Dr. Stefan Richter, Timo Dörzbach
Wahrscheinlichkeitstheorie 2
Sommersemester 2023

Abgabeblatt 7

Aufgabe A1 (Konvergenz der emp. Verteilungsfunktion der Residuen einer Regression, 10 Punkte).
Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Betrachtet wird das einfache Regressionsmodell

$$Y_i = b_0 X_i + Z_i, \quad i = 1, ..., n$$

wobei $b_0 \in \mathbb{R}$ ein unbekannter Parameter ist, und Z_i, X_i i.i.d. Zufallsvariablen mit $\mathbb{E}Z_1 = 0$, $\mathbb{E}[Z_1^2] < \infty$. Die X_i seien auch unabhängig von den Z_i. Sei $P = \mathbb{P}^{(X_1, Z_1)}$, und es bezeichne $\hat{P}_n$, $\mathbb{G}_n$ die empirischen Maße bzw. Prozesse bzgl. $(X_1, Z_1), ..., (X_n, Z_n)$.

(a) Zeigen Sie, dass der Kleinste-Quadrate-Schätzer

$$\hat{b}_n = \arg\min_{b \in \mathbb{R}} \hat{R}_n(b), \qquad \hat{R}_n(b) = \frac{1}{n} \sum_{i=1}^{n} (Y_i - bX_i)^2$$

durch $\hat{b} = \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2}$ gegeben ist.

(b) Betrachtet werden nun die empirischen Residuen

$$\hat{Z}_i = Y_i - \hat{b}_n X_i, \quad i = 1, ..., n$$

Definiere

$$\tilde{F}_Z(z) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{\hat{Z}_i \leq z\}},$$

und für $z, x \in \mathbb{R}, b \in \mathbb{R}$:

$$f_{z,b}(x, y) = \mathbb{1}_{\{y \leq z - (b_0 - b)x\}}$$

Zeigen Sie, dass gilt:

$$\tilde{F}_Z(z) = \hat{P}_n f_{z, \hat{b}_n}.$$

(c) Begründen Sie, dass

$$\mathcal{F} = \{f_{z,b} : z \in \mathbb{R}, b \in \mathbb{R}\}$$

erfüllt: $N_2(\varepsilon, \mathcal{F}, P) \leq \frac{2}{\varepsilon^2}$ und $J_2(1, \mathcal{F}, P) < \infty$.

Ab nun sei $z_0 \in \mathbb{R}$ fest gewählt. Es soll gezeigt werden, dass

$$\sqrt{n}\big(\tilde{F}_Z(z_0) - F_Z(z_0)\big) \xrightarrow{D} N(0, \sigma^2), \qquad (*)$$

wobei σ^2 zu bestimmen ist. Sei dazu angenommen, dass F_Z in z_0 differenzierbar ist.

(d) Sei $\hat{f}_n := f_{z_0,\hat{b}_n}$ und $f_0 := f_{z_0,b_0}$. Zeigen Sie, dass

$$\int |\hat{f}_n(x,y) - f_0(x,y)|^2 dP(x,y) \xrightarrow{\mathbb{P}} 0.$$

(e) Folgern Sie mit Lemma 1.72: $\mathbb{G}_n(\hat{f}_n - f_0) \xrightarrow{\mathbb{P}} 0$.

(f) Zeigen Sie (*) und ermitteln Sie σ^2.
Hinweis: Beispiel 1.71 / 1.73 und Delta-Methode in $\mathbb{R}$.

Aufgabe A2 (Brackets von monotonen Funktionen, 6 Bonuspunkte).

Sei P ein beliebiges Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ und F die zugehörige Verteilungsfunktion. Sei

$$\mathcal{F} = \{f : \mathbb{R} \to [0,1] \text{ monoton wachsend}\}$$

Zeigen Sie, dass für beliebiges $r \geq 1$ gilt:

$$N_r(\varepsilon, \mathcal{F}, P) \leq C_r \cdot \frac{1}{\varepsilon} \cdot \log\Big(\frac{1}{\varepsilon}\Big),$$

wobei $C_r > 0$ eine Konstante ist, die nur von r abhängt.
Hinweis: Gehen Sie ähnlich vor wie bei der Abschätzung der Bracketing-Zahlen der Funktionenklasse der empirischen Verteilungsfunktion. Teilen Sie für $n \in \mathbb{N}$ also zunächst $\mathbb{R}$ in Intervalle $[x_{i-1}, x_i]$ mit $x_0 = -\infty$, $x_n = \infty$ auf, so dass $F(x_n-) - F_{x_{n-1}} \leq \frac{1}{n}$. Nun teilen Sie den y-Bereich in jedem Intervall in geeignete Abschnitte auf. Dadurch erhalten Sie insgesamt (mindestens) $(n+1)^n$ stückweise konstante Funktionen, die Sie als Brackets verwenden können.

Abgabe bis Donnerstag 11:00 Uhr, 15.06.2023, im Zettelkasten 01 im Mathematikon INF205, 1. Etage.
Vorlesungswebseite:
`https://ssp.math.uni-heidelberg.de/wt2-ss23/`