

Springer-Lehrbuch

Helge Toutenburg
Michael Schomaker
Malte Wißmann

Arbeitsbuch zur deskriptiven und induktiven Statistik

Mit 58 Abbildungen

 Springer

Professor Dr. Dr. Helge Toutenburg
Michael Schomaker
Institut für Statistik der
Universität München
Akademiestraße 1
80799 München
toutenburg@stat.uni-muenchen.de
michaelschomaker@gmx.de

Dipl.-Volkswirt Malte Wißmann
Universität Basel
WWZ
Petersgraben 51
CH-4003 Basel
malte.wissmann@unibas.ch

ISBN-10 3-540-32141-1 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-32141-5 Springer Berlin Heidelberg New York

Bibliografische Information Der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <<http://dnb.ddb.de>> abrufbar.

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland vom 9. September 1965 in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechtsgesetzes.

Springer ist ein Unternehmen von Springer Science+Business Media
springer.de

© Springer-Verlag Berlin Heidelberg 2006
Printed in Germany

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Umschlaggestaltung: Design & Production, Heidelberg

SPIN 11669418 154/3100YL 5 4 3 2 1 0 – Gedruckt auf säurefreiem Papier

Vorwort

Statistik ist die wichtigste Methode zur Datenanalyse - kombiniert mit statistischer Software.

Das Fach Statistik gehört zum Grundstudium in vielen Fachrichtungen. Wegen des zum Teil abstrakten und mathematisch begründeten Vorgehens haben Studenten häufig Probleme im Verständnis der statistischen Methoden. Die Autoren - ein Professor, ein Student der Statistik kurz vor dem Diplom und ein Assistent für Statistik - bieten mit diesem Arbeitsbuch **eine Ergänzung - keinen Ersatz!** zu den beiden Lehrbüchern

H. Toutenburg : "Deskriptive Statistik", Springer Verlag 2004

H. Toutenburg : "Induktive Statisti", Springer Verlag 2005,

deren voller Stoffumfang klausurrelevant für Haupt- und Nebenfachstudenten an deutschsprachigen Universitäten ist.

Dieses Arbeitsbuch soll eine effektive Lernhilfe für die Statistik I und II Vorlesungen sein.

Das didaktische Anliegen des Buches wird durch eine Vielzahl neuer und - wie wir hoffen - origineller Beispiele unterstützt, die durch Fortsetzung den Stoff mehrerer Kapitel umfassen können. Dazu kommen Datensätze auf der Homepage, die zur Übung allgemein und zu speziellen Aufgaben mit SPSS genutzt werden können. Sie finden Sie unter

<http://www.stat.uni-muenchen.de> (Index -> AG Toutenburg)

Wir hoffen, dass dieses Buch Anklang bei den Studenten findet. Für Verbesserungsvorschläge und Fehlermeldungen sind wir dankbar (E-mail: toutenburg@stat.uni-muenchen.de).

Wir danken den Studenten, die das Manuskript gegengelesen haben.

Die Autoren

München und Basel im Januar 2006

Inhaltsverzeichnis

Vorwort	V
1. Grundlagen	1
1.1 Merkmal oder statistische Variable	1
1.1.1 Qualitative und Quantitative Merkmale	2
1.1.2 Diskrete und Stetige Merkmale	2
1.1.3 Skalierung von Merkmalen	3
1.2 Aufgaben	3
2. Häufigkeitsverteilungen	7
2.1 Absolute und relative Häufigkeiten	7
2.2 Empirische Verteilungsfunktion	8
2.3 Grafische Darstellungen	8
2.3.1 Stab- oder Balkendiagramme	9
2.3.2 Kreisdiagramme	9
2.3.3 Stamm-und-Blatt-Diagramme	10
2.3.4 Histogramme	11
2.4 Aufgaben	12
3. Maßzahlen für eindimensionale Merkmale	17
3.1 Lagemaße	17
3.1.1 Modus oder Modalwert	17
3.1.2 Median und Quantile	17
3.1.3 Arithmetisches Mittel	20
3.1.4 Geometrisches Mittel	21
3.2 Streuungsmaße	22
3.2.1 Spannweite und Quartilsabstand	22
3.2.2 Varianz und Standardabweichung	23
3.2.3 Variationskoeffizient	24
3.3 Box-Plots	25
3.4 Konzentrationsmaße	26
3.4.1 Lorenzkurven	27
3.4.2 Gini-Koeffizient	27
3.5 Aufgaben	29

4. Analyse des Zusammenhangs zweier Merkmale	45
4.1 Darstellung der Verteilung zweidimensionaler Merkmale	45
4.1.1 Kontingenztafeln bei diskreten Merkmalen	45
4.1.2 Grafische Darstellung bei diskreten Merkmalen	46
4.1.3 Grafische Darstellung der Verteilung stetiger bzw. gemischt stetig-diskreter Merkmale	47
4.2 Maßzahlen für den Zusammenhang zweier kategorialer Merkmale	49
4.2.1 Pearsons Chi-Quadrat-Statistik	50
4.2.2 Phi-Koeffizient	50
4.2.3 Kontingenzmaß von Cramer	51
4.2.4 Kontingenzkoeffizient C	51
4.2.5 Der Odds-Ratio	52
4.2.6 Rangkorrelationskoeffizient von Spearman	53
4.3 Zusammenhang zwischen zwei stetigen Merkmalen	54
4.4 Aufgaben	56
5. Lineare Regression	71
5.1 Einleitung	71
5.2 Plots	72
5.3 Prinzip der kleinsten Quadrate	73
5.3.1 Eigenschaften der Regressionsgeraden	76
5.4 Güte der Anpassung	77
5.4.1 Varianzanalyse	77
5.4.2 Korrelation	79
5.5 Lineare Regression mit kategorialen Regressoren	81
5.6 Aufgaben	84
6. Kombinatorik	91
6.1 Einleitung	91
6.2 Permutationen	92
6.2.1 Permutationen ohne Wiederholung	92
6.2.2 Permutationen mit Wiederholung	93
6.3 Kombinationen	93
6.3.1 Kombinationen ohne Wiederholung und ohne Berücksichtigung der Reihenfolge	94
6.3.2 Kombinationen ohne Wiederholung, aber mit Berücksichtigung der Reihenfolge	94
6.3.3 Kombinationen mit Wiederholung, aber ohne Berücksichtigung der Reihenfolge	95
6.3.4 Kombinationen mit Wiederholung und mit Berücksichtigung der Reihenfolge	96
6.4 Zusammenfassung	96
6.5 Aufgaben	97

7.	Elemente der Wahrscheinlichkeitsrechnung	101
7.1	Einleitung	101
7.2	Zufällige Ereignisse	101
7.3	Relative Häufigkeit und Laplacesche Wahrscheinlichkeit	105
7.4	Axiome der Wahrscheinlichkeitsrechnung	106
7.4.1	Folgerungen aus den Axiomen	107
7.4.2	Rechenregeln für Wahrscheinlichkeiten	108
7.5	Bedingte Wahrscheinlichkeit	108
7.5.1	Motivation und Definition	108
7.5.2	Der Satz von Bayes	109
7.6	Unabhängigkeit	110
7.7	Aufgaben	112
8.	Zufallsvariablen	117
8.1	Einleitung	117
8.2	Verteilungsfunktion einer Zufallsvariablen	117
8.3	Diskrete Zufallsvariablen und ihre Verteilungsfunktion	119
8.4	Stetige Zufallsvariablen und ihre Verteilungsfunktion	120
8.5	Erwartungswert und Varianz einer Zufallsvariablen	123
8.5.1	Erwartungswert	123
8.5.2	Rechenregeln für den Erwartungswert	124
8.5.3	Varianz	124
8.5.4	Rechenregeln für die Varianz	125
8.5.5	Standardisierte Zufallsvariablen	126
8.5.6	Erwartungswert und Varianz des arithmetischen Mittels	126
8.5.7	Ungleichung von Tschebyschev	127
8.6	Zweidimensionale Zufallsvariablen	127
8.6.1	Zweidimensionale diskrete Zufallsvariablen	128
8.6.2	Zweidimensionale stetige Zufallsvariablen	129
8.6.3	Momente von zweidimensionalen Zufallsvariablen	131
8.6.4	Korrelationskoeffizient	133
8.7	Aufgaben	133
9.	Diskrete und stetige Standardverteilungen	143
9.1	Spezielle diskrete Verteilungen	143
9.1.1	Die diskrete Gleichverteilung	143
9.1.2	Die Einpunktverteilung	144
9.1.3	Die Null-Eins-Verteilung	145
9.1.4	Die Binomialverteilung	146
9.1.5	Die hypergeometrische Verteilung	147
9.2	Spezielle stetige Verteilungen	148
9.2.1	Die stetige Gleichverteilung	148
9.2.2	Die Normalverteilung	149
9.3	Weitere Verteilungen	153
9.3.1	Die Poissonverteilung	153

9.3.2	Die Multinomialverteilung	154
9.3.3	Die Exponentialverteilung	155
9.4	Prüfverteilungen	156
9.4.1	Die χ^2 -Verteilung	156
9.4.2	Die t -Verteilung	157
9.4.3	Die F -Verteilung	157
9.5	Aufgaben	158
10.	Schätzung von Parametern	167
10.1	Einleitung	167
10.2	Punktschätzung von Parametern	168
10.2.1	Punktschätzung für μ bei einer normalverteilten Zufallsvariable	168
10.2.2	Punktschätzung für σ^2 bei einer normalverteilten Zufallsvariable	169
10.2.3	Punktschätzung von p bei einer binomialverteilten Zufallsvariable	169
10.3	Konfidenzschätzungen von Parametern	170
10.3.1	Grundlagen	170
10.3.2	Konfidenzschätzung des Erwartungswerts einer Normalverteilung	170
10.3.3	Konfidenzschätzung einer Binomialwahrscheinlichkeit ..	172
10.4	Aufgaben	173
11.	Prüfen statistischer Hypothesen	177
11.1	Einleitung	177
11.2	Grundlegende Begriffe	177
11.2.1	Ein- und Zweistichprobenprobleme	177
11.2.2	Ein- und Zweiseitige Tests	178
11.2.3	Allgemeines Vorgehen	179
11.2.4	Fehler 1. und 2. Art	179
11.3	Einstichprobenprobleme	180
11.3.1	Prüfen des Mittelwerts bei bekannter Varianz (einfacher Gauss-Test)	180
11.3.2	Prüfung des Mittelwertes bei unbekannter Varianz (einfacher t -Test)	183
11.4	Zweistichprobenprobleme	184
11.4.1	Prüfen der Gleichheit der Varianzen (F -Test)	184
11.4.2	Prüfen der Gleichheit der Mittelwerte zweier unabhängiger normalverteilter Zufallsvariablen	187
11.4.3	Prüfen der Gleichheit der Mittelwerte aus einer verbundenen Stichprobe (paired t -Test)	189
11.5	Prüfen von Hypothesen über Binomialverteilungen	190
11.5.1	Prüfen der Wahrscheinlichkeit für das Auftreten eines Ereignisses (Binomialtest für p)	190

11.5.2 Prüfen der Gleichheit zweier Binomialwahrscheinlichkeiten	191
11.6 Testentscheidung mit p -values	193
11.7 Aufgaben	193
12. Nichtparametrische Tests	209
12.1 Einleitung	209
12.2 Anpassungstests	209
12.2.1 Chi-Quadrat-Anpassungstest	210
12.2.2 Kolmogorov-Smirnov-Anpassungstest	212
12.3 Homogenitätstests für zwei unabhängige Stichproben	213
12.3.1 Kolmogorov-Smirnov-Test im Zweistichprobenproblem	213
12.3.2 Mann-Whitney- U -Test	214
12.4 Aufgaben	216
13. Multiple lineare Regression	227
13.1 Einleitung	227
13.2 Modellannahmen der multiplen Regression	227
13.3 Schätzung der Parameter	228
13.4 Prüfen von linearen Hypothesen	228
13.5 Aufgaben	234
14. Analyse von Kontingenztafeln	241
14.1 Einleitung	241
14.2 Zweidimensionale kategoriale Zufallsvariablen	241
14.3 Unabhängigkeit	243
14.4 χ^2 -Unabhängigkeitstest	244
14.5 Die Vierfeldertafel	245
14.6 Aufgaben	249
A. Tabellenanhang	253
Literatur	265
Sachverzeichnis	267