

Übungen zu Angewandte Statistik

(Abgabe am Montag, den 06.06.2016, 12:15h)

1. Lade den Datensatz `Mord.txt` von der Homepage herunter. Er enthält einen Datensatz der folgenden Form:

Einwohnerzahl	Armenanteil (%)	Arbeitslosenanteil (%)	Mordrate (%)
$X_{1,1}$	$X_{2,1}$	$X_{3,1}$	Y_1
\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot
$X_{1,n}$	$X_{2,n}$	$X_{3,n}$	Y_n

- (a) Passe ein (multivariates) lineares Modell der Form

$$Y_i = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \beta_3 X_{3,i} + \varepsilon_i$$

an den Datensatz an (um in R ein multivariates lineares Modell anzupassen benutze `lm(Y~X1+X2+X3)`).

- (b) Teste zum Signifikanzniveau $\alpha = 0.05$, ob die Einwohnerzahl einen signifikanten Einfluss auf die Mordrate hat, d.h. teste

$$H_0 : \beta_1 = 0 \text{ vs } H_1 : \beta_1 \neq 0.$$

Benutze dafür die Teststatistik

$$T = \frac{\hat{\beta}_i - \beta_i}{\sqrt{\frac{Q_r}{n-r} (A^T A)^{-1}_{ii}}}$$

Q_r kann dabei über den Befehl `deviance(lm())` ausgegeben werden. Berechne den p -Wert und vergleiche ihn mit dem Ergebnis aus `summary(lm())`.

- (c) Teste zum Signifikanzniveau $\alpha = 0.05$, ob die Hypothese

$$H_0 : 4 \cdot \beta_2 = \beta_3$$

plausibel ist. Benutze dafür die Teststatistik

$$W = \frac{Q_{R|H_0} - Q_R}{Q_R} \frac{n-r}{q}$$

(5 Punkte)

2. Lade den Datensatz `data7.txt` von der Homepage herunter. Die erste Spalte X enthält Messwerte an den Messpunkten t in der zweiten Spalte.

- (a) Passe ein lineares Modell

$$X_i = \beta_0 + \beta_1 t_i + \varepsilon_i$$

an den Datensatz an und berechne den sich daraus resultierenden Wert für das Streuungsmaß R^2 . Vergleiche ihn mit dem Ergebnis von `summary(lm())`.

(b) Plote die Punkte (t_i, x_i) zusammen mit der geschätzten Regressionsfunktion, sowie einem punktwisen Konfidenzintervall für $X(t)$ zum Niveau 0.99. Überprüfe das Ergebnis von `predict()` an der Stelle $t = 0$, indem du das Konfidenzintervall mit der Formel aus der Vorlesung berechnest.

(4 Punkte)

3. Betrachte das lineare Modell

$$\vec{X} = A\vec{\beta} + \vec{\epsilon}.$$

Sei $\widehat{\beta}$ ein Schätzer für $\vec{\beta}$ und sei

$$\widehat{X} = A\widehat{\beta}.$$

Zeige, dass

$$\vec{X}^T \vec{X} = \widehat{X}^T \widehat{X} + (X - \widehat{X})^T (X - \widehat{X}).$$

(3 Punkte)