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

Schreiben Sie für Aufgabe 2-4 stets alle Zwischenschritte und -rechnungen sowie Begründungen auf. Vereinfachen Sie die Resultate so weit wie möglich. 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 Anwort 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 und B zwei Ereignisse mit $A \subset B$ und $P[A] > 0$.

1. $P[A^c] \leq P[B^c]$.
2. $P[A] = P[A|B]P[B]$.
3. $P[B^c|A] = P[B^c]$.

b) Seien A und B beliebige Ereignisse mit $0 < P[B] < 1$.

1. $P[A|B^c] + P[A|B] = 1$.
2. $P[A|B] + P[A^c|B] = 1$.
3. $P[A|B^c] = P[A]$, falls A und B disjunkt sind.

c) Seien X und Y identisch verteilt mit $\mathrm{Var}(X) = 2$, $E[X] = 0$ und $E[XY] = 1$.

1. $\mathrm{Var}(X - 3Y + 2) = 20$.
2. $\mathrm{Var}(X - 3Y + 2) = 14$.
3. $\mathrm{Var}(X - 3Y + 2) = 17$.

Bitte wenden!

d) Die gemeinsame Wahrscheinlichkeitsfunktion von X und Y ist in folgender Tabelle gegeben.

X\Y	-1	0	2	$p_X(x)$
0	$\frac{1}{6}$	0	$\frac{1}{12}$	$\frac{1}{4}$
1	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{12}$	$\frac{3}{4}$
$p_Y(y)$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$	

1. X und Y sind unkorreliert und unabhängig.
2. X und Y sind unkorreliert aber abhängig.
3. X und Y sind korreliert und abhängig.

e) Es sei $X \sim \mathcal{N}(\mu, \sigma^2)$ mit $\mu = 3$ und $\sigma^2 = 4$.

1. $P[X \leq 1] \leq P[X \geq 4]$.
2. $E[X^2] = 7$.
3. $Y = e^X \sim \mathcal{N}(e^\mu, e^{\sigma^2})$.

f) Ein System bestehend aus zwei unabhängigen Komponenten X_1 und X_2 funktioniert solange mindestens eine Komponente funktioniert. Die Lebensdauern der Komponenten seien identisch verteilt mit Verteilungsfunktion F_X. Die Verteilungsfunktion der Lebensdauer des Systems sei $F(x)$.

1. $F(x) = 1 - (1 - F_X(x))^2$.
2. $F(x) = \big(1 - F_X^2(x)\big)$.
3. $F(x) = F_X^2(x)$.

g) Seien X und Y unabhängige Poisson-verteilte Zufallsvariablen mit Parametern λ beziehungsweise μ.

1. $X + Y$ ist normalverteilt.
2. $-X$ ist Poisson-verteilt.
3. $\mathrm{Var}(X - Y) = \lambda + \mu$.

h) Seien $X_1, X_2, \ldots, X_n \sim$ i.i.d. Pois(λ) und $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$.

1. $\frac{1}{n}\sum_{i=1}^n \big(X_i - \overline{X}_n\big)^2$ ist ein erwartungstreuer Schätzer für λ.
2. $\frac{1}{n-1}\sum_{i=1}^n \big(X_i - \overline{X}_n\big)^2$ ist ein erwartungstreuer Schätzer für λ^2.
3. $\frac{1}{n-1}\sum_{i=1}^n \big(X_i - \overline{X}_n\big)^2$ ist ein erwartungstreuer Schätzer für λ.

Siehe nächstes Blatt!

i) Wir betrachten einen z-Test zum Niveau α und das zugehörige $100 \times (1-\alpha)\%$-Vertrauensintervall für μ.

1. Wenn der Stichprobenumfang n erhöht wird, dann wird die Länge des Vertrauensintervalls kürzer.
2. Der Verwerfungsbereich hängt von den Beobachtungen ab, ist also eine zufällige Grösse.
3. Das Vertrauensintervall fängt den wahren Parameter mit Wahrscheinlichkeit α ein.

j) Wir betrachten einen statistischen Test zum Niveau α und einen P-Wert p.

1. Wenn das Niveau vergrössert wird, wird die Wahrscheinlichkeit für einen Fehler 1. Art vergrössert.
2. Wenn das Niveau vergrössert wird, wird das zugehörige Vertrauensintervall grösser.
3. Wenn das Niveau vergrössert wird, wird der P-Wert grösser.

2. (8 Punkte)

In einer Vorlesung wird aus hundert Studenten jede Woche zufällig einer ausgewählt, der an der Wandtafel eine Aufgabe vorlösen soll. Da dieser Test keinen Einfluss auf die Prüfungsnote hat, bereiten sich 70% der Studenten nicht auf diese Aufgabe vor, die restlichen 30% repetieren zur Vorbereitung jeweils den Vorlesungsstoff der Vorwoche. Ein fleissiger Student, der sich ausreichend vorbereitet, kann die Testaufgabe mit einer Wahrscheinlichkeit von p_1 richtig lösen, während die Erfolgsquote für einen faulen Studenten bei p_2 liegt. Bemerke, dass $p_1, p_2 \in (0,1)$ sind. Wir nehmen an, dass alle hundert Studenten die Vorlesung immer besuchen.

a) Wie gross ist die Wahrscheinlichkeit (in Abhängigkeit von p_1 und p_2), dass die Testaufgabe falsch gelöst wird?

b) Angenommen der ausgewählte Student löst die Aufgabe richtig. Wie gross ist die Wahrscheinlichkeit (in Abhängigkeit von p_1 und p_2), dass sich dieser nicht vorbereitet hat?

Der Dozent realisiert, dass eine Testaufgabe nicht ausreicht um zu entscheiden, ob der Student den Vorlesungsstoff verstanden hat oder nicht. Deshalb entscheidet er sich dafür, dem zufällig ausgewählten Studenten neu fünf unabhängige Aufgaben zu verschiedenen Vorlesungsinhalten zu stellen. Analog zu den Teilaufgaben **a)** und **b)** nehmen wir an, dass ein fleissiger Student jede einzelne Aufgabe mit Wahrscheinlichkeit p_1 richtig löst, während die Erfolgsquote für einen faulen Student bei p_2 liegt. Der Anteil fleissiger Studenten beträgt nach wie vor 30%.

c) Wie gross ist die Wahrscheinlichkeit (in Abhängigkeit von p_1 und p_2), dass ein fleissiger Student ausgewählt wird und dieser mindestens eine der fünf Aufgaben richtig löst?

d) Angenommen ein Student löst genau zwei Aufgaben richtig. Wie gross ist die Wahrscheinlichkeit (in Abhängigkeit von p_1 und p_2), dass dieser sich nicht vorbereitet hat?

Alice und Bob sind die ersten beiden Studenten, die die Übungsaufgaben nach dem neuen Format lösen müssen. Wir nehmen an, dass Alice und Bob unabhängige Aufgaben lösen mussten.

e) Sei $k \in \{0, 1, \cdots, 10\}$ eine natürliche Zahl. Wie gross ist die Wahrscheinlichkeit (in Abhängigkeit von p_1 und p_2), dass Alice und Bob *zusammen* genau k Aufgaben richtig lösen, falls Alice eine fleissige Studentin ist und Bob sich nicht auf die Fragen vorbereitet hat?

Siehe nächstes Blatt!

3. (9 Punkte)

In einem Labor werden Nanodrähte von 10μm Länge erzeugt, die $1\text{pg} = 10^{-12}$g wiegen und aus Platin und Nickel bestehen. Aufgrund des Herstellungsverfahrens sind der Nickelinhalt X (in pg) und die Dauer der Funktionsfähigkeit T (in Sekunden) der Drähte zufällig, mit gemeinsamer Dichte

$$f_{X,T}(x,t) = \begin{cases} cx\cos\left(\frac{\pi}{2}x\right)\mathrm{e}^{-xt}, & \text{falls } x \in (0,1),\ t > 0, \\ 0, & \text{sonst.} \end{cases}$$

a) Zeigen Sie, dass die Randdichte $f_X(x)$ von X gegeben ist durch

$$f_X(x) = \begin{cases} c\cos\left(\frac{\pi}{2}x\right), & \text{falls } x \in (0,1), \\ 0, & \text{sonst.} \end{cases}$$

b) Berechnen Sie c, so dass $f_{X,T}$ eine Wahrscheinlichkeitsdichte ist.

c) Berechnen Sie die bedingte Dichte von T gegeben, dass der Nickelinhalt $X = x$ eines Nanodrahts bekannt ist. Wie heisst diese Verteilung?

d) Berechnen Sie die Wahrscheinlichkeit, dass der Nickelinhalt weniger als 0.5pg beträgt.

e) Berechnen Sie den Erwartungswert und die Varianz von X (ohne den numerischen Wert von π einzusetzen).

Wir nehmen nun an, dass $E[X] = 0.3634$ und $\mathrm{Var}(X) = 0.0574$.

f) Berechnen Sie approximativ die Wahrscheinlichkeit, dass 100 unabhängige Nanodrähte insgesamt weniger als 35pg Nickel enthalten.
Hinweis: Normal-Approximation.

4. (8 Punkte)

Für die nächste Wahl zum Präsidenten gibt es zwei Kandidaten A und B. Christoph behauptet, dass nur 10 % der Bevölkerung den Kandidaten A wählen. Lena ist skeptisch und vermutet, dass es mehr als 10% sind. Diese Vermutung möchte sie mit einem statistischen Test prüfen. Dazu hat sie 10 Leute auf der Strasse gefragt, ob sie den Kandidaten A wählen. Von den 10 befragten Leuten unterstützen 2 den Kandidaten A. Wir nehmen an, dass die Antworten der einzelnen Leute unabhängig sind.

Verwenden Sie für die Rechnungen die untenstehenden Tabellen mit den Wahrscheinlichkeiten für $X \sim \text{Bin}(n = 10, p)$ für $p = 0.1$ und $p = 0.5$.

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

b) Welche Teststatistik verwenden Sie? Geben Sie auch die Verteilung der Teststatistik unter der Nullhypothese an.

c) Berechnen Sie den Verwerfungbereich auf dem 5% Niveau. Wie entscheidet der Test?

d) Berechnen Sie die Macht des Tests unter der Annahme, dass 50% der Bevölkerung für den Kandidaten A stimmt.

e) Berechnen Sie den P-Wert.

$X \sim \text{Bin}(10, p)$	$i = 0$	$i = 1$	$i = 2$	$i = 3$	$i = 4$
$P_{p=0.1}[X = i]$	0.3487	0.3874	0.1937	0.0574	0.0112
$P_{p=0.1}[X \leq i]$	0.3487	0.7361	0.9298	0.9872	0.9984
$P_{p=0.5}[X = i]$	0.0010	0.0098	0.0439	0.1172	0.2051
$P_{p=0.5}[X \leq i]$	0.0010	0.0108	0.0547	0.1719	0.3770

$X \sim \text{Bin}(10, p)$	$i = 5$	$i = 6$	$i = 7$	$i = 8$	$i = 9$	$i = 10$
$P_{p=0.1}[X = i]$	0.0015	0.0001	0	0	0	0
$P_{p=0.1}[X \leq i]$	0.9999	1	1	1	1	1
$P_{p=0.5}[X = i]$	0.2461	0.2051	0.1172	0.0439	0.0098	0.0010
$P_{p=0.5}[X \leq i]$	0.6231	0.8282	0.9454	0.9893	0.9991	1