Dieses PDF ist das Gedächtnisprotokoll der Klausur zum Zweittermin im Modul Stochastik für Informatik (9LP) im Sommersemester 2024. Die Klausur dauerte 90 Minuten. **Für die Richtigkeit der Aufgaben wird keine Gewähr übernommen**. Manche Wortlaute können Fehlerhaft sein, wurden aber nach bestem Wissen und Gewissen niedergeschrieben.

Aufgabe 1

$10 = 2 + 2 + 2 + 2 + 2$ Punkte

Es seien X und Y Zufallsvariablen, deren gemeinsame Verteilung gemäß folgender Tabelle gegeben ist:

X \ Y	-1	0	1	$\mathbb{P}(X = x)$
0	$\frac{1}{16}$	$\frac{6}{16}$	$\frac{1}{16}$	
1	$\frac{1}{8}$	0	$\frac{1}{8}$	
2	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	
$\mathbb{P}(Y = y)$				

(a) Berechnen Sie die Randverteilung von X und Y.

(b) Berechnen Sie $\mathbb{E}[X]$, $\mathbb{E}[Y]$ sowie die Standardabweichung $\sigma(X)$.

(c) Berechnen Sie die $\mathrm{Cov}(X, Y)$. Entscheiden Sie mithilfe der Kovarianz, ob X und Y korreliert sind.

(d) Sind X und Y unabhängig? Begründen Sie Ihre Antwort mit einem kurzen Beweis oder einem Gegenbeispiel.

(e) Berechnen Sie die bedingte Verteilung von X gegeben $Y = 0$.

Aufgabe 2

$10 = 1 + 3 + 4 + 2$ Punkte

Es sei eine homogene Markov-Kette auf dem Zustandsraum $S = \{1, 2, 3\}$ mit der folgenden Übergangsmatrix gegeben:

$$P = (p_{i,j})_{i,j \in S} := \begin{pmatrix} 0 & p & 1-p \\ 1-p & 0 & p \\ p & 0 & 1-p \end{pmatrix}$$

(a) Stellen Sie den Übergangsgraphen der Markov-Kette dar.

(b) Ist die Markov-Kette irreduzibel? Ist sie aperiodisch? Begründen Sie Ihre Antwort.

Ab jetzt betrachten wir $p = \frac{1}{3}$

(c) Bestimmen Sie die invariante Verteilung der Markov-Kette. Ist diese eindeutig?

(d) Bestimmen Sie für alle Knoten $i \in S$ die durchschnittliche Anzahl an Schritten bei Start in einem Knoten i um wieder in i zu landen. D.h. konkret $\mathbb{E}[X_i \mid X_0 = i]$ wobei $X_i = \inf\{n \in \mathbb{N} \mid X_n = i\}$.

(*Wortlaut war etwas anders*)

Aufgabe 3 **10 = 2 + 4 + 4 Punkte**

Sie steigen in das Aktiengeschäft ein. Sei X eine Zufallsvariable welche für jede Minute angibt, wie sich der Preis (in Cent) entwickelt. D.h. $\mathbb{P}(X = 1) = 0.5$ bedeutet, dass die Wahrscheinlichkeit einer Preisänderung (in einer beliebigen Minute) von +1 Cent 0.5 beträgt.

$\mathbb{P}(X = 1) = 0.5$, $\mathbb{P}(X = 0) = 0.2$, $\mathbb{P}(X = -1) = 0.3$

(a) Bestimmen Sie $\mathbb{E}[X]$ sowie $\mathbb{V}[X]$.

(b) Verwenden Sie den zentralen Grenzwertsatz um die Wahrscheinlichkeit zu approximieren, dass in 3 Stunden der Preis um mindestens 50 Cent steigt.

(c) Bestimmen Sie das Konfidenzintervall, in welchem nach 3 Stunden 99% der Aktiengewinne enthalten sind.

(*Wortlaut war etwas anders*)

Aufgabe 4 **10 = 2 + 3 + 2 + 3 Punkte**

Eine kontinuierliche Zufallsvariable X habe die folgende Dichte mit $c \in \mathbb{R}$:

$$f(x) = \begin{cases} c \cdot x^3, & \text{falls } 0 \leq x \leq 1 \\ 0 & \text{sonst.} \end{cases}$$

(a) Bestimmen Sie den Wert der Konstanten c.

Nutzen Sie für die nachfolgende Berechnung den in (a) berechneten Wert für c. Falls Sie (a) nicht gelöst haben, dürfen Sie c weiterhin als Konstante ohne konkreten Wert nutzen.

(b) Berechnen Sie die Verteilungsfunktion von X.

(c) Bestimmen Sie $\mathbb{E}[X]$.

(d) Es sei Y eine Zufallsvariable mit $Y = X^4$. Bestimmen Sie die Verteilung von Y sowie $\mathbb{E}[Y]$.

Aufgabe 5 **10 = 1 + 2 + 2 + 1 + 4 Punkte**

Bei einer Untersuchung der Verunreinigung eines Wasserlaufs wird die Schadstoffkonzentration an 5 verschiedenen Stellen gemessen. Die Standorte befinden sich in unterschiedlichen Entfernungen zur Verschmutzungsquelle. In der folgenden Tabelle sind diese Entfernungen (x, in Km) und die Konzentrationen (y) gegeben.

Entfernung (x)	2	4	6	8	10
Konzentration (y)	11,5	10,2	10,3	9,68	9,32

(a) Berechnen Sie das empirische Mittel für die Entferung und Konzentrationen.

(b) Aus den Dtaen können wir zudem die empirische Varianz $\overline{\sigma}_5^2(y) = 0.68$ für die Konzentration, sowie die empirische Covarianz $\overline{c}_5(x, y) = -2.44$ berechnen. Bestimmen Sie den empirischen Korrelationskoeffizienten $\overline{r}_5(x, y)$.

(c) Bestimmen Sie die Koeffizienten der Regressionsgeraden der Form $y = ax + b$.

(d) Schätzen Sie mithilfe der Regressionsgeraden die Konzentration bei einer Entfernung von 3 Km.

(e) Ab einer mittleren Schadstoffkonzentration von 11 Einheiten müssen die Bewohner im Umkreis des Wasserlaufs über die Verunreinigung informiert werden. Die Untersucher müssen daher anhand der 5 Stichproben ermitteln, ob der wahre Wert der mittleren Schadstoffkonzentration echt kleiner als 11 ist und zwar für ein Signifikanzniveau von $\alpha = 0.01$.