

Springer-Lehrbuch

Judith Eckle-Kohler • Michael Kohler

# Eine Einführung in die Statistik und ihre Anwendungen

Zweite Auflage

 Springer

Dr. Judith Eckle-Kohler  
TU Darmstadt  
FB Mathematik  
Schloßgartenstr. 7  
64289 Darmstadt  
Deutschland  
eckle-kohler@t-online.de

Prof. Dr. Michael Kohler  
TU Darmstadt  
FB Mathematik  
Schloßgartenstr. 7  
64289 Darmstadt  
Deutschland  
kohler@mathematik.tu-darmstadt.de

ISSN 0937-7433

ISBN 978-3-642-15225-2

e-ISBN 978-3-642-15226-9

DOI 10.1007/978-3-642-15226-9

Springer Heidelberg Dordrecht London New York

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

© Springer-Verlag Berlin Heidelberg 2009, 2011

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.

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.

*Einbandentwurf:* WMXDesign GmbH, Heidelberg

Gedruckt auf säurefreiem Papier

Springer ist Teil der Fachverlagsgruppe Springer Science+Business Media ([www.springer.com](http://www.springer.com))

*Für Iris und Julius*

# Vorwort

Die Statistik beschäftigt sich mit der Analyse von Phänomenen, die im mathematischen Sinne als zufällig aufgefasst werden können. Dabei kann die Einführung des Zufalls in verschiedener Hinsicht nützlich sein: Der Zufall kann einerseits zur Vereinfachung der Datenerhebung eingeführt werden: indem zufällig Daten herausgegriffen werden, muss nicht die Gesamtzahl der Daten untersucht werden (was z.B. bei einer Wahlumfrage ausgenützt wird). Andererseits kann der Zufall künstlich eingeführt werden zur Vereinfachung der Modellierung deterministischer Vorgänge: Dabei werden sehr komplexe Teile als unbestimmt angesehen und durch einen einfachen zufälligen Prozess modelliert.

Das vorliegende Buch gibt eine umfassende Einführung in die Grundprinzipien der Statistik und die zugrundeliegende mathematische Theorie des Zufalls. Dabei wird bewusst auf allzu viele Details verzichtet. Vielmehr sollen Leser ohne Vorkenntnisse in diesem Bereich die grundlegenden Ideen und den Nutzen dieser Theorie kennenlernen. Diese kann dann später bei Bedarf durch weiterführende Literatur wie z.B. Bauer (1992) oder Witting (1985) vertieft werden.

Das Buch ist in 6 Kapitel unterteilt. Kapitel 1 macht deutlich, dass die Statistik ein Gebiet mit vielfältigen Anwendungsmöglichkeiten ist, und dass Statistikwissen auch im alltäglichen Leben immer wieder benötigt wird. Die dafür notwendige Theorie wird in den folgenden Kapiteln beschrieben. Kapitel 2 stellt zunächst die Erhebung von Daten im Rahmen von Studien und Umfragen vor. Kenntnisse darüber sind deshalb wichtig, weil sich oft beobachten lässt, dass die Erhebung von Daten zu Qualitätseinbußen bei den Daten und damit zu starken Einschränkungen in Bezug auf die Analyse der Daten führt. In Kapitel 3 werden Verfahren der beschreibenden Statistik beschrieben. Diese legen kein mathematisches Modell der Entstehung der Daten zugrunde, lassen aber andererseits auch keine Rückschlüsse zu, die über den beobachteten Datensatz hinaus gültig sind. Um dies zu erreichen, muss man Modellannahmen an die Entstehung der Daten machen. Dazu wird in den Kapiteln 4 und 5 das mathematische Modell des Zufalls eingeführt. Während sich Kapitel 4 mit dem mathematischen Begriff der Wahrscheinlichkeit beschäftigt und einfache Schlussfolgerungen daraus vorstellt, werden in Kapitel 5 Zufallsvariablen eingeführt, die eine besonders elegante Beschreibung zufälliger Phänomene ermöglichen. Neben

Kennzahlen dieser Zufallsvariablen wie Erwartungswert und Varianz werden dort auch die Gesetze der großen Zahlen sowie der zentrale Grenzwertsatz vorgestellt, letzterer aber ohne Beweis. Die darauf aufbauenden Verfahren der so genannten schließenden Statistik sind dann Inhalt von Kapitel 6. Mit diesen Verfahren und mit Hilfe von Annahmen an die Entstehung der Daten lassen sich Schlussfolgerungen ziehen, die über den vorliegenden Datensatz hinaus gültig sind. Im Anhang sind die wichtigsten zum Verständnis des Buches benötigten Grundlagen aus der Mathematik kurz dargestellt.

Das Buch ist gedacht für Studenten, die ohne Vorwissen aus der Wahrscheinlichkeitstheorie und der Statistik einen Überblick über dieses doch sehr umfangreiche Gebiet bekommen wollen. Es entstand aus einer Reihe von Vorlesungen, die der zweite Autor innerhalb der letzten 10 Jahre an den Universitäten Stuttgart, Jena, Saarbrücken und Darmstadt abgehalten hat. Diese Veranstaltungen richteten sich zum einen an Studierende des Faches Mathematik und fanden in Diplom- bzw. Bachelor- bzw. Lehramtsstudiengängen innerhalb des Grundstudiums statt. Zum anderen wurde Material dieses Buches in Vorlesungen des ersten Semesters für Studierende der Fächer Biologie, Pädagogik, Psychologie, Soziologie, Volkswirtschaftslehre und Wirtschaftswissenschaften eingesetzt.

Den Studierenden in diesen Vorlesungen gebührt unser Dank für Kommentare, die immer wieder zur Verbesserung dieses Buches beigetragen haben.

Darmstadt,  
Dezember 2008

*Judith Eckle-Kohler*  
*Michael Kohler*

Bei der vorliegenden Auflage handelt es sich um eine durchgesehene und korrigierte Version der Erstauflage des Buches. Dabei wurden insbesondere die Kommentare und Verbesserungsvorschläge der Studentinnen und Studenten der Vorlesung *Einführung in die Stochastik* im Sommersemester 2009 berücksichtigt, für die wir uns recht herzlich bedanken.

Ergänzend zu diesem Buch gibt es mittlerweile im Internet eine frei zugängliche Vorlesungsaufzeichnung. Dabei handelt es sich um die Aufzeichnung einer Vorlesung an der Technischen Universität Darmstadt, die primär für Studierende des Faches Mathematik im Bachelor- oder Lehramtsstudium im 4. Semester angeboten worden ist. Weitere Informationen zu dieser Vorlesungsaufzeichnung gibt es auf der Homepage dieses Buches unter

<http://www.mathematik.tu-darmstadt.de/statistik-und-anwendungen>

Darmstadt,  
August 2010

*Judith Eckle-Kohler*  
*Michael Kohler*

# Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b> .....	1
1.1	Übungsteilnahme und Statistik-Note .....	1
1.2	Sex und Herzinfarkt .....	2
1.3	Die Challenger-Katastrophe .....	3
1.4	Präsidentchaftswahl in den USA, Herbst 2000 .....	5
1.5	Positionsbestimmung mittels GPS .....	6
1.6	Analyse von DNA-Microarray-Daten .....	7
1.7	Berechnung von Prämien in der Schadenversicherung .....	8
1.8	Bewertung des Risikos von Kapitalanlagen bei Banken .....	8
1.9	Vorhersage des Verschleißes von Kfz-Bauteilen .....	9
1.10	Nutzen der Statistik in verschiedenen Studiengängen .....	9
1.11	Weiterer Aufbau dieses Buches .....	10
<b>2</b>	<b>Erhebung von Daten</b> .....	13
2.1	Kontrollierte Studien .....	13
2.2	Beobachtungsstudien .....	17
2.3	Probleme bei der Durchführung von Studien .....	21
2.4	Umfragen .....	24
	Aufgaben .....	27
<b>3</b>	<b>Deskriptive und explorative Statistik</b> .....	31
3.1	Typen von Messgrößen .....	31
3.2	Histogramme .....	32
3.3	Dichteschätzung .....	35
3.4	Statistische Maßzahlen .....	41
3.5	Regressionsrechnung .....	48
3.6	Nichtparametrische Regressionsschätzung .....	56
3.7	Probleme bei der Interpretation der bisher eingeführten Verfahren ..	57
	Aufgaben .....	59

<b>4</b>	<b>Das mathematische Modell des Zufalls</b> .....	63
4.1	Der Begriff der Wahrscheinlichkeit .....	63
4.2	Grundaufgaben der Kombinatorik .....	70
4.3	Der Begriff des Wahrscheinlichkeitsraumes .....	78
4.4	Der Begriff der $\sigma$ -Algebra .....	86
4.5	Der Laplacesche Wahrscheinlichkeitsraum .....	90
4.6	Wahrscheinlichkeitsräume mit Zähldichten .....	93
4.7	Wahrscheinlichkeitsräume mit Dichten .....	98
4.8	Bedingte Wahrscheinlichkeit .....	104
	Aufgaben .....	108
<b>5</b>	<b>Zufallsvariablen und ihre Eigenschaften</b> .....	113
5.1	Der Begriff der Zufallsvariablen .....	113
5.2	Der Begriff der Verteilungsfunktion .....	121
5.3	Der Begriff der Unabhängigkeit .....	125
5.4	Der Erwartungswert einer Zufallsvariable .....	131
5.5	Die Varianz einer Zufallsvariable .....	151
5.6	Gesetze der großen Zahlen .....	157
5.7	Der Beweis des starken Gesetzes der großen Zahlen .....	163
5.8	Der zentrale Grenzwertsatz .....	167
	Aufgaben .....	174
<b>6</b>	<b>Induktive Statistik</b> .....	179
6.1	Fragestellungen .....	179
6.2	Punktschätzverfahren .....	183
6.3	Bereichsschätzungen .....	194
6.4	Statistische Testverfahren .....	203
6.5	Tests zur Überprüfung von Verteilungsmodellen .....	217
6.6	Die einfaktorielle Varianzanalyse .....	228
	Aufgaben .....	232
<b>A</b>	<b>Mathematische Grundlagen</b> .....	235
A.1	Mengen und Mengenoperationen .....	235
A.2	Das Summenzeichen .....	238
A.3	Folgen und Reihen .....	239
A.4	Differentialrechnung .....	244
A.5	Integralrechnung .....	247
	<b>Anmerkungen</b> .....	251
	<b>Literaturverzeichnis</b> .....	261
	<b>Sachverzeichnis</b> .....	263