



Contents

Preface	v
Typographical Conventions	xi
1 Introduction	1
1.1 A Quick Overview of S	3
1.2 Using S	5
1.3 An Introductory Session	6
1.4 What Next?	12
2 Data Manipulation	13
2.1 Objects	13
2.2 Connections	20
2.3 Data Manipulation	27
2.4 Tables and Cross-Classification	37
3 The S Language	41
3.1 Language Layout	41
3.2 More on S Objects	44
3.3 Arithmetical Expressions	47
3.4 Character Vector Operations	51
3.5 Formatting and Printing	54
3.6 Calling Conventions for Functions	55
3.7 Model Formulae	56
3.8 Control Structures	58
3.9 Array and Matrix Operations	60
3.10 Introduction to Classes and Methods	66
4 Graphics	69
4.1 Graphics Devices	71
4.2 Basic Plotting Functions	72

4.3	Enhancing Plots	77
4.4	Fine Control of Graphics	82
4.5	Trellis Graphics	89
5	Univariate Statistics	107
5.1	Probability Distributions	107
5.2	Generating Random Data	110
5.3	Data Summaries	111
5.4	Classical Univariate Statistics	115
5.5	Robust Summaries	119
5.6	Density Estimation	126
5.7	Bootstrap and Permutation Methods	133
6	Linear Statistical Models	139
6.1	An Analysis of Covariance Example	139
6.2	Model Formulae and Model Matrices	144
6.3	Regression Diagnostics	151
6.4	Safe Prediction	155
6.5	Robust and Resistant Regression	156
6.6	Bootstrapping Linear Models	163
6.7	Factorial Designs and Designed Experiments	165
6.8	An Unbalanced Four-Way Layout	169
6.9	Predicting Computer Performance	177
6.10	Multiple Comparisons	178
7	Generalized Linear Models	183
7.1	Functions for Generalized Linear Modelling	187
7.2	Binomial Data	190
7.3	Poisson and Multinomial Models	199
7.4	A Negative Binomial Family	206
7.5	Over-Dispersion in Binomial and Poisson GLMs	208
8	Non-Linear and Smooth Regression	211
8.1	An Introductory Example	211
8.2	Fitting Non-Linear Regression Models	212
8.3	Non-Linear Fitted Model Objects and Method Functions	217
8.4	Confidence Intervals for Parameters	220
8.5	Profiles	226

8.6	Constrained Non-Linear Regression	227
8.7	One-Dimensional Curve-Fitting	228
8.8	Additive Models	232
8.9	Projection-Pursuit Regression	238
8.10	Neural Networks	243
8.11	Conclusions	249
9	Tree-Based Methods	251
9.1	Partitioning Methods	253
9.2	Implementation in <code>rpart</code>	258
9.3	Implementation in <code>tree</code>	266
10	Random and Mixed Effects	271
10.1	Linear Models	272
10.2	Classic Nested Designs	279
10.3	Non-Linear Mixed Effects Models	286
10.4	Generalized Linear Mixed Models	292
10.5	GEE Models	299
11	Exploratory Multivariate Analysis	301
11.1	Visualization Methods	302
11.2	Cluster Analysis	315
11.3	Factor Analysis	321
11.4	Discrete Multivariate Analysis	325
12	Classification	331
12.1	Discriminant Analysis	331
12.2	Classification Theory	338
12.3	Non-Parametric Rules	341
12.4	Neural Networks	342
12.5	Support Vector Machines	344
12.6	Forensic Glass Example	346
12.7	Calibration Plots	349
13	Survival Analysis	353
13.1	Estimators of Survivor Curves	355
13.2	Parametric Models	359
13.3	Cox Proportional Hazards Model	365

13.4 Further Examples	371
14 Time Series Analysis	387
14.1 Second-Order Summaries	389
14.2 ARIMA Models	397
14.3 Seasonality	403
14.4 Nottingham Temperature Data	406
14.5 Regression with Autocorrelated Errors	411
14.6 Models for Financial Series	414
15 Spatial Statistics	419
15.1 Spatial Interpolation and Smoothing	419
15.2 Kriging	425
15.3 Point Process Analysis	430
16 Optimization	435
16.1 Univariate Functions	435
16.2 Special-Purpose Optimization Functions	436
16.3 General Optimization	436
 Appendices	
A Implementation-Specific Details	447
A.1 Using S-PLUS under Unix / Linux	447
A.2 Using S-PLUS under Windows	450
A.3 Using R under Unix / Linux	453
A.4 Using R under Windows	454
A.5 For Emacs Users	455
B The S-PLUS GUI	457
C Datasets, Software and Libraries	461
C.1 Our Software	461
C.2 Using Libraries	462
References	465
Index	481