UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

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

7. Übungsblatt

Aufgabe 20	Aufgabe 21	Aufgabe 22	Summe:

Übungsgruppe: dienstags Tutorin: Maybritt Schillinger
Namen:

Aufgabe 20 (4 Punkte).

Sei $\mathbb{P}_\beta = \mathrm{Par}(\beta, 1)$, $\beta > 2$, eine Familie von Pareto-Verteilungen mit Lebesgue-Dichten $p_\beta(x) = \frac{\beta}{x^{\beta+1}} \mathbb{1}_{\{x>1\}}$ auf $\mathbb{R}$. Wir beobachten nun $X = (X_1, ..., X_n) \sim \mathbb{P}_{n,\beta} := \mathbb{P}_\beta^n$ und beschäftigen uns mit folgendem Testproblem:

$$H : \beta \leq \beta_0 := 3 \text{ gegen } K : \beta > \beta_0.$$

Wir nutzen die Teststatistik

$$\phi_n(X) = \begin{cases} 1, & -\overline{X}_n > d_n, \\ 0, & -\overline{X}_n \leq d_n. \end{cases}$$

(a) Finden Sie $(d_n)_{n\in\mathbb{N}}$ derart, dass der Test ϕ_n asymptotisch das Niveau $\alpha > 0$ einhält.

(b) Bestimmen Sie die asymptotisch relative Effizienz von ϕ_n bzgl. des optimalen Tests ϕ_n^*.

(c) Wie viele Beobachtungen benötigen Sie, damit ϕ_n dieselbe asymptotische Güte aufweisen kann wie $\phi_{n'}^*$ mit $n' = 100$ Beobachtungen?

Hinweis: Sie dürfen ohne Beweis verwenden, dass $(\mathbb{P}_\beta)_{\beta>2}$ *LAN bzgl.* $\delta_n = \frac{1}{\sqrt{n}}$, $I(\beta_0) = \frac{1}{\beta_0^2}$ *und der zentralen Folge*

$$Z_n := \frac{I(\beta_0)^{-1}}{\sqrt{n}} \sum_{i=1}^n \left(-\log(X_i) + \frac{1}{\beta_0} \right)$$

ist. Weiter haben wir $\mathbb{E}[X_i \log(X_i)] = \frac{\beta_0}{(\beta_0-1)^2}$.

Aufgabe 21 (4 Punkte).

Seien $X_1, ..., X_n$ i.i.d. Zufallsvariablen mit Lebesgue-Dichte f und Verteilung F.

(a) Zeigen Sie, dass die gemeinsame Dichte von $X_{r:n}$ und $X_{s:n}$, $r < s$, gegeben ist durch:

$$f(x, y) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} f(x)f(y) \cdot F^{r-1}(x)(F(y) - F(x))^{s-r-1}(1 - F(y))^{n-s} \mathbb{1}_{\{x<y\}}.$$

Hinweis: Sie dürfen davon ausgehen, dass $(\sum_{i=1}^n \mathbb{1}_{\{X_i<x\}}, \sum_{i=1}^n \mathbb{1}_{\{x<X_i<y\}}, \sum_{i=1}^n \mathbb{1}_{\{y<X_i\}})$, $x < y$, *multinomialverteilt mit den Parametern* $(n, (F(x), F(y) - F(x), 1 - F(y)))$ *ist.*

(b) Bestimmen Sie die Dichte der „sample range“ $R := X_{n:n} - X_{1:n}$.

Aufgabe 22 (6 Punkte).

Seien $X_1, ..., X_n$ i.i.d. Zufallsvariablen einer stetigen Verteilung mit Dichte f und Verteilungsfunktion F. Wir definieren mit $R_n(X) := (R_{ni}(X))_{i=1,...,n}$ den Rang-Vektor und untersuchen folgende lineare Rang-Statistik:

$$T_n := \sum_{i=1}^{n} c_{ni} \cdot a_{n,R_{ni}}, \tag{1}$$

wobei $(c_{ni})_{i=1,...,n}$ und $(a_{ni})_{i=1,...,n}$ deterministisch und reell-wertig sind. Wir nehmen weiter an, dass für eine deterministische, messbare Funktion $\phi : [0,1] \to \mathbb{R}$ gerade $a_{ni} = \phi\left(\frac{i}{n+1}\right)$ gilt.

(a) Zeigen Sie, dass

$$\mathbb{E}(T_n) = n \cdot \bar{c}_n \cdot \bar{a}_n \text{ und } \operatorname{Var}(T_n) = \frac{1}{n-1} \sum_{i=1}^{n} (c_{ni} - \bar{c}_n)^2 \cdot \sum_{i=1}^{n} (a_{ni} - \bar{a}_n)^2,$$

wobei $\bar{c}_n := \frac{1}{n} \sum_{i=1}^{n} c_{ni}$ und $\bar{a}_n := \frac{1}{n} \sum_{i=1}^{n} a_{ni}$ für jeweils $(c_{ni})_{i=1,...,n}$ und $(a_{ni})_{i=1,...,n}$.
Hinweis: Sie benötigen keine spezifische Kenntnis über a_{ni}.

(b) Sei $n = n_1 + n_2$ und T_n die Median-Teststatistik:

$$T_n := \frac{1}{n} \sum_{i=n_1+1}^{n} \mathbb{1}_{\left\{R_{ni} < \frac{n+1}{2}\right\}}.$$

Bestimmen Sie $(c_{ni})_{i=1,...,n}$ und ϕ so, dass die Form aus Gleichung (1) erreicht wird; damit ist die Median-Teststatistik eine lineare Rang-Statistik.

(c) Angenommen $\frac{n_1}{n} \to \lambda \in (0,1)$. Wir definieren

$$\tilde{T}_n := \mathbb{E}(T_n) - \frac{n_2}{n^2} \sum_{i=1}^{n_1} \mathbb{1}_{\left\{F(X_i) < \frac{1}{2}\right\}} + \frac{n_1}{n^2} \sum_{i=n_1+1}^{n} \mathbb{1}_{\left\{F(X_i) < \frac{1}{2}\right\}}.$$

Zeigen Sie, dass $\sqrt{n}(T_n - \tilde{T}_n) \xrightarrow{p} 0$.

Hinweis: Sie dürfen folgende Aussage ohne Beweis verwenden: Sei ϕ wie oben gegeben; fast überall stetig, nicht konstant und $\frac{1}{n} \sum_{i=1}^{n} \phi\left(\frac{i}{n+1}\right)^2 \to \int_0^1 \phi(x)^2 \, \mathrm{d}x < \infty$. Dann gilt für

$$\tilde{T}_n := n \cdot \bar{c}_n \cdot \bar{a}_n + \sum_{i=1}^{n} (c_{ni} - \bar{c}_n) \cdot \phi(F(X_i)),$$

gerade $\frac{\operatorname{Var}(T_n - \tilde{T}_n)}{\operatorname{Var}(T_n)} \to 0$ und $\mathbb{E}(T_n) = \mathbb{E}(\tilde{T}_n)$.

Abgabe:
In Zweiergruppen, bis spätestens Freitag, den **10.12.2021, 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`