DAV

DEUTSCHE AKTUARVEREINIGUNG e.V.

# Zulassungsprüfung Stochastik

der Deutschen Aktuarvereinigung e. V.

am 12. Mai 2018

*Hinweise:*

- Als Hilfsmittel ist ein Taschenrechner zugelassen.
- Die Gesamtpunktzahl beträgt 90 Punkte. Die Klausur ist bestanden, wenn mindestens 45 Punkte erreicht werden.
- Bitte prüfen Sie die Ihnen vorliegende Prüfungsklausur auf Vollständigkeit. Die Klausur besteht aus 3 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 $(\Omega, \mathcal{A}, \mathbb{P})$ aus. Die Borel $\sigma$-Algebra auf $\mathbb{R}^n$ wird mit $\mathcal{B}^n$ bezeichnet, das Lebesgue Maß auf $\mathbb{R}^n$ wird mit $\lambda^n$ bezeichnet.

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

**Aufgabe 1.** *[20 Punkte]*

Sei $f : \Omega \longrightarrow [0, \infty)$ eine messbare Funktion mit

$$\int_\Omega f \, d\mu = 0$$

gilt. Beweisen Sie, dass dann $f = 0$ fast überall gilt.
*Hinweis: Beweisen Sie die Behauptung zunächst für Treppenfunktionen.*

**Aufgabe 2.** *[20 Punkte]*

Gegeben seien zwei Zufallsvariablen $X$ und $\Theta$. Dabei sei $\Theta \sim \Gamma(\alpha, \lambda)$ verteilt mit $\alpha > 2, \lambda > 0$ und die bedingte Verteilung von $X$ unter $\Theta$ sei $\Gamma(\gamma, \gamma\Theta)$ mit $\gamma > 0$.

(a) *[5 Punkte]* Bestimmen Sie $E(X|\Theta)$ und $\mathrm{Var}(X|\Theta)$.

(b) *[15 Punkte]* Bestimmen Sie $E(X)$ und $\mathrm{Var}(X)$.

*Hinweise: Verwenden Sie die Parametrisierung der Gammaverteilung aus der Formelsammlung. Für die $\Gamma$-Funktion gilt $x\Gamma(x) = \Gamma(x+1)$ für alle $x > 0$. Außerdem gilt* $\int_0^\infty x^{\alpha-1} e^{-\lambda x} \, dx = \dfrac{\Gamma(\alpha)}{\lambda^\alpha}$ *für $\alpha, \lambda > 0$.*

**Aufgabe 3.** *[25 Punkte]*

Seien $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathrm{Exp}(\lambda)$ und $N_i := 1_{\{X_i > 1\}}$, $i = 1, \ldots, n$.

(a) *[4 Punkte]* Begründen Sie $N_1, \ldots, N_n \overset{\text{iid}}{\sim} B(1, p)$ mit $p := e^{-\lambda}$.

(b) *[12 Punkte]* Bestimmen Sie einen ML-Schätzer $\hat{p}$ für $p$.

(c) *[5 Punkte]* Gegeben sei folgende Stichprobe $x_1, \ldots, x_{10}$:

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | $\Sigma x_i$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 4 | 0,5 | 3,3 | 8,2 | 0,4 | 0,6 | 1,6 | 0,2 | 0,9 | 0,2 | 19,9 |

Bestimmen Sie die sich ergebenden Schätzwerte $\hat{p}$, $\hat{\lambda} := -\ln(\hat{p})$.

(d) *[4 Punkte]* Wieso kann es sinnvoll sein $\hat{\lambda}$ aus (c) als Schätzer für $\lambda$ zu verwenden?

**Aufgabe 4** *[25 Punkte] Beachten Sie auch die unten stehenden Quantile*

In einer Studie zum Lernerfolg wurden zwei Lehrmethoden untersucht, A (aktiv) und P (passiv). Verglichen werden die Anzahl der erzielten Punkte $X$ und $Y$ und die Anzahl der bestandenen Prüfungen $N_A$ und $N_P$. Folgende Daten liegen vor:

$$\sum_{i=1}^{25} x_i = 1.316, \qquad \sum_{i=1}^{25} x_i^2 = 71.540, \qquad n_A = 17,$$
$$\sum_{i=1}^{25} y_i = 1.223, \qquad \sum_{i=1}^{25} y_i^2 = 61.762, \qquad n_P = 10.$$

(a) *[10 Punkte]* Die Anzahl der erzielten Punkte sei normalverteilt $N(\mu_A, \sigma^2)$ bzw. $N(\mu_P, \sigma^2)$. Untersuchen Sie, ob sich die Lehrmethoden auf Basis der erzielten Punkte unterscheiden. Formulieren Sie $H_0$ und verwenden Sie das Signifikanzniveau von 5 %.

(b) *[10 Punkte]* Sei $p_A$ bzw. $p_P$ die Wahrscheinlichkeit zu bestehen. Bestimmen Sie mit den gegebenen Daten einen approximativen Intervallschätzer für $p_A$ und $p_P$ zum Niveau 95 %. Die Bestehensgrenze liegt bei 50 Punkten.

(c) *[5 Punkte]* Kommentieren Sie die Ergebnisse von (a) und (b).

Quantile der $t_n$ Verteilung

| $n$ | 60,0% | 66,7% | 75,0% | 80,0% | 87,5% | 90,0% | 95,0% | 97,5% | 99,0% | 99,5% | 99,9% |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 22 | 0,256 | 0,437 | 0,686 | 0,858 | 1,182 | 1,321 | 1,717 | 2,074 | 2,508 | 2,819 | 3,505 |
| 23 | 0,256 | 0,436 | 0,685 | 0,858 | 1,180 | 1,319 | 1,714 | 2,069 | 2,500 | 2,807 | 3,485 |
| 24 | 0,256 | 0,436 | 0,685 | 0,857 | 1,179 | 1,318 | 1,711 | 2,064 | 2,492 | 2,797 | 3,467 |
| 25 | 0,256 | 0,436 | 0,684 | 0,856 | 1,178 | 1,316 | 1,708 | 2,060 | 2,485 | 2,787 | 3,450 |
| 44 | 0,255 | 0,457 | 0,680 | 0,850 | 1,166 | 1,301 | 1,680 | 2,015 | 2,414 | 2,692 | 3,286 |
| 45 | 0,255 | 0,457 | 0,680 | 0,850 | 1,165 | 1,301 | 1,679 | 2,014 | 2,412 | 2,690 | 3,281 |
| 46 | 0,255 | 0,457 | 0,680 | 0,850 | 1,165 | 1,300 | 1,679 | 2,013 | 2,410 | 2,687 | 3,277 |
| 47 | 0,255 | 0,457 | 0,680 | 0,849 | 1,165 | 1,300 | 1,678 | 2,012 | 2,408 | 2,685 | 3,273 |
| 48 | 0,255 | 0,457 | 0,680 | 0,849 | 1,164 | 1,299 | 1,677 | 2,011 | 2,407 | 2,682 | 3,269 |
| 49 | 0,255 | 0,457 | 0,680 | 0,849 | 1,164 | 1,299 | 1,677 | 2,010 | 2,405 | 2,680 | 3,265 |
| 50 | 0,255 | 0,457 | 0,679 | 0,849 | 1,164 | 1,299 | 1,676 | 2,009 | 2,403 | 2,678 | 3,261 |
| $\infty$ | 0,253 | 0,431 | 0,674 | 0,842 | 1,150 | 1,282 | 1,645 | 1,960 | 2,326 | 2,576 | 3,090 |

# Lösungsvorschläge

**Aufgabe 1** [20 Punkte]

**1. Schritt** Ist $f \geq 0$ eine Treppenfunktion, d.h. es gibt $\alpha_i \geq 0$, $A_i \in \mathcal{A}$, $i = 1, \ldots, n$ mit $f = \sum_{i=1}^{n} \alpha_i 1_{A_i}$ und

$$0 = \int_\Omega f \, d\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i).$$

Somit gilt

$$\alpha_i \mu(A_i) = 0 \text{ folglich } \alpha_i = 0 \text{ oder } \mu(A_i) = 0 \text{ für alle } i = 1, \ldots, n.$$

Damit ist

$$A := \bigcup_{\{i:\alpha_i \neq 0\}} A_i$$

als endliche Vereinigung von Nullmengen eine Nullmenge, und es gilt $\{f \neq 0\} \subset A$. Damit gilt $f = 0$ f.ü.

**2. Schritt** Ist nun $f \geq 0$, dann sei $(f_n)_{n \in \mathbb{N}}$ eine Folge von Treppenfunktionen mit $f_n \uparrow f$, $f_n \geq 0$, $n \in \mathbb{N}$. Es folgt

$$0 \leq \int_\Omega f_n \, d\mu \leq \int_\Omega f \, d\mu = 0 \text{ und somit } \int_\Omega f_n \, d\mu = 0 \text{ für alle } n \in \mathbb{N}.$$

Damit gilt mit dem ersten Schritt $f_n = 0$ fast überall für alle $n \in \mathbb{N}$. Ferner gilt

$$\{f \neq 0\} \subset \bigcup_{n \in \mathbb{N}} \{f_n \neq 0\}.$$

Auf der rechten Seite steht eine abzählbare Vereinigung von Nullmengen, also eine Nullmenge. Die Menge auf der linken Seite ist messbar und somit ihrerseits eine Nullmenge.

**Aufgabe 2** [(a) 5, (b) 15]

Zu (a) Laut Voraussetzung gilt

$$E(X|\Theta = \vartheta) = \frac{\gamma}{\gamma\vartheta} = \frac{1}{\vartheta}$$
$$\mathrm{Var}(X|\Theta = \vartheta) = \frac{\gamma}{(\gamma\vartheta)^2} = \frac{1}{\gamma\vartheta^2}.$$

Ersetzt man $\vartheta$ mit $\Theta$ folgt

$$E(X|\Theta) = \frac{1}{\Theta} \text{ und } \mathrm{Var}(X|\Theta) = \frac{1}{\gamma\Theta^2}.$$

Zu (b)

Mit dem Satz von der iterierten Erwartung ergibt sich

$$\begin{aligned} E(X) &= E(E(X|\Theta)) = E\left(\frac{1}{\Theta}\right) \\ \mathrm{Var}(X) &= \mathrm{Var}(E(X|\Theta)) + E(\mathrm{Var}(X|\Theta)) = \mathrm{Var}\left(\frac{1}{\Theta}\right) + E\left(\frac{1}{\gamma\Theta^2}\right) \\ &= E\left(\frac{1}{\Theta^2}\right) - E\left(\frac{1}{\Theta}\right)^2 + \frac{1}{\gamma}E\left(\frac{1}{\Theta^2}\right). \end{aligned}$$

Nun gilt wegen $\Theta \sim \Gamma(\alpha, \lambda)$ und den Hinweisen für $m \in \{1, 2\}$

$$\begin{aligned} E\left(\frac{1}{\Theta^m}\right) &= \int_0^\infty \frac{1}{\vartheta^m}\frac{\lambda^\alpha}{\Gamma(\alpha)}\vartheta^{\alpha-1}e^{-\lambda\vartheta}\,d\vartheta = \frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty \vartheta^{\alpha-1-m}e^{-\lambda\vartheta}\,d\vartheta = \frac{\lambda^\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha-m)}{\lambda^{\alpha-m}} \\ &= \frac{\lambda^m\Gamma(\alpha-m)}{\Gamma(\alpha)} = \frac{\lambda^m\Gamma(\alpha-m)}{(\alpha-1)\ldots(\alpha-m)\Gamma(\alpha-m)} = \frac{\lambda^m}{(\alpha-1)\ldots(\alpha-m)}. \end{aligned}$$

Somit gilt

$$E\left(\frac{1}{\Theta}\right) = \frac{\lambda}{\alpha-1} \text{ und } E\left(\frac{1}{\Theta^2}\right) = \frac{\lambda^2\Gamma(\alpha-2)}{(\alpha-2)(\alpha-1)\Gamma(\alpha-2)} = \frac{\lambda^2}{(\alpha-2)(\alpha-1)}.$$

Schließlich folgt

$$\begin{aligned} E(X) &= E(E(X|\Theta)) = E\left(\frac{1}{\Theta}\right) = \frac{\lambda}{\alpha-1}, \\ \mathrm{Var}(X) &= E\left(\frac{1}{\Theta^2}\right)\left(1+\frac{1}{\gamma}\right) - E\left(\frac{1}{\Theta}\right)^2 = \left(1+\frac{1}{\gamma}\right)\frac{\lambda^2}{(\alpha-2)(\alpha-1)} - \left(\frac{\lambda}{\alpha-1}\right)^2 \\ &= \frac{\lambda^2}{\alpha-1}\left(\frac{1+\gamma}{\gamma(\alpha-2)} - \frac{1}{\alpha-1}\right) = \frac{\lambda^2}{\alpha-1}\left(\frac{(1+\gamma)(\alpha-1)-\gamma(\alpha-2)}{\gamma(\alpha-1)(\alpha-2)}\right) \\ &= \frac{\lambda^2}{\gamma(\alpha-1)^2(\alpha-2)}(\alpha-1+\gamma\alpha-\gamma-\gamma\alpha+2\gamma)) = \frac{\lambda^2(\alpha+\gamma-1))}{\gamma(\alpha-1)^2(\alpha-2)}. \end{aligned}$$

**Aufgabe 3** [(a) 4 (b) 12, (c) 5, (d) 4 ]

Zu (a)

$N_i$, $i = 1, \ldots, n$ nimmt die Werte 0 und 1 an, ist also Bernoulli verteilt. Außerdem gilt

$$P(N_i = 1) = P(X_i > 1) = 1 - P(X_i \leq 1) = 1 - (1 - e^{-\lambda}) = e^{-\lambda}.$$

Die $N_i$ sind unabhängig, wegen $N_i = 1_{(1,\infty)}(X_i)$ und der Unabhängigkeit der $X_i$.

Zu (b)

Die gemeinsame (Zähl-) Dichte $f : \{0,1\}^n \longrightarrow [0,\infty)$ von $(N_1,\ldots,N_n)$ ist auf ihrem Träger wegen der Unabhängigkeit gegeben durch

$$f(n_1,\ldots,n_n) = \prod_{i=1}^{n} p^{n_i}(1-p)^{1-n_i} = p^{\sum_{i=1}^n n_i}(1-p)^{n-\sum_{i=1}^n n_i}.$$

Geht man zur Log-Likelihood-Funktion über folgt

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

Damit ergibt sich das Maximum $\hat{p}$

$$\begin{aligned}
\frac{1}{\hat{p}}\sum_{i=1}^{n} n_i &= \frac{1}{1-\hat{p}}\left(n - \sum_{i=1}^{n} n_i\right)\\
(1-\hat{p})\sum_{i=1}^{n} n_i &= \hat{p}\left(n - \sum_{i=1}^{n} n_i\right)\\
\sum_{i=1}^{n} n_i &= \hat{p}n\\
\hat{p} &= \frac{1}{n}\sum_{i=1}^{n} n_i.
\end{aligned}$$

Somit ergibt sich der Schätzer $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} N_i$.

Zu (c) Nach Definition von $N_i$ ergeben sich die Realisierungen $n_i$:

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | $\Sigma x_i$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 4 | 0,5 | 3,3 | 8,2 | 0,4 | 0,6 | 1,6 | 0,2 | 0,9 | 0,2 | 19,9 |
| $n_i$ | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | |

und mit (b) $\hat{p} = \frac{4}{10} = 0,4$, $\hat{\lambda} = -\ln 0,4 = 0,916$.

Zu (d) Aus $p = e^{-\lambda}$ ergibt sich durch Auflösen $\lambda = -\ln p$, was dem angegebenen Schätzer entspricht.

**Aufgabe 4**[(a) 10 (b) 10 (c) 5]

Zu (a)

Es wird der zwei Stichproben $t$-Test durchgeführt, die Nullhypothese lautet

$H_0$: $\mu_A = \mu_P$.

Zur Bestimmung der Testgröße. Die beiden Stichproben weisen denselben Umfang $n = m = 25$ auf

$$\overline{x} = \frac{1316}{25} = 52,64 \qquad \overline{y} = \frac{1223}{25} = 48,92$$

$$(n-1)s_x^2 = \sum_{i=1}^{n} x_i^2 - n\overline{x}^2 = 2265,76 \qquad (n-1)s_y^2 = \sum_{i=1}^{n} y_i^2 - n\overline{y}^2 = 1932,84$$

$$s_x^2 = 94,406 \qquad s_y^2 = 80,535$$

$$s^2 = \frac{2265,76 + 1932,84}{48} = 87,47 \qquad s = 9,35$$

$$t = \frac{52,64 - 48,92}{\sqrt{87,47}} \sqrt{\frac{25 \cdot 25}{50}} = 1,406.$$

Damit wird $H_0$ nicht abgelehnt, wegen $t_{48,1-\frac{975}{1000}} = 2,011$ und $|t| = 1,406 < 2,011$.

*Bemerkung: in der Prüfung war aus der Angabe heraus nicht klar, ob es sich bei „." um ein Tausendertrennzeichen oder um ein Dezimalkomma handelte. Man kann auch mit dieser Variante rechnen und erhält*

$$\overline{x} = \frac{1,316}{25} = 0,05264 \qquad \overline{y} = \frac{1,223}{25} = 0,04892$$

$$s_x^2 = 2,9779 \qquad s_y^2 = 2,570$$

$$s^2 = \frac{24 \cdot 2,9779 + 24 \cdot 2,570}{48} = 2,7744$$

$$t = \frac{0,05264 - 0,04892}{\sqrt{2,7744}} \sqrt{\frac{25 \cdot 25}{50}} = 0,007896.$$

*Auch hier wird $H_0$ nicht abgelehnt.*

Zu (b)

Die Schätzwerte sind $\hat{p}_A = \frac{17}{25} = 0,68$, $\hat{p}_P = \frac{10}{25} = 0,4$. Die approximativen Schätzintervalle sind wegen

$$u_{1-\frac{975}{1000}} \cdot \sqrt{\frac{0,68(1-0,68)}{25}} \approx 1,96 \cdot 0,0933 \approx 0,18 \text{ bzw.} \tag{A}$$

$$u_{1-\frac{975}{1000}} \cdot \sqrt{\frac{0,6(1-0,6)}{25}} \approx 1,96 \cdot 0,0980 \approx 0,19 \tag{P}$$

gegeben durch

$$\left(\frac{50}{100}, \frac{86}{100}\right) \text{ für } p_A \text{ und } \left(\frac{21}{100}, \frac{59}{100}\right) \text{ für } p_P.$$

Zu (c)

Die beiden Intervalle in (b) schneiden sich, so dass das Ergebnis in (b) nicht dem Ergebnis in (a) widerspricht. Jedoch ist keiner der beiden Schätzwerte im jeweils anderen Schätzintervall enthalten, so dass bei der Bestehensquote ein Unterschied der beiden Methoden denkbar ist, umso mehr, weil hier die Normalverteilungsapproximation verwendet wird.

Auch könnte es eine Rolle spielen, dass die Bestehensgrenze nahe der gemessenen Mittelwerte liegt und bei der vorgenommenen Diskretisierung (Bestehen, nicht Bestehen) $\hat{p}_A$ und $\hat{p}_P$ weiter auseinander liegen als die Mittelwerte der erreichten Punkte.