Wahrscheinlichkeit und Statistik
BSc D-INFK

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) Sei X eine Zufallsvariable mit Erwartungswert 3 und Varianz 5. Ferner sei $Z = 2X - 7$. Dann gilt:

1) $\mathrm{Var}(Z) = 20$.
2) $\mathrm{Var}(Z) = 10$.
3) $\mathrm{Var}(Z) = 3$.

b) Seien A und B Ereignisse mit $P(A) = \frac{1}{3}$ und $P(A \cup B) = \frac{3}{4}$. Ferner sei $P(A|B) = \frac{1}{2}$. Wie gross ist $P(B)$?

1) $P(B) = \frac{2}{3}$.
2) $P(B) = \frac{5}{6}$.
3) $P(B) = \frac{7}{12}$.

c) Sei X eine Zufallsvariable mit Verteilungsfunktion F_X. Ferner sei $Y := e^X$ und F_Y die zugehörige Verteilungsfunktion. Dann gilt:

1) $F_Y(y) = F_X(e^y)$.
2) $F_Y(y) = e^{F_X(y)}$.
3) $F_Y(y) = F_X(\log(y))$.

d) Sei X eine positive Zufallsvariable mit Dichte f und es sei $y > 0$. Welche ist die korrekte Formel, um $P(X^2 \geq y)$ zu berechnen?

1) $P(X^2 \geq y) = \int_{\sqrt{y}}^{\infty} f(x)dx$.
2) $P(X^2 \geq y) = \int_{y}^{\infty} \sqrt{f(x)}dx$.
3) $P(X^2 \geq y) = \int_{y}^{\infty} f(x)^2dx$.

Bitte wenden!

e) Seien X und Y Zufallsvariablen mit Var$(X) = 2$, Var$(Y) = 3$, und Var$(X + Y) = 7$. Ferner sei $Z = 4X - 3Y + 2$. Dann gilt:

1) Cov$(Y, Z) = -5$.
2) Cov$(Y, Z) = -3$.
3) Cov$(Y, Z) = 13$.

f) Sei $X \sim$ Bin$(n, \frac{1}{3})$ für ein $n \in \mathbb{N}$. Ferner nehmen wir an, dass $E[X] = 1$. Wie gross ist $P(X \geq 2)$?

1) $P(X \geq 2) = \frac{7}{3^3}$.
2) $P(X \geq 2) = \frac{3}{3^3}$.
3) $P(X \geq 2) = \frac{21}{3^3}$.

g) Sei $X \sim \mathcal{N}(0, 1)$ und sei Y eine Zufallsvariable unabhängig von X, so dass $X + Y \sim \mathcal{N}(1, 5)$. Dann gilt:

1) $E[Y^2] = 5$.
2) $E[Y^2] = 4$.
3) $E[Y^2] = 3$.

h) Seien $T_1 \sim$ Exp(λ_1) und $T_2 \sim$ Exp(λ_2) mit $\lambda_1, \lambda_2 > 0$. Dann gilt:

1) $\max(T_1, T_2) \sim$ Exp$(\lambda_1 + \lambda_2)$, falls T_1 und T_2 unabhängig sind.
2) $\max(T_1, T_2) \sim$ Exp$(\lambda_1 + \lambda_2)$ allgemein (egal, ob T_1, T_2 unabhängig sind oder nicht)
3) Keine der beiden Aussagen stimmt.

i) Sei $X_1, X_2, \ldots$ eine Folge von unabhängigen und identisch verteilten Zufallsvariablen mit Erwartungswert $E[X_i] = \mu$ und Varianz $\mathrm{Var}(X_i) = \sigma^2$. Setze $S_n := \sum_{i=1}^{n} X_i$ für $n \in \mathbb{N}$. Wie kann man S_n für "grosse" n approximieren?

1) $S_n \sim \mathcal{N}(\mu, \sigma^2/n)$.
2) $S_n \sim \mathcal{N}(n\mu, n\sigma^2)$.
3) $S_n \sim \mathcal{N}(n\mu, \sqrt{n}\sigma^2)$.

j) Bei einem statistischen Test und dem zugehörigen Vertrauensintervall gilt:

1) Je grösser das Testniveau α, desto grösser das Vertrauensintervall.
2) Das Testniveau α entspricht der Wahrscheinlichkeit eines Fehlers 2. Art.
3) Ein Fehler 1. Art kann nur auftreten, wenn der Wert der Teststatistik im Verwerfungsbereich liegt.

Siehe nächstes Blatt!

2. (7 Punkte)

a) In eine Urne liegen drei faire Münzen. Zwei sind normale Münzen (Kopf/Zahl), die dritte jedoch hat zwei Köpfe. Nun wird eine der drei Münzen zufällig gewählt und diese Münze wird ein Mal geworfen. Berechnen Sie die Wahrscheinlichkeit, dass das Resultat des Münzwurfes “Kopf” ist?

b) Aus einer Untersuchung einer Krankheit sei folgendes bekannt. Die Wahrscheinlichkeit, dass eine zufällig gewählte Person diese Krankeit hat betrage zehn Prozent. Die Diagnose gestaltet sich jedoch als schwierig und das Untersuchungsergebnis ist darum in ein Prozent der Fälle falsch. Wie gross ist dann die Wahrscheinlichkeit, dass eine zufällig gewählte Person diese Krankheit hat, gegeben dass sie positiv getestet wurde?

c) Ein Casinobetreiber bietet das folgende Spiel an. Die Spielerin würfelt mit einen (fairen) Würfel und das Casino würfelt mit zwei (fairen) Würfeln. Die Spielerin gewinnt, falls der Wert ihres Würfels plus Eins grösser oder gleich der Summe der Augenzahlen der beiden Würfeln des Casinos ist. Berechnen Sie die Wahrscheinlichkeit, dass die Spielerin den Wert “Eins” auf ihrem Würfel hat, gegeben dass sie gewonnen hat.
Hinweis: Sie dürfen verwenden, dass die Spielerin mit einer Wahrscheinlichkeit von 7/27 gewinnt (muss nicht gezeigt werden).

3. (10 Punkte) Eine Fluglinie fliegt jeden Morgen von London nach Zürich. Die Stornierungsrate bei den Buchungen betrage 1/5, d.h. die Wahrscheinlichkeit, dass ein Passagier zwar den Flug bucht, aber das Ticket dann nicht bezahlt ist 20%.

a) Nehmen Sie an, dass die Passagiere unabhängig voneinander ihre Entscheidungen treffen und dass 100 Passagiere diesen Flug buchen am 1. Februar 2010. Wie ist die Anzahl X der Stornierungen verteilt?

b) Wie gross ist die Wahrscheinlichkeit, dass von 100 Passagieren höchstens 25 die Buchung stornieren?

i) Geben Sie die Formel an, mit welcher sich diese Wahrscheinlichkeit exakt berechnen lässt (ohne die eigentliche Berechnung).

ii) Verwenden Sie nun den zentralen Grenzwertsatz um diese Wahrscheinlichkeit approximativ zu berechnen.

c) Wir führen nun einen Binomialtest auf dem Signifikanzniveau $\alpha = 10\%$ durch. Dabei seien die Nullhypothese $H_0 : p = p_0 = 0.1$ und die Alternativhypothese $H_A : p > p_0$ gegeben. Der Verwerfungsbereich dieses Tests bei allgemein n Buchungen ist als $X \geq c(n)$ gegeben, d.h. $P_{H_0}(X \geq c(n)) = \alpha$. Zeigen Sie dass sich $c(n)$ im Falle der Normalapproximation aus **b)** ii) schreiben lässt als

$$c(n) = \Phi^{-1}(1-\alpha)\sqrt{n}\sqrt{p_0(1-p_0)} + np_0.$$

d) Bestimmen Sie die kleinste Anzahl $n \in \mathbb{N}$ der Buchungen, sodass der in **c)** durchgeführte Test bei einem wahren Wert $p_A = 0.2$ einen Fehler 2. Art von maximal 10% hat.

Hinweis: Verwenden Sie den in **c)** angegebenen Verwerfungsbereich und benützen Sie die Näherung $\Phi^{-1}(0.9) \approx 9/7$.

4. (10 Punkte) Zwei Typen von Handyakkus werden bezüglich ihrer Laufzeit untersucht. Dazu wurden für beide Akkus die jeweiligen Laufzeiten (in Stunden) an 9 verschiedenen Testhandys gemessen. In der untenstehenden Tabelle sind diese Akkulaufzeiten für Akkutyp A (Werte a_i) und Akkutyp B (Werte b_i) für die 9 Handys angegeben.

i	1	2	3	4	5	6	7	8	9
a_i	41	51	44	46	20	36	46	53	28
b_i	34	41	38	24	23	42	35	31	23
$d_i = a_i - b_i$	7	10	6	22	-3	-6	11	22	5

Die folgenden aus den Daten berechneten Mittelwerte und Standardabweichungen stehen zur Verfügung: $\bar{a} = 40.5, \bar{b} = 32.3, S_a = 10.8, S_b = 7.5, S_d = 9.6, S_{pool} = 9.3$.

a) Handelt es sich hier um gepaarte oder ungepaarte Stichproben?

Es soll nun getestet werden, ob es einen signifikanten Unterschied begüglich Laufzeiten zwischen den beiden Akkutypen gibt. Nehmen Sie dazu an, dass die Daten normalverteilt sind.

b) Führen Sie einen t-Test zum Niveau $\alpha = 5\%$ durch und gehen Sie dabei wie folgt vor:

i) Wählen Sie eine geeignete Null- und Alternativhypothese.
ii) Ist der Test ein- oder zweiseitig durchzuführen? (kurze Begründung)
iii) Berechnen Sie den Verwerfungsbereich K des Tests.
iv) Wie entscheidet der Test? (Benutzen Sie, dass gilt $\frac{246}{96} \approx 2.56$)

Der Hersteller eines weiteren Akkutyps C behauptet, dass seine Akkus eine längere Laufzeit habe als die Akkus vom Typ A. Wir sind skeptisch und führen darum einen einseitigen t-Test durch. Der berechnete Wert der Teststatistik sei dabei $T_{obs} = -2.54$ und der Verwerfungsbereich von der Form $(-\infty, c_\alpha]$ für einen Wert $c_\alpha \in \mathbb{R}$.

c) Geben Sie für diesen Fall eine möglichst genaue ganzzahlige Abschätzung für den P-Wert an (z.B. “P-Wert liegt zwischen 6% und 7%”).