Vorlesung *Statistik 1*
Wintersemester 2015/16
Ruprecht-Karls-Universität Heidelberg

Prof. Dr. Jan JOHANNES

2. Übungsblatt

Aufgabe 1. *(Frisch-Waugh Theorem)* Sei $Y \circledsim \{\mathfrak{L}(X\beta, \sigma^2 \operatorname{Id}_n), \beta \in \mathbb{R}^p, \sigma > 0\}$ durch ein gewöhnliches lineares Modell mit $X = (X_1, X_2)$ und $\beta^t = (\beta_1^t, \beta_2^t)$ für $\beta_i \in \mathbb{R}^{p_i}$, $p_i > 0$, und $X_i \in \mathbb{R}^{n \times p_i}$, $i = 1, 2$, sowie $n \geqslant p_1 + p_2$ adäquat beschrieben. Die Matrix X besitze den Rang $\operatorname{rg}[(X_1, X_2)] = p_1 + p_2$ und die Koordinaten der Störgröße $\varepsilon := Y - X\beta$ seien unabhängig und identisch verteilt. Wir untersuchen den gewöhnlichen Kleinste-Quadrate-Schätzer $\widehat{\beta}^t = (\widehat{\beta}_1^t, \widehat{\beta}_2^t)$ von $\beta^t = (\beta_1^t, \beta_2^t)$. Dazu sei $\Pi_{\mathcal{R}(X_1)}$ die Orthogonalprojektion auf $\mathcal{R}(X_1) = \{X_1 b : b \in \mathbb{R}^{p_1}\}$ und $\widetilde{X}_2 := (\operatorname{Id}_n - \Pi_{\mathcal{R}(X_1)})X_2$ sowie $\widetilde{Y} = (\operatorname{Id}_n - \Pi_{\mathcal{R}(X_1)})Y$. Zeigen Sie die folgenden Aussagen

(i) $\widehat{\beta}_2 = (\widetilde{X}_2^t \widetilde{X}_2)^{-1} \widetilde{X}_2^t \widetilde{Y}$;

(ii) $\widehat{\beta}_2$ ändert sich nicht, wenn $\widetilde{Y}$ durch Y ersetzt wird, so dass $\widehat{\beta}_2 = (\widetilde{X}_2^t \widetilde{X}_2)^{-1} \widetilde{X}_2^t Y$ gilt.

Hinweis: Zeigen Sie, dass $\widetilde{X}_2 (\widetilde{X}_2^t \widetilde{X}_2)^{-1} \widetilde{X}_2^t$ eine Orthogonalprojektion ist (auf welchen Raum?) und folgern Sie, dass $\widetilde{X}_2 \widehat{\beta}_2 = \widetilde{X}_2 (\widetilde{X}_2^t \widetilde{X}_2)^{-1} \widetilde{X}_2^t \widetilde{Y}$ bzw. $\widetilde{X}_2 \widehat{\beta}_2 = \widetilde{X}_2 (\widetilde{X}_2^t \widetilde{X}_2)^{-1} \widetilde{X}_2^t Y$ gilt.

(4 Punkte)

Aufgabe 2. Beweisen Sie den Korrespondenzsatz zwischen Tests und Konfidenzbereichen: Ist für jeden Parameter $\theta_o \in \Theta$ ein Test φ_{θ_o} auf die Hypothese $H_0 : \theta = \theta_o$ (gegen eine beliebige Alternative) vom Niveau $\alpha \in (0, 1)$ gegeben, so bildet

$$C := \{\theta_o \in \Theta : \varphi_{\theta_o} \text{ lehnt die Hypothese } H_0 : \theta = \theta_o \text{ nicht ab }\}$$

eine $(1 - \alpha)$ Konfidenzmenge. Andersherum kann aus einer $(1 - \alpha)$ Konfidenzmenge für jeden Parameter $\theta_o \in \Theta$ ein Test φ_{θ_o} auf die Hypothese $H_0 : \theta = \theta_o$ zum Niveau α konstruiert werden (wie?).

(2 Punkte)

Aufgabe 3. Es seien X und Y reellwertige Zufallsvariablen mit $\mathbb{E}(|Y|) < \infty$ und gemeinsamer Lebesgue-Dichte $f_{X,Y}$. Als bedingte Dichte von Y gegeben $X = x$ definiert man

$$f_{Y|X=x}(y) := \frac{f_{X,Y}(x,y)}{\int_{\mathbb{R}} f_{X,Y}(x,z)dz}.$$

(i) Zeigen Sie, dass $f_{Y|X=x}(y)$ für P^X-fast alle x wohl definiert ist.

(ii) Betrachten Sie die Funktion $g(x) := \int_{\mathbb{R}} y f_{Y|X=x}(y)dy$, falls die rechte Seite wohldefiniert ist, und $g(x) := 0$ andernfalls. Weisen Sie nach, dass $g(X)$ eine Version der bedingten Erwartung $\mathbb{E}[Y|X]$ ist.

(iii) Sind X und Y zusätzlich gemeinsam normalverteilt mit Erwartungswert $\mu \in \mathbb{R}^2$ und Kovarianzmatrix $\Sigma \in \mathbb{R}^{2\times 2}$. Bestimmen Sie $\mathbb{E}[Y|X]$.

(iv) Sei $\mathbb{E}(|X|) < \infty$, berechnen Sie die bedingte Erwartung $\mathbb{E}\big[X \,\big|\, |X|\big]$.

(5 Punkte)

Aufgabe 4. Es seien X und Y reellwertige Zufallsvariablen mit $\mathbb{E}(|Y|) < \infty$, gemeinsamer Lebesgue-Dichte und bedingter Dichte von Y gegeben $X = x$ wie in Aufgabe 3. $m(x)$ ist ein bedingter Median von Y gegeben $X = x$, sofern (falls wohldefiniert)

$$\int_{-\infty}^{m(x)} f_{Y|X=x}(y)dy = 1/2$$

gilt. Zeigen Sie, dass m die Minimalitätseigenschaft

$$\mathbb{E}(|Y - m(X)|) = \inf_h \mathbb{E}(|Y - h(X)|)$$

besitzt, wobei sich das Infimum über alle Borel-messbaren $h : \mathbb{R} \to \mathbb{R}$ erstreckt. Welches Kriterium minimieren die entsprechend definierten bedingten α-Quantile, $\alpha \in (0, 1)$?

Hinweis: Jeder Wert m mit $P(Z \leqslant m) \geqslant 1/2$ und $P(Z \geqslant m) \geqslant 1/2$ heißt Median der reellwertigen Zufallsvariable Z. Zeigen und benuzten Sie, dass für $c > m$ gilt $\mathbb{E}(|Z - c|) - \mathbb{E}(|Z - m|) = (c - m)[P(Z \leqslant m) - P(Z > m)] + 2\mathbb{E}([c - Z]\mathbb{1}_{\{m<Z<c\}})$.

(5 Punkte)