Einführung in die
Wahrscheinlichkeitstheorie und Statistik
Prof. Dr. Jan Johannes
Sergio Brenner Miguel
Wintersemester 2020/21

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

5. Übungsblatt

Aufgabe 17 (Neyman-Pearson-Lemma für stetige Verteilungen, 4 Punkte).

Beweisen Sie das Neyman-Pearson-Lemma aus der Vorlesung:
Sei $(\mathbb{R}^n, \mathscr{B}^n, (\mathbb{P}_0, \mathbb{P}_1))$ ein (binäres) stetiges statistisches Experiment. Für das Testproblem der einfachen Nullhypothese $H_0 : \{\mathbb{P}_0\}$ gegen die einfache Alternativhypothese $H_1 : \{\mathbb{P}_1\}$ ist jeder Neyman-Pearson-Test $\varphi = \mathbb{1}_{A_k}$ mit kritischem Wert $k \in \mathbb{R}^+$ und Ablehnbereich

$$A_k = \{x \in \mathbb{R}^n \mid \mathbb{f}_1(x) \geq k\mathbb{f}_0(x)\}$$

ein **bester Test zum Niveau** $\mathbb{P}_0(A_k) \in [0, 1]$.

Aufgabe 18 (Neyman-Pearson-Lemma, Konfidenzbereiche, 4 = 4 × 1 Punkte).

Eine Forschungsgruppe untersucht die Halbwertszeit von ^{60}Co. Die bisher vermutete Halbwertszeit beträgt $\lambda_0 = 5.2714$ Jahre. Die Forschenden beobachten nun in einem Experiment eine Halbwertszeit von $X = 13.0$ Jahren.
In ihrem Modell nehmen die Forscher an, dass die Halbwertszeit exponentialverteilt ist mit Parameter $\frac{1}{\lambda}$ (wobei $\lambda > 0$), d.h. $X \sim \mathrm{Exp}_{\frac{1}{\lambda}} =: \mathbb{P}_\lambda$. Die Forschenden wollen basierend auf ihrer Messung nachweisen, dass die wahre Halbwertszeit größer als die bisher vermutete ist.

(a) Sei zunächst $\lambda_1 > \lambda_0 > 0$ fest gewählt. Formulieren Sie den Neyman-Pearson-Test $\varphi : \mathbb{R} \to \{0, 1\}$ für die $\mathbb{P}_\lambda$-Verteilung zu den Hypothesen

$$H_0 : \lambda = \lambda_0 \quad \text{gegen} \quad H_1 : \lambda = \lambda_1$$

zum Niveau $\alpha \in (0, 1)$. Vereinfachen Sie dann diesen Test mittels der Technik des monotonen Likelihood-Quotienten.

(b) Begründen Sie, warum φ aus (a) auch ein gleichmäßig bester Test zum Niveau α für die Hypothesen

$$H_0 : \lambda \leq \lambda_0 \quad \text{gegen} \quad H_1' : \lambda > \lambda_0$$

ist.

(c) Die Forschende wollen den peinlichen Fehler, falsch zu liegen, nur mit einer Wahrscheinlichkeit von 5% begehen. Werden sie sich auf Basis ihrer Beobachtung X und dem Test φ dazu entscheiden, ihre Ergebnisse zu publizieren?

(d) Geben Sie einen gleichmäßig besten Konfidenzbereich zum Niveau $1 - \alpha$ für die falschen Parameter $\mathcal{F}_\lambda = (0, \lambda)$ an.

Aufgabe 19 (Berechnung MLS, 4 = 2 + 2 Punkte).

Seien $X_1, ..., X_n$ identisch verteilte Zufallsvariablen mit gemeinsamer Produktdichte $\mathbb{f}(x_1, \ldots, x_n) = \prod_{i=1}^n \mathbb{f}_i(x_i)$, wobei $\mathbb{f}_i$ die Wahrscheinlichkeitsdichte von X_i sei. Berechnen Sie den zugehörigen Maximum-Likelihood-Schätzer $\hat{\theta}_n$ für θ, falls

(a) $X_1 \sim \mathrm{N}_{(\mu,\sigma^2)}$ normalverteilt mit Parametervektor $\theta = (\mu, \sigma^2)$,
Hinweis: Sie müssen nicht nachrechnen, dass $\hat{\theta}_n$ ein globales Minimum/Maximum ist.

(b) $X_1 \sim \mathrm{Bin}_{(m,p)}$ binomialverteilt mit Parameter $\theta = p \in (0, 1)$, $m \in \mathbb{N}$ sei hier bekannt.

Aufgabe 20 (Beispiel: Nichtexistenz MLS, 4 = 3 + 1 Punkte).

Wir wollen die durchschnittliche Anzahl der Zerfälle $\lambda > 0$ pro Stunde von ^{60}Co bestimmen. Wir wissen, dass diese Anzahl Poi(λ)-verteilt ist. Wir haben n-mal das exakt gleiche ^{60}Co-Präparat beschafft (d.h. wir können annehmen, dass die gemeinsame Wahrscheinlichkeitsdichte durch die Produktdichte gegeben ist) und führen nun mit jedem Präparat den folgenden Versuch durch ($i = 1, ..., n$):
Wir halten das Präparat eine Stunde an den Geigerzähler und notieren die Anzahl der Zerfälle X_i, die er uns anzeigt. Leider entdecken wir erst am Ende unserer Versuchsreihe, dass der Geigerzähler defekt ist und nur noch erkennt, ob in der Stunde überhaupt ein Zerfall geschehen ist oder nicht (d.h. wir beobachten nur $X_i = 0$ oder $X_i = 1$).
Wir müssen also gezwungenermaßen mit unseren Beobachtungen $X_1, ..., X_n$ auskommen. Auch hier können wir annehmen, dass die gemeinsame Dichte der Produktdichte entspricht.

(a) Zeigen Sie, dass der Maximum-Likelihood-Schätzer $\hat{\lambda}_n$ für λ basierend auf $X_1, ..., X_n$ durch

$$\hat{\lambda}_n = -\log(1 - \overline{X_n})$$

gegeben ist, sofern er existiert. Hierbei bezeichnet $\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$.
Hinweis: Ist $X \sim Bin_{(1,p)}$, so ist die Zähldichte gegeben durch $\mathbb{p}_X(k) = p^k \cdot (1-p)^{1-k}$ für $k \in \{0, 1\}$.

(b) Zeigen Sie, dass der Maximum-Likelihood-Schätzer mit positiver Wahrscheinlichkeit nicht existiert.

Abgabe:
In Zweiergruppen, bis spätestens Montag, den **14. Dezember 2020, 09:00 Uhr**.

Homepage der Vorlesung:
`https://sip.math.uni-heidelberg.de/vl/ews-ws20/`