Stochastik - Musterlösung
(BSc D-MAVT / BSc D-MATH / BSc D-MATL)

1. **a)** 1. **b)** 2. **c)** 1. **d)** 3. **e)** 1. **f)** 1. **g)** 2. **h)** 2. **i)** 3. **j)** 2.

2. Sei K das Ereignis, dass die ausgewählte Wasserprobe mit Nitraten kontaminiert ist. Weiter bezeichne R das Ereignis, dass der Test durch die Reagenzmethode positiv ausfällt, d.h. dass sich die Wasserprobe rot verfärbt.

a) Dann gilt

$$\begin{aligned}\mathbb{P}[R] &= \mathbb{P}[R \cap K] + \mathbb{P}[R \cap K^c] \\ &= \mathbb{P}[R|K] \times \mathbb{P}[K] + \mathbb{P}[R|K^c] \times \mathbb{P}[K^c] \\ &= 0.95 \times 0.3 + 0.10 \times 0.7 \\ &= 0.355\end{aligned}$$

für die Wahrscheinlichkeit, dass sich eine Zufällig ausgewählte Wasserprobe rot verfärben wird.

b) Mit der Formel von Bayes erhalten wir

$$\begin{aligned}\mathbb{P}[K|R] &= \frac{\mathbb{P}[R|K]\mathbb{P}[K]}{\mathbb{P}[R]} \\ &= \frac{0.95 \times 0.3}{0.355} \\ &\approx 0.803\end{aligned}$$

für die Wahrscheinlichkeit, dass die getestete Wasserprobe nitrathaltig war.

c) Mit $\tilde{K}$ bezeichnen wir das Ereignis, dass eine im Labor zufällig ausgewählte Wasserprobe nitrathaltig ist, *nachdem* der Fehler des Praktikanten passiert ist. Ausserdem bezeichnet K weiterhin das Ereignis, dass eine im Labor zufällig ausgewählte Wasserprobe nitrathaltig ist, *bevor* der Fehler des Praktikanten passiert ist. Dann gilt

$$\begin{aligned}\mathbb{P}[\tilde{K}] &= \mathbb{P}[K] + (1 - \mathbb{P}[K])p \\ &= 0.3 + 0.7p.\end{aligned}$$

Bitte wenden!

d) Die Wahrscheinlichkeit für ein positives Testergebnis berechnet sich ähnlich wie in Teilaufgabe **a)**. Daher liegt die Wahrscheinlichkeit für einen positiven Test unter 50% solange die Ungleichung

$$\begin{aligned}\mathbb{P}[R \cap \tilde{K}] + \mathbb{P}[R \cap \tilde{K}^c] &= \mathbb{P}[R|\tilde{K}]\mathbb{P}[\tilde{K}] + \mathbb{P}[R|\tilde{K}^c]\mathbb{P}[\tilde{K}^c] \\ &= 0.95(0.3 + 0.7p) + 0.1(1 - (0.3 + 0.7p)) \\ &\leq 0.5,\end{aligned}$$

erfüllt ist. Dies ist der Fall, wenn $p \leq 0.244$.

3. a) Die Wartezeit $T \in [0, \infty)$ ist Exponentialverteilt mit Parameter $\lambda = \frac{2}{3}$. Die Verteilungsfunktion von $T \sim \text{Exp}[\lambda]$ ist gegeben durch

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

wobei $F(x) = \mathbb{P}[T \leq x]$ die Wahrscheinlichkeit beschreibt, dass der erste Anbiss eines Fisches vor dem Zeitpunkt x eintritt. Andreas möchte den Angel-Tagesschein nur dann erwerben, falls mit mindestens 90% Wahrscheinlichkeit der erste Fisch nach spätestens $x = 4$ Stunden anbeisst. Dies lässt sich durch die Ungleichung

$$\mathbb{P}[T \leq 4] \geq 0.9$$

ausdrücken. Mit der Verteilungsfunktion und mit Einsetzen des Parameters $\lambda = \frac{2}{3}$ erhält man

$$\begin{aligned}\mathbb{P}[T \leq 4] = F(4) = 1 - e^{-\frac{2}{3}4} &\geq 0.9 \\ \frac{8}{3} &\geq \ln(10).\end{aligned}$$

Da $\ln(10) \approx 2.303$, ist die obige Ungleichung erfüllt, d.h. die Wahrscheinlichkeit, dass Andreas nicht länger als 4 Stunden warten muss bis der erste Fisch anbeisst liegt über 90%. Andreas sollte daher den Angel-Tagesschein erwerben.

b) Die Wahrscheinlichkeitsdichte von $T \sim \text{Exp}[\lambda]$ ist gegeben durch

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{sonst.} \end{cases}$$

Aus der Dichte der Exponentialverteilung erhalten wir die folgende log-Likelihood Funktion

$$l(\lambda) = n \log(\lambda) - \lambda \sum_{i=1}^{n} t_i.$$

Siehe nächstes Blatt!

Ableiten und gleich null setzen ergibt den folgenden Maximum-Likelihood Schätzer für den Parameter λ

$$\widehat{\lambda} = \left(\frac{1}{n}\sum_{i=1}^{n} t_i\right)^{-1}.$$

Anhand der Angaben aus der Tabelle haben wir $n = 3$ und

$$\frac{1}{\widehat{\lambda}} = \frac{1}{3}(1.90 + 1.17 + 1.45) = \frac{4.52}{3},$$

deshalb ist $\widehat{\lambda} = \frac{3}{4.52} \approx 0.664$.

c) Das erste Moment einer Poi$[\theta]$-verteilten Zufallsvariablen N ist

$$\mathbb{E}[N] = \theta.$$

Aus den Realisierungen $x_1, \ldots, x_n$ der Zufallsvariable N erhalten wir den Schätzer für das erste Moment für den Parameter θ:

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Aus den Angaben aus der Tabelle erhält man $n = 7$ und

$$\frac{1}{7}\sum_{i=1}^{7} x_i = \frac{1}{7}(10 + 11 + 7 + 7 + 6 + 8 + 9) \approx 8.286$$

d) Sei $N(t)$ die Anzahl Fische, die in t Stunden gefangen wurden. Dann gilt

$$N(t) = \text{Poi}[\lambda t],$$

und $\mathbb{E}[N(t)] = \lambda t$. Der Preis des Angel-Tagesscheins beträgt 40 CHF und Fische werden für 7 CHF pro Stück zurückgehauft, deshalb suchen wir das kleinste t, sodass

$$7 \cdot \mathbb{E}[N(t)] = 7 \cdot \frac{2}{3} t \geq 40.$$

Dies gilt wenn $t \geq \frac{60}{7} \approx 8.571$. Andreas müsste deswegen mindestens 8.572 Stunden angeln, damit sich der Angel-Tagesschein in Erwartung auszahlt.

4. a) Die Stichprobe ist gepaart. Der Test ist zweiseitig zum Niveau $\alpha = 20\%$. Zu einem Münzwurf kommt es wenn der Test, unter der Annahme beide seien gleich schnell, nicht verwirft, die Wahrscheinlichkeit dafür ist definitionsgemäß $1 - \alpha = 1 - 0.2 = 80\%$.

Bitte wenden!

b) Die für den Test relevanten Daten ist die Stichprobe $(d_1, \ldots, d_8)$, die Punkte des entsprechenden QQ-Plots liegen für die verhältnismässig kleine Stichprobengrösse von $n = 8$ sehr schön auf einer Geraden, somit ist die Annahme der Normalverteilung gerechtfertigt. Da die Varianz unbekannt ist, werden wir unter den üblichen Annahmen, dass die $D_1, \ldots, D_8$ i.i.d. sind, einen t-Test durchführen.

c) Wir haben $D \sim \mathcal{N}\left(\mu, \sigma^2\right)$ mit μ unbekannt (wird getestet) und σ unbekannt (wird geschätzt). Wir haben $\mu_0 := 0$ mit den Hypothesen

$$H_0 : \mu = \mu_0 \qquad\qquad H_A : \mu \neq \mu_0.$$

Die Teststatistik ist $T = \sqrt{8}\frac{\overline{D}_8}{S_d}$, sie ist unter der Nullhypothese t_7-verteilt. Der Verwerfungsbereich ist daher gegeben durch $V\!B_{20\%} = (-\infty, -t_{7,0.9}] \cup [t_{7,0.9}, \infty) = (-\infty, -1.415] \cup [1.415, \infty)$.

d) Je größer die d_i's, desto schlechter für Carlo. Liegt die Teststatistik in $[1.415, \infty)$, dann wird ohne Münzwurf gegen ihn entschieden.

e) Die Realisierung der Teststatistik ist $t = \sqrt{8}\frac{\overline{d}_8}{s_d} = \sqrt{8}\frac{-0.33}{0.75} = -1.24$. Wegen $-1.24 \notin V\!B_{20\%}$ kann die Nullhypothese nicht verworfen werden. Es muss die Münze entscheiden.