



SABINE FROMM
 2010: Datenanalyse
 mit SPSS für
 Fortgeschrittene 2:
 Multivariate
 Verfahren für
 Querschnittsdaten.
 Wiesbaden:
 VS Verlag für
 Sozialwissenschaften.
 ISBN: 978-3-531-
 14792-5, 257 Seiten,
 24,95 EUR.

Der Band „Datenanalyse mit SPSS für Fortgeschrittene 2: Multivariate Verfahren für Querschnittsdaten“ von Sabine Fromm ist der zweite Band einer Reihe. Teil 1 dieser SPSS Lehrbuchreihe beschäftigt sich mit den grundlegenden Problemen der Vorbereitung quantitativer Auswertungen (z. B. die Schritte vom ausgefüllten Fragebogen zum analysefähigen Datensatz, wie werden Daten bereinigt und wie können neue Variablen berechnet werden, etc.), während Teil 2 weiterführende Verfahren für die Analyse von Querschnittsdaten beinhaltet.

Das erste Kapitel führt in Mittelwertvergleiche ein. Das einführende Beispiel zur Varianzanalyse beschreibt einen Anwendungsfall mit drei Faktoren (16 Untergruppen), der relativ gut erläutert wird. Besonders gelungen ist die Veranschaulichung kleiner und großer Varianzen in Teilgruppen (S. 32 und 33). Ob ein Beispiel mit drei Faktoren unnötig komplex und damit als Einstieg in die Thematik weniger geeignet ist, kann bei der Darstellung der Aufteilung der Gesamtstreuung gefragt werden. Die Vielzahl an Effekten scheint eine handhabbare Menge leider deutlich zu überschreiten. Ein weiteres Beispiel ist zum Glück einfacher gehalten und illustriert in durchaus gelungener Weise die Funktionsweise der Varianzanalyse.

Anschließend wird die Faktoren- und Reliabilitätsanalyse vorgestellt (Kapitel 2).

Die Faktorenanalyse wird als „Verallgemeinerung der Dimensionsanalyse nach dem Modell der Likert-Skalierung“ dargestellt (S. 59). Die rotierte Matrix stellt das Anwendungsbeispiel anschaulich dar. Anschließend wird Cronbachs Alpha zur Reliabilitätsanalyse erläutert. Ob man Variablen, die nach einer Faktorenanalyse zu einem Faktor zusammengestellt wurden, zusätzlich einer Reliabilitätsanalyse unterziehen muss, kann zwar hinterfragt werden; prinzipiell sind die Erläuterungen zur Reliabilitätsanalyse jedoch gut gelungen.

Kapitel drei stellt die lineare Regression vor. Die Einführung in die statistischen Grundlagen der linearen Regression erfolgt hierbei routiniert. Der Schwerpunkt der Diskussion der Modellvoraussetzungen liegt auf der Multikollinearität. Bei der Vorstellung des Variance Inflation Factor (bzw. Toleranz) wäre es allerdings hilfreich gewesen zu erfahren, wo sich die Grenzwerte befinden, die bedenklich hohe Multikollinearität anzeigen. In der Durchführung der Regressionsanalyse wird den LeserInnen zunächst das schrittweise Prüfen auf Signifikanz einzelner Variablen nahegelegt. Hier wäre ein Hinweis darauf, dass diese Option im Zweifelsfall zum „theorielosen Auswerten“ verleitet, durchaus sinnvoll gewesen. Anschließend wird die gleichzeitige Aufnahme von Variablen in die Berechnung besprochen (Method = Enter).

Kapitel vier behandelt die logistische Regression. Auch hier erfolgt die Einführung in das Thema fachkundig. Gut gelungen ist auch der Abschnitt zur Transformation kategorialer unabhängiger Variablen. Die Erläuterungen zu den einzelnen Optionen der Syntax logistischer Regressionen sind vielseitig, fallen aber recht knapp aus. Leider ist das Beispiel der Wahlbeteiligung nicht glücklich gewählt, da NichtwählerInnen durch das Modell fast nicht vorhergesagt werden können (siehe S. 130). So gesehen „nützt“ es auch nichts, wenn der Hosmer-Lemeshow Test akzeptable Werte ausweist. Zum HL-Test hätte man sich zu-

dem eine etwas detailliertere Ausführung gewünscht, was genau hier auf „Goodness of Fit“ geprüft wird. Im folgenden Schritt wird zur Erläuterung der multinominalen logistischen Regression die Parteipräferenz mit fünf Ausprägungen untersucht. Anschließend folgt die Einführung in die Diskriminanzanalyse, wofür dieselben beiden abhängigen Variablen verwendet werden wie im vorausgegangenen Abschnitt zur logistischen Regression. Dies macht durchaus Sinn, da beide Verfahrenstypen in der vorgeschlagenen inhaltlichen Anwendung ähnlich sind.

Das nächste Kapitel ist der Clusteranalyse gewidmet. Hier wird zwischen Verfahren zur vorgegebenen Anzahl von Clustern (Austauschverfahren) und hierarchischen Verfahren, die die geeigneten Clusteranzahlen ermitteln, unterschieden. Im Anwendungsbeispiel wird die Klassifikation von Ländern vorgenommen. Anschließend wird der interpretative Schritt der Klassendiagnose durchgeführt und die Cluster werden über einen Mittelwertvergleich inhaltlich näher beschrieben.

Im letzten Kapitel wird die Korrespondenzanalyse behandelt. Es handelt sich hierbei um „ein exploratives Verfahren zur Visualisierung der Datenstruktur einer Kontingenztafel“ (S. 223). Die Beispieldaten sind hier allerdings etwas schwierig in der Handhabung, ebenso die Syntax, die nicht klar auf die Datensätze zugreift. In der graphischen Darstellung werden Länder und Einstellungen auf zwei Dimensionen dargestellt. Je geringer die Distanz von Land und Einstellung, desto stärker die Assoziation.

Allgemein ist anzumerken, dass Studierende es eventuell verwirren könnte, wenn Effekte auf dem Niveau von z.B. $\alpha = 0,053$ oder $0,099$ an der einen Stelle als signifikant bezeichnet werden und an einer anderen Stelle der T-Wert von 1,96 als Grenzwert kommuniziert wird, der bekanntlich dem 5%-Niveau entspricht. Natürlich ist das nicht falsch, da man prinzipiell Signifikanzniveaus

frei wählen kann. Dennoch hätte man zu sonst üblichen Konventionen klar Stellung beziehen können.

Generell ist das Layout des Buches zu beanstanden. Tabellen sind abgeschnitten über verschiedene Seiten verteilt, Zahlenwerte in Tabellen sind nicht am Komma ausgerichtet oder sehr stark verkleinert und einige Spaltenbeschriftungen sind unleserlich. Einerseits liegt das natürlich an den SPSS Outputs, andererseits hätte man diese Tabellen durchaus nachbearbeiten können, um die Lesbarkeit zu verbessern. Leider sind auch Formeln oder formelhafte Erläuterungen optisch überwiegend wenig überzeugend gesetzt. Von Lesbarkeit und Ästhetik abgesehen sollten zumindest Abbildungen den laufenden Text oder andere Abbildungen nicht verdecken (S. 175). Durch die Verwendung alternativer Software, wie beispielsweise LaTeX, hätte man das Layout des Buches sicherlich verbessern können.

Die Beispieldatensätze, die von <http://www.vs-verlag.de/> herunter geladen werden können, stellen eine gelungene Ergänzung zum Buch dar und regen an, nicht nur die besprochenen Analysen zu replizieren, sondern diese auch zu variieren und eigene Analysen zu erstellen. Zu Kapitel 2 bis 7 sind auch die Syntaxdateien bereitgestellt.

Alles in Allem kann das Buch unter Anleitung in der Lehre eingesetzt werden. Für das Selbststudium ist das Lehrbuch meiner Meinung nach weniger gut geeignet, da sich zum einen einige Bereiche zu unübersichtlich gestalten und oft in einer Vielzahl knapp beschriebener Details verlieren, und zum anderen wurden bestimmte Aspekte zu oberflächlich dargestellt. Trotz des relativ geringen Umfangs von ca. 250 Seiten wird dennoch ein angemessener Überblick zu multivariaten statistischen Verfahren sozialwissenschaftlicher Analyse für Querschnittsdaten geboten.