

Book Reviews

Ulrich KOHLER und Frauke KREUTER (2008). **Datenanalyse mit Stata. Allgemeine Konzepte der Datenanalyse und ihre praktische Anwendung** (3. aktualisierte und überarbeitete Auflage) München, Oldenbourg Verlag, 398 S., ISBN 978-3-486-58456-1. (broschiert € 34,80)

Das vorliegende Buch trägt als Untertitel “Allgemeine Konzepte der Datenanalyse und ihre praktische Anwendung” und es ist wohl auch der Anspruch der AutorInnen mehr zu liefern als ein einfaches Handbuch zum Gebrauch von Stata. Dies Ansinnen ist nur zum Teil geglückt. Einerseits werden durchaus Aspekte behandelt, die in herkömmlichen Statistiklehrbüchern selten vorkommen, etwa die Ausreißerdiagnose und -behandlung, andererseits ist die Vermittlung von spezifischen Konzepten oft zu kochrezeptartig, der Exkurs zur Erklärung des Maximum-Likelihood-Prinzips anlässlich der logistischen Regression steht exemplarisch dafür.

Inhaltlich spannt das Werk einen durchaus großen Bogen, wobei die ersten 5 Kapitel auf 105 Seiten für Stata-Einführung und Syntax aufgewendet werden. Danach wendet es sich deskriptiven Techniken, Grafiken, Verteilungen, etc. auf weiteren knapp 80 Seiten in zwei Kapiteln zu. Das eigentliche Kernstück bilden die Kapitel 8 über die Regression für metrische abhängige Variablen (70 Seiten) und für kategoriale abhängige Variablen (50 Seiten), worauf wieder drei Stata-spezifische Kapitel zum Umgang mit Datensätzen, der Programmierung und abschließend dem Umfeld (Ressourcen, community, etc.) folgen, was einen Gesamtumfang von etwa 400 Seiten ergibt.

Grundsätzlich liest sich der Text recht locker und wenn man die zahlreichen angegebenen Beispiele gleich mit ausführt, kann man auf diese Weise den Umgang mit Stata sicher sehr gut erlernen. Für diesen Zweck hat sich das Buch anscheinend im deutschsprachigen Raum - durchaus zu Recht - als Referenzwerk etabliert. Um die grundlegenden Konzepte der Datenanalyse verstehen zu lernen, sollte man ihm allerdings Ergänzendes zur Seite stellen.

*Werner Müller
IFAS, Institut für Angewandte Statistik
Universität Linz*

Gerhard MARINELL und Gabriele STECKEL-BERGER (2008). **Einführung in die Statistik** (2. Auflage) München, Oldenbourg Verlag, 236 S., ISBN 978-3-486-58713-5. (broschiert € 29,80)

Dieses Buch ist keinesfalls eine Einführung in die Statistik, sondern setzt fundierte Kenntnisse von ihr voraus. Ausnahmen davon sind wiederholte Erklärungen wie z.B. n Fakultät, Stichprobenmittel und -varianz berechnet werden. Im Gegensatz dazu werden Kenntnisse über die Bedeutung von Unabhängigkeit, über Anwendungsmöglichkeiten bestimmter Wahrscheinlichkeitsverteilungen, Bilden von Hypothesen und Testentscheidungen vorausgesetzt.

Laut Vorwort wird dem Leser Hilfe angeboten, statistische Verfahren zu verstehen, wobei die Erklärungen des Rechengangs im Vordergrund stehen. Entgegen der Erklärung des Verlags richtet sich das Buch nicht an Studierende der Betriebs- oder Volkswirtschaft. Diese würden nur wenige der gestellten Aufgaben lösen können – und diese auch nur deswegen, weil es sich durch die Zugehörigkeit zum betreffenden Unterkapitel ergibt, welches Verfahren anzuwenden ist. Wünschenswert wäre auch die Angabe der Lösungen, die in den meisten Fällen nur aus einer einzelnen Zahl oder aus einer Ja/Nein-Entscheidung bestehen würden.

Die schwierige Lösung der Frage, welches der vielen von einem Programmpaket angebotenen Verfahren richtigerweise anzuwenden ist, bleibt dem Fachstatistiker vorbehalten und kann einem Laien nur schwer vermittelt werden. Wünschenswert wären jedenfalls Erwähnungen von Gegenbeispielen, wann und warum ein Verfahren nicht anzuwenden ist. Etwa in Aufgabe 4 des Kapitels 2_4 müsste darauf hingewiesen werden, dass es sich nur um eine Rechenübung handelt und es gerade bei einer Stichprobe von Lebensdauern bedenklich ist, eine normalverteilte Grundgesamtheit anzunehmen.

Im Kapitel 9_1 müsste man andeuten, bei einem Stichprobenumfang von 5 den kritischen Wert einer Tabelle zu entnehmen, und sich nicht auf die Asymptotik zu verlassen. Weiters müsste vom KS-Test abgeraten werden, weil die Verteilungsfunktion mit geschätzten Parametern berechnet worden ist. Vom Leser wird das Wissen vorausgesetzt, den KS-Test bei jeder stetigen Verteilung und auch bei kleinsten Stichproben anwenden zu können und nicht auf Vor- und Nachteile gegenüber einem anderen Anpassungstest hingewiesen. Weiters müsste bei der Berechnung der Prüfgröße erklärt werden, von welchen Differenzen richtigerweise das Maximum zu suchen ist. Um keine Verwirrung zu stiften, müsste das in der Tabelle falsch eingetragene Maximum durch das am Ende richtig angegebene ersetzt werden.

Die Vorgangsweise, zu den Beispielen über Konfidenzintervalle oder Testentscheidungen auch eine Dichte abzubilden, könnte zum Verständnis beitragen. Sie ist aber nur dann hilfreich, wenn beschrieben wäre, um die Dichte welcher Verteilung es sich handelt. Aus den ersten 6 Abbildungen ist nicht abzulesen, welche Dichte gemeint ist. Die unmittelbar unter der Beschreibung “Dichtefunktion” angegebenen Konfidenzgrenzen sind keineswegs Quantile aus dieser Verteilung. In der 7. Abbildung ist die Konfidenzuntergrenze ein Quantil dieser Verteilung, die Obergrenze aber eines aus einer anderen.

Die Gliederung des Buches ist übersichtlich und gut gelungen. Etwas störend sind die an die Nummern der Hauptkapitel angehängten $_0$, die sich zusätzlich noch auf jeder Seite in der Kopfzeile wiederholen.

Das Buch ist für fortgeschrittene Statistikstudenten empfehlenswert. Für Studenten anderer Fächer müsste auf der Rückseite des Einbands hingewiesen werden, dass ihnen nur der Rechengang verschiedener Verfahren erklärt wird, sie diese aber ohne Vorbildung nur schwer verstehen werden.

Heinrich Potuschak
IFAS, Institut für Angewandte Statistik
Universität Linz