

# まえがき

R は統計解析と作図 (graphics) のための人気のあるオープンソースソフトウェアツールです。本書は R の中でも、とくに作図のしくみを紹介するものです。R の作図はとても強力なので、論文など出版に使えるような品質の図・統計グラフを作ることができます。

## この本は何を説明しているか

本書は R の作図システムについて説明しています。最初の章では R の作図のしくみの全体像を俯瞰します。ここでは R を使って生成したさまざまな複雑な図や統計グラフを掲載しています。R 作図で使えるいくつかの出力フォーマット、そして R 作図システムの全体構成の説明も含まれていますので、ユーザーの皆さんはある特定の目的をもつ関数がどこにあるのかわかるでしょう。

R 作図のもっとも重要な特徴は、R の中にまったく異なる作図システムが二つあることです：traditional 作図システムと grid 作図システムです。1.2.2 項でどちらのシステムを使うべきなのかを説明しています。本書の第 I 部では traditional 作図システムをあつかいます。これは S 言語 [11][5] の「伝統的な (traditional)」作図のしくみを実装したものです (S 言語はもともと Bell 研究所で開発されたもので S-PLUS など商用ソフトウェアでも使えます)。本書の執筆時点において、R 作図関数の多くはこのシステムを使っています。本書の第 I 部の各章では traditional 作図関数の使いかたを説明します。とくに最終的な出力を修正したり加工する方法に重点をおいています。第 2 章では完備した図を作るときに利用できる関数、第 3 章では図の詳細を変更する方法、複数の図

を組みあわせる方法，そして図への追加出力について説明します．

第 II 部は grid 作図システムについて説明しています．これは R だけで使えるもので，traditional システムより格段に強力なものです．執筆時点では，grid を使っている統計グラフを描く作図関数の種類はそんなに多くはありませんが，grid にはもっといろいろな種類のグラフを作る能力があります．この grid を使って完全なグラフを作ることのできる関数のほとんどは Deepayan Sarkar の lattice package に入っています．これは Bill Cleveland の Trellis 作図システムを R 用に実装したものです．この lattice については第 4 章で説明します．残りの各章では，grid で空白のページに作図シーンを作っていく方法を説明します．とくに，他の人たちにとって使いやすく部品として利用しやすい新しい作図関数を開発する方法を検討してみます．

付録 A は R システム一般の短い入門です．付録 B では traditional と grid の両作図システムを組みあわせる方法を検討しています．

本書のほとんどの部分では，ユーザーの皆さんが R 言語と環境の基本をある程度はわかっている，と想定して書いています．R についてもっと詳しい情報が必要ならば R プロジェクトの home page (URL は後述) にあたってみてください．ネット上の (英語で書かれた) 関連文書と印刷物の一覧があります．

R のグラフィカルユーザーインターフェイス (GUI) を作ろうとするプロジェクトは多数あります．しかしながら，コマンドラインを使った R の操作はどんな OS でも使える基本となる方法です．本書では，R で作図するときコマンドラインインターフェイスから対話的に R コードを入力したりスクリプトに書かれたコードを R に読ませる，つまりバッチジョブとして走らせる方法をあつかっています．

## 本書は何を説明していないか

本書ではこういったデータにはこういう統計グラフが最適です，といったことは検討していませんし，正しい図の提示のしかたのガイドラインも含んでいません．実際のところ，円グラフとか交叉斜線で図をおおうといった，一般にあまり賛成できないような統計グラフや作図要素についても説明しています．

本書では，すでにどういう形式の統計グラフを使うのが決まっているときに，どうやって図を作るのか，すでにあるデータセットをいろいろな方法で図にし

てみる実験方法について説明しています。特定の統計グラフの形式の説明を意図的には排除していません。というのも、どんな統計グラフもまったくダメということはありえません(たとえば、円グラフは単純な比率を示すのに効果的な方法です)、本や論文の出版社があまり推奨できない作図要素、たとえば交叉斜線を使えと指示することがあるからです。

従来の統計グラフにだけユーザーをしばりつけないようにする、これが R 作図における柔軟性によって可能になります。本書の目的はたくさんの便利な道具を示すこと、そしてその使いかたを説明することです。作図のガイドラインと推奨される統計グラフの種類をあつかった本などはたくさんあるので、以下にそのいくつかを挙げてみます。

多くの統計学入門書には適切な統計グラフの種類を選ぶガイドラインが書かれているでしょう。専門家ではない人たちに見せる効果的な統計グラフの作りかたをあつまっている本としては Naomi Robbins[50] の “*Creating More Effective Graphs*”, Edward Tufte の “*Visual Display of Quantitative Information*”[59], “*Envisioning Information*”[60] などがあります。この問題についてのもっと専門的な考察には、Bill Cleveland による “*Visualizing Data*” [12] と “*Elements of Graphing Data*”[13], そして Leland Wilkinson による “*The Grammar of Graphics*”[66] を読んでみてください。

ある特定のデータ解析やデータ種類をどう適切に図示するかについては、まず読むべき本がいくつかあって John Maindonald and John Braun[36] の “*Data analysis and graphics using R*”, John Fox の “*An R and S-Plus Companion to Applied Regression*”[19], Richard Heiberger and Burt Holland の “*Statistical Analysis and Data Display*”[28], そして Michael Friendly の “*Visualizing Categorical Data*”[24] などです。

本書は完全な R リファレンスでもありません。付録 A はとても短い R の入門にすぎないものですが、世の中には入門的であるだけでなく R システムについてもっとつっこんで説明している無料の文書がたくさんあります。まずは R プロジェクトウェブサイトの home page の “Documentation” セクションから始めるのがいいでしょう(ix ページの「ネット上の情報」をみてください)。入門的な教科書を二つほどあげると、Peter Dalgaard による “*Introductory Statistics with R*” [17], そして John Verzani “*Using R for Introductory Statistics*” [64] があります。また、より進んだ内容のよく読まれている教科書としては Bill Venables

と Brian Ripley による “*Modern Applied Statistics with S*” [63] があります。

最後に、本書は R の標準的なインストールではインストールされない追加パッケージ (add-on package) の多くの作図関数については詳しくは説明していません。本書では、デフォルトで R に入っている作図のしくみ、とくに `grDevices`, `graphics`, `grid`, そして `lattice` パッケージをあつかっています。R に入っている全ての作図関数の一覧はもちろん、何か作図関数を含む追加パッケージの一覧すら作るうとはしませんでした；それはとても長くなってしまいうし、常に増えつづけているからです。ただし、とくに明示している場合をのぞいて、本書にでてくる全ての追加パッケージは、R の中心的なダウンロードサイトである CRAN\* から入手できます。

## S-PLUS との違い

R の traditional 作図システムは S 言語の traditional 作図システムの再実装です。つまり、本書の第 I 部で traditional システムについて述べていることは、S の商用版、つまり S-PLUS の traditional 作図のしくみにもあてはまるということです。しかしながら、R と S それぞれの traditional 作図システムには大きな違いがいくつかあり、たとえば色や線のタイプを文字列で指定できるとか、統計グラフを配置するレイアウトの基本概念とか、テキスト中で数学的な表記が使えるかどうかといったことなどです。このような違いがあるので R 用に書かれたコードが S-PLUS で同じ結果を返すかどうか（そもそも走るかどうか）は保証されません。また、第 II 部で説明している `grid` 作図システムに関しては、そもそも S-PLUS では使えません（ちょうど S-PLUS の対話的に編集可能な作図システムが R では使えないのと同じです）。

本書では R で使える作図システムを重視しているので、S-PLUS との詳しい相違は本文中では強調しません。しかしながら、第 I 部で述べていることの多くは S-PLUS の traditional 作図システムにも適用できるでしょう。

---

\*The Comprehensive R Archive Network; <http://cran.r-project.org>

## この本をどう読めばいいか

この本はさまざまな R ユーザーの興味をひくはずです。R の初心者の皆さんは、本書によって R の作図システムの全体像、つまり、作図関数でこんな図が作れそうだと、いったことがわかってくるでしょう。そして、図を追加したり修正する方法を理解する助けになるでしょう。こういった目的のために、第 1 章と第 2 章を読んで標準的な統計グラフを作り始めるのがよいでしょう。ただし、そのグラフを微調整するために、すぐに第 3 章のあちこちを読んでみることになるかもしれません。また、第 4 章をちょっと見てもらうと、`trellis` 作図システムでできることが理解できます。

R 中級者にとって、本書は R で作った統計グラフの洗練されたカスタマイズに必要な全ての情報を提供しています。多くのアプリケーションソフトウェアでもそうであるように、R を数年間つかっていていともなお、重要かつ便利な特性を知らずにいることがあります。この本では、R 作図の全貌に関する知識、そして R で図を作るための正しいモデルを説明しているので役にたつだろうと思います。1.2 節と 1.3 節、さらに第 3 章と第 4 章を最初に読んでください。新しい作図関数を作ってみようとする人たちは、第 5, 6, 7 章を読んでみるべきです。

R 上級者たちは、本書から首尾一貫した再利用可能で拡張可能な作図関数を作るのに必要不可欠な情報が得られるでしょう。上級ユーザーたちはとくに第 II 部を注意して読んでみてください。

## 本書で使われている表記法

本書にはたくさんの R 関数とコード例があります。R のコマンドラインで対話的に入力するコード例は、以下のような書式で表わします：

```
> 1:10
```

ここで `>` は R のコマンドラインプロンプトで、それ以外はユーザーが入力すべきものです。もし入力が行におさまらないほど長いときには：

```
> plot(1:10, 1:10, col="blue", lty="dashed",  
      axes=FALSE, type="l")
```

このようにインデントされた追加行として示します。

本書で説明されている関数は、図を出力するという副作用 (side-effect) のために使われることが多いので、関数を実行させたときの結果は図で示します。関数の結果が何か知っておくべき値である場合には、その結果は実行したコードの下に、このような書式で示します：

```
[1] 1 2 3 4 5 6 7 8 9 10
```

まったく新しい R 関数を定義してるところもあります。そういうコードは (コマンドラインから一行ずつ入力するのではなく) スクリプトファイルを記述して R に一度に読みこむのがふつうでしょうから、新しい R 関数のコードは図として次のような書式で示しています：

```
1 myfun <- function(x, y) {  
2   plot(x , y)  
3 }
```

本文からその部分を参照するときに便利なので行番号をつけています。

本文の中である関数を参照するときは、関数名を typewriter 体で示し、さらに plot() のように関数名のあとにかっこをつけています。

関数の引数 (argument) やその引数の値を参照するときも typewriter フォントを使いますが、その場合には `x`, `y`, あるいは `col="red"`, というように末尾にかっこをつけません。

S3 クラスを参照するときは「"classname" クラス」といった書式で typewriter フォントと二重引用符をつけて示します。しかしながらあるクラスのインスタンスであるオブジェクトを参照する場合には「classname オブジェクト」というふうに typewriter 体を使いますが、クラス名に二重引用符はつけません。

## ネット上の情報

下の URL に本書のウェブサイトがあり、誤植訂正と本書の全ての図の PNG 版、それらを生成した R コードへのページのリンクがあります。

```
http://www.stat.auckland.ac.nz/~paul/RGraphics/rgraphics.html
```

さらに RGraphics パッケージ も準備しています。これには本書の図を作るのに必要な関数全て、そして本書で定義されたメソッド（第 7 章参照）が含まれています。このパッケージは CRAN で入手できます（vi ページの脚注を参照）。

## R のバージョンについて

ソフトウェア開発というものはどんどん継続されていくものなので、本書で紹介している R の作図のしくみは現在のスナップショットにすぎません。本書の中の説明とコード例は R の version 2.1.0 かそれより新しいもので正しいものです。さらに二箇所を別にすれば、第 7 章のほとんどの例は R version 2.0.1 でも正しいコードです。これらの個別の事例ごとに新旧の相違、さらに可能な場合には R version 2.0.1 用にどう修正すればよいのかといった情報についての脚注をつけました。第 I 部のほとんどの内容はもっと古いバージョンの R でも正しいものですが、互換性がない特定部分に関しては本文でも示していません。

だいたい 6 ヶ月ごとに R の新しい“マイナー”バージョンがリリースされています。もっとも最近の R と grid はオンラインヘルプページと R プロジェクトと grid パッケージの home page で最新の情報を入手できます：

```
http://www.R-project.org/
```

```
http://www.stat.auckland.ac.nz/~paul/grid/grid.html
```

## 謝辞

R による作図は R それ自身の存在なくしてありえませんが、何もかも全てを始めてくれた Ross Ihaka と Robert Gentleman に最初に感謝したいと思います。R 開発コアチームにはかくも信頼性が高く、高品質なソフトウェアを作ってくれることについて、幅広い層をもつ R コミュニティーには R で仕事す

ることでのやりがいと楽しみを与えてくれることについて感謝します。

R の traditional 作図システムの成功と人気はもともとの S 作図システムの卓越した設計のおかげです。R に特有の traditional システムの拡張の名声のほとんどは Ross Ihaka のおかげです。grid システムに関しては、まあ開発者である私がほとんど全ての責めを負うべきなのでしょう。

本書について言えば、原稿に対して有益なフィードバックをくれた John Chambers, Ross Ihaka, Duncan Murdoch, Stefano Iacus, Deepayan Sarkar, そして匿名の査読者たちに感謝したいと思います\*。

最後に Ju に最大の感謝を。

*Paul Murrell*

オークランド、  
ニュージーランド、  
2005 年 5 月

---

\*この原稿は L<sup>A</sup>T<sub>E</sub>X 組版システムと Friedrich Leisch の Sweave パッケージ、多数の GNU ソフトウェアツール、そしてもちろん R を Fedora Core 1 Linux システム上で動かして作られました。