

# Vorwort

Gegenstand dieses Lehrbuchs sind statistische Modelle der Regressionsanalyse, die in den modernen Sozialwissenschaften eine zunehmende Bedeutung erlangen bzw. teilweise bereits zum fortgeschrittenen Standardrepertoire gehören. Es bietet eine Einführung in Verfahren für kategoriale, nicht normal verteilte abhängige Variablen, die für die Analyse von Zuständen verwendet werden und aus denen sich das Verfahren der Ereignisanalyse (synonym: Hazardratenmodelle, Survivalanalyse, Failure-time-models) herleiten lässt. Als Einführung in diese anspruchsvolle Materie setzt das Buch daher solide Grundkenntnisse der deskriptiven und der schließenden sozialwissenschaftlichen Statistik sowie der linearen Regression voraus, die mindestens auf dem Niveau eines gründlichen Studiums der Lehrbücher von Urban und Mayerl (2006) und Diaz-Bone (2006) liegen sollten.

In der Ökonometrie werden die hier dargestellten Verfahren häufig unter dem Sammelbegriff der „Categorical and Limited Dependent Variables“-Modelle behandelt (Maddala 1983; Long 1997; Greene 2000, S. 896). Aus einer anderen Perspektive lässt sich zeigen, dass nahezu alle der hier besprochenen Analyseverfahren einer Familie von Modellen angehören, die als *generalisierte lineare Modelle* bezeichnet wird.

Wenngleich moderne Lehrbücher der Ökonometrie auch didaktisch auf hohem Niveau sind (Cameron und Trivedi 2005; Cameron und Trivedi 2009; Greene 2000; Wooldridge 2005), ist die Art der Darstellung manchen Studierenden und Forschenden außerhalb der Ökonomie und der Statistik schwer zugänglich, weil die formale Darstellung die verdichtete Matrixschreibweise nutzt. Darauf wird in diesem einführenden Lehrbuch weitgehend verzichtet. Das anwendungsorientierte Buch soll eine Lücke schließen zwischen den modernen mikroökonomischen Verfahren und den Lehrinhalten guter BA-Statistikurse in der Soziologie.

Neben der Vermittlung von Kompetenzen, die immer das primäre Ziel eines Lehrbuches darstellen soll, wird durch den hoffentlich nachvollziehbaren Umgang mit Formeln und durch die möglichst anschaulichen Beispiele versucht, in die Materie einsteigenden Studierenden und Forschenden Begeisterung die Möglichkeiten dieser Methoden nahezulegen. Eine Idealvorstellung für ein Einführungsbuch in fortgeschrittene Methoden besteht darin, dass auch erfahrene Forscherinnen und Forscher während der praktischen Arbeit davon profitieren.

Ein Lehrbuch über formale Methoden setzt sich in besonderem Maße dem Risiko fehlerhafter Darstellungen aus. Gewiss können auch inhaltlich-substanzielle sozialwissenschaftliche Darstellungen in Lehrbüchern richtig oder falsch sein, aber das Wesen der formalen Methoden besteht gerade darin, möglichst präzise zwischen richtig oder falsch unterscheiden zu können. Beim Autor ruft diese Tatsache natürlich Unbehagen hervor, zumal Rezensionen von Lehrbüchern durchaus streng sein können.

Eine kontinuierlich aktualisierte Liste mit Korrekturen wird für dieses Buch erstellt und ist dem begleitenden Datenpaket beigelegt. Dieses Paket ist online abrufbar unter:

<http://www.barkhof.uni-bremen.de/~mwindzio/ereignisse.zip>

Für alle in dem Buch besprochenen Analyseverfahren werden praktische Anwendungen durchgeführt. Die Analysen basieren auf der mittlerweile sehr prominenten Software *Stata*. Grundkenntnisse in *Stata* sind für ein Verständnis der praktischen Übungen notwendig, wenngleich die präsentierte Syntax fast immer sehr einfach strukturiert ist. Die in den Übungen verwendeten Daten können zwar auch mit der *Stata* Version 9 eingelesen werden, aber nicht alle (jedoch die weit überwiegende Zahl) der Modelle lassen sich mit dieser Version schätzen. Empfohlen werden Versionen ab *Stata* 10. Im Text wird darauf hingewiesen, welche Zusatzmodule (*ado-files*) jeweils zu installieren sind. Aufgrund unterschiedlicher Versionen kann es hier zu Komplikationen kommen. Mit meiner 12er Version von *Stata* läuft alles problemlos, eine Aktualisierung der *ados* könnte bei Fehlermeldungen helfen. Sollten ältere Versionen der *ados* in *Stata* 12 partout nicht funktionieren – z.B. *mlogtest* oder *prgen* –, sollte jeweils *vor* dem letzten Schätzmodell mit dem Befehl `version 9` auf die ältere *Stata* Version umgestellt werden (anschließend zurück mit `version 12`).

Vor jeder Übungssitzung muss ein sogenanntes globales Makro `pfad` definiert werden, damit man ohne viel Schreibarbeit die jeweils zu verwendenden Datensätze über das Internet einlesen kann. Dies geschieht mit dem folgenden Befehl:

```
global pfad http://www.barkhof.uni-bremen.de/~mwindzio/
```

Ist das Makro einmal definiert, kann man z.B. mit dem Befehl `use $pfad/logit.dta, clear` den Analysedatensatz für die binäre logistische Regression einlesen. Ist der PC nicht mit dem Internet verbunden, sollten alle in dem Archiv *ereignisse.zip* enthaltenen Dateien nach `c:\temp\` kopiert werden. Anschließend ist das globale Makro folgendermaßen zu definieren:

global pfad c:\temp\

Die im Buch dargestellte Stata Syntax ist als do-file `models.do` in `ereignisse.zip` enthalten, so dass man die Syntax nicht eigenhändig schreiben muss. Ob „copy & paste“ empfehlenswert ist, oder ob das eigenhändige Abschreiben der Befehle nicht doch eine didaktische Funktion erfüllt, bleibt der Leserin und dem Leser überlassen.

Großer Dank gebührt Reinhold Sackmann für seine schier unerschöpfliche Geduld, die er mit mir als säumigen Autor hatte. Gleiches gilt für den VS-Verlag. Sarah Oldenburg las den Text Korrektur und leistete zudem großartige Unterstützung bei der Transkription des Manuskriptes. Sehr wertvolle inhaltliche Kommentare zu vorläufigen Fassungen des Buches und wichtige Korrekturhinweise kamen von Timm Fulge, Katharina Groß, Michael Hanslmeier, Reinhold Sackmann und insbesondere von Alexander Gattig. Ausgewählte Kapitel kommentierten Johannes Huinink, Thorsten Schneider, Janna Teltemann und Maximilian Trommer. Ihnen allen gebührt großer Dank. Sie alle trugen maßgeblich zur Verbesserung des Buches bei. Selbstverständlich sind verbleibende Mängel allein von mir zu verantworten. Zu danken habe ich auch den Studierenden des Masterprogramms *Soziologie und Sozialforschung* an der Universität Bremen, die in der Vorlesung das eine oder andere didaktische Experiment über sich ergehen lassen mussten. Hoffentlich geschah dies überwiegend unauffällig und ohne irreparable Folgeschäden.

Bei der Arbeit an diesem Buch musste ich zwei Dinge einsehen: *Erstens* sollte man sich niemals auf einen Abgabetermin einlassen, wenn das Buch nicht bereits in weiten Teilen geschrieben ist. *Zweitens* bleibt aufgrund der aufgescheuchten Stimmung in einer zertifiziert exzellenten, zugleich aber chronisch unterfinanzierten Universität kaum Zeit für das Verfassen von Lehrbüchern. Die Tatsache, dass dieses Buch dennoch fertig wurde, ist der Inspiration durch einige wunderbare Lehrbücher und Aufsätze der angewandten Statistik und Ökonometrie zu verdanken. Schön wäre es, wenn zumindest ein Teil dieser Inspiration auch in dieses Buch hinüber gerettet werden konnte.

Bremen, im Mai 2013

Michael Windzio

Regressionsmodelle für Zustände und Ereignisse

Eine Einführung

Windzio, M.

2013, X, 282 S. 46 Abb., Softcover

ISBN: 978-3-531-15554-8