

STATISTIK 2 (107.325) WS 2008
COMPUTERSTATISTIK (107.258) WS 2008

Übung 10

15. Jänner 2009
Dutter

37. Betrachten Sie die einfache, lineare Abhängigkeit von Cholesterin vom Alter. Zeichnen Sie die Werte, die geschätzte Gerade, den Konfidenzbereich für die Gerade und den Toleranzbereich für weitere (zukünftige, unabhängige) Beobachtungen (in Form von Hyperbeln).¹
38. Führen Sie die gleiche Aufgabe wie in Punkt 35. mit elimierten Ausreißern (Cholesterin > 350 und < 100) durch.
39. Betrachten Sie als Modell Cholesterin linear abhängig von den anderen Variablen. Welche Variable sind in erster Linie von Einfluss und welche nicht?
40. Betrachten Sie als Modell Cholesterin linear abhängig von den anderen Variablen. Berechnen und zeichnen Sie die *Mallows – C_p*-Werte. Diskutieren Sie die Resulte.

Bitte senden Sie die Ausarbeitung in Form eines pdf-Files (nicht mehr als 3 Seiten) mit den Resultaten (Outputs plus textliche Kommentare) und Listing des Programmcodes (Funktion) an

R.Dutter@tuwien.ac.at

bis zum Dienstag-Abend, 20. Jänner 2009.

Bitte den Namen des pdf-File folgendermaßen:

name_exer_9.pdf

wobei ‘name’ für den Familiennamen steht.

Empfehlenswert ist es, den Bericht mit „Sweave“ zu erstellen. (Vorlage siehe <http://www.statistik.tuwien.ac.at/public/dutt/vorles/> .)

¹Ein Konfidenzintervall mit der Konfidenzzahl α für den Mittelwert $\mu_{y.x}$ an der Stelle x erhält man mit der Formel

$$\hat{y}_x - t_{n-2;1-\frac{\alpha}{2}} S \sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{(n-1)s_x^2}} < \mu_{y.x} < \hat{y}_x + t_{n-2;1-\frac{\alpha}{2}} S \sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{(n-1)s_x^2}}.$$

Ein Toleranzintervall ist folgendermaßen:

$$\hat{y}_x - t_{n-2;1-\frac{\alpha}{2}} S \sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{(n-1)s_x^2}} < y < \hat{y}_x + t_{n-2;1-\frac{\alpha}{2}} S \sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{(n-1)s_x^2}}.$$