

Inhaltsverzeichnis

1	Erste Schritte	1
1.1	Vorstellung	1
1.1.1	Was ist R?	1
1.1.2	Typographische Konventionen	3
1.1.3	Wo erhalte ich R und Dokumentation zu R?	4
1.1.4	Installation von R unter Windows	5
1.2	Grundlegende Elemente	6
1.2.1	Starten und beenden, die Konsole	6
1.2.2	Einstellungen	9
1.2.3	Umgang mit dem Workspace	10
1.2.4	Einfache Arithmetik	11
1.2.5	Funktionen mit Argumenten aufrufen	13
1.2.6	Hilfe-Funktionen	14
1.2.7	Zusatzpakete verwenden	15
1.3	Datenstrukturen: Klassen, Objekte, Datentypen	17
1.3.1	Objekte benennen	18
1.3.2	Zuweisungen an Objekte	19
1.3.3	Objekte ausgeben	20
1.3.4	Objekte anzeigen lassen und entfernen	20
1.3.5	Datentypen	20
1.3.6	Logische Werte, Operatoren und Verknüpfungen	22
2	Elementare Dateneingabe und -verarbeitung	25
2.1	Vektoren	25
2.1.1	Vektoren erzeugen	25
2.1.2	Elemente auswählen und verändern	26
2.1.3	Datentypen in Vektoren	28
2.1.4	Reihenfolge von Elementen kontrollieren	29
2.1.5	Elemente benennen	30
2.1.6	Elemente löschen	31
2.1.7	Rechenoperationen mit Vektoren	31
2.2	Logische Operatoren	35

2.2.1	Logische Operatoren zum Vergleich von Vektoren	35
2.2.2	Logische Indexvektoren	37
2.2.3	Werte ersetzen oder recodieren.	38
2.3	Mengen.	40
2.3.1	Duplizierte Werte behandeln.	40
2.3.2	Mengenoperationen	41
2.3.3	Kombinatorik	42
2.4	Numerische Sequenzen und feste Wertefolgen erzeugen	44
2.4.1	Numerische Sequenzen erstellen	44
2.4.2	Wertefolgen wiederholen	46
2.5	Zufallszahlen und zufällige Reihenfolgen generieren	46
2.5.1	Zufällig aus einer Urne ziehen	47
2.5.2	Zufallszahlen aus bestimmten Verteilungen erzeugen . .	47
2.5.3	Unterauswahl einer Datenmenge bilden	48
2.5.4	Zufällige Reihenfolgen erstellen	49
2.6	Deskriptive Kennwerte numerischer Vektoren	49
2.6.1	Summen, Differenzen und Produkte	50
2.6.2	Extremwerte	51
2.6.3	Mittelwert, Median und Modalwert	52
2.6.4	Quartile, Quantile, Interquartilabstand	53
2.6.5	Varianz, Streuung, Schiefe und Wölbung	54
2.6.6	Kovarianz, Korrelation und Partialkorrelation	55
2.6.7	Funktionen auf geordnete Paare von Werten anwenden.	57
2.7	Gruppierungsfaktoren	57
2.7.1	Ungeordnete Faktoren	58
2.7.2	Faktorstufen hinzufügen, entfernen und zusammenfassen	60
2.7.3	Geordnete Faktoren	61
2.7.4	Reihenfolge von Faktorstufen.	61
2.7.5	Faktoren nach Muster erstellen.	63
2.7.6	Quantitative Variablen in Faktoren umwandeln	64
2.7.7	Funktionen getrennt nach Gruppen anwenden	64
2.8	Matrizen	66
2.8.1	Datentypen in Matrizen	67
2.8.2	Dimensionierung, Zeilen und Spalten	68
2.8.3	Elemente auswählen und verändern	70
2.8.4	Weitere Wege, um Elemente auszuwählen und zu verändern	71
2.8.5	Matrizen verbinden	72
2.8.6	Randkennwerte	73
2.8.7	Beliebige Funktionen auf Matrizen anwenden	73
2.8.8	Matrix zeilen- oder spaltenweise mit Kennwerten verrechnen	74
2.8.9	Kovarianz- und Korrelationsmatrizen	75
2.8.10	Matrizen sortieren	76

2.9	Lineare Algebra	78
2.9.1	Matrix-Algebra	78
2.9.2	Lineare Gleichungssysteme lösen	81
2.9.3	Norm und Abstand von Vektoren und Matrizen	81
2.9.4	Orthogonale Projektion	85
2.9.5	Kennwerte und Zerlegungen von Matrizen	86
2.10	Arrays	90
2.11	Häufigkeitsauszählungen	91
2.11.1	Einfache Tabellen absoluter und relativer Häufigkeiten	91
2.11.2	Häufigkeiten natürlicher Zahlen	93
2.11.3	Iterationen zählen	94
2.11.4	Absolute, relative und bedingte relative Häufigkeiten in Kreuztabellen	94
2.11.5	Randkennwerte von Kreuztabellen	97
2.11.6	Kumulierte relative Häufigkeiten und Prozentrang	98
2.12	Codierung, Identifikation und Behandlung fehlender Werte	100
2.12.1	Fehlende Werte codieren und ihr Vorhandensein prüfen	101
2.12.2	Fehlende Werte ersetzen oder umcodieren	102
2.12.3	Behandlung fehlender Werte bei der Berechnung einfacher Kennwerte	103
2.12.4	Behandlung fehlender Werte in Matrizen	104
2.12.5	Behandlung fehlender Werte beim Sortieren von Daten	106
2.13	Zeichenketten verarbeiten	106
2.13.1	Objekte in Zeichenketten umwandeln	107
2.13.2	Zeichenketten erstellen und ausgeben	107
2.13.3	Zeichenketten manipulieren	109
2.13.4	Zeichenfolgen finden	111
2.13.5	Zeichenfolgen ersetzen	112
2.13.6	Zeichenketten als Befehl ausführen	113
2.14	Datum und Uhrzeit	113
2.14.1	Datum	114
2.14.2	Uhrzeit	115
3	Datensätze	117
3.1	Listen	117
3.1.1	Komponenten auswählen und verändern	117
3.1.2	Listen mit mehreren Ebenen	120
3.2	Datensätze	121
3.2.1	Datentypen in Datensätzen	123
3.2.2	Elemente auswählen und verändern	124
3.2.3	Datensätze in den Suchpfad einfügen	126
3.2.4	Namen von Variablen und Beobachtungen	127
3.2.5	Variablen einem Datensatz hinzufügen oder aus diesem entfernen	128

3.2.6	Teilmengen von Daten auswählen	130
3.2.7	Organisationsform von Daten ändern	133
3.2.8	Organisationsform eines Datensatzes ändern	134
3.2.9	Datensätze teilen	138
3.2.10	Datensätze zusammenfügen	139
3.2.11	Funktionen auf Variablen anwenden	143
3.2.12	Funktionen für mehrere Variablen anwenden	145
3.2.13	Funktionen getrennt nach Gruppen anwenden	146
3.2.14	Doppelte und fehlende Werte	147
3.2.15	Datensätze sortieren	149
4	Befehle und Daten verwalten	151
4.1	Befehlssequenzen im Editor bearbeiten	151
4.2	Daten importieren und exportieren	153
4.2.1	Daten in der Konsole einlesen	153
4.2.2	Daten im Editor eingeben	154
4.2.3	Im Textformat gespeicherte Daten	155
4.2.4	R-Objekte	157
4.2.5	Daten mit anderen Programmen austauschen	157
5	Hilfsmittel für die Inferenzstatistik	163
5.1	Lineare Modelle formulieren	163
5.2	Funktionen von Zufallsvariablen	165
5.2.1	Dichtefunktionen	166
5.2.2	Verteilungsfunktionen	167
5.2.3	Quantilfunktionen	168
5.3	Behandlung fehlender Werte in inferenzstatistischen Tests	169
6	Nonparametrische Methoden	171
6.1	Anpassungstests	171
6.1.1	Binomialtest	172
6.1.2	Test auf Zufälligkeit (Runs Test)	174
6.1.3	Kolmogorov-Smirnov-Anpassungstest	176
6.1.4	χ^2 -Test auf eine feste Verteilung	179
6.1.5	χ^2 -Test auf eine Verteilungsklasse	180
6.2	Analyse von gemeinsamen Häufigkeiten kategorialer Variablen	181
6.2.1	χ^2 -Test auf Unabhängigkeit	182
6.2.2	χ^2 -Test auf Gleichheit von Verteilungen	183
6.2.3	χ^2 -Test für mehrere Auftretenswahrscheinlichkeiten	184
6.2.4	Fishers exakter Test auf Unabhängigkeit	185
6.2.5	Fishers exakter Test auf Gleichheit von Verteilungen	186
6.2.6	Kennwerte für 2×2 Kontingenztafeln	187
6.3	Maße für Zusammenhang und Übereinstimmung kategorialer Daten	189

6.3.1	Zusammenhang ordinaler Variablen: Spearmans ρ und Kendalls τ	189
6.3.2	Weitere Zusammenhangsmaße: φ , Cramérs V , Kontingenzkoeffizient, Goodman und Kruskals γ , Somers' d , ICC, r_{WG}	191
6.3.3	Inter-Rater-Übereinstimmung: Cohens κ , Fleiss' κ , Kendalls W	192
6.4	Tests auf Übereinstimmung von Verteilungen	199
6.4.1	Vorzeichen-Test	199
6.4.2	Wilcoxon-Vorzeichen-Rang-Test für eine Stichprobe	200
6.4.3	Wald-Wolfowitz-Test für zwei Stichproben	203
6.4.4	Kolmogorov-Smirnov-Test für zwei Stichproben	203
6.4.5	Wilcoxon-Rangsummen-Test/Mann-Whitney-U-Test für zwei unabhängige Stichproben	205
6.4.6	Wilcoxon-Test für zwei abhängige Stichproben	207
6.4.7	Kruskal-Wallis-H-Test für unabhängige Stichproben	207
6.4.8	Friedman-Rangsummen-Test für abhängige Stichproben	208
6.4.9	Cochran-Q-Test für abhängige Stichproben	210
6.4.10	Bowker-Test für zwei abhängige Stichproben	211
6.4.11	McNemar-Test für zwei abhängige Stichproben	213
6.4.12	Stuart-Maxwell-Test für zwei abhängige Stichproben	214
7	Korrelation und Regressionsanalyse	217
7.1	Test auf Korrelation	217
7.2	Einfache lineare Regression	218
7.2.1	Regressionsanalyse	221
7.2.2	Regressionsdiagnostik	224
7.2.3	Vorhersage bei Anwendung auf andere Daten	225
7.2.4	Partialkorrelation	227
7.2.5	Kreuzvalidierung	227
7.3	Multiple lineare Regression	228
7.3.1	Modell verändern	230
7.3.2	Modell auswählen	231
7.3.3	Auf Multikollinearität prüfen	232
7.4	Logistische Regression	233
8	Parametrische Tests für Dispersions- und Lageparameter von Verteilungen	239
8.1	Tests auf Varianzhomogenität	239
8.1.1	F -Test auf Varianzhomogenität bei zwei Stichproben	239
8.1.2	Fligner-Killeen-Test und Bartlett-Test	241
8.1.3	Levene-Test	241
8.2	t -Tests	242
8.2.1	t -Test für eine Stichprobe	242

8.2.2	<i>t</i> -Test für zwei unabhängige Stichproben	243
8.2.3	<i>t</i> -Test für zwei abhängige Stichproben	245
8.3	Einfaktorielle Varianzanalyse (CR- <i>p</i>)	246
8.3.1	Regression und Varianzanalyse als lineare Modelle	246
8.3.2	Auswertung mit <code>oneway.test()</code>	248
8.3.3	Auswertung mit <code>aov()</code>	249
8.3.4	Graphische Prüfung der Voraussetzungen	251
8.3.5	Auswertung mit <code>anova()</code>	252
8.3.6	Einzelvergleiche (Kontraste)	254
8.4	Einfaktorielle Varianzanalyse mit abhängigen Gruppen (RB- <i>p</i>)	261
8.4.1	Univariat formulierte Auswertung mit <code>aov()</code>	262
8.4.2	Zirkularität der Kovarianzmatrix prüfen	264
8.4.3	Multivariat formulierte Auswertung mit <code>Anova()</code>	266
8.4.4	Einzelvergleiche (Kontraste)	268
8.5	Zweifaktorielle Varianzanalyse (CRF- <i>pq</i>)	268
8.5.1	Auswertung mit <code>aov()</code>	269
8.5.2	Quadratsummen vom Typ I und III	271
8.5.3	Beliebige a-priori Kontraste	275
8.5.4	Beliebige post-hoc Kontraste nach Scheffé	278
8.6	Zweifaktorielle Varianzanalyse mit zwei Intra-Gruppen Faktoren (RBF- <i>pq</i>)	279
8.6.1	Univariat formulierte Auswertung mit <code>aov()</code>	280
8.6.2	Zirkularität der Kovarianzmatrizen prüfen	283
8.6.3	Multivariat formulierte Auswertung mit <code>Anova()</code>	284
8.6.4	Einzelvergleiche (Kontraste)	285
8.7	Zweifaktorielle Varianzanalyse mit Split-Plot-Design (SPF- <i>p · q</i>)	286
8.7.1	Univariat formulierte Auswertung mit <code>aov()</code>	286
8.7.2	Voraussetzungen und Prüfen der Zirkularität	287
8.7.3	Multivariat formulierte Auswertung mit <code>Anova()</code>	288
8.7.4	Einzelvergleiche (Kontraste)	289
8.7.5	Erweiterung auf dreifaktorielles SPF- <i>p · qr</i> Design	290
8.7.6	Erweiterung auf dreifaktorielles SPF- <i>pq · r</i> Design	291
8.8	Kovarianzanalyse	292
8.8.1	Test der Effekte von Gruppenzugehörigkeit und Kovariaten	292
8.8.2	Beliebige a-priori Kontraste	297
8.8.3	Beliebige post-hoc Kontraste nach Scheffé	299
8.9	Power und notwendige Stichprobengrößen berechnen	299
8.9.1	Binomialtest	300
8.9.2	<i>t</i> -Test	301
8.9.3	Einfaktorielle Varianzanalyse	304
9	Multivariate Verfahren	309
9.1	Multivariate Multiple Regression	309

9.2	Hauptkomponentenanalyse	311
9.3	Faktorenanalyse	317
9.4	Multidimensionale Skalierung	323
9.5	Hotellings T^2	325
9.5.1	Test für eine Stichprobe	325
9.5.2	Test für zwei Stichproben	327
9.6	Multivariate Varianzanalyse (MANOVA)	329
9.6.1	Einfaktorielle MANOVA	329
9.6.2	Mehr faktorielle MANOVA	332
10	Diagramme erstellen	335
10.1	Graphik-Devices	335
10.1.1	Aufbau und Verwaltung von Graphik-Devices	335
10.1.2	Graphiken speichern	337
10.2	Streu- und Liniendiagramme	339
10.2.1	Streudiagramme mit <code>plot()</code>	339
10.2.2	Datenpunkte eines Streudiagramms identifizieren	342
10.2.3	Streudiagramme mit <code>matplotlib()</code>	342
10.3	Diagramme formatieren	343
10.3.1	Graphikelemente formatieren	343
10.3.2	Farben spezifizieren	346
10.3.3	Achsen formatieren	348
10.4	Säulen- und Punktdiagramme	348
10.4.1	Einfache Säulendiagramme	349
10.4.2	Gruppierte und gestapelte Säulendiagramme	350
10.4.3	Cleveland Dotchart	352
10.5	Elemente einem bestehenden Diagramm hinzufügen	354
10.5.1	Koordinaten in einem Diagramm identifizieren	354
10.5.2	Punkte	355
10.5.3	Linien	355
10.5.4	Polygone	358
10.5.5	Funktionsgraphen	362
10.5.6	Text und mathematische Formeln	362
10.5.7	Achsen	365
10.5.8	Fehlerbalken	366
10.6	Verteilungsdiagramme	369
10.6.1	Histogramm und Schätzung der Dichtefunktion	369
10.6.2	Stamm-Blatt-Diagramm	371
10.6.3	Box-Whisker-Plot	372
10.6.4	<code>stripchart()</code>	374
10.6.5	Quantil-Quantil-Diagramme	375
10.6.6	Empirische kumulative Häufigkeitsverteilungen	376
10.6.7	Kreisdiagramm	377
10.6.8	Gemeinsame Verteilung zweier Variablen	378

10.7	Datenpunkte interpolieren	381
10.7.1	Lineare Interpolation und polynomiale Glätter	381
10.7.2	Splines	382
10.8	Multivariate Daten visualisieren	383
10.8.1	Höhenlinien und variable Datenpunktssymbole	384
10.8.2	Dreidimensionale Gitter und Streudiagramme	387
10.8.3	Simultane Darstellung mehrerer Diagramme gleichen Typs	388
10.8.4	Matrix aus Streudiagrammen	390
10.9	Mehrere Diagramme in einem Graphik-Device darstellen	393
10.9.1	<code>layout()</code>	393
10.9.2	<code>par(mfrow, mfcoll, fig)</code>	395
10.9.3	<code>split.screen()</code>	398
11	R als Programmiersprache	401
11.1	Eigene Funktionen erstellen	401
11.1.1	Funktionskopf mit Argumentliste	402
11.1.2	Funktionsrumpf mit Befehlen	403
11.1.3	Rückgabewert	403
11.1.4	Eigene Funktionen verwenden	404
11.1.5	Generische Funktionen	405
11.2	Kontrollstrukturen	407
11.2.1	Bedingungen zur Fallunterscheidung prüfen	407
11.2.2	Schleifen	409
	Literaturverzeichnis	413
	Sachverzeichnis	419