Prof. Dr. Volker Schmidt
Matthias Neumann

Sommersemester 2016

Stochastik I - Übungsblatt 12

Abgabe am 5.7.2016 vor Beginn der Übung

Aufgabe 1 (1 + 4 + 1 + 2 + 4 Punkte)
Betrachte folgende Daten

i	1	2	3	4	5	6	7	8	9	10
x_i	0.52	0.51	0.69	0.12	0.81	0.76	0.29	0.54	0.07	0.91
y_i	11.3	10.7	15.3	4.7	18.3	16.7	8.3	12.7	3.4	22.0

An die Daten soll ein lineares Regressionsmodel der Form

$$Y_i = \alpha + \beta x_i + \varepsilon_i \qquad \text{für jedes } i \in \{1, \ldots, 10\}$$

angepasst werden. Bearbeite die folgende Aufgabe mit **R**.

(a) Plotte die Daten als Punktwolke.

(b) Berechne die KQ-Schätzwerte $\widehat{\alpha}$ und $\widehat{\beta}$. Zeichne die geschätzte Regressionsgerade mit rot in den Plot aus Teilaufgabe (a) ein.

(c) Nutze das Regressionsmodel, um für eine zusätzliche Beobachtung $x_{11} = 2$ den zugehörigen Wert y_{11} vorherzusagen.

(d) Schätze Mittelwert und Varianz der Residuen.

(e) Bestimme Konfidenzintervalle für den Mittelwert und die Varianz der Residuen, unter der Annahme, dass die Residuen sowohl unabhängig und identisch normalverteilt sind.

Aufgabe 2 (2 + 4 Punkte)

Sei $n \geq 2$. Sei $(Y_1, \ldots, Y_n)$ eine Folge von Zufallsvariablen und $x_1, \ldots, x_n \in \mathbb{R}$ derart, dass nicht alle x_i den gleichen Wert annehmen. Betrachte das Regressionsmodel

$$Y_i = \alpha + \beta x_i + \varepsilon_i \qquad \text{für jedes } i \in \{1, \ldots, n\},$$

wobei die $\varepsilon_1, \ldots, \varepsilon_n$ eine Folge von unkorrelierten Zufallsvariablen ist mit $\mathbb{E}\,\varepsilon_i = 0$ und $\operatorname{Var} \varepsilon_i = \sigma^2 > 0$ für jedes $i \in \{1, \ldots, n\}$. Betrachte für $c_1, \ldots, c_n \in \mathbb{R}$ den linearen Schätzer $\widehat{\alpha} = c_1 Y_1 + \ldots + c_n Y_n$ für α. Zeige die folgenden Aussagen.

(a) Der Schätzer $\widehat{\alpha}$ ist genau dann erwartungstreu für α, wenn

$$\sum_{i=1}^{n} c_i = 1 \qquad \text{und} \qquad \sum_{i=1}^{n} x_i c_i = 0.$$

(b) Die Varianz des Schätzers $\widehat{\alpha}$ ist minimal, wenn für jedes $i \in \{1, \ldots, n\}$ gilt:

$$c_i = \frac{1}{n} - \frac{\overline{x}_n (x_i - \overline{x}_n)}{(n-1) s_{xx}^2}.$$

Aufgabe 3 (7 Punkte)

Seien $(X_1, Y_1), \ldots, (X_n, Y_n)$ eine Folge von $n \geq 2$ i.i.d Zufallsvektoren mit $\mathbb{E}\,X_1^2, \mathbb{E}\,Y_1^2 < \infty$. Betrachte das Regressionsmodel

$$Y_i = \alpha + \beta X_i + \varepsilon_i \qquad \text{für jedes } i \in \{1, \ldots, n\},$$

wobei $\varepsilon_1, \varepsilon_2, \ldots$ eine Folge von i.i.d. Zufallsvariablen ist mit $\mathbb{E}\,\varepsilon_i = 0$ und $\operatorname{Cov}(X_i, \varepsilon_i) = 0$ für jedes $i \in \mathbb{N}$. Bezeichne mit $\widehat{\alpha}_n$ und $\widehat{\beta}_n$ die KQ-Schätzer für α bzw. β bei einem Stichprobenumfang von $n \in \mathbb{N}$. Zeige, dass $\alpha_n \xrightarrow{\text{f.s.}} \alpha$ und $\beta_n \xrightarrow{\text{f.s.}} \beta$ für $n \longrightarrow \infty$.