Wahrscheinlichkeit und Statistik - Lösung
BSc D-INFK

1. (11 Punkte)

a) (3 Punkte) Die Funktion $F_\alpha(x) = e^{-e^{-(x-\alpha)}}$ ist für alle x differenzierbar mit der Ableitung

$$f_\alpha(x) = F'_\alpha(x) = e^{-\alpha}e^{-x}e^{-e^{-\alpha}}e^{-e^{-x}} = e^{-(x-\alpha)}e^{-e^{-(x-\alpha)}}.$$

Sie ist also insbesondere stetig und wegen $f_\alpha(x) > 0$ monoton steigend. Für $x \to \infty$ konvergiert $s = e^{-x}$ gegen Null und damit $F_\alpha(x) = e^{-e^{\alpha}s}$ gegen 1. Für $x \to -\infty$ konvergiert $s = e^{-x}$ gegen ∞ und damit $F_\alpha(x) = e^{-e^{\alpha}s}$ gegen 0. $F_\alpha(x)$ besitzt also alle Eigenschaften einer Verteilungsfunktion und f_α ist die zugehörige Dichte.

b) (4 Punkte) Für den Erwartungswert dieser Verteilung erhält man

$$\begin{aligned} E[X] &= \int_{\mathbb{R}} x f_\alpha(x)dx = \int_{\mathbb{R}} (x-\alpha) f_\alpha(x)dx + \alpha \int_{\mathbb{R}} f_\alpha(x)dx \\ &= \int_{\mathbb{R}} (x-\alpha)e^{-(x-\alpha)}e^{-e^{-(x-\alpha)}}dx + \alpha = \int_{\mathbb{R}} se^{-s}e^{-e^{-s}}ds + \alpha \,. \end{aligned}$$

Durch die Variablensubstitution $u = e^{-s}$ mit $s = -\log u$ und $\frac{ds}{du} = -\frac{1}{u}$ erhält man für das verbleibende Integral

$$\int_{\mathbb{R}} se^{-s}e^{-e^{-s}}ds = \int_{\infty}^{0} (-\log u)ue^{-u}\left(-\frac{1}{u}\right)du = -\int_{0}^{\infty} \log u e^{-u}du = \gamma\,.$$

Also $E[X] = \gamma + \alpha$.

c) (4 Punkte) Es ist $Z(\omega) \leq z$ genau dann, wenn $X(\omega) \leq z$ und $Y(\omega) \leq z$, woraus $\{Z \leq z\} = \{X \leq z\} \cap \{Y \leq z\}$ folgt. Wegen der Unabhängigkeit der Zufallsvariablen X und Y erhält man daraus

$$P[Z \leq z] = P\left[\{X \leq z\} \cap \{Y \leq z\}\right] = P[X \leq z]P[Y \leq z]$$

Damit lautet die Berechnungsformel für die Verteilungsfunktion des Maximums zweier Zufallsvariabler also

$$F_\delta(z) = F_\alpha(z)F_\beta(z)$$

Man erhält

$$F_\delta(z) = e^{-e^{-(z-\alpha)}}e^{-e^{-(z-\beta)}} = e^{-\left(e^\alpha + e^\beta\right)e^{-z}} = e^{-e^\delta e^{-z}} = e^{-e^{-(z-\delta)}}$$

mit $\delta = \log\left(e^\alpha + e^\beta\right)$, also wieder eine Verteilung vom gleichen Typ.

2. (9 Punkte)

a) (3 Punkte) Die erwartete Lebensdauer und die Streuung von der Lebensdauer des i-ten Lamas ist gegeben durch $E[L_i] = \mu$ und $\sigma_{L_i} = 6$. Man erhält $E[\overline{L}_n] = \mu$ und $\sigma_{\overline{L}_n} = \frac{6}{\sqrt{n}}$.

$$\begin{aligned}
P\left[\left|\overline{L}_n - \mu\right| \le 1\right] &= P\left[-1 \le \overline{L}_n - \mu \le 1\right] \\
&= P\left[\mu - 1 \le \overline{L}_n \le \mu + 1\right] \\
&= P\left[\overline{L}_n \le \mu + 1\right] - P\left[\overline{L}_n \le \mu - 1\right] \\
&= F_{\overline{L}_n}[\mu + 1] - F_{\overline{L}_n}[\mu - 1].
\end{aligned}$$

b) (3 Punkte) Wegen des zentralen Grenzwertsatzes nehmen wir nun an, dass $\overline{L}_n$ normalverteilt ist (mit Mittelwert μ und Streuung $\frac{6}{\sqrt{n}}$, wie oben ausgerechnet) und erhalten

$$\begin{aligned}
P[\left|\overline{L}_n - \mu\right| \le 1] &= F_{\overline{L}_n}[\mu + 1] - F_{\overline{L}_n}[\mu - 1] \\
&\approx \Phi\left[\frac{\mu + 1 - \mu}{\frac{6}{\sqrt{n}}}\right] - \Phi\left[\frac{\mu - 1 - \mu}{\frac{6}{\sqrt{n}}}\right] = \Phi\left[\frac{\sqrt{n}}{6}\right] - \Phi\left[-\frac{\sqrt{n}}{6}\right] \\
&= \Phi\left[\frac{\sqrt{n}}{6}\right] - \left[1 - \Phi\left[\frac{\sqrt{n}}{6}\right]\right] = 2\Phi\left[\frac{\sqrt{n}}{6}\right] - 1.
\end{aligned}$$

c) (3 Punkte) Gesucht ist das kleinste n für welches

$$P[\left|\overline{L}_n - \mu\right| \le 1] \ge 0.90.$$

Wir erhalten

$$\begin{aligned}
P[\left|\overline{L}_n - \mu\right| \le 1] \ge 0.90 \quad &\Leftrightarrow \quad 2\Phi\left[\frac{\sqrt{n}}{6}\right] - 1 \ge 0.90 \\
&\Leftrightarrow \quad \Phi\left[\frac{\sqrt{n}}{6}\right] \ge \frac{1.90}{2} = 0.95 \\
&\Leftrightarrow \quad \frac{\sqrt{n}}{6} \ge \Phi^{-1}[0.95] \\
&\Leftrightarrow \quad n \ge 36\left(\Phi^{-1}[0.95]\right)^2 .
\end{aligned}$$

Aus der Tabelle der Standardnormalverteilung lesen wir das 0.95-Quantil ab und erhalten $(\Phi^{-1}[0.95])^2 \approx 1.65.$ und daraus $36\,(\Phi^{-1}[0.95])^2 \approx 36 \times 1.65^2 = 98.01$. Somit muss $n \ge 99$ sein.

3. (12 Punkte)

a) (2 Punkte) Der Erwartungswert von X_i, für jedes $i \in \{1, \ldots, n\}$, ist gegeben durch

$$E[X_i] = \theta E[X_{i-1}] + \underbrace{E[\varepsilon_i]}_{0} = \theta E[X_{i-1}] = \theta^2 E[X_{i-2}] = \cdots = \theta^i E[\underbrace{X_0}_{0}] = 0\,.$$

b) **(2 Punkte)** Die Varianz von X_i, für jedes $i \in \{1, \dots, n\}$, ist gegeben durch

$$\begin{aligned}
\mathrm{Var}[X_i] &= \mathrm{Var}\left[\theta X_{i-1} + \varepsilon_i\right] = \theta^2 \mathrm{Var}\left[X_{i-1}\right] + \underbrace{\mathrm{Var}[\varepsilon_i]}_{\sigma^2} = \theta^2\left(\theta^2 \mathrm{Var}\left[X_{i-2}\right] + \sigma^2\right) + \sigma^2 \\
&= \theta^4\, \mathrm{Var}\left[X_{i-2}\right] + \sigma^2(1+\theta^2) = \theta^{2i}\, \underbrace{\mathrm{Var}\left[X_0\right]}_{0} + \sigma^2(1+\theta^2+\dots+\theta^{2(i-1)}) \\
&= \sigma^2 \sum_{j=0}^{i-1} \theta^{2j} = \sigma^2 \frac{1-\theta^{2i}}{1-\theta^2}\,.
\end{aligned}$$

c) **(2 Punkte)** Da ε_i für $i \in \{1, \dots, n\}$ unabhängig sind, ist die gemeinsame Dichte gegeben durch

$$f_{(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)}(e_1, e_2, \dots, e_n) = f_{\varepsilon_1}(e_1) f_{\varepsilon_2}(e_2) \dots f_{\varepsilon_n}(e_n)\,.$$

Da ε_i für $i \in \{1, \dots, n\}$ die gleiche Verteilung haben, die $\mathcal{N}(0, \sigma^2)$-Verteilung, bekommen wir, dass

$$f_{\varepsilon_1} = f_{\varepsilon_2} = \dots = f_{\varepsilon_n} = f_{\mathcal{N}(0,\sigma^2)}\,.$$

d) **(4 Punkte)** Die Likelihoodfunktion ist gegeben durch

$$L(x_1, \dots, x_n, \theta) = f(x_1) f(x_2 - \theta x_1) \dots f(x_n - \theta x_{n-1}) = f(x_1) \prod_{i=2}^n f(x_i - \theta x_{i-1})\,.$$

Die log-Likelihoodfunktion ist

$$\begin{aligned}
l(x_1, \dots, x_n, \theta) &= \log L(x_1, \dots, x_n, \theta) = \log\left(f(x_1) \prod_{i=2}^n f(x_i - \theta x_{i-1}) \right) \\
&= \log f(x_1) + \sum_{i=2}^n \log f(x_i - \theta x_{i-1}) \\
&= \log\left(\frac{1}{\sqrt{2\pi}\sigma} e^{-x_1^2/(2\sigma^2)} \right) + \sum_{i=2}^n \log\left(\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \theta x_{i-1})^2}{2\sigma^2}} \right) \\
&= -\log\left(\sqrt{2\pi}\sigma\right) - \frac{x_1^2}{2\sigma^2} - (n-1)\log\left(\sqrt{2\pi}\sigma\right) - \sum_{i=2}^n \frac{(x_i - \theta x_{i-1})^2}{2\sigma^2} \\
&= -n\log\left(\sqrt{2\pi}\sigma\right) - \frac{x_1^2}{2\sigma^2} - \sum_{i=2}^n \frac{(x_i - \theta x_{i-1})^2}{2\sigma^2}
\end{aligned}$$

Die Ableitung der log-Likelihoodfunktion ist

$$\frac{\partial l(x_1, \dots, x_n, \theta)}{\partial \theta} = -\frac{\partial}{\partial \theta} \sum_{i=2}^n \frac{(x_i - \theta x_{i-1})^2}{2\sigma^2} = \frac{1}{\sigma^2} \sum_{i=2}^n (x_i - \theta x_{i-1}) x_{i-1}\,.$$

Also haben wir, dass

$$\begin{aligned}
\frac{\partial l(x_1, \dots, x_n, \theta^*)}{\partial \theta} = 0 \quad &\Rightarrow \quad \frac{1}{\sigma^2} \sum_{i=2}^n (x_i - \theta^* x_{i-1}) x_{i-1} = 0 \\
&\Rightarrow \quad \frac{1}{\sigma^2} \sum_{i=2}^n x_i x_{i-1} = \theta^* \frac{1}{\sigma^2} \sum_{i=2}^n x_{i-1}^2 \Rightarrow \theta^* = \frac{\sum_{i=2}^n x_i x_{i-1}}{\sum_{i=2}^n x_{i-1}^2}\,.
\end{aligned}$$

Da die zweite Abteilung

$$\frac{\partial l^2(x_1,\ldots,x_n,\theta)}{\partial\theta^2} = -\frac{1}{\sigma^2}\sum_{i=2}^{n} x_{i-1}^2$$

negativ ist, schliessen wir daraus, dass bei θ^* ein lokales Maximum ist. Das lokale Maximum entspricht dem globalen Maximum. Der gesuchte Maximum-Likelihood-Schätzer ist also

$$\hat{\theta}_{MLE} = \frac{\sum_{i=2}^{n} X_i X_{i-1}}{\sum_{i=2}^{n} X_{i-1}^2}\,.$$

e) **(2 Punkte)** Die lineare Regression ist $X_i = \theta X_{i-1} + \varepsilon_i$. Der Ordinary Least Squares-Schätzer ist gegeben durch $\hat{\theta}_{OLS} = \arg\min_\theta \varphi(\theta)$ mit $\varphi(\theta) = \sum_{i=2}^{n} \varepsilon_n^2(\theta) = \sum_{i=2}^{n} (x_i - \theta x_{i-1})^2$. Die Ableitung der φ Funktion ist

$$\frac{\partial\varphi(x_1,\ldots,x_n,\theta^*)}{\partial\theta} = \frac{\partial}{\partial\theta}\sum_{i=2}^{n}(x_i-\theta x_{i-1})^2 = -2\sum_{i=2}^{n}(x_i-\theta x_{i-1})\,x_{i-1}\,.$$

Also haben wir, dass $\frac{\partial\varphi(x_1,\ldots,x_n,\theta^*)}{\partial\theta} = 0 \Rightarrow \theta^* = \frac{\sum_{i=2}^{n} x_i x_{i-1}}{\sum_{i=2}^{n} x_{i-1}^2}$. Da die zweite Abteilung $\frac{\partial\varphi^2(x_1,\ldots,x_n,\theta)}{\partial\theta^2} = 2\sum_{i=2}^{n} x_{i-1}^2$ positiv ist, schliessen wir daraus, dass bei θ^* ein lokales Minimum ist. Das lokale Minimum entspricht dem globalen Minimum. Der gesuchte Ordinary Least Squares-Schätzer ist also

$$\hat{\theta}_{OLS} = \frac{\sum_{i=2}^{n} X_i X_{i-1}}{\sum_{i=2}^{n} X_{i-1}^2} = \hat{\theta}_{MLE}\,.$$

4. **(8 Punkte)** Seien $X_i \sim \mathcal{N}(\mu,\sigma^2)$ die gemessenen Geschwindigkeiten und wie immer die tatsächlich gemessenen Werte $x_1,\ldots,x_{10}$ Realisierungen davon. Der Parameter μ soll getestet werden. Da die Standardabweichung σ unbekannt ist, muss sie aus den Daten geschätzt werden, was uns zum t-Test führt.

a) **(2 Punkte)** Mit $t_{n-1;1-\frac{\alpha}{2}} = t_{15;0.995} = 2.947$ ergibt sich die Realisierung des 99%-Vertrauensintervalls zu

$$\begin{aligned}
I_{99\%} &= \left[\overline{x}_{16} - t_{15,0.995}\frac{s_{16}}{\sqrt{16}}\ ,\ \overline{x}_{16} + t_{15,0.995}\frac{s_{16}}{\sqrt{16}}\right] \\
&= \left[\overline{x}_{16} - 2.947\frac{s_{16}}{4}\ ,\ \overline{x}_{16} + 2.947\frac{s_{16}}{4}\right] \\
&= \left[98 - 2.947\frac{4.0}{4}\ ,\ 98 + 2.947\frac{4.0}{4}\right] \\
&= [95.053, 100.947]
\end{aligned}$$

b) **(4 Punkte)** Wir wollen wissen, ob das μ der Geschwindigkeit bei 90 km/h liegt. Dazu führen wir einen *zweiseitigen* **t-Test** zum Niveau 5% mit Null- und Alternativhypothese

$H_0 : \mu = \mu_0 := 90$ und $H_A : \mu \neq \mu_0$ durch. Der Verwerfungsbereich ist somit gegeben durch

$$\begin{aligned} K &= (-\infty, t_{n-1,\frac{\alpha}{2}}] \cup [t_{n-1,1-\frac{\alpha}{2}} \infty) = (-\infty, -t_{n-1,1-\frac{\alpha}{2}}] \cup [t_{n-1,1-\frac{\alpha}{2}}, \infty) \\ &= (-\infty, t_{15,0.975}] \cup [t_{15,0.975}, \infty) = (-\infty, -2.131] \cup [2.131, \infty) \,. \end{aligned}$$

Die Realisierung der Teststatistik ist

$$t = \sqrt{16}\frac{\overline{x}_{16} - \mu_0}{s_{16}} = \sqrt{16}\frac{98 - 90}{4} = 8 \in K \,,$$

weshalb die Nullhypothese verworfen wird.

c) **(2 Punkte)** Wir suchen das Niveau α, das gerade den Verwerfungsbereich $[8, \infty)$ hat; das heisst, wir suchen das Niveau α, für das $t_{n-1,1-\alpha} = 8$ gilt. Durch Rückwärtsauslesen aus der Tabelle der t-Quantile sieht man lediglich, dass der Wert zwischen 0.5% und 0% liegen muss. Dieses Niveau heisst P-Wert.