DAV

DEUTSCHE AKTUARVEREINIGUNG e.V.

Zulassungsprüfung Stochastik

der Deutschen Aktuarvereinigung e. V.

am 19. Juni 2020

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 9 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 $g : \Omega \longrightarrow \Omega$ messbar. Es gelte

$$\forall A \in \mathcal{A}: \quad \mu(g^{-1}(A)) = \mu(A).$$

(a) Beweisen Sie, dass $1_A(g(\omega)) = 1_{g^{-1}(A)}(\omega)$ für alle $A \in \mathcal{A}$ und $\omega \in \Omega$ gilt. *Hierbei wird mit 1_A die Indikatorfunktion von A bezeichnet.*

(b) Sei $f : \Omega \longrightarrow [0, \infty)$ messbar. Beweisen Sie ohne Verwendung der Transformationsformel für Bildmaße, dass $\int_\Omega f \, d\mu = \int_\Omega f \circ g \, d\mu$ gilt.

Aufgabe 2. *[24 Punkte]*

Gegeben sei ein großer Versicherungsbestand, der aus Verträgen mit unterschiedlichen Vertragsdauern beseht. Für einen zufällig ausgewählten Vertrag bezeichne T die Vertragsdauer und N die Anzahl der gemeldeten Schäden. Es wird angenommen, dass T die Exponentialverteilung $\mathrm{Exp}(\lambda)$ mit $\lambda > 0$ besitzt und dass N unter der Bedingung $\{T = t\}$ mit $t \in [0, \infty)$ die Poisson-Verteilung $\mathrm{Poi}(\alpha t)$ mit $\alpha > 0$ besitzt.

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

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

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

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

Aufgabe 3. *[24 Punkte]*

Eine Person P_1 infiziert sich zum Zeitpunkt A. Sie weist zum Zeitpunkt $A + T$ erste Symptome auf und infiziert zum Zeitpunkt $A + S$ eine weitere Person P_2. Die Zufallsvariablen T und S seien unabhängig und normalverteilt mit $T \sim N(\mu_T, \sigma_T^2)$ und $S \sim N(\mu_Z, \sigma_S^2)$.

(a) Begründen Sie, dass

$$S - T \sim N(\mu_S - \mu_T, \sigma_T^2 + \sigma_S^2)$$

gilt.

(b) Angenommen es gilt $\mu_S = 5,8$ Tage, $\sigma_S^2 = 10,5$, $\mu_T = 5,2$ Tage, $\sigma_T^2 = 14,2$. Bestimmen Sie die Wahrscheinlichkeit, dass P_2 sich

(i) vor Beginn der Symptome von P_1 ansteckt.

(ii) mehr als sieben Tage nach Beginn der Symptome von P_1 ansteckt.

(c) Diskutieren Sie die Verteilungsannahmen der Modellierung.

Aufgabe 4. *[24 Punkte]* Seien $X_1, \ldots, X_m$ Zufallsvariablen, $p \in (0,1)$, $n_1, \ldots, n_m \in \mathbb{N}$ und $X_i \sim B(n_i, p)$, $i = 1, \ldots, m$. Seien $x_1, \ldots, x_m$ mit $x_i \in \{0, \ldots, n_i\}$ unabhängige Realisierungen von $X_1, \ldots, X_m$.

(a) Bestimmen Sie einen Maximum Likelihood Schätzer $\hat{p}$ für p.

(b) Folgende Realisationen sind gegeben:

i	1	2	3	4	5	6	7	8	9	10	$\sum$
n_i	13	8	5	11	12	11	5	7	5	10	87
x_i	1	1	1	1	2	2	2	1	1	1	13

Welcher Schätzwert ergibt sich mit dem Schätzer aus (a)?

(c) Ist der Schätzer $\hat{\hat{p}} := \frac{1}{m} \sum_{i=1}^{m} \frac{X_i}{n_i}$ für p erwartungstreu?

Aufgabe 5. *[24 Punkte]* Aus einer Studie über die Verbreitung von Viren betrachten wir zwei Gruppen. Bei den voneinander unabhängigen Personen wurde ein Virentest mit dem Ergebnis positiv bzw. negativ durchgeführt, es ergaben sich die folgenden Ergebnisse:

Gruppe	Anzahl positiv	Anzahl negativ
1	51	2180
2	430	5520

Bei den positiv getesteten Personen wurde die Virenkonzentration gemessen. Wir gehen davon aus, dass die Virenkozentration X_i einer positiv getesteten Person in der Gruppe $i = 1, 2$ gemäß $X_i \sim N(\mu_i, \sigma_i^2)$ verteilt ist. Die Ergebnisse sind in der folgenden Tabelle zu finden:

Gruppe	Anzahl	Mittelwert $\overline{x}_i$	Empirische Standardabweichung s_i
1	51	4,6	$1,8$
2	430	5,3	$1,9$

Beachten Sie auch die angegebenen Quantile am Ende der Aufgabe!

(a) Wir betrachten die postiv getesteten Personen.

(i) Wie beurteilen Sie bei den positiv getesteten Personen die Hypothese $H_0 : \sigma_1^2 = \sigma_2^2$ zum Niveau von 95 %?

(ii) Wir gehen nun von $\sigma_1^2 = \sigma_2^2$ aus. Testen Sie die Hypothese, dass die Erwartungswerte der Konzentration der positiv getesteten Personen in beiden Gruppen gleich sind, auf dem Niveau von $\alpha = 0,05$.

(b) Sei p_i, $i = 1,2$ die Wahrscheinlichkeit, dass der Test in Gruppe i positiv ausfällt. Bestimmen Sie einen approximativen um $\hat{p}_i$ symmetrischen Intervallschätzer für p_i, $i = 1,2$ zum Niveau 95%. Kommentieren Sie die Ergebnisse hinsichtlich der Frage ob $p_1 = p_2$ gilt.

Sie können folgende q-Quantile der F-Verteilung verwenden:

q	m	n	$F_{m,n,q}$
0,025	50	429	0,6358
0,025	51	430	0,6387
0,025	52	431	0,6416
0,05	50	429	0,685
0,05	51	430	0,6876
0,05	52	431	0,6901

q	m	n	$F_{m,n,q}$
0,95	50	429	1,3805
0,95	51	430	1,3769
0,95	52	431	1,3735
0,975	50	429	1,4671
0,975	51	430	1,4626
0,975	52	431	1,4583

Lösungsvorschläge

Aufgabe 1 [24 Punkte]

Zu (a)

Es gilt für alle $A \in \mathcal{A}$ und $\omega \in \Omega$

$$1_A(g(\omega)) = 1 \Longleftrightarrow g(\omega) \in A \Longleftrightarrow \omega \in g^{-1}(A) \Longleftrightarrow 1_{g^{-1}(A)}(\omega) = 1.$$

Zu (b)

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 nicht negative Treppenfunktion. Dann gilt wegen (a)

$$f \circ g = \sum_{i=1}^n \alpha_i 1_{A_i} \circ g = \sum_{i=1}^n \alpha_i 1_{g^{-1}(A_i)}$$

und somit folgt aus

$$\begin{aligned} \int_\Omega f \circ g \, d\mu &= \sum_{i=1}^n \alpha_i \int_\Omega 1_{g^{-1}(A_i)} \, d\mu = \sum_{i=1}^n \alpha_i \mu(g^{-1}(A_i)) = \sum_{i=1}^n \alpha_i \mu(A_i) \\ &= = \sum_{i=1}^n \alpha_i \int_{A_i} d\mu = \int_\Omega \sum_{i=1}^n \alpha_i 1_{A_i} d\mu = \int_\Omega f \, d\mu \end{aligned}$$

die Behauptung für nicht negative 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 nicht negative Treppenfunktion ist, gilt wie oben gezeigt

$$\int_\Omega f_n \, d\mu = \int_\Omega f_n \circ g \, d\mu.$$

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

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

und somit insgesamt :

$$\int_\Omega f \circ g \, d\mu = \int_\Omega f \, d\mu.$$

Aufgabe 2 *[24 Punkte]* **Lösung**

Zu (a)

Es gilt $E(N|T=t)=\alpha t$ und $\text{Var}(N|T=t)=\alpha t$ und somit

$$E(N|T)=\alpha T, \quad \text{Var}(N|T)=\alpha T.$$

Zu (b)

Mit dem Satz vom iterierten Erwartungswert gilt

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

Somit folgt mit (a)

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

Zu (c)

Laut Annahme gilt $\mathbb{P}(N=0|T=t)=\exp(-\alpha t)$. Somit folgt

$$P(N=0) \quad = \quad \lambda\int_0^\infty P(N=0|T=t)e^{-\lambda t}\,dt=\lambda\int_0^\infty e^{-\alpha t-\lambda t}dt=\frac{\lambda}{\alpha+\lambda}.$$

Damit erhält man $P(N\geq 1)=1-P(N=0)=1-\dfrac{\lambda}{\alpha+\lambda}=\dfrac{\alpha}{\alpha+\lambda}$.

Aufgabe 3 *[24 Punkte]*

Zu (a)

Es gilt $-T\sim N(-\mu_T,\sigma_T^2)$, wegen $T\sim N(\mu_T,\sigma_T^2)$. Ferner sind T und S unabhängig, also auch $-T$ und S. Folglich ist auch deren Summe normalverteilt und

$$E(S-T)=\mu_S-\mu_T \text{ und}$$
$$Var(S-T)=Var(S)+Var(-T)=Var(S)+(-1)^2Var(T)=\sigma_S^2+\sigma_T^2.$$

Zu (b)

(i) Zu bestimmen ist

$$P(A+S\leq A+T)=P(S\leq T)=P(S-T\leq 0).$$

Wegen (a) ist $S-T$ normalverteilt mit Erwartungswert $0,6$ und Varianz $24,7$. Somit gilt

$$P(A+S \leq A+T) = \Phi\left(\frac{-0,6}{\sqrt{24,7}}\right) = 1-\Phi(0,12) = 1-0,5478 = 0,4522.$$

(ii) Wie in (i) rechnet man

$$\begin{aligned} P(A+S > A+T+7) = P(S-T>7) = 1-P(S-T \leq 7) = 1-\Phi\left(\frac{6,4}{\sqrt{24,7}}\right) \\ = \Phi(1,29) = 1-0,9015 = 0,0985. \end{aligned}$$

Zu (c)

Die Normalverteilungsannahme ist wegen möglicher negativer Realisierung nicht geeignet. So ist in (b)

$$P(T \leq 0) = \Phi\left(\frac{-5,8}{\sqrt{14,2}}\right) = 1-0,9382 = 0,0618.$$

Somit würde mit einer positiven Wahrscheinlichkeit P_1 vor Ansteckung Symptome zeigen.

Aufgabe 4 *[24 Punkte]*

Zu (a)

Die gemeinsame Dichte der $X_1, \ldots, X_m$ ist gegeben durch

$$L(p) = \prod_{i=1}^{m} \binom{n_i}{x_i} p^{x_i}(1-p)^{n_i-x_i}.$$

Mit Hilfe der Loglikelihood folgt weiter

$$\begin{aligned} \ell(p) &= \sum_{i=1}^{m} \left\{ \ln\left[\binom{n_i}{x_i}\right] + x_i \ln p + (n_i - x_i)\ln(1-p) \right\} \\ \frac{d}{dp}\ell(p) &= \frac{1}{p}\sum_{i=1}^{m} x_i - \frac{1}{1-p}\sum_{i=1}^{n}(n_i - x_i) \\ \frac{d^2}{dp^2}\ell(p) &= -\frac{1}{p^2}\sum_{i=1}^{m} x_i - \frac{1}{(1-p)^2}\sum_{i=1}^{n}(n_i - x_i) < 0. \end{aligned}$$

Nullsetzen der ersten Ableitung und Auflösen ergibt

$$\frac{1}{\hat{p}}\sum_{i=1}^{m} x_i = \frac{1}{1-\hat{p}}\sum_{i=1}^{n}(n_i - x_i)$$
$$(1-\hat{p})\sum_{i=1}^{m} x_i = \hat{p}\sum_{i=1}^{n}(n_i - x_i)$$
$$\hat{p} = \frac{\sum_{i=1}^{m} X_i}{\sum_{i=1}^{m} n_i}.$$

Zu (b)

Einsetzen führt zu

$$\hat{p} = \frac{13}{87} = 0,15.$$

Zu (c)

Es gilt

$$E(\hat{\hat{p}}) = \frac{1}{m}\sum_{i=1}^{m} E\left(\frac{X_i}{n_i}\right) = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{n_i}E\left(X_i\right) = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{n_i}n_i p = p,$$

also ist der Schätzer ewartungstreu.

Aufgabe 5*[24 Punkte]*

Zu (a)

(i) Wir verwenden den F-Test. Die Testgröße ist

$$t = \frac{1,8^2}{1,9^2} = 0,8975.$$

H_0 wird nicht verworfen, wegen $t \notin [F_{50,429;0,025}, F_{50,429;0,975}] = [0,6358,\ 1,4671]$.

(ii) Wir verwenden den zwei Stichproben t-Test. Für die Testgröße wird benötigt

$$s^2 = \frac{50 \cdot 1,8^2 + 429 \cdot 1,9^2}{479} = 3,5714$$
$$t = \frac{4,6 - 5,3}{\sqrt{3,5714}}\sqrt{\frac{51 \cdot 430}{51 + 430}} = -2,5.$$

Für das benötigte 97,5 % Quantil der t-Verteilung mit 479 ist näherungsweise $u_{0,975} = 1,96$. Da $|t| > 1,96$ gilt, wird H_0 verworfen.

Zu (b)

Die Schätzer sind

$$\hat{p}_1 = \frac{51}{2180+51} = 0,0229$$
$$\hat{p}_2 = \frac{430}{5520+430} = 0,0723.$$

Mit $u_{0,975} = 1,96$ ergibt sich

$$\left[\hat{p}_1 - 1,96\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{2231}}, \hat{p}_1 + 1,96\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{2231}}\right] = [0,0167, 0,029]$$
$$\left[\hat{p}_2 - 1,96\sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{5950}}, \hat{p}_2 + 1,96\sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{5950}}\right] = [0,0657, 0,0789]$$

Die beiden Intervalle sind disjunkt, also kann man die Hypothese $p_1 = p_2$ verwerfen, die beiden Wahrscheinlichkeiten sind nicht gleich.