



Karlsruher Institut für Technologie

Institut für Stochastik

Prof. Dr. Mathias Trabs

M. Sc. Sebastian Höfer

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

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?

$$\binom{k+5}{k}$$

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

$$f_X(x) = \mathbb{1}_{[0,1]}(x)$$

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

$$\text{Var}(X) = 1.$$

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:

$$\mathbb{E}X_i = \mathbb{E}X_1 \quad \forall i \in \mathbb{N}$$

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:

$$s_n^2(X_1, \dots, X_n) = \frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X}_n)^2.$$

Lösungsvorschlag:

a) klar.

b) 1. Fächermodell mit 6 Fächern und n ununterscheidbaren Murmeln.

2. Gemeinsame Dichte integrieren nach einer Variable.

3. $\mathbb{E}[X] = \mathbb{P}(A) - \mathbb{P}(A^c) = 0$ und damit $\text{Var}(X) = \mathbb{E}[X^2] = \mathbb{E}[\mathbb{1}_A^2 + (-\mathbb{1}_{A^c})^2] = 1$.

4. Aussage des schwachen Gesetzes großer Zahlen.

5. Klar.

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 Stichprobenquartil $\tilde{y}_{3/4}$ und die empirische Standardabweichung s_y von $(y_1, \dots, y_{17})$ an: (1+2+1P)

$$\bar{y} = \boxed{1.7059} \quad \tilde{y}_{3/4} = \boxed{2} \quad s_y = \boxed{1.1045}$$

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

$$\bar{y}_{0.1} = \boxed{\frac{5}{3}}$$

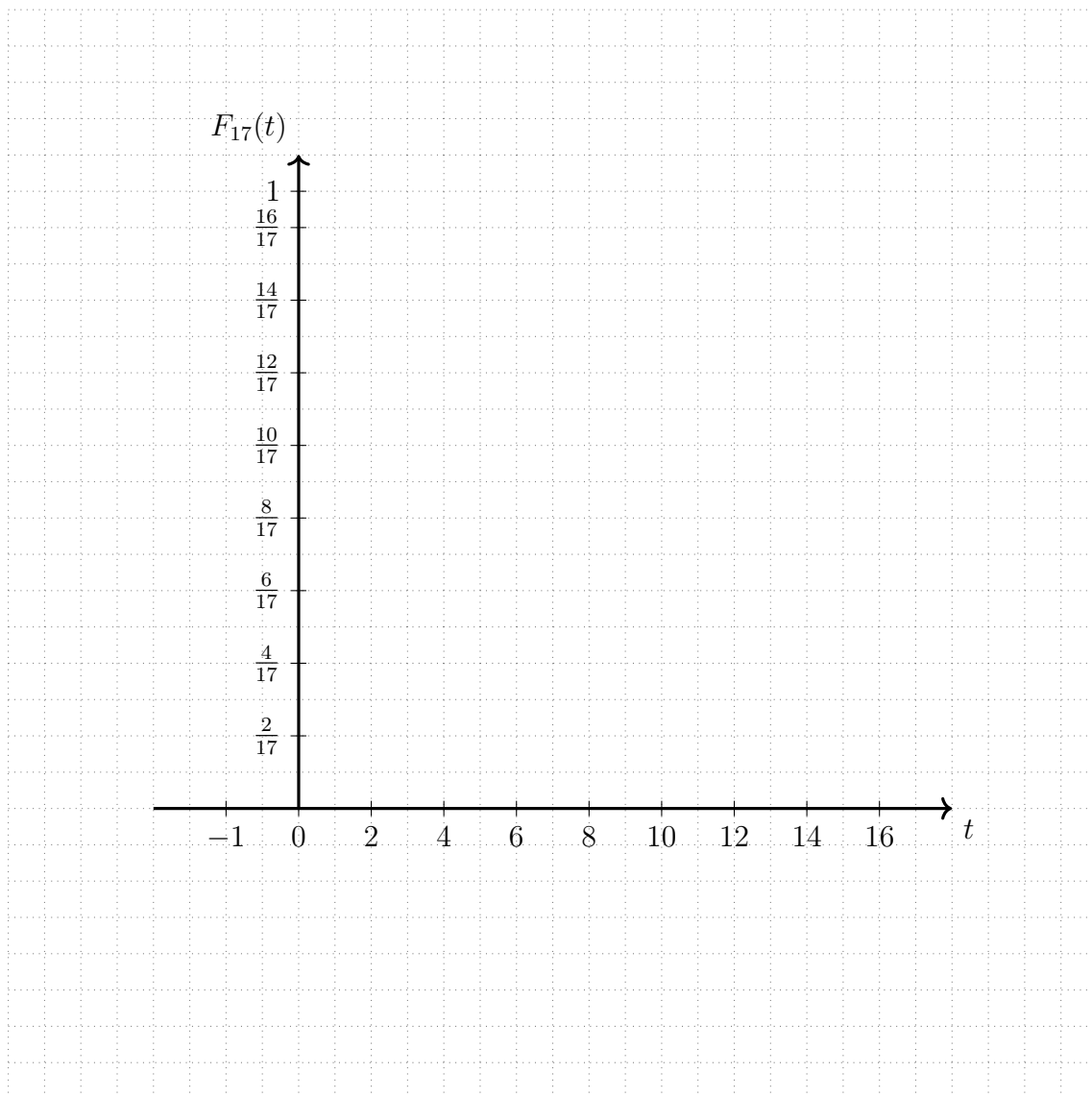
- c) Bestimmen Sie die Koeffizienten der Regressionsgeraden $f(x) = a^* + b^*x$ des Datensatzes, sowie den empirischen Korrelationskoeffizienten r_{xy} . (3P)

$$a^* = \boxed{1.8521} \quad b^* = \boxed{-0.0289} \quad r_{xy} = \boxed{-0.0518}$$

- 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)

$$\text{Vorhergesagte Tore} = \boxed{1.6787}$$

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



Lösungsvorschlag:

- a) – Stichprobenmittel: $\bar{y} = \frac{1}{17} \sum_{j=1}^{17} y_j = 1.7059$
– Geordnete Stichprobe:

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
$y_{(j)}$	0	0	1	1	1	1	1	1	2	2	2	2	2	3	3	3	4

$$\tilde{y}_{3/4} = y_{(13)} = 2$$

- Standardabweichung: $s_y = \sqrt{s_y^2} = 1.1045$

- b) 0.1-getrimmtes Stichprobenmittel: $k = \lfloor 0.1 \cdot 17 \rfloor = 1$.

$$\bar{y}_{0.1} = \frac{1}{15} \sum_{j=2}^{16} y_{(j)} = \frac{5}{3}$$

- c) – $b^* = \frac{\sum_{j=1}^{17} (x_j - \bar{x})(y_j - \bar{y})}{16 \cdot s_x^2} = -0.0289$
– $a^* = \bar{y} - b^* \cdot \bar{x} = 1.8521$
– $r_{xy} = \frac{\frac{1}{16} \sum_{j=1}^{17} (x_j - \bar{x})(y_j - \bar{y})}{s_x s_y} = -0.0518$

- d) Vorhergesagte Tore bei 6 Eckbällen $= a^* + b^* \cdot 6 = 1.6787$

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: $(0.5 + 0.5 + 0.5 + 0.5\text{P})$

$$\mathbb{E}[X] = \boxed{\frac{3}{2}}$$

$$\text{Cov}(X, Y) = \boxed{0}$$

$$\mathbb{E}[Y] = \boxed{\frac{1}{5}}$$

$$\text{Var}(Y) = \boxed{\frac{24}{25}}$$

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)

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

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)

$$\frac{2}{9}$$

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

$$\frac{7}{18}$$

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

$$\frac{4}{7}$$

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

$$\frac{13}{18}$$

- 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 $\mathbb{P}(\text{„Vorhersage eines Tages korrekt“}) = \frac{2}{3}$.

$$\left(\frac{13}{18}\right)^3 \approx 0.3767$$

Lösungsvorschlag:

a) Wir haben:

$j \backslash k$	0	1	2	3	$\mathbb{P}(Y = j)$
-1	$\frac{1}{20}$	$\frac{1}{40}$	$\frac{3}{40}$	$\frac{1}{20}$	$\frac{1}{5}$
0	$\frac{3}{20}$	$\frac{3}{40}$	$\frac{9}{40}$	$\frac{3}{20}$	$\frac{3}{5}$
2	$\frac{1}{20}$	$\frac{1}{40}$	$\frac{3}{40}$	$\frac{1}{20}$	$\frac{1}{5}$
$\mathbb{P}(X = k)$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{4}$	1

b) Für die Erwartungswerte von X und Y haben wir

$$\mathbb{E}[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{8} + 2 \cdot \frac{1}{2} + 3 \cdot \frac{1}{8} = \frac{3}{2}$$

$$\mathbb{E}[Y] = -1 \cdot \frac{1}{5} + 0 \cdot \frac{3}{5} + 2 \cdot \frac{1}{5} = \frac{1}{5}.$$

Weiter ist wegen der Unabhängigkeit von X und Y

$$\text{Cov}(X, Y) = 0.$$

Für die Varianz von Y ergibt sich

$$\text{Var}(Y) = 1 \cdot \frac{1}{5} + 0 \cdot \frac{3}{5} + 4 \cdot \frac{1}{5} - \frac{1}{5^2} = \frac{24}{25}.$$

c) Nach dem schwachen Gesetz großer Zahlen konvergiert $\frac{1}{n} \left| \sum_{i=1}^n \left(Y_i - \frac{1}{5} \right) \right|$ stochastisch gegen 0, für den Grenzwert gilt also

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

Zum zweiten Teil der Aufgabe:

Sei R das Ereignis, dass es regnet und V das Ereignis, dass Regen vorhergesagt wird. Es ist gegeben, dass

$$\mathbb{P}(R) = \frac{1}{3}, \quad \mathbb{P}(V|R) = \frac{2}{3}, \quad \mathbb{P}(V^c|R^c) = \frac{3}{4}.$$

d)

$$\mathbb{P}(R, V) = \mathbb{P}(V | R) \cdot \mathbb{P}(R) = \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}$$

e) Mit der Formel der totalen Wahrscheinlichkeit gilt

$$\begin{aligned} \mathbb{P}(V) &= \mathbb{P}(V|R) \cdot \mathbb{P}(R) + \mathbb{P}(V|R^c) \cdot \mathbb{P}(R^c) = \frac{2}{3} \cdot \frac{1}{3} + \left(1 - \frac{3}{4}\right) \cdot \left(1 - \frac{1}{3}\right) \\ &= \frac{2}{3} \cdot \frac{1}{3} + \frac{1}{4} \cdot \frac{2}{3} = \frac{7}{18} \approx 0.3889. \end{aligned}$$

f) Mit der Formel von Bayes und a) gilt

$$\mathbb{P}(R|V) = \frac{\mathbb{P}(R \cap V)}{\mathbb{P}(V)} = \frac{\mathbb{P}(V|R) \cdot \mathbb{P}(R)}{\mathbb{P}(V)} = \frac{\frac{2}{3} \cdot \frac{1}{3}}{\frac{7}{18}} = \frac{4}{7} \approx 0.5714.$$

g) Sei K das Ereignis, dass die Vorhersage für einen bestimmten Tag richtig ist. Dieses tritt ein, falls es regnet und Regen vorhergesagt wurde oder falls es nicht regnet und kein Regen vorhergesagt wurde. Daher gilt

$$\begin{aligned} \mathbb{P}(K) &= \mathbb{P}(R \cap V) + \mathbb{P}(R^c \cap V^c) = \mathbb{P}(V|R) \cdot \mathbb{P}(R) + \mathbb{P}(V^c|R^c) \cdot \mathbb{P}(R^c) \\ &= \frac{2}{3} \cdot \frac{1}{3} + \frac{3}{4} \cdot \frac{2}{3} = \frac{13}{18} \approx 0.722. \end{aligned}$$

h) Sei nun K_j , $j = 1, 2, 3$, das Ereignis, dass die Regenvorhersage an Tag j korrekt ist. Nach Konstruktion sind die Ereignisse unabhängig und besitzen jeweils dieselbe Wahrscheinlichkeit, d.h. $\mathbb{P}(K_j) = \mathbb{P}(K)$ für $j = 1, 2, 3$. Daraus folgt

$$\mathbb{P}(K_1 \cap K_2 \cap K_3) = \prod_{j=1}^3 \mathbb{P}(K_j) = \mathbb{P}(K)^3 = \left(\frac{13}{18}\right)^3 \approx 0.377.$$

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)

$$\begin{aligned} \mathbb{E}[Z_1] &= \boxed{-2}, & \text{Var}(Z_1) &= \boxed{9}, \\ \mathbb{E}[Z_2] &= \boxed{0}, & \text{Var}(Z_2) &= \boxed{40}. \end{aligned}$$

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

$$f(x) = \boxed{\frac{1}{\sqrt{2\pi}9} \exp\left(-\frac{1}{18}(x+2)^2\right)}$$

c) Berechnen Sie folgende Wahrscheinlichkeit: (1.5P)

$$\mathbb{P}(-1.04 \leq Z_1 \leq 0.1) = \boxed{0.1325}$$

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

$$\mathbb{E}[|Z_1|] = \boxed{\int_{\mathbb{R}} |3x-2| \varphi(x) dx}$$

e) Berechnen Sie für $x > 0$: (1P)

$$\mathbb{P}(Z_2 > x) - \mathbb{P}(Z_2 < -x) = \boxed{0}$$

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

$$m(Z_1) = \boxed{-2}, \quad m(Z_2) = \boxed{0}$$

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

$$\text{Cov}(Z_1, Z_2) = \boxed{-16}, \quad \text{Corr}(Z_1, Z_2) = \boxed{-\frac{8}{3\sqrt{10}}} \approx -0.8433$$

Lösungsvorschlag:

a)

$$\begin{aligned}\mathbb{E}[Z_1] &= \mathbb{E}[2X + Y] = 2\mathbb{E}[X] + \mathbb{E}[Y] = -2 \\ \mathbb{E}[Z_2] &= \mathbb{E}[2XY] = 2\mathbb{E}[X]\mathbb{E}[Y] = 0 \\ \text{Var}(Z_1) &= \text{Var}(2X + Y) \stackrel{Unabh}{=} 4\text{Var}(X) + \text{Var}(Y) = 9 \\ \text{Var}(Z_2) &= \text{Var}(2XY) = 4\text{Var}(XY) = 4(\mathbb{E}[X^2Y^2] - \mathbb{E}[XY]^2) \\ &= 4(\mathbb{E}[X^2]\mathbb{E}[Y^2] - \underbrace{\mathbb{E}[X]^2\mathbb{E}[Y]^2}_{=0}) \\ &= 4\text{Var}(X)(\text{Var}(Y) + \mathbb{E}[Y]^2) = 4 \cdot 2(1 + 4) = 40.\end{aligned}$$

b) Nach dem Additionsgesetz der Normalverteilung gilt $Z_1 \sim N(-2, 9)$.

c) Für $Z \sim N(0, 1)$ ist

$$\begin{aligned}\mathbb{P}(-1.04 \leq Z_1 \leq 0.1) &= \mathbb{P}(-1.04 \leq 3Z - 2 \leq 0.1) = \mathbb{P}(0.32 \leq Z \leq 0.7) \\ &= \Phi(0.7) - \Phi(0.32) = 0.7580 - 0.6255 = 0.1325\end{aligned}$$

d) Sei $Z \sim N(0, 1)$. Dann ist $Z_1 \sim 3Z - 2$. Damit ergibt sich obige Darstellung.

e) Aufgrund der Symmetrie der Normalverteilung gilt:

$$\mathbb{P}(Z_2 > x) - \mathbb{P}(Z_2 < -x) = \mathbb{P}(2XY > x) - \mathbb{P}(-2XY > x) = 0$$

f) Da die Normalverteilungen symmetrisch sind, stimmen die Mediane mit den Erwartungswerten überein.

g) Es gilt

$$\text{Cov}(Z_1, Z_2) = \mathbb{E}[Z_1, Z_2] - \mathbb{E}[Z_1]\mathbb{E}[Z_2] = \mathbb{E}[(2X + Y)2XY] = 4\mathbb{E}[X^2Y] + 2\mathbb{E}[XY^2] = -16$$

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)

$$\vartheta \in \boxed{(-\infty, 0)}$$

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)

$$\hat{p}(X_1, \dots, X_n) = \boxed{\frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{X_j=5\}}}$$

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

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

- 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)

$$L_x(\vartheta) = \boxed{e^{-n\vartheta} \prod_{j=1}^n \frac{\vartheta^{x_j}}{x_j!}}$$

$$\ell_x(\vartheta) = \boxed{-n\vartheta + \sum_{j=1}^n (x_j \log \vartheta - \log(x_j!))}$$

$$\ell'_x(\vartheta) = \boxed{-n + \frac{1}{\vartheta} \sum_{j=1}^n x_j}$$

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

$$\hat{\vartheta}(x) = \boxed{\frac{1}{n} \sum_{j=1}^n x_j}$$

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

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

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

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

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

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

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(-\frac{1}{2} \sqrt{2n} \leq \sum_{j=1}^n (X_j - \mathbb{E}[X_1]) \leq \sqrt{2n} \right) = \boxed{0.5328}$$

Lösungsvorschlag:

- a) Für $\vartheta > 0$ entspricht $\mathbb{P}$ gerade der Zähldichte der Poisson-Verteilung. Auch für $\vartheta = 0$ ergibt sich eine Zähldichte. Also ist die richtige Antwort $\vartheta \in (-\infty, 0)$.
- b) Nach Vorlesung wählen wir als Schätzer für die Wahrscheinlichkeit die empirische Verteilung der Stichprobe.
- c)

$$\begin{aligned}\mathbb{E}_{\vartheta}[X] &= \sum_{k=0}^{\infty} k \cdot e^{-\vartheta} \cdot \frac{\vartheta^k}{k!} \\ &= \vartheta \cdot e^{-\vartheta} \sum_{k=1}^{\infty} \frac{\vartheta^{k-1}}{(k-1)!} = \vartheta \cdot e^{-\vartheta} \cdot e^{\vartheta} = 1.\end{aligned}$$

- d) Die Likelihood-Funktion lautet für die Stichprobe $x = (x_1, \dots, x_n)$

$$L_{\vartheta}(x) = \prod_{j=1}^n P_{\vartheta}(X_j = x_j) = e^{-n\vartheta} \prod_{j=1}^n \frac{\vartheta^{x_j}}{x_j!}.$$

Die Log-Likelihood-Funktion ist dann

$$\ell_x(\vartheta) = \log(L_{\vartheta}(x)) = -n\vartheta + \sum_{j=1}^n x_j \log \vartheta - \log(x_j!)$$

mit der Ableitung nach ϑ

$$\ell'_x(\vartheta) = -n + \frac{1}{\vartheta} \sum_{j=1}^n x_j \stackrel{!}{=} 0.$$

Also ist

$$\vartheta = \frac{1}{n} \sum_{j=1}^n x_j$$

ein stationärer Punkt von $\ell'_x(\vartheta)$ und wegen

$$\ell''_x(\vartheta) = -\frac{1}{\vartheta^2} \sum_{j=1}^n x_j < 0$$

ist es auch ein Maximum. Der Maximum-Likelihood-Schätzer lautet also

$$\hat{\vartheta}(x) = \frac{1}{n} \sum_{j=1}^n x_j.$$

e) Für alle $\vartheta > 0$ gilt wegen der Linearität des Erwartungswerts

$$\mathbb{E}(\widehat{\vartheta}_n(X_1, \dots, X_n)) = \frac{1}{n} \sum_{j=1}^n \mathbb{E}(X_j) = \vartheta.$$

f) Nach der Bias-Varianz-Zerlegung entspricht der quadratische Verlust gerade der Varianz des Schätzers, da es sich um einen erwartungstreuen Schätzer handelt.

g) Nach dem zentralen Grenzwertsatz gilt

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P} \left(-\frac{1}{2} \sqrt{2n} \leq \sum_{j=1}^n (X_j - \mathbb{E}[X_1]) \leq \sqrt{2n} \right) &= \lim_{n \rightarrow \infty} \mathbb{P} \left(-\frac{1}{2} \leq \frac{1}{\sqrt{2n}} \sum_{j=1}^n (X_j - \mathbb{E}[X_1]) \leq 1 \right) \\ &= \Phi(1) - \Phi\left(-\frac{1}{2}\right) = 0.5328. \end{aligned}$$