

Angewandte Statistik für Biometrie

(Abgabe: Di., 12.07.2011, vor den Übungen)

1. Betrachte das Random-Effekt-Modell aus der Vorlesung:

$$X_{ij\nu} = \mu + \alpha_i + b_j + c_{ij} + \varepsilon_{ij\nu}$$

wobei $1 \leq i \leq k$, $1 \leq j \leq l$, $1 \leq \nu \leq m$, $n = k \cdot l \cdot m$ und

$$\varepsilon \sim N(0, \sigma_\varepsilon^2 I_n), \quad b \sim N(0, \sigma_b^2 I_l), \quad c \sim N(0, \Sigma_c)$$

unabhängig sind. Die Kovarianzmatrix Σ_c hat die Einträge $\text{var}(c_{i,j}) = \frac{k-1}{k} \sigma_c^2$ und $\text{cov}(c_{i_1 j_1}, c_{i_2 j_2}) = 0$, falls $j_1 \neq j_2$ und $\text{cov}(c_{i_1 j_1}, c_{i_2 j_2}) = -\frac{1}{k} \sigma_c^2$, falls $i_1 \neq i_2$ und $j_1 = j_2$.

Zeige dass für $i_1 \neq i_2$ gilt $\text{cov}(\bar{X}_{i_1 \bullet \bullet}, \bar{X}_{i_2 \bullet \bullet}) = \frac{1}{l} (\sigma_b^2 - \frac{1}{k} \sigma_c^2)$.

(6 Punkte)

2. Der Datensatz `bryk.dat` enthält die Ergebnisse eines Mathetest unter 7185 Schülern durchgeführt an amerikanischen Highschools im Jahre 1982. Neben der Nummer der Schule, und den Ergebnissen (`mathach`) enthält der Datensatz auch einen sogenannten Sozioökonomischen Faktor (`ses`, 0-zentriert), sowie die Variable (`sector`), die angibt, ob es sich um eine Schule handelt, die vom Staat getragen wird, oder ob die katholische Kirche die Trägerschaft der Schule inne hat. Lade den Datensatz `bryk.dat` von der Veranstaltungshomepage.

- Passe für jeden Sektor je ein lineares Modell an den Datensatz an. Hierbei soll der Faktor `ses` als zufällige und der Faktor `school` als deterministische Einflussgröße ins Modell einfließen.
- Erzeuge und veranschauliche mittels zweier Plots (einen für jeden Sektor) die Konfidenzintervalle der geschätzten y -Achsenabschnitte und Steigungen für alle Schulen. Interpretiere und kommentiere die resultierenden Graphiken.
- Erzeuge für die geschätzten y -Achsenabschnitte und Steigung für jeden Sektor jeweils einen Boxplot und vergleiche die Ergebnisse zwischen den Sektoren.

Hinweis: Das Paket `nlme`, die Funktionen `lmList`, `lme` und `intervals`, sowie ihre Dokumentationen könnten hilfreich sein.

(6 Punkte)