Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

D–INFK

Prüfung Wahrscheinlichkeit und Statistik

401-0614-00S

Lösungsvorschlag (A)

Multiple Choice

1. Aufgabe **[30 Punkte]**

1.MC1 **[2 Punkte]** Betrachten Sie zwei Ereignisse C und D mit $\mathbb{P}[D] > 0$. Welche der folgenden Aussagen ist im Allgemeinen wahr?

(A) $\mathbb{P}[C \cup D] = \mathbb{P}[C] + \mathbb{P}[D]$,

(B) $\mathbb{P}[C \cup D] = 1 - \mathbb{P}[C \cap D]$,

(C) $\mathbb{P}[C \cup D] = \mathbb{P}[C \mid D]\,\mathbb{P}[D]$,

(D) **True:** $\mathbb{P}[C \cup D] = \mathbb{P}[C] + \mathbb{P}[D] - \mathbb{P}[C \mid D]\mathbb{P}[D]$.

Lösung:

Für zwei Ereignisse C und D mit $\mathbb{P}[D] > 0$ gilt im Allgemeinen:

$$\mathbb{P}[C \cap D] = \mathbb{P}[C|D]\mathbb{P}[D].$$

Da $\mathbb{P}[C \cup D] = \mathbb{P}[C] + \mathbb{P}[D] - \mathbb{P}[C \cap D]$, folgt:

$$\mathbb{P}[C \cup D] = \mathbb{P}[C] + \mathbb{P}[D] - \mathbb{P}[C|D]\mathbb{P}[D].$$

1.MC2 **[3 Punkte]** Sei Y eine Zufallsvariable mit $\mathbb{P}[Y = -2] = \frac{1}{5}$, $\mathbb{P}[Y = 1] = \frac{2}{5}$ und $\mathbb{P}[Y = 3] = \frac{2}{5}$. Was ist die Verteilungsfunktion F_Y von Y?

(A)

$$F_Y(y) = \begin{cases} 0 & \text{für } y < -2, \\ \frac{1}{5} & \text{für } y = -2, \\ \frac{2}{5} & \text{für } y = 1, \\ 1 & \text{für } y = 3, \\ 0 & \text{für } y \notin \{-2, 1, 3\}, \end{cases}$$

(B) **True:**

$$F_Y(y) = \begin{cases} 0 & \text{für } y < -2, \\ \frac{1}{5} & \text{für } -2 \leq y < 1, \\ \frac{3}{5} & \text{für } 1 \leq y < 3, \\ 1 & \text{für } y \geq 3, \end{cases}$$

(C)

$$F_Y(y) = \begin{cases} 0 & \text{für } y < -2, \\ \frac{1}{5} & \text{für } y = -2, \\ \frac{3}{5} & \text{für } y = 1, \\ 1 & \text{für } y = 3, \\ 0 & \text{für } y \notin \{-2, 1, 3\}, \end{cases}$$

(D)

$$F_Y(y) = \begin{cases} 0 & \text{für } y \leq -2, \\ \frac{1}{5} & \text{für } -2 < y \leq 1, \\ \frac{3}{5} & \text{für } 1 < y \leq 3, \\ 1 & \text{für } y > 3. \end{cases}$$

Lösung:

Die Verteilungsfunktion $F_Y(y)$ ist definiert als die Wahrscheinlichkeit, dass Y einen Wert kleiner oder gleich y annimmt:

$$F_Y(y) = \mathbb{P}[Y \leq y].$$

Wir haben die Wahrscheinlichkeiten:

$$\mathbb{P}[Y = -2] = \frac{1}{5}, \quad \mathbb{P}[Y = 1] = \frac{2}{5}, \quad \mathbb{P}[Y = 3] = \frac{2}{5}.$$

Die Verteilungsfunktion $F_Y(y)$ ist daher:

$$F_Y(y) = \begin{cases} 0 & \text{für } y < -2, \\ \frac{1}{5} & \text{für } -2 \leq y < 1, \\ \frac{3}{5} & \text{für } 1 \leq y < 3, \\ 1 & \text{für } y \geq 3. \end{cases}$$

1.MC3 **[3 Punkte]** Seien A, B und C Ereignisse. Welche der folgenden Aussagen ist allgemein **falsch**?

(A) Falls A, B und C unabhängig sind, so sind $A \cup B$ und C unabhängig,

(B) Falls $\mathbb{P}[A \cup B \cup C] = \mathbb{P}[A] + \mathbb{P}[B] + \mathbb{P}[C]$, folgt daraus nicht, dass sie gemeinsam unabhängig sind,

(C) **True:** Falls A und B sowie B und C unabhängig sind, so sind auch A und $B \cup C$ unabhängig,

(D) $\mathbb{P}[A \cup B \cup C] \geq \mathbb{P}[A] + \mathbb{P}[A \cap B \cap C] - \mathbb{P}[B \cap C]$.

Lösung:

Betrachten wir die Unabhängigkeit der Ereignisse A, B und C.

Aussage (B) ist korrekt, da $\mathbb{P}[A \cup B \cup C] = \mathbb{P}[A] + \mathbb{P}[B] + \mathbb{P}[C]$ nur gilt, wenn die Ereignisse paarweise disjunkt sind, was nicht impliziert, dass sie unabhängig sind.

Aussage (A) ist korrekt, weil die Unabhängigkeit der drei Ereignisse A, B und C bedeutet, dass jede Kombination der Vereinigung und der Schnittmengen der Ereignisse ebenfalls unabhängig ist.

Aussage (C) ist falsch, weil die Unabhängigkeit von A und B sowie B und C nicht notwendigerweise die Unabhängigkeit von A und $B \cup C$ bedeutet.

Aussage (D) ist korrekt, da $P[A \cup B \cup C] \geq P[A]$ und $P[A \cap B \cap C] \leq P[B \cap C]$, so

$$P[A \cup B \cup C] \geq P[A] + P[A \cap B \cap C] - P[B \cap C].$$

1.MC4 **[3 Punkte]** Sei X eine reellwertige Zufallsvariable mit Dichte f_X, und sei F_X ihre Verteilungsfunktion. Welche der folgenden Aussagen ist im Allgemeinen **falsch**?

(A) $F_{X^2}(4) = \int_{-2}^{2} f_X(t)dt$,

(B) $\mathbb{P}[a \leq X \leq b] = \int_a^b f_X(t)\,dt - \mathbb{P}[X = a]\,\mathbb{P}[X = b]$ für $-\infty < a \leq b < \infty$,

(C) **True:** $F_{X^4}(1) = 3\int_{-1}^{1} f_X(t)dt$,

(D) $F_X(b) = \int_{-\infty}^{b} f_X(t)dt$.

Lösung:

$\mathbb{P}[a \leq X \leq b] = \int_a^b f_X(t)\,dt$ und $\mathbb{P}[X = a] = \mathbb{P}[X = b] = 0$ gilt.

$F_{X^2}(4) = P[X^2 \leq 4] = P[-2 \leq X \leq 2] = \int_{-2}^{2} f_X(t)\,dt$.

$F_{X^4}(1) = P[X^4 \leq 1] = P[-1 \leq X \leq 1] = \int_{-1}^{1} f_X(t)\,dt$.

$F_X(b) = P[X \leq b] = \int_{-\infty}^{b} f_X(t)\,dt$.

1.MC5 [**2 Punkte**] Welche von den folgenden Funktionen ist **keine** Dichte?

(A) $h(x) = \frac{4}{3}(1 - x^3) \cdot \mathbf{1}_{0 \le x \le 1}$,

(B) $g(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$,

(C) $f(x) = \frac{\pi}{2} \sin(\pi x) \cdot \mathbf{1}_{0 \le x \le 1}$,

(D) **True:** $k(x) = C \log\left(x + \frac{1}{2}\right) \cdot \mathbf{1}_{0 \le x \le 10}$ für einige konstante C.

Lösung:

Aussage (B) ist korrekt. Diese Funktion ist positiv und ihr Integral über den gesamten Definitionsbereich ergibt 1:

$$\int_{-\infty}^{\infty} g(x)\,dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx = 1.$$

Aussage (C) ist korrekt. Diese Funktion ist positiv für $0 \le x \le 1$ und ihr Integral ergibt 1:

$$\int_0^1 f(x)\,dx = \int_0^1 \frac{1}{\pi} \sin(\pi x)\,dx = 1.$$

Aussage (A) ist korrekt. Diese Funktion ist positiv für $0 \le x \le 1$ und ihr Integral ergibt 1:

$$\int_0^1 h(x)\,dx = \int_0^1 \frac{4}{3}(1 - x^3)\,dx = 1.$$

Aussage (D) ist falsch, da $\log(x + \frac{1}{2})$ für $x < \frac{1}{2}$ negativ ist und daher keine Dichtefunktion darstellen kann.

1.MC6 [**3 Punkte**] Seien X und Y zwei reellwertige und unabhängige Zufallsvariablen. Welche der folgenden Aussagen ist allgemein wahr?

(A) **True:** Falls $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ und $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$, so ist $2X - Y \sim \mathcal{N}(2\mu_1 - \mu_2, 4\sigma_1^2 + \sigma_2^2)$,

(B) Falls $X \sim \text{Exp}(\lambda)$ und $Y \sim \text{Exp}(\mu)$ für $\lambda > \mu > 0$, so ist $X - Y \sim \text{Exp}(\lambda - \mu)$,

(C) Falls $X \sim \mathcal{U}(0, 1)$ und $Y \sim \mathcal{U}(1, 2)$, so ist $X + Y \sim \mathcal{U}(0, 2)$,

(D) Falls $X \sim \text{Poisson}(\lambda)$ und $Y \sim \text{Poisson}(\mu)$ für $\lambda, \mu > 0$, so ist $X + 2Y \sim \text{Poisson}(\lambda + 2\mu)$.

Lösung:

Aussage (B) ist falsch, da $P[X - Y < 0] > 0$ gilt.

Aussage (D) ist falsch, da

$$P[X + 2Y = 0] = P[X = 0] \cdot P[Y = 0] = e^{-\lambda} e^{-\mu} \neq e^{-(\lambda + 2\mu)}$$

gilt.

Aussage (A) ist korrekt.

Aussage (D) ist falsch, da $P[X + Y > 2] \geq P[X > 1/2, Y > 3/2] = \frac{1}{4} > 0$.

1.MC7 **[3 Punkte]** Seien X und Y zwei nicht-negative Zufallsvariablen mit den folgenden Eigenschaften:

$$\mathrm{Var}(X) = 9, \quad \mathrm{Var}(Y) = 11, \quad \mathbb{E}[X] = 4, \quad \mathbb{E}[Y^2] = 20, \quad \mathrm{Cov}(X, Y) = 0.$$

Dann ist $\mathrm{Var}\left(\frac{1}{3}(X - Y)\right)$ gleich:

(A) $\frac{44}{9}$,
(B) **True:** $\frac{20}{9}$,
(C) $\frac{44}{3}$,
(D) $\frac{20}{3}$.

Lösung:

Um die Varianz von $\frac{1}{3}(X - Y)$ zu berechnen, verwenden wir die Eigenschaft der Varianz bei Linearkombinationen von Zufallsvariablen:

$$\mathrm{Var}\left(\frac{1}{3}(X - Y)\right) = \left(\frac{1}{3}\right)^2 \mathrm{Var}(X - Y).$$

Da X und Y unkorreliert sind, ist die Kovarianz $\mathrm{Cov}(X, Y) = 0$ und daher:

$$\mathrm{Var}(X - Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Es folgt:

$$\mathrm{Var}(X - Y) = 9 + 11 = 20.$$

Somit erhalten wir:

$$\mathrm{Var}\left(\frac{1}{3}(X - Y)\right) = \left(\frac{1}{3}\right)^2 \cdot 20 = \frac{20}{9}.$$

1.MC8 [4 Punkte]

Sei (X_n) i.i.d. $\mathcal{U}(0,5)$-verteilt. Wir betrachten den P-f.s. grenzwert

$$A := \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k^2.$$

Welche der folgenden Aussagen ist korrekt?

(A) 0,

(B) **True:** $\frac{25}{3}$,

(C) $+\infty$,

(D) $\frac{25}{4}$.

Lösung:

Nach dem Gesetz der großen Zahlen konvergiert der Durchschnitt der Quadrate der X_k gegen den Erwartungswert von X_k^2:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k^2 = E[X_k^2] = \int_0^5 x^2 \cdot \frac{1}{5}\, dt = \frac{25}{3}.$$

1.MC9 **[4 Punkte]** Sei $T_1, T_2, \ldots$ eine Folge von unabhängigen, identisch verteilten Zufallsvariablen mit $T_1 \sim \mathcal{U}(0,1)$. Sei Z eine standardnormalverteilte Zufallsvariable, also $Z \sim \mathcal{N}(0,1)$. Welche Aussage ist korrekt?

(A) $\mathbb{P}\left[\sum_{i=1}^n T_i \leq \frac{n+a\sqrt{n}}{0.5}\right] \xrightarrow{n\to\infty} \mathbb{P}[Z \leq a]$, $\forall a \in \mathbb{R}$,

(B) $\mathbb{P}\left[\sum_{i=1}^n T_i \leq n + a\sqrt{n/12}\right] \xrightarrow{n\to\infty} \mathbb{P}[Z \leq a]$, $\forall a \in \mathbb{R}$,

(C) **True:** $\mathbb{P}\left[\sum_{i=1}^n T_i \leq n/2 + a\sqrt{n/12}\right] \xrightarrow{n\to\infty} \mathbb{P}[Z \leq a]$, $\forall a \in \mathbb{R}$,

(D) $\mathbb{P}\left[\sum_{i=1}^n T_i \leq n + a\sqrt{n}\right] \xrightarrow{n\to\infty} \mathbb{P}[Z \leq a]$, $\forall a \in \mathbb{R}$.

Lösung:

Verwenden wir den zentralen Grenzwertsatz (ZGS) für eine Folge von unabhängigen, identisch verteilten Zufallsvariablen T_i:

$$\mathbb{P}\left[\frac{\sum_{i=1}^n T_i - nE[T_i]}{\sqrt{n\mathrm{Var}(T_i)}} \leq a\right] \xrightarrow{n\to\infty} \mathbb{P}[Z \leq a], \text{ für } Z \sim \mathcal{N}(0,1).$$

Wir berechnen den Erwartungswert und die Varianz von T_i:

$$E[T_i] = \frac{1}{2}, \quad \mathrm{Var}(T_i) = \frac{1}{12}.$$

Einsetzen in den ZGS:

$$\mathbb{P}\left[\sum_{i=1}^n T_i \leq n/2 + a\sqrt{n/12}\right] \xrightarrow{n\to\infty} \mathbb{P}[Z \leq a].$$

1.MC10 [**3 Punkte**] Sei $\Theta = (0, \infty)$, und $X_1, \ldots, X_n$ unter $\mathbb{P}_\theta$ unabhängig, identisch verteilt mit $X_1 \sim \mathcal{N}(\mu, 4)$, wobei μ ein unbekannter Parameter ist. Was ist der Maximum-Likelihood-Schätzer T_{ML} für μ?

(A) **True:** $\frac{X_1+\ldots+X_n}{n}$,
(B) $\frac{X_1+\ldots+X_n}{4n}$,
(C) $\frac{2n}{X_1+\ldots+X_n}$,
(D) $\frac{X_1+\ldots+X_n}{2n}$.

Lösung:

Für die Zufallsvariablen $X_i \sim \mathcal{N}(\mu, 4)$ ist die Likelihood-Funktion:

$$L(\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{8\pi}} \exp\left(-\frac{(x_i - \mu)^2}{8}\right).$$

Die logarithmische Likelihood-Funktion ist:

$$\ell(\mu) = -\frac{n}{2}\log(8\pi) - \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{8}.$$

Maximieren wir $\ell(\mu)$ durch Ableitung und Nullsetzen:

$$\frac{\partial \ell(\mu)}{\partial \mu} = \sum_{i=1}^{n} \frac{x_i - \mu}{4} = 0 \implies \mu = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Der Maximum-Likelihood-Schätzer für μ ist:

$$T_{ML} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

1.MC11 [**3 Punkte**] Sei $\Theta = \mathbb{R}$. Seien $X_1, \ldots, X_n$ Zufallsvariablen. Wir betrachten die Modellfamilie $(\mathbb{P}_\theta)_{\theta \in \Theta}$, sodass $X_1, \ldots, X_n$ unter $\mathbb{P}_\theta$ unabhängig, identisch verteilt sind mit $X_i \sim \mathcal{N}(\mu, \sigma^2)$, wobei $\sigma^2 = 4$ bekannt und μ ein unbekannter Parameter ist. Welcher der folgenden Schätzer ist **nicht** erwartungstreu für μ?

(A) $T_3^n = \frac{1}{2}(X_1 + X_n), \quad n > 1,$

(B) $T_4^n = X_1 + X_1 X_2 - X_n^2 + 4, \quad n > 1,$

(C) **True:** $T_1^n = \sqrt[n]{|X_1 \cdot \ldots \cdot X_n|},$

(D) $T_2^n = \frac{1}{2n} \sum_{i=1}^{2n} X_i.$

Lösung:

$$E[T_1^1] = E[|X|] > E[X] = \mu.$$

$$E[T_2^n] = \frac{1}{2n} \sum_{i=1}^{2n} E[X_i] = \frac{1}{2n} \cdot 2n\mu = \mu.$$

$$E[T_3^n] = \frac{1}{2}(E[X_1] + E[X_n]) = \frac{1}{2} \cdot \mu = \mu.$$

$$E[T_4^n] = E[X_1] + E[X_1 X_2] - E[X_n^2] + 4 = \mu + \mu^2 - (\sigma^2 + \mu^2) + 4 = \mu.$$

Aufgaben

2. Aufgabe **[10 Punkte]**

Seien X, Y i.i.d. Zufallsvariablen mit $X, Y \sim \mathrm{Exp}(1)$. Wir definieren $L = \min(X, Y)$ und $M = \max(X, Y)$.

(a) **[3 Punkte]** Berechnen Sie die Verteilungsfunktion von M.

Lösung:

Für eine Exponentialverteilung mit Parameter $\lambda = 1$ ist die Verteilungsfunktion $F_X(x) = (1 - e^{-x}) \cdot \mathbf{1}_{x \geq 0}$ und die Dichtefunktion $f_X(x) = e^{-x} \cdot \mathbf{1}_{x \geq 0}$. Die Verteilungsfunktion von $M = \max(X, Y)$ ist:

$$F_M(y) = \mathbb{P}[M \leq y] = \mathbb{P}[X \leq y, Y \leq y] = \Big(F_X(y)\Big)^2 = (1 - e^{-y})^2 \cdot \mathbf{1}_{y \geq 0}.$$

(b) **[4 Punkte]** Berechnen Sie $\mathbb{P}[L \geq x, M \leq y]$ für $0 \leq x \leq y$.

Lösung:

Für $L = \min(X, Y)$ und $M = \max(X, Y)$ betrachten wir die Wahrscheinlichkeit $\mathbb{P}[L \geq x, M \leq y]$:

$$\mathbb{P}[L \geq x, M \leq y] = \mathbb{P}[x \leq \min(X, Y) \leq y, \max(X, Y) \leq y].$$

Da X und Y unabhängig und identisch verteilt sind:

$$\mathbb{P}[L \geq x, M \leq y] = \mathbb{P}[X \geq x \text{ und } Y \geq x, X \leq y \text{ und } Y \leq y].$$

$$= \mathbb{P}[x \leq X \leq y \text{ und } x \leq Y \leq y] = \Big(\mathbb{P}[x \leq X \leq y]\Big)^2.$$

Die Wahrscheinlichkeit $\mathbb{P}[x \leq X \leq y]$ ist:

$$\mathbb{P}[x \leq X \leq y] = F_X(y) - F_X(x) = (1 - e^{-y}) - (1 - e^{-x}) = e^{-x} - e^{-y}.$$

Also ist:

$$\mathbb{P}[L \geq x, M \leq y] = (e^{-x} - e^{-y})^2.$$

(c) **[3 Punkte]** Berechnen Sie $\mathbb{P}[L \geq x \mid M \leq y]$ für $x = \log(2)$ und $y = \log(4)$.

Lösung:

Die bedingte Wahrscheinlichkeit $\mathbb{P}[L \geq x \mid M \leq y]$ ergibt sich aus dem Verhältnis der gemeinsamen Wahrscheinlichkeit und der Marginalwahrscheinlichkeit:

$$\mathbb{P}[L \geq x \mid M \leq y] = \frac{\mathbb{P}[L \geq x, M \leq y]}{\mathbb{P}[M \leq y]}.$$

Berechnen wir zunächst $\mathbb{P}[M \leq y]$:

$$\mathbb{P}[M \leq y] = F_M(y) = (1 - e^{-y})^2.$$

Setzen wir $y = \log(4)$ ein:

$$\mathbb{P}[M \leq \log(4)] = \left(1 - e^{-\log(4)}\right)^2 = \left(1 - \frac{1}{4}\right)^2 = \left(\frac{3}{4}\right)^2 = \frac{9}{16}.$$

Nun berechnen wir $\mathbb{P}[L \geq x, M \leq y]$ für $x = \log(2)$ und $y = \log(4)$:

$$\mathbb{P}[L \geq \log(2), M \leq \log(4)] = \left(e^{-\log(2)} - e^{-\log(4)}\right)^2 = \left(\frac{1}{2} - \frac{1}{4}\right)^2 = \left(\frac{1}{4}\right)^2 = \frac{1}{16}.$$

Somit ist:

$$\mathbb{P}[L \geq \log(2) \mid M \leq \log(4)] = \frac{\mathbb{P}[L \geq \log(2), M \leq \log(4)]}{\mathbb{P}[M \leq \log(4)]} = \frac{\frac{1}{16}}{\frac{9}{16}} = \frac{1}{9}.$$

3. Aufgabe **[10 Punkte]**

Sei $X_1, X_2, \ldots$ eine unendliche Folge von unabhängigen, identisch verteilten Zufallsvariablen, wobei X_1 kontinuierlich ist mit

$$f_{X_1}(x) = \begin{cases} 2x & \text{für } 0 \leq x \leq 1, \\ 0 & \text{sonst.} \end{cases}$$

(a) **[2 Punkte]** Bestimmen Sie den Erwartungswert $E[X_1]$ und die Varianz $\sigma^2_{X_1}$ von X_1.

Lösung:

$$E[X_1] = \int_0^1 x \cdot 2x \, dx = 2\int_0^1 x^2 \, dx = 2\left[\frac{x^3}{3}\right]_0^1 = \frac{2}{3}.$$

$$E[X_1^2] = \int_0^1 x^2 \cdot 2x \, dx = 2\int_0^1 x^3 \, dx = 2\left[\frac{x^4}{4}\right]_0^1 = \frac{1}{2}.$$

$$\mathrm{Var}(X_1) = E[X_1^2] - (E[X_1])^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{9}{18} - \frac{8}{18} = \frac{1}{18}.$$

$$\sigma^2_{X_1} = \frac{1}{18}.$$

Für jedes $n \in \mathbb{N}$ betrachten wir die Partialsummen

$$T_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i.$$

(b) **[2 Punkte]** Bestimmen Sie den Erwartungswert und die Varianz von T_n.

Lösung:

$$E[T_n] = E\left[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i\right] = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} E[X_i] = \frac{1}{\sqrt{n}} \cdot n \cdot \frac{2}{3} = \frac{2\sqrt{n}}{3}.$$

$$\mathrm{Var}(T_n) = \mathrm{Var}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i\right) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{1}{n} \cdot n \cdot \frac{1}{18} = \frac{1}{18}.$$

(c) **[2 Punkte]** Zeigen Sie, dass

$$\mathbb{P}\left[T_n \geq \frac{5}{4}\sqrt{n}\right] < \frac{3}{5}.$$

Lösung:

Um das zu zeigen, verwenden wir das Markov-Ungleichung:

$$\mathbb{P}[T_n \geq k] \leq \frac{E[T_n]}{k}.$$

Setzen wir $k = \frac{5}{4}\sqrt{n}$ ein:

$$\mathbb{P}\left[T_n \geq \frac{5}{4}\sqrt{n}\right] \leq \frac{E[T_n]}{\frac{5}{4}\sqrt{n}}.$$

Einsetzen des Erwartungswerts von T_n ergibt:

$$E[T_n] = \frac{2\sqrt{n}}{3}.$$

Einsetzen ergibt:

$$\mathbb{P}\left[T_n \geq \frac{5}{4}\sqrt{n}\right] \leq \frac{\frac{2\sqrt{n}}{3}}{\frac{5}{4}\sqrt{n}} = \frac{\frac{2}{3}}{\frac{5}{4}} = \frac{2}{3} \cdot \frac{4}{5} = \frac{8}{15} < \frac{9}{15} = \frac{3}{5}.$$

Daher ist:

$$\mathbb{P}\left[T_n \geq \frac{5}{4}\sqrt{n}\right] < \frac{3}{5}.$$

(d) **[1 Punkt]** Bestimmen Sie den Grenzwert in Wahrscheinlichkeit:

$$\lim_{n\to\infty} \frac{T_n}{\sqrt{n}}$$

Lösung:

Wir betrachten den Grenzwert: $\frac{T_n}{\sqrt{n}} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{S_n}{n}$.
Für große n: $\lim_{n\to\infty} \frac{T_n}{\sqrt{n}} = E[X_1] = \frac{2}{3}$

(e) [**3 Punkte**] Zeigen Sie, dass

$$\frac{\frac{3}{2}T_n - \sqrt{n}}{\sigma} \xrightarrow{d} \mathcal{N}(0,1),$$

und finden Sie den Parameter σ^2, sodass diese Bedingung erfüllt ist.

Lösung:

Nach dem zentralen Grenzwertsatz (ZGS) gilt für eine Sequenz unabhängiger, identisch verteilter Zufallsvariablen Y_i mit endlichem Erwartungswert μ_Y und endlicher Varianz σ_Y^2:

$$\frac{\sum_{i=1}^n Y_i - n\mu_Y}{\sqrt{n}\sigma_Y} \xrightarrow{d} \mathcal{N}(0,1).$$

Für $Y_i = \frac{3}{2}X_i$ haben wir $\mu_Y = \frac{3}{2}E[X_1] = 1$ und $\sigma_Y^2 = \left(\frac{3}{2}\right)^2 \cdot \sigma_{X_1}^2 = \frac{9}{4} \cdot \frac{1}{18} = \frac{1}{8}$.
Das lässt sich umschreiben zu:

$$\frac{\frac{3}{2}T_n - \sqrt{n}}{\sqrt{\frac{1}{8}}} = \frac{\sum_{i=1}^n Y_i - n}{\sqrt{n} \cdot \sqrt{\frac{1}{8}}} \xrightarrow{d} \mathcal{N}(0,1).$$

Der Parameter σ^2, der diese Bedingung erfüllt, ist:

$$\sigma^2 = \frac{1}{8}.$$

4. Aufgabe [10 Punkte]

Ein Milliardär zieht eine Investition in der Schweiz in Betracht. Dazu beobachtet er die Renditen von 6 Schweizer Hedgefonds über ein Jahr. Es wird angenommen, dass die Renditen $X_1, ..., X_6$ (in Prozent, d.h. $X_1 = 7$, falls die Rendite des 1. Hedgefonds im Jahr 7% beträgt) dieser 6 Hedgefonds unabhängig und normalverteilt seien mit unbekanntem Mittelwert μ und bekannter Varianz $\sigma^2 = 16$. Das bedeutet

$$X_i \overset{i.i.d.}{\sim} \mathcal{N}(\mu, \sigma^2) \quad \text{für alle } 1 \leq i \leq 6.$$

Am Ende des ersten Jahres beobachtet der Milliardär die folgende Stichprobe:

x_1	x_2	x_3	x_4	x_5	x_6
6	6	12	8	4	0

(a) **[2 Punkte]** Berechnen Sie das empirische Stichprobenmittel $\bar{x}_6$ und die empirische Stichprobenvarianz s^2.

Lösung:

Das empirische Stichprobenmittel ist gegeben durch

$$\bar{x}_6 = \frac{1}{6}(6 + 6 + 12 + 8 + 4 + 0) = 6.$$

Die empirische Stichprobenvarianz ist gegeben durch

$$s^2 = \frac{1}{6-1}\Big((6-6)^2 + (6-6)^2 + (12-6)^2 + (8-6)^2 + (4-6)^2 + (0-6)^2\Big) = \frac{80}{5} = 16.$$

(b) **[2 Punkte]** Bestimmen Sie den Maximum-Likelihood-Schätzer für μ.

Lösung:

Die Likelihood-Funktion ergibt sich aus dem Produkt der Dichten. Für $x_1, \ldots, x_n$ erhalten wir

$$L(x_1, \ldots, x_6; \mu) = \prod_{i=1}^{6} f(x_i),$$

wobei

$$f(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x_i - \mu)^2}{2\sigma^2}\right).$$

Die Log-Likelhood-Funktion erhalten wir duch Logarithmieren obiger Former als

$$\ell(x_1, \ldots, x_6; \mu) = \sum_{i=1}^{n} \log\left(f(x_6)\right) = \frac{6}{\sqrt{2\pi\sigma^2}} + \sum_{i=1}^{6} \frac{-(x_i - \mu)^2}{2\sigma^2}.$$

Ableiten und Nullsetzen der log-Likelihood-Funktion ergibt

$$\frac{\partial}{\partial \mu} \ell(x_1, \ldots, x_6; \mu) = \sum_{i=1}^{6} \frac{-2(x_i - \mu)}{2\sigma^2} = 0.$$

Also ist der Maximum-Likelihood-Schätzer gegeben durch

$$T_{ML} = \frac{1}{6} \sum_{i=1}^{6} x_i = \bar{x}_6 = 6.$$

Wir nehmen nun an, dass die Renditen $X_1, \ldots X_6$ unabhänging und gleichverteilt auf einem Intervall $[a, b]$ sind. Das bedeutet

$$X_i \overset{i.i.d.}{\sim} \mathcal{U}(a, b) \quad \text{für alle} \quad 1 \leq i \leq 6 \quad \text{und für} \quad a < b \in \mathbb{R}.$$

(c) **[3 Punkte]** Bestimmen Sie die Verteilungsfunktion der Ordnungsstatistik $X_{(n)}$, wobei $X_{(n)} = \max(X_1, \ldots, X_n)$.

Lösung:

$$F_{X_{(n)}}(x) = P(X_{(n)} \leq x) = P(\forall i \leq n \; X_i \leq x) = P(X_1 \leq x, \ldots, X_n \leq x) = P(X_1 \leq x)^n.$$

Da $X_i \sim \mathcal{U}(a, b)$, ist $P(X_i \leq x) = \frac{x-a}{b-a}$ für $a \leq x \leq b$, daher

$$F_{X_{(n)}}(x) = \left(\frac{x-a}{b-a}\right)^n \quad \text{für} \quad a \leq x \leq b.$$

(d) [**3 Punkte**] Bestimmen Sie den Maximum-Likelihood-Schätzer für $\vartheta = (a, b)$.

Hinweis: Die Likelihood-Funktion ist nicht differenzierbar. Das Maximum der Likelihood-Funktion muss daher auf eine andere Art bestimmt werden.

Lösung:

Seien $X_i \overset{i.i.d.}{\sim} \mathcal{U}(a, b)$. Mit $\vartheta = (a, b)$ ist die Likelihood-Funktion gegeben durch

$$L(x_1, \dots, x_6; \vartheta) = \frac{1}{(b-a)^6} \prod_{i=1}^{6} 1_{[a,b]}(x_i). \tag{1}$$

Nun muss $L(x_1, \dots, x_6; \vartheta)$ für feste $(x_1, \dots, x_6)$ bezüglich a und b maximiert werden. Seien

$$x_* := \min_{1 \le i \le 6} x_i \quad \text{und} \quad x^* := \max_{1 \le i \le 6} x_i.$$

Falls x_* und x^* ausserhalb $[a, b]$ liegen, dann verschwindet die rechte Seite von (1). Es genügt also, den Maximierer $\vartheta = (a, b)$ in der Menge $\mathcal{D} := \{(a, b) \in \mathbb{R}^2 : a \le x_*, x^* \le b\}$ zu suchen. Für $(a, b) \in \mathcal{D}$ gilt jedoch

$$\frac{1}{(b-a)^6} \prod_{i=1}^{6} 1_{[a,b]}(x_i) \le \frac{1}{(x^* - x_*)^6} \prod_{i=1}^{6} 1_{[x_*, x^*]}(x_i),$$

d.h. $T_{ML} = (x_*, x^*)$.