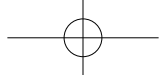


# 目 录

## 上篇 使用R语言进行统计建模

<b>第一章 R语言数据科学</b> .....	3
1.1 基础操作 .....	3
1.1.1 简介及安装 .....	3
1.1.2 设置工作目录 .....	7
1.1.3 数据导入和保存 .....	8
1.1.3.1 R的数据结构 .....	8
1.1.3.2 R的数据 .....	9
1.1.3.3 数据的导入和保存 .....	11
1.2 数据管理 .....	14
1.2.1 传统数据框的操作和管理 .....	14
1.2.1.1 一些常规操作 .....	14
1.2.1.2 对数据框信息进行修改 .....	19
1.2.1.3 数据操作的常用函数 .....	21
1.2.2 tibble简单数据框的操作和管理 .....	26
1.2.2.1 与传统的数据框比较 .....	26
1.2.2.2 tibble数据的导入 .....	27
1.2.2.3 数据管理最重要的五大函数 .....	30
1.2.2.4 长、宽数据的相互转换 .....	37
练习 .....	39
<b>第二章 数据探索：描述性统计和数据可视化</b> .....	41
2.1 使用R对数据进行描述 .....	41
2.1.1 趋中度和变异性：平均数、方差和标准差 .....	41
2.1.1.1 概念 .....	41
2.1.1.2 使用R计算平均数和标准差 .....	44
2.1.2 趋中度和变异性：中位数 (median) 和四分位数 (quartile) .....	49
2.1.3 趋中度和变异性：众数和幅度 .....	51
2.1.4 其他描述统计方法 .....	53



2.1.5	频数表和列联表.....	54
2.2	数据可视化.....	57
2.2.1	基础知识：R基础图形方法.....	58
2.2.1.1	图形的标题和坐标轴标签.....	59
2.2.1.2	符号、线条、颜色、文本属性.....	61
2.2.1.3	图形尺寸和边界尺寸.....	62
2.2.1.4	添加图例、自定义坐标轴、添加参考线、 文本标注.....	63
2.2.2	几个常用绘图函数使用实例.....	66
2.2.2.1	plot ( ) 绘图函数.....	66
2.2.2.2	直方图、密度图和箱体图.....	70
2.2.3	Lattice包的几个绘图函数.....	73
2.2.4	ggplot 2绘图.....	75
2.2.4.1	几何对象函数geom_histogram ( ).....	77
2.2.4.2	几何对象函数geom_density ( ).....	80
2.2.4.3	几何对象函数geom_boxplot ( ).....	81
2.2.4.4	同时绘制多个几何对象.....	81
2.2.4.5	统计变换.....	83
2.2.4.6	总结.....	85
	练习.....	86

### 第三章 从样本估计总体：概率分布和假设检验..... 88

3.1	$z$ 分布.....	88
3.2	$t$ 分布、 $F$ 分布和 $\chi^2$ 分布.....	95
3.3	二项分布.....	100
3.4	假设检验 (Hypothesis Testing).....	102
3.5	标准误和置信区间.....	107
	练习.....	110

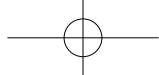
### 第四章 使用R进行统计建模..... 112

4.1	回归分析的概念.....	114
4.1.1	斜率和截距.....	116
4.1.2	最小二乘法以及模型拟合度指标 $R^2$ .....	118
4.2	简单回归分析.....	121
4.2.1	因变量和自变量都是数值型变量.....	121

4.2.1.1	实例一 .....	121
4.2.1.2	实例二 .....	126
4.2.2	因变量是数值型变量，自变量是分类变量 .....	129
4.2.2.1	因变量是连续型变量、自变量是二元变量 .....	130
4.2.2.2	分类变量有多个水平 .....	136
4.3	多元回归分析 .....	143
4.3.1	两个自变量都是数值变量 .....	144
4.3.2	交互效应 .....	147
4.3.3	自变量：一个数值型变量加一个分类变量 .....	149
4.3.4	数值型自变量做趋中处理 (Centering) .....	158
4.3.5	自变量：两个分类变量 .....	160
4.3.6	回归分析要满足的统计假设的前提以及模型诊断 .....	166
4.3.7	比较编码方案与多重比较 .....	173
4.3.7.1	Treatment Coding .....	174
4.3.7.2	sum coding .....	177
4.3.7.3	treatment coding与sum coding对比 .....	179
4.3.8	事先计划比较和事后比较 .....	188
4.3.8.1	事先计划比较 .....	188
4.3.8.2	事后比较 .....	192
4.3.9	变量的选择和模型比较 .....	194
4.3.9.1	变量进入模型 .....	194
4.3.9.2	模型比较：anova ( ) , AIC ( ) 和drop 1 ( ) .....	198
4.4	广义线性模型: Logistic Regression .....	200
4.4.1	介绍 .....	200
4.4.2	例子一：被试正误判断数据 .....	201
4.4.3	例子二：let还是allow .....	206
4.5	广义线性模型: 泊松回归 .....	211
	练习 .....	215

## 第五章 重复测量和混合设计 .....

5.1	一个自变量两个水平的数据模型： $t$ 检验 .....	219
5.1.1	独立样本 $t$ 检验 .....	219
5.1.2	配对样本 $t$ 检验 .....	221
5.2	一个自变量多个水平的统计模型 .....	223
5.3	两个自变量混合设计的统计模型 .....	229



<b>第六章 混合效应模型</b> .....	236
6.1 引言 .....	236
6.2 对比传统方差分析和混合效应模型：一个具体的研究案例 .....	237
6.2.1 反应时数据：平均数与个体差异的矛盾 .....	238
6.2.2 准确率数据：求比例与二元变量的矛盾 .....	240
6.3 混合效应模型：概念及内涵 .....	241
6.4 翻译判断实验的混合效应模型 .....	246
6.4.1 导入数据，并描述、探索 .....	246
6.4.2 拟合、比较、选择和解释模型 .....	248

## 下篇 二语加工主题下的R应用

<b>第七章 二语语音加工：中国英语学习者元音感知中的“范畴合并”现象</b> .....	257
7.1 研究背景 .....	258
7.2 实验设计 .....	260
7.2.1 被试 .....	260
7.2.2 实验材料 .....	261
7.2.3 程序 .....	262
7.3 R语言数据分析 .....	262
7.3.1 A'分数 .....	262
7.3.2 结论 .....	274
练习 .....	275
<b>第八章 二语词汇加工：中国英语学习者词汇与概念表征发展研究</b> .....	276
8.1 研究背景 .....	276
8.2 实验设计 .....	279
8.2.1 被试 .....	279
8.2.2 材料 .....	279
8.2.3 程序 .....	280
8.3 R语言数据分析 .....	281

8.3.1 反应时.....	281
8.3.2 准确率的模型拟合.....	289
练习.....	291
<b>第九章 二语指称加工与格赖斯“量”的原理.....</b>	<b>292</b>
9.1 研究背景.....	292
9.2 实验设计.....	295
9.2.1 被试.....	295
9.2.2 材料.....	296
9.2.3 程序.....	296
9.3 R语言数据分析.....	297
9.3.1 介绍.....	297
9.3.2 区域一.....	298
9.3.3 区域二.....	304
练习.....	309
<b>参考文献.....</b>	<b>310</b>

