Prof. Dr. U. Stadtmüller
M. Weyhmüller

Sommer 2011
14.06.2011
Blatt 10

Angewandte Statistik für Biometrie

(Abgabe: Di., 21.06.2011, vor den Übungen)

1. Es soll analysiert werden ob die verschiedenen Ausgabeschalter der Mensa einen Einfluss auf die Essenswahl der Studenten haben. Lies den Datensatz `mensa.dat` von der Homepage mit R ein, er enthält die Anzahl der verkauften Essen der jeweiligen Ausgabeschalter („Gut und Günstig“, „Lecker und Fein“, „Vegetarisch“, „Bio“) an verschiedenen Tagen.

   (a) Führe eine Varianzanalyse durch, mit der Nullhypothese dass die Studenten keinen der vier Ausgabeschalter signifikant (zum Niveau $\alpha = 5\%$) bevorzugen. Wie lautet das Ergebnis?

   (b) Führe eine Varianzanalyse durch, mit der Nullhypothese dass die Studenten keinen der beiden Ausgabeschalter „Gut und Günstig“ und „Lecker und Fein“ signifikant (zum Niveau $\alpha = 5\%$) bevorzugen. Überprüfe die Hypothese ebenso für „Vegetarisch“ und „Bio“. Wie lautet das Ergebnis?

   Hinweis: Die Funktionen `aov()` und `anova()` sowie ihre Dokumentationen könnten hilfreich sein.

   (2+2 Punkte)

2. Wir betrachten das Modell der einfaktoriellen Varianzanalyse aus der Vorlesung, also

$$X_{i,j} = \mu + \alpha_i + \varepsilon_{i,j},$$

   wobei $1 \leq j \leq n_i, 1 \leq i \leq k$ und $\varepsilon \sim \mathrm{N}(0, \sigma^2 I_n)$ mit $n := \sum_{i=1}^k n_i$, sowie $\sum_{i=1}^k n_i \alpha_i = 0$. Wie in der Vorlesung sei $\overline{X}_{\bullet\bullet} := \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} X_{i,j}$ und $\overline{X}_{i\bullet} := \frac{1}{n_i} \sum_{j=1}^{n_i} X_{i,j}$.

   (a) Finde je eine idempotente Matrix A um die Summen $\sum_{i=1}^k \sum_{j=1}^{n_i} (X_{i,j})^2$, $\sum_{i=1}^k \sum_{j=1}^{n_i} (\overline{X}_{\bullet\bullet})^2$ und $\sum_{i=1}^k \sum_{j=1}^{n_i} (\overline{X}_{i\bullet})^2$ in der Form $X^\top AX$ darzustellen.

   (b) Finde je eine idempotente Matrix A um die Größen $Q_R := \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{i,j} - \overline{X}_{i\bullet})^2$, $Q_{R|H_0} := \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{i,j} - \overline{X}_{\bullet\bullet})^2$ und $Q_0 := Q_{R|H_0} - Q_R = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{i\bullet} - \overline{X}_{\bullet\bullet})^2$ in der Form $X^\top AX$ darzustellen.

   (c) Bestimme die Verteilung der Größen Q_R und Q_0 unter $H_0 : \alpha_1 = \ldots = \alpha_k$ und zeige dass sie unabhängig sind.

   Hinweis: Eventuell ist Lemma 1.1, sowie Lemma 1.2. über die Verteilung bzw. Unabhängigkeit von quadratischen Formen $X^\top AX$ hilfreich.

   (2+3+3 Punkte)