Zulassungsprüfung Stochastik, 11.05.13

Wir gehen stets von einem Maßraum $(\Omega, \mathcal{A}, \mu)$ bzw. 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.

Sollten Ihnen in Teilaufgaben Ergebnisse fehlen, dann treffen Sie eine plausible Annahme dafür.

Aufgabe 1 (18 Punkte)

(a) Sei $f : [0,1] \longrightarrow [0,\infty]$ stetig und monoton fallend. Beweisen Sie

$$\lim_{n\to\infty} \int_{[0,1]} f(x^n) d\lambda(x) = f(0).$$

(b) Seien $A_n \in \mathcal{A}$, $n \in \mathbb{N}$ paarweise disjunkt, $A := \bigcup_{n\geq 1} A_n$ und $f : \Omega \longrightarrow [0,\infty)$ messbar. Beweisen Sie

$$\int_A f d\mu = \sum_{n=1}^{\infty} \int_{A_n} f d\mu.$$

Aufgabe 2 (18 Punkte)

Ein Stab der Länge 1 werde zufällig in zwei Teile zerbrochen. Die Bruchstelle wird mit einer Zufallsvariablen $X \sim U(0,1)$ modelliert.

(a) Sei $Y := \min(X, 1-X)$ die Länge des kürzeren Bruchstücks. Bestimmen Sie den Erwartungswert von Y.

(b) Sei Z die Länge des längeren Bruchstücks. Bestimmen Sie $P(Z \leq z)$ für $z \in (1/2, 1)$ und schließen Sie daraus, dass $Z \sim U(1/2, 1)$ gilt.

(c) Bestimmen Sie den Erwartungswert von $\frac{Y}{Z}$ (kürzeres durch längeres Stück).

Aufgabe 3 (18 Punkte)

Gegeben seien die Zufallsvariablen $N : \Omega \longrightarrow \mathbb{N}_0$ und $\Theta : \Omega \longrightarrow [0,1]$. Es gelte $P(\Theta \leq \vartheta) = \vartheta^2$ für $\vartheta \in (0,1)$.
Für gegebenes $\vartheta \in (0,1)$, sei $N \sim NB(1,\vartheta)$, d.h. N gegeben $\Theta = \vartheta$ besitzt die die bedingte Dichte $f(\cdot|\vartheta) : \mathbb{N}_0 \longrightarrow [0,1)$ mit $f(k|\vartheta) = \vartheta(1-\vartheta)^k$.

(a) Bestimmen Sie die gemeinsame Dichte von N und Θ.

(b) Bestimmen Sie die bedingte Dichte von Θ gegeben $N = k$.

Hinweis: Für $p, q \in \mathbb{N}_0$ gilt $\displaystyle\int_0^1 x^p(1-x)^q\, dx = \frac{p!q!}{(p+q+1)!}$.

(c) Bestimmen Sie $\mathbb{E}(\Theta|N)$.

Aufgabe 4 (18 Punkte)

Ein Bestand mit Volumen $v \in \mathbb{N}$ (z.B. Anzahl Versicherungsnehmer) wird durch die Schadenzahl $N : \Omega \longrightarrow \mathbb{N}_0$ und die Frequenz $Z := \dfrac{N}{v}$ modelliert. Dabei wird $N \sim \mathrm{Poi}(\lambda v)$ mit $\lambda > 0$ angenommen.

(a) Bestimmen Sie $E(N), E(Z), \mathrm{Var}(N)$ und $\mathrm{Var}(Z)$.

(b) Bestimmen Sie die Wahrscheinlichkeitsfunktion von Z.

(c) Beobachtet werden unabhängige Realisierungen von $Z_1, \ldots, Z_n$ (Frequenzen) und Volumina $v_1, \ldots v_n$, insbesondere gilt also $Z_i v_i \sim \mathrm{Poi}(\lambda v_i)$ für dasselbe $\lambda > 0$.

  (i) Bestimmen Sie die gemeinsame Wahrscheinlichkeitsfunktion (also die gemeinsame Zähldichte) von $(Z_1, \ldots, Z_n)$.

  (ii) Bestimmen Sie einen Maximum-Likelihood-Schätzer für λ.

  (iii) Gegeben seien die folgenden Beobachtungen:

i	1	2	3	4	5	6	Σ
z_i	$2/7$	$1/3$	$3/11$	$1/10$	$1/3$	$1/8$	
v_i	14	12	11	10	9	8	64
$v_i z_i$	4	4	3	1	3	1	16

  Bestimmen Sie einen ML-Schätzwert für λ.

Aufgabe 5 (18 Punkte)
In der folgenden Tabelle sind Körpergröße x_i und Gewicht y_i voneinander unabhängiger Personen aufgezeichnet.

179,0	181,4	152,5	173,5	178,4	178,6	175,6	180,8	186,2	171,0	182,2
82,8	86,1	66,4	76,1	82,1	86,2	79,4	81,8	82,3	75,1	82,0

Es wird für die Zufallsvariablen Y_i das Modell

$$Y_i = a + bx_i + \varepsilon_i,$$

untersucht. Die ε_i seien unabhängig und $N(0, \sigma^2)$-verteilt. Ein Statistikprogramm berechnet die folgenden Größen, die Sie verwenden können:

$$\hat{a} = -21,8;\ \hat{b} = 0,58; \quad \sum_i (y_i - \hat{y}_i)^2 = 58,9$$

$$\overline{x} = 176,3; \quad \sum_i (x_i - \overline{x})^2 = 796,7$$

$$\overline{y} = 80; \quad \sum_i (y_i - \overline{y})^2 = 324,9;\ \sum_i (y_i - \overline{y})(x_i - \overline{x}) = 460,3.$$

(a) Stützen die Daten die Behauptung, dass das Gewicht von der Körpergröße abhängt? Die Aussage soll zu einem Niveau von $\alpha = 10$ % erfolgen.

(b) Ein Statistiker teilt Ihnen mit, er hätte die Daten durch Simulation mit $b = 0,4$ erzeugt. Wie beurteilen Sie seine Behauptung zu einem Niveau von $\alpha = 10$ %? Kommentieren Sie das Ergebnis.

Lösungsvorschläge

Aufgabe 1
Zu (a)
Für $x \in [0,1)$ gilt $0 \leq x^{n+1} \leq x^n$ für alle $n \in \mathbb{N}$. Damit folgt $0 \leq f(x^n) \leq f(x^{n+1})$, da f monoton fällt und $\lim_{n\to\infty} f(x^n) = f(0)$, da f stetig ist. Seien $g_n, g : [0,1] \longrightarrow [0,\infty),\ g_n(x) = f(x^n),\ g(x) = \begin{cases} f(0) & \text{falls } x < 1 \\ 0 & x = 1 \end{cases}$. Laut Definition gilt $0 \leq g_n \uparrow g$ fast überall und mit dem Satz von der monotonen Konvergenz folgt

$$\begin{aligned}\lim_{n\to\infty} \int_{[0,1)} f(x^n) d\lambda(x) &= \lim_{n\to\infty} \int_{[0,1)} g_n(x) d\lambda(x) = \int_{[0,1)} g(x) d\lambda(x) \\ &= \int_{[0,1)} f(0) d\lambda(x) = f(0).\end{aligned}$$

Zu (b)
Sei $g_n := \sum_{k=1}^{n} 1_{A_k},\ g := \sum_{k=1}^{\infty} 1_{A_k}$. Es gilt $0 \leq g_n \uparrow g$ und wegen $f \geq 0$ auch $0 \leq f g_n \uparrow f g$. Damit folgt mit dem Satz von der monotonen Konvergenz

$$\lim_{n\to\infty} \int_\Omega f g_n \, d\mu = \int_\Omega f g \, d\mu.$$

Die Behauptung folgt nun wegen

$$\begin{aligned}\int_\Omega f g_n \, d\mu &= \int_\Omega f \sum_{k=1}^{n} 1_{A_k} \, d\mu = \sum_{k=1}^{n} \int_\Omega f 1_{A_k} \, d\mu = \sum_{k=1}^{n} \int_{A_k} f \, d\mu \\ \int_\Omega f g d\mu &= \int_A f \, d\mu.\end{aligned}$$

Die letzte Gleichung gilt wegen $g = 1_A$:

$$1_A(x) = 1 \iff x \in A \iff \underset{k_0 \in \mathbb{N}}{\exists!}\, x \in A_{k_0} \iff g(x) = \sum_{k=1}^{\infty} 1_A(x) = 1_{A_{k_0}}(x) = 1.$$

Aufgabe 2
Zu (a)

$$E(Y) = \int_0^1 \min(x, 1-x) \, dx = \int_0^{1/2} x \, dx + \int_{1/2}^1 (1-x) \, dx = 2 \int_0^{1/2} x \, dx = \frac{1}{4}$$

Zu (b)
Es gilt $Z = \max(X, 1-X)$. Für $z \in (0.5, 1)$ erhalten wir

$$\begin{aligned}\mathbb{P}(Z \leq z) &= \mathbb{P}(\max(X, 1-X) \leq z) = \mathbb{P}(1-z \leq X \leq z) \\ &= \mathbb{P}(X \leq z) - P(X \leq 1-z) = z - 1 + z = 2z - 1\end{aligned}$$

also die Verteilungsfunktion einer $U(1/2, 1)$-verteilten Zufallsvariablen.

Zu (c)
Es gilt $Y = 1 - Z$ und somit

$$E\left(\frac{Y}{Z}\right) = E\left(\frac{1-Z}{Z}\right) = 2\int_{1/2}^{1} \frac{1-z}{z}\,dz = 2\ln 2 - 1 \approx 0,39.$$

Aufgabe 3
Zu (a)
Die Dichte f_Θ von Θ ist gegeben durch $f_\Theta(\vartheta) = 2\vartheta \cdot 1_{(0,1)}(\vartheta)$. Für $k \in \mathbb{N}_0$, $\vartheta \in (0,1)$ gilt

$$f(k|\vartheta) = \frac{f(k,\vartheta)}{f_\Theta(\vartheta)} \Longrightarrow f(k,\vartheta) = 2\vartheta^2(1-\vartheta)^k 1_{(0,1)}(\vartheta).$$

Zu (b)
Die Wahrscheinlichkeitsfunktion von N ist gegeben durch

$$P(N=k) = \int_{\mathbb{R}} f(k|\vartheta)\,d\vartheta = 2\int_0^1 \vartheta^2(1-\vartheta)^k\,d\vartheta = \frac{4k!}{(k+3)!}.$$

Damit folgt für $\vartheta \in (0,1)$, $k \in \mathbb{N}_0$

$$f(\vartheta|N=k) = \frac{f(k,\vartheta)}{P(N=k)} = 2\vartheta^2(1-\vartheta)^k \frac{(k+3)!}{4k!}.$$

Zu (c)

$$\begin{aligned} E(\Theta|N=k) &= \int \vartheta f(\vartheta|N=k)\,d\vartheta = \frac{(k+3)!}{2k!}\int_0^1 \vartheta^3(1-\vartheta)^k\,d\vartheta \\ &= \frac{(k+3)!}{2k!}\cdot\frac{3!k!}{(k+4)!} = \frac{3}{k+4}. \end{aligned}$$

Damit folgt $E(\Theta|N) = \dfrac{3}{N+4}$.

Aufgabe 4
Zu (a)

$$\begin{aligned} E(N) &= \mathrm{Var}(N) = \lambda v && \text{(da } N \sim \mathrm{Poi}(\lambda v)) \\ E(Z) &= E\left(\frac{N}{v}\right) = \frac{E(N)}{v} = \lambda \\ \mathrm{Var}(Z) &= \mathrm{Var}\left(\frac{N}{v}\right) = \frac{\mathrm{Var}(N)}{v^2} = \frac{\lambda}{v}. \end{aligned}$$

Zu (b) Sei $k \in \mathbb{N}_0$.

$$P(Z = k/v) = P(Zv = k) = e^{-\lambda v}\frac{(\lambda v)^k}{k!}.$$

Zu (c)

Für $z_i = \dfrac{k_i}{v_i}$, $k_i \in \mathbb{N}_0$ gilt

$$\begin{aligned}
L(\lambda) &= \prod_{i=1}^{n} e^{-\lambda v_i} \frac{(\lambda v_i)^{k_i}}{k_i!} \\
\ell(\lambda) &:= \ln L(\lambda) = \sum_{i=1}^{n} (-\lambda v_i + k_i \ln(\lambda v_i) - \ln k_i!) \\
\frac{d}{d\lambda}\ell(\lambda) &= -\sum_{i=1}^{n} v_i + \frac{1}{\lambda}\sum_{i=1}^{n} k_i = -\sum_{i=1}^{n} v_i + \frac{1}{\lambda}\sum_{i=1}^{n} v_i z_i \\
\ell''(\lambda) &= -\frac{1}{\lambda^2}\sum_{i=1}^{n} v_i z_i < 0 \\
\ell'(\hat{\lambda}) &= 0 \Longrightarrow \hat{\lambda} = \frac{\sum\limits_{i=1}^{n} v_i z_i}{\sum\limits_{i=1}^{n} v_i}.
\end{aligned}$$

Somit ist $\hat{\lambda} = \dfrac{\sum\limits_{i=1}^{n} v_i Z_i}{\sum\limits_{i=1}^{n} v_i}$ ML-Schätzer. Mit den Daten ergibt sich der Schätzwert $\hat{\lambda} = \dfrac{16}{64} = \dfrac{1}{4}$.

Aufgabe 5

Zu (a)

Zu prüfen ist die Nullhypothese $H_0 : b = 0$ gegen $H_1 : b \neq 0$. Die Prüfstatistik $t = \frac{\hat{b}-0}{\mathrm{se}(\hat{b})}$ ist laut Formelsammlung t_{n-2}-verteilt, es ergibt sich

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{n-2}\sum_i (y_i - \hat{y}_i)^2 = \frac{58,9}{9} \\
\mathrm{se}(\hat{b}) &= \sqrt{\frac{\hat{\sigma}^2}{\sum\limits_{i=1}^{n}(x_i - \bar{x})^2}} = \sqrt{\frac{58,9}{9 \cdot 796,7}} = 0,0906 \\
t &= 6,40.
\end{aligned}$$

H_0 wird verworfen, denn $|t| > t_{n-2,1-\alpha} = 1,833$ mit $n = 11, \alpha = 10\ \%$.

Zu (b)

Analog zu (a) ergibt sich mit $H_0 : b = 0,4$

$$t = \frac{0,58 - 0,4}{\mathrm{se}(\hat{b})} = 1,987 > 1,833.$$

Auch in diesem Fall wird verworfen. Auch wenn H_0 zutrifft, kann es zur Ablehnung kommen (Fehler 1. Art). Bei dem Signifikanzniveau von $\alpha = 10\ \%$ kann es in 10 % der Fälle einer solchen Untersuchung zur Verwerfung kommen.