# Wahrscheinlichkeit und Statistik
# BSc D-INFK

**1. a)** (i) **b)** (iii) **c)** (i) **d)** (ii) **e)** (iii) **f)** (ii) **g)** (ii) **h)** (ii) **i)** (i) **j)** (iii)

**2.** Bezeichne $A$ (bzw. $A^c$) das Ereignis, dass der Bewerber Fähigkeit $A$ (nicht) besitzt und $E$ (bzw. $E^c$) das Ereignis, dass der Bewerber (nicht) eingestellt wird.

**a)** Wir betrachten die Ereignisse $A$ und $E$ von oben. Dann ist

$$P[A \cap E] = P[E|A]P[A] = 0.8 \cdot 0.1 = 0.08.$$

**b)** Es gilt

$$\begin{aligned} P[E^c] &= P[A \cap E^c] + P[A^c \cap E^c] \\ &= P[E^c|A]P[A] + P[E^c|A^c]P[A^c] \\ &= 0.2 \cdot 0.1 + \tfrac{48}{90} \cdot 0.9 \\ &= 0.5. \end{aligned}$$

Daraus erhalten wir mit der Bayes'schen Formel

$$\begin{aligned} P[A|E^c] = &\frac{P[A \cap E^c]}{P[E^c]} \\ = &\frac{0.02}{0.5} \quad = 0.04 \end{aligned}$$

für die gefragte Wahrscheinlichkeit.

**c)** Bezeichne $Z$ die Zufallsvariable, die die Anzahl der Kandidaten mit Fähigkeit $A$ unter den drei vorgeschlagenen Kandidaten angibt. Dann ist $Z$ $Bin(3, 0.1)$-verteilt, und daher ist die gefragte Wahrscheinlichkeit für $k \in \{0, 1, 2, 3\}$

$$P[Z = k] = \binom{3}{k} 0.1^k \cdot 0.9^{3-k}.$$

**Bitte wenden!**

**d)** Bezeichne $B$ das Ereignis, dass der einwandfreie Kandidat den Test besteht. Dann ist nach den Angaben $P[A|B] = 0.8$ und $P[B|A] = 1$. Daher gilt

$$0.8 = P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[B|A]P[A]}{P[B]} = \frac{0.1}{P[B]}.$$

Daraus erhalten wir

$$P[B] = \frac{1}{8}$$

für die gefragte Wahrscheinlichkeit.

**e)** Bezeichne nun $\widetilde{Z}$ die Zufallsvariable, die die Anzahl der Kandidaten mit Fähigkeit $A$ unter den drei vorgeschlagenen Kandidaten angibt. $\widetilde{Z}$ ist hypergeometrisch verteilt mit Parametern $n = 5, r = 4, m = 3$ und insbesondere kann $\widetilde{Z}$ nur die Werte 2 oder 3 annehmen. Deshalb ist

$$P[\widetilde{Z} = k] = \begin{cases} 0 & \text{für } k = 0, 1, \\ \frac{\binom{4}{k}\binom{1}{3-k}}{\binom{5}{3}} & \text{für } k = 2, 3, \end{cases}$$

und es ergibt sich

$$P[\widetilde{Z} = 3] = \frac{\binom{4}{3}}{\binom{5}{3}} = \frac{2}{5}$$

für die gefragte Wahrscheinlichkeit.

**3. a)** Es gilt

$$\begin{aligned} E[X_A] &= \int_1^3 \frac{x}{2}(3-x)dx = \frac{1}{2}\left[\frac{3x^2}{2} - \frac{x^3}{3}\right]_1^3 = \frac{1}{2}\left(\frac{27}{2} - \frac{27}{3} - \frac{3}{2} + \frac{1}{3}\right) \\ &= \frac{12}{2} - \frac{13}{3} = \frac{5}{3} \ (\widehat{=}\ \text{1h 40min}). \end{aligned}$$

**b)** Sei $n$ die gesuchte Lösung. Die Anzahl der Kerzen, die nach zwei Stunden noch brennen, folgt einer Binomialverteilung mit Parametern $n$ und $p = P[X_A \geq 2]$. Deren Erwartungswert ist $np$ und somit haben wir

$$n = \frac{5}{p},$$

wobei

$$p = P[X_A \geq 2] = \int_2^3 \frac{1}{2}(3-x)dx = \frac{1}{2}\left(3 - \frac{9}{2} + 2\right) = \frac{1}{4}.$$

**Siehe nächstes Blatt!**

Daher gilt $n = 5 \cdot 4 = 20$.

Alternativer Lösungsweg: Gesucht ist die Zahl $n$, so dass

$$E\left[\sum_{i=1}^{n} I_{\left\{X_A^{(i)} \geq 2\right\}}\right] = 5,$$

wobei die $X_A^{(i)}$, $i = 1, \ldots, n$ unabhängige Kopien von $X_A$ sind. Es gilt

$$E\left[\sum_{i=1}^{n} I_{\left\{X_A^{(i)} \geq 2\right\}}\right] = nP\left[X_A \geq 2\right] = n\int_2^3 \frac{1}{2}(3-x)dx = n\frac{1}{2}\left(3 - \frac{9}{2} + 2\right) = \frac{n}{4}.$$

Für die Lösung $n$ erhalten wir $n = 5 \cdot 4 = 20$.

**c)** Gesucht ist

$$P\left[X_A \geq 2.5 \middle| X_A \geq 1.5\right] = \frac{P\left[X_A \geq 2.5\right]}{P\left[X_A \geq 1.5\right]}.$$

Es gilt

$$P\left[X_A \geq 2.5\right] = \int_{2.5}^3 \frac{1}{2}(3-x)dx = \frac{1}{2}\left(\frac{3}{2} - \frac{9}{2} + \frac{25}{8}\right) = \frac{1}{16}.$$

Analog gilt

$$P\left[X_A \geq 1.5\right] = \int_{1.5}^3 \frac{1}{2}(3-x)dx = \frac{1}{2}\left(\frac{9}{2} - \frac{9}{2} + \frac{9}{8}\right) = \frac{9}{16}.$$

Wir erhalten

$$P\left[X_A \geq 2.5 \middle| X_A \geq 1.5\right] = \frac{P\left[X_A \geq 2.5\right]}{P\left[X_A \geq 1.5\right]} = \frac{1}{9}.$$

**d)** Wir berechnen die Wahrscheinlichkeit von $\{X_A \geq 2.5,\ X_B \geq 2.5\}$. Wegen der Unabhängigkeit gilt

$$P\left[X_A \geq 2.5,\ X_B \geq 2.5\right] = P\left[X_A \geq 2.5\right] P\left[X_B \geq 2.5\right].$$

Es gilt (wie oben)

$$P\left[X_A \geq 2.5\right] = \int_{2.5}^3 \frac{1}{2}(3-x)dx = \frac{1}{2}\left(\frac{3}{2} - \frac{9}{2} + \frac{25}{8}\right) = \frac{1}{16}.$$

Analog gilt

$$P\left[X_B \geq 2.5\right] = \frac{1}{4}\int_{2.5}^6 dx = \frac{7}{8}.$$

Es folgt

$$P\left[X_A \geq 2.5,\ X_B \geq 2.5\right] = P\left[X_A \geq 2.5\right] P\left[X_B \geq 2.5\right] = \frac{1}{16} \cdot \frac{7}{8} = \frac{7}{128}.$$

**Bitte wenden!**

**e)** Wir berechnen die Wahrscheinlichkeit von $\{X_B < X_A\}$. Wir brauchen die gemeinsame Dichte $f_{X_A,X_B}(a,b)$. Wegen der Unabhängigkeit gilt

$$f_{X_A,X_B}(a,b) = f_{X_A}(a) f_{X_B}(b) = \frac{1}{8}(3-a) I_{\{(a,b)\in[1,3]\times[2,6]\}}.$$

Um die Wahrscheinlichkeit von $\{X_B < X_A\}$ zu bestimmen, müssen wir die gemeinsame Dichte über den Bereich $\{(a,b) \in \mathbb{R}^2 : b < a\}$ integrieren. Da die gemeinsame Dichte für $b < 2$ und für $a > 3$ verschwindet, genügt es über den Bereich $\{(a,b) \in \mathbb{R}^2 : 2 \le b < a \le 3\}$ zu integrieren. Wir erhalten

$$\begin{aligned}
P\left[X_B < X_A\right] &= \frac{1}{8}\int_2^3 \int_2^a (3-a) db\, da \\
&= \frac{1}{8}\int_2^3 (3-a)(a-2) da \\
&= \frac{1}{8}\int_2^3 \left(-a^2 + 5a - 6\right) da \\
&= \frac{1}{8}\left[-\frac{a^3}{3} + \frac{5a^2}{2} - 6a\right]_2^3 \\
&= \frac{1}{8}\cdot\frac{1}{6} = \frac{1}{48}.
\end{aligned}$$

Alternativ,

$$\begin{aligned}
P\left[X_B < X_A\right] &= \frac{1}{8}\int_2^3 \int_b^3 (3-a) da\, db \\
&= \frac{1}{8}\int_2^3 \left(3(3-b) - \frac{9}{2} + \frac{b^2}{2}\right) db \\
&= \frac{1}{16}\int_2^3 \left(b^2 - 6b + 9\right) db \\
&= \frac{1}{16}\left[\frac{b^3}{3} - 3b^2 + 9b\right]_2^3 \\
&= \frac{1}{16}\cdot\frac{1}{3} = \frac{1}{48}.
\end{aligned}$$

**4.** (a) Laut der Aufgabestellung ist die Varianz bekannt. Somit entscheiden wir uns für den einseitigen $z$-Test.

(b) Sei $\mu_0 = 84$.

(i) Wir führen nun einen einseitigen Test durch:

$$H_0 : \mu = \mu_0, \quad H_A : \mu > \mu_0. \tag{1}$$

($H_0 : \mu \le \mu_0$ ist ebenfalls korrekt, aber dann ist die Begründung für die Wahl des Verwerfungsbereiches etwas schwieriger.)

**Siehe nächstes Blatt!**

(ii) Die Teststatistik ist somit

$$T = \frac{\overline{X}_9 - \mu_0}{\sigma/\sqrt{9}} = \overline{X}_9 - 84.$$

(iii) Der Verwerfungsbereich zum Niveau $\alpha$ ist also gegeben durch $K = [c(\alpha), \infty)$, wobei $c$ das $(1-\alpha)$-Quantil der Standardnormalverteilung ist. Für $\alpha = 10\%$ bekommen wir $c(\alpha) = 1.28$. Somit ist der Verwerfungsbereich $K = [1.28, \infty)$.

(iv) Der beobachtete Wert der Teststatistik lautet $T(\omega) = \overline{x}_9 - 84 = -1 \notin K$. Der Testentscheid lautet also, die Nullhypothese auf dem $10\%$-Niveau nicht zu verwerfen. Ronaldos Behauptung ist also aufgrund der Daten nicht gerechtfertigt.

(c) Per Definition gibt die Macht des Tests an, mit welcher Wahrscheinlichkeit ein Signifikanztest zugunsten einer konkreten Alternativhypothese $H_A$ entscheidet, falls diese richtig ist.

$$\begin{aligned} P_{\mu=85}\left[T \in K\right] &= P_{\mu=85}\left[T \geq 1.28\right] \\ &= P_{\mu=85}\left[\frac{\overline{X}_9 - 85}{\sigma/\sqrt{9}} - \frac{\mu_0 - 85}{\sigma/\sqrt{9}} \geq 1.28\right] \\ &= P_{\mu=85}\left[\frac{\overline{X}_9 - 85}{\sigma/\sqrt{9}} \geq 0.28\right] \\ &= 1 - \Phi(0.28) \approx 38.97\%. \end{aligned}$$

(d) Sei $Y$ die Zufallsvariable, welche die Wertung von L. Messi in einem Spiel beschreibt. Der Aufgabestellung entnehmen wir, dass $Y \sim \mathcal{N}(\mu_{LM}, \sigma^2)$. Das heisst,

$$p = P[Y \geq 95] = 1 - P[Y \leq 95] = 1 - \Phi\left(\frac{95 - \mu_{LM}}{\sigma}\right) = 1 - \Phi\left(\frac{95 - \mu_{LM}}{3}\right).$$

(e) Da die Zufallsvariablen $G_i$, $i = 1, \dots, 5$, i.i.d. geometrisch verteilt sind mit Erfolgswahrscheinlichkeit $p$, ist die Likelihood-Funktion $L$ gegeben durch

$$L(g_1, \dots, g_5; p) = \prod_{i=1}^{5} \left((1-p)^{g_i - 1} p\right).$$

Somit ist die log-Likelihood-Funktion

$$l(g_1, \dots, g_5; p) = \sum_{i=1}^{5} \left((g_i - 1)\log(1-p) + \log p\right).$$

Die partielle Ableitung der log-Likelihood-Funktion nach $p$ lautet

$$\frac{\partial}{\partial p} l(g_1, \dots, g_5; p) = \sum_{i=1}^{5} \left((g_i - 1)\frac{-1}{1-p} + \frac{1}{p}\right) = \frac{5}{p} - \frac{1}{1-p}\left(\sum_{i=1}^{5} g_i - 5\right)$$

**Bitte wenden!**

und diese ist gleich $0$ für $p = \frac{1}{\frac{1}{5}\sum_{i=1}^{5} g_i}$. Der Maximum-Likelihood-Schätzer für $p$ lautet daher $T = \frac{1}{\frac{1}{5}\sum_{i=1}^{5} G_i}$.