

12. Übungsblatt zur Vorlesung
Schließende Statistik WS 2015/16

Aufgabe 46

Gegeben seien n Punkte (y_i, x_i) , durch die eine Gerade $y = a + b \cdot x$ gelegt werden soll.

- (a) Wie lautet das Kleinst-Quadrate-Prinzip?
(b) Zeigen Sie, dass das Kleinst-Quadrate-Prinzip auf die beiden Normalgleichungen

$$(1) \quad na + b \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

$$(2) \quad a \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i \cdot y_i$$

führt.

- (c) Geben Sie die Kleinst-Quadrate-Lösungen \hat{a} und \hat{b} an.
(d) Es sei $\hat{u}_i = y_i - \hat{y}_i$ mit $\hat{y}_i = \hat{a} + \hat{b} \cdot x_i$. Zeigen Sie die Gültigkeit der Beziehung $\sum_{i=1}^n \hat{u}_i = 0$.
(e) Zeigen Sie, dass die Beziehung $\bar{y} = \bar{\hat{y}}$ gilt, wobei $\bar{\hat{y}} = \frac{1}{n} \sum_{i=1}^n \hat{y}_i$.
(f) Zeigen Sie die Gültigkeit der Beziehung $\sum_{i=1}^n x_i \cdot \hat{u}_i = 0$.
(g) Zeigen Sie die Gültigkeit der Beziehung $\sum_{i=1}^n \hat{y}_i \cdot \hat{u}_i = 0$.
(h) Beweisen Sie die folgende Zerlegungsformel:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2 + \sum_{i=1}^n \hat{u}_i^2.$$

Aufgabe 47

Man gehe davon aus, dass sich die Abhängigkeit des systolischen Blutdrucks y_i vom Lebensalter x_i eines Menschen durch das einfache lineare Regressionsmodell

$$y_i = \beta_1 + \beta_2 x_i + u_i \quad \text{mit} \quad u_i \stackrel{\text{iid}}{\sim} N(0, \sigma^2), \quad i = 1, \dots, n$$

erklären lässt.

In einer medizinischen Studie wurden bei $n = 62$ Personen die Merkmale Alter x_i und systolischer Blutdruck y_i erhoben und daraus für die Durchführung einer einfachen linearen Regressionsanalyse bereits die folgenden Zwischenwerte errechnet:

$$\sum_{i=1}^{62} y_i = 9193; \quad \sum_{i=1}^{62} y_i^2 = 1384977; \quad \sum_{i=1}^{62} x_i = 2882;$$

$$\sum_{i=1}^{62} x_i^2 = 148292; \quad \sum_{i=1}^{62} x_i y_i = 441517$$

- (a) Schätzen Sie β_1 und β_2 mit Hilfe der Kleinst-Quadrate-Methode.
- (b) Berechnen Sie das zugehörige Bestimmtheitsmaß R^2 .
- (c) Schätzen Sie die Varianz σ^2 mit einer erwartungstreuen Schätzfunktion und berechnen Sie $\widehat{\sigma^2_{\hat{\beta}_1}}$ sowie $\widehat{\sigma^2_{\hat{\beta}_2}}$.
- (d) Testen Sie zum Signifikanzniveau $\alpha = 0.001$ (!), ob das Alter einer Person einen signifikanten Einfluss auf den systolischen Blutdruck hat.
- (e) Geben Sie ein zweiseitiges Konfidenzintervall zum Niveau $1 - \alpha = 0.95$ für β_1 an.
- (f) Testen Sie zum Signifikanzniveau $\alpha = 0.05$, ob sich β_1 signifikant von 100 unterscheidet.
- (g) Steigt der systolische Blutdruck eher mit zunehmendem oder eher mit abnehmendem Alter? Begründen Sie Ihre Antwort.