DAV

DEUTSCHE
AKTUARVEREINIGUNG e.V.

Zulassungsprüfung Stochastik

der Deutschen Aktuarvereinigung e. V.

am 15. Oktober 2021

Hinweise:

- Alle Hilfsmittel sind zugelassen (open book).
- Die Gesamtpunktzahl beträgt 120 Punkte. Die Klausur ist bestanden, wenn mindestens 60 Punkte erreicht werden.
- Bitte prüfen Sie die Ihnen vorliegende Prüfungsklausur auf Vollständigkeit. Die Klausur besteht aus 10 Seiten.
- Alle Antworten sind zu begründen, und bei Rechenaufgaben muss der Lösungsweg ersichtlich sein.

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. *[24 Punkte]*

Sei $B \in \mathcal{A}$ fest, sei $\eta : \mathcal{A} \longrightarrow [0, \infty]$,

$$\eta(A) := \mu(A \cap B).$$

(a) *[8 Punkte]* Beweisen Sie, dass η ein Maß auf $\mathcal{A}$ ist.

(b) *[16 Punkte]* Sei $f \geq 0$ messbar. Beweisen Sie

$$\int_\Omega f \, d\eta = \int_\Omega 1_B f \, d\mu.$$

Aufgabe 2. *[26 Punkte]*

Seien $X_1, X_2, \ldots$ unabhängige Zufallsvariablen mit

$$E(X_i) = 0, \; Var(X_i) = 1 \text{ für alle } i \in \mathbb{N}.$$

Für $n \in \mathbb{N}$ sei

$$S_n = \sum_i^n X_i.$$

Seien $n, k \in \mathbb{N}$.

(a) *[3 Punkte]* Beweisen Sie, dass S_n und $S_{n+k} - S_n$ unabhängig sind.

(b) *[8 Punkte]* Bestimmen Sie Erwartungswert und Varianz von S_n und $S_{n+k} - S_n$.

(c) *[6 Punkte]* Beweisen Sie, dass $E(S_{n+k}|S_n) = S_n$ gilt.

(d) *[9 Punkte]* Beweisen Sie, dass $Var(S_{n+k}|S_n) = k$ gilt. Zeigen Sie hierzu zunächst, dass für unabhängige Zufallsvariablen A, B mit $E(A^2) < \infty$

$$Var(A + B|B) = Var(A)$$

gilt.

Aufgabe 3. *[28 Punkte]*

Sei $Y \sim N(0, \sigma^2)$ mit $\sigma > 0$ und $X = |Y|$.

(a) *[4 Punkte]* Bestimmen Sie die Verteilungsfunktion von X.

(b) *[4 Punkte]* Beweisen Sie, dass für die Dichte f von X gilt:

$$f(x) = \begin{cases} \frac{2}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) & x > 0, \\ 0 & \text{sonst.} \end{cases} \tag{1}$$

(c) *[12 Punkte]* Bestimmen Sie Erwartungswert und Varianz von X.

(d) *[8 Punkte]* Beweisen Sie, dass $X^2 \sim \Gamma\left(\frac{1}{2}, \frac{1}{2\sigma^2}\right)$ gilt.

Aufgabe 4. *[22 Punkte]*

In dieser Aufgabe greifen wir die Verteilung der Aufgabe 3 wieder auf.

(a) *[14 Punkte]* Bestimmen Sie einen Maximum-Likelihood Schätzer für σ^2 für unabhängige und identisch verteilte $X_1, \ldots, X_n$ mit Dichte (1) in Aufgabe 3.

(b) *[4 Punkte]* Seien $X_1, \ldots, X_n$ wie in (a) und sei

$$\widehat{\sigma^2} = \frac{1}{n} \sum_{i=1}^{n} X_i^2$$

ein Schätzer für σ^2. Beweisen Sie, dass $\widehat{\sigma^2} \sim \Gamma\left(\frac{n}{2}, \frac{n}{2\sigma^2}\right)$ gilt.

(c) *[4 Punkte]* Ist der Schätzer $\widehat{\sigma^2}$ in (b) erwartungstreu? Ist er konsistent?

Aufgabe 5. *[20 Punkte]*

Es wird das Auftreten der Krankheit M in Zusammenhang mit der Infektion mit dem Virus C untersucht. Dabei werden die folgenden Gruppen unterschieden:

G1: Personen, die mit C infiziert sind
G2: Personen, die nicht mit C infiziert sind.

Sei p_i die bedingte Wahrscheinlichkeit für eine Person an M zu erkranken, gegeben sie ist aus Gruppe G_i, $i = 1, 2$. In einer Stichprobe von 110000 voneinander unabhängigen Personen werden folgende Daten erhoben:

Gruppe	an M erkrankt	nicht an M erkrankt	Anzahl
G_1	250	9750	10000
G_2	150	99850	100000

(a) *[6 Punkte]* Als Schätzer $\hat{p}_i$ für p_i, $i = 1, 2$ wird die relative Häufigkeit verwendet. Bestimmen Sie Intervallschätzer für p_i, die symmetrisch um die Schätzwerte $\hat{p}_i$ zum Niveau 95 % sind.

(b) *[4 Punkte]* Kommentieren Sie die Ergebnisse in (a) hinsichtlich der Frage über die Gleichheit und Anordnung der beiden Wahrscheinlichkeiten.

(c) *[4 Punkte]* In der Folge sei

$$p_1 = 0,025; \; p_2 = 0,0015.$$

Sei θ die Wahrscheinlichkeit für eine Person mit C infiziert zu sein. Bestimmen Sie die Wahrscheinlichkeit an M zu erkranken, in Abhängigkeit von θ.

(d) *[6 Punkte]* Es wird eine Impfung gegen die Infektion mit C eingeführt. Geimpfte Personen erkranken an M mit der Wahrscheinlichkeit $\rho = 0,005$. Es soll nun eine Risikoeinschätzung in Abhängigkeit von θ aus (c) erfolgen, ob man sich impfen lassen soll. Dabei wird nur die Erkrankung M betrachtet. Beschreiben Sie einen Ansatz mit Hilfe des Ergebnisses von (c).

Sollten Sie (c) nicht gelöst haben, verwenden Sie $0,0015 + 0,02\theta$ *als Wahrscheinlichkeit des Auftretens von M.*

Lösungsvorschläge

Aufgabe 1 [24 Punkte]

Zu (a) [8 Punkte]

Es gilt

$$\eta(\emptyset) = \mu(B \cap \emptyset) = \mu(\emptyset) = 0.$$

Seien nun $A_i \in \mathcal{A}$, $i \in \mathbb{N}$ paarweise disjunkt. Es gilt

$$\left(\bigcup_{i\in\mathbb{N}} A_i\right) \cap B = \bigcup_{i\in\mathbb{N}} (A_i \cap B)$$

und die $B \cap A_i$, $i \in \mathbb{N}$, sind paarweise disjunkt. Damit folgt aus der σ-Additivität von μ

$$\eta\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \mu\left(\left(\bigcup_{i\in\mathbb{N}} A_i\right) \cap B\right) = \mu\left(\bigcup_{i\in\mathbb{N}} (A_i \cap B)\right) = \sum_{i\in\mathbb{N}} \mu(A_i \cap B) = \sum_{i\in\mathbb{N}} \eta(A_i).$$

Zu (b) [16 Punkte]

Sei $f = \sum_{i=1}^{n} \alpha_i 1_{A_i}$, $\alpha_i \geq 0$, $A_i \in \mathcal{A}$, $i = 1, \ldots, n$, eine Treppenfunktion. Dann gilt laut Konstruktion des Integrals

$$\int_\Omega f\, d\eta = \sum_{i=1}^{n} \alpha_i \eta(A_i) = \sum_{i=1}^{n} \alpha_i \mu(A_i \cap B).$$

Andererseits gilt wegen $1_B 1_{A_i} = 1_{B\cap A_i}$

$$\begin{aligned}\int_\Omega 1_B f\, d\mu &= \int_\Omega 1_B \sum_{i=1}^{n} \alpha_i 1_{A_i} dP = \sum_{i=1}^{n} \alpha_i \int_\Omega 1_B 1_{A_i} dP = \sum_{i=1}^{n} \alpha_i \int_\Omega 1_{A_i \cap B} dP = \sum_{i=1}^{n} \alpha_i \mu(A_i \cap B) \\ &= \sum_{i=1}^{n} \alpha_i \eta(A_i).\end{aligned}$$

und somit die Behauptung für Treppenfunktionen.

Sei nun $f \geq 0$. Es gibt eine Folge $(f_n)_{n\in\mathbb{N}}$ von Treppenfunktionen mit $f_n \geq 0$ und $f_n \nearrow f$. Da f_n eine Treppenfunktion ist, gilt wie oben gezeigt

$$\int_\Omega f_n\, d\eta = \int_\Omega 1_B f_n\, d\mu.$$

Aus $f_n \nearrow f$ folgt $f_n 1_B \nearrow f 1_B$ und mit dem Satz von der monotonen Konvergenz folgt für die beiden Seiten

$$\begin{aligned}\lim_{n\to\infty} \int_\Omega f_n\, d\eta &= \int_\Omega f\, d\eta, \\ \lim_{n\to\infty} \int_\Omega 1_B f_n\, d\mu &= \int_\Omega 1_B f\, d\mu\end{aligned}$$

und somit insgesamt:

$$\int_\Omega f\, d\eta = \int_\Omega 1_B f\, d\mu.$$

Aufgabe 2 *[26 Punkte]* **Lösung**

Zu (a) *[3 Punkte]*

Es gilt

$$S_n = \sum_{i=1}^{n} X_i,\ S_{n+k} - S_n = \sum_{i=n+1}^{n+k} X_i.$$

Die Summanden von S_n und $S_{n+k} - S_n$ sind also jeweils verschieden. Da diese unabhängig sind, folgt die Behauptung.

Zu (b) *[8 Punkte]*

$$E(S_n) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = 0$$
$$E(S_{n+k} - S_n) = E\left(\sum_{i=n+1}^{n+k} X_i\right) = \sum_{i=n+1}^{n+k} E(X_i) = 0.$$

Wegen der Unabhängigkeit der X_i gilt ferner

$$Var(S_n) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) = n$$
$$Var(S_{n+k} - S_n) = Var\left(\sum_{i=n+1}^{n+k} X_i\right) = \sum_{i=n+1}^{n+k} Var(X_i) = k.$$

Zu (c) *[6 Punkte]*

Es gilt

$$\begin{aligned} E(S_{n+k}|S_n) &= E(S_{n+k} - S_n + S_n|S_n) = E(S_{n+k} - S_n|S_n) + E(S_n|S_n) \\ &= E(S_{n+k} - S_n) + S_n = S_n \end{aligned}$$

Das dritte Gleichheitszeichen gilt, weil $S_{n+k} - S_n, S_n$ unabhängig sind und S_n $\sigma(S_n)$-messbar ist.

Zu (d) *[9 Punkte]*

Wir beweisen zunächst den Hinweis.

$$\begin{aligned}
Var(A+B|B) &= E\left((A+B)^2|B\right) - E\,(A+B|B)^2 \\
&= E\left(A^2+2AB+B^2|B\right) - (E(A|B)+E(B|B)))^2 \\
&= E(A^2|B)+2BE(A|B)+E(B^2|B) - (E(A)+B))^2 \qquad (A,\, B \text{ unabhängig}) \\
&= E(A^2)+2BE(A)+B^2 - \left(E(A)^2+2E(A)B+B^2\right) \\
&= Var(A).
\end{aligned}$$

Daraus erhält man mit $A = S_{n+k} - S_n$ und $B = S_n$

$$Var(S_{n+k}|S_n) = Var(S_{n+k} - S_n + S_n|S_n) = Var(S_{n+k} - S_n) = k.$$

Aufgabe 3 *[28 Punkte]*

Zu (a) *[4 Punkte]*

Sei $x \geq 0$ und Φ die Verteilungsfunktion von $N(0,1)$. Dann gilt wegen $X = |Y|$ mit $Y \sim N(0,\sigma^2)$

$$\begin{aligned}
P(X \leq x) &= P(-x \leq Y \leq x) = P(Y \leq x) - P(Y \leq -x) \\
&= \Phi\left(\frac{x}{\sigma}\right) - \Phi\left(-\frac{x}{\sigma}\right) = \Phi\left(\frac{x}{\sigma}\right) - \left(1 - \Phi\left(\frac{x}{\sigma}\right)\right) = 2\Phi\left(\frac{x}{\sigma}\right) - 1.
\end{aligned}$$

Für $x < 0$ gilt wegen $X \geq 0$ offensichtlich $P(X \leq x) = 0$

Zu (b) *[4 Punkte]*

Wir leiten die Verteilungsfunktion in (a) nach x ab. Sei φ die Dichte von $N(0,1)$. Es folgt für $x > 0$

$$f(x) = \frac{2}{\sigma}\varphi\left(\frac{x}{\sigma}\right) = \frac{2}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{2}$$

und $f(x) = 0$ für $x < 0$.

Zu (c) *[12 Punkte]* (8 E, 4 Var)

Es gilt mit den Bezeichnungen aus (b)

$$\begin{aligned}
E(X) &= \int_0^\infty x f(x)\,dx = \int_0^\infty x \frac{2}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \\
&= \frac{2}{\sqrt{2\pi\sigma^2}} \int_0^\infty x \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \\
&= \frac{2}{\sqrt{2\pi\sigma^2}} \left[-\sigma^2 \exp\left(-\frac{x^2}{2\sigma^2}\right)\right]_0^\infty = \frac{2\sigma^2}{\sqrt{2\pi\sigma^2}} = \frac{2\sigma}{\sqrt{2\pi}}.
\end{aligned}$$

Für die Varianz verwenden wir $X = |Y|,\ Y \sim N(0,1)$

$$\sigma^2 = Var(Y) = E(Y^2) - E(Y)^2 = E(Y^2) = E(X^2).$$

Es folgt

$$Var(X) = E(X^2) - E(X)^2 = \sigma^2 - \frac{4\sigma^2}{2\pi} = \sigma^2\left(1 - \frac{2}{\pi}\right)$$

Zu (d) *[8 Punkte]*

Wir betrachten erneut $X = |Y|$ mit $Y \sim N(0,\sigma^2)$. Die Standardisierung $\frac{Y}{\sigma}$ von Y ist standardnormalverteilt und somit gilt $\frac{X^2}{\sigma^2} = \left(\frac{Y}{\sigma}\right)^2 \sim \chi_1^2 = \Gamma\left(\frac{1}{2},\frac{1}{2}\right)$. Mit den Eigenschaften der Gamma-Verteilung (vgl. Formelsammlung) gilt

$$\frac{X^2}{\sigma^2} \sim \Gamma\left(\frac{1}{2},\frac{1}{2}\right) \Longrightarrow X^2 = \sigma^2\frac{X^2}{\sigma^2} \sim \Gamma\left(\frac{1}{2},\frac{1}{2\sigma^2}\right).$$

Alternative Lösung über die Verteilungsfunktion die Dichte von X^2 bestimmen.

Aufgabe 4 *[22 Punkte]*

Zu (a) *[14 Punkte]*

Die gemeinsame Dichte der unabhängigen Zufallsvariablen $X_i \sim X$, $i = 1,\dots,n$ führt zu einer Likelihood L und ihrer Loglikelihood ℓ:

$$\begin{aligned}
L(\sigma^2) &= \prod_{i=1}^{n} \frac{2}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{X_i^2}{2\sigma^2}\right) \\
&= \left(\frac{2}{\sqrt{2\pi}}\right)^n (\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n} X_i^2\right) \\
\ell(\sigma^2) &= n\ln\left(\frac{2}{\sqrt{2\pi}}\right) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} X_i^2 \\
\frac{d\ell(\sigma^2)}{d\sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n} X_i^2 \\
\frac{d^2\ell(\sigma^2)}{d(\sigma^2)^2} &= \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{i=1}^{n} X_i^2 = \frac{n}{\sigma^4}\left(\frac{1}{2} - \frac{1}{\sigma^2}\cdot\frac{1}{n}\sum_{i=1}^{n} X_i^2\right).
\end{aligned}$$

Für die Nullstelle $\widehat{\sigma^2}$ der ersten Ableitung ergibt sich

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \text{ und } \frac{d^2\ell\left(\widehat{\sigma^2}\right)}{d(\sigma^2)^2} = -\frac{n}{2\sigma^4} < 0$$

und somit der ML-Schätzer

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

Zu (b) *[4 Punkte]*

Nach Aufgabe 3 (d) gilt $X_i^2 \sim \Gamma\left(\frac{1}{2}, \frac{1}{2\sigma^2}\right)$, $i = 1, \ldots, n$. Aus der Unabhängigkeit der X_i folgt die Unabhängigkeit der X_i^2 und somit mit den Eigenschaften der Γ-Verteilung

$$\sum_{i=1}^{n} X_i^2 \sim \Gamma\left(\frac{n}{2}, \frac{1}{2\sigma^2}\right) \text{ und } \frac{1}{n}\sum_{i=1}^{n} X_i^2 \sim \Gamma\left(\frac{n}{2}, \frac{n}{2\sigma^2}\right).$$

Zu (c) *[4 Punkte]* (1 Erwartungswert, 1 Varianz, 2 Konsistenz)

Es gilt

$$E\left(\widehat{\widehat{\sigma^2}}\right) = \frac{\frac{n}{2}}{\frac{n}{2\sigma^2}} = \sigma^2$$

$$Var\left(\widehat{\widehat{\sigma^2}}\right) = \frac{\frac{n}{2}}{\left(\frac{n}{2\sigma^2}\right)^2} = \frac{2\sigma^2}{n}.$$

Damit ist der Schätzer $\widehat{\widehat{\sigma^2}}$ erwartungstreu. Wegen $\lim_{n\to\infty} \dfrac{2\sigma^2}{n} = 0$ und der Erwartungstreue ist er auch konsistent.

Aufgabe 5 *[20 Punkte]*

Zu (a) *[6 Punkte]*

Die Schätzwerte sind

$$\hat{p}_1 = \frac{250}{10000} = 0,025$$

$$\hat{p}_2 = \frac{150}{100000} = 0,0015.$$

Mit $u_{0,975} = 1,96$, $n_1 = 10000$ und $n_2 = 100000$ ergibt sich als Intervallschätzer

$$I_i = \left[\hat{p}_i - 1,96\sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{n_i}}, \hat{p}_i + 1,96\sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{n_i}}\right], \; i = 1,2$$

und somit

$$I_1 = [0,02194,\ 0,02806]; \quad I_2 = [0,001260,00174].$$

Zu (b) *[4 Punkte]*

Die Intervalle sind disjunkt, also kann man die Hypothese $p_1 = p_2$ verwerfen, die Wahrscheinlichkeiten sind nicht gleich. Man kann auch auf $p_1 > p_2$ schließen. Die Wahrscheinlichkeit an M zu erkranken ist für G_1 höher als bei G_2.

Zu (c) *[4 Punkte]*

Sei M bzw. C das Ereignis, dass eine Person am M bzw. C erkrankt. Es gilt mit dem Satz von der totalen Wahrscheinlichkeit

$$\begin{aligned} P(M) &= P(M|C)P(C) + P(M|\overline{C})P(\overline{C}) = p_1\theta + p_2(1-\theta) = (p_1 - p_2)\theta + p_1 \\ &= 0,0235\theta + 0,0015. \end{aligned}$$

Zu (d) *[6 Punkte]*

Eine Imfpung ist sinnvoll, wenn die Wahrscheinlichkeit, mit Impfung an M zu erkranken, kleiner ist als die Wahrscheinlichkeit in (c) also, wenn

$$0,0235\theta + 0,0015 > 0,005 \iff \theta > \frac{7}{47} \approx 0,149$$

gilt.