Zulassungsprüfung Stochastik, 01.10.10

Wir gehen stets von einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ aus. Die Borel σ-Algebra auf $\mathbb{R}^n$ wird mit $\mathcal{B}^n$ bezeichnet, das Lebesgue Maß auf $\mathbb{R}^n$ wird mit λ^n bezeichnet.

Aufgabe 1 (18 Punkte)
Sei $\mathbb{N} = \{1, 2, \dots\}$ die Menge der natürlichen Zahlen. Für $A \subset \mathbb{N}$ sei $\mu(A) := \sum_{n=1}^{\infty} \frac{1}{n} 1_A(n)$, also die Summe über alle $\frac{1}{n}$ mit $n \in A$.

(a) Zeigen Sie, dass für paarweise disjunkte Mengen $A_1, A_2, \dots$ und $A := \bigcup_{n\in\mathbb{N}} A_n$ gilt
$$1_A = \sum_{n=1}^{\infty} 1_{A_n}.$$

(b) Zeigen Sie mit Hilfe von (a), dass μ ein Maß auf $\mathcal{P}(\mathbb{N})$ ist.
Hinweis: Sie können ohne Beweis verwenden, dass für doppelt indizierte Folgen $(a_{ij})_{i,j\in\mathbb{N}} \subset (0,1)$ nicht negativer reeller Zahlen $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij}$ gilt.

(c) Bestimmen Sie alle Nullmengen bezüglich μ.

(d) Sei $f : \mathbb{N} \longrightarrow [0, \infty)$.

(i) Zeigen Sie, dass für alle $k \in \mathbb{N}$ gilt: $f(k) = \sum_{n=1}^{\infty} f(n) \cdot 1_{\{n\}}(k)$ gilt.

(ii) Beweisen Sie: $\int f \, d\mu = \sum_{n=1}^{\infty} \frac{f(n)}{n}$.

Lösung
Zu (a)
$1_A(\omega) = 1 \iff \omega \in A \iff \exists_1 i_0 \in \mathbb{N} : \omega \in A_{i_0} \iff 1_{A_{i_0}}(\omega) = 1, 1_{A_i}(\omega) = 0$ für alle $i \neq i_0 \iff \sum_{n=1}^{\infty} 1_{A_n}(\omega) = 1_{A_{i_0}}(\omega) = 1$.
Zu (b)
$\mu(\emptyset) = 0$ ist klar. Seien $A_1, A_2, \dots \subset \mathbb{N}$ paarweise disjunkt und $A := \bigcup_{n\in\mathbb{N}} A_n$. Dann gilt

$$\begin{aligned} \mu(A) &= \sum_{n=1}^{\infty} \frac{1}{n} 1_A(n) \overset{(a)}{=} \sum_{n=1}^{\infty} \frac{1}{n} \sum_{i=1}^{\infty} 1_{A_i}(n) = \sum_{n=1}^{\infty}\sum_{i=1}^{\infty} \frac{1}{n} 1_{A_i}(n) \\ &= \sum_{i=1}^{\infty} \underbrace{\sum_{n=1}^{\infty} \frac{1}{n} 1_{A_i}(n)}_{=\mu(A_i)} = \sum_{i=1}^{\infty} \mu(A_i). \end{aligned}$$

Zu (c)
Sei $A \subset \mathbb{N}$. Es gilt:
$\mu(A) = 0 \iff \sum_{n \in \mathbb{N}} 1_A(n)\frac{1}{n} = 0 \iff 1_A(n) = 0$ für alle $n \in \mathbb{N} \iff n \notin A$ für alle $n \in \mathbb{N} \iff A = \emptyset$. Damit ist $\emptyset$ die einzige Nullmenge.
Zu (d)
(i) Für $k \in \mathbb{N}$ gilt $\displaystyle\sum_{n=1}^{\infty} f(n)1_{\{n\}}(k) = 1_{\{k\}}f(k) = f(k)$.

(ii) Sei $f_k := \displaystyle\sum_{n=1}^{k} f(n)1_{\{n\}}$. Dann ist die Folge $(f_k)_{k\in\mathbb{N}}$ eine monoton steigende Folge messbarer, nicht negativer Funktionen mit $\lim_{k\to\infty} f_k = f$. Mit dem Satz von der monotonen Konvergenz folgt $\displaystyle\lim_{k\to\infty} \int f_k \, d\mu = \int f \, d\mu$ und somit wegen

$$\forall k \in \mathbb{N} : \quad \int f_k \, d\mu = \sum_{n=1}^{k} f(n)\mu(\{n\}) = \sum_{n=1}^{k} f(n)\frac{1}{n}$$

das Gewünschte.

Aufgabe 2 (18 Punkte)
Ein Produzent von PC-Festplatten hat durch umfangreiche Test festgestellt, dass die Lebensdauer X einer Festplatte in Jahren exponential-verteilt ist mit Erwartungswert vier Jahre.

(a) Bestimmen Sie die Standardabweichung σ und die Verteilungsfunktion von X.

(b) Mit welcher Wahrscheinlichkeit funktioniert eine Festplatte länger als vier Jahre?

(c) Bestimmen Sie die Wahrscheinlichkeit, dass eine Festplatte mindestens sechs Jahre lang funktioniert, vorausgesetzt sie funktioniert bereits fünf Jahre lang.

(d) Mit welcher Wahrscheinlichkeit ist die Lebensdauer X im Intervall $[\mathbb{E}(X)-\sigma, \mathbb{E}(X)+\sigma]$ enthalten?

(e) Während welcher Zeitdauer funktionieren 50 % der Festplatten?

(f) Geht während der Garantiezeit eine Festplatte kaputt, dann muss sie ersetzt werden. Welches wäre die optimale Garantiezeit um maximal 20 % der Festplatten ersetzen zu müssen?

Lösung
Zu (a)
Aus den Angaben folgt $X \sim \mathrm{Exp}(\lambda)$ wegen $\mathbb{E}(X) = \dfrac{1}{\lambda}$ sofort $\lambda = \dfrac{1}{4}$, $\sigma = \sqrt{\dfrac{1}{\lambda^2}} = 4$. Außerdem gilt für die Verteilungsfunktion F von X

$$F(x) = \begin{cases} 1 - e^{-x/4} & x \geq 0 \\ 0 & \text{sonst} \end{cases}.$$

Zu (b)

$$\mathbb{P}(X \geq 4) = 1 - F(4) = e^{-1} = 0,37.$$

Zu (c)

$$\mathbb{P}(X \geq 6|X \geq 5) = \frac{e^{-1,5}}{e^{-1,25}} = 0,79.$$

Zu (d)

$$\mathbb{P}(4 - 4 \leq X \leq 4 + 4) = F(8) - F(0) = 1 - e^{-2} = 0,86$$

Zu (e)

$$\mathbb{P}(X \geq t) = 0,5 \iff e^{-t/4} = 0,5 \iff t = -4 \ln 0,5 = 2,77.$$

Während 2,77 Jahren funktionieren 50 % der Geräte einwandfrei.
Zu (f)

$$\mathbb{P}(X \leq t) = 0,20 \iff 1 - e^{-t/4} = 0,20 \iff t = -4 \ln 0,8 = 0,89.$$

Die Garantiezeit sollte 0,89 Jahre betragen.

Aufgabe 3 (18 Punkte)
Für einen Versicherungsnehmer mit Vertragsdauer $t > 0$ wird seine Schadenzahl N_t in Abhängigkeit von t modelliert. Es wird angenommen, dass $N_t \sim \mathrm{Poi}(\lambda(t))$ gilt, mit $\lambda(t) := ae^{-t}$, $a > 0$. In einem großen Versicherungsbestand wird angenommen, dass die Vertragsdauer $T \sim \mathrm{Exp}(1)$ verteilt ist.
Sei N die Schadenzahl eines beliebigen Versicherungsnehmers des Bestands (also ohne Kenntnis der Vertragsdauer).

(a) Bestimmen Sie $\mathbb{E}\left(e^{-kT}\right)$ für $k \in \mathbb{N}$.

(b) Bestimmen Sie $\mathbb{E}(N|T)$ und $\mathrm{Var}(N|T)$.

(c) Bestimmen Sie $\mathbb{E}(N)$ und $\mathrm{Var}(N)$.

(d) Bestimmen Sie die Wahrscheinlichkeit, dass ein zufällig ausgewählter Versicherungsnehmer mindestens einen Schaden meldet.

Hinweis: $\exp(-ae^{-t})' = \exp(-ae^{-t})ae^{-t}$.

Lösung
Zu (a)
Für $k \in \mathbb{N}$ gilt

$$\mathbb{E}\left(e^{-kT}\right) = \int_0^\infty e^{-(k+1)t}\, dt = \frac{1}{k+1}.$$

Zu (b)
Es gilt $\mathbb{E}(N|T = t) = \lambda(t) = ae^{-t}$ und $\mathrm{Var}(N|T = t) = \lambda(t) = ae^{-t}$ und somit

$$\mathbb{E}(N|T) = ae^{-T}, \quad \mathrm{Var}(N|T) = ae^{-T}.$$

Zu (c)

Mit dem Satz vom iterierten Erwartungswert gilt

$$\mathbb{E}(N) = \mathbb{E}(\mathbb{E}(N|T)), \quad \mathrm{Var}(N) = \mathrm{Var}(\mathbb{E}(N|T)) + \mathbb{E}(\mathrm{Var}(N|T)).$$

Somit folgt mit (a) und (b)

$$\begin{aligned}
\mathbb{E}(N) &= \mathbb{E}(ae^{-T}) = \frac{a}{2} \\
\mathrm{Var}(N) &= \mathrm{Var}(ae^{-T}) + \mathbb{E}(ae^{-T}) \\
&= a^2\mathbb{E}(e^{-2T}) - a^2\mathbb{E}(e^{-T})^2 + a\mathbb{E}(e^{-T}) \\
&= a^2 \cdot \frac{1}{3} - a^2 \cdot \frac{1}{4} + a \cdot \frac{1}{2} \\
&= \frac{a^2}{12} + \frac{a}{2}
\end{aligned}$$

Zu (d)
Laut Annahme gilt $\mathbb{P}(N = 0|T = t) = \exp(-\lambda(t)) = \exp(-ae^{-t})$. Somit folgt

$$\begin{aligned}
\mathbb{P}(N = 0) &= \int_0^\infty \mathbb{P}(N = 0|T = t)e^{-t}\,dt = \int_0^\infty \underbrace{\exp(-ae^{-t})e^{-t}}_{=\frac{1}{a}\exp(-ae^{-t})'}\,dt \\
&= \frac{1}{a}(1 - e^{-a}).
\end{aligned}$$

Damit erhält man $\mathbb{P}(N \geq 1) = 1 - \mathbb{P}(N = 0) = 1 - \dfrac{1}{a}(1 - e^{-a})$.

Aufgabe 4 (18 Punkte)
Während der Fußball WM 2010 errang Krake Paul den Ruf eines Fußballorakels. Es sagte unter Anderem den Ausgang aller sieben Spiele der deutschen Nationalmannschaft (Sieg oder Niederlage) richtig voraus.

(a) Angenommen der Krake Paul tippt zufällig.

   (i) Wie groß ist die Wahrscheinlichkeit, dass er alle sieben Spiele richtig tippt?

   (ii) Wie beurteilen Sie die Hypothese $p = \frac{1}{2}$, wobei p die Wahrscheinlichkeit sei, dass Paul richtig tippt?

(b) Weltweit gab es einige „tierische Fußballexperten“ (z. B. Papagei Mani in Singapur, pulpo Pepe in Spanien usw.).

   (i) Angenommen n tierische Fußballexperten geben jeweils ihren Tipp für sieben Fußballspiele zufällig ab (nur Sieg oder Niederlage). Wie groß ist die Wahrscheinlichkeit, dass mindestens eines der n Tiere alle sieben Spiele richtig tippt?

   (ii) Wieviele Experten muss man am Anfang tippen lassen, damit mit einer Wahrscheinlichkeit von wenigstens 90 % mindestens einer alle sieben Spiele richtig tippt.

(c) Wie interpretieren Sie die Ergebnisse des Kraken Paul?

Lösung
Zu (a)
(i) Sei N die Anzahl der richtigen Antworten. Dann ist $N \sim B(n,p)$-verteilt mit $n = 7$ und $p = 0,5$.
Somit gilt $\mathbb{P}(N = 7) = 0,5^7 = 0,78\%$.
(ii) Die Hypothese wird abgelehnt bis zum Niveau $0,78\%$.

Zu (b)
(i) Sei M die Anzahl der Experten, die alle sieben Spiele richtig tippen. Dann gilt $M \sim B(n,p)$ mit $p = 0,5^7$. Dann gilt

$$\mathbb{P}(M \geq 1) = 1 - \mathbb{P}(M = 0) = 1 - (1 - 0,5^7)^n.$$

(ii) Somit folgt $\mathbb{P}(M \geq 1) \geq 0,9 \iff 1 - (1 - 0,5^7)^n \geq 0,9 \iff n \geq \dfrac{\ln 0,1}{\ln(1 - 0,5^7)} = 293,6$. Man muss mindestens 294 Experten befragen.
Zu (c)
Auch wenn Pauls Ergebnisse beeindrucken, können sie darauf zurückzuführen sein, dass über die anderen Tiere mit weniger beeindruckenden Ergebnissen nicht mehr berichtet wurde, also nur über erwünschte Ergebnisse berichtet wird.

Aufgabe 5 (18 Punkte) Seien $X \sim N(1, \theta)$ und $Y \sim N\left(1, \frac{1}{\theta}\right)$ mit $\theta > 0$ unabhängig. Es seien jeweils Stichproben vom Umfang n durch unabhängige Realisierungen $x_1, \ldots, x_n$ bzw. $y_1, \ldots, y_n$ von X bzw. Y gegeben. Bestimmen Sie einen Maximum-Likelihood-Schätzer für θ mit Hilfe beider Stichproben.
Lösung Wegen der Unabhängigkeit ist die Likelihood gegeben durch

$$\begin{aligned} L(\theta) &= \left[\prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{1}{2}\frac{(x_i - 1)^2}{\theta}\right)\right] \cdot \left[\prod_{i=1}^n \frac{\sqrt{\theta}}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y_i - 1)^2\theta\right)\right] \\ &= \left(\frac{1}{2\pi}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^n \frac{(x_i - 1)^2}{\theta} - \frac{1}{2}\sum_{i=1}^n (y_i - 1)^2\theta\right). \end{aligned}$$

Somit folgt für $\ell := \ln L$

$$\begin{aligned} \ell(\theta) &= n\ln(2\pi) - \frac{1}{2}\sum_{i=1}^n \left(\frac{(x_i - 1)^2}{\theta} + (y_i - 1)^2\theta\right) \\ \ell'(\theta) &= \frac{1}{2}\sum_{i=1}^n \left(\frac{(x_i - 1)^2}{\theta^2} - (y_i - 1)^2\right) \\ \ell''(\theta) &= -\frac{1}{2}\sum_{i=1}^n \frac{(x_i - 1)^2}{\theta^3}. \end{aligned}$$

Aus der zweiten Gleichung ergibt sich durch Nullsetzen und Auflösen

$$\hat{\theta} = \sqrt{\frac{\sum\limits_{i=1}^{n} (x_i - 1)^2}{\sum\limits_{i=1}^{n} (y_i - 1)^2}}$$

und mit ℓ'' folgt, dass in $\hat{\theta}$ ein Maximum vorliegt. Dies ist der gesuchte Schätzer.