

Join the discussion @ p2p.wrox.com



Wrox Programmer to Programmer™

大数据应用与技术丛书

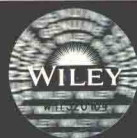


Beginning R: The Statistical Programming Language

R编程入门经典

——大数据时代的统计分析语言

[英] Mark Gardener 著
蒲成 译



清华大学出版社

大数据应用与技术丛书

R 编程入门经典

——大数据时代的统计分析语言

[英] Mark Gardener 著
蒲 成 译

清华大学出版社

北 京

Mark Gardener

Beginning R: The Statistical Programming Language

EISBN: 978-1-118-16430-3

Copyright © 2012 by John Wiley & Sons, Inc., Indianapolis, Indiana

All Rights Reserved. This translation published under license.

Trademarks: Wiley, the Wiley logo, Wrox, the Wrox logo, Programmer to Programmer, and related trade dress are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates, in the United States and other countries, and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc., is not associated with any product or vendor mentioned in this book.

本书中文简体字版由 Wiley Publishing, Inc. 授权清华大学出版社出版。未经出版者书面许可，不得以任何方式复制或抄袭本书内容。

北京市版权局著作权合同登记号 图字：01-2015-2197

Copies of this book sold without a Wiley sticker on the cover are unauthorized and illegal.

本书封面贴有 Wiley 公司防伪标签，无标签者不得销售。

版权所有，侵权必究。侵权举报电话：010-62782989 13701121933

图书在版编目(CIP)数据

R 编程入门经典——大数据时代的统计分析语言/(英)加德纳(Gardener, M.)著；蒲成译。—北京：清华大学出版社，2015

(大数据应用与技术丛书)

书名原文：Beginning R: The Statistical Programming Language

ISBN 978-7-302-40676-1

I. ①R… II. ①加… ②蒲… III. ①统计分析—应用软件 IV. ①C819

中国版本图书馆 CIP 数据核字(2015)第 157134 号

责任编辑：王 军 韩宏志

装帧设计：孔祥峰

责任校对：成凤进

责任印制：刘海龙

出版发行：清华大学出版社

网 址：<http://www.tup.com.cn>，<http://www.wqbook.com>

地 址：北京清华大学学研大厦 A 座 邮 编：100084

社 总 机：010-62770175 邮 购：010-62786544

投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn

质 量 反 馈：010-62772015，zhiliang@tup.tsinghua.edu.cn

印 刷 者：清华大学印刷厂

装 订 者：三河市漂源装订厂

经 销：全国新华书店

开 本：185mm×260mm 印 张：27.5 字 数：704 千字

版 次：2015 年 9 月第 1 版 印 次：2015 年 9 月第 1 次印刷

印 数：1~3500

定 价：59.80 元

产品编号：062552-01

译者序

随着计算机在各行各业中的广泛使用，如何汇总、分析、统计专业的信息化数据成为各领域专业工作人员日常需要应对的问题。人们迫切需要一整套工具来帮助他们高效、快捷地完成信息处理、数据分析的任务。由此，R 应运而生；并且它是免费的。所以 R 的风靡也就顺理成章了。

R 是一套完整的数据处理、计算和制图软件系统。其功能包括数据存储和处理系统、数组运算工具(其矢量、矩阵运算方面功能尤其强大)、完整连贯的统计分析工具、优秀的统计制图功能等。R 是简便而强大的编程语言，可操纵数据的输入和输出，可实现分支、循环，用户可自定义其功能。

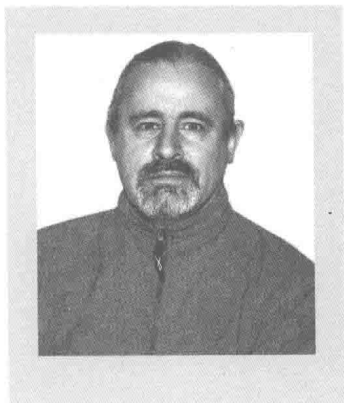
本书作为一本关于 R 的入门级书籍，内容涵盖了从 R 的获取和安装到使用 R 进行自定义编程的整个应用过程。能逐步引领读者一步一个台阶地提升 R 技能和经验，最终踏上 R 高阶应用的平台。相信在通览并深刻理解了本书的内容后，读者就能熟练运用 R 轻松应对日常工作所需的数据分析任务了。本书列举大量 R 应用示例，借助这些示例为读者详细讲解了 R 方方面面的功能，并在示例之后加入步骤详解和练习实践，以期让读者通过练习将书本内容转化成自己的知识。

本书前半部分重点介绍 R 自身的基础特性，其中包括 R 的对象以及数据转换和处理，这对于读者快速理解 R 的应用基础知识是很有必要的。本书的后半部分则重点介绍在 R 中如何进行各种统计分析方法的应用以及数据提炼和制图，这部分内容体现了 R 的重要价值，也是各行业专业人士借助 R 所能产出的关键结果，对于日常工作的总结分析、汇总报告都非常有用。相信通过阅读本书，读者能对数据统计和分析以及 R 和相关工具的使用有一个全新的体会和认知。

在此要特别感谢清华大学出版社的编辑们，在本书翻译过程中他们为译者提供了巨大帮助，没有他们的热情付出，本书将难以顺利付梓。

本书全部章节由蒲成翻译，参与翻译活动的还有王佳、何东武、杨达辉、申成龙、李文强、杨帆、王滨、李鹏、负书谦、赵栋、林超、陈世佳。对于这本经典之作，译者本着“诚惶诚恐”的态度，在翻译过程中力求“信、达、雅”，但鉴于译者水平有限，错误和失误在所难免，望请读者指正，我们将不胜感激。

作者简介



Mark Gardener(<http://www.gardenersown.co.uk>)是在英国工作的一位生态学家、讲师和作者。他对于自然界以及学习新事物有很大的热情。最初他从事的是光学方面的工作，但于1996年返回校园学习，并最终获得了生态学和进化生物学的博士学位。这一工作涉及大量的数据分析，因而他对R逐渐产生兴趣，并将R当成了帮助研究的一项工具。目前他是自由职业者，并为各种组织提供生态学、数据分析和R方面的课程。Mark和妻子Christine(一位生物化学家)居住在德文郡的乡村，并且仍然喜爱自然界以及学习新事物。

致 谢

首先感谢R项目团队以及为让R成为无与伦比的程序而不知疲倦地工作的许多作者和编程人员。我还要感谢我的妻子Christine，她在我写作本书期间给了我最大的包容，在此期间，她很大程度上变成了R造成的寡居主妇！谢谢Wiley出版公司让我来编著本书，其中包括Paul Reese、Carol Long和Victoria Swider。没有你们，我无法完成本书的编著。还要感谢技术评论家Richard Rowe，是他最初将我带入R的世界，让我感受到R令人信服的(更确切地说是令人着迷的)能力。

最后但同样重要的，总体而言我要感谢R社区。很大程度上我是通过反复试错来学习使用R的，并且享用了R社区中巨大的知识财富。我希望我已成功地将这一知识提炼成有价值的封装包，以便未来的R爱好者使用。

——Mark Gardener

前 言

本书主要介绍有关数据分析的内容以及称为 R 的编程语言。它正迅速变成各专业事实上的标准，并且被应用到了从科学和医学到商业和工程学的每一个可以想象到的知识领域中。

R 不仅是一个计算机程序，它还是一种统计编程环境和语言。R 是免费和开源的，因此可供每一位使用计算机的用户所用。它非常强大且具有灵活性，但它也不同于你可能使用过的大多数计算机程序。你必须将命令直接输入程序中，并让它可用。由于这一点及其复杂性，掌握 R 是一项难度很大的任务。

本书会深入探究 R 语言并使用简单数据示例让其可用于探究其能力和多用性。在学习如何“使用 R 进行表述”的过程中，你要释放其潜力，并且即便在应对最复杂的数据分析任务时也能得到更好的见解。

本书读者对象

本书适合那些需要进行数据分析的读者，无论他们的专业和工作是什么。不管你是否在科学、商业、医学或工程学领域工作，你都需要分析数据并呈现结果。R 功能强大且具有灵活性，并且是完全跨平台的。这意味着你可以将数据和结果分享给任何人。有一个规模庞大的项目团队支持 R，所以免费并不意味着低劣！

如果你完全没有学习过 R，那么本书将让你能够获得它并开始你的 R 学习之旅。开始学习 R 并不需要你预先具备与该程序有关的某方面的知识。如果你已经了解了 R，那么你会发现本书是一本有用的参考书，你可以反复使用本书中的知识；第 1 章主要与安装 R 有关，所以你可能会希望跳转到第 2 章。

本书的重点并非统计分析，因此熟悉一些基础分析方法会有所助益(但这并非必需的)。本书主要介绍让 R 为你所用的方法；这意味着要学习的是 R 语言而非统计方法。一旦你熟悉了 R，你将能使用它来执行各式各样的分析任务，而这并非一本书的内容能够全面介绍的。R 还可以生成展示用的高质量图形，这本书将引导你全面掌握其复杂处理。

本书内容

R 是一种计算机程序和统计编程语言/环境。它允许你使用各种分析方法并且会生成展

示用的高质量图形。本书将全面介绍 R 语言并引导你更好地理解如何使用 R 完成所需的任务。本书不会过于着重介绍实际的统计检验；相反，R 的灵活性是如此之大，它能够执行的检验列表非常长，要在像这样一本入门级书籍中完全介绍其检验方法是不现实的。更准确地说，本书的目的是让读者熟悉 R 语言并能执行一些常用的统计方法。这样一来，你就能进行自我学习并亲自探究 R 的全部潜力。

所以，本书重点放在 R 自身的操作上。跟随本书的学习过程，你将了解如何执行一系列通常会用到的统计方法，其中包括方差分析(ANOVA)以及线性回归分析，这两种分析方法被广泛应用于各行各业，因此掌握它们非常重要。你还将学习生成各种各样满足需求的图形的一系列方法。

本书涵盖了各个最新的 R 版本。随着新版本的发布，R 程序时时刻刻都在保持着更新。不过，你需要掌握的大多数命令并没有变化，即便是较旧的版本也能运行良好。

本书的组织结构

本书具有循序渐进的特征，后一章的内容往往构建在你之前学习到的技能的基础上。因此如果你是一个初学者，那么你很可能会发现从本书开头开始学习并渐进式逐章阅读本书的方式是最有用的。如果你是一个经验较丰富的读者，那么你可能会希望使用所选的章节作为参考资料，以便提升你的技能。

没有任何一种学习 R 的方法是完全通用的，但我尽可能尝试提供了最合理的学习路径。例如，学习生成图形非常重要，但除非你知道可能需要呈现的分析类型，否则制作这些图形可能会比较乏味。因此，主要介绍图形的章节一般会放在主要介绍分析的章节之后。

总体而言，本书从如何获得并安装 R 以及如何访问帮助系统开始介绍。接着介绍数据的基础知识——例如，如何将数据放入 R 中。之后，你将了解如何处理数据，执行一些基础统计分析，并开始处理图形。其后你将学习一些更高级的分析方法并回到图形的内容介绍上面。最后，将介绍使用 R 创建你自己的程序的方法。

每一章的开头都会对你将要学习的主题进行概述。章节内容包含许多示例，并且是以“跟我学”的样式编写的。贯穿全文，所有的概念都在简单示例的引导下进行阐释。你可以从配套学习网站下载数据并在你阅读本书的过程中使用这些数据(稍后将详细介绍这一点)。本书包含各种建议你进行练习的活动；每一个练习活动都旨在帮助你理解一个重要主题。所有的章节结尾处都带有一系列练习题，它们会帮助你强化理解所学的内容(附录 A 中提供了答案)。最后，所有章节结尾都提供了你所学内容的简要汇总，以及一个揭示主题和一些要点的表格，这可用作参考材料。下面是每一章内容的简述。

第 1 章“R 介绍：R 是什么以及如何获得它”——在该章中，将介绍如何获得 R 并在你的计算机上安装它。你还将学习如何访问内置的帮助系统以及找出有用分析例程的附加包，你可以将这些附加包添加到 R 中。

第2章“出发：开始熟悉R”——该章将让你建立起使用R进行工作的一些了解，从介绍一些简单数学运算开始，到导出和制作你可以使用的数据对象(以及将数据保存到硬盘以供后续使用)为止。

第3章“出发：处理对象”——该章将介绍如何操作你已经创建或导入的数据。这些都是重要的任务，能够构成许多后续练习的基础。你在此处学到的技能将反复投入使用。

第4章“数据：描述统计和制表”——该章的内容都是关于数据汇总的。此处你将学习与基础汇总方法有关的内容，其中包括累计统计。你还将学习如何进行交叉制表以及如何创建汇总表格。

第5章“数据：分布”——该章将介绍使用图形方法可视化显示数据——例如直方图——以及数学方法。该章还会介绍一些关于随机数和不同分布类型(例如正态分布和泊松分布)方面的内容。

第6章“简单假设检验”——在该章中你将学习如何执行一些基础统计方法，比如t检验、相关分析和关联关系检验。学习如何执行这些检验有助于你执行更复杂的分析，并帮助你了解使用R的一系列技术。

第7章“图形分析介绍”——在该章中你将学习如何生成一系列图形，其中包括柱状图、散点图和饼状图。这是本书在制作图形方面的“首次介绍”，第11章将回到这一主题进行详细介绍，其中你将学习如何将你的图形从刚好合格提升为极佳效果。

第8章“公式符号与复杂统计”——当你的分析变得更为复杂时，你就需要一种更为复杂的方法告知R你希望做什么。该章主要关注R的一项重要元素：如何定义复杂情况。该章有两个主要部分。第一部分会介绍如何使用公式符号处理简单情况。第二部分会使用方差分析这一重要的分析方法作为内容阐释。该章的其余部分主要介绍ANOVA。该章很重要，因为定义复杂分析情况的能力在某些时候是必备的。

第9章“操作数据并提取组成部分”——该章构建在前一章的基础上。现在你已经了解了如何定义更复杂的分析情况，学习了如何制作并重新整理数据以便能够更轻易地对其进行分析。这些内容同样构建于你在第3章中所获得的知识的基础上。许多情况下，当你已经执行了一项分析时，你需要为特定分组提取数据；该章还将介绍更多你需要的工具，以便轻松执行复杂分析。

第10章“回归分析(线性建模)”——该章的内容都与回归分析有关。它构建在之前章节的基础上，涵盖了这一重要分析方法的方方面面。你将了解如何执行基础回归分析，以及复杂模型的构建和曲线回归分析。它也非常重要，因为它揭示了R的一些有用特性(例如，如何分析结果)。该章后半部分将介绍回归分析图形方面的内容，比如如何添加最佳拟合线和置信区间。

第11章“图形的更多相关介绍”——该章基于介绍图形的第7章以及关于回归分析的第10章的内容，介绍如何从你的数据中生成更符合自己需要的图形。例如，你要学习如何将文本添加到绘图和坐标轴，以及如何制作上标和下标文本以及数学符号。你将了解如何将图例添加到绘图，以及如何将误差线添加到柱状图或散点图。最后，你将学习如何将图

形作为高质量图形文件导出到硬盘，以便适用于发布。

第 12 章“编写你自己的脚本：开始编程”——在该章中，你将学习如何开始生成能够自动化处理工作流程的自定义函数和简单脚本，以及让复杂和重复性任务的处理变得更为简单。

使用本书的前提条件

要学习本书的知识，你唯一需要的东西就是一台计算机和学习热情！R 程序适用于所有操作系统，所以你可以使用(任何版本的)Windows、Macintosh 或 Linux。R 甚至可以完美运行在古老的计算机上，因此你不需要特别专用的工具。有时候需要互联网连接，因为你需从 R 项目网站获得 R。不过，完全可以在另一台计算机上下载安装文件，并将这些安装文件转移到你的工作计算机上。

如果你已经拥有了某个版本的 R，就不必非得获得最新版本了。R 在持续更新和改进，但 R 的旧版本也非常有可能适用于本书介绍的内容，因为基础命令集的变化很小。话虽如此，我还是建议你将 R 更新到 2009 年之后的版本。

约定

为帮助你尽可能理解本书的内容并且能随时注意到当前正在介绍哪方面的内容，我们在本书中使用了很多约定。

显示由用户输入的命令行的每个示例都以>符号开始，它模拟了 R 的指针，就像这样：

```
>help()
```

以除>符号之外的内容开头的行表示来自 R 的输出(但要注意所输入的那些很长且分布于多行的内容)，所以在以下示例中，第一行是由用户输入的，而第二行是结果：

```
> data1  
[1] 3 5 7 5 3 2 6 8 5 6 9
```

试一试

“试一试”是你应该进行处理的练习题，请遵循书中的内容进行练习。

- (1) 它们通常由多个步骤构成。
- (2) 每个步骤都有一个编号。
- (3) 使用你的数据库副本遵循这些步骤进行练习。

示例说明

在每一个“试一试”之后，都会详尽说明你所输入的代码。

警告：

此类特殊段落表示非常重要、不可遗忘的信息，这些信息与其周围的内容密切相关。

提示：

此类特殊段落表示提示、贴士、暗示、技巧以及当前内容的旁注。

代码有两种不同的表示方式：

- 我们使用非高亮的等线体类型来表示大多数代码示例。
- 我们使用粗体来强调在当前上下文中特别重要的代码。

勘误表

尽管我们已经尽了各种努力来保证文章或代码中不出现错误，但错误总是难免的，如果您在本书中找到了错误，例如拼写错误或代码错误，请告诉我们，我们将非常感激。通过勘误表，可以让其他读者避免受挫，当然，这还有助于提供更高质量的信息。

请给 wkservice@vip.163.com 发电子邮件，我们会检查您的反馈信息，如果是正确的，我们将在本书的后续版本中采用。

要在网站上找到本书英文版的勘误表，可以登录 <http://www.wrox.com>，通过 Search 工具或书名列表查找本书，然后在本书的细目页面上，点击 **Book Errata** 链接。在这个页面上可以查看到 Wrox 编辑已提交和粘贴的所有勘误项。完整的图书列表还包括每本书的勘误表，网址是 www.wrox.com/misc-pages/booklist.shtml。

p2p.wrox.com

要与作者和同行讨论，请加入 p2p.wrox.com 上的 P2P 论坛。这个论坛是一个基于 Web 的系统，便于您张贴与 Wrox 图书相关的消息和相关技术，与其他读者和技术用户交流心得。该论坛提供了订阅功能，当论坛上有新的消息时，它可以给您传送感兴趣的论题。Wrox 作者、编辑和其他业界专家和读者都会到这个论坛上来探讨问题。

在 <http://p2p.wrox.com> 上，有许多不同的论坛，它们不仅有助于阅读本书，还有助于开发自己的应用程序。要加入论坛，可以遵循下面的步骤：

- (1) 进入 p2p.wrox.com，点击 **Register** 链接。
- (2) 阅读使用协议，并单击 **Agree** 按钮。
- (3) 填写加入该论坛所需要的信息和自己希望提供的其他信息，单击 **Submit** 按钮。
- (4) 您会收到一封电子邮件，其中的信息描述了如何验证账户，完成加入过程。

提示：

不加入 P2P 也可以阅读论坛上的消息，但要发布自己的消息，就必须加入该论坛。

加入论坛后,就可以张贴新消息,响应其他用户张贴的消息。可以随时在 Web 上阅读消息。如果要想让该网站给自己发送特定论坛中的消息,可以单击论坛列表中该论坛名旁边的 **Subscribe to this Forum** 图标。

关于使用 Wrox P2P 的更多信息,可阅读 **P2P FAQ**,了解论坛软件的工作情况以及 P2P 和 Wrox 图书的许多常见问题。要阅读 FAQ,可以在任意 P2P 页面上单击 **FAQ** 链接。

源代码

在你理解本书中的示例时,你可能会选择手动输入所有的数据和代码或者使用本书提供的源代码和数据对象文件。本书中使用的所有数据和源代码都可以从 <http://www.wrox.com> 下载。另外,也可访问 www.tupwk.com/downpage,输入本书中文书名或 ISBN,下载源代码和文件。你会发现每个示例练习活动所需的数据集都具有一个下载图标和表明数据文件名称的注释,这样你就能知道它可供下载并能在下载文件中轻易找到。打开该站点后,只要找到本书的书名(使用搜索框或书名列表中的一个)并点击本书详情页面上的 **Download Code** 链接,以获得本书的所有源代码。

只有一个要下载的文件,名为 **Beginning.RData**。这一文件包含全书所需的所有示例数据集和脚本。当你将文件下载到计算机后,就可以用几个方法中的一个将它加载到 R 中:

- 对于 Windows 或 Mac,可将 **Beginning.RData** 文件图标拖放到 R 程序图标上;如果 R 还没有运行和加载该数据的话,这样就能打开 R。如果已经打开了 R,则该数据将附加到 R 中已经有的任意内容上;否则只会加载文件中的数据。
- 如果使用的是 Windows 或 Macintosh,也可以使用菜单命令或在 R 中输入一个命令来加载该文件:
 - 对于 Windows,使用 **File | Load Workspace**,或者在 R 中输入以下命令:

```
> load(file.choose())
```

- 对于 Mac,使用 **Workspace | Load Workspace File**,或者在 R 中输入以下命令(与 Windows 相同):

```
> load(file.choose())
```

- 如果使用的是 Linux,则可以使用 **load()**命令,但必须(在引号中)准确指定文件名,例如:

```
> load("Beginning.RData")
```

Beginning.RData 文件必须位于你的默认工作目录中,如果不是这样,则必须将文件位置指定为文件名的一部分。

提示:

由于许多书都具有类似的书名,所以你可能会发现根据 EISBN 搜索是最简单的;本书的 EISBN 是 978-1-118-16430-3。

或者,你也可以在 <http://www.wrox.com/dynamic/books/download.aspx> 上打开 Wrox 的代码下载主页面,以查看本书及所有其他 Wrox 书籍的可用代码。

目 录

第 1 章 R 介绍: R 是什么以及 如何获得它..... 1	
1.1 学习 R 的诀窍..... 2	
1.1.1 R 网站..... 2	
1.1.2 从 CRAN 处下载和安装 R..... 3	
1.2 运行 R 程序..... 7	
1.3 找到使用 R 的合适方法..... 8	
1.3.1 通过 CRAN 网站和 互联网获得帮助..... 9	
1.3.2 R 中的帮助命令..... 9	
1.3.3 解析 R 中的帮助项..... 12	
1.4 命令包..... 14	
1.4.1 标准命令包..... 14	
1.4.2 额外的包能为你完成 什么任务..... 15	
1.4.3 如何获得 R 命令的额外包..... 16	
1.4.4 包的运行和处理..... 18	
1.5 本章小结..... 20	
1.6 练习题..... 20	
1.7 本章概要..... 20	
第 2 章 出发: 开始熟悉 R..... 23	
2.1 一些简单的数学运算..... 23	
2.1.1 将 R 用作计算器..... 24	
2.1.2 保存计算结果..... 26	
2.2 将数据读取并放入 R 中..... 27	
2.2.1 使用组合命令制作数据..... 27	
2.2.2 使用 scan 命令制作数据..... 29	
2.2.3 读取较大的数据文件..... 34	
2.3 查看命名对象..... 38	
2.3.1 查看之前加载的命名对象..... 39	
2.3.2 从 R 中移除对象..... 41	
2.4 数据项的类型..... 41	
2.4.1 数字数据..... 41	
2.4.2 文本项..... 42	
2.4.3 在数字和文本数据之间 进行转换..... 42	
2.5 数据项的结构..... 44	
2.5.1 矢量项..... 44	
2.5.2 数据帧..... 45	
2.5.3 矩阵对象..... 45	
2.5.4 列表对象..... 46	
2.6 检查数据结构..... 46	
2.7 使用历史命令..... 47	
2.7.1 使用历史文件..... 48	
2.7.2 编辑历史文件..... 49	
2.8 在 R 中保存工作内容..... 50	
2.8.1 退出时保存工作区..... 50	
2.8.2 将数据文件保存到硬盘..... 50	
2.8.3 从硬盘读取数据文件..... 51	
2.8.4 将数据作为文本文件 保存到硬盘..... 53	
2.9 本章小结..... 57	
2.10 练习题..... 57	
2.11 本章概要..... 58	
第 3 章 出发: 处理对象..... 61	
3.1 对象处理..... 61	
3.1.1 矢量处理..... 62	
3.1.2 矩阵和数据帧的处理..... 66	

3.1.3	列表处理	71	4.3.4	验证表格对象	126
3.2	查看对象内部的对象	72	4.3.5	复杂(无格式)表格	126
3.2.1	查看复杂数据对象的内部	72	4.3.6	验证“无格式”表格对象	131
3.2.2	打开复杂数据对象	73	4.3.7	用于表格的汇总命令	132
3.2.3	快速查看复杂数据对象	75	4.3.8	交叉制表	134
3.2.4	查看及设置名称	77	4.4	本章小结	139
3.2.5	旋转数据表	80	4.5	练习题	139
3.3	构造数据对象	81	4.6	本章概要	140
3.3.1	制作列表	81	第 5 章	数据: 分布	143
3.3.2	制作数据帧	82	5.1	查看数据分布	143
3.3.3	制作矩阵对象	84	5.1.1	茎叶图	144
3.3.4	对数据帧和矩阵对象重新排序	86	5.1.2	直方图	146
3.4	数据对象的形式: 验证和转换	90	5.1.3	密度函数	149
3.4.1	验证你拥有的对象是什么类型	90	5.1.4	数据分布类型	152
3.4.2	将一种对象形式转换成另一种形式	91	5.1.5	用于正态性检验的夏皮洛-威尔克检验法	161
3.5	本章小结	98	5.1.6	柯尔莫哥洛夫-斯米尔诺夫检验	162
3.6	练习题	98	5.1.7	分位数-分位数图	163
3.7	本章概要	99	5.2	本章小结	167
第 4 章	数据: 描述统计和制表	101	5.3	练习题	167
4.1	汇总命令	101	5.4	本章概要	167
4.2	汇总样本	104	第 6 章	简单假设检验	169
4.2.1	矢量的汇总统计	104	6.1	使用学生 t 检验	169
4.2.2	累计统计	108	6.1.1	具有不等方差的 双样本 t 检验	170
4.2.3	数据帧的汇总统计	111	6.1.2	具有等方差的 双样本 t 检验	170
4.2.4	矩阵对象的汇总统计	113	6.1.3	单样本 t 检验	171
4.2.5	列表的汇总统计	114	6.1.4	使用定向假设	171
4.3	汇总表格	115	6.1.5	t 检验中的公式语法和 取子集样本	172
4.3.1	制作列联表	116			
4.3.2	选择表格对象的各个部分	123			
4.3.3	将对象转换成表格	125			

6.2 威尔克森 U 检验 (曼-惠特尼).....	175	7.4 折线图.....	213
6.2.1 双样本 U 检验.....	176	7.4.1 使用数值数据的折线图.....	213
6.2.2 单样本 U 检验.....	176	7.4.2 使用分类数据的折线图.....	214
6.2.3 使用定向假设.....	177	7.5 饼状图.....	217
6.2.4 U 检验中的公式语法和 取子集样本.....	177	7.6 克里夫兰点状图.....	220
6.3 配对 t 检验和 U 检验.....	180	7.7 柱状图.....	225
6.4 相关性和协方差.....	182	7.7.1 单类别柱状图.....	225
6.4.1 简单相关性.....	183	7.7.2 多类别柱状图.....	229
6.4.2 协方差.....	185	7.7.3 水平柱.....	232
6.4.3 相关性检验中的显著性 检验.....	185	7.7.4 汇总数据的柱状图.....	233
6.4.4 公式语法.....	185	7.8 将图形复制到其他应用程序.....	236
6.5 关联关系检验.....	188	7.8.1 使用 copy / paste 来 复制图形.....	236
6.5.1 多重分类: 卡方检验.....	189	7.8.2 将图形保存到硬盘.....	236
6.5.2 单一分类: 拟合优度 检验.....	191	7.9 本章小结.....	238
6.6 本章小结.....	195	7.10 练习题.....	238
6.7 练习题.....	195	7.11 本章概要.....	239
6.8 本章概要.....	196	第 8 章 公式符号与复杂统计.....	241
第 7 章 图形分析介绍.....	197	8.1 将公式语法用于基础 检验的示例.....	242
7.1 盒须图.....	197	8.2 图形中的公式符号.....	244
7.1.1 基础盒须图.....	198	8.3 方差分析(ANOVA).....	246
7.1.2 自定义盒须图.....	198	8.3.1 单向 ANOVA.....	246
7.1.3 水平盒须图.....	201	8.3.2 简单事后检验.....	248
7.2 散点图.....	204	8.3.3 从 aov()模型中提取 平均值.....	249
7.2.1 基础散点图.....	204	8.3.4 双向 ANOVA.....	250
7.2.2 添加坐标轴标签.....	205	8.3.5 提取平均值和汇总统计.....	258
7.2.3 绘图符号.....	205	8.3.6 相互作用图.....	260
7.2.4 设置坐标轴界限.....	206	8.3.7 更复杂的 ANOVA 模型.....	264
7.2.5 使用公式语法.....	207	8.3.8 aov()的其他选项.....	265
7.2.6 将最佳拟合线添加到 散点图.....	207	8.4 本章小结.....	267
7.3 配对图(多重相关图).....	210	8.5 练习题.....	268
		8.6 本章概要.....	268

第 9 章 操作数据并提取组成部分269	第 11 章 图形的更多相关介绍 333
9.1 创建用于复杂分析的数据..... 269	11.1 将元素添加到现有绘图..... 333
9.1.1 数据帧..... 270	11.1.1 误差线..... 334
9.1.2 矩阵对象..... 272	11.1.2 将图例添加到图形..... 338
9.1.3 创建并设置因子数据..... 274	11.1.3 将文本添加到图形..... 341
9.1.4 制作重复处理方式因子..... 277	11.1.4 将数据点添加到现有 图形..... 350
9.1.5 添加行或列..... 280	11.1.5 将各类线条添加到图形.. 353
9.2 数据汇总..... 285	11.2 矩阵图(一个图形上的 多个序列)..... 362
9.2.1 简单列与行汇总..... 285	11.3 一个窗口中的多个绘图..... 364
9.2.2 复杂汇总函数..... 287	11.3.1 将绘图窗口划分成 相等的区域..... 364
9.3 本章小结..... 296	11.3.2 将绘图窗口划分成 不相等的区域..... 367
9.4 练习题..... 296	11.4 图形导出..... 370
9.5 本章概要..... 297	11.4.1 使用复制粘贴移动图形.. 370
第 10 章 回归分析(线性建模)299	11.4.2 将图形保存到文件..... 371
10.1 简单线性回归分析..... 299	11.4.3 使用设备驱动程序将 图形保存到硬盘..... 371
10.1.1 线性模型结果对象..... 301	11.5 本章小结..... 374
10.1.2 lm()和 aov()之间的 相似性..... 305	11.6 练习题..... 375
10.2 多元回归分析..... 307	11.7 本章概要..... 375
10.2.1 公式和线性模型..... 307	第 12 章 编写你自己的脚本:
10.2.2 建模..... 308	开始编程 377
10.3 曲线回归分析..... 314	12.1 复制和粘贴脚本..... 377
10.3.1 对数回归分析..... 315	12.1.1 将自己的帮助文件制作成 普通文本..... 378
10.3.2 多项式回归分析..... 316	12.1.2 使用带有#字符的注解..... 378
10.4 绘制线性模型和曲线 拟合图..... 318	12.2 创建简单函数..... 379
10.4.1 最佳拟合线..... 319	12.2.1 单行函数..... 379
10.4.2 拟合线上的置信区间..... 322	12.2.2 在函数中使用默认值..... 380
10.5 回归分析模型汇总..... 326	12.2.3 带有多行命令的简单 自定义函数..... 380
10.5.1 诊断图..... 326	12.2.4 存储自定义函数..... 381
10.5.2 拟合汇总..... 328	
10.6 本章小结..... 329	
10.7 练习题..... 330	
10.8 本章概要..... 330	

12.3 制作源代码.....	383	12.4 本章小结.....	389
12.3.1 显示自定义函数和 脚本的结果.....	383	12.5 练习题.....	390
12.3.2 将消息作为脚本输出 的一部分进行显示.....	383	12.6 本章概要.....	390
		附录 A 练习题答案.....	393