

# Inhaltsverzeichnis

|                                                                      |           |
|----------------------------------------------------------------------|-----------|
| <b>Vorwort.....</b>                                                  | <b>X</b>  |
| <b>1. Grundlegende SAS-Befehle und Datenhandling .....</b>           | <b>1</b>  |
| 1.1. Eine Vorbemerkung zum Programmpaket SAS .....                   | 1         |
| 1.2. SAS-Fenster .....                                               | 1         |
| 1.2.1. Der Editor .....                                              | 1         |
| 1.2.2. Speichern des Inhalts des Editors.....                        | 3         |
| 1.3. Datenerfassung, Datenimport und Datenexport.....                | 3         |
| 1.3.1. Datenerfassung .....                                          | 3         |
| 1.3.2. Datenimport.....                                              | 3         |
| 1.3.3. Datenexport .....                                             | 4         |
| 1.4. Die Library.....                                                | 4         |
| 1.4.1. Anlegen einer Library .....                                   | 4         |
| 1.4.2. Arbeiten mit einer Library.....                               | 4         |
| 1.5. Die wichtigsten Data Steps zur Bearbeitung von Datenfiles ..... | 5         |
| 1.5.1. Kopieren eines Datenfiles .....                               | 5         |
| 1.5.2. Einlesen ausgewählter Variablen .....                         | 6         |
| 1.5.3. Beibehalten bestimmter Beobachtungen einer Variable .....     | 6         |
| 1.5.4. Eliminieren ausgewählter Variablen .....                      | 6         |
| 1.5.5. Eliminieren von Beobachtungen mittels if.....                 | 6         |
| 1.5.6. Eliminieren von Beobachtungen mittels where.....              | 6         |
| 1.5.7. Dateien sortieren.....                                        | 7         |
| 1.5.8. Dateien aufspalten nach Beobachtungen.....                    | 7         |
| 1.5.9. Dateien aufspalten nach Variablen.....                        | 8         |
| 1.5.10. Vertikales Zusammenführen von Dateien.....                   | 8         |
| 1.5.11. Horizontales Zusammenführen (mergen) von Dateien .....       | 8         |
| 1.5.12. Vergabe von Labels.....                                      | 8         |
| 1.5.13. Erzeugen von neuen Variablen.....                            | 9         |
| 1.6. Die Verwendung von do-Schleifen.....                            | 10        |
| 1.7. Die Erzeugung von Zufallszahlen.....                            | 11        |
| 1.8. Übungsaufgaben .....                                            | 12        |
| <b>2. Datenbeschreibung und graphische Darstellung.....</b>          | <b>13</b> |
| 2.1. Die Analyse von Häufigkeitsverteilungen mit SAS .....           | 13        |
| 2.1.1. Datenanalyse mit der Prozedur Means.....                      | 13        |
| 2.1.2. Datenanalyse mit der Prozedur Univariate.....                 | 18        |
| 2.1.2.1. Der Vorzeichentest .....                                    | 19        |
| 2.1.2.2. Der Vorzeichenrangtest nach Wilcoxon .....                  | 19        |
| 2.1.2.3. Der Shapiro-Wilk-Test .....                                 | 22        |
| 2.1.2.4. Der Kolmogorov-Smirnow-Test.....                            | 24        |
| 2.2. Graphische Darstellungen mit SAS .....                          | 25        |
| 2.2.1. Die Prozedur Gchart.....                                      | 25        |
| 2.2.2. Die Prozedur Gplot.....                                       | 28        |
| 2.3. Übungsaufgaben .....                                            | 29        |

|                                                                                                |           |
|------------------------------------------------------------------------------------------------|-----------|
| <b>3. Schätzen und Testen.....</b>                                                             | <b>31</b> |
| 3.1. Wichtige Wahrscheinlichkeitsverteilungen .....                                            | 31        |
| 3.1.1. Die Hypergeometrische Verteilung.....                                                   | 31        |
| 3.1.2. Die Binomialverteilung .....                                                            | 32        |
| 3.1.3. Die Normalverteilung.....                                                               | 32        |
| 3.1.4. Veranschaulichung der Normalverteilung mit SAS.....                                     | 33        |
| 3.1.5. Der zentrale Grenzwertsatz .....                                                        | 35        |
| 3.2. Grundidee von Schätzen und Testen.....                                                    | 37        |
| 3.2.1. Schätzen im Ein-Stichprobenfall.....                                                    | 37        |
| 3.2.1.1. Schätzen im heterograden Fall.....                                                    | 37        |
| 3.2.1.2. Schätzen mit SAS: heterograder Fall.....                                              | 39        |
| 3.2.1.3. Schätzen im homograden Fall.....                                                      | 39        |
| 3.2.1.4. Schätzen mit SAS: homograder Fall.....                                                | 40        |
| 3.2.2. Testen im Ein-Stichprobenfall.....                                                      | 41        |
| 3.2.2.1. Testen im heterograden Fall .....                                                     | 41        |
| 3.2.2.2. Testen im homograden Fall .....                                                       | 42        |
| 3.2.3. Mittelwertdifferenz im Zwei-Stichprobenfall .....                                       | 42        |
| 3.2.3.1. Die Schätz- und Testprozedur.....                                                     | 42        |
| 3.2.3.2. Der Test auf Mittelwertgleichheit mit SAS .....                                       | 45        |
| 3.2.3.2.1. Die Prozedur Ttest.....                                                             | 45        |
| 3.2.3.2.2. Der Test auf Mittelwertgleichheit mit Data Steps.....                               | 46        |
| 3.3. Übungsaufgaben .....                                                                      | 47        |
| <b>4. Eine Simulation mit SAS .....</b>                                                        | <b>49</b> |
| 4.1. Erwartungswert und Standardabweichung der Kosten .....                                    | 49        |
| 4.1.1. Methodische Erläuterungen.....                                                          | 49        |
| 4.1.2. Die Berechnung mit SAS .....                                                            | 51        |
| 4.2. Eine Simulation zur Ermittlung der Wahrscheinlichkeitsverteilung<br>der Gesamtkosten..... | 53        |
| 4.2.1. Die grundlegende Idee der Simulation.....                                               | 53        |
| 4.2.2. Die Durchführung mit SAS.....                                                           | 53        |
| 4.3. Übungsaufgaben .....                                                                      | 57        |
| <b>5. Häufigkeitstabellen und die Prozedur Frequencies .....</b>                               | <b>58</b> |
| 5.1. Die allgemeine Befehlssyntax.....                                                         | 58        |
| 5.1.1. Der Befehl tables .....                                                                 | 58        |
| 5.1.2. Das by-statement .....                                                                  | 60        |
| 5.1.3. Das weight-statement .....                                                              | 60        |
| 5.1.4. Die Ausgabe von Datenfiles.....                                                         | 60        |
| 5.1.5. Weitere Optionen der Tabellenausgabe .....                                              | 61        |
| 5.2. Die statistischen Assoziationsmaße .....                                                  | 61        |
| 5.2.1. Die Notation für Vierfeldertafeln und $R \times C$ -Feldertafeln.....                   | 61        |
| 5.2.2. Der Chi-Quadrat-Test.....                                                               | 62        |
| 5.2.3. Das kontinuieritätsadjustierte Chi-Quadrat .....                                        | 63        |
| 5.2.4. Der exakte Fisher-Test .....                                                            | 64        |

|           |                                                                         |            |
|-----------|-------------------------------------------------------------------------|------------|
| 5.2.5.    | Der Phi-Koeffizient .....                                               | 67         |
| 5.2.6.    | Der Kontingenzkoeffizient P .....                                       | 68         |
| 5.2.7.    | Cramer's V .....                                                        | 68         |
| 5.2.8.    | Der Rangkorrelationskoeffizient nach Spearman .....                     | 68         |
| 5.2.9.    | Kendall's $\tau$ .....                                                  | 71         |
| 5.3.      | Beispiele für Zusammenhangsanalysen mit SAS .....                       | 72         |
| 5.4.      | Übungsaufgaben .....                                                    | 74         |
| <b>6.</b> | <b>Varianzanalyse mit der Prozedur Anova .....</b>                      | <b>76</b>  |
| 6.1.      | Die einfache Varianzanalyse .....                                       | 76         |
| 6.1.1.    | Die Modellierung .....                                                  | 76         |
| 6.1.2.    | Signifikanzbeurteilung der Effekte .....                                | 77         |
| 6.1.3.    | Simultaner Mittelwertvergleich nach Scheffe .....                       | 81         |
| 6.2.      | Die mehrfaktorielle Varianzanalyse .....                                | 82         |
| 6.2.1.    | Signifikanzbeurteilung bei der zweifaktoriellen<br>Varianzanalyse ..... | 84         |
| 6.3.      | Die Prozedur Anova .....                                                | 87         |
| 6.3.1.    | Das model-statement .....                                               | 88         |
| 6.3.2.    | Das means-statement .....                                               | 88         |
| 6.3.3.    | Die SAS-Ausgabe .....                                                   | 88         |
| 6.4.      | Übungsaufgaben .....                                                    | 91         |
| <b>7.</b> | <b>Die einfache lineare Regressionsanalyse .....</b>                    | <b>92</b>  |
| 7.1.      | Die einfache lineare Regression als deskriptives Instrument .....       | 92         |
| 7.2.      | Das stochastische Regressionsmodell .....                               | 96         |
| 7.2.1.    | Die Annahmen des stochastischen Regressionsmodells .....                | 96         |
| 7.2.2.    | Signifikanzbeurteilung der geschätzten Parameter .....                  | 97         |
| 7.2.3.    | Der F-Test zur Signifikanzbeurteilung der Regression .....              | 99         |
| 7.3.      | Die lineare Einfachregression mit Data Steps .....                      | 101        |
| 7.4.      | Die Prozedur Reg .....                                                  | 104        |
| 7.5.      | Übungsaufgaben .....                                                    | 105        |
| <b>8.</b> | <b>Einführung in die Matrixsprache IML .....</b>                        | <b>107</b> |
| 8.1.      | Einfache Matrixoperationen in IML .....                                 | 107        |
| 8.2.      | Ein Beispiel mit weiteren Matrixbefehlen .....                          | 107        |
| 8.3.      | Die Verwendung von Modulen und Funktionsmodulen .....                   | 109        |
| 8.3.1.    | Module .....                                                            | 109        |
| 8.3.2.    | Funktionsmodule .....                                                   | 110        |
| 8.4.      | Das multiple Regressionsmodell in Matrixnotation .....                  | 111        |
| 8.5.      | Die Berechnung einer linearen Regression in IML .....                   | 112        |
| 8.6.      | Eine Monte Carlo-Simulation mit IML .....                               | 113        |
| 8.7.      | Die wichtigsten Befehle in IML .....                                    | 116        |
| 8.7.1.    | Matrixoperationen .....                                                 | 116        |
| 8.7.2.    | Wichtige Matrixfunktionen in IML .....                                  | 117        |
| 8.7.3.    | Wichtige skalare Funktionen .....                                       | 117        |
| 8.8.      | Übungsaufgaben .....                                                    | 118        |

|                                                                        |            |
|------------------------------------------------------------------------|------------|
| <b>9. Regressionsanalyse: Lags, Dummies und Modellselektion .....</b>  | <b>119</b> |
| 9.1. Ein Exkurs: Zur Theorie der Phillips-Kurve.....                   | 119        |
| 9.2. Die Verwendung von verzögerten exogenen Variablen (Lags).....     | 120        |
| 9.3. Die Verwendung von Dummyvariablen .....                           | 122        |
| 9.4. Die Erzeugung von Ergebnisfiles, Outputfiles und Graphiken .....  | 124        |
| 9.4.1. Die Erzeugung von Ergebnisfiles.....                            | 124        |
| 9.4.2. Die Erzeugung von Outputfiles mit dem output-statement.....     | 125        |
| 9.4.3. Die Erzeugung von Graphiken mit dem plot-statement .....        | 126        |
| 9.5. Verfahren der automatisierten Modellselektion.....                | 126        |
| 9.5.1. Forward .....                                                   | 127        |
| 9.5.2. Backward.....                                                   | 127        |
| 9.5.3. Stepwise.....                                                   | 127        |
| 9.5.4. Maximum $R^2$ Improvement (maxr).....                           | 127        |
| 9.5.5. Minimum $R^2$ Improvement (minr) .....                          | 128        |
| 9.5.6. $R^2$ Selection (rsquare).....                                  | 128        |
| 9.5.7. Adjusted $R^2$ Selection (adjrsq) .....                         | 128        |
| 9.5.8. Mallows' $C_p$ Selection (cp) .....                             | 128        |
| 9.5.9. Modellselektion mit SAS an dem Beispiel der Phillips-Kurve..... | 129        |
| 9.5.9.1. Eine Modellselektion mit Forward .....                        | 129        |
| 9.5.9.2. Eine Modellselektion mit Adjusted $R^2$ .....                 | 130        |
| 9.6. Übungsaufgaben .....                                              | 131        |
| <b>10. Regressionsanalyse und Verletzung von Modellannahmen .....</b>  | <b>133</b> |
| 10.1. Partielle Regressionskoeffizienten.....                          | 133        |
| 10.2. Partielle Bestimmtheitsmaße .....                                | 135        |
| 10.3. Konfidenzintervalle um die Regressionslinie .....                | 138        |
| 10.3.1. Das Konfidenzintervall für Regressionswerte.....               | 138        |
| 10.3.2. Das Konfidenzintervall für individuelle Werte .....            | 139        |
| 10.3.3. Die Berechnung von Konfidenzintervallen mit SAS .....          | 139        |
| 10.4. Verletzungen der Modellannahmen.....                             | 140        |
| 10.4.1. Multikollinearität.....                                        | 140        |
| 10.4.2. Autokorrelation .....                                          | 142        |
| 10.4.3. Heteroskedastie .....                                          | 143        |
| 10.4.3.1. Der Goldfeld-Quandt-Test.....                                | 143        |
| 10.4.3.2. Der White-Test .....                                         | 144        |
| 10.4.4. Überprüfung der Modellannahmen mit SAS.....                    | 145        |
| 10.4.4.1. Die Prüfung auf Kollinearität .....                          | 145        |
| 10.4.4.2. Die Prüfung auf Autokorrelation .....                        | 147        |
| 10.4.4.3. Die Prüfung auf Heteroskedastie .....                        | 147        |
| 10.5. Übungsaufgaben .....                                             | 148        |
| <b>11. Nichtlineare Regression .....</b>                               | <b>149</b> |
| 11.1. Das allgemeine Schema der iterativen Minimierung.....            | 149        |
| 11.1.1. Die Steepest Gradient-Methode .....                            | 150        |
| 11.1.2. Die Newton-Raphson-Methode.....                                | 150        |
| 11.1.3. Die Gauss-Newton-Methode.....                                  | 151        |
| 11.1.4. Der Gauss-Newton Algorithmus: Ein einfaches Beispiel .....     | 152        |

|                                                                                              |            |
|----------------------------------------------------------------------------------------------|------------|
| 11.2. Die Prozedur Nlin .....                                                                | 158        |
| 11.2.1. Die Befehlssyntax.....                                                               | 158        |
| 11.2.2. Ein Beispiel .....                                                                   | 160        |
| 11.3. Übungsaufgaben .....                                                                   | 162        |
| <b>12. Regression mit kategorialen endogenen Variablen .....</b>                             | <b>164</b> |
| 12.1. Das lineare Wahrscheinlichkeitsmodell.....                                             | 164        |
| 12.1.1. Die Schätzung mit SAS.....                                                           | 166        |
| 12.2. Das Logit-Modell.....                                                                  | 168        |
| 12.2.1. Die Maximum Likelihood-Schätzung .....                                               | 170        |
| 12.2.1.1. Maßzahlen der Anpassungsgüte .....                                                 | 172        |
| 12.2.1.2. Ein kleines Zahlenbeispiel.....                                                    | 172        |
| 12.2.2. Die Schätzung mit SAS.....                                                           | 175        |
| 12.2.2.1. Die Prozedur Logistic .....                                                        | 175        |
| 12.2.2.2. Die logistische Regression mit der Prozedur Probit .....                           | 177        |
| 12.3. Das Probit-Modell.....                                                                 | 178        |
| 12.4. Ein Vergleich der Ergebnisse der drei Modelle.....                                     | 180        |
| 12.5. Übungsaufgaben .....                                                                   | 181        |
| <b>13. Gemischte Modelle und die Prozedur Mixed.....</b>                                     | <b>183</b> |
| 13.1. Einführung in die Methoden der Panelanalyse .....                                      | 183        |
| 13.1.1. Längs- und Querschnittsregressionen versus Pooled<br>Regression .....                | 183        |
| 13.1.2. Die Längsschnittsregression für einzelne Einheiten .....                             | 184        |
| 13.1.3. Die Querschnittsregression für einzelne Jahre .....                                  | 185        |
| 13.1.4. Die einfache Pooled Regression.....                                                  | 185        |
| 13.1.5. Parametersignifikanzen und partielle Bestimmtheitsmaße .....                         | 186        |
| 13.1.6. Verallgemeinerte Kleinste Quadrate Methode und Maximum<br>Likelihood-Schätzung ..... | 187        |
| 13.2. Der Zusammenhang von offenen Stellen und Arbeitslosen .....                            | 188        |
| 13.2.1. Die empirische makroökonomische Entwicklung.....                                     | 188        |
| 13.2.2. Theoretische Überlegungen zur Beveridge-Kurve.....                                   | 188        |
| 13.3. Die Prozedur Mixed .....                                                               | 189        |
| 13.3.1. Pooled Regression .....                                                              | 191        |
| 13.3.2. Die Darstellung als Fehlerkomponentenmodell .....                                    | 193        |
| 13.3.3. Fixed Effects-Modelle.....                                                           | 194        |
| 13.3.3.1. Fixe Effekte der Einheiten .....                                                   | 194        |
| 13.3.3.2. Fixe Effekte der Jahre .....                                                       | 194        |
| 13.3.4. Modelle mit nichttrivialer Varianz-Kovarianz-Matrix.....                             | 195        |
| 13.3.4.1. Zufällige Effekte der Einheiten.....                                               | 195        |
| 13.3.4.2. Ein Modell autokorrelierter Residuen.....                                          | 197        |
| 13.3.4.3. Random Coefficient-Modelle .....                                                   | 197        |
| 13.3.4.4. Eine Beispielschätzung .....                                                       | 198        |
| 13.4. Übungsaufgaben .....                                                                   | 200        |
| <b>Literaturverzeichnis.....</b>                                                             | <b>221</b> |
| <b>Sachwortverzeichnis .....</b>                                                             | <b>225</b> |