

Vorwort

Dieses Buch ist aus dem Fortbildungskurs „Statistik in Theorie und Praxis“ entstanden, welcher vom 31. Oktober bis zum 6. November 2010 am Mathematischen Forschungsinstitut Oberwolfach durchgeführt wurde und sich an Gymnasiallehrer richtete. Hauptanliegen des Buches ist dem Leser zu vermitteln, wie vielseitig die oft als trockene Materie dargestellte Statistik sein kann.

In vielen Theorie-Büchern zur angewandten Statistik werden praktische Beispiele oft nur am Rande behandelt, ohne thematisch in die Tiefe zu gehen. Praxisorientierte Werke zur Statistik hingegen setzen beim Leser häufig ein fundiertes Hintergrundwissen der Statistik-Theorie zum Verständnis voraus. Das vorliegende Buch versucht den Brückenschlag zwischen beiden Gegensätzen. Zum einen wird Wert auf eine saubere Herleitung der gängigen statistischen Verfahren gelegt. Zum anderen beinhaltet der zweite Themenschwerpunkt reale Anwendungen der Methoden in der Praxis. Darüber hinaus werden Projekte an der Schnittstelle zwischen Schul- und Hochschulunterricht vorgestellt. Als unterstützende Software dient das kostenlose Programm R mit der zugehörigen grafischen Oberfläche R-Commander; die zahlreichen Programmbeispiele ermöglichen es, die vorgestellten Themen eigenständig nachzuvollziehen.

Zielgruppe

Das Buch soll Leser mit Interesse an den Problemstellungen von Mathematik und Statistik ansprechen. Dazu gehören Anwender statistischer Methoden, zum Beispiel im wissenschaftlichen Bereich, oder Studierende (der Mathematik) mit einer tiefergehenden Neugier an statistischen Fragestellungen. Genauso aber sollten sich auch solche Leser angesprochen fühlen, die bisher wenig in Kontakt mit wissenschaftlichen Methoden der Statistik gekommen sind, aber an diesen Interesse haben. Insbesondere ist das vorliegende Buch für Lehrer und Studierende des Lehramts an Gymnasien gedacht, die den Schülern im Rahmen des Unterrichts oder in Unterrichtsprojekten die Vielfalt der Statistik näher bringen wollen.

Vorkenntnisse

Für die Lektüre werden lediglich elementare Stochastik-Kenntnisse vorausgesetzt, wie sie beispielsweise in einer Einführungsvorlesung im Rahmen eines Bachelor-Studiums vermittelt werden. Für Leser ohne Vorwissen in Stochastik werden die wichtigsten nicht im Text erläuterten Grundbegriffe in einem Glossar am Ende des Buchs zusammengefasst,

weshalb die Inhalte des Buchs auch für Neulinge zugänglich sein sollten. Darüber hinaus findet man am Ende jedes Kapitels Hinweise auf weiterführende Literatur.

Kenntnisse des Programmpakets R, mit dem die Beispiele im Buch vorgestellt werden, sind zum Verständnis nicht vorausgesetzt. Alle verwendeten Funktionen von R werden ausführlich erläutert; im letzten Teil des Buchs wird im Rahmen einer R-Einführung zusätzlich die Arbeitsweise der Software vorgestellt.

Inhalt

Die unterschiedlichen Thematiken des Buchs werden in fünf einzelnen Teilen vorgestellt. Statistik zu verstehen ohne die theoretischen Grundlagen zu kennen, ist nicht möglich. Aus diesem Grund behandelt der erste Teil des Buchs fundamentale statistische Grundlagen („Statistik in der Theorie“). Dabei wird der mathematische Hintergrund beleuchtet, ohne an allen Stellen ins Detail zu gehen.

Im Unterschied zu den meisten anderen Lehrbüchern wird auf die praktischen Anwendungen nicht nur in einzelnen kurzen und gekünstelten Beispielen eingegangen. Vielmehr wird diesem Aspekt der Statistik ein kompletter Teil des Buchs gewidmet („Statistik in der Praxis“). In diesem werden reale Anwendungsbeispiele ausführlich und verständlich erläutert und die im ersten Teil besprochenen Verfahren angewandt.

In den letzten etwa 20-30 Jahren haben in der Statistik rechenintensive, computergestützte Simulationsverfahren eine zentrale Rolle eingenommen. In einem Buch, das die Vielfalt der Statistik vorstellen möchte, darf dieser Aspekt natürlich nicht fehlen. Im dritten Teil des Buchs („Statistik mittels Simulationen“) wird diesen Entwicklungen Rechnung getragen, indem einige Beispiele gesondert und ausführlich präsentiert werden.

Da das Buch insbesondere auch Lehrende an Schulen und Hochschulen ansprechen soll, wird im vierten Teil des Buchs („Statistik als Projekt im Unterricht“) darauf eingegangen, wie man mittels einfacher Fragestellungen Lernende zum eigenständigen Bearbeiten statistischer Probleme motivieren kann. Die Themen in diesem Teil sind das Ergebnis von tatsächlichen Projekten, die in der Vergangenheit von Schüler- und auch Lehrergruppen erfolgreich durchgeführt wurden. Wir hoffen, dass die hier gegebenen Anregungen einen Beitrag zur Einbeziehung praxisnaher Unterrichtsprojekte in Schule und Hochschule geben können.

Alle vorgestellten Beispiele in den Teilen I bis IV werden mit der Software R durchgeführt. Da aber gerade Statistik-Anfänger Schwierigkeiten mit der Komman-

dozeilenorientierung von R haben, versuchen wir den Einstieg in R zu erleichtern, indem wir – wo immer es möglich ist – den R-Commander als grafische Oberfläche zu benutzen. Der letzte Teil des Buchs stellt daher eine Einführung sowohl in R als auch in den R-Commander dar.

Organisation des Textes, Notation und technische Details

Im ersten Teil des Buchs werden die theoretischen Grundlagen besprochen, die Teile II bis IV bauen auf diesen Grundlagen auf, können selbst aber unabhängig voneinander durchgearbeitet werden. Durch das komplette Buch ziehen sich Programmbeispiele, in denen die besprochenen Verfahren in R umgesetzt werden. Lesern, die noch keine Erfahrung mit R haben und gleich die entsprechenden R-Beispiele am Rechner mitverfolgen wollen, empfehlen wir mit der Lektüre des Teils „Statistik für Einsteiger mit R und R-Commander“ zu beginnen.

Die Durchführung der Beispiele kann im kompletten Buch mit dem R-Commander erfolgen. Sollten die Befehle zu komplex werden und der R-Commander an seine Grenzen stoßen, werden die nötigen R-Befehle erläutert, die aber weiterhin im Skriptfenster des R-Commanders ausgeführt werden können.



Programmbeispiele werden – wie hier zu sehen – im Text immer mit dieser Abtrennung kenntlich gemacht. Möchte man auf die eigene Bearbeitung der Beispiele verzichten, können diese Teile des Buchs übersprungen werden.

Für ein besseres Verständnis wollen wir kurz die verschiedenen Notationsweisen innerhalb der Programmbeispiele erläutern. Pfadverzeichnisse (z.B. **C:\R-Buch**) oder Paketnamen (z.B. **Rcmdr**) erkennt man an der fetten Schrift. Genauso werden Menüpunkte in fetter Schrift gedruckt, das Klicken von mehreren Menüpunkten hintereinander wird durch das Zeichen \longrightarrow angezeigt. So bedeutet etwa **Grafiken** \longrightarrow **Histogramm** ..., dass zuerst der Menüpunkt **Grafiken** und darin das Untermenü **Histogramm** ... angeklickt werden soll. Schaltflächen in Dialogfeldern, wie beispielsweise **OK** oder **Datenmatrix betrachten** sind stets umrahmt, Optionen zum (De-)Aktivieren in Dialogfeldern sind an der *kursiven* Schrift zu erkennen. Einzelne Tastenbefehle sind in Großbuchstaben gesetzt (z.B. ENTER), wobei ein +-Zeichen zwischen zwei Tastennamen das gemeinsame Drücken beider Tasten bedeutet (z.B. STRG+V). Müssen in Dialogfeldern freie Texte in spezielle Felder eingegeben werden, steht der Text in Anführungszeichen (z.B. „Größe des Manns“), genauso wie Karteinamen innerhalb eines Dialogfelds (z.B. „Optionen“). Längere Befehle (z.B. bei der Spezifizierung einer Fallauswahlbedingung) oder Befehle, die im Skriptfenster eingegeben werden müssen, werden abgesetzt vom normalen Fließtext in *Festbreitenschrift* gesetzt.



Ganz allgemein werden Definitionen und neu eingeführte Begriffe im Text mit **fetter** Schrift hervorgehoben. Datensatz- und Variablenamen sind an der

Festbreitenschrift zu erkennen. Das Ende eines Beweises wird mit dem Symbol \square gekennzeichnet.

Das Betriebssystem des Rechners, mit dem alle Berechnungen und Diagramme durchgeführt wurden, war Windows 7. Die Version von R war 3.0.3, die des R-Commanders lautete 2.0-0. Sämtliche in den Programmbeispielen verwendeten Datensätze stehen unter

<http://www.springer.com/>

zum Download allen Nutzern zur Verfügung.

Danksagungen

Wir danken

- ▶ Andreas Gegg für die vielfältige Unterstützung insbesondere bei spezifischen R-Problemen
- ▶ Thomas Mauch für die Mitautorschaft an den Kapiteln 8 und 9
- ▶ Norbert Krämer, Herbert Michel und Stefan Englert für die in Kapitel 10 vorgestellte gemeinsame Projektarbeit
- ▶ Bertram Gerber, Kirsä Neuser und Birgit Michels für die Unterstützung bei Kapitel 12 und die Erlaubnis zur Verwendung der Bilder in diesem Kapitel
- ▶ Christina Zube und Nicole Saverschek für die fachliche Hilfestellung in Kapitel 14
- ▶ Nicole Vornberger für die medizinischen Hintergrundinformationen zu Kapitel 15
- ▶ Sabine Sigloch für die Anmerkungen zu den Projektkapiteln in Teil IV
- ▶ den Teilnehmern an den Schülerprojekttagen und der Lehrerfortbildung in Oberwolfach für die engagierte Teilnahme an den in Teil IV beschriebenen Projekten
- ▶ Hassan Humeida für die Bereitstellung des Datensatzes `sudan.csv`
- ▶ Stefanie Linder für die Bereitstellung der Daten in Kapitel 16
- ▶ Diana Tichy und Oka Daiji für die Bereitstellung der Daten in Kapitel 15
- ▶ Peter Zimmermann von der Katholischen Universität Eichstätt-Ingolstadt für die \TeX -nische Hilfe
- ▶ Christian Weiß für die Durchsicht des kompletten Manuskripts
- ▶ Kerstin Werler von der Katholischen Universität Eichstätt-Ingolstadt für die vielfältige Hilfe bei der Texterstellung
- ▶ dem Springer Verlag für die Bereitschaft, dieses Buch zu veröffentlichen. Insbesondere danken wir Agnes Herrmann und Clemens Heine für die fruchtbare Kooperation.

Würzburg
Eichstätt
Frankfurt
März 2014

*Michael Falk, Johannes Hain, Frank Marohn
Hans Fischer
René Michel*

Statistik in Theorie und Praxis

Mit Anwendungen in R

Falk, M.; Hain, J.; Marohn, F.; Fischer, H.; Michel, R.

2014, XV, 507 S. 140 Abb., Softcover

ISBN: 978-3-642-55252-6