

<<R语言实战>>

图书基本信息

书名：<<R语言实战>>

13位ISBN编号：9787115299901

10位ISBN编号：7115299900

出版时间：2013-1

出版时间：人民邮电出版社

作者：卡巴科弗

页数：388

字数：602000

译者：高涛,肖楠,陈钢

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：<http://www.tushu007.com>

<<R语言实战>>

内容概要

R是一个开源项目，具有强大的统计计算及制图能力，是从大数据中获取有用信息的绝佳工具，在各种主流操作系统上都可以安装使用，其基本安装就提供了数以百计的数据管理、统计和图形函数。

另外，社区开发的数以千计的扩展(包)为R增加了更多强大功能。

《R语言实战》注重实用性，是一本全面而细致的R指南，高度概括了该软件和它的强大功能，展示了实用的统计示例，且对于难以用传统方法处理的凌乱、不完整和非正态的数据给出了优雅的处理方法。

作者不仅仅探讨统计分析，还阐述了大量探索和展示数据的图形功能。

《R语言实战》适合数据分析人员及R用户学习参考。

<<R语言实战>>

作者简介

Robert I. Kabacoff

R语言社区著名学习网站Quick-R(<http://www.statmethods.net/>)的幕后维护者，现为全球化开发与咨询公司Management研究集团研发副总裁。

此前，Kabacoff博士是佛罗里达诺瓦东南大学的教授，讲授定量方法和统计编程的研究生课程。

Kabacoff还是临床心理学博士、统计顾问，擅长数据分析，在健康、金融服务、制造业、行为科学、政府和学术界有20余年的研究和统计咨询经验。

<<R语言实战>>

书籍目录

目 录

第一部分 入门

第1章 R语言介绍

1.1 为何要使用R?

1.2 R的获取和安装

1.3 R的使用

1.3.1 新手上路

1.3.2 获取帮助

1.3.3 工作空间

1.3.4 输入和输出

1.4 包

1.4.1 什么是包

1.4.2 包的安装

1.4.3 包的载入

1.4.4 包的使用方法

1.5 批处理

1.6 将输出用为输入——结果的重用

1.7 处理大数据集

1.8 示例实践

1.9 小结

第2章 创建数据集

2.1 数据集的概念

2.2 数据结构

2.2.1 向量

2.2.2 矩阵

2.2.3 数组

2.2.4 数据框

2.2.5 因子

2.2.6 列表

2.3 数据的输入

2.3.1 使用键盘输入数据

2.3.2 从带分隔符的文本文件导入数据

2.3.3 导入Excel数据

2.3.4 导入XML数据

2.3.5 从网页抓取数据

2.3.6 导入SPSS数据

2.3.7 导入SAS数据

2.3.8 导入Stata数据

2.3.9 导入netCDF数据

2.3.10 导入HDF5数据

2.3.11 访问数据库管理系统

2.3.12 通过Stat/Transfer导入数据

2.4 数据集的标注

2.4.1 变量标签

<<R语言实战>>

- 2.4.2 值标签
- 2.5 处理数据对象的实用函数
- 2.6 小结
- 第3章 图形初阶
 - 3.1 使用图形
 - 3.2 一个简单的例子
 - 3.3 图形参数
 - 3.3.1 符号和线条
 - 3.3.2 颜色
 - 3.3.3 文本属性
 - 3.3.4 图形尺寸与边界尺寸
 - 3.4 添加文本、自定义坐标轴和图例
 - 3.4.1 标题
 - 3.4.2 坐标轴
 - 3.4.3 参考线
 - 3.4.4 图例
 - 3.4.5 文本标注
 - 3.5 图形的组合
 - 3.6 小结
- 第4章 基本数据管理
 - 4.1 一个示例
 - 4.2 创建新变量
 - 4.3 变量的重编码
 - 4.4 变量的重命名
 - 4.5 缺失值
 - 4.5.1 重编码某些值为缺失值
 - 4.5.2 在分析中排除缺失值
 - 4.6 日期值
 - 4.6.1 将日期转换为字符型变量
 - 4.6.2 更进一步
 - 4.7 类型转换
 - 4.8 数据排序
 - 4.9 数据集的合并
 - 4.9.1 添加列
 - 4.9.2 添加行
 - 4.10 数据集取子集
 - 4.10.1 选入(保留)变量
 - 4.10.2 剔除(丢弃)变量
 - 4.10.3 选入观测
 - 4.10.4 subset()函数
 - 4.10.5 随机抽样
 - 4.11 使用SQL语句操作数据框
 - 4.12 小结
- 第5章 高级数据管理
 - 5.1 一个数据处理难题
 - 5.2 数值和字符处理函数
 - 5.2.1 数学函数

<<R语言实战>>

- 5.2.2 统计函数
- 5.2.3 概率函数
- 5.2.4 字符处理函数
- 5.2.5 其他实用函数
- 5.2.6 将函数应用于矩阵和数据框
- 5.3 数据处理难题的一套解决方案
- 5.4 控制流
 - 5.4.1 重复和循环
 - 5.4.2 条件执行
- 5.5 用户自编函数
- 5.6 整合与重构
 - 5.6.1 转置
 - 5.6.2 整合数据
 - 5.6.3 reshape包
- 5.7 小结
- 第二部分 基本方法
- 第6章 基本图形
 - 6.1 条形图
 - 6.1.1 简单的条形图
 - 6.1.2 堆砌条形图和分组条形图
 - 6.1.3 均值条形图
 - 6.1.4 条形图的微调
 - 6.1.5 棘状图
 - 6.2 饼图
 - 6.3 直方图
 - 6.4 核密度图
 - 6.5 箱线图
 - 6.5.1 使用并列箱线图进行跨组比较
 - 6.5.2 小提琴图
 - 6.6 点图
 - 6.7 小结
- 第7章 基本统计分析
 - 7.1 描述性统计分析
 - 7.1.1 方法云集
 - 7.1.2 分组计算描述性统计量
 - 7.1.3 结果的可视化
 - 7.2 频数表和列联表
 - 7.2.1 生成频数表
 - 7.2.2 独立性检验
 - 7.2.3 相关性的度量
 - 7.2.4 结果的可视化
 - 7.2.5 将表转换为扁平格式
 - 7.3 相关
 - 7.3.1 相关的类型
 - 7.3.2 相关性的显著性检验
 - 7.3.3 相关关系的可视化
 - 7.4 t检验

<<R语言实战>>

- 7.4.1 独立样本的t检验
- 7.4.2 非独立样本的t检验
- 7.4.3 多于两组的情况
- 7.5 组间差异的非参数检验
- 7.5.1 两组的比较
- 7.5.2 多于两组的比较
- 7.6 组间差异的可视化
- 7.7 小结
- 第三部分 中级方法
- 第8章 回归
- 8.1 回归的多面性
- 8.1.1 OLS回归的适用情境
- 8.1.2 基础回顾
- 8.2 OLS回归
- 8.2.1 用lm()拟合回归模型
- 8.2.2 简单线性回归
- 8.2.3 多项式回归
- 8.2.4 多元线性回归
- 8.2.5 有交互项的多元线性回归
- 8.3 回归诊断
- 8.3.1 标准方法
- 8.3.2 改进的方法
- 8.3.3 线性模型假设的综合验证
- 8.3.4 多重共线性
- 8.4 异常观测值
- 8.4.1 离群点
- 8.4.2 高杠杆值点
- 8.4.3 强影响点
- 8.5 改进措施
- 8.5.1 删除观测点
- 8.5.2 变量变换
- 8.5.3 增删变量
- 8.5.4 尝试其他方法
- 8.6 选择“最佳”的回归模型
- 8.6.1 模型比较
- 8.6.2 变量选择
- 8.7 深层次分析
- 8.7.1 交叉验证
- 8.7.2 相对重要性
- 8.8 小结
- 第9章 方差分析
- 9.1 术语速成
- 9.2 ANOVA模型拟合
- 9.2.1 aov()函数
- 9.2.2 表达式中各项的顺序
- 9.3 单因素方差分析
- 9.3.1 多重比较

<<R语言实战>>

- 9.3.2 评估检验的假设条件
- 9.4 单因素协方差分析
 - 9.4.1 评估检验的假设条件
 - 9.4.2 结果可视化
- 9.5 双因素方差分析
- 9.6 重复测量方差分析
- 9.7 多元方差分析
 - 9.7.1 评估假设检验
 - 9.7.2 稳健多元方差分析
- 9.8 用回归来做ANOVA
- 9.9 小结
- 第10章 功效分析
 - 10.1 假设检验速览
 - 10.2 用pwr包做功效分析
 - 10.2.1 t检验
 - 10.2.2 方差分析
 - 10.2.3 相关性
 - 10.2.4 线性模型
 - 10.2.5 比例检验
 - 10.2.6 卡方检验
 - 10.2.7 在新情况中选择合适的效应值
 - 10.3 绘制功效分析图形
 - 10.4 其他软件包
 - 10.5 小结
- 第11章 中级绘图
 - 11.1 散点图
 - 11.1.1 散点图矩阵
 - 11.1.2 高密度散点图
 - 11.1.3 三维散点图
 - 11.1.4 气泡图
 - 11.2 折线图
 - 11.3 相关图
 - 11.4 马赛克图
 - 11.5 小结
- 第12章 重抽样与自助法
 - 12.1 置换检验
 - 12.2 用coin包做置换检验
 - 12.2.1 独立两样本和K样本检验
 - 12.2.2 列联表中的独立性
 - 12.2.3 数值变量间的独立性
 - 12.2.4 两样本和K样本相关性检验
 - 12.2.5 深入探究
 - 12.3 lmPerm包的置换检验
 - 12.3.1 简单回归和多项式回归
 - 12.3.2 多元回归
 - 12.3.3 单因素方差分析和协方差分析
 - 12.3.4 双因素方差分析

<<R语言实战>>

- 12.4 置换检验点评
- 12.5 自助法
- 12.6 boot包中的自助法
 - 12.6.1 对单个统计量使用自助法
 - 12.6.2 多个统计量的自助法
- 12.7 小结
- 第四部分 高级方法
- 第13章 广义线性模型
 - 13.1 广义线性模型和glm()函数
 - 13.1.1 glm()函数
 - 13.1.2 连用的函数
 - 13.1.3 模型拟合和回归诊断
 - 13.2 Logistic回归
 - 13.2.1 解释模型参数
 - 13.2.2 评价预测变量对结果概率的影响
 - 13.2.3 过度离势
 - 13.2.4 扩展
 - 13.3 泊松回归
 - 13.3.1 解释模型参数
 - 13.3.2 过度离势
 - 13.3.3 扩展
 - 13.4 小结
- 第14章 主成分和因子分析
 - 14.1 R中的主成分和因子分析
 - 14.2 主成分分析
 - 14.2.1 判断主成分的个数
 - 14.2.2 提取主成分
 - 14.2.3 主成分旋转
 - 14.2.4 获取主成分得分
 - 14.3 探索性因子分析
 - 14.3.1 判断需提取的公共因子数
 - 14.3.2 提取公共因子
 - 14.3.3 因子旋转
 - 14.3.4 因子得分
 - 14.3.5 其他与EFA相关的包
 - 14.4 其他潜变量模型
 - 14.5 小结
- 第15章 处理缺失数据的高级方法
 - 15.1 处理缺失值的步骤
 - 15.2 识别缺失值
 - 15.3 探索缺失值模式
 - 15.3.1 列表显示缺失值
 - 15.3.2 图形探究缺失数据
 - 15.3.3 用相关性探索缺失值
 - 15.4 理解缺失数据的来由和影响
 - 15.5 理性处理不完整数据
 - 15.6 完整实例分析(行删除)

<<R语言实战>>

- 15.7 多重插补
- 15.8 处理缺失值的其他方法
 - 15.8.1 成对删除
 - 15.8.2 简单(非随机)插补
- 15.9 小结
- 第16章 高级图形进阶
 - 16.1 R中的四种图形系统
 - 16.2 lattice包
 - 16.2.1 条件变量
 - 16.2.2 面板函数
 - 16.2.3 分组变量
 - 16.2.4 图形参数
 - 16.2.5 页面摆放
 - 16.3 ggplot2包
 - 16.4 交互式图形
 - 16.4.1 与图形交互：鉴别点
 - 16.4.2 playwith
 - 16.4.3 latticist
 - 16.4.4 iplots包的交互图形
 - 16.4.5 rggobi
 - 16.5 小结
- 后记：探索R的世界
- 附录A 图形用户界面
- 附录B 自定义启动环境
- 附录C 从R中导出数据
- 附录D 制作出版级品质的输出
- 附录E R中的矩阵运算
- 附录F 本书中用到的扩展包
- 附录G 处理大数据
- 附录H 更新R
- 参考文献

<<R语言实战>>

章节摘录

版权页： 插图： 这种方法的交互性很强。

他拟合了一系列模型，检验它们是否符合相应的统计假设，探索了所有异常地发现，最终从许多可能的模型中选择了“最佳”的模型。

如果成功，那么结果将会帮助他完成以下任务。

在众多变量中判断哪些对预测桥梁退化是有用的，得到它们的相对重要性，从而关注重要的变量。

根据回归所得的等式预测新的桥梁的退化情况（预测变量的值已知，但是桥梁退化程度未知），找出那些可能会有麻烦的桥梁。

利用对异常桥梁的分析，获得一些意外的信息。

比如他发现某些桥梁的退化速度比预测的更快或更慢，那么研究这些“离群点”可能会有重大的发现，能够帮助理解桥梁退化的机制。

可能桥梁的例子并不能引起你的兴趣。

而我是从事临床心理学和统计的，对土木工程也是一无所知，但是这其中蕴含的一般性思想适用于物理、生物和社会科学的许多问题。

以下问题都可以通过OLS方法进行处理。

<<R语言实战>>

媒体关注与评论

“ 本书从务实的角度出发，清晰阐释了R的基本知识及统计数据分析，为我提供了很大帮助。
——读者评论

<<R语言实战>>

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:<http://www.tushu007.com>