
Andreas Handl · Torben Kuhlenkasper

Einführung in die Statistik

Theorie und Praxis mit R

 Springer Spektrum

Inhaltsverzeichnis

Teil I Grundlagen

1	Grundbegriffe	3
1.1	Erhebungsarten	5
1.2	Merkmale	8
2	Eine kleine Einführung in R	13
2.1	R als mächtiger Taschenrechner	13
2.2	Datenstrukturen	16
2.3	Einlesen von Daten aus externen Dateien	24
2.4	Selektion unter Bedingungen	26
2.5	Grafiken in R	30
2.6	Pakete	38

Teil II Deskriptive Statistik

3	Univariate Analyse	43
3.1	Darstellung univariater Datensätze	43
3.1.1	Darstellung qualitativer Merkmale	44
3.1.2	Darstellung qualitativer Merkmale in R	54
3.1.3	Darstellung quantitativer Merkmale	56
3.1.4	Darstellung quantitativer Merkmale in R	72
3.1.5	Wie schreibt man eine Funktion in R?	76
3.2	Beschreibung univariater Datensätze	81
3.2.1	Maßzahlen für die Lage	82
3.2.2	Maßzahlen für die Lage in R	97
3.2.3	Quantile	99
3.2.4	Maßzahlen für die Variabilität	102
3.2.5	Konzentrationsmessung	111
3.2.6	Boxplot	115
3.3	Mathematischer Anhang und Beweise	120
3.3.1	Summenzeichen	120

3.3.2	Wie bestimmt man eine Gerade aus zwei Punkten?	123
3.3.3	Beweise	125
3.4	Datensätze	127
4	Multivariate Analyse	129
4.1	Quantitativ und qualitativ	129
4.1.1	Analyse mit R	135
4.2	Zwei qualitative Merkmale	138
4.2.1	Kontingenztafel	138
4.2.2	Bedingte relative Häufigkeiten	140
4.2.3	Kontingenzkoeffizient	146
4.2.4	Analyse in R	152
4.3	Korrelationen	158
4.3.1	Streudiagramm	158
4.3.2	Korrelationskoeffizient von Bravais-Pearson	161
4.3.3	Rangkorrelationskoeffizient von Spearman	168
4.3.4	Zur Interpretation von Korrelation	172
4.3.5	Analyse in R	174
5	Wahrscheinlichkeitsrechnung	179
5.1	Zufallsvorgänge und Ereignisse	179
5.1.1	Operationen zwischen Ereignissen	181
5.2	Wahrscheinlichkeit	184
5.2.1	Klassischer Ansatz	184
5.2.2	Frequentistischer Ansatz	185
5.2.3	Axiomatische Definition	187
5.2.4	Kombinatorik	193
5.2.5	Bedingte Wahrscheinlichkeit	206
5.2.6	Multiplikationssätze	209
5.2.7	Satz von der totalen Wahrscheinlichkeit	211
5.2.8	Satz von Bayes	215
5.2.9	Unabhängigkeit	217
5.3	Wahrscheinlichkeitsrechnung mit R	220
 Teil III Induktive Statistik		
6	Univariate Zufallsvariablen	223
6.1	Diskrete Zufallsvariablen	223
6.2	Stetige Zufallsvariablen	231
7	Verteilungsparameter	237
7.1	Erwartungswert	237
7.1.1	Diskrete Zufallsvariablen	237
7.1.2	Stetige Zufallsvariablen	238

7.1.3	Erwartungswerte von Funktionen von Zufallsvariablen	239
7.1.4	Eigenschaften des Erwartungswerts	240
7.2	Varianz	242
7.3	Tschebyscheff-Ungleichung	246
7.4	Quantile	248
8	Multivariate Zufallsvariablen	251
8.1	Diskrete Zufallsvariablen	251
8.2	Stetige Zufallsvariablen	255
8.3	Unabhängigkeit	256
8.4	Funktionen von Zufallsvariablen	259
9	Parameter multivariater Verteilungen	261
9.1	Erwartungswerte	261
9.2	Kovarianz und Korrelationskoeffizient	263
10	Verteilungsmodelle	273
10.1	Diskrete Verteilungsmodelle	273
10.1.1	Gleichverteilung	273
10.1.2	Vom Bernoulli-Prozess abgeleitete Verteilungen	275
10.1.3	Hypergeometrische Verteilung	280
10.1.4	Poisson-Verteilung	283
10.2	Stetige Verteilungsmodelle	287
10.2.1	Gleichverteilung	287
10.2.2	Exponentialverteilung	289
10.2.3	Normalverteilung	291
10.2.4	Prüfverteilungen	297
10.2.5	Spezielle Verteilungen in \mathbb{R}	299
11	Stichproben	305
12	Stichprobenfunktionen	311
12.1	Stichprobenfunktion \bar{X}	317
12.1.1	Erwartungswert und Varianz von \bar{X}	318
12.1.2	Normalverteilte Zufallsvariablen	319
12.1.3	Bernoulli-verteilte Zufallsvariablen	320
12.1.4	Schwaches Gesetz der Großen Zahlen	322
12.1.5	Zentraler Grenzwertsatz	323
12.2	Verteilung von Maximum und Minimum	332
12.3	Simulation	333
12.4	Simulation in \mathbb{R}	339
13	Schätzung von Parametern	343
13.1	Eigenschaften von Schätzfunktionen	345
13.1.1	Erwartungstreue	345

13.1.2	Konsistenz	349
13.1.3	Effizienz	354
13.2	Konstruktionsprinzipien	355
13.2.1	Momentenschätzer	355
13.2.2	Maximum-Likelihood-Methode	356
13.3	Intervallschätzung	365
13.3.1	Konfidenzintervalle	365
13.3.2	Prognoseintervalle	376
13.3.3	Toleranzintervalle	380
13.4	Geschichtete Stichproben	382
13.5	Schätzen in R	386
14	Grundbegriffe statistischer Tests	389
15	Einstichprobenproblem	399
15.1	Tests auf einen Lageparameter	399
15.1.1	t -Test	401
15.1.2	Vorzeichentest	405
15.1.3	Wilcoxon-Vorzeichen-Rangtest	408
15.1.4	Praktische Aspekte	414
15.2	Anpassungstests	415
15.3	Einstichprobenproblem in R	420
16	Zweistichprobenproblem	429
16.1	Verbundene Stichproben	431
16.1.1	t -Test	432
16.1.2	Vorzeichentest	433
16.1.3	Wilcoxon-Vorzeichen-Rangtest	434
16.1.4	Praktische Aspekte	436
16.2	Unverbundene Stichproben	436
16.2.1	t -Test	437
16.2.2	Welch-Test	439
16.2.3	F -Test	440
16.2.4	Wilcoxon-Rangsummentest	441
16.3	Zweistichprobenproblem in R	446
17	Einfaktorielle Varianzanalyse	451
17.1	Varianzanalyse bei Normalverteilung	452
17.2	Kruskal-Wallis-Test	458
17.3	Varianzanalyse in R	461
18	Unabhängigkeit und Homogenität	463
18.1	Unabhängigkeit	463
18.2	Homogenität	466
18.3	Unabhängigkeit und Homogenität in R	471

19 Lineares Modell	475
19.1 Modell	475
19.2 Methode der Kleinsten Quadrate	478
19.3 Güte der Anpassung	490
19.4 Tests und Konfidenzintervalle	494
19.5 Ausreißer und einflussreiche Beobachtungen	497
19.6 Linearisierbare Zusammenhänge	500
19.7 Regressionsanalyse in R	507
Anhang A Tabellen	511
Literatur	527
Sachverzeichnis	529