# Übungen zur Wahrscheinlichkeitstheorie und Statistik

Prof. Dr. C. Löh/M. Blank — Blatt 13 vom 12. Juli 2012

---

**Aufgabe 1** (Exponentialfamilien)**.** Welche der folgenden Aussagen sind wahr? Begründen Sie jeweils kurz Ihre Antwort.

1. Ist $\big(\mathbb{N}, \mathrm{Pot}(\mathbb{N}), (\mathrm{Poi}(\vartheta))_{\vartheta\in\mathbb{R}_{>0}}\big)$ eine Exponentialfamilie?
2. Sei $a \in \mathbb{R}$. Ist $\big(\mathbb{R}, B(\mathbb{R}), (N(a,\vartheta^2))_{\vartheta\in\mathbb{R}_{>0}}\big)$ eine Exponentialfamilie?

*Lösungshinweise.*

1. Ja, denn: Für jedes $\vartheta \in \mathbb{R}_{>0}$ ist
$$\begin{aligned} f_\vartheta \colon \mathbb{N} &\longrightarrow [0,1] \\ k &\longmapsto e^{-\vartheta} \cdot \frac{\vartheta^k}{k!} = e^{-\vartheta + k\cdot\ln\vartheta} \cdot \frac{1}{k!} \end{aligned}$$
eine Zähldichte für $\mathrm{Poi}(\vartheta)$. [Also ist $\big(\mathbb{N}, \mathrm{Pot}(\mathbb{N}), (\mathrm{Poi}(\vartheta))_{\vartheta\in\mathbb{R}_{>0}}\big)$ eine Exponentialfamilie bzgl. des Zählmaßes auf $\mathbb{N}$ und der Inklusion $\mathbb{N} \hookrightarrow \mathbb{R}$ (und $a = \ln$, $b = (\mathbb{R}_{>0} \hookrightarrow \mathbb{R})$ und $h \colon k \longmapsto 1/k!$ in der Notation der Vorlesung. Die Abbildung $h$ ist natürlich eine Zufallsvariable bzgl. $\mathrm{Pot}(\mathbb{N})$)].
2. Ja, denn: Für jedes $\vartheta \in \mathbb{R}_{>0}$ ist
$$\begin{aligned} f_\vartheta \colon \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{\sqrt{2\cdot\pi\cdot\vartheta^2}} \cdot e^{-\frac{(x-a)^2}{2\cdot\vartheta^2}} = \frac{1}{\sqrt{2\cdot\pi}} \cdot e^{-\frac{(x-a)^2}{2\cdot\vartheta^2} - \ln\vartheta}. \end{aligned}$$
eine $\lambda^1$-Wahrscheinlichkeitsdichte für $N(a,\vartheta^2)$.
Also ist $\big(\mathbb{R}, B(\mathbb{R}), (N(a,\vartheta^2))_{\vartheta\in\mathbb{R}_{>0}}\big)$ eine Exponentialfamilie bzgl. $\lambda^1$ und
$$\begin{aligned} T \colon \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto 1/2 \cdot (x-a)^2. \end{aligned}$$
(In der Notation der Vorlesung ist hier $b = \ln$ und
$$\begin{aligned} \text{„}a\text{“} \colon \mathbb{R}_{>0} &\longrightarrow \mathbb{R} \\ \vartheta &\longmapsto \frac{1}{\vartheta^2} \end{aligned} \qquad \begin{aligned} h \colon \mathbb{R} &\longrightarrow \mathbb{R} \\ \omega &\longmapsto 1/\sqrt{2\cdot\pi}. \end{aligned}$$
Als konstante Abbildung ist $h$ eine Zufallsvariable.)

**Aufgabe 2** (Eindeutigkeit gleichmäßig bester Schätzer)**.** Sei $(\Omega, S, (P_\vartheta)_{\vartheta\in\Theta})$ ein statistisches Modell und seien $T, \widetilde{T} \colon (\Omega, S) \longrightarrow (\mathbb{R}, B(\mathbb{R}))$ gleichmäßig beste erwartungstreue Schätzer für eine Abbildung $\tau \colon \Theta \longrightarrow \mathbb{R}$. Zeigen Sie: Für alle $\vartheta \in \Theta$ gilt dann $P_\vartheta(T = \widetilde{T}) = 1$.

*Lösungshinweise.* Sei $\vartheta \in \Theta$. Nach Voraussetzung sind $T$ und $\widetilde{T}$ (und damit auch $T - \widetilde{T}$ und $T + \widetilde{T}$) bzgl. $P_\vartheta$ quadratintegrierbar. Es gilt $\mathrm{Var}_{P_\vartheta}(T - \widetilde{T}) = 0$, denn: Es gilt

$$\begin{aligned}
&\mathrm{Var}_{P_\vartheta}(T) + \mathrm{Var}_{P_\vartheta}(1/2(T - \widetilde{T})) \\
&\quad \leq \mathrm{Var}_{P_\vartheta}(1/2 \cdot (T + \widetilde{T})) + \mathrm{Var}_{P_\vartheta}(1/2 \cdot (T - \widetilde{T})) \\
&\quad = \frac{1}{4} \cdot (\mathrm{Var}_{P_\vartheta}(T) + \mathrm{Var}_{P_\vartheta}(\widetilde{T}) + \mathrm{Cov}(T, \widetilde{T}) + \mathrm{Var}_{P_\vartheta}(T) + \mathrm{Var}_{P_\vartheta}(\widetilde{T}) + \mathrm{Cov}(T, -\widetilde{T})) \\
&\quad = \frac{1}{4} \cdot (\mathrm{Var}_{P_\vartheta}(T) + \mathrm{Var}_{P_\vartheta}(\widetilde{T}) + \mathrm{Var}_{P_\vartheta}(T) + \mathrm{Var}_{P_\vartheta}(\widetilde{T})) \\
&\quad = \mathrm{Var}_{P_\vartheta}(T),
\end{aligned}$$

wobei wir in der ersten Ungleichung

$$\forall_{\vartheta' \in \Theta} \quad E_{P_{\vartheta'}}(1/2 \cdot (T + \widetilde{T})) = 1/2 \cdot (E_{P_{\vartheta'}}(T) + E_{P_{\vartheta'}}(\widetilde{T})) = E_{P_{\vartheta'}}(T)$$

und die Minimalität von $T$ benutzt haben. Somit gilt

$$0 \leq \mathrm{Var}_{P_\vartheta}((T - \widetilde{T})) = 4 \cdot \mathrm{Var}_{P_\vartheta}(1/2(T - \widetilde{T})) \leq 0.$$

Es ist $T - \widetilde{T}$ also $P_\vartheta$-fast sicher konstant, und daher gilt $P_\vartheta$-fast sicher

$$T - \widetilde{T} = E_\vartheta(T - \widetilde{T}) = 0.$$

[Mit anderen Worten $P_\vartheta(T = \widetilde{T}) = 1$.]

**Aufgabe 3** (Maximum-Likelihood-Schätzer for Poisson-Verteilungen)**.** Wir betrachten das statistische Modell $(\Omega, S, (P_\vartheta)_{\vartheta \in \Theta})$, wobei $\Omega := \mathbb{N}$, $S := \mathrm{Pot}(\Omega)$, $\Theta := \mathbb{R}_{>0}$ und $P_\vartheta := \mathrm{Poi}(\vartheta)$ für $\vartheta \in \Theta$ ist.

1. Zeigen Sie, dass der nachfolgende „Beweis“ nicht korrekt ist, indem Sie erklären, welcher Schritt nicht korrekt ist und erklären, warum die Behauptung nicht korrekt ist.

   *Behauptung.* Dann ist

   $$\begin{aligned} T \colon \Omega &\longrightarrow \Theta \\ \omega &\longmapsto \omega \end{aligned}$$

   ein Maximum-Likelihood-Schätzer für $\mathrm{id}_\Theta$ (bezüglich der offensichtlichen Zähldichten und der Borel-$\sigma$-Algebra auf $\Theta = \mathbb{R}_{>0}$).

   *Beweis.* Offensichtlich ist $T$ bezüglich $S$ und $B(\mathbb{R})$ messbar. Es bleibt also zu zeigen, dass $T$ das entsprechende Maximierungsproblem löst: Die entsprechende Likelihoodfunktion ist

   $$\begin{aligned} L \colon \Omega \times \Theta &\longrightarrow \mathbb{R}_{\geq 0} \\ (\omega, \vartheta) &\longmapsto e^{-\vartheta} \cdot \frac{\vartheta^\omega}{\omega!}. \end{aligned}$$

   Sei $\omega \in \Omega$. Indem man $\ln \circ L(\omega, \cdot\,)$ differenziert und die Methoden der reellen Analysis zur Lösung von Extremwertproblemen anwendet, sieht man, dass die einzige Maximalstelle von $\ln \circ L(\omega, \cdot\,)$ bzw. $L(\omega, \cdot\,)$ an der Stelle $\vartheta = \omega$ liegt. Damit folgt die Behauptung. □

2. Wie kann man die obige Situation so modifizieren, dass die der in der Behauptung angegebenen entsprechende Abbildung ein Maximum-Likelihood-Schätzer für die Identitätsabbildung ist?

*Lösungshinweise.*

1. Der Fehler liegt darin, dass die Funktion $L(0,\cdot)$ kein Maximum auf $\Theta = \mathbb{R}_{>0}$ annimmt. Anders formuliert: $\omega = 0 \in \Omega$ liegt nicht in $\Theta$, also lässt sich $T\colon \Omega \longrightarrow \Theta$ nicht wie oben definieren.

2. Wir definieren $\mathrm{Poi}_0 := \delta_0$ und betrachten das modifizierte statistische Modell $(\Omega, S, (P_\vartheta)_{\vartheta\in\mathbb{R}_{\geq 0}})$. Dann ist die Likelihoodfunktion dieses Modells gegeben durch

$$\begin{aligned} L\colon \Omega \times \Theta &\longrightarrow \mathbb{R}_{\geq 0} \\ (\omega,\vartheta) &\longmapsto \begin{cases} e^{-\vartheta}\cdot\frac{\vartheta^\omega}{\omega!} & \text{für } \vartheta > 0 \\ 1 & \text{für } \vartheta = 0 \text{ und } \omega = 0 \\ 0 & \text{sonst.} \end{cases} \end{aligned}$$

Dann ist

$$\begin{aligned} T\colon \Omega &\longrightarrow \Theta \\ \omega &\longmapsto \omega \end{aligned}$$

der gesuchte Maximum-Likelihood-Schätzer, denn: Für $\omega = 0$ nimmt $L(\omega,\cdot)$ ein eindeutiges Maximum in 0 an. Für alle $\omega > 0$ gilt (wie im fehlerhaften Beweis): $L(\omega,\cdot)\colon \mathbb{R}_{\geq 0} \longmapsto \mathbb{R}_{>0}$ ist (zweimal) differenzierbar auf $\mathbb{R}_{>0}$, also ist auch $\varphi := \ln\circ L(\omega,\cdot)|_{\mathbb{R}_{>0}}$ (zweimal) differenzierbar, und es gilt für alle $\vartheta \in \mathbb{R}_{>0}$

$$\begin{aligned} \varphi'(\vartheta) &= -1 + \frac{\omega}{\vartheta} \\ \varphi''(\omega) &= -1 - \frac{1}{\omega} < 0 \end{aligned}$$

Also ist $\vartheta = \omega$ die einzige lokale Maximalstelle von $\varphi$ und nach dem Satz von Rolle auch die einzige globale Maximalstelle. Wegen der Monotonie des Logarithmus ist $\omega$ auch die einzige Maximalstelle von $L(\omega,\cdot)|_{\mathbb{R}>0}$. Wegen $L(\omega,0) = 0$ für $\omega > 0$ folgt daher die Behauptung.

**Aufgabe 4** (German Tank Problem)**.** Wir betrachten das statistische Modell $(\Omega, S, (P_\vartheta)_{\vartheta\in\Theta})$ für das German Tank Problem, das wie folgt gegeben ist: Sei $n \in \mathbb{N}_{>0}$ (die beobachtete Anzahl der Seriennummern), sei

$$\Omega := \binom{\mathbb{N}_{>0}}{n}, \qquad S := \mathrm{Pot}(\Omega), \qquad \Theta := \mathbb{N}_{\geq n},$$

und für alle $\vartheta \in \Theta$ bezeichne $P_\vartheta$ die Gleichverteilung auf $\binom{\{1,\dots,\vartheta\}}{n}$. Hierbei schreiben für eine Menge $A$ kurz $\binom{A}{n} := \left\{B \subset A \;\middle|\; |B| = n\right\}$ für die Menge der $n$-elementigen Teilmengen von $A$.

1. Zeigen Sie, dass es für das Modell $(\Omega, S, (P_\vartheta)_{\vartheta\in\Theta})$ einen Maximum-Likelihood-Schätzer (bezüglich der offensichtlichen Zähldichten und der $\sigma$-Algebra $\mathrm{Pot}(\Theta)$ auf $\Theta$) gibt und bestimmen Sie einen solchen.

2. Zeigen Sie durch vollständige Induktion: Für alle $r \in \mathbb{N}$, $s \in \mathbb{N}_{\geq r}$ gilt
$$\sum_{m=r}^{s} m \cdot \frac{\binom{m-1}{r-1}}{\binom{s}{r}} = \frac{r\cdot(s+1)}{r+1}.$$

3. Ist Ihr Maximum-Likelihood-Schätzer aus dem ersten Aufgabenteil ein erwartungstreuer Schätzer für die Inklusion $\Theta \hookrightarrow \mathbb{R}$ ?

4. Bestimmen Sie einen (weiteren) erwartungstreuen Schätzer für die Inklusion $\Theta \hookrightarrow \mathbb{R}$.

*Lösungshinweise.*

1. Wir erhalten als Likelihood-Funktion
$$\begin{aligned} L\colon \Omega\times\Theta &\longrightarrow \mathbb{R} \\ (\omega,\vartheta) &\longmapsto \begin{cases} \binom{\vartheta}{n}^{-1} & \text{für } \max\omega \leq \vartheta \\ 0 & \text{sonst.} \end{cases} \end{aligned}$$
Damit ist
$$\begin{aligned} T\colon (\Omega, S) &\longrightarrow (\Theta, \mathrm{Pot}(\Theta)) \\ \omega &\longmapsto \max\omega \end{aligned}$$
ein Maximum-Likelihood-Schätzer, denn: Die Abbildung $T$ ist natürlich messbar bzgl $S = \mathrm{Pot}(\Omega)$. Sei $\omega \in \Omega$ und $\vartheta \in \Theta$. Für $\vartheta > \max\omega = T(\omega)$ gilt $L(\omega,\vartheta) = 0 \leq L(\omega, T(\omega))$. Für $\vartheta \leq \max\omega$ gilt
$$L(\omega,\vartheta) = \binom{\vartheta}{n}^{-1} \leq \binom{T(\omega)}{n}^{-1} = L(\omega, T(\omega)).$$

2. Sei $r \in \mathbb{N}$. Wir zeigen die Behauptung mit vollständiger Induktion nach $s \in \mathbb{N}_{\geq r}$. Für $s = r$ erhalten wir [wobei wir etwa $\binom{a}{-1} := 0$ setzen für alle $a \in \mathbb{Z}$]
$$\sum_{m=r}^{r} m \cdot \frac{\binom{m-1}{r-1}}{\binom{r}{r}} = r = \frac{r\cdot(r+1)}{r+1}.$$
Sei die Aussage für $s \in \mathbb{N}_{\geq r}$ bereits gezeigt. Dann folgt
$$\begin{aligned} \sum_{m=r}^{s+1} m \cdot \frac{\binom{m-1}{r-1}}{\binom{s+1}{r}} &= (s+1)\cdot\frac{\binom{s}{r-1}}{\binom{s+1}{r}} + \sum_{m=r}^{s} m \cdot \frac{\binom{m-1}{r-1}\cdot r!\cdot(s+1-r)!}{(s+1)!} \\ &= (s+1)\cdot\frac{s!\cdot r!\cdot(s+1-r)!}{(s+1)!\cdot(r-1)!\cdot(s+1-r)!} + \frac{s+1-r}{s+1}\cdot\sum_{m=r}^{s} m\cdot\frac{\binom{m-1}{r-1}}{\binom{r}{s}} \\ &\overset{\text{IV}}{=} r + \frac{s+1-r}{s+1}\cdot\frac{r\cdot(s+1)}{r+1} \\ &= \frac{r\cdot(r+1)+r\cdot(s+1-r)}{r+1} \\ &= \frac{r\cdot(s+2)}{r+1}. \end{aligned}$$

3. Nein, denn: Wir erhalten für alle $\vartheta \in \Theta$: Der Schätzer $T$ ist ausserhalb einer $P_\vartheta$-Nullmenge beschränkt, ist also bzgl $P_\vartheta$ integrierbar, und es gilt

$$\begin{aligned} E_{P_\vartheta}(T) &= \sum_{\omega \in \binom{\{1,\dots,\vartheta\}}{n}} \max \omega \cdot \frac{1}{\binom{\vartheta}{n}} \\ &= \sum_{m=n}^{\vartheta} \sum_{\omega \in \binom{\{1,\dots,m-1\}}{n-1}} m \cdot \frac{1}{\binom{\vartheta}{n}} \\ &= \sum_{m=n}^{\vartheta} m \cdot \frac{\binom{m-1}{n-1}}{\binom{\vartheta}{n}} \\ &= \frac{n \cdot (\vartheta+1)}{n+1}. \end{aligned} \qquad \text{(Teil 2)}$$

   Für alle $\vartheta \in \Theta_{>n}$ gilt also $E_{P_\vartheta}(T) \neq \vartheta$. Daher ist $T$ nicht erwartungstreu.

4. Nach Teil 3 [und der Linearität des Erwartungswertes] ist $T \cdot \frac{n+1}{n} - 1$ ein erwartungstreuer Schätzer für $\Theta \hookrightarrow \mathbb{R}$.

**Bonusaufgabe** (Modellbau)**.** Da Commander Blorx nicht jünger wird und demnächst in Rente geht, möchte er sich ein bizarres Hobby für seinen Ruhestand zulegen. Seine Wahl fällt dabei auf Modellbau. Helfen Sie Commander Blorx bei der Modellierung der folgenden Probleme!

1. Commander Blorx möchte sich mit einem Freund aus der Studienzeit treffen. Beide nehmen es mit der Pünktlichkeit nicht so genau und werden daher (jeweils unabhängig voneinander, gleichverteilt) innerhalb von einer Stunde nach der verabredeten Zeit am vereinbarten Treffpunkt eintrudeln. Da beide etwas ungeduldig sind, werden sie aber nicht länger als 10 Minuten auf den jeweils anderen warten. Wie groß ist die Wahrscheinlichkeit, dass sie sich tatsächlich treffen?

2. Was ist eine vernünftige Approximation für die Anzahl der Tipfehler pro Seite in der Encyclopedia Blorxica?

3. Ein von Commander Blorx ausgesetzter Schachkönig steht auf einem unendlich breiten und unendlich langen (ansonsten leeren) Schachbrett. In jedem Zug bewegt er sich zufällig gleichverteilt ein Feld in eine der zulässigen acht Richtungen. Mit welcher Wahrscheinlichkeit kehrt er nach einer endlichen Anzahl von Zügen zu seinem Ausgangsfeld zurück?

4. Wie kann Commander Blorx abschätzen, ob der „garantiert faire“ Würfel, den er von einem dubiosen Casinobetreiber erhalten hat, tatsächlich fair ist?

*Hinweis.* Es genügt, die jeweiligen wahrscheinlichkeitstheoretischen bzw. statistischen Modelle anzugeben und kurz zu begründen; es geht *nicht* darum, die entsprechenden Probleme tatsächlich zu lösen.

*Lösungshinweise.*

1. Seien $T, T' \colon (\Omega, S, P) \longrightarrow ([0,1], B([0,1]))$ unabhängige gleichverteilte Zufallsvariablen. Dann lässt sich das gesuchte Ereignis beschreiben durch

$$\text{(Blorx und sein Freund treffen sich)} = \{|T - T'| \leq 1/6\}$$

und die gesuchte Wahrscheinlichkeit ist $P(\{|T - T'| \leq 1/6\})$. Dabei beschreibe $T$ den Ankunftszeitpunkt von Blorx und $T'$ den seines Freundes und $[0, 1]$ das Zeitintervall in welchem sich die beiden treffen wollten.

Alternativ betrachte man das Maß $P_{|T-T'|}$ auf $([0, 1], B([0, 1]))$ und das Ereignis $[0, 1/6] \subset [0, 1]$.

2. Wir betrachten das folgende (unrealistische) Modell: Wir nehmen an, dass es ein (globales) $p \in [0, 1]$ gibt, so dass jedes Zeichen in der Encyclopedia Blorxica unabhängig von den anderen Zeichen mit Wahrscheinlichkeit $p$ falsch ist. [Dabei nehmen wir insbesondere an, dass die Zeichen unabhängig voneinander entweder richtig oder falsch sind.]

   Sei $n \in \mathbb{N}_{>0}$ die Anzahl der Zeichen in der Encyclopedia Blorxica. Wir erhalten dann das Binomialmodell $(\{0, \dots, n\}, \mathrm{Pot}(\{0, \dots, n\}), B(n, p))$ für die Anzahl der Tippfehler in der Encyclopedia Blorxica. Für große Werte von $n$ sind die entsprechenden Wahrscheinlichkeit numerisch schwierig zu berechen. Wir betrachten daher stattdessen als „Approximation“ das Poisson-Modell

   $$(\mathbb{N}, \mathrm{Pot}(\mathbb{N}), \mathrm{Poi}(n \cdot p)).$$

   [Die Fehler „pro Seite“ modellieren wir dann als Gesamtfehlerzahl geteilt durch die Anzahl der Seiten.]

3. Sei $\Omega := \{-1, 0, 1\}^2 \setminus \{(0, 0)\}$. Wir verwenden als Modell den Wahrscheinlichkeitsraum

   $$(\Omega, \mathrm{Pot}(\Omega), U_\Omega^d)^{\otimes \mathbb{N}_{>0}}$$

   Wir beschreiben also das Schachbrett durch $\mathbb{Z}^2$ und modellieren dann die Folge der Züge des Königs durch $(\pi_i)_{i \in \mathbb{N}_{>0}}$. Das gesuchte Ereignis lässt sich dann schreiben als

   $$(\text{Der König kehrt zum Ausgangsfeld zurück}) = \bigcup_{n \in \mathbb{N}_{>0}} \Big\{ \sum_{i=1}^{n} \pi_i = (0, 0) \Big\}$$

   und die gesuchte Wahrscheinlichkeit ist

   $$P\Big( \bigcup_{n \in \mathbb{N}_{>0}} \Big\{ \sum_{i=1}^{n} \pi_i = (0, 0) \Big\} \Big).$$

4. Sei $\Theta_6 = \{p \in [0, 1]^6 \mid \sum_{i=1}^{6} p_i = 1\}$ die Menge der möglichen Zähldichten auf $\{1, \dots, 6\}$ und $(P_\vartheta)_{\vartheta \in \Theta_6}$ die zugehörige Familie von Wahrscheinlichkeitsmaßen auf $(\{1, \dots, 6\}, \mathrm{Pot}(\{1, \dots, 6\}))$. Wir betrachten das statistische Modell

   $$(\{1, \dots, 6\}, \mathrm{Pot}(\{1, \dots, 6\}), (P_\vartheta)_{\vartheta \in \Theta_6})^{\otimes \mathbb{N}}.$$

   Für $\vartheta \in \Theta_6$ und $i \in \{1, \dots, 6\}$ beschreibt $\vartheta_i$ die Wahrscheinlichkeit mit dem Würfel $i$ Augen zu werfen.

   Man lege nun geeignete Hypothesenmengen fest, etwa (willkuerlich)

   $$\Theta_6 = ((1/12, 3/12)^6 \cap \Theta_6) \cup (([0, 1]^6 \setminus (1/12, 3/12)^6) \cap \Theta_6)$$

   mit Nullhypothese $([0, 1]^6 \setminus (1/12, 3/12)^6) \cap \Theta_6$, und wähle einen passenden Test.

keine Abgabe; Lösungshinweise werden zu Beginn der vorlesungsfreien Zeit zur Verfügung gestellt