Aufgaben und Lösungsvorschlag

Für jede der Subfragen von 1., 2. und 3. ist genau eine Antwortmöglichkeit richtig.

1. Multiple choice [3 Punkte]

1.1. [**0.5 Punkte**] Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Welche der Folgenden Aussagen ist für alle Zufallsvariablen $X, Y : \Omega \to \mathbb{R}$ mit $\mathbb{E}[X], \mathbb{E}[Y] < \infty$ wahr:

(A) $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ gilt immer.

(B) $\mathbb{E}[X+Y] \geq \mathbb{E}[X] + \mathbb{E}[Y]$ gilt immer, aber (A) gilt im allgemeinen nicht.

(C) $\mathbb{E}[X+Y] \leq \mathbb{E}[X] + \mathbb{E}[Y]$ gilt immer, aber (A) gilt im allgemeinen nicht.

(D) Keine dieser (Un-)gleichungen gilt immer, aber $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ gilt, wenn X und Y unabhängig sind.

Lösung:

(A) ist korrekt, wegen der Linearität des Erwartungswert (siehe Theorem 2.34 und Remark 2.35 im Skript). □

1.2. [**1.5 Punkte**] Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Welche der Folgenden Aussagen ist für alle Zufallsvariablen $X, Y : \Omega \to \mathbb{R}$ mit $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$ wahr:

(A) $\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$ gilt immer.

(B) $\sigma^2_{X+Y} \geq \sigma^2_X + \sigma^2_Y$ gilt immer, aber (A) gilt im allgemeinen nicht.

(C) $\sigma^2_{X+Y} \leq \sigma^2_X + \sigma^2_Y$ gilt immer, aber (A) gilt im allgemeinen nicht.

(D) Keine dieser (Un-)gleichungen gilt immer. Aber $\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$ gilt, wenn X und Y unabhängig sind.

Lösung:

(D). Beweis:

Beispiel für “<”: Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Laplace-Modell auf $\Omega = \{-1, 1\}$ und $X(\omega) = \omega$ und $Y = -X$ (i.e., $Y(\omega) = -\omega$), dann gilt $X + Y = 0$ und somit

$$\sigma^2_{X+Y} = 0 < 1 + 1 = \sigma^2_X + \sigma^2_Y.$$

Daher sind die Antwortmöglichkeiten (A) and (B) falsch.

Beispiel für “>”: Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Laplace-Modell auf $\Omega = \{-1, 1\}$ und $X(\omega) = \omega$ und $Y = X$ (i.e., $Y(\omega) = \omega$), dann gilt $X + Y = 2X$ und somit

$$\sigma^2_{X+Y} = \sigma^2_{2X} = 2^2\sigma^2_X = 4 > 1 + 1 = \sigma^2_X + \sigma^2_Y.$$

Daher sind die Antwortmöglichkeiten (A) and (C) falsch.

(D) ist korrekt, da die (Un-)gleichungen aus (A)–(C) im allgemeinen falsch sind aber $\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$ gilt, wenn X und Y unabhängig sind (siehe Proposition 2.42 im Skript). Im allgemeinen würde $\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$ auch schon gelten wenn man nur fordern würde dass X und Y unkorreliert[a] sind. □

[a] Zwei Zufallsvariablen X und Y werden als unkorreliert bezeichnet genau dann wenn die Kovarianz $\mathrm{Cov}[X, Y] := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = 0$ ist. Unabhängige Zufallsvariablen sind immer unkorreliert, wie

direkt aus Theorem 2.38(iii) aus dem Skript für $f = id - \mathbb{E}[X]$ und $g = id - \mathbb{E}[Y]$ folgt. Allerdings gibt es auch unkorrelierte Zufallsvariablen, die nicht unabhängig sind. Im allgemeinen gilt $\sigma^2_{X+Y} = \sigma^2_X + 2\mathrm{Cov}[X,Y] + \sigma^2_Y$. Somit hängt die Richtung der Ungleichung vom Vorzeichen der Kovarianz $\mathrm{Cov}[X,Y]$ ab.

1.3. **[1 Punkt]** Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Laplace-Modell auf $\Omega := \{1,2,3,4,5\}$. Wie viele Elemente hat die Menge $\mathcal{F}$?

(A) $|\mathcal{F}| = 1$

(B) $|\mathcal{F}| = 5$

(C) $|\mathcal{F}| = 10$

(D) $|\mathcal{F}| = 32$

(E) $|\mathcal{F}| = 55$

Lösung:

(D). Per Definition ist die Sigma-Algebra eines Laplace-Modells immer gegeben als Potenzmenge von Ω, i.e.,

$$\mathcal{F} = \mathcal{P}(\Omega) = \Big\{\varnothing, \{1\}, \{2\}, \{1,2\}, \{3\}, \{1,3\}, \{2,3\}, \{1,2,3\}, \{4\}, \{1,4\}, \{2,4\}, \{1,2,4\}, \{3,4\}, \{1,3,4\}, \{2,3,4\}, \{1,2,3,4\}, \{5\}, \{1,5\}, \{2,5\}, \{1,2,5\}, \{3,5\}, \dots, \Omega\Big\}.$$

Die Potenzmenge $\mathcal{P}(\Omega)$ von Ω ist die Menge aller Teilmengen von Ω. Im allgemeinen gilt $|\mathcal{P}(\Omega)| = 2^{|\Omega|}$. Die Formel $|\mathcal{P}(\Omega)| = 2^{|\Omega|}$ kann hergeleitet werden indem man einsieht, dass jede Teilmenge von Ω einem $|\Omega|$-dimensionalem binären Vektor $x \in \{0,1\}^{|\Omega|}$ entspricht, wobei die k-te Koordinate x_k des binren Vektors x angibt ob das k-te Element von Ω in der Menge enthalte ist: Zum Beispiel: $\varnothing \cong (0,0,0,0,0)$, $\{1\} \cong (1,0,0,0,0)$, $\{2\} \cong (0,1,0,0,0)$, $\{1,2\} \cong (1,1,0,0,0)$. Und $\{0,1\}^{|\Omega|}$ hat offensichtlich $2^{|\Omega|}$ Elemente.

2. Multiple-Choice: Unsicherheits-Quantifizierung [9 Punkte]

Sei $Y_1, Y_2, \dots, Y_n \sim \mathcal{N}(m, \sigma^2)$ eine Folge von i.i.d. normalverteilten Zufallsvariablen mit unbekanntem Mittelwert m und bekannter Standardabweichung $\sigma = 0.5$. Wir betrachten das Intervall $I_n = [\frac{1}{n}\sum_{i=1}^n Y_i - a_n, \frac{1}{n}\sum_{i=1}^n Y_i + a_n]$ mit $a_n := \sigma\Phi^{-1}(97.5\%)$.

2.1. **[2 Punkte]** Wie hoch ist die Wahrscheinlichkeit, dass eine neue unabhängige identisch verteilte Zufallsvariable $Y_{n+1} \sim \mathcal{N}(m, \sigma^2)$ im Intervall I_n liegt, im Limes $n \to \infty$?

(A) $\lim_{n\to\infty} \mathbb{P}[Y_{n+1} \in I_n] = 0\%$.
(B) $\lim_{n\to\infty} \mathbb{P}[Y_{n+1} \in I_n] = 5\%$.
(C) $\lim_{n\to\infty} \mathbb{P}[Y_{n+1} \in I_n] = 50\%$.
(D) $\lim_{n\to\infty} \mathbb{P}[Y_{n+1} \in I_n] = 95\%$.
(E) $\lim_{n\to\infty} \mathbb{P}[Y_{n+1} \in I_n] = 100\%$.

Lösung:

(D) Intuitive Begründung: Im Limes $n \to \infty$ konvergiert $\hat{m} = \frac{1}{n}\sum_{i=1}^n Y_i$ gegen m. Somit kovergiert das Intervall $I_n = [\hat{m} - a_n, \hat{m} + a_n]$ gegen das Intervall $I_\infty := [m - a_1, m + a_1]$, weil $a_n = a_1 = \sigma\Phi^{-1}(97.5\%)$ konstant ist. Die folgende einfache Rechnung ergibt

$$\mathbb{P}[Y_{n+1} \in [m - a_1, m + a_1]] = \mathbb{P}\left[\frac{Y_{n+1} - m}{\sigma} \in [-\frac{a_1}{\sigma}, \frac{a_1}{\sigma}]\right] = \Phi(\frac{a_1}{\sigma}) - \Phi(-\frac{a_1}{\sigma}),$$

weil $\frac{Y_{n+1}-m}{\sigma} \sim \mathcal{N}(0,1)$ standard-normalverteilt ist. Aus Symmetrie-Gründen gilt $\Phi(-\frac{a_1}{\sigma}) = 1 - \Phi(\frac{a_1}{\sigma})$. Durch einsetzen der Definition von a_1 erhalten wir, dass $\Phi(\frac{a_1}{\sigma}) = \Phi(\frac{\sigma\Phi^{-1}(97.5\%)}{\sigma}) = \Phi(\Phi^{-1}(97.5\%)) = 97.5\%$. Somit gilt

$$\mathbb{P}[Y_{n+1} \in [m - a_1, m + a_1]] = 97.5\% - (1 - 97.5\%) = 95\%.$$

Intuitiv ist es relativ offensichtlich, dass $\lim_{n\to\infty} \mathbb{P}[Y_{n+1} \in I_n] = \mathbb{P}[Y_{n+1} \in [m - a_1, m + a_1]]$ gilt. Für einen vollständigen Beweise müsste man mit der Stetigkeit von Φ argumentieren.

Alternativer Beweis: Die obige Beweisidee funktioniert auch so ähnlich für andere stetige Verteilungen. Alternativ dazu kann man sich im Fall der Normalverteilung die Wahrscheinlichkeiten einfach explizit berechnen:

$$\begin{aligned}
Y_{n+1} \in I_n &\iff Y_{n+1} \in \left[\frac{1}{n}\sum_{i=1}^n Y_i - a_n, \frac{1}{n}\sum_{i=1}^n Y_i + a_n\right] \\
&\iff Y_{n+1} - \frac{1}{n}\sum_{i=1}^n Y_i \in [-a_n, a_n] \\
&\iff \frac{Y_{n+1} - \frac{1}{n}\sum_{i=1}^n Y_i}{\sqrt{\sigma^2 + n\frac{\sigma^2}{n^2}}} \in \left[-\frac{a_n}{\sqrt{\sigma^2 + n\frac{\sigma^2}{n^2}}}, \frac{a_n}{\sqrt{\sigma^2 + n\frac{\sigma^2}{n^2}}}\right]
\end{aligned}$$

Weil $\frac{Y_{n+1} - \frac{1}{n}\sum_{i=1}^n Y_i}{\sqrt{\sigma^2 + n\frac{\sigma^2}{n^2}}} \sim \mathcal{N}(0,1)$ standard-Normalverteilt ist (wegen den Properties of normal random variables aus Seite 50 des Skripts), gilt

$$\begin{aligned}\mathbb{P}\left[Y_{n+1} \in I_n\right] &= \Phi\left(\frac{a_n}{\sqrt{\sigma^2 + n\frac{\sigma^2}{n^2}}}\right) - \Phi\left(-\frac{a_n}{\sqrt{\sigma^2 + n\frac{\sigma^2}{n^2}}}\right) \\ &= \Phi\left(\frac{\sigma\Phi^{-1}(97.5\%)}{\sqrt{\sigma^2(1+\frac{1}{n})}}\right) - \Phi\left(-\frac{\sigma\Phi^{-1}(97.5\%)}{\sqrt{\sigma^2(1+\frac{1}{n})}}\right) \\ &= \Phi\left(\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{1}{n})}}\right) - \Phi\left(-\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{1}{n})}}\right) \\ &= \Phi\left(\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{1}{n})}}\right) - 1 + \Phi\left(\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{1}{n})}}\right) \\ &= 2\Phi\left(\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{1}{n})}}\right) - 1,\end{aligned}$$

wobei wir im vorletzten Schritt die Symmetrie der Dichte der Normalverteilung verwendet haben. Wegen der Stetigkeit von Φ können wir dem Limes in Φ hineinziehen:

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}\left[Y_{n+1} \in I_n\right] = \lim_{n\to\infty} 2\Phi\left(\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{1}{n})}}\right) - 1 &= 2\Phi\left(\lim_{n\to\infty}\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{1}{n})}}\right) - 1 \\ &= 2\Phi\left(\Phi^{-1}(97.5\%)\right) - 1 = 2 \cdot 97.5\% - 1 = 95\%.\end{aligned}$$

□

Diskussion: I_n ist nicht das 95%-Konfidenzintervall für m, weil in a_n der Faktor $\frac{1}{\sqrt{n}}$ fehlt. Das Konfidenzintervall $\tilde{I}_n := [\hat{m} - \frac{\sigma\Phi^{-1}(97.5\%)}{\sqrt{n}}, \hat{m} + \frac{\sigma\Phi^{-1}(97.5\%)}{\sqrt{n}}]$ würde nicht $\lim_{n\to\infty} \mathbb{P}\left[Y_{n+1} \in \tilde{I}_n\right] = 95\%$ sondern $\lim_{n\to\infty} \mathbb{P}\left[Y_{n+1} \in \tilde{I}_n\right] = 0\%$ erfüllen (siehe Aufgabe 10.4 aus der 10. Serie). Dafür würde das Konfidenzintervall $\mathbb{P}\left[m \in \tilde{I}_n\right] = 95\%$ erfüllen im Gegensatz zu I_n (siehe Frage 2.4). Das Prognoseintervall $\lim_{n\to\infty} I_n$ sollte nicht mit dem Konfidenzintervall $\tilde{I}_n$ verwechselt werden, da beide unterschiedliche Bedeutung haben. Das Prognoseintervall $\lim_{n\to\infty} I_n$ gibt die Unsicherheit bezüglich neuer Datenpunkte Y_{n+1} an, während das Konfidenzintervall $\tilde{I}_n$ die Unsicherheit bezüglich dem Mittelpunkt m beschreibt.

2.2. [**2 Punkte**] Wir fragen uns wie sich die Antwort zu Frage 2.1 ändert, wenn wir nur $\tilde{n} < \infty$ Beobachtungen haben anstatt den Limes $n \to \infty$ zu betrachten. Welche der Aussagen ist wahr?[1]

(A) $\forall \tilde{n} \in \mathbb{N} : \mathbb{P}\left[Y_{\tilde{n}+1} \in I_{\tilde{n}}\right] = \lim_{n\to\infty} \mathbb{P}\left[Y_{n+1} \in I_n\right]$.

(B) $\forall \tilde{n} \in \mathbb{N} : \mathbb{P}\left[Y_{\tilde{n}+1} \in I_{\tilde{n}}\right] < \lim_{n\to\infty} \mathbb{P}\left[Y_{n+1} \in I_n\right]$.

(C) $\forall \tilde{n} \in \mathbb{N} : \mathbb{P}\left[Y_{\tilde{n}+1} \in I_{\tilde{n}}\right] > \lim_{n\to\infty} \mathbb{P}\left[Y_{n+1} \in I_n\right]$.

(D) Keine der anderen Antwortmöglichkeiten ist korrekt, da die Richtung der Ungleichung von $\tilde{n}$ und/oder m abhängt.

Lösung:

[1] Wir definieren die Menge der natürlichen Zahlen $\mathbb{N} := \{1, 2, 3, \dots\}$ ohne 0.

(B) Intuitive Begründung: "Wenn nur endlich viele Datenpunkte zum berechnen des Schätzers $\hat{m}$ des Mittelwerts m verwendet werden, erhält man dadurch zusätzliche Unsicherheit, dies es tendenziell unwahrscheinlicher macht in dem Intervall zu liegen."

Diese Begründung ist für allgemeine Verteilungsfunktionen noch nicht ganz korrekt, aber im Fall der Normalverteilung (und vieler anderer Verteilungen) gilt diese Begründung wie das folgende präzisere Argument zeigt: Für jedes n hat das Intervall die gleiche Länge $2a_1$ unabhängig von n. Weil die Dichte der Normalverteilung $\mathcal{N}(m, \sigma^2)$ ihr Maximum bei m annimmt, unimodal und symmetrisch ist, ist unter allen Intervallen der Länge $2a_1$ das um m zentrierte Intervall $[m - a_1, m + a_1] = I_\infty$ das Intervall mit der größten Wahrscheinlichkeit. Weil die Dichte der Normalverteilung $\mathcal{N}(m, \sigma^2)$ links von m streng monoton steigend und rechts von m streng monoton fallend ist, sieht man sofort, dass es kein anderes Intervall der gleichen Länge geben kann mit gleicher Wahrscheinlichkeit. Somit führt jede Abweichung von $\hat{m}$ zu m zu einer kleineren Wahrsheinlichkeit $\mathbb{P}[Y_{\tilde{n}+1} \in I_{\tilde{n}}]$ als die Wahrscheinlichkeit $\mathbb{P}[Y_{\tilde{n}+1} \in I_\infty]$.

Alternativer Beweis: Die obige Beweisidee funktioniert auch so ähnlich für andere unimodale Verteilungen. Alternativ dazu kann man sich im Fall der Normalverteilung die Wahrscheinlichkeiten einfach explizit berechnen: Im Beweis der vorigen Aufgabe haben wir berechnet:

$$\mathbb{P}[Y_{\tilde{n}+1} \in I_{\tilde{n}}] = 2\Phi\left(\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{1}{\tilde{n}})}}\right) - 1.$$

Man sieht sofort, dass $\frac{1}{\sqrt{(1+\frac{1}{\tilde{n}})}} < 1$ gilt für jedes endliche $\tilde{n} \in \mathbb{N}$. Weil Φ eine streng monoton steigende Funktion ist gilt somit $\Phi\left(\frac{1}{\sqrt{(1+\frac{1}{\tilde{n}})}} a\right) < \Phi(a)$.

Daher gilt

$$\mathbb{P}[Y_{\tilde{n}+1} \in I_{\tilde{n}}] = 2\Phi\left(\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{1}{\tilde{n}})}}\right) - 1 < 2\Phi\left(\Phi^{-1}(97.5\%)\right) - 1 = 95\%.$$

□

Diskussion: Im Limes $n \to \infty$ deckt I_n schon die gesamte Unsicherheit für neue Beobachtungen Y_{n+1} ab. Für endliche n, deckt I_n aber nur die aleatorische Unsicherheit ab, aber nicht die epistemische Unsicherheit. Im Gegensatz dazu beschäftigt sich $\widetilde{I}_n$ (aus der Diskussion der vorigen Aufgabe) nur mit der epistemischen Unsicherheit (auch Modell-Unsicherheit genannt). Im Fall von endlichem n muss man beide Typen von Unsicherheit beachten, wenn man die gesamte Unsicherheit von Y_{n+1} abdecken will. Die aleatorische Unsichheit ist oft straight-forward berechenbar. Eine korrekte Abschätzung der epistemischen Unsicherheit ist oft subjektiver und hängt stark von den gemachten Annahmen ab.

2.3. [**2 Punkte**] Welche der folgenden Aussagen ist wahr?[1]

(A) $\forall \tilde{n} \in \mathbb{N} : \mathbb{P}[Y_{\tilde{n}} \in I_{\tilde{n}}] = \mathbb{P}[Y_{\tilde{n}+1} \in I_{\tilde{n}}]$.
(B) $\forall \tilde{n} \in \mathbb{N} : \mathbb{P}[Y_{\tilde{n}} \in I_{\tilde{n}}] < \mathbb{P}[Y_{\tilde{n}+1} \in I_{\tilde{n}}]$.
(C) $\forall \tilde{n} \in \mathbb{N} : \mathbb{P}[Y_{\tilde{n}} \in I_{\tilde{n}}] > \mathbb{P}[Y_{\tilde{n}+1} \in I_{\tilde{n}}]$.
(D) Keine der anderen Antwortmöglichkeiten ist korrekt, da die Richtung der Ungleichung von $\tilde{n}$ und/oder m abhängt.

Lösung:

(C) Intuitive Begründung: Weil $Y_{\tilde{n}}$ zieht den Schätzer $\hat{m}$ in Richtung $Y_{\tilde{n}}$. Somit ist $|Y_{\tilde{n}} - \hat{m}|$ in Erwartung kleiner als $|Y_{\tilde{n+1}} - \hat{m}|$. Alternativer Beweis: Die obige Beweisidee funktioniert auch so ähnlich für andere unimodale Verteilungen. Alternativ dazu kann man sich im Fall der Normalverteilung die Wahrscheinlichkeiten einfach explizit berechnen:

$$\begin{aligned}
Y_n \in I_n \iff & Y_n \in \left[\frac{1}{n}\sum_{i=1}^{n} Y_i - a_n, \frac{1}{n}\sum_{i=1}^{n} Y_i + a_n\right] \\
\iff & Y_n - \frac{1}{n}\sum_{i=1}^{n} Y_i \in [-a_n, a_n] \\
\iff & Y_n\left(1+\frac{1}{n}\right) - \frac{1}{n}\sum_{i=1}^{n-1} Y_i \in [-a_n, a_n] \\
\iff & \frac{Y_n\left(1+\frac{1}{n}\right) - \frac{1}{n}\sum_{i=1}^{n-1} Y_i}{\sqrt{\sigma^2\left(1+\frac{1}{n}\right)^2 + (n-1)\frac{\sigma^2}{n^2}}} \in \left[-\frac{a_n}{\sqrt{\sigma^2\left(1+\frac{1}{n}\right)^2 + (n-1)\frac{\sigma^2}{n^2}}}, \frac{a_n}{\sqrt{\sigma^2\left(1+\frac{1}{n}\right)^2 + (n-1)\frac{\sigma^2}{n^2}}}\right]
\end{aligned}$$

Weil $\frac{Y_n\left(1+\frac{1}{n}\right) - \frac{1}{n}\sum_{i=1}^{n-1} Y_i}{\sqrt{\sigma^2\left(1+\frac{1}{n}\right)^2 + (n-1)\frac{\sigma^2}{n^2}}} \sim \mathcal{N}(0,1)$ standard-Normalverteilt ist (wegen der Properties of normal random variables aus Seite 50 des Skripts), gilt[a]

$$\begin{aligned}
\mathbb{P}\left[Y_n \in I_n\right] &= \Phi\left(\frac{a_n}{\sqrt{\sigma^2\left(1+\frac{1}{n}\right)^2 + (n-1)\frac{\sigma^2}{n^2}}}\right) - \Phi\left(-\frac{a_n}{\sqrt{\sigma^2\left(1+\frac{1}{n}\right)^2 + (n-1)\frac{\sigma^2}{n^2}}}\right) \\
&= \Phi\left(\frac{\sigma\Phi^{-1}(97.5\%)}{\sqrt{\sigma^2(1+\frac{3}{n})}}\right) - \Phi\left(-\frac{\sigma\Phi^{-1}(97.5\%)}{\sqrt{\sigma^2(1+\frac{3}{n})}}\right) \\
&= \Phi\left(\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{3}{n})}}\right) - \Phi\left(-\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{3}{n})}}\right) \\
&= \Phi\left(\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{3}{n})}}\right) - 1 + \Phi\left(\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{3}{n})}}\right) \\
&= 2\Phi\left(\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{3}{n})}}\right) - 1,
\end{aligned}$$

wobei wir im vorletzten Schritt die Symmetrie der Dichte der Normalverteilung verwendet haben. Man sieht sofort, dass $\frac{1}{\sqrt{(1+\frac{3}{\tilde{n}})}} < \frac{1}{\sqrt{(1+\frac{1}{\tilde{n}})}}$ gilt für jedes endliche $n \in \mathbb{N}$. Weil Φ eine streng monoton steigende Funktion ist gilt somit $\Phi\left(\frac{1}{\sqrt{(1+\frac{3}{\tilde{n}})}}a\right) < \Phi\left(\frac{1}{\sqrt{(1+\frac{1}{\tilde{n}})}}a\right)$.

Daher gilt

$$\mathbb{P}\left[Y_{\tilde{n}} \in I_{\tilde{n}}\right] = 2\Phi\left(\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{3}{\tilde{n}})}}\right) - 1 < 2\Phi\left(\frac{\Phi^{-1}(97.5\%)}{\sqrt{(1+\frac{1}{\tilde{n}})}}\right) - 1 = \mathbb{P}\left[Y_{\tilde{n}+1} \in I_{\tilde{n}}\right].$$

□

[a] Für den zweiten Schritt in der folgenden Rechnung verwenden wir, dass

$$\left(1+\frac{1}{n}\right)^2 + (n-1)\frac{1}{n^2} = 1 + \frac{2}{n} + \frac{1}{n^2} + \frac{1}{n} - \frac{1}{n^2} = 1 + \frac{3}{n}$$

2.4. **[2 Punkte]** Wie hoch ist die Wahrscheinlichkeit, dass der unbekannte Mittelwert m im Intervall I_n liegt, im Limes $n \to \infty$?

(A) $\lim_{n\to\infty} \mathbb{P}\left[m \in I_n\right] = 0\%$.
(B) $\lim_{n\to\infty} \mathbb{P}\left[m \in I_n\right] = 5\%$.
(C) $\lim_{n\to\infty} \mathbb{P}\left[m \in I_n\right] = 50\%$.
(D) $\lim_{n\to\infty} \mathbb{P}\left[m \in I_n\right] = 95\%$.
(E) $\lim_{n\to\infty} \mathbb{P}\left[m \in I_n\right] = 100\%$.

Hinweis: $\mathbb{P}\left[m \in I_n\right] = p\%$ bedeutet dass I_n ein $p\%$-Konfidenzintervall für den Mittelwert m ist.

Lösung:

(E) Intuitive Begründung: $\hat{m}$ konvergiert gegen m und $a_n = a_1 > 0$.

Alternativer Beweis: Die obige Beweisidee funktioniert auch so ähnlich für sehr allgemeine andere Verteilungen. Alternativ dazu kann man sich im Fall der Normalverteilung die Wahrscheinlichkeiten einfach explizit berechnen:

$$\begin{aligned} m \in I_n &\iff m \in \left[\frac{1}{n}\sum_{i=1}^{n} Y_i - a_n, \frac{1}{n}\sum_{i=1}^{n} Y_i + a_n\right] \\ &\iff m - \frac{1}{n}\sum_{i=1}^{n} Y_i \in [-a_n, a_n] \\ &\iff \frac{m - \frac{1}{n}\sum_{i=1}^{n} Y_i}{\sqrt{n\frac{\sigma^2}{n^2}}} \in \left[-\frac{a_n}{\sqrt{n\frac{\sigma^2}{n^2}}}, \frac{a_n}{\sqrt{n\frac{\sigma^2}{n^2}}}\right] \end{aligned}$$

Weil $\frac{m - \frac{1}{n}\sum_{i=1}^{n} Y_i}{\sqrt{n\frac{\sigma^2}{n^2}}} \sim \mathcal{N}(0,1)$ standard-Normalverteilt ist (wegen den Properties of normal random variables aus Seite 50 des Skripts), gilt

$$\begin{aligned}\mathbb{P}\left[m \in I_n\right] &= \Phi\left(\frac{a_n}{\sqrt{n\frac{\sigma^2}{n^2}}}\right) - \Phi\left(-\frac{a_n}{\sqrt{n\frac{\sigma^2}{n^2}}}\right) \\ &= \Phi\left(\frac{\sigma\Phi^{-1}(97.5\%)}{\sqrt{\sigma^2(\frac{1}{n})}}\right) - \Phi\left(-\frac{\sigma\Phi^{-1}(97.5\%)}{\sqrt{\sigma^2(\frac{1}{n})}}\right) \\ &= \Phi\left(\sqrt{n}\Phi^{-1}(97.5\%)\right) - \Phi\left(-\sqrt{n}\Phi^{-1}(97.5\%)\right) \\ &= \Phi\left(\sqrt{n}\Phi^{-1}(97.5\%)\right) - 1 + \Phi\left(\sqrt{n}\Phi^{-1}(97.5\%)\right) \\ &= 2\Phi\left(\sqrt{n}\Phi^{-1}(97.5\%)\right) - 1,\end{aligned}$$

wobei wir im vorletzten Schritt die Symmetrie der Dichte der Normalverteilung verwendet haben. Weiters erhalten wir:

$$\lim_{n\to\infty} \mathbb{P}\left[Y_{n+1} \in I_n\right] = \lim_{n\to\infty} 2\Phi\left(\sqrt{n}\Phi^{-1}(97.5\%)\right) - 1 = 2 - 1 = 100\%.$$

□

2.5. **[1 Punkt]** Wir fragen uns wie sich die Antwort zu Frage 2.4 ändert, wenn wir nur $\tilde{n} < \infty$ Beobachtungen haben anstatt den Limes $n \to \infty$ zu betrachten. Welche der Aussagen ist wahr?[1]

(A) $\forall \tilde{n} \in \mathbb{N} : \mathbb{P}\left[m \in I_{\tilde{n}}\right] = \lim_{n\to\infty} \mathbb{P}\left[m \in I_n\right]$.

(B) $\forall \tilde{n} \in \mathbb{N} : \mathbb{P}\left[m \in I_{\tilde{n}}\right] < \lim_{n\to\infty} \mathbb{P}\left[m \in I_n\right]$.

(C) $\forall \tilde{n} \in \mathbb{N} : \mathbb{P}\left[m \in I_{\tilde{n}}\right] > \lim_{n\to\infty} \mathbb{P}\left[m \in I_n\right]$.

(D) Keine der anderen Antwortmöglichkeiten ist korrekt, da die Richtung der Ungleichung von $\tilde{n}$ und/oder m abhängt.

Lösung:

(B) Intuitiv kann man ähnlich argumentieren wie bei Frage 2.2.

Eine alternative einfachere Begründung lautet: $\lim_{n\to\infty} \mathbb{P}\left[m \in I_n\right] = 1$ (siehe Frage 2.4) und für jedes endliche $\tilde{n} < \infty$ ist $\mathbb{P}\left[m \in I_{\tilde{n}}\right] < 1$.

Alternativer Beweis: Die obige Beweisidee funktioniert auch so ähnlich für für sehr allgemeine andere Verteilungen. Alternativ dazu kann man sich im Fall der Normalverteilung die Wahrscheinlichkeiten einfach explizit berechnen: Im Beweis der vorigen Aufgabe haben wir berechnet:

$$\mathbb{P}\left[m \in I_{\tilde{n}}\right] = 2\Phi\left(\sqrt{\tilde{n}}\Phi^{-1}(97.5\%)\right) - 1.$$

Man sieht sofort, dass $\sqrt{\tilde{n}}\Phi^{-1}(97.5\%) < \infty$ gilt für jedes endliche $\tilde{n} \in \mathbb{N}$. Weil Φ eine streng monoton steigende Funktion ist gilt somit $\Phi\left(\sqrt{\tilde{n}}\Phi^{-1}(97.5\%)\right) < \lim_{a\to\infty} \Phi(a) = 1$.
Daher gilt

$$\mathbb{P}\left[m \in I_{\tilde{n}}\right] = 2\Phi\left(\sqrt{\tilde{n}}\Phi^{-1}(97.5\%)\right) - 1 < 2 - 1 = 100\%.$$

□

3. Multiple-Choice: Z **[3.5 Punkte]**

Gegeben sei eine Zufallsvariable Z mit der Verteilungsfunktion

$$F_Z(a) = \begin{cases} 0, & \text{falls } a < 0, \\ 2a, & \text{falls } 0 \leq a < \frac{1}{4} \\ \frac{1}{2} + \frac{1}{2}(a - \frac{1}{4})^2, & \text{falls } \frac{1}{4} \leq a < \frac{5}{4} \\ 1, & \text{falls } a \geq \frac{5}{4}. \end{cases}$$

3.1. **[0.5 Punkte]** Sei $E_1 = \{Z < \frac{1}{8}\}$. Was ist die Wahrscheinlichkeit von E_1?

(A) $\mathbb{P}[E_1] = 0$ (B) $\mathbb{P}[E_1] = \frac{1}{64}$ (C) $\mathbb{P}[E_1] = \frac{1}{4}$

Lösung:

(C). Beweis: $\mathbb{P}[E_1] = \mathbb{P}\left[\{Z < \frac{1}{8}\}\right] = F_X(\frac{1}{8}-) = 2\frac{1}{8} = \frac{1}{4}$, wobei $F_X(\frac{1}{8}-) = F_X(\frac{1}{8})$ wegen der Stetigkeit von F_X an $\frac{1}{8}$. In diesem Beispiel haben wir den Sonderfall, dass F_X stetig auf ganz $\mathbb{R}$ ist. Somit folgt für diese stetige Funktion F_X, dass $F_X(a-) = F_X(a)$ für alle $a \in \mathbb{R}$, weil in diesem Beispiel F_X an jeder Stelle $a \in \mathbb{R}$ stetig ist.

3.2. **[0.5 Punkte]** Sei $E_2 = \{Z = \frac{1}{6}\}$. Was ist die Wahrscheinlichkeit von E_2?

(A) $\mathbb{P}[E_2] = 0$ (B) $\mathbb{P}[E_2] = \frac{1}{6}$ (C) $\mathbb{P}[E_2] = \frac{1}{3}$

Lösung:

(A). Beweis: $\mathbb{P}[E_2] = \mathbb{P}\left[\{Z = \frac{1}{6}\}\right] = \mathbb{P}\left[\{Z \leq \frac{1}{6}\}\right] - \mathbb{P}\left[\{Z < \frac{1}{6}\}\right] = F_X(\frac{1}{6}) - F_X(\frac{1}{6}-) = 2\frac{1}{6} - 2\frac{1}{6} = 0$.

3.3. **[0.5 Punkte]** Sei $E_3 = \{\frac{1}{16} \leq Z < \frac{3}{4}\}$. Was ist die Wahrscheinlichkeit von E_3?

(A) $\mathbb{P}[E_3] = 0$ (B) $\mathbb{P}[E_3] = \frac{1}{3}$ (C) $\mathbb{P}[E_3] = \frac{1}{2}$

Lösung:

(C). Beweis: $\mathbb{P}[E_3] = \mathbb{P}\left[\{\frac{1}{16} \leq Z < \frac{3}{4}\}\right] = \mathbb{P}\left[\{Z < \frac{3}{4}\}\right] - \mathbb{P}\left[\{Z < \frac{1}{16}\}\right] = F_X(\frac{3}{4}-) - F_X(\frac{1}{16}-) = \frac{1}{2} + \frac{1}{2}\left(\frac{3}{4} - \frac{1}{4}\right)^2 - 2\frac{1}{16} = \frac{1}{2} + \frac{1}{2}\left(\frac{1}{2}\right)^2 - \frac{1}{8} = \frac{1}{2} + \frac{1}{8} - \frac{1}{8} = \frac{1}{2}$.

3.4. **[0.5 Punkte]** Sei $E_4 = \{Z > \frac{1}{5}\}$. Was ist die Wahrscheinlichkeit von E_4?

(A) $\mathbb{P}[E_4] = 0$ (B) $\mathbb{P}[E_4] = \frac{2}{5}$ (C) $\mathbb{P}[E_4] = \frac{3}{5}$

Lösung:

(C). Beweis: $\mathbb{P}[E_4] = \mathbb{P}\left[\{Z > \frac{1}{5}\}\right] = 1 - \mathbb{P}\left[\{Z \leq \frac{1}{5}\}\right] = 1 - F_X(\frac{1}{5}) = 1 - 2\frac{1}{5} = \frac{3}{5}$.

3.5. **[0.5 Punkte]** Sind E_1 und E_2 unabhängig?

(A) Nein (B) Ja

Lösung:

(B). Beweis:

$$\mathbb{P}[E_1 \cap E_2] = \mathbb{P}[\varnothing] = 0 = \frac{1}{4}0 = \mathbb{P}[E_1]\,\mathbb{P}[E_2]\,.$$

E_1 und E_2 sind unabhängig.

Intuition: E_2 ist ein deterministisches Ereignis, da $\mathbb{P}[E_2] = 0$. Deterministische Ereignisse sind immer unabhängig. In anderen Worten: Wenn E_2 sowieso nie eintritt kann es auch keine Informationen über andere Ereignisse offenbaren.

3.6. **[0.5 Punkte]** Sind E_1 und E_3 unabhängig?

(A) Nein (B) Ja

Lösung:

(B). Beweis:

$$\begin{aligned}\mathbb{P}[E_1 \cap E_3] = \mathbb{P}\left[\{\frac{1}{16} \leq Z < \frac{1}{8}\}\right] = F_X(\frac{1}{8}-) - F_X(\frac{1}{16}-) = 2\frac{1}{8} - 2\frac{1}{16} = \frac{2}{8} - \frac{1}{8} = \frac{1}{8} \\ = \frac{1}{4}\frac{1}{2} = \mathbb{P}[E_1]\,\mathbb{P}[E_3]\,.\end{aligned}$$

E_1 und E_3 sind unabhängig.

3.7. **[0.5 Punkte]** Sind E_1 und E_4 unabhängig?

(A) Nein (B) Ja

Lösung:

(A). Beweis:

$$\mathbb{P}[E_1 \cap E_4] = \mathbb{P}[\varnothing] = 0 \neq \frac{1}{4}\frac{3}{5} = \mathbb{P}[E_1]\,\mathbb{P}[E_4]\,.$$

E_1 und E_4 sind nicht unabhängig.

4. Fortsetzung von Aufgabe 3.: Z **[2 Punkte]**

Hat Z aus der 3. Aufgabe eine Dichte? Wenn ja, dann berechne die Dichtefunktion f_Z. Wenn nein, begründe warum.

Lösung:

Ja, Z hat eine Dichte (weil $F_Z(0-) = 0 = 2 \cdot 0 = F_Z(0)$, $F_Z(\frac{1}{4}-) = 2\frac{1}{4} = \frac{1}{2} + \frac{1}{2}(\frac{1}{4} - \frac{1}{4})^2 = F_Z(\frac{1}{4})$, $F_Z(\frac{5}{4}-) = \frac{1}{2} + \frac{1}{2}(\frac{5}{4} - \frac{1}{4})^2 = 1 = F_Z(\frac{5}{4})$ und weil F_Z auch sonst überall absolut stetig[a] ist.[b]):

$$f_Z(a) = F'_Z(a) = \begin{cases} 0, & \text{falls } a < 0, \\ 2, & \text{falls } 0 \leq a < \frac{1}{4} \\ a - \frac{1}{4}, & \text{falls } \frac{1}{4} \leq a < \frac{5}{4} \\ 0, & \text{falls } a \geq \frac{5}{4}. \end{cases}$$

Beweis: Man könnte leicht mithilfe einer Fallunterscheidung mit 4 Fällen nachrechnen, dass $\forall z \in \mathbb{R} : F_Z(z) = \int_{-\infty}^{z} f_Z(a)\, da$ gilt.

Diskussion: Falls F_Z eine Sprungstelle hätte, dann hätte Z keine Dichte. Eine Dichte muss weder monoton steigend noch stetig sein.

Die Dichte f_Z ist nicht eindeutig definiert: Wenn man den Wert der Dichte f_Z an endlich vielen Werten endet, ist es immer noch eine Dichte von Z (deshalb ist es kein Problem, dass F_Z an 0, $\frac{1}{4}$ und $\frac{5}{4}$ nicht differenzierbar ist. Beispielsweise ist

$$\tilde{f}_Z(a) = \begin{cases} 0, & \text{falls } a \leq 0, \\ 2, & \text{falls } 0 < a \leq \frac{1}{4} \\ a - \frac{1}{4}, & \text{falls } \frac{1}{4} < a \leq \frac{5}{4} \\ 0, & \text{falls } a > \frac{5}{4}. \end{cases}$$

auch eine gültige Lösung.

[a] Jede Lipschitz-stetige Funktion ist auch absolut stetig.

[b] Falls in der Angabe von Aufgabe 3. noch nicht explizit gegeben worden wäre, dass F_Z eine Verteilungsfunktion einer Zufallsvariable ist, müsste man noch überprüfen. dass F_Z monoton steigend ist und $\lim_{a\to-\infty} F_Z = 0$ und $\lim_{a\to\infty} F_Z = 1$ gilt.

5. Eddy **[6.5 Punkte]**

Eddy bekommt 2 Säcke mit Münzen. Der erste Sack enthält 50 faire[2] Münzen, die er zum Casino bringen soll. Der zweite Sack enthält 50 defekte Münzen, die mit 75% Wahrscheinlichkeit "Kopf" anzeigen, wenn sie geworfen werden (und somit entsorgt werden sollen).

Leider passiert Eddy ein Unfall bei dem alle 100 Münzen vermischt werden. Darum zieht Eddy zufällig eine Münze aus den 100 gemischten Münzen und will testen ob sie fair oder defekt ist.

Die Nullhypothese $\mathcal{H}_0$ ist, dass die Münze defekt ist. Zum testen wirft Eddy die Münze 2 Mal und verwirft die Nullhypothese nur falls keiner der beiden Würfe "Kopf" ergibt.

Notation: Wir definieren die beiden Ereignisse:
H_0 :="Die Münze ist defekt." und
T :="Der Test verwirft die Nullhypothese."="Keiner der beiden Würfe ergibt 'Kopf'.".

Hinweis: Verwende diese Notation um Terme wie $\mathbb{P}[T|H_0]$ oder $\mathbb{P}[T \cap H_0]$ übersichtlich aufzuschreiben.

5.1. **[1.5 Punkte]** Berechne das Signifikanz-Niveau α des Tests.

Hinweis: Das Signifikanz-Niveau (englisch: "relevance level") α ist die Wahrscheinlichkeit eines Fehlers 1. Art. Man spricht von einem Fehler 1. Art (englisch: "Type I error") wenn die Nullhypothese fälschlicherweise verworfen wird.

5.2. **[1 Punkt]** Berechne die Stärke $1-\beta$ des Tests.

Hinweis: β ist die Wahrscheinlichkeit eines Fehlers 2. Art. Die Stärke $1-\beta$ des Tests wird auch "Trennschärfe" oder (im englischen) als "Power" bezeichnet.

5.3. **[2 Punkte]** Berechne die Wahrscheinlichkeit, dass die Münze fair ist, wenn der Test die Nullhypotehse verwirft.

5.4. Eddy wendet diesen Test auf alle 100 Münzen an. Eddy gibt alle Münzen die bei keinem der beiden Würfe "Kopf" ergeben an das Casino weiter und entsorgt alle anderen Münzen.

(i) **[1 Punkt]** Wie viele defekte Münzen werden dem Casino gegeben in Erwartung?

(ii) **[1 Punkt]** Wie viele faire Münzen werden dem Casino gegeben in Erwartung?

Hinweis: Wir definieren die Ereignisse $H_{0,i}$ und T_i analog zu H_0 und T für die i-te Münze. Das bedeutet, die i-te Münze wird genau dann dem Casino gegeben, wenn T_i eintritt.

Lösung:

5.1. Wenn wir $\mathbb{P}_{\mathcal{H}_0}$ in Bayesianische Sprache übersetzen erhalten wir $\mathbb{P}[\cdot|H_0]$.[a] Mathematisch sind diese beiden Wahrscheinlichkeitsmasse identisch. Somit ist

$$\alpha := \mathbb{P}_{\mathcal{H}_0}[T] = \mathbb{P}[T|H_0] = (1-0.75)^2 = \frac{1}{4^2} = \frac{1}{16}.$$

Diskussion: Eine präzisere Formulierung des Hinweises, wäre: "α ist die Wahrscheinlichkeit eines Fehlers 1. Art *unter der Annahme, dass $\mathcal{H}_0$ wahr ist.*", denn $\alpha = \mathbb{P}_{\mathcal{H}_0}[T] = \mathbb{P}[T|H_0] \neq \mathbb{P}[T \cap H_0]$.

5.2. In diesem Fall gibt es nur eine mögliche alternative Hypothese, nämlich $\mathcal{H}_1$ ="Die Münze

[2] Faire Münzen ergeben mit 50% Wahrscheinlichkeit "Kopf", wenn sie geworfen werden.

ist fair.".

$$1-\beta := 1-\mathbb{P}_{\mathcal{H}_1}[T^{\complement}] = 1-\mathbb{P}\left[T^{\complement}|H_0^{\complement}\right] = \mathbb{P}\left[T|H_0^{\complement}\right] = 0.5^2 = \frac{1}{4}.$$

Diskussion: Eine präzisere Formulierung des Hinweises, wäre: "β ist die Wahrscheinlichkeit eines Fehlers 2. Art *unter der Annahme, dass $\mathcal{H}_1$ wahr ist*.", denn $\beta = \mathbb{P}_{\mathcal{H}_1}[T^{\complement}] = \mathbb{P}\left[T^{\complement}|H_0^{\complement}\right] \neq \mathbb{P}\left[T^{\complement} \cap H_0^{\complement}\right]$.

5.3. Weil in beiden Säcken 50 Münzen waren und die Münze zufällig aus den 100 gemischten Münzen gezogen wird, gilt $\mathbb{P}[H_0] = \mathbb{P}\left[H_0^{\complement}\right] = \frac{1}{2}$. Somit kann

$$\mathbb{P}\left[H_0^{\complement}|T\right] = \frac{\mathbb{P}\left[H_0^{\complement}\cap T\right]}{\mathbb{P}[T]} = \frac{\mathbb{P}\left[T|H_0^{\complement}\right]\mathbb{P}\left[H_0^{\complement}\right]}{\mathbb{P}[T|H_0]\,\mathbb{P}[H_0] + \mathbb{P}\left[T|H_0^{\complement}\right]\mathbb{P}\left[H_0^{\complement}\right]} = \frac{\frac{1}{4}\frac{1}{2}}{\frac{1}{16}\frac{1}{2}+\frac{1}{4}\frac{1}{2}} = \frac{1}{\frac{1}{4}+1} = \frac{1}{\frac{5}{4}} = \frac{4}{5}$$

berechnet werden.

Alternative Lösung: Es kann auch zuerst

$$\mathbb{P}[H_0|T] = \frac{\mathbb{P}[H_0\cap T]}{\mathbb{P}[T]} = \frac{\mathbb{P}[T|H_0]\,\mathbb{P}[H_0]}{\mathbb{P}[T|H_0]\,\mathbb{P}[H_0] + \mathbb{P}\left[T|H_0^{\complement}\right]\mathbb{P}\left[H_0^{\complement}\right]} = \frac{\frac{1}{16}\frac{1}{2}}{\frac{1}{16}\frac{1}{2}+\frac{1}{4}\frac{1}{2}} = \frac{\frac{1}{4}}{\frac{1}{4}+1} = \frac{\frac{1}{4}}{\frac{5}{4}} = \frac{1}{5}$$

berechnet werden um dann $\mathbb{P}\left[H_0^{\complement}|T\right] = 1-\mathbb{P}[H_0|T] = \frac{4}{5}$ zu erhalten.

Diskussion: In schlechten (frequentistischen) Lehrbüchern wird oft fälschlicherweise behauptet, dass das verwerfen der Nullhypothese auf Signifikanz-Niveau α bedeutet, dass die Nullhypothese mit Wahrscheinlichkeit α wahr ist. Dieses Beispiel zeigt, dass dem nicht so ist ($\frac{1}{16} \neq \frac{1}{5}$). Im allgemeinen ist $\mathbb{P}[H_0|T] \neq \mathbb{P}[T|H_0]$, daher sollten in guten Lehrbüchern $\mathbb{P}[H_0|T]$ und $\mathbb{P}[T|H_0]$ nicht miteinander verwechselt werden. In diesem Setting sollte die Entscheidung ob eine Münze mit ausreichender Wahrscheinlichkeit fair ist nachdem sie getestet wurde nur aufgrund von $\mathbb{P}[H_0|T]$ (und nicht auf Grund von $\mathbb{P}[T|H_0]$) getroffen werden, denn $1-\mathbb{P}[H_0|T]$ gibt exakt die Wahrscheinlichkeit an, dass die Münze fair ist gegeben dem Test-Resultat.

In anderen Settings, wo $\mathbb{P}[H_0]$ nicht bekannt ist (zum Beispiel, wenn man keine Information darüber hat ob in einem der beiden Säcke mehr Münzen waren als im anderen Sack), dann ist es nicht möglich $\mathbb{P}[H_0|T]$ zu berechnen. Frequentistische Hypothesentests wurde genau für solche Settings geschaffen in denen keine Informationen über $\mathbb{P}[H_0]$. Dort wird $\mathbb{P}[T|H_0]$ als Heuristik verwendet um $\mathbb{P}[H_0|T]$ zu "approximieren". Diese "Approximation" kann im allgemeinen beliebig schlecht sein, aber (mangels besserer Alternativen) ist diese Heuristik weit verbreitet im Fall von sehr kleinen Werten von $\alpha = \mathbb{P}[T|H_0]$. Wenn wir $\alpha = \mathbb{P}[T|H_0]$ gegen null gehen lassen ohne $\mathbb{P}\left[T|H_0^{\complement}\right]\mathbb{P}\left[H_0^{\complement}\right] > 0$ zu verkleinern, dann konvergiert $\mathbb{P}[H_0|T]$ auch gegen null. Somit ist es oft eine sinnvolle Heuristik $\mathbb{P}[H_0|T]$ als vermutlich sehr klein zu betrachten, wenn $\alpha = \mathbb{P}[T|H_0]$ extrem klein ist und man vermutet, dass $\mathbb{P}\left[T|H_0^{\complement}\right]\mathbb{P}\left[H_0^{\complement}\right] > 0$ nicht allzu klein ist. Grosse Werte von $\mathbb{P}[T|H_0]$ geben allerdings fast gar keine Informationen darüber wie gross $\mathbb{P}[H_0|T]$ ist.

5.4. (i) Es gibt 50 defekte Münzen. Jede defekte Münze wird mit Wahrscheinlichkeit $\alpha = \mathbb{P}[T|H_0]$ an das Casino weitergegeben. Somit resultiert jede defekte Münze in Erwartung

in α defekten Münzen an das Casino. Wegen Linearität der Erwartung erhält das Casino insgesamt $50\alpha = \frac{50}{16} = 3.125$ defekte Münzen in Erwartung.
Im folgenden formulieren wir dieses Argument mathematischer: Sei D die (zufällige) Menge von defekten Münzen, i.e., $D(\omega) := \{i \in \{1, \ldots, 100\} : \omega \in H_{0,i}\}$ und sei C die (zufällige) Menge von Münzen, die dem Casino weitergegeben werden, i.e., $C(\omega) := \{i \in \{1, \ldots, 100\} : \omega \in T_i\}$. Dann ist die (zufällige) Anzahl Z_{dC} an defekten Münzen, die an das Casino weitergegeben wird gegeben durch:

$$Z_{\mathrm{dC}} := |D \cap C| = \sum_{i \in D} \mathbb{1}_{T_i}.$$

Daher gilt

$$\mathbb{E}\left[Z_{\mathrm{dC}}\right] = \mathbb{E}\left[\sum_{i \in D} \mathbb{1}_{T_i}\right] = \sum_{i \in D} \mathbb{E}\left[\mathbb{1}_{T_i} | i \in D\right] = \sum_{i \in D} \mathbb{E}\left[\mathbb{1}_{T_i} | H_{0,i}\right] = \sum_{i \in D} \alpha = |D|\alpha = 50\alpha,$$

wegen der Linearität des Erwartungswert und weil die Menge D zufällig ist muss auf bedingte Erwartungen gewechselt werden.[b]
Alternative Lösung 1 für (i):

$$\begin{aligned}\mathbb{E}\left[Z_{\mathrm{dC}}\right] := \mathbb{E}\left[|D \cap C|\right] = \mathbb{E}\left[\sum_{i=1}^{100} \mathbb{1}_{T_i \cap H_{0,i}}\right] &= \sum_{i=1}^{100} \mathbb{E}\left[\mathbb{1}_{T_i \cap H_{0,i}}\right] = \sum_{i=1}^{100} \mathbb{P}\left[T_i \cap H_{0,i}\right] \\ &= \sum_{i=1}^{100} \mathbb{P}\left[T_i | H_{0,i}\right] \mathbb{P}\left[H_{0,i}\right] = \sum_{i=1}^{100} \alpha \frac{1}{2} = 100\alpha\frac{1}{2} = 50\alpha,\end{aligned}$$

wegen der Linearität der Erwartung (gilt auch für nicht unabhängige Zufallsvariablen).
Alternative Lösung 2 für (i): Jede der 50 defekten Münzen hat eine Wahrscheinlichkeit $\alpha = \mathbb{P}\left[T | H_0\right]$ an das Casino weitergegeben zu werden. Die Tests der 50 defekten Münzen ist unabhängig somit ist $Z_{\mathrm{dC}} \sim \mathrm{Bin}(50, \alpha)$ verteilt. Für eine $\mathrm{Bin}(n, p)$-Verteilte Zufallsvariable ist der Erwartungswert np, somit gilt $\mathbb{E}\left[Z_{\mathrm{dC}}\right] = 50\alpha$.

(ii) Analog zu (i) berechnen wir die erwartete Anzahl an fairen Münzen, die dem Casino weiter gegeben werden

$$\begin{aligned}\mathbb{E}\left[Z_{\mathrm{fC}}\right] = \mathbb{E}\left[|(\{1, \ldots, 100\} \setminus D) \cap C|\right] = |\{1, \ldots, 100\} \setminus D| \; \mathbb{P}\left[T | H_0^{\complement}\right] \\ = 50(1-\beta) = \frac{50}{4} = 12.5.\end{aligned}$$

[a] Aus frequentistischer Perspektive wird $\mathcal{H}_0$ als deterministisch in einem hypothetischen Universum mit hypothetischem Wahrscheinlichkeitsmass $\mathbb{P}_{\mathcal{H}_0}$ betrachtet. Im Bayesianischen wird H_0 als Zufälliges Ereignis gesehen. Daher braucht man im Bayesianischen keine verschiedenen Wahrscheinlichkeitsmasse $\mathbb{P}_{\mathcal{H}_0}$ und $\mathbb{P}_{\mathcal{H}_1}$ für verschiedenen hypothetische Universen sondern nur ein Wahrscheinlichkeitsmass $\mathbb{P}$. Im Bayesianischen kann man dann verschiedene bedingte Wahrscheinlichkeitsmasse über die Standard-Notation von bedingten Wahrscheinlichkeiten $\mathbb{P}\left[\cdot | H_0\right]$ und $\mathbb{P}\left[\cdot | H_1\right]$ beschreiben.

[b] Die Zufällige Indexmenge D beschränkt sich nur auf Münzen i, die defekt sind, darum muss man darauf bedingen, dass die Münzen defekt sind. In der Vorlesung wurde nicht behandelt, wie man mit Summen über zufällige Index-Mengen umgeht und bedingte Erwartungswerte wurden nicht definiert. Darum sind im folgenden 2 alternative Lösungen gegeben, die nur Konzepte aus dem Skript verwenden.

6. X und Y [6 Punkte]

Betrachten Sie die Zufallsvariable (X, Y) mit Werten in $\mathbb{R}^2$, deren gemeinsame Dichte gegeben ist durch:

$$f_{X,Y}(x, y) = \begin{cases} \frac{3}{4}(x^2 + 2y) & \text{falls } 0 \leq x \leq 1 \text{ und } 0 \leq y \leq 1, \\ 0 & \text{sonst.} \end{cases}$$

6.1. [**2 Punkte**] Zeigen Sie, dass die Randdichten f_X und f_Y durch

$$f_X(x) = \mathbb{1}_{0\leq x\leq 1}\frac{3}{4}(x^2 + 1)$$
$$f_Y(y) = \mathbb{1}_{0\leq y\leq 1}\left(\frac{1}{4} + \frac{3}{2}y\right)$$

gegeben sind.

Lösung:

Randdichte f_X:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy = \mathbb{1}_{0\leq x\leq 1}\frac{3}{4}\int_0^1 (x^2+2y)dy = \mathbb{1}_{0\leq x\leq 1}\frac{3}{4}(x^2y + y^2)\Big|_{y=0}^1 = \mathbb{1}_{0\leq x\leq 1}\frac{3}{4}(x^2+1)$$

$\square$

Randdichte f_Y:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dx = \mathbb{1}_{0\leq y\leq 1}\frac{3}{4}\int_0^1 (x^2 + 2y)dx = \mathbb{1}_{0\leq y\leq 1}\frac{3}{4}(\frac{1}{3}x^3 + 2xy)\Big|_{x=0}^1$$
$$= \mathbb{1}_{0\leq y\leq 1}\frac{3}{4}\left(\frac{1}{3} + 2y\right) = \mathbb{1}_{0\leq y\leq 1}\left(\frac{1}{4} + \frac{3}{2}y\right).$$

$\square$

6.2. [**1 Punkt**] Sind X und Y unabhängig? Begrüde die Antwort.

Lösung:

Es gilt $f_X(x) \cdot f_Y(y) \neq f_{X,Y}(x, y)$, folglich sind X und Y nicht unabhängig.

6.3. Berechne

(i) [**1.5 Punkte**] den Erwartungswert von X

(ii) [**1.5 Punkte**] den Erwartungswert von XY.

Lösung:

(i)

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf_X(x)dx = \int_{-\infty}^{\infty} x\mathbb{1}_{0\leq x\leq 1}\frac{3}{4}(x^2+1)dx = \int_0^1 \frac{3}{4}(x^3+x)dx = \frac{3}{4}(\frac{1}{4}x^4 + \frac{1}{2}x^2)\Big|_{x=0}^1$$
$$= \frac{9}{16}.$$

Alternative Lösung für (i):

$$
\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X,Y}(x,y)dxdy \\
&= \int_0^1\int_0^1 x\frac{3}{4}(x^2+2y)dxdy \\
&= \int_0^1\int_0^1 \frac{3}{4}(x^3+2xy)dxdy \\
&= \int_0^1 \frac{3}{4}(\frac{1}{4}x^4+x^2y)\Big|_{x=0}^{1} dy \\
&= \int_0^1 \frac{3}{4}(\frac{1}{4}+y)dy \\
&= \frac{3}{4}(\frac{1}{4}y+\frac{1}{2}y^2)\Big|_{y=0}^{1} \\
&= \frac{3}{4}(\frac{1}{4}+\frac{1}{2}) \\
&= \frac{3}{4}(\frac{1+2}{4}) = \frac{3^2}{4^2} = \frac{9}{16}.
\end{aligned}
$$

(ii)

$$
\begin{aligned}
\mathbb{E}[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x,y)dxdy \\
&= \int_0^1\int_0^1 xy\frac{3}{4}(x^2+2y)dxdy \\
&= \int_0^1\int_0^1 \frac{3}{4}(x^3y+2xy^2)dxdy \\
&= \int_0^1 \frac{3}{4}(\frac{1}{4}x^4y+x^2y^2)\Big|_{x=0}^{1} dy \\
&= \int_0^1 \frac{3}{4}(\frac{1}{4}y+y^2)dy \\
&= \frac{3}{4}(\frac{1}{8}y^2+\frac{1}{3}y^3)\Big|_{y=0}^{1} \\
&= \frac{3}{4}(\frac{1}{8}+\frac{1}{3}) \\
&= \frac{3}{4}(\frac{3+8}{3\cdot 8}) = \frac{11}{4\cdot 8} = \frac{11}{32}.
\end{aligned}
$$

Alternative Lösung für (ii):

$$
\begin{aligned}
\mathbb{E}[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x,y)dydx \\
&= \int_0^1\int_0^1 xy\frac{3}{4}(x^2+2y)dydx \\
&= \int_0^1\int_0^1 \frac{3}{4}(x^3y+2xy^2)dydx \\
&= \int_0^1 \frac{3}{4}(\frac{1}{2}x^3y^2+\frac{2}{3}xy^3)\Big|_{y=0}^{1} dx \\
&= \int_0^1 \frac{3}{4}(\frac{1}{2}x^3+\frac{2}{3}x)dx \\
&= \frac{3}{4}(\frac{1}{8}x^4+\frac{1}{3}x^2)\Big|_{x=0}^{1} \\
&= \frac{3}{4}(\frac{1}{8}+\frac{1}{3}) \\
&= \frac{3}{4}(\frac{3+8}{3\cdot 8}) = \frac{11}{4\cdot 8} = \frac{11}{32}.
\end{aligned}
$$

Diskussion: Da X und Y nicht unabhängig sind, muss $\mathbb{E}[XY]$ nicht das gleiche sein, wie $\mathbb{E}[X]\,\mathbb{E}[Y]$. Hier in diesem Beispiel gilt $\mathbb{E}[XY] = \frac{11}{32} \neq \frac{9}{16}\frac{5}{8} = \mathbb{E}[X]\,\mathbb{E}[Y]$.

Tabelle der Standardnormalverteilung

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Zum Beispiel ist $\mathbb{P}[Z \leq 1.96] = \Phi(1.96) \approx 0.975$.