

# Inhalt

|                      |           |
|----------------------|-----------|
| <b>Vorwort</b> ..... | <b>xi</b> |
|----------------------|-----------|

## Teil I: R-Grundlagen

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Beschaffung und Installation von R</b> .....        | <b>3</b>  |
|          | R-Versionen  | 3         |
|          | Gebrauchsfertige R-Distribution laden und installieren | 4         |
| <b>2</b> | <b>Die R-Benutzerschnittstelle</b> .....               | <b>7</b>  |
|          | Die grafische Benutzerschnittstelle von R              | 7         |
|          | Die R-Konsole  | 11        |
|          | Stapelverarbeitung                                     | 13        |
|          | Mit Microsoft Excel auf R zugreifen                    | 14        |
|          | Andere Methoden, R auszuführen                         | 15        |
| <b>3</b> | <b>Kurzeinführung in R</b> .....                       | <b>17</b> |
|          | Einfache Operationen in R                              | 17        |
|          | Funktionen   | 19        |
|          | Variablen  | 20        |
|          | Einführung in Datenstrukturen                          | 22        |
|          | Objekte und Klassen                                    | 26        |
|          | Modelle und Formeln                                    | 27        |
|          | Diagramme und Grafiken                                 | 29        |
|          | Hilfe abrufen  | 34        |
| <b>4</b> | <b>R-Pakete</b> .....                                  | <b>37</b> |
|          | Überblick über Pakete                                  | 37        |
|          | Pakete in lokalen Bibliotheken auflisten               | 38        |

|  |            |
|--|------------|
| Pakete laden                                     | 40         |
| Durchsuchen von Paket-Repositories               | 40         |
| Eigene Pakete                                    | 45         |
| <b>Teil II: Die Programmiersprache R</b>         |            |
| <b>5 Überblick über die Programmiersprache R</b> | <b>51</b>  |
| Ausdrücke  | 51         |
| Objekte  | 52         |
| Symbole  | 52         |
| Funktionen                                       | 53         |
| Objekte werden in Zuweisungen kopiert            | 54         |
| In R ist alles ein Objekt                        | 55         |
| Spezielle Werte                                  | 55         |
| Typumwandlung                                    | 56         |
| Der R-Interpreter                                | 57         |
| R-Sprachelemente inspizieren                     | 59         |
| <b>6 R-Syntax</b>                                | <b>63</b>  |
| Konstanten                                       | 63         |
| Operatoren                                       | 66         |
| Ausdrücke  | 69         |
| Kontrollstrukturen                               | 71         |
| Zugriff auf Datenstrukturen                      | 74         |
| R-Codekonventionen                               | 80         |
| <b>7 R-Objekte</b>                               | <b>83</b>  |
| Basisdatentypen                                  | 83         |
| Vektoren   | 86         |
| Listen   | 87         |
| Weitere Datentypen                               | 88         |
| Attribute  | 97         |
| <b>8 Symbole und Umgebungen</b>                  | <b>103</b> |
| Symbole  | 103        |
| Mit Umgebungen arbeiten                          | 104        |
| Die globale Umgebung                             | 105        |
| Umgebungen und Funktionen                        | 106        |
| Ausnahmen  | 110        |
| <b>9 Funktionen</b>                              | <b>113</b> |
| Funktionen deklarieren                           | 113        |
| Argumente  | 113        |

|   |            |
|---|------------|
| Rückgabewerte   | 115        |
| Funktionen als Argumente                                  | 115        |
| Benannte Argumente und Auswertungsreihenfolge             | 119        |
| Seiteneffekte   | 120        |
| <b>10 Objektorientierte Programmierung . . . . .</b>      | <b>123</b> |
| Überblick über objektorientierte Programmierung in R      | 124        |
| Objektorientierte Programmierung in R: S4-Klassen         | 131        |
| OOP nach alter Schule in R: S3                            | 138        |
| <b>11 Hochperformantes R . . . . .</b>                    | <b>143</b> |
| Interne Mathematikfunktionen benutzen                     | 143        |
| Auswertungsumgebungen als Look-up-Tabellen                | 144        |
| Große Datensätze in einer Datenbank verwalten             | 145        |
| Speicher vorbelegen                                       | 145        |
| Arbeitsspeicher verwalten                                 | 145        |
| Funktionen für große Datensätze                           | 147        |
| Paralleles Rechnen mit R                                  | 148        |
| Hochperformante R-Binaries                                | 148        |
| <b>Teil III: Daten aufbereiten und darstellen</b>         |            |
| <b>12 Daten speichern, laden und bearbeiten . . . . .</b> | <b>157</b> |
| Dateneingabe in R   | 157        |
| R-Objekte speichern und laden                             | 161        |
| Daten aus externen Dateien importieren                    | 163        |
| Daten exportieren   | 174        |
| Daten aus Datenbanken importieren                         | 174        |
| <b>13 Datenaufbereitung . . . . .</b>                     | <b>193</b> |
| Datensätze kombinieren                                    | 194        |
| Transformationen  | 200        |
| Klasseneinteilung   | 205        |
| Teilmengen  | 207        |
| Daten aggregieren   | 210        |
| Datenbereinigung  | 224        |
| Dubletten identifizieren und beseitigen                   | 225        |
| Daten sortieren   | 225        |
| <b>14 Grafik . . . . .</b>                                | <b>229</b> |
| Überblick über Grafiken in R                              | 229        |
| Grafik-Devices  | 266        |
| Grafiken anpassen   | 267        |

|   |            |
|---|------------|
| <b>15 Trellis-Grafik mit Lattice . . . . .</b>      | <b>291</b> |
| Geschichte  | 291        |
| Überblick über das lattice-Paket                    | 292        |
| High-Level-Grafikfunktionen in Lattice              | 297        |
| Lattice-Grafiken anpassen                           | 340        |
| Low-Level-Grafikfunktionen in Lattice               | 352        |
| <b>Teil IV: Statistik mit R</b>                     |            |
| <b>16 Datenexploration . . . . .</b>                | <b>357</b> |
| Lage- und Streuungsmaße                             | 357        |
| Korrelation und Kovarianz                           | 362        |
| Hauptkomponentenanalyse                             | 366        |
| Faktorenanalyse                                     | 370        |
| Bootstrap-Resampling                                | 371        |
| <b>17 Wahrscheinlichkeitsverteilungen . . . . .</b> | <b>375</b> |
| Die Normalverteilung                                | 375        |
| Argumente von Verteilungsfunktionen                 | 379        |
| Wahrscheinlichkeitsverteilungsfamilien              | 379        |
| <b>18 Statistische Testverfahren . . . . .</b>      | <b>383</b> |
| Parametrische Tests                                 | 384        |
| Verfahren für Rangdaten                             | 401        |
| Verfahren für Häufigkeitsdaten                      | 406        |
| <b>19 Poweranalyse . . . . .</b>                    | <b>417</b> |
| Beispiel: Planung einer Medikamentenstudie          | 417        |
| Planung eines <i>t</i> -Tests                       | 418        |
| Planung eines Anteilstests                          | 419        |
| Planung einer Varianzanalyse                        | 421        |
| <b>20 Regressionsmodelle . . . . .</b>              | <b>423</b> |
| Beispiel: Ein einfaches lineares Regressionsmodell  | 424        |
| Die klassische lineare Regression im Detail         | 433        |
| Prädiktorselektion und Schrumpfungsmethoden         | 440        |
| Regressionsverfahren für andere Modellannahmen      | 445        |
| Überlebensanalyse                                   | 451        |
| Glättungsverfahren                                  | 457        |
| Regression und maschinelles Lernen                  | 462        |

|  |            |
|--|------------|
| <b>21 Klassifikationsmodelle</b> . . . . .   | <b>491</b> |
| Lineare Klassifikationsmodelle               | 491        |
| Klassifikation und maschinelles Lernen       | 503        |
| <b>22 Maschinelles Lernen</b> . . . . .      | <b>513</b> |
| Warenkorbanalyse                             | 513        |
| Verfahren der Clusteranalyse                 | 519        |
| <b>23 Zeitreihenanalyse</b> . . . . .        | <b>525</b> |
| Autokorrelationsmaße                         | 525        |
| Zeitreihenmodelle                            | 526        |
| <b>24 Bioconductor</b> . . . . .             | <b>531</b> |
| Ein Beispiel                                 | 532        |
| Wichtige Bioconductor-Pakete                 | 546        |
| Datenstrukturen                              | 548        |
| Wie geht es weiter?                          | 554        |
| <b>Anhang A: Funktionsreferenz</b> . . . . . | <b>557</b> |
| <b>Anhang B: Datensatzreferenz</b> . . . . . | <b>705</b> |
| <b>Literatur</b> . . . . .                   | <b>729</b> |
| <b>Index</b> . . . . .                       | <b>731</b> |