

Übungsblatt 3

Abgabe: Bis **Montag, den 10.06.2024, bis 23:59 Uhr** über Moodle. Die Übungsblätter sind in Gruppen von drei (in Ausnahmefällen zwei) Studierenden zu bearbeiten. Die Lösungen sind, sofern nicht anders angegeben, auf nach Aufgaben getrennten **PDFs** über Moodle abzugeben. Alle Teilaufgaben einer Aufgabe sind in einem PDF hochzuladen. Pro Gruppe genügt es, wenn eine Person die Lösung der Gruppe abgibt. Zur Bewertung wird der zuletzt hochgeladene Stand herangezogen. Vermerken Sie auf allen Abgaben Ihre **Namen**, Ihre **CMS-Benutzernamen** und Ihre **Abgabegruppe (z.B. Gruppe 123)** aus Moodle. Benennen Sie die hochgeladenen PDF-Dateien nach dem Schema: A<Aufgabe>-<Person1>-<Person2>-<Person3>.pdf, bspw. A3-Musterfrau-Mustermann-Beispiel.pdf für Aufgabe 3 von Lisa Musterfrau, Peter Mustermann und Karla Beispiel. Die Auflistung der Namen kann in beliebiger Reihenfolge erfolgen. Beachten Sie die Informationen im Moodle-Kurs <https://hu.berlin/sds24>.

Aufgabe 1 (Kombinatorik)

1 + 2 + 2 + 2 + 3 = 10 Punkte

In dieser Aufgabe beschäftigen wir uns mit verschiedenen Fragestellungen aus der Kombinatorik. Sie sollen jeweils die Anzahl von Zusammenstellungen berechnen. Geben Sie je Teilaufgabe Ihren Rechenweg sowie die Klasse des Problems an (Permutation, Variation, Kombination jeweils mit oder ohne Wiederholung).

- (a) Ein Zahlenschloss verwendet 4 Ziffern. Die dritte Ziffer hakt und kann lediglich die Werte 3 oder 4 annehmen. Jede der anderen Ziffern kann zwischen 0 und 9 liegen. Wie viele verschiedene Zahlenschlüssel sind möglich?
- (b) Sie haben eine Urne mit 2 roten, 3 grünen und 4 blauen Kugeln. Wie viele Möglichkeiten gibt es, 3 Kugeln herauszuziehen, wenn die Reihenfolge der Ziehung (1) berücksichtigt wird und (2) nicht berücksichtigt wird?
- (c) Sie haben 6 verschiedene Bücher und möchten 3 davon auswählen und in einer Reihe aufstellen. Wie viele mögliche Reihen gibt es dafür?
- (d) Sie haben eine Menge $S = \{a, b, c, d\}$ sowie $R = \{1, 2, 3, 4\}$. Bestimmen Sie alle möglichen ungeordneten Zusammenstellungen von 2 Elementen aus S sowie 2 Elementen aus R , wobei die Elemente aus S nicht wiederholt und die Elemente aus R wiederholt werden dürfen.
- (e) Sie haben die Ziffern 1, 2, 3, 4, 5 und sollen eine dreistellige Zahl bilden, die größer als 320 ist. Wie viele solcher Zahlen können gebildet werden, ohne Ziffern zu wiederholen?

Aufgabe 2 (Stetige Zufallsvariablen)**5 + 5 = 10 Punkte**

Sei eine Funktion $f : \mathbb{R} \mapsto \mathbb{R}$ gegeben durch

$$f(x) = \begin{cases} \frac{1}{16}x + \frac{1}{4} & \text{falls } -4 \leq x \leq 0 \\ -\frac{1}{16}x + \frac{1}{4} & \text{falls } 0 < x \leq 4 \\ 0 & \text{sonst} \end{cases}$$

- (a) Zeigen Sie, dass f eine Dichtefunktion einer stetigen Zufallsvariable X ist.
- (b) Bestimmen Sie den Erwartungswert und die Varianz von X .

Aufgabe 3 (Gesetz der großen Zahlen)

2 + 4 + 4 = 10 Punkte

Sie sind Data Scientist und beraten die Betreiber eines Casinos. In der vergangenen Nacht wurden Spieler des Casinos verdächtigt, bei einem Würfelspiel betrogen zu haben. Analysieren Sie die Spielprotokolle der vergangenen Nacht hinsichtlich Informationen zu einem möglichen Betrug.

In Moodle finden Sie die Datei 'casino.csv'. Dieser Datensatz enthält ein Protokoll über Würfelspiele in dem Casino. Jeder Eintrag (Zeile) besteht aus einem Zeitstempel, der ID des Spieltisches, der ID des Spielers, sowie dem Würfelergebnis. Nutzen Sie den Datensatz, um die folgenden Fragen zu beantworten.

- (a) Vor der Analyse des Datensatzes, analysieren Sie zunächst die theoretischen Erwartungen an einen fairen Spielwürfel. Im Spiel kommen spezielle 12-seitige Würfel zum Einsatz, deren Seiten die folgenden Werte haben: 1, 1, 1, 1, 2, 2, 2, 4, 4, 8, 16, 32. Implementieren Sie dazu (siehe Python-Template) die kumulierte Wahrscheinlichkeitsfunktion (CDF) und berechnen Sie den Erwartungswert eines solchen, fairen Würfels.
- (b) Nutzen Sie nun den Datensatz, um die Würfe von Spieler 1 und Tisch B auf Fairness zu untersuchen. Jeder Spieler hat seinen eigenen Würfel und es wird vermutet, dass Spieler 1 um 21 Uhr seinen Würfel mit einem manipulierten Würfel ausgetauscht haben könnte. Berechnen Sie den Sample Mean ab diesem Zeitpunkt und visualisieren Sie die relativen Häufigkeiten der Würfelergebnisse von Spieler 1 an Tisch B ab 21 Uhr in einem Histogramm. Es soll je möglichem Ergebnis (horizontal, kategorische Skala) die relative Häufigkeit (vertikal) dargestellt werden. Beschriften Sie die Achsen. Beschreiben Sie stichhaltig in je 1-2 Sätzen etwaige Unterschiede zwischen (1) dem beobachteten Sample Mean und dem Erwartungswert eines fairen Würfels (Teilaufgabe a), sowie (2) den visualisierten relativen Häufigkeiten und der kumulativen Wahrscheinlichkeit eines fairen Würfels (Teilaufgabe a) für Ergebnisse ≤ 2 .
- (c) Visualisieren Sie den zeitlichen Ablauf der Würfelergebnisse. Implementieren Sie dazu eine Funktion, die den Sample Mean der Würfelergebnisse eines Spielers im Verlauf der Zeit visualisiert. Die Abbildung soll zu jeder Uhrzeit (horizontal) den bisherigen Sample Mean (vertikal) darstellen. Zudem soll der Erwartungswert eines fairen Würfels aus Teilaufgabe (a) als horizontale, grau gestrichelte Linie dargestellt werden. Weiterhin soll der Zeitpunkt des möglichen Würfeltauschs als vertikale, grau gepunktete Linie markiert werden. Erstellen Sie eine Legende zu allen drei Linien und beschriften Sie die Achsen. Beschreiben Sie kurz und stichhaltig, wie Sie die Visualisierungen der Ergebnisse von Spieler 1 und 2 hinsichtlich eines möglicherweise manipulierten Würfels deuten. Nehmen Sie dabei Bezug auf das Gesetz der großen Zahlen.

Nutzen Sie zur Beantwortung der Fragen aus (a), (b) und (c) das in Moodle bereitgestellte Python-Template "aufgabe3.py". Im Template gibt es jeweils eine Funktion je Teilaufgabe, in welcher die Aufgabe programmatisch gelöst werden soll. Die erwarteten Datentypen der Rückgabe entnehmen Sie der Beschreibung zu Beginn jeder Funktion.

Verändern Sie **nicht** die Signatur der definierten Funktionen. Darüber hinaus können Sie beliebige Hilfsfunktionen definieren. Sie dürfen lediglich die Open-Source-Bibliotheken NumPy, pandas und matplotlib nutzen.