Blatt Nr. 14 Prof. F. Merkl

Tutorium zur Stochastik für Lehramt
SS 2016

Aufgabe 1

Ein Beispiel zum Neyman–Pearson–Lemma. In einer schriftlichen Prüfung, an der sehr viele Kandidaten teilnahmen, gab es drei multiple–choice–Aufgaben. Aufgabe Nr. n wurde von folgendem Anteil p_n der Kandidaten richtig beantwortet:

n	1	2	3
p_n	10%	20%	40%

Wir wollen testen, ob der Kandidat Fritz signifikant besser (z. B. “doppelt so gut”) war. Dazu machen wir die Modellannahme, dass Fritz eine (unbekannte) Wahrscheinlichkeit q_n hatte, die Aufgabe Nr. n richtig zu beantworten, unabhängig für die verschiedenen Aufgaben.
Nullhypothese: $q_n = p_n$ für $n = 1, 2, 3$.
Alternativhypothese: $q_n = 2p_n$ für $n = 1, 2, 3$.

a) Konstruieren Sie hierzu einen randomisierten Test zum 10%–Niveau mit maximaler Macht, der als Eingabedaten verwenden soll, *welche* Aufgaben Fritz richtig beantwortete. Welche Macht hat dieser Test?

b) Konstruieren Sie einen weiteren randomisierten Test zum 10%–Niveau mit maximaler Macht, der aber nur verwenden darf, *wie viele* Aufgaben Fritz richtig beantwortete. Berechnen Sie auch die Macht dieses Tests. Vergleichen Sie die Macht der beiden Tests.

c) Fritz hat genau die ersten beiden Aufgaben richtig beantwortet. Wie lautet das Ergebnis der beiden Tests? Interpretieren Sie die Resultate!

Aufgabe 2

Gegeben sei ein Datensatz $\omega = (\{A_i, B_i\})_{i=1,\dots,n}$ aus n zufälligen ungerichteten Sehnen im Einheitskreis. Betrachten Sie folgende Hypothesen:

H_0: Die Sehnen wurden unabhängig voneinander nach dem Verfahren von Übungsblatt 5, Aufgabe 2 gewonnen.

H_A: Die Sehnen wurden unabhängig voneinander nach dem Verfahren von Übungsblatt 5, Aufgabe 3 gewonnen.

Entwickeln Sie eine Klasse von Tests, die die Nullhypothese H_0 gegen die Alternative H_A mit verschiedenen Signifikanzniveaus testen. Die Tests sollen im Sinne des Neyman-Pearson Lemmas maximale Macht besitzen. Wie lautet eine geeignete Teststatistik?

Aufgabe 3

Ein Modellproblem zur Bayesschen Statistik. In einer Schachtel befinden sich n Kugeln, davon sind eine unbekannte Anzahl K rot, die übrigen $n-K$ sind grün. Wir setzen an, dass $K=k$ mit der "a–priori"–Wahrscheinlichkeit $p_k > 0$ vorliegt, $k = 0, \ldots, n$. Nun werden m Kugeln aus der Schachtel zufällig mit Zurücklegen entnommen; darunter sind L rote Kugeln. Das Wahrscheinlichkeitsmass P_m beschreibe dieses Zufallsexperiment. Wir wollen die relative Häufigkeit $\frac{K}{n}$ mit Hilfe der beobachteten relativen Häufigkeit $\frac{L}{m}$ schätzen:

a) Berechnen Sie die "a–posteriori"–Verteilung von K, d. h. $P_m[K=k|L=l]$.

b) Zeigen Sie:

$$\exists \alpha > 0 \; \exists C > 0 \; \forall m \in \mathbb{N} \; \forall l \in \{0, \ldots, m\}: \; P_m\left[\frac{L}{m} - \frac{K}{n} \geq \frac{1}{n} \;\middle|\; L=l\right] \leq Ce^{-\alpha m}$$

Hinweis: Es gelte $\frac{l}{m} - \frac{k}{n} \geq \frac{1}{n}$. Bemerken Sie zunächst

$$P_m[K=k|L=l] = P_m[K=k+1|L=l] \frac{p_k}{p_{k+1}} f\left(\frac{k}{n}, \frac{k+1}{n}, \frac{l}{m}\right)^m ,$$

wobei $f(p,1,1) := p$ und

$$f(p,p',\pi) := \left(\frac{p}{p'}\right)^{\pi} \left(\frac{1-p}{1-p'}\right)^{1-\pi} \quad \text{für } 0 < p' < 1.$$

Beweisen Sie dann $f(p,p',\pi) \leq f(p,p',p') < 1$ für $0 \leq p < p' \leq \pi \leq 1$. Leiten Sie hieraus eine in m exponentiell abfallende obere Schranke für $P_m[K=k|L=l]$ her.

c) Folgern Sie:

$$\exists \alpha > 0 \; \exists C > 0 \; \forall m \in \mathbb{N} \; \forall l \in \{0, \ldots, m\}: \; P_m\left[\left|\frac{L}{m} - \frac{K}{n}\right| \geq \frac{1}{n} \;\middle|\; L=l\right] \leq Ce^{-\alpha m}$$

Hinweis: Verwenden Sie ein Symmetrieargument, das rote und grüne Kugeln vertauscht.