Wahrscheinlichkeit und Statistik (BSc D-INFK)

1. (7 Punkte)

Bei den folgenden 7 Fragen wird nur die Lösung verlangt. Der Lösungsweg wird nicht bewertet. Es gibt pro richtig beantwortete Frage 1 Punkt.

a) Es seien $P[B] = 3P[A]$, $P[A \cup B] = \frac{5}{6}$ und $P[A \cap B] = \frac{1}{6}$. Wie gross ist $P[A^c]$?

b) Sei

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

Wie ist c zu wählen, damit f eine Wahrscheinlichkeitsdichte darstellt?

c) Wir nehmen an, es seien drei (faire) Münzen geworfen worden. Wir wissen, dass mindestens zwei Würfe Kopf ergeben haben. Was ist dann die Wahrscheinlichkeit, dass alle drei Würfe Kopf ergeben haben?

d) Es seien X, Y Zufallsvariablen mit $E[X] = 5$, $E[Y] = 1$ und $E[(X+1)(Y+2)] = 20$. Bestimmen Sie $\mathrm{Cov}(X, Y)$.

e) Für gewisses $n \in \mathbb{N}$ seien $X_1, \dots, X_n$ unabhängige, identisch verteilte Zufallsvariablen. X_1 habe Verteilungsfunktion F. Bestimmen Sie die Verteilungsfunktion der Zufallsvariablen $Y = \max\{X_1, \dots, X_n\}$ in Abhängigkeit von F.

f) Sei $X_1, X_2, \dots$ eine Folge von unabhängigen, identisch verteilten Zufallsvariablen mit $E[X_i] = \mu$ und $\mathrm{Var}(X_i) = \sigma^2$ für alle i. Wir definieren

$$Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i, \qquad \text{für } n \in \mathbb{N}.$$

Durch welche Verteilung kann man die Verteilung von Y_n für "grosse n" approximieren?

g) Für gewisses $\vartheta > 1$ habe die Zufallsvariable X die Dichte

$$f_\vartheta(x) = \begin{cases} (\vartheta - 1)x^{-\vartheta} & \text{falls } x \geq 1, \\ 0 & \text{sonst.} \end{cases}$$

Es seien n Beobachtungen $x_1, \dots, x_n$ von X gegeben. Bestimmen Sie den Maximum-Likelihood Schätzer für ϑ.

Bitte wenden!

2. (9 Punkte) Student Christoph fährt jeden Tag mit dem Velo an die ETH. Dabei kommt er an genau zwei Kreuzungen mit Ampeln vorbei. Aus Erfahrung weiss Christoph, dass er an der ersten Ampel mit Wahrscheinlichkeit 25% warten muss. Auf Grund der Ampelsteuerung und Christophs Fahrgeschwindigkeit hängt die Wahrscheinlichkeit, dass Christoph an der zweiten Ampel warten muss davon ab, ob er an der ersten Ampel warten musste oder nicht. Falls er an der ersten Ampel nicht warten musste, ist die Wahrscheinlichkeit an der zweiten auch nicht warten zu müssen $\frac{2}{3}$. Hatte er an der ersten Ampel hingegen "rot", muss er mit Wahrscheinlichkeit $\frac{1}{7}$ auch an der zweiten Ampel warten. Christoph kommt genau dann zu spät zur Vorlesung, wenn er an wenigstens einer Ampel warten muss.

a) Mit welcher Wahrscheinlichkeit muss Christoph an genau einer Ampel warten?

b) Mit welcher Wahrscheinlichkeit kommt Christoph zu spät zur Vorlesung?

c) Mit welcher Wahrscheinlichkeit hatte Christoph an der ersten Ampel "grün", gegeben, dass er an der zweiten Ampel nicht warten musste?

Hinweis: Falls Sie Teil **b)** nicht gelöst haben, benutzen Sie im Folgenden für die Wahrscheinlichkeit, dass Christoph zu spät zur Vorlesung kommt den Wert $\frac{1}{2}$.

d) Mit welcher Wahrscheinlichkeit musste Christoph an der ersten Ampel warten, gegeben, dass er zu spät zur Vorlesung kam? Sind die Ereignisse "Christoph kommt zu spät zur Vorlesung" und "Christoph muss an der ersten Ampel warten" unabhängig?

e) Christoph fährt im Februar 10-mal an die ETH. Nehmen Sie an, die Ereignisse, dass er an einem der 10 Tage zu spät zur Vorlesung kommt, seien unabhängig. Wie gross ist die Wahrscheinlichkeit, dass er im Februar mindestens zweimal zu spät zur Vorlesung kommt?

Hinweis: Potenzen wie beispielsweise 2^{10} müssen nicht weiter vereinfacht werden.

3. (12 Punkte)

Wir betrachten einen Eisstand am Zürichsee im späten Frühling.

a) Der Stand öffnet um 10 Uhr. An einem Tag, an dem die Temperatur t Grad beträgt, ist die Zeit Z_1 (in Stunden), bis der erste Kunde kommt, exponentialverteilt mit Parameter t. Was ist die Dichte von Z_1? Was ist die Verteilungsfunktion von Z_1?

b) Z_1 ist die Zeit *in Stunden*, bis der erste Kunde kommt. Sei M dieselbe Zeit, aber in Minuten. Welche Verteilung hat M?

c) Der Stand ist nur in Betrieb, falls die Temperatur zumindest 20 Grad ist. An einem Tag, an dem dies eintrifft, bezeichen wir die Temperatur mit T. Wir nehmen an, dass T eine Zufallsvariable mit folgender Dichte ist.

$$f_T(t) = \begin{cases} \frac{c}{t} & 20 \leq t \leq 30 \\ 0 & \text{sonst} \end{cases}, \tag{1}$$

wobei c eine Konstante ist. Berechne c. (Falls die Antwort so weit wie möglich vereinfacht ist, muss sie nicht explizit ausgerechnet werden.)

d) Was ist die erwartete Temperatur? Was ist die Varianz? (Falls die Antworten so weit wie möglich vereinfacht sind, müssen sie nicht explizit ausgerechnet werden.)

e) In dieser Teilaufgabe und der nächsten, nehmen wir an, dass die Temperatur T die Dichte von *d*) hat, und, dass die Wartezeit Z_1 *gegeben die Temperatur* $t = T$, die Dichte in (a) hat. Was ist in diesem Fall die gemeinsame Dichte von T und Z_1?

f) Wenn T und Z_1 diese gemeinsame Dichte haben, was ist die Randdichte von Z_1? (Die Randdichte ist nicht gleich der Dichte von (a), weil wir hier eine zufällige Temperatur betrachten.)

g) Jetzt betrachten wir wieder den Fall, dass die Temperatur *nicht* zufällig ist, sondern gleich t ist für t konstant. Also hat Z_1 wieder die Dichte von (a). Wir bezeichnen die Zeit, nachdem der erste Kunde gegangen ist, bis der zweite Kunde kommt, mit Z_2. Wir nehmen an, dass auch Z_2 exponentialverteilt ist mit Paramter t. Ferner nehmen wir an, dass Z_1 und Z_2 unabhängig sind. Bezeichne mit X die Zeit, in Stunden nach 10 Uhr, bis der zweite Kunde ankommt. Was ist die Dichte von X?

Bitte wenden!

4. (9 Punkte) Ihre Webseite wird auf einem entfernten Server gehostet. Der Host behauptet, dass die durchschnittliche Antwortzeit des Servers gleich $400ms$ ist. Da das Signal von Ihrem Haus zum Server über k Router gesendet wird, modellieren wir die Antwortzeit X mit der Gamma(k, θ) Verteilung. Die entsprechende Dichte ist:

$$f(x) = \frac{x^{k-1}}{(k-1)!} \cdot \frac{e^{-x/\theta}}{\theta^k}$$

a) Sei $X_i \overset{iid}{=} X$.

i) Bestimmen Sie die Verteilung von $Y = \sum_{i=1}^{n} X_i$.

ii) Berechnen Sie $\mathbb{E}[X]$ und $\text{Var}(X)$

Hinweis: Falls $X = \sum_{j=1}^{k} E_j$, wobei $E_j \overset{iid}{\sim} \text{EXP}(1/\theta)$, $j = 1, \ldots, k$, dann gilt $X \sim \text{Gamma}(k, \theta)$.

b) Sie haben zehn Experimente durchgeführt, und die folgenden Ergebnisse (in ms) erhalten:

i	1	2	3	4	5	6	7	8	9	10
x_i	700	244	696	495	1164	864	113	427	510	1211

Kennzahl: $\bar{x} = 642.4$

Schätzen Sie θ mittels der Maximum-Likelihood-Methode wenn $k = 4$ gegeben ist. Geben sie dazu sowohl den Schätzer $\hat{\theta}(\underline{X})$ wie auch den Schätzwert $\hat{\theta}(\underline{x})$ an.

c) Warum ist es vernünftig anzunehmen, dass $\hat{\theta}(\underline{X})$ näherungsweise normalverteilt ist? Begründen Sie kurz.

d) Sie vermuten, dass die durchschnittliche Antwortzeit des Servers von Ihrem Computer eigentlich länger als $400ms$ ist.

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

ii) Berechnen Sie der Verwerfungsbereich für $\hat{\theta}$ zum Signifikanzniveau $\alpha = 5\%$. Wird die Null-Hypothese auf diesem Niveau abgelehnt?

iii) Berechnen Sie die ungefähre Macht des Tests unter der Annahme, dass die wahre durchschnittliche Antwortzeit $600ms$ beträgt.

Hinweis: $50/\sqrt{10} = 15.8$ und $2\sqrt{10}/25 = 1.01$. Benutzen Sie $\hat{\theta} \approx$ normalverteilt.