Wahrscheinlichkeit und Statistik
BSc D-INFK

1. **a)** 3. **b)** 2. **c)** 3. **d)** 3. **e)** 1. **f)** 2. **g)** 3. **h)** 2. **i)** 1. **j)** 3.

2. Seien die folgenden Ereignisse mit den unten angegebenen Symbolen notiert:

W={Webanwendung funktioniert.}
W^c={Webanwendung funktioniert nicht.}
X={Browser X wird verwendet.}
X^c={Nicht Browser X wird verwendet.}
Y={Browser Y wird verwendet.}
Y^c={Nicht Browser Y wird verwendet.}

Folgende Wahrscheinlichkeiten werden in der Aufgabenstellung gegeben:
$P[W|X] = \frac{9}{10}$, $P[X] = \frac{10}{11}$ und $P[W|X^c] = \frac{1}{3}$.

a) Gefragt wird nach den Wahrscheinlichkeit $P[W]$. Es gilt nach der Formel für die bedingte Wahrscheinlichkeit

$$P[W \cap X] = P[W|X] \cdot P[X] = \frac{9}{10} \cdot \frac{10}{11} = \frac{9}{11}$$

und nach der Formel für die totale Wahrscheinlichkeit

$$\begin{aligned} p = P[W] &= P[W \cap X^c] + P[W \cap X] = P[W|X^c]P[X^c] + P[W \cap X] \\ &= \frac{1}{3} \cdot \frac{1}{11} + \frac{9}{11} = \frac{1}{33} + \frac{9}{11} = \frac{28}{33}. \end{aligned}$$

b) Folgende Angaben werden noch in der Aufgabenstellung gemacht: $P[W|Y] = \frac{1}{4}$, $P[Y] = \frac{1}{22}$ und $P[W] = p$ ist ein Parameter.
Gefragt wird nach einer Formel für $P[Y|W]$ in Abhängigkeit von p. Es gilt

$$P[Y|W] = \frac{P[W \cap Y]}{P[W]} = \frac{P[W|Y]\ P[Y]}{P[W]} = \frac{1}{88p}.$$

Bitte wenden!

c) Es gilt $P[Y] = \frac{1}{22}$. Gefragt wird nach der Wahrscheinlichkeit $P[W|X^c \cap Y^c]$. Es gilt $\{Y\} \subset \{X^c\}$, daher

$$P[W|X^c \cap Y^c] = \frac{P[W \cap X^c \cap Y^c]}{P[X^c \cap Y^c]} = \frac{P[W \cap X^c] - P[W \cap Y]}{P[X^c] - P[Y]},$$

und diese berechnet sich als

$$\frac{P[W|X^c]\ P[X^c] - P[W|Y]\ P[Y]}{P[X^c] - P[Y]} = \frac{\frac{1}{33} - \frac{1}{88}}{\frac{1}{22}} = \frac{5}{12}.$$

Achtung: im Allgemeinen gilt *nicht* $P[W|X^c] = P[W|X^c \cap Y^c] + P[W|X^c \cap Y]$!

3. **a)** Die Randdichte ist definiert als $f_V(v) = \int_{-\infty}^{\infty} f(v,m)\,\mathrm{d}m$. Für $v < 0$ ist also $f_V(v) = 0$. Für $v \geq 0$ gilt

$$\begin{aligned} f_V(v) &= \int_v^\infty k v^k m^{-(1+k)} \lambda \exp(-\lambda v)\,\mathrm{d}m = k v^k \lambda \exp(-\lambda v) \left(-\frac{1}{k} m^{-k} \right) \Bigg|_{m=v}^{\infty} \\ &= \lambda \exp(-\lambda v). \end{aligned}$$

V ist also $\mathrm{Exp}(\lambda)$-verteilt.

b)

$$\begin{aligned} E[VM] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} v m f(v,m)\,\mathrm{d}m\mathrm{d}v = \int_0^\infty \int_v^\infty k v^{k+1} m^{-k} \lambda \exp(-\lambda v)\,\mathrm{d}m\mathrm{d}v \\ &= k\lambda \int_0^\infty v^{k+1} \exp(-\lambda v) \int_v^\infty m^{-k}\,\mathrm{d}m\mathrm{d}v \\ &= k\lambda \int_0^\infty v^{k+1} \exp(-\lambda v) \frac{1}{k-1} v^{1-k} \mathrm{d}v \\ &= \frac{k}{k-1} \int_0^\infty \lambda v^2 \exp(-\lambda v)\,\mathrm{d}v = \frac{2k}{(k-1)\lambda^2}. \end{aligned}$$

c)

$$\begin{aligned} P[M \geq 2V] &= \int_0^\infty \int_{2v}^\infty k v^k m^{-(1+k)} \lambda \exp(-\lambda v)\,\mathrm{d}m\mathrm{d}v \\ &= \int_0^\infty k v^k \lambda \exp(-\lambda v) \left(-\frac{1}{k} m^{-k} \right) \Bigg|_{m=2v}^{\infty} \mathrm{d}v \\ &= 2^{-k} \int_0^\infty \lambda \exp(-\lambda v) \mathrm{d}v = 2^{-k}. \end{aligned}$$

Siehe nächstes Blatt!

d) Die Verteilungsfunktion von W ist

$$F_W(t) = P[W \leq t] = P[\sqrt{V} \leq t] = P[V \leq t^2] = F_V(t^2)$$

für $t > 0$ und $F_W(t) = 0$ für $t \leq 0$. Es gilt $f_W(t) = F'_W(t)$. Also ist $f_W(t) = 0$ für $t \leq 0$ und für $t > 0$ gilt

$$f_W(t) = \frac{\mathrm{d}}{\mathrm{d}t} F_V(t^2) = F'_V(t^2) 2t = 2t f_V(t^2) = 2\lambda t \exp(-\lambda t^2).$$

(Für die alternativ angegebene Dichte f_V ist $f_W(t) = 4\lambda t \exp(-2\lambda t^2)$ für $t > 0$.)

e) Setze $\mu = E[V_i] = 5$, $\sigma = \sqrt{\mathrm{Var}[V_i]} = 5$ und $n = 100$. Es gilt $E[\overline{V}] = \mu$ und $\mathrm{Var}[\overline{V}] = \sigma^2/n$. Mit dem zentralen Grenzwertsatz approximieren wir

$$P[\overline{V} > 6] = P\left[\frac{\overline{V} - \mu}{\sigma/\sqrt{n}} > \frac{6-5}{5/\sqrt{100}}\right] \approx 1 - \Phi(2) \approx 1 - 0.97725 = 0.02275.$$

4. **a)** Da die Daten normalverteilt sind mit bekannter Varianz und unbekanntem Mittelwert, bietet sich ein z-Test an. Da wir den Zusatzstoff nur verwenden wollen, wenn sich die mittlere Trocknungszeit dadurch *verbessert*, ist der Test einseitig durchzuführen.

b) Sei n der Stichprobenumfang und für $i = 1, 2, \ldots, n$ sei X_i die Trocknungszeit der i-ten Probe. Wir nehmen an, dass $X_1, \ldots, X_n$ unter P_μ i.i.d. $\mathcal{N}(\mu, \sigma^2)$-verteilt sind mit $\sigma^2 = 10^2 = 100$ bekannt und $\mu \in \mathbb{R}$ unbekannt.

(i) Nach unseren Überlegungen aus **a)** lauten Nullhypothese und Alternative

$$H_0 : \mu = \mu_0 := 90 \quad \text{und} \quad H_A : \mu \leq \mu_0.$$

(ii) Die Teststatistik des z-Tests lautet $T = \frac{\overline{X}_n - \mu_0}{\sigma/\sqrt{n}} = \frac{\overline{X}_n - 90}{10}\sqrt{n}$. Unter H_0 ist $T \sim \mathcal{N}(0, 1)$.

(iii) Der Verwerfungsbereich hat die Form $K = (-\infty, c]$ für ein $c \in \mathbb{R}$. Damit der Test das Niveau $\alpha = 1\%$ einhält, muss man für c das α-Quantil z_α der Standardnormalverteilung wählen. Wegen Symmetrie der Standardnormalverteilung ist $z_\alpha = -z_{1-\alpha}$. Aus der Tabelle finden wir $z_{1-\alpha} = z_{0.99} \approx 2.33$. Somit lautet der gesuchte Verwerfungsbereich $K = (-\infty, -2.33]$.

c) Der beobachtete Wert der Teststatistik beträgt $T(\omega) = \frac{\overline{x}_{25} - 90}{10}\sqrt{25} = \frac{84-90}{2} = -3 \in K$. Der Testentscheid lautet also, die Nullhypothese auf dem 1%-Niveau zu verwerfen. Auf dem 1%-Niveau kann man sagen, dass der Zusatzstoff die mittlere Trocknungszeit reduziert.

d) Ein Fehler 2. Art für die Alternative $\mu_A = 85$ [Minuten] tritt auf, falls die Nullhypothese nicht verworfen wird, obwohl die mittlere Trocknungszeit in Wirklichkeit 85 Minuten ist. Die Wahrscheinlichkeit für so einen Fehler ist

$$
\begin{aligned}
P_{\mu_A}[T \notin K] &= P_{\mu_A}\left[\frac{\overline{X}_n - \mu_0}{\sigma/\sqrt{n}} > -2.33\right] \\
&= P_{\mu_A}\left[\frac{\overline{X}_n - \mu_A}{\sigma/\sqrt{n}} - \frac{\mu_0 - \mu_A}{\sigma/\sqrt{n}} > -2.33\right] \\
&= P_{\mu_A}\left[\frac{\overline{X}_n - \mu_A}{\sigma/\sqrt{n}} > -2.33 + \frac{5}{10}\sqrt{25}\right] = 1 - \Phi(0.17) \\
&\approx 1 - 0.5675 = 0.4325.
\end{aligned}
$$

e) Da die Varianz der Daten nun nicht mehr bekannt ist, führen wir einen t-Test durch mit der selben Nullhypothese und Alternative wie oben. Die Teststatistik lautet nun $T' = \frac{\overline{X}_n - \mu_0}{S/\sqrt{n}} = \frac{\overline{X}_{25} - 90}{S/5}$ und ist unter der Nullhypothese t-verteilt mit 24 Freiheitsgraden. Der Verwerfungsbereich ist analog wie oben $K' = (-\infty, c]$ wobei $c = t_{24,\alpha} = -t_{24,1-\alpha} = -t_{24,0.99} = -2.492$ nun das α-Quantil der t-Verteilung mit 24 Freiheitsgraden ist. Der Verwerfungsbereich ist also $K' = (-\infty, -2.492]$. Da der beobachtete Wert der Teststatistik $T'(\omega) = \frac{\overline{x}_{25} - 90}{s/5} = -\frac{30}{15} = -2$ nicht im Verwerfungsbereich liegt, lautet der Testentscheid nun, die Nullhypothese auf dem 1%-Niveau *nicht* zu verwerfen.