Prof. Dr. E. Spodarev
D. Meschenmoser

Wintersemester 2009 / 2010
03. Dezember 2009

Statistik II – Übungsblatt 8

Abgabe: 10. Dezember 2009, vor den Übungen

Aufgabe 1
In dieser Aufgabe soll der Datensatz `wasser.txt` untersucht werden. Er enthält die experimentell bestimmten Siedepunkte von Wasser bei unterschiedlichem Luftdruck. Dabei bezeichnet die Spalte `Siedepunkt` den Siedepunkt in Grad Fahrenheit und die Spalte `Luftdruck` den Luftdruck in „inches of mercury“, einem amerikanischen Luftdruckmaß. Als Modell für diese Aufgabe nehmen wir an, dass der Siedepunkt linear in Abhängigkeit vom Luftdruck zunimmt. Du kannst von identisch normalverteilten und unabhängigen Störgrößen ausgehen.

a) Führe in R die lineare Regression durch ohne den Befehl `lm()` zu verwenden (d.h. mit Matrixoperationen). Gib die Regressionsparameter aus und plotte die Daten zusammen mit der Regressionsgerade in ein gemeinsames Schaubild. (2)

b) Teste mit R, ob der Siedepunkt doppelt so stark wächst wie der Luftdruck, und ob der Achsenabschnitt (d.h. der Funktionswert der Gerade bei Luftdruck 0) bei 150 liegt (in einem gemeinsamen Test). (3)

c) Leite eine Formel für die Schranken des Konfidenzbands für die Regressionsgerade her und plotte sie zusätzlich in das Schaubild. (4)

Hinweis: Beachte hierzu auch Abschnitt 6.3.1 im R-Skript.

Aufgabe 2
Die Anzahl der Autos pro 100 Einwohner möge vom Pro-Kopf-Einkommen (in 10000 €) und dem Benzinpreis (in €) abhängen. Auf der Homepage der Vorlesung ist die Datei `autos.txt` verfügbar, die folgende Daten enthält:

Land	Autos	Einkommen	Benzinpreis
Belgien	30	29.4	136
Dänemark	28	33.0	129
Deutschland	35	31.2	113
Finnland	24	21.3	113
Frankreich	33	26.4	140
Griechenland	8	10.2	129
Irland	20	11.4	92
Island	34	29.4	131
Kanada	42	26.1	39
Österreich	27	23.1	113

Bearbeite Teil a) und b) mit R und Teil c) - e) von Hand (d.h. benutze R höchstens als Taschenrechner). Nimm für Teil c) - e) zusätzlich an, dass der Vektor der Störgrössen $\mathcal{N}(0, \sigma^2 I)$-verteilt ist. In Norwegen sei zusätzlich zu obigen Werten das Pro-Kopf-Einkommen $x_{0,2} = 27.9$ und der Benzinpreis $x_{0,3} = 138$ (für Teil d) und e)).

a) Lege ein lineares Regressionsmodell zugrunde und schätze die Regressionsebene mit dem Befehl `lm()` in R. (1)

b) Plotte die Regressionsebene mit `persp()` in ein Schaubild und beschrifte die Achsen. (2)

c) Bestimme zum Niveau $\alpha = 0.05$ ein Konfidenzellipsoid für den Parametervektor β. (3)

d) Bestimme zum Niveau $\alpha = 0.05$ ein Konfidenzintervall für den erwarteten Zielwert $\varphi(1, x_{0,2}, x_{0,3}) = \beta_1 + x_{0,2}\beta_2 + x_{0,3}\beta_3$. (2)

e) Bestimme zum Niveau $\alpha = 0.05$ ein Prognoseintervall für die Zielvariable Y_0. (2)

Aufgabe 3

a) Sei A eine $n \times m$ Matrix mit $\operatorname{rg}(A) = r$, die in der Form (4)

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

dargestellt werden kann, wobei A_{11} eine $r \times r$ Matrix mit $\operatorname{rg}(A_{11}) = r$ ist. Zeige, dass eine verallgemeinerte Inverse A^- von A durch die $m \times n$ Matrix

$$A^- = \begin{pmatrix} A_{11}^{-1} & o \\ o & o \end{pmatrix}$$

gegeben ist.

b) Berechne eine verallgemeinerte Inverse von (2)

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 0 & 1 \\ 3 & 7 & 3 & 5 \end{pmatrix}$$

c) Berechne eine verallgemeinerte Inverse von (3)

$$A = \begin{pmatrix} 3 & 5 \\ 6 & 10 \\ 1 & 2 \end{pmatrix}.$$