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

1. (jeweils: 1 Punkt für richtig, 0 Punkt für falsch)

a) 2.

b) 2.

c) 2.

d) 1.

e) 1.

f) 1.

g) 3.

h) 3.

i) 2.

j) 3.

2. **a)** Wir führen folgende Ereignisse ein: V - die Wetterstation hat Sturmböen vorhergesagt, B - es gibt Sturmböen. Aus der Aufgabestellung kommen folgende Angaben hervor: $P[B|V] = 0.7$, $P[B^C|V^C] = 0.9$. Daraus erhalten wir mithilfe des Satzes von Bayes

$$p = P[B] = P[V]P[B|V] + P[V^C]P[B|V^C] = q \cdot 0.7 + (1-q) \cdot (1 - 0.9) = 0.1 + 0.6q.$$

(1 Punkt)

b) Aus den Axiomen der Wahrscheinlichkeitsrechnung wissen wir dass $q \in [0, 1]$. **(1/2 Punkt)** Somit ist $p \in [0.1, 0.7]$. Deswegen ist $p = 0.6$ möglich aber $p = 0.85$ ist nicht möglich. Jan hat also sicher unrecht. **(1/2 Punkt)**

c) Wir möchten $P[V|B]$ berechnen. Aus der Aufgabestellung und aus **a)** kommen folgende Angaben hervor: $P[B|V] = 0.7$, $P[B] = p = 0.5$ und $P[V] = q = (p - 0.1)/0.6 = 2/3$. Mithilfe des Satzes von Bayes lässt sich dann $P[V|B]$ wie folgt berechnen:

$$P[V|B] = \frac{P[B|V]P[V]}{P[B]} = \frac{0.7 \cdot 2/3}{0.5} = \frac{14}{15} \approx 0.93. \qquad \textbf{(2 Punkte)}$$

d) X_1 ist $Ber(p)$-verteilt. **(1 Punkt)**

e) Da die Realisierungen unabhängig sind, können wir die Likelihoodfunktion wie folgt berechnen:

$$L(p) = \Pi_{i=1}^{50} P(X_i = x_i) = p^2(1-p)^{48}. \qquad \textbf{(1 Punkt)}$$

Bemerke, dass L weder in $p = 0$ noch in $p = 1$ ein Maximum einnimmt. **(kein Abzug wenn fehlt)** Wir setzten die Ableitung der log-Likelihoodfunktion gleich 0:

$$0 = \frac{d}{dp} l(\lambda) = \frac{d}{dp}(48 \log(1-p) + 2 \log p) = -48 \frac{1}{1-p} + 2\frac{1}{p}.$$

Daraus ergibt sich nach multiplizieren mit $p(1-p)$ eine lineare Gleichung, deren Lösung $p = 1/25$ ist. Da die zweite Ableitung

$$\frac{d^2}{dp^2} l(\lambda) = \frac{d}{dp}\left(-48\frac{1}{1-p} + 2\frac{1}{p}\right) = -\frac{48}{(1-p)^2} - \frac{2}{p^2}$$

insbesondere an der Stelle $p = 1/25$ negativ ist **(kein Abzug wenn fehlt)**, schliessen wir daraus, dass $p = 1/25$ ein Maximum ist. Das gesuchte MLE ist also $\hat{p} = 1/25$. **(1 Punkt)**

f) Die inverse Wahrscheinlichkeit lässt sich dank der Unabhängigkeit via Multiplikation der einzelnen Wahrscheinlichkeiten berechnen: $(1-p)^{50} = 0.95^{50} \approx 0.077$. Somit ist die gefragte Wahrscheinlichkeit $1 - 0.077 = 0.923$. **(1 Punkt)**

3. **a)** Zur Normierung der Dichte $f_{X,Y}$ berechnen wir

$$\int_1^\infty \int_1^\infty f_{X,Y}(x,y)dxdy = c\int_1^\infty \int_1^\infty \frac{1}{x^4y^4}dxdy = c\int_1^\infty \frac{1}{x^4}dx \int_1^\infty \frac{1}{y^4}dy = \frac{c}{9}. \qquad \textbf{(1 Punkt)}$$

Somit erhalten wir $c = 9$.

b) Die Randdichte f_X lässt sich durch Integrieren nach y berechnen. Für $x \geq 1$ haben wir

$$f_X(x) = \int_{-\infty}^\infty f_{X,Y}(x,y)dy = c\frac{1}{x^4}\int_1^\infty \frac{1}{y^4}dy = 9\frac{1}{x^4} \cdot \frac{1}{3} = \frac{3}{x^4}, \qquad \textbf{(1 Punkt)}$$

für $x < 1$ gilt $f_X(x) = 0$. Nun berechnen wir die gefragte Wahrscheinlichkeit

$$P[X > 3] = \int_3^\infty f_X(x)dx = \int_3^\infty \frac{3}{x^4}dx = \frac{1}{3^3} = \frac{1}{27} = 0.037. \qquad \textbf{(1 Punkt)}$$

c) Wegen der Symmetrie von x und y in der gemeinsamen Dichte haben wir die Randdichte von Y sofort: $f_Y(y) = 3/y^4$ für $y \geq 1$ **(1/2 Punkt)** und $f_Y(y) = 0$ sonst. Daraus folgt $f_{X,Y} = f_X \cdot f_Y$. Das beweist die Unabhängigkeit von X und Y. **(1/2 Punkt)**

d) Erwartungswert:

$$E[X] = \int_{-\infty}^\infty xf_X(x)dx = 3\int_1^\infty \frac{1}{x^3}dx = \frac{3}{2}. \qquad \textbf{(1 Punkt)}$$

Zur Berechnung der Varianz berechnen wir zuerst

$$E[X^2] = \int_{-\infty}^\infty x^2 f_X(x)dx = 3\int_1^\infty \frac{1}{x^2}dx = 3.$$

Daraus erhalten wir

$$Var(X) = E[X^2] - E[X]^2 = 3 - \frac{9}{4} = \frac{3}{4}. \qquad \textbf{(1 Punkt)}$$

e) Wir nehmen an, dass die Sturmschäden in unterschiedlichen Jahren unabhängig sind. **(1/2 Punkt)** Eine zweite Annahme ist, dass 30 schon gross genug ist, damit die Normalverteilung via CLT schon eine gute Approximation liefert. **(1/2 Punkt)**

Wir wenden also den CLT an. Die Sturmschäden aus den jeweiligen Jahren bezeichnen wir mit X_1, X_2, ... , X_{30}. Wir definieren $\bar{X} := \sum X_i/30$. Laut dem CLT und d) ist $\bar{X} \approx \mathcal{N}(1.5, 0.75/30) = \mathcal{N}(1.5, 0.025)$. **(1 Punkt)** Das ergibt mithilfe der Normalverteilungstabelle

$$P[\bar{X} > 2] = P\left[\frac{\bar{X} - 1.5}{\sqrt{0.025}} > \frac{2 - 1.5}{\sqrt{0.025}}\right] = P\left[\frac{\bar{X} - 1.5}{\sqrt{0.025}} > 3.1622\right] \approx 1 - 0.9992 = 0.0008.$$

(1 Punkt) Die Wahrscheinlichkeit, dass der durchschnittliche jährliche Sturmschaden 2 Millionen CHF übersteigen ist also ungefähr 0.08%.

f) Wir berechnen die Verteilungsfunktion F, die zu der Dichte f_X gehört. Für $z < 1$, $F(z) = 0$, für $z \geq 1$

$$F(z) = \int_{-\infty}^{z} f(x)dx = \int_{1}^{z} \frac{3}{x^4}dx = 1 - \frac{1}{z^3}. \qquad \textbf{(1/2 Punkt)}$$

Wir möchten bestimmen, für welche z ist $F(z) = 0.31$ und $F(z) = 0.94$. Dazu berechnen wir die inverse Funktion von F auf $[1, \infty)$: $F^{-1}(v) = 1/(1-v)^{1/3}$. **(1/2 Punkt)** Daraus ergibt sich

$$F^{-1}(0.31) = 1/(1-0.31)^{1/3} \approx 1.132 \qquad \textbf{(1/2 Punkt)}$$
$$F^{-1}(0.94) = 1/(1-0.94)^{1/3} \approx 2.554. \qquad \textbf{(1/2 Punkt)}$$

4. Messdaten:

Messung Nr.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
Messwerte (x_i)	10.7	6.1	8.1	8.9	10.9	13.6	10.4	12.1	12.3	8.6

a) Bonaprate AG kann nur den Vorzeichen-Test durchführen, **(1/2 Punkt)** da alle andere Tests Symmetrie voraussetzen. **(1/2 Punkt)**

b) Sei μ_{med} der Median der Zufallsvariable X_1. **(1/2 Punkt)**

H_0: $\mu_{med} = 10$.

H_A: $\mu_{med} > 10$. **(1/2 Punkt)**

c) Unsere Teststatistik, $T = \sum 1_{[10,\infty)}(X_i)$ ist unter der Nullhypothese $Bin(10, 0.5)$-verteilt. **(1/2 Punkt)** Die Realisierung der Teststatistik ist $t = \sum 1_{[10,\infty)}(x_i) = 6$. **(1/2 Punkt)** Unter der Nullhypothese haben wir folgende Wahrscheinlichkeiten:

$$P_{H_0}(T = 10) = \binom{10}{0}(1/2)^{10} = 0.001$$
$$P_{H_0}(T = 9) = \binom{10}{1}(1/2)^{10} = 0.010$$
$$P_{H_0}(T = 8) = \binom{10}{2}(1/2)^{10} = 0.044$$

Da das Signifikanzniveau 5% Prozent beträgt, erhalten wir daher den Verwerfungsbereich $\mathcal{K} = \{9, 10\}$. Da $T \notin \mathcal{K}$, wird die Nullhypothese beibehalten. **(1 Punkt)**

d) Unter der Annahme $P[X_1 > 10] = 0.75$ hat T die Verteilung $Bin(10, 0.75)$. Der Verwerfungsbereich für das 5%-Niveau ist $\mathcal{K} = \{9, 10\}$. Die Wahrscheinlichkeit unter der Alternativhypothese $P[X_1 > 0] = 0.75$, dass T in den Verwerfungsbereich fällt lässt sich folgendermassen berechnen:

$$P_{H_A}[T \in \mathcal{K}] = \binom{10}{0}0.75^{10} + \binom{10}{1}0.75^{9} \cdot 0.25^{1} \approx 0.244.$$

Die Macht ist also 0.244. **(2 Punkte)**

e) Der QQ-Plot zeugt von einer sehr guten Vereinbarkeit der Daten mit der Normalverteilung. Unter Annahme der Normalverteilung soll man den t-Test wählen, **(1/2 Punkt)** da er eine bessere Macht hat (unter dieser Annahme) als der Wilcoxon-Test oder der Vorzeichen-Test. Der z-Test ist keine Option, denn die Varianz ist nicht bekannt. **(1/2 Punkt)**

f) Sei $\mu := E[X_1]$ der Erwartungswert von X_1. **(1/2 Punkt)**

H_0: $\mu = \mu_0 := 10$.

H_A: $\mu > 10$. **(1/2 Punkt)**

g) Die Teststatistik, $T = \frac{\bar{X}_n - \mu_0}{S_X/\sqrt{n}}$ ist unter der Nullhypothese t_9-verteilt. **(1/2 Punkt)** Die Realisierung ist $t = \frac{\bar{x} - \mu_0}{s_X/\sqrt{n}} = \frac{10.17 - 10}{\sqrt{2.256}/\sqrt{10}} \approx 0.358$. **(1/2 Punkt)** Aus der Tabelle können wir den p-Wert zwar nicht genau ablesen, aber es ist ersichtlich, dass er zwischen 0.3 und 0.4 liegt. Der p-Wert ist demzufolge grösser als das Signifikanzniveau 0.05, deshalb wird die Nullhypothese beibehalten. **(1 Punkt)**