

まえがき

高度情報化社会といわれる今日、身の周りのあらゆる「もの」がデータとして扱うことができるだけでなく、自動的にデータとして記録されている。そのようなデータは、大量の数値情報からなることもあり、その場合にはデータを理解することも非常に困難である。データマイニングと呼ばれる方法などにより、そのような大量のデータを扱うことが容易になってきているが、その中で特に重要な役割を果たすのが統計データの視覚化である。ビックデータと呼ばれる新たなチャレンジにおいても、大量のデータの振る舞いを把握するために視覚化は重要である。

本書ではまず第1章で、統計データについて説明し、データの要約方法から解説して基本的な視覚化の方法について紹介する。その後、統計解析環境 R における視覚化の基本として、標準的なグラフ表示の方法について紹介する。R はコマンド入力により解析やグラフの作成を行うため、R の主要なコマンドや R の仕組みについても理解することが必要なので、基本的な R の使い方について、特にグラフを作成するための最低限の使い方について解説する。

第2章から第5章では、R で作成できるグラフとその作成方法について紹介する。第2章では、R の初心者や R のコマンドに慣れていない読者でも、コマンドを使わずにメニューからさまざまなグラフを作成することが可能な、R コマンダーを用いたグラフ作成について解説する。どのようなグラフが R で作成できるか、各グラフではどのような変更が可能かについて理解できるだろう。第3章では、R の使い方として一般的なコマンド入力により、R の古典的な関数を利用してグラフを作成する方法について紹介する。さらに、R におけるグラフのオプションの詳細についても紹介し、さまざまなグラフを作成する関数を紹介する。R でのコマンドによる作業に慣れていない読者は、第2章の方法で実際にグラフを作成し、ある程度理解が高まったら第3章でのコマンドによるグラフ作成方法に慣れてほしい。なぜなら、R コマンダーで実現可能なグラフは表現の制限が多いが、コマンドにより R の関数を用いてグラフを作成する際には、より柔軟にグラフの詳細が変更可能となる。R の機能を十分活用するなら、コマンドラインからの利用について理解する必要がある。しかし、関数の利用方法のすべてを覚える必要はなく、どのようなことができるかさえ知っていれば、ヘルプ等を参照することで必要な設定は可能である。本書では、R における新しいグラフについても紹介するため、古典的グラフについてあまり詳細に解説することは避け、よく使うグラフとそのグラフを実現するために必要な変更方法を重点的に解説する。

引き続き、第4章では、質的データの視覚化法としてまとめられた **vcd** パッケージで実現可能なグラフについて、作成方法とそのグラフを利用したデータ解析方法について解説する。第5

章では、3次元データを描画するための、**scatterplot3d** パッケージ、**rgl** パッケージ、そして **misc3d** パッケージについて紹介する。

第6章以降は、高度なグラフを作成するためのパッケージを利用したグラフ作成方法について紹介する。第6章では、**lattice** パッケージと **ggplot2** パッケージなど高度な統計グラフを作成するための環境を提供する **grid** について紹介する。**grid** について知らなくても、**lattice** パッケージや **ggplot2** パッケージを使うことはできるので、高度なグラフィックスの仕組みについて知りたい、独自のグラフを作成したい、などの場合を除いてこの章は読み飛ばしてもかまわない。第7章および第8章では、古典的なグラフに比べデザインも含めより使いやすい、高度なグラフ表現が可能な **lattice** パッケージと **ggplot2** パッケージによるグラフ作成について紹介する。Rで利用できる統計解析のためのパッケージが、最近では **lattice** や **ggplot2** を用いるものも増えてきており、今後のRにおける標準的なグラフになっていくと思われる。これらのパッケージについて詳細を理解することは、統計解析のためのパッケージで得られるグラフについて詳細を変更できることにもつながるので、必要に応じてより深く理解するようにするとよいだろう。

グラフでは、一部を確認したり、他のグラフとの関連について確認したいこともある。その要望に応えるグラフもある。第9章では、グラフ作成後に対話的にグラフの表示を変更することが可能なインタラクティブグラフを提供する **iplots** パッケージについて紹介する。

さらに、グラフの応用として第10章では、地理情報などの位置情報を伴うデータの視覚化について、**mapproj** パッケージ、**GeoXp** パッケージなどを用いた方法について紹介する。

Rのパッケージは非常に多いため、どのようなことができるのか、自分がやりたいことがどのRパッケージでできるのかについて知らないと使いようがない。第11章では、Task Viewとしてグラフ作成のためのパッケージをまとめた **Graphics** ビューを紹介し、これまでの章で紹介していないいくつかのパッケージについて紹介する。本書の内容を実施するにあたり、必要となるパッケージは **Graphics** ビューでインストールされるので、まずこの章を読み **Graphics** ビューをインストールしておくとう便利だろう。

さらに、グラフを作成するために理解するとよい補足としてデータの種類についての解説と、古典的なグラフの詳細となるグラフィックパラメータと低水準作図関数の解説を付録に与えた。

本書においては、統計グラフについての節の見出しとして、その節で紹介するグラフ作成関数の関数名を示している。これは、関数の項目別の索引として目次で確認できることを意図している。実際には関数を入力することのないRコマンドーについての第2章についても、内部で用いられている関数名を紹介するために示している。

本書で扱っている、Rのパッケージとしては提供されていないデータや本書でのRコードをはじめ、本書についての情報については、本書のWebサイト

<http://rbookvisjp.web.fc2.com/>

に公開しているので活用いただきたい。統計グラフではカラー表示は重要であるが、本書ではすべてをカラー表示としては提示できないため特に重要なグラフについて口絵として示している。Webサイトにはコードと一緒にすべてのグラフ見本をカラーで提示している。

2013年4月

著者一同