

はじめに

本書は、『Rで学ぶ確率統計学(一変量統計編)』の続編である。本書ではタイトル通り主として多変量の統計学について学ぶ。一変量統計編と同じく、比較的数学は得意だがプログラミングとなるとちょっと…という人に向けて書いたもので、数理統計学とRの操作の説明を同時に行うのが本書の基本的な姿勢である。

統計学は、自然科学、社会科学、医学、工学など様々な分野で利用される「言語」の1つである。これらの分野では、実験や調査票の結果や、臨床で得られたデータを扱う。統計学は、これらのデータを要約し、ノイズをふるい落とし、法則をつかみ出す。例えば、「気温や熟成期間、雨量からボルドーワインの価格を予想する公式」、「労働力と資本からGDPがどうなるかを予想する公式」、「風速と気温から大気中のオゾンの量を予想する公式」は全て「重回帰分析」と呼ばれる多変量統計の技術を使って求めることができる。このように、統計学を学べば、様々な分野を横断することができるのだ。本書では、多数の例題・問題を通して統計学が様々な分野で役立っていることが実感できるように最大限配慮した。予備知識としては、『Rで学ぶ確率統計学(一変量統計編)』の内容を仮定している。

本書は、自習書としても、教科書としても使うことができる。各章ごとにいくつかの練習問題があり、読者はこれらの問題を解きながら、Rの操作を確認するとともに本文の内容の理解を深めることができる。解答は一部を除いて完全なものである。Rスクリプトの動作チェックは、『Rで学ぶ確率統計学(一変量統計編)』と同様に、64bit版R(ver.3.5.1)とR Studio(ver.1.1.456)で行った。演習に使うデータはエクセル形式のファイルまたはCSVファイルであり、比較的長いRスクリプト(スクリプト1~9, A1~A5)と併せて出版社のサイトからダウンロードできる*1。

本書を出版するにあたり、内田老鶴圃社長の内田学氏ならびに編集者の方々には大変お世話になった。また、統計技術の使い方については東北学院大学の神林博史先生にご助言いただいた。記して感謝したい。

本書は、東北学院大学における講義資料に大幅に加筆したものである。一変量統計編と同様に、神永の講義ノートをもとに木下が数式のチェック、Rスクリプトの動作チェックなどを行う形で執筆した。統計学上の誤りがあれば、全て神永の責任である。

2019年11月

著者

*1 http://www.rokakuho.co.jp/data/04_support.html