# Übungen zur Vorlesung "Mathematische Statistik"

## Blatt 12

**Abgabetermin:** Montag, 22.1.2018, bis 14:00 Uhr im Briefkasten im UG Eckerstraße 1
(Geben Sie auf jedem Lösungsblatt Ihren Namen an.
Sie dürfen maximal zu zweit abgeben.)

### Aufgabe 1 (4 Punkte)

Betrachten Sie die Situation von Aufgabe 11.4. Nun sei aber $p_\xi = \frac{1}{2}\mathbb{1}_{[-1,1]}$. Zeigen Sie, dass für ein festes $x_0 \in [0,1]$ gilt, dass

$$\liminf_{n\to\infty} \inf_{\hat\theta} \sup_{f\in\Sigma(\beta,L)} \mathbb{E}_f \left[ n^{\frac{2\beta}{\beta+1}} |\hat\theta(x_0) - f(x_0)|^2 \right] \geq c',$$

für ein $c' > 0$. Vergleichen Sie die Rate mit der aus Aufgabe 11.4.

### Aufgabe 2 (4 Punkte)

Betrachten Sie das nicht-parametrische Regressionsmodell

$$Y_i = f\left(\frac{i}{n}\right) + \xi_i, \qquad \xi_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1), \quad i = 1,\dots,n,$$

mit $f \in \mathcal{H}(\beta, L)$ wie in Beispiel 4.33 aus der Vorlesung. Vervollständigen Sie dieses Beispiel. Leiten Sie also mit Hilfe der Methode von Assouad eine untere Schranke an das Minimax-Risiko der Schätzung von $f$ in $L_2$ her, die von der Ordnung $n^{-\frac{\beta}{2\beta+1}}$ ist.

Wir betrachten nun eine ganze Klasse von verschiedenen Schätzproblemen und leiten eine abstrakte untere Schranke an die zugehörigen Minimax-Risiken her.
Es sei $\mathbb{P}$ eine Menge von W-Maßen auf einem messbaren Raum $(\mathcal{X}, \mathcal{F})$ und $\mathbb{P}^{(n)}$ die Menge aller $n$-fachen Produktmaße mit identischen marginalen Verteilungen aus $\mathbb{P}$. $\mathbb{P}^{(n)}$ ist also unser Modell. Wir möchten nun ein Funktional $\theta : \mathbb{P} \to \mathbb{R}$ schätzen, also den Wert von $\theta(P)$ falls $P \in \mathbb{P}$ die datengenerierende Verteilung ist. Es sei $l : \mathbb{R}_+ \to \mathbb{R}_+$ eine monoton nicht-fallende Verlustfunktion, dann ist das zugehörige Minimax-Risiko gegeben durch

$$\mathcal{M}_n(\mathbb{P}) := \inf_{\hat\theta_n} \sup_{P\in\mathbb{P}} \mathbb{E}_P \left[ l\left( |\hat\theta_n - \theta(P)| \right) \right].$$

Mit $l(x_0^-)$ bezeichnen wir den linksseitigen Grenzwert von $l$ bei $x_0$. Weiter definieren wir für $P_0, P_1 \in \mathbb{P}$, die *Test-Affinität* $\pi(P_0,P_1) := \inf_{\text{Tests } \phi} \mathbb{E}_{P_0}[\phi] + \mathbb{E}_{P_1}[1-\phi]$, wobei das Infimum über alle messbaren Funktionen $\phi : \mathcal{X} \to [0,1]$ läuft (alle 'randomisierten' Tests). Für Teilmengen $\mathbb{P}_0, \mathbb{P}_1 \subseteq \mathbb{P}$ schreiben wir auch $\pi(\mathbb{P}_0, \mathbb{P}_1) = \sup_{P_0\in\mathbb{P}_0, P_1\in\mathbb{P}_1} \pi(P_0,P_1)$ und wir definieren $\mathbb{P}_{\le t} := \{P \in \mathbb{P} : \theta(P) \le t\}$ und $\mathbb{P}_{\ge t+\Delta} := \{P \in \mathbb{P} : \theta(P) \ge t + \Delta\}$. Mit $\operatorname{conv}(\mathbb{P})$ bezeichnen wir die konvexe Hülle von $\mathbb{P}$, also die Menge aller Konvexkombinationen $\lambda P_0 + (1-\lambda)P_1$, $\lambda \in [0,1]$, $P_0, P_1 \in \mathbb{P}$. Schließlich definieren wir noch für $\Delta \ge 0$ und $\eta \in [0,1)$, die Größen

$$\eta_A^{(n)}(\Delta) := \sup_{t\in\mathbb{R}} \pi\left( \operatorname{conv}\left(\mathbb{P}_{\le t}^{(n)}\right), \operatorname{conv}\left(\mathbb{P}_{\ge t+\Delta}^{(n)}\right) \right),$$

und

$$\Delta_A^{(n)}(\eta) := \sup\{\Delta \ge 0 : \eta_A^{(n)} > \eta\}.$$

**Aufgabe 3** (4 Punkte)

Angenommen die Menge $\mathbb{P}$ is dominiert durch ein $\sigma$-endliches Maß. Zeigen Sie, dass für jedes $\eta \in (0,1)$ und $n \in \mathbb{N}$, gilt

$$\mathcal{M}_n(\mathbb{P}) \quad \geq \quad \frac{\eta}{2} \cdot l\left(\left[\frac{1}{2}\Delta_A^{(n)}(\eta)\right]^-\right),$$

sofern $\Delta_A^{(n)}(\eta) < \infty$, und $\mathcal{M}_n(\mathbb{P}) \ \geq \ \frac{\eta}{2} \cdot \sup_{x \in [0,\infty)} l(x)$, falls $\Delta_A^{(n)}(\eta) = \infty$.

**Hinweis:** Betrachten Sie Mengen $S = \{x \in \mathcal{X}^n : |\hat{\theta}_n(x) - \theta(P)| \geq \Delta\}$, $S_1 = \{x \in \mathcal{X}^n : \hat{\theta}_n(x) \geq t + \Delta, \theta(P) \leq t\}$ und $S_2 = \{x \in \mathcal{X}^n : \hat{\theta}_n(x) < t + \Delta, \theta(P) > t + 2\Delta\}$, so dass $S_j \subseteq S$, für $j = 1, 2$. Verwenden Sie die folgende Identität (ohne Beweis), welche nur gilt wenn $\mathbb{P}_0, \mathbb{P}_1 \subseteq \mathbb{P}$ und wenn $\mathbb{P}$ dominiert ist.

$$\inf_{\text{Tests } \phi} \sup_{\substack{P_0 \in \mathbb{P}_0^{(n)} \\ P_1 \in \mathbb{P}_1^{(n)}}} \mathbb{E}_{P_0}[\phi] + \mathbb{E}_{P_1}[1-\phi] = \sup_{\substack{P_0 \in \mathrm{conv}(\mathbb{P}_0^{(n)}) \\ P_1 \in \mathrm{conv}(\mathbb{P}_1^{(n)})}} \inf_{\text{Tests } \phi} \mathbb{E}_{P_0}[\phi] + \mathbb{E}_{P_1}[1-\phi].$$

Auf dem nächsten Übungszettel werden wir die untere Schranke $\Delta_A^{(n)}(\eta)$ noch weiter vereinfachen. Dafür benötigen wir aber zunächst ein Hilfsresultat. Für W-Maße $P$ und $Q$, definiere die Hellinger-Affinität $\rho(P,Q) = \int \sqrt{pq}\, d\mu$, wobei $p$ und $q$ Dichten bezüglich eines $P$ und $Q$ dominierenden Maßes $\mu$ sind.

**Aufgabe 4** (4 Punkte)

Zeigen Sie die folgenden Relationen.

$$\begin{aligned}
&H^2(P,Q) = 2(1-\rho), \quad \rho(P^n, Q^n) = \rho(P,Q)^n, \quad \pi(P,Q) = 1 - d_{TV}(P,Q), \\
&\pi(P,Q) \leq \rho(P,Q) \leq \sqrt{\pi(P,Q)(2-\pi(P,Q))} \quad \text{und} \\
&g_H(\eta) := \inf\{H(P_0,P_1) : \pi(P_0^n, P_1^n) \leq \eta, P_0, P_1 \in \mathbb{P}\} \geq \sqrt{2\left(1 - (\eta[2-\eta])^{\frac{1}{2n}}\right)},
\end{aligned}$$

für jedes $\eta \in (0,1)$.