
Inhaltsverzeichnis

1	Grundbegriffe und Elementare Methoden	1
1.1	Einführung	1
1.1.1	Stichprobe und Grundgesamtheit	1
1.1.2	Beispiele	2
1.1.3	Schätzer, Test, Konfidenzintervall	4
1.1.4	Einteilen der Verfahren. Skalen	9
1.2	Ein-Stichproben Situation	11
1.2.1	Deskription: Histogramm, Geordnete Stichprobe	12
1.2.2	Deskription: Lage, Streuung, Schiefe	16
1.2.3	Inferenz für Lageparameter	19
1.2.4	Anpassung	22
1.3	Zwei-Stichproben Situation	28
1.3.1	Tests und Konfidenzintervalle für Normalverteilungsparameter	29
1.3.2	Mann-Whitney U-Test	30
1.4	Bivariate Stichprobe	33
1.4.1	Scattergramm, Korrelationskoeffizient	34
1.4.2	Einfache lineare Regression	37
1.4.3	Partielle Korrelation	42
1.4.4	Rangkorrelation	44
1.4.5	Kontingenztafel	45
1.5	Weiterführende Verfahren	49
1.5.1	Simultane Verfahren	49
1.5.2	Asymptotische Verfahren	50
1.5.3	Bootstrap Verfahren	51
1.6	Bestimmungsschlüssel	54
1.6.1	Gleiche Funktion, gleiche Skalen	54
1.6.2	Kriteriums- und Kovariable	55
1.6.3	Multivariate Verfahren	56

2	Varianzanalyse	57
2.1	Einfache Klassifikation	58
2.1.1	Lineares Modell und Parameterschätzung	58
2.1.2	Testen der globalen Nullhypothese	60
2.1.3	Simultane Tests und Konfidenzintervalle	63
2.1.4	Varianzhomogenität	68
2.2	Zweifache Varianzanalyse, Kreuzklassifikation $A \times B$	72
2.2.1	Lineares Modell und Parameterschätzung	72
2.2.2	Testen von Hypothesen	74
2.2.3	Wechselwirkung	75
2.2.4	Simultane Konfidenzintervalle, Paarvergleiche	77
2.3	Varianzanalyse mit 3 Faktoren	81
2.3.1	Kreuzklassifikation $A \times B \times C$	81
2.3.2	Lateinisches Quadrat	83
2.4	Ein Faktor mit korrelierten Beobachtungen	85
2.4.1	Lineares Modell	86
2.4.2	Schätz- und Testgrößen	87
2.4.3	Simultane Konfidenzintervalle	89
2.5	Rang-Varianzanalysen	94
2.5.1	k unabhängige Stichproben	94
2.5.2	k verbundene Stichproben	98
3	Lineare Regressionsanalyse	105
3.1	Multiple lineare Regression	106
3.1.1	Lineares Modell und Parameterschätzung	107
3.1.2	Testen von Hypothesen	108
3.2	Standardfehler, Konfidenzintervalle	114
3.2.1	Konfidenzintervalle	115
3.2.2	Prognoseintervalle	116
3.2.3	Spezialfall der einfachen Regression	118
3.3	Variablenselektion	119
3.3.1	Schrittweise Regression	119
3.3.2	Best-subset Selektion	121
3.4	Prüfen der Voraussetzungen	124
3.4.1	Residuenanalyse	125
3.4.2	Fishers Linearitätstest	129
3.5	Korrelationsanalyse	131
3.5.1	Multipler Korrelationskoeffizient	132
3.5.2	Partieller Korrelationskoeffizient	134
3.6	Kovarianzanalyse	136
3.6.1	Lineares Modell und Schätzer der einfachen Kovarianzanalyse	137
3.6.2	F-Tests der einfachen Kovarianzanalyse	138
3.6.3	Lineares Modell und Schätzer der zweifachen Kovarianzanalyse	140

3.6.4	F-Tests der zweifachen Kovarianzanalyse	145
4	Kategoriale Datenanalyse	149
4.1	Binäre logistische Regression	149
4.1.1	Modell und Parameterschätzung	151
4.1.2	Residuen, Goodness-of-fit	153
4.1.3	Asymptotische χ^2 Test-Statistiken	154
4.2	Multikategoriale logistische Regression	160
4.2.1	Multikategoriales Modell	160
4.2.2	Inferenz im multikategorialen Modell	162
4.2.3	Kumulatives Modell	163
4.2.4	Kumulatives Modell: Parameterschätzung	164
4.2.5	Kumulatives Modell: Diagnose und Inferenz	165
4.3	Zweidimensionale Tafel: Unabhängigkeitsproblem	170
4.3.1	Unabhängigkeitshypothese	170
4.3.2	Cross-product ratios	171
4.3.3	Strukturelle Nullen	174
4.4	Zweidimensionale Tafel: Homogenitätsproblem	177
4.4.1	Homogenitätshypothese	177
4.4.2	Simultane Konfidenzintervalle und Tests	179
4.5	Mehrdimensionale Kontingenztafeln	181
4.5.1	Dreidimensionale Kontingenztafel	181
4.5.2	Saturiertes Modell und hierarchische Unter-Modelle	183
4.5.3	Schätzen und Testen	184
4.5.4	Übersicht: Modelle, Hypothesen, Schätzer	187
4.5.5	Schätzen und Testen der λ -Terme	190
4.5.6	Vierdimensionale Kontingenztafel	192
4.6	Logit-Modelle	194
4.6.1	Logit-Modell mit 2 Regressoren A und B	194
4.6.2	Spezielle Logit-Modelle mit 2 Regressoren	195
4.6.3	Logit-Modell mit 3 Regressoren A, B und C	198
5	Nichtlineare, nichtparametrische Regression	199
5.1	Nichtlineare Regression	200
5.1.1	Modell und MQ-Methode	202
5.1.2	Konfidenzintervalle und Tests	204
5.1.3	Beispiele	205
5.2	Nichtparametrische Regression: Kernschätzer	208
5.2.1	Nichtparametrisches Regressionsmodell	209
5.2.2	Kerne	209
5.2.3	Kernschätzer	210
5.2.4	Asymptotische Eigenschaften	211
5.3	Nichtparametrische Regression: Splineschätzer	215
5.3.1	Natürliche Splinefunktionen	216
5.3.2	Penalisiertes MQ-Kriterium, Splineschätzer	217

5.3.3	Hat-Matrix, Matrizenrechnung	219
5.3.4	Rechengang zur Bestimmung des Splineschätzers	221
5.4	Additive Modelle	222
5.4.1	Smoothing Operator, smoothing Matrix	223
5.4.2	Backfitting Algorithmus	225
5.4.3	Semiparametrisches lineares Modell	228
6	MANOVA und Diskriminanzanalyse	231
6.1	Einfache MANOVA	232
6.1.1	Lineares Modell und Parameterschätzung	232
6.1.2	Produktsummen-Matrizen, Testkriterien	234
6.1.3	Simultane Tests und Konfidenzintervalle	239
6.2	Zweifache MANOVA	243
6.2.1	Lineares Modell und Parameterschätzung	244
6.2.2	Testen von Hypothesen	245
6.3	Diskriminanzanalyse	249
6.3.1	Geometrische Beschreibung	250
6.3.2	Spezialfall zweier Gruppen	251
6.3.3	Diskriminanzfunktionen	253
6.3.4	Fishers Klassifikationsfunktionen	259
6.3.5	Schrittweise Diskriminanzanalyse	263
7	Hauptkomponenten- und Faktoranalyse	267
7.1	Hauptkomponentenanalyse	268
7.1.1	Hauptkomponenten aus der Kovarianzmatrix	269
7.1.2	Hauptkomponenten aus der Korrelationsmatrix	271
7.1.3	Tests und Konfidenzintervalle	276
7.2	Faktoranalyse	277
7.2.1	Darstellung des Beobachtungsvektors	278
7.2.2	Zerlegung der Korrelationsmatrix	279
7.2.3	Schritte der Faktoranalyse	280
7.2.4	Kommunalitäten, Extraktion der Faktoren	281
7.2.5	Rotation	285
7.2.6	Faktorwerte	287
8	Clusteranalyse	289
8.1	Probleme, Begriffe, Methodik	291
8.1.1	Partitionen und Enumeration	291
8.1.2	Distanzmaße	291
8.1.3	Gütemaße	293
8.1.4	Clusterbewertungen	294
8.1.5	Einteilung der Clusterverfahren	296
8.2	Hierarchische Verfahren	297
8.2.1	Agglomerative Verfahren	297
8.2.2	Die agglomerativen Verfahren im Überblick	303

8.2.3	Divisive Verfahren	305
8.3	Nicht-hierarchische Verfahren	305
8.3.1	Totale Enumeration	305
8.3.2	Hill-climbing Verfahren	306
8.3.3	k-means Verfahren	307
8.4	Clustern bei kategorialen Daten	308
8.4.1	Transinformation als Heterogenitätsmaß	309
8.4.2	Transinformation einer Partition	310
8.4.3	Agglomeratives hierarchisches Verfahren	312
8.4.4	Clusteranalyse in einer Übergangsmatrix	314
9	Zeitreihenanalyse	317
9.1	Einführung	317
9.1.1	Aufgaben der Zeitreihenanalyse	317
9.1.2	Bestimmung eines Trends	320
9.1.3	Saisonkomponente	323
9.2	Kenngößen stationärer Prozesse	326
9.2.1	Stationarität, Kovarianzfunktion	326
9.2.2	Spektraldichte	328
9.3	Schätzen und Testen der Kenngrößen	329
9.3.1	Empirische Autokorrelation	329
9.3.2	Asymptotische Eigenschaften des Korrelogramms	330
9.3.3	Empirische partielle Autokorrelation	332
9.3.4	Periodogramm einer Zeitreihe	333
9.3.5	Periodogramm-Analyse	338
9.3.6	Spektraldichteschätzer	342
9.3.7	Asymptotisches Verhalten des Spektraldichteschätzers	344
9.4	Zeitreihenmodelle	346
9.4.1	Moving average Prozesse	347
9.4.2	Autoregressive Prozesse	348
9.4.3	ARMA und ARIMA Prozesse	351
9.4.4	Schätzen von ARMA-Parametern (aus den Residuen)	354
9.4.5	Schätzen von AR-Parametern (aus dem Korrelogramm)	355
9.5	Modelldiagnostik und Prognose	360
9.5.1	Identifikation, Residuenanalyse	360
9.5.2	Prognoseverfahren	361
9.5.3	Box-Jenkins Forecast-Formel	362
9.5.4	Prognoseintervalle	363
9.6	Bivariate Zeitreihen	364
9.6.1	Kenngößen einer bivariaten Zeitreihe	365
9.6.2	Schätzen der Kenngrößen	367

A	Fallstudien zur Statistik	371
A.1	Waldzustand Spessart [Spessart]	371
A.2	Baumwollsamens-Ertrag [Cotton]	375
A.3	Porphyrgestein [Porphy]	376
A.4	Insektenfallen [Insekten]	377
A.5	Stylometrie in Texten [Texte]	380
A.6	Gesteinsproben Toskana [Toskana]	382
A.7	Bodenproben Höglwald [Höglwald]	383
A.8	Pädiatrischer Längsschnitt [Laengs]	385
A.9	Primaten-Taxonomie [Primaten]	387
A.10	Klima Hohenpeißenberg [Hohenpeißenberg]	390
A.11	Sonnenfleckenzahlen [Sunspot]	392
A.12	Qualität pflanzlicher Nahrungsmittel [VDLUFA]	393
A.13	Verhalten von Primaten [Verhalten]	394
B	Quantil-Tabellen	395
	Literaturverzeichnis	403
	Index	407

Statistisches Methodenbuch
Verfahren, Fallstudien, Programmcodes
Pruscha, H.
2006, XVI, 412 S. 73 Abb., Softcover
ISBN: 978-3-540-26006-6