Wahrscheinlichkeit und Statistik (BSc D-INFK)

1. (9 Punkte) Bei den folgenden 9 Fragen wird nur die Lösung verlangt. Der Lösungsweg wird nicht bewertet. Es gibt pro richtig beantwortete Frage 1 Punkt.

a) Seien A und B Ereignisse in einem Wahrscheinlichkeitsraum.
Es gelte $P[A \cup B^c] = \frac{2}{3}$, $P[B] = \frac{2}{3}$ und $P[A \cap B^c] = \frac{1}{6}$. Berechnen Sie $P[A]$.

b) Sei die Zufallsvariable $X \sim \text{Binom}(n, \frac{1}{2})$ für ein $n \in \mathbb{N}$ und es gelte $P[X \leq 1] = \frac{5}{16}$. Bestimmen Sie $P[X \geq 3]$.

c) Seien X_1 und X_2 Zufallsvariablen mit

$$E[X_1] = 2,\ E[X_2] = 4,\ \text{Var}[X_1] = 12,\ \text{Var}[X_2] = 9,\ \text{Var}[1 + \frac{1}{2}X_1 + \frac{1}{3}X_2] = 3.$$

Bestimmen Sie $E[X_1 X_2]$.

d) Seien X_1 bzw. X_2 unabhängige Zufallsvariablen mit Verteilung $\mathcal{N}(\mu_1, \sigma_1^2)$ bzw. $\mathcal{N}(\mu_2, \sigma_2^2)$. Bestimmen Sie die Verteilung von $X_3 := 1 + aX_1 + bX_2$ für $a, b \in \mathbb{R}$.

e) Seien X_1, X_2 unabhängige, exponentialverteilte Zufallsvariablen mit Parameter $\lambda > 0$, d.h. die Dichte von X_i, $i \in \{1, 2\}$ ist gegeben durch $f_{X_i}(x) = \lambda e^{-\lambda x},\ x \geq 0$. Sei die Zufallsvariable Y gegeben als $Y := X_1 + X_2$. Bestimmen Sie die Verteilung von Y.

f) Seien nun $X_1, \ldots, X_n,\ n \in \mathbb{N}$ wie in Teilaufgabe **e)**, und seien mit $\bar{X}_n,\ n \in \mathbb{N}$ die Zufallsvariablen $\bar{X}_n := \frac{1}{n}\sum_{i=1}^n X_i$ bezeichnet. Bestimmen Sie diejenige Verteilung, mit der sich die Verteilung von $\bar{X}_n$ für grosse n approximieren lässt.

g) Sei $X > 0$ eine Zufallsvariable mit Dichtefunktion f_X und sei $Y := \frac{1}{X}$. Bestimmen Sie die Verteilungsfunktion $F_Y(t)$ und die Dichtefunktion $f_Y(t)$, $t > 0$ von Y in Abhängigkeit von F_X und f_X.

h) Welcher Typ von Fehler kann bei einem statistischen Test vorliegen, wenn die Nullhypothese akzeptiert wurde?

i) Die Zufallsvariable X hat die Dichte

$$f_X(x) = \begin{cases} \theta e^{-\theta x}, & x \geq 0 \\ 0, & \text{sonst.} \end{cases}$$

Es liegen folgende Beobachtungen von X vor: $x_1 = 0.27, x_2 = 0.22, x_3 = 0.23, x_4 = 0.25, x_5 = 0.28$. Bestimmen Sie den Maximum-Likelihood Schätzer $\hat{\theta}(x_1, \ldots, x_5)$ für θ.

Bitte wenden!

2. (5 Punkte) Wir betrachten eine Eishockey-Play-Off-Serie der Kloten Flyers gegen die ZSC Lions. Die Serie besteht aus maximal drei Runden. Jede Runde hat einen eindeutigen Sieger, es gibt also keine Unentschieden. Die Mannschaft, die zuerst zwei Runden für sich entscheiden kann, gewinnt die Serie. Falls eine Mannschaft die Serie bereits nach zwei Runden gewonnen hat, wird keine dritte Runde mehr gespielt.

Im Folgenden bezeichne F_i für $i = 1, 2, 3$ das Ereignis, dass die Kloten Flyers die i-te Runde gewinnen. Ferner sei R die Gesamtzahl der gespielten Runden und X die Anzahl der Runden, die die Kloten Flyers gewinnen. Folgende Wahrscheinlichkeiten seien gegeben:

$$P[F_1] = \frac{1}{3},\ P[R = 2, X = 0] = \frac{1}{4},\ P[R = 2, X = 2] = \frac{1}{6},\ P[R = 3, X = 1] = \frac{1}{3}.$$

Beachten Sie: Setzen Sie erst am Ende Ihrer Berechungen konkrete Zahlen ein. Es muss klar werden, woher die Zahlen stammen.

a) Mit welcher Wahrscheinlichkeit gewinnen die Flyers die Serie in insgesamt drei Runden?

b) Berechnen Sie $P[F_2]$.

Wir nehmen nun an, die zwei Mannschaften treffen jede Spielsaison in einer Serie dieser Form aufeinander. Ausserdem nehmen wir an, dass die Ergebnisse der Serien unabhängig voneinander sind. Es bezeichne N die Saison (mit $N = 1$ falls die ZSC Lions in der ersten Saison gewinnen), in der die ZSC Lions zum ersten Mal eine Serie für sich entscheiden.

c) Benennen Sie die Verteilung von N und bestimmen Sie den/die zugehörigen Parameter. Berechnen Sie ausserdem den Erwartungswert $E[N]$.

Siehe nächstes Blatt!

3. (8 Punkte)

In einer Versicherung wird die Schadenshöhe eines Schadensfalls als eine Zufallsvariable S mit Dichte

$$f_S(x) = \begin{cases} 0 & \text{für } x < 0, \\ \frac{1}{2000} & \text{für } x \in [0, 1000], \\ \frac{c}{x^4} & \text{für } x > 1000, \end{cases}$$

modelliert, wobei c eine Konstante ist.

a) Wie muss man c wählen, damit f_S eine Dichte ist?

b) Berechnen Sie die Verteilungsfunktion der Schadenshöhe in diesem Modell.

c) Berechnen Sie die erwartete Schadenshöhe in diesem Modell. Berechnen Sie die Varianz.

d) Die Versicherung hat ein System, wobei Schadensfälle nach Schadenshöhe kategorisiert werden, und entsprechend unterschiedlich behandelt werden. Die erste Kategorie besteht aus Schadensfällen mit Schadenshöhen von weniger als 500 Franken, und die zweite Kategorie besteht aus Schadensfällen mit Schadenshöhen zwischen 500 und 2000 Franken. Wie viel Prozent der Schadensfällen gehören in diesem Modell im Durchschnitt zur *zweiten* Kategorie?

e) Die höchste Kategorie besteht aus den Schadensfällen, deren Schadenshöhen zu den 0.4 Prozent höchsten aller Schadenshöhen gehören. Wie hoch muss die Schadenshöhe (in diesem Modell) sein, damit ein Schadensfall zur höchsten Kategorie gehört?

Bitte wenden!

4. (7 Punkte) Die Post hat festgestellt, dass normalerweise 5% aller Sendungen auf dem Postweg verloren gehen. Der Online-Shop *Azamon.com* möchte diese Information benutzen, um betrügerische Kunden zu erkennen.

a) Welche Verteilung können wir benutzen, um die Anzahl X der verlorenen Pakete für einen Kunde, der n Bestellungen gemacht hat, zu modellieren?

b) *Shopper99* hat 15 Bestellungen gemacht und zwei von ihnen als "auf dem Postweg verloren gegangen" angezeigt. *Azamon.com* möchte testen, ob dieser Kunde betrügerisch ist.

i) Formulieren Sie eine geeignete Null- und Alternativ-Hypothese.

ii) Bestimmen Sie den Verwerfungsbereich K für X zum Niveau $\alpha = 5\%$. Ist *Shopper99* nach Auswertung dieses Tests als Betrüger anzusehen?

iii) Wie gross ist die Wahrscheinlichkeit, unter H_0, dass zwei oder mehr Pakete von 15 verloren gehen?

Hinweis: Verwenden Sie die folgende Binomial-Tabelle für $n = 15$ und $p = 0.05$:

k	0	1	2	3	4
$P(X \leq k)$	0.463	0.829	0.964	0.995	0.999

c) Der Einfachheit halber nehmen wir an, dass ein Test mit Signifikanzniveau von genau $\alpha = 5\%$ und Macht gleich $1 - \beta = 20\%$ zur Verfügung steht. Das Geschäft hat $N = 100$ Kunden getestet und festgestellt, dass nach diesem Test Y von ihnen als betrügerisch gelten. Wir möchten den Anteil f aller Kunden, die wirklich betrügerisch sind, abschätzen.

i) Was ist die Wahrscheinlichkeit f^* für einen Kunden als betrügerisch betrachtet zu werden? *(Schreiben Sie diese als Ausdruck von α, β und f.)*

ii) Geben Sie den Momenten-Schätzer $\hat{f}(Y)$ für f an.

iii) Berechnen Sie $\hat{f}(8)$.