

Statistik II – Übungsblatt 10

Abgabe: 14. Januar 2010, vor den Übungen

Aufgabe 1

(3)

Im Modell der einfaktoriellen Varianzanalyse wurde folgende Reparametrisierung der Erwartungswerte μ_1, \dots, μ_k betrachtet:

$$\mu_i = \mu + \alpha_i \quad \forall i = 1, \dots, k \quad \text{mit} \quad \sum_{i=1}^k n_i \alpha_i = 0.$$

Zeige, dass diese Darstellung eindeutig ist.

Aufgabe 2

Betrachte das (reparametrisierte) Modell der einfaktoriellen Varianzanalyse, d.h.

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

mit $i = 1, \dots, k$, $j = 1, \dots, n_i$, $\sum_{i=1}^k n_i = n$, $\mathbb{E}(\varepsilon) = 0$, $\text{Cov}(\varepsilon) = \sigma^2 I$, $\sigma^2 > 0$.

- Bestimme die MKQ-Schätzer $\hat{\mu}$ für μ und $\hat{\alpha}_1, \dots, \hat{\alpha}_k$ für $\alpha_1, \dots, \alpha_k$, d.h. bestimme die Lösung der Normalgleichung $X^\top X \beta = X^\top Y$. (3)
- Formuliere die Nullhypothese $H_0 : \alpha_1 = \dots = \alpha_k$ in Matrixschreibweise, d.h. finde eine Matrix H und einen Vektor d , so dass H_0 äquivalent ist zu $H\beta = d$. (3)
- ~~Zeige, dass H_0 eine testbare Hypothese ist. Kommt auf die Lösung von Teil b) an!~~ Überprüfe, ob die Hypothese aus Teil b) testbar ist. (Bonuspunkte) (3*)

Hinweise: Auf Grund der Nebenbedingung $\sum_{i=1}^k n_i \alpha_i = 0$ besitzt die Normalgleichung eine eindeutige Lösung, d.h. es gibt genau eine verallgemeinerte Inverse von $X^\top X$, die zum Ziel führt. Deshalb ist es einfacher, sich die einzelnen Gleichungen in der Normalgleichung genau anzuschauen, als genau diese verallgemeinerte Inverse von $X^\top X$ zu finden.

Aufgabe 3

Ein Ulmer Bäckereibetrieb misst die Verkaufszahlen seiner Filialen in den drei Stadtteilen Weststadt, Eselsberg und Stadtmitte, wobei er in jedem Stadtteil sechs Filialen unterhält. Folgende Tabelle enthält die Anzahlen der verkauften Brötchen an einem bestimmten Tag.

Weststadt	Eselsberg	Stadtmitte
703	770	819
788	816	784
715	797	801
832	774	939
890	867	833
782	884	786

- a) Bestimme die Schätzwerte für die Parameter μ und α_i des reparametrisierten Modells der einfaktoriellen Varianzanalyse. (3)
- b) Teste zum Niveau $\alpha = 0.05$ die ANOVA-Nullhypothese mit Hilfe der oben berechneten Schätzer für μ und α_i , ohne die in R eingebauten Methoden zu verwenden. Gib alle Zwischenergebnisse an. (3)

Hinweis: Man kann zeigen, dass für die Testgröße aus Abschnitt 3.3.5 im Skript folgendes gilt:

$$\begin{aligned} T &= \frac{(H\bar{\beta} - d)^\top \left(H (X^\top X)^{-1} H^\top \right)^{-1} (H\bar{\beta} - d)}{s\bar{\sigma}^2} \\ &= \frac{(n - k) \sum_{i=1}^k n_i (\hat{\alpha}_i - \hat{\mu})^2}{(k - 1) \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \hat{\alpha}_i)^2} \end{aligned}$$

Aufgabe 4

Beweise folgende Quadratsummenzerlegung für das (reparametrisierte) Modell der einfaktoriellen Varianzanalyse:

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \hat{\mu})^2 = \sum_{i=1}^k n_i \hat{\alpha}_i^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - (\hat{\alpha}_i + \hat{\mu}))^2$$

Dabei bezeichnen $\hat{\mu}, \hat{\alpha}_1, \dots, \hat{\alpha}_k$ die MKQ-Schätzer für $\mu, \alpha_1, \dots, \alpha_k$ aus Aufgabe 1.

Aufgabe 5

Gegeben sei das (reparametrisierte) Modell der zweifaktoriellen Varianzanalyse ohne Wechselwirkungen, d.h.

$$Y_{i_1 i_2 j} = \mu + \alpha_{i_1} + \beta_{i_2} + \varepsilon_{i_1 i_2 j}, \quad 1 \leq i_1, i_2, j \leq 2,$$

mit $\mathbb{E}(\varepsilon) = 0$, $\text{Cov}(\varepsilon) = \sigma^2 I$, $\sigma^2 > 0$ und Parametervektor $\beta = (\mu, \alpha_1, \alpha_2, \beta_1, \beta_2)^\top$. Entscheide, ob folgende Linearkombinationen von β erwartungstreu schätzbar sind. Falls ja, bestimme einen Schätzer mit minimaler Varianz in der Klasse der linearen, erwartungstreuen Schätzfunktionen:

a) $f_1(\beta) = \alpha_1 - \beta_1$ (2)

b) $f_2(\beta) = \mu + \alpha_1 + \beta_1$ (4)

c) $f_3(\beta) = \mu + 2\alpha_1 - \alpha_2 + \beta_2$ (4)

Frohe Weihnachten und einen Guten Rutsch!

