UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

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

Vorbereitungsblatt

Aufgabe 1.

Seien f und g zwei Lebesgue-Dichten. Die Kullback-Leibler-Divergenz zwischen f und g definieren wir dann als

$$K(f,g) = \mathbb{E}_f \left[\log \frac{f(X)}{g(X)} \right],$$

wobei $\mathbb{E}_f$ sich auf die Dichte f von X bezieht.

(a) Bezüglich der Modellklasse $\mathcal{M} := \{f_\theta : \theta \in \Theta\}$ betrachten wir die negative log-Likelihood $L_n^{\mathcal{M}}(\theta)$ von n unabhängigen Beobachtungen $X_1, ..., X_n$, wobei jede Zufallsvariable X_i die Dichte f besitzt,

$$L_n^{\mathcal{M}}(\theta) = -\frac{1}{n} \log \left(\prod_{i=1}^{n} f_\theta(X_i) \right).$$

Beachten Sie, dass f nicht notwendigerweise in $\mathcal{M}$ enthalten sein muss. Zeigen Sie, dass unter der Annhame (A1) aus der Vorlesung

$$\hat{\theta}_n := \text{argmin}_{\theta \in \Theta} L_n^{\mathcal{M}}(\theta) \xrightarrow{p} \text{argmin}_{\theta \in \Theta} K(f, f_\theta).$$

(b) Sei f_j die Dichte einer k-dimensionalen Normalverteilung $\mathcal{N}(\mu_j, \Sigma_j)$, wobei $j \in \{0,1\}$ und $k \in \mathbb{N}$. Zeigen Sie, dass

$$K(f_0, f_1) = \frac{1}{2} \left(\text{tr}(\Sigma_1^{-1}\Sigma_0) + (\mu_1 - \mu_0)^T \Sigma_1^{-1} (\mu_1 - \mu_0) - k - \log \left(\frac{\det \Sigma_0}{\det \Sigma_1} \right) \right)$$

Hinweis: Sie dürfen ohne Beweis verwenden, dass

$$\mathbb{E}_{f_0}((X - \mu_0)^T \Sigma_1^{-1} (X - \mu_0)) = \text{tr}(\Sigma_1^{-1}\Sigma_0)$$

für beliebige Matrizen Σ_0, Σ_1 gilt.

(c) Sei nun $X_1, ..., X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \Sigma)$ eine k-dimensionale Stichprobe. Bestimmen Sie die Parameter so, dass der Maximum-Likelihood-Schätzer bezüglich der Modellklasse

$$\mathcal{M} := \left\{ \mathcal{N}(\nu, \Lambda) : \nu \in \mathbb{R}^k, \Lambda = \text{diag}(\lambda_1, ..., \lambda_k) \right\}$$

konvergiert. Sie dürfen dabei annehmen, dass Satz 1.2 gilt, obwohl der Parameterraum nicht kompakt ist.

Aufgabe 2.

Seien $X_1, X_2, ..., X_n$ unabhängig und identisch verteilte Zufallsvariablen, die auf $[a,b]$ gleichverteilt sind.

(a) Zeigen Sie, dass der Maximum-Likelihood-Schätzer für $\theta = (a, b)$ durch

$$\hat{\theta}_n := (\hat{a}_n, \hat{b}_n) = (\min\{X_1, ..., X_n\}, \max\{X_1, ..., X_n\})$$

gegeben ist.

(b) Zeigen Sie weiter, dass $\hat{\theta}_n$ asymptotisch konsistent ist.

Aufgabe 3.

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 λ_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 4.

Sei $\mathbb{P}_\lambda = \mathrm{Poi}(\lambda)$ die Poisson-Verteilung für den Parameter $\lambda > 0$ und sei $\mathbb{P}_{n,\lambda} = \mathbb{P}_\lambda^n$ die gemeinsame Verteilung von n unabhängigen Poisson-verteilten Zufallsvariablen. Zeigen Sie, dass $\mathbb{P}_{n,\lambda}$ LAN in $\lambda_0 > 0$ für eine geeignete Folge δ_n und zentrale Folge Z_{n,λ_0} ist.

Aufgabe 5.

Gegeben sei eine Familie $\{\mathbb{P}_\sigma : \sigma > 0\}$ von Verteilungen auf $\mathbb{R}$ mit Lebesgue-Dichten $p_\sigma(x) = g\left(\frac{x}{\sigma}\right) \cdot \frac{1}{\sigma}$, wobei g stetig differenzierbar und $g > 0$ ist.

(a) Formulieren Sie eine Integrabilitätsbedingung für g derart, dass $\{\mathbb{P}_\sigma : \sigma > 0\}$ differenzierbar im quadratischen Mittel ist. Folgern Sie, dass die Folge $\mathbb{P}_{n,\sigma} = \mathbb{P}_\sigma^n$ LAN für alle $\sigma_0 > 0$ ist. Geben Sie die zugehörige zentrale Folge an.

(b) Zeigen Sie anhand von Teilaufgabe (a), dass $\mathbb{P}_{n,\sigma} = \mathcal{N}(0, \sigma^2)^n$ LAN für alle $\sigma_0 > 0$ ist und ermitteln Sie die zugehörige zentrale Folge.
Hinweis: Sie dürfen ohne Beweis die Identität $\int_{\mathbb{R}} (-x^2 + 1)^2 \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x = 2$ *verwenden.*

Aufgabe 6.

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 7.

Gegeben sei eine Familie von Verteilungen $\{\mathbb{P}_\theta : \theta \in \Theta\}$, $\Theta \subseteq \mathbb{R}^d$, mit Dichten f_θ bezüglich eines σ-endlichen Maßes μ. Wir bezeichnen mit $I(\theta)$ die Fisher-Informationsmatrix, wobei $\theta = (\theta_1, ..., \theta_d)$. Unter der Annahme, dass alle Regularitätsbedingungen aus Satz 6.1 (Informations-Ungleichung; Frechet-Cramer-Rao) erfüllt werden, zeigen Sie:

(a) Für jeden erwartungstreuen Schätzer S der ersten k Parameter $(\theta_1, ..., \theta_k)$ mit $k \leq d$ gilt

$$\operatorname{Var}(S) \geq \left(I(\theta)^{-1}\right)_{i,j=1,\dots,k}.$$

(b) Unter der Kenntnis von $\theta_{k+1}, ..., \theta_d$ gilt für jeden erwartungstreuen Schätzer S von $(\theta_1, ..., \theta_k)$ mit $k < d$,

$$\operatorname{Var}(S) \geq \left(I(\theta)_{i,j=1,\dots,k}\right)^{-1}.$$

Aufgabe 8 (6 Punkte).

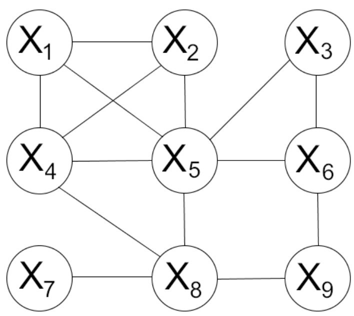
Sei $(X_1, ..., X_n) \sim U[0,\theta]^n := \mathbb{P}_\theta^n =: \mathbb{P}_{n,\theta}$ und p_θ bezeichne die Dichte von $\mathbb{P}_{n,\theta}$ bzgl. des Lebesguemaßes auf $\mathbb{R}^n$.

(a) Bestimmen Sie den Maximum-Likelihood-Schätzer $\hat{\theta}_n$ für θ und bestimmen Sie die zugehörige Varianz $\operatorname{Var}(\hat{\theta}_n)$.

(b) Zeigen Sie, dass $\hat{\theta}_n$ nicht für beliebiges $n \in \mathbb{N}$ die Cramer-Rao-Schranke einhält. Welche Voraussetzung von Satz 6.1 wird verletzt?

Aufgabe 9.

(a) Sei $\mathbb{P}^X$ eine R-Verteilung, welche bzgl. dem Graphen



die Eigenschaft (P) besitzt. Zeigen Sie, dass $\mathbb{P}^X$ bzgl. des Graphens die Eigenschaft (F) erfüllt und geben Sie eine Faktorisierung der Dichte bzgl. der (maximalen) Cliquen an.

(b) Sei $\mathbb{P}^X$ eine Verteilung auf $[0,1]^4$ mit Dichte

$$f(x_1, x_2, x_3, x_4) = \exp(c + x_1 + x_3 + x_1 x_2 + x_2 x_3 x_4)$$

bzgl. λ^4 (λ als Lebesguemaß, c eine Normierungskonstante). Geben Sie einen Graphen mit möglichst wenig Kanten an bzgl. dem $\mathbb{P}^X$ die Eigenschaft (F) besitzt. Geben Sie mittels (G) alle nichttrivialen bedingten Unabhängigkeiten an, welche direkt aus dem Graphen gefolgert werden können.

Hinweis:
Dies ist ein (Klausur-) Vorbereitungsblatt und ist nicht abzugeben bzw. wird nicht bewertet.

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