

Inhalt

Vorwort — V

Zu diesem Buch — 1

1 „Das erste Mal“ — 9

- 1.1 Aufruf von Stata — 9**
- 1.2 Gestalten der Bildschirmansicht — 10**
- 1.3 Erste Analysen — 11**
- 1.4 Do-Files — 29**
- 1.5 Stata verlassen — 31**
- 1.6 Übungen — 32**

2 Arbeiten mit Do-Files — 34

- 2.1 Von der interaktiven Arbeit zum Do-File — 34**
- 2.2 Do-Files sinnvoll gestalten — 40**
 - 2.2.1 Kommentare — 41**
 - 2.2.2 Zeilenwechsel — 41**
 - 2.2.3 Befehle, die in keinem Do-File fehlen sollten — 43**
- 2.3 Arbeitsorganisation — 46**
- 2.4 Übungen — 51**

3 Die Stata-Grammatik — 52

- 3.1 Elemente der Stata-Kommandos — 52**
 - 3.1.1 Der Befehl — 52**
 - 3.1.2 Die Variablenliste — 54**
 - 3.1.3 Optionen — 57**
 - 3.1.4 Die in-Bedingung — 58**
 - 3.1.5 Die if-Bedingung — 60**
 - 3.1.6 Ausdrücke — 62**
 - 3.1.7 Die Nummernliste — 68**
 - 3.1.8 Dateinamen — 69**
- 3.2 Wiederholung ähnlicher Befehle — 70**
 - 3.2.1 Das by-Präfix — 71**
 - 3.2.2 Die foreach-Schleife — 72**
 - 3.2.3 Die forvalues-Schleife — 76**
- 3.3 Die Gewichtungsanweisung — 77**
- 3.4 Übungen — 82**

| | |
|----------|--|
| 4 | Eine allgemeine Bemerkung zu den Statistik-Kommandos — 84 |
| 4.1 | Herkömmliche Statistikbefehle — 84 |
| 4.2 | Modellbefehle — 87 |
| 4.3 | Übungen — 89 |
| 5 | Erstellen und Verändern von Variablen — 91 |
| 5.1 | Die Befehle generate und replace — 92 |
| 5.1.1 | Variablennamen — 93 |
| 5.1.2 | Einige Beispiele — 94 |
| 5.1.3 | Nützliche Funktionen — 98 |
| 5.2 | Missings zuweisen und aufheben — 101 |
| 5.3 | Beschriftung von Variablen — 105 |
| 5.4 | Spezielle Recodierungs-Befehle — 108 |
| 5.4.1 | recode — 108 |
| 5.4.2 | egen — 109 |
| 5.5 | Recodieren für Fortgeschrittene — 111 |
| 5.5.1 | Recodieren mit by, _n und _N — 112 |
| 5.5.2 | Explizite Subscripte — 114 |
| 5.6 | Recodieren von String-Variablen — 116 |
| 5.7 | Recodierung von Datums- und Zeitangaben — 121 |
| 5.7.1 | Datumsangaben — 121 |
| 5.7.2 | Zeit — 126 |
| 5.8 | Storage-Types oder: der Geist in der Maschine — 128 |
| 5.9 | Übungen — 130 |
| 6 | Erstellen und Verändern von Grafiken — 131 |
| 6.1 | Eine Vorbemerkung zur Syntax — 131 |
| 6.2 | Typen von Grafiken — 132 |
| 6.2.1 | Beispiele — 133 |
| 6.2.2 | Spezielle Grafiken — 133 |
| 6.3 | Elemente der Grafiken — 135 |
| 6.3.1 | Erscheinungsbild der Daten — 137 |
| 6.3.2 | Grafik- und Plotregion — 146 |
| 6.3.3 | Informationen innerhalb der Plotregion — 149 |
| 6.3.4 | Informationen außerhalb der Plotregion — 154 |
| 6.4 | Multiple Grafiken — 161 |
| 6.4.1 | Überlagerung mehrerer twoway-Grafiken — 162 |
| 6.4.2 | Befehlsoption by() — 163 |
| 6.4.3 | Zusammenführung von Grafiken — 164 |
| 6.5 | Speichern und Drucken von Grafiken — 166 |
| 6.6 | Übungen — 169 |

| | |
|----------|--|
| 7 | Die Beschreibung von Verteilungen — 170 |
| 7.1 | Wenige oder viele Ausprägungen? — 171 |
| 7.2 | Variablen mit wenigen Ausprägungen — 172 |
| 7.2.1 | Tabellarische Darstellungen — 172 |
| 7.2.2 | Grafische Verfahren — 177 |
| 7.3 | Variablen mit vielen Ausprägungen — 183 |
| 7.3.1 | Häufigkeitsverteilung gruppiertter Daten — 184 |
| 7.3.2 | Beschreibung durch Maßzahlen — 187 |
| 7.3.3 | Grafische Verfahren — 198 |
| 7.4 | Übungen — 210 |
| 8 | Grundlagen statistischer Inferenz — 211 |
| 8.1 | Zufallsstichproben und Stichprobenverteilungen — 212 |
| 8.1.1 | Erzeugung von Zufallszahlen — 212 |
| 8.1.2 | Erzeugung fiktiver Datensätze — 213 |
| 8.1.3 | Ziehung von Stichproben — 217 |
| 8.1.4 | Die Stichprobenverteilung — 219 |
| 8.2 | Deskriptive Inferenz — 223 |
| 8.2.1 | Standardfehler für einfache Zufallsstichproben — 224 |
| 8.2.2 | Standardfehler für komplexe Stichproben — 225 |
| 8.2.3 | Standardfehler bei fehlenden Daten — 233 |
| 8.2.4 | Verwendungen für Standardfehler — 242 |
| 8.3 | Kausale Inferenz — 254 |
| 8.3.1 | Grundlegende Konzepte — 254 |
| 8.3.2 | Der Effekt der dritten Klasse — 259 |
| 8.3.3 | Einige Probleme der kausalen Inferenz — 261 |
| 8.4 | Übungen — 263 |
| 9 | Einführung in die Regressionstechnik — 265 |
| 9.1 | Lineare Einfachregression — 268 |
| 9.1.1 | Das Grundprinzip — 268 |
| 9.1.2 | Lineare Regression mit Stata — 272 |
| 9.2 | Die multiple Regression — 282 |
| 9.2.1 | Multiple lineare Regression mit Stata — 283 |
| 9.2.2 | Spezielle Kennzahlen der multiplen Regression — 286 |
| 9.2.3 | Was bedeutet eigentlich „unter Kontrolle“? — 288 |
| 9.3 | Regressionsdiagnostik — 290 |
| 9.3.1 | Die Verletzung von $E(\epsilon_i) = 0$ — 291 |
| 9.3.2 | Heteroskedastizität — 307 |
| 9.3.3 | Autokorrelation — 309 |
| 9.4 | Verfeinerte Modelle — 310 |
| 9.4.1 | Kategoriale unabhängige Variablen — 311 |

| | | |
|-----------|---|------------|
| 9.4.2 | Interaktionseffekte — | 314 |
| 9.4.3 | Regressionsmodelle mit transformierten Daten — | 319 |
| 9.5 | Darstellung von Regressionsergebnissen — | 324 |
| 9.5.1 | Tabellen ähnlicher Regressionsmodelle — | 324 |
| 9.5.2 | Koeffizienten-Plots — | 327 |
| 9.5.3 | Conditional-Effects-Plots — | 332 |
| 9.6 | Weiterführende Verfahren — | 335 |
| 9.6.1 | Median-Regression — | 335 |
| 9.6.2 | Regressionsmodelle für Paneldaten — | 337 |
| 9.7 | Übungen — | 348 |
| 10 | Regressionsmodelle für kategoriale abhängige Variablen — | 350 |
| 10.1 | Das lineare Wahrscheinlichkeitsmodell — | 351 |
| 10.2 | Grundkonzepte — | 355 |
| 10.2.1 | Odds, Log-Odds und Odds-Ratios — | 355 |
| 10.2.2 | Exkurs: Das Maximum-Likelihood-Prinzip — | 360 |
| 10.3 | Logistische Regression mit Stata — | 364 |
| 10.3.1 | Der Koeffizientenblock — | 366 |
| 10.3.2 | Der Iterationsblock — | 372 |
| 10.3.3 | Der Modellfit-Block — | 373 |
| 10.4 | Diagnostik der logistischen Regression — | 379 |
| 10.4.1 | Linearität — | 379 |
| 10.4.2 | Einflussreiche Fälle — | 383 |
| 10.5 | Likelihood-Ratio-Test — | 387 |
| 10.6 | Verfeinerte Modelle — | 389 |
| 10.7 | Weiterführende Verfahren — | 394 |
| 10.7.1 | Probit-Modelle — | 395 |
| 10.7.2 | Multinomiale logistische Regression — | 397 |
| 10.7.3 | Ordinale Logit-Modelle — | 401 |
| 10.8 | Übungen — | 404 |
| 11 | Daten lesen und schreiben — | 406 |
| 11.1 | Das Ziel: Die Datenmatrix — | 406 |
| 11.2 | Import maschinenlesbarer Daten — | 408 |
| 11.2.1 | Einlesen von System-Files anderer Programme — | 409 |
| 11.2.2 | Einlesen von Textdateien — | 413 |
| 11.3 | Dateneingabe — | 421 |
| 11.3.1 | Dateneingabe über den Editor — | 422 |
| 11.3.2 | Der input-Befehl — | 423 |
| 11.4 | Zusammenführung von Datensätzen — | 427 |
| 11.4.1 | Die Datenstruktur des GSOEP — | 428 |
| 11.4.2 | Der Befehl merge — | 430 |

| | | |
|-----------|---|------------|
| 11.4.3 | Der Befehl append — | 441 |
| 11.5 | Datensätze speichern und exportieren — | 444 |
| 11.6 | Zum Umgang mit großen Datensätze — | 445 |
| 11.6.1 | Regeln zum Umgang mit dem Arbeitsspeicher — | 445 |
| 11.6.2 | Die Verwendung zu großer Datensätze — | 446 |
| 11.7 | Unicode — | 447 |
| 11.7.1 | Datencodierungen — | 448 |
| 11.7.2 | Kompatibilitätsprobleme — | 451 |
| 11.7.3 | Babylon — | 455 |
| 11.8 | Übungen — | 459 |
| 12 | Do-Files für Fortgeschrittene und eigene Programme — | 460 |
| 12.1 | Zwei Anwendungsbeispiele — | 460 |
| 12.2 | Vier Programmierwerkzeuge — | 462 |
| 12.2.1 | Makros — | 462 |
| 12.2.2 | Do-Files — | 466 |
| 12.2.3 | Programme — | 466 |
| 12.2.4 | Ado-Files — | 469 |
| 12.3 | Selbst programmierte Stata-Befehle — | 473 |
| 12.3.1 | Konzept der Syntax — | 475 |
| 12.3.2 | Erstellen eines ersten Ado-Files — | 475 |
| 12.3.3 | Weitergabe von Variablenlisten — | 477 |
| 12.3.4 | Weitergabe von Optionen — | 478 |
| 12.3.5 | Weitergabe von if und in — | 480 |
| 12.3.6 | Bilden von Variablen unbekannter Anzahl — | 481 |
| 12.3.7 | Voreinstellungen — | 484 |
| 12.3.8 | Erweiterte Makrofunktionen — | 486 |
| 12.3.9 | Veränderungen am Datensatz vermeiden — | 488 |
| 12.3.10 | Help-Files — | 489 |
| 12.4 | Übungen — | 491 |
| 13 | Rund um Stata — | 492 |
| 13.1 | Ressourcen mit Informationen — | 492 |
| 13.2 | Pflege von Stata — | 493 |
| 13.3 | Zusätzliche Prozeduren — | 494 |
| 13.3.1 | Stata Journal-Ados — | 494 |
| 13.3.2 | SSC-Ados — | 497 |
| 13.3.3 | Andere Ados — | 497 |
| 13.4 | Bezugsquellen — | 499 |

XII — Inhalt

Literatur — 501

Stichwortverzeichnis — 505