Wahrscheinlichkeitstheorie und Statistik

Lösungen Lernkontrolle 2

Version 2.1 (21. Mai 2021): Footnote 2 hinzugefügt und weiteren Zwischenschritt in letzter Gleichung hinzugefügt.

Uploade deine Lösung nach 45 Minuten Bearbeitungszeit bis spätestens um 14:15 auf Moodle hoch: https://moodle-app2.let.ethz.ch/mod/assign/view.php?id=593957

Mehr Details sind auf der Kurswebseite: https://metaphor.ethz.ch/x/2021/fs/401-0604-00L/#Lernkontrolle

Aufgabe 2.1 Betrachte die gemeinsame Dichte

$$f_{X,Y}(x,y) = \begin{cases} cxy, & 1 \leq x \leq 3 \text{ und } 1 \leq y \leq 3, \\ 0, & \text{sonst.} \end{cases}$$

(a) Berechne die Normalisierungs-Konstante c.

(b) Sind X und Y unabhängig? Warum? Sind X, Y i.i.d.? Warum?

(c) Berechne $\mathbb{E}[X]$, $\mathbb{E}[Y]$, $\mathbb{E}[XY]$ und $\mathbb{E}\left[X^2\right]$.

(d) Berechne die Varianz und die Standardabweichung von X.

(e) Berechne $\mathbb{E}\left[\frac{Y}{X}\right]$ und $\mathbb{E}\left[\frac{Y}{X} + Y\right]$.

(f) Berechne die Dichte $f_{\frac{1}{X}}$ von $\frac{1}{X}$.

Lösung 2.1

(a) We find the normalising constant by integrating via

$$1 = \int_{\mathbb{R}^2} f_{X,Y}(x,y)dxdy = \int_1^3 \int_1^3 cxy\, dxdy = c\int_1^3 xdx \int_1^3 ydy = c \cdot 4 \cdot 4 = 16c,$$

so $c = \frac{1}{16}$.

(b) Yes, X and Y are independent. We can see this immediately since the joint density can be factorized as

$$f_{X,Y}(x,y) = \frac{xy}{16}\mathbb{1}_{[1,3]\times[1,3]}(x,y) = \left(\frac{x}{4}\mathbb{1}_{[1,3]}(x)\right)\left(\frac{y}{4}\mathbb{1}_{[1,3]}(y)\right).$$

To give some more detail, note that the marginal density of X is

$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)dy = \int_1^3 \frac{xy}{16}\mathbb{1}_{[1,3]}(x)dy = \frac{x}{4}\mathbb{1}_{[1,3]}(x),$$

and by a similar calculation

$$f_Y(y) = \frac{y}{4}\mathbb{1}_{[1,3]}(y),$$

and indeed $f_{X,Y}(x,y) = f_X(x)f_Y(y)$, showing that they are independent.

Yes, they are i.i.d., since we have already shown that they are indpendent and they have the same cdf, because[1] they have the same density $f_X = f_Y : \mathbb{R} \to \mathbb{R}, t \mapsto \frac{t}{4}\mathbb{1}_{[1,3]}(t)$.

[1] If $f_X = f_Y$, one can directly see that $F_X(a) = \int_{-\infty}^a f_X(t)dt = \int_{-\infty}^a f_Y(t)dt = F_Y(a)\ \forall a \in \mathbb{R}$.

(c) By using the densities f_X and f_Y from (b), we compute

$$\mathbb{E}[X] = \int_{\mathbb{R}} x f_X(x) dx = \int_1^3 x \frac{x}{4} dx = \frac{13}{6} \approx 2.166667$$

and analogously

$$\mathbb{E}[Y] = \frac{13}{6} \approx 2.166667.$$

Alternatively one could directly compute more tediously

$$\begin{aligned}
\mathbb{E}[X] &= \int_{\mathbb{R}^2} x f_{X,Y}(x,y) dx dy \\
&= \int_1^3 \int_1^3 x \frac{xy}{16} \, dx dy \\
&= \int_1^3 \frac{x^2}{16} dx \int_1^3 y dy \\
&= \frac{3^3 - 1^3}{3 \cdot 16} \frac{3^2 - 1^2}{2} = \frac{27-1}{3 \cdot 16} \frac{9-1}{2} = \frac{26}{3 \cdot 16} 4 = \frac{13}{6} \approx 2.166667.
\end{aligned}$$

By independence,

$$\mathbb{E}[XY] \overset{\text{ind.}}{=} \mathbb{E}[X] \, \mathbb{E}[Y] = \left(\frac{13}{6}\right)^2 = \frac{169}{36} \approx 4.694444.$$

Alternatively one could directly compute more tediously

$$\mathbb{E}[XY] = \int_{\mathbb{R}^2} xy f_{X,Y}(x,y) dx dy = \int_1^3 \int_1^3 xy \frac{xy}{16} \, dx dy = \int_1^3 \frac{x^2}{16} dx \int_1^3 y^2 dy = \cdots = \frac{13^2}{6^2}.$$

By using the density f_X from (b), we compute

$$\mathbb{E}\left[X^2\right] = \int_{\mathbb{R}} x^2 f_X(x) dx = \int_1^3 x^2 \frac{x}{4} dx = \frac{3^4}{16} - \frac{1^4}{16} = \frac{80}{16} = 5.$$

Alternatively one could directly compute

$$\mathbb{E}\left[X^2\right] = \int_{\mathbb{R}^2} x^2 f_{X,Y}(x,y) dx dy = \int_1^3 \int_1^3 x^2 \frac{xy}{16} \, dx dy = \int_1^3 \frac{x^3}{16} dx \int_1^3 y dy = \cdots = 5.$$

(d) Unter Verwendung der Resultate aus (c) berechnen wir die Varianz

$$\sigma_X^2 = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2 = 5 - \frac{169}{36} = \frac{11}{36} \approx 0.305556$$

und die Standardabweichung

$$\sigma_X = \sqrt{\sigma_X^2} = \sqrt{\frac{11}{36}} = \frac{\sqrt{11}}{6} \approx 0.5528.$$

Alternativ kann man auch direkt aber aufwendiger die Varianz berechnen:

$$\sigma_X^2 = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[\left(X - \frac{13}{6}\right)^2\right] = \int_1^3 \int_1^3 \left(x - \frac{13}{6}\right)^2 \frac{xy}{16} \, dx dy = \cdots = \frac{11}{36} \quad \text{or}$$

$$\sigma_X^2 = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[\left(X - \frac{13}{6}\right)^2\right] = \int_{\mathbb{R}} \left(x - \frac{13}{6}\right)^2 f_X(x) dx = \int_1^3 \left(x - \frac{13}{6}\right)^2 \frac{x}{4} \, dx = \cdots = \frac{11}{36}.$$

(e) Zuerst berechnen wir

$$\mathbb{E}\left[\frac{1}{X}\right] = \int_{\mathbb{R}} \frac{1}{x} f_X(x)dx = \int_1^3 \frac{1}{x}\frac{x}{4}dx = \int_1^3 \frac{1}{4}dx = \frac{3-1}{4} = \frac{1}{2}.$$

Weil X und Y unabhängig sind, sind auch $\frac{1}{X}$ und Y unabhängig, daher gilt

$$\mathbb{E}\left[\frac{Y}{X}\right] = \mathbb{E}\left[Y\frac{1}{X}\right] \overset{\text{unabh.}^2}{=} \mathbb{E}[Y]\,\mathbb{E}\left[\frac{1}{X}\right] = \frac{13}{6}\frac{1}{2} = \frac{13}{12}.$$

Alternativ kann man auch direkt

$$\begin{aligned}\mathbb{E}\left[\frac{Y}{X}\right] &= \int_{\mathbb{R}^2} \frac{y}{x} f_{X,Y}(x,y)dxdy \\ &= \int_1^3 \int_1^3 \frac{y}{x}\frac{xy}{16}dxdy \\ &= \int_1^3 \frac{1}{16}dx \int_1^3 y^2 dy \\ &= \frac{3-1}{16}\frac{3^3-1^3}{3} = \frac{2}{16}\frac{27-1}{3} = \frac{1}{8}\frac{26}{3} = \frac{13}{12}\end{aligned}$$

berechnen.

Mit der Linearität des Erwartungswert erhalten wir

$$\mathbb{E}\left[\frac{Y}{X} + Y\right] = \mathbb{E}\left[\frac{Y}{X}\right] + \mathbb{E}[Y] = \frac{13}{12} + \frac{13}{6} = \frac{3 \cdot 13}{12} = \frac{13}{4} = 3.25.$$

(f) Wir wollen

$$f_{\frac{1}{X}}(a) = \frac{d}{da}F_{\frac{1}{X}}(a) = \frac{d}{da}\mathbb{P}\left[\frac{1}{X} \le a\right] \tag{1}$$

berechnen. Da wir wissen, dass X $\mathbb{P}$-fast sicher positiv ist, können wir im Fall $a > 0$ die Äquivalenzumformung

$$\frac{1}{X} \le a \quad \overset{\mathbb{P}\text{-f.s.}}{\Longleftrightarrow} \quad \frac{1}{a} \le X \tag{2}$$

anwenden. Im anderen Fall ist $\mathbb{P}\left[\frac{1}{X} \le a\right] = 0 \ \forall a \in \mathbb{R}_{\le 0}$ und somit $f_{\frac{1}{X}}(a) = 0 \ \forall a \in \mathbb{R}_{\le 0}$.

Im folgenden können wir immer $a > 0$ annehmen und somit die vorher beschriebene Äquivalenzumformung (2) verwenden um $F_{\frac{1}{X}}(a)$ aus (1) zu berechnen:

$$\mathbb{P}\left[\frac{1}{X} \le a\right] = \mathbb{P}\left[\frac{1}{a} \le X\right] = 1 - \mathbb{P}\left[X < \frac{1}{a}\right] = 1 - F_X\left(\frac{1}{a}-\right) = 1 - F_X\left(\frac{1}{a}\right), \tag{3}$$

wobei wir im letzten Schritt verwendet haben, dass X eine kontinuierliche Zufallsvariable ist. Man könnte sich entweder F_X explizit ausrechnen, ähnlich wie es in der alternativen Lösung gemacht wird, oder man spart sich diesen Schritt, weil man im Folgenden mit der Kettenregel nur die Ableitung $F'_X = f_X$ benötigt anstatt einer expliziten Darstellung von F_X. Somit können wir (3) in (1) einsetzen und erhalten mit der Kettenregel:

$$\begin{aligned} f_{\frac{1}{X}}(a) = \frac{d}{da}\left(1 - F_X\left(\frac{1}{a}\right)\right) &= 0 - F'_X\left(\frac{1}{a}\right)\frac{-1}{a^2} \\ &= f_X\left(\frac{1}{a}\right)\frac{1}{a^2} \\ &= \frac{\frac{1}{a}}{4}\mathbb{1}_{[1,3]}\left(\frac{1}{a}\right)\frac{1}{a^2} = \frac{1}{4a^3}\mathbb{1}_{[1,3]}\left(\frac{1}{a}\right) \overset{a\ge 0}{=} \frac{1}{4a^3}\mathbb{1}_{[\frac{1}{3},1]}(a).\end{aligned}$$

[2] Anstatt zu argumentieren, dass $\frac{1}{X}$ und Y unabhängig sind, kann man auch direkt Theorem 3.14 aus dem Skript verwenden mit $\phi : \mathbb{R} \to \mathbb{R}, y \mapsto \phi(x) := y$ und $\psi : \mathbb{R} \to \mathbb{R}, x \mapsto \psi(x) := \begin{cases} \frac{1}{x} & , \text{für } x \neq 0 \\ 0 & , \text{sonst} \end{cases}$, wobei $\psi(0)$ kein Problem darstellt, da $\mathbb{P}[X=0] = 0$ und somit $\mathbb{E}\left[Y\frac{1}{X}\right] = \mathbb{E}[\phi(Y)\psi(X)]$.

Durch zusammensetzen der beiden Fälle erhalten wir für alle $a \in \mathbb{R}$:

$$f_{\frac{1}{X}}(a) = \frac{1}{4a^3}\mathbb{1}_{[\frac{1}{3},1]}(a)\,.$$

Alternativ kann man auch $F_{\frac{1}{X}}$ bzw F_X explizit ausrechnen (wobei die folgende Rechnung nur den Fall $a > 0$ behandelt):

Zuerst berechnen wir die (kumulierte) Verteilungsfunktion

$$\begin{aligned}
F_{\frac{1}{X}}(a) = \mathbb{P}\left[\frac{1}{X} \le a\right] \overset{X>0\ \mathbb{P}\text{-f.s.}}{=} \mathbb{P}\left[\frac{1}{a} \le X\right] &= 1 - \mathbb{P}\left[X < \frac{1}{a}\right] \\
&= 1 - \int_{-\infty}^{\frac{1}{a}} f_X(x)dx \\
&= 1 - \begin{cases} 0 & \text{, falls } \frac{1}{a} < 1 \\ \int_1^{\frac{1}{a}} \frac{x}{4}dx & \text{, falls } 1 \le \frac{1}{a} \le 3 \\ 1 & \text{, falls } \frac{1}{a} > 3 \end{cases} \\
&= 1 - \begin{cases} 0 & \text{, falls } \frac{1}{a} < 1 \\ \dfrac{\frac{1}{a^2} - 1^2}{8} & \text{, falls } 1 \le \frac{1}{a} \le 3 \\ 1 & \text{, falls } \frac{1}{a} > 3 \end{cases} \\
&= \begin{cases} 1 & \text{, falls } \frac{1}{a} < 1 \\ \dfrac{9 - \frac{1}{a^2}}{8} & \text{, falls } 1 \le \frac{1}{a} \le 3 \\ 0 & \text{, falls } \frac{1}{a} > 3 \end{cases} \\
&= \begin{cases} 1 & \text{, falls } 1 < a \\ \dfrac{9 - \frac{1}{a^2}}{8} & \text{, falls } \frac{1}{3} \le a \le 1 \\ 0 & \text{, falls } \frac{1}{3} > a \end{cases} \\
&= \begin{cases} 0 & \text{, falls } a < \frac{1}{3} \\ \dfrac{9 - \frac{1}{a^2}}{8} & \text{, falls } \frac{1}{3} \le a \le 1 \\ 1 & \text{, falls } a > 1. \end{cases}
\end{aligned}$$

Damit können wir durch Ableiten direkt die Dichte

$$f_{\frac{1}{X}}(a) = F'_{\frac{1}{X}}(a) = \begin{cases} 0 & \text{, falls } a < \frac{1}{3} \\ \dfrac{2\frac{1}{a^3}}{8} & \text{, falls } \frac{1}{3} \le a \le 1 \\ 0 & \text{, falls } a > 1 \end{cases} = \begin{cases} 0 & \text{, falls } a < \frac{1}{3} \\ \frac{1}{4a^3} & \text{, falls } \frac{1}{3} \le a \le 1 \\ 0 & \text{, falls } a > 1 \end{cases} = \frac{1}{4a^3}\mathbb{1}_{[\frac{1}{3},1]}(a)$$

erhalten.

Aufgabe 2.2 Seien $X_1, \dots, X_n$ i.i.d. $\mathcal{N}(m, \sigma^2)$-normalverteilte Zufallsvariablen mit $m = 1.85$ und $\sigma = 0.14$.

(a) Was ist der Erwartungswert, die Varianz und die Standardabweichung von X_i? (Die Ergebnisse können direkt ohne Beweis angegeben werden.)

(b) Welcher Verteilung folgt der Stichprobenmittelwert $\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n}$?

(c) Berechne die Wahrscheinlichkeit, dass der Stichprobenmittelwert $\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n}$ kleiner als 1.76 ist, wobei du die Lösung mithilfe von Φ ausdrücken kannst. Du kannst also $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-\frac{t^2}{2}}dt$ verwenden ohne das Integral explizit auszurechnen.

Lösung 2.2

(a) Für alle $i \in \{1, \dots, n\}$ gilt $X_i \sim \mathcal{N}(m, \sigma^2)$ und somit:

Erwartungswert:	$\mathbb{E}[X_i] = m = 1.85$
Varianz:	$\sigma^2_{X_i} = \sigma^2 = 0.14^2 = 0.0196$
Standardabweichung:	$\sigma_{X_i} = \sigma = 0.14.$

(b) Zuerst können wir den Stichprobenmittelwert als linearkombination unabhängiger Normalverweitlen Zufallsvariablen darstellen

$$\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n} = \sum_{i=1}^n \frac{1}{n} X_i.$$

Wegen der Unabhängigkeit der X_i gilt

$$\bar{X}_n \sim \mathcal{N}\left(\sum_{i=1}^n \frac{1}{n} m, \sum_{i=1}^n \left(\frac{1}{n}\right)^2 \sigma^2\right)$$

weil $X_i \sim \mathcal{N}(m, \sigma^2)$. Durch weitere Vereinfachungen der Summen über n gleiche Summanden erhalten wir

$$\bar{X}_n \sim \mathcal{N}\left(n\frac{1}{n}m, n\left(\frac{1}{n}\right)^2 \sigma^2\right) = \mathcal{N}\left(m, \frac{1}{n}\sigma^2\right).$$

Somit gilt $\bar{X}_n \sim \mathcal{N}\left(1.85, \frac{0.0196}{n}\right)$.

(c) Wir können die (kumulierte) Verteilungsfunktion Φ einer standardnormalverteilten[3] Zufallsvariable verwenden, haben aber nicht direkt die Verteilungsfunktion einer $\mathcal{N}(m, \tilde{\sigma}^2)$-verteilten Zufallsvariable gegeben. Im folgenden bezeichnet $\tilde{\sigma} = \frac{0.14}{\sqrt{n}}$. Wenn wir von $\bar{X}_n \sim \mathcal{N}\left(m, \tilde{\sigma}^2\right)$ den Mittelwert m abziehen erhalten wir eine zentrierte normalverteilte Zufallsvariable $\bar{X}_n - m \sim \mathcal{N}\left(0, \tilde{\sigma}^2\right)$. Wenn wir diese Zufallsvariable als nächstes multiplizieren mit $\frac{1}{\tilde{\sigma}}$ erhalten wir eine standardnormalverteilte ZufallsVaraible $Z := \frac{\bar{X}_n - m}{\tilde{\sigma}} \sim \mathcal{N}(0,1)$. Jetzt können wir eine Äquivalenzumformung auf die Ungleichung

$$\begin{aligned} &\bar{X}_n < 1.76 \\ \iff &\frac{\bar{X}_n - m}{\tilde{\sigma}} < \frac{1.76 - m}{\tilde{\sigma}} \end{aligned}$$

anwenden. Schließlich können wir die Wahrscheinlichkeit

$$\mathbb{P}\left[\bar{X}_n < 1.76\right] = \mathbb{P}\left[\frac{\bar{X}_n - m}{\tilde{\sigma}} < \frac{1.76 - m}{\tilde{\sigma}}\right] = \mathbb{P}\left[Z < \frac{1.76 - 1.85}{\tilde{\sigma}}\right] = \mathbb{P}\left[Z < \frac{-0.09}{\tilde{\sigma}}\right]$$

berechnen. Weil $Z \sim \mathcal{N}(0,1)$ standardnormalverteilt ist hat es die Dichte $f_Z(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$. Somit ist

$$\mathbb{P}[Z < a] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{t^2}{2}} dt = \Phi(a).$$

Durch einsetzen erhalten wir das Ergebnis:

$$\mathbb{P}\left[\bar{X}_n < 1.76\right] = \Phi\left(\frac{-0.09}{\tilde{\sigma}}\right) = \Phi\left(\frac{-0.09\sqrt{n}}{0.14}\right) = \Phi\left(\frac{-9\sqrt{n}}{14}\right) \approx \Phi\left(-0.643\sqrt{n}\right).$$

[3] Die Standardnormalverteilung bezeichnet $\mathcal{N}(0,1)$, also eine Normalverteilung mit Mittelwert 0 und Varianz 1.