



Karlsruher Institut für Technologie

Institut für Stochastik

Prof. Dr. Mathias Trabs

M. Sc. Sebastian Höfer

Nachname:

Vorname:

Matr.-Nr.:

**Klausur zur Vorlesung
Wahrscheinlichkeitstheorie und Statistik
für die Fachrichtungen Maschinenbau und Informatik**

Datum: 26. Juli 2022

Bearbeitungszeit: 90 Minuten

- **Zugelassene Hilfsmittel:** Ein handbeschriebenes DIN A4 Blatt, Taschenrechner (nicht vernetzbar), Wörterbuch.
- Bei dieser Klausur werden nur diejenigen Ergebnisse gewertet, die in die **vorgesehenen Kästchen** eingetragen sind! Eine Begründung bzw. **Herleitung** der Ergebnisse ist **nicht erforderlich**.
- Vereinfachen Sie Ergebnisse soweit wie möglich. Geben Sie Ergebnisse so exakt wie möglich an, z.B. als Bruch. Runden Sie **Endergebnisse** auf 4 **Nachkommastellen** genau, wenn nicht anders angegeben.
- Hinreichend zum Bestehen der Klausur sind 20 Punkte.

Viel Erfolg!

Aufgabe	1 (10 P)	2 (10 P)	3 (10 P)	4 (10 P)	5 (10 P)	Σ (50 P)
Punkte						
Korrektor						

bestanden	nicht bestanden	Note

Verteilungsfunktion $\Phi(x)$ der Standard – Normalverteilung $\mathcal{N}(0, 1)$

[illegible]

Aufgabe 1 (10 Punkte)

a) Kreuzen Sie die jeweils zutreffende Aussage an. Dabei ist jeweils nur **genau** ein Kreuz pro Teilaufgabe zu setzen.

1. f_1 und f_2 seien zwei Zähldichten auf $\mathbb{N}$. Dann ist folgende Funktion ebenfalls eine Zähldichte auf $\mathbb{N}$:	
$f_1 - f_2$	☐
$f_1 + f_2$	☐
$f_1 * f_2$	☐
$f_1 \cdot f_2$	☐
2. Auf einem diskreten Wahrscheinlichkeitsraum $(\Omega, \mathbb{P})$ gilt für zwei Ereignisse $A, B \subset \Omega$ nicht :	
$\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B)$	☐
$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(A B)$	☐
$\mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A \setminus B)$	☐
$\mathbb{P}(A \cap B) = \mathbb{E}[\mathbb{1}_A \cdot \mathbb{1}_B]$	☐
3. Ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ besitzt stets...	
... eine Verteilungsfunktion.	☐
... eine Zähldichte.	☐
... eine Dichte.	☐
... einen endlichen Mittelwert.	☐
4. Ist $X: \Omega \rightarrow \mathbb{N}$ eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit Zähldichte f_X , dann gilt ...	
$\mathbb{E}[X] = \sum_{\omega \in \Omega} \omega f_X(\omega)$	☐
$\mathbb{E}[X] = \int_{\Omega} \omega f_X(\omega) d\omega$	☐
$\mathbb{E}[X] = \sum_{\omega \in \mathbb{N}} \omega f_X(\omega)$	☐
$\mathbb{E}[X] = \int_{\mathbb{N}} \omega f_X(\omega) d\omega$	☐
5. Ist C eine Konfidenzmenge zum Niveau $1 - \alpha$, dann ist ...	
... $\mathbb{1}_{C(x)}(\vartheta_0)$ ein Niveau- α -Test der Hypothese $H_0: \vartheta = \vartheta_0$.	☐
... $\mathbb{1}_{C(\vartheta_0)}(x)$ ein Niveau- α -Test der Hypothese $H_0: \vartheta = \vartheta_0$.	☐
... $1 - \mathbb{1}_{C(x)}(\vartheta_0)$ ein Niveau- α -Test der Hypothese $H_0: \vartheta = \vartheta_0$.	☐
... $1 - \mathbb{1}_{C(\vartheta_0)}(x)$ ein Niveau- α -Test der Hypothese $H_0: \vartheta = \vartheta_0$.	☐

b) Füllen Sie folgende Lücken aus.

1. Es gibt 6 Farben von Gummibärchen. Wir ziehen aus einer vollen Tüte gleichzeitig k Gummibärchen, wobei mindestens k Gummibärchen von jeder Farbe in der Tüte sind. Wie viele verschiedene Farbkombination gibt es?

2. Hat der Zufallsvektor (X, Y) eine Gleichverteilung auf $[0, 1]^2$, dann ist die Randdichte von X gegeben durch:

3. Für ein Ereignis A mit Wahrscheinlichkeit $\mathbb{P}(A) = 0.5$ sei $X = \mathbb{1}_A - \mathbb{1}_{A^c}$. Dann hat X die Varianz

$$\text{Var}(X) = \boxed{}.$$

4. Es seien $(X_i)_{i \in \mathbb{N}}$ reellwertige und unkorrelierte Zufallsvariablen mit endlichem Erwartungswert und $\text{Var}(X_i) = 1$ für alle $i \in \mathbb{N}$. Die stochastische Konvergenz von $\frac{1}{n} \sum_{i=1}^n X_i$ gegen $\mathbb{E}[X_1]$ für $n \rightarrow \infty$ folgt, wenn:

5. Geben Sie für eine unabhängige Stichprobe $X_1, \dots, X_n \in \mathbb{R}$ einen unverzerrten Schätzer der Varianz $\text{Var}(X_1)$ an:

Aufgabe 2 (10 Punkte) Der folgende Datensatz umfasst die Anzahl der Eckbälle und Tore des KSC in der Hinrunde der Saison 21/22. Dabei ist

x_j = Anzahl der Eckbälle des KSC in Spiel j

y_j = Anzahl der Tore des KSC in Spiel j

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
x_j	3	5	4	5	7	6	8	7	5	3	1	5	9	5	5	3	5
y_j	3	3	0	0	1	2	2	1	2	2	1	2	1	1	4	1	3

Sie können ohne Überprüfung annehmen:

$$\bar{x} = 5.06, \quad s_x^2 = 3.93, \quad \sum_{j=1}^{17} (x_j - \bar{x})(y_j - \bar{y}) = -1.8141, \quad s_y^2 = 1.22.$$

- a) Geben Sie das Stichprobenmittel $\bar{y}$, das obere Stichprobenquantil $\tilde{y}_{3/4}$ und die empirische Standardabweichung s_y von $(y_1, \dots, y_{17})$ an: (1+2+1P)

$\bar{y} =$ $\tilde{y}_{3/4} =$ $s_y =$

- b) Geben Sie das $\alpha = 0.1$ -getrimmte Stichprobenmittel $\bar{y}_{0.1}$ von $(y_1, \dots, y_{17})$ an. (1P)

$$\bar{y}_{0.1} =$$

- c) Bestimmen Sie die Koeffizienten der Regressionsgeraden $f(x) = a^* + b^*x$ des Datensatzes, sowie den empirischen Korrelationskoeffizienten r_{xy} . (3P)

$a^* =$

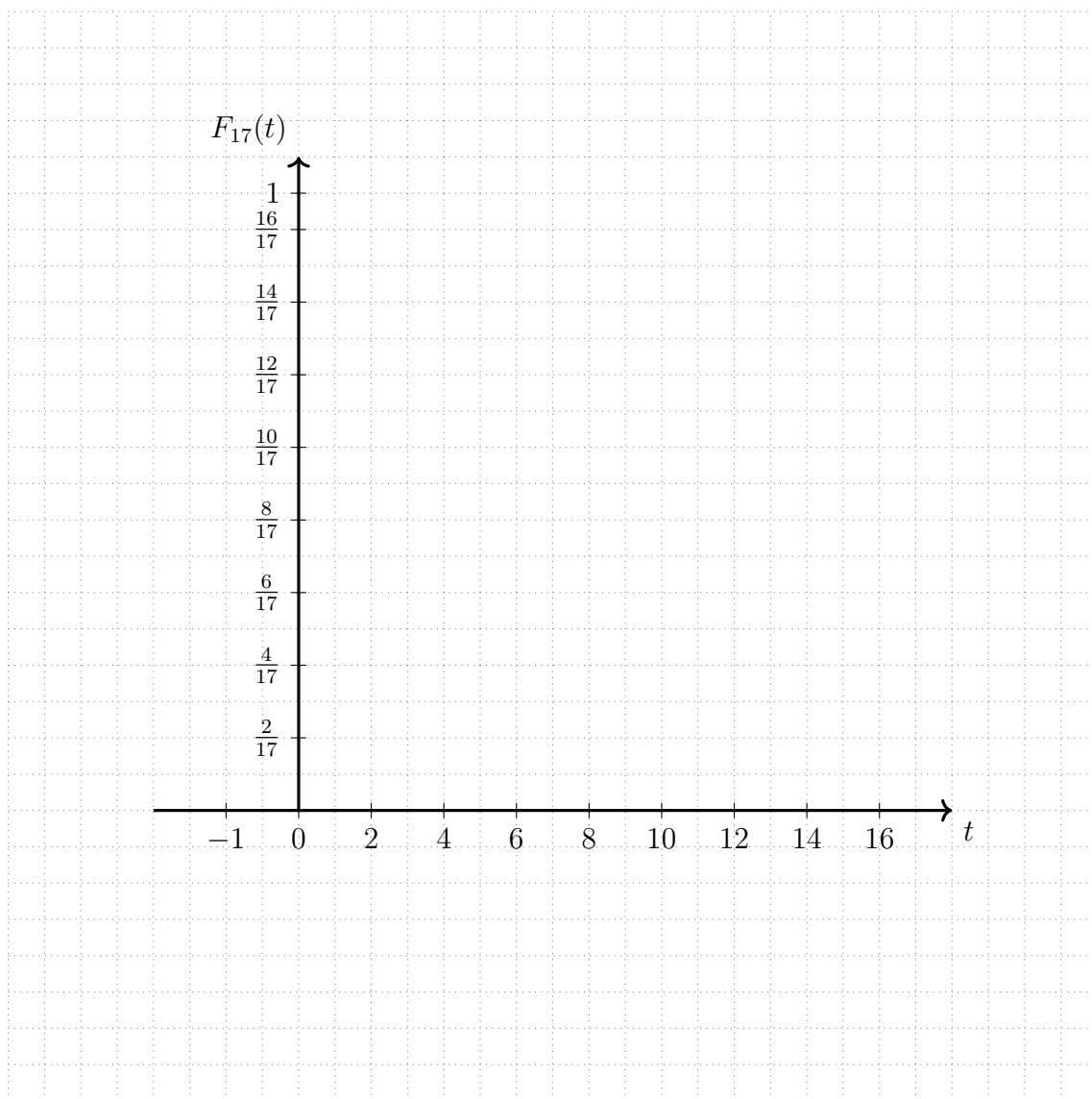
$b^* =$

$r_{xy} =$

- d) Was wäre nach der linearen Regression die Vorhersage für die Anzahl der Tore des KSC am 18. Spieltag, wenn bekannt wäre, dass der KSC in diesem Spiel 6 Eckbälle erhalten hat? Geben Sie das Ergebnis auf 4 Nachkommastellen genau an. (1P)

Vorhergesagte Tore =

e) Zeichnen Sie die empirische Verteilungsfunktion F_{17} von $(x_1, \dots, x_{17})$. (1P)



Aufgabe 3 (10 Punkte)

Gegeben seien zwei **unabhängige** Zufallsvariablen X, Y , deren gemeinsame Verteilung durch die unvollständige Tabelle

$j \backslash k$	0	1	2	3	$\mathbb{P}(Y = j)$
-1	$\frac{1}{20}$				$\frac{1}{5}$
0		$\frac{3}{40}$			
2				$\frac{1}{20}$	$\frac{1}{5}$
$\mathbb{P}(X = k)$		$\frac{1}{8}$			

festgelegt ist.

a) Vervollständigen Sie die Tabelle. (2P)

b) Berechnen Sie: (0.5 + 0.5 + 0.5 + 0.5P)

$$\mathbb{E}[X] = \boxed{}$$

$$\text{Cov}(X, Y) = \boxed{}$$

$$\mathbb{E}[Y] = \boxed{}$$

$$\text{Var}(Y) = \boxed{}$$

c) Betrachten Sie nun unabhängige und identisch verteilte $Y_1, \dots, Y_n$, wobei Y_i dieselbe Verteilung wie Y hat. Berechnen Sie den folgenden Grenzwert: (1P)

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{1}{n} \left| \sum_{i=1}^n \left(Y_i - \frac{1}{5} \right) \right| \leq \frac{1}{8} \right) = \boxed{}$$

In Karlsruhe sind $\frac{1}{3}$ der Tage eines Jahres verregnet. Die Zuverlässigkeit der Regenvorhersage (Regen ja oder nein) hängt von der jeweiligen Wetterlage ab. An den Tagen, an denen es regnet, stimmt die Vorhersage für diesen Tag in $\frac{2}{3}$ aller Fälle, an den anderen Tagen in $\frac{3}{4}$ aller Fälle.

Sie können annehmen, dass die Regenvorhersagen verschiedener Tage stochastisch unabhängig sind. Sie können zudem annehmen, dass die Ereignisse, dass es an unterschiedlichen Tagen regnet, ebenfalls unabhängig sind.

- d) Wie groß ist die Wahrscheinlichkeit, dass es an einem Tag regnet und an diesem Tag auch Regen vorhergesagt ist? (1P)

- e) Mit welcher Wahrscheinlichkeit wird an einem (rein zufälligen) Tag des Jahres Regen vorhergesagt? (1P)

- f) Wie groß ist die Wahrscheinlichkeit, dass es wirklich regnet, wenn Regen vorhergesagt ist? (1P)

- g) Mit welcher Wahrscheinlichkeit ist die Regenvorhersage eines Tages korrekt? (1P)

- h) Mit welcher Wahrscheinlichkeit ist die Regenvorhersage an 3 Tagen in Folge korrekt? (1P)
Hinweis: Falls Sie den Wert aus Teil g) nicht berechnen können, dann verwenden Sie im Folgenden den (nicht korrekten) Wert $\mathbb{P}(„\text{Vorhersage eines Tages korrekt}“) = \frac{2}{3}$.

Aufgabe 4 (10 Punkte)

Für zwei unabhängige normalverteilte Zufallsvariablen $X \sim N(0, 2)$ und $Y \sim N(-2, 1)$ seien

$$Z_1 := 2X + Y \quad \text{und} \quad Z_2 := 2X \cdot Y.$$

- a) Bestimmen Sie die Erwartungswerte und Varianzen von Z_1 und Z_2 . (2P)

$$\begin{aligned}\mathbb{E}[Z_1] &= \boxed{}, & \text{Var}(Z_1) &= \boxed{}, \\ \mathbb{E}[Z_2] &= \boxed{}, & \text{Var}(Z_2) &= \boxed{}.\end{aligned}$$

- b) Geben Sie die Dichte von Z_1 an. (1.5P)

$$f(x) =$$

- c) Berechnen Sie folgende Wahrscheinlichkeit: (1.5P)

$$\mathbb{P}(-1.04 \leq Z_1 \leq 0.1) = \boxed{}$$

- d) Drücken Sie folgenden Erwartungswert mit Hilfe der Dichte φ der Standardnormalverteilung aus: (1P)

$$\mathbb{E}[|Z_1|] =$$

- e) Berechnen Sie für $x > 0$: (1P)

$$\mathbb{P}(Z_2 > x) - \mathbb{P}(Z_2 < -x) =$$

- f) Bestimmen Sie die Mediane von Z_1 und Z_2 . (1P)

$$m(Z_1) = \boxed{}, \quad m(Z_2) = \boxed{}$$

g) Wie groß ist die Kovarianz und Korrelation von Z_1 und Z_2 ? (2P)

$$\text{Cov}(Z_1, Z_2) = \boxed{}, \quad \text{Corr}(Z_1, Z_2) = \boxed{}$$

Aufgabe 5 (10 Punkte)

Seien $X, X_1, \dots, X_n, \dots$ unabhängig identisch verteilte Zufallsvariablen mit Verteilung

$$\mathbb{P}_\vartheta(X = k) = e^{-\vartheta} \cdot \frac{\vartheta^k}{k!}, \quad k = 0, 1, 2, \dots,$$

für einen unbekannten Parameter ϑ .

- a) Für welche $\vartheta \in \mathbb{R}$ ist $\mathbb{P}_\vartheta$ **keine** Zähldichte? (1P)

$$\vartheta \in \boxed{}$$

Nun sei $\vartheta > 0$.

- b) Geben Sie einen erwartungstreuen Schätzer $\hat{p}(X_1, \dots, X_n)$ für $\mathbb{P}_\vartheta(X = 5)$ an. (1P)

$$\hat{p}(X_1, \dots, X_n) = \boxed{}$$

- c) Bestimmen Sie den Erwartungswert von X . (1P)

$$\mathbb{E}_\vartheta[X] = \boxed{}$$

- d) Der unbekannte Parameter ϑ soll basierend auf einer unabhängigen Stichprobe $x = (x_1, \dots, x_n)$ mit Hilfe der Maximum-Likelihood-Methode geschätzt werden.

Bestimmen Sie die Likelihood-Funktion $L_x(\vartheta)$, berechnen Sie die Loglikelihood-Funktion $\ell_x(\vartheta)$ und deren Ableitung $\ell'_x(\vartheta)$. (3P)

$$L_x(\vartheta) = \boxed{}$$

$$\ell_x(\vartheta) = \boxed{}$$

$$\ell'_x(\vartheta) = \boxed{}$$

Geben Sie einen Maximum-Likelihood-Schätzer $\hat{\vartheta}(x)$ für ϑ an. (1P)

$$\hat{\vartheta}(x) = \boxed{}$$

e) Bestimmen Sie den Erwartungswert von $\hat{v}(X_1, \dots, X_n)$. (1P)

$$\mathbb{E}_{\vartheta}[\widehat{\vartheta}(X_1, \dots, X_n)] =$$

f) Berechnen Sie den quadratischen Verlust von $\hat{\vartheta}(X_1, \dots, X_n)$. (1P)

$$\mathbb{E}_{\vartheta}[|\widehat{\vartheta}(X_1, \dots, X_n) - \vartheta|^2] =$$

Hinweis: Sie können ohne Beweis verwenden, dass $\text{Var}(X) = \vartheta$.

g) Nun sei $\vartheta = 2$. Berechnen Sie den folgenden Grenzwert. (1P)

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(-\frac{1}{2}\sqrt{2n} \leq \sum_{j=1}^n (X_j - \mathbb{E}[X_1]) \leq \sqrt{2n} \right) = \boxed{}$$