

# Le logiciel R

## R软件教程与统计分析： 入门到精通

Pierre Lafaye de Micheaux  
Rémy Drouilhet  
Benoit Liquet 著  
潘东东 李启寨 唐年胜 译

□ 应用统

Le logiciel R

# R软件教程与统计分析： 入门到精通

Pierre Lafaye de Micheaux  
Rémy Drouilhet  
Benoit Liquet 著  
潘东东 李启寨 唐年胜 译

R RUANJI JIAOCHENG YU TONGJI FENXI  
RUMEN DAO JINGTONG

高等教育出版社·北京

图字号 : 01-2013-0442

Translation from French language edition: *Le logiciel R*  
by Pierre Lafaye de Micheaux, Rémy Drouilhet and Benoit Liquet  
Copyright © 2011 Springer Paris  
Springer Paris is a part of Springer Science+Business Media  
All Rights Reserved

### 图书在版编目 (C I P) 数据

R 软件教程与统计分析 : 入门到精通 / (法) 麦考斯 ,  
(法) 德鲁伊特 , (法) 利凯著 ; 潘东东 , 李启寨 , 唐年  
胜译 . -- 北京 : 高等教育出版社 , 2015.4

ISBN 978-7-12-041950-4

I . ① R . . II . ①麦… ②德… ③利… ④潘… ⑤李…  
⑥唐… III . ①统计分析—应用软件 IV . ① C819

中国版本图书馆 CIP 数据核字 (2015) 第 024176 号

策划编辑 王丽萍  
责任编辑 李华英  
责任校对 王雨

责任印制 刘思涵

封面设计 姜磊

版式设计 马敬茹

出版发行 高等教育出版社  
社址 北京市西城区德外大街4号  
邮政编码 100120  
印 刷 山东鸿君杰文化发展有限公司  
开 本 787 mm×1092 mm 1/16  
印 张 38.5  
字 数 880 千字  
购书热线 010-58581118

咨询电话 400-810-0598  
网 址 <http://www.hep.edu.cn>  
<http://www.hep.com.cn>  
网上订购 <http://www.landraco.com>  
<http://www.landraco.com.cn>  
版 次 2015 年 4 月第 1 版  
印 次 2015 年 4 月第 1 次印刷  
定 价 69.00 元

本书如有缺页、倒页、脱页等质量问题, 请到所购图书销售部门联系调换

版权所有 侵权必究

物 料 号 41950-00

献给开放源代码社区

献给所有曾经、现在及将来帮助我们的人们

献给我们  
在蒙彼利埃大学、格勒诺布尔大学、  
波尔多大学及公共卫生与流行病学研究所的同事们

# 前言

本书已被法国巴黎第九大学数学决策研究中心 (CEREMADE) 的罗宾·赖德 (Robin Ryder) 博士由法文翻译为英文。作为法语版的原作者，我们对赖德博士的此项翻译工作感到由衷的高兴并表示感谢！

本书是作者根据近年来在法国格勒诺布尔第二大学 (Institut Universitaire de Technologie Grenoble 2) 统计与商业智能系 (STID) 开设的一系列 R 课程的讲义整理而成。听课学生的好反映和高度评价是促成该书得以完成的原始动力。在此我们要对 STID 的学生们表示感谢，同时也要感谢我们的同事及朋友米歇尔·勒琼 (Michel Lejeune)，他建议我们去写一部手稿并投给 Springer 出版社公开出版。三位作者因为机缘走到了一起并共事多年，每位作者所具有专业知识的互补使得本书编写过程中碰到的巨大困难得以一一克服。最后，我们衷心地感谢我们的同事及朋友马修·杜布瓦 (Matthieu Dubois)，他是一名实验心理学领域的研究员，并对 R 和苹果系统 (Mac) 有很深的造诣，他也是该书几近定稿的法语版的第一位读者并提出了许多宝贵的意见。

本书的内容是在精心挑选和组织的基础上形成的，同时在力争尽可能详尽的基础上也希望让读者容易理解。它既可用作 R 初学者的课程教材或自学材料，对高级学者和研究人员也有一定的参考价值。三位作者为本书的撰写付出了大量的心血和劳动，特别在内容的选取和结构组织等方面更是凝聚了三位作者的大量心血。值得注意的是，本书的大部分内容对多种操作系统的用户都是适用的且有一定的帮助，但也有一小部分章节主要是针对微软 Windows 用户。此外，我们针对 Linux 或 Mac 用户适时给出的补充说明也是大有裨益的。

本书的每一章都以同样的架构来编排：开头列出该章的预备知识和内容提要；所有的理论概念都给出了例子来进行解释并提供代码供读者在计算机上自己动手练习；章末是一个回顾小节（列出了该章的要点）和一个习题小节（可用作考试习题）；并给出一个实践操作工作簿（只需利用前面章节的知识即可完成）。

全书按以下顺序结构来编排。第一部分在一个简短的介绍（目的是激发读者的兴趣）之后给出了几个数据集（贯穿全书用以说明 R 的使用）的详细描述。第二部分是介绍 R 的基本概念，包括数据的结构及其输入和输出、画图、编程及维护等内容，这部分可帮助读者对 R 有一个整体的了解。第三部分介绍 R 在一些数学或统计专业知识背景下的使用，建议读者先阅读第二部分的内容后再转至此部分。该部分覆盖了本科阶段大学三年级数学和统计专业课程的主要内容及其在 R 环境下的实现（比如，它覆盖了加拿大蒙特利尔大学统计学及精算科学专业的学士学位课程，以及法国大学理工学院 (IUT) 统计及商业智能专业的学士学位课程），具体包括矩阵运算、积分、优化、描述性

统计、随机模拟、置信区间及假设检验、简单及多重线性回归、方差分析等内容。第三部分的每一个统计专题章节都依赖于一个或几个实际数据集，它们都由法国波尔多大学公共卫生与流行病学研究所 (ISPED) 提供并在本书的开头进行了描述。所有的这些都将使本书变得更具体、更有吸引力。借此机会，我们向 ISPED 公共卫生学院的全体教学人员表示感谢！这些数据集以及几个专门为本书而编写的若干函数都包含在一个与本书配套的并被命名为 **TheRSoftware** 的 R 程序包中。我们也要感谢穆罕默德埃尔·梅斯利 (Mohamed El Methni) 和塔吉·巴鲁玛扎德 (Taghi Barumandzadeh) 提供给我们的方差分析 (ANOVA) 章节的相关材料，以及休伯特·雷蒙德 (Hubert Raymondaud) 为本书法语版提供的大量评论意见，这些都使本书的若干章节得到显著地改进和完善。

皮埃尔·拉菲米修斯 (Pierre Lafaye de Micheaux)

雷米·德鲁耶 (Rémy Drouilhet)

伯努瓦·理克威特 (Benoît Liquet)

## 译者的话

2011年初,国内统计学界迎来了统计学学科建设史上的一件大事,在国务院学位委员会审议通过的新版《学位授予和人才培养学科目录(2011)》中,增设了统计学一级学科。毫无疑问,这一重大变革和飞跃将对中国统计理论与实践的发展和进步产生巨大而深远的影响。这一切来之不易,它离不开近几十年来老一辈统计学家的无私奉献以及活跃在国内统计学界的一大批中青年统计学家的共同努力,特此向所有致力于中国统计教育事业发展与进步的前辈和同仁致以崇高的敬意!

R软件是伴随着统计学的发展而逐步兴起的一种计算语言。由于具有自由、免费、源代码开放等特点,自1993年诞生以来R得到了越来越多的统计学家和工程技术人员的使用。R内置多种统计学及数字分析功能,其功能还可以通过安装数以千计的并仍在不断更新的程序包来得到扩展,国际统计学术期刊上新近发表的论文中介绍的高引用统计方法大部分都可以从Comprehensive R Archive Network (CRAN)上自由下载,这极大地促进了统计学新思维、新方法的传播和应用。鉴于此,R被国外大量的学术和科研机构采用,欧美国家的许多著名高校的统计系或生物统计系都将R语言作为学生的必修或选修课程,其应用范围涵盖了几乎所有的统计分支方向和许多的交叉学科(比如:统计遗传学、生物信息学、实证金融学等)。近年来,国内也有不少的统计学家一直在推广R语言在国内高校与科研院所中的普及和应用,并已取得了广泛的影响和明显的成效。

2012年7月,加拿大蒙特利尔大学数学与统计系的Pierre Lafaye de Micheaux教授到访云南大学,他告诉我们他与在法国工作时的两位同事合著的法语版R软件教程已由Springer出版社于2010年12月出版发行,同时Springer出版社也即将推出其英文版本(具体时间为2014年1月)。这本书凝集了他们多年来从事R语言教学和科研的心血,他们非常期望在中国有该书的中文版面世以供感兴趣的读者参阅。由于我们近年来一直都在使用R语言进行统计模拟和数据分析等科学计算工作,也希望能有一本适合给统计学专业本科生和研究生开设R软件及统计计算课程的优秀教材或参考书。因此,我们满怀诚意地向Pierre Lafaye de Micheaux教授自荐来承担这项翻译工作并得到了该书三位原作者的欣然应允。

在阅读该专著的英文版后,我们发现,该书不仅内容丰富、结构合理、重点突出、层次分明、兼顾说理性和实用性,而且文字通俗易懂,是一本难得的介绍R软件与统计分析的好教材。我们希望本书中文版的出版能够进一步补充和丰富国内R语言教学的教材和参考资料,也希望该书的出版能使国内高校和科研院所以及实际应用者进一步推广对此软件的使用。需要说明的是,我

们遵从原著的风格,在**R**的英文版本的图形用户界面(GUI)下陈述各项操作和运算,因此,所有的截图或绘图中的说明文字都是英文(与英文版原著一致),我们希望并相信这一点不会给读者带来理解上的困难.事实上,在安装**R**的时候,可以将“语言”选择为简体中文,启动**R**后的控制台中的菜单键以及命令窗口中的说明文字都会是中文.

本书的翻译工作由唐年胜负责组织协调,是集体协作的结晶.其中潘东东翻译了前言、第1~6章、附录、索引和习题解答,唐年胜翻译了第7~10章,李启寨翻译了第11~15章.潘东东负责全书的统稿工作和初核工作,最后由李启寨和唐年胜对全书的内容进行核校.在本书的翻译过程中我们得到了许多人的热情帮助.在此,我们要特别感谢原著的三位作者(Pierre Lafaye de Micheaux教授、Rémy Drouilhet教授、Benoît Liquet教授)自始至终对本书翻译工作的关心和支持!特别感谢高等教育出版社的编辑们对本译著的出版所付出的辛勤劳动,尤其要感谢自然科学学术出版事业部王丽萍分社长的大力支持和帮助.另外,云南大学统计系硕士研究生黄艳彩、李壁涛和王淑贤帮助我们完成了部分章节的初译工作,特此表示衷心的感谢!

由于译者水平有限,本译著中肯定存在不少的缺点和不妥之处,恳请同行专家、学者以及广大读者批评指正.

译者

2015年2月

# 本书另外一种阅读顺序

我们已使用符号 † 来对某些难度较高或较低基础性的章节做了明确的标记, 在初次阅读的时候读者可以先跳过去而不会有碍对 R 的理解和掌握.

需要注意的是, 本书是为数学或统计学专业背景的学生量身打造的. 但是, 它对于那些更具“应用”专业背景的学生或学者也是适用的: 针对这部分读者, 我们提出如下所示的另外一种阅读顺序, 其中比较难的章节也应当先略过.

## 第一部分: R 的基础知识

- a) 基本概念与数据结构 (第 3 章)
- b) 输入、输出及生成数据 (第 4 章)
- c) 数据操作及函数 (第 5 章)
- d) R 及其帮助文件 (第 6 章)
- e) 绘制曲线与图像 (第 7 章)
- f) 管理会话 (第 9 章)

## 第二部分: 统计基本知识

- a) 利用 R 的特性来更好地理解随机变量、分布及模拟 (第 12 章)
- b) 描述性统计 (第 11 章)
- c) 置信区间与假设检验 (第 13 章)
- d) 简单及多重线性回归 (第 14 章)
- e) 方差分析基础 (第 15 章)

## 第三部分: 高级概念

- a) 数学基础: 矩阵运算、积分及最优化 (第 10 章)
- b) R 中编程 (第 8 章)

# 插入语的说明

我们对本书的结构形式进行了细致安排,期许其内容或信息更易被读者所理解和消化.在各技术要点处放置了不同的插入语以便引入一些重要信息使相关概念更易于理解,这些插入语的类型可由其边界处的图标来识别.

## 小窍门

给出一些关于当前讨论的主题的补充信息.



## 提醒

应该注意的要点.



## 注释

有关的建议及实用技巧.



## 另见

指向另外一个章节或一个网址.



## 高级用户

高级元素(可以先略过它们).



## Linux

针对 Linux 用户的信息.





Mac

针对苹果 (Mac) 用户的信息.

# 理论性习题和实践操作题的参考答案

理论性习题和实践操作题的参考答案在本书的网站上给出 (<http://www.biostatisticien.eu/springeR>).

比实践操作题规模更大的其他练习题将会陆续上传至该网站供用户下载使用.

## 字体规范

- 字母 **R** 指代 R 软件.
- 固定宽度字体 (**Verbatim** 环境) 表示 R 指令.
- **SMALL CAPS** (小型大写) 表示数据集, **sans serif** 字体表示包含这些数据集的文件名, 而且在本书中所有的文件名及文件夹名都采用此种字体.

# 数学符号

$\coloneqq$	左右两边不同的符号指代同一个对象
$\cup$	集合的并
$a \in A$	元素 $a$ 属于集合 $A$
$A \subset B$	集合 $A$ 包含于集合 $B$ 中
$A \supset B$	集合 $A$ 包含集合 $B$
$A \cap B$	集合 $A$ 与 $B$ 的交集
$A \cup B$	集合 $A$ 与 $B$ 的并集
$A \setminus B$	集合 $B$ 在集合 $A$ 中的补集
$(A \cup B) \setminus (A \cap B)$	集合 $A$ 与 $B$ 的对称差分
$f_i$	一种形态或模式的频率
$ x $	数 $x$ 的绝对值
$x!$	数 $x$ 的阶乘
$\binom{n}{p}$	二项系数: 从 $n$ 个元素中任意选择 $p$ 个的组合数
$\Gamma(\cdot)$	伽玛函数
$\gamma$	Euler 常数
$\psi(\cdot)$	双伽玛函数
$\pi$	圆周率 $\pi$
$\lambda$	标量
$\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$	矩阵
$I$	单位矩阵
$n \times p$	矩阵的维数
$\mathcal{A}^T$	矩阵 $\mathcal{A}$ 的转置
$\mathcal{B}^{-1}$	矩阵 $\mathcal{B}$ 的逆
$\bar{\mathcal{C}}$	复数矩阵 $\mathcal{C}$ 的共轭
$\mathbf{x} = (x_1, \dots, x_n)^T$	列向量
$\mathbf{x}^T$	向量 $\mathbf{x}$ 的转置
$\mathcal{A} \otimes \mathcal{B}$	矩阵 $\mathcal{A}$ 与矩阵 $\mathcal{B}$ 的 Kronecker 积
$vec(\mathcal{A})$	矩阵 $\mathcal{A}$ 按列拉直
$vech(\mathcal{A})$	矩阵 $\mathcal{A}$ 去掉主对角线上方的元素后按列拉直
$\mathcal{A}^*$	矩阵 $\mathcal{A}$ 的伴随矩阵(共轭转置)
$\mathcal{A}^{1/2}$	矩阵 $\mathcal{A}$ 的平方根
$\mathbb{1}_{[A]}(x)$	示性函数: 当 $x \in A$ 时取值为 1, 否则为 0

$[a, b]$	$a$ 与 $b$ 构成的闭区间内的值
$\det(\mathcal{A})$	矩阵 $\mathcal{A}$ 的行列式
$\Phi(\cdot)$	标准正态随机变量 $N(0, 1)$ 的累积分布函数
$\hat{\mathbf{X}}$	矩阵 $\mathbf{X}$ 按列中心化后形成的矩阵
$\mathbb{1}_n$	长度为 $n$ 的向量 $(1, \dots, 1)^T$
$X, Y$	非随机变量 (描述性统计量)
$N$	总体大小
$n$	样本大小 (样本量)
$m_e := q_{1/2}$	中位数
$PFC_X(\cdot)$	$X$ 的累积频数多边形的值
$\mu_x$	随机变量 $x$ 的期望值或描述性统计中的总体均值
$q_p$ 或 $x_p$	一个随机变量的 $p$ 阶分位数
$q_{1/4}, q_{3/4}$	第一个和第三个四分位数 (也记作 $q_1$ 和 $q_3$ )
$\sigma_{Pop}^2(\mathbf{x})$	总体方差 (描述性统计量)
$\sigma_{Pop}(\mathbf{x})$	总体标准误差 (描述性统计量)
$c_v$	总体变异系数 (描述性统计量)
$\gamma_1$	偏度系数
$\beta_2$	峰度系数
$\mu_3$	三阶中心距
$\mu_4$	四阶中心距
$\chi^2$	Pearson $\chi^2$ 统计量
$\Phi^2, V^2$	Cramer $\Phi^2$ 和 $V^2$
$\tau, \tau_b$	Kendall $\tau$ 和 $\tau_b$
$\rho$	理论上的 Pearson 相关系数
$\eta_{Y X}^2$	相关比率
$X, Y, \epsilon$	随机变量
$x_i, y_i, \epsilon_i$	随机变量 $X, Y, \epsilon$ 的实现值
$\mathbf{X}, \mathbf{Y}, \boldsymbol{\epsilon}$	随机向量
$\mathbf{X}_n$	样本 (随机的)
$\mathbf{x}_n$	样本 (观测的)
$\mathbf{X}$	随机矩阵
$\mathcal{L}$	一个随机变量的分布类
$N(0, 1)$	标准正态分布
$N(\mu, \sigma^2)$	均值为 $\mu$ , 方差为 $\sigma^2$ 的正态分布
$\mathcal{U}(a, b)$	区间 $[a, b]$ 上的均匀分布
$\mathcal{B}in(n, p)$	参数为 $n$ 和 $p$ 的二项分布
$\mathcal{E}(\lambda)$	参数为 $\lambda$ 的指数分布
$\mathcal{P}(\lambda)$	参数为 $\lambda$ 的 Poisson 分布
$\mathcal{T}(n)$	自由度为 $n$ 的 $t$ 分布
$\chi^2(n)$ 或 $\chi_n^2$	自由度为 $n$ 的 $\chi^2$ 分布
$\mathcal{F}(n, m)$	自由度为 $n$ 和 $m$ 的 $F$ 分布
$f_X(\cdot)$	随机变量 $X$ 的概率密度函数
$F_X(\cdot)$	随机变量 $X$ 的累积分布函数

$F_X^{-1}(\cdot)$	随机变量 $X$ 的逆累积分布函数
$\mu$	一个随机变量的期望
$\sigma^2$	一个随机变量的方差
$E(Y)$	随机变量 $Y$ 的理论期望
$Var(Y)$	随机变量 $Y$ 的理论方差
$\bar{X}_n$	样本 $\mathbf{X}_n = (X_1, \dots, X_n)^T$ 的经验均值 $\frac{1}{n} \sum_{i=1}^n X_i$ , 作为 $\mu_X$ 的估计值
$\bar{x}_n$	样本 $\mathbf{X}_n = (X_1, \dots, X_n)^T$ 的经验均值 $\frac{1}{n} \sum_{i=1}^n X_i$ 的实现值, 用以估计 $\mu_X$
$\xrightarrow{P}$	依概率收敛
$\hat{F}_n(\cdot) := \hat{F}_{\mathbf{X}_n}(\cdot)$	样本 $\mathbf{X}_n$ 的经验累积分布函数
$\theta$	未知参数 (有时将参数的未知真值表示为 $\theta^*$ )
$\hat{\theta}(X_1, \dots, X_n)$ 或 $\hat{\theta}$	未知参数 $\theta$ 基于样本 $\mathbf{X}_n = (X_1, \dots, X_n)^T$ 的理论估计量
$\hat{\theta}(x_1, \dots, x_n)$ 或 $\hat{\theta}$	未知参数 $\theta$ 基于样本观测值 $\mathbf{x}_n = (x_1, \dots, x_n)^T$ 的估计值
$B(\hat{\theta}(X_1, \dots, X_n); \theta)$	用 $\hat{\theta}(X_1, \dots, X_n)$ 来估计未知参数 $\theta$ 的偏差
$P[A]$	事件 $A$ 发生的概率
$\mathcal{V}(\theta; X_1, \dots, X_n)$	样本 $\mathbf{X}_n$ 的似然函数在 $\theta$ 处计算的值
$\mathbf{x}^* = (x_1^*, \dots, x_n^*)^T$	由观测样本 $\mathbf{x}_n = (x_1, \dots, x_n)^T$ 生成的 bootstrap 样本
$\hat{\sigma}$	$\sigma$ 的理论估计量
$\hat{\sigma}$	$\sigma$ 的实际估计值
$p$	比率
$\hat{p}$	比率 (或概率) 的理论估计量
$\hat{p}$	比率 (或概率) 的实际估计值
$\widehat{m}_e$	中位数的理论估计量
$\widehat{m}_e$	中位数的实际估计值
$M$	在蒙特卡罗 (Monte Carlo) 模拟中产生样本的迭代次数
$B$	产生 bootstrap 样本的数目
$B(\cdot, \cdot)$	贝塔函数
$I'_x(\cdot, \cdot)$	不完全贝塔函数的导数
$I(\cdot)$	修正 Bessel 函数
$I_\alpha(\cdot)$	修正 Bessel 函数族
$u_p$	标准正态分布 $N(0, 1)$ 的 $p$ 分位数
$t_p^n$	$T(n)$ 的 $p$ 分位数
$q_p^n$	$\chi^2(n)$ 的 $p$ 分位数
$f_p^{n,m}$	$\mathcal{F}(n, m)$ 的 $p$ 分位数
$CI_{1-\alpha}(\theta)$	$\theta$ 的水平为 $1 - \alpha$ 的随机置信区间
$ci_{1-\alpha}(\theta)$	$\theta$ 的水平为 $1 - \alpha$ 的实现置信区间
$1 - \alpha$	置信区间的水平
$(x_{(1)}, \dots, x_{(n)})$	按从小到大排序的观测样本

$\mathcal{H}_1$	假设检验中感兴趣的断言(对立假设)
$\mathcal{H}_0$	零(Null)假设,与 $\mathcal{H}_1$ 相对立
$\alpha$	显著性水平或假设检验中犯第一类错误的概率
$R$	随机的 Pearson 经验相关系数
$r$	实现的 Pearson 经验相关系数
$\beta_0, \beta_1$	简单线性回归模型中的未知系数
$\hat{\beta}_0, \hat{\beta}_1$	简单线性回归模型中未知系数的估计值
$\hat{\epsilon}_i$	简单线性回归模型中的观测残差
$\hat{y}_i$	简单线性回归模型中因变量的拟合值
$R^2$	回归模型中随机的判决系数
$r^2$	回归模型中判决系数的实现值
$R_a^2$	回归模型中随机的调整判决系数
$r_a^2$	回归模型中调整判决系数的实现值
$\hat{Y}^p$	回归模型中解释变量 $X$ 取一个新值时因变量 $Y$ 的预测值
$PI_{1-\alpha}(Y_0, x_0)$	因变量 $Y_0$ 在解释变量新值 $x_0$ 处的水平为 $1 - \alpha$ 的预测区间
$\boldsymbol{\beta} = (\beta_0, \dots, \beta_p)^T$	具有 $p$ 个解释变量的多重线性回归模型中 $p + 1$ 维的未知系数向量
$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$	多重线性回归模型中基于观测的解释变量矩阵 $\mathbf{X}$ 以及因变量的观测向量得到的未知参数向量 $\boldsymbol{\beta}$ 的估计量
$\hat{\boldsymbol{\beta}}$	$\boldsymbol{\beta}$ 的估计值
$VIF$	方差膨胀因子
$AIC$	Akaike 信息准则
$BIC$	Bayes 信息准则
$h_{ii}$	回归模型中第 $i$ 个观测的杠杆值
$t_i$	标准化残差
$t_i^*$	学生化残差
$\hat{\sigma}_{(-i)}$	剔除第 $i$ 个观测后 $\sigma$ 的估计值
$C_i$	Cook 距离
$\hat{y}_j^{(-i)}$	不使用第 $i$ 个观测得到的 $y_j$ 的预测值
$\hat{\beta}_j^{(-i)}$	不使用第 $i$ 个观测得到的 $\beta_j$ 的估计值
$I, J$	方差分析(ANOVA) 中某个因子的水平数
$\mu_{\bullet\bullet}$	方差分析中的总平均效应
$\mu_{i\bullet}$	方差分析中某个因子第 $i$ 水平的效应
$\mu_{\bullet j}$	方差分析中某个因子第 $j$ 水平的效应

# 目录

插图目录

表格目录

数学符号

<b>I 预备知识</b>	<b>1</b>
<b>1 R 软件概况</b>	<b>3</b>
1.1 R 软件基本情况介绍 . . . . .	3
1.1.1 起源 . . . . .	3
1.1.2 为什么要使用 R? . . . . .	3
1.2 R 与统计学 . . . . .	4
1.3 R 与绘图 . . . . .	5
1.4 R 的图形用户界面 (GUI) . . . . .	6
1.5 R 的第一步 . . . . .	7
1.5.1 使用 RCommander . . . . .	7
1.5.1.1 启动 RCommander . . . . .	7
1.5.1.2 使用 RCommander 来处理数据 . . . . .	9
1.5.1.3 使用 RCommander 完成一些统计分析任务 . . . . .	13
1.5.1.4 给 RCommander 界面添加功能 . . . . .	17
1.5.2 通过控制台 (console) 来使用 R . . . . .	18
1.5.2.1 R 在一个实例中展现出的优势 . . . . .	18
1.5.2.2 键入一些指令来对 R 的语法做一个简介 . . . . .	21
<b>2 若干数据集和研究问题</b>	<b>27</b>
2.1 儿童的体重指数 . . . . .	27
2.2 婴儿出生时的体重 . . . . .	28
2.3 内膜 – 中膜厚度 . . . . .	28
2.4 老年人的饮食及营养 . . . . .	29