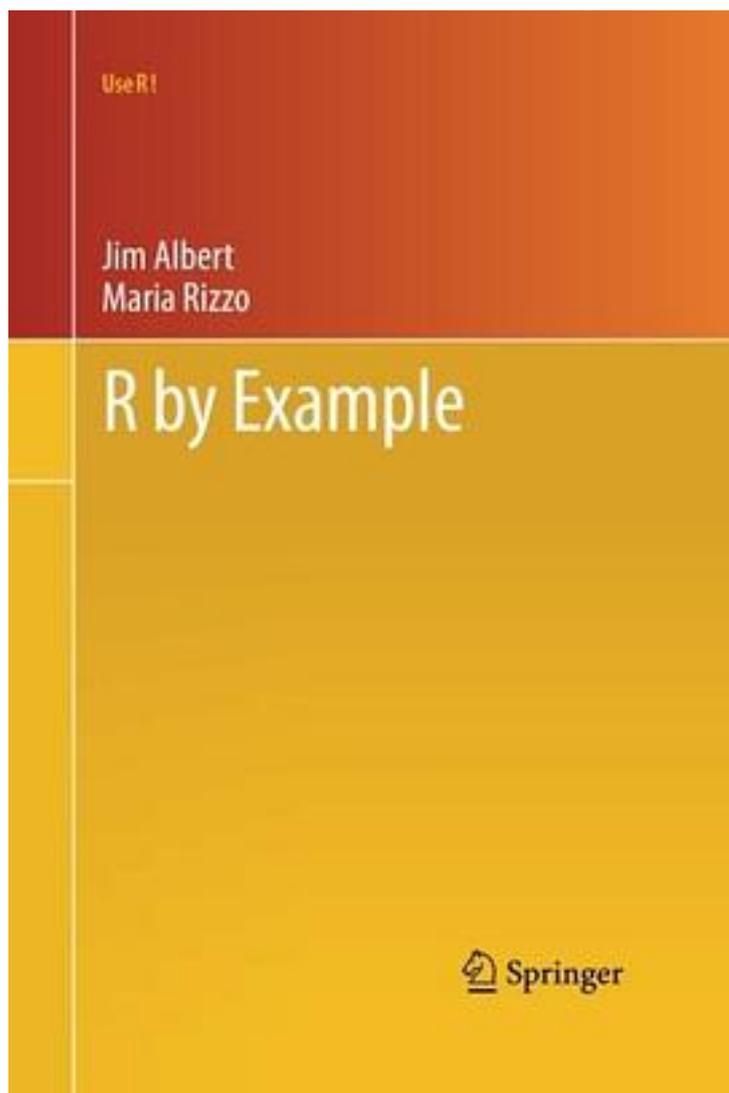


# R by Example



[R by Example\\_ 下载链接1](#)

著者:Jim Albert

出版者:Springer

出版时间:2011-11-23

装帧:Paperback

isbn:9781461413646

R by Example is an example-based introduction to the statistical computing environment that does not assume any previous familiarity with R or other software packages. R functions are presented in the context of interesting applications with real data. The purpose of this book is to illustrate a range of statistical and probability computations using R for people who are learning, teaching, or using statistics. Specifically, this book is written for users who have covered at least the equivalent of (or are currently studying) undergraduate level calculus-based courses in statistics. These users are learning or applying exploratory and inferential methods for analyzing data and this book is intended to be a useful resource for learning how to implement these procedures in R.

作者介绍:

目录: 1 Introduction .....	1
1.1 Getting Started .....	1
1.1.1 Preliminaries .....	2
1.1.2 Basic operations .....	3
1.1.3 R Scripts .....	11
1.1.4 The R Help System .....	13
1.2 Functions .....	14
1.3 Vectors and Matrices .....	16
1.4 Data Frames .....	19
1.4.1 Introduction to data frames .....	20
1.4.2 Working with a data frame .....	22
1.5 Importing Data .....	26
1.5.1 Entering data manually .....	26
1.5.2 Importing data from a text file .....	28
1.5.3 Data available on the internet .....	30
1.6 Packages .....	33
1.7 The R Workspace .....	35
1.8 Options and Resources .....	36
1.9 Reports and Reproducible Research .....	39
Exercises .....	39
2 Quantitative Data .....	43
2.1 Introduction .....	43
2.2 Bivariate Data: Two Quantitative Variables .....	43
2.2.1 Exploring the data .....	44
2.2.2 Correlation and regression line .....	46
2.2.3 Analysis of bivariate data by group .....	48
2.2.4 Conditional plots .....	50
2.3 Multivariate Data: Several Quantitative Variables .....	52
2.3.1 Exploring the data .....	53
2.3.2 Missing values .....	54
2.3.3 Summarize by group .....	54
2.3.4 Summarize pairs of variables .....	55
2.3.5 Identifying missing values .....	58
2.4 Time Series Data .....	59
2.5 Integer Data: Draft Lottery .....	61
2.6 Sample Means and the Central Limit Theorem .....	65
2.7 Special Topics .....	68
2.7.1 Adding a new variable .....	68
2.7.2 Which observation is the maximum? .....	69

2.7.3	Sorting a data frame	70
2.7.4	Distances between points	71
2.7.5	Quick look at cluster analysis	73
	Exercises	75
3	Categorical data	79
3.1	Introduction	79
3.1.1	Tabulating and plotting categorical data	79
3.1.2	Character vectors and factors	80
3.2	Chi-square Goodness-of-Fit Test	82
3.3	Relating Two Categorical Variables	85
3.3.1	Introduction	85
3.3.2	Frequency tables and graphs	86
3.3.3	Contingency tables	88
3.4	Association Patterns in Contingency Tables	90
3.4.1	Constructing a contingency table	90
3.4.2	Graphing patterns of association	92
3.5	Testing Independence by a Chi-square Test	93
	Exercises	96
4	Presentation Graphics	101
4.1	Introduction	101
4.2	Labeling the Axes and Adding a Title	102
4.3	Changing the Plot Type and Plotting Symbol	103
4.4	Overlaying Lines and Line Types	105
4.5	Using Different Colors for Points and Lines	108
4.6	Changing the Format of Text	110
4.7	Interacting with the Graph	111
4.8	Multiple Figures in a Window	112
4.9	Overlaying a Curve and Adding a Mathematical Expression	113
4.10	Multiple Plots and Varying the Graphical Parameters	116
4.11	Creating a Plot using Low-Level Functions	119
4.12	Exporting a Graph to a Graphics File	120
4.13	The lattice Package	121
4.14	The ggplot2 Package	126
	Exercises	127
5	Exploratory Data Analysis	133
5.1	Introduction	133
5.2	Meet the Data	134
5.3	Comparing Distributions	135
5.3.1	Stripcharts	135
5.3.2	Identifying outliers	136
5.3.3	Five-number summaries and boxplots	137
5.4	Relationships Between Variables	139
5.4.1	Scatterplot and a resistant line	139
5.4.2	Plotting residuals and identifying outliers	140
5.5	Time Series Data	141
5.5.1	Scatterplot, least-squares line, and residuals	141
5.5.2	Transforming by a logarithm and fitting a line	143
5.6	Exploring Fraction Data	145
5.6.1	Stemplot	145
5.6.2	Transforming fraction data	146
	Exercises	149
6	Basic Inference Methods	153
6.1	Introduction	153
6.2	Learning About a Proportion	154

6.2.1 Testing and estimation problems	154
6.2.2 Creating group variables by the ifelse function	154
6.2.3 Large-sample test and estimation methods	154
6.2.4 Small sample methods	156
6.3 Learning About a Mean	158
6.3.1 Introduction	158
6.3.2 One-sample t statistic methods	158
6.3.3 Nonparametric methods	161
6.4 Two Sample Inference	163
6.4.1 Introduction	163
6.4.2 Two sample t-test	163
6.4.3 Two sample Mann-Whitney-Wilcoxon test	165
6.4.4 Permutation test	166
6.5 Paired Sample Inference Using a t Statistic	167
Exercises	170
7 Regression	173
7.1 Introduction	173
7.2 Simple Linear Regression	174
7.2.1 Fitting the model	174
7.2.2 Residuals	176
7.2.3 Regression through the origin	177
7.3 Regression Analysis for Data with Two Predictors	178
7.3.1 Preliminary analysis	178
7.3.2 Multiple regression model	181
7.3.3 The summary and anova methods for lm	184
7.3.4 Interval estimates for new observations	185
7.4 Fitting a Regression Curve	187
Exercises	196
8 Analysis of Variance I	199
8.1 Introduction	199
8.1.1 Data entry for one-way ANOVA	200
8.1.2 Preliminary data analysis	202
8.2 One-way ANOVA	204
8.2.1 ANOVA F test using oneway.test	204
8.2.2 One-way ANOVA model	205
8.2.3 ANOVA using lm or aov	206
8.2.4 Fitting the model	207
8.2.5 Tables of means or estimated effects	207
8.2.6 ANOVA Table	208
8.3 Comparison of Treatment Means	210
8.3.1 Fisher Least Significant Difference (LSD)	210
8.3.2 Tukey's multiple comparison method	212
8.4 A Statistical Reference Dataset from NIST	215
8.5 Stacking Data	217
Exercises	222
8.6 Chapter 8 Appendix: Exploring ANOVA calculations	223
9 Analysis of Variance II	227
9.1 Introduction	227
9.2 Randomized Block Designs	227
9.2.1 The randomized block model	229
9.2.2 Analysis of the randomized block model	230
9.3 Two-way ANOVA	234
9.3.1 The two-way ANOVA model	235
9.3.2 Analysis of the two-way ANOVA model	236

Exercises .....	240
10 Randomization Tests .....	243
10.1 Introduction .....	243
10.2 Exploring Data for One-way Analysis .....	243
10.3 Randomization Test for Location .....	246
10.4 Permutation Test for Correlation .....	249
Exercises .....	252
11 Simulation Experiments .....	255
11.1 Introduction .....	255
11.2 Simulating a Game of Chance .....	255
11.2.1 The sample function .....	256
11.2.2 Exploring cumulative winnings .....	256
11.2.3 R function to implement a Monte Carlo experiment ...	258
11.2.4 Summarizing the Monte Carlo results .....	258
11.2.5 Modifying the experiment to learn about new statistics	260
11.3 Random Permutations .....	262
11.3.1 Using sample to simulate an experiment .....	262
11.3.2 Comparing two permutations of a sample .....	263
11.3.3 Writing a function to perform simulation .....	263
11.3.4 Repeating the simulation .....	264
11.4 The Collector's Problem .....	265
11.4.1 Simulating experiment using the sample function .....	266
11.4.2 Writing a function to perform the simulation .....	267
11.4.3 Buying an optimal number of cards .....	268
11.5 Patterns of Dependence in a Sequence .....	270
11.5.1 Writing a function to compute streaks .....	270
11.5.2 Writing a function to simulate hitting data .....	271
Exercises .....	273
12 Bayesian Modeling .....	277
12.1 Introduction .....	277
12.2 Learning about a Poisson Rate .....	278
12.3 A Prior Density .....	278
12.4 Information Contained in the Data: the Likelihood .....	279
12.5 The Posterior and Inferences .....	281
12.5.1 Computation of the posterior .....	281
12.5.2 Exact summarization of the posterior .....	282
12.5.3 Summarizing a posterior by simulation .....	284
12.6 Simulating a Probability Distribution by a Random Walk ...	285
12.6.1 Introduction .....	285
12.6.2 The Metropolis-Hastings random walk algorithm .....	286
12.6.3 Using an alternative prior .....	289
12.7 Bayesian Model Checking .....	290
12.7.1 The predictive distribution .....	291
12.7.2 Model checking .....	292
12.8 Negative Binomial Modeling .....	295
12.8.1 Overdispersion .....	295
12.8.2 Fitting the Negative Binomial model .....	296
Exercises .....	302
13 Monte Carlo Methods .....	307
13.1 The Monte Carlo Method of Computing Integrals .....	307
13.1.1 Introduction .....	307
13.1.2 Estimating a probability .....	308
13.1.3 Estimating an expectation .....	309
13.2 Learning about the Sampling Distribution of a Statistic .....	311

13.2.1 Simulating the sampling distribution by the Monte Carlo method .....	312
13.2.2 Constructing a percentile confidence interval .....	312
13.3 Comparing Estimators .....	314
13.3.1 A simulation experiment .....	315
13.3.2 Estimating bias .....	315
13.3.3 Estimating mean distance from the target .....	316
13.4 Assessing Probability of Coverage .....	318
13.4.1 A Monte Carlo experiment to compute a coverage probability .....	318
13.5 Markov Chain Monte Carlo .....	321
13.5.1 Markov Chains .....	321
13.5.2 Metropolis-Hastings algorithm .....	324
13.5.3 Random walk Metropolis-Hastings algorithm .....	328
13.5.4 Gibbs sampling .....	331
13.6 Further Reading .....	333
Exercises .....	333
A Vectors, Matrices, and Lists .....	337
A.1 Vectors .....	337
A.1.1 Creating a vector .....	337
A.1.2 Sequences .....	337
A.1.3 Extracting and replacing elements of vectors .....	338
A.2 The sort and order functions .....	339
A.3 Matrices .....	340
A.3.1 Creating a matrix .....	340
A.3.2 Arithmetic on matrices .....	342
A.4 Lists .....	344
A.5 Sampling from a data frame .....	347
References .....	349
Index .....	353
• • • • • <a href="#">(收起)</a>	

[R by Example\\_ 下载链接1](#)

标签

R

统计

数据分析

计算机

Statistics

编程

评论

把大致的统计步骤跑了一遍，可参照薛毅的编的那本一起读。

-----  
[R by Example\\_下载链接1](#)

书评

-----  
[R by Example\\_下载链接1](#)