

Reiner Hellbrück

# **Angewandte Statistik mit R**

Eine Einführung für Ökonomen  
und Sozialwissenschaftler



# Inhaltsverzeichnis

<b>Vorwort</b>	<b>v</b>
<b>Abbildungsverzeichnis</b>	<b>xv</b>
<b>Tabellenverzeichnis</b>	<b>xvii</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Gegenstand . . . . .	1
1.2 Aufbau . . . . .	4
<b>2 Datenerhebung - ganz praktisch</b>	<b>7</b>
2.1 Einleitung . . . . .	7
2.2 Erhebungsplan . . . . .	7
2.2.1 Grundlagen . . . . .	7
2.2.2 Beispiel . . . . .	8
2.3 Software . . . . .	10
2.3.1 Moodle und LimeSurvey . . . . .	10
2.3.2 Statistikpaket R . . . . .	11
2.4 Ziehen einer Stichprobe . . . . .	12
2.4.1 Grundlagen . . . . .	12
2.4.2 Beispiel . . . . .	13
2.5 Rohdaten auslesen . . . . .	13
2.5.1 Grundlagen . . . . .	13
2.5.2 Beispiel . . . . .	14
2.6 Daten in Statistikprogramm einlesen . . . . .	15
2.6.1 Grundlagen . . . . .	15
2.6.2 Beispiel . . . . .	15
2.7 Plausibilitätsprüfung . . . . .	17
2.7.1 Grundlagen . . . . .	17
2.7.2 Beispiel 1 . . . . .	18
2.7.3 Einfache Datensätze . . . . .	19
2.7.4 Beispiel 2 . . . . .	20
2.7.5 komplexe Datensätze . . . . .	21
2.7.6 Beispiel 3 . . . . .	23

2.8	Abschließende Bemerkungen . . . . .	24
2.9	Kontrollfragen . . . . .	25
2.10	Aufgaben . . . . .	27
<b>3</b>	<b>Datenaufbereitung</b>	<b>29</b>
3.1	Einleitung . . . . .	29
3.2	Graphische Methoden . . . . .	30
3.2.1	Grundlagen . . . . .	30
3.2.2	Beispiele . . . . .	32
3.3	Absolute Häufigkeitsverteilung . . . . .	36
3.3.1	Grundlagen . . . . .	36
3.3.2	Beispiel 1 . . . . .	36
3.3.3	Maßzahlen . . . . .	37
3.3.4	Beispiel 2 . . . . .	38
3.4	Relative Häufigkeitsverteilung . . . . .	40
3.4.1	Grundlagen . . . . .	40
3.4.2	Beispiel 1 . . . . .	41
3.4.3	Maßzahlen . . . . .	42
3.4.4	Beispiel 2 . . . . .	43
3.5	Verteilungsfunktion und Quantile . . . . .	43
3.5.1	Verteilungsfunktion . . . . .	43
3.5.2	Quantile . . . . .	45
3.5.3	Verteilungsfunktion und Quantile . . . . .	48
3.6	Histogramme . . . . .	50
3.6.1	Absolute Häufigkeit . . . . .	50
3.6.2	Durchschnittliche Häufigkeitsdichte . . . . .	51
3.7	Kontingenztabelle . . . . .	53
3.7.1	Gemeinsame Verteilung . . . . .	53
3.7.2	Randverteilungen . . . . .	55
3.7.3	Bedingte Verteilung und statistische Unabhängigkeit . . . . .	56
3.8	Lorenz-Kurve . . . . .	57
3.8.1	Grundlagen . . . . .	57
3.8.2	Beispiel . . . . .	58
3.8.3	Gini-Koeffizienten . . . . .	60
3.9	Abschließende Bemerkungen . . . . .	63
3.10	Kontrollfragen . . . . .	63
3.11	Aufgaben . . . . .	64
3.A	Nützliches zu Maßzahlen* . . . . .	67
<b>4</b>	<b>Statistisches Testen</b>	<b>69</b>
4.1	Einleitung . . . . .	69
4.2	Binomialverteilung . . . . .	70
4.2.1	Grundlagen . . . . .	70
4.2.2	Beispiel . . . . .	72
4.3	Test . . . . .	73
4.3.1	Zweiseitige Fragestellung . . . . .	73

4.3.2	Einseitige Fragestellung - Version 1 . . . . .	76
4.3.3	Einseitige Fragestellung - Version 2 . . . . .	78
4.3.4	Fehler 1. Art . . . . .	79
4.3.5	Beispiel . . . . .	79
4.4	Abschließende Bemerkungen . . . . .	82
4.5	Kontrollfragen . . . . .	83
4.6	Aufgaben . . . . .	84
4.A	Wirkungsanalyse* . . . . .	87
4.A.1	Grundlagen . . . . .	87
4.A.2	Test . . . . .	88
4.A.3	Beispiel . . . . .	89
4.A.4	Abschließende Bemerkungen . . . . .	90
<b>5</b>	<b>Chi-Quadrat Tests</b>	<b>91</b>
5.1	Einleitung . . . . .	91
5.2	Unabhängigkeitstest . . . . .	92
5.2.1	Grundlagen . . . . .	92
5.2.2	Beispiel . . . . .	94
5.3	Anpassungstest . . . . .	97
5.3.1	Grundlagen . . . . .	97
5.3.2	Beispiel . . . . .	98
5.4	Homogenitätstest . . . . .	99
5.4.1	Grundlagen . . . . .	99
5.4.2	Beispiel . . . . .	101
5.5	Abschließende Bemerkungen . . . . .	102
5.6	Kontrollfragen . . . . .	103
5.7	Aufgaben . . . . .	104
<b>6</b>	<b>Wahrscheinlichkeitsräume</b>	<b>107</b>
6.1	Einleitung . . . . .	107
6.2	Definitionsmenge . . . . .	108
6.3	Wahrscheinlichkeitsraum der Grundgesamtheit . . . . .	110
6.4	Wahrscheinlichkeitsraum der Stichprobe . . . . .	113
6.5	Wichtige Zusammenhänge und Begriffe . . . . .	115
6.5.1	Rechenregeln . . . . .	115
6.5.2	Bedingte Wahrscheinlichkeit . . . . .	116
6.5.3	Stochastische Unabhängigkeit . . . . .	116
6.5.4	Multiplikationssatz . . . . .	117
6.5.5	Satz von der totalen Wahrscheinlichkeit . . . . .	117
6.5.6	Satz von Bayes . . . . .	118
6.5.7	Diskreter Wahrscheinlichkeitsraum . . . . .	120
6.6	Abschließende Bemerkungen . . . . .	120
6.7	Kontrollfragen . . . . .	121
6.8	Aufgaben . . . . .	122

<b>Abbildungen von Ergebnisräumen</b>	<b>125</b>
7.1 Einleitung . . . . .	125
7.2 Meßbarkeit und Zufallsvariable . . . . .	126
7.2.1 Meßbarkeit . . . . .	126
7.2.2 Zufallsvariablen . . . . .	127
7.3 Verteilungsfunktion und Dichte . . . . .	128
7.3.1 Verteilungsfunktion . . . . .	128
7.3.2 Dichte . . . . .	130
7.4 Maßzahlen . . . . .	131
7.4.1 Erwartungswert . . . . .	131
7.4.2 Kovarianz, Varianz und Standardabweichung . . . . .	132
7.4.3 Standardisierung . . . . .	133
7.5 Abschließende Bemerkungen . . . . .	133
7.6 Kontrollfragen . . . . .	134
7.7 Aufgaben . . . . .	135
<b>Einfache Korrelationsanalyse</b>	<b>137</b>
8.1 Einleitung . . . . .	137
8.2 Korrelation . . . . .	138
8.2.1 Wahrscheinlichkeitstheorie . . . . .	138
8.2.2 Empirische Korrelation . . . . .	139
8.2.3 Berechnung bei Wertepaaren . . . . .	140
8.2.4 Beispiele . . . . .	140
8.3 Tests bei kardinalen Merkmalen . . . . .	143
8.3.1 Stetige normalverteilte Zufallsvariablen . . . . .	143
8.3.2 Stetige nicht-normalverteilte Zufallsvariablen . . . . .	147
8.4 Test bei ordinalen Merkmalen: Bell-Doksum Test . . . . .	154
8.4.1 Test . . . . .	154
8.4.2 Beispiel . . . . .	155
8.5 Abschließende Bemerkungen . . . . .	158
8.6 Kontrollfragen . . . . .	159
8.7 Aufgaben . . . . .	160
8.A Weitere Tests* . . . . .	161
<b>Multivariate Korrelationsanalyse*</b>	<b>165</b>
9.1 Einleitung . . . . .	165
9.2 Vergleich zweier Korrelationen . . . . .	166
9.2.1 Grundlagen . . . . .	166
9.2.2 Beispiel . . . . .	167
9.3 Partielle Korrelation . . . . .	168
9.3.1 Grundlagen . . . . .	168
9.3.2 Beispiel 1 . . . . .	169
9.3.3 Test . . . . .	169
9.3.4 Beispiel 2 . . . . .	170
9.4 Zusammenhang zwischen mehreren Merkmalen . . . . .	170
9.4.1 Grundlagen . . . . .	170

9.4.2	Beispiel . . . . .	172
9.5	Globaltest . . . . .	173
9.5.1	Test . . . . .	173
9.5.2	Beispiel . . . . .	173
9.6	Multiple Vergleiche . . . . .	174
9.6.1	Test . . . . .	174
9.6.2	Beispiel . . . . .	176
9.7	Multiple Korrelation . . . . .	179
9.7.1	Grundlagen . . . . .	179
9.7.2	Beispiel 1. . . . .	179
9.7.3	Test . . . . .	180
9.7.4	Beispiel 2. . . . .	181
9.8	Kanonische Korrelation . . . . .	182
9.8.1	Grundlagen . . . . .	182
9.8.2	Beispiel 1. . . . .	183
9.8.3	Test . . . . .	184
9.8.4	Beispiel 2. . . . .	185
9.9	Abschließende Bemerkungen . . . . .	186
9.10	Kontrollfragen . . . . .	187
9.11	Aufgaben . . . . .	188
<b>10</b>	<b>Daten- und Distanzmatrix</b>	<b>189</b>
10.1	Einleitung . . . . .	189
10.2	Distanzmatrizen . . . . .	191
10.2.1	Definition und Eigenschaften . . . . .	191
10.2.2	Skalierung . . . . .	192
10.3	Kardinale Merkmale . . . . .	192
10.3.1	Intervall- und Verhältnisskala . . . . .	192
10.3.2	Manhattan-Distanz . . . . .	194
10.4	Ordinale Merkmale . . . . .	198
10.4.1	Grundlagen . . . . .	198
10.4.2	Beispiel . . . . .	199
10.5	Nominale Merkmale . . . . .	202
10.5.1	Grundlagen . . . . .	202
10.5.2	Beispiel . . . . .	202
10.6	Binäre Merkmale . . . . .	203
10.6.1	Grundlagen . . . . .	203
10.6.2	Beispiel . . . . .	205
10.7	Abschließende Bemerkungen . . . . .	206
10.8	Kontrollfragen . . . . .	207
10.9	Aufgaben . . . . .	207

<b>11 Clusteranalyse</b>	<b>211</b>
11.1 Einleitung	211
11.2 Klassifikation	214
11.2.1 Klassifikationstypen	214
11.2.2 Konstruktionsverfahren	215
11.3 PAM	216
11.3.1 Grundlagen	216
11.3.2 Beispiel 1	216
11.3.3 Bestimmung der Medoiden*	220
11.3.4 Beispiel 2	223
11.3.5 Isolierte Cluster	224
11.3.6 Beispiel 3	225
11.3.7 Überprüfung der Klassenbildung	226
11.3.8 Beispiel 4	228
11.3.9 Bestimmung der Klassenzahl	229
11.3.10 Beispiel 5	229
11.4 FANNY	229
11.4.1 Grundlagen	229
11.4.2 Beispiel 1	231
11.4.3 Partition und Überdeckung	232
11.4.4 Beispiel 2	233
11.4.5 Überprüfung der Klassenbildung und Klassenanzahl	236
11.4.6 Beispiel 3	237
11.5 MONA	237
11.5.1 Grundlagen	237
11.5.2 Beispiel 1	238
11.5.3 Assoziationsmaß	240
11.5.4 Beispiel 2	241
11.5.5 Missings	244
11.5.6 Beispiel 3	244
11.6 Abschließende Bemerkungen	245
11.7 Kontrollfragen	246
11.8 Aufgaben	247
<b>12 Einfache Regression</b>	<b>249</b>
12.1 Einleitung	249
12.2 Einfaches klassisches Regressionsmodell	250
12.2.1 Grundlagen	250
12.2.2 Beispiel	253
12.3 Regressionsfunktion	255
12.3.1 Grundlagen	255
12.3.2 Beispiel	256
12.4 Prognose	258
12.4.1 Grundlagen	258
12.4.2 Beispiel	258
12.5 Bestimmtheitsmaß	259

12.5.1	Grundlagen	259
12.5.2	Beispiel	262
12.6	Vollständiges Modell	263
12.7	Tests	264
12.7.1	Grundlagen	264
12.7.2	Beispiel	266
12.8	Abschließende Bemerkungen	269
12.9	Kontrollfragen	269
12.10	Aufgaben	271
12. A	Beweis der Streuungszzerlegungsformel*	272
12.B	Erwartungswerte der KQ-Koeffizienten*	274
12.G	Standardisierung*	274
12.C.1	Erwartungswert	274
12.C.2	Varianz	275
12.D	Partielle Korrelation*	277
<b>A</b>	<b>Theoretische Verteilungen</b>	<b>279</b>
A.1	Einleitung	279
A.2	Diskrete Verteilungen	280
A.2.1	Gleichverteilung*	280
A.2.2	Bernoulli- und Binomialverteilung	281
A.2.3	Hypergeometrische Verteilung*	282
A.2.4	Poisson-Verteilung*	283
A.2.5	Geometrische Verteilung*	285
A.3	Stetige Verteilungen	287
A.3.1	Rechteckverteilung	287
A.3.2	Exponentialverteilung*	289
A.3.3	Normalverteilung	290
A.3.4	Chi-Quadrat-Verteilung	295
A.3.5	t-Verteilung	296
A.3.6	F-Verteilung	299
<b>B</b>	<b>Matrizenrechnung</b>	<b>303</b>
B.1	Einleitung	303
B.2	Matrizen	304
B.2.1	Definition	304
B.2.2	Vektoren	304
B.2.3	Typen	305
B.3	Verknüpfungen	306
B.3.1	Gleichheitsrelation	306
B.3.2	Addition	307
B.3.3	Skalare Multiplikation	308
B.3.4	Produkt zweier Matrizen	310
B.3.5	Multiplikation von Vektoren	312
B.4	Unabhängigkeit, Rang, Determinante, Inverse	313
B.4.1	Lineare Unabhängigkeit	313



B.4.2	Rang	313
B.4.3	Determinante	314
B.4.4	Inverse	316
B.5	Eigenwerte, Eigenvektoren und Spur	319
B.5.1	Definitionen	319
B.5.2	Rechenregel	319
B.5.3	Beispiele	319
<b>C</b>	<b>Befehle in R</b>	<b>321</b>
C.1	Einleitung	321
C.2	Grundlagen	322
C.3	Daten einlesen, Objekte speichern und laden	323
C.4	Dateneigenschaften	325
C.5	Manipulation eingelesener Datensätze	327
C.6	Graphik	328
C.7	Suchen und Finden <sup>1</sup>	329
C.8	Besonderheiten in Windows	331
C.9	Fehlermeldungen	332
	<b>Lösungen</b>	<b>334</b>
	<b>Glossar</b>	<b>347</b>
	<b>Literaturverzeichnis</b>	<b>351</b>
	<b>Stichwortverzeichnis</b>	<b>353</b>