Prof. Dr. U. Stadtmüller
C. Hering

Sommer 2010
27.04.2010
Blatt 2

Angewandte Statistik für Biometrie

(Due: Di., 03.05.2010, 13:15 Uhr, vor den Übungen)

1. In dieser Aufgabe sollen Daten gemäß einem linearen Modell $X_i = \beta_0 + \beta_1 Y_i + \beta_2 Y_i^2 + \varepsilon_i$, $i = 1, \dots, 50$ erzeugt werden. Hierbei soll $\beta_0 = 3$, $\beta_1 = 0.01$ und $\beta_2 = 0.005$ sein. Außerdem seien Y_i bzw. ε_i iid verteilte Zufallsvariablen mit $Y_1 \sim U[1, 25]$ und $\varepsilon_1 \sim N(0, 1)$. Gehe hierzu wie folgt vor:

   - Initialisiere den Zufallszahlengenerator mit `set.seed(13)`. Dieser Befehl bewirkt, dass die Ergebnisse reproduzierbar sind.
   - Erzeuge 50 auf [1,25] gleichverteilte und 50 $N(0, 1)$-verteilte Zufallsvariablen.
   - Definiere den Vektor X gemäß dem linearen Modell.
   - Erzeuge eine Matrix, so dass die erste Spalte der Matrix X und die zweite Spalte Y ist.

   Hinweis: Die Befehle `runif,` `rnorm` *und* `cbind` *können verwendet werden.*

   Nach dem Erzeugen der Daten wollen wir nun zwei verschiedene lineare Modell anpassen. Erstens ein Modell der Form: $X_i = \beta_0 + \beta_1 Y_i$ und zweitens ein Modell der Form $X_i = \beta_0 + \beta_1 Y_i + \beta_2 Y_i^2$. Um ein Gefühl für die Schätzungen zu bekommen wollen wir für jedes dieser Modelle jeweils einen Plot erzeugen, der die Daten, die angepasste Regressionskurve und die Konfidenzbereich der Regressionskurve zeigt. Führe dazu folgende Schritte durch:

   - Ordne die Zeilen der Matrix so, dass die zweite Spalte aufsteigend sortiert ist.
   - Passe die beiden linearen Modelle an und führe für jedes die folgenden Schritte durch
     - Plotte die Daten
     - Erzeuge einen Konfidenzbereich mit Konfidenzlevel 0.99.
     - Verwende den Befehl `matlines`, um den Konfidenzbereich und die Regressionskurve im Plot anzuzeigen.

   Hinweis: Die Befehle `order,` `lm` *und* `predict` *können verwendet werden.*

   Bemerkung: Die theoretische Interpretation eines Konfidenzbereiches wird in der Vorlesung noch genauer erläutert.

   (8 Punkte)

2. Gegeben sei das lineares Modell $\vec{X} = A\vec{\beta} + \vec{\varepsilon}$, wobei $\vec{\varepsilon} \sim N(\vec{0}, I_n)$ sei. Man beweise dann die Unabhängigkeit der Schätzer

$$\hat{\beta} = (A^T A)^- A^T \vec{X} \quad \text{und} \quad \hat{\sigma}^2 = \frac{1}{n-r} ||\vec{X} - A(A^T A)^- A^T \vec{X}||^2,$$

   wobei $r = \text{Rang}(A)$.

   (4 Punkte)

http://www.uni-ulm.de/mawi/zawa/lehre/sommer2010/asb2010.html