D–MATH

Prüfung Wahrscheinlichkeit und Statistik

401-2604-00L

Nachname **XX** *Vorname* **XX**

Legi-Nr. **XX-000-000**

Prüfungs-Nr. ***000***

Bitte noch nicht umblättern!

Beachten Sie auch die Hinweise auf dem Antwortheft.

- Fachbegriffe, die in der Vorlesung auf Englisch eingeführt wurden, werden auch im Examen auf Englisch benutzt (ohne deutsche Übersetzung).
- Wir benutzen die Abkürzungen *pmf* und *pdf* für, respektive, *probability mass function* und *probability density function*.
- Der Satz "Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein *probability space*." wird manchmal weggelassen und implizit angenommen.
- "Sei (...)" : Mathematische Aussagen sollten in der Regel quantifiers ("es gibt" oder "für alle") beinhalten. Allerdings werden solche Sätze manchmal lang und kompliziert. Wir benutzen daher die folgende, übliche Konvention. Aussagen der Form "Sei (...). Dann gilt (...)" oder "Seien (...). Dann gilt (...)" sind äquivalent zu "für alle"-Aussagen. Z.B.: "Sei X eine *random variable*. Dann gilt $\mathbb{E}(X) = 2$." muss verstanden werden als "Für alle *random variables* X gilt $\mathbb{E}(X) = 2$." (Diese Aussage ist falsch.)
- Es ist jeweils **genau eine** Antwort korrekt. Punktabzug bei falschen Antworten gibt es **nicht**.

Aufgabe 1 (Wahrscheinlichkeitstheorie)

1.MC1 [1 Punkt] Beurteilen Sie die folgende Aussage. Sei $\Omega := \{-27, 5, 19\}$. Dann ist $\mathcal{F} := \{\emptyset, \{5\}, \{-27, 19\}, \Omega\}$ eine σ-*algbra* über Ω.

(A) wahr

(B) falsch

1.MC2 [1 Punkt] Beurteilen Sie die folgende Aussage. Sei $(\mathbb{R}, \mathcal{F})$ ein *measurable space*. Dann ist $\#\mathcal{F} \neq 3$ ($\#\mathcal{F}$ bezeichnet die Anzahl der Elemente von $\mathcal{F}$).

(A) falsch

(B) wahr

1.MC3 [1 Punkt] Beurteilen Sie die folgende Aussage. Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein *probability space* und $E_1, E_2 \in \mathcal{F}$ mit $E_2 \neq \emptyset$. Dann gilt $\mathbb{P}(E_1 \cup E_2) \neq \mathbb{P}(E_1)$.

(A) wahr

(B) falsch

1.MC4 [1 Punkt] Sei X eine *random variable* mit $\mathbb{P}(X < 4) = 0.2$ und $\mathbb{P}(X < 5) = 0.7$. Welches der folgenden Intervalle enthält den wahren Wert von $\mathbb{P}(X \in [4, 5))$?

(A) $[0.5, 0.67)$

(B) $[0.83, 1]$

(C) $[0.33, 0.5)$

(D) $[0.67, 0.83)$

(E) $[0.17, 0.33)$

(F) $[0, 0.17)$

1.MC5 [1 Punkt] Sei X eine *discrete random variable* mit *pmf*

$$p_X(x) = \begin{cases} 0.3 & \text{falls } x = -3 \\ 0.7 & \text{falls } x = 7 \\ 0 & \text{sonst.} \end{cases}$$

Welches der folgenden Intervalle enthält den wahren Wert von $\mathbb{E}(X)$?

(A) $[2, 3)$
(B) $[0, 1)$
(C) $(-\infty, 0)$
(D) $[1, 2)$
(E) $[3, 4)$
(F) $[4, \infty)$

1.MC6 [1 Punkt] Sei X eine *discrete random variable* mit $\mathbb{P}(X = 20) = 0.5 = \mathbb{P}(X = -20)$. Welches der folgenden Intervalle enthält den wahren Wert von $\mathrm{var}(X)$?

(A) $[300, 1000)$
(B) $[0, 10)$
(C) $[30, 100)$
(D) $[1000, \infty)$
(E) $[100, 300)$
(F) $[10, 30)$

1.MC7 [1 Punkt] **Setting:** Nehmen Sie an, dass Sie 142 Mal ein Ei in die Luft werfen und versuchen dieses zu fangen. Sei X die Anzahl an Versuchen (von 142), bei denen es Ihnen gelingt, das Ei zu fangen. Nehmen Sie an, dass $X \sim \mathrm{Bin}(142, 0.5)$. Hier ist das Setting zu Ende.

Beurteilen Sie die folgende Aussage. Sei p die *pmf* von X. Diese erfüllt für alle $x \in \mathbb{R}$

$$p(x) = \begin{cases} \binom{142}{x} 0.5^x & \text{falls } x \in \{1, 2, \dots, 142\} \\ 0 & \text{sonst.} \end{cases}$$

(A) falsch
(B) wahr

1.MC8 [1 Punkt] Sei das **Setting** von Aufgabe 1.MC7 gegeben. Welche der folgenden Möglichkeiten entspricht R code, der $\mathbb{P}(X = 100)$ berechnet?

(A) `pbinom(100, size = 142, prob = 0.5)`
(B) `pbinom(142, size = 100, prob = 0.5)`
(C) `dbinom(100, size = 142, prob = 0.5)`
(D) `dbinom(142, size = 100, prob = 0.5)`

1.MC9 [1 Punkt] Sei das **Setting** von Aufgabe 1.MC7 gegeben. Welche der folgenden Möglichkeiten entspricht R code, der $\mathbb{P}(90 \leq X \leq 100)$ berechnet?

(A) `pbinom(100, size = 142, prob = 0.5) - pbinom(89, size = 142, prob = 0.5)`
(B) `dbinom(100, size = 142, prob = 0.5) - dbinom(90, size = 142, prob = 0.5)`
(C) `pbinom(100, size = 142, prob = 0.5) - pbinom(90, size = 142, prob = 0.5)`
(D) `dbinom(100, size = 142, prob = 0.5) - dbinom(89, size = 142, prob = 0.5)`

1.MC10 [1 Punkt] Sei X eine *continuous random variable* mit *pdf*

$$f_X(x) = \begin{cases} x & \text{falls } 0 < x < \sqrt{2} \\ 0 & \text{sonst.} \end{cases}$$

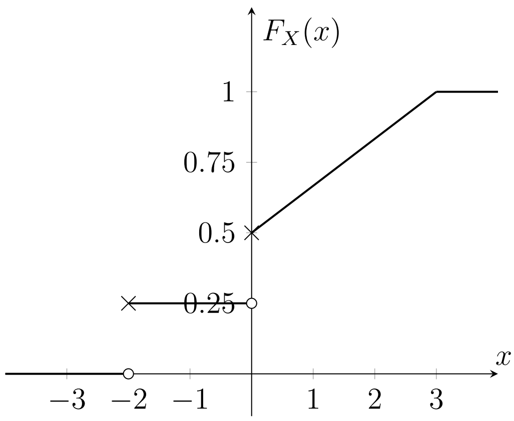
Sei F_X die zugehörige *cdf*. Welches der folgenden Intervalle enthält den wahren Wert von $F_X(0.5)$?

(A) $[0.5, 0.67)$
(B) $[0, 0.16)$
(C) $[0.16, 0.33)$
(D) $[0.83, 1]$
(E) $[0.67, 0.83)$
(F) $[0.33, 0.5)$

1.MC11 [1 Punkt] Sei X eine *continuous random variable* und $f_X : \mathbb{R} \to \mathbb{R}$ eine zugehörige *pdf* mit *support* $\text{supp}(f_X) = \mathbb{R}$. Welche der folgenden Möglichkeiten beschreibt eine *pdf* von $2X + 4$?

(A) $x \mapsto 2f_X(2x + 4)$
(B) $x \mapsto \frac{1}{2} \cdot f_X(2x + 4)$
(C) $x \mapsto f_X\left(\frac{x-4}{2}\right)$
(D) $x \mapsto f_X(2x + 4)$
(E) $x \mapsto \frac{1}{2} \cdot f_X\left(\frac{x-4}{2}\right)$
(F) $x \mapsto 2f_X\left(\frac{x-4}{2}\right)$

1.MC12 [1 Punkt] Sei X eine *random variable* mit folgender *cdf* F_X.



Welche der folgenden Aussagen ist wahr?

(A) Die *random variable* X ist weder *continuous* noch *discrete*.

(B) Die *random variable* X ist sowohl *continuous* als auch *discrete*.

(C) Die *random variable* X ist *discrete*.

(D) Die *random variable* X ist *continuous*.

1.MC13 [1 Punkt] Sei X eine *random variable* mit *cdf*

$$F_X(x) = \begin{cases} 0, & \text{falls } x < -\pi/2, \\ \frac{1+\sin(x)}{2}, & \text{falls } x \in [-\pi/2, \pi/2], \\ 1, & \text{falls } x > \pi/2. \end{cases}$$

Welches der folgenden Intervalle enthält $\mathbb{E}(X)$?

(A) $[0, 0.25)$

(B) $(-\infty, 0)$

(C) $[1, \infty)$

(D) $[0.25, 0.5)$

(E) $[0.5, 0.75)$

(F) $[0.75, 1)$

1.MC14 [1 Punkt] Sei X eine *continuous random variable* mit *pdf*

$$f_X(x) = \begin{cases} 4x^3, & \text{falls } x \in [0,1], \\ 0, & \text{sonst.} \end{cases}$$

Welches der folgenden Intervalle enthält $\text{var}(X)$?

(A) $[1, \infty)$
(B) $[0.1, 0.3)$
(C) $[0.03, 0.1)$
(D) $[0.01, 0.03)$
(E) $[0, 0.01)$
(F) $[0.3, 1)$

1.MC15 [1 Punkt] Sei X eine *random variable* mit $\mathbb{E}(X) = 0$ und $\text{var}(X) = 1$. Welches ist das kleinste $c \in \mathbb{R}$, sodass $c \in \left\{\frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1\right\}$ und die Aussage $\mathbb{P}(X \geq 2) \leq c$ wahr ist?

(A) $c = \frac{1}{8}$
(B) $c = 1$
(C) $c = \frac{1}{2}$
(D) $c = \frac{1}{4}$

1.MC16 [1 Punkt] Sei (X, Y) ein *discrete random vector* mit folgender *joint pmf* p.

$p(x,y)$	$y=-1$	$y=1$
$x=1$	0.1	0
$x=2$	0.3	0
$x=3$	0.4	0.2

und $p(x, y) = 0$ für alle anderen (x, y). Welches der folgenden Intervalle enthält den wahren Wert von $\text{cov}(X, Y)$?

(A) $(-\infty, 0)$
(B) $[0.5, 0.75)$
(C) $[1, \infty)$
(D) $[0.25, 0.5)$
(E) $[0.75, 1)$
(F) $[0, 0.25)$

1.MC17 [1 Punkt] Sei für alle $c > 0$, die Funktion $f_c : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ gegeben durch

$$f_c(x, y) = \begin{cases} c, & \text{falls } x > 0 \text{ und } y > 0 \text{ und } x^2 + y^2 < 1, \\ 0, & \text{sonst.} \end{cases}$$

Welches der folgenden Intervalle enthält einen Wert von c, sodass f_c *joint pdf-like* ist?

(A) $[2, 2.5)$

(B) $[1.5, 2)$

(C) $(0, 0.5)$

(D) $[2.5, \infty)$

(E) $[1, 1.5)$

(F) $[0.5, 1)$

1.MC18 [1 Punkt] Sei (X, Y) ein *continuous random vector*, sodass $\text{var}(X) = \text{var}(Y) = 1$ und $\text{var}(X + Y) = 1.3$. Welches der folgenden Intervalle enthält den wahren Wert von $\text{cov}(X, Y)$?

(A) $[2, \infty)$

(B) $[0.5, 1)$

(C) $[0, 0.5)$

(D) $[1, 1.5)$

(E) $[1.5, 2)$

(F) $(-\infty, 0)$

1.MC19 [1 Punkt] **Setting:** Sei (X, Y) ein *discrete random vector*, dessen *support* gegeben ist durch $\{(0, 0), (0, 1), (1, 0), (1, 1)\}$. Es gelte $\mathbb{P}(X = 0) = 0.6$, $\mathbb{P}(Y = 1) = 0.7$ und $\mathbb{P}(X = 0 \,|\, Y = 1) = 0.5$. Hier ist das Setting zu Ende.

Beurteilen Sie folgende Aussage. X und Y sind *independent*.

(A) wahr

(B) falsch

1.MC20 [1 Punkt] Sei das **Setting** von Aufgabe 1.MC19 gegeben. Welches der folgenden Intervalle enthält den wahren Wert von $\mathbb{P}(Y = 1 \mid X = 1)$?

(A) $[0, 0.17)$

(B) $[0.33, 0.5)$

(C) $[0.83, 1.0]$

(D) $[0.17, 0.33)$

(E) $[0.5, 0.67)$

(F) $[0.67, 0.83)$

1.MC21 [1 Punkt] Sei (X, Y, Z) ein *random vector* mit *joint pdf* f, sodass $X^2 + YZ$ *finite mean* hat. Welche der folgenden Aussagen ist wahr?

(A) $\mathbb{E}[X^2 + YZ] = \int_{-\infty}^{\infty} x^2 \cdot f(x, y, z)\ dx + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yz \cdot f(x, y, z)\, dy\, dz$

(B) $\mathbb{E}[X^2 + YZ] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x^2 + yz) \cdot f(x, y, z)\ dx\, dy\, dz$

(C) $\mathbb{E}[X^2 + YZ] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z)\ dx\, dy\, dz$

(D) $\mathbb{E}[X^2 + YZ] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 yz \cdot f(x, y, z)\ dx\, dy\, dz$

1.MC22 [1 Punkt] Seien X und Y zwei *i.i.d. random variables* mit $\mathrm{var}(X) = \mathrm{var}(Y) = 2$. Welches der folgenden Intervalle enthält den wahren Wert von $\mathrm{var}(2 - X + Y + Y)$?

(A) $[0, 3)$

(B) $[12, 15)$

(C) $[3, 6)$

(D) $[9, 12)$

(E) $[6, 9)$

(F) $[15, \infty)$

1.MC23 [1 Punkt] Sei (X, Y) ein *random vector*, sodass $\mathbb{E}(X) = 0$, $\text{var}(X) = 4$ und $\text{var}(Y) = 9$ und XY *finite mean* hat. Welches ist das kleinste $c \in \mathbb{R}$, sodass $c \in \{0, 3, 6, 9\}$ und die Aussage $\mathbb{E}(XY) \leq c$ wahr ist?

(A) $c = 6$
(B) $c = 9$
(C) $c = 3$
(D) $c = 0$

1.MC24 [1 Punkt] Beurteilen Sie folgende Aussage. Sei $X \sim \mathrm{U}([-2, 2])$. Es gilt

$$(-1)^n X \xrightarrow{\mathbb{P}} X.$$

(A) falsch
(B) wahr

1.MC25 [1 Punkt] Beurteilen Sie folgende Aussage. Sei $X \sim \mathrm{U}([-2, 2])$. Es gilt

$$(-1)^n X \xrightarrow{\mathcal{L}} X.$$

(A) falsch
(B) wahr

1.MC26 [1 Punkt] Welche der folgenden Aussagen ist wahr?

(A) Sei $X_1, X_2, ...$ eine *i.i.d.* Folge von *random variables* mit *finite mean*. Dann gilt

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \xrightarrow{\mathcal{L}} X_1.$$

(B) Sei $X_1, X_2, ...$ eine *i.i.d.* Folge von *random variables* mit *finite mean* und *finite variance*. Dann gilt

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \xrightarrow{\mathbb{P}} \mathbb{E}(X_1).$$

(C) Sei $X_1, X_2, ...$ eine *i.i.d.* Folge von *random variables* mit *finite mean* und *finite variance*. Dann gilt

$$\frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{\mathbb{P}} \mathbb{E}(X_1).$$

(D) Sei $X_1, X_2, ...$ eine *i.i.d.* Folge von *random variables* mit *finite mean*. Dann gilt

$$\frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{\mathcal{L}} X_1.$$

1.MC27 [1 Punkt] Sei $X_1, X_2, ...$ eine Folge von *i.i.d. random variables* mit $X_1 \sim \text{Exp}(\lambda)$ mit $\lambda \in \mathbb{R}_{>0}$. Sei weiter, für alle $k \in \{1, 2, ...\}$, $Y_k := \sqrt{2}(-1)^k X_k$ und sei $Z \sim \text{N}(0, 1)$. (Hinweis 1: Es gilt $\text{var}(X_1) = \lambda^{-2}$. Hinweis 2: Wenn $X_1, X_2, ...$ eine Folge von *independent random variables* ist, dann ist $X_1 + X_2, X_3 + X_4, ...$ ebenfalls eine Folge von *independent random variables*.)
Welche der folgenden Aussagen ist wahr?

(A)
$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} Y_k \xrightarrow{\mathcal{L}} 2\lambda^{-2} Z.$$

(B)
$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} Y_k \xrightarrow{\mathcal{L}} 0.$$

(C)
$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} Y_k \xrightarrow{\mathcal{L}} \sqrt{2}\lambda^{-1} Z.$$

(D)
Es gibt keine *random variable* Y, sodass $\frac{1}{\sqrt{n}} \sum_{k=1}^{n} Y_k \xrightarrow{\mathcal{L}} Y$

1.MC28 [1 Punkt] Beurteilen Sei folgende Aussage. Sei X eine *random variable* und sei a_n eine Folge von reellen Zahlen mit

$$\lim_{n \to \infty} a_n = 0.$$

Dann gilt

$$a_n X \xrightarrow{\mathbb{P}} 0.$$

(A) wahr

(B) falsch

Aufgabe 2 (Statistik)

2.MC1 [1 Punkt] Sei $(X_1, \dots, X_{10})$ ein *random vector*, sodass $X_1, \dots, X_{10}$ *i.i.d.* sind mit $X_1 \sim \mathrm{Bin}(n, 1/3)$. Welcher der folgenden *estimators* ist kein *unbiased estimator* für n?

(A) $(x_1, \dots, x_{10}) \mapsto \max(x_1, \dots, x_{10})$

(B) $(x_1, \dots, x_{10}) \mapsto 3x_1$

(C) $(x_1, \dots, x_{10}) \mapsto x_1 + x_2 + x_3$

(D) $(x_1, \dots, x_{10}) \mapsto \frac{3}{10} \sum_{i=1}^{10} x_i$

2.MC2 [1 Punkt] **Setting:** Seien $X_1, \dots, X_n$ *i.i.d. discrete random variables.* Sei $\theta_0 \in [-10, 10]$ und habe X_1 eine *pmf* p mit

$$p(-10) = \frac{1}{2} - \frac{\theta_0}{20} \qquad \text{und} \qquad p(10) = \frac{1}{2} + \frac{\theta_0}{20} \qquad \text{und} \qquad p(x) = 0 \text{ für alle } x \in \mathbb{R} \setminus \{-10, 10\}.$$

Sei $\hat{\theta} : \mathbb{R}^n \to \mathbb{R}$, definiert als

$$\hat{\theta}(x_1, \dots, x_n) := \frac{1}{n} \sum_{i=1}^{n} x_i,$$

ein *estimator* für θ_0. Hier ist das Setting zu Ende.

Beurteilen Sie die folgende Aussage. $\hat{\theta}$ ist ein *unbiased estimator* für θ_0.

(A) wahr

(B) falsch

2.MC3 [1 Punkt] Sei das **Setting** von Aufgabe 2.MC2 gegeben. Seien nun $\theta_0 = 5$ und $n = 15$. Welches der folgenden Intervalle enthält $\mathrm{var}(\hat{\theta})$?

(A) $[4, 6)$

(B) $[0, 2)$

(C) $[2, 4)$

(D) $[6, 8)$

(E) $[8, 10)$

(F) $[10, \infty)$

2.MC4 [1 Punkt] Beurteilen Sie die folgende Aussage. Seien $X_1, \ldots, X_n$ *i.i.d. random variables* mit *distribution* abhängig von unbekanntem $\theta_0 \in \mathbb{R}$ und sei weiter $\hat{\theta}$ ein *estimator* für θ_0. Es gilt

$$\mathrm{MSE}(\hat{\theta}) = 0 \quad \Longrightarrow \quad \mathrm{BIAS}(\hat{\theta}) = 0.$$

(A) wahr

(B) falsch

2.MC5 [1 Punkt] Seien $X_1, \ldots, X_5$ *i.i.d. random variables* mit *pdf*

$$f_{\theta_0}(x) = \begin{cases} \frac{2}{\theta_0^2} x, & \text{für } 0 < x < \theta_0, \\ 0, & \text{sonst,} \end{cases}$$

wobei $\theta_0 > 0$ unbekannt ist. Sei $\hat{\theta}_{\mathrm{ML}} : \mathbb{R}^5 \to \mathbb{R}$ ein *maximum likelihood estimator* für θ_0. Welches der folgenden Intervalle enthält den wahren Wert von $\hat{\theta}_{\mathrm{ML}}(1.1, 2.4, 0.9, 2.7, 1.3)$?

(A) $[2, 3)$

(B) $[0, 1)$

(C) $[1, 2)$

(D) $[3, 4)$

(E) $[4, 5)$

(F) $[5, \infty)$

2.MC6 [1 Punkt] Beurteilen Sie die folgende Aussage. Seien $X_1, \ldots, X_{12}$ *i.i.d. random variables* mit *distribution* $\mathrm{U}([0, \theta_0])$, wobei $\theta_0 > 0$ unbekannt ist. Dann ist die Funktion $\hat{\theta} : \mathbb{R}^{12} \to \mathbb{R}$, gegeben durch

$$\hat{\theta}(x_1, \ldots, x_{12}) = \begin{cases} \frac{1}{2x_1} & \text{für } x_1 \neq 0, \\ 0 & \text{sonst,} \end{cases}$$

ein *unbiased estimator* für $\frac{1}{\theta_0}$.

(A) falsch

(B) wahr

2.MC7 [1 Punkt] Seien $X_1, \dots, X_n$ *i.i.d.* mit einer *distribution*, die von einem unbekanntem $\theta_0 \in \mathbb{R}$ abhängt und sei $\hat{\theta}$ ein *estimator* für θ_0. Sei weiter $f : \mathbb{N} \to \mathbb{N}$ so, dass für alle $n \in \mathbb{N}$ und alle $\theta_0 \in \mathbb{R}$ gilt

$$f(n)(\hat{\theta} - \theta_0) \sim \mathrm{N}(0, 1)$$

Welche der folgenden Möglichkeiten beschreibt ein 0.95-*confidence interval* für θ_0?

(A) $\left[\hat{\theta} - \frac{z_{0.975}}{f(n)}, \hat{\theta} + \frac{z_{0.975}}{f(n)}\right]$

(B) $\left[\hat{\theta} - \frac{z_{0.95}}{f(n)}, \hat{\theta} + \frac{z_{0.95}}{f(n)}\right]$

(C) $\left[\hat{\theta} - \frac{z_{0.975}}{(f(n))^2}, \hat{\theta} + \frac{z_{0.975}}{(f(n))^2}\right]$

(D) $\left[\hat{\theta} - \frac{z_{0.95}}{(f(n))^2}, \hat{\theta} + \frac{z_{0.95}}{(f(n))^2}\right]$

2.MC8 [1 Punkt] Sei Z die *random variable*, die dem Output des folgenden R codes entspricht:

```
Z <- sum((rnorm(100) + rnorm(100))^2)/99
```

Welches der folgenden Intervalle enthält den wahren Wert von $\mathbb{E}(Z)$?

(A) $[2, 3)$

(B) $(-\infty, 0)$

(C) $[0, 1)$

(D) $[1, 2)$

(E) $[3, 4)$

(F) $[4, \infty)$

2.MC9 [1 Punkt] **Setting:** In dieser Aufgabe berufen wir uns auf die Schritte T1–T7 aus der Vorlesung. Die Pflanze Fiverage wird, im Mittel, 5cm gross. Ein Biologe behauptet, wenn die Fiverage während des Wachstums mit Heavy Metal Musik beschallt wird, wird sie, im Mittel grösser als 5cm. Der Biologe probiert dies aus mit drei Pflanzen und misst folgende Grössen (in cm).

$$5.2, 4.8, 5.3.$$

Nehmen Sie an, dass $X_1, X_2, X_3 \sim \mathcal{N}(\mu, \sigma^2)$ (Schritt T1). Wir möchten nun einen statistischen Test durchführen $H_0 : \mu = 5$ gegen $H_1 : \mu > 5$ (Schritt T2) mit *significance level* $\alpha = 0.05$ (Schritt T4). Seien $\hat{\mu}$ und $\hat{\sigma}$ definiert als

$$\hat{\mu} := \frac{1}{3}\sum_{i=1}^{3} X_i,$$
$$\hat{\sigma} := \sqrt{\frac{1}{2}\sum_{i=1}^{3}(X_i - \hat{\mu})^2}.$$

Hier ist das Setting zu Ende.

Welche der folgenden Möglichkeiten beschreibt eine *test statistic* T, sodass ein $k \in \mathbb{N}$ existiert, sodass under H_0 gilt: $T \sim t_k$?

(A) $T = \sqrt{3} \cdot \frac{\hat{\mu}-5}{\hat{\sigma}}$

(B) $T = 3 \cdot \frac{\hat{\mu}-5}{\hat{\sigma}}$

(C) $T = \sqrt{3} \cdot \frac{\hat{\mu}-5}{\hat{\sigma}^2}$

(D) $T = 3 \cdot \frac{\hat{\mu}-5}{\hat{\sigma}^2}$

2.MC10 [1 Punkt] Sei das **Setting** aus Aufgabe 2.MC9 gegeben. Sei T die zu einem *t-test* gehöhrende *test statistic*. Welche der folgenden Möglichkeiten beschreibt eine *rejection region* $\mathcal{R}$, sodass die Aussage "Der *test is level*." wahr ist?

(A) $\mathcal{R} = (t_{2;0.95}, \infty)$

(B) $\mathcal{R} = (t_{2;0.05}, \infty)$

(C) $\mathcal{R} = (t_{3;0.95}, \infty)$

(D) $\mathcal{R} = (t_{3;0.05}, \infty)$

2.MC11 [1 Punkt] Sei das **Setting** aus Aufgabe 2.MC9 gegeben. Sei T die zu einem *t-test* gehörende *test statistic*. Welches der folgenden Intervalle enthält den wahren Wert von T, wenn man die *observations* einsetzt?

(A) $(-\infty, 0.81)$

(B) $[0.81, 1.32)$

(C) $[1.32, 1.91)$

(D) $[1.91, 2.55)$

(E) $[2.55, 3.39)$

(F) $[3.39, \infty)$

2.MC12 [1 Punkt] Seien $X_1, \ldots, X_6$ *i.i.d. random variables* mit $X_1 \sim \mathrm{Ber}(\theta_0)$, wobei θ_0 unbekannt ist. Wir betrachten die *null hypothesis* $H_0 : \theta_0 = 1/2$. Welche der folgenden Möglichkeiten beschreibt die *probability* eines *type I errors* für eine *decision function*, die definiert wird mithilfe der *test statistic*

$$T = \sum_{i=1}^{6} X_i$$

und *rejection region* $\mathcal{R} = (-\infty, 0]$?

(A) $(1/2)^6$

(B) $1 - (1/2)^6$

(C) $6 \cdot (1/2)$

(D) $1 - 6 \cdot (1/2)$

2.MC13 [1 Punkt] **Setting:** Sei $\mathbf{X}$ eine reellwertige $n \times (d+1)$ Matrix and $\mathbf{Y}$ eine reellwertige $n \times 1$ Matrix und sei $f : \mathbb{R}^{d+1} \to \mathbb{R}$ eine Funktion definiert durch $b \mapsto (\mathbf{Y} - \mathbf{X}b)^\top(\mathbf{Y} - \mathbf{X}b)$. Hier ist das Setting zu Ende.

Beurteilen Sie folgende Aussage. Für den Gradienten $\nabla_b f$ von f und für alle $b \in \mathbb{R}^{d+1}$ gilt, dass $\nabla_b f(b) = -2\mathbf{X}^\top\mathbf{Y} + \mathbf{X}^\top\mathbf{X}b$.

(A) falsch

(B) wahr

2.MC14 [1 Punkt] Sei das **Setting** von Aufgabe 2.MC13 gegeben. Beurteilen Sie folgende Aussage. Es existiert ein $b_0 \in \mathbb{R}^{d+1}$, sodass für alle $b \in \mathbb{R}^{d+1} \setminus \{b_0\}$ gilt:

$$f(b) > f(b_0).$$

(A) falsch

(B) wahr

2.MC15 [1 Punkt] **Setting:** Sei eine 13×2 Matrix $\mathbf{X}$ und ein 13×1 Vektor $\mathbf{Y}$ gegeben, sodass $\mathbf{X}^\top \mathbf{X}$ invertierbar ist. Seien diese in R gegeben durch `X` und `Y`. Sei $\hat{\beta}_{OLS} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y}$ der *OLS estimator*. Eine *linear regression* in R liefert folgende Resultate.

```
X1 <- X[,1]
X2 <- X[,2]
linmod <- lm(Y~X1+X2)
linmod$coefficients
#(Intercept) X1 X2
#-0.1700000 1.4252873 0.4188439
```

Hier ist das Setting zu Ende.

Welches der folgenden Intervalle enthält den wahren Wert von $\hat{\beta}_{OLS}^\top x_{\text{new}}$, wobei $x_{\text{new}} = (1, 0)^\top$?

(A) $[1, 1.5)$

(B) $(-\infty, 0)$

(C) $[0, 0.5)$

(D) $[0.5, 1)$

(E) $[1.5, 2)$

(F) $[2, \infty)$

2.MC16 [1 Punkt] Sei das **Setting** von Aufgabe 2.MC15 gegeben. Beurteilen Sie folgende Aussage.

$$\|(I_n - \mathbf{P})\mathbf{Y}\|_2^2,$$

wobei I_n eine 13×13-dimensionale Identitätsmatrix ist und $\mathbf{P} := \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$, kann mit R berechnet werden durch
`sum((linmod$residuals)^2)`.

(A) wahr

(B) falsch

2.MC17 [1 Punkt] Seien $Y_1, Y_2 \overset{i.i.d.}{\sim} \mathrm{N}(0,1)$. Welche der folgenden Aussagen ist wahr?

(A) $\frac{(Y_1+Y_2)^2}{2}$ und $\frac{(Y_1-Y_2)^2}{2}$ sind χ_1^2 *distributed* und *independent*.

(B) $\frac{(Y_1+Y_2)^2}{2}$ und $\frac{(Y_1-Y_2)^2}{2}$ sind χ_1^2 *distributed* und nicht *independent*.

(C) $(Y_1+Y_2)^2$ und $(Y_1-Y_2)^2$ sind χ_1^2 *distributed* und *independent*.

(D) $(Y_1+Y_2)^2$ und $(Y_1-Y_2)^2$ sind χ_1^2 *distributed* und nicht *independent*.