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 für die Aufgaben benötigten Tabellen (Normalverteilung, Quantile der t-Verteilung) wurden mit der Prüfung ausgeteilt.

1. **(10 Punkte)** Geben Sie jeweils die richtige Aussage an, indem Sie **einen Kreis um die entsprechende Ziffer** machen! Pro Teilaufgabe ist genau eine Aussage korrekt! Eine richtig gelöste Teilaufgabe gibt 1 Punkt, ansonsten 0 Punkte.

   a) Die Varianz der Zufallsvariablen $Y = 2X$ ist im Vergleich zur Varianz der Zufallsvariablen X

   i) gleich,
   ii) doppelt so gross,
   iii) viermal so gross.

   b) Die Varianz der Zufallsvariablen $Y = X - 10$ ist im Vergleich zur Varianz der Zufallsvariablen X

   i) kleiner,
   ii) gleich,
   iii) grösser.

   c) In welchem Fall darf man annehmen, dass die Ereignisse $A_1, \ldots, A_n$, bzw. die Zufallsvariablen $X_1, \ldots, X_n$, *unabhängig* sind?

   i) A_k = "Am 1. März 2008 gibt es einen Stromausfall beim k-ten Zürcher Haushalt" (n = Anzahl Haushalte in Zürich).
   ii) X_k = "Körpergrösse des k-ten Dozenten der ETH Zürich" (n = Anzahl Dozenten an der ETH Zürich).
   iii) A_k = "Das k-te Bio-Menü, serviert am 3. März 2008 in der Hauptmensa, wurde zu lange gekocht" (n = Anzahl servierter Bio-Menüs an diesem Tag).

Bitte wenden!

d) Sei T_1, T_2 i.i.d. $\sim \text{Exp}(\lambda)$. Dann gilt:

i) $\min(T_1, T_2) \sim \text{Exp}(2\lambda)$,
ii) $\min(T_1, T_2) \sim \text{Exp}(\lambda/2)$,
iii) keine der beiden Aussagen stimmt.

e) Sei T_1, T_2 i.i.d. $\sim \text{Exp}(\lambda)$. Dann gilt:

i) $\max(T_1, T_2) \sim \text{Exp}(2\lambda)$,
ii) $\max(T_1, T_2) \sim \text{Exp}(\lambda/2)$,
iii) keine der beiden Aussagen stimmt.

f) Die Aussage $P(A|B^c) = 1 - P(A|B)$ ist im Allgemeinen für beliebige Ereignisse A und B

i) richtig,
ii) richtig, falls A und B unabhängig sind,
iii) falsch.

g) Man hat 1000 Werte einer Zufallsvariablen mit stetiger Verteilung simuliert. Was stellt das aus diesen Daten resultierende Histogramm dar?

i) Die empirische Dichte der Verteilung von X.
ii) Die empirische Verteilungsfunktion der Verteilung von X.
iii) Die Abweichungen zum Mittelwert von X.

h) Bei einer Prüfung war die Durchschnittsnote 4.6 und der Median lag bei 5. Sie haben die Note 4.7 erhalten.

i) Sie gehören ganz klar zur besseren Hälfte Ihrer Klasse, weil sie über dem Notendurchschnitt liegen.
ii) Obwohl Sie über dem Notendurchschnitt liegen, gehören Sie zur schlechteren Hälfte Ihrer Klasse.
iii) Um eine der obigen Aussagen machen zu können, braucht man Angaben über die Streuung der Noten (empirische Varianz).

i) Vertrauensintervall

i) Wenn das Niveau eines statistischen Tests kleiner wird, dann wird das Vertrauensintervall kleiner.
ii) Wenn das Niveau eines statistischen Tests kleiner wird, dann wird das Vertrauensintervall grösser.
iii) Das Niveau und das Vertrauensintervall haben keine direkte Beziehung zueinander.

j) P-Wert

i) Der P-Wert ist ein absoluter Wert, der nur vom Test-Typ abhängt.
ii) Den P-Wert gibt man sich vor bevor man einen statistischen Test durchführt.
iii) Der P-Wert hängt von den beobachteten Daten ab.

Siehe nächstes Blatt!

2. **(8 Punkte)** Gemäss Schätzungen befindet sich etwa 1 Terrorist unter 1 Million Flugpassagieren. An einem Flughafen wird deshalb eine neuartige Anlage zur Terrorismusbekämpfung eingeführt. Eine Lampe an der Anlage leuchtet grün, falls der Flugpassagier kein Sicherheitsrisiko darstellt, und sie leuchtet rot, falls es sich beim Flugpassagier um einen Terroristen handelt. Gemäss Angaben der Behörde gilt der Test als sehr zuverlässig. Falls es sich nämlich bei einem Flugpassagier um einen Terroristen handelt, zeigt die Anlage dies mit einer Wahrscheinlichkeit von 99.9% richtig an. Falls es sich aber um einen Passagier ohne Sicherheitsrisiko handelt, so ist die Wahrscheinlichkeit, dass die Lampe fälschlicherweise rot anzeigt, der folgende Wert:

$$\frac{0.999 \cdot 10^{-4}}{1 - P(\text{ein zufällig ausgewählter Passagier ist ein Terrorist})} (< 0.01\%).$$

a) Herr Meier wird am Flughafen mit der neuen Anlage kontrolliert. Die Anlage leuchtet rot. Mit welcher Wahrscheinlichkeit ist Herr Meier ein Terrorist? Geben Sie das Resultat als Bruch und vollständig gekürzt an.

b) Welche der folgenden Aussagen ist richtig? Die Wahrscheinlichkeit, dass es sich um einen Terroristen handelt, falls die Lampe grün leuchtet, ist:

i) 100/101,

ii) 0.1%,

iii) keine der beiden Aussagen stimmt.

c) Wie würde das Resultat aus **a)** lauten, wenn sich 1 Terrorist unter 10'000 Flugpassagieren befinden würde? Geben Sie das Resultat als Bruch und vollständig gekürzt an.

An einem bestimmten Tag möchte Herr Meier vom Flughafen abfliegen. Um einen Passagier abzufertigen, benötigt man Z Minuten, wobei Z exponentialverteilt ist mit Parameter $\lambda = 4$, d.h. die Dichte von Z ist gegeben durch $f(x) = \lambda e^{-\lambda x}$, für $x \geq 0$.

d) Bestimmen Sie den Median der Wartezeit Z, d.h. bestimmen Sie x (auf drei Dezimalstellen hinter dem Komma gerundet) so, dass $P(Z > x) = 1/2$. *Hinweis: Benutzen Sie die folgende Näherung:* $\ln(2) \approx 0.693$.

3. **(12 Punkte)** Sie sind zuständig für den Strassenunterhalt. Der Verkehr verursacht immer wieder Schlaglöcher, die Sie auffüllen müssen. Die Löcher haben eine zylindrische Form mit Durchmesser D und Tiefe H. Der Durchmesser D ist verteilt gemäss der Dichtefunktion

$$f(x) = \begin{cases} c(1-x)^2, & 0 \leq x \leq 1, \\ 0, & \text{sonst.} \end{cases}$$

a) Bestimmen Sie die Konstante c und den Erwartungswert von D.

b) Bestimmen Sie die Verteilungsfunktion von D.

Bitte wenden!

c) Die Tiefe H der Löcher sei uniform verteilt zwischen 0 m und 0.2 m, und unabhängig vom Durchmesser. Wie gross ist der erwartete Materialverbrauch (in m^3) pro Loch? *Hinweis: Das Volumen eines Loches ist gegeben durch* $V = \frac{\pi}{4} H D^2$.

d) Nehmen wir an, dass alle Löcher genau 0.1 m tief sind. Wie gross ist die Wahrscheinlichkeit, dass Sie für ein Loch mehr als $x = \pi/160$ m^3 Material brauchen?

e) Ihr Chef verlässt sich eher auf seine Erfahrung als auf Ihre Rechenkünste. Aus Erfahrung weiss er, dass der mittlere Materialverbrauch pro Loch $\mu = 1/100$ m^3 ist, und dass die Abweichung des Materialverbrauchs vom Mittel etwa durch die Standardabweichung $\sigma = 1/200$ m^3 beschrieben wird. Sie sollen einen neuen Lastwagen kaufen, der gerade so gross ist, dass Sie mit einer Ladung mit einer Wahrscheinlichkeit von 95% ganze 1000 Löcher füllen können. Wie gross wählen Sie die Ladekapazität (in m^3)? (Verwenden Sie die Normalapproximation!)

4. (12 Punkte) Die Firma Paraffin hat kürzlich eine neue Kerze auf den Markt gebracht, die sich nur durch eine andere Kerzenwachsmischung von der alten Kerze unterscheidet. Für die neue Kerze wird mit dem Argument geworben, dass sie länger brenne. Mit einem Test soll nun überprüft werden, ob dieses Argument, mit dem Werbung für die neue Kerze gemacht wird, gerechtfertigt ist. Dazu nimmt man jeweils 9 neue und 9 alte Kerzen und lässt sie im Labor kontrolliert niederbrennen. Man misst folgende Brenndauern (in Minuten):

Kerze Nr.	1	2	3	4	5	6	7	8	9
alte Kerze: X_i	36	38	39	40	43	51	52	53	92
neue Kerze: Y_i	43	43	49	53	59	63	69	81	83

Aus der obigen Tabelle kann man folgende Werte errechnen:

$$\bar{X}_9 = 49.33,\ \bar{Y}_9 = 60.33,\ s_X^2 = 298,\ s_Y^2 = 226,\ s_{X-Y}^2 = 101,\ s_{pool}^2 = 262.$$

Man kann davon ausgehen, dass die Brenndauer gut durch eine Normalverteilung beschrieben werden kann, wobei sich die Varianzen nicht unterscheiden, d.h. $X_i \sim \mathcal{N}(\mu_X, \sigma^2)$ und $Y_i \sim \mathcal{N}(\mu_Y, \sigma^2)$ (und σ unbekannt).

a) Geben Sie ein zweiseitiges 95%-Vertrauensintervall für die erwartete Brenndauer der *neuen* Kerze an.

Führen Sie nun einen geeigneten Test zum Niveau 5% durch, um festzustellen, ob der Werbespruch ('längere Brenndauer') für die neue Kerze gerechtfertigt ist.

b) Wie lauten die Null- und Alternativhypothese?

c) Handelt es sich um einen gepaarten oder ungepaarten Vergleich?

d) Ist der Test ein- oder zweiseitig?

e) Geben Sie den Verwerfungsbereich für den obigen Test an (Niveau 5%).

f) Wie entscheidet der Test?