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

1. (10 Punkte) Wir definieren die Ereignisse

$$\begin{aligned} S &= \{\text{ Anruf aus der Schweiz }\}; \\ A &= \{\text{ Anruf aus dem Ausland }\} = S^c, \end{aligned}$$

mit $P[S] = 0.7$ und $P[A] = 0.3$. Wir definieren die Ereignisse

$$\begin{aligned} G &= \{\text{ geschäftlicher Anruf }\}; \\ Pr &= \{\text{ privater Anruf }\}; \\ W &= \{\text{ Werbeanruf }\}. \end{aligned}$$

Laut Aufgabenstellung gilt

$$P[G|S] = 0.3; \quad P[Pr|S] = 0.5; \quad P[W|S] = 0.2.$$

und

$$P[G|A] = 0.6; \quad P[Pr|A] = 0.3; \quad P[W|A] = 0.1.$$

a) Der Satz der totalen Wahrscheinlichkeit liefert

$$P[Pr] = P[Pr|S]P[S] + P[Pr|A]P[A] = \frac{5}{10}\frac{7}{10} + \frac{3}{10}\frac{3}{10} = \frac{44}{100} = 0.44.$$

b) Der Satz von Bayes liefert

$$\begin{aligned} P[A|G] &= \frac{P[G|A]P[A]}{P[G]} = \frac{P[G|A]P[A]}{P[G|S]P[S] + P[G|A]P[A]} \\ &= \frac{\frac{6}{10}\frac{3}{10}}{\frac{3}{10}\frac{7}{10} + \frac{6}{10}\frac{3}{10}} = \frac{6}{7+6} = \frac{6}{13} = 0.46. \end{aligned}$$

c) Wegen der Unabhängigkeit gilt

$$\begin{aligned} E[X_1 I_1] &= E[X_1]E[I_1] = 2P[I_1 = 1] = 2p; \\ E[(X_1 I_1)^2] &= E[X_1^2]E[I_1^2] = 8P[I_1 = 1] = 8p. \end{aligned}$$

Somit haben wir $E[X_1 I_1] = 2p$ und $\mathrm{Var}(X_1 I_1) = 8p - 4p^2 = 4p(2-p)$.

Bitte wenden!

d) Sei $X_i \sim \text{Exp}(0.5)$ die Dauer des i-ten Anrufes aus der Schweiz, $i = 1, \ldots, 25$. Wir führen unabhängige Zufallsvariablen $I_i \sim \text{Ber}(0.2)$, $i = 1, \ldots, 25$, ein, die auch unabhängig von den X_i sind. Sei $I_i = 1$, falls der i-te Anruf aus der Schweiz ein Werbeanruf ist und $I_i = 0$ sonst. $X_i I_i$ beschreibt also die Zeit, die Franz beim i-ten Anruf mit einem Werbeanruf verbringt. Gesucht ist

$$P\left[\sum_{i=1}^{25} X_i I_i \geq 20\right].$$

Wegen der Unabhängigkeit gilt wie in **c)** mit $p = 0.2$,

$$\begin{aligned} E[X_i I_i] &= E[X_i]E[I_i] = 2P[I_i = 1] = 2p = 0.4; \\ E[(X_i I_i)^2] &= E[X_i^2]E[I_i^2] = 8P[I_i = 1] = 8p = 1.6. \end{aligned}$$

Somit haben wir $E[X_i I_i] = 0.4$ und $\text{Var}(X_i I_i) = 1.6 - (0.4)^2 = 1.6 - 0.16 = 1.44$. Nach dem Zentralen Grenzwertsatz gilt also

$$\begin{aligned} P\left[\sum_{i=1}^{25} X_i I_i \geq 20\right] &= P\left[\frac{\sum_{i=1}^{25} X_i I_i - 25E[X_1 I_1]}{\sqrt{25\text{Var}(X_1 I_1)}} \geq \frac{20 - 25E[X_1 I_1]}{\sqrt{25\text{Var}(X_1 I_1)}}\right] \\ &\approx 1 - \Phi\left(\frac{20 - 25E[X_1 I_1]}{5\sqrt{\text{Var}(X_1 I_1)}}\right) \\ &= 1 - \Phi\left(\frac{2}{\sqrt{1.44}}\right) \\ &= 1 - \Phi\left(\frac{2}{1.2}\right) = 1 - \Phi(1.667) = 1 - 0.95 = 0.05. \end{aligned}$$

2. (10 Punkte)

a) Für den Erwartungswert von Z gilt

$$\begin{aligned} E[Z] &= \int_0^\infty z \frac{2}{(1+z)^3} dz \\ &= \int_0^\infty (1+z) \frac{2}{(1+z)^3} dz - \int_0^\infty \frac{2}{(1+z)^3} dz \\ &= \int_0^\infty \frac{2}{(1+z)^2} dz - 1 \\ &= 2 - 1 = 1. \end{aligned}$$

b) Es gilt $\log(1+Z) > 0$, also $F(t) = 0$ für $t \leq 0$, wobei F die Verteilungsfunktion von $\log(1+Z)$ ist. Für $t > 0$ gilt

$$\begin{aligned} F(t) &= P[\log(1+Z) \leq t] = P\left[Z \leq e^t - 1\right] = \int_0^{e^t - 1} \frac{2}{(1+z)^3} dz \\ &= \left.\frac{-1}{(1+z)^2}\right|_0^{e^t - 1} \\ &= 1 - \frac{1}{e^{2t}} = 1 - e^{-2t}. \end{aligned}$$

Siehe nächstes Blatt!

$\log(1+Z)$ hat also eine Exponentialverteilung mit Parameter 2.

c) Für die Randdichte von X gilt

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \frac{1}{4}\int_{-1}^{1}(1+xy^2)dy = \frac{1}{4}\left(2+x\frac{2}{3}\right) \\ &= \frac{1}{2}+\frac{x}{6}, \end{aligned}$$

für $|x| \le 1$. Für die Randdichte von Y gilt

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \frac{1}{4}\int_{-1}^{1}(1+xy^2)dx = \frac{2}{4} = \frac{1}{2},$$

für $|y| \le 1$.

d) Es gilt

$$E[X] = \int_{-\infty}^{\infty} xf_X(x)dx = \int_{-1}^{1}\frac{x}{2}+\frac{x^2}{6}dx = \int_{-1}^{1}\frac{x^2}{6}dx = \frac{1}{9},$$

und

$$E[Y] = \int_{-\infty}^{\infty} yf_Y(y)dy = \int_{-1}^{1}\frac{y}{2}dy = 0.$$

Weiter gilt

$$E[XY] = \int_{-\infty}^{\infty} xyf_{X,Y}(x,y)dxdy = \frac{1}{4}\int_{-1}^{1}\left(\int_{-1}^{1} xy+x^2y^3dy\right)dx = 0.$$

Für die Kovarianz von X und Y gilt daher

$$\mathrm{Cov}(X,Y) = E[XY]-E[X]E[Y] = 0.$$

X und Y sind aber nicht unabhängig, da z.B. $f_X(1)f_Y(1) = \frac{2}{3}\frac{1}{2} \neq \frac{2}{4} = f_{X,Y}(1,1)$.

3. (10 Punkte)

a) Seien $x_1,\ldots,x_{10} > 0$ unabhängige Beobachtungen von X. Die Likelihood-Funktion von θ ist dann

$$L(\theta;x_1,\ldots,x_{10}) = \prod_{i=1}^{10} x_i^2\theta e^{-\theta x_i^3/3}.$$

Die Log-Likelihood-Funktion ist also

$$l(\theta;x_1,\ldots,x_{10}) = \sum_{i=1}^{10}\left(2\log(x_i)+\log(\theta)-\theta x_i^3/3\right).$$

Bitte wenden!

Die Nullstelle der Ableitung ist gegeben durch

$$0 \quad = \quad \frac{10}{\theta} - \frac{1}{3}\sum_{i=1}^{10} x_i^3.$$

Der Maximum-Likelihood Schätzer von θ basierend auf $x_1, \ldots, x_{10} > 0$ ist also

$$\hat{\theta}_{MLE} \quad = \quad \left(\frac{1}{30}\sum_{i=1}^{10} x_i^3\right)^{-1}.$$

b) Die Apriori-Dichte von Θ ist

$$f(\theta) \quad = \quad e^{-\theta}, \qquad \text{für } \theta > 0.$$

Die Posteriori-Dichte von Θ gegeben die Beobachtung $x = 3$ ist

$$\begin{aligned}
f(\theta|1) \quad &= \quad \frac{f_\theta(x)f(\theta)}{\int_{-\infty}^{\infty} f_\theta(3)f(\theta)d\theta} \quad = \quad \frac{9\theta e^{-9\theta}e^{-\theta}}{\int_{-\infty}^{\infty} f_\theta(3)f(\theta)d\theta} \\
&= \quad \frac{9\theta e^{-10\theta}}{\int_0^\infty 9\theta e^{-10\theta}d\theta} \\
&= \quad \frac{9\theta e^{-10\theta}}{9/100} \quad = \quad 100\theta e^{-10\theta},
\end{aligned}$$

für $x > 0$. Die Posteriori-Verteilung ist also die Gammaverteilung mit Parametern $\gamma = 2$ und $c = 10$.

c) Sei $X \sim \mathcal{N}(\mu, \sigma^2)$ der Eisengehalt pro Flasche mit $\sigma = 0.25$ und $\mu \in \mathbb{R}$ unbekannt. Da die Daten normalverteilt sind und σ bekannt ist, führen wir einen z-Test zum Niveau $\alpha = 2.5\%$ durch. Die Hypothesen sind

$$\begin{aligned}
H_0 : \mu \quad &= \quad \mu_0 = 1 \\
H_A : \mu \quad &> \quad \mu_0.
\end{aligned}$$

Mit $q_{1-\alpha} = \Phi^{-1}(1-\alpha) = \Phi^{-1}(0.975) = 1.96$ ist der Verwerfungsbereich gegeben durch

$$K \quad = \quad [1.96, \infty).$$

Die Realisierung z der Teststatistik Z errechnet sich zu

$$z = \frac{\overline{x}_{25} - \mu_0}{\sigma/\sqrt{25}} = \frac{1.1 - 1}{\sqrt{0.25}/5} = \frac{0.1}{(1/2)/5} = \frac{0.1}{0.1} = 1 \notin K,$$

weshalb die Nullhypothese *nicht* verworfen wird.