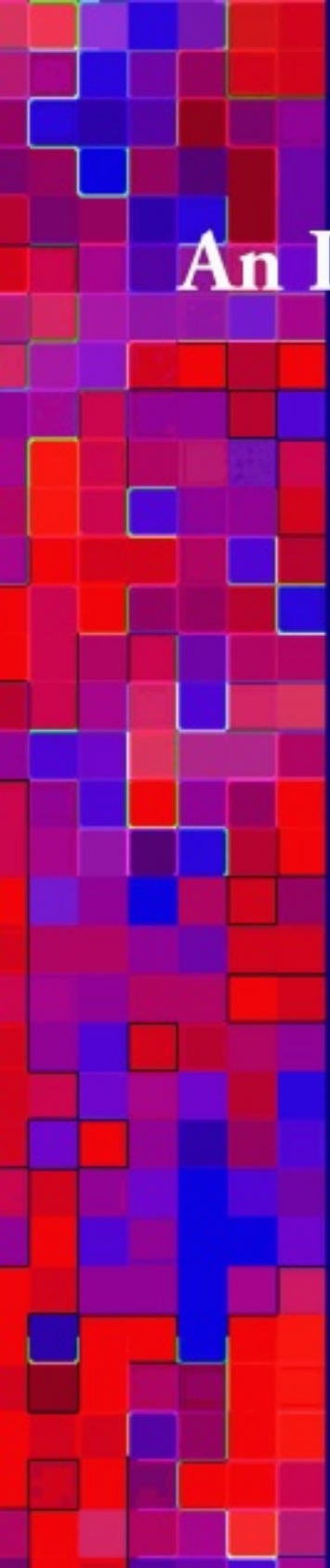

An Introduction to R

Second Edition

W. N. Venables, D. M. Smith and the
R Development Core Team

Copyright © 1990 W. N. Venables
Copyright © 1992 W. N. Venables & D. M. Smith
Copyright © 1997 R. Gentleman & R. Ihaka
Copyright © 1997, 1998 M. Maechler
Copyright © 1999–2006 R Development Core Team

Permission is granted to make and distribute verbatim copies of this manual provided the copyright notice and this permission notice are preserved on all copies.

Permission is granted to copy and distribute modified versions of this manual under the conditions for verbatim copying, provided that the entire resulting derived work is distributed under the terms of a permission notice identical to this one.

Permission is granted to copy and distribute translations of this manual into another language, under the above conditions for modified versions, except that this permission notice may be stated in a translation approved by the R Development Core Team.

ISBN 3-900051-12-7

Table of Contents

Preface 1

1 Introduction and preliminaries 2
- 1.1 The R environment 2
- 1.2 Related software and documentation 2
- 1.3 R and statistics 2
- 1.4 R and the window system 3
- 1.5 Using R interactively 3
- 1.6 An introductory session 4
- 1.7 Getting help with functions and features 4
- 1.8 R commands, case sensitivity, etc. 4
- 1.9 Recall and correction of previous commands 5
- 1.10 Executing commands from or diverting output to a file 5
- 1.11 Data permanency and removing objects 5

2 Simple manipulations; numbers and vectors 7
- 2.1 Vectors and assignment 7
- 2.2 Vector arithmetic 7
- 2.3 Generating regular sequences 8
- 2.4 Logical vectors 9
- 2.5 Missing values 9
- 2.6 Character vectors 10
- 2.7 Index vectors; selecting and modifying subsets of a data set 10
- 2.8 Other types of objects 11

3 Objects, their modes and attributes 13
- 3.1 Intrinsic attributes: mode and length 13
- 3.2 Changing the length of an object 14
- 3.3 Getting and setting attributes 14
- 3.4 The class of an object 14

4 Ordered and unordered factors 16
- 4.1 A specific example 16
- 4.2 The function `tapply()` and ragged arrays 16
- 4.3 Ordered factors 17

5 Arrays and matrices 18
- 5.1 Arrays 18
- 5.2 Array indexing. Subsections of an array 18
- 5.3 Index matrices 19
- 5.4 The `array()` function 20
  - 5.4.1 Mixed vector and array arithmetic. The recycling rule 20
- 5.5 The outer product of two arrays 21
- 5.6 Generalized transpose of an array 21
- 5.7 Matrix facilities 22
  - 5.7.1 Matrix multiplication 22

5.7.2 Linear equations and inversion . . . 22
5.7.3 Eigenvalues and eigenvectors . . . 23
5.7.4 Singular value decomposition and determinants . . . 23
5.7.5 Least squares fitting and the QR decomposition . . . 23
5.8 Forming partitioned matrices, `cbind()` and `rbind()` . . . 24
5.9 The concatenation function, `c()`, with arrays . . . 24
5.10 Frequency tables from factors . . . 25

6 Lists and data frames . . . 26
6.1 Lists . . . 26
6.2 Constructing and modifying lists . . . 26
6.2.1 Concatenating lists . . . 27
6.3 Data frames . . . 27
6.3.1 Making data frames . . . 27
6.3.2 `attach()` and `detach()` . . . 27
6.3.3 Working with data frames . . . 28
6.3.4 Attaching arbitrary lists . . . 28
6.3.5 Managing the search path . . . 29

7 Reading data from files . . . 30
7.1 The `read.table()` function . . . 30
7.2 The `scan()` function . . . 31
7.3 Accessing builtin datasets . . . 31
7.3.1 Loading data from other R packages . . . 31
7.4 Editing data . . . 32

8 Probability distributions . . . 33
8.1 R as a set of statistical tables . . . 33
8.2 Examining the distribution of a set of data . . . 33
8.3 One- and two-sample tests . . . 36

9 Grouping, loops and conditional execution . . . 40
9.1 Grouped expressions . . . 40
9.2 Control statements . . . 40
9.2.1 Conditional execution: `if` statements . . . 40
9.2.2 Repetitive execution: `for` loops, `repeat` and `while` . . . 40

10 Writing your own functions . . . 42
10.1 Simple examples . . . 42
10.2 Defining new binary operators . . . 43
10.3 Named arguments and defaults . . . 43
10.4 The '`...`' argument . . . 44
10.5 Assignments within functions . . . 44
10.6 More advanced examples . . . 44
10.6.1 Efficiency factors in block designs . . . 44
10.6.2 Dropping all names in a printed array . . . 45
10.6.3 Recursive numerical integration . . . 45
10.7 Scope . . . 46
10.8 Customizing the environment . . . 48
10.9 Classes, generic functions and object orientation . . . 48

11 Statistical models in R ... 50

11.1 Defining statistical models; formulae ... 50
11.1.1 Contrasts ... 52
11.2 Linear models ... 53
11.3 Generic functions for extracting model information ... 53
11.4 Analysis of variance and model comparison ... 54
11.4.1 ANOVA tables ... 54
11.5 Updating fitted models ... 54
11.6 Generalized linear models ... 55
11.6.1 Families ... 56
11.6.2 The `glm()` function ... 56
11.7 Nonlinear least squares and maximum likelihood models ... 58
11.7.1 Least squares ... 58
11.7.2 Maximum likelihood ... 59
11.8 Some non-standard models ... 60

12 Graphical procedures ... 62

12.1 High-level plotting commands ... 62
12.1.1 The `plot()` function ... 62
12.1.2 Displaying multivariate data ... 63
12.1.3 Display graphics ... 63
12.1.4 Arguments to high-level plotting functions ... 64
12.2 Low-level plotting commands ... 65
12.2.1 Mathematical annotation ... 66
12.2.2 Hershey vector fonts ... 66
12.3 Interacting with graphics ... 66
12.4 Using graphics parameters ... 67
12.4.1 Permanent changes: The `par()` function ... 67
12.4.2 Temporary changes: Arguments to graphics functions ... 68
12.5 Graphics parameters list ... 68
12.5.1 Graphical elements ... 69
12.5.2 Axes and tick marks ... 70
12.5.3 Figure margins ... 70
12.5.4 Multiple figure environment ... 72
12.6 Device drivers ... 73
12.6.1 PostScript diagrams for typeset documents ... 73
12.6.2 Multiple graphics devices ... 74
12.7 Dynamic graphics ... 75

13 Packages ... 76

13.1 Standard packages ... 76
13.2 Contributed packages and CRAN ... 76
13.3 Namespaces ... 76

Appendix A A sample session ... 78

Appendix B Invoking R ... 81

B.1 Invoking R from the command line ... 81
B.2 Invoking R under Windows ... 84
B.3 Invoking R under Mac OS X ... 85
B.4 Scripting with R ... 85

Appendix C The command-line editor 87
C.1 Preliminaries 87
C.2 Editing actions 87
C.3 Command-line editor summary 87

Appendix D Function and variable index 89

Appendix E Concept index 92

Appendix F References 94