Wahrscheinlichkeit und Statistik - Musterlösung (BSc D-INFK)

1. (7 Punkte) **a)** $\frac{3}{4}$ **b)** $1/\log(2)$ **c)** $\frac{1}{4}$ **d)** 2 **e)** F^n **f)** $\mathcal{N}\Big(\mu, \frac{\sigma^2}{n}\Big)$ **g)** $1 + \frac{n}{\sum_{i=1}^n \log(x_i)}$

2. Wir definieren $W_i :=$ "Christoph muss an der i-ten Ampel warten" für $i = 1, 2$. Aus der Aufgabenstellung erhält man $P[W_1] = \frac{1}{4}$, $P[W_2^c \mid W_1^c] = \frac{2}{3}$, $P[W_2 \mid W_1] = \frac{1}{7}$.

a) (2 Punkte) Die gesuchte Wahrscheinlichkeit ist $P[(W_1 \cap W_2^c) \cup (W_1^c \cap W_2)]$. Wir berechnen

$$\begin{aligned} P[W_1 \cap W_2^c] &= P[W_2^c \mid W_1]P[W_1] = (1 - P[W_2 \mid W_1])P[W_1] \\ &= (1 - \frac{1}{7}) \cdot \frac{1}{4} = \frac{6}{7} \cdot \frac{1}{4} = \frac{6}{28} \end{aligned}$$

und

$$\begin{aligned} P[W_1^c \cap W_2] &= P[W_2 \mid W_1^c]P[W_1^c] = (1 - P[W_2 \mid W_1^c])(1 - P[W_1]) \\ &= (1 - \frac{2}{3}) \cdot (1 - \frac{1}{4}) = \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{4}. \end{aligned}$$

Da $W_1 \cap W_2^c$ und $W_1^c \cap W_2$ disjunkt sind haben wir also

$$P[(W_1 \cap W_2^c) \cup (W_1^c \cap W_2)] = P[W_1 \cap W_2^c] + P[W_1^c \cap W_2] = \frac{6}{28} + \frac{1}{4} = \frac{13}{28}.$$

b) (1 Punkt) $S := W_1 \cup W_2$ beschreibt das Ereignis, dass Christoph zu spät zur Vorlesung kommt.

$$\begin{aligned} P[S] &= 1 - P[S^c] = 1 - P[W_1^c \cap W_2^c] \\ &= 1 - P[W_2^c \mid W_1^c]P[W_1^c] = 1 - \frac{2}{3} \cdot (1 - \frac{1}{4}) = \frac{1}{2}. \end{aligned}$$

c) (2 Punkte) Mit dem Satz von Bayes erhält man

$$\begin{aligned} P[W_1^c \mid W_2^c] &= \frac{P[W_2^c \mid W_1^c]P[W_1^c]}{P[W_2^c \mid W_1^c]P[W_1^c] + P[W_2^c \mid W_1]P[W_1]} \\ &= \frac{\frac{2}{3} \cdot (1 - \frac{1}{4})}{\frac{2}{3} \cdot (1 - \frac{1}{4}) + (1 - \frac{1}{7}) \cdot \frac{1}{4}} = \frac{\frac{1}{2}}{\frac{1}{2} + \frac{6}{28}} = \frac{\frac{1}{2}}{\frac{20}{28}} = \frac{7}{10}. \end{aligned}$$

Bitte wenden!

d) (2 Punkte) Mit der Definition der bedingten Wahrscheinlichkeit folgt

$$P[W_1 \mid S] = \frac{P[W_1 \cap S]}{P[S]} = \frac{P[W_1 \cap (W_1 \cup W_2)]}{P[S]} = \frac{P[W_1]}{P[S]}$$
$$= \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}.$$

Da $P[W_1 \mid S] = \frac{1}{2} \neq \frac{1}{4} = P[W_1]$, sind die Ereignisse W_1 und S abhängig.

e) (2 Punkte) Sei E_i, $i = 1, \ldots, 10$, das Ereignis, dass Christoph am i-ten Tag zu spät zur Vorlesung kommt. Wir wissen $P[E_i] = P[S] = \frac{1}{2}$. Da die E_i unabhängig sind, ist die Anzahl der Verspätungen X binomialverteilt mit Parametern $n = 10$ und $p = \frac{1}{2}$.

$$P[X \geq 2] = 1 - P[X = 1] - P[X = 0] = 1 - \binom{10}{1}\left(\frac{1}{2}\right)^{10} - \left(\frac{1}{2}\right)^{10}$$
$$= 1 - \frac{11}{2^{10}} \left(= \frac{2^{10} - 11}{2^{10}}\right).$$

3. a) (1 Punkt)

$$f_{Z_1}(z) = te^{-tz}$$
$$P[Z_1 \leq 1 - z] = 1 - e^{-tz}.$$

b) (2 Punkte)

$$P[M \leq z] = P[60Z_1 \leq z] = P[Z_1 \leq \frac{z}{60}] = 1 - e^{-\frac{tz}{60}} \implies Exp(\frac{t}{60}).$$

c) (1 Punkt)

$$\int_{20}^{30} \frac{1}{t} dt = \ln t|_{20}^{30} = \ln \frac{30}{20} = \ln \frac{3}{2}$$
$$c = \frac{1}{\ln \frac{3}{2}}.$$

d) (2 Punkte)

$$E[T] = \frac{1}{\ln \frac{3}{2}} \int_{20}^{30} \frac{1}{t} t dt = \frac{10}{\ln \frac{3}{2}}.$$
$$E[T^2] = \frac{1}{\ln \frac{3}{2}} \int_{20}^{30} t dt = \frac{1}{\ln \frac{3}{2}} \frac{t^2}{2}|_{20}^{30} = \frac{30^2 - 20^2}{2 \ln \frac{3}{2}} = \frac{250}{\ln \frac{3}{2}}.$$
$$Var[T] = \frac{250}{\ln \frac{3}{2}} - \frac{100}{(\ln \frac{3}{2})^2}.$$

Siehe nächstes Blatt!

e) (2 Punkte)

$$f_{T,Z_1}(t,z) = \frac{1}{\ln\frac{3}{2}} 1_{[20,30]}(t) \cdot e^{-tz} 1_{[0,\infty)}(z).$$

f) (2 Punkte)

$$f_{Z_1}(z) = \frac{1}{\ln\frac{3}{2}} \int_{20}^{30} e^{-tz} dz = \frac{1}{\ln\frac{3}{2}} \frac{e^{-tz}}{z}\Big|_{20}^{30} = \frac{e^{-20z} - e^{-30z}}{z \ln\frac{3}{2}} \qquad \text{für } z \geq 0$$

g) (2 Punkte)

$$\begin{aligned} f_X(x) &= \\ (f_{Z_1} * f_{Z_1})(x) &= \\ t^2 \int_0^x e^{-t(x-u)} e^{-tu} du &= \\ t^2 \int_0^x e^{-tx} du &= \\ t^2 x e^{-tx} & \qquad \text{für } x \geq 0 \end{aligned}$$

4. a) (3 Punkte)

i) $Y = \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} \sum_{j=1}^{k} E_{ij} \sim \text{Gamma}(nk, \theta)$, wobei $E_{ij} \overset{\text{iid}}{\sim} \text{EXP}(1/\theta)$.

ii) Falls $E_j \sim \text{EXP}(1/\theta)$, dann $\mathbb{E}[E_j] = \theta$ und $\text{Var}(E_j) = \theta^2$.
Also $\mathbb{E}[X] \overset{\text{lin.}}{=} \sum_{j=1}^{k} \mathbb{E}[E_j] = k\theta$ und $\text{Var}(X) \overset{\text{indep.}}{=} \sum_{j=1}^{k} \text{Var}(E_j) = k\theta^2$

b) (2 Punkte)

$$\begin{aligned} L(x_1, \ldots, x_n; \theta) &= \prod_{i=1}^{n} f(x_i; \theta) = \frac{(\prod x_i)^{k-1}}{((k-1)!)^n} \cdot \frac{e^{-\sum x_i/\theta}}{\theta^{nk}} \\ l(x_1, \ldots, x_n; \theta) &= (k-1) \sum_{i=1}^{n} \log x_i - n \sum_{j=1}^{k-1} \log j - \sum_{i=1}^{n} x_i/\theta - nk \log \theta \\ \frac{dl}{d\theta} &= \sum_{i=1}^{n} x_i/\theta^2 - nk/\theta \overset{!}{=} 0 \\ \hat{\theta} &= \sum_{i=1}^{n} x_i/nk = \bar{x}/k \end{aligned}$$

Wenn $k = 4$, dann $\hat{\theta}(\underline{X}) = \bar{X}/4$ und $\hat{\theta}(\underline{x}) = \bar{x}/4 = 160.6$.

c) (1 Punkt) Da $\hat{\theta} = Y/nk$ der Mittelwert von 40 iid exponentialverteilten Zufallsvariablen ist, können wir mit Hilfe des zentralen Grenzwertsatzes annehmen, dass $\hat{\theta}$ näherungsweise normalverteilt ist.

Bitte wenden!

d) (3 Punkte)
i)

$$H_0 : \theta k \leq 400, \text{ d.h. } \theta \leq 100.$$
$$H_A : \theta k > 400, \text{ d.h. } \theta > 100.$$

ii) ZGS $\Rightarrow \hat{\theta}(\underline{\mathrm{X}}) \approx \mathrm{N}(\theta, \frac{n \cdot k\theta^2}{n^2k^2}) = \mathrm{N}(\theta, \frac{\theta^2}{nk})$. Unter H_0, $\theta = 100$.
Aus der Tabelle der Standardnormalverteilung, $\Phi(0.95)^{-1} = 1.64$.
Der Verwerfungsbereich ist $K = (\mu + 1.64\sigma, \infty) = (100 + 1.64 \cdot 15.8, \infty) = (126, \infty)$.
Da $\hat{\theta}(\underline{\mathrm{x}}) = 160.6 \in K$, wird die Null-Hypothese abgelehnt.
iii) Falls $\theta = 600/4 = 150$,

$$\mathbb{P}(\hat{\theta} \in K) \approx \bar{\Phi}\left(\frac{126 - 150}{150/\sqrt{40}}\right) = \bar{\Phi}(-1.01) = \Phi(1.01) = 0.84$$

Also, wenn die wahre durchschnittliche Antwortzeit $600ms$ ist, wird die Null-Hypothese mit Wahrscheinlichkeit 84% (=Macht des Tests) abgelenht.