Dr. Stefan Richter, Timo Dörzbach
Wahrscheinlichkeitstheorie 2
Sommersemester 2023

Präsenzblatt 7 - Lösungen

Aufgabe P1 (Konvergenz der Verteilungsfkt. der Residuen bei einer Regression). Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Betrachtet wird das einfache Regressionsmodell

$$Y_i = a_0 + Z_i, \quad i = 1, ..., n$$

wobei $a_0 \in \mathbb{R}$ ein unbekannter Parameter ist, und Z_i i.i.d. Zufallsvariablen mit $\mathbb{E}Z_1 = 0$, $\mathbb{E}[Z_1^2] < \infty$. Sei $P = \mathbb{P}^{Z_1}$, und es bezeichne $\hat{P}_n$, $\mathbb{G}_n$ die empirischen Maße bzw. Prozesse bzgl. $Z_1, ..., Z_n$.

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

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

durch $\hat{a} = \overline{Y_n} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ gegeben ist.

(b) Betrachtet werden nun die empirischen Residuen

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

Definiere

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

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

$$f_{z,a}(y) = \mathbb{1}_{\{y \le z - (a_0 - a)\}}$$

Zeigen Sie, dass gilt:

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

(c) Begründen Sie, dass

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

erfüllt: $N_2(\varepsilon, \mathcal{F}, P) \le \frac{2}{\varepsilon^2}$ 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 $\hat{f}_n := f_{z_0,\hat{a}_n}$ und $f_0 := f_{z_0,a_0}$. Zeigen Sie, dass

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

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

(f) Zeigen Sie, dass

$$\sqrt{n}\big(\tilde{F}_Z(z_0) - F_Z(z_0)\big) = \mathbb{G}_n(\hat{f}_n - f_0) + \sqrt{n}\big(F_Z\big(z_0 - (a_0 - \hat{a}_n)\big) - F_Z(z_0)\big) + \sqrt{n}\Big(\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Z_i \le z_0\}} - F_Z(z_0)\Big)$$

(g) Zeigen Sie (*) und ermitteln Sie σ^2.
Hinweis: Delta-Methode in $\mathbb{R}$.

Lösung:

(a) Es gilt

$$\hat{R}_n'(a) = -\frac{2}{n}\sum_{i=1}^n (Y_i - a), \qquad \hat{R}_n''(a) = \frac{2}{n}\sum_{i=1}^n 1 = 2 > 0.$$

Damit ist $\hat{R}_n$ strikt konvex auf $\mathbb{R}$, hat also entweder kein Minimum oder ein globales, eindeutiges Minimum. Lösen von

$$0 = \hat{R}_n'(a) = -\frac{2}{n}\Big(\sum_{i=1}^n Y_i - an\Big) = 2 \cdot \Big(-\overline{Y_n} + a\Big)$$

liefert $\hat{a}_n = \overline{Y_n}$ als lokales und damit globales Minimum von $\hat{R}_n$.

(b) Hier geht es nur um Formalia. Es ist

$$\hat{Z}_i = Y_i - \hat{a}_n = Z_i + a_0 - \hat{a}_n$$

$$\hat{P}_n f_{z,\hat{a}_n} = \frac{1}{n}\sum_{i=1}^n f_{z,\hat{a}_n}(Z_i) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Z_i \le z - (a_0 - \hat{a}_n)\}} = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Z_i + a_0 - \hat{a}_n \le z\}} = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{\hat{Z}_i \le z\}} = \tilde{F}_Z(z).$$

(c) Der Werte $z - (a_0 - a)$ variiert für $z, a \in \mathbb{R}$ ebenfalls nur über $\mathbb{R}$. Damit gilt

$$\mathcal{F} = \{y \mapsto \mathbb{1}_{\{y \le z - (a_0 - a)\}} : z \in \mathbb{R}, a \in \mathbb{R}\} = \{y \mapsto \mathbb{1}_{\{y \le x\}} : x \in \mathbb{R}\}.$$

Damit sind die Bracketing-Zahlen der Funktionenklasse $\mathcal{F}$ dieselben wie bei der empirischen Verteilungsfunktion. Diese wurden in der Vorlesung abgeschätzt mittels $N_2(\varepsilon, \mathcal{F}, P) \le \frac{2}{\varepsilon^2}$. Daraus ergibt sich unmittelbar (mit $\log(1+x) \le x$):

$$\begin{aligned} J(1, \mathcal{F}, P) &= \int_0^1 \sqrt{1 \vee \log N_2(\varepsilon, \mathcal{F}, P)}\,d\varepsilon \le \int_0^1 \sqrt{1 + \log(1 + \frac{1}{\varepsilon})}\,d\varepsilon \\ &\le \int_0^1 \sqrt{1 + \frac{1}{\varepsilon}}\,d\varepsilon \le \int_0^1 \sqrt{\frac{2}{\varepsilon}}\,d\varepsilon = \Big[2\sqrt{2}\sqrt{\varepsilon}\Big]_0^1 = 2\sqrt{2} < \infty. \end{aligned}$$

(d) Es gilt mit den Definitionen aus der Aufgabe:

$$\begin{aligned} \int |\hat{f}_n(x) - f_0(x)|^2 dP(x) &= \int |f_{z_0,\hat{a}_n}(x) - f_{z_0,a_0}(x)| dP(x) \\ &= \int \Big|\mathbb{1}_{\{x \le z_0 - (a_0 - \hat{a}_n)\}} - \mathbb{1}_{\{x \le z_0\}}\Big|^2 dP(x) \\ &= \int \underbrace{\Big|\mathbb{1}_{\{z_0 < x \le z_0 - (a_0 - \hat{a}_n) \text{ oder } z_0 - (a_0 - \hat{a}_n) < x \le z_0\}}\Big|^2}_{= \mathbb{1}_{\{z_0 < x \le z_0 - (a_0 - \hat{a}_n) \text{ oder } z_0 - (a_0 - \hat{a}_n) < x \le z_0\}}} dP(x) \\ &= P\Big(z_0 < Z_1 \le z_0 - (a_0 - a) \text{ oder } z_0 - (a_0 - \hat{a}_n) < x \le z_0\Big)\Big|_{a = \hat{a}_n} \\ &= \Big|F_Z(z_0 - (a_0 - \hat{a}_n)) - F_Z(z_0)\Big|. \end{aligned}$$

Da F_Z stetig in z_0 und $\hat{a}_n = \overline{Y_n} = \left(\frac{1}{n}\sum_{i=1}^n Z_i\right) + a_0 \to \mathbb{E}Z_1 + a_0 = a_0$ f.s. nach dem starken Gesetz der großen Zahlen, folgt

$$\int |\hat{f}_n(x) - f_0(x)|^2 dP(x) \to 0 \quad f.s.$$

und damit auch stochastisch.

(e) Laut (c) ist $\mathcal{F}$ P-Donsker. Laut (d) gilt $\int |\hat{f}_n(x) - f_0(x)|^2 dP(x) \xrightarrow{\mathbb{P}} 0$. Damit sind die Voraussetzungen von Lemma 1.72 erfüllt und es folgt $\mathbb{G}_n(\hat{f}_n - f_0) \xrightarrow{\mathbb{P}} 0$.

(f) Es wurde bereits in (b) gesehen, dass $\tilde{F}_Z(z_0) = \hat{P}_n f_{z_0,\hat{a}_n}$. Damit ist

$$\begin{aligned}
& \sqrt{n}(\tilde{F}_Z(z_0) - F_Z(z_0)) \\
= \; & \sqrt{n}\big(\hat{P}_n f_{z_0,\hat{a}_n} - P f_{z_0,a_0}\big) \\
= \; & \sqrt{n}\big(\hat{P}_n f_{z_0,\hat{a}_n} - P f_{z_0,\hat{a}_n}\big) + \sqrt{n}\big(P f_{z_0,\hat{a}_n} - P f_{z_0,a_0}\big) \\
= \; & \underbrace{\sqrt{n}\big(\hat{P}_n(f_{z_0,\hat{a}_n} - f_{z_0,a_0}) - P(f_{z_0,\hat{a}_n} - f_{z_0,a_0})\big)}_{=\mathbb{G}_n(\hat{f}_n - f_0)} + \underbrace{\sqrt{n}\big(\hat{P}_n f_{z_0,a_0} - P f_{z_0,a_0}\big)}_{=\sqrt{n}(\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Z_i \le z_0\}} - F_Z(z_0))} \\
& + \underbrace{\sqrt{n}\big(P f_{z_0,\hat{a}_n} - P f_{z_0,a_0}\big)}_{=\sqrt{n}\big(F_Z(z_0 - (a_0 - \hat{a}_n)) - F_Z(z_0)\big)}
\end{aligned}$$

(g) Wir nutzen die Darstellung aus (f). Da F_Z differenzierbar in z_0 ist, gilt mit der Delta-Methode in $\mathbb{R}$:

$$\sqrt{n}\big(F_Z(z_0 - (a_0 - \hat{a}_n)) - F_Z(z_0)\big) = F'_Z(z_0)\cdot\sqrt{n}(\hat{a}_n - a_0) + o_P(1) \overset{Y_i = Z_i + a_0}{=} F'_Z(z_0)\cdot\sqrt{n}\Big(\frac{1}{n}\sum_{i=1}^n Z_i - 0\Big) + o_P(1)$$

wobei $o_P(1)$ ein Summand ist, der stochastisch gegen 0 konvergiert. Damit ist zusammen mit (e):

$$\begin{aligned}
\sqrt{n}(\tilde{F}_Z(z_0) - F_Z(z_0)) \; &= \; \frac{1}{\sqrt{n}}\sum_{i=1}^n \big(\mathbb{1}_{\{Z_i \le z_0\}} + F'_Z(z_0) Z_i - F_Z(z_0)\big) + o_P(1) \\
&\xrightarrow{D} \; N(0, \sigma^2), \qquad \sigma^2 = \mathrm{Var}\big(\mathbb{1}_{\{Z_1 \le z_0\}} + F'_Z(z_0) Z_1\big)
\end{aligned}$$

(Letzter Schritt ist normaler Zentraler Grenzwertsatz in $\mathbb{R}$).

Aufgabe P2 (Abschätzung eines Supremums).

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, $(\mathcal{X}, d)$ ein metrischer Raum und $X_i : \Omega \to \mathcal{X}$, $i \in \mathbb{N}$ i.i.d. Zufallsvariablen. Sei $P = \mathbb{P}^{X_1}$. Sei $\mathcal{F} = \{f : \mathcal{X} \to \mathbb{R} \text{ messbar}\}$ und $L : \mathcal{X} \to \mathbb{R}$ eine Einhüllende von $\mathcal{F}$. Zeigen Sie zunächst, dass

$$\mathbb{E}^* \sup_{f,g \in \mathcal{F}, P((f-g)^2) \le \delta^2} |\mathbb{G}_n(f - g)| \lesssim J_2\Big(\frac{\delta}{2}, \mathcal{F}, P\Big) + \sqrt{n} P\big[L \mathbb{1}_{\{2L > \sqrt{n} a(\frac{\delta}{2})\}}\big],$$

wobei $a(\delta) = \frac{\delta}{\sqrt{Log N_2(\delta, \mathcal{F}, P)}}$.

Es bezeichne $\hat{F}_n$ die empirische Verteilungsfunktion von X_1 und $F(x) = \mathbb{P}(X \le x)$. Zeigen Sie, dass für $\delta \le 1$:

$$\mathbb{E} \sup_{|F(x) - F(y)| \lesssim \delta^2} \big|\big(\hat{F}_n(x) - F(x)\big) - \big(\hat{F}_n(y) - F(y)\big)\big| \lesssim \frac{1}{\sqrt{n}} \cdot \Big(\delta\sqrt{Log(1/\delta)} + \frac{\sqrt{Log(1/\delta)}}{\delta}\Big)$$

Hinweis: Nutzen Sie ohne Beweis, dass für $\delta \leq 1$ gilt: $\int_0^\delta \sqrt{Log(1/\varepsilon)}d\varepsilon \lesssim \delta\sqrt{Log(1/\delta)}$.

Lösung: Der erste Teil folgt dem Beweis von Satz 1.64 (Donsker) der Vorlesung. Definiere die Klasse

$$\mathcal{G} = \{f - g : f, g \in \mathcal{F}, P((f-g)^2) \leq \delta^2\}.$$

Die Bracketing-Zahlen dieser Klasse können wir folgt abgeschätzt werden: Sind $[l_i, u_i]$, $i = 1, ..., N$, $\frac{\delta}{2}$-Brackets von $\mathcal{F}$, so sind

$$[l_i - u_j, u_i - l_j], \quad i, j \in \{1, ..., N\}$$

δ-Brackets von $\mathcal{G}$, denn:

$$\|(u_i - l_j) - (l_i - u_j)\|_2 \leq \|u_i - l_i\|_2 + \|u_j - l_j\|_2 \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta,$$

und für $f, g \in \mathcal{F}$ gibt es jeweils Brackets $[l_i, u_i], [l_j, u_j]$ mit $l_i \leq f \leq u_i$, $l_j \leq g \leq u_j$, und damit

$$l_i - u_j \leq f - g \leq u_i - l_j,$$

das heißt, $f - g \in [l_i - u_j, u_i - l_j]$.
Damit ist also

$$N_2(\varepsilon, \mathcal{G}, P) \leq N_2(\frac{\varepsilon}{2}, \mathcal{F}, P)^2$$

Es folgt

$$\log N_2(\varepsilon, \mathcal{G}, P) \leq 2 \log N_2(\frac{\varepsilon}{2}, \mathcal{F}, P)$$

und damit

$$\begin{aligned} J_2(\delta, \mathcal{G}, P) &= \int_0^\delta \sqrt{1 \vee N_2(\frac{\varepsilon}{2}, \mathcal{F}, P)} d\varepsilon \leq 2 \int_0^\delta \sqrt{1 \vee \log N_2(\frac{\varepsilon}{2}, \mathcal{F}, P)} d\varepsilon \\ &\overset{u=\varepsilon/2}{=} 4 \int_0^{\delta/2} \sqrt{1 \vee \log N_2(u, \mathcal{F}, P)} du = 4 J_2(\frac{\delta}{2}, \mathcal{F}, P). \end{aligned}$$

Zuletzt ist eine Einhüllende von G gegeben durch $2L$, da für $f, g \in \mathcal{F}$ gilt: $|f - g| \leq |f| + |g| \leq L + L = 2L$.

Es folgt nun aus Lemma 1.67 (Achtung: Auch $a(\delta)$ muss noch entsprechend abgeschätzt werden, da das Lemma auf $\mathcal{G}$ angewandt wird):

$$\begin{aligned} \mathbb{E}^* \sup_{f,g \in \mathcal{F}, P((f-g)^2) \leq \delta^2} |\mathbb{G}_n(f-g)| &= \mathbb{E}^* \sup_{g \in \mathcal{G}} |\mathbb{G}_n(g)| \lesssim J_2(\delta, \mathcal{G}, P) + \sqrt{n} P\big[L \mathbb{1}_{\{2L > \sqrt{n} a(\delta)\}}\big] \\ &\lesssim J_2(\frac{\delta}{2}, \mathcal{F}, P) + \sqrt{n} P\big[L \mathbb{1}_{\{2L > \sqrt{n} a(\delta/2)\}}\big]. \quad (*) \end{aligned}$$

Für die Klasse $\mathcal{F} = \{f_x(y) = \mathbb{1}_{\{y \leq x\}} : x \in \mathbb{R}\}$ erhält man zunächst für $f_x, f_y \in \mathcal{F}$, oBdA $x < y$:

$$P((f_x - f_y)^2) = \mathbb{E}[(\mathbb{1}_{\{X_1 \leq x\}} - \mathbb{1}_{\{X_1 \leq y\}})^2] = \mathbb{E}[\mathbb{1}_{\{x < X_1 \leq y\}}] = F(y) - F(x).$$

Damit ist $P((f-g)^2) \leq \delta^2$ äquivalent zu

$$|F(x) - F(y)| \leq \delta^2,$$

und

$$
\begin{aligned}
\mathbb{E}^* \sup_{f,g\in\mathcal{F}, P((f-g)^2)\le\delta^2} |\mathbb{G}_n(f-g)| &= \mathbb{E}^* \sup_{x,y\in\mathbb{R}, |F(x)-F(y)|\le\delta^2} |\mathbb{G}_n(f_x - f_y)| \\
&= \sqrt{n}\mathbb{E}^* \sup_{x,y\in\mathbb{R}, |F(x)-F(y)|\le\delta^2} |(\hat{F}_n(x)-F(x)) - (\hat{F}_n(y)-F(y))|.
\end{aligned}
$$

Eine Einhüllende ist $L = 1$ (konstante Funktion 1). Zusammen mit der Abschätzung aus (*) folgt:

$$
\begin{aligned}
& \mathbb{E}^* \sup_{x,y\in\mathbb{R}, |F(x)-F(y)|\le\delta^2} |(\hat{F}_n(x)-F(x)) - (\hat{F}_n(y)-F(y))| \\
\lesssim\ & \frac{1}{\sqrt{n}} J_2(\frac{\delta}{2}, \mathcal{F}, P) + P\big[\mathbb{1}_{\{2>\sqrt{n}a(\delta/2)\}}\big] \\
\overset{Markov}{\lesssim}\ & \frac{1}{\sqrt{n}} J_2(\frac{\delta}{2}, \mathcal{F}, P) + \frac{1}{\sqrt{n}a(\delta/2)} \qquad (**)
\end{aligned}
$$

Nun ist bekanntermaßen $N_2(\varepsilon, \mathcal{F}, P) \le \frac{2}{\varepsilon^2}$, mit Hinweis folgt $J_2(\delta, \mathcal{F}, P) \lesssim \delta\sqrt{Log(1/\delta)}$ und außerdem $N_2(\delta, \mathcal{F}, P) \lesssim Log(1/\delta)$. Damit ist

$$
\frac{1}{a(\delta/2)} = \frac{\sqrt{Log N_2(\delta/2, \mathcal{F}, P)}}{\delta/2} \lesssim \frac{\sqrt{Log(1/\delta)}}{\delta},
$$

und

$$
J_2(\frac{\delta}{2}, \mathcal{F}, P) \lesssim \delta\sqrt{Log(1/\delta)}.
$$

Einsetzen in (**) liefert die Behauptung.

Vorlesungswebseite:
`https://ssp.math.uni-heidelberg.de/wt2-ss23/`