

LEHRSTUHL FÜR STATISTIK UND ÖKONOMETRIE  
ÜBUNG ZUR DATENANALYSE

**Aufgabenserie 3: Lineares Regressionsmodell**

**Aufgabe 6**

Das Jahresgehalt des CEO ( $Y$ ) eines deutschen Aktienunternehmens soll durch die Variablen  $X_1 :=$  "wirtschaftlicher Erfolg des Unternehmens, gemessen durch die Eigenkapitalrendite (return on equity, ROE)" und  $X_2 :=$  "Umsatz" erklärt werden. Zudem wird ein Einfluss des Wirtschaftssektors, in dem das Unternehmen agiert, vermutet, wobei nur zwischen dem primären/sekundären Sektor einerseits und dem tertiären Sektor andererseits unterschieden werden soll. Für  $n = 6$  Unternehmen liegen Ihnen folgende Beobachtungen der endogenen und exogenen Variablen vor:

Unternehmen	CEO-Gehalt (in Mio. EUR)	ROE	Umsatz (in Mrd. EUR)	Sektor
1	5.1	0.27	71	primär/sekundär
2	6.6	0.20	67	primär/sekundär
3	4.5	0.09	36	tertiär
4	1.3	0.10	5	primär/sekundär
5	0.5	0.02	18	tertiär
6	8.0	0.13	103	primär/sekundär

1. Gehen Sie zunächst von einem linearen Modell  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$  aus, in dem das CEO-Jahresgehalt einzig durch den Wirtschaftssektor erklärt wird. Zeigen Sie, dass ein Modell, welches neben dem Achsenabschnitt auch Dummyvariablen für beide Ausprägungen der qualitativen Variable enthält, nicht geschätzt werden kann.
2. Formulieren Sie zwei schätzbare Varianten des linearen Modells, welches das CEO-Gehalt durch Eigenkapitalrendite, Umsatz und Wirtschaftssektor erklärt. Stellen Sie jeweils die Datenmatrix  $\mathbf{X}$  auf.
3. Zeigen Sie, dass sich für die beiden in der letzten Teilaufgabe formulierten Varianten des linearen Regressionsmodells aus dem Ansatz der OLS-Schätzung jeweils eine Residuensumme von null ergibt.
4. Bestimmen Sie für das in Teilaufgabe 2 formulierte Modell mit Achsenabschnitt die Matrizen  $\mathbf{X}'\mathbf{X}$  und  $(\mathbf{X}'\mathbf{X})^{-1}$  sowie den Wert des OLS-Schätzvektors  $\hat{\boldsymbol{\beta}}$ .
5. Schätzen Sie ausgehend von den Annahmen des klassischen linearen Modells die Varianz von  $\hat{\beta}_2$  geeignet, und testen Sie bei einer Irrtumswahrscheinlichkeit  $\alpha$  von 10% die Hypothese, dass der Umsatz eines Unternehmens keinen Einfluss auf das gezahlte CEO-Gehalt hat.
6. Zeigen Sie allgemein, dass die Matrizen  $\mathbf{H} \equiv \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$  und  $\mathbf{M} \equiv \mathbf{I} - \mathbf{H}$  idempotent und symmetrisch sind und zudem gilt:  $\mathbf{H}\mathbf{X} = \mathbf{X}$ ,  $\mathbf{M}\mathbf{X} = \mathbf{0}$ ,  $\mathbf{H}\mathbf{M} = \mathbf{M}\mathbf{H} = \mathbf{0}$ . Was bedeutet dies für die Vektoren  $\mathbf{y}$ ,  $\hat{\mathbf{y}}$  und  $\hat{\boldsymbol{\epsilon}}$ ?
7. Zeigen Sie für ein lineares Modell mit Achsenabschnitt allgemein die Gültigkeit der Streuungszerlegung.

## Aufgabe 7

Zur Erklärung des CEO-Jahresgehalts von DAX30-Unternehmen sollen nun in das in Aufgabe 6, Teilaufgabe 4 verwendete Modell die Variablen  $X_4 :=$  "Gewinn vor Zinsen und Steuern (earnings before interest and taxes, EBIT)" und  $X_5 :=$  "Eigenkapitalquote" als zusätzliche exogene Variablen aufgenommen werden. Das unterstellte Modell lautet also:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \epsilon.$$

Zur empirischen Untersuchung des Modells stehen Ihnen in der Datei `Unternehmen.csv` für alle DAX30-Unternehmen die Beobachtungen aus dem Jahr 2011 zur Verfügung.

1. Lesen Sie den Datensatz in R ein, und machen Sie sich ein Bild von den Daten.
2. Schätzen Sie das Modell mittels OLS, und interpretieren Sie die Werte  $\hat{\beta}_2$  und  $\hat{\beta}_3$ .
3. Beurteilen Sie die Anpassungsgüte des Modells.
4. Stellen Sie die Verteilung der geschätzten Residuen grafisch dar. Überprüfen Sie mithilfe eines geeigneten statistischen Tests, ob die Annahme normalverteilter Störgrößen gehalten werden kann.
5. Führen Sie ausgehend von den Annahmen des klassischen linearen Modells folgende Analysen durch:
  - (a) Welche der untersuchten Variablen haben bei einer Irrtumswahrscheinlichkeit  $\alpha$  von 10% einen signifikanten Einfluss auf das CEO-Jahresgehalt?
  - (b) Sind bei einer Irrtumswahrscheinlichkeit  $\alpha$  von 1% die Variablen gemeinsam signifikant?
6. Schätzen Sie nun
  - (a) das lineare Modell ohne Achsenabschnitt, welches eine Dummyvariable für nur eine der beiden Ausprägungen des Wirtschaftssektors enthält, und
  - (b) das lineare Modell ohne Achsenabschnitt, welches Dummyvariablen für beide Ausprägungen des Wirtschaftssektors enthält,und interpretieren Sie die Ergebnisse.
7. Was passiert, wenn die Fremdkapitalquote als weitere exogene Variable in das Modell aufgenommen wird?