Stochastik
(BSc D-MAVT / BSc D-MATH / BSc D-MATL)

Schreiben Sie für Aufgabe 2-4 stets alle Zwischenschritte und -rechnungen sowie Begründungen auf. Aufgabe 1 ist eine Multiple Choice Aufgabe (keine Begründungen notwendig).

Die Tabellen der Normalverteilung, t-Verteilung und Verwerfungsbereiche für den Wilcoxon-Test wurden mit der Prüfung ausgeteilt.

1. (10 Punkte)

Bei den folgenden 10 Fragen ist jeweils genau eine Antwort richtig. Es gibt pro richtig beantwortete Frage 1 Punkt und pro falsche Antwort 1/2 Punkt Abzug. Minimal erhält man für die gesamte Aufgabe 0 Punkte. **Bitte benützen Sie das beiliegende Antwortblatt.**

a) Seien A, B und C beliebige Ereignisse.

1. $P[A \cup B \cup C] = P[A] + P[B] + P[C] - P[A \cap B \cap C]$.
2. $P[A \cap B] = P[A \cap B \cap C] + P[A \cap B \cap C^c]$.
3. $P[A \cap B \cap C] = P[A]P[B]P[C]$.

b) Seien X und Y unabhängig und identisch verteilt mit Varianz $\sigma^2 > 0$.

1. $\mathrm{Var}(X + X - Y) = 5\sigma^2$.
2. $\mathrm{Var}(X + Y) = \mathrm{Var}(2X)$.
3. $\mathrm{Var}(X - Y) \leq \sigma^2$.

c) Sei $X \sim \mathrm{Pois}(\lambda)$ mit $\lambda > 0$.

1. $P[X > 5] = 1 - P[X < 5]$.
2. $P[X \geq 1 | X \leq 1] = \frac{\lambda}{\lambda+1}$.
3. $2X \sim \mathrm{Pois}(2\lambda)$.

Bitte wenden!

d) Seien T_1 und T_2 unabhängig und geometrisch verteilt mit Erfolgsparameter p.

1. $\min(T_1, T_2)$ ist geometrisch verteilt mit Erfolgsparameter $1 - (1-p)^2$.
2. $\min(T_1, T_2)$ ist geometrisch verteilt mit Erfolgsparameter $(1-p)^2$.
3. $\min(T_1, T_2)$ ist nicht geometrisch verteilt.

e) Sei X die Anzahl Einsen, die Fritz in 10 unabhängigen Würfen mit einem fairen Würfel wirft. Welche Verteilung kommt für X in Frage?

1. Geometrische Verteilung mit Parameter $p = \frac{1}{6}$.
2. Geometrische Verteilung mit Parameter $p = \frac{1}{10}$.
3. Binomialverteilung mit Parametern $n = 10$ und $p = \frac{1}{6}$.

f) Sei U eine auf $(0, 1)$ uniform verteilte Zufallsvariable, das heisst $U \sim \text{Uni}(0, 1)$.

1. $U/2 \sim \text{Uni}(0, 2)$.
2. $\frac{1}{2}\log(1-U) \sim \text{Exp}(2)$.
3. $E\left[\frac{1}{2}\log(1-U)\right] = -\frac{1}{2}$.

g) Seien $X_1, X_2, \ldots, X_n$ unabhängig und identisch verteilt mit $E[X_1] = \mu$ und $\text{Var}(X_1) = \sigma^2 > 0$. Wir betrachten $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

1. $\text{Var}(\overline{X}_n) = \sigma^2$.
2. $\overline{X}_n$ ist für grosse n approximativ normalverteilt mit $E[\overline{X}_n] = \mu$ und $\text{Var}(\overline{X}_n) = n\sigma^2$.
3. $T = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$ ist ein erwartungstreuer Schätzer für σ^2, das heisst $E[T] = \sigma^2$.

h) Es gelten die folgenden Aussagen für Schätzer.

1. Für $X_1, X_2, \ldots, X_n$ i.i.d. $\sim \mathcal{N}(\mu, \sigma^2)$ liefert die Momentenmethode den gleichen Schätzer für μ wie die Maximum-Likelihood-Methode.
2. Falls $X_1, X_2, \ldots, X_n$ i.i.d. $\sim \text{Exp}(\lambda)$-verteilt sind, so ist der Momentenschätzer für λ gegeben durch $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.
3. Momentenschätzer sind immer erwartungstreu.

i) Bei einem statistischen Test gilt:

1. Das Signifikanzniveau kontrolliert die Wahrscheinlichkeit für einen Fehler 2. Art.
2. Falls der P-Wert kleiner als das Niveau ist, so verwirft der Test die Nullhypothese.
3. Wenn das Niveau verkleinert wird, so wird der Verwerfungsbereich grösser.

Siehe nächstes Blatt!

j) In einer Eisenabsorptionsstudie soll untersucht werden, ob Eisen von einem Präparat im Körper aufgenommen wird. Dazu wurden bei 10 Frauen vor und nach der Einnahme des Präparats Blutproben entnommen und auf den Eisenstatus untersucht. Uns interessiert, ob der Eisenstatus nach Einnahme des Präparats höher ist.

1. Es handelt sich um eine gepaarte Stichprobe und der Test ist zweiseitig durchzuführen.
2. Es handelt sich um eine gepaarte Stichprobe und der Test ist einseitig durchzuführen.
3. Es handelt sich um eine ungepaarte Stichprobe und der Test ist zweiseitig durchzuführen.

Bitte wenden!

2. (8 Punkte)

Alice ist auf dem Weg zur Arbeit und wartet an einer Bushaltestelle, an der die drei Buslininen 1, 2 und 3 halten. Da alle drei Buslinien Alices Arbeitsort bedienen, steigt sie einfach in den ersten Bus ein, der vorbeifährt. Wir nehmen an, dass gleich viele Busse der Linien 1, 2 und 3 unterwegs sind und in zufälliger Reihenfolge an der Bushaltestelle vorbeikommen. Da die Linie 1 ein beliebtes Einkaufszentrum anfährt und deshalb viele Leute mit diesem Bus fahren, wird bei der Linie 1 auf jeder fünften Fahrt eine Billettkontrolle durchgeführt. Bei den Linien 2 und 3 wird hingegen nur jede sechste Fahrt kontrolliert.

a) Wie gross ist die Wahrscheinlichkeit, dass Alice auf ihrem Arbeitsweg nicht kontrolliert wird?

b) Wie gross ist die Wahrscheinlichkeit, dass Alice auf ihrem Arbeitsweg nicht mit der Linie 1 fährt oder nicht kontrolliert wird?

Während der Weihnachtszeit wird Linie 1 besonders stark beansprucht. Infolgedessen entscheidet sich der Busbetreiber, mehr Busse auf der Linie 1 einzusetzen. Neu werden auf Linie 1 j mal so viele Busse eingesetzt wie auf Linie 2. Die Anzahl Fahrten Linien 2 und 3 bleiben gleich. Alle Busse kommen weiterhin zufällig an der Bushaltestelle vorbei und kontrolliert wird wie oben.

c) Angenommen Alice wurde auf dem Arbeitsweg kontrolliert. Wie gross darf j maximal sein, damit die Wahrscheinlichkeit, dass Alice mit der Linie 2 oder 3 zur Arbeit gefahren ist, mindestens 25% beträgt?

Nach der schokoladenreichen Weihnachtszeit entscheidet sich Alice dazu, mehr Sport zu treiben und fährt ab sofort zwei Mal pro Arbeitswoche (5 Tage) mit dem Fahrrad zur Arbeit. Die übrigen drei Tage fährt sie nach wie vor mit dem Bus. Die Wahrscheinlichkeit beim Fahrradfahren von einem Polizisten kontrolliert zu werden, liegt bei p_1, während die Chance der Billettkontrolle im Bus bei p_2 liegt. Bemerke, dass $p_1, p_2 \in (0, 1)$.

d) Wie gross ist die Wahrscheinlichkeit (in Abhängigkeit von p_1 und p_2), dass Alice innert einer Arbeitswoche (5 Tage) auf ihrem Arbeitsweg genau 4 Mal kontrolliert wird, wenn die Kontrollen unabhängig voneinander stattfinden?

Siehe nächstes Blatt!

3. (9 Punkte)

Bei einem Laborexperiment werden Mäuse mit Grippe angesteckt. Danach wird die Zeit X (in Tagen), bis sie krank werden, und die Zeit Y, bis sie wieder gesund werden, gemessen. Die gemeinsame Verteilung von X, Y kann mit folgender Dichte modelliert werden:

$$f_{X,Y}(x,y) = \begin{cases} cx^2(y-x)\mathrm{e}^{-y} & \text{falls } 0 < x < y, \\ 0 & \text{sonst.} \end{cases}$$

a) Berechnen Sie die Randdichte $f_X(x)$ von X und die Randdichte $f_Y(y)$ von Y (ohne die Normierungskonstante c zu bestimmen).

b) Berechnen Sie c, so dass $f_{X,Y}(x,y)$ eine Wahrscheinlichkeitsdichte ist.
Hinweis: $\int_0^\infty u^n \mathrm{e}^{-u} du = n!$.

c) Berechnen Sie $E[X]$, $E[Y]$ und $\mathrm{Cov}(X,Y)$.

Wir definieren nun die Zufallsvariable $Z = Y - X$ (die Erholungszeit nach der Erkrankung). Die gemeinsame Dichte von X und Z ist gegeben durch (nicht zu beweisen)

$$f_{X,Z}(x,z) = \begin{cases} cx^2 z\mathrm{e}^{-(x+z)} & \text{falls } x,z > 0, \\ 0 & \text{sonst.} \end{cases}$$

d) Sind X und Z unabhängig?

Ein neues Medikament, das die Erholungszeit reduzieren soll, wird getestet. Deshalb modellieren wir nun die Erholungszeit Z mit folgender Dichte:

$$f_Z(z) = \begin{cases} \frac{z}{\theta^2}\mathrm{e}^{-z/\theta} & \text{falls } z > 0, \\ 0 & \text{sonst.} \end{cases}$$

e) Bestimmen Sie den Maximum-Likelihood Schätzer $\hat{\theta}$ für θ basierend auf n unabhängigen Beobachtungen $Z_1, \ldots, Z_n \overset{\text{i.i.d.}}{\sim} Z$.

4. (8 Punkte)

Eine Maschine produziert Aluminiumrohre, deren Qualität durch den Durchmesser des Rohrs charakterisiert ist. Bei der Maschine kann der Durchmesser eingestellt werden. Der ideale Durchmesser der Rohre beträgt 9.75 mm. Um nachzuprüfen, ob die Maschine korrekt kalibriert ist, wurde der Durchmesser bei 5 Rohren gemessen, mit folgenden Ergebnissen (in mm):

$$9.4 \qquad 9.7 \qquad 9.05 \qquad 8.95 \qquad 9.5.$$

Die einzelnen Durchmesser seien unabhängig und normalverteilt mit $\sigma = 0.3$ mm. Wir möchten mit einem statistischen Test beurteilen, ob die Maschine richtig kalibriert ist.

a) Geben Sie ein geeignetes Modell an und formulieren Sie die Null- und Alternativhypothese.

b) Geben Sie die Teststatistik an und bestimmen Sie ihre Verteilung unter der Nullhypothese.

c) Bestimmen Sie den Verwerfungsbereich zum Signifikanzniveau $\alpha = 0.05$ und geben Sie den Testentscheid an.

d) Bestimmen Sie den P-Wert.

e) Berechnen Sie die Macht des Tests unter der Annahme, dass die Maschine auf einen Durchmesser von 9.9 mm kalibriert ist.