Wahrscheinlichkeit und Statistik - Lösung
BSc D-INFK

1. (10 Punkte)

a) (2 Punkte) Die Randdichte f_X ist gegeben durch

$$\begin{aligned} f_X(x) &= \int_{\mathbb{R}} f_{X,Y}(x,y)\, dy = \int_{\mathbb{R}} 4\frac{y}{x^3} 1_{\{0<x\leq 1\}} 1_{\{0<y\leq x^2\}}\, dy \\ &= \frac{4}{x^3} 1_{\{0<x\leq 1\}} \int_{\mathbb{R}} y 1_{\{0<y\leq x^2\}}\, dy = \frac{4}{x^3} 1_{\{0<x\leq 1\}} \int_0^{x^2} y\, dy \\ &= \frac{4}{x^3} 1_{\{0<x\leq 1\}} \left[\frac{y^2}{2}\right]_0^{x^2} = \frac{4}{x^3} 1_{\{0<x\leq 1\}} \frac{x^4}{2} \\ &= 2x 1_{\{0<x\leq 1\}}\,. \end{aligned}$$

b) (2 Punkte) Die Wahrscheinlichkeit, dass der Küchenchef weniger als a kg Tomaten für einen Salat verwendet, ist gegeben durch

$$P(X \leq a) = F(a) = \int_{-\infty}^{a} f_X(x)dx = \int_0^a 2x 1_{\{0<x\leq 1\}} dx = a^2 1_{\{0\leq a\leq 1\}} + 1_{\{a>1\}}\,.$$

Also ist die Wahrscheinlichkeit, dass der Küchenchef weniger als 0.5 kg Tomaten für einen Salat verwendet, gegeben durch $P(X \leq 0.5) = (0.5)^2 = 0.25 = 25\%$.

c) (2 Punkte) Der erwartete Quotient zwischen den Mengen an Tomaten und Feta für einen Salat ist gegeben durch

$$\begin{aligned} E\left(\frac{X}{Y}\right) &= \int_{\mathbb{R}^2} \frac{x}{y} f_{X,Y}(x,y) dxdy = \int_{\mathbb{R}^2} \frac{x}{y} 4\frac{y}{x^3} 1_{\{0<x\leq 1\}} 1_{\{0<y\leq x^2\}} dxdy \\ &= \int_{\mathbb{R}^2} \frac{4}{x^2} 1_{\{0<x\leq 1\}} 1_{\{0<y\leq x^2\}} dxdy \\ &= \int_0^1 \frac{4}{x^2} \left(\int_0^{x^2} dy \right) dx = \int_0^1 4dx = 4\,. \end{aligned}$$

Im Durchschnitt verwendet der Küchenchef viermal soviel Tomaten wie Feta.

d) (2 Punkt) Wir berechnen die bedingte Dichte von Y gegeben $X = x$ für $x > 0$. Das ist

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{4\frac{y}{x^3}1_{\{0<x\leq 1\}}1_{\{0<y\leq x^2\}}}{2x1_{\{0<x\leq 1\}}} = 2\frac{y}{x^4}1_{\{0<y\leq x^2\}}$$

für $x > 0$. Wir berechnen den bedingten Erwartungswert von Y gegeben $X = \alpha \leq 1$ mit $\alpha > 0$ als

$$E(Y|X=\alpha) = \int_{\mathbb{R}} y f_{Y|X=\alpha}(y)dy = \int_{\mathbb{R}} 2\frac{y^2}{\alpha^4}1_{\{0\leq y\leq \alpha^2\}}dy = \frac{2}{\alpha^4}\int_0^{\alpha^2} y^2 dy = \frac{2}{\alpha^4}\left[\frac{y^3}{3}\right]_0^{\alpha^2} = \frac{2}{\alpha^4}\frac{\alpha^6}{3} = \frac{2}{3}\alpha^2 .$$

Wenn der Küchenchef schon 300 g = 0.3 kg Tomaten für einen Salat verwendet hat, so ist die erwartete Menge von Feta $E(Y|X=0.3) = \frac{2}{3}\times 0.3^2$ kg $= 0.06$ kg $= 60$ g.

e) (2 Punkte) Die Dichte von Y ist

$$\begin{aligned}
f_Y(y) &= \int_{\mathbb{R}} f_{X,Y}(x,y)\, dx = \int_{\mathbb{R}} 4\frac{y}{x^3}1_{\{0<x\leq 1\}}1_{\{0<y\leq x^2\}}\, dx \\
&= 4y\int_{\mathbb{R}} \frac{1}{x^3}1_{\{0<x\leq 1\}}1_{\{0<y\leq x^2\}}\, dx = 4y1_{\{0\leq y\leq 1\}}\int_{\sqrt{y}}^{1} x^{-3}\, dx \\
&= 4y1_{\{0\leq y\leq 1\}}\left[-\frac{x^{-2}}{2}\right]_{\sqrt{y}}^{1} = 4y1_{\{0\leq y\leq 1\}}\left(\frac{1}{2y}-\frac{1}{2}\right) \\
&= 2y\left(\frac{1}{y}-1\right)1_{\{0\leq y\leq 1\}} .
\end{aligned}$$

Also gilt

$$f_{X,Y}(x,y) = 4\frac{y}{x^3}1_{\{0<x\leq 1\}}1_{\{0<y\leq x^2\}} \neq 2x1_{\{0<x\leq 1\}}\, 2y\left(\frac{1}{y}-1\right)1_{\{0\leq y\leq 1\}} = f_X(x)f_Y(y) ,$$

d.h. X und Y sind nicht unabhängig.

2. (10 Punkte)

a) (2 Punkte) Die momenterzeugende Funktion von X ist gegeben durch:

$$\begin{aligned}
M_X(t) &= E\left[e^{tX}\right] = \int_{\mathbb{R}} e^{tx}f_X(x)dx = \int_{\mathbb{R}} e^{tx}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x^2-2tx)}dx \\
&= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x^2-2tx+t^2-t^2)}dx = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x^2-2tx+t^2)}e^{\frac{1}{2}t^2}dx = e^{\frac{1}{2}t^2}\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-t)^2}dx \\
&= e^{\frac{1}{2}t^2}\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du = e^{\frac{1}{2}t^2}\underbrace{\int_{\mathbb{R}} f_X(u)du}_{=1} = e^{\frac{1}{2}t^2} .
\end{aligned}$$

b) (2 Punkte) Sei φ eine Funktion, die gegeben ist durch $\varphi(t) = \frac{n}{2}t^2 - tb$. Wir berechnen die Ableitung der Funktion φ als $\frac{\partial \varphi}{\partial t}(t) = nt - b$. Aus $\frac{\partial \varphi}{\partial t}(t_*) = 0$ folgt also $t_* = \frac{b}{n}$. Da die zweite Abteilung $\frac{\partial^2 \varphi}{\partial t^2}(t) = n$ positiv ist, schliessen wir, dass sich bei t_* ein lokales Minimum befindet. Das lokale Minimum entspricht dem globalen Minimum. Das gesuchte Infimum ist also

$$\varphi(t_*) = \frac{n}{2}t_*^2 - t_* b = \frac{n}{2}\left(\frac{b}{n}\right)^2 - \left(\frac{b}{n}\right) b = \frac{b^2}{2n} - \frac{b^2}{n} = -\frac{b^2}{2n}\,.$$

c) (1 Punkt) Der Erwartungswert und die Varianz von S_n sind gegeben durch

$$\begin{aligned} E\left[S_n\right] &= E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \underbrace{E\left[X_i\right]}_{=0} = n \times 0 = 0\,, \\ \operatorname{Var}\left[S_n\right] &= \operatorname{Var}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \underbrace{\operatorname{Var}\left[X_i\right]}_{=1} + 2\sum_{i<j} \underbrace{\operatorname{Cov}(X_i, X_j)}_{=0} = n \times 1 = n\,. \end{aligned}$$

d) (1 Punkt) Der Erwartungswert und die Varianz von S_n sind $E[S_n] = 0$ und $\operatorname{Var}[S_n] = n$. Die Varianz ist endlich. Nach der *Chebyshev-Ungleichung* bekommen wir, dass

$$P\left[|S_n - E[S_n]| \geq b\right] \leq \frac{\operatorname{Var}[S_n]}{b^2}\,,$$

für jedes $b > 0$. Daraus folgt

$$P\left[|S_n| \geq b\right] \leq \frac{n}{b^2}\,.$$

e) (2 Punkte) Die $X_1, \ldots, X_n$ sind i.i.d und die momenterzeugende Funktion $M_X(t)$ ist für alle $t \in \mathbb{R}$ endlich. Nach *Satz 5.6* bekommen wir, dass für jedes $b \in \mathbb{R}$

$$\begin{aligned} P\left[S_n \geq b\right] &\leq \exp\left(\inf_{t\in\mathbb{R}} \left(n \log M_X(t) - bt\right)\right) \\ &\leq \exp\left(\inf_{t\in\mathbb{R}} \left(\frac{n}{2}t^2 - bt\right)\right) \\ &= \exp\left(-\frac{b^2}{2n}\right)\,. \end{aligned}$$

f) (2 Punkte) Für jede Konstante $\varepsilon > 0$,

$$\begin{aligned} P\left[\left|\frac{S_n}{n}\right| > \varepsilon\right] &= P\left[|S_n| > \varepsilon n\right] \\ &= P\left[S_n > \varepsilon n\right] + P\left[S_n < -\varepsilon n\right] \\ &= 2P\left[S_n > \varepsilon n\right] \\ &\leq 2\exp\left(-\frac{(\varepsilon n)^2}{2n}\right) \\ &= 2\exp\left(-\frac{\varepsilon^2}{2}n\right)\,. \end{aligned}$$

Also ist

$$\lim_{n\to\infty} P\left[\left|\frac{S_n}{n}\right| > \varepsilon\right] \;\leq\; \lim_{n\to\infty}\left(2\exp\left(-\frac{\varepsilon^2}{2}n\right)\right) = 0\,.$$

Daraus folgt, dass $\lim\limits_{n\to\infty} P\left[\left|\frac{S_n}{n}\right| > \varepsilon\right] = 0$ für jede Konstante $\varepsilon > 0$, d.h. $\frac{S_n}{n}$ konvergiert für $n \to \infty$ in Wahrscheinlichkeit gegen 0.

3. (12 Punkte)

a) (1 Punkt) Damit $f_{\alpha,\beta}$ eine Dichte ist, muss gelten $\int_{\mathbb{R}} f_{\alpha,\beta}(x)dx = 1$. Also berechnen wir $\int_{\mathbb{R}} f_{\alpha,\beta}(x)dx = \int_{\mathbb{R}} \alpha x^{\beta} 1_{\{0<x\leq 1\}}dx = \alpha\int_0^1 x^{\beta}dx = \alpha\left[\frac{x^{\beta+1}}{\beta+1}\right]_0^1 = \frac{\alpha}{\beta+1}$. Somit erhalten wir $\beta = \alpha - 1$.

b) (1 Punkt) Das s-te Moment von X ist

$$E[X^s] = \int\limits_{\mathbb{R}} \alpha x^{\alpha+s-1} 1_{\{0<x\leq 1\}}dx = \alpha\left[\frac{x^{\alpha+s}}{\alpha+s}\right]_0^1 = \frac{\alpha}{\alpha+s}$$

für $s > -\alpha$.

c) (1 Punkt) Insbesondere ist $E[X] = \frac{\alpha}{\alpha+1}$ und $E(X^2) = \frac{\alpha}{\alpha+2}$. Daraus folgt

$$\begin{aligned}\mathrm{Var}(X) &= E[X^2] - E[X]^2 \\ &= \frac{\alpha}{\alpha+2} - \left(\frac{\alpha}{\alpha+1}\right)^2 \\ &= \frac{\alpha}{(\alpha+2)\,(\alpha+1)^2}\end{aligned}$$

für $\alpha > 0$.

d) (2 Punkte) Wir suchen α so, dass das s-te theoretische Moment von X, definiert als $\mu_s = E(X^s)$, und das s-te empirische Moment von $x_1, \ldots, x_n$, definiert als $m_s = \frac{1}{n}\sum_{i=1}^n x_i^s$, übereinstimmen. Also haben wir für $s = 1$

$$\mu_1(\alpha) = m_1 \quad\Leftrightarrow\quad E(X) = m_1 \Leftrightarrow \frac{\alpha}{\alpha+1} = m_1 \Leftrightarrow \alpha = \frac{m_1}{1-m_1}$$

Dies führt zum Momentenschätzer

$$\hat{\alpha}_{MoM} = \frac{\frac{1}{n}\sum_{i=1}^n X_i}{1-\frac{1}{n}\sum_{i=1}^n X_i} = \frac{\overline{X}_n}{1-\overline{X}_n}\,.$$

e) (3 Punkte) Da die Beobachtungen unabhängig sind, können wir die Likelihoodfunktion schreiben als

$$\begin{aligned}
L(x_1,\ldots,x_n,\alpha) &= \prod_{i=1}^{n} f_\alpha(x_i) = \prod_{i=1}^{n} \alpha x_i^{\alpha-1} 1_{\{0<x_i\le 1\}} \\
&= \prod_{i=1}^{n} \alpha x_i^{\alpha-1} \prod_{i=1}^{n} 1_{\{0<x_i\le 1\}} \\
&= \prod_{i=1}^{n} \alpha x_i^{\alpha-1} 1_{\{0<\min_{1\le i\le n} x_i\}} 1_{\{\max_{1\le i\le n} x_i \le 1\}} \\
&= \begin{cases} \prod_{i=1}^{n} \alpha x_i^{\alpha-1} & \text{wenn} \quad 0<\min_{1\le i\le n} x_i \quad \text{und} \quad \max_{1\le i\le n} x_i \le 1, \\ 0 & \text{sonst}\,. \end{cases}
\end{aligned}$$

Die log-Likelihoodfunktion ist

$$\begin{aligned}
l(x_1,\ldots,x_n,\alpha) &= \log L(x_1,\ldots,x_n,\alpha) \\
&= \log\left(\prod_{i=1}^{n} \alpha x_i^{\alpha-1} 1_{\{0<\min_{1\le i\le n} x_i\}} 1_{\{\max_{1\le i\le n} x_i \le 1\}} \right) \\
&= \sum_{i=1}^{n} \log\left(\alpha x^{\alpha-1}\right) + \log\left(1_{\{0<\min_{1\le i\le n} x_i\}}\right) + \log\left(1_{\{\max_{1\le i\le n} x_i \le 1\}}\right) \\
&= n\log\alpha + (\alpha-1)\sum_{i=1}^{n}\log x_i + \log\left(1_{\{0<\min_{1\le i\le n} x_i\}}\right) + \log\left(1_{\{\max_{1\le i\le n} x_i \le 1\}}\right).
\end{aligned}$$

Die Ableitung der log-Likelihoodfunktion ist $\dfrac{\partial l(x_1,\ldots,x_n,\alpha)}{\partial\alpha} = \dfrac{n}{\alpha} + \sum_{i=1}^{n}\log x_i$. Also haben wir, dass $\dfrac{\partial l(x_1,\ldots,x_n,\alpha^*)}{\partial\alpha} = 0 \Rightarrow \alpha^* = -\dfrac{1}{\frac{1}{n}\sum_{i=1}^{n}\log x_i}$. Da die zweite Abteilung $\dfrac{\partial l^2(x_1,\ldots,x_n,\alpha)}{\partial\alpha^2} = -\dfrac{n}{\alpha^2}$ negativ ist, schliessen wir daraus, dass bei α^* ein lokales Maximum ist. Das lokale Maximum entspricht dem globalen Maximum. Der gesuchte Maximum-Likelihood-Schätzer ist also

$$\hat{\alpha}_{MLE} = -\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\log X_i} = -\frac{1}{\overline{\log X}_n}\,.$$

f) (4 Punkte) Wir nehmen an, dass die 12 Windparks unabhängig voneinander sind. Eine zweite Annahme ist, dass 12 gross genug ist, damit die Normalverteilung unter Anwendung des zentralen Grenzwertsatzes eine gute Approximation liefert. Wir wenden also den zentralen Grenzwertsatz an. Laut dem ZGS ist $\overline{X}_{12}$ approximativ $\mathcal{N}(\mu_{X_1}, \frac{\sigma^2_{X_1}}{12})$-verteilt mit

$$\begin{aligned}
\mu_{X_1} &= E(X) = \frac{\alpha}{\alpha+1} = \frac{3}{3+1} = \frac{3}{4}\,, \\
\sigma^2_{X_1} &= \mathrm{Var}(X) = \frac{\alpha}{(\alpha+2)(\alpha+1)^2} = \frac{3}{(3+2)(3+1)^2} = \frac{3}{5\times 4^2}
\end{aligned}$$

Daraus folgt, dass

$$\begin{aligned}\mu_{\bar{X}_{12}} &= \mu_{X_1} = \frac{3}{4}\,,\\ \sigma^2_{\bar{X}_{12}} &= \frac{\sigma^2_{X_1}}{12} = \frac{1}{12}\frac{3}{5\times 4^2} = \frac{1}{5\times 4\times 4^2}\end{aligned}$$

und

$$\sigma_{\bar{X}_{12}} = \frac{1}{\sqrt{5\times 4\times 4^2}} = \frac{1}{8\sqrt{5}}\,.$$

Die Wahrscheinlichkeit, dass die durchschnittliche Lebensdauer der 12 Windparks die Höchstgrenze 0.5 überschreitet, ist also approximativ gegeben durch [1]

$$\begin{aligned}P\left(\overline{X}_{12} > \frac{1}{2}\right) &= 1 - P\left(\overline{X}_{12} \le \frac{1}{2}\right) = 1 - P\left(\frac{\overline{X}_{12} - \mu_{\bar{X}_{12}}}{\sigma_{\bar{X}_{12}}} \le \frac{\frac{1}{2} - \mu_{\bar{X}_{12}}}{\sigma_{\bar{X}_{12}}}\right)\\ &\approx 1 - P\left(Z \le \frac{\frac{1}{2} - \mu_{\bar{X}_{12}}}{\sigma_{\bar{X}_{12}}}\right) \qquad \text{für} \quad Z \sim \mathcal{N}(0,1)\,,\\ &= 1 - \Phi\left(\frac{\frac{1}{2} - \mu_{\bar{X}_{12}}}{\sigma_{\bar{X}_{12}}}\right) = 1 - \Phi\left(\frac{\frac{1}{2} - \frac{3}{4}}{\frac{1}{8\sqrt{5}}}\right)\\ &= 1 - \Phi\left(-\frac{8\sqrt{5}}{4}\right) = 1 - \Phi\left(-2\sqrt{5}\right) = \Phi\left(2\sqrt{5}\right) \approx \Phi\,(4.5) \approx 99\%\,.\end{aligned}$$

4. (8 Punkte)

a) (3 Punkte) Wir wollen wissen, ob mehr als 1% der Spargel verfault sind, d.h. ob μ grösser als 10 ist. Dazu führen wir einen *nach oben einseitigen* Test zum Niveau 10% mit Null- und Alternativhypothese $H_0 : \mu = \mu_0 = 10$ und $H_A : \mu > \mu_0$ durch. Der Verwerfungsbereich ist somit gegeben durch $K = [t_{n-1,1-\alpha}, \infty) = [t_{15,0.90}, \infty) = [1.341, \infty)$. Die Realisierung der Teststatistik ist

$$t = \sqrt{16}\frac{\overline{x}_{16} - \mu_0}{s_{16}} = \sqrt{16}\frac{11-10}{3.00} = \frac{4}{3} = 1 + \frac{1}{3} = 1.33 \notin K\,,$$

weshalb die Nullhypothese *nicht* verworfen wird.

b) (3 Punkte) Wenn $\sigma = 3$ bekannt ist, verwenden wir jetzt den z-Test. Der einzige Unterschied ist, dass die Quantile der t-Verteilung durch jene der Standardnormalverteilung und die geschätzte Streuung durch die Standardabweichung ersetzt werden. $K = [z_{1-\alpha}, \infty) = [z_{0.90}, \infty) = [1.29, \infty)$.

Die Teststatistik lautet nun $Z = \sqrt{16}\frac{\overline{X}_{16} - \mu_0}{\sigma}$, und ihre Realisierung ist gegeben durch

$$z = \sqrt{16}\frac{\overline{x}_{16} - \mu_0}{\sigma} = \sqrt{16}\frac{11-10}{3} = \frac{4}{3} = 1 + \frac{1}{3} = 1.33 \in K\,.$$

Also wird H_0 hier verworfen.

c) (2 Punkte) Wir suchen das Niveau α, das gerade den Verwerfungsbereich $[1.33, \infty)$ hat; das heisst, wir suchen das Niveau α, für das $z_{1-\alpha} = 1.33$ gilt. Durch Rückwärtsauslesen aus der Tabelle der z-Quantile sieht man lediglich, dass der Wert $z_{0.9082} = 1.33$.So $\alpha = 0.0918 = 9.18\%$. Dieses Niveau heisst P-Wert.

[1] $\sqrt{5} = \sqrt{4+1} = \sqrt{4\left(1+\frac{1}{4}\right)} = 2\left(1+\frac{1}{4}\right)^{1/2} \approx 2\left(1+\frac{1}{2}\frac{1}{4}\right) = 2 + \frac{1}{4} = 2.25$