TU Ilmenau
Jun.-Prof. Dr. Thomas Hotz
Philipp Schmitz

Sommersemester 2017

Blatt 4

Inferenz im GLM

Aufgabe 4.1 Ein- und Zweistichprobenproblem

a) Leiten Sie für das Einstichprobenproblem

$$y \sim \mathcal{N}(X\beta, \sigma^2 \mathbf{I}_n), \quad \beta \in \mathbf{R}, X = \mathbf{1}_n = (1, \ldots, 1)^t \in \mathbf{R}^n,$$

den Einstichproben-t-Test mit einem entsprechenden Konfidenzintervall her.

b) Betrachten Sie das Zweistichprobenproblem

$$y \sim \mathcal{N}(X\beta, \sigma^2 \mathbf{I}_n), \quad \beta \in \mathbf{R}^2, X = \begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{0}_{n_1} \\ \mathbf{0}_{n_2} & \mathbf{1}_{n_2} \end{pmatrix}, n = n_1 + n_2$$

und leiten Sie den Zweistichproben-t-Test sowie ein Konfidenzintervall für die Differenz der Erwartungswerte her. Wie müssen n_1 und n_2 gewählt werden, damit die Varianz des Schätzers für die Differenz der Erwartungswerte minimal wird?

Aufgabe 4.2 Körpergrößen

Betrachten Sie den Datensatz `Heights` aus dem Paket `library(alr4)`. Führen Sie einen F-Test hinsichtlich der Hypothese, dass die Körpergröße der Töchter linear mit Steigung 1 von der Körpergröße der Mütter abhängt, zum Signifikanzniveau $\alpha = 0.05$ durch und stellen Sie den Konfidenzbereich graphisch dar.

Aufgabe 4.3 Fallzahlberechnung

Geplant ist eine Studie zur Ermittelung des Zusammenhangs der Körpergröße von Töchtern und der Körpergröße ihrer Mütter. Es soll getestet werden, ob die Körpergröße der Töchter linear von der Körpergröße der Mütter abhängt, wobei $\beta_2 = 1$ der Anstieg sei. Gesucht ist die Anzahl n an Stichproben so, dass der Anstieg β_2 bis auf ± 0.05 genau zum Signifikanzniveau $\alpha = 0.05$ geschätzt werden kann.
Wir nehmen an, dass die Körpergrößen der Mütter unabhängig identisch $\mathcal{N}(\mu, \tau^2)$ verteilt sind, wobei die Parameter μ und τ^2 zu Beginn unbekannt sind. Durch eine Pilotstudie mit 20 Stichproben sollen daher diese und weitere benötigte Parameter geschätzt werden. Betrachten Sie den Datensatz `Heights` aus dem Paket `library(alr4)` und entnehmen Sie diesem zufällig 20 Daten. Machen Sie sich dazu mit dem Befehl `sample.int` vertraut. Bestimmen Sie nun eine Näherung $\tilde{C}$ der Matrix $C = (X^t X)^{-1}$ mit Hilfe des starken Gesetzes der großen Zahl (vgl. Bemerkung 4.10).

Nächste Übung: Donnerstag 18. Mai 2017 um 13:00 Uhr im Sr C 115