Buchbesprechung: Datenanalyse mit SAS

• Rainer Muche¹

Bibliographische Angaben

Krämer, W., Schoffer, O., Tschiersch, L.:

Datenanalyse mit SAS - Statistische Verfahren und ihre grafischen Aspekte

Springer Verlag, Berlin, 2005

376 Seiten, € 27,95, ISBN 3-540-20787-2

Rezension

Das Buch "Datenanalyse mit SAS" ist ein neues Lehrbuch zum Einsatz von SAS im Umfeld der Statistik. Es wendet sich "vor allem an Praktiker, die mittels des Programmpaketes SAS selbst umfangreiche Datenbestände analysieren wollen und setzt nur die wichtigsten statistischen Verfahren und Begriffe als bekannt voraus." Entsprechend sind die Schwerpunkte des Buches gesetzt: Der erste Teil beschreibt die Oberfläche des Systems und die Hilfskomponenten, bevor ausführlich auf den DATA- und PROC-Step der SAS-Syntax eingegangen wird. Anschließend werden erst Möglichkeiten zur Präsentation der Ergebnisse (grafisch und tabellarisch, inklusive ODS-System) erläutert, bevor die Umsetzung grundlegender statistischer Methoden (deskriptive Statistik, statistische Tests und Regressionsanalyse) und Spezialgebiete der Statistik mit SAS angegeben werden. Die Spezialgebiete umfassen die Varianzanalyse, Nichtparametrische Verfahren, Multivariate Verfahren, Zeitreihenanalyse und Ökonometrische Analysen. Darüber hinaus gibt es Kapitel zur Matrixsprache IML, zur Makroprogrammierung und zu verschiedenen Assistenten im SAS-System.

Das Buch ist entstanden aus einer Vorlesungsreihe an der Abteilung Statistik der Universität Dortmund. Dies kommt dem Leser zugute, da so viele, schon bewährte Beispiele, Aufgaben (inklusive Lösungen) und die entsprechenden Daten (im Internet verfügbar) dem Leser das Erlernen der SAS-Syntax erleichtern. Die Beispiele sind sehr anschaulich, speziell im ersten einleitenden Teil des Buches. Zum Beispiel gibt es hier die Beschreibung zur Erzeugung von Zufallszahlen

mit Beispielprogrammen, was selten in Lehrbüchern über SAS präsentiert wird.

Der Umfang der beschriebenen statistischen Methoden ist sehr groß, wobei nicht jeder Leser des Buches alle Auswertungsmethoden für seine eigenen Auswertungen benötigen wird. Aber die Breite der unter dem Kapitel "Spezialgebiete und Anwendungen" präsentierten Verfahren ist sicher ein Plus des Buches. Allerdings kann dann natürlich nicht jedes Verfahren bis in die Tiefe erläutert werden. Im Bereich der grundlegenden statistischen Verfahren hätten die Autoren aber an der einen oder anderen Stelle eine umfangreichere Beschreibung wählen können. Das Kapitel Signifikanztests ist z.B. mit 5 Seiten und der Beschreibung der beiden t-Tests äußerst knapp ausgefallen.

Die Autoren gehen in ihrer Beschreibung der Nutzung von SAS, speziell im Abschnitt des Datenmanagements, von einer alleinigen Nutzung der SAS-Syntax aus, ohne auf Möglichkeiten hinzuweisen, die durch die Nutzung des verbreiteten Betriebssystems MS-Windows mitgeliefert werden. So wird z.B. auf mehreren Seiten das Einlesen von ASCII-Datensätzen mit dem Input-/infile-Statement beschrieben, ohne auf die Einlesemöglichkeiten von z.B. Excel-Dateien (mit oder ohne Menüsteuerung) einzugehen. Speziell diese Möglichkeiten helfen aber dem Neueinsteiger, baldmöglichst mit dem Auswerten der vorliegenden Daten beginnen zu können. Weitere Möglichkeiten, die das System unter dem Betriebssystem MS-Windows bietet, werden ebenfalls nicht erwähnt in den zentralen Kapiteln (s. Untertitel des Buches) zur "Präsentation und Aufbereitung von Ergebnissen". Einfache Möglichkeiten des Transfers der Ergebnisse (z.B. über Copy/Paste) in Berichte oder Präsentationen sind ein wichtiges Hilfsmittel, was man Anfängern mit der Software mitgeben sollte. Ebenfalls hilfreich sind hier die Assistenten von SAS, die über eine maus- und menügesteuerte Oberfläche SAS-Syntax erzeugen können. Beide von SAS dafür angebotenen Assistenten (SAS-Analyst und Enterprise Guide) laufen nur unter der Windows-Oberfläche, können aber beim Erlernen der umfangreichen Syntax extrem hilfreich sein. SAS-Analyst wird in dem Buch kurz erwähnt (ohne die Möglichkeit der

¹ Abteilung Biometrie und Medizinische Dokumentation, Universität Ulm, Ulm, Deutschland

Syntax-Generierung), der Enterprise Guide erst gar nicht.

Ein weiteres Manko des Buches ist die Einschränkung der Beschreibung der SAS-Syntax auf die Version 8.2. Seit längerem ist die SAS Version 9 erhältlich. Ein Buch, was als Erscheinungsjahr das Jahr 2005 trägt sollte mindestens auf den Verlagsseiten und/oder im Klappentext und Vorwort explizit darauf hinweisen, dass die neueste Version nicht Grundlage des Inhalts ist. Es gibt genügend Veränderungen zwischen den beiden SAS-Versionen, so dass der Leser nicht sicher sein kann, ob die beschriebenen Prozeduren und Beispielprogramme auf der aktuellen Version 9 noch lauffähig sind.

Als Fazit kann das Buch etwas fortgeschrittenen Anwendern empfohlen werden, die bereit sind, sich von Beginn an in die umfangreiche SAS-Syntax einzulernen. Ebenfalls hilfreich wird das Buch denjenigen Personen sein, die auf Basis ihrer bisherigen rudimen-

tären SAS-Kenntnisse einiges dazulernen wollen in Bezug auf Datenmanagement, Präsentation und spezielle Auswertungsmethoden. Komplette Neuanfänger mit dem System sollten mit entsprechenden Hilfsmitteln wie maus- und menügesteuerten Systemen, Betriebssystemhilfen etc. vertraut gemacht und vorsichtiger an SAS herangeführt werden. Dies hängt natürlich vom Kenntnisstand und Ausbildungsstand der Leser ab, sollte aber nicht so unberücksichtigt gelassen werden. SAS-Anwender, die sich neueste Infos zu Auswertungen mit der Version 9 erhoffen, werden enttäuscht sein.

Korrespondenzadresse:

• Dr. biol. hum. Rainer Muche, Abteilung Biometrie und Medizinische Dokumentation, Universität Ulm, Schwabstr. 13, 89075 Ulm, Tel: ++49-731-502-6903, Fax: ++49-731-502-6902
rainer.muche@medizin.uni-ulm.de