LMU München **Prof. F. Merkl**
Mathematisches Institut
Sommersemester 2018

Übungen zur Mathematischen Statistik
Blatt 7

1. **Satz von Glivenko-Cantelli in 2d.** Beweisen Sie die folgende zweidimensionale Version des Satzes von Glivenko-Cantelli:
Es sei $(Z_n)_{n\in\mathbb{N}}$ eine i.i.d. Folge von Zufallsvektoren mit Werten in $\mathbb{R}^2$ mit Komponenten $Z_n = (X_n, Y_n)$ und $F : \mathbb{R}^2 \to [0,1]$, $F(x,y) = P[X_n \leq x,\ Y_n \leq y]$ (mit $n \in \mathbb{N}$ beliebig und P dem zugrundeliegenden Wahrscheinlichkeitsmaß) die zugehörige Verteilungsfunktion. Es bezeichne
$$\hat{F}_n : \mathbb{R}^2 \to [0,1], \quad F_n(x,y) = \frac{1}{n}\sum_{k=1}^{n} 1_{\{X_k \leq x,\ Y_k \leq y\}}$$
die empirische Verteilungsfunktion. Dann gilt P-fast sicher:
$$\sup_{(x,y)\in\mathbb{R}^2} |\hat{F}_n(x,y) - F(x,y)| \stackrel{n\to\infty}{\longrightarrow} 0.$$

2. Beweisen Sie in der Situation der vorhergehenden Aufgabe, dass die endlichdimensionalen Randverteilungen von $\sqrt{n}(\hat{F}_n - F)$ (also die Verteilungen von $(\sqrt{n}(\hat{F}_n(z_j) - F(z_j)))_{j=1,\ldots,k}$ mit gegebenen $z_1, \ldots, z_k \in \mathbb{R}^2$ für $n \to \infty$ schwach gegen die entsprechenden endlichdimensionalen Randverteilungen eines zentrierten Gaußschen Prozesses konvergieren. Bestimmen Sie die Kovarianz dieses Gaußschen Prozesses in Abhängigkeit von der Verteilung $Q := \mathcal{L}_P(Z_k)$ der Zufallsvektoren Z_k. Welche Kovarianz erhält man im Spezialfall, dass die Z_k uniform auf dem Einheitsquadrat $[0,1]^2$ verteilt sind?

3. **Nichtexistenz supereffizienter Schätzer in 2d.**
   (a) Für $a > 1$ sei die stetige, mit Ausnahme der 1 differenzierbare Funktion
   $$f_a(r) = \left(\log\frac{\max(r,1)}{a}\right)^2, \quad r \geq 0$$
   gegeben. Für $R > 0$ bezeichne $B_R = \{\theta \in \mathbb{R}^2 | \ \|\theta\| < R\}$ die offene Kreisscheibe um 0 mit Radius R. Weiter sei $c_a > 0$ diejenige Konstante, für die $\mathbb{P}_a(d\theta) = \rho_a(\theta)\,d\theta := c_a f_a(\|\theta\|) 1_{B_a}(\theta)\,d\theta$ ein Wahrscheinlichkeitsmaß auf B_a ist; dx bedeutet das 2-dimensionale Lebeguemaß. Zeigen Sie, dass für die Spur der “a priori Informationsmatrix” gilt:
   $$\text{Spur}\, I_{\mathbb{P}} = E_{\mathbb{P}}[\|\nabla \log\rho_a\|^2] = \text{const}\cdot c_a \log a,$$
   z.B. unter Verwendung von Polarkoordinaten. Die Nullmenge, auf der $\log\rho_a$ nicht differenzierbar ist, sollen Sie dabei ignorieren.
   (b) Überzeugen Sie sich, dass $\mathbb{P}_a$ bedingt auf B_1 die Gleichverteilung auf B_1 ist, unabhängig vom Wert von $a > 1$, und dass $\mathbb{P}_a(B_1) = \pi c_a(\log a)^2$ gilt.
   (c) Nun sei X ein Zufallsvektor mit Werten in $\mathbb{R}^2$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P_\theta)$, so dass $X - \theta$ zweidimensional standardnormalverteilt ist, mit einem unbekannten Parameter $\theta \in \mathbb{R}^2$. Angenommen, $\hat{\theta}(X)$ wäre ein supereffizienter Schätzer dafür. Zeigen Sie, dass dann eine Konstante $C > 0$ existiert, so dass für alle $a > 1$ gilt:
   $$\int_{B_a} E_{P_\theta}[\|\hat{\theta}(X) - \theta\|^2]\,\mathbb{P}_a(d\theta) \leq 2 - Cc_a(\log a)^2.$$
   Teilaufgabe (b) kann Ihnen dabei helfen.
   (d) Überzeugen Sie sich davon, dass die van Trees Ungleichung auch hier anwendbar ist, trotz der Nichtdifferenzierbarkeit von ρ_a auf einer Nullmenge. (Alternativ können Sie auch die Dichte ρ_a nahe an den Nichtdifferenzierbarkeitsstellen etwas glätten.)
   (e) Folgern Sie, dass es keinen supereffizienten Schätzer für θ gibt, indem Sie große a betrachten.

Keine Abgabe. Studierende sollen ihre Lösungen in der Übungsstunde 8 präsentieren.