Wahrscheinlichkeit & Statistik

Musterlösung 9

1. Die log-Likelihood-Funktion ist gegeben durch

$$\begin{aligned}\log L(x_1,\dots,x_n;\vartheta) &= \log\left(\frac{1}{\vartheta^n}\prod_{i=1}^n (1+x_i)^{-(1+\frac{1}{\vartheta})}\right)\\ &= -n\log\vartheta - (1+\frac{1}{\vartheta})\sum_{i=1}^n \log(1+x_i).\end{aligned}$$

Ableiten nach ϑ und Nullsetzen ergibt

$$\frac{\partial}{\partial\vartheta}\log L(x_1,\dots,x_n;\vartheta) = -\frac{n}{\vartheta} + \frac{1}{\vartheta^2}\sum_{i=1}^n \log(1+x_i) = 0$$

für

$$\vartheta = \frac{1}{n}\sum_{i=1}^n \log(1+x_i).$$

Der Maximum-Likelihood-Schätzer für ϑ ist also

$$T_{ML} = \frac{1}{n}\sum_{i=1}^n \log(1+X_i) = T^{(n)}.$$

2. **a)** Wir fixieren zuerst ein $c > 0$ und setzen

$$T^* := \frac{1}{\sigma^2 c}T^{(n)}(c) = \frac{1}{\sigma^2}\sum_{i=1}^n (X_i - \overline{X}_n)^2.$$

Mit dem Hinweis folgt, dass

$$E_\vartheta[T^*] = n-1 \quad \text{und} \quad \mathrm{Var}_\vartheta[T^*] = 2(n-1).$$

Also ist

$$\begin{aligned}f(c) &= \mathrm{Var}_\vartheta[T^{(n)}(c)-\sigma^2] + \big(E_\vartheta\big[T^{(n)}(c)-\sigma^2\big]\big)^2\\ &= \mathrm{Var}_\vartheta[\sigma^2 cT^*] + \big(E_\vartheta[\sigma^2 cT^* - \sigma^2]\big)^2\\ &= \sigma^4c^2 2(n-1) + \sigma^4\big(c(n-1)-1\big)^2\\ &= \sigma^4c^2\big((n-1)^2 + 2(n-1)\big) - \sigma^4 c2(n-1) + \sigma^4.\end{aligned}$$

Siehe nächste Seite!

Für beliebiges $c > 0$ ist also

$$f'(c) = 2c\left(2(n-1) + (n-1)^2\right)\sigma^4 - 2(n-1)\sigma^4$$

und

$$f''(c) = \left(4(n-1) + 2(n-1)^2\right)\sigma^4 > 0.$$

Insbesondere ist f strikt konvex und nimmt das globale Minimum bei $c^* = \frac{1}{n+1}$ an. Der gesuchte Schätzer ist also

$$T^{(n)}(c^*) = \frac{1}{n+1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2.$$

b) Nach der Markov-Ungleichung gilt für $\varepsilon > 0$

$$P_\vartheta[|T^{(n)}(c_n^*) - \sigma^2| \geq \varepsilon] \leq \frac{1}{\varepsilon^2}E_\vartheta\left[\left(T^{(n)}(c_n^*) - \sigma^2\right)^2\right] = \frac{f(c_n^*)}{\varepsilon^2}.$$

Für $n \to \infty$ gilt $c_n^* = \frac{1}{n+1} \to 0$, also $f(c_n^*) \to f(0) = 0$, und damit ist die Folge von Schätzern konsistent.

3. Die Likelihood-Funktion ist

$$\begin{aligned} L(x_1,\dots,x_n;\vartheta_1,\vartheta_2) = \prod_{i=1}^{n} f(x_i;\vartheta_1,\vartheta_2) &= \vartheta_2^n \exp\left(n\vartheta_1\vartheta_2 - \vartheta_2\sum_{i=1}^{n} x_i\right)\prod_{i=1}^{n} I_{[\vartheta_1,\infty)}(x_i) \\ &= \vartheta_2^n \exp\left(n\vartheta_1\vartheta_2 - \vartheta_2\sum_{i=1}^{n} x_i\right) I_{\{\min_{1\leq i\leq n} x_i \geq \vartheta_1\}} \end{aligned}$$

und sie ist genau dann positiv, wenn alle x_i grösser oder gleich ϑ_1 sind. Unter diesen Nebenbedingungen müssen wir $n(\log\vartheta_2 + \vartheta_1\vartheta_2) - \vartheta_2\sum_{i=1}^{n} x_i$ maximieren. Daraus folgt sofort, dass

$$T_1 = \min_{1\leq i\leq n} X_i \quad \text{und} \quad T_2 = \frac{n}{\sum_{i=1}^{n} X_i - nT_1} = \frac{1}{\overline{X}_n - T_1}$$

die Maximum-Likelihood-Schätzer von ϑ_1 und ϑ_2 sind. Für die gegebenen Daten erhalten wir die realisierten Schätzwerte

$$T_1(\omega) = t_1(x_1,\dots,x_{10}) = 3.1 \quad \text{und} \quad T_2(\omega) = t_2(x_1,\ \dots,x_{10}) \approx 0.7634.$$

4. a) Seien $X_1,\dots,X_n$ i.i.d. $\sim \mathcal{U}(a,b)$ unter P_ϑ. Mit $\vartheta = (a,b)$ ist die Likelihood-Funktion gegeben durch

$$L(x_1,\dots,x_n;\vartheta) = \frac{1}{(b-a)^n}\prod_{i=1}^{n} 1_{[a,b]}(x_i). \tag{1}$$

Nun muss $L(x_1,\dots,x_n;\vartheta)$ für feste $(x_1,\dots,x_n)$ bezüglich a und b maximiert werden. Seien $x_* := \min_{1\leq i\leq n} x_i$ und $x^* := \max_{1\leq i\leq n} x_i$. Falls x_* oder x^* ausserhalb

Siehe nächste Seite!

$[a,b]$ liegen, dann verschwindet die rechte Seite von (1). Es genügt also, den Maximierer $\vartheta = (a,b)$ in der Menge $D := \{(a,b) \in \mathbb{R}^2 : a \le x_*, x^* \le b\}$ zu suchen. Für $(a,b) \in D$ gilt jedoch

$$\begin{aligned} L(x_1,\dots,x_n;\vartheta) = \frac{1}{(b-a)^n} \prod_{i=1}^n 1_{[a,b]}(x_i) &= \frac{1}{(b-a)^n} \\ &\le \frac{1}{(x^*-x_*)^n} = L(x_1,\dots,x_n;(x_*,x^*)), \end{aligned}$$

d.h. $T_{ML} = (\min_{1\le i\le n} X_i, \max_{1\le i\le n} X_i)$ ist der ML-Schätzer für $\vartheta = (a,b)$.

Der realisierte Schätzwert ist

$$T_{ML}(\omega) = t_{ML}(x_1,\dots,x_{10}) = (9.2, 18.85).$$

Sarah vermutet also, dass Benno zwischen 9:12 und 18:51 Uhr lernt.

b) Es gilt $E_\vartheta[X_i] = \frac{a+b}{2} =: m_1(\vartheta)$ und $E_\vartheta[X_i^2] = \frac{1}{3}\frac{b^3-a^3}{b-a} = \frac{b^2+ab+a^2}{3} =: m_2(\vartheta)$. Aus

$$\frac{a+b}{2} = m_1(\vartheta) = \tilde{m}_1(x_1,\dots,x_n) := \frac{1}{n}\sum_{i=1}^n x_i =: \overline{x}_n$$

und

$$\frac{b^2+ab+a^2}{3} = m_2(\vartheta) = \tilde{m}_2(x_1,\dots,x_n) := \frac{1}{n}\sum_{i=1}^n x_i^2 =: \overline{(x^2)}_n$$

folgt dann, dass

$$2\overline{x}_n = a + b \quad \text{und} \quad 3\overline{(x^2)}_n = b^2 + ab + a^2.$$

Auflösen der ersten Gleichung nach $b = 2\overline{x}_n - a$ und quadrieren ergibt dann $b^2 = 4\overline{x}_n^2 - 4\overline{x}_n a + a^2$. Setzen wir b und b^2 in die zweite obige Gleichung ein, erhalten wir

$$0 = a^2 - 2a\overline{x}_n + 4\overline{x}_n^2 - 3\overline{(x^2)}_n.$$

Auflösen nach a liefert

$$a_{1,2} = \frac{2\overline{x}_n \pm \sqrt{4\overline{x}_n^2 - 16\overline{x}_n^2 + 12\overline{(x^2)}_n}}{2} = \overline{x}_n \pm \sqrt{3\left(\overline{(x^2)}_n - \overline{x}_n^2\right)}$$

und damit ist

$$b_{1,2} = 2\overline{x}_n - a_{1,2} = \overline{x}_n \mp \sqrt{3\left(\overline{(x^2)}_n - \overline{x}_n^2\right)}.$$

Wir erhalten somit die Schätzer

$$T_{MM}^a = \overline{X}_n - \sqrt{3\left(\overline{(X^2)}_n - \overline{X}_n^2\right)} \quad \text{und} \quad T_{MM}^b = \overline{X}_n + \sqrt{3\left(\overline{(X^2)}_n - \overline{X}_n^2\right)}.$$

Einsetzen ergibt die realisierten Schätzwerte

$$T_{MM}^a(\omega) = 7.86 \quad \text{and} \quad T_{MM}^b(\omega) = 18.02.$$

Bei dieser Schätzmethode vermutet Sarah also, dass Benno zwischen 7:52 und 18:01 Uhr lernt.

ETH Zürich FS 2021
D-INFK
Prof. Dr. Martin Schweizer

Koordinator
Marco Rodrigues

Wahrscheinlichkeit & Statistik

Musterlösung 8

1. **a)** Aus der Definition von Y folgt

$$E[Y] = \sum_{i=1}^{n} E[X_i^2] = \sum_{i=1}^{n} 1 = n.$$

Andererseits ist wegen Unabhängigkeit und identischer Verteilung

$$E[Y^2] = E[(X_1^2 + X_2^2 + \cdots + X_n^2)^2] = nE[X_1^4] + n(n-1)(E[X_1^2])^2.$$

Mit $E[X_i^2] = 1$ und

$$\begin{aligned} E[X_i^4] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^4 e^{-x^2/2}\,dx \\ &= \frac{1}{\sqrt{2\pi}} \left[-x^3 e^{-x^2/2}\right]_{-\infty}^{\infty} + \frac{3}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2}\,dx = 3 \end{aligned}$$

erhalten wir schliesslich $E[Y^2] = 3n + n(n-1) = n^2 + 2n$, und daraus

$$\mathrm{Var}[Y] = (n^2 + 2n) - n^2 = 2n.$$

Alternativ: Mit $X_1, \ldots, X_n$ sind auch $X_1^2, \ldots, X_n^2$ unabhängig und damit ist

$$\mathrm{Var}[Y] = \sum_{i=1}^{n} \mathrm{Var}[X_i^2] = n\mathrm{Var}[X_1^2].$$

Wie oben ist

$$\mathrm{Var}[X_1^2] = E[X_1^4] - (E[X_1^2])^2 = 3 - 1 = 2,$$

also $\mathrm{Var}[Y] = 2n$.

b) Mit der Chebyshev-Ungleichung erhalten wir

$$\begin{aligned} P\left[\left|\frac{Y}{n} - 1\right| \le 0.75\right] &= P\left[\left|\frac{Y-n}{n}\right| \le \frac{3}{4}\right] = 1 - P\left[|Y - E[Y]| > \frac{3n}{4}\right] \\ &\ge 1 - \frac{\mathrm{Var}[Y]}{9n^2/16} = 1 - \frac{32}{9n} \quad \left(= \frac{19}{27} \approx 0.7037 \text{ für } n = 12\right). \end{aligned}$$

Siehe nächste Seite!

c) Seien $Z_i := X_i^2$ für $i = 1, \ldots, n$. Dann gilt $Z_i \sim \chi_1^2$, also $E[Z_i] = 1$ und $\mathrm{Var}[Z_i] = 2$. Mit dem zentralen Grenzwertsatz folgt, dass

$$\begin{aligned} P\left[\left|\frac{Y}{n} - 1\right| \leq 0.75\right] &= P\left[\left|\frac{Y-n}{n}\right| \leq \frac{3}{4}\right] \\ &= P\left[\left|\frac{Y-n}{\sqrt{2n}}\right| \leq \frac{3}{4}\sqrt{\frac{n}{2}}\right] \\ &\approx \Phi\left(\frac{3}{4}\sqrt{\frac{n}{2}}\right) - \Phi\left(-\frac{3}{4}\sqrt{\frac{n}{2}}\right) \\ &= 2\Phi\left(\frac{3}{4}\sqrt{\frac{n}{2}}\right) - 1. \end{aligned}$$

Für $n = 12$ ist

$$P\left[\left|\frac{Y}{n} - 1\right| \leq 0.75\right] \approx 2\Phi\left(\frac{3}{4}\sqrt{6}\right) - 1 = 2\Phi(1.837) - 1 \approx 0.9338.$$

2. a) Sei $\vartheta \in \mathbb{R}$ fixiert. Aus der Gleichverteilung folgt, dass

$$E_\vartheta[T^{(n)}] = \left(\frac{1}{n}\sum_{i=1}^n E_\vartheta[X_i]\right) + \frac{1}{2} = E_\vartheta[X_1] + \frac{1}{2} = \vartheta - \frac{1}{2} + \frac{1}{2} = \vartheta.$$

Um die Rechnung für den anderen Schätzer zu vereinfachen, führen wir die Zufallsvariablen $Y_i = X_i - (\vartheta - 1)$ ein. Unter P_ϑ sind $Y_1, \ldots, Y_n$ unabhängig und $\mathcal{U}(0,1)$-verteilt. Insbesondere ist

$$Y_{(n)} := \max\{Y_1, \ldots, Y_n\} = \max\{X_1, \ldots, X_n\} - (\vartheta - 1) = \hat{T}^{(n)} - (\vartheta - 1).$$

In Aufgabe 3 der Serie 7 haben wir gezeigt, dass

$$f_{Y_{(n)}}(t) = nF(t)^{n-1} f(t),$$

wobei f die Dichte und F die Verteilungsfunktion von Y_1 sind. Es folgt, dass

$$\begin{aligned} E_\vartheta[Y_{(n)}] &= \int_0^1 tnF(t)^{n-1} f(t)dt \\ &= n\int_0^1 t^n dt \\ &= n\left[\frac{1}{n+1}t^{n+1}\right]_{t=0}^{t=1} \\ &= \frac{n}{n+1} = 1 - \frac{1}{n+1}. \end{aligned}$$

Also ist

$$E_\vartheta[\hat{T}^{(n)}] = E_\vartheta[Y_{(n)}] + (\vartheta - 1) = 1 - \frac{1}{n+1} + (\vartheta - 1) = \vartheta - \frac{1}{n+1}.$$

Der zweite Schätzer ist somit nicht erwartungstreu.

Siehe nächste Seite!

b) Wir erinnern zuerst daran, dass $\mathrm{Var}[X+b]=\mathrm{Var}[X]$ für eine beliebige Zufallsvariable Z und $b\in\mathbb{R}$ gilt. Also folgt wegen der Unabhängigkeit von $X_1,\dots,X_n$ und $Y_1\sim\mathcal{U}(0,1)$, dass

$$\mathrm{Var}_\vartheta[T^{(n)}]=\frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}_\vartheta[X_i-(\vartheta-1)]=\frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}_\vartheta[Y_i]=\frac{1}{n}\mathrm{Var}_\vartheta[Y_1]=\frac{1}{12n}.$$

Um die Varianz des anderen Schätzers zu bestimmen, berechnen wir zuerst $\mathrm{Var}_\vartheta[Y_{(n)}]$. Auf ähnliche Weise wie in a) erhalten wir

$$\begin{aligned}E_\vartheta[Y_{(n)}^2]&=n\int_0^1 t^2t^{n-1}dt\\&=n\int_0^1 t^{n+1}dt\\&=n\left[\frac{1}{n+2}t^{n+2}\right]_{t=0}^{t=1}\\&=\frac{n}{n+2}.\end{aligned}$$

Also ist

$$\begin{aligned}\mathrm{Var}_\vartheta[\hat{T}^{(n)}]&=\mathrm{Var}_\vartheta[Y_{(n)}+(\vartheta-1)]=\mathrm{Var}_\vartheta[Y_{(n)}]\\&=E_\vartheta[Y_{(n)}^2]-(E_\vartheta[Y_{(n)}])^2=\frac{n}{n+2}-\left(\frac{n}{n+1}\right)^2=\frac{n}{(n+1)^2(n+2)}.\end{aligned}$$

c) Sei $\varepsilon>0$. Aus der Chebyshev-Ungleichung folgt, dass

$$P_\vartheta[|T^{(n)}-\vartheta|>\varepsilon]\le\frac{\mathrm{Var}_\vartheta[T^{(n)}]}{\varepsilon^2}=\frac{1}{12n\varepsilon^2}\xrightarrow{n\to\infty}0,$$

weil $T^{(n)}$ erwartungstreu ist. Die Folge von Schätzern $T^{(n)}$, $n\in\mathbb{N}$, ist also konsistent. Da $\hat{T}^{(n)}$ nicht erwartungstreu ist, müssen wir bei diesem Schätzer anders vorgehen. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $\varepsilon\in(0,1)$. Wegen $\hat{T}^{(n)}\le\vartheta$ ist

$$\begin{aligned}P_\vartheta[|\hat{T}^{(n)}-\vartheta|>\varepsilon]&=P_\vartheta[\vartheta-\hat{T}^{(n)}>\varepsilon]\\&=P_\vartheta[\vartheta-\varepsilon>\hat{T}^{(n)}]\\&=P_\vartheta[\vartheta-\varepsilon>\max\{X_1,\dots,X_n\}]\\&=(P_\vartheta[\vartheta-\varepsilon>X_1])^n\\&=(\vartheta-\varepsilon-(\vartheta-1))^n\\&=(1-\varepsilon)^n\xrightarrow{n\to\infty}0.\end{aligned}$$

Also ist auch $\hat{T}^{(n)}$, $n\in\mathbb{N}$, eine Folge konsistenter (aber nicht erwartungstreuer) Schätzer.

Siehe nächste Seite!

3. **a)** Die Linearität des Erwartungswertes liefert

$$E_\vartheta[T^{(n)}] = E_\vartheta\left[\frac{1}{n}\sum_{i=1}^n \log(1+X_i)\right] = \frac{1}{n}\sum_{i=1}^n E_\vartheta[\log(1+X_i)] = E_\vartheta[\log(1+X_1)].$$

Nach dem Hinweis ist $Y_1 := \log(1+X_1) \sim Exp(\frac{1}{\vartheta})$. Also ist $E_\vartheta[Y_1] = \vartheta$ und $\mathrm{Var}_\vartheta[Y_1] = \vartheta^2$. Daher gilt $E_\vartheta[T^{(n)}] = E_\vartheta[Y_1] = \frac{1}{\frac{1}{\vartheta}} = \vartheta$.

Die Varianz rechnen wir auf eine ähnliche Weise aus:

$$\begin{aligned}\mathrm{Var}_\vartheta[T^{(n)}] &= \mathrm{Var}_\vartheta\left[\frac{1}{n}\sum_{i=1}^n \log(1+X_i)\right] = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}_\vartheta[\log(1+X_i)] \\ &= \frac{1}{n}\mathrm{Var}_\vartheta[\log(1+X_1)] = \frac{1}{n}\mathrm{Var}_\vartheta[Y_1] = \frac{1}{n}\frac{1}{\frac{1}{\vartheta^2}} = \frac{1}{n}\vartheta^2.\end{aligned}$$

b) Aus a) folgt, dass $E_\vartheta[T^{(n)}] = \vartheta$, d.h. $T^{(n)}$ ist erwartungstreu für ϑ.

c) Die Chebyshev-Ungleichung liefert für jedes $n \in \mathbb{N}$ und für beliebiges $\varepsilon > 0$

$$P_\vartheta[|T^{(n)} - \vartheta| > \varepsilon] \leq \frac{1}{\varepsilon^2}\mathrm{Var}_\vartheta[T^{(n)}] = \frac{1}{\varepsilon^2}\frac{1}{n}\vartheta^2 \overset{n\to\infty}{\longrightarrow} 0.$$

Also ist $T^{(n)}$, $n \in \mathbb{N}$, eine Folge konsistenter Schätzer.

4. **a)** Die Likelihood-Funktion ergibt sich aus dem Produkt der Dichten als

$$L(x_1, \ldots x_n; \vartheta) = \prod_{i=1}^n f(x_i) = \prod_{i=1}^n \frac{\vartheta}{x_i^{\vartheta+1}} = \vartheta^n \frac{1}{(\prod_{i=1}^n x_i)^{\vartheta+1}} .$$

Beachten Sie: wir können immer die Formel für $x_i \geq 1$ verwenden; der Fall $x_i < 1$ tritt nur mit Wahrscheinlichkeit 0 ein.

Die log-Likelihood-Funktion erhalten wir durch Logarithmieren obiger Formel als

$$\ell(x_1, \ldots x_n; \vartheta) = \log L(x_1, \ldots x_n; \vartheta) = n \log \vartheta - (\vartheta + 1)\sum_{i=1}^n \log x_i.$$

b) Ableiten und Nullsetzen der log-Likelihood-Funktion ergibt

$$\frac{\partial}{\partial \vartheta}\ell(x_1, \ldots x_n; \vartheta) = \frac{n}{\vartheta} - \sum_{i=1}^n \log x_i = 0$$

für $\vartheta = \frac{n}{\sum_{i=1}^n \log x_i}$. Also ist der Maximum-Likelihood-Schätzer

$$T_{ML} = \frac{n}{\sum_{i=1}^n \log X_i} = \frac{1}{\frac{1}{n}\sum_{i=1}^n \log X_i}.$$

Der realisierte Schätzwert für die gegebenen Daten ist dann

$$t_{ML} = \frac{5}{\sum_{i=1}^5 \log x_i} = 0.4214.$$

Wahrscheinlichkeit & Statistik

Musterlösung 5

1. Sei $Z := X + Y$. Aus der Vorlesung ist bekannt, dass

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = \int_{-\infty}^{\infty} 1_{[0,1]}(x)1_{[0,1]}(z-x)dx = \int_0^1 1_{[0,1]}(z-x)dx.$$

Zuerst bemerken wir, dass $f_Z(z) = 0$ für $z > 2$ oder $z < 0$. Betrachten wir nun den Fall, dass $0 \leq z \leq 1$. Für $x \in [0,1]$ gilt $0 \leq z - x \leq 1$ genau dann, wenn $x \leq z$. Also ist

$$f_Z(z) = \int_0^1 1_{[0,1]}(z-x)dx = \int_0^z dx = z \quad \text{für } 0 \leq z \leq 1.$$

Sei nun $1 < z \leq 2$. Für $x \in [0,1]$ gilt $0 \leq z - x \leq 1$ genau dann, wenn $x \geq z - 1$. Also ist

$$f_Z(z) = \int_0^1 1_{[0,1]}(z-x)dx = \int_{z-1}^1 dx = 1 - (z-1) = 2 - z \quad \text{für } 1 < z \leq 2.$$

Zusammengefasst ist

$$f_Z(z) = \begin{cases} 0 & \text{für } z < 0 \text{ oder } z > 2, \\ z & \text{für } 0 \leq z \leq 1, \\ 2 - z & \text{für } 1 < z \leq 2. \end{cases}$$

Mit Integration und einer Fallunterscheidung ergibt sich die Verteilungsfunktion

$$F_Z(t) = \begin{cases} 0 & \text{für } t < 0, \\ \frac{t^2}{2} & \text{für } 0 \leq t \leq 1, \\ 1 - \frac{(2-t)^2}{2} & \text{für } 1 < t \leq 2, \\ 1 & \text{für } t > 2. \end{cases}$$

Siehe nächste Seite!

2. Wegen $X, Y > 0$ ist klar, dass $F_Z(t) = 0$ für $t \leq 0$. Für $t > 0$ ist

$$\begin{aligned}
F_Z(t) &= P[X/Y \leq t] \\
&= P[X \leq tY] \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) 1_{\{x \leq ty\}} dx dy \\
&= \int_0^{\infty} \int_0^{\infty} f_X(x) f_Y(y) 1_{\{x \leq ty\}} dx dy \\
&= \int_0^{\infty} f_Y(y) \int_0^{ty} f_X(x) dx dy \\
&= \int_0^{\infty} \lambda e^{-\lambda y} \int_0^{ty} \lambda e^{-\lambda x} dx dy \\
&= \lambda^2 \int_0^{\infty} e^{-\lambda y} \left[-\frac{1}{\lambda} e^{-\lambda x} \right]_{x=0}^{x=ty} dy \\
&= \lambda^2 \int_0^{\infty} e^{-\lambda y} \left(-\frac{1}{\lambda} e^{-\lambda t y} + \frac{1}{\lambda} \right) dy \\
&= -\lambda \int_0^{\infty} \left(e^{-\lambda(1+t)y} - e^{-\lambda y} \right) dy \\
&= -\lambda \int_0^{\infty} e^{-\lambda(1+t)y} dy + \int_0^{\infty} \lambda e^{-\lambda y} dy \\
&= \frac{1}{1+t} \left[e^{-\lambda(1+t)y} \right]_{y=0}^{y=\infty} - 1 \\
&= -\frac{1}{1+t} + 1 \\
&= \frac{t}{1+t}.
\end{aligned}$$

3. a) Wir zeigen zuerst, dass Y und Z unkorreliert sind. Nach Definition der Kovarianz ist

$$\begin{aligned}
\mathrm{Cov}(Y, Z) &= E[YZ] - E[Y]E[Z] \\
&= E[\cos(X)\sin(X)] - E[\cos(X)]E[\sin(X)] \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(x)\sin(x) dx - \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(x) dx \right) \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \sin(x) dx \right) \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(x)\sin(x) dx,
\end{aligned}$$

und mit partieller Integration erhalten wir

$$\int_{-\pi}^{\pi} \cos(x)\sin(x) dx = \left[\sin^2(x)\right]_{x=-\pi}^{x=\pi} - \int_{-\pi}^{\pi} \sin(x)\cos(x) dx = - \int_{-\pi}^{\pi} \sin(x)\cos(x) dx.$$

Also ist

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(x)\sin(x) dx = 0$$

Siehe nächste Seite!

und

$$\operatorname{Cov}(Y,Z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(x)\sin(x)dx = 0.$$

Alternativ: Aus $\sin(x)\cos(x) = \frac{1}{2}\sin(2x)$ folgt, dass

$$\int_{-\pi}^{\pi} \cos(x)\sin(x)dx = \frac{1}{2}\int_{-\pi}^{\pi} \sin(2x)dx = -\frac{1}{4}[\cos(2x)]_{x=-\pi}^{x=\pi} = 0$$

wegen $\cos(2\pi) = \cos(-2\pi) = 1$.

b) Wir zeigen nun, dass Y und Z nicht unabhängig sind. Wegen

$$Y^2 + Z^2 = \cos^2(X) + \sin^2(X) = 1$$

ist

$$0 \leq P[Y^2 < 1/2, Z^2 < 1/2] \leq P[Y^2 + Z^2 < 1] = 0.$$

Jedoch ist

$$\begin{aligned} P[Y^2 < 1/2] &= \frac{1}{2\pi}\int_{-\pi}^{\pi} 1_{\{\cos^2(y)<1/2\}} dy \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} 1_{\{-1/\sqrt{2}<\cos(y)<1/\sqrt{2}\}} dy \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} 1_{\{\pi/4<y<3\pi/4\}\cup\{-3\pi/4<y<-\pi/4\}} dy \\ &= \frac{1}{2\pi}\left(\frac{3\pi}{4} - \frac{\pi}{4} + \frac{-\pi}{4} - \frac{-3\pi}{4}\right) \\ &= \frac{1}{2}. \end{aligned}$$

Analog erhält man

$$P[Z^2 < 1/2] = \frac{1}{2}.$$

Angenommen Y und Z sind unabhängig. Dann sind auch Y^2 und Z^2 unabhängig. Wegen

$$P[Y^2 < 1/2, Z^2 < 1/2] = 0 \neq \frac{1}{4} = \frac{1}{2}\cdot\frac{1}{2} = P[Y^2 < 1/2]P[Z^2 < 1/2]$$

ist dies jedoch ein Widerspruch. Also sind Y und Z unkorreliert, aber nicht unabhängig.

4. a) Die Inverse von Σ ist

$$\Sigma^{-1} = \frac{1}{\det(\Sigma)}\begin{pmatrix} \Sigma_{2,2} & -\Sigma_{1,2} \\ -\Sigma_{2,1} & \Sigma_{1,1} \end{pmatrix}.$$

Siehe nächste Seite!

Damit folgt, dass die Randdichte von X_1

$$
\begin{aligned}
f_{X_1}(x_1) &= \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)dx_2 \\
&= \frac{1}{2\pi\sqrt{\det(\Sigma)}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right) dx_2 \\
&= \frac{1}{2\pi\sqrt{\det(\Sigma)}} \int_{-\infty}^{\infty} \exp\Bigg(- \frac{1}{2\det(\Sigma)}\big((x_1-\mu_1)^2\Sigma_{2,2} \\
&\qquad - 2(x_1-\mu_1)(x_2-\mu_2)\Sigma_{1,2} + (x_2-\mu_2)^2\Sigma_{1,1}\big)\Bigg)dx_2 \\
&= \frac{1}{2\pi\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2\det(\Sigma)}\big((x_1-\mu_1)^2\Sigma_{2,2}\right) \\
&\qquad \times \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\det(\Sigma)}\left(-2(x_1-\mu_1)(x_2-\mu_2)\Sigma_{1,2} + (x_2-\mu_2)^2\Sigma_{1,1}\big)\right)\right) dx_2 \\
&= \frac{1}{2\pi\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2\det(\Sigma)}\big((x_1-\mu_1)^2\Sigma_{2,2}\right) \\
&\qquad \times \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\det(\Sigma)}\left(-2(x_1-\mu_1)y_2\Sigma_{1,2} + y_2^2\Sigma_{1,1}\big)\right)\right) dy_2 \\
&= \frac{1}{2\pi\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2\det(\Sigma)}\big((x_1-\mu_1)^2\Sigma_{2,2}\right) \int_{-\infty}^{\infty} \exp\left(-ay_2^2 - by_2\right) dy_2 \\
&= \frac{1}{2\pi\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2\det(\Sigma)}(x_1-\mu_1)^2\Sigma_{2,2}\right) \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{4a}\right) \\
&= \frac{1}{2\pi\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}\frac{(x_1-\mu_1)^2}{\Sigma_{1,1}\det(\Sigma)}\Sigma_{1,1}\Sigma_{2,2}\right) \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{4a}\right),
\end{aligned}
$$

wobei wir nach der dritten Gleichheit die Variablentransformation $y_2 := x_2 - \mu_2$ gemacht haben, ab der fünften Gleichheit

$$
a := \frac{\Sigma_{1,1}}{2\det(\Sigma)} \quad \text{und} \quad b := -(x_1-\mu_1)\frac{\Sigma_{1,2}}{\det(\Sigma)}
$$

gesetzt haben und bei der siebten Gleichheit den Hinweis benutzt haben. Insbesondere ist

$$
\sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{4a}\right) = \sqrt{\frac{2\pi\det(\Sigma)}{\Sigma_{1,1}}} \exp\left(\frac{1}{2}\frac{(x_1-\mu_1)^2\Sigma_{1,2}^2}{\Sigma_{1,1}\det(\Sigma)}\right)
$$

und damit ist

$$
\begin{aligned}
f_{X_1}(x_1) &= \frac{1}{\sqrt{2\pi\Sigma_{1,1}}} \exp\left(-\frac{1}{2}\frac{(x_1-\mu_1)^2(\Sigma_{1,1}\Sigma_{2,2}-\Sigma_{1,2}^2)}{\Sigma_{1,1}\det(\Sigma)}\right) \\
&= \frac{1}{\sqrt{2\pi\Sigma_{1,1}}} \exp\left(-\frac{1}{2}\frac{(x_1-\mu_1)^2\det(\Sigma)}{\Sigma_{1,1}\det(\Sigma)}\right) \\
&= \frac{1}{\sqrt{2\pi\Sigma_{1,1}}} \exp\left(-\frac{1}{2}\frac{(x_1-\mu_1)^2}{\Sigma_{1,1}}\right).
\end{aligned}
$$

Siehe nächste Seite!

Wir erkennen nun, dass $X_1 \sim \mathcal{N}(\mu_1, \Sigma_{1,1})$. Analog zeigt man $X_2 \sim \mathcal{N}(\mu_2, \Sigma_{2,2})$. (Das ist auch klar aus der Symmetrie des Problems.) Also ist $\mathrm{Var}[X_1] = \Sigma_{1,1}$ und $\mathrm{Var}[X_2] = \Sigma_{2,2}$ sowie $E[X_1] = \mu_1$, $E[X_2] = \mu_2$.

Nach Definition der Kovarianz ist

$$\begin{aligned}\mathrm{Cov}(X_1, X_2) &= E\big[(X_1 - E[X_1])(X_2 - E[X_2])\big] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1 - \mu_1)(x_2 - \mu_2) f_{X_1,X_2}(x_1, x_2) dx_1 dx_2.\end{aligned}$$

Mit der Variablentransformation $y_1 := x_1 - \mu_1$ und $y_2 := x_2 - \mu_2$ folgt, dass

$$\begin{aligned}\mathrm{Cov}(X_1, X_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_2 y_1 f_{X_1,X_2}(y_1 + \mu_1, y_2 + \mu_2) dy_2 dy_1 \\ &= \frac{1}{2\pi\sqrt{\det(\Sigma)}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_2 y_1 \exp\bigg(- \frac{1}{2\det(\Sigma)} \big(y_1^2 \Sigma_{2,2} \\ &\qquad - 2\Sigma_{1,2} y_1 y_2 + y_2^2 \Sigma_{1,1}\big)\bigg) dy_1 dy_2 \\ &= \frac{1}{2\pi\sqrt{\det(\Sigma)}} \int_{-\infty}^{\infty} y_1 \int_{-\infty}^{\infty} y_2 \exp\bigg(- \frac{1}{2\det(\Sigma)} \big(y_1^2 \Sigma_{2,2} \\ &\qquad - 2\Sigma_{1,2} y_1 y_2 + y_2^2 \Sigma_{1,1}\big)\bigg) dy_2 dy_1 \\ &= \frac{1}{2\pi\sqrt{\det(\Sigma)}} \int_{-\infty}^{\infty} y_1 \int_{-\infty}^{\infty} y_2 \exp\Bigg(- \frac{\Sigma_{1,1}}{2\det(\Sigma)} \Bigg(\bigg(y_2 - \frac{\Sigma_{1,2}}{\Sigma_{1,1}} y_1\bigg)^2 \\ &\qquad + \frac{\det(\Sigma)}{\Sigma_{1,1}^2} y_1^2 \Bigg)\Bigg) dy_2 dy_1 \\ &= \frac{1}{2\pi\sqrt{\det(\Sigma)}} \int_{-\infty}^{\infty} y_1 \exp\bigg(-\frac{y_1^2}{2\Sigma_{1,1}}\bigg) \int_{-\infty}^{\infty} y_2 \exp\Bigg(- \frac{\Sigma_{1,1}}{2\det(\Sigma)} \\ &\qquad \times \bigg(y_2 - \frac{\Sigma_{1,2}}{\Sigma_{1,1}} y_1\bigg)^2 \Bigg) dy_2 dy_1,\end{aligned}$$

wobei wir in der vierten Gleichheit

$$y_1^2 \Sigma_{2,2} - \bigg(\frac{\Sigma_{1,2}}{\Sigma_{1,1}} y_1\bigg)^2 \Sigma_{1,1} = y_1^2 \frac{1}{\Sigma_{1,1}} \cdot (\Sigma_{1,1}\Sigma_{2,2} - \Sigma_{1,2}^2) = \frac{\det(\Sigma)}{\Sigma_{1,1}} y_1^2$$

Siehe nächste Seite!

benutzt haben. Nun ist

$$
\begin{aligned}
&\int_{-\infty}^{\infty} y_2 \exp\left(-\frac{\Sigma_{1,1}}{2\det(\Sigma)}\left(y_2 - \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1\right)^2\right) dy_2 \\
&\quad = \int_{-\infty}^{\infty} \left(y_2 - \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1 + \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1\right) \exp\left(-\frac{\Sigma_{1,1}}{2\det(\Sigma)}\left(y_2 - \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1\right)^2\right) dy_2 \\
&\quad = \int_{-\infty}^{\infty} \left(y_2 - \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1\right) \exp\left(-\frac{\Sigma_{1,1}}{2\det(\Sigma)}\left(y_2 - \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1\right)^2\right) dy_2 \\
&\qquad + \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1 \int_{-\infty}^{\infty} \exp\left(-\frac{\Sigma_{1,1}}{2\det(\Sigma)}\left(y_2 - \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1\right)^2\right) dy_2 \\
&\quad = \int_{-\infty}^{\infty} z_2 \exp\left(-\frac{\Sigma_{1,1}}{2\det(\Sigma)} z_2^2\right) dz_2 \\
&\qquad + \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1 \int_{-\infty}^{\infty} \exp\left(-\frac{\Sigma_{1,1}}{2\det(\Sigma)}\left(y_2 - \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1\right)^2\right) dy_2 \\
&\quad = 0 + \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1 \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\frac{\det(\Sigma)}{\Sigma_{1,1}}}\left(y_2 - \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1\right)^2\right) dy_2 \\
&\quad = \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1 \frac{\sqrt{2\pi\frac{\det(\Sigma)}{\Sigma_{1,1}}}}{\sqrt{2\pi\frac{\det(\Sigma)}{\Sigma_{1,1}}}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\frac{\det(\Sigma)}{\Sigma_{1,1}}}\left(y_2 - \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1\right)^2\right) dy_2 \\
&\quad = \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1 \sqrt{2\pi\frac{\det(\Sigma)}{\Sigma_{1,1}}}.
\end{aligned}
$$

Die dritte Gleichheit folgt durch die Variablentransformation $z_2 := y_2 - \frac{\Sigma_{1,2}}{\Sigma_{1,1}}y_1$, die vierte Gleichheit gilt, weil der Integrand im ersten Integral ungerade ist, und bei der letzten Gleichheit haben wir benutzt, dass das Integral und der Nenner vor dem Integral das Integral über die Dichte einer Normalverteilung ist. Somit ist

$$
\begin{aligned}
\mathrm{Cov}(X_1, X_2) &= \frac{\Sigma_{1,2}}{\Sigma_{1,1}} \frac{1}{\sqrt{2\pi\Sigma_{1,1}}} \int_{-\infty}^{\infty} y_1^2 \exp\left(-\frac{y_1^2}{2\Sigma_{1,1}}\right) dy_1 \\
&= \frac{\Sigma_{1,2}}{\Sigma_{1,1}} \frac{1}{\sqrt{2\pi\Sigma_{1,1}}} \int_{-\infty}^{\infty} (z_1 - \mu_1)^2 \exp\left(-\frac{(z_1-\mu_1)^2}{2\Sigma_{1,1}}\right) dz_1 \\
&= \frac{\Sigma_{1,2}}{\Sigma_{1,1}} \mathrm{Var}[X_1] \\
&= \frac{\Sigma_{1,2}}{\Sigma_{1,1}} \Sigma_{1,1} \\
&= \Sigma_{1,2}.
\end{aligned}
$$

Hier haben wir bei der zweiten Gleichheit die Variablentransformation $z_1 := y_1 + \mu_1$ verwendet.

Siehe nächste Seite!

b) Aus der Vorlesung ist schon bekannt, dass $\mathrm{Cov}(X_1, X_2) = 0$, falls X_1 und X_2 unabhängig sind. Nehmen wir also an, dass $\mathrm{Cov}(X_1, X_2) = \Sigma_{1,2} = \Sigma_{2,1} = 0$. Man rechnet leicht nach, dass dann

$$f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2) \quad \text{für alle } (x_1, x_2) \in \mathbb{R}$$

gilt. Also sind X_1 und X_2 unabhängig.

ETH Zürich FS 2021
D-INFK
Prof. Dr. Martin Schweizer

Koordinator
Marco Rodrigues

Wahrscheinlichkeit & Statistik

Musterlösung 4

1. **a)** Aus $\frac{1}{20} = P[Y \leq 1]$ folgt $0.95 = P[Y > 1] = e^{-\lambda}$. Also ist $\lambda = -\ln(0.95)$.

b) Die Wahrscheinlichkeit, dass die Sonde mehr als 10 Minuten überlebt, ist

$$P[Y > 10] = 1 - P[Y \leq 10] = 1 - F_Y(10) = 1 - (1 - e^{\ln(0.95)10}) = 0.95^{10} \approx 0.5987.$$

Alternativ ist $P[Y > 10] = e^{-\lambda \cdot 10} = e^{10\ln(0.95)} = (0.95)^{10}$.

c) Aus der Vorlesung ist bekannt, dass die Exponentialverteilung gedächtnislos ist. Also folgt, dass

$$P[Y > 30 | Y > 20] = P[Y > 10] = 0.95^{10}.$$

d) Wir möchten die Wahrscheinlichkeit $P[Z > Y]$ berechnen. Sei

$$A := \{(z, y) \in \mathbb{R}^2 : z > y\}.$$

Dann ist

$$\begin{aligned}
P[Z > Y] &= \int\int f_{Z,Y}(z, y) 1_A(z, y) dzdy \\
&= \int_{-\infty}^{\infty} \int_{y}^{\infty} f_{Z,Y}(z, y) dzdy \\
&= \int_{-\infty}^{\infty} \int_{y}^{\infty} f_Z(z) f_Y(y) dzdy \\
&= \int_{0}^{\infty} \int_{y}^{\infty} \mu e^{-\mu z} \lambda e^{-\lambda y} dzdy \\
&= \int_{0}^{\infty} \lambda e^{-\lambda y} \int_{y}^{\infty} \mu e^{-\mu z} dzdy \\
&= \int_{0}^{\infty} \lambda e^{-\lambda y} e^{-\mu y} dy \\
&= \int_{0}^{\infty} \lambda e^{-(\lambda+\mu) y} dy \\
&= \frac{\lambda}{\lambda + \mu}.
\end{aligned}$$

Siehe nächste Seite!

2. **a)** Wir bemerken zuerst, dass der Wertebereich von M das Einheitsintervall $[0,1]$ ist. Für $m \in [0,1]$ gilt mit der Unabhängigkeit und der Gleichverteilung von U_1, U_2 und U_3, dass

$$P[M \leq m] = P[U_1 \leq m, U_2 \leq m, U_3 \leq m] = (P[U_1 \leq m])^3 = m^3.$$

Also ist die Verteilungsfunktion F_M von M gegeben durch

$$F_M(m) = \begin{cases} 0 & \text{für } m < 0, \\ m^3 & \text{für } 0 \leq m \leq 1, \\ 1 & \text{für } m > 1. \end{cases}$$

Für $m \in (0,1)$ gilt also

$$f_M(m) = F_M'(m) = 3m^2,$$

und somit ist die Dichte von M gegeben durch

$$f_M(m) = \begin{cases} 0 & \text{für } m < 0, \\ 3m^2 & \text{für } 0 \leq m \leq 1, \\ 0 & \text{für } m > 1. \end{cases}$$

b) Die gemeinsame Verteilungsfunktion von (M, L) ist

$$F_{M,L}(m,\ell) = P[M \leq m, L \leq \ell] \quad \text{für } (m,\ell) \in \mathbb{R}^2.$$

Falls eine Dichte existiert, dann muss

$$f_{M,L}(m,\ell) = \frac{\partial^2 F_{M,L}}{\partial m \partial \ell}(m,\ell)$$

dort gelten, wo $F_{M,L}$ differenzierbar ist. Offenbar ist immer $L \leq M$, und die Wahrscheinlichkeit $P[M \leq m, L \leq \ell] = P[M \leq m] - P[\ell < L \leq M \leq m]$ hängt nur dann von beiden Argumenten ℓ und m ab, wenn $0 < \ell < m < 1$ gilt. Dann ist sie wegen der Unabhängigkeit und identischen Verteilung von U_1, U_2, U_3

$$\begin{aligned} P[\max(U_1,U_2,U_3) \leq m] - P[\ell < \min(U_1,U_2,U_3) \leq \max(U_1,U_2,U_3) \leq m] \\ = (P[U_1 \leq m])^3 - (P[\ell < U_1 \leq m])^3 = m^3 - (m-\ell)^3 \end{aligned}$$

wegen $U_1 \sim \mathcal{U}(0,1)$. Also ist für $0 < \ell < m < 1$

$$f_{M,L}(m,\ell) = \frac{\partial^2 F_{M,L}}{\partial m \partial \ell}(m,\ell) = 6(m-\ell).$$

Die gemeinsame Dichte von (M,L) ist also

$$f_{M,L}(m,\ell) = \begin{cases} 6(m-\ell) & \text{für } 0 < \ell < m < 1, \\ 0 & \text{sonst.} \end{cases}$$

Siehe nächste Seite!

c) Für $m \in (0,1]$ ist $f_M(m) > 0$. Also ist

$$f_{L|M}(\ell|m) = \frac{f_{M,L}(m,\ell)}{f_M(m)} = \frac{6(m-\ell)}{3m^2} 1_{\{0<\ell<m<1\}}.$$

Zusammengefasst ergibt sich also die bedingte Dichte

$$f_{L|M}(\ell|m) = \begin{cases} \frac{6(m-\ell)}{3m^2} & \text{für } 0 < \ell < m < 1, \\ 0 & \text{sonst.} \end{cases}$$

3. **a)** Aus der Vorlesung ist bekannt, dass

$$F_{X_1}(t_1) = P[X_1 \le t_1] = \lim_{t_2 \to \infty} P[X_1 \le t_1, X_2 \le t_2].$$

Wegen

$$\begin{aligned} P[X_1 \le t_1, X_2 \le t_2] &= \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} f_{X_1,X_2}(x_1,x_2) dx_2 dx_1 \\ &= \frac{\sqrt{2}}{\pi} \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} \exp\left(-\frac{3}{2}x_1^2 - x_1x_2 - \frac{3}{2}x_2^2\right) dx_2 dx_1 \\ &= \frac{\sqrt{2}}{\pi} \int_{-\infty}^{t_1} \exp\left(-\frac{3}{2}x_1^2\right) \int_{-\infty}^{t_2} \exp\left(-x_1x_2 - \frac{3}{2}x_2^2\right) dx_2 dx_1 \end{aligned}$$

folgt mit dem Hinweis, dass

$$\begin{aligned} F_{X_1}(t_1) &= \frac{\sqrt{2}}{\pi} \int_{-\infty}^{t_1} \exp\left(-\frac{3}{2}x_1^2\right) \int_{-\infty}^{\infty} \exp\left(-x_1x_2 - \frac{3}{2}x_2^2\right) dx_2 dx_1 \\ &= \frac{\sqrt{2}}{\pi} \sqrt{\frac{2\pi}{3}} \int_{-\infty}^{t_1} \exp\left(-\frac{3}{2}x_1^2\right) \exp\left(\frac{x_1^2}{6}\right) dx_1 \\ &= \frac{1}{\sqrt{2\pi \cdot \frac{3}{8}}} \int_{-\infty}^{t_1} \exp\left(-\frac{4}{3}x_1^2\right) dx_1 \\ &= \frac{1}{\sqrt{2\pi \cdot \frac{3}{8}}} \int_{-\infty}^{t_1} \exp\left(-\frac{1}{2} \cdot \frac{x_1^2}{3/8}\right) dx_1. \end{aligned}$$

Also ist $X_1 \sim \mathcal{N}(0, 3/8)$. Analog dazu berechnet man, dass $X_2 \sim \mathcal{N}(0, 3/8)$. (Das ist auch klar aus der Symmetrie des Problems.)

b) Die Zufallsvariablen X_1 und X_2 sind nicht unabhängig, da $f_{X_1,X_2} \neq f_{X_1} f_{X_2}$.

4. **a)** Dass $X = (X_1, X_2)$ uniform auf $[-1/2, 1/2]^2$ verteilt ist, bedeutet, dass

$$f_X(x_1,x_2) = 1_{[-1/2,1/2]^2}(x_1,x_2) = 1_{[-1/2,1/2]}(x_1) \cdot 1_{[-1/2,1/2]}(x_2).$$

Siehe nächste Seite!

Die Dichten von X_1 und X_2 sind also

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_X(x_1, x_2) dx_2 = 1_{[-1/2,1/2]}(x_1),$$
$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_X(x_1, x_2) dx_1 = 1_{[-1/2,1/2]}(x_2).$$

Wegen

$$f_X(x_1, x_2) = 1_{[-1/2,1/2]}(x_1) \cdot 1_{[-1/2,1/2]}(x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$$

für alle $(x_1, x_2) \in \mathbb{R}^2$ sind X_1 und X_2 unabhängig.

b) Das um $\pi/4$ gedrehte Quadrat $[-1/2, 1/2]^2$ ist gegeben durch die Menge

$$A := \left\{ (x_1, x_2) \in \mathbb{R}^2 : |x_1| + |x_2| \leq 1/\sqrt{2} \right\}.$$

Eine Diagonale im Quadrat A hat die Länge $\sqrt{2}$. Die Dichte von $X = (X_1, X_2)$ ist

$$f_{X_1,X_2}(x_1, x_2) = 1_A(x_1, x_2) = 1_{\{|x_1|+|x_2| \leq 1/\sqrt{2}\}},$$

und insbesondere hat $f_{X_1,X_2}(x_1, x_2)$ für alle $(x_1, x_2) \in A$ den gleichen Wert (nämlich 1). Ferner folgt, dass

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2) dx_2 = \begin{cases} 0 & \text{für } |x_1| > \frac{\sqrt{2}}{2}, \\ \sqrt{2} - 2|x_1| & \text{für } |x_1| \leq \frac{\sqrt{2}}{2} \end{cases}$$

die Dichte der Koordinate X_1 und

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2) dx_1 = \begin{cases} 0 & \text{für } |x_2| > \frac{\sqrt{2}}{2}, \\ \sqrt{2} - 2|x_2| & \text{für } |x_2| \leq \frac{\sqrt{2}}{2} \end{cases}$$

die Dichte der Koordinate X_2 ist. Damit ist das Produkt $f_{X_1}(x_1) f_{X_2}(x_2)$ nicht konstant auf dem Quadrat A, kann also nicht mit der gemeinsamen Dichte $f_{X_1,X_2}(x_1, x_2)$ übereinstimmen, und damit sind X_1 und X_2 nicht unabhängig.

Wahrscheinlichkeit & Statistik

Musterlösung 6

1. a) Für jedes $i \in \mathbb{N}$ kann die Zufallsvariable Y_i nur die Werte 1 und 0 annehmen. Insbesondere ist

$$P[Y_i = 0] = P[X_i > t] = P[X_1 > t]$$

und

$$P[Y_i = 1] = P[X_i \leq t] = P[X_1 \leq t],$$

da $X_1, X_2, \ldots$ identisch verteilt sind. Also ist auch die Folge $Y_1, Y_2, \ldots$ identisch verteilt, und ihre Unabhängigkeit folgt mit dem Hinweis.

b) Sei $t \in \mathbb{R}$ beliebig und seien $Y_1, Y_2, \ldots$ wie in der vorherigen Teilaufgabe definiert. Wir wissen, dass $Y_1, Y_2, \ldots$ eine Folge unabhängiger und identisch verteilter Zufallsvariablen ist. Insbesondere ist $E[Y_1] = P[X \leq t] = F(t)$. Mit dem starken Gesetz der grossen Zahlen folgt, dass

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \leq t\}} = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow[n\to\infty]{} E[Y_1] = F(t) \quad P\text{-fastsicher.}$$

2. a) Nach Definition des Erwartungswertes für diskrete Zufallsvariablen ist

$$E[X] = \sum_{k=0}^{\infty} k p_X(k) = \sum_{k=1}^{\infty} k e^{-\lambda}\frac{\lambda^k}{k!} = \lambda \sum_{k=1}^{\infty} e^{-\lambda}\frac{\lambda^{k-1}}{(k-1)!} = \lambda \sum_{j=0}^{\infty} e^{-\lambda}\frac{\lambda^j}{j!} = \lambda.$$

Analog erhält man

$$E[X(X-1)] = \sum_{k=0}^{\infty} k(k-1)e^{-\lambda}\frac{\lambda^k}{k!} = \lambda^2 \sum_{k=2}^{\infty} e^{-\lambda}\frac{\lambda^{k-2}}{(k-2)!} = \lambda^2.$$

Also ist

$$\mathrm{Var}[X] = E[X^2] - (E[X])^2 = E[X(X-1)] + E[X] - (E[X])^2 = \lambda.$$

b) Seien $X_1, X_2, \ldots$ unabhängige und $\mathcal{P}(1)$-verteilte Zufallsvariablen. Aus der Teilaufgabe a) ist bekannt, dass $\mu := E[X_1] = 1 = \mathrm{Var}[X_1] =: \sigma^2$. Ausserdem ist

Siehe nächste Seite!

$S_n = X_1 + \cdots + X_n$ wieder Poisson-verteilt mit Parameter n (Hinweis). Mit dem zentralen Grenzwertsatz folgt

$$\begin{aligned} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} &= \sum_{k=0}^{n} e^{-n} \frac{n^k}{k!} = \sum_{k=0}^{n} P[S_n = k] = P[S_n \leq n] \\ &= P\left[\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq \frac{n - n\mu}{\sqrt{n}\sigma}\right] = P\left[\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq 0\right] \overset{n\to\infty}{\longrightarrow} \Phi(0) = \frac{1}{2}. \end{aligned}$$

3. **a)** Für $f(x) = I_{[-3,1]}(x)$ ist

$$A = E[f(U)] = P[U \in [-3,1]] = P[f(U) = 1].$$

Insbesondere ist $f(U) \sim Be(A)$.

b) Da $f(U_1), f(U_2), \ldots$ unabhängig und $Be(A)$-verteilt sind (Hinweis aus Aufgabe 1), ist $nS_n \sim Bin(n, A)$. Also ist

$$P[nS_n = k] = \binom{n}{k} A^k (1-A)^{n-k} \quad \text{für } k \in \mathbb{N}.$$

c) Aus der Linearität des Erwartungswertes folgt, dass

$$E[S_n] = \frac{1}{n} \sum_{i=1}^{n} E[f(U_i)] = A,$$

und aus der Vorlesung ist bekannt, dass

$$\mathrm{Var}[S_n] = \frac{1}{n^2} \mathrm{Var}[nS_n] = \frac{1}{n^2} nA(1-A) = \frac{A(1-A)}{n}.$$

d) Für jedes $\varepsilon > 0$ folgt mit der Chebyshev-Ungleichung, dass

$$P[|S_n - A| \geq \varepsilon] \leq \frac{\mathrm{Var}(S_n)}{\varepsilon^2} = \frac{A(1-A)}{n\varepsilon^2} \leq \frac{1}{n\varepsilon^2}.$$

Insbesondere ist damit ersichtlich, dass $(S_n)_{n\in\mathbb{N}}$ in Wahrscheinlichkeit/stochastisch gegen A konvergiert.

e) Die Konvergenz würde auch mit einer Anwendung des schwachen Gesetzes der grossen Zahlen folgen. Zudem erfüllt $S_1, S_2, \ldots$ auch alle Voraussetzungen, unter denen das starke Gesetz der grossen Zahlen gilt. Also konvergiert die Folge sogar P-fastsicher gegen A.

f) Um das Integral A numerisch zu approximieren, generiert man sich zufällige Zahlen $u_1, u_2, \ldots$ aus der Standardnormalverteilung, also $u_1 = U_2(\omega), u_2 = U_2(\omega), \ldots$ für $U_1, U_2, \ldots$ unabhängig und $\mathcal{N}(0,1)$-verteilt. Für ω ausserhalb einer Menge mit Wahrscheinlichkeit 0, haben wir $u_1, u_2, \ldots$ so generiert, dass

$$\frac{1}{n} \sum_{i=1}^{n} f(u_i) \to A \quad \text{für } n \to \infty.$$

Siehe nächste Seite!

4. a) Wegen $E[X] = 50$ ist die Wahrscheinlichkeit, dass die Anzahl Turbinen in dieser Woche zwischen 40 und 60 liegt, gleich

$$P[40 \leq X \leq 60] = P[-10 \leq X - E[X] \leq 10] = P[|X - E[X]| \leq 10],$$

und mit der Chebyshev-Ungleichung folgt, dass

$$P[|X - 50| > 10] \leq P[|X - 50| \geq 10] \leq \frac{\sigma^2}{10^2} = 0.25.$$

Also ist

$$P[40 \leq X \leq 60] = P[|X - E[X]| \leq 10] \geq 0.75.$$

b) Mit der Annahme der Normalverteilung folgt

$$\begin{aligned}
P[|X - 50| \leq 10] &= P[40 \leq X \leq 60] \\
&= P[X \leq 60] - P[X < 40] \\
&= P\left[\frac{X - 50}{5} \leq \frac{60 - 50}{5}\right] - P\left[\frac{X - 50}{5} < \frac{40 - 50}{5}\right] \\
&= P\left[\frac{X - 50}{5} \leq 2\right] - P\left[\frac{X - 50}{5} < -2\right] \\
&= \Phi(2) - \Phi(-2) \\
&= \Phi(2) - (1 - \Phi(2)) \\
&= 2\Phi(2) - 1 \\
&= 2 \cdot 0.97725 - 1 \\
&= 0.9545,
\end{aligned}$$

wobei wir in der dritten Gleichheit benutzt haben, dass $(X - 50)/5 \sim \mathcal{N}(0, 1)$ wegen $X \sim \mathcal{N}(50, 25)$. Die Chebyshev-Ungleichung liefert also eine korrekte Abschätzung der Wahrscheinlichkeit (0.9545 ist grösser als 0.75), aber sie ist nicht sehr genau (0.75 liegt nicht sehr nahe an 0.9545). Dies war allerdings zu erwarten, weil die Chebyshev-Ungleichung für **jede** Zufallsvariable mit endlicher Varianz gilt.

ETH Zürich FS 2021
D-INFK
Prof. Dr. Martin Schweizer

Koordinator
Marco Rodrigues

Wahrscheinlichkeit & Statistik

Musterlösung 7

1. **a)** Sei $t \in \mathbb{R}$ beliebig. Wegen der Unabhängigkeit von $e^{tX_1}, \dots, e^{tX_n}$ ist

$$\begin{aligned} M_{X_1+\cdots+X_n}(t) = E[e^{t(X_1+\cdots+X_n)}] &= E[e^{tX_1} \cdots e^{tX_n}] \\ &= E[e^{tX_1}] \cdots E[e^{tX_n}] = M_{X_1}(t) \cdots M_{X_n}(t). \end{aligned}$$

b) Sei $t \in \mathbb{R}$ beliebig. Wegen

$$\begin{aligned} E[e^{tX_i}] = \sum_{k \in \mathcal{W}(X)} e^{tk} P[X = k] &= \sum_{k=0}^{\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t-1)} \end{aligned}$$

folgt mit Teilaufgabe a), dass

$$M_{X_1+\cdots+X_n}(t) = M_{X_1}(t) \cdots M_{X_n}(t) = \left(e^{\lambda(e^t-1)}\right)^n = e^{n\lambda(e^t-1)}.$$

Alternativ kann man benutzen, dass $X_1 + \cdots + X_n$ gemäss einem Hinweis aus Serie 6 Poisson-verteilt ist mit Parameter $n\lambda$. Dann folgt das Resultat sofort.

2. Sei $S_n := X_1 + \cdots + X_n$. Aus Aufgabe 1 wissen wir, dass $M_{S_n}(t)$ für alle $t \in \mathbb{R}$ endlich ist. Mit Satz 5.6 aus dem Skript folgt, dass

$$P\left[\frac{1}{n}\sum_{i=1}^{n} X_i \geq a\right] = P\left[\frac{1}{n} S_n \geq a\right] = P\left[S_n \geq na\right] \leq \exp\left(\inf_{t \in \mathbb{R}} \left(n \log M_{X_1}(t) - tna\right)\right).$$

In Aufgabe 1 haben wir gezeigt, dass $M_{X_1}(t) = e^{\lambda(e^t-1)}$ ist. Also ist

$$f(t) := n \log M_{X_1}(t) - tna = n \log e^{\lambda(e^t-1)} - tna = n\lambda(e^t - 1) - tna.$$

Falls t_0 ein kritischer Punkt von f ist, muss

$$f'(t_0) = n\lambda e^{t_0} - na = 0$$

Siehe nächste Seite!

gelten. Also ist $t_0 = \log \frac{a}{\lambda}$ der einzige kritische Punkt von f. Wegen $f''(t) > 0$ für alle $t \in \mathbb{R}$ ist f eine strikt konvexe Funktion und t_0 also ein globales Minimum. Damit ist

$$\inf_{t\in\mathbb{R}} \big(n \log M_{X_1}(t) - tna\big) = \inf_{t\in\mathbb{R}} f(t) = f(t_0) = n\,(a-\lambda) - na \log \frac{a}{\lambda}$$

und

$$P\left[\frac{1}{n}\sum_{i=1}^{n} X_i \geq a\right] \leq \exp\Big(n(a-\lambda) - na \log \frac{a}{\lambda}\Big) = e^{-nh(a;\lambda)}.$$

3. **a)** Wir berechnen zuerst die Dichte von $X_{(n)}$. Für die Verteilungsfunktion $F_{(n)}$ von $X_{(n)}$ gilt wegen Unabhängigkeit

$$F_{(n)}(t) = P[X_{(n)} \leq t] = P[X_1 \leq t, \dots, X_n \leq t] = \prod_{k=1}^{n} P[X_k \leq t] = \big(F(t)\big)^n.$$

Demzufolge erhalten wir

$$f_{(n)}(t) = \frac{d}{dt} F_{(n)}(t) = nF^{n-1}(t) f(t).$$

Analog gehen wir beim Minimum vor. Wir erhalten

$$F_{(1)}(t) = 1 - P[X_{(1)} > t] = 1 - P[X_1 > t, \dots, X_n > t] = 1 - \big(1 - F(t)\big)^n,$$

und deshalb ist

$$f_{(1)}(t) = \frac{d}{dt} F_{(1)}(t) = n\big(1 - F(t)\big)^{n-1} f(t).$$

b) Einsetzen ergibt gemäss a)

$$\begin{aligned}
P\left[\frac{X_{(n)} - a_n}{b_n} \leq t\right] &= P\left[X_{(n)} \leq a_n + b_n t\right] \\
&= \big(F(a_n + b_n t)\big)^n \\
&= \Big(1 - e^{-\lambda(a_n + b_n t)}\Big)^n \\
&= \left(1 - \frac{1}{n} e^{-t}\right)^n.
\end{aligned}$$

Das bedeutet, dass

$$\lim_{n\to\infty} P\left[\frac{X_{(n)} - a_n}{b_n} \leq t\right] = \lim_{n\to\infty} \left(1 - \frac{1}{n} e^{-t}\right)^n = e^{-e^{-t}}.$$

4. Sei $X_i := I_{\{\text{Passagier } i \text{ erscheint zum Flug}\}}$. Nach den gemachten Annahmen sind dann die X_i unabhängig und Bernoulli-verteilt mit Parameter $p = 0.96$. Ist n die Anzahl der

Siehe nächste Seite!

Buchungen, so ist $S_n = \sum_{i=1}^n X_i$ die Anzahl der erscheinenden Passagiere, und S_n ist $Bin(n,p)$-verteilt. Gesucht ist das grösste n, sodass $P[S_n \geq 527] \leq 0.05$. Mit

$$u := \frac{526 + \frac{1}{2} - np}{\sqrt{np(1-p)}} \tag{1}$$

gilt nach dem zentralen Grenzwertsatz (inklusive Kontinuitätskorrektur) für grosses n, dass

$$P[S_n \geq 527] = 1 - P[S_n \leq 526] = 1 - P\left[\frac{S_n - np}{\sqrt{np(1-p)}} \leq \frac{526 - np}{\sqrt{np(1-p)}}\right] \approx 1 - \Phi(u).$$

Die Lösung n ergibt sich also approximativ aus der Gleichung $u = \Phi^{-1}(0.95) = 1.645$. Quadriert man die u definierende Gleichung, so erhält man

$$u^2 = \frac{(526 + \frac{1}{2} - np)^2}{np(1-p)},$$

und wenn man die nach Multiplikation mit dem Term $np(1-p)$ entstehende quadratische Gleichung

$$np(1-p)u^2 = \left(526 + \frac{1}{2} - np\right)^2$$

nach n auflöst und dabei $u = 1.645$ einsetzt, so ergeben sich die Lösungen $n_1 \approx 553.71$ und $n_2 \approx 540.63$. Das liefert $u_1 = -1.098$ und $u_2 = 1.645$, und wegen $u > 0$ ergibt sich also approximativ die Antwort „540 Tickets dürfen verkauft werden“.

ETH Zürich FS 2021
D-INFK
Prof. Dr. Martin Schweizer

Koordinator
Marco Rodrigues

Wahrscheinlichkeit & Statistik

Musterlösung 13

1. a) Sei $X_1, \dots, X_{10}$ die Stichprobe, welche die gemessenen Temperaturen realisiert. Unter den gemachten Annahmen sind $X_1, \dots, X_{10}$ unabhängig und je $\mathcal{N}(\mu, \sigma^2)$-verteilt unter P_μ, wobei μ ein unbekannter Parameter und $\sigma = 0.5$ ist. Als Hypothese und Alternative wählen wir

$$H_0 : \mu = \mu_0 = 20 \quad \text{und} \quad H_A : \mu \neq \mu_0.$$

Unter den gemachten Annahmen ist es naheliegend einen z-Test durchzuführen. Wir wählen als Teststatistik also

$$T = \frac{\overline{X}_n - \mu_0}{\sigma/\sqrt{n}} = \frac{\overline{X}_{10} - 20}{0.5/\sqrt{10}}.$$

Unter der Hypothese H_0 gilt $T \sim \mathcal{N}(0,1)$, und wir wählen als Verwerfungsbereich $K_{\neq} = (-\infty, -c_{\neq}) \cup (c_{\neq}, \infty)$. Wir verwerfen also H_0, falls $|T| > c_{\neq}$ für ein zu bestimmendes $c_{\neq}$ ist. Da wir auf dem 5%-Niveau testen möchten, wählen wir $c_{\neq}$ so, dass

$$0.05 = \alpha = P_{\mu_0}[T \in K_{\neq}] = 2\big(1 - \Phi(c_{\neq})\big)$$

gilt. Das ergibt den Wert $c_{\neq} = \Phi^{-1}(1 - \alpha/2) = z_{0.975} = 1.96$. Der realisierte Wert der Teststatistik T ist

$$T(\omega) = t(x_1, \dots, x_{10}) = 3.4342.$$

Wegen $|T(\omega)| > 1.96$ verwerfen wir also die Hypothese, dass die Klimaanlage die angestrebte Raumtemperatur von 20.00 Grad Celsius im Mittel erreicht.

b) Der realisierte p-Wert ist

$$p\text{-Wert}(\omega) = P_{H_0}[|T| > 3.43] = 2P_{H_0}[T > 3.43] = 2\big(1 - \Phi(3.43)\big) \approx 0.0006.$$

2. a) Unter den Angaben aus der Fragestellung ist es nun naheliegend, einen einseitigen z-Test durchzuführen. Sei $X_1, \dots, X_{12}$ die Stichprobe, welche die Daten $x_1, \dots, x_{12}$ realisiert. Unter den Annahmen aus der Aufgabenstellung sind $X_1, \dots, X_{12}$ unabhängig und $\mathcal{N}(\mu, \sigma^2)$-verteilt unter P_μ, wobei μ ein unbekannter Parameter und $\sigma = 1.5$ ist. Als Hypothese und Alternative wählen wir

$$H_0 : \mu = \mu_0 = 70 \quad \text{und} \quad H_A : \mu < \mu_0.$$

Siehe nächste Seite!

Die Teststatistik lautet

$$T = \frac{\overline{X}_{12} - \mu_0}{\sigma/\sqrt{12}}$$

und der realisierte Wert ist

$$T(\omega) = \frac{70.25 - 70}{1.5/\sqrt{12}} = 0.5774.$$

Unter H_0 gilt $T \sim \mathcal{N}(0,1)$. Sei $K_< = (-\infty, c_<)$ unser Verwerfungsbereich. Um das 5%-Niveau einzuhalten, wählen wir

$$c_< = \Phi^{-1}(0.05) = -\Phi^{-1}(1-0.05) = -z_{0.95} = -1.645.$$

Wegen $T(\omega) > -1.645$ wird die Hypothese auf dem 5%- Niveau beibehalten.

b) Die Wahrscheinlichkeit für einen Fehler 2.Art für den obigen Test mit Alternative $\mu_A = 69.5$ ist

$$\begin{aligned}
P_{\mu_A}[T \in K_<^c] &= P_{\mu_A}[T \geq -1.645] \\
&= P_{\mu_A}\left[\frac{\overline{X}_{12} - \mu_0}{\sigma/\sqrt{12}} \geq -1.645\right] \\
&= P_{\mu_A}\left[\frac{\overline{X}_{12} - \mu_A}{\sigma/\sqrt{12}} - \frac{\mu_0 - \mu_A}{\sigma/\sqrt{12}} \geq -1.645\right] \\
&= P_{\mu_A}\left[\frac{\overline{X}_{12} - \mu_A}{\sigma/\sqrt{12}} \geq \frac{\mu_0 - \mu_A}{\sigma/\sqrt{12}} - 1.645\right] \\
&= 1 - \Phi\left(\frac{\mu_0 - \mu_A}{\sigma/\sqrt{12}} - 1.645\right) \\
&= 1 - \Phi\left(\frac{0.5}{1.5/\sqrt{12}} - 1.645\right) \\
&= 1 - \Phi(-0.49) = \Phi(0.49) = 0.6879,
\end{aligned}$$

wobei wir bei der fünften Gleichheit

$$\frac{\overline{X}_{12} - \mu_A}{\sigma/\sqrt{12}} \sim \mathcal{N}(0,1) \quad \text{unter } P_{\mu_A}$$

benutzt haben.

3. a) Die Stichproben sind ungepaart. Wir haben keine Informationen darüber, wie die Stichproben entstanden sind. Eine Paarung beispielsweise nach Reihenfolge oder Grösse kann nicht begründet werden.

b) Sei $X_1, \ldots, X_9$ bzw. $Y_1, \ldots, Y_9$ die Stichprobe, welche die Zeiten des Cisalpino bzw. des neuen Zugs realisiert. Es gilt also, dass $X_1, \ldots, X_9, Y_1, \ldots, Y_9$ unabhängig sind, wobei $X_i \sim \mathcal{N}(\mu_X, \sigma^2)$ und $Y_i \sim \mathcal{N}(\mu_Y, \sigma^2)$ für $i = 1, \ldots, n$. Wir wollen die beiden unbekannten Mittelwerte μ_X und μ_Y vergleichen. Da die Stichproben

Siehe nächste Seite!

ungepaart sind und die Streuung σ unbekannt ist, werden wir einen ungepaarten Zweistichproben-t-Test durchführen. Die Hypothese und Alternative lauten

$$H_0 : \mu_X = \mu_Y, \sigma^2 > 0 \qquad \text{und} \qquad H_A : \mu_X \neq \mu_Y, \sigma^2 > 0.$$

Die Teststatistik (vgl. Skript S. 154)

$$T = \frac{(\overline{X}_n - \overline{Y}_m) - (\mu_X - \mu_Y)}{S\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

ist unter der Hypothese t-verteilt mit $n + m - 2 = 16$ Freiheitsgraden (wir haben $n = m = 9$). Der Verwerfungsbereich zum 5%-Niveau ist somit

$$K = (-\infty, -t_{16,0.975}) \cup (t_{16,0.975}, \infty) \approx (-\infty, -2.12) \cup (2.12, \infty).$$

Der realisierte Wert der Teststatistik ist

$$T(\omega) = \frac{\overline{x}_9 - \overline{y}_9}{s\sqrt{\frac{1}{9} + \frac{1}{9}}} = \frac{143.33 - 142}{\sqrt{24.875}\sqrt{\frac{2}{9}}} \approx 0.54 \notin K.$$

Die Hypothese kann somit nicht verworfen werden, d.h. auf dem 5%-Niveau kann mit unserem Test nicht nachgewiesen werden, dass die Züge unterschiedlich schnell sind.

c) Der realisierte p-Wert ist

$$p\text{-Wert}(\omega) = P_{H_0}[|T| > 0.54] = 2(1 - P_{H_0}[T \leq 0.54]) = 2(1 - 0.7017) = 0.5966.$$

Wir könnten also die Hypothese erst verwerfen, wenn wir für die Wahrscheinlichkeit eines Fehlers 1. Art rund 60% akzeptieren würden.

4. **a)** Da wir in der Situation sind, wo eine Gruppe (von Fahrzeugen) zwei verschiedene Dinge ausprobiert (Reifen vom Typ A und B), haben wir gepaarte Daten. Unter den gemachten Annahmen und Daten ist es also naheliegend, einen gepaarten Zweistichproben-t-Test durchzuführen.

b) Sei $U_1, \ldots, U_8$ die Stichprobe, welche die Daten $u_1, \ldots, u_8$ realisiert. Es gilt, dass $U_1, \ldots, U_8$ unabhängig und $\mathcal{N}(\mu, \sigma^2)$-verteilt sind unter P_ϑ, wobei $\vartheta = (\mu, \sigma^2)$ ein unbekannter Parameter ist. Da wir testen möchten, ob sich die Bremszeit verändert hat, wählen wir als Hypothese und Alternative

$$H_0 : \mu = \mu_0 = 0 \quad \text{und} \quad H_A : \mu \neq \mu_0$$

und als Verwerfungsbereich also $K = (-\infty, -c_{\neq}) \cup (c_{\neq}, \infty)$ für ein zu bestimmendes $c_{\neq}$. Der Test ist somit zweiseitig.

Siehe nächste Seite!

c) Die Teststatistik ist

$$T = \frac{\overline{U}_8 - \mu_0}{S/\sqrt{8}}$$

wobei

$$S^2 = \frac{1}{8-1} \sum_{i=1}^{8} (U_i - \overline{U}_8)^2.$$

Unter P_{μ_0} ist $T \sim t_{8-1}$. Um auf dem 2%-Niveau zu testen, wählen wir also

$$c_{\neq} = t_{n-1,1-\frac{\alpha}{2}} = t_{7,0.99} = 2.998.$$

Es gilt $\overline{u}_8 = 2.4625$ und $s^2 = 13.9598$. Also ist der realisierte Wert der Teststatistik

$$T(\omega) = \frac{\overline{u}_8 - \mu_0}{s/\sqrt{8}} = 1.8642.$$

Wegen $|T(\omega)| < c_{\neq}$ wird die Hypothese nicht abgelehnt.

ETH Zürich FS 2021
D-INFK
Prof. Dr. Martin Schweizer

Koordinator
Marco Rodrigues

Wahrscheinlichkeit & Statistik

Musterlösung 3

1. Nach Definition ist A unabhängig von sich selbst genau dann, wenn

$$P[A \cap A] = P[A]P[A],$$

also genau für $P[A] = (P[A])^2$ wegen $A \cap A = A$. Wegen $P[A] \geq 0$ ist die letzte Bedingung aber äquivalent zu $P[A] \in \{0, 1\}$

2. **a)** Sei $Z := a + (b-a)X$. Dann ist

$$F_Z(t) = P[Z \leq t] = P\left[X \leq \frac{t-a}{b-a}\right] = \begin{cases} 0 & \text{für } \frac{t-a}{b-a} < 0, \\ \frac{t-a}{b-a} & \text{für } 0 \leq \frac{t-a}{b-a} \leq 1, \\ 1 & \text{für } \frac{t-a}{b-a} > 1. \end{cases}$$

Durch Umformen folgt, dass

$$F_Z(t) = \begin{cases} 0 & \text{für } t < a, \\ \frac{t-a}{b-a} & \text{für } a \leq t \leq b, \\ 1 & \text{für } t > b. \end{cases}$$

Also ist $Z \sim \mathcal{U}(a, b)$.

Alternative Lösung: $F_Z(t) = F_X\left(\frac{t-a}{b-a}\right)$. Weil F_X absolutstetig ist, folgt für $t \neq a, b$

$$\begin{aligned} f_Z(t) &= \frac{d}{dt} F_Z(t) \\ &= F_X'\left(\frac{t-a}{b-a}\right) \cdot \frac{1}{b-a} \\ &= \frac{1}{b-a} f_X\left(\frac{t-a}{b-a}\right) \\ &= \frac{1}{b-a} I_{[0,1]}\left(\frac{t-a}{b-a}\right) \\ &= \frac{1}{b-a} I_{[a,b]}(t), \end{aligned}$$

und das ist die Dichtefunktion einer $\mathcal{U}(a, b)$-Verteilung.

Siehe nächste Seite!

b) Die Verteilungsfunktion von cY hat die Form

$$F_{cY}(t) = P[cY \leq t] = P[Y \leq t/c] = \begin{cases} 1 - e^{-\lambda t/c} & \text{für } t/c \geq 0, \\ 0 & \text{für } t/c < 0. \end{cases}$$

Wegen $c > 0$ folgt, dass

$$F_{cY}(t) = \begin{cases} 1 - e^{-\lambda t/c} & \text{für } t \geq 0, \\ 0 & \text{für } t < 0, \end{cases}$$

also ist $cY \sim Exp(\lambda/c)$. Daraus folgt

$$E[cY] = \frac{1}{\lambda/c} = \frac{c}{\lambda}.$$

Alternativ dazu ist $E[cY] = cE[Y] = \frac{c}{\lambda}$ wegen $E[Y] = \frac{1}{\lambda}$.

3. **a)** Wir bemerken zuerst, dass $f \geq 0$. Damit f zu einer Dichtefunktion wird, muss

$$1 = \int_{-\infty}^{+\infty} f(x)dx = \int_{1}^{+\infty} cx^{-r}dx = c\left[\frac{1}{-r+1}x^{-r+1}\right]_{x=1}^{x=+\infty} = \frac{c}{r-1}$$

gelten. Mit $c = r - 1$ wird somit f zu einer Dichtefunktion.

b) Die Verteilungsfunktion von X ist

$$F_X(t) = \int_{-\infty}^{t} f_X(x)dx \quad \text{für } t \in \mathbb{R}.$$

Für $t \leq 1$ ist

$$F_X(t) = \int_{-\infty}^{t} f_X(x)dx = 0,$$

da $f_X(x) = 0$ für $x \leq 1$. Für $t > 1$ ist

$$\begin{aligned} F_X(t) &= \int_{1}^{t} f_X(x)dx \\ &= \int_{1}^{t} (r-1)x^{-r}dx \\ &= (r-1)\left[\frac{1}{-r+1}x^{-r+1}\right]_{x=1}^{x=t} \\ &= -(t^{-r+1} - 1) = 1 - t^{-r+1}. \end{aligned}$$

Die Verteilungsfunktion F_X ist demnach gegeben durch

$$F_X(t) = \begin{cases} 0 & \text{für } t \leq 1, \\ 1 - t^{-r+1} & \text{für } t > 1. \end{cases}$$

Siehe nächste Seite!

c) Der Erwartungswert von X existiert, falls

$$E[|X|] = \int_{-\infty}^{+\infty} |x| f_X(x) dx < \infty.$$

Das gilt in der Tat, denn

$$\begin{aligned}
\int_{-\infty}^{+\infty} |x| f_X(x) dx &= \int_{1}^{+\infty} |x| f_X(x) dx \\
&= \int_{1}^{+\infty} x f_X(x) dx \\
&= \int_{1}^{+\infty} (r-1) x^{-r+1} dx \\
&= (r-1) \left[\frac{1}{-r+2} x^{-r+2} \right]_{x=1}^{x=\infty} \\
&= \frac{r-1}{r-2}.
\end{aligned}$$

Wegen

$$\int_{-\infty}^{+\infty} x f_X(x) dx = \int_{1}^{+\infty} x f_X(x) dx = \int_{1}^{+\infty} |x| f_X(x) dx = \frac{r-1}{r-2}$$

ist der Erwartungswert von X also

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx = \frac{r-1}{r-2}.$$

4. Für $s, t > 0$ beliebig ist

$$P[X > t] = P[X > t + s \,|\, X > s] = \frac{P[X > t + s, X > s]}{P[X > s]} = \frac{P[X > t + s]}{P[X > s]}.$$

Sei nun $g : (0, \infty) \to \mathbb{R}$ gegeben durch $g(u) := P[X > u]$. Die Funktion g erfüllt also die Eigenschaft

$$g(t + s) = g(t) g(s) \quad \text{für alle } s, t > 0.$$

Insbesondere ist g rechtsstetig, da

$$g(t) = P[X > t] = 1 - P[X \leq t] = 1 - F_X(t) \quad \text{für alle } t > 0$$

und wir aus der Vorlesung wissen, dass Verteilungsfunktionen rechtsstetig sind. Sei $a = p/q$ eine beliebige positive rationale Zahl mit $p, q \in \mathbb{N}$. Wegen

$$g(1)^p = g(p) = g(pq/q) = g(p/q)^q$$

ist

$$g(p/q) = g(1)^{p/q} = e^{\ln(g(1)) p/q}.$$

Siehe nächste Seite!

Mit dem Hinweis gilt somit, dass

$$P[X > t] = g(t) = e^{\ln(g(1))t} \quad \text{für alle } t > 0.$$

Sei nun $\lambda := -\ln(g(1))$. Dann ist $\lambda > 0$ denn $g(1) = P[X > 1] < 1$ wegen der Annahme $P[X \leq 1] > 0$, und X hat die Verteilungsfunktion

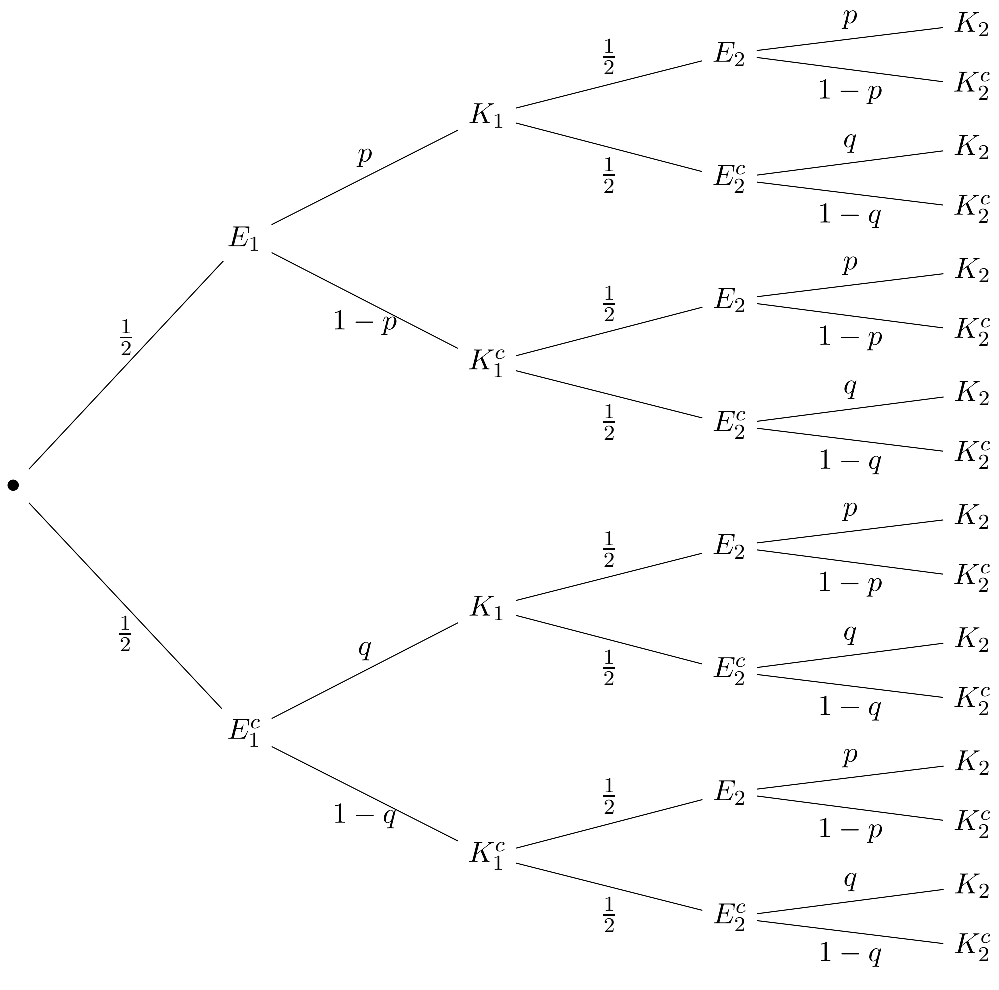
$$P[X \leq t] = \begin{cases} 1 - e^{-\lambda t} & \text{für } t > 0, \\ 0 & \text{für } t \leq 0. \end{cases}$$

Somit ist $X \sim Exp(\lambda)$ mit Parameter $\lambda > 0$.

Wahrscheinlichkeit & Statistik

Musterlösung 2

1. **a)** Seien $E_i = \{i\text{-te gezogene Münze ist } M_1\}$ und $K_i = \{i\text{-ter Wurf ist Kopf}\}$ für $i = 1, 2$. Der Ereignisbaum sieht dann folgendermassen aus:



b) Wir wollen die Wahrscheinlichkeit $P[K_1 \cap K_2]$ berechnen. Da wir das gleiche Experiment zweimal hintereinander ausführen, sind einerseits die Ereignisse K_1 und

Siehe nächste Seite!

K_2 unabhängig, und andererseits gilt $P[K_1] = P[K_2]$. Also ist

$$P[K_1 \cap K_2] = P[K_1]^2.$$

Mit dem Satz der totalen Wahrscheinlichkeit erhalten wir

$$P[K_1] = P[K_1|E_1]P[E_1] + P[K_1|E_1^c]P[E_1^c] = p \cdot \frac{1}{2} + q \cdot \frac{1}{2} = \frac{p+q}{2},$$

und demnach ist

$$P[K_1 \cap K_2] = P[K_1]^2 = \frac{(p+q)^2}{4}.$$

Alternative Lösung: Für jeden Pfad der K_1 und K_2 enthält, berechnen wir seine Wahrscheinlichkeit durch Multiplizieren der sukzessiven Übergangswahrsch. und summieren dann über alle Pfade. Das liefert

$$P[K_1 \cap K_2] = \frac{1}{4}p^2 + \frac{1}{4}pq + \frac{1}{4}pq + \frac{1}{4}q^2 = \frac{(p+q)^2}{4}.$$

Analog kann man zeigen, dass

$$P[K_1] = \frac{p+q}{2} = P[K_2].$$

Somit haben wir auch die Unabhängigkeit von K_1 und K_2 bewiesen.

2. Sei $K_i = \{\text{Person } i \text{ ist infiziert}\}$ und $N_i = \{\text{Person } i \text{ wird vom Arzt genannt}\}$ für $i = A, B, C$. Da in Wirklichkeit nur eine Person infiziert ist und wir keine weiteren Informationen haben, ist

$$P[K_A] = P[K_B] = P[K_C] = \frac{1}{3}.$$

Nehmen wir nun an, dass der Arzt die Auskunft erteilt. Dann gilt

- $P[N_B|K_A] = \frac{1}{2}$ (nach Annahme)
- $P[N_B|K_B] = 0$ (der Arzt soll A eine gesunde Person nennen)
- $P[N_B|K_C] = 1$ (der Arzt soll A eine gesunde Person nennen)

und

- $P[N_C|K_A] = \frac{1}{2}$ (nach Annahme)
- $P[N_C|K_B] = 1$ (der Arzt soll A eine gesunde Person nennen)
- $P[N_C|K_C] = 0$ (der Arzt soll A eine gesunde Person nennen)

Siehe nächste Seite!

a) Mit dem Satz der totalen Wahrscheinlichkeit folgt, dass

$$\begin{aligned} P[N_B] &= P[N_B|K_A] \cdot P[K_A] + P[N_B|K_B] \cdot P[K_B] + P[N_B|K_C] \cdot P[K_C] \\ &= \frac{1}{2} \cdot \frac{1}{3} + 0 + 1 \cdot \frac{1}{3} = \frac{1}{2} \end{aligned}$$

und

$$\begin{aligned} P[N_C] &= P[N_C|K_A] \cdot P[K_A] + P[N_C|K_B] \cdot P[K_B] + P[N_C|K_C] \cdot P[K_C] \\ &= \frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 0 = \frac{1}{2}. \end{aligned}$$

Dieses Ergebnis ist auch schon rein wegen der Symmetrie der Situation plausibel.

b) Mit dem Satz von Bayes ist

$$P[K_A|N_B] = \frac{P[N_B|K_A] \cdot P[K_A]}{P[N_B]} = \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

und

$$P[K_A|N_C] = \frac{P[N_C|K_A] \cdot P[K_A]}{P[N_C]} = \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}.$$

Die Aussage des Arztes stimmt also nicht; die bedingte Wahrscheinlichkeit bleibt auch nach der Auskunft bei $\frac{1}{3}$.

3. Wir zeigen induktiv, dass die Inklusions- und Exklusionsformel gilt. Die Induktionsverankerung bei $n = 2$ ist gerade die Additionsregel. Nehmen wir also an, dass die Formel für eine natürliche Zahl $n \geq 2$ gilt. Seien $A_1, \ldots, A_{n+1}$ Elemente in $\mathcal{F}$. Mit der Additionsregel folgt

$$\begin{aligned} P\left[\bigcup_{i=1}^{n+1} A_i\right] &= P\left[\left(\bigcup_{i=1}^{n} A_i\right) \cup A_{n+1}\right] \\ &= P[A_1 \cup \cdots \cup A_n] + P[A_{n+1}] - P[(A_1 \cup \cdots \cup A_n) \cap A_{n+1}], \end{aligned} \tag{1}$$

und mit der Induktionsvoraussetzung erhalten wir

$$\begin{aligned} P\left[A_1 \cup \cdots \cup A_n\right] &= \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} P\left[\bigcap_{j=1}^{k} A_{i_j}\right] \\ &= \sum_{i=1}^{n} P[A_i] + \sum_{k=2}^{n} (-1)^{k+1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} P\left[\bigcap_{j=1}^{k} A_{i_j}\right] \end{aligned} \tag{2}$$

und

$$\begin{aligned} P[(A_1 \cup \cdots \cup A_n) \cap A_{n+1}] &= P[(A_1 \cap A_{n+1}) \cup \cdots \cup (A_n \cap A_{n+1})] \\ &= \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} P\left[\bigcap_{j=1}^{k} (A_{i_j} \cap A_{n+1})\right] \\ &= -\sum_{k'=2}^{n+1} (-1)^{k'+1} \sum_{1 \leq i_1 < \cdots < i_{k'} = n+1} P\left[\bigcap_{j=1}^{k'} A_{i_j}\right]. \end{aligned} \tag{3}$$

Siehe nächste Seite!

Die letzte Gleichung folgt durch Umbenennung der Laufvariable $k =: k' - 1$. Nun benutzen wir (1), setzen dort (2) und (3) ein, und erhalten so

$$\begin{aligned}
P\left[\bigcup_{i=1}^{n+1} A_i\right] &= \sum_{i=1}^{n} P[A_i] + \sum_{k=2}^{n} (-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n} P\left[\bigcap_{j=1}^{k} A_{i_j}\right] \\
&\quad + P[A_{n+1}] + \sum_{k'=2}^{n+1} (-1)^{k'+1} \sum_{1 \le i_1 < \cdots < i_{k'} = n+1} P\left[\bigcap_{j=1}^{k'} A_{i_j}\right] \\
&= \sum_{i=1}^{n+1} P[A_i] + \sum_{k=2}^{n+1} (-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n+1} P\left[\bigcap_{j=1}^{k} A_{i_j}\right] \\
&= \sum_{k=1}^{n+1} (-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n+1} P\left[\bigcap_{j=1}^{k} A_{i_j}\right].
\end{aligned}$$

4. a) Das Zurückgeben der Hüte kann man sich als eine Permutation von n Plätzen vorstellen. Darum ist es naheliegend, als Grundraum die Menge S_n aller Permutationen von $\{1, \dots, n\}$ zu wählen. Aufgrund von Symmetrie-Erwägungen gehen wir davon aus, dass alle Elementarereignisse gleichwahrscheinlich sind; also ist

$$P[\{\sigma\}] = \frac{1}{n!} \quad \text{für alle } \sigma \in S_n.$$

b) Betrachte das Ereignis

$$\begin{aligned}
A_i &= \{\sigma \in S_n : \sigma(i) = i\} \\
&= \{\text{die } i\text{-te Person bekommt ihren Hut zurück}\}.
\end{aligned}$$

Das Ereignis, dass niemand seinen Hut zurückbekommt, ist $A^{(n)} := (A_1 \cup \ldots \cup A_n)^c$. Mit der Inklusions- und Exklusionsformel erhalten wir

$$\begin{aligned}
P_n := P[A^{(n)}] &= 1 - P[A_1 \cup \cdots \cup A_n] \\
&= 1 - \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n} P[A_{i_1} \cap \cdots \cap A_{i_k}].
\end{aligned}$$

Nun gilt

$$\begin{aligned}
P[A_{i_1} \cap \cdots \cap A_{i_k}] &= \frac{\text{Anzahl der Permutationen, die } i_1, \dots, i_k \text{ unverändert lassen}}{n!} \\
&= \frac{(n-k)!}{n!},
\end{aligned}$$

und man hat $\binom{n}{k}$ verschiedene Möglichkeiten, die Menge $\{i_1, \dots, i_k\}$ aus $\{1, \dots, n\}$ auszuwählen. Also ist

$$P[A^{(n)}] = 1 - \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} \frac{(n-k)!}{n!} = 1 - 1 + \frac{1}{2!} + \cdots + (-1)^n \frac{1}{n!}.$$

Siehe nächste Seite!

c) Wir wählen zuerst eine Gruppe von k Personen aus den n Personen aus. Dazu haben wir $\binom{n}{k}$ Möglichkeiten. Die Wahrscheinlichkeit, dass diese k Personen ihren Hut zurückbekommen, ist $\frac{(n-k)!}{n!}$. Um die Wahrscheinlichkeit zu erhalten, dass *genau* diese k Personen ihren Hut zurückbekommen, müssen wir mit der Wahrscheinlichkeit multiplizieren, dass von den restlichen $n-k$ Personen niemand seinen Hut zurückbekommt, also mit P_{n-k} (siehe oben). Die Wahrscheinlichkeit, dass genau k Personen ihren Hut zurückbekommen, ist also

$$\begin{aligned} Q_n^k &:= \binom{n}{k} \frac{(n-k)!}{n!} P_{n-k} \\ &= \frac{1}{k!}\left(1 - 1 + \frac{1}{2!} + \ldots + (-1)^{n-k} \frac{1}{(n-k)!}\right) \\ &= \frac{1}{k!} \sum_{\ell=0}^{n-k} \frac{(-1)^\ell}{\ell!}. \end{aligned}$$

d) Aus der vorherigen Teilaufgabe ist ersichtlich, dass

$$\begin{aligned} \lim_{n\to\infty} k! Q_n^k &= \lim_{n\to\infty} \sum_{\ell=0}^{n-k} \frac{(-1)^\ell}{\ell!} \\ &= \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{\ell!} \\ &= \exp(-1), \end{aligned}$$

und deshalb ist

$$\lim_{n\to\infty} Q_n^k = \frac{\exp(-1)}{k!}.$$

Wahrscheinlichkeit & Statistik

Musterlösung 12

1. a) Aus der Vorlesung ist bekannt, dass bei Unabhängigkeit von X und Y die Dichte der Summe $X+Y$ die Faltung von f_X mit f_Y ist. Also ist für $z \geq 0$

$$\begin{aligned} f_{X+Y}(z) &= f_X * f_Y(z) \\ &= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx \\ &= \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^z x^{\alpha-1}(z-x)^{\beta-1} e^{-\lambda x} e^{-\lambda(z-x)} dx \\ &= \frac{\lambda^{\alpha+\beta} e^{-\lambda z}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^z x^{\alpha-1}(z-x)^{\beta-1} dx \\ &= \frac{\lambda^{\alpha+\beta} z^{\alpha+\beta-2} e^{-\lambda z}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} z dt \\ &= \frac{1}{\Gamma(\alpha+\beta)} \lambda^{\alpha+\beta} z^{\alpha+\beta-1} e^{-\lambda z} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt, \end{aligned}$$

wobei wir in der fünften Gleichheit die Variablensubstitution $t = x/z$, $dx = zdt$ benutzt haben. Andererseits ist für $z < 0$ klar, dass $f_{X+Z}(z) = 0$. Also ist die Dichte von $X+Y$ gegeben durch die Dichte einer $Ga(\alpha+\beta, \lambda)$-verteilten Zufallsvariablen multipliziert mit einem konstanten Faktor. Da aber die Faltung eine Dichte ist, muss der konstante Faktor

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt = 1$$

sein. Also ist $X+Y \sim Ga(\alpha+\beta, \lambda)$.

b) Seien $X_1, \ldots, X_n$ unabhängig und je $Exp(\lambda)$-verteilt mit $\lambda > 0$. Wegen $\Gamma(1) = 1$ folgt $X_i \sim Ge(1, \lambda)$ für $i = 1, \ldots, n$. Aus **a)** können wir also induktiv schliessen, dass $\sum_{i=1}^n X_i \sim Ge(n, \lambda)$.

2. Da die Varianz $\sigma^2 = 9$ bekannt ist, liegt es nahe, in diesem Fall einen (zweiseitigen) z-Test durchzuführen. Wir möchten also die Hypothese $H_0 : \mu = \mu_0 = 146$ gegen die

Siehe nächste Seite!

Alternative $H_A : \mu \neq \mu_0$ testen und dabei die Teststatistik

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

verwenden. Unter P_{μ_0} ist $T \sim \mathcal{N}(0,1)$, und den kritischen Bereich wählen wir von der Form $K_{\neq} = (-\infty, c_{\neq}) \cup (c_{\neq}, \infty)$. Wir verwerfen also H_0, falls $|T| > c_{\neq}$ für ein zu bestimmendes $c_{\neq}$ ist. Da wir auf dem 5%-Niveau testen möchten, wählen wir $c_{\neq}$ so, dass

$$0.05 = \alpha = P_{\mu_0}[T \in K_{\neq}] = 2\big(1 - \Phi(c_{\neq})\big)$$

gilt. Das ergibt den Wert $c_{\neq} = \Phi^{-1}(1 - \alpha/2) = z_{0.975} = 1.96$. Der realisierte Wert der Teststatistik T ist

$$T(\omega) = t(x_1, \ldots, x_n) = -2.67.$$

Wegen $|T(\omega)| > 1.96$ verwerfen wir also die Hypothese, dass die mittlere Fahrzeit des Cisalpino von jener des Intercity nicht abweicht. Würden wir stattdessen die Alternative $H'_A : \mu < \mu_0$ testen, so wäre $K_< = (-\infty, c_<)$ mit $c_< = z_{0.05} = -z_{0.95} = -1.645$. Wegen $T(\omega) \in K_<$ wird auch hier die Hypothese verworfen; die Daten weisen also darauf hin (aber sie beweisen nicht), dass der Cisalpino im Mittel eine kürzere Fahrzeit hat.

3. **a)** Für $i = 1, \ldots, 6$ sei X_i die Zugriffszeit auf die Festplatte (in ms). Nach der Aufgabenstellung sind $X_1, \ldots, X_n$ unabhängig und je $\mathcal{N}(\mu, \sigma^2)$-verteilt unter P_ϑ, wobei $\vartheta = (\mu, \sigma^2)$ ein unbekannter Parameter ist. Da Sarah vermutet, dass sich die Zugriffszeit verändert hat, wählt sie als Hypothese und Alternative

$$H_0 : \mu = \mu_0 = 5.8 \quad \text{und} \quad H_A : \mu \neq 5.8.$$

Unter den obigen Annahmen liegt es nun nahe, einen t-Test durchzuführen, also als Teststatistik für den Erwartungswert

$$T := \frac{\overline{X}_6 - 5.8}{S_6/\sqrt{6}}$$

zu wählen. Unter P_{ϑ_0} folgt T einer t-Verteilung mit 5 Freiheitsgraden. Nach der Alternative hat der kritische Bereich die Form $K_{\neq} = (-\infty, c_{\neq}) \cup (c_{\neq}, \infty)$, d.h. die Hypothese wird verworfen, falls $|T| > c_{\neq}$ für ein zu bestimmendes $c_{\neq}$ ist. Für $\alpha = 0.01$ wählen wir $c_{\neq} = t_{5,1-\frac{\alpha}{2}} = 4.032$. Da nun $|T(\omega)| = 2.12 \leq c_{\neq}$ gilt, können wir die Hypothese nicht verwerfen.

b) Mit der geänderten fünften Beobachtung erhalten wir $|T(\omega)| = 4.648 > c_{\neq}$, d.h. hier wird die Hypothese verworfen.

4. **a)** Als Modell für die Anzahl S_n falsch eingeordneter Bücher eignet sich die Binomialverteilung. Wir nehmen also an, dass $S_n \sim Bin(n, \vartheta)$ mit $n = 100$ und unbekanntem Parameter $\vartheta \in \Theta = [0, 1]$.

Siehe nächste Seite!

i) Wir wählen $T := S_n$ als plausible Teststatistik. Als Hypothese und Alternative wählen wir

$$H_0 : \vartheta = \vartheta_0 = 0.03 \quad \text{und} \quad H_A : \vartheta < \vartheta_0.$$

ii) Mit Hilfe des Hinweises erhalten wir

$$P_{\vartheta_0}[T = 0] = (1 - 0.03)^{100} = 0.97^{100} = 0.0475 < 0.05$$

und

$$\begin{aligned} P_{\vartheta_0}[T \leq 1] &= P_{\vartheta_0}[T = 0] + P_{\vartheta_0}[T = 1] \\ &= 0.97^{100} + 100 \cdot 0.03 \cdot 0.97^{100} = 0.1882 > 0.05. \end{aligned}$$

Also ist der Verwerfungsbereich $K = (-\infty, 1)$.

iii) Die Hypothese wird nicht verworfen, da ein falsch eingeordnetes Buch gefunden wurde. Die Bestandsaufnahme wird also nicht verschoben.

b) Die Bestandaufnahme wird unnötigerweise verfrüht durchgeführt, falls die Hypothese angenommen wird, obwohl H_A: $\vartheta = \vartheta_A = 0.01$ richtig ist. Wir berechnen deshalb die Wahrscheinlichkeit eines Fehlers 2. Art als

$$P_{\vartheta_A}[T > 0] = 1 - P_{\vartheta_A}[T = 0] = 1 - (0.99)^{100} = 0.634,$$

d.h. die Bestandsaufnahme wird mit einer Wahrscheinlichkeit von 63.4% unnötig durchgeführt. Die Macht an der Stelle ϑ_A ist die Wahrscheinlichkeit, H_0 zu verwerfen unter der Annahme, dass H_A korrekt ist, und beträgt $\beta(\vartheta_A) = 1 - 0.634 = 0.3660$.

c) i) Als Hypothese und Alternative wählen wir wieder

$$H_0 : \vartheta = \vartheta_0 = 0.03 \quad \text{und} \quad H_A : \vartheta < \vartheta_0,$$

und als Tesetstatistik wählen wir nun

$$T = \frac{S_n/n - \vartheta_0}{\sqrt{\vartheta_0(1 - \vartheta_0)/n}}$$

mit $n = 1000$, da T dann nach dem zentralen Grenzwertsatz approximativ standardnormalverteilt ist. Als approximativen Test benutzen wir hier also einen einseitigen z-Test.

ii) Die Hypothese wird verworfen, wenn der realisierte Wert

$$T(\omega) < -z_{1-\alpha} = -1.645$$

ist, wobei $\alpha = 0.05$ ist.

iii) In unserem Fall ist der realisierte Wert

$$T(\omega) = t(x_1, \ldots, x_{1000}) = \frac{0.01 - 0.03}{\sqrt{0.03 \cdot 0.97/1000}} = -3.71 < -1.645$$

(bzw. mit dem Hinweis $\sqrt{\frac{0.03 \cdot 0.97}{1000}} \approx 0.005$ ist $T(\omega) \approx -4$). Die Hypothese wird also verworfen und die Bestandsaufnahme kann verschoben werden.

ETH Zürich FS 2021
D-INFK
Prof. Dr. Martin Schweizer

Koordinator
Marco Rodrigues

Wahrscheinlichkeit & Statistik

Musterlösung 10

1. Wegen $E[R_i] = 0$ und $\sigma^2 := \mathrm{Var}[R_i] = 1/12$ folgt mit dem zentralen Grenzwertsatz, dass approximativ gilt

$$\begin{aligned} P\left[\left|\sum_{i=1}^{1200} R_i\right| \leq 20\right] &= P\left[\left|\frac{\sum_{i=1}^{1200} R_i}{\sigma\sqrt{1200}}\right| \leq \frac{20}{\sigma\sqrt{1200}}\right] \\ &\approx \Phi\left(\frac{20}{\sigma\sqrt{1200}}\right) - \Phi\left(-\frac{20}{\sigma\sqrt{1200}}\right) \\ &= \Phi(2) - \Phi(-2) \\ &= 2\Phi(2) - 1 = 0.9534. \end{aligned}$$

2. Es gilt

$$\begin{aligned} S_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}\big((X_i-\mu) - (\overline{X}_n - \mu)\big)^2 \\ &= \frac{1}{n-1}\left(\sum_{i=1}^{n}(X_i-\mu)^2 - n(\overline{X}_n-\mu)^2\right) \\ &= \frac{n}{n-1}\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 - \frac{n}{n-1}(\overline{X}_n-\mu)^2. \end{aligned}$$

Mit dem starken Gesetz der grossen Zahlen folgt mit $\mu := E[X_1]$, dass

$$\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 \overset{n\to\infty}{\longrightarrow} E[(X_1-\mu)^2] = \sigma^2 \quad P\text{-fastsicher} \tag{1}$$

und

$$\overline{X}_n - \mu \overset{n\to\infty}{\longrightarrow} 0 \quad P\text{-fastsicher.}$$

Wegen $n/(n-1) \to 1$ für $n \to \infty$ folgt also, dass

$$S_n^2 \overset{n\to\infty}{\longrightarrow} \sigma^2 \quad P\text{-fastsicher.}$$

Siehe nächste Seite!

3. **a)** Damit $f_X(x;\vartheta_1,\vartheta_2)$ eine Dichte ist, muss

$$1=\int_{-\infty}^{\infty} f_X(x;\vartheta_1,\vartheta_2)=\int_{\vartheta_2}^{\infty} cx^{-\vartheta_1}dx=c\left[\frac{x^{-\vartheta_1+1}}{-\vartheta_1+1}\right]_{\vartheta_2}^{\infty}=c\frac{\vartheta_2^{-\vartheta_1+1}}{\vartheta_1-1}$$

gelten. Somit erhalten wir $c=(\vartheta_1-1)\vartheta_2^{\vartheta_1-1}$.

b) Es gilt

$$\begin{aligned} E_\vartheta[X^s]=\int_{-\infty}^{\infty} x^s f_X(x;\vartheta_1,\vartheta_2)dx &= \int_{\vartheta_2}^{\infty}(\vartheta_1-1)\vartheta_2^{\vartheta_1-1}x^{s-\vartheta_1}dx \\ &= (\vartheta_1-1)\vartheta_2^{\vartheta_1-1}\left[\frac{x^{s-\vartheta_1+1}}{s-\vartheta_1+1}\right]_{\vartheta_2}^{\infty}=\frac{(\vartheta_1-1)\vartheta_2^s}{\vartheta_1-1-s}\end{aligned}$$

für $\vartheta_1>s+1$. Insbesondere ist $E_\vartheta[X]=\frac{(\vartheta_1-1)\vartheta_2}{\vartheta_1-2}$ für $\vartheta_1>2$ und $E_\vartheta[X^2]=\frac{(\vartheta_1-1)\vartheta_2^2}{\vartheta_1-3}$ für $\vartheta_1>3$. Daraus folgt, dass

$$\mathrm{Var}_\vartheta[X]=E_\vartheta[X^2]-(E_\vartheta[X])^2=\frac{(\vartheta_1-1)\vartheta_2^2}{(\vartheta_1-3)(\vartheta_1-2)^2}$$

für $\vartheta_1>3$.

c) Wir wählen ϑ_1 so, dass das s-te Moment $m_s(\vartheta)=E_\vartheta[X^s]$ von X unter P_ϑ und das s-te empirische Mittel $\tilde m_s(x_1,\dots,x_n)=\frac{1}{n}\sum_{i=1}^n x_i^s$ von $x_1,\dots,x_n$ übereinstimmen. Für $s=1$ gilt

$$\begin{aligned} m_1(\vartheta_1)=\tilde m_1(x_1,\dots,x_n) &\Leftrightarrow \frac{(\vartheta_1-1)\vartheta_2}{\vartheta_1-2}=\tilde m_1(x_1,\dots,x_n)\\ &\Leftrightarrow \vartheta_1=\frac{\vartheta_2-2\tilde m_1(x_1,\dots,x_n)}{\vartheta_2-\tilde m_1(x_1,\dots,x_n)}=1-\frac{\tilde m_1(x_1,\dots,x_n)}{\vartheta_2-\tilde m_1(x_1,\dots,x_n)}\\ &\Leftrightarrow \vartheta_1=1-\frac{\frac{1}{n}\sum_{i=1}^n x_i}{\vartheta_2-\frac{1}{n}\sum_{i=1}^n x_i}.\end{aligned}$$

Dies führt zum Momentenschätzer

$$T_{\mathrm{MM}}=1-\frac{\frac{1}{n}\sum_{i=1}^n X_i}{\vartheta_2-\frac{1}{n}\sum_{i=1}^n X_i}=1-\frac{\bar X_n}{\vartheta_2-\bar X_n}.$$

Für den Maximum-Likelihood-Schätzer gehen wir folgendermassen vor: Da die Beobachtungen unabhängig sind, können wir die Likelihood-Funktion schreiben als $L(x_1,\dots,x_n;\vartheta_1,\vartheta_2)=\prod_{i=1}^n f_X(x_i;\vartheta_1,\vartheta_2)$. Die log-Likelihood-Funktion $\ell:=\log L$

Siehe nächste Seite!

ist dann

$$\begin{aligned}
\ell(x_1, \ldots, x_n; \vartheta_1, \vartheta_2) &= \log\left(\prod_{i=1}^n f_X(x_i; \vartheta_1, \vartheta_2)\right) \\
&= \sum_{i=1}^n \log f_X(x_i; \vartheta_1, \vartheta_2) \\
&= \sum_{i=1}^n \log\left((\vartheta_1 - 1)\vartheta_2^{\vartheta_1 - 1} x_i^{-\vartheta_1}\right) \\
&= n\log(\vartheta_1 - 1) + n(\vartheta_1 - 1)\log\vartheta_2 - \vartheta_1 \sum_{i=1}^n \log x_i \,.
\end{aligned}$$

Ableiten nach ϑ_1 und Nullsetzen der log-Likelihood-Funktion ergibt

$$\frac{\partial \ell(x_1, \ldots, x_n; \vartheta_1, \vartheta_2)}{\partial \vartheta_1} = \frac{n}{\vartheta_1 - 1} + n\log\vartheta_2 - \sum_{i=1}^n \log x_i = \frac{n}{\vartheta_1 - 1} + \sum_{i=1}^n \log\frac{\vartheta_2}{x_i} = 0$$

für $\vartheta_1 = 1 - \frac{n}{\sum_{i=1}^n \log\frac{\vartheta_2}{x_i}}$. Da die zweite Abteilung

$$\frac{\partial^2 \ell(x_1, \ldots, x_n; \vartheta_1, \vartheta_2)}{\partial \vartheta_1^2} = -\frac{n}{(\vartheta_1 - 1)^2}$$

negativ ist, schliessen wir, dass ϑ_1 ein Maximum ist. Der Maximum-Likelihood-Schätzer von ϑ_1 ist also

$$T_{\mathrm{ML}}^1 = 1 - \frac{n}{\sum_{i=1}^n \log\frac{\vartheta_2}{X_i}} = 1 + \frac{1}{\overline{(\log X)}_n - \log\vartheta_2}\,,$$

wobei $\overline{(\log X)}_n = \frac{1}{n}\sum_{i=1}^n \log X_i$.

d) Wir nehmen an, dass die N strategischen Ölreserven unabhängig und identisch verteilt sind. Eine zweite Annahme ist, dass N gross genug ist, damit die Normalverteilung via den zentralen Grenzwertsatz eine gute Approximation liefert. Laut dem zentralen Grenzwertsatz ist S_N approximativ $\mathcal{N}(N\mu, N\sigma^2)$-verteilt für $\mu := E_\vartheta[X_1] = \frac{(\vartheta_1-1)\vartheta_2}{\vartheta_1-2}$ und $\sigma^2 := \mathrm{Var}_\vartheta[X_1] = \frac{(\vartheta_1-1)\vartheta_2^2}{(\vartheta_1-3)(\vartheta_1-2)^2}$. Die Wahrscheinlichkeit, dass das Gesamtniveau der N strategischen Ölreserven die Mindestgrenze δ unterschreitet, ist approximativ gegeben durch

$$\begin{aligned}
P_\vartheta[S_N \le \delta] &= P\left[\frac{S_N - N\mu}{\sqrt{N}\sigma} \le \frac{\delta - N\mu}{\sqrt{N}\sigma}\right] \approx P\left[Z \le \frac{\delta - N\mu}{\sqrt{N}\sigma}\right] \\
&= \Phi\left(\frac{\delta - N\mu}{\sqrt{N}\sigma}\right) = \Phi\left(\frac{\delta - N\frac{(\vartheta_1-1)\vartheta_2}{\vartheta_1-2}}{\sqrt{N\frac{(\vartheta_1-1)\vartheta_2^2}{(\vartheta_1-3)(\vartheta_1-2)^2}}}\right).
\end{aligned}$$

4. **a)** Aus den Annahmen folgt, dass $X_1, \ldots, X_{10}$ unabhängig und jeweils Bernoulli-verteilt sind mit unbekanntem Erfolgsparameter $\vartheta := P[X_1 = 1] \in \Theta = [0, 1]$.

Siehe nächste Seite!

b) Unsere Hypothese ist, dass der Würfel nicht gezinkt ist; also ist

$$H_0 : \vartheta = \frac{1}{6}, \quad \text{d.h. } \Theta_0 = \left\{\frac{1}{6}\right\}.$$

Die Alternative, dass der Würfel gezinkt ist, ist dann

$$H_A : \vartheta > 1/6, \quad \text{d.h. } \Theta_A = \left(\frac{1}{6}, 1\right].$$

c) Da $X_1, \ldots, X_{10}$ unabhängig und $Be(\vartheta)$-verteilt sind unter P_ϑ, ist $T \sim Bin(10, \vartheta)$ unter P_ϑ.

d) Die Wahrscheinlichkeit für einen Fehler 1. Art ist

$$\begin{aligned} P_{\frac{1}{6}}[T \in K] &= P_{\frac{1}{6}}[T \geq 5] \\ &= \sum_{k=5}^{10} \binom{10}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{10-k} \\ &= \frac{1}{6^{10}} \Bigg(\binom{10}{5} 5^5 + \binom{10}{6} 5^4 + \binom{10}{7} 5^3 \\ &\qquad + \binom{10}{8} 5^2 + \binom{10}{9} 5^1 + \binom{10}{10} 5^0 \Bigg) \\ &= 0.0155. \end{aligned}$$

e) Da $t(x_1, \ldots, x_{10}) = t(X_1(\omega), \ldots, X_{10}(\omega)) = T(\omega) = 4$ nicht im Verwerfungsbereich liegt, verwerfen wir die Hypothese nicht.

Wahrscheinlichkeit & Statistik

Musterlösung 1

1. a) $A \cup B =$ „Federer gewinnt den ersten Satz oder gewinnt den Match."
$A^c \cap B =$ „Federer verliert den ersten Satz und gewinnt den Match."
$A \cap B^c = A \setminus B =$ „Federer gewinnt den ersten Satz und verliert den Match."

Zur Berechnung der bedingten Wahrscheinlichkeiten führen wir zusätzlich die Ereignisse $S_i =$ „Federer gewinnt den i-ten Satz", $i = 1, 2, 3$, ein. Aus der Aufgabenstellung geht hervor, dass $P[S_i] = \frac{1}{3}$, $i = 1, 2, 3$, und dass S_1, S_2, S_3 unabhängig sind. Somit gilt

$$\begin{aligned}
P[B^c \mid A] &= \frac{P[B^c \cap A]}{P[A]} = \frac{P[S_1 \cap S_2^c \cap S_3^c]}{P[S_1]} \overset{\text{Unabh.}}{=} \frac{P[S_1]P[S_2^c]P[S_3^c]}{P[S_1]} \\
&= P[S_2^c]P[S_3^c] = \frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9}, \\
P[B \mid A] &= 1 - P[B^c \mid A] = \frac{5}{9}, \\
P[B \mid A^c] &= \frac{P[B \cap A^c]}{P[A^c]} = \frac{P[S_1^c \cap S_2 \cap S_3]}{P[S_1^c]} \overset{\text{Unabh.}}{=} P[S_2]P[S_3] = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}.
\end{aligned}$$

b) Wir berechnen $P[B]$ mit dem Satz der totalen Wahrscheinlichkeit:

$$P[B] = P[B|A]\,P[A] + P[B|A^c]\,P[A^c] = \frac{5}{9} \cdot \frac{1}{3} + \frac{1}{9} \cdot \frac{2}{3} = \frac{7}{27}.$$

Alternativ können wir B schreiben als

$$B = (S_1 \cap S_2) \,\dot\cup\, (S_1 \cap S_2^c \cap S_3) \,\dot\cup\, (S_1^c \cap S_2 \cap S_3)$$

und erhalten dann mit Unabhängigkeit

$$\begin{aligned}
P[B] &= P[S_1 \cap S_2] + P[S_1 \cap S_2^c \cap S_3] + P[S_1^c \cap S_2 \cap S_3] \\
&= P[S_1]P[S_2] + P[S_1]P[S_2^c]P[S_3] + P[S_1^c]P[S_2]P[S_3] \\
&= \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} + \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \\
&= \frac{7}{27}.
\end{aligned}$$

Siehe nächste Seite!

c) $A|B$ = „Federer hat den ersten Satz gewonnen, gegeben, dass er den Match gewonnen hat.“
$A|B^c$ = „Federer hat den ersten Satz gewonnen, gegeben, dass er den Match verloren hat.“
$P[A|B]$ und $P[A|B^c]$ berechnen wir mit dem Satz von Bayes:

$$P[A|B] = \frac{P[B|A]\,P[A]}{P[B]} = \frac{\frac{5}{9}\cdot\frac{1}{3}}{\frac{7}{27}} = \frac{\frac{5}{27}}{\frac{7}{27}} = \frac{5}{7}$$

und

$$P[A|B^c] = \frac{P[B^c|A]\,P[A]}{P[B^c]} = \frac{\frac{4}{9}\cdot\frac{1}{3}}{\frac{20}{27}} = \frac{\frac{4}{27}}{\frac{20}{27}} = \frac{1}{5}.$$

2. **a)** Nehmen wir an, ein Würfel sei rot, der andere schwarz. Falls mit dem roten und dem schwarzen Würfel je eine Eins geworfen wird, haben wir das Ereignis “zwei Einsen”. Das Ereignis “eine Eins und eine Zwei” kann auf zwei Arten erzeugt werden: es kann sein, dass mit dem roten Würfel eine Eins geworfen wird und mit dem schwarzen eine Zwei, oder umgekehrt. Um diese Ereignisse formal zu erfassen, benutzen wir Klammern. Der erste Eintrag ist das Ergebnis des roten Würfels und der zweite das Ergebnis des schwarzen. Für die obigen Ereignisse haben wir also: $(1,1),(1,2)$ und $(2,1)$. Mit diesen Überlegungen erhalten wir den Grundraum

$$\Omega = \{(1,1),(1,2),\ldots,(1,6),(2,1),(2,2),\ldots,(2,6),(3,1)\ldots,(6,6)\}.$$

Der Grundraum Ω besteht aus 36 Elementarereignissen.

b) Die Zufallsvariable S ist die Augensumme der Würfel. Ihr Wertebereich ist also gegeben durch: $\mathcal{W}(S) = \{2,3,4,\ldots,11,12\}$.

c) In der folgenden Tabelle ist die Wahrscheinlichkeits-Verteilung von S beschrieben:

k	2	3	4	5	6	7	8	9	10	11	12
$P[S=k]$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Um das zu erhalten, nehmen wir an, dass alle 36 Elementarereignisse gleich wahrscheinlich sind. Dann ist

$$P[S=k] = \frac{1}{36}\#\big(\{\omega\in\Omega : S(\omega)=k\}\big),$$

und wir müssen nur noch die Menge $\{S=k\}$ abzählen.

3. Sei $p := P[X=1]$. Wir zeigen induktiv, dass

$$P[X=\ell] = p(1-p)^{\ell-1} \quad \text{für alle } \ell\in\mathbb{N}, \tag{1}$$

also dass X geometrisch verteilt ist mit Parameter p.

Siehe nächste Seite!

Für $\ell = 1$ folgt obige Gleichung durch die Definition von p. Nehmen wir also an, dass die obige Gleichung bis und mit zu einem $\ell \in \mathbb{N}$ stimmt. Wegen

$$P[X = \ell + 1, X \leq 1] = 0$$

gilt

$$\begin{aligned} P[X = \ell + 1, X > 1] &= P[X = \ell + 1] - P[X = \ell + 1, X \leq 1] \\ &= P[X = \ell + 1]. \end{aligned}$$

Somit ist

$$\begin{aligned} P[X = \ell + 1 | X > 1] &= \frac{P[X = \ell + 1]}{P[X > 1]} \\ &= \frac{P[X = \ell + 1]}{1 - P[X \leq 1]} \\ &= \frac{P[X = \ell + 1]}{1 - P[X = 1]} \\ &= \frac{P[X = \ell + 1]}{1 - p}. \end{aligned}$$

Da nach Induktionsvoraussetzung $P[X = \ell] = p(1-p)^{\ell-1}$ gilt, folgt mit der Annahme in der Aufgabenstellung, dass

$$\begin{aligned} P[X = \ell + 1] &= (1-p)P[X = \ell + 1 | X > 1] \\ &= (1-p)P[X = \ell] \\ &= p(1-p)^{\ell}. \end{aligned}$$

Dies zeigt, dass (1) gilt, und somit ist $X \sim Geom(p)$.

4. Wir zeigen zuerst, dass die Implikation von (i) nach (ii) gilt. Sei n also eine Primzahl und seien A, B zwei unabhängige Teilmengen von Ω. Dann gilt

$$\frac{|A \cap B|}{n} = \frac{|A|}{n} \cdot \frac{|B|}{n},$$

wobei $|\cdot|$ die Anzahl Elemente in einer endlichen Menge bezeichnet. Aus der obigen Gleichung folgt durch Umformen

$$n \cdot |A \cap B| = |A| \cdot |B|.$$

Nehmen wir nun an, dass weder A noch B leer oder die gesamte Grundmenge Ω sind. Dann gilt $1 \leq |A|, |B|, |A \cap B| \leq n - 1$. Insbesondere gilt also, dass n ein Primfaktor von $|A|$ oder $|B|$ ist. Wegen $|A|, |B| \leq n - 1$ ist dies ein Wiederspruch. Also muss eine der beiden Mengen A, B leer oder die gesamte Grundmenge Ω sein.

Wir zeigen nun, dass die Implikation von (ii) nach (i) gilt. Nehmen wir an, dass (i) nicht gilt, also $n = k \cdot \ell$ für zwei ganze Zahlen $1 < k, \ell < n$, und nehmen wir ohne Beschränkung der Allgemeinheit an, dass $\ell \leq k$. Daraus folgt

$$\frac{1}{n} = \frac{k}{n} \cdot \frac{\ell}{n}$$

Siehe nächste Seite!

und

$$n - k = k \cdot \ell - k = k(\ell - 1) \geq k \geq \ell > 1.$$

Somit gilt insbesondere $n \geq k + \ell$. Seien nun

$$A := \{1, \ldots, k\} \quad \text{und} \quad B := \{k, k+1, \ldots, k+\ell-1\}.$$

Dann sind A und B unabhängig, da

$$P[A \cap B] = \frac{|A \cap B|}{n} = \frac{1}{n} = \frac{k}{n} \cdot \frac{\ell}{n} = \frac{|A|}{n} \cdot \frac{|B|}{n} = P[A] \cdot P[B],$$

und weder A noch B sind leer oder die gesamte Grundmenge Ω. Also gilt (ii) nicht. Damit haben wir die Implikation von (ii) nach (i) bewiesen.

Wahrscheinlichkeit & Statistik

Musterlösung 11

1. Wir wenden das Neyman-Pearson-Lemma an (Satz 9.1 im Skript). Der Likelihood-Quotient ist

$$
\begin{aligned}
R(x_1,\dots,x_{10};\vartheta_0,\vartheta_A) = \frac{L(x_1,\dots,x_{10};\vartheta_A)}{L(x_1,\dots,x_{10};\vartheta_0)} &= \frac{e^{-10\vartheta_A}\prod_{i=1}^{10}\frac{\vartheta_A^{x_i}}{x_i!}}{e^{-10\vartheta_0}\prod_{i=1}^{10}\frac{\vartheta_0^{x_i}}{x_i!}} \\
&= e^{-10(\vartheta_A-\vartheta_0)}\left(\frac{\vartheta_A}{\vartheta_0}\right)^{\sum_{i=1}^{10}x_i}.
\end{aligned}
$$

Statt des komplizierten Likelihood-Quotienten, können wir als Teststatistik auch

$$
T := \sum_{i=1}^{10} X_i
$$

wählen; wegen $\vartheta_A > \vartheta_0$ ist nämlich R gross genau dann, wenn T gross ist. Wir bestimmen nun das kleinste $c \in \mathbb{N}_0$ so, dass für $K = (c,\infty)$

$$
0.8 = 1-\alpha \le P_{\vartheta_0}[T \notin K] = P_{\vartheta_0}[T \le c]
$$

gilt. Für jedes $\vartheta \in \Theta$ ist $T \sim \mathcal{P}(10\vartheta)$ unter P_ϑ, weil Summen von unabhängigen Poisson-verteilten Zufallsvariablen wieder Poisson-verteilt sind. Wegen $T \sim \mathcal{P}(10\vartheta_0)$ unter P_{ϑ_0} ist also

$$
P_{\vartheta_0}[T \le c] = \sum_{k=0}^{c} P_{\vartheta_0}[T=k] = \sum_{k=0}^{c} e^{-10\vartheta_0}\frac{(10\vartheta_0)^k}{k!} = e^{-5}\sum_{k=0}^{c}\frac{5^k}{k!},
$$

und wir erhalten $c = 7$.

Sei nun (T',K') ein anderer Test. Falls $P_{\vartheta_0}[T' \in K'] > P_{\vartheta_0}[T \in K]$, so ist die Wahrscheinlichkeit für einen Fehler 1. Art bei (T',K') grösser. Andererseits gilt für den Fall $P_{\vartheta_0}[T' \in K'] \le P_{\vartheta_0}[T \in K]$ mit dem Neyman-Pearson-Lemma, dass die Macht von (T',K') kleiner ist als die von (T,K), das heisst, dass (T',K') eine grössere Wahrscheinlichkeit für einen Fehler 2. Art hat.

Siehe nächste Seite!

2. **a)** Für $i = 1, \dots, 6$ sei X_i die Anzahl der Ausfälle in der i-ten Betriebsstunde. Nach der Aufgabenstellung sind $X_1, \dots, X_6$ unabhängig und $\mathcal{P}(\vartheta)$-verteilt unter P_ϑ. Unter P_ϑ hat X_i den Erwartungswert $E_\vartheta[X_i] = \vartheta$.

b) Da wir vermuten, dass im Mittel mehr als 0.5 Ausfälle pro Stunde eintreten, wählen wir als Hypothese und Alternative

$$H_0 : \vartheta = \vartheta_0 = 0.5 \quad \text{und} \quad H_A : \vartheta = \vartheta_A = 1.$$

c) Um eine passende Teststatistik zu finden, berechnen wir den Likelihood-Quotienten wie in Aufgabe 1:

$$\begin{aligned} R(x_1, \dots, x_6; \vartheta_0, \vartheta_A) &= \frac{L(x_1, \dots, x_6; \vartheta_A)}{L(x_1, \dots, x_6; \vartheta_0)} = \frac{e^{-6\vartheta_A} \prod_{i=1}^6 \frac{\vartheta_A^{x_i}}{x_i!}}{e^{-6\vartheta_0} \prod_{i=1}^6 \frac{\vartheta_0^{x_i}}{x_i!}} \\ &= e^{-6(\vartheta_A - \vartheta_0)} \left(\frac{\vartheta_A}{\vartheta_0} \right)^{\sum_{i=1}^6 x_i} = e^3 2^{\sum_{i=1}^6 x_i}. \end{aligned}$$

$R(x_1, \dots, x_6; \vartheta_0, \vartheta_A)$ ist genau dann gross, wenn $\sum_{i=1}^6 x_i$ gross ist. Statt des komplizierten Quotienten wählen wir als Teststatistik also

$$T = \sum_{i=1}^{6} X_i.$$

d) Unter der Hypothese ist T die Summe von 6 unabhängigen $\mathcal{P}(\vartheta_0)$-verteilten Zufallsvariablen, also $\mathcal{P}(3)$-verteilt.

e) Es liegt nahe, als kritischen Bereich $K := (c, \infty)$ zu wählen, wenn man ϑ_0 gegen ϑ_A testen will. Man verwirft die Hypothese also, wenn T gross wird. Um das Signifikanzniveau von 2.5% einzuhalten, muss $P_{\vartheta_0}[T > c] \leq 0.025$ gelten. Das ist äquivalent zur Bedingung $P_{\vartheta_0}[T \leq c] \geq 0.975$. Wir erhalten als Verwerfungsbereich $K = (7, \infty)$.

f) Da wir insgesamt 6 Ausfälle beobachten ist der beobachtete Wert der Teststatistik $T(\omega) = t(x_1, \dots, x_6) = 6$.

g) Da 6 nicht im Verwerfungsbereich liegt, verwerfen wir die Hypothese nicht. Die Angabe des Herstellers, dass im Mittel 0.5 Ausfälle pro Stunde eintreten, kann also auf dem 2.5%-Niveau nicht verworfen werden.

3. **a)** Wegen $\sum_{i=1}^{12} X_i \sim \mathcal{N}(12\mu_0, 12\sigma^2)$ unter P_{μ_0} ist

$$T = \frac{\sum_{i=1}^{12} X_i - 12\mu_0}{\sigma\sqrt{12}} \sim \mathcal{N}(0, 1)$$

unter P_{μ_0}. Wir wählen also $a = \sigma\sqrt{12}$ und $b = -12\mu_0$.

Siehe nächste Seite!

b) Sei $\alpha = 0.05$. Wir führen einen Test mit der Teststatistik T und dem Verwerfungsbereich K durch, verwerfen also die Hypothese, falls $|T| > c_{\neq}$ für ein noch zu bestimmendes $c_{\neq}$. Die Definition des Niveaus ergibt

$$\begin{aligned}\alpha = P_{\mu_0}[T \in K] = P_{\mu_0}[T < -c_{\neq}, T > c_{\neq}] &= P_{\mu_0}[T < -c_{\neq}] + P_{\mu_0}[T > c_{\neq}] \\ &= \Phi(-c_{\neq}) + 1 - \Phi(c_{\neq}) = 2 - 2\Phi(c_{\neq}),\end{aligned}$$

da $T \sim \mathcal{N}(0,1)$-verteilt ist unter P_{μ_0}. Aus

$$\Phi(c_{\neq}) = 1 - \frac{\alpha}{2} = 1 - \frac{0.05}{2} = 0.975$$

folgt, dass $c_{\neq} = \Phi^{-1}(0.975) = z_{0.975} = 1.96$. Der realisierte Schätzwert ist

$$T(\omega) = t(x_1, \dots, x_{10}) = -0.0598;$$

also verwerfen wir die Hypothese nicht.

c) Die Macht des Tests an der Stelle μ_A ist

$$\begin{aligned}P_{\mu_A}[T \in K] &= P_{\mu_A}[T < -c_{\neq}, T > c_{\neq}] \\ &= P_{\mu_A}\left[\frac{\sum_{i=1}^{12} X_i - 12\mu_1}{\sigma\sqrt{12}} < -c_{\neq} + \frac{\sqrt{12}}{\sigma}(\mu_0 - \mu_1)\right] \\ &\quad + P_{\mu_A}\left[\frac{\sum_{i=1}^{12} X_i - 12\mu_1}{\sigma\sqrt{12}} > c_{\neq} + \frac{\sqrt{12}}{\sigma}(\mu_0 - \mu_1)\right] \\ &= \Phi\left(-c_{\neq} + \frac{\sqrt{12}}{\sigma}(\mu_0 - \mu_1)\right) + 1 - \Phi\left(c_{\neq} + \frac{\sqrt{12}}{\sigma}(\mu_0 - \mu_1)\right) \\ &= \Phi(-1.93) + 1 - \Phi(1.99) = 0.0501.\end{aligned}$$

Das ist ein niedriger Wert, und der Grund dafür ist, dass μ_A sehr nahe bei μ_0 liegt.

4. **a)** Die Binomialverteilung (mit $n = 50$) ist ein geeignetes Modell, um die Daten zu beschreiben. Wir nehmen dabei an, dass die Melonen unabhängig voneinander Fäulnis aufweisen oder nicht. Wir vernachlässigen so beispielsweise den Effekt, dass sich bei der Lagerung benachbarte Melonen gegenseitig mit Fäulnis anstecken könnten. Seien also $X_1, \dots, X_{50}$ unabhängig und $Be(\vartheta)$-verteilt mit unbekanntem Parameter $\vartheta \in (0,1)$, und $X_i = 1$ bzw. $X_i = 0$ bedeutet, dass die i-te Melone faul bzw. gut ist.

b) Die Modellannahme ist, dass die Anzahl $S_{50} = X_1 + \cdots + X_{50}$ fauler Melonen $Bin(n, \vartheta)$-verteilt ist unter P_ϑ. Die Hypothese und Alternative lauten

$$H_0 : \vartheta = \vartheta_0 := 0.04 \qquad \text{und} \qquad H_A : \vartheta > \vartheta_0.$$

Siehe nächste Seite!

Als Teststatistik wählen wir S_{50} (vgl. das Beispiel der "tea testing lady" im Skript). Man erhält dann

$$P_{\vartheta_0}[S_{50} > 3] = 1 - P_{\vartheta_0}[S_{50} \leq 3] \approx 1 - 0.861 = 0.139 > 0.05,$$
$$P_{\vartheta_0}[S_{50} > 4] = 1 - P_{\vartheta_0}[S_{50} \leq 4] \approx 1 - 0.951 = 0.049 < 0.05.$$

Also ist der Verwerfungsbereich $K = (4, \infty)$. Da sich nur 4 faule Melonen unter den 50 untersuchten befinden, behalten wir die Hypothese bei; wir können auf dem 5%-Niveau also nicht sicher sein, dass der Händler lügt.

c) Die Wahrscheinlichkeit für einen Fehler 2. Art ist $P_{\vartheta_A}[S_{50} \leq 4] \approx 0.431$. Diese Wahrscheinlichkeit ist sehr hoch, d.h. wir können dem Händler nur schwer eine Lüge nachweisen. Die Macht des Tests für ϑ_A beträgt $\beta(\vartheta_A) = 1 - P_{\vartheta_A}[S_{50} \leq 4] \approx 0.569$. Um die Situation zu verbessern, müsste der Grossverteiler die Stichprobengrösse erhöhen (z.B. 100 Melonen entnehmen).