Übungen zur Vorlesung „Mathematische Statistik“

Blatt 12

Abgabetermin: Mittwoch, 23.01.2019, bis 12.00 Uhr im zugehörigen Briefkasten im UG des Mathematischen Instituts, Ernst-Zermelo-Straße 1.
(Geben Sie auf jedem Lösungsblatt Ihren Namen an.
Sie dürfen maximal zu zweit abgeben.)

Aufgabe 1 (4 Punkte)

Seien $X_1, \dots, X_n$ reelle, unabhängige und identisch nach $\mathbb{P}_{X_1}$ verteilte Zufallsvariablen, wobei die Verteilung $\mathbb{P}_{X_1}$ absolutstetig bezüglich des Lebesguemaßes auf $(\mathbb{R}, \mathcal{B})$ sei, d.h. $\mathbb{P}_{X_1} \ll \lambda$. Ferner sei F_n die zugehörige empirische Verteilungsfunktion.

a) Berechnen Sie für alle $x \in \mathbb{R}$ und $1 \le k \le n$ die Wahrscheinlichkeit $\mathbb{P}\big(F_n(x) = \frac{k}{n}\big)$.

b) Wie groß ist die Wahrscheinlichkeit, dass F_n einen Sprung der Höhe $\frac{2}{n}$ hat?

c) Seien nun speziell $X_1, \dots, X_n$ unabhängig und identisch normalverteilt, d.h. es liege ein n-faches Gaußmodell $\mathscr{M}^{\otimes n}$ mit $\mathscr{M} = (\mathbb{R}, \mathcal{B}(\mathbb{R}), (\mathbb{P}_\vartheta = \mathrm{N}(\mu, \sigma^2) : \vartheta = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times (0, \infty)))$ zugrunde. Zeigen Sie, dass die Größe $D_n(F)$ aus Definition 4.2.13 der Vorlesung in diesem Fall unabhängig von $\overline{X}$ und $s^2(X)$ ist.

   Hinweis: Eine Aufgabe von Übungsblatt 8 kann hierzu sehr hilfreich sein.

Aufgabe 2 (4 Punkte)

Sei $\mathcal{F} := \{F_{n,x} \mid x \in \mathbb{R}^n\}$ die Menge aller empirischen Verteilungsfunktionen, d.h.

$$F_{n,x} : \mathbb{R} \to [0,1], \quad F_{n,x}(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{(-\infty,t]}(x_i).$$

a) Zeigen Sie, dass die Abbildung $F_n : (\mathbb{R}^n, \mathcal{B}^{\otimes n}) \to (\mathcal{F}, \mathcal{F} \cap \mathcal{B}^{\mathbb{R}})$ mit $F_n(x) = F_{n,x}$ messbar ist.

   Hinweise: $\mathcal{B}^{\mathbb{R}}$ ist die σ-Algebra auf dem Raum $\mathbb{R}^{\mathbb{R}}$ der reellwertigen Funktionen, die durch die Projektionen $(\pi_t)_{t \in \mathbb{R}}$ mit $\pi_t(f) = f(t)$ erzeugt wird.

   Zeigen Sie zunächst die folgende Aussage: Seien $I \neq \emptyset$ eine Indexmenge, $(\Omega, \mathcal{A})$, $(\Omega', \mathcal{A}')$ und $(\Omega_i, \mathcal{A}_i)$ Messräume für alle $i \in I$, $X_i : \Omega' \to \Omega_i$ messbare Abbildungen und $\mathcal{A}' = \sigma(X_i, i \in I)$. Dann ist eine Abbildung $Y : \Omega \to \Omega'$ $\mathcal{A}$-$\mathcal{A}'$-messbar genau dann, wenn $X_i \circ Y$ $\mathcal{A}$-$\mathcal{A}_i$-messbar ist für alle $i \in I$.

b) Sei $O : \mathbb{R}^n \to \mathbb{R}^n$ die Ordnungsstatistik aus Aufgabe 4 von Übungsblatt 11. Zeigen Sie, dass gilt $\sigma(O) = \sigma(F_n)$.

   Bemerkung: Da die Ordnungsstatistik vollständig und suffizient ist, folgt aus der Gleichheit der von O und F_n erzeugten σ-Algebren, dass auch F_n eine vollständige und suffiziente Statistik für $\Theta = \{\vartheta \mid \vartheta \text{ stetiges oder diskretes Wahrscheinlichkeitsmaß auf } (\mathbb{R}, \mathcal{B})\}$ ist.

(bitte wenden)

Aufgabe 3 (4 Punkte)

Seien $X_1, \ldots, X_n$ reelle, unabhängige und identisch nach $\mathbb{P}_{X_1}$ verteilte Zufallsvariablen und $\mathcal{F}$ eine Klasse von Funktionen $f : \mathbb{R} \to \mathbb{R}$, so dass für alle $f \in \mathcal{F}$ gilt $\mathbb{E}[|f(X_1)|] < \infty$. Für Funktionen $l, u \in \mathscr{L}^1(\mathbb{P}_{X_1})$ bezeichnet man die Menge von Funktionen

$$[l, u] := \{f : \mathbb{R} \to \mathbb{R} \mid l(x) \leq f(x) \leq u(x) \text{ für alle } x \in \mathbb{R}\}$$

als zugehörige *Klammer*. Insbesondere heißt $[l, u]$ eine ϵ-*Klammer*, falls $\mathbb{E}[|u(X_1) - l(X_1)|] \leq \epsilon$. Zeigen Sie: Falls für jedes $\epsilon > 0$ endlich viele ϵ-Klammern existieren, die $\mathcal{F}$ überdecken, so gilt die *Glivenko-Cantelli-Eigenschaft*

$$\sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^{n} f(X_i) - \mathbb{E}[f(X_1)] \right| \longrightarrow 0 \quad \mathbb{P}\text{-fast sicher für } n \to \infty.$$

Aufgabe 4 (4 Punkte)

Sei $X = (X_1, \ldots, X_d)^\top \sim \mathrm{N}(\mathbf{0}, \Sigma^2)$ d-dimensional normalverteilt mit Mittelwertvektor $\mu = \mathbf{0} = (0, \ldots, 0)$ und Covarianzmatrix $\Sigma^2 = \mathrm{Id}_d - \sqrt{\pi}\sqrt{\pi}^\top$, wobei Id_d die $d \times d$-Einheitsmatrix sei und $\pi = (\pi_1, \ldots, \pi_d)^\top$ ein Spaltenvektor mit $\sum_{i=1}^d \pi_i = 1$ ($\sqrt{\pi}$ ist dann komponentenweise zu verstehen, d.h. $\sqrt{\pi} = \big(\sqrt{\pi_1}, \ldots, \sqrt{\pi_d}\big)^\top$).
Zeigen Sie, dass $\sum_{i=1}^d X_i^2 \sim \chi^2_{d-1}$.