Übungen zur Vorlesung "Mathematische Statistik"

Blatt 1

Abgabetermin: Donnerstag, 26.10.2017, in der ersten Übungseinheit
(Geben Sie auf jedem Lösungsblatt Ihren Namen an.
Sie dürfen nur einzeln abgeben.)

Erinnern Sie sich, dass ein *statistisches Modell* gegeben ist durch ein Tripel $(\mathcal{X}, \mathcal{F}, \mathscr{P})$, wobei $\mathscr{P} = \{P_\theta : \theta \in \Theta\}$ eine Menge von Wahrscheinlichkeitsmaßen auf dem messbaren Raum $(\mathcal{X}, \mathcal{F})$ ist. Dabei heißt $\mathcal{X}$ *Stichprobenraum* und Θ *Parameterraum*. Die Abbildung $\theta \mapsto P_\theta$ wird auch *Parametrisierung* des Modells genannt. Ist diese Abbildung injektiv, so sprechen wir auch von einer *identifizierten* oder *identifizierbaren* Parametrisierung.

Aufgabe 1 (4 Punkte)

Angenommen wir beobachten n Realisierungen von unabhängig und identisch verteilten p-dimensionalen Zufallsvektoren $Y_1, \ldots, Y_n$, wobei Y_1 gegeben ist durch $Y_1 = LF + u$. Hier ist L eine nicht zufällige $p \times k$ Matrix $(k < p)$ von unbekannten Parametern, $F \sim \mathcal{N}_k(0, I_k)$ und $u \sim \mathcal{N}_p(0, \sigma^2 I_p)$, F und u sind unabhängig und $\sigma^2 \in (0, \infty)$ ist ebenfalls ein unbekannter Parameter. Den Wert von k setzen wir als bekannt voraus. Somit können wir den Parameterraum wählen als $\Theta = \mathbb{R}^{p \times k} \times (0, \infty)$. Das vorliegende Modell nennt man übrigens auch *Faktormodell*.

a) Beschreiben Sie den Stichprobenraum und die zugehörige Parametrisierung des Modells.

b) Ist die gewählte Parametrisierung identifiziert?

c) Betrachten Sie stattdessen

$$\Theta = \{(\Sigma, \sigma^2) : \sigma^2 \in (0, \infty), \Sigma \in \mathbb{R}^{p \times p}, \Sigma' = \Sigma, \mathrm{Rang}(\Sigma) \leq k, \Sigma \text{ ist positiv semidefinit}\}$$

und die Parametrisierung

$$(\Sigma, \sigma^2) \mapsto P_{\Sigma, \sigma^2} := \mathcal{N}_p(0, \Sigma + \sigma^2 I_p)^{\otimes n}.$$

Zeigen Sie, dass diese Parametrisierung identifiziert ist.

d) Warum ist es in der Statistik wichtig identifizierte Parametrisierungen zu betrachten?

Aufgabe 2 (4 Punkte)

Seien $X_1, \ldots, X_n$ i.i.d. Zufallsvariable mit einer $\mathcal{N}(0, \sigma^2)$-Verteilung mit $0 < \sigma^2 < \infty$. Die gemeinsame Verteilung der Daten ist also gegeben durch $P_{\sigma^2} = \mathcal{N}_n(0, \sigma^2 I_n)$. Berechnen Sie den Maximum-Likelihood-Schätzer $\hat{\sigma}_n^2$ für σ^2. Zeigen Sie, dass für $\varepsilon > 0$

$$\sup_{\sigma^2 \in (0,\infty)} \lim_{n \to \infty} P_{\sigma^2}(|\hat{\sigma}_n^2 - \sigma^2| > \varepsilon) = 0.$$

Zeigen Sie weiter, dass

$$\sup_{\sigma^2 \in (0,\infty)} P_{\sigma^2}(|\hat{\sigma}_n^2 - \sigma^2| > \varepsilon) = 1.$$

Hier unterscheidet sich das punktweise asymptotische Verhalten des Schätzers dramatisch vom Verhalten in endlichen Stichproben (beliebiger Größe)!