

Contents

1	The Basics of R.	1
1.1	Introduction	1
1.2	R on Your Computer.	1
1.2.1	History of the R Language	1
1.2.2	Installing and Updating R.	2
1.2.3	Packages	3
1.3	First Steps	5
1.3.1	“Hello World !!!”	5
1.3.2	Getting Help.	5
1.3.3	Working Space	7
1.4	Basics of the R Language.	8
1.4.1	R as a Calculator.	8
1.4.2	Variables	10
1.4.3	Arrays	13
1.4.4	Data Frames.	20
1.4.5	Lists	24
1.4.6	Programming in R.	26
1.4.7	Date Types.	29
1.4.8	Reading and Writing Data from and to Files.	30
2	Numerical Techniques.	33
2.1	Matrix Algebra	33
2.1.1	Characteristics of Matrices	34
2.1.2	Matrix Operations	37
2.1.3	Eigenvalues and Eigenvectors	41
2.1.4	Spectral Decomposition	43
2.1.5	Norm.	44
2.2	Numerical Integration	46
2.2.1	Integration of Functions of One Variable	46
2.2.2	Integration of Functions of Several Variables	50

2.3	Differentiation	54
2.3.1	Analytical Differentiation	54
2.3.2	Numerical Differentiation	56
2.3.3	Automatic Differentiation	59
2.4	Root Finding	62
2.4.1	Solving Systems of Linear Equations	62
2.4.2	Solving Systems of Nonlinear Equations	64
2.4.3	Maximisation and Minimisation of Functions	66
3	Combinatorics and Discrete Distributions	77
3.1	Set Theory	77
3.1.1	Creating Sets	77
3.1.2	Basics of Set Theory	78
3.1.3	Base Package	79
3.1.4	Sets Package	81
3.1.5	Generalised Sets	82
3.2	Probabilistic Experiments with Finite Sample Spaces	85
3.2.1	R Functionality	87
3.2.2	Sample Space and Sampling from Urns	87
3.2.3	Sampling Procedure	91
3.2.4	Random Variables	92
3.3	Binomial Distribution	94
3.3.1	Bernoulli Random Variables	94
3.3.2	Binomial Distribution	95
3.3.3	Properties	98
3.4	Multinomial Distribution	99
3.5	Hypergeometric Distribution	101
3.6	Poisson Distribution	103
3.6.1	Summation of Poisson Distributed Random Variables	106
4	Univariate Distributions	109
4.1	Continuous Distributions	109
4.1.1	Properties of Continuous Distributions	110
4.2	Uniform Distribution	112
4.3	Normal Distribution	113
4.4	Distributions Related to the Normal Distribution	114
4.4.1	χ^2 Distribution	115
4.4.2	Student's t -distribution	117
4.4.3	F -distribution	119
4.5	Other Univariate Distributions	121
4.5.1	Exponential Distribution	121
4.5.2	Stable Distributions	123
4.5.3	Cauchy Distribution	127

5	Univariate Statistical Analysis	129
5.1	Descriptive Statistics	129
5.1.1	Graphical Data Representation	130
5.1.2	Empirical (Cumulative) Distribution Function	132
5.1.3	Histogram	133
5.1.4	Kernel Density Estimation	135
5.1.5	Location Parameters	137
5.1.6	Dispersion Parameters	140
5.1.7	Higher Moments	143
5.1.8	Box-Plot	144
5.2	Confidence Intervals and Hypothesis Testing	146
5.2.1	Confidence Intervals	146
5.2.2	Hypothesis Testing	149
5.3	Goodness-of-Fit Tests	160
5.3.1	General Tests	161
5.3.2	Tests for Normality	164
5.3.3	Wilcoxon Signed Rank Test and Mann–Whitney U Test	167
5.3.4	Kruskal–Wallis Test	169
6	Multivariate Distributions	171
6.1	The Distribution Function and the Density Function of a Random Vector	171
6.1.1	Moments	173
6.2	The Multinormal Distribution	178
6.2.1	Sampling Distributions and Limit Theorems	182
6.3	Copulae	183
6.3.1	Copula Families	185
6.3.2	Archimedean Copulae	189
6.3.3	Hierarchical Archimedean Copulae	191
6.3.4	Estimation	193
7	Regression Models	197
7.1	Idea of Regression	197
7.2	Linear Regression	198
7.2.1	Model Selection Criteria	200
7.2.2	Stepwise Regression	201
7.3	Nonparametric Regression	205
7.3.1	General Form	206
7.3.2	Kernel Regression	207
7.3.3	k -Nearest Neighbours (k -NN)	209
7.3.4	Splines	212
7.3.5	LOESS or Local Regression	213

8	Multivariate Statistical Analysis.	219
8.1	Principal Components Analysis.	219
8.2	Factor Analysis.	224
8.2.1	Maximum Likelihood Factor Analysis	225
8.3	Cluster Analysis	229
8.3.1	Proximity of Objects	230
8.3.2	Clustering Algorithms	231
8.4	Multidimensional Scaling	234
8.4.1	Metric Multidimensional Scaling.	235
8.4.2	Non-metric Multidimensional Scaling	236
8.5	Discriminant Analysis	238
9	Random Numbers in R.	243
9.1	Generating Random Numbers.	243
9.1.1	Pseudorandom Number Generators	244
9.1.2	Uniformly Distributed Pseudorandom Numbers.	248
9.1.3	Uniformly Distributed True Random Numbers	249
9.2	Generating Random Variables	250
9.2.1	General Principles for Random Variable Generation	251
9.2.2	Random Variables.	253
9.2.3	Random Variable Generation for Continuous Distributions.	253
9.2.4	Random Variable Generation for Discrete Distributions.	259
9.2.5	Random Variable Generation for Multivariate Distributions.	261
9.3	Tests for Randomness	265
9.3.1	Birthday Spacings.	266
9.3.2	k -Distribution Test	266
10	Advanced Graphical Techniques in R	269
10.1	Package <code>lattice</code>	269
10.1.1	Getting Started with <code>lattice</code>	270
10.1.2	<code>formula</code> Argument.	270
10.1.3	<code>panel</code> Argument and Appearance Settings	272
10.1.4	Conditional and Grouped Plots	273
10.1.5	Concept of <code>shingle</code>	275
10.1.6	Time Series Plots	278
10.1.7	Three- and Four-Dimensional Plots	279
10.2	Package <code>rgl</code>	281
10.2.1	Getting Started with <code>rgl</code>	281
10.2.2	Shape Functions	282
10.2.3	Export and Animation Functions.	287

Contents	xvii
10.3 Package <code>rpanel</code>	289
10.3.1 Getting Started with <code>rpanel</code>	289
10.3.2 Application Functions in <code>rpanel</code>	293
Bibliography	297
Index	303

Basic Elements of Computational Statistics

Härdle, W.; Okhrin, O.; Okhrin, Y.

2017, XXI, 305 p. 97 illus., 66 illus. in color., Hardcover

ISBN: 978-3-319-55335-1