UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

**Statistik II**
Prof. Dr. Rainer Dahlhaus
Wintersemester 2021/2022

## 3. Übungsblatt

| Aufgabe 7 | Aufgabe 8 | Aufgabe 9 | Aufgabe 10 | Summe: |
|---|---|---|---|---|
| | | | | |

Übungsgruppe (Mo oder Di): Tutorin: Maybritt Schillinger
Namen:

### Aufgabe 7 (4 Punkte).

Seien $X_1, ..., X_n$ unabhängig exponentialverteilte Zufallsvariablen mit unbekanntem Parameter $\lambda_0 \in [a, b] \subseteq \mathbb{R}$.

(a) Bestimmen Sie den Maximum-Likelihood-Schätzer $\hat{\lambda}_n$ für $\lambda_0$ und zeigen Sie, dass er asymptotisch normal ist.

(b) Wir beobachten 1000 Realisierungen einer Exponentialverteilung mit unbekanntem Parameter $\lambda_0 \in [a, b]$. Der Durchschnitt aller Beobachtungen ist 0.1955922. Wir wollen mit diversen Statistiken die Hypothese $H_0 : \lambda_0 = 5$ testen. Führen Sie dazu den

(i) Likelihood-Ratio-Test,

(ii) Wald-Test,

(iii) Score-Test

aus und entscheiden Sie anhand der Daten, ob die Hypothese $H_0$ zum Niveau $\alpha = 0.05$ in den jeweiligen Fällen abgelehnt wird.

### Aufgabe 8 (2 Punkte + 2 Bonuspunkte).

Sei $(\Omega, \mathcal{A})$ ein messbarer Raum und $(\mathbb{P}_\theta)_{\theta\in\Theta}$, $\Theta \subseteq \mathbb{R}^d$ kompakt, eine Familie von Wahrscheinlichkeitsmaßen darauf. Weiter seien i.i.d. Zufallsvariablen $X_1, ..., X_n : \Omega \to \mathbb{R}$ gegeben. Für die negative log-Likelihood $L_n(\theta)$ von $X_1, ..., X_n$ zum Parameter $\theta$ nehmen wir an, dass

$$\begin{aligned} \nabla^2 L_n(\theta_{0,n}) &\quad\to\quad I(\theta_0) \\ \sqrt{n}\nabla L_n(\theta_{0,n}) &\quad\overset{d,\mathbb{P}_{\theta_{0,n}}}{\to}\quad \mathcal{N}(0, I(\theta_0)), \end{aligned}$$

wobei $I(\theta_0)$ eine invertierbare Matrix ist, für einen ausgezeichneten Punkt $\theta_0 \in \mathrm{Int}(\Theta)$ und jede Folge $\theta_{0,n} \to \theta_0$ gilt. Der Ausdruck „$d$, $\mathbb{P}_{\theta_{0,n}}$“ bezeichnet hierbei die Konvergenz in Verteilung bezüglich $\mathbb{P}_{\theta_{0,n}}$, d.h., die Verteilung von $(X_1, ..., X_n)$ wird bezüglich $\mathbb{P}^n_{\theta_{0,n}}$ betrachtet. Sei

$$R_n := n\nabla L_n(\theta_0)I(\theta_0)^{-1}\nabla L_n(\theta_0)$$

die Score-Statistik für $\theta_0$. Wir definieren den Test $\varphi_n := \mathbb{1}_{\{R_n \geq \chi^2_{d,1-\alpha}\}}$, der die Hypothese $H_0 : \theta = \theta_0$ ablehnt, wenn $\varphi_n = 1$, wobei $\chi^2_{d,1-\alpha}$ so gegeben ist, dass $\mathbb{P}(\xi \geq \chi^2_{d,1-\alpha}) = \alpha$ für $\xi \sim \chi^2_d$. Zeigen Sie, dass

(a) $\varphi_n$ asymptotisch ein Test zum Niveau $\alpha$ ist, also $\lim_{n\to\infty} \mathbb{P}^n_{\theta_0}(\varphi_n = 1) = \alpha$.

(b) für $\theta_{0,n} = \theta_0 + n^{-\frac{1}{2}}h$, $h \in \mathbb{R}^d$ fest:

$$\alpha < \lim_{n\to\infty} \mathbb{P}^n_{\theta_{0,n}}(\varphi_n = 1) < 1.$$

(c) für $\sqrt{n}(\theta_{0,n} - \theta_0) \to 0$:

$$\lim_{n\to\infty} \mathbb{P}^n_{\theta_{0,n}}(\varphi_n = 1) = \alpha$$

(d) für $\sqrt{n}(\theta_{0,n} - \theta_0) \to \infty$:

$$\lim_{n\to\infty} \mathbb{P}^n_{\theta_{0,n}}(\varphi_n = 1) = 1$$

*Bemerkung: Wir haben gezeigt, dass $\frac{1}{\sqrt{n}}$ die Rate ist, mit der Hypothese und Alternative unterschieden werden kann: Ist die Alternative näher an der Hypothese (Teilaufgabe (c)), so wird der Test nicht abgelehnt. Falls aber die Entfernung zwischen Alternative und Nullhypothese größer wird (Teilaufgabe (d)), führt der Test immer zu einer Ablehnung.*

**Aufgabe 9 (4 Punkte).**

In dieser Aufgabe betrachten wir ein allgemeines lineares Modell

$$Y = X\beta + \varepsilon,$$

wobei $X \in \mathbb{R}^{n\times d}$ mit $\text{Rang}(X) = d, \beta \in \mathbb{R}^d$ und $\varepsilon = (\varepsilon_1, ..., \varepsilon_n)$ mit unabhängigen und normalverteilten $\varepsilon_i \sim N(0, \sigma^2)$, $\sigma > 0$.

(a) Bestimmen Sie den Maximum-Likelihood-Schätzer $\hat{\beta}$ für $\beta$.

Wir wollen zusammengesetzte Hypothesen (mit Spezifikation durch Gleichungen) der Form $H_0 : H\beta = 0$, $\text{Rang}(H) = k$, testen und verwenden die $F$-Statistik (s. Statistik 1).

(b) Unter der Hypothese ist es möglich den $F$-Test asymptotisch mit den Quantilen der $\chi^2$-Verteilung auszuwerten. Zeigen Sie zunächst, dass die $F_{k,n-d}$-Verteilung durch eine $\chi^2_k$-Verteilung angenähert werden kann, also $F_{k,n-d} \xrightarrow{\mathcal{D}} k^{-1}\chi^2_k$. Folgern Sie dann, dass $F_{k,n-d,1-\alpha} \to k^{-1}\chi^2_{k,1-\alpha}$, $\alpha > 0$, für die Quantile gilt.

**Aufgabe 10 (2 Punkte + 2 Bonuspunkte).**

Sei $\{\mathbb{P}_{n,\theta} : \theta \in \Theta\}$ eine Familie von Wahrscheinlichkeitsmaßen auf $\mathbb{R}^n$ (bzgl. der Borel'schen-$\sigma$-Algebra), $\Theta \subseteq \mathbb{R}^n$. Für eine Folge von Schätzern $(\check{\theta}_n)_{n\in\mathbb{N}}$ und $\theta_0 \in \Theta$ gelte unter $\mathbb{P}_{n,\theta_0}$:

$$\|\check{\theta}_n - \theta_0\| = O_p(n^{-1/2}).$$

Für eine beliebige messbare Funktion $L_n : \mathbb{R}^n \times \Theta \to \mathbb{R}$ sei $\hat{\theta}_n := \arg\min_{\theta\in\Theta} L_n(\theta)$. Wir setzen die Gültigkeit der Bedingungen (A1) und (A2) voraus. Zeigen Sie, dass

$$\tilde{\theta}_n := \check{\theta}_n - \nabla^2 L_n(\check{\theta}_n)^{-1} \nabla L_n(\check{\theta}_n)$$

unter $\mathbb{P}_{n,\theta_0}$,

$$\tilde{\theta}_n = \hat{\theta}_n + o_p(n^{-1/2})$$

erfüllt, und $\sqrt{n}(\check{\theta}_n - \theta_0)$ den gleichen asymptotisch normalen Grenzwert hat wie $\sqrt{n}(\hat{\theta}_n - \theta_0)$.
*Bemerkung: Oft ist es nicht möglich, einen Maximum-Likelihood-Schätzer explizit zu berechnen. Wir haben gezeigt, dass bereits ein asymptotisch normaler Schätzer und das Wissen um die Existenz eines asymptotisch effizienten Maximum-Likelihood-Schätzers genügt, um einen asymptotisch effizienten Schätzer zu konstruieren.*

---

**Abgabe:**
In Zweiergruppen, bis spätestens Freitag, den **12.11.2021, 11:00 Uhr**.
(Der Zettelkasten ist im 1. OG, INF 205, vor dem Dekanat, Zettelkasten-Nr. 14.)

**Homepage der Vorlesung:**
`https://ssp.math.uni-heidelberg.de/stat2-ws2122/index.html`