

Book Reviews

Editor: P. Hackl

Weak Convergence and Empirical Processes (A.W. van der Vaart, J.A. Wellner)

Klaus Pötzelberger

Quasi-Likelihood and its Application. A General Approach to Parameter Estimation
(C.C. Heyde)

Werner G. Müller

Wahrscheinlichkeitsrechnung und Statistik in Beispielen und Aufgaben (V. Nollau,
L. Partzsch, R. Storm, C. Lange)

Hans Kellerer

A First Course in Multivariate Statistics (B. Flury)

Mushtaq Hussain

Empirische Forschungsmethoden (W. Stier)

Michael G. Schimek

Applying Generalized Linear Models (J.K. Lindsey)

Reinhold Hatzinger

Analyse von Tabellen und kategorialen Daten (H.J. Andreß, J.A. Hagenaars, S. Kühnel)

Walter Katzenbeisser

Elements of Multivariate Time Series Analysis (G.C. Reinsel)

Friedrich Leisch

Nonparametric Smoothing and Lack-of-Fit Tests (J.D. Hart)

Herwig Friedl

Modelling Extremal Events for Insurance and Finance (P. Embrechts, C. Klüppelberg,
T. Mikosch)

Franz Konecny

Statistical Analysis of Extreme Values (R.D. Reiss, M. Thomas)

Klemens Fuchs

Das Quotenverfahren (A. Quatember)

Alfred Franz

Prophetentheorie (F. Harten, A. Meyerthole, N. Schmitz)

Wolfgang Wertz

Advances in Combinational Methods and Applications to Probability and Statistics.
(N. Balakrishnan)

Jürgen Pilz

A.W. van der VAART und Jon A. WELLNER, **Weak Convergence and Empirical Processes**. Berlin: Springer-Verlag, 1996, 459 S., öS 540.20, ISBN 0-387-94640-3.

Das Buch ist eine moderne Darstellung schwacher Konvergenz und der Theorie empirischer Prozesse. Es besteht aus drei Teilen. Der erste beschäftigt sich mit der Konvergenz stochastischer Größen, insbesondere der Konvergenz in Verteilung im Rahmen von Familien beschränkter Funktionen, die mit der Supremumsnorm versehen sind. Die stochastischen Größen werden dabei nicht als Borel-meßbar vorausgesetzt. Dieser Zugang vermeidet übliche technische Schwierigkeiten. So ist z.B. der empirische Prozeß als Funktion mit Werten im Skorohodraum, der mit der Supremumsnorm versehen ist, nicht Borel-meßbar. Der zentrale zweite Teil ist empirischen Prozessen gewidmet, die mit Klassen von Abbildungen oder meßbaren Mengen indiziert sind. Gleichmäßige Glivenko-Cantelli Sätze und gleichmäßige zentrale Grenzverteilungssätze bilden hier die Höhepunkte. Im dritten Teil werden Anwendungen der in den ersten beiden Teilen vorgestellten Ergebnisse in der Statistik behandelt, wobei naturgemäß Grenzwertsätze der asymptotischen Statistik im Vordergrund stehen. Die Themen umfassen parametrische und nichtparametrische Schätzung, *bootstrap*, Kolmogorov-Smirnov Statistik im Zwei-Stichprobenproblem und den Faltungssatz. Die Kapitel sind mit Aufgaben versehen, die nicht nur die vorgestellten Resultate illustrieren, sondern teilweise Erweiterungen und technische Einzelheiten beinhalten.

Das Buch ist leicht verständlich geschrieben. Es macht eine Fülle von Resultaten, die bisher meist nur in Zeitschriften veröffentlicht wurden, einem weiten Leserkreis zugänglich. Es stellt neben Pollard (1984) und Shorack und Wellner (1986) einen Meilenstein seit Billingsley (1968) dar und wird in den nächsten Jahren ein oft gelesenes und vielzitiertes Werk sein.

*Klaus Pötzelberger
Institut für Statistik
Wirtschaftsuniversität Wien*

Christopher C. HEYDE, **Quasi-Likelihood and Its Application. A General Approach to Parameter Estimation**. Berlin: Springer-Verlag, 1997, ix+240 S., öS 624.–, ISBN 0-387-98225-6.

Das Buch gibt eine gelungene Übersicht zum Stand des Wissens auf dem Gebiet der *optimal estimating functions* (in der Folge von mir mit "optimale Schätzungsfunktionen" etwas holprig übersetzt), wiewohl ganz aktuelle Publikationen wie z.B. Chaganty (1997) keine Erwähnung finden. Der Schreibstil ist mathematisch dicht, jedoch trotzdem gut lesbar. Als Einführung in die Materie ist aber wohl eher der Überblicksartikel von Godambe und Kale (1991) geeignet.

Inhaltlich spannt sich der Bogen von einer allgemeinen Beschreibung der Theorie der Schätzungsfunktionen über die Behandlung von Konfidenzintervallen und Hypothesentests bis hin zu asymptotischen Eigenschaften. Das in dem Buch verwendete Hauptkonstruktionsprinzip von optimalen Schätzungsfunktionen besteht darin, die Martingaldifferenzenfolge aus der Scorefunktion einer beliebigen Verteilung (üblicherweise aus der

Exponentialfamilie) zu bilden, und anschließend die Gewichte a_t für die Klasse $H = \sum_{t=1}^T a_t h_t$ geeignet zu wählen. Wie aus dem Titel des Buchs ersichtlich, wird viel Platz den Beziehungen zur *quasi-likelihood* Theorie von Wedderburn gewidmet, vor allem deren Darstellung als Spezialfall der optimalen Schätzungsgleichungen und deren Limitationen, wenn die Klasse erweitert wird.

Die ebenfalls im Titel angekündigten Anwendungen beschränken sich allerdings auf solche theoretischer Natur. Echte Datensätze werden kaum und wenn dann nur knapp analysiert. Auch im Kapitel "Miscellaneous Applications" läßt sich die Bedeutung der Theorie für den Anwender bestenfalls erahnen. Das ist etwas schade, ließen sich doch sicher zu vielen behandelten Problemfeldern gute Anwendungsbeispiele finden - zum Thema Heteroskedastie ist das etwa in der Geostatistik der Fall.

Zusammenfassend soll gesagt sein, daß das besprochene Buch inhaltlich dem guten Standard der Serie sicher gerecht wird und in den Fachbibliotheken nicht fehlen darf. Für den mathematischen Statistiker, der auf den genannten Gebieten forscht, ist es bestimmt unverzichtbar, für den angewandten Statistiker, der sich konkrete Ratschläge für die Datenanalyse erwartet, hingegen vermutlich schon.

Literatur

- Chaganty, N.R. (1997). An alternative approach to the analysis of longitudinal data via generalized estimating equations. *Journal of Statistical Planning and Inference* 63, 39-54.
 Godambe, V.P. and Kale, B.K. (1991). "Estimating functions: an overview", 3-20 in Godambe, V.P. (Ed.), *Estimating Functions*. Oxford: Science Publications.

Werner G. Müller
 Institut für Statistik
 Wirtschaftsuniversität Wien

Volker NOLLAU, Lothar PARTZSCH, Regina STORM und Claus LANGE, **Wahrscheinlichkeitsrechnung und Statistik in Beispielen und Aufgaben**. Stuttgart: B.G. Teubner Verlag, 1997, 271 S., öS 254.-, ISBN 3-8154-2073-3.

Das vorliegende Buch richtet sich an Studierende, die ein Grundwissen aus Statistik erwerben wollen, und kann sowohl als straffes einführendes Lehrbuch als auch als Beispielsammlung betrachtet werden. Beiden Zielsetzungen gerecht zu werden, ist eine schwierige Aufgabe, die hier mit einiger Bravour gelöst wurde. Ohne Definitionen und Sätze zu formulieren, wird das nötige Rüstzeug zur Behandlung statistischer Probleme bereitgestellt. Jeder Begriff und jedes Verfahren wird an durchgerechneten Beispielen sehr gut verständlich und anschaulich erläutert. Nach jedem Abschnitt folgen insgesamt über 250 Aufgaben, deren Praxisbezug nur manchmal zugunsten der rechnerischen Kürze etwas vernachlässigt wurde. Alle Beispiele sind ohne PC und Statistikprogramm lösbar. Die Lösungen sind am Ende des Buches in kürzest möglicher Form angegeben. Dort findet man auch alle nötigen Tabellen.

Die Voraussetzungen an den Leser beschränken sich in erster Linie auf die prinzipielle Fähigkeit, Sachverhalte in formaler Schreibweise darzustellen. Grundkenntnisse aus der Analysis bis zum partiellen Differenzieren und bis zum zweidimensionalen uneigentlichen Integral sind vorteilhaft und ausreichend. Der Inhalt reicht von beschreibender

Statistik für ein- und zweidimensionale Daten und Zeitreihen über Wahrscheinlichkeitsrechnung, Verteilungen und Grenzwertsätze bis zu Punkt- und Intervallschätzung und Statistischen Testverfahren.

Nicht nur Studierende werden es als vorteilhaft empfinden, daß auf die Bezeichnung “Zufallsvariable” zugunsten des leichter verständlichen “Zufallsgröße” verzichtet wurde. Man könnte auch noch “Stichprobenraum” durch “Ergebnismenge” ersetzen. Sehr schön ist Beispiel 3.1.2, eine teils stetige, teils diskrete Zufallsgröße und deren Verteilungsfunktion. Das Kapitel über Verteilungen enthält erfreulicherweise auch die logarithmische Normalverteilung sowie Weibull-, Beta- und Gammaverteilung. Das Kapitel über Testverfahren beginnt mit Parametertests für Ein- und Zweistichprobenprobleme. Hier wäre etwas mehr Ausführlichkeit im Text und eine Trennung hilfreich. Es folgen Anpassungstests (Chiquadrat und Kolmogoroff) und nichtparametrische Tests für das Zweistichprobenproblem. Letzere sind - obwohl klar erklärt - eher als Alibikapitel zu betrachten. In einer neuen Auflage sollte dieser Abschnitt in Hinblick auf die Praxis erweitert werden. Wie in allen einführenden Büchern entsteht ganz von selbst ein Übergewicht der wahrscheinlichkeitstheoretischen Grundlagen über die - in der Anwendung eigentlich gefragte - beurteilende Statistik. Dieser Mangel wird hier allerdings durch die hohe Anzahl an Aufgaben gut wettgemacht.

Nicht nur der Inhalt, sondern auch die Druckqualität machen das Buch gut lesbar und man kann es jedem Studienanfänger und jedem an Statistik am Rande interessierten Praktiker empfehlen. Für die Erstgenannten ist es eine ausgezeichnete Einführung, für die Zweitgenannten eine Motivation, zu weiterführender Literatur zu greifen.

*Hans Kellerer
Institut für Statistik
Universität Graz*

Bernard FLURY, **A First Course in Multivariate Statistics**. Berlin: Springer-Verlag, 1997, xiii+735 S., öS 1080.40, ISBN 0-387-98206-X.

Das vorliegende Buch beinhaltet eine sehr gute Beschreibung sowohl der theoretischen als auch der praktischen Aspekte der multivariaten Statistik. Obwohl das Buch mehr als 700 Seiten umfaßt, werden relativ wenige Methoden beschrieben, was zu einer sehr detaillierten Behandlung der Themen führt. Nach einem einführenden Kapitel, welches den Leser für die multivariate Statistik motivieren soll, werden im Kapitel 2 bi- und multivariate Verteilungen und im Kapitel 3 die multivariate Normalverteilung besprochen. Kapitel 4 beinhaltet Schätzverfahren wie *plug-in*, *maximum likelihood* sowie *quasi-likelihood* Schätzer. Der Thematik Diskriminanz- und Clusteranalyse wurden zwei Kapitel gewidmet – einführende Themen in Kapitel 5 sowie fortgeschrittene Themen in Kapitel 7. Im Kapitel 6 wurden multivariate statistische Tests für den Mittelwert behandelt. Die Hauptkomponentenanalyse wurde in Kapitel 8 unorthodox, jedoch auf wirksame Weise präsentiert. Das letzte Kapitel behandelt das Thema “Mischverteilungen”. Das Buch hat also hauptsächlich parametrische Modelle zum Inhalt, beschreibt aber auch computerintensive Verfahren wie *bootstrap* oder randomisierte Tests.

Eine Reihe multivariater Verfahren wie z.B. Faktorenanalyse, MANOVA, multivariate Regressionsanalyse, werden nicht behandelt. Der Autor bemerkt dazu, daß es sein Ziel

war, relativ wenige Themen, aber diese dafür sehr detailliert zu behandeln, was ihm auch sehr gut gelungen ist. Jedes Kapitel beinhaltet zahlreiche Beispiele, graphische Darstellungen sowie fundierte theoretische Abhandlungen. Alle Begriffe sind sehr ausführlich erklärt.

Das Buch fängt auf einem einfachen Niveau an, erreicht in Kapitel 3-4 mittleres und in Kapitel 7-9 eher fortgeschrittenes Niveau. Aus diesem Grund ist das Buch für *graduate* als auch für *postgraduate* Studenten geeignet. Die Druckqualität des Buches ist sehr gut. Auf Seite 244 (Definition 4.3.4 Pkt (b)) entdeckte ich einen Druckfehler, und zwar lauten die richtigen Indizes für die Informationsfunktion (j,h) und nicht wie angegeben (j,k).

Insgesamt vermittelt das Buch einen sehr guten Einblick über die behandelten Themen. Die Beispiele erleichtern das Verständnis der teilweise anspruchsvollen Themen. Das Buch kann für eine fundierte Ausbildung im Bereich der multivariaten Statistik wertvolle Dienste leisten.

Mushtaq Hussain
Institut für angewandte Statistik
Joanneum Research Graz

Winfried STIER, **Empirische Forschungsmethoden**. Berlin: Springer-Verlag, 1996, xi+409 S., öS 401.50, ISBN 3-540-61393-5.

Das vorliegende Buch wendet sich an Studierende der Wirtschafts- und Sozialwissenschaften und orientiert sich an den Anforderungen der Universität St. Gallen (Schweiz). Unter "empirischen Forschungsmethoden" versteht sein Autor die Grundlagen empirischen Arbeitens, Messens und Skalierens, Auswahlverfahren, Methoden der Datenerhebung (mit Betonung von Fragebögen), Untersuchungsdesigns und eine Einführung in wichtige multivariate Verfahren. Im Prinzip handelt es sich um ein Lehrbuch der klassischen multivariaten Datenanalyse wie beispielsweise jenes von Backhaus *et al.* (1986), das weite Verbreitung erfahren hat.

Was sind die Unterschiede zwischen diesen beiden Werken? Oberflächlich betrachtet, unterscheiden sie sich bezüglich der Auswahl der behandelten statistischen Methoden und des statistischen Programmsystems. Stier demonstriert seine Anwendungen mittels SYSTAT. Backhaus *et al.* beziehen sich auf SPSS. Regression, Varianzanalyse, Diskriminanzanalyse, Faktorenanalyse und Clusteranalyse werden in beiden Werken ausführlich behandelt. Stier geht zusätzlich detailliert auf Log-lineare und Logit-Modelle ein, berücksichtigt jedoch Kovarianzstrukturmodelle und multidimensionale Skalierungsverfahren nicht.

Setzt man sich mit Stiers Lehrbuch näher auseinander, wird man schnell an den einführungsfähigen Kapiteln Gefallen finden, beginnend mit einigen wissenschaftstheoretischen Überlegungen. Viele in anderen Büchern häufig vernachlässigte Aspekte des Messens, Skalierens und der Stichprobenwahl werden kurz, nichtsdestoweniger informativ abgehandelt. Ohne Vorkenntnisse in Fragebogenkonstruktion wird eine brauchbare Instruktion gegeben, die die schlimmsten Fehler vermeiden läßt. Bei den speziellen Untersuchungsdesigns wird auch auf die in der Praxis wichtigen quasi-experimentellen Untersuchungsanordnungen eingegangen.

Wie sieht es nun mit der Behandlung der primär statistischen Themen aus? Sowohl Stier als auch Backhaus *et al.* verzichten auf mathematisches Detail zugunsten leichter Verständlichkeit des formal durchaus anspruchsvollen Inhalts. Stier gelingt es meines Erachtens jedoch besser, relevante statistische Aspekte herauszuarbeiten, beziehungsweise kritisch zu durchleuchten (z.B. die stufenweise Variablenselektion in der multivariaten Regression). Auch ein gewisser intellektueller Einfluß von Fahrmeir und Hamerles (1984) Standardwerk ist nicht zu verleugnen. In der Tat gib es wiederholt Hinweise auf einzelne Kapitel dieses Werks, wenn es um kompliziertere Betrachtungen geht, die in dem vorliegenden Einführungstext zurecht nicht diskutiert werden.

Insgesamt ist die Darstellung der multivariaten Verfahren gut lesbar und praxisorientiert, ohne ein Zuviel an Sprache, was manche Einführungstexte für Anwender charakterisiert, dargestellt. Die Referenzliteratur ist auf dem Stand der Forschung. Der Zugang zu SYSTAT ist für die Lektüre des Lehrbuches, das sich auch zum Selbststudium eignet, entbehrlich. Ich zögere nicht, dieses Lehrbuch meinen Studierenden der Sozial- und Wirtschaftswissenschaften und der Psychologie ab dem dritten Semester als erste Einführung in die multivariaten Analyseverfahren zu empfehlen.

Literatur

Backhaus et al. (1986). *Multivariate Analysemethoden*. Berlin: Springer-Verlag.

Fahrmeir L. und Hamerles (1984). *Multivariate statistische Verfahren*. Berlin: de Gruyter.

Michael Schimek
Institut für Medizinische Informatik, Statistik und Dokumentation
Karl-Franzens-Universität Graz

James K. LINDSEY, **Applying Generalized Linear Models**. Berlin: Springer-Verlag, 1997, ix+305 S., öS 642.40, ISBN 0-387-94876-7.

Das Konzept verallgemeinerter linearer Modelle (GLM) gewann in den letzten beiden Jahrzehnten zunehmend Einfluß darauf, wie angewandte Statistik betrieben wird. Als Autor einer Reihe von Büchern mit dem Schwerpunkt statistischen Modellierens hat Lindsey sicherlich dazu beigetragen, eine Reihe von Methoden durch ihre Formulierung als GLM (und damit der Möglichkeit, Standardpakete wie z.B. S-Plus oder SAS verwenden zu können) einem weiteren Anwenderkreis zugänglich zu machen. Sein neuestes Buch versucht einen Überblick über Modelle aus verschiedensten Bereichen zu geben, wobei der Autor betont, weniger Methoden statistischer Inferenz behandeln zu wollen, als vielmehr sein Hauptaugenmerk auf den Modellierungsansatz und Gemeinsamkeiten dieser Methoden zu richten.

Im ersten Kapitel wird eine kurze Einführung in den theoretischen Hintergrund von GLM und die Idee statistischen Modellierens gegeben. Kapitel 2 und 3 behandeln log-lineare und logistische Modelle, sowie die Verwendung und den Vergleich von (hauptsächlich diskreten) Wahrscheinlichkeitsverteilungen. Beginnend mit sehr einfachen Beispielen erweitert Lindsey seine Darstellung sehr rasch zu spezialisierteren Methoden, wie Symmetrie-, *mover-stayer*- oder Rasch-Modelle, beschäftigt sich mit *random effects* und Überdispersion, aber auch mit multivariaten und Mischverteilungen. Kapitel 4 ist *growth*

curves gewidmet. Im fünften Kapitel werden (diskrete) Zeitreihen behandelt, insbesondere Poisson- und Markov-Prozesse. Die nächsten beiden Kapitel, 6 und 7, erörtern die Anwendung von GLM auf Überlebenszeit und *event-history* Daten. Eine Darstellung von Analysemöglichkeiten räumlicher Daten gibt Kapitel 8, mit Hauptaugenmerk auf *clustering* und dem Auffinden räumlicher Muster. Erst im neunten Kapitel werden kurz Normalverteilungsmodelle wie z.B. ANOVA oder nichtlineare Regression gestreift. Das zehnte und letzte Kapitel behandelt dynamische Modelle, wobei wieder der Aspekt einer einheitlichen Darstellung verschiedener Daten- und Modelltypen im Rahmen von GLM im weitesten Sinn in den Vordergrund gestellt wird. In zwei Anhängen, A und B, werden allgemeinere Prinzipien der Inferenz (insbesondere direkte *likelihood*-Inferenz) und Modelldiagnostik präsentiert.

Das Buch hinterläßt einen ambivalenten Eindruck. Es umfaßt auf insgesamt 280 Seiten etwa 150 Kapitel und Unterkapitel, deren Inhalte schon allein auf Grund des Buchumfangs nur sehr oberflächlich behandelt werden können. Trotz der Angabe von Literatur zu den einzelnen Themen (deren Studium für ein tiefergehendes Verständnis wohl unumgänglich ist) ist man geneigt zu sagen, daß weniger mehr gewesen wäre. Andererseits gibt das Buch einen Überblick über die erstaunliche Vielfalt von Methoden, die im Rahmen von GLM einheitlich darstellbar sind und dem Praktiker durchaus behilflich sein können, bei bestimmten Fragestellungen innovative Problemlösungen zu finden. Eine Vielzahl von interessanten Beispielen (im Text und in eigenen Unterkapiteln) mit der Angabe von Daten und auch konkreten Hinweisen zur Berechnung illustrieren die behandelten Methoden sehr gut. Das Buch richtet sich wohl in erster Linie an angewandte Statistiker, die sich einen Überblick über neuere Methoden und Möglichkeiten statistischen Modellierens im Rahmen von GLM verschaffen wollen, ohne sich durch allzu viele technische Details durcharbeiten zu müssen.

*Reinhold Hatzinger
Institut für Statistik
Wirtschaftsuniversität Wien*

Hans-Jürgen ANDRESZ, Jacques A. HAGENAARS und Steffen KÜHNEL, **Analyse von Tabellen und kategorialen Daten. Log-lineare Modelle, latente Klassenanalyse, logistische Regression und GSK-Ansatz.** Berlin: Springer-Verlag, 1997, xx+455 S., öS 436.60, ISBN 3-540-62515-1.

Motivation für vorliegendes Buch war die Feststellung der Autoren, daß es zwar eine Reihe englischsprachiger Bücher über die Analyse kategorialer Daten, aber keine deutschsprachige, primär anwendungsorientierte Einführung in diese Fragestellung gibt. Entsprechend dieser Motivation ist der Aufbau des Buches gestaltet. In den ersten fünf Kapitel, dem Grundlagenteil, werden die wesentlichsten Methoden zur Analyse kategorialer Daten vorgestellt, im anschließenden zweiten Teil (Kapitel 6 bis 9) werden die besprochenen Verfahren an Hand von interessanten realen Problemstellungen ausführlich dargestellt; dabei ist neben dem "Handwerk" der kategoriellen Datenanalyse auch die Darstellung der sozialwissenschaftlichen Fragestellung ein wesentlicher Schwerpunkt.

Im Kapitel 1 werden die wichtigsten Begriffe und statistische Methoden, wie z.B. Modelle für kategoriale Daten, Spezifikation statistischer Modelle, ML-Schätzung, *goodness-of-fit*

Tests u.ä., besprochen. Im Kapitel 2 wird der von Grizzle, Starmer und Koch (GSK) vorgeschlagene Ansatz zur Analyse kategorialer Daten vorgestellt. Kapitel 3 befaßt sich mit der *maximum-likelihood* Schätzung der Parameter von Log-linearen Modellen. Im Rahmen dieser Darstellung werden u.a. Modellspezifikation, Modellauswahl und das Aggregierungstheorem besprochen und auf die Äquivalenz von Log-linearen und Logitmodellen für Kontingenztafel mit einer binären Responsevariablen eingegangen. Kapitel 4 befaßt sich mit *latent class* Analyse und Log-linearen Modellen mit latenten Klassen, Kapitel 5 mit der Logistischen Regression. In den anwendungsbezogenen Abschnitten bietet Kapitel 6 eine "GSK-Analyse über den Zusammenhang von objektiven Lebensbedingungen und subjektivem Wohlbefinden"; Kapitel 7 eine "Log-lineare Kohortenanalyse der Religiosität"; Kapitel 8 eine "Log-lineare Panelanalyse politischer Präferenzen". Kapitel 9 versucht schließlich die Frage zu beantworten, ob es den "rationalen Wähler" gibt.

Das Buch gibt einen sehr breit und sehr kompetent geschriebenen Überblick über die wichtigsten Verfahren zur Analyse kategorialer Daten, ein gerade in den Sozialwissenschaften wesentlicher Zweig der Statistik. Es wendet sich primär an den statistisch arbeitenden und interessierten Anwender; daher liegt der Schwerpunkt auf der anwendungsbezogenen Darstellung; z.B. wird auf die Interpretation der Modellparameter besonderer Wert gelegt. Die Anwendung der besprochenen Verfahren wird durch die ausführliche Darstellung realistischer Fragestellungen schmackhaft gemacht, während die mathematisch-technischen Voraussetzungen bewußt niedrig gehalten sind, niedriger als z.B. jene im konkurrierenden Agresti (1990). Als ein kleiner Nachteil ist zu werten, daß die rechen-technische Darstellung etwas in den Hintergrund tritt; für den Anwender wäre vielleicht der Hinweis auf die Handhabung von Computerprogrammen, zumindest im Bereich der *latent class* Analyse basierend auf dem EM-Algorithmus nützlich gewesen.

Literatur

Agresti, A. (1990). *Categorical Data Analysis*. New York: Wiley.

Walter Katzenbeisser
 Institut für Statistik
 Wirtschaftsuniversität Wien

Gregory C. REINSEL, **Elements of Multivariate Time Series Analysis**. Springer-Verlag, 1997, xvii+385 S., öS 715.40, ISBN 0-387-94918-6.

Das vorliegende Buch stellt eine Einführung in die klassische, multivariate Zeitreihenanalyse dar. Der Schwerpunkt liegt auf der Analyse von vektorwertigen ARMA-Modellen im Zeitbereich. Die nunmehr vorliegende 2. Auflage behandelt daneben auch noch Erweiterungen des ARMA-Ansatzes wie ARMAX oder integrierte Modelle. Analysen im Frequenzbereich werden (fast) nicht behandelt.

Kapitel 1 definiert Grundbegriffe der Zeitreihenanalyse (ARMA Modelle, Stationarität) und multivariater Daten (Matrizenoperation, multivariate Normalverteilung). Kapitel 2 behandelt die Eigenschaften vektorwertiger ARMA Modellen im Detail (Bedingungen für Stationarität, Prognose, etc.), Kapitel 3 stellt kanonische Strukturen für die Modelle vor (Kronecker Indizes, kanonische Korrelation). Die folgenden beiden Kapiteln behandeln Parameterschätzung für AR und ARMA Modelle inklusive verschiedener

Anpassungstest und Modellselektion auf Basis des AIC. Die abschließenden Kapitel 6–8 sind einigen Erweiterungen des klassischen Modells gewidmet: Kointegration und *unit roots*, Zustandsraummodelle, Kalman Filter sowie lineare Modelle mit exogenen Variablen (ARMAX). Im Anhang finden sich die Daten der im Buch verwendeten Beispiele.

Das Buch richtet sich laut Eigendefinition an Leser mit Erfahrung in der univariaten Zeitreihentheorie, was voll und ganz zutrifft. Alle Modelle werden von Beginn an sofort multivariat angeschrieben; die deshalb zwangsweise komplexere Notation und Vielfalt ist als Einstieg in die Zeitreihenanalyse sicher nicht zu empfehlen (dafür werden Doppelgleisigkeiten mit den univariaten Standardwerken vermieden). Der Stil des Autors ist insgesamt sehr unübersichtlich. Viele (auch längere) Formeln sind einfach im laufenden Text gesetzt und nicht abgesetzt in einer eigenen Zeile. Dasselbe trifft auf Definitionen und Sätze zu, die ebenfalls nicht getrennt gesetzt und nummeriert sind. Dies schmälert den Wert des Buches als Nachschlagewerk beträchtlich. Gerade ein Buch, das den Schwerpunkt auf die systematische Abhandlung klassischer Ansätze legt (und diese Anforderung inhaltlich auch voll erfüllt), hätte größeres Augenmerk auf die Qualität der Präsentation des Materials legen sollen.

Friedrich Leisch
Institut für Statistik und Wahrscheinlichkeitstheorie
Technische Universität Wien

Jeffrey D. HART, **Nonparametric Smoothing and Lack-of-Fit Tests**. Berlin: Springer-Verlag, 1997, xii+300 S., öS 576,70, ISBN 0-387-94980-1.

Die sorgsame Überprüfung eines gewählten Wahrscheinlichkeitsmodells sollte in keiner statistischen Analyse fehlen. Im wesentlichen ist man daran interessiert, wie gut sich ein Modell den beobachteten Daten anpaßt. Nichtparametrische Glättungsverfahren stellen in sehr vielen Bereichen ein mächtiges Instrumentarium dar, um diese Modellanpassung zu prüfen. Das vorliegende Buch untersucht die Einsatzmöglichkeiten diverser Glättungsverfahren, um die Güte der Anpassung parametrischer Regressionsmodelle zu testen. Dabei werden zahlreiche Methoden im Detail diskutiert und auch neue Verfahren vorgestellt.

Ist man mit dieser Thematik noch nicht vertraut, so bieten die ersten vier Kapitel eine allgemeine Einleitung anhand von Regressionsmodellen mit einer einzelnen Designvariablen. Dieser Buchteil könnte auch sehr gut als Unterlage für eine Lehrveranstaltung über nichtparametrisches Schätzen von Funktionen dienen. Hier wird gezeigt, wie Regressionskurven geschätzt werden können, wobei hauptsächlich auf Kern- und Fourierreihen-Schätzer eingegangen wird und der *lack-of-fit* Aspekt kaum behandelt wird. Verschiedene Ansätze für eine datengesteuerte Wahl des Glättungsparameters werden vorgestellt und die theoretischen (asymptotischen) Eigenschaften des resultierenden *smoothers* werden in einem übersichtlichen Satz/Beweis-Stil diskutiert. Die nächsten beiden Kapitel geben einen Überblick über klassische *lack-of-fit* Tests, wobei der Glättungsparameter vorerst als feste Größe betrachtet wird. Im folgenden, mathematisch anspruchsvollen Abschnitt wird dieser Parameter als stochastische Größe betrachtet und das Verhalten der darauf basierenden Teststatistiken studiert.

Die vom Leser für das Verständnis verlangten Voraussetzungen beschränken sich auf ein Grundwissen in der mathematischen und asymptotischen Statistik sowie der Analysis. Da sowohl auf theoretische als auch auf angewandte Aspekte eingegangen wird, sollte dieses Buch auch für den Praktiker interessant sein. Die gelungene Mischung aus mathematischer Theorie mit wichtigen Interpretationen und zum Verständnis hilfreichen Beispielen mit übersichtlichen Abbildungen lädt den Leser zum Studium und zur Anwendung dieser modernen Methoden ein. Offen bleibt jedoch für den Anwender die Frage, mit welcher Statistiksoftware die angeführten Beispiele nachvollzogen werden können.

Herwig Friedl
Institut für Statistik
Technische Universität Graz

Paul EMBRECHTS, Claudia KLÜPPELBERG und Thomas MIKOSCH, **Modelling Extremal Events for Insurance and Finance**. Berlin: Springer-Verlag, 1997, xv+644 S., öS 861.40, ISBN 3-540-60931-8.

Das Studium von extremalen und seltenen Ereignissen ist von jeher ein wichtiges Thema der Versicherungsmathematik gewesen. Das steigende Interesse an diesen Fragen ist einerseits durch die Erfordernisse der Praxis des Versicherungs- und Finanzwesens, andererseits durch die Fortschritte in der stochastischen Theorie der Extremalereignisse begründet. Zielsetzung der Autoren war es, eine anwendungsorientierte Darstellung jener Modelle und Methoden zu geben, welche eine präzise, mathematische Beschreibung von Extremalereignissen der Versicherungs- und Finanzmathematik erlauben. Das Buch richtet sich in erster Linie an Studierende der Versicherungsmathematik, an aktive Versicherungsmathematiker und Finanzexperten sowie an Theoretiker auf diesen Gebieten. Die *reader guidelines* geben in ca. 20 Seiten einen guten Überblick über das Gesamtvorhaben.

Der Haupttext beginnt mit einem Kapitel über die Riskentheorie. Besonderes Augenmerk findet dabei die Bedeutung von Subexponentialverteilungen. Die Kapitel 3-5 beinhalten die mathematischen Grundlagen zur Modellierung von Extremalereignissen: Grenzwertsätze für Summen, Grenzwertsätze für Maxima, Grenzverteilungen und deren Einzugsgebiete, Verhalten von Ordnungsstatistiken, Punktprozesse. Das Kapitel 6 ist der statistischen Analyse gewidmet: EDA von Extremalereignissen, Schätztheorie von Extremwertverteilungen, Inferenz für Überschreitungen von kritischen Werten. Letzteres ist ein Standardproblem der stochastischen Hydrologie und wird in Hochwasser-Risikoanalysen unter der Bezeichnung *peaks over threshold (POT)*-Methode seit 1970 erfolgreich verwendet. Die Kapitel 7 und 8 sind spezielleren Themen gewidmet: Zeitreihen mit starken Fluktuationen (*heavy tailed innovations*), Extremalindizes, *ruin*- und *run*-Studien, ARCH-Prozesse, stabile Prozesse.

Insgesamt kann gesagt werden, daß das Vorhaben, ein Werk zu schaffen, welches die Bedürfnisse der Versicherungs- und Finanzmathematik befriedigt, zumindest aus der Sicht des Stochastikers hervorragend gelungen ist. Es wird nicht nur die Theorie vorbildlich dargestellt, sondern es wird auch den Problemen des Anwenders breiter Raum eingeräumt. Die Druckqualität des Buches ist sehr gut, Inhaltsverzeichnis und Register sind

sehr detailliert. Im Anhang kann sich der Leser über die Grundlagen aus Wahrscheinlichkeits- und Prozeßtheorie informieren. Zweifellos wird das Buch eine Standardreferenz für Versicherungen und Finanzwelt werden. Darüber hinaus ist das Buch für alle Interessenten der Extremwertstatistik, insbesondere in den Erd- und Umweltwissenschaften, empfehlenswert.

Franz Konecny
Institut für Mathematik und Angewandte Statistik
Universität für Bodenkultur Wien

R.D. REISS und M. THOMAS, **Statistical Analysis of Extreme Values**. Basel: Birkhäuser-Verlag, 1997, xvi+336 S., öS 634.–, ISBN 3-7643-5768-1.

Das Buch beschäftigt sich mit der statistischen Modellierung und Analyse von "extremen" Ereignissen aus dem Umweltbereich (Flutwellen, Hurrikans oder hohe Ozonkonzentrationen) sowie der Finanz- und Versicherungswelt. Der Umstand, daß die Wahrscheinlichkeit einer zukünftigen Katastrophe nicht vernachlässigbar klein ist, kann Präventivmaßnahmen auslösen, die ein größeres Desaster verhindern sollen. Aus diesem Grund ist das Buch als Hilfestellung für Statistiker in allen Bereichen der Umweltforschung und der Ingenieurstätigkeiten sowie des Finanz- und Versicherungswesens gedacht, die sich mit Problemen extremer Werte beschäftigen. Das Buch ist in fünf Teile untergliedert und beinhaltet zusätzlich die Statistiksoftware XTREMES sowohl für Windows-NT und für Windows 3.1. Im ersten Teil werden der statistische Hintergrund für die Modellierung von extremen Werten und die wichtigsten parametrischen Modelle und nichtparametrischen Diagnosetools erläutert. Teil 2 beginnt mit inferenzstatistischen Methoden für den parametrischen Fall und für Poissonmodelle. Darauf aufbauend wird die für die Extremwertanalyse benötigte Methodologie wie etwa die generalisierten Paretomodelle entwickelt. Teil 3 beschreibt die Analyse multivariater Daten, wobei parametrische und nichtparametrische Extremwertmodelle vorgestellt und Schätzmethoden für multivariate Extremwertprobleme diskutiert werden. Teil 4, der spezifische Fragestellungen aus den Bereichen Versicherungswesen, Finanz und Hydrogeologie bespricht, geht direkt in Teil 5 über, wo einige Fallstudien zur Extremwertanalyse behandelt werden. Als Beispiele seien genannt: Hochwasserereignisse in Venedig, Lebensdauerabschätzungen westdeutscher Frauen, die Ozonbelastung von Houston, oder der Treibhauseffekt in Relation zur Clusterung von extremen Werten. Jeder der fünf Teile ist in mehrere Kapitel unterteilt, wobei am Ende jedes Kapitels Demos, Implementationsbeispiele und Übungen in XTREMES abgedruckt sind. Im Anhang ist eine Einführung in XTREMES angeschlossen, die das Menüsystem sowie die Programmiersprache XPL vorstellt. Neben einem Nachschlagewerk zum praktischen Arbeiten mit Daten extremer Ereignisse eignet sich das Buch durchaus auch als Textbuch für Kurse an Universitäten, wobei für das Verständnis eines Großteil des Buches nur ein statistisches Basiswissen notwendig ist. Nähere Informationen zur Software XTREMES findet man auch im Internet unter <http://www.extrems.math.uni-siegen.de>.

Klemens Fuchs
Institut für angewandte Statistik und Systemanalyse
Joanneum Research Graz

Andreas QUATEMBER, **Das Quotenverfahren**. Universitätsverlag Rudolf Trauner, 1996, 115 S., öS 248.–, ISBN 3-85320-792-8.

Die Freude des Statistikers über den emsigen Gebrauch der Errungenschaften seiner Disziplin ist getrübt, wenn praxisbedingte Kompromisse sich dogmatisch verfestigen; wenn gewisse Einschränkungen gar nicht mehr zur Sprache kommen; und also letztlich sogar Irreführung durch solchen Umgang mit den Ergebnissen nicht mehr ausgeschlossen ist. Das Quotenverfahren als die methodische Basis der allgegenwärtigen Markt- und v.a. der Meinungsforschung ist dafür ein vorzügliches Beispiel, und auch dafür, daß sogar die theoretische Disziplin vor soviel Wucht der Praxis zu kapitulieren scheint.

Diesen Punkt hat nun A. Quatember mit seiner Arbeit in einer Weise aufgearbeitet, die nicht nur den akademischen Statistiker anzusprechen vermag. Mit einem kurzen historischen Überblick über statistische Stichproben im allgemeinen und das Quotenverfahren im besonderen schafft der Autor zunächst ein Problembewußtsein. Er hält sich dabei mit dem formalen Apparat zurück und ermöglicht somit einen auch intuitiv noch leicht faßlichen Einstieg.

Hinterfragt wird zunächst das Konzept der Repräsentativität, das bei den (normalerweise gegebenen) kombinierten Merkmals-Strukturen eines solchen Erhebungsvorganges auf der Grundlage der Randquoten sicher nicht hält. Vielmehr müßte dabei die kombinierte Felderstruktur der gedachten Merkmalsmatrix angesprochen werden, was natürlich aufwendig ist, oder es müßten wenigsten die Assoziationswahrscheinlichkeiten der Randverteilungen mit den gesuchten Variablen bekannt sein. Anderenfalls bleibt eben das Problem keineswegs trivialer Varianzen der solcherart hochgerechneten Beobachtungen bestehen, wie Quatember formal und anhand von Beispielsmaterial durchaus überzeugend demonstriert. Diese Diskussion führt der Autor geschickt ab, anhand von Alternativ-Szenarien unter Idealbedingungen. Eine Quotenauswahl nach Randquoten wäre nur für eindimensionale Quotenmerkmale repräsentativ, bei mehrdimensionalen (kombinierten) Merkmalen nur im Falle der Unabhängigkeit untereinander unter gleichzeitig erfüllten Gleichverteilungsbedingungen. Die mangelnde Beachtung dieser Bedingungen vor allem für die einzelnen Teilstichproben im kombinierten Fall ist der Hauptgrund der hohen Varianz und Verzerrungsgefahr von Randquotenstichproben.

Das zweite Problem des Quotenverfahrens ist die Erhebungs-Situation, wo dem Interviewer beträchtliche Freiheitsgrade der Auswahl der Interroganden überlassen sind - ein anderer Umstand, der formal erstaunlicherweise überhaupt noch gar nicht angegangen worden ist, wie Quatember richtig bemerkt. Der letztere Effekt kann formal betrachtet, und wohl auch eingeeengt (Mehrdimensionalisierung der Quoten), aber nicht wirklich beseitigt werden.

Insgesamt wird damit der Auswahlfehler der Zentralpunkt: formal bezüglich der nicht erfaßten Varianz der Randquoten im Hinblick auf das Untersuchungsmerkmal; materiell bezüglich der arbiträren, nicht normalisierten Vorgangsweise der Interviewer. Die Schlußfolgerungen Quatembers können unter solchen Verhältnissen nur dahingehen, daß praktisch das Randquotenverfahren womöglich zu vermeiden wäre, oder zumindest die obigen zwei Bedingungen möglichst angenähert werden sollen, also ein Plädoyer für das mehrdimensionale schichtenspezifische Verfahren. Was die Interviewseite anbelangt, wird jede Empfehlung naturgemäß umso weniger klar ausfallen können. Auch der Ausblick auf die neue Praxis der Telefoninterviews ist da nicht unbedingt überzeugend.

Insgesamt ist dem kleinen Werk zu bescheinigen, daß ihm theoretisches und praktisches Interesse durchaus zukommen sollte. Angesprochen müßten sich nicht nur die akademischen Kreise sehen, sondern auch die Praktiker bis hin zur amtlichen Statistik, die immer mehr in Abgrenzungsschwierigkeiten zur kommerziellen Markt- und Meinungsforschung kommt; ja bis hin zu den Medien und ihrem ungenierten Umgang mit mancher Statistik sowieso. Das auszudeutschen bleibt freilich ein Anliegen an das statistische Fach.

Die hiermit mitangesprochene Interpretationsproblematik hätte vielleicht etwas mehr Beachtung verdient und wäre sicher ein lohnendes Feld für weiterführende Untersuchungen. Etwas mehr Ausführlichkeit hätte vielleicht auch den resümierenden Schlußfolgerungen zukommen können. Aber möglicherweise steht das alles ohnehin in einer (nicht eben so leicht zugänglichen) Langfassung. Etwas zu kurz gekommen scheint vielleicht auch die klassische soziologische Erhebungspraxis (zumindest als Alternativ-Referenz).

Ein Nachtrag zum Schluß: unbeschadet der Rechtschreibreform würde es genügen, "Hypotenuse" (im Gegensatz zur Hypothese) mit einem, und "kartesisch" ganz ohne h (im Gegensatz zum Katheter oder zur Katharsis) zu schreiben (S. 56/57), und hätten wir auch der Allozierung (gegenüber der Allokierung, S. 85) den Vorzug gegeben. Damit sei aber nichts anderes gesagt, als daß ein gediegener Beitrag zur akademischen Statistik nicht länger all den humanistischen Bildungsballast voraussetzt.

Alfred Franz
Österreichisches Statistisches Zentralamt, Wien

Friedrich HARTEN, Andreas MEYERTHOLE und Norbert SCHMITZ, **Prophetentheorie**. Stuttgart: B.G. Teubner-Verlag, 1997, 210 S., öS 394.-, ISBN 3-519-02737-2.

Bei der Prophetentheorie, einem jungen Teilgebiet der Wahrscheinlichkeitstheorie, geht es um Fragen folgender Art: Gegeben sind zwei Akteure (der "Statistiker" und der "Prophet") und eine Folge von Zufallsgrößen, (X_i) , $i = 1, 2, \dots$, die sich etwa als mögliche Auszahlungen an die Akteure interpretieren lassen. Während der Statistiker seine Entscheidung zu einem bestimmten Zeitpunkt i nur aufgrund seiner bisherigen Information vornehmen kann, also mittels einer geeigneten Stopzeit seinen Gewinn zu maximieren versuchen muß, kennt der Prophet die ganze Folge und kann somit die für ihn optimale Entscheidung treffen. Die Prophetentheorie beschäftigt sich mit dem Vergleich des möglichen Gewinnes des Statistikers und des Gewinnes des Propheten, im konkreten Beispiel läuft dies auf einen Vergleich von $V := \sup\{EX_\tau\}$, wobei τ die Stopzeit ist, und $M := E(\sup X_i)$ hinaus (tatsächlich wird also nicht mit der Folge selbst gearbeitet, sondern mit deren Verteilung). Zwischen M und V ist eine Reihe von Ungleichungen bekannt, etwa vom Typ $M \leq C.V$ oder $M - V \leq C'$ mit geeigneten Konstanten C bzw. C' , die sich für manche Modelle sogar scharf wählen lassen. Die beschriebene Problematik läßt sich im Rahmen einer spieltheoretischen Formulierung weitgehend verallgemeinern, und dieser Standpunkt liegt dem Buch zugrunde. Die gegebenen Folgen (oder allgemeiner: stochastischen Prozesse) entsprechen dann den Strategien des Propheten, während die Strategien des Spielers Stopzeiten sind. Insbesondere wird mit Hilfe geeigneter definierter Auszahlungsfunktionen des Spieles eine präzise Definition des Begriffes "Prophetenungleichung" u. dgl. möglich. Ein Fragenkomplex sind Reduktionsmöglichkeiten

für die Prophetenstrategien, z.B. von Folgen mit beliebigen Abhängigkeiten auf Supermartingale. In diesem Zusammenhang treten bisweilen auf den ersten Blick ungewohnte Phänomene auf. Während etwa die Annahme der Unabhängigkeit der Folge die Reduktion erleichtert, tritt bei zusätzlich vorausgesetzter identischer Verteilung das Gegenteil ein.

Das Buch liefert einen ausgezeichneten Überblick über das Gebiet. Die Martingaltheorie, die Theorie des optimalen Stoppens sowie die Spieltheorie spielen naturgemäß eine tragende Rolle in diesem Buch, und vom Leser werden einschlägige Kenntnisse vorausgesetzt. Der Text ist präzise verfaßt, die Beweise sind durchwegs ausgeführt, und zahlreiche Beispiele erläutern die allgemeine Theorie. Das Buch ist als Pionierleistung anzusehen, da es kaum zusammenfassende Darstellungen der Prophetentheorie gibt, sieht man von einem ohnedies schwer zugänglichen Vorabdruck von R.P. Kertz (1987) und einem kürzeren Überblicksartikel von T.P. Hill und R.P. Kertz (1992) ab, also Veröffentlichungen, die nicht mehr den neuesten Stand der Theorie wiedergeben. Der Gegenstand erweist sich in vielen Situationen als wichtig, wo Stoppen eine wesentliche Rolle spielt, beispielsweise in der Finanzmathematik. Das gut und ansprechend ausgestattete Buch schließt eine Lücke in der Literatur und kann bestens empfohlen werden.

*Wolfgang Wertz
Institut für Statistik und Wahrscheinlichkeitstheorie
Technische Universität Wien*

N. BALAKRISHNAN, *Advances in Combinatorial Methods and Applications to Probability and Statistics*. Basel: Birkhäuser-Verlag, 1997, xxxiv+596 S., öS 1300.—, ISBN 3-7643-3908-X.

The volume was written in honor of Sri Gopal Mohanty and it starts with an overview of his life and works, showing his pioneering contributions to lattice path counting and its application to probability and statistics. The book contains 32 chapters, each written by leading experts in theory and/or applications of combinatorial methods in probability and statistics. The articles are carefully organized in seven parts:

- I. Lattice paths and combinatorial methods. This part centers around counting paths, enumeration of lattice paths and their common features with random walks.
- II. Applications to probability. Here ballot theorems, planar random walks, generalized random walks and problems of network reliability are dealt with.
- III. Applications to urn models
- IV. Applications to queueing theory
- V. Applications to waiting time problems, including e.g. success and failure runs.
- VI. Applications to distribution theory. This part includes articles on problems of statistical physics and (Bernoulli) learning models.

VII. Applications to nonparametric statistics. This part is certainly most interesting to statisticians and deals e.g. with nonparametric tests of homogeneity of distributions against ordered alternatives and with testing the ratio of variance components in one-way random effects models.

The present volume clearly achieves its goal to bring together the work of theoreticians and applied researchers, to show recent advances in lattice paths, combinatorial and enumerative techniques and display their impact and significant role in probability and statistics. Fortunately, the book is not a mere collection of seemingly unrelated papers, which happens too often with collections of invited papers, but represents a good example of editorial art and taste. However, as a volume in the series “Statistics for Industry and Technology” it would have deserved a more direct focus on applications, e.g. on the exciting areas of graphical modeling and (telecommunication) network modeling.

*Jürgen Pilz
Institut für Mathematik und Statistik
Universität Klagenfurt*