

# 《R语言统计入门(第2版)》

## 图书基本信息

书名：《R语言统计入门(第2版)》

13位ISBN编号：9787115348634

出版时间：2014-5-9

作者：Peter Dalgaard

译者：郝智恒,何通,邓一硕,刘旭华

版权说明：本站所提供下载的PDF图书仅提供预览和简介以及在线试读，请支持正版图书。

更多资源请访问：[www.tushu000.com](http://www.tushu000.com)

# 《R语言统计入门(第2版)》

## 内容概要

这《R语言统计入门(第2版)》以最恰当的方式向初学者介绍了R的全貌，内容涵盖基本的R语言编程方法、基本数据处理和一些高级数据操作的技巧，有助于读者理解R向量化编程的特点。此外，作者在本书中还详细描述了包含回归分析、假设检验、广义线性模型、非线性拟合等常用统计方法的原理。虽然本书以实际案例解析居多，但是并非不重视理论，作者恰当而到位地描述了理论方面的内容，既不晦涩，也非浅白，而是向读者打开了一扇窗。作者希望这本书可以作为一道“开胃菜”引导更多的人投入到对统计和R的研究之中。

# 《R语言统计入门(第2版)》

## 作者简介

Peter Dalgaard, 哥本哈根大学生物统计系的副教授, 有着丰富的教学经验。自1997年起, 便加入了R核心小组至今。译者简介: 郝智恒, 兰州人, 南开大学数理统计专业毕业, 目前在淘宝做数据挖掘工程师。推崇开源, 认为唯有共享和交流才能让这个世界变好, 固执的认为一切技术上的封闭和试图垄断都是纸老虎。何通, 生长于桂林, 求学于广州, 喜统计, 好编程, 结交众友, 不亦乐乎。邓一硕, 毕业于中央财经大学统计学院统计专业, 现效力于首钢总公司计财部。国内最大的统计学专业社区“统计之都”理事人员, 感兴趣的领域是时间序列分析及数据挖掘技术在金融投资分析中的应用。个人网站为<http://yishuo.org>, 微博为<http://weibo.com/dengyishuo>。刘旭华, 现为中国农业大学理学院应用数学系副教授, 北京理工大学博士, 美国北卡莱罗纳大学教堂山分校 (University of North Carolina at Chapel Hill) 访问学者, 主要从事数理统计方向的研究工作。主持和参与多项国家自然科学基金及其他科研基金项目, 发表科研论文多篇。

## 书籍目录

### 第1章 基础知识 1

#### 1.1 初始步骤 1

##### 1.1.1 大型计算器 2

##### 1.1.2 赋值 3

##### 1.1.3 向量运算 4

##### 1.1.4 标准过程 5

##### 1.1.5 作图 6

#### 1.2 r语言基础 8

##### 1.2.1 表达式和对象 8

##### 1.2.2 函数和参数 9

##### 1.2.3 向量 10

##### 1.2.4 引用和转义序列 10

##### 1.2.5 缺失值 11

##### 1.2.6 生成向量的函数 11

##### 1.2.7 矩阵和数组 13

##### 1.2.8 因子 15

##### 1.2.9 列表 16

##### 1.2.10 数据框 17

##### 1.2.11 索引 17

##### 1.2.12 条件选择 18

##### 1.2.13 数据框的索引 19

##### 1.2.14 分组数据和数据框 20

##### 1.2.15 隐式循环 21

##### 1.2.16 排序 23

#### 1.3 练习题 24

### 第2章 r语言环境 25

#### 2.1 会话管理 25

##### 2.1.1 工作空间窗口 25

##### 2.1.2 文本输出 26

##### 2.1.3 脚本 27

##### 2.1.4 获取帮助 27

##### 2.1.5 包 28

##### 2.1.6 内置数据 29

##### 2.1.7 attach和detach 29

##### 2.1.8 subset, transform和within 31

#### 2.2 作图系统 32

##### 2.2.1 图形布局 32

##### 2.2.2 利用部分构造图形 33

##### 2.2.3 par的使用 34

##### 2.2.4 组合图形 35

#### 2.3 r编程 36

##### 2.3.1 流程控制 37

##### 2.3.2 类和类函数 37

#### 2.4 数据输入 38

##### 2.4.1 读取文本文件 39

##### 2.4.2 read.table 的进一步讨论 41

##### 2.4.3 数据编辑器 42

2.4.4 其他程序的接口	43
2.5 练习题	44
第3章 概率和分布	45
3.1 随机抽样	45
3.2 概率计算和排列组合	46
3.3 离散分布	47
3.4 连续分布	47
3.5 r中的内置分布	48
3.5.1 密度	48
3.5.2 累积分布函数	50
3.5.3 分位数	51
3.5.4 随机数字	52
3.6 练习题	53
第4章 描述性统计和图形	54
4.1 单组的汇总统计量	54
4.2 分布的图形展示	58
4.2.1 直方图	58
4.2.2 经验累积分布	59
4.2.3 q-q图	59
4.2.4 箱式图	60
4.3 分组数据的汇总统计量	61
4.4 分组数据作图	64
4.4.1 直方图	64
4.4.2 并联箱式图	65
4.4.3 带状图	66
4.5 表格	68
4.5.1 生成表格	68
4.5.2 边际表格和相对频数	71
4.6 表格的图形显示	72
4.6.1 条形图	72
4.6.2 点图	74
4.6.3 饼图	75
4.7 练习题	76
第5章 单样本与双样本检验	77
5.1 单样本t检验	77
5.2 wilcoxon符号秩检验	80
5.3 两样本t检验	82
5.4 比较方差	83
5.5 两样本wilcoxon检验	84
5.6 配对t检验	85
5.7 配对wilcoxon检验	86
5.8 练习题	87
第6章 回归与相关性	88
6.1 简单线性回归	88
6.2 残差与回归值	92
6.3 预测与置信带	95
6.4 相关性	98
6.4.1 皮尔逊相关系数	98
6.4.2 斯皮尔曼相关系数	99

- 6.4.3 肯德尔等级相关系数? 100
- 6.5 练习题 100
- 第7章 方差分析与kruskal-wallis检验 102
  - 7.1 单因素方差分析 102
    - 7.1.1 成对比较和多重检验 106
    - 7.1.2 放宽对方差的假设 107
    - 7.1.3 图像表示 108
    - 7.1.4 bartlett检验 109
  - 7.2 kruskal-wallis检验 110
  - 7.3 双因素方差分析 110
  - 7.4 friedman检验 114
  - 7.5 回归分析中的方差分析表 114
  - 7.6 练习题 115
- 第8章 表格数据 117
  - 8.1 单比例 117
  - 8.2 两个独立的比例 118
  - 8.3 k比例, 检验趋势 120
  - 8.4 r<sup>2</sup>c表格 122
  - 8.5 练习题 124
- 第9章 功效与样本容量的计算 126
  - 9.1 功效计算原则 126
    - 9.1.1 单样本t及配对样本t检验的功效 127
    - 9.1.2 两样本t检验的功效 128
    - 9.1.3 近似方法 128
    - 9.1.4 比较比例的功效 129
  - 9.2 两样本问题 129
  - 9.3 单样本问题及配对样本检验 131
  - 9.4 比例的比较 131
  - 9.5 练习题 132
- 第10章 数据处理的高级技术 133
  - 10.1 变量的重编码 133
    - 10.1.1 cut函数 133
    - 10.1.2 处理因子 135
    - 10.1.3 日期的使用 136
    - 10.1.4 多变量重编码 139
  - 10.2 条件计算 140
  - 10.3 合并与重构数据框 141
    - 10.3.1 追加数据框 141
    - 10.3.2 合并数据框 142
    - 10.3.3 重塑数据框 144
  - 10.4 数据的分组及分案例操作 146
  - 10.5 时间分割 148
  - 10.6 练习题 152
- 第11章 多元回归 153
  - 11.1 多维数据绘图 153
  - 11.2 模型设定和模型输出 155
  - 11.3 模型筛选 157
  - 11.4 练习题 161
- 第12章 线性模型 162

- 12.1 多项式回归 163
- 12.2 过原点的回归分析 165
- 12.3 设计矩阵与虚拟变量 166
- 12.4 组间的共线性 168
- 12.5 交互效应 172
- 12.6 可重复的双因素方差分析 172
- 12.7 协方差分析 173
  - 12.7.1 图形描述 174
  - 12.7.2 比较回归线 177
- 12.8 模型诊断 183
- 12.9 练习题 187
- 第13章 逻辑回归 189
  - 13.1 广义线性模型 190
  - 13.2 表格化数据的逻辑回归 190
    - 13.2.1 偏差表分析 195
    - 13.2.2 与趋势检验之间的关联 196
  - 13.3 似然剖面分析 197
  - 13.4 让步比估计的表达 199
  - 13.5 原始数据的逻辑回归 199
  - 13.6 预测 201
  - 13.7 模型检查 202
  - 13.8 练习题 206
- 第14章 生存分析 208
  - 14.1 重要概念 208
  - 14.2 生存对象 209
  - 14.3 kaplan-meier估计 210
  - 14.4 对数秩检验 213
  - 14.5 cox比例风险模型 214
  - 14.6 练习题 216
- 第15章 比率和泊松回归 217
  - 15.1 基本思想 217
    - 15.1.1 泊松分布 217
    - 15.1.2 带有常数风险的生存分析 218
  - 15.2 泊松模型的拟合 219
  - 15.3 计算比率 223
  - 15.4 带有常数强度的模型 226
  - 15.5 练习题 230
- 第16章 非线性曲线拟合 231
  - 16.1 基本用法 232
  - 16.2 寻找初值 233
  - 16.3 自启动模型 238
  - 16.4 剖面分析 240
  - 16.5 更好地控制拟合算法 241
  - 16.6 练习题 242
- 附录a 获取并安装r以及iswr包 243
- 附录b iswr中的数据 246
- 附录c 摘要 272
- 附录d 练习题答案 283



# 《R语言统计入门(第2版)》

## 精彩短评

- 1、无功无过。
- 2、16018 ( 0313-0322 )。
- 3、书很不错。只是想吐槽一下自己，都要毕业了还在挣扎着啃一些入门。难怪求职处处碰壁呢
- 4、没有薛毅那本好.....
- 5、近似小白 基础差 这本书我竟然觉得有一些看不懂啊啊啊
- 6、现在网上对应的R语言资料也很多了，这本书是真的入门级的。如果，对统计学知识了解，而且有操作过R语言的基本经验，那么会发现看这本书很快就能看完了。如果对R没有任何了解的话，还是可以仔细读一读的。

# 《R语言统计入门(第2版)》

## 精彩书评

1、这是一本2008年的老作品，英文原版我看过部分章节。这是一本类似于工具书的介绍R语言的书籍。总体来说，因为写得比较详细、代码也给得很充足，这本书适合想学习R语言初学者。如果你是一名搞统计的但苦于学习R语言的，这本书能很好地作为范例来教你如何轻松使用R语言；另一方面如果你了解R语言的基本语法和数据类型（比如这一章里介绍过的），只是想调用统计包而已，那么R肯定是你的不二之选，这本书也能很好地给予参考。我个人认为这本书最大的优点但同时也是最大的弱点，是全书的R语言操作都是基于内置函数。这在初学R时很有帮助，因为简单的代码很容易重复出来。但实际工作中，尤其是绘图以及数据前处理时，需要很多业界大神编写的R语言包（比如ggplot2、reshape2）。R语言变化更新快，即便是去年那几位翻译的《R语言实战》，也落后了有些包的实际发展。何况这终究是一本2008年的书。我感觉，R语言书籍在国内的翻译越来越多，越来越去翻旧账了，比如《R数据可视化手册》。我读的书不多，但始终更喜欢《An Introduction to Statistical Learning with Applications in R》，希望哪天有大神去翻译。最后说说翻译。因为是几位大大译作，所以代码传承以及翻译质量都是有保证的。

# 《R语言统计入门(第2版)》

## 版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:[www.tushu000.com](http://www.tushu000.com)