**Probability and Statistics**
**FS 2018**
**Prüfung**
**13.08.2018**
**Dauer: 180 Minuten**

**Name:** ______________________

**Legi-Nummer:** ______________________

---

- Diese Prüfung enthält 14 Seiten (zusammen mit dem Deckblatt) und 10 Aufgaben. Das Formelblatt wird separat verteilt.
- Begründen Sie ihre Lösungen sorgsam. Ergebnisse ohne Rechenwege und Begründungen geben KEINE Punkte.
- Beginnen Sie jede Aufgabe auf einem neuen Blatt.

Punktetabelle (wird von den Korrektoren beschriftet)

| Aufgabe | Punkte | erreichte Punktzahl |
|---|---|---|
| 1 | 10 | |
| 2 | 10 | |
| 3 | 10 | |
| 4 | 10 | |
| 5 | 10 | |
| 6 | 10 | |
| 7 | 10 | |
| 8 | 10 | |
| 9 | 10 | |
| 10 ($*$) | 10 | |
| Gesamt | 90 | |

Die maximal erreichbare Punktzahl ist 90. Aufgabe 10 ist eine **Bonusaufgabe**.

---

# Viel Erfolg

1. (10 Punkte) Die Zufallsvariable $X$ beschreibt die tägliche Arbeitszeit eines Ingenieurs in Stunden und hat folgende Dichte:

$$f(x) = \begin{cases} c(x-7)^\alpha & : \quad 7 \le x \le 10, \\ 0 & : \quad \text{sonst}, \end{cases}$$

wobei $\alpha > -1$.

a) Bestimme die Konstante $c$ in Abhängigkeit von $\alpha$ so, dass $f(x)$ eine Wahrscheinlichkeitsdichte definiert.

b) Wähle $c = \frac{1}{9}$ und $\alpha = 2$. Berechne die Verteilungsfunktion von $X$.

c) Wähle $c = \frac{1}{9}$ und $\alpha = 2$. Berechne die Wahrscheinlichkeit, dass $X$ einen Wert zwischen 8 und 9 Stunden einnimmt.

**Lösung**

a) Die Dichte muss zu eins aufintegrieren, deshalb gilt:

$$\int_7^{10} c\,(x-7)^\alpha\,dx \stackrel{!}{=} 1.$$

Mit Substitution:

$$\begin{aligned} \int_7^{10} c\,(x-7)^\alpha\,dx \quad &\stackrel{y:=x-7}{=} \quad c \cdot \int_0^3 y^\alpha dx \\ &= \quad c \cdot \frac{1}{\alpha+1} y^{\alpha+1} \Big|_0^3 \\ &= \quad c \cdot \frac{1}{\alpha+1} \left(3^{\alpha+1} - 0\right) \\ &= \quad c \cdot \frac{1}{\alpha+1} \cdot 3^{\alpha+1} \end{aligned}$$

d.h. $c = (\alpha+1) \cdot 3^{-(\alpha+1)} > 0$.

b) Für $7 \le x < 10$ gilt: $F(x) = \mathbb{P}(X \le x) = \int_7^x c(z-7)^2 dz = c\frac{(x-7)^3}{3} = \frac{(x-7)^3}{27}$. Insgesamt erhalten wir:

$$F(x) = \begin{cases} 0 & : \quad x < 7, \\ \frac{(x-7)^3}{27} & : \quad 7 \le x < 10, \\ 1 & : \quad x \ge 10. \end{cases}$$

c) $\mathbb{P}(8 < X < 9) = \mathbb{P}(8 < X \le 9) = F(9) - F(8) = \frac{2^3}{3c} - \frac{1^3}{3c} = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$, wobei wir verwendet haben, dass $\mathbb{P}(\{x\}) = 0$ für alle $x \in \mathbb{R}$ gilt.

2. (10 Punkte) Student Christoph fährt jeden Tag mit dem Velo an die ETH. Dabei kommt er an genau zwei Kreuzungen mit Ampeln vorbei. Aus Erfahrung weiss Christoph, dass er an der ersten Ampel mit Wahrscheinlichkeit 25% warten muss. Auf Grund der Ampelsteuerung und Christophs Fahrgeschwindigkeit hängt die Wahrscheinlichkeit, dass Christoph an der zweiten Ampel warten muss davon ab, ob er an der ersten Ampel warten musste oder nicht. Falls er an der ersten Ampel nicht warten musste, ist die Wahrscheinlichkeit an der zweiten auch nicht warten zu müssen $\frac{2}{3}$. Hatte er an der ersten Ampel hingegen "rot", muss er mit Wahrscheinlichkeit $\frac{1}{7}$ auch an der zweiten Ampel warten. Christoph kommt genau dann zu spät zur Vorlesung, wenn er wenigstens an einer Ampel warten muss.

   a) Mit welcher Wahrscheinlichkeit muss Christoph an genau einer Ampel warten?
   b) Mit welcher Wahrscheinlichkeit kommt Christoph zu spät zur Vorlesung?
   c) Mit welcher Wahrscheinlichkeit hatte Christoph an der ersten Ampel "grün", gegeben, dass er an der zweiten Ampel nicht warten musste?

      *Hinweis:* Falls Sie Teil b) nicht gelöst haben, benutzen Sie im Folgenden für die Wahrscheinlichkeit, dass Christoph zu spät zur Vorlesung kommt, den Wert $\frac{1}{2}$.
   d) Mit welcher Wahrscheinlichkeit musste Christoph an der ersten Ampel warten, gegeben, dass er zu spät zur Vorlesung kam? Sind die Ereignisse "Christoph kommt zu spät zur Vorlesung" und "Christoph muss an der ersten Ampel warten" unabhängig?
   e) Christoph fährt im Februar 10-mal an die ETH. Nehmen Sie an, die Ereignisse, dass er an einem der 10 Tage zu spät zur Vorlesung kommt, seien unabhängig. Wie gross ist die Wahrscheinlichkeit, dass er im Februar mindestens zweimal zu spät zur Vorlesung kommt?
      *Hinweis:* Potenzen wie beispielsweise $2^{10}$ müssen nicht weiter vereinfacht werden.

**Lösung** Wir definieren $W_i :=$ "Christoph muss an der $i$-ten Ampel warten" für $i = 1, 2$. Aus der Aufgabenstellung erhält man $\mathbb{P}[W_1] = \frac{1}{4}$, $\mathbb{P}[W_2^c \mid W_1^c] = \frac{2}{3}$, $\mathbb{P}[W_2 \mid W_1] = \frac{1}{7}$.

   a) Die gesuchte Wahrscheinlichkeit ist $\mathbb{P}[(W_1 \cap W_2^c) \cup (W_1^c \cap W_2)]$. Wir berechnen

$$\begin{aligned}\mathbb{P}[W_1 \cap W_2^c] &= \mathbb{P}[W_2^c \mid W_1]\mathbb{P}[W_1] = (1 - \mathbb{P}[W_2 \mid W_1])\mathbb{P}[W_1] \\ &= (1 - \frac{1}{7}) \cdot \frac{1}{4} = \frac{6}{7} \cdot \frac{1}{4} = \frac{6}{28}\end{aligned}$$

   und

$$\begin{aligned}\mathbb{P}[W_1^c \cap W_2] &= \mathbb{P}[W_2 \mid W_1^c]\mathbb{P}[W_1^c] = (1 - \mathbb{P}[W_2 \mid W_1^c])(1 - \mathbb{P}[W_1]) \\ &= (1 - \frac{2}{3}) \cdot (1 - \frac{1}{4}) = \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{4}.\end{aligned}$$

   Da $W_1 \cap W_2^c$ und $W_1^c \cap W_2$ disjunkt sind haben wir also

$$\mathbb{P}[(W_1 \cap W_2^c) \cup (W_1^c \cap W_2)] = \mathbb{P}[W_1 \cap W_2^c] + \mathbb{P}[W_1^c \cap W_2] = \frac{6}{28} + \frac{1}{4} = \frac{13}{28}.$$

b) $S := W_1 \cup W_2$ beschreibt das Ereignis, dass Christoph zu spät zur Vorlesung kommt.

$$\begin{aligned}\mathbb{P}[S] &= 1 - \mathbb{P}[S^c] = 1 - \mathbb{P}[W_1^c \cap W_2^c] \\ &= 1 - \mathbb{P}[W_2^c \mid W_1^c] P[W_1^c] = 1 - \frac{2}{3} \cdot (1 - \frac{1}{4}) = \frac{1}{2}.\end{aligned}$$

c) Mit dem Satz von Bayes erhält man

$$\begin{aligned}\mathbb{P}[W_1^c \mid W_2^c] &= \frac{\mathbb{P}[W_2^c \mid W_1^c]\mathbb{P}[W_1^c]}{\mathbb{P}[W_2^c \mid W_1^c]\mathbb{P}[W_1^c] + \mathbb{P}[W_2^c \mid W_1]\mathbb{P}[W_1]} \\ &= \frac{\frac{2}{3} \cdot (1 - \frac{1}{4})}{\frac{2}{3} \cdot (1 - \frac{1}{4}) + (1 - \frac{1}{7}) \cdot \frac{1}{4}} = \frac{\frac{1}{2}}{\frac{1}{2} + \frac{6}{28}} = \frac{\frac{1}{2}}{\frac{20}{28}} = \frac{7}{10}.\end{aligned}$$

d) Mit der Definition der bedingten Wahrscheinlichkeit folgt

$$\begin{aligned}\mathbb{P}[W_1 \mid S] &= \frac{\mathbb{P}[W_1 \cap S]}{\mathbb{P}[S]} = \frac{\mathbb{P}[W_1 \cap (W_1 \cup W_2)]}{\mathbb{P}[S]} = \frac{\mathbb{P}[W_1]}{\mathbb{P}[S]} \\ &= \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}.\end{aligned}$$

Da $\mathbb{P}[W_1 \mid S] = \frac{1}{2} \neq \frac{1}{4} = \mathbb{P}[W_1]$, sind die Ereignisse $W_1$ und $S$ abhängig.

e) Sei $E_i$, $i = 1, \ldots, 10$, das Ereignis, dass Christoph am $i$-ten Tag zu spät zur Vorlesung kommt. Wir wissen $\mathbb{P}[E_i] = \mathbb{P}[S] = \frac{1}{2}$. Da die $E_i$ unabhängig sind, ist die Anzahl der Verspätungen $X$ binomialverteilt mit Parametern $n = 10$ und $p = \frac{1}{2}$.

$$\begin{aligned}\mathbb{P}[X \geq 2] &= 1 - \mathbb{P}[X = 1] - \mathbb{P}[X = 0] = 1 - \binom{10}{1}\left(\frac{1}{2}\right)^{10} - \left(\frac{1}{2}\right)^{10} \\ &= 1 - \frac{11}{2^{10}} \; \left(= \frac{2^{10} - 11}{2^{10}}\right).\end{aligned}$$

3. (10 Punkte) Sei $X$ geometrisch verteilt mit Parameter $p \in (0,1)$, das heisst $\mathbb{P}(X = k) = (1-p)^k p$, wobei $k \in \mathbb{N}_0 = \{0, 1, 2, ...\}$. Die wahrscheinlichkeitserzeugende Funktion ist definiert als

$$G_X(z) = \mathbb{E}[z^X] = \sum_{k=0}^{\infty} z^k \mathbb{P}(X = k).$$

a) Berechne $G_X(z)$. Für welche $z \in \mathbb{R}$ ist $G_X(z)$ definiert?

b) Was ist $G_X(0)$ und $G_X(1)$?

c) Berechne $G'_X(z)|_{z=1}$. Was ist das?

**Lösung**

a) $G_X(z) = \sum_{k=0}^{\infty} z^k (1-p)^k p = \frac{p}{1-z(1-p)}$ (geometrische Reihe).
Diese Reihe konvergiert genau dann, wenn

$$|z(1-p)| < 1,$$

d.h.

$$z \in \left( \frac{-1}{1-p}, \frac{1}{1-p} \right).$$

Aus $0 < p < 1$ folgt $\frac{1}{1-p} > 1$; insbesondere gilt, dass $G_X(1)$ und $G'_X(1)$ definiert sind.

b) Da wir wollen, dass $G_X(z)$ stetig in $z = 0$ ist, benutzen wir die Konvention $0^0 \equiv 1$. Folglich gilt $G_X(0) = \mathbb{P}(X = 0) = p$. Beachte, dass im Allgemeinen gilt $G_X^{(k)}(0) = k!\mathbb{P}(X = k)$. Klarerweise gilt $G_X(1) = \frac{p}{1-(1-p)} = 1 = \sum_{k=0}^{\infty} \mathbb{P}(X = k)$.

c)

$$G'_X(z) = \left( \frac{p}{1-z(1-p)} \right)' = \frac{(1-p)p}{(1-z(1-p))^2},$$
$$G'_X(1) = \frac{(1-p)p}{p^2} = \frac{1-p}{p}.$$

Das Differenzieren der Reihe innerhalb des Konvergenzradius führt zu

$$G'_X(z)|_{z=1} = \sum_{k=1}^{\infty} k z^{k-1} \mathbb{P}(X = k)|_{z=1} = \sum_{k=1}^{\infty} k \mathbb{P}(X = k) = \mathbb{E}[X].$$

4. (10 Punkte) a) Berechne mit Hilfe des zentralen Grenzwertsatzes

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!}.$$

*Hinweis:* Betrachte i.i.d. Zufallsvariablen $X_1, X_2, \ldots$ mit $X_i \sim \mathrm{POIS}(1)$ und setze $S_n = X_1 + \cdots + X_n$.

**Lösung**

Betrachte $X_1, X_2, \ldots$ i.i.d. mit $X_i \sim \mathrm{POIS}(1)$. Insbesondere gilt $\mu = \mathbb{E}[X_1] = 1 = \mathrm{Var}[X_1] = \sigma^2$. Aussderdem ist $S_n = X_1 + \cdots + X_n$ wieder Poisson-verteilt mit Parameter $n$. Wegen des zentralen Grenzwertsatzes folgt

$$\begin{aligned} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} &= \sum_{k=0}^{n} e^{-n} \frac{n^k}{k!} = \sum_{k=0}^{n} \mathbb{P}[S_n = k] = \mathbb{P}[S_n \leq n] \\ &= \mathbb{P}\left[\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq \frac{n - n\mu}{\sqrt{n}\sigma}\right] = \mathbb{P}\left[\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq 0\right] \xrightarrow{n\to\infty} \Phi(0) = \frac{1}{2}. \end{aligned}$$

5. (10 Punkte) Sei $X$ eine Zufallsvariable mit momenterzeugender Funktion $M(t) = \mathbb{E}[e^{tX}]$.

a) Zeigen Sie, dass
$$\mathbb{P}(X \geq a) \leq e^{-ta} M(t), \text{ für } t > 0$$
und
$$\mathbb{P}(X \leq a) \leq e^{-ta} M(t), \text{ für } t < 0$$
gilt.

b) Betrachten Sie eine Normalverteilte Zufallsvariable $Z \sim \mathcal{N}(0,1)$. Zeigen Sie, dass
$$\mathbb{P}(Z \geq a) \leq e^{-a^2/2}, \text{ für } a > 0$$
und
$$\mathbb{P}(Z \leq a) \leq e^{-a^2/2}, \text{ für } a < 0$$
gilt.

**Lösung**

a) Für $t > 0$ gilt:
$$\mathbb{P}(X \geq a) = \mathbb{P}(e^{tX} \geq e^{ta}) \overset{Chebyshev}{\leq} \mathbb{E}(e^{tX})e^{-ta} = M(t)e^{-ta}.$$
Für $t < 0$ gilt:
$$\mathbb{P}(X \leq a) = \mathbb{P}(e^{tX} \geq e^{ta}) \overset{Chebyshev}{\leq} \mathbb{E}(e^{tX})e^{-ta} = M(t)e^{-ta}.$$

b) Die momenterzeugende Funktion von $Z$ ist

$$\begin{aligned}
M(t) &= \mathbb{E}[e^{tZ}] \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tz} e^{-z^2/2} dz \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(z^2-2tz/2)} dz \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-((z-t)^2/2+t^2/2)} dz \\
&= e^{t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-((z-t)^2/2)} dz \\
&= e^{t^2/2}.
\end{aligned}$$

Durch Teil a) und $M(t) = e^{t^2/2}$ erhalten wir:
$$\mathbb{P}(Z \geq a) \leq e^{-ta} e^{t^2/2}$$
für $t > 0$. Setzen wir $t = a$, dann gilt für $a > 0$
$$\mathbb{P}(Z \geq a) \leq e^{-a^2/2}.$$
Da $Z \sim \mathcal{N}(0,1)$ symmetrisch um 0 ist, gilt für $a < 0$ auch
$$\mathbb{P}(Z \leq a) \leq e^{-(-a)^2/2} = e^{-a^2/2}.$$

6. (10 Punkte) Sei $(\Omega, \mathbb{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $(Z_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen.

   a) Zeige: Falls $Z_n \xrightarrow{\mathbb{P}} c \in \mathbb{R}$, dann gilt für alle beschränkten und stetigen Funktionen $f : \mathbb{R} \to \mathbb{R}$, dass

   $$\mathbb{E}\left(f(Z_n)\right) \to f(c).$$

   b) Zeige: Falls $Z_n \to c \in \mathbb{R}$ in Verteilung konvergiert, dann gilt $Z_n \xrightarrow{\mathbb{P}} c$.

**Lösung**

a) Wähle $\epsilon > 0$. Wir folgern aus der Stetigkeit der Funktion $f$, dass es ein $\delta > 0$ gibt, so dass für alle $x \in [c-\delta, c+\delta]$, $|f(x) - f(c)| \leq \epsilon$. Folglich

$$\begin{aligned}
|\mathbb{E}\left(f(Z_n) - f(c)\right)| &\leq \mathbb{E}\left(|f(Z_n) - f(c)|\right) \\
&\leq \mathbb{E}\left(|f(Z_n) - f(c)|\mathbf{1}_{\{|Z_n - c| \leq \delta\}}\right) + \mathbb{E}\left(|f(Z_n) - f(c)|\mathbf{1}_{\{|Z_n - c| > \delta\}}\right) \\
&\leq \epsilon + \|f\|_\infty \mathbb{P}(|Z_n - c| > \delta) \xrightarrow[n\to\infty]{} \epsilon.
\end{aligned}$$

b) Wähle $\epsilon > 0$ und definiere

$$f_\epsilon : x \mapsto \min\left\{\frac{1}{\epsilon} d(x, [c-\epsilon, c+\epsilon]), 1\right\}.$$

$f_\epsilon$ ist klarerweise eine stetige Funktion. Beachte, dass $f_\epsilon(x) = 0$, falls $x \in [c-\epsilon, c+\epsilon]$ und $f(x) = 1$, falls $|x - c| \geq 2\epsilon$. Also haben wir:

$$\mathbb{P}(|X_n - c| \geq 2\epsilon) \leq f_\epsilon(X_n) \to f_\epsilon(c) = 0.$$

7. (10 Punkte) Seien $X$ und $Y$ unabhängige, $\mathcal{N}(0,1)$ - verteilte Zufallsvariablen. Wir betrachten die Zufallsvariable $Z$, welche definiert ist als

$$Z := \operatorname{sgn}(Y) \cdot X = \begin{cases} X & \text{für } Y > 0, \\ -X & \text{für } Y \leq 0. \end{cases}$$

a) Bestimmen Sie die Verteilung von $Z$.

b) Berechnen Sie die Korrelation von $X$ und $Z$.

c) Bestimmen Sie die Wahrscheinlichkeit $\mathbb{P}(X + Z = 0)$.

d) Sind $X$ und $Z$ unabhängig? Begründen Sie Ihre Antwort durch ein präzises mathematisches Argument.

**Lösung**

a) Verteilungsfunktion von $Z$:

$$\begin{aligned} F_Z(z) &= \mathbb{P}[Z \leq z] = \mathbb{P}[Z \leq z | Y > 0]\mathbb{P}[Y > 0] + \mathbb{P}[Z \leq z | Y \leq 0]\mathbb{P}[Y \leq 0] \\ &\overset{(*)}{=} 1/2\mathbb{P}[X \leq z | Y > 0] + 1/2\mathbb{P}[-X \leq z | Y \leq 0] \\ &\overset{(**)}{=} 1/2\mathbb{P}[X \leq z] + 1/2\mathbb{P}[X \geq -z] = 1/2\left(\Phi(z) + 1 - \Phi(-z)\right) = \Phi(z). \end{aligned}$$

Also ist $Z$ ebenfalls $\mathcal{N}(0,1)$-verteilt.
(*) falls $Y > 0$ ist, so ist $Z = X$ und falls $Y \leq 0$ ist, so ist $Z = -X$.
(**) da $X$ und $Y$ unabhängig sind.

b) Mit $\rho(X, Z) = \frac{\operatorname{Cov}(X,Z)}{\sigma_X \, \sigma_Z}$ folgt aus

$$\begin{aligned} \operatorname{Cov}(X, Z) &= \mathbb{E}[(X - \mathbb{E}[X])(Z - \mathbb{E}[Z])] = \mathbb{E}[XZ] = \mathbb{E}[\operatorname{sign}(Y)\, X^2] \\ &= \mathbb{E}[\operatorname{sign}(Y)]\, \mathbb{E}[X^2] = 0, \end{aligned}$$

dass die Korrelation von $X$ und $Z$ gleich 0 ist.

c) $\mathbb{P}[X + Z = 0] = \mathbb{P}[Z = -X] = \mathbb{P}[Y \leq 0] = 1/2$.

d) Nein, denn sonst wäre $X + Z \sim \mathcal{N}(0, 2)$ und damit $\mathbb{P}[X + Z = 0] = 0$, im Widerspruch zu c).
Alternative Begründung: wenn $X$ und $Z$ unabhängig wären, müsste für alle Intervalle $A$ und $B \subseteq \mathbb{R}$ gelten

$$\mathbb{P}[X \in A, Z \in B] = \mathbb{P}[X \in A]\mathbb{P}[Z \in B],$$

aber z.B. gilt dies nicht für $A = [-1, 1]$ und $B = [2, \infty)$.

8. (10 Punkte) Wir betrachten i.i.d. Zufallsvariablen $X_1, X_2, ...,$ wobei $X_i$ die uniforme Verteilung auf $[0, \vartheta]$ mit $\vartheta > 0$ besitze.

   a) Bestimme den Maximum-Likelihood Schätzer $T_n$ für $\vartheta$ basierend auf den Beobachtungen $(X_1, ..., X_n)$. Warum ist dieser nicht erwartungstreu? (Begründe ohne zu rechnen.)

      **Vorsicht:** Die Likelihood-Funktion ist nicht differenzierbar.

   b) Berechne die Dichtefunktion für $T_n$. Benutze dies, um einen Faktor $c_n$ zu bestimmen, so dass $c_n T_n$ ein erwartungstreuer Schätzer für $\vartheta$ ist.

   c) Wie verhält sich die Varianz von $c_n T_n$ für $n \to \infty$?

   d) Schätze nun, basierend auf den Beobachtungen $(X_1, ..., X_n)$, $\vartheta$ mit der Momenten-Methode, d.h. finde denjenigen Parameter $\vartheta$, für welchen der Erwartungswert gleich dem empirischen Mittel ist.

**Lösung**

a) Wir behandeln gleich den allgemeinen Fall der uniformen Verteilung auf $[a, b]$ und bestimmen den Maximum-Likelihood Schätzer für die Parameter $a$ und $b$. Die gemeinsame Dichte von $X_1, \dots, X_n$ ist gleich

$$f_X^{a,b}(x_1, \dots, x_n) = \frac{1}{(b-a)^n} \prod_{i=1}^{n} 1_{[a,b]}(x_i). \tag{1}$$

Nun muss $f_X^{a,b}(x_1, \dots, x_n)$ für feste $(x_1, \dots, x_n)$ bezüglich $a$ und $b$ maximiert werden. Seien $x_* := \min_{1 \le i \le n} x_i$ und $x^* := \max_{1 \le i \le n} x_i$. Falls $x_*$ oder $x^*$ ausserhalb $[a, b]$ liegen, dann verschwindet die rechte Seite von (1). Also muss die Maximumstelle $(\hat{a}_{\mathrm{ML}}, \hat{b}_{\mathrm{ML}})$ die Bedingungen $\hat{b}_{\mathrm{ML}} \ge x^* \ge x_* \ge \hat{a}_{\mathrm{ML}}$ erfüllen. Für $\hat{a}_{\mathrm{ML}} := x_* = \min_{1 \le i \le n} x_i$ und $\hat{b}_{\mathrm{ML}} := x^* = \max_{1 \le i \le n} x_i$ ist $f_X^{a,b} < f_X^{\hat{a}_{\mathrm{ML}}, \hat{b}_{\mathrm{ML}}}$ für alle $a$, $b$. Somit sind dies die Maximum-Likelihood Schätzungen für $a$ und $b$.

Der Maximum-Likelihood Schätzer für $\vartheta$ ist somit gegeben durch $T_n := X^* := \max(X_1, \dots, X_n)$, diese Zufallsvariable ist nach oben beschränkt durch $\vartheta$, falls also die Varianz nicht verschwindet, muss der Erwartungswert von $T_n$ echt kleiner als $\vartheta$ sein.

b) Die Verteilungsfunktion von $T_n$ ist

$$F_{T_n}(t) = \mathbb{P}(X^* \le t) = \mathbb{P}(X_i \le t \quad \forall i = 1 \dots n) \overset{\text{i.i.d.}}{=} \left(\frac{t}{\vartheta}\right)^n \quad \text{für} \quad t \in [0, \vartheta],$$

Durch Ableiten erhält man die Dichtefunktion $n \frac{1}{\vartheta^n} t^{n-1} 1_{[0,\vartheta]}(t)$. Damit $c_n T_n$ erwartungstreu wird, muss gelten

$$\begin{aligned} \vartheta = \mathbb{E}[c_n T_n] &= \int_0^{\vartheta} c_n t n \frac{1}{\vartheta^n} t^{n-1} \, dt = c_n n \frac{1}{\vartheta^n} \int_0^{\vartheta} t^n \, dt = c_n n \frac{1}{\vartheta^n} \left. \frac{t^{n+1}}{n+1} \right|_{t=0}^{\vartheta} \\ &= c_n \frac{1}{\vartheta^n} \frac{n}{n+1} \vartheta^{n+1} = c_n \frac{n}{n+1} \vartheta, \end{aligned}$$

woraus sich sofort $c_n = \frac{n+1}{n}$ ergibt.

c) Wir berechnen zuerst

$$\begin{aligned}\mathbb{E}[c_n^2 T_n^2] &= \int_0^\vartheta c_n^2 t^2 n \frac{1}{\vartheta^n} t^{n-1}\, dt = c_n^2 n \frac{1}{\vartheta^n} \int_0^\vartheta t^{n+1}\, dt = c_n^2 n \frac{1}{\vartheta^n} \left. \frac{t^{n+2}}{n+2} \right|_0^\vartheta \\ &= c_n^2 \frac{1}{\vartheta^n} \frac{n}{n+2} \vartheta^{n+2} = c_n^2 \frac{n}{n+2} \vartheta^2,\end{aligned}$$

und somit

$$\begin{aligned}\mathrm{Var}[c_n T_n] &= \mathbb{E}[c_n^2 T_n^2] - \mathbb{E}[c_n T_n]^2 = c_n^2 \frac{n}{n+2}\vartheta^2 - c_n^2 \left(\frac{n}{n+1}\right)^2 \vartheta^2 \\ &= c_n^2 \vartheta^2 \left(\frac{n}{n+2} - \frac{n^2}{(n+1)^2}\right) = c_n^2 \vartheta^2 \left(\frac{n(n+1)^2 - n^2(n+2)}{(n+2)(n+1)^2}\right) \\ &= c_n^2 \vartheta^2 \left(\frac{n^3 + 2n^2 + n - (n^3 + 2n^2)}{(n+2)(n+1)^2}\right) = c_n^2 \vartheta^2 \left(\frac{n}{(n+2)(n+1)^2}\right),\end{aligned}$$

geht also mit $\frac{1}{n^2}$ gegen 0, für $c_n = \frac{n+1}{n}$ ergibt sich $\mathrm{Var}[c_n T_n] = \frac{\vartheta^2}{n(n+2)}$.

d) $\mathbb{E}[X_i] = \vartheta/2$ also ist der Moment-Schätzer gegeben durch $\hat{\vartheta}_{\mathrm{MM}} = 2\bar{X}_n$.

9. (10 Punkte) Einer Gruppe von 20 Tennisspielern mittleren Niveaus werden je zwei Tennisschläger zum Testen ausgehändigt. Einer der Schläger ist jeweils mit einer Nylon-Saite bespannt, der andere mit einer synthetischen Darm-Saite. Nach einigen Wochen Testzeit wird jeder Spieler gefragt, ob er Nylon- oder Darm-Saiten bevorzugt. Es sei $p$ der Anteil aller Tennisspieler mittleren Niveaus, die Darm-Saiten bevorzugen und $X$ sei die Anzahl der Spieler unter den 20 Testspielern, die Darm-Saiten bevorzugen. Da Darm-Saiten teurer sind als Nylon-Saiten, betrachten wir die Nullhypothese, dass höchstens die Hälfte der Spieler Darm-Saiten bevorzugt. Wir vereinfachen dies zu $H_0 : p = 0.5$ und werden $H_0$ nur ablehnen, falls der Versuchsausgang eindeutig Darm-Saiten bevorzugt.

   a) Welcher der beiden Verwerfungsbereiche $\{15, 16, 17, 18, 19, 20\}$ oder $\{0, 1, 2, 3, 4, 5\}$ ist eher für den Test geeignet?

   b) Was ist die Wahrscheinlichkeit eines Fehlers erster Art für den gewählten Bereich in a)? Erhält man mit diesem Bereich einen Test zum Level $\alpha = 0.05$? Ist dies der beste Test zum Niveau $\alpha$?

   c) Berechnen Sie die Wahrscheinlichkeit eines Fehlers zweiter Art für den Verwerfungsbereich, den Sie in a) gewählt haben, falls $p = p_1 = 0.6$ und falls $p = p_2 = 0.8$.

   d) Berechnen Sie den P-Wert, falls $X = 13$. Wird $H_0$ auf einem Signifikanzniveau von $\alpha = 0.10$ abgelehnt?

Unten finden Sie eine Tabelle der *kumulierten* Binomialverteilung für $n = 20$; Für Werte $p \geq 0.50$ (über dem grau hinterlegten Bereich) ist abzulesen: $\mathbb{P}[X \leq l] = 1 -$ abgelesener Wert.

| $n$ | $k$ | $p$ | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 | 0.40 | 0.45 | 0.50 | $l$ | $n$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 0 | | 0.1216 | 0.0388 | 0.0115 | 0.0032 | 0.0008 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 19 | 20 |
| | 1 | | 0.3917 | 0.1756 | 0.0692 | 0.0243 | 0.0076 | 0.0021 | 0.0005 | 0.0001 | 0.0000 | 18 | |
| | 2 | | 0.6769 | 0.4049 | 0.2061 | 0.0913 | 0.0355 | 0.0121 | 0.0036 | 0.0009 | 0.0002 | 17 | |
| | 3 | | 0.8670 | 0.6477 | 0.4114 | 0.2252 | 0.1071 | 0.0444 | 0.0160 | 0.0049 | 0.0013 | 16 | |
| | 4 | | 0.9568 | 0.8298 | 0.6296 | 0.4148 | 0.2375 | 0.1182 | 0.0510 | 0.0189 | 0.0059 | 15 | |
| | 5 | | 0.9887 | 0.9327 | 0.8042 | 0.6172 | 0.4164 | 0.2454 | 0.1256 | 0.0553 | 0.0207 | 14 | |
| | 6 | | 0.9976 | 0.9781 | 0.9133 | 0.7858 | 0.6080 | 0.4166 | 0.2500 | 0.1299 | 0.0577 | 13 | |
| | 7 | | 0.9996 | 0.9941 | 0.9679 | 0.8982 | 0.7723 | 0.6010 | 0.4159 | 0.2520 | 0.1316 | 12 | |
| | 8 | | 0.9999 | 0.9987 | 0.9900 | 0.9591 | 0.8867 | 0.7624 | 0.5956 | 0.4143 | 0.2517 | 11 | |
| | 9 | | 1.0000 | 0.9998 | 0.9974 | 0.9861 | 0.9520 | 0.8782 | 0.7553 | 0.5914 | 0.4119 | 10 | |
| | 10 | | 1.0000 | 1.0000 | 0.9994 | 0.9961 | 0.9829 | 0.9468 | 0.8725 | 0.7507 | 0.5881 | 9 | |
| | 11 | | 1.0000 | 1.0000 | 0.9999 | 0.9991 | 0.9949 | 0.9804 | 0.9435 | 0.8692 | 0.7483 | 8 | |
| | 12 | | 1.0000 | 1.0000 | 1.0000 | 0.9998 | 0.9987 | 0.9940 | 0.9790 | 0.9420 | 0.8684 | 7 | |
| | 13 | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9997 | 0.9985 | 0.9935 | 0.9786 | 0.9423 | 6 | |
| | 14 | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9997 | 0.9984 | 0.9936 | 0.9793 | 5 | |
| | 15 | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9997 | 0.9985 | 0.9941 | 4 | |
| | 16 | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9997 | 0.9987 | 3 | |
| | 17 | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9998 | 2 | |
| | 18 | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1 | |
| $n$ | $k$ | $p$ | 0.90 | 0.85 | 0.80 | 0.75 | 0.37 | 0.65 | 0.60 | 0.55 | 0.50 | $l$ | $n$ |

**Lösung**

a) Wir nehmen an, dass die Präferenzen der einzelnen Tennisspieler unabhängig voneinander und gleichverteilt sind, also ist $X$ binomialverteilt mit $n = 20$ und $p$. Die Nullhypothese und Alternative sind

$$H_0 : p = 0.5 \quad H_A : p > 0.5.$$

Wir suchen also einen Verwerfungsbereich, der die Nullhypothese verwirft, falls viel mehr als die Hälfte der Tennisspieler Darm-Saiten bevorzugen, also falls $X$ gross ist. Der einzige Verwerfungsbereich der dies erfüllt ist der erste, also $\{15, 16, 17, 18, 19, 20\}$.

b) Die Wahrscheinlichkeit eines Fehlers erster Art, also die Nullhypothese abzulehnen, obwohl sie stimmt, ist

$$\mathbb{P}_{H_0}[X \geq 15] = \sum_{k=15}^{20} \binom{20}{k} 0.5^k 0.5^{(20-k)} \approx 0.0207.$$

Also liefert der in a) gewählte Verwerfungsbereich einen Test zum Niveau $\alpha = 0.05 \geq 0.0207$. Wir erhalten

$$\mathbb{P}_{H_0}[X \geq 14] \approx 0.0577 > 0.05,$$

also ist der gewählte Verwerfungsbereich der grösstmögliche zum Niveau $\alpha$.

c) Ein Fehler zweiter Art passiert, wenn wir die Nullhypothese nicht verwerfen obwohl sie falsch ist, also

$$\mathbb{P}_{p=p_1=0.6}[X \leq 14] \approx 0.8744 \quad \text{und} \quad \mathbb{P}_{p=p_2=0.8}[X \leq 14] \approx 0.1958.$$

d) Der P-Wert ist

$$\mathbb{P}_{H_0}[X \geq 13] \approx 0.1316,$$

die Nullhypothese wird also auf dem Signifikanzniveau 0.10 nicht verworfen.

10. (10 Punkte) ($*$) Sei $(A_n)_{n\geq 1}$ eine Folge von Ereignissen in einem Wahrscheinlichkeitsraum. Nehme an

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty \,.$$

Beweise dass

$$\mathbb{P}\big( \bigcap_{k\in\mathbb{N}} \bigcup_{n\geq k} A_n \big) = 0 \,.$$

**Lösung**

Da $\sum\limits_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, gilt

$$\lim_{k\to\infty} \sum_{n=k}^{\infty} \mathbb{P}(A_n) = 0.$$

Folglich

$$\mathbb{P}(\bigcap_{k\in\mathbb{N}} \bigcup_{n\geq k} A_n) \leq \mathbb{P}(\bigcup_{n\geq k'} A_n) \leq \sum_{n=k'}^{\infty} \mathbb{P}(A_n), \quad \forall k' \in \mathbb{N},$$

woraus $\mathbb{P}(\bigcap_{k\in\mathbb{N}} \bigcup_{n\geq k} A_n) = 0$ folgt. b)

b) Es bezeichne $\mathcal{A}$ eine nicht leere Menge von Teilmengen von $\Omega \neq \emptyset$. Zeige, dass

$$\sigma(\mathcal{A}) := \bigcap_{\substack{\mathcal{B}\supseteq\mathcal{A} \\ \mathcal{B} \text{ ist } \sigma\text{-Algebra}}} \mathcal{B}$$

die kleinste $\sigma$-Algebra ist, die alle Teilmengen von $\mathcal{A}$ enthält.

**Bemerkung**: Diese $\sigma$-Algebra heisst die von $\mathcal{A}$ erzeugte $\sigma$-Algebra. "Kleinste $\sigma$-Algebra" bedeutet hier, dass jede $\sigma$-Algebra, die die Mengen von $\mathcal{A}$ enthält, auch die von $\sigma(\mathcal{A})$ enthalten muss.

**Lösung** Es bezeichne $\mathcal{M}$ die Menge aller $\sigma$-Algebren, die die Mengen von $\mathcal{A}$ enthalten:

$$\mathcal{M} = \{\mathcal{B} \,|\, \mathcal{B} \subseteq \mathcal{P}(\Omega) \text{ ist } \sigma\text{-Algebra und } A \in \mathcal{B} \text{ für alle } A \in \mathcal{A} \,\}$$

Der Durchschnitt $\mathcal{F} = \cap_{\mathcal{B}\in\mathcal{M}}\mathcal{B}$ von $\sigma$-Algebren ist wieder eine $\sigma$-Algebra: $\Omega \in \mathcal{F}$, falls $A \in \mathcal{F}$ muss auch $A^c \in \mathcal{F}$ und falls $A_1, A_2, \ldots \in \mathcal{F}$ muss auch die Vereinigung $\cup_n A_n \in \mathcal{F}$ sein. Ist $\mathcal{G}$ eine $\sigma$-Algebra, die $\mathcal{A}$ enthält, gilt

$$\sigma(\mathcal{A}) = \bigcap_{\substack{\mathcal{B}\supseteq\mathcal{A} \\ \mathcal{B} \text{ ist } \sigma\text{-Algebra}}} \mathcal{B} = \left( \bigcap_{\substack{\mathcal{B}\supseteq\mathcal{A}, \mathcal{B}\neq\mathcal{G} \\ \mathcal{B} \text{ ist } \sigma\text{-Algebra}}} \mathcal{B} \right) \cap \mathcal{G} \subseteq \mathcal{G}$$

und $\sigma(\mathcal{A})$ ist somit die kleinste $\sigma$-Algebra, die $\mathcal{A}$ enthält.