DAV

DEUTSCHE
AKTUARVEREINIGUNG e.V.

Zulassungsprüfung Stochastik

der Deutschen Aktuarvereinigung e. V.

am 15. Oktober 2022

Hinweise:

- Als Hilfsmittel ist ein Taschenrechner zugelassen.
- 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 4 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]*

(a) Gegeben sei eine monoton fallende Folge $(f_n)_{n\in\mathbb{N}}$ von messbaren Funktionen $f_n : \Omega \longrightarrow [0,\infty)$. Ferner gebe es ein $n_0 \in \mathbb{N}$ so, dass f_{n_0} integrierbar ist. Beweisen Sie: Es gibt $f : \Omega \longrightarrow [0,\infty)$ integrierbar mit

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

(b) Für $n \in \mathbb{N}$ sei $f_n : \mathbb{R} \longrightarrow [0,\infty)$ gegeben durch $f_n = 1 - 1_{[-n,n]}$.

(i) Zeigen Sie, dass f_n messbar ist.

(ii) Zeigen Sie, dass die Folge $(f_n)_{n\in\mathbb{N}}$ monoton fallend ist mit $\lim_{n\to\infty} f_n = 0$.

(iii) Gilt $\lim\limits_{n\to\infty} \int_{\mathbb{R}} f_n(x) d\lambda(x) = 0$? Ist das Ergebnis von (a) anwendbar?

Aufgabe 2. *[24 Punkte]*

Ein Produzent von LED-Lampen hat durch umfangreiche Tests festgestellt, dass die Lebensdauer X einer LED-Lampe in Stunden exponential-verteilt ist mit Erwartungswert 15000.

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

(b) Mit welcher Wahrscheinlichkeit funktioniert eine LED-Lampe mindestens 15000 Stunden lang?

(c) Bestimmen Sie die Wahrscheinlichkeit, dass eine LED-Lampe mindestens 25000 Stunden lang funktioniert, wenn bekannt ist, dass sie mindestens 20000 Stunden funktioniert hat.

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

(e) Wie lange funktionieren mindestens 50 % aller LED-Lampen?

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

Aufgabe 3. *[24 Punkte]*

Ein Versicherungsunternehmen verfügt über einen großen Bestand von Verträgen, die alle gleichartig sind, aber nicht alle dieselbe zukünftige Vertragsdauer besitzen. Für die zukünftige Anzahl N_t der Schäden eines Vertrages mit zukünftiger Laufzeit $t > 0$ wird angenommen, dass $N_t \sim Poi(\lambda t)$ gilt, wobei $\lambda > 0$ eine Konstante ist, die für alle Verträge gleich ist. Des Weiteren wird angenommen, dass für die zufällige Laufzeit T eines zufällig ausgewählten Vertrages $T \sim Exp(\theta)$ mit $\theta > 0$ gilt.

Sei N die Schadenzahl eines beliebigen Versicherungsnehmers des Bestands (also ohne Kenntnis der Vertragsdauer).

(a) Bestimmen Sie $E(N|T)$ und $Var(N|T)$.

(b) Bestimmen Sie $E(N)$ und $Var(N)$.

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

(d) Wählt man $\theta = \lambda$, dann hängen die in (b) bestimmten Größen nicht von den Parametern λ, θ ab. Wie ist das zu erklären ohne auf die Berechnungen in (b) zurückzugreifen?

Aufgabe 4. *[32 Punkte]*

Sei $\mu \in \mathbb{R}$ und für alle $t > 0$ sei $Y_t \sim N(\mu t, t)$. Gegeben seien $t_1, \dots, t_n > 0$ und unabhängige Beobachtungen Y_{t_i}, $i = 1, \dots n$ wobei $t_i > 0$ gelte.

(a) Bestimmen Sie einen ML-Schätzer $\hat{\mu}$.

(b) Sei $\hat{\hat{\mu}} := \frac{1}{n} \sum_{i=1}^{n} \frac{Y_{t_i}}{t_i}$ ein Schätzer für μ. Bestimmen Sie den Erwartungswert und die Verteilung von $\hat{\hat{\mu}}$.

(c) Sei $t > 0$. Bestimmen Sie für $\alpha \in (0,1)$ das Konfidenzintervall $I_\alpha = [t\hat{\hat{\mu}} - \delta_\alpha, \hat{\hat{\mu}} t + \delta_\alpha]$ für das $P(\mu t \in I_\alpha) = 1 - \alpha$ gilt.

Wenn Sie (b) nicht gelöst haben, dann verwenden Sie die (falsche) Annahme, dass $\hat{\hat{\mu}} \sim N\left(\mu, \frac{1}{\sum_{i=1}^{n} t_i}\right)$ gilt.

Aufgabe 5. *[16 Punkte]*

Es liegen Daten von 47 US-Bundesstaaten vor:

- Anzahl der Todesopfer bei Verkehrsunfällen in Tausend und
- Bevölkerung in Millionen.

Bezeichnet man mit x_i die Bevölkerung in Millionen und mit Y_i die Anzahl der Todesopfer bei Verkehrsunfällen, so wird für die Zufallsvariable Y_i das Modell

$$Y_i = a + bx_i + \varepsilon_i,\ i = 1, \ldots, 47$$

untersucht. Die ε_i seien unabhängig und $N(0, \sigma^2)$-verteilt.

(a) Bestimmen Sie Schätzwerte für a, b und σ^2.

(b) Stützen die Daten die Behauptung, dass die Anzahl der Verkehrstoten in einem Bundesstaat nicht von der Bevölkerung abhängt? Die Aussage soll zu einem Niveau von $\alpha = 10$ % erfolgen. Sie können verwenden:

$$\overline{x} = 4,1 \qquad \overline{y} = 0,642$$

$$\sum_{i=1}^{47}(x_i - \overline{x})^2 = 473 \qquad \sum_{i=1}^{47}(y_i - \overline{y})^2 = 9,5$$

$$\sum_{i=1}^{47}(y_i - \overline{y})(x_i - \overline{x}) = 56.$$

Lösungsvorschläge

Aufgabe 1 [24 Punkte]

Zu (a) [8 Punkte]

Wegen der Monotonie und Nichtnegativität konvergiert die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ punktweise gegen eine Funktion $f \geq 0$. Ferner folgt wegen $f_n \geq 0$ und der Monotonie für alle $n \geq n_0$

$$0 \leq f_n \leq f_{n_0}$$

wobei f_{n_0} integrierbar ist. Damit folgt mit dem Satz von der dominierten Konvergenz die Integrierbarkeit von f und die Gleichung

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

Zu (b) [16 Punkte]

(i) Sei $n \in \mathbb{N}$. $1_{[-n,n]}$ ist als Indikatorfunktion der messbaren Menge $[-n, n]$ messbar. Die konstante Funktion $g : \mathbb{R} \longrightarrow \mathbb{R}$, $g(x) = 1$ ist stetig, also messbar. Damit ist f_n messbar als Differenz zweier messbarer Funktionen.

(ii) Die Monotonie der Folge ergibt sich aus

$$f_n - f_{n+1} = 1_{[-(n+1),n+1]} - 1_{[-n,n]} = 1_{[-(n+1),n]} + 1_{[n,n+1]} \geq 0.$$

Sei $x \in \mathbb{R}$. Dann gilt für alle $n \in \mathbb{N}$ mit $|x| \leq n$ also $-n \leq x \leq n$

$$f_n(x) = 1 - 1_{[-n,n]}(x) = 1 - 1 = 0$$

also $\lim_{n\to\infty} f_n(x) = 0$.

(iii) Die Gleichung $\displaystyle\int_\mathbb{R} f_n d\lambda = 0$ gilt nicht wegen

$$\int_\mathbb{R} f_n(x) d\lambda(x) = \infty$$

für alle $n \in \mathbb{N}$. Das Ergebnis von (a) ist nicht anwendbar, da f_n nicht integrierbar ist, für alle $n \in \mathbb{N}$.

Aufgabe 2 *[24 Punkte]*

Zu (a)

Aus den Angaben folgt wegen $X \sim \text{Exp}(\lambda)$ und $E(X) = \frac{1}{\lambda}$ sofort $\lambda = \frac{1}{15000}$, $\sigma = \sqrt{\frac{1}{\lambda^2}} = 15000$. Außerdem gilt für die Verteilungsfunktion F von X

$$F(x) = \begin{cases} 1 - e^{-\frac{x}{15000}} & x \geq 0 \\ 0 & \text{sonst} . \end{cases}$$

Zu (b)

$$P(X \geq 15000) = 1 - F(15000) = e^{-1} = 0,37.$$

Zu (c)

$$P(X \geq 25000 | X \geq 20000) = \frac{e^{-\frac{5}{3}}}{e^{-\frac{4}{3}}} = e^{-\frac{1}{3}} \approx 0,72.$$

Zu (d)

$$P(15000 - 15000 \leq X \leq 15000 + 15000) = F(30000) - F(0) = 1 - e^{-2} = 0,86$$

Zu (e)

$$P(X \geq t) = 0,5 \iff e^{-t/15000} = 0,5 \iff t = -15000 \ln(0,5) \approx 10397.$$

Während 10397 Stunden funktionieren 50 % der LED-Lampen einwandfrei.

Zu (f)

$$P(X \leq t) = 0,20 \iff 1 - e^{-t/15000} = 0,20 \iff t = -15000 \ln(0,8) \approx 3347.$$

Die Garantiezeit sollte 3347 Stunden betragen.

Aufgabe 3 *[24 Punkte]*

Zu (a) Wegen der Poissonverteilung ergibt sich sofort

$$E(N|T) = \lambda T = Var(N|T)$$

Zu (b) Wir verwenden die iterierten Erwartungen und $E(T) = \frac{1}{\theta}$, $Var(T) = \frac{1}{\theta^2}$:

$$\begin{aligned} E(N) &= E(E(N|T)) = E(\lambda T) = \lambda E(T) = \frac{\lambda}{\theta} \\ Var(N) &= Var(E(N|T)) + E(Var(N|T)) = Var(\lambda T) + E(\lambda T) \\ &= \lambda^2 Var(T) + \lambda E(\lambda T) = \frac{\lambda^2}{\theta^2} + \frac{\lambda}{\theta} \end{aligned}$$

Zu (c) Laut Annahme gilt $P(N=0|T=t)=e^{-\lambda t}$. Somit folgt

$$\begin{aligned} P(N=0) &= \int_0^\infty P(N=0|T=t)\theta e^{-\theta t}\,dt = \int_0^\infty e^{-\lambda t}\theta e^{-\theta t}\,dt \\ &= \theta \int_0^\infty e^{-t(\lambda+\theta)}\,dt = \frac{\theta}{\lambda+\theta} \end{aligned}$$

Damit gilt

$$P(N\geq 1) = 1-P(N=0) = 1-\frac{\theta}{\lambda+\theta} = \frac{\lambda}{\lambda+\theta}$$

Zu (d) Wählt man $\theta=\lambda$, so gilt $\lambda T \sim \mathrm{Exp}(1)$ verteilt. Damit hängt λT von keinem Parameter ab und somit auch die in (b) berechneten Größen.

Aufgabe 4 *[32 Punkte]*

Zu (a) Die Dichte von Y_t ist gegeben durch

$$f(y) = \frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{1}{2}\left(\frac{y-t\mu}{\sqrt{t}}\right)^2\right) = \frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{(y-t\mu)^2}{2t}\right).$$

Da die Y_{t_i}, $i=1,...,n$ unabhängig sind, ergibt sich die gemeinsame Dichte

$$\prod_{i=1}^n \frac{1}{\sqrt{2\pi t_i}}\exp\left(-\frac{(y_{t_i}-t_i\mu)^2}{2t_i}\right).$$

Daraus ergibt sich eine Likelihood

$$L(\mu) = \exp\left(-\sum_{i=1}^n \frac{(y_{t_i}-t_i\mu)^2}{2t_i}\right).$$

Somit folgt für $\ell := \ln(L)$

$$\begin{aligned} \ell(\mu) &= -\sum_{i=1}^n \frac{(y_{t_i}-t_i\mu)^2}{2t_i} \\ \ell'(\mu) &= \sum_{i=1}^n \frac{y_{t_i}-t_i\mu}{t_i}t_i = \sum_{i=1}^n (y_{t_i}-t_i\mu) = \sum_{i=1}^n y_{t_i} - \mu\sum_{i=1}^n t_i \\ \ell''(\mu) &= -\sum_{i=1}^n t_i < 0 \end{aligned}$$

Aus der zweiten Gleichung ergibt sich durch Nullsetzen und Auflösen

$$\hat{\mu} = \frac{\sum_{i=1}^n y_{t_i}}{\sum_{i=1}^n t_i}$$

und mit $\ell'' < 0$ folgt, dass in $\hat{\mu}$ ein Maximum vorliegt. Der gesuchte Schätzer ist also

$$\hat{\mu} = \frac{\sum_{i=1}^{n} Y_{t_i}}{\sum_{i=1}^{n} t_i}$$

(b) Da Y_t normalverteilt ist, ist auch $\frac{Y_t}{t}$ normalverteilt. Wegen

$$E\left(\frac{Y_t}{t}\right) = \frac{1}{t}E(Y_t) = \frac{1}{t}\mu t = \mu$$
$$Var\left(\frac{Y_t}{t}\right) = \frac{1}{t^2}Var(Y_t) = \frac{t}{t^2} = \frac{1}{t}$$

gilt also $\frac{Y_t}{t} \sim N\left(\mu, \frac{1}{t}\right)$. Wegen der Unabhängigkeit der Y_{t_i} ist $\hat{\hat{\mu}} = \frac{1}{n}\sum_{i=1}^{n} \frac{Y_{t_i}}{t_i}$ als Linearkombination von unabhängigen, normalverteilten Zufallsvariable ebenfalls normalverteilt und es gilt

$$E\left(\hat{\hat{\mu}}\right) = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu$$
$$Var\left(\hat{\hat{\mu}}\right) = \frac{1}{n^2}\sum_{i=1}^{n} \frac{1}{t_i}$$

also

$$\hat{\hat{\mu}} \sim N\left(\mu, \frac{1}{n^2}\sum_{i=1}^{n} \frac{1}{t_i}\right).$$

Zu (c) Mit $\sigma := Var\left(\hat{\hat{\mu}}\right)$ und Φ der Verteilungsfunktion von $N(0,1)$ gilt

$$\begin{aligned}
1-\alpha &= P\left(\hat{\hat{\mu}}t - \delta_\alpha \leq \mu t \leq t\hat{\hat{\mu}} + \delta_\alpha\right) \\
&= P\left(\mu t - \delta_\alpha \leq \hat{\hat{\mu}}t \leq t\mu + \delta_\alpha\right) = P\left(\mu - \frac{\delta_\alpha}{t} \leq \hat{\hat{\mu}} \leq \mu + \frac{\delta_\alpha}{t}\right) \\
&= P\left(\hat{\hat{\mu}} \leq \mu + \frac{\delta_\alpha}{t}\right) - P\left(\hat{\hat{\mu}} \leq \mu - \frac{\delta_\alpha}{t}\right) = \Phi\left(\frac{\delta_\alpha}{t\sigma}\right) - \Phi\left(-\frac{\delta_\alpha}{t\sigma}\right) \\
&= 2\Phi\left(\frac{\delta_\alpha}{t\sigma}\right) - 1
\end{aligned}$$

Auflösen nach δ_α ergibt

$$\delta_\alpha = t u_{1-\frac{\alpha}{2}} \sigma$$

wobei $u_{1-\frac{\alpha}{2}}$ das $1-\frac{\alpha}{2}$-Quantil von $N(0,1)$ bezeichnet. Somit ergibt sich das Konfidenzintervall

$$\left[\hat{\mu}t - \frac{t u_{1-\frac{\alpha}{2}}}{n}\sqrt{\sum_{i=1}^{n}\frac{1}{t_i}},\ \hat{\mu}t + \frac{t u_{1-\frac{\alpha}{2}}}{n}\sqrt{\sum_{i=1}^{n}\frac{1}{t_i}}\right].$$

Mit der falschen Annahme der Aufgabenstellung ergibt sich

$$\left[\hat{\mu}t - \frac{tu_{1-\frac{\alpha}{2}}}{\sqrt{\sum_{i=1}^n t_i}}, \hat{\mu}t - \frac{tu_{1-\frac{\alpha}{2}}}{\sqrt{\sum_{i=1}^n t_i}}\right].$$

Es gilt übrigens mit dem ML-Schätzer aus (a) $Var(\hat{\mu}) = \frac{1}{\sum_{i=1}^n t_i}$

Aufgabe 5 (16 Punkte)

Bemerkung: in diesem Lösungsvorschlag wird jeweils mit den auf drei Nachkommastellen gerundeten Werte weitergerechnet.

Zu (a)

Es ergeben sich folgende Schätzwerte:

$$\begin{aligned}
\hat{b} &= \frac{56}{473} = 0,118 \\
\hat{a} &= \overline{y} - \hat{b}\overline{x} = 0,642 - 0,118 \cdot 4,1 = 0,158 \\
\hat{\sigma}^2 &= \frac{1}{n-2}\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \frac{1}{n-2}\left(\sum_{i=1}^n (y_i - \overline{y})^2 - \frac{\left(\sum\limits_{i=1}^n (y_i - \overline{y})(x_i - \overline{x})\right)^2}{\sum\limits_{i=1}^n (x_i - \overline{x})^2}\right) \\
&= \frac{1}{45}\left(9,5 - \frac{56^2}{473}\right) = 0,064
\end{aligned}$$

Zu (b) Zu prüfen ist die Hypothese $b = 0$.

$$\begin{aligned}
\mathrm{se}(\hat{b}) &= \hat{\sigma}\sqrt{\frac{1}{\sum\limits_{i=1}^n (x_i - \overline{x})^2}} = \sqrt{\frac{0,064}{473}} = 0,012 \\
T &= \frac{\hat{b} - 0}{\mathrm{se}(\hat{b})} = \frac{0,118}{0,011} = 10,727.
\end{aligned}$$

Die Hypothese $b = 0$ wird abgelehnt, denn $|T| > t_{45;0,95} = 1,679$. Somit wird die Hypothese, dass die Anzahl der Unfalltoten nicht vom Alter abhängt, abgelehnt.