤池情報量規準の原点、本書の第4章で（多くの人は少し難しいと思われるが）詳しく議論されている。現時点で実用的かつ正統的な統計的時系列分析を学びたいと希望する学生・院生も少なくないのでは、とは年配者の希望的観測である。

現代日本の社会・経済では確かにデータ・サイエンスは進展している。統計的時系列分析の初級から最前線まで駆け足で到達したいと希望している読者には適切な教科書として本書を推奨したい。

評者：国友 直人・くにとも なおと  
(統計数理研究所特任教授、東京大学名誉教授)