

序

本書はR言語を楽しく使いながら統計学（主として古典的な部分）の要点を勉強しようという趣旨で書いたものである。できるだけ類書にない書き方を心がけた。

R言語の予備知識や統計学の知識は仮定していない。高校程度の数式は出てくるが、苦手なかたは無視していただいてもかまわない。また、「☞」マークがついた「脇道の話題」の部分も、難しそうなところは飛ばして読んでいただいてもかまわない。

第1章はRの基本である。すでにRをご存じのかたはこの章を飛ばすか、あるいは「☞」（脇道の話題）だけ拾って読んでいただければよいであろう。逆に、R初心者のかたは、最初は「☞」をすべて無視されることをお勧めする。ちなみに、本書で使っているRは、ほとんどが15年以上変わっていない部分、おそらく今後も当分は変わらないと思われる部分である。また、プログラミングの初心者を悩ませる制御構造（ifやforの類）は、本質的でないプログラム例でほんの少し使っているだけで、実質的な計算は1行ごとに完結するよう心がけた。1行プログラムの積み重ねでほとんどのことができるのはRの強みである。

第2章は、記述統計の基本から、Rで実験的に導いた中心極限定理と正規分布および関連するいくつかの分布を扱う。このあたりはRに慣れるための部分でもあるが、適当に読み飛ばし、後で必要になったら参照していただいてもよい。

第3章からが本書の中心となる部分である。第3章では、強いて正規分布から始めず、高校生でも十分に数学が理解できる2項分布から出発して、古典的な統計学の考え方の基本を述べる。特に、仮説検定の考え方、 p 値と信頼区間については、ここで十分に理解していただきたい。ベイズ流の考え方については、本書では触れるにとどめた（もう1冊本が必要である）。最後に、2項分布の極限あるいは近似として、再び正規分布に立ち戻る。

第4章は、2項分布の特殊な場合（一つの極限）として、ポアソン分布を導入する。ポアソン分布は統計学の入門書では詳しく扱わないものが多いと思うが、個数（カウント）データに広く適用できる重要な分布であり、本書ではかなり詳しく扱った。Feldman-Cousinsの信頼区間など、他書にあまりないことも書いた。

第5章はおもに 2×2 の分割表の話で、フィッシャーの正確検定や、疫学で重要となるオッズ比などの話をかなり詳しく扱った。

第6章は、連続量に立ち戻り、測定値の誤差（不確かさ）、 t 検定、簡単な分散分析の話を買った。

第7章は、古い統計学を勉強した人が陥りがちな「有意か否か」という二分法的な考え方を脱して、いわゆる効果量（他と比較できる方法で測った量）とその不確かさをちゃんと報告すれば

(できれば生データを公開すれば), メタアナリシスによって次第に不確かさを減らせるという話である。

第8章は相関係数, 第9章は回帰分析で, いずれも重要な話である。

第10章は特に理系向けのピークフィットの話で, 度数分布を使わない (unbinned な) ポアソンフィッティングの話まで書いた。

第11章は主成分分析と因子分析の話で, こちらはおもに文系向きであるが, 理系っぽい例も挙げて説明した。

第12章は, 段階で答えるアンケートのようないわゆるリッカート型データを例として, ノンパラメトリック検定やブートストラップ法を紹介した。

第13章では, 臨床研究でよく使う生存時間解析の考え方と計算法を紹介した。

これらの内容は, 三重大学教育学部での授業や, 全学対象の講習会などで扱ったことを中心に, その際に用いた私の Web サイト [83] を再構成・詳細化したものが大部分を占めるが, 一部の話題は30年前の拙著 [57] (絶版) から採った。あまり他の本を参考にしないで書く癖があるので, 独特な書き方になってしまったところも多いと思われるが, このシリーズの編者である尊敬する石田基広先生に全ページをチェックしていただいたおかげで, 何とか読める形になった。原稿共有に快適な GitHub を使うことができたのも石田先生のリーダーシップのおかげである。

共立出版の石井徹也さん, 大谷早紀さんを初めとする皆様には, たいへんお世話になった。大谷さんには L^AT_EX のコマンドレベルの誤記もたくさん直していただいた。

思い返せば, 私が R 言語の前身の S 言語を知ったきっかけは, 25 年前に共立出版から出た S 言語の本の訳書 [2] であった。Bell 研究所で作られた C 言語に次ぐ S 言語ということで, たいへん興味を持って勉強した覚えがある。今回, S の後継である R を使った統計の本を同じ共立出版から出すことができることを嬉しく思う。

組版については, すばらしい L^AT_EX スタイルファイルを作ってください, 一部の字形を仮想フォントで置き換えるという難しい要求にも応じてくださった啓文堂電算部の宮川憲欣さんに感謝したい。

欧文の書体については, コード (プログラム) の類は `sapply()` のようなタイプライタ体, R のパッケージ名は `readxl` のようなボールド体, それ以外の索引語は CP932 のようなサンセリフ体を使っている。

本書で用いたサンプルデータやコード例は本書サポートページ <https://github.com/okumuralab/RforFun> からダウンロードできる。本シリーズ全体のサポートページは <http://www.kyoritsu-pub.co.jp/series/205/> である。

2016 年 8 月

奥村晴彦