

 现代统计学系列丛书


# 统计计算

李东风

---

**Statistical Computing**

高等教育出版社

 现代统计学系列丛书

# 统计计算

Statistical Computing



李东风

高等教育出版社·北京

## 内容提要

本书讲述统计计算的基本概念和统计计算中最常用的算法,内容涵盖了误差、描述统计、随机数产生、随机模拟、逼近、插值、数值积分与数值微分、矩阵计算、最优化与方程求根等各个方面。本书的讲解比较系统,提供了大量的例题和习题,使用应用广泛的 R 语言进行算法描述与编程,可以用作统计学专业本科生“统计计算”课程的教材,也可以作为统计学以及其他专业的本科生、研究生和研究人员关于统计计算算法的参考书。

## 图书在版编目(CIP)数据

统计计算 / 李东风编. --北京: 高等教育出版社, 2017. 3

(现代统计学系列丛书)

ISBN 978-7-04-047070-3

I. ①统… II. ①李… III. ①概率统计计算法-高等学校-教材 IV. ①O242.28

中国版本图书馆 CIP 数据核字(2016)第 321525 号

|          |          |          |          |
|----------|----------|----------|----------|
| 策划编辑 张晓丽 | 责任编辑 张晓丽 | 封面设计 赵阳  | 版式设计 杜微言 |
| 插图绘制 杜晓丹 | 责任校对 杨凤玲 | 责任印制 赵义民 |          |

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 中国农业出版社印刷厂

开 本 787mm×960mm 1/16  
印 张 20  
字 数 360 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.hepmall.com.cn>  
<http://www.hepmall.com>  
<http://www.hepmall.cn>  
版 次 2017 年 3 月第 1 版  
印 次 2017 年 3 月第 1 次印刷  
定 价 37.30 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 47070-00

# 现代统计学系列丛书编委会

(按姓氏笔画排序)

主 编: 方开泰

副主编: 史宁中 何书元 陈 敏 耿 直

编 委: 马 洪 方开泰 史宁中 杨 虎 何书元 何晓群

张爱军 张崇岐 陈 敏 郑 明 赵彦云 耿 直

曾五一 缪柏其

# 总 序

统计学是一门收集、整理和分析数据的科学和艺术。这里的“数据”泛指“信息的载体”，涵盖了大千世界中的文本、图像、视频、时空数据、基因数据等。统计学是一个独立的学科，在历史上曾隶属于数学，但统计学与数学有着本质的区别，因此统计学教育有其自身的特点和要求，这些特点表现为：(1) 统计学研究的是随机现象，而数学研究的是确定性的规律；(2) 统计学是一门应用性很强的学科，许多概念和原理来自于实际的需要，不是数理逻辑的产物；(3) 数据在统计学中扮演了重要的角色。目前，统计学已被列为一级学科。

在过去的 30 年中，随着生命科学、信息科学、物质科学、资源环境、认知科学、工程技术、经济金融和人文科学等众多学科的发展，产生了许多新的统计学分支，如风险管理、数据挖掘、基因芯片分析等。此外，计算机及其有关软件在统计教育和应用中扮演了越来越重要的角色，它们提供了越来越多的图形表达和分析的方法，使得许多原来教科书中重要的内容，现在已变得无足轻重。统计教育必须要改革才能适应高速发展的形势。

大学的统计教育可分为两大类，一类是非统计学专业的课程，另一类是统计学专业的教学设计。非统计学专业的学生学习统计的目的是为了应用，在大学阶段，课程不多，主要是学习基础的统计概念和方法，学会使用统计软件，培养其解决实际问题的能力。统计学专业的课程设置十分重要，应向国际靠拢，对教师队伍的要求也较高。虽然这两类学生的教育有很多共同点，但在课程设置中必须加以区分。

我国的统计教育在过去受苏联的影响很深，把统计学作为数学的一个分支，在内容上偏理论，少应用，过于强调概率论在统计中的作用。统计学是一门应用性很强的学科，应从实际问题、从数据出发，通过统计的工具来揭示数据内部的规律。用“建模”的思路来教统计，使学生能更加容易理解统计的概念和方法，知道如何将实际问题抽象为统计模型，反过来又指导实践。对非统计学专业的学生，要强调统计的应用。学生要能熟练地使用至少一个统计软件包。对于统计学专业的学生，要培养学生对实际问题的建模能力。有些实际问题可直接应用现有的统计方法来解决，如问卷调查的统计分析。有些问题在初次接触时并不像一个统计问题，必须有坚实的统计基础和对实际问题的洞察力，才

能从中发掘出统计模型。要培养学生的这种能力及统计思想(统计思想是统计文化的一部分,是用统计学的逻辑思考问题)。教师在授课中要结合较多的应用例子,要求学生做案例研究,鼓励学生参加建模比赛,参加企业的实际项目。

为满足我国统计教育发展的需要,我们计划编写一套面向高校本科生、特别是一般院校,适用于统计学专业和非统计学专业的系列教材。系列教材的编写宗旨是:突出教学内容的现代化,重视统计思想的介绍,适应现代统计教育的特点及时代发展的新要求;以统计软件为支撑,注重统计知识的应用;内容简明扼要,生动活泼,通俗易懂。编写原则为:(1)从数据出发,不是从假设、定理出发;(2)从归纳出发,不是从演绎出发;(3)强调案例分析;(4)重统计思想的阐述,弱化数学证明的推导。系列教材分为两个方向,一个面对统计学专业,另一个面对非统计学专业和应用统计工作者。

系列教材是适应形势的要求,由高等教育出版社邀请专家组成“现代统计学系列丛书编委会”负责选题、审稿,由高等教育出版社出版。

以上是我们编写这套教材的背景和理念,希望得到读者的支持,特别是高校领导和教学一线教师的支持。我们希望使用这套教材的师生和读者多提宝贵意见,使教材不断完善。

现代统计学系列丛书编委会



扫描二维码,获取更多丛书信息

# 前 言

---

统计学是研究如何有效地搜集、整理和分析带有随机性的数据,以及对所考察的问题作出推断或预测,直至为采取一定的决策和行动提供依据和建议的学科。统计学用于实际数据的建模和分析,需要把统计建模过程变成适当的计算机算法,并将其用编程进行实现。这样,学习统计学专业的学生,不仅需要学习高等数学、概率论、数理统计、回归分析、多元统计分析、时间序列分析、生存分析等许多数学和概率统计专业课程,还需要了解统计学有关的计算机算法的理论和实现细节,这样才能了解统计建模计算中的各种问题,不至于滥用、错用统计方法,在自己研究新的统计模型以及进行数据分析建模时才能开发出正确、可靠、高效的算法和程序。

本书内容主要是传统的统计计算基本概念与算法,如误差分析、矩阵计算、最优化、随机数生成等。在统计学相关的算法中,有很多算法可以归入所谓“计算统计”范畴,这些算法主要是借助于现代计算机的强大性能进行统计推断,比如 Bootstrap 法、MCMC 方法、序贯重要抽样方法等。本书对这类方法也进行了少量的论述。

本书第一章讲述了算法的基本概念和描述统计的一些方法,并在 1.1.3 节给出了全书比较详细的内容提要。第二章主要介绍了随机数生成算法,其中包括随机数的产生,随机向量和随机过程的产生。第三章叙述了随机模拟方法如随机模拟积分、方差缩减方法等。第四章描述了函数逼近、插值、数值积分、数值微分等近似计算方法。第五章阐释了矩阵计算的多种方法。第六章讲述了最优化与方程求根问题。

本书使用 R 软件作为算法的实现和描述语言。选用 R 软件的原因是 R 软件特别适用于统计计算方法的开发,统计学家普遍地使用 R 软件作为统计计算和统计算法的开发语言,而且 R 软件也很容易自学。教材的 1.2 节给出了一个简短的 R 软件编程入门讲解,这一部分在授课时可以让学生自学。本书作为教材使用时,教师也可以选用其他的计算机语言,如 JAVA、C、C++、FORTRAN、Python、Matlab 等。

本书包含大量的例题和习题,其中大部分需要编程实现。读者应该多动手编写程序,并与标准的算法程序进行比较。R 软件中包含了许多种标准的算法,可以用来检查读者的编程计算结果是否正确。

本书可以作为高校统计学本科专业统计计算课程的教材,也可以作为统计学以及其他专业的本科生、研究生和研究人员关于统计计算算法的参考书。在作为本科生教材使用时,目录中标有星号的部分可以有选择地讲授。

耿直教授、何书元教授对本书的编写给予了热情的关心与指导,高等教育出版社李蕊编辑、张晓丽编辑、安琪在成书过程中付出了大量的心血和汗水,作者在此表示衷心的感谢。书中使用了其他教材的一些例子和习题,没有一一指明出处,在此感谢原作者。

李东风

2016年7月3日



# 目 录

|                           |    |
|---------------------------|----|
| 第一章 绪论 .....              | 1  |
| 1.1 介绍 .....              | 1  |
| 1.1.1 统计计算的范畴 .....       | 1  |
| 1.1.2 算法和计算机语言 .....      | 2  |
| 1.1.3 内容概述 .....          | 4  |
| 1.2 R 软件基础 .....          | 5  |
| 1.2.1 向量 .....            | 6  |
| 1.2.2 向量运算 .....          | 9  |
| 1.2.3 矩阵 .....            | 11 |
| 1.2.4 分支和循环 .....         | 14 |
| 1.2.5 函数 .....            | 17 |
| 1.3 误差 .....              | 18 |
| 1.3.1 误差的种类 .....         | 18 |
| 1.3.2 数值计算误差 .....        | 19 |
| 1.3.3 随机误差的度量 .....       | 23 |
| 1.3.4 问题的适定性与算法稳定性 .....  | 25 |
| 1.4 描述统计量 .....           | 26 |
| 1.4.1 总体和样本 .....         | 26 |
| 1.4.2 样本的描述统计量 .....      | 28 |
| 1.5 统计图形 .....            | 33 |
| 1.5.1 直方图 .....           | 33 |
| 1.5.2 核密度估计 .....         | 37 |
| 1.5.3 盒形图 .....           | 40 |
| 1.5.4 茎叶图 .....           | 41 |
| 1.5.5 正态 QQ 图和正态概率图 ..... | 42 |
| 1.5.6 散点图和曲线图 .....       | 44 |
| 1.5.7 三维图 .....           | 47 |
| 习题一 .....                 | 50 |
| 第二章 随机数 .....             | 54 |
| 2.1 均匀分布随机数的产生 .....      | 54 |
| 2.1.1 线性同余发生器 (LCG) ..... | 55 |
| 2.1.2 FSR 发生器 * .....     | 63 |

|            |                       |            |
|------------|-----------------------|------------|
| 2.1.3      | 组合发生器法                | 64         |
| 2.1.4      | 随机数的检验 *              | 64         |
| <b>2.2</b> | <b>非均匀分布随机数的产生</b>    | <b>66</b>  |
| 2.2.1      | 逆变换法                  | 66         |
| 2.2.2      | 离散型随机数                | 67         |
| 2.2.3      | 用变换方法生成连续型分布的随机数      | 74         |
| 2.2.4      | 舍选法                   | 78         |
| 2.2.5      | 复合法                   | 85         |
| <b>2.3</b> | <b>随机向量和随机过程的生成</b>   | <b>89</b>  |
| 2.3.1      | 条件分布法                 | 89         |
| 2.3.2      | 多元正态分布模拟              | 90         |
| 2.3.3      | 用 copula 描述多元分布 *     | 90         |
| 2.3.4      | Poisson 过程模拟 *        | 91         |
| 2.3.5      | 平稳时间序列模拟 *            | 92         |
|            | 习题二                   | 93         |
| <b>第三章</b> | <b>随机模拟</b>           | <b>99</b>  |
| <b>3.1</b> | <b>概述</b>             | <b>99</b>  |
| <b>3.2</b> | <b>随机模拟积分</b>         | <b>102</b> |
| 3.2.1      | 随机投点法                 | 102        |
| 3.2.2      | 平均值法                  | 104        |
| 3.2.3      | 高维定积分                 | 106        |
| 3.2.4      | 重要抽样法                 | 110        |
| 3.2.5      | 分层抽样法                 | 117        |
| <b>3.3</b> | <b>方差缩减方法</b>         | <b>121</b> |
| 3.3.1      | 控制变量法                 | 121        |
| 3.3.2      | 对立变量法                 | 124        |
| 3.3.3      | 条件期望法                 | 127        |
| 3.3.4      | 随机数复用                 | 130        |
| <b>3.4</b> | <b>随机服务系统模拟 *</b>     | <b>131</b> |
| <b>3.5</b> | <b>统计研究与随机模拟</b>      | <b>136</b> |
| <b>3.6</b> | <b>Bootstrap 方法 *</b> | <b>140</b> |
| 3.6.1      | 标准误差                  | 140        |
| 3.6.2      | Bootstrap 方法的引入       | 141        |
| 3.6.3      | Bootstrap 偏差校正        | 144        |
| 3.6.4      | Bootstrap 置信区间        | 145        |

|          |                       |     |
|----------|-----------------------|-----|
| 3.7      | MCMC                  | 146 |
| 3.7.1    | Markov 链和 MCMC 介绍     | 146 |
| 3.7.2    | Metropolis-Hasting 抽样 | 148 |
| 3.7.3    | Gibbs 抽样              | 152 |
| 3.7.4    | MCMC 计算软件 *           | 155 |
| 3.8      | 序贯重要抽样 *              | 159 |
| 3.8.1    | 非线性滤波平滑               | 160 |
| 3.8.2    | 再抽样                   | 162 |
|          | 习题三                   | 164 |
| 第四章 近似计算 |                       | 172 |
| 4.1      | 函数逼近 *                | 172 |
| 4.1.1    | 多项式逼近                 | 172 |
| 4.1.2    | 连分式逼近                 | 176 |
| 4.1.3    | 逼近技巧                  | 179 |
| 4.2      | 插值                    | 180 |
| 4.2.1    | 多项式插值                 | 180 |
| 4.2.2    | 样条插值介绍                | 188 |
| 4.3      | 数值积分和数值微分             | 190 |
| 4.3.1    | 数值积分的用途               | 190 |
| 4.3.2    | 一维数值积分                | 191 |
| 4.3.3    | 多维数值积分                | 199 |
| 4.3.4    | 数值微分                  | 199 |
|          | 习题四                   | 202 |
| 第五章 矩阵计算 |                       | 204 |
| 5.1      | 介绍                    | 204 |
| 5.2      | 线性方程组求解               | 207 |
| 5.2.1    | 三角形线性方程组求解            | 207 |
| 5.2.2    | Gauss 消元法和 LU 分解      | 207 |
| 5.2.3    | Cholesky 分解           | 211 |
| 5.2.4    | 线性方程组求解的稳定性           | 214 |
| 5.3      | 线性方程组的特殊解法 *          | 217 |
| 5.3.1    | 带状矩阵                  | 217 |
| 5.3.2    | Toeplitz 矩阵           | 219 |
| 5.3.3    | 稀疏系数矩阵方程组求解           | 221 |
| 5.3.4    | 用迭代法求解线性方程组           | 221 |

|              |                       |     |
|--------------|-----------------------|-----|
| 5.4          | QR 分解                 | 223 |
| 5.4.1        | Gram-Schmidt 正交化方法    | 224 |
| 5.4.2        | Householder 变换 *      | 226 |
| 5.4.3        | Givens 变换 *           | 227 |
| 5.5          | 特征值、奇异值               | 229 |
| 5.5.1        | 定义                    | 229 |
| 5.5.2        | 对称阵特征值分解的 Jacobi 算法 * | 230 |
| 5.5.3        | 用 QR 分解方法求对称矩阵特征值分解 * | 233 |
| 5.5.4        | 奇异值分解的计算 *            | 233 |
| 5.6          | 广义逆矩阵                 | 235 |
|              | 习题五                   | 238 |
| 第六章 最优化与方程求根 |                       | 242 |
| 6.1          | 最优化问题和求解              | 242 |
| 6.1.1        | 优化问题的类型               | 243 |
| 6.1.2        | 一元函数的极值               | 244 |
| 6.1.3        | 凸函数 *                 | 246 |
| 6.1.4        | 无约束极值点的条件             | 250 |
| 6.1.5        | 约束极值点的条件 *            | 252 |
| 6.1.6        | 迭代收敛                  | 257 |
| 6.2          | 一维搜索与求根               | 257 |
| 6.2.1        | 二分法求根                 | 258 |
| 6.2.2        | Newton 法              | 260 |
| 6.2.3        | 一维搜索的区间 *             | 262 |
| 6.2.4        | 0.618 法               | 263 |
| 6.2.5        | 抛物线法 *                | 264 |
| 6.2.6        | Wolf(沃尔夫) 准则 *        | 265 |
| 6.3          | 无约束优化方法               | 267 |
| 6.3.1        | 分块松弛法                 | 267 |
| 6.3.2        | 最速下降法                 | 267 |
| 6.3.3        | Newton 法              | 269 |
| 6.3.4        | 拟 Newton 法            | 270 |
| 6.3.5        | Nelder-Mead 方法 *      | 272 |
| 6.4          | 约束优化方法 *              | 274 |
| 6.4.1        | 约束的化简                 | 274 |
| 6.4.2        | 仅含线性等式约束的情形           | 275 |
| 6.4.3        | 线性约束最优化方法             | 276 |
| 6.4.4        | 二次规划问题                | 278 |
| 6.4.5        | 非线性约束优化问题             | 282 |

---

|                 |     |
|-----------------|-----|
| 6.5 统计计算中的优化问题* | 285 |
| 6.5.1 最大似然估计    | 285 |
| 6.5.2 EM 算法     | 289 |
| 6.5.3 非线性回归     | 294 |
| 习题六             | 296 |
| 参考文献            | 301 |

### 1.1 介绍

#### 1.1.1 统计计算的范畴

统计计算是现代统计的重要组成部分。从上个世纪三四十年代起,数理统计的理论和方法得到跨越式的发展,统计推断理论、回归分析、试验设计、方差分析、序贯分析、时间序列分析、随机过程等理论和方法在这个时期逐渐成熟。但是,直到上个世纪八十年代,统计学作为一门学科才真正得到了广泛的普及,其应用深入到了我们的学术研究和社会生活的每一个方面,只要需要分析数据的地方就需要用到统计学。这种普及很大程度上要归功于电子信息技术的高速发展。

统计计算就是统计方法和实际计算的结合。统计计算有以下两个方面的内容:

- 把统计方法变成可靠、高效的算法,并编程实现。这是经典的统计计算要解决的问题,比如计算分布函数值、分位数函数值、计算线性回归参数估计和检验、求解最大似然估计等。
- 借助于现代计算机的强大处理能力,发展新的统计方法。这是计算技术对统计学的贡献,比如用随机模拟方法求解 Bayes 模型、Bootstrap 置信区间,等等。这个方面有时被称为“计算统计”(computational statistics)。

第二个方面的内容包括计算密集的统计方法及相应的理论工作。随机模拟方法是其中一个重要内容。随机模拟的基本思想是在计算机上模拟生成一个统计问题的数据并进行大量的重复,这样相当于获得了此问题的海量的样本。如果我们的目的是评估某种建模方法,我们可以对每个样本建模,最后从所有建模的结果评估这种方法的性能;可以从理论模型产生海量模拟样本后对此模型

进行理论推断,如 Monte Carlo 检验;可以对观测数据进行多次重复再抽样生成许多新样本,如 Bootstrap 方法. 在 Bayes 统计框架下,我们可以从先验分布抽样,并按照模型产生大量样本并结合观测数据计算其似然估计,从而获得参数后验分布的大量样本,以此进行贝叶斯推断. 借助于随机模拟,我们可以试验各种各样的模型与方法,发现表现优秀的模型和方法后再进行深入研究.

另外,各行各业中数据收集越来越广泛,在迅速积累的海量数据中包含了许多以前无法触摸的现象和规律,对海量数据进行探索性分析,从海量数据中发现规律,已经成为统计学和信息科学的热门研究方法,通常称为机器学习、数据挖掘等. 这也是统计计算第二个方面的重要内容.

现在已经有了许多专用的统计软件,比如 R, SAS 等,为我们平常遇到的许多问题提供了现成的解决办法,那么,我们为什么还需要学习统计计算呢?

我们遇到的具体应用问题常常是没有现成的方法可以套用的. 即使有现成的统计软件可用,我们也需要理解这些软件的工作原理以避免错误使用;在遇到新问题时,需要能够修改原有代码或编写新代码,把计算工具结合在一起解决自己的数据分析问题,而不是修改自己的问题以适应现成的软件.

### 1.1.2 算法和计算机语言

算法是完成某项任务步骤的精确的描述. 比如,制作水果沙拉的一种算法为

1. 准备一个盘子;
2. 取一个苹果削皮后把果肉切成小块放入盘中;
3. 取一个香蕉去皮后切成小块放入盘中;
4. 取一调羹沙拉酱放入盘中并与果料一起搅拌均匀.

当然,算法主要针对在计算机上的计算而设计. 好的算法应该满足如下要求:

- 结果正确,即要求算法的最后结果是我们问题的正确解,最好能够验证结果的正确性.
- 指令可行,即指令含义明确无歧义,指令可以执行并且在现有的计算条件下算法能在允许的时间内计算结束.
- 高效,尽可能少地消耗时间和内存、外存资源.

计算机由 CPU、内存、大容量外存、输入 / 输出装置等硬件构成,但是依赖于软件完成任务. 软件在执行时是一系列机器指令进行取值、存储、加法等操作. 操作系统是最基本的计算机软件,用来管理内存位置、软件指令、输入 / 输出和其他程序的运行调度.

计算机软件可以用于完成特定任务,如字处理、记账,也可以适用于较宽的范围,比如电子表格软件可以用来记账、试算、作图,而计算机语言则是用来编制新的软件的工具.

计算机语言根据其运行方式可以分为解释型和编译型两种,解释型语言逐句解释程序并逐句执行,编译型语言把整个程序编译为二进制代码后再执行. 按照计算机语言的抽象程度区分,包括二进制形式的机器语言,仅能用在特定的硬件中;汇编语言,用与 CPU 指令对应的命令编写,主要用于底层功能;面向细节的通用语言,如 Pascal, C, Fortran, Lisp, Cobol, C++, Java 等,优点是通用性和可复用性.

更高级的计算机语言有 R, Matlab 等,提供了包括向量、矩阵等高级数据类型,代码与统计的数学公式相似,直接支持求和、向量、矩阵等运算,易写易读,用户不用自己实现如解线性方程、求特征根这样的基础操作,但是这样的语言一般是解释执行的,执行效率难以改进,不利于使用循环或迭代算法. 本书使用 R 作为配套的编程语言.

R 软件是一个统计计算软件,同时也是一种计算机编程语言,与 S 语言基本兼容. S 语言是 Rick Becker, John Chambers 等人在贝尔实验室开发的一个进行数据分析和交互作图的计算机语言,可以使用向量、矩阵、对象等数据进行编程,功能强大,程序简单. R 是 GPL 授权的自由软件,最初由新西兰 Auckland 大学的 Ross Ihaka 和 Robert Gentleman 于 1997 年发布,现在由 R 核心团队开发,全世界的用户贡献了数千个软件包,功能涵盖了绝大部分的经典和现代统计方法,是世界上许多顶尖的统计学家进行统计研究和发表算法的工具(见 R 的网站: <http://www.r-project.org/>). 我们将讲授 R 的基本使用,在算法示例和习题中使用 R 作为编程语言,并在讲到具体统计计算方法时提及 R 中有关的函数.

本书目的主要是要求学生掌握统计计算方法、理解统计计算思想,但是这些算法的正确、高效实现也是很重要的. 我们设置了足够的习题让学生通过实际编程了解算法实现中的各种问题.

在进行程序设计时,应注意以下几点:

- 要对程序抱有怀疑态度. 逐步验证每一个模块.



- 大的任务要分解为小的模块. 对每个模块进行详细测试. 像 S 这样的高级语言的优点是许多模块已经由系统本身提供了, 如矩阵计算、函数优化等. 但即使这样也应进行测试, 因为系统提供的功能有可能恰恰不适用于你的特殊情况.
- 测试为一系列测试问题, 涵盖应用的不同情况. 最开始用最简单的测试. 测试也应该包括超出范围的问题, 这时程序应该能正确地判别错误输入.
- 好的程序不应该是用了很多高明的技巧以至于别人很难看懂, 而是越直接越好, 这样只要出错误就是明显的错误.
- 要有程序文档.
- 只要对程序的每一模块都进行了详细验证和检查, 就可以确信程序的正确性.

### 1.1.3 内容概述

本书包括基本的统计计算方法, 如计算分布函数、分位数函数的一般方法, 矩阵计算方法、最优化方法、随机数生成算法. 另外还用较大篇幅讲述了随机模拟方法, 包括随机模拟的基本思想、改进精度的方法及重要应用. 最后, 还介绍了完全由计算方法发展出来的统计方法, 如 Bootstrap, EM 算法, MCMC 方法等.

第一章后面部分先介绍 R 语言的基础, 由于篇幅所限, 仅包括数据类型、程序结构、函数的基础介绍, 进一步的用法还要参考其他教材和 R 用户手册. 然后, 讲解误差的来源和分类以及避免和减少误差的方法, 这对我们了解算法的局限并在算法实现时避免产生有缺陷的算法有重要意义. 最后, 介绍描述统计量计算和简单的统计图形用法.

第二章为随机数产生与检验, 包括均匀随机数的产生方法和检验方法, 非均匀随机数的各种生成方法, 包括函数变换、舍选抽样、重要抽样等. 关于随机向量和随机过程的产生方法也进行了简单介绍.

第三章为随机模拟方法. 首先, 用随机模拟积分作为例子, 讲解随机模拟方法的基本思想, 包括减小随机模拟误差的技术. 后面介绍了离散随机事件模拟中随机服务系统模拟问题. 随机模拟方法对新统计方法的研究比较具有广泛应用, 本章用一个例子演示了随机模拟在统计方法研究中的用法. 然后, 本章讲述了 Bootstrap 方法, 这是完全利用随机模拟方法解决统计推断问题的一个具体示例. MCMC 是现代统计计算的重要工具, 尤其是在 Bayes 建模中起到关键