Prof. Dr. E. Spodarev
D. Meschenmoser

Wintersemester 2009 / 2010
26. November 2009

Statistik II – Übungsblatt 7

Abgabe: 03. Dezember 2009, vor den Übungen

Aufgabe 1

a) Sei $Z = (Z_1, Z_2, Z_3)^\top \sim \mathcal{N}(\mu, I_3)$ ein normalverteilter Zufallsvektor mit Erwartungswertvektor $\mu = (1, 7, -5)^\top$. Bestimme die Verteilung von $\frac{1}{2}Z_1^2 + Z_2^2 + \frac{1}{2}Z_3^2 - Z_1 Z_3$. (3)

b) Sei $Z = (Z_1, Z_2, Z_3)^\top \sim \mathcal{N}(\mu, K)$ ein normalverteilter Zufallsvektor mit Erwartungswertvektor $\mu = (1, -3, 2)^\top$ und Kovarianzmatrix (4)

$$K = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Bestimme die Verteilung von $Z_1^2 + Z_2^2 + Z_3^2 + 2Z_1Z_2 - 2Z_1Z_3 - 2Z_2Z_3$.

Aufgaben 2

Sei $X = (X_1, \ldots, X_n) \sim \mathcal{N}(\mu, I_n)$ multivariat normalverteilt mit Erwartungswertvektor μ und Kovarianzmatrix I_n.

a) Bestimme die charakteristische Funktion von $X^\top X$. (2)

b) Berechne den Erwartungswert von $X^\top X$. (2)

c) Berechne die Varianz von $X^\top X$. (2)

Aufgabe 3 (4)

Betrachte den Spezialfall $m = 2$ des multiplen linearen Regressionsmodells, d.h. $Y = X\beta + \varepsilon$ mit

$$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Zeige, dass dann der Kleinste-Quadrate-Schätzer $\left(\widehat{\beta}_1, \widehat{\beta}_2\right)^\top = \left(X^\top X\right)^{-1} X^\top Y$ dem aus Statistik I / Stochastik I bekannten Schätzer im einfachen linearen Regressionsmodell entspricht, d.h.

$$\widehat{\beta}_2 = \frac{s_{xy}^2}{s_{xx}^2}, \quad \widehat{\beta}_1 = \overline{y}_n - \widehat{\beta}_2 \overline{x}_n,$$

wobei $\overline{x}_n$ bzw. $\overline{y}_n$ die Stichprobenmittel bezeichnen, d.h.

$$\overline{x}_n = \frac{1}{n}\sum_{i=1}^n x_i \quad \text{bzw.} \quad \overline{y}_n = \frac{1}{n}\sum_{i=1}^n y_i,$$

und die Stichprobenvarianz s_{xx}^2 bzw. die Stichprobenkovarianz s_{xy}^2 gegeben sind durch

$$s_{xx}^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \overline{x}_n)^2 \quad \text{bzw.} \quad s_{xy}^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \overline{x}_n)(y_i - \overline{y}_n).$$

Aufgabe 4

Das Produktionsvolumen der USA zwischen 1932 und 1953 lässt sich mit Hilfe der Cobb-Douglas Produktionsfunktion

$$Y_t = c \cdot K_t^{\beta_1} \cdot A_t^{\beta_2} \cdot \varepsilon_t, \quad t \in \{1, \dots, 22\}$$

mit unbekannten Konstanten c, β_1, β_2 beschreiben. Dabei bezeichnet Y die Produktion (in Mrd. Dollar), K den Kapitaleinsatz (in Mrd. Dollar) und A den Arbeitseinsatz (in Mill. Arbeitskräften). Die Datei *production.txt* auf der Homepage der Vorlesung enthält folgende Daten:

Jahr	t	Produktion	Kapital	Arbeit
1932	1	60.3	297.1	39.3
1933	2	58.2	290.1	39.6
⋮	⋮	⋮	⋮	⋮

a) Führe den Modellansatz in ein geeignetes lineares Modell über und bestimme mit R den Kleinste-Quadrate-Schätzer für $\beta = (\log c, \beta_1, \beta_2)$. (3)

b) Gib die geschätzten Werte $\widehat{Y}_t$ sowie die Residuen $\widehat{\varepsilon}_t$ an. (2)

c) Erstelle *ein* Schaubild, das die tatsächliche Entwicklung des Produktionsvolumens Y über die Jahre 1932 – 1953 darstellt, sowie die „Verlaufskurve“ der geschätzten Daten, d.h. eine stückweise lineare Funktion durch die Punkte $(t, \widehat{Y}_t)$. (2)

   Hinweis: Mit dem Befehl `lines()` kann ein weiterer Polygonzug zu einem Schaubild hinzugefügt werden.