Statistik 1
Prof. Dr. Jan Johannes
Sergio Brenner Miguel
Sommersemester 2019

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

07. Übungsblatt

Aufgabe 1	Aufgabe 2	Aufgabe 3	Aufgabe 4	Vorbereitung	Σ

Aufgabe 1 (4 Punkte)

Überprüfen Sie, ob die folgenden Verteilungsfamilien Exponentialfamilien bilden. Bestimmen Sie gegebenfalls den natürlichen Parameterraum.

a) Multinomialverteilung $\{\mathrm{M}_{(\theta,n)}, \theta = (\theta_0, \dots, \theta_s) \in (0,1)^{s+1}, \sum_{i=0}^s \theta_i = 1\}$ mit Zähldichten

$$\mathbb{p}_\theta(x) = \frac{n!}{x_0! \cdots x_s!} \theta_0^{x_0} \cdots \theta_s^{x_s}, \quad x \in [\![0, n]\!]^{s+1}, \sum_{i=0}^{s} x_i = n;$$

b) Poissonverteilung $\{\mathrm{Poi}_\lambda, \lambda > 0\}$;

c) Gleichverteilung $\{U_{[0,\theta]}, \theta > 0\}$;

d) Betaverteilung $\{\mathrm{Beta}_{(a,b)}, a, b > 0\}$ mit Lebesguedichten

$$\mathbb{f}_{a,b}(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \quad x \in (0,1).$$

Hinweis zu d): Benutzen Sie ohne Beweis, dass $B(a,b) := \int_0^1 x^{a-1}(1-x)^{b-1}dx < \infty \Leftrightarrow a, b > 0$.

Hinweis zur Aufgabe 1: Verwenden Sie Definition 13.01 und geben Sie das dominierende Maß μ und die Funktionen η, C, S, h explizit an oder versuchen Sie die Nichtexistenz solcher Funktionen zu begründen mittels eines Widerspruchsbeweis. Verwenden Sie auch die Definition 13.03.

Aufgabe 2 (4 Punkte)

Sei $\mathbb{P}_{\Theta_{\text{nat}}}$ eine Exponentialfamilie mit natürlichem Parameterraum $\Theta_{\text{nat}} \subset \mathbb{R}^k$ und Darstellung

$$\frac{d\,\mathbb{P}_\theta}{d\mu}(x) = C(\theta)h(x)\exp(\langle\theta, S(x)\rangle) = h(x)\exp(\langle\theta, S(x)\rangle - A(\theta)),$$

mit $A(\theta) = \log(C(\theta)^{-1}) = \log(\int_{\mathcal{X}} h(x)\exp(\langle\theta, S(x)\rangle)\mu(dx))$. Sei $\boldsymbol{\theta}_0$ ein innerer Punkt von Θ_{nat}. Zeigen Sie:

a) Die erzeugende Funktion $\psi_{\boldsymbol{\theta}_0}(s) = \mathbb{E}_{\boldsymbol{\theta}_0}[\exp(\langle S, s\rangle)]$ ist in einer Umgebung $U(0)$ des Nullvektors wohldefiniert.

b) Es gilt für $s \in U(0)$ die Darstellung $\psi_{\boldsymbol{\theta}_0}(s) = \exp(A(\boldsymbol{\theta}_0 + s) - A(\boldsymbol{\theta}_0))$.

Nun möchten wir zeigen, dass auf $U(0)$ die Funktion $\psi_{\boldsymbol{\theta}_0}$ in jeder Variable θ_i unendlich differenzierbar ist. Wir definieren für $s \in U(0)$ die Funktion $\varphi(\boldsymbol{\theta}_0+s) = \int h(x)\exp(\langle S(x), \boldsymbol{\theta}_0+s\rangle)\mu(dx)$ und sehen, dass $\psi_{\boldsymbol{\theta}_0}(s) = \exp(-A(\boldsymbol{\theta}_0))\varphi(\boldsymbol{\theta}_0+s)$. Folglich reicht es die Aussage für $\varphi(\boldsymbol{\theta}_0 + s)$ zu zeigen. Da die Argumentation für beliebige Richtungen und höhere Ableitungen analog ist, lässt sich das Problem reduzieren. Zeigen Sie nun:

c) Für $i \in [\![1, k]\!]$, $s \in U(0)$ existiert $\frac{\partial}{\partial\theta_i}\varphi(\boldsymbol{\theta}_0 + s)$ und es gilt

$$\frac{\partial\varphi(\boldsymbol{\theta}_0 + s)}{\partial\theta_i} = \int h(x)\exp(\langle S(x), \boldsymbol{\theta}_0 + s\rangle)S_i(x)\mu(dx).$$

Anmerkung: Insbesondere gilt dann auch $\frac{\partial^2}{\partial\theta_i\partial\theta_j}\exp(A(\boldsymbol{\theta}_0)) = \int h(x)\exp(\langle S(x), \boldsymbol{\theta}_0 + s\rangle)S_i(x)S_j(x)\mu(dx)$ für $i, j \in [\![1, k]\!]$.

d) Für $i, j \in [\![1, k]\!]$ folgt außerdem $\mathbb{E}_{\boldsymbol{\theta}_0}[S_i] = \frac{\partial A}{\partial\theta_i}(\boldsymbol{\theta}_0)$ und $\mathbb{C}\mathrm{ov}_{\boldsymbol{\theta}_0}[S_i, S_j] = \frac{\partial^2 A}{\partial\theta_i\partial\theta_j}(\boldsymbol{\theta}_0)$.

Aufgabe 3 (4 Punkte)

Für eine Familie $\mathcal{P}$ von Verteilungen wird eine erschöpfende (suffiziente) Statistik S^* *minimalerschöpfend* (minimalsuffizient) genannt, falls für jede weitere erschöpfende Statistik S eine messbare Funktion g existiert, so dass $S^* = g \circ S$ $\mathbb{P}$-fast sicher für alle $\mathbb{P} \in \mathcal{P}$.

Sei $\mathbb{P}_\Theta$ eine Familie von Verteilungen auf $(\mathcal{X}, \mathscr{X})$ mit Dichten $\mathbb{f}_\theta$ bezüglich eines σ-endlichen Maßes μ. Ferner sei der Träger der Dichten identisch.

a) Eine Statistik S ist erschöpfend für $\mathbb{P}_\Theta$ genau dann, wenn für feste θ und θ' der Quotient $\mathbb{f}_\theta \,/\, \mathbb{f}_{\theta'}$ eine Funktion von $S(x)$ ist.

b) Ist $\Theta = [\![0, k]\!]$, so ist $S(x) = \left(\frac{\mathbb{f}_1(x)}{\mathbb{f}_0(x)}, \ldots, \frac{\mathbb{f}_k(x)}{\mathbb{f}_0(x)}\right)$ minimalerschöpfend für $\mathbb{P}_\Theta$.

c) Sei $\Theta_0 \subset \Theta$, T minimalerschöpfend für $\mathbb{P}_{\Theta_0}$ und erschöpfend für $\mathbb{P}_\Theta$. Beweisen Sie, dass dann T minimalerschöpfend für $\mathbb{P}_\Theta$ ist.
Hinweis: Zeigen Sie zunächst, dass alle Maße aus $\mathbb{P}_\Theta$ zueinander äquivalent sind.

d) Zeigen Sie, dass $\overline{X} := \frac{1}{n}\sum_{i=1}^n x_i$ minimalerschöpfend für $\{\mathrm{N}^n_{(\mu,1)}, \mu \in \mathbb{R}\}$ ist.
Hinweis: Verwenden Sie Aufgabenteil a)-c).

Aufgabe 4 (4 Punkte)

Sei $\mathbb{P}_\Theta$ eine Familie von Verteilungen auf $(\mathcal{X}, \mathscr{X})$, die von einem σ-endlichen Maß μ dominiert werde. Ferner sei A ein festes Element von $\mathscr{X}$ mit $\mathbb{P}_\theta[A] > 0$ für alle $\theta \in \Theta$, und $\mathbb{P}^*_\theta$ die auf A eingeschränkte Verteilung, d.h. $\mathbb{P}^*_\theta[B] = \frac{\mathbb{P}_\theta[B\cap A]}{\mathbb{P}_\theta[A]}$. Zeigen Sie:

a) Eine Dichte $\mathbb{f}^*_\theta$ von $\mathbb{P}^*_\theta$ bezüglich μ für beliebige $\theta \in \Theta$ ist durch $\mathbb{f}^*_\theta = \frac{1}{\mathbb{P}_\theta[A]}\mathbb{1}_A \frac{d\mathbb{P}_\theta}{d\mu}$ gegeben.

b) Ist $T : (\mathcal{X}, \mathscr{X}) \to (\mathcal{Y}, \mathscr{Y})$ erschöpfend für $\mathbb{P}_\Theta$, dann ist T auch erschöpfend für $\mathbb{P}^*_\Theta := \{\mathbb{P}^*_\theta : \theta \in \Theta\}$.

c) Ist T außerdem vollständig für $\mathbb{P}_\Theta$, dann ist T auch vollständig für $\mathbb{P}^*_\Theta$.
Hinweis: Verwenden Sie die Notation $\varphi(T) := \mathbb{E}[\mathbb{1}_A|T]$ und den Satz 03.41.

Klausurvorbereitung:
Die folgende Aufgabe dient zur Wiederholung des Vorlesungstoffes und als Hilfsmittel zum gezielten Lernen auf die Klausur.

Aufgabe 5 (Vorbereitung; 4 Bonuspunkte)

Gegeben sei das lineare Modell für $X \in \mathbb{R}^{(n,p)}, \beta \in \mathbb{R}^p$

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

wobei $X^tX = \mathrm{E}_p$ und $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^t \sim \mathrm{N}_{(0,\sigma^2\mathrm{E}_n)}$ mit $\sigma^2 > 0$.

a) Bestimmen Sie den gewöhnlichen Kleinste Quadrate Schätzer $\widehat{\beta} = (\widehat{\beta}_1, \ldots, \widehat{\beta}_p)^t$ von $\beta = (\beta_1, \ldots, \beta_p)^t$ und geben Sie dessen Verteilung an. Begründen Sie die Wohldefiniertheit.

b) Geben Sie die Verteilung von $\lambda_1\widehat{\beta}_1 + \lambda_2\widehat{\beta}_2$ für $\lambda_1, \lambda_2 \in \mathbb{R}$ an.

Seien nun $Z_1, \ldots, Z_n$ unabhängig und identisch $\mathrm{N}_{(\mu,\sigma^2)}$-verteilte Zufallsvariablen mit $\sigma^2 > 0$. Weiter sei $\overline{Z}_n := \frac{1}{n}\sum_{i=1}^n Z_i$.

c) Zeigen Sie, dass der Schätzer $\widetilde{\sigma^2_n} := \frac{1}{n}\sum_{i=1}^n (Z_i - \overline{Z}_n)^2 \xrightarrow{\mathbb{P}} \sigma^2$ ist und zeigen Sie, dass $\mathbb{E}[\widetilde{\sigma^2_n}] = \frac{n-1}{n}\sigma^2$ um zu zeigen, dass der Schätzer nicht erwartungstreu ist.

d) Sei nun $\widehat{\sigma_n^2} := \frac{1}{n-1}\sum_{i=1}^n (Z_i - \overline{Z}_n)^2$. Zeigen Sie:

i) $\widehat{\sigma_n^2}$ ist erwartungstreu;

ii) $\widehat{\sigma_n^2} - \widetilde{\sigma_n^2} \xrightarrow{\mathbb{P}} 0$ und

Hinweis: Verwenden Sie die Aussage der Aufgabe 40 b) des Zettels 11 der Vorlesung "Einführung in die Wahrscheinlichkeitstheorie und Statistik" WS 18/19. https://sip.math.uni-heidelberg.de/vl/ews-ws18/src/uebung11.pdf

iii) $\widehat{\sigma_n^2} \xrightarrow{\mathbb{P}} \sigma^2$.

Anmerkung: Sie haben nun gezeigt, dass $\widehat{\sigma_n^2}$ ein konsistenter, erwartungstreuer Schätzer von σ^2 ist.

Abgabe:
In Zweiergruppen, bis spätestens Dienstag, den **11. Juni 2019, 11:15 Uhr**.
(Die Übungszettelkästen sind im 1. OG, INF 205, vor dem Dekanat.)

Homepage:
https://sip.math.uni-heidelberg.de/vl/st1-ss19/