Übungen zur Vorlesung "Mathematische Statistik"

Blatt 11

Abgabetermin: Montag, 15.1.2018, 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)

Beweisen Sie die Varshamov-Gilbert-Schranke.

Satz 0.1 (Varshamov-Gilbert). *Es sei $m \geq 8$. Mit $\rho(\omega, \omega') = \sum_{k=1}^{m} \mathbb{1}_{\omega_k \neq \omega'_k}$, $\omega, \omega' \in \Omega = \{0,1\}^m$, bezeichne man die* Hamming-Distanz *zwischen ω und ω'. Dann gibt es eine Teilmenge $\{\omega^{(0)}, \dots, \omega^{(M)}\}$ von Ω, so dass $M \geq 2^{m/8}$, $\omega^{(0)} = (0, \dots, 0)$, und*

$$\rho(\omega^{(j)}, \omega^{(k)}) \geq \frac{m}{8}, \quad \forall\, 0 \leq j < k \leq M.$$

Hinweis: Verwenden Sie die Hoeffding-Ungleichung.

Aufgabe 2 (4 Punkte)

Es seien P und Q zwei W-Maße mit Dichten p und q bezüglich des Lebesguemaßes auf $[0,1]$, so dass $0 < c_1 \leq p(x), q(x) \leq c_2 < \infty$, für alle $x \in [0,1]$. Zeigen Sie, dass die Kullback-Leibler-Divergenz $K(P,Q)$ und die quadrierte L_2-Distanz zwischen den Dichten p und q äquivalent sind in dem Sinne, dass

$$k_1 \int (p(x) - q(x))^2 \, dx \;\leq\; K(P,Q) \;\leq\; k_2 \int (p(x) - q(x))^2 \, dx,$$

für Konstante $k_1, k_2 > 0$ die nicht von P und Q abhängen.

Aufgabe 3 (4 Punkte)

Für $M \geq 1$ und $x \in [0,1]$, setze $\mathcal{H}(x) = -x \log x - (1-x) \log(1-x)$ und $g(x) = x \log M + \mathcal{H}(x)$, mit der Konvention, dass $0 \log 0 = 0$. Beweisen Sie das folgende Lemma.

Lemma 0.2. *Für alle $j_0 \in \{0, 1, \dots, M\}$ und alle reellen Zahlen $p_0, \dots, p_M$, so dass $\sum_{j=0}^{M} p_j = 1$ und $p_j \geq 0$, gilt*

$$g\left(\sum_{j \neq j_0} p_j\right) \geq -\sum_{j=0}^{M} p_j \log p_j.$$

(bitte wenden)

Aufgabe 4 (4 Punkte)

Betrachten Sie das folgende nicht-parametrische Regressionsmodell mit zufälligem Design.

$$Y_i = f(X_i) + \xi_i, \qquad i = 1, \dots, n,$$

wobei die X_i i.i.d. Zufallsvariablen mit Dichte p_X auf $[0,1]$ sind, so dass $p_X(x) \leq p_0 < \infty$, $\forall x \in [0,1]$, die ξ_i sind i.i.d. mit Dichte p_ξ auf $\mathbb{R}$ und die Zufallsvektoren $(X_1, \dots, X_n)$ und $(\xi_1, \dots, \xi_n)$ sind unabhängig. Die Beobachtungen sind hier die Paare (Y_i, X_i), $i = 1, \dots, n$. Der Stichprobenraum ist somit gegeben durch $\mathcal{X} = (\mathbb{R} \times [0,1])^n$. Weiter sei $f \in \Sigma(\beta, L)$ der unbekannte Parameter und $\Sigma(\beta, L)$, mit $\beta > 0$, $L > 0$, die Hölder-Klasse der Funktionen auf $[0,1]$ mit

$$|f^{(\ell)}(x) - f^{(\ell)}(x')| \leq L|x - x'|^{\beta - \ell}, \qquad \forall x, x' \in [0,1],$$

und $\ell = \lfloor \beta \rfloor$.

Zeigen Sie, dass für ein festes $x_0 \in [0,1]$ und für p_ξ mit der Eigenschaft

$$\int \left(\sqrt{p_\xi(y)} - \sqrt{p_\xi(y+t)} \right)^2 dy \; \leq \; p_* t^2, \qquad \forall t \in \mathbb{R},$$

$0 < p_* < \infty$, gilt, dass

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

für ein $c > 0$.