Zulassungsprüfung Stochastik, 16.10.15

Wir gehen stets von einem Maßraum $(\Omega, \mathcal{A}, \mu)$ bzw. einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ 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)
Sei P ein Wahrscheinlichkeitsmaß, $C \in \mathcal{A}$, $X : \Omega \longrightarrow (0, \infty)$ integrierbar und $Q : \mathcal{A} \longrightarrow [0, \infty)$,

$$Q(A) = \int_\Omega X 1_A 1_C \, dP.$$

(a) Angenommen es gilt $\int_\Omega X \, dP = 1$. Ist Q ein Wahrscheinlichkeitsmaß?

(b) Beweisen Sie für jede Zufallsvariable $Y \geq 0$

$$\int_\Omega Y \, dQ = \int XY 1_C \, dP.$$

Aufgabe 2 (26 Punkte) Sei R_F die von einem Fondsmanager erwirtschaftete Rendite, R_M die Marktrendite, und es gelte:

$$\text{Der Zufallsvektor } (R_F, R_M)^\top \text{ ist zweidimensional normalverteilt,} \tag{1}$$

$$E(R_F) = E(R_M) = \mu, \tag{2}$$

$$\operatorname{Var}(R_F) = \operatorname{Var}(R_M) = \sigma^2, \tag{3}$$

$$\operatorname{Cov}(R_F, R_M) = \frac{1}{2}\sigma^2. \tag{4}$$

(a) Bestimmen Sie Erwartungswert, Varianz und Verteilung von $R_F - R_M$.

(b) Sei $x_\alpha > 0$. Ein Fondsmanager wird als gut bewertet, wenn $R_F - R_M > x_\alpha$ gilt. Bestimmen Sie die Wahrscheinlichkeit $\alpha \in (0,1)$, dass ein Fondsmanager als gut bewertet wird.

Für den Rest der Aufgabe betrachten wir n Fondsmanager mit Renditen $R_{F_i} \sim R_F$, $i = 1, \ldots, n$, R_{F_i} und R_M erfüllen jeweils (1)-(4), $R_{F_i} - R_M$ seien voneinander unabhängig. Sei N die Anzahl der Fondsmanager, die als gut bewertet werden.

(c) Bestimmen Sie

- (i) die Verteilung von N (in Abhängigkeit von n und α),
- (ii) die Wahrscheinlichkeit β, dass mindestens ein Fondsmanager als gut bewertet wird (in Abhängigkeit von n und α),
- (iii) für $n = 100$ und $\alpha = 5\,\%$ den Erwartungswert von N und die Wahrscheinlichkeit β.

(d) (i) Bestimmen Sie α aus (b) in Abhängigkeit von n so, dass für β aus (c) (ii) gilt: $\beta = 5\,\%$.

(ii) Sei $n = 100$. Geben Sie für die Angaben in (i) x_α in Abhängigkeit von σ so an, dass $P(N \geq 1) = 5\,\%$ näherungsweise gilt. Verwenden Sie die unten stehenden Quantile u_{1-t} der Standardnormalverteilung

t	$5 \cdot 10^{-1}$	$5 \cdot 10^{-2}$	$5 \cdot 10^{-3}$	$5 \cdot 10^{-4}$	$5 \cdot 10^{-5}$	$5 \cdot 10^{-6}$	$5 \cdot 10^{-7}$
u_{1-t}	0	1,645	2,576	3,291	3,891	4,417	4,892

(e) Von einem guten Fondsmanager würde man $E(R_F) > E(R_M)$ fordern und als Indiz eine Regel wie in (b) zugrunde legen. Ein Anlegermagazin schlägt vor, zur Bestimmung von x in (b) $\alpha = 5\,\%$ zu wählen. Ist dies sinnvoll? Begründen Sie Ihre Antwort.

Aufgabe 3 (13 Punkte) Seien X_1, X_2 Zufallsvariablen mit $E(X_i) = \mu_i \in \mathbb{R}$, $\operatorname{Var}(X_i) = \sigma_i^2 > 0$, $\sigma_i = \sqrt{\sigma_i^2}$, $i = 1,2$. Sei $Z_i = \dfrac{X_i - \mu_i}{\sigma_i}$ die Standardisierung von X_i, $i = 1,2$.

(a) Angenommen für $\rho \in \mathbb{R}$ gilt $E(X_1|X_2) = \rho X_2$. Beweisen Sie, dass $\operatorname{Cov}(X_1, X_2) = \rho\sigma_2^2$ gilt.

(b) Angenommen es gilt $E(Z_1|Z_2) = \rho Z_2$. Bestimmen Sie $E(X_1|X_2)$ in Abhängigkeit von X_2. Verwenden Sie ohne Beweis, dass $E(Z_1|Z_2) = E(Z_1|X_2)$ gilt.

Aufgabe 4 (33 Punkte) Sei X eine Zufallsvariable mit der Verteilungsfunktion

$$F(x) = \begin{cases} 0 & \text{falls } x \leq 0 \\ 1 - \exp\left(-\frac{\sqrt{x}}{\theta}\right) & \text{sonst} \end{cases}$$

mit $\theta > 0$.

(a) Bestimmen Sie die Dichte von X.

(b) Beweisen Sie für unabhängige $X_i \sim X$, $i = 1, \ldots, n$, dass $\hat{\theta} = \frac{1}{n}\sum_{i=1}^n \sqrt{X_i}$ ein Maximum-Likelihood Schätzer für θ ist.

(c) Beweisen Sie, dass $\sqrt{X} \sim \operatorname{Exp}\left(\frac{1}{\theta}\right)$ gilt.

(d) Beweisen Sie, dass $\frac{2n\hat{\theta}}{\theta} \sim \chi^2_{2n}$ gilt und bestimmen Sie Erwartungswert und Varianz von $\hat{\theta}$.
Hinweis: Verwenden Sie die Eigenschaften der Gammaverteilung in der Formelsammlung!

(e) Wir betrachten die Nullhypothese

$$H_0 : \theta \leq \theta_0.$$

Beweisen Sie, dass unter H_0 (also bei Gültigkeit von H_0) für $\alpha \in (0,1)$

$$\left\{\omega \in \Omega \middle| \frac{2n\hat{\theta}(\omega)}{\theta_0} > \chi^2_{2n,1-\alpha}\right\} \subset \left\{\omega \in \Omega \middle| \frac{2n\hat{\theta}(\omega)}{\theta} > \chi^2_{2n,1-\alpha}\right\}$$

und schließen Sie daraus, dass

$$P\left(\hat{\theta} > \frac{\chi^2_{2n,1-\alpha}\theta_0}{2n}\right) \leq \alpha$$

gilt.

(f) Gegeben seien die Daten

i	1	2	3	4	5	6	7	8	9	10	Summe
x_i	0,1	1,05	0,37	0,96	7,5	0,3	0,46	0,18	2,55	0,57	14,04
x_i^2	0,01	1,11	0,14	0,93	56,31	0,09	0,21	0,03	6,48	0,32	65,63
$x_i^{1/2}$	0,31	1,03	0,61	0,98	2,74	0,54	0,68	0,43	1,6	0,75	9,66

Wie beurteilen Sie die Hypothese $H_0 : \theta \leq 1$ zu einem Signifikanzniveau von 5 %?

Lösungsvorschläge

Aufgabe 1 [(a) 5 Punkte, (b) 13 Punkte]

Zu (a)
Q ist ein Maß aber kein Wahrscheinlichkeitsmaß, denn wenn C eine Nullmenge ist, gilt $Q(\Omega) = 0 \neq 1$.
Zu (b)
Sei zunächst Y eine Treppenfunktion, also $Y = \sum_{i=1}^{n} \alpha_i 1_{A_i},\ \alpha_i \geq 0,\ A_i \in \mathcal{B}^1$. Dann gilt

$$\begin{aligned}
\int_\Omega Y\,dQ &= \sum_{i=1}^{n} \alpha_i Q(A_i) = \sum_{i=1}^{n} \alpha_i \int_\Omega X 1_{A_i} 1_C\,dP = \int_\Omega X \left(\sum_{i=1}^{n} \alpha_i 1_{A_i} \right) 1_C\,dP \\
&= \int_\Omega XY1_C\,dQ.
\end{aligned}$$

Sei nun $Y \geq 0$ messbar. Dann gibt es eine monoton steigende Folge von Treppenfunktionen $Y_n \geq 0$ mit $\lim_{n\to\infty} Y_n = Y$. Wegen $X1_C \geq 0$ ist $X1_C Y_n$ eine monoton steigende Folge mit $X1_C Y_n \geq 0$ und $\lim_{n\to\infty} X1_C Y_n = X1_C Y$. Da Y_n Treppenfunktionen sind, gilt

$$\int_\Omega Y_n\,dQ = \int_\Omega XY_n 1_C\,dP \text{ für alle } n \in \mathbb{N}.$$

Laut Satz von der monotonen Konvergenz konvergieren beide Seiten obiger Gleichung

$$\begin{aligned}
\lim_{n\to\infty} \int_\Omega Y_n\,dQ &= \int_\Omega Y\,dQ, \\
\lim_{n\to\infty} \int_\Omega XY_n 1_C\,dP &= \int_\Omega XY1_C\,dP
\end{aligned}$$

und es folgt die Behauptung.

Aufgabe 2 [(a) 6, (b) 3, (c) 7, (d) 5, (e) 5]
Zu (a)
Es gilt

$$\mathrm{Var}(R_F - R_M) = \mathrm{Var}(R_F + (-R_M)) = \mathrm{Var}(R_F) + 2\mathrm{Cov}(R_F, -R_M) + \mathrm{Var}(-R_M).$$

Wegen

$$\begin{aligned}
\mathrm{Cov}(R_F, -R_M) &= -\mathrm{Cov}(R_F, R_M) = -\frac{1}{2}\sigma^2 \\
\mathrm{Var}(-R_M) &= (-1)^2 \mathrm{Var}(R_M) = \sigma^2
\end{aligned}$$

folgt weiter

$$\begin{aligned}
\mathrm{Var}(R_F - R_M) &= \sigma^2 - 2 \cdot \frac{1}{2}\sigma^2 + \sigma^2 = \sigma^2 \\
E(R_F - R_M) &= E(R_F) - E(R_M) = \mu - \mu = 0.
\end{aligned}$$

Da $(R_F, R_M)^\top$ zweidimensional normalverteilt ist, ist $\left\langle (1,-1)^\top, (R_F, R_M)^\top \right\rangle = R_F - R_M$ normalverteilt $N(0, \sigma^2)$.
Zu (b)
Sei $\alpha \in (0,1)$. Wegen (a) gilt $R_F - R_M \sim N(0, \sigma^2)$ und somit

$$\begin{aligned} \alpha &= P(R_F - R_M > x_\alpha) = 1 - P(R_F - R_M \leq x_\alpha) \\ &= 1 - \Phi\left(\frac{x_\alpha}{\sigma}\right). \end{aligned}$$

Zu (c)
(i) Es gilt $N \sim B(n, \alpha)$.
(ii) Gesucht ist $P(N \geq 1)$. Da N binomialverteilt ist, gilt

$$\beta = P(N \geq 1) \quad = \quad 1 - P(N = 0) = 1 - \alpha^0 (1-\alpha)^n = 1 - (1-\alpha)^n.$$

(iii) Da N binomialverteilt ist gilt für $n = 100$ und $\alpha = 0,05$

$$\begin{aligned} E(N) &= n\alpha = 5, \\ P(N \geq 1) &= 1 - 0,95^{100} = 0,994. \end{aligned}$$

Zu (d)
(i) Zu bestimmen ist $\alpha \in (0,1)$ so, dass $1 - (1-\alpha)^n \leq 0,05$.

$$1 - (1-\alpha)^n \leq 0,05 \Longleftrightarrow (1-\alpha)^n \geq 0,95 \Longleftrightarrow 1 - \alpha \geq 0,95^{1/n} \Longleftrightarrow \alpha \leq 1 - 0,95^{1/n}.$$

(ii) Mit $n = 100$ ergibt sich in (i) $\alpha = 1 - 0,95^{1/100} \approx 5 \cdot 10^{-4}$. Aus (b) ergibt sich

$$\alpha \quad = \quad 1 - \Phi\left(\frac{x_\alpha}{\sigma}\right) \Longrightarrow u_{1-\alpha} = \frac{x_\alpha}{\sigma}.$$

Somit ergibt sich $x_\alpha = \sigma u_{1-\alpha}$ wobei $u_{1-\alpha}$ das $(1-\alpha)$-Quantil der Standardnormalverteilung ist. Nach Tabelle erhalten wir $x_\alpha = 3,291 \cdot \sigma$.
Zu (e)
Laut Annahmen ist keiner der Fondsmanager besser als der Markt. Wählt man $\alpha = 0,05$, dann ist bei $n = 100$ die Wahrscheinlichkeit mindestens einen Fondsmanager als gut einzustufen 99,4 %, obwohl kein guter Fondsmanager dabei ist. Also ist $\alpha = 0,05$ nicht streng genug. Angemessen ist im Allgemeinen $\alpha < 1 - 0,95^{1/n}$ und das sich daraus ergebende x_α.

Aufgabe 3 [(a) 7, (b) 6]

Zu (a)
Mit der iterierten Erwartung gilt

$$\begin{aligned} E(X_1 X_2) &= E(E(X_1 X_2 | X_2)) = E(X_2 E(X_1 | X_2)) = E(X_2 \rho X_2) \\ &= \rho E(X_2^2), \\ E(X_1)E(X_2) &= E(E(X_1|X_2))E(X_2) = E(\rho X_2)E(X_2) = \rho E(X_2)^2, \\ \mathrm{Cov}(X_1, X_2) &= E(X_1 X_2) - E(X_1)E(X_2) = \rho\left(E(X_2^2) - E(X_2)^2\right) \\ &= \rho \mathrm{Var}(X_2) = \rho \sigma_2^2. \end{aligned}$$

Zu (b)

Mit (b) und der Linearität der bedingten Erwartung gilt

$$
\begin{aligned}
E(Z_1|Z_2) = E(Z_1|X_2) = E\left(\frac{X_1-\mu_1}{\sigma_1}\big|X_2\right) &= \frac{1}{\sigma_1}E\left(X_1-\mu_1\big|X_2\right) \\
&= \frac{1}{\sigma_1}\left(E\left(X_1|X_2\right)-\mu_1\right).
\end{aligned}
$$

Mit der Voraussetzung $E(Z_1|Z_2) = \rho Z_2$ folgt durch Ersetzten des ersten Terms der oberen Gleichung und Auflösen nach $E(X_1|X_2)$

$$
\begin{aligned}
\rho Z_2 &= \frac{1}{\sigma_1}\left(E\left(X_1|X_2\right)-\mu_1\right), \\
\rho\cdot\frac{X_2-\mu_2}{\sigma_2} &= \frac{1}{\sigma_1}\left(E\left(X_1|X_2\right)-\mu_1\right), \\
E\left(X_1|X_2\right) &= \mu_1 + \frac{\rho\sigma_1}{\sigma_2}\cdot(X_2-\mu_2).
\end{aligned}
$$

Aufgabe 4 [(a) 3, (b) 8, (c) 3, (d) 8, (e) 8, (f) 4]

Zu (a)
Bezeichne mit f die gesuchte Dichte. Es gilt $f(x) = 0$ für $x \leq 0$. Sei $x > 0$. Differenzieren von F nach x ergibt

$$\frac{1}{2\theta\sqrt{x}}\exp\left(-\frac{\sqrt{x}}{\theta}\right).$$

Zu (b)
Sei $x_1, \ldots, x_n > 0$ eine Realisierung von $(X_1, \ldots, X_n)$. Aufgrund der Unabhängigkeit der X_i, $i = 1, \ldots, n$, gilt für die Likelihoodfunktion

$$L(\theta; x_1, \ldots, x_n) = \prod_{i=1}^n \frac{1}{2\theta\sqrt{x_i}}\exp\left(-\frac{\sqrt{x_i}}{\theta}\right) = \left(\frac{1}{2\theta}\right)^n \exp\left(-\sum_{i=1}^n\frac{\sqrt{x_i}}{\theta}\right)\prod_{i=1}^n\frac{1}{\sqrt{x_i}}.$$

Es ergibt sich

$$
\begin{aligned}
\ell(\theta) &:= \ln L(\theta) = -n\ln\theta - n\ln 2 - \sum_{i=1}^n\frac{\sqrt{x_i}}{\theta} - \frac{1}{2}\sum_{i=1}^n \ln x_i, \\
\ell'(\theta) &= -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^n\sqrt{x_i}.
\end{aligned}
$$

Nullsetzen der ersten Ableitung ergibt

$$\hat\theta = \frac{1}{n}\sum_{i=1}^n\sqrt{x_i}$$

und wegen

$$
\begin{aligned}
\ell''(\theta) &= \frac{n}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^n\sqrt{x_i}, \\
\ell''(\hat\theta) &= \frac{n}{\hat\theta^2} - \frac{2n}{\hat\theta^2} = -\frac{n}{\hat\theta^2} < 0
\end{aligned}
$$

liegt in $\hat{\theta}$ ein Maximum der Likelihoodfunktion vor. Damit ist die Zufallsvariable

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} \sqrt{X_i}$$

ein Maximum Likelihood Schätzer.
Zu (c)
Sei $y > 0$. Es gilt mit (a)

$$P(\sqrt{X} \leq y) = P(X \leq y^2) = 1 - \exp\left(\frac{\sqrt{y^2}}{\theta}\right) = 1 - \exp\left(\frac{y}{\theta}\right)$$

und es folgt die Behauptung.
Zu (d)
Mit dem Ergebnis aus (b) folgt

$$\frac{2n\hat{\theta}}{\theta} = \frac{2n\frac{1}{n}\sum\limits_{i=1}^{n}\sqrt{X_i}}{\theta} = \frac{2}{\theta}\sum_{i=1}^{n}\sqrt{X_i}.$$

Mit den Eigenschaften der Gamma-Verteilung gilt: Die Summanden $\sqrt{X_i}$ des letzten Terms sind unabhängig und wegen (c) identisch $\mathrm{Exp}\left(\frac{1}{\theta}\right) = \Gamma\left(1, \frac{1}{\theta}\right)$ verteilt. Somit ist deren Summe $\Gamma\left(n, \frac{1}{\theta}\right)$ und schließlich folgt bei der Multiplikation mit $\dfrac{2}{\theta}$

$$\frac{2}{\theta}\sum_{i=1}^{n}\sqrt{X_i} \sim \Gamma\left(n, \frac{\theta}{2} \cdot \frac{1}{\theta}\right) = \Gamma\left(n, \frac{1}{2}\right) = \chi^2_{2n}.$$

Damit ergeben sich

$$\begin{aligned} E\left(\frac{2n}{\theta}\hat{\theta}\right) &= \frac{n}{1/2} = 2n, \text{ also } E(\hat{\theta}) = \theta, \\ \mathrm{Var}\left(\frac{2n}{\theta}\hat{\theta}\right) &= \frac{n}{1/4} = 4n, \text{ also } \mathrm{Var}(\hat{\theta}) = \frac{\theta^2}{n}. \end{aligned}$$

Zu (e)
Sei $\omega \in \Omega$ mit $\frac{2n\hat{\theta}(\omega)}{\theta_0} > \chi^2_{2n,1-\alpha}$. Unter $H_0 : \theta \leq \theta_0$ gilt

$$\frac{2n\hat{\theta}(\omega)}{\theta} \geq \frac{2n\hat{\theta}(\omega)}{\theta_0} > \chi^2_{2n,1-\alpha}$$

und somit die behauptete Teilmengenbeziehung, bzw. durch Auflösen nach $\hat{\theta}$

$$\left\{\hat{\theta} > \frac{\theta_0\chi^2_{2n,1-\alpha}}{2n}\right\} \subset \left\{\hat{\theta} > \frac{\theta\chi^2_{2n,1-\alpha}}{2n}\right\}.$$

Da P ein Wahrscheinlichkeitsmaß ist, folgt mit (d)

$$P\left(\hat{\theta} > \frac{\theta_0 \chi^2_{2n,1-\alpha}}{2n}\right) \leq P\left(\hat{\theta} > \frac{\theta \chi^2_{2n,1-\alpha}}{2n}\right) = P\left(\frac{2n\hat{\theta}}{\theta} > \chi^2_{2n,1-\alpha}\right)$$
$$= 1 - P\left(\frac{2n\hat{\theta}}{\theta} \leq \chi^2_{2n,1-\alpha}\right) = 1 - (1-\alpha) = \alpha.$$

Zu (f)

Es ergibt sich $\hat{\theta} = 0,966$. H_0 wird abgelehnt, wenn $\hat{\theta} > \frac{\chi^2_{20,1-5/100} \cdot 1}{20}$ gilt. Wegen $\chi^2_{20,1-5/100} = 31,410$ und $0,966 < 1,5705$ wird H_0 nicht abgelehnt.