



Induktive Statistik

Musterlösungen zum Lernheft

Copyright © 2021 StudyHelp
StudyHelp GmbH, Paderborn
WWW.STUDYHELP.DE

Autor: Dr. Andreas Stahl

Redaktion & Satz: Carlo Oberkönig
Kontakt: verlag@studyhelp.de
Umschlaggestaltung, Illustration: StudyHelp GmbH

Das Werk, einschließlich seiner Teile, ist urheberrechtlich geschützt. Jede Verwertung ist ohne Zustimmung des Verlages und des Autors unzulässig. Dies gilt insbesondere für die elektronische oder sonstige Vervielfältigung, Übersetzung, Verbreitung und öffentliche Zugänglichmachung. Dies gilt auch für Intranets von Schulen und sonstigen Bildungseinrichtungen.

Auf verschiedenen Seiten dieses Buches befinden sich Verweise (Links) auf Internet-Adressen. Haftungshinweis: Trotz sorgfältiger inhaltlicher Kontrolle wird die Haftung für die Inhalte der externen Seiten ausgeschlossen. Für den Inhalt dieser externen Seiten sind ausschließlich deren Betreiber verantwortlich. Sollten Sie bei dem angegebenen Inhalt des Anbieters dieser Seite auf kostenpflichtige, illegale oder anstößige Inhalte treffen, so bedauern wir dies ausdrücklich und bitten Sie, uns umgehend per E-Mail davon in Kenntnis zu setzen, damit beim Nachdruck der Verweis gelöscht wird.

E-Book

Inhalt

1	Lösungen zu Basics	5
2	Lösungen zu Wahrscheinlichkeitstheorie	9
2.1	zu Grundbegriffe	9
2.2	zu Kombinatorik	9
2.3	zu Der Laplace'sche Wahrscheinlichkeitsraum	14
2.4	zu Wahrscheinlichkeitsmaß und Zufallsvariablen	18
2.5	zu Bedingte Wahrscheinlichkeit	18
2.6	zu Stochastische Unabhängigkeit	22
2.7	zu Diskreten Zufallsvariablen	25
2.8	zu Spezielle diskrete Verteilungen	30
2.9	zu Stetige Zufallsvariablen	34
2.10	zu Spezielle stetige Verteilungen	38
2.11	zu Mehrdimensionale Verteilungen	41
2.12	zu Abschätzungen und Grenzwertsätze	45
3	Lösungen zu Inferenzstatistik	49
3.1	zu Grundbegriffe	49
3.2	zu Parameterschätzung	49
3.3	zu Hypothesentests	56
3.4	zu Parametrische Einstichproben-Tests	58
3.5	zu Parametrische Zweistichproben-Tests	61
3.6	zu Nichtparametrische Tests	68

1 Lösungen zu Basics

zu A.1.1:

Bei der Menge A ist jedes Element voneinander zu unterscheiden, sodass es sich um eine Menge mit insgesamt 9 Elementen handelt. B hingegen besteht nicht aus wohlunterscheidbaren Elementen. Dies liegt am zweifach auftretenden Element ab . Keine Begründung gegen eine Menge ist das Wiederholen des Buchstabens a . Denn sowohl a , aa als auch aaa sind unterscheidbar. Dies ist gut mit Zahlen vergleichbar, denn die Zahlen 1, 11 und 111 bedeuteten jeweils etwas anderes.

zu A.1.2:

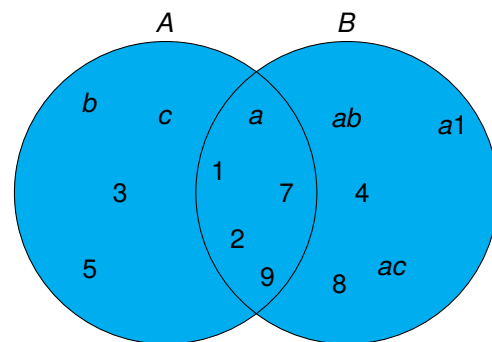
Vereinigung:

Beginnen wir mit der Vereinigung $A \cup B$. Die Vereinigung enthält alle Mengen, die in A oder in B liegen, wir müssen nur beachten, kein Element doppelt zu nennen.

Das Ergebnis lautet

$$C = \{a, b, c, 1, 2, 3, 5, 7, 9, ab, ac, a1, 4, 8\}.$$

Visuell lässt sich dies wie folgt darstellen, die blaue Fläche verdeutlicht dabei die Menge C :



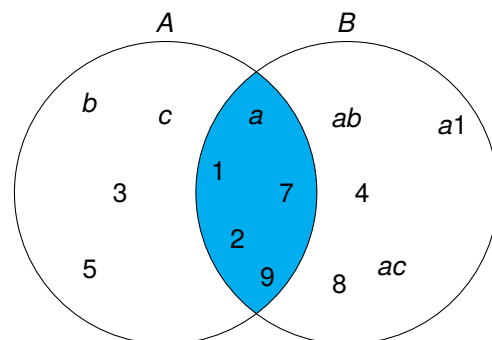
Schnitt:

Den Schnitt $A \cap B$ erhalten wir, indem wir Elemente wählen, die in A und in B enthalten sind.

Das Ergebnis lautet

$$C = \{a, 1, 2, 7, 9\}.$$

Visuell lässt sich der Schnitt durch folgenden blauen Bereich darstellen:



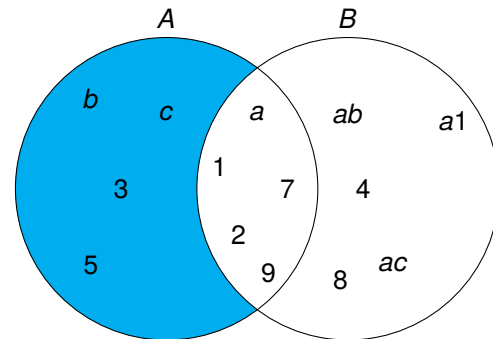
Differenz A ohne B:

Die Menge $A \setminus B =$ „A ohne B“ enthält alle Elemente von A, die nicht gleichzeitig in B liegen.

Hier lautet das Ergebnis

$$C = \{b, c, 3, 5\}.$$

Visuell entspricht C dieser blauen Fläche:

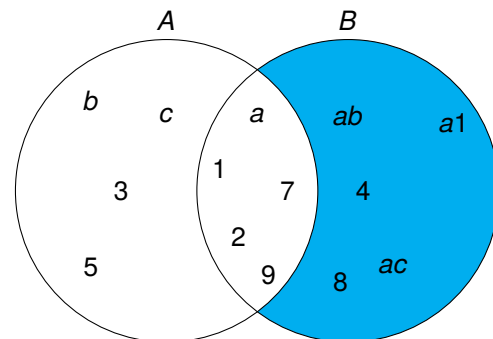
Differenz B ohne A:

Ebenso erhalten wir die Menge $B \setminus A =$ „B ohne A“ durch die Elemente aus der Menge B, die nicht gleichzeitig in A liegen.

Das Ergebnis lautet in diesem Fall

$$C = \{ab, ac, a1, 4, 8\}.$$

Visuell entspricht C dieser blauen Fläche:

**zu A.1.3:**

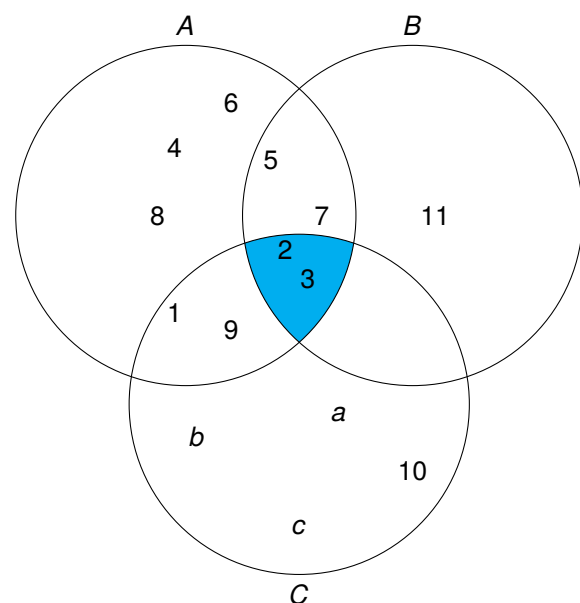
- a) Für den Schnitt $A \cap B \cap C$ können wir entweder alle Elemente suchen, die sowohl in A, in B als auch in C vorkommen oder wir gehen sukzessive vor, indem wir zuerst den Schnitt $A \cap B$ bestimmen und das Ergebnis anschließend mit C schneiden, denn es gilt folgende Rechenregel:

$$A \cap B \cap C = (A \cap B) \cap C$$

Da die Mengen übersichtlich sind, suchen wir diejenigen Elemente, die in jeder Menge enthalten sind. Dabei starten wir mit der Menge B, da diese die wenigsten Elemente besitzt und erhalten als Lösung D:

$$D = A \cap B \cap C = \{2, 3\}$$

Visuell lässt sich der Schnitt über alle Mengen darstellen durch:

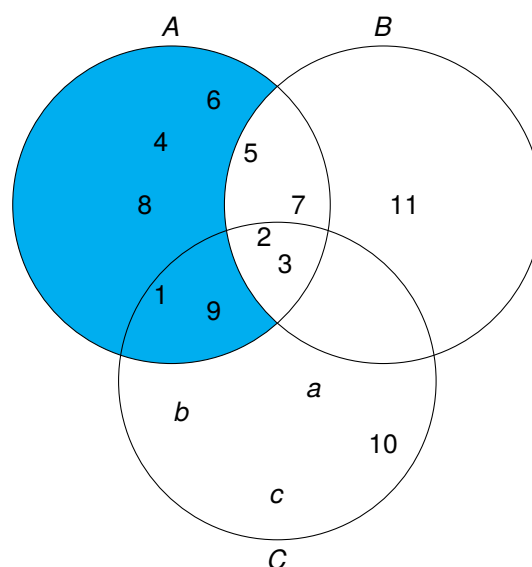


- b) Diese Aufgabe lässt sich ebenfalls dadurch lösen, dass wir alle Elemente suchen, die in der entstehenden Menge enthalten sind. Da der Ausdruck etwas komplizierter als in der ersten Teilaufgabe ist, gehen wir hier sukzessive vor.

Wir erhalten zunächst die Menge

$$A \setminus B = \{1, 4, 6, 8, 9\}.$$

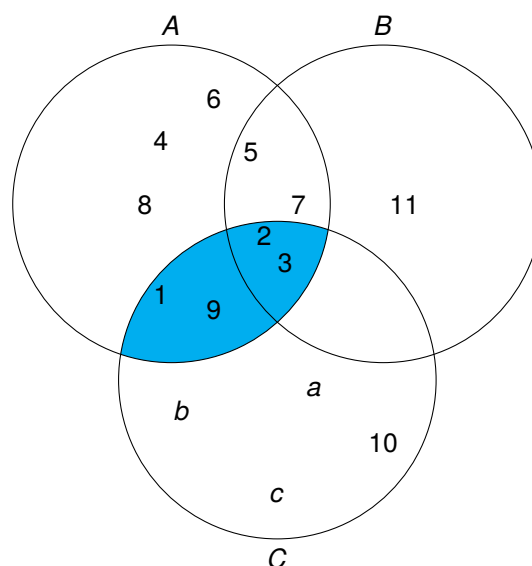
Visuell erhalten wir diese blaue Fläche:



Als weitere Menge erhalten wir den Schnitt

$$A \cap C = \{1, 2, 3, 9\}.$$

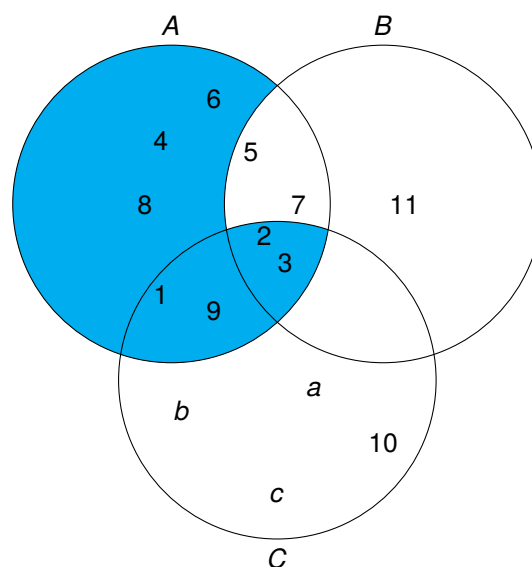
Diesen können wir visuell durch folgende blaue Fläche darstellen:



Vereinigen wir diese beiden Mengen erhalten wir die Menge D

$$D = (A \setminus B) \cup (A \cap C) = \{1, 2, 3, 4, 6, 8, 9\}.$$

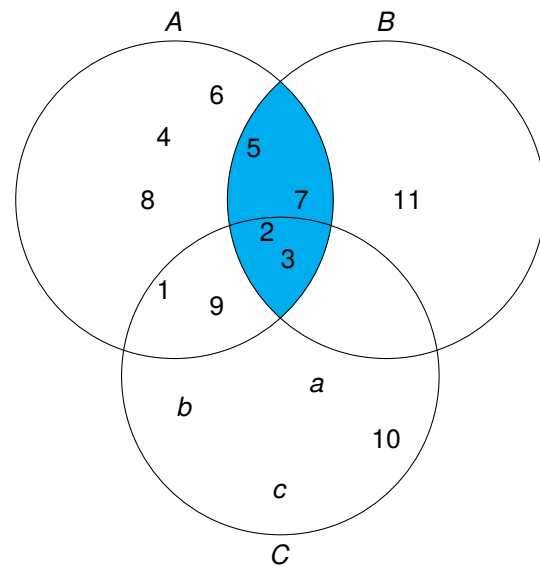
Visuell stellen wir D in folgendem Bild durch die blaue Fläche als Vereinigung der beiden Mengen dar, die wir soeben berechnet haben:



- c) In diesem Beispiel bestimmen wir sukzessive alle Elemente der beiden Teile der Vereinigung und fügen diese zusammen. Dazu bestimmen wir als Erstes den Schnitt der Mengen A und B :

$$A \cap B = \{2, 3, 5, 7\}$$

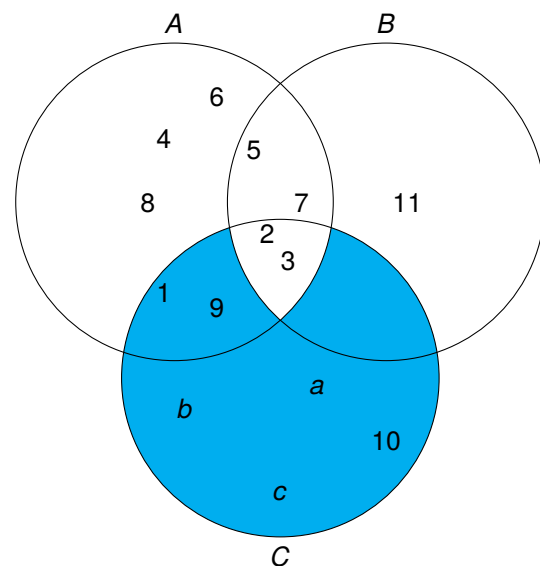
Dargestellt in blau in folgender Grafik:



Zudem erhalten wir für den zweiten Teil der Vereinigung

$$C \setminus (A \cap B) = \{10, a, b, c\}.$$

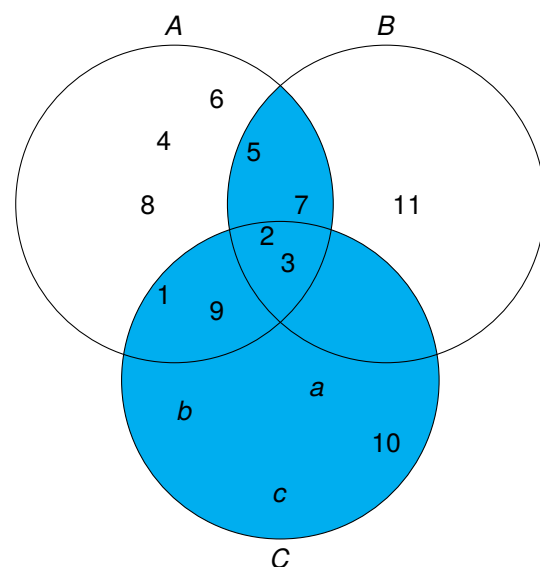
Diese stellen wir in folgender Grafik in blau dar:



Insgesamt ergibt sich demnach als Lösung

$$\begin{aligned} D &= (A \cap B) \cup (C \setminus (A \cap B)) \\ &= \{1, 2, 3, 5, 7, 9, 10, a, b, c\}. \end{aligned}$$

Visuell ergibt sich D als die beiden obigen Flächen durch:



2 Lösungen zu Wahrscheinlichkeitstheorie

2.1 zu Grundbegriffe

keine Aufgaben vorhanden

2.2 zu Kombinatorik

zu A.2.1:

Wir bestimmen die Anzahl der Menge Ω im vorliegenden Fall:

a) Für eine k -Kombination ohne Wiederholung gilt die Formel

$$|\Omega| = \binom{n}{k}.$$

Wir erhalten mit $k = 5$ und $n = 7$ den Wert

$$|\Omega| = \binom{7}{5} = \frac{7!}{(7-5)! \cdot 5!} = \frac{7!}{2! \cdot 5!} = \frac{\overbrace{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7}^{=7!}}{\underbrace{2! \cdot 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}_{=5!}} = \frac{6 \cdot 7}{1 \cdot 2} = 21.$$

b) Für eine k -Variation ohne Wiederholung gilt die Formel

$$|\Omega| = \frac{n!}{(n-k)!}.$$

Im Falle einer 3 -Variation ohne Wiederholung mit $n = 12$ erhalten wir

$$|\Omega| = \frac{12!}{(12-3)!} = \frac{12!}{9!} = \frac{1 \cdot 2 \cdot \dots \cdot 8 \cdot 9 \cdot 10 \cdot 11 \cdot 12}{1 \cdot 2 \cdot \dots \cdot 8 \cdot 9} = 10 \cdot 11 \cdot 12 = 1320.$$

c) Für eine k -Kombination mit Wiederholung gilt die Formel

$$|\Omega| = \binom{n+k-1}{k}.$$

Im Falle von $k = 7$ und $n = 9$ erhalten wir

$$|\Omega| = \binom{9+7-1}{7} = \binom{15}{7} = \frac{15!}{(15-7)! \cdot 7!} = \frac{15!}{8! \cdot 7!} = \frac{1 \cdot 2 \cdot \dots \cdot 7 \cdot 8 \cdot 9 \cdot \dots \cdot 15}{1 \cdot 2 \cdot \dots \cdot 7 \cdot 8 \cdot 7!} = \frac{9 \cdot \dots \cdot 15}{7!} = 6435.$$

d) Für eine $k = n$ -Permutation mit Wiederholung gilt die Formel

$$|\Omega| = \frac{n!}{k_1! \cdot \dots \cdot k_n!}.$$

Mit $k = n = 6$ folgt für $k_1 + \dots + k_n = 6$

$$|\Omega| = \frac{6!}{k_1! \cdot \dots \cdot k_n!} = \frac{720}{k_1! \cdot \dots \cdot k_n!}.$$

zu A.2.2:

Wir berechnen jeweils die Anzahl der Möglichkeiten:

- a) Ein Skatblatt besteht aus den Werten 7, 8, 9, 10, Bube, Dame, König und Ass, die jeweils in vier verschiedenen Farben vorhanden sind. Die Unterteilung „Bild“ und „Zahl“ unterteilt die Karten in zwei Stapel zu je 16 Karten.

Mit Zurücklegen bedeutet in diesem Fall, dass es eine Wiederholung geben kann. Ein bereits gezogenes Bild kann durch das Zurücklegen erneut gezogen werden. Es handelt sich sowohl beim Ziehen der Bildkarten als auch beim Ziehen der Zahlkarten damit um eine Kombination mit Wiederholung, da aus der Aufgabenstellung nicht hervor geht, dass die Reihenfolge wichtig ist. Wir erhalten mit $n = 16$ und $k = 2$ die Anzahl an Möglichkeiten, zwei Bildkarten zu ziehen durch

$$\binom{n+k-1}{k} = \binom{16+2-1}{2} = \binom{17}{2} = \frac{17!}{15! \cdot 2!} = 136.$$

Die Anzahl, eine Karte mit einer Zahl zu ziehen, beträgt 16, da es 16 verschiedene Karten mit einer Zahl gibt. Wir erhalten das gleiche Ergebnis, wenn wir das Ziehen ebenso durch eine Kombination mit $k = 1$ und $n = 16$ interpretieren durch

$$\binom{n+k-1}{k} = \binom{16+1-1}{1} = \binom{16}{1} = \frac{16!}{15! \cdot 1!} = 16.$$

Insgesamt müssen wir die Möglichkeiten nun noch miteinander multiplizieren, da für jede Möglichkeit, zwei Bildkarten zu ziehen, wiederum jeweils 16 Möglichkeiten existieren, eine Zahl zu ziehen. Damit erhalten wir insgesamt als Lösung

$$\binom{17}{2} \cdot \binom{16}{1} = 136 \cdot 16 = 2176.$$

- b) Die Information, dass genau zwei der Würfe eine 6 ergeben sollen, schränkt die Anzahl der Möglichkeiten stark ein. Denn nun sind diese Ergebnisse nicht mehr variabel. Betrachten wir den verbleibenden Wurf, so können dort noch die Werte von 1 bis 5 gewürfelt werden. Aus der Aufgabenstellung ist sprachlich nicht ersichtlich, ob die Reihenfolge entscheidend ist oder nicht, daher präsentieren wir für beide Interpretationen eine Lösung:

ohne Beachtung der Reihenfolge:

Ist die Reihenfolge irrelevant, so betrachten wir den Fall, dass die ersten beiden Würfe bereits eine 6 ergeben haben. In diesem Fall gibt es noch 5 Möglichkeiten, dass wir genau zwei 6-en würfeln, denn genau dann, wenn der letzte Wurf keine 6 ergibt. Da eine andere Reihenfolge keinen Einfluss hat, erhalten wir als Lösung diese 5 Möglichkeiten.

mit Beachtung der Reihenfolge:

Ist es indes für uns interessant, welcher Wurf welches Ergebnis erzielt hat, betrachten wir die Anzahl an Möglichkeiten, den variablen Wurf, der keine 6 ergibt zu verteilen. Dieser kann an erster, zweiter oder dritter Stelle auftreten. Da für jede dieser Stellen mit obiger Begründung 5 Möglichkeiten einhergehen, erhalten wir insgesamt $5 \cdot 5 \cdot 5 = 125$ Möglichkeiten.

- c) Die Frage lässt sich übersetzen in „Wie viele Möglichkeiten gibt es, aus 20 Studierenden zwei ohne Wiederholung und unter Beachtung der Reihenfolge zu wählen?“. Die Reihenfolge ist in diesem Fall interessant, da sich der Preis für den ersten und zweiten Platz unterscheidet. Wir befinden uns demnach im Falle einer Variation ohne Wiederholung mit $n = 20$ und $k = 2$. Damit ergibt sich die Anzahl der Konstellationen durch

$$\frac{n!}{(n-k)!} = \frac{20!}{(20-2)!} = \frac{20!}{18!} = \frac{1 \cdot 2 \cdot \dots \cdot 17 \cdot 18 \cdot 19 \cdot 20}{1 \cdot 2 \cdot \dots \cdot 17 \cdot 18} = 20 \cdot 19 = 380.$$

- d) Bei dieser Frage unterstellen wir, dass alle Läufer ans Ziel kommen und niemand disqualifiziert wird, sodass wir stets allen Teilnehmern einen Platz zuordnen können. Daher ergibt sich die Anzahl durch eine Permutation. Da ein Läufer nicht gleichzeitig an zwei Plätzen sein kann, liegt eine Permutation ohne Wiederholung vor und erhalten die Anzahl an Möglichkeiten der Platzierungen mit $n = 8$ durch

$$n! = 8! = 40320.$$

- e) Durch die Aufgabenstellung erhalten wir eine sechsstellige Zahl, bestehend aus den Ziffern 1 bis 6, bei der keine Ziffer mehrfach auftreten darf. Wir befinden uns erneut im Modell der Permutation ohne Wiederholung, da alle Zahlen von 1 bis 6 zwingend vergeben werden müssen, um eine sechsstellige Zahl zu erhalten. Die Reihenfolge ist dabei entscheidend, sodass wir für die Anzahl der möglichen Zahlen mit $n = 6$ den Wert

$$n! = 6! = 720$$

erhalten.

- f) Bei diesem Experiment erkennen wir direkt, dass eine Wiederholung möglich ist, da beim Werfen des Würfels häufiger die gleiche Zahl geworfen werden kann. Wir verteilen allerdings nicht zwingend alle 6 Werte des Würfels, sodass wir die Anzahl nicht mit Hilfe der Formel der Permutation berechnen können. Aus der Beschreibung wird nicht direkt ersichtlich, dass die Reihenfolge wichtig ist, jedoch können wir annehmen, dass diese bei einem Vorgang wie dem „Loseziehen“ eine Rolle spielt. Daher können wir die Anzahl durch eine Variation mit Wiederholung und $n = 6 =$ Anzahl der zu würfelnden Möglichkeiten sowie $k = 5 =$ Anzahl der Würfe berechnen:

$$n^k = 6^5 = 7776.$$

Eine Schwierigkeit besteht hier in der korrekten Zuweisung von k und n . Wir können dies ebenfalls so zählen:

Der erste Wurf hat 6 Möglichkeiten, der zweite Wurf hat 6 Möglichkeiten, ..., , der fünfte Wurf hat 6 Möglichkeiten. Dies ergibt insgesamt $6 \cdot 6 \cdot 6 \cdot 6 \cdot 6 = 6^5$ Möglichkeiten.

zu A.2.3:

Wir berechnen jeweils die Anzahl der Möglichkeiten:

- a) Die Anzahl an Möglichkeiten von 10 verschiedenen Personen eines Gremiums erhalten wir, indem wir die Anzahl an Möglichkeiten der 5 aus 50 Mädchen mit der Anzahl der 5 aus 40 Jungen multiplizieren. Bei beiden Vorgängen handelt es sich um eine Kombination ohne Wiederholung, da die Reihenfolge, in der die Personen im Gremium aufgezählt werden, irrelevant ist. Es ist nur nach der Konstellation der 10 Personen im Gremium gefragt und nicht, welcher zuerst oder zuletzt bestimmt wurde. Eine Wiederholung können wir zudem ausschließen, da keine Person doppelt für das Gremium ausgewählt wird.

Im Falle der Mädchen erhalten wir daher mit $n = 50$ und $k = 5$ die Anzahl durch

$$\binom{n}{k} = \binom{50}{5} = \frac{50!}{(50-5)! \cdot 5!} = \frac{50!}{45! \cdot 5!} = \frac{\overbrace{1 \cdot 2 \cdot \dots \cdot 44 \cdot 45 \cdot 46 \cdot \dots \cdot 50}^{=50!}}{\underbrace{1 \cdot 2 \cdot \dots \cdot 44 \cdot 45 \cdot 5!}_{=45!}} = \frac{46 \cdot 47 \cdot 48 \cdot 49 \cdot 50}{120} = 2.118.760.$$

Für die Anzahl, 5 Jungen ohne Wiederholung und ohne Beachtung der Reihenfolge aus 40 auszuwählen, erhalten wir analog

$$\binom{n}{k} = \binom{40}{5} = \frac{40!}{(40-5)! \cdot 5!} = \frac{40!}{35! \cdot 5!} = \frac{\overbrace{1 \cdot 2 \cdot \dots \cdot 34 \cdot 35 \cdot 36 \cdot \dots \cdot 40}^{=40!}}{\underbrace{1 \cdot 2 \cdot \dots \cdot 34 \cdot 35 \cdot 5!}_{=35!}} = \frac{36 \cdot 37 \cdot 38 \cdot 39 \cdot 40}{120} = 658.008.$$

Insgesamt erhalten wir demnach $2.118.760 \cdot 658.008 = 1.394.161.030.080$ verschiedene Möglichkeiten das Gremium zu besetzen.

- b) Ein Themenabend ist in der Beschreibung nicht genauer charakterisiert, sodass auch ein Themenabend aus einer Mischung von mehreren Musikrichtungen existieren kann. Beispielsweise könnten zwei Musikrichtungen durch „RnB“ und „House“ gegeben sein und einen „RnB and House“ Themenabend bilden.

Für eine feste Anzahl k an Musikrichtungen eines Abends, lässt sich die Anzahl der Möglichkeiten, diese aus ($n =$) 5 Musikrichtungen zu bilden, durch eine k -Kombination modellieren, da die Reihenfolge der Musikrichtungen für den Themenabend keine Rolle spielt. Wir setzen einen „RnB and House“ Themenabend mit einem „House and RnB“ Themenabend gleich. Zudem ergibt ein Themenabend, der eine Musikrichtung mehrfach enthält, wie beispielsweise ein „House and House“ Themenabend, wenig Sinn, sodass wir keine Wiederholung zulassen. Summieren wir nun die Anzahlen der Möglichkeiten für jedes k auf, erhalten wir die Gesamtanzahl an möglichen Themenabenden. Zunächst erhalten wir:

$$\begin{aligned} k=1 & : \binom{5}{1} = \frac{5!}{(5-1)! \cdot 1!} = \frac{5!}{4!} = 5 \\ k=2 & : \binom{5}{2} = \frac{5!}{(5-2)! \cdot 2!} = \frac{5!}{3! \cdot 2!} = 10 \\ k=3 & : \binom{5}{3} = \frac{5!}{(5-3)! \cdot 3!} = \frac{5!}{2! \cdot 3!} = 10 \end{aligned}$$

$$k=4 : \binom{5}{4} = \frac{5!}{(5-4)! \cdot 4!} = \frac{5!}{1! \cdot 4!} = 5$$

$$k=5 : \binom{5}{5} = \frac{5!}{(5-5)! \cdot 5!} = \frac{5!}{0! \cdot 5!} = 1$$

Dabei haben wir benutzt, dass $0! = 1$ gilt und es bedeutet beispielsweise $k=3 : \binom{5}{3} = 10$, dass es 10 Möglichkeiten gibt, aus 5 Musikrichtungen einen Themenabend mit drei verschiedenen dieser Musikrichtungen zu bilden. Die Gesamtanzahl erhalten wir nun durch

$$\sum_{k=1}^5 \binom{5}{k} = \binom{5}{1} + \binom{5}{2} + \binom{5}{3} + \binom{5}{4} + \binom{5}{5} = 5 + 10 + 10 + 5 + 1 = 31.$$

c) Die Anzahl dieses Experiments berechnen wir mit zwei verschiedenen Methoden.

Variante 1:

Als erste Möglichkeit berechnen wir die Anzahl in zwei Schritten. Wir bestimmen zunächst die Anzahl an Möglichkeiten, von 8 Personen 4 auszuwählen und anschließend die Anzahl der Möglichkeiten, 4 Personen an der Kasse anzuordnen. Die Gesamtanzahl erhalten wir abschließend durch Multiplikation der beiden zuvor berechneten Zahlen.

Aus einer Gruppe von 8 Personen 4 auszuwählen, modellieren wir erneut mit einer Kombination mit $n=8$ und $k=4$, da die Reihenfolge, in der wir die Freunde auswählen, für diesen Teil des Experiments keine Rolle spielt, da die Anordnung an der Kasse vorgenommen wird. Wir erhalten demnach

$$\binom{n}{k} = \binom{8}{4} = \frac{8!}{(8-4)! \cdot 4!} = \frac{8!}{4! \cdot 4!} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 4!} = \frac{5 \cdot 6 \cdot 7 \cdot 8}{24} = 70$$

Möglichkeiten.

Die Anzahl an Möglichkeiten, eine Gruppe von 4 Personen an einer Kasse anzuordnen, entspricht einer Permutation ohne Wiederholung, da allen Personen ein Platz zugeordnet wird und keine Person mehrfach in der Schlange auftritt. Die Anzahl erhalten wir in diesem Fall durch

$$n! = 4! = 24.$$

Insgesamt ergibt sich somit eine Anzahl von $70 \cdot 24 = 1680$.

Variante 2:

Als zweite Möglichkeit interpretieren wir das Experiment so, dass wir 4 aus einer Gruppe von 8 Personen unter Berücksichtigung der Reihenfolge und ohne Wiederholung auswählen. Die Reihenfolge entspricht dann der Anordnung an der Kasse. Somit können wir den Vorgang auch durch eine Variation ohne Wiederholung modellieren und erhalten die Gesamtanzahl durch

$$\frac{n!}{(n-k)!} = \frac{8!}{(8-4)!} = \frac{8!}{4!} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8}{1 \cdot 2 \cdot 3 \cdot 4} = 5 \cdot 6 \cdot 7 \cdot 8 = 1680.$$

2.3 zu Der Laplace'sche Wahrscheinlichkeitsraum

In diesem Abschnitt präsentieren wir die Lösungen des Kapitels 2.3. Manche der Aufgaben lassen sich ebenfalls mit Hilfe von Methoden anderer Kapitel lösen, beispielsweise durch Anwenden spezieller diskreter Verteilungen. Darauf werden wir bei den Lösungen verzichten, da wir explizit das elementare Vorgehen im Laplace'schen Wahrscheinlichkeitsraum betrachten möchten.

zu A.2.4:

Wir bestimmen die jeweilige Wahrscheinlichkeit:

- a) Zum Lösen dieser Aufgabe berechnen wir zunächst die möglichen Ausgänge des Experiments. Diese können wir in diesem Fall gesammelt aufzählen. Wir bezeichnen mit „K“ den Ausgang „Kopf“ sowie mit „Z“ den Ausgang „Zahl“ und erhalten beim zweifachen Münzwurf

$$\Omega = \{(K, K), (K, Z), (Z, K), (Z, Z)\}.$$

Somit folgt $|\Omega| = 4$. Die Menge A beschreibe die möglichen Ausgänge für das gewünschte Ereignis, zweimal Kopf zu werfen. Damit folgt

$$A = \{(K, K)\}$$

und somit $|A| = 1$. Insgesamt erhalten wir

$$\mathbb{P}(\text{„zweimal Kopf“}) = \mathbb{P}(A) = \mathbb{P}(\{(K, K)\}) = \frac{|A|}{|\Omega|} = \frac{1}{4} = 25\%.$$

- b) Bei diesem Experiment bestimmen wir erneut die Anzahl an möglichen Ausgängen:

$$\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

Damit erhalten wir $|\Omega| = 10$. Anschließend betrachten wir die Anzahl der günstigen Möglichkeiten. Eine Möglichkeit ist nach der Aufgabenstellung für uns günstig, wenn es eine Primzahl ist. Somit erhalten wir

$$A = \{2, 3, 5, 7\}$$

und somit $|A| = 4$. Insgesamt erhalten wir die Wahrscheinlichkeit

$$\mathbb{P}(\text{„Drehen einer Primzahl“}) = \mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{4}{10} = 40\%.$$

zu A.2.5:

Wir bestimmen die jeweilige Wahrscheinlichkeit:

- a) Die möglichen Ereignisse erhalten wir durch alle Paare, bei denen jeder Eintrag zwischen 1 und 6 liegen kann. Dies formulieren wir mathematisch durch

$$\Omega = \{(\omega_1, \omega_2) : \omega_i \in \{1, 2, 3, 4, 5, 6\} \text{ für } i \in \{1, 2\}\} = \{(1, 1), \dots, (1, 6), (2, 1), \dots, (2, 6), \dots, (6, 6)\}.$$

Für die Anzahl erhalten wir

$$|\Omega| = 36 = 6 \cdot 6.$$

Das Produkt $6 \cdot 6$ zeigt, dass man die Anzahl in diesem Fall auch durch Multiplikation der Anzahlen der einzelnen Würfel bekommt. Eine Begründung ist die zugrundeliegende kombinatorische Natur. Wir können das Werfen zweier Würfel als Variation mit $k = 2$ und $n = 6$ interpretieren. Damit ergeben sich $6^2 = 6 \cdot 6$ Möglichkeiten.

Für die Menge der möglichen Ereignisse A gilt mathematisch

$$A = \{(\omega_1, \omega_2) \in \Omega : \omega_1 \cdot \omega_2 > 10\}.$$

Diese Menge gilt es nun zu zählen. Dafür gehen wir systematisch wie folgt vor:

$$\begin{aligned} &(2, 6), (3, 6), (4, 6), (5, 6), (6, 6) \\ &(3, 5), (4, 5), (5, 5), (6, 5) \\ &(3, 4), (4, 4), (5, 4), (6, 4) \\ &(4, 3), (5, 3), (6, 3) \\ &(6, 2) \end{aligned}$$

Somit erhalten wir $|A| = 17$ und damit die Wahrscheinlichkeit

$$\mathbb{P}(\text{„Produkt größer 10“}) = \mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{17}{36} \approx 47,22\%.$$

b) Ähnlich wie in der vorigen Aufgabe können wir Ω mathematisch durch

$$\Omega = \{(\omega_1, \omega_2) : \omega_1 \in \{1, 2, 3, 4, 5, 6\} \text{ und } \omega_2 \in \{1, \dots, 10\}\}$$

definieren. Die Anzahl erhalten wir erneut durch Multiplikation der Augenzahlen der Würfel durch

$$|\Omega| = 6 \cdot 10 = 60.$$

Die Menge A der günstigen Möglichkeiten definieren wir durch

$$A = \{(\omega_1, \omega_2) \in \Omega : \omega_1 > \omega_2\}.$$

Die Möglichkeiten lassen sich wieder strukturiert zählen:

$$\begin{aligned} &(6, 1), (6, 2), (6, 3), (6, 4), (6, 5) \\ &(5, 1), (5, 2), (5, 3), (5, 4) \\ &(4, 1), (4, 2), (4, 3) \\ &(3, 1), (3, 2) \\ &(2, 1) \end{aligned}$$

Damit ergibt sich

$$|A| = 5 + 4 + 3 + 2 + 1 = 15.$$

Insgesamt erhalten wir die Wahrscheinlichkeit

$$\mathbb{P}(\text{„Würfel 1} > \text{Würfel 2“}) = \mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{15}{60} = \frac{1}{4} = 25\%.$$

c) Zunächst beschreiben wir Ω mathematisch durch

$$\Omega = \{\omega_1, \omega_2, \omega_3 : \omega_i \in \{A, \dots, Z\} \text{ für } i \in \{1, 2, 3\}\}.$$

Wir benutzen an dieser Stelle erneut Erkenntnisse aus dem vorigen Kapitel der Kombinatorik. Für die möglichen Ereignisse ziehen wir 3 Buchstaben mit Wiederholung und unter Beachtung der Reihenfolge aus dem Alphabet mit 26 möglichen Buchstaben. Daher erhalten wir die Anzahl der möglichen Ereignisse durch

$$|\Omega| = 26^3 = 17576.$$

Die Menge der günstigen Möglichkeiten definieren wir durch

$$A = \{\omega_1, \omega_2, \omega_3 \in \Omega : \omega_i \in \{a, e, i, o, u\} \text{ für } i \in \{1, 2, 3\}\}.$$

Die Anzahl der Elemente von A beschreiben wir erneut durch eine Variation. Dieses Mal ist $k = 3$ und $n = 5$. Somit erhalten wir

$$|A| = 5^3 = 125.$$

Damit ergibt sich insgesamt die Wahrscheinlichkeit

$$\mathbb{P}(\text{Wort aus Vokalen}) = \mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{125}{17576} \approx 0,71\%.$$

zu A.2.6:

Wir bestimmen die jeweilige Wahrscheinlichkeit:

a) Bei dieser Aufgabe fokussieren wir uns auf die Lage der beiden unterschriebenen Karten. Insgesamt enthält ein Kartenspiel 4 Assen, sodass wir bei n Kartenspielen auf eine Anzahl von $4n$ Assen kommen. Statt die Assen gedanklich auf zwei Spieler aufzuteilen, nummerieren wir alle ausgeteilten Karten von 1 bis $4n$ durch und merken uns, dass die ersten $2n$ Karten, also die mit den Positionen 1 bis $2n$ zu Person A gehören und die zweiten $2n$ Karten, mit den Positionen $2n + 1$ bis $4n$ zu Person B . Wir beschreiben die Menge aller Möglichkeiten also durch

$$\Omega = \{(\omega_1, \omega_2) : \omega_i \in \{1, 2, \dots, 4n\} \text{ für } i \in \{1, 2\} \text{ und } \omega_1 \neq \omega_2\}.$$

Ein Paar (ω_1, ω_2) beschreibt nun die Lage der unterschriebenen Assen. Wir fordern in der Definition $\omega_1 \neq \omega_2$, da sich nicht beide unterschriebenen Karten an der gleichen Stelle befinden können. $(1, 2n + 1)$ würde beispielsweise bedeuten, dass die erste Karte und damit die erste Karte der Person A eine Unterschrift trägt und die Karte an Position $2n + 1$ und damit die erste Karte der Person B ebenfalls unterschrieben wurde. Die Anzahl von Ω erhalten wir durch eine Variation ohne Wiederholung mit $4n$ und $k = 2$. Eine Modellierung ohne Berücksichtigung der Reihenfolge ist denkbar, allerdings wird das Zählen im Laufe der Aufgabe komplizierter. Damit ergibt sich die Anzahl aller Möglichkeiten durch

$$|\Omega| = \frac{(4n)!}{(4n-2)!} = \frac{(4n)!}{(4n-2)!} = \frac{1 \cdot \dots \cdot (4n-2) \cdot (4n-1) \cdot (4n)}{1 \cdot \dots \cdot (4n-2)} = 4n \cdot (4n-1).$$

Für die günstigen Möglichkeiten gilt nun, dass sich entweder beide unterschriebenen Karten unter den ersten $2n$ Assen befinden und damit Positionen zwischen 1 und $2n$ besitzen oder unter den zweiten $2n$ Karten, also die Positionen $2n + 1$ bis $4n$ besitzen. Mathematisch sieht dies wie folgt aus:

$$A = \{(\omega_1, \omega_2) \in \Omega : \omega_i \in \{1, \dots, 2n\} \text{ für } i \in 1, 2 \text{ oder } \omega_i \in \{2n + 1, \dots, 4n\} \text{ für } i \in 1, 2 \text{ mit } \omega_1 \neq \omega_2\}$$

Betrachten wir die erste Möglichkeit, dass sich beide unterschriebenen Karten bei Person A befinden, so besitzt dies erneut den Charakter einer Variation ohne Wiederholung mit $2n$ und $k = 2$. Da wir für Person B die gleiche Anzahl an Möglichkeiten erhalten, gilt insgesamt

$$|A| = \frac{\overbrace{(2n)!}^{\text{beide bei Person 1}}}{(2n-2)!} + \frac{\overbrace{(2n)!}^{\text{beide bei Person 2}}}{(2n-2)!} = 2 \cdot \frac{(2n)!}{(2n-2)!} = 2 \cdot \frac{1 \cdot \dots \cdot (2n-2) \cdot (2n-1) \cdot (2n)}{1 \cdot \dots \cdot (2n-2)}$$

$$= 2 \cdot (2n-1) \cdot 2n = 4n \cdot (2n-1).$$

Damit erhalten wir die Wahrscheinlichkeit, dass eine der beiden Personen beide unterschriebenen Asse in der Hand hält durch

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{4n \cdot (2n-1)}{4n \cdot (4n-1)} = \frac{2n-1}{4n-1}.$$

Für ein Kartenspiel ergibt sich beispielsweise die Wahrscheinlichkeit $\frac{2 \cdot 1 - 1}{4 \cdot 1 - 1} = \frac{1}{3}$.

- b) Die Anzahl aller möglichen Ereignisse erhalten wir durch eine Variation mit Wiederholung und $n = 6$ sowie $k = 3$, da wir 3 Elemente mit Wiederholung aus 6 Möglichen wählen. Damit folgt mathematisch

$$\Omega = \{(\omega_1, \omega_2, \omega_3) : \omega_i \in \{1, \dots, 6\} \text{ für } i \in \{1, 2, 3\}\}$$

und

$$|\Omega| = n^k = 6^3 = 216.$$

Die mathematisch genaue Beschreibung der Menge A der möglichen Elemente lassen wir an dieser Stelle weg. Für die Bestimmung der Anzahl betrachten wir zunächst nur das erste Element ω_1 und überlegen, wie viele Möglichkeiten es gibt, dass die anderen beiden Ergebnisse ω_2 und ω_3 per Summation den Wert ω_1 ergeben. Anschließend können wir die Anzahl dieser Möglichkeiten mit 3 multiplizieren, da für jeden der drei Einträge ω_1 , ω_2 und ω_3 die gleiche Anzahl existiert, während wir für die Multiplikation alle Fälle zählen müssen, da wir sonst manche Ereignisse mehr als einmal zählen.

Für die Summation erhalten wir die möglichen 15 Elemente

$$(2, 1, 1)$$

$$(3, 1, 2), (3, 2, 1)$$

$$(4, 1, 3), (4, 2, 2), (4, 3, 1)$$

$$(5, 1, 4), (5, 2, 3), (5, 3, 2), (5, 4, 1)$$

$$(6, 1, 5), (6, 2, 4), (6, 3, 3), (6, 4, 2), (6, 5, 1).$$

Für die Multiplikation erhalten wir derweil

$$(1, 1, 1)$$

$$(2, 1, 2), (2, 2, 1), (1, 2, 2)$$

$$(3, 1, 3), (3, 3, 1), (1, 3, 3)$$

$$(4, 1, 4), (1, 4, 4), (4, 4, 1), (4, 2, 2), (2, 4, 2), (2, 2, 4)$$

$$(5, 1, 5), (5, 5, 1), (1, 5, 5)$$

$$(6, 1, 6), (6, 6, 1), (1, 6, 6), (6, 2, 3), (6, 3, 2), (2, 6, 3), (3, 6, 2), (2, 3, 6), (3, 2, 6).$$

Insgesamt erhalten wir also $3 \cdot 15 = 45$ Elemente für die Summation und 25 für das Produkt. Somit erhalten wir die Gesamtanzahl

$$|A| = 3 \cdot 15 + 25 = 70.$$

Die Wahrscheinlichkeit berechnet sich somit durch

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{70}{216} \approx 32,41\%.$$

2.4 zu Wahrscheinlichkeitsmaß und Zufallsvariablen

keine Aufgaben vorhanden

2.5 zu Bedingte Wahrscheinlichkeit

zu A.2.7:

- a) Wir berechnen die Wahrscheinlichkeiten direkt mit der Formel der bedingten Wahrscheinlichkeit. Für eine bessere Verständlichkeit bestimmen wir dazu im ersten Schritt

$$\begin{aligned} \mathbb{P}(A \cap B_1) &= \frac{|A \cap B_1|}{|\Omega|} = \frac{| \{1, 2\} |}{|\Omega|} = \frac{2}{9}, \\ \mathbb{P}(A \cap B_2) &= \frac{|A \cap B_2|}{|\Omega|} = \frac{| \{3, 4, 5\} |}{|\Omega|} = \frac{3}{9} = \frac{1}{3}, \\ \mathbb{P}(B_1) &= \frac{|B_1|}{|\Omega|} = \frac{4}{9} \quad \text{sowie} \\ \mathbb{P}(B_2) &= \frac{|B_2|}{|\Omega|} = \frac{5}{9}. \end{aligned}$$

Abschließend benutzen wir diese Werte für

$$\begin{aligned} \mathbb{P}(A | B_1) &= \frac{\mathbb{P}(A \cap B_1)}{\mathbb{P}(B_1)} = \frac{\frac{2}{9}}{\frac{4}{9}} = \frac{2}{4} = \frac{1}{2} \quad \text{und} \\ \mathbb{P}(A | B_2) &= \frac{\mathbb{P}(A \cap B_2)}{\mathbb{P}(B_2)} = \frac{\frac{1}{3}}{\frac{5}{9}} = \frac{1}{3} \cdot \frac{9}{5} = \frac{3}{5}. \end{aligned}$$

- b) Obwohl wir die Wahrscheinlichkeit von A wie im ersten Teil der Aufgabe direkt mit Hilfe des Quotienten aus der Anzahl der günstigen Fälle, also $|A|$, durch die Anzahl der möglichen Fälle, $|\Omega|$, berechnen können, da wir uns im Laplace'schen Wahrscheinlichkeitsraum befinden, möchten wir die Wahrscheinlichkeit über den Satz der totalen Wahrscheinlichkeit (**S.d.t.W.**) berechnen. Da $B_1 \cup B_2 = \{1, 2, 6, 7\} \cup \{0, 3, 4, 5, 8\} = \{0, 1, \dots, 7, 8\} = \Omega$ gilt, erhalten wir mit den bereits berechneten Zwischenwerten aus obiger Teilaufgabe

$$\mathbb{P}(A) \stackrel{\text{S.d.t.W.}}{=} \mathbb{P}(A | B_1) \cdot \mathbb{P}(B_1) + \mathbb{P}(A | B_2) \cdot \mathbb{P}(B_2) = \frac{1}{2} \cdot \frac{4}{9} + \frac{3}{5} \cdot \frac{5}{9} = \frac{2}{9} + \frac{3}{9} = \frac{5}{9}.$$

Das Ergebnis leuchtet ein, da wir mit $|A| = 5$ und $|\Omega| = 9$ das gleiche Ergebnis bei direkter Berechnung erhalten.

c) Mit Hilfe des Satzes von Bayes berechnen wir direkt

$$\mathbb{P}(B_1 | A) = \frac{\mathbb{P}(A | B_1) \cdot \mathbb{P}(B_1)}{\mathbb{P}(A)} = \frac{\frac{2}{9} \cdot \frac{4}{9}}{\frac{5}{9}} = \frac{2}{9} \cdot \frac{4}{9} \cdot \frac{9}{5} = \frac{8}{45}.$$

zu A.2.8:

In dieser Aufgabe bezeichnen wir mit der Menge A das Ereignis „Kind schaffte einen 1er Schnitt im Abitur“. Mit der Menge B beschreiben wir das Ereignis „das Kind wartete auf einen zweiten Marshmallow“. Insgesamt betrachten wir nur Kinder, die am Test teilgenommen haben.

a) In der ersten Teilaufgabe sind wir an der Wahrscheinlichkeit $\mathbb{P}(A)$ interessiert, dass ein Kind, unabhängig vom Ausgang des Marshmallow-Tests, einen 1er Schnitt im Abitur schafft. Wenn eine Aufgabe so gestellt ist, dass viele bedingte Ereignisse gegeben sind und in der Aufgabe die Wahrscheinlichkeit ohne Bedingung gefragt ist, wird dies in der Regel mit dem Satz der totalen Wahrscheinlichkeit (**S.d.t.W.**) gelöst. Dazu benötigen wir zwei bedingende Ereignisse, die den gesamten Raum erfassen. In diesem Fall ergibt sich:

B = Kind wartete auf zweiten Marshmallow

B^c = Kind wartete nicht auf zweiten Marshmallow

Insgesamt erhalten wir mit $B \cup B^c$ alle teilnehmenden Kinder.

Generell gilt immer: $B \cup B^c = B \cup (\Omega \setminus B) = \Omega$.

Aus dem Text erhalten wir die Wahrscheinlichkeiten

$$\mathbb{P}(A | B) = \mathbb{P}(\text{1er Schnitt} | \text{wartete auf zweiten M.}) = \frac{1}{2},$$

$$\mathbb{P}(A | B^c) = \mathbb{P}(\text{1er Schnitt} | \text{wartete nicht auf zweiten M.}) = \frac{1}{10},$$

$$\mathbb{P}(B^c) = \mathbb{P}(\text{Wartete nicht auf zweiten M.}) = \frac{6}{10} \quad \text{sowie}$$

$$\mathbb{P}(B) = \mathbb{P}(\text{Wartete auf zweiten M.}) = \frac{4}{10}.$$

Dabei erhalten wir die $\frac{4}{10}$ in der letzten Zeile komplementär zu der Aussage, dass 6 von 10 nicht auf einen zweiten Marshmallow gewartet haben (also 4 von 10, die tatsächlich gewartet haben).

Damit erhalten wir

$$\mathbb{P}(A) \stackrel{\text{S.d.t.W.}}{=} \mathbb{P}(A | B) \cdot \mathbb{P}(B) + \mathbb{P}(A | B^c) \cdot \mathbb{P}(B^c) = \frac{1}{2} \cdot \frac{4}{10} + \frac{1}{10} \cdot \frac{6}{10} = \frac{1}{5} + \frac{6}{100} = \frac{26}{100} = 26\%.$$

b) Für die zweite Teilaufgabe ist die Wahrscheinlichkeit $\mathbb{P}(B^c | A)$ gesucht. Wir wissen bereits, dass die Person einen 1er Schnitt im Abitur geschafft hat (bedingen also nach A) und möchten die Wahrscheinlichkeit wissen, dass sie den Marshmallow-Test nicht bestanden hat (daher B^c). Diese lösen wir mit Hilfe der Regel von Bayes durch

$$\mathbb{P}(B^c | A) = \frac{\mathbb{P}(A | B^c) \cdot \mathbb{P}(B^c)}{\mathbb{P}(A)} = \frac{\frac{1}{10} \cdot \frac{6}{10}}{\frac{26}{100}} = \frac{6}{100} \cdot \frac{100}{26} = \frac{3}{13} \approx 23,08\%.$$

zu A.2.9

Diese Aufgabe lösen wir mit Hilfe des allgemeinen Multiplikationssatzes (*a.M.*). Zunächst übersetzen wir obige Aufgabenstellung in mathematische Sprache. Dazu definieren wir die Mengen A_i für $i \in \{1, 2, 3, 4\}$ als jeweils das Ereignis, dass der i -te Karton Äpfel enthält. Für die Aufgabenstellung möchten wir somit den Wert

$$\mathbb{P}(\text{„Die ersten vier Kartons sind mit Äpfeln gefüllt“}) = \mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4)$$

bestimmen, da durch den Schnitt jedes der vier Ereignisse erfüllt ist. Es ergibt sich

$$\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4) \stackrel{\text{a.M.}}{=} \mathbb{P}(A_1) \cdot \mathbb{P}(A_2 | A_1) \cdot \mathbb{P}(A_3 | A_1 \cap A_2) \cdot \mathbb{P}(A_4 | A_1 \cap A_2 \cap A_3).$$

Anschließend berechnen wir die einzelnen Faktoren:

$$\mathbb{P}(A_1) = \frac{95}{100} \quad \text{Es sind noch 95 von 100 möglichen Kartons mit Äpfeln befüllt.}$$

$$\mathbb{P}(A_2 | A_1) = \frac{94}{99} \quad \text{Es sind noch 94 von 99 möglichen Kartons mit Äpfeln befüllt.}$$

$$\mathbb{P}(A_3 | A_1 \cap A_2) = \frac{93}{98} \quad \text{Es sind noch 93 von 98 möglichen Kartons mit Äpfeln befüllt.}$$

$$\mathbb{P}(A_4 | A_1 \cap A_2 \cap A_3) = \frac{92}{97} \quad \text{Es sind noch 92 von 97 möglichen Kartons mit Äpfeln befüllt.}$$

Durch das Bedingen wissen wir, dass in den vorigen Kartons bereits Äpfel enthalten sind, sodass wir sowohl die Anzahl der möglichen Kartons als auch die Anzahl der möglichen Kartons mit Äpfeln verringern. Damit folgt

$$\mathbb{P}(\text{„Die ersten vier Kartons sind mit Äpfeln gefüllt“}) = \frac{95}{100} \cdot \frac{94}{99} \cdot \frac{93}{98} \cdot \frac{92}{97} \approx 0,8119 = 81,19\%.$$

zu A.2.10

Wir übersetzen die Aufgabenstellung zunächst in mathematische Sprache. Zum Einen beschreiben wir mit der Menge A das Ereignis „Person ist Brillenträger“ und zum mit der Menge B das Ereignis „Person schaute früh viel Fernsehen“. Dazu erhalten wir die Mengen A^c , also dass eine Person keine Brille trägt und B^c , dass eine Person früher wenig oder kein Fernsehen geschaut hat. Aus dem Aufgabentext erhalten wir zudem die Wahrscheinlichkeiten

$$\mathbb{P}(A) = \frac{40}{100} = \frac{2}{5}$$

$$\mathbb{P}(B | A) = \frac{20}{40} = \frac{1}{2}$$

$$\mathbb{P}(B | A^c) = \frac{15}{60} = \frac{1}{4}$$

Gesucht ist nun die Wahrscheinlichkeit, dass eine Person, die früher viel Fernsehen geschaut hat (Bedingung), heute eine Brille trägt? Mit Hilfe der Formel von Bayes und des Satzes der totalen Wahrscheinlichkeit (*S.d.t.W.*) erhalten wir

$$\mathbb{P}(A | B) \stackrel{\text{Bayes}}{=} \frac{\mathbb{P}(B | A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)} \stackrel{\text{S.d.t.W.}}{=} \frac{\mathbb{P}(B | A) \cdot \mathbb{P}(A)}{\mathbb{P}(B | A) \cdot \mathbb{P}(A) + \mathbb{P}(B | A^c) \cdot \mathbb{P}(A^c)} = \frac{\frac{1}{2} \cdot \frac{2}{5}}{\frac{1}{2} \cdot \frac{2}{5} + \frac{1}{4} \cdot \frac{3}{5}} = \frac{\frac{1}{5}}{\frac{2}{5} + \frac{3}{20}} = \frac{\frac{1}{5}}{\frac{4}{5}} = \frac{1}{4}$$

zu A.2.11:

Bevor wir in die Beantwortung der einzelnen Fragestellungen einstieg, übersetzen wir den Text zunächst in mathematische Sprache. Wir bezeichnen mit

- A = Person ist unter 25 Jahre alt.
 A^c = Person ist über 24 Jahre alt.
 B = Person nimmt Gesangsunterricht.
 B^c = Person nimmt keinen Gesangsunterricht.
 C = Person kommt eine Runde weiter.

Aus dem Text können wir zudem folgende Wahrscheinlichkeiten ablesen:

$$\begin{array}{ll}
 \mathbb{P}(C | A \cap B^c) = \frac{3}{10} & \mathbb{P}(A \cap B^c) = \frac{80}{200} \\
 \mathbb{P}(C | A \cap B) = \frac{5}{10} & \mathbb{P}(A \cap B) = \frac{40}{200} \\
 \mathbb{P}(C | A^c \cap B^c) = \frac{1}{10} & \mathbb{P}(A^c \cap B^c) = \frac{60}{200} \\
 \mathbb{P}(C | A^c \cap B) = \frac{6}{10} & \mathbb{P}(A^c \cap B) = \frac{20}{200}
 \end{array}$$

Damit beantworten wir nun die beiden Fragestellungen:

- a) Gefragt ist nach der Anzahl der Personen, die eine Runde weiterkommen. Dazu bestimmen wir zunächst die Wahrscheinlichkeit, dass eine Person weiterkommt und wenden diese anschließend auf die Anzahl an. Für diese Wahrscheinlichkeit benutzen wir erneut den Satz der totalen Wahrscheinlichkeit (S.d.t.W.). Im Gegensatz zu den vorigen Aufgaben liegen in diesem Fall zwei Eigenschaften vor, die das Weiterkommen (empirisch) beeinflussen, das Alter und der Gesangsunterricht. Wir erhalten jedoch alle Teilnehmer der Show durch die Vereinigung der möglichen Kombinationen dieser beiden Eigenschaften. Es gilt also

$$\Omega = \text{„Alle Teilnehmer“} = (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B) \cup (A^c \cap B^c).$$

Damit folgt

$$\begin{aligned}
 \mathbb{P}(C) &= \mathbb{P}(C | A \cap B^c) \cdot \mathbb{P}(A \cap B^c) + \mathbb{P}(C | A \cap B) \cdot \mathbb{P}(A \cap B) \\
 &\quad + \mathbb{P}(C | A^c \cap B^c) \cdot \mathbb{P}(A^c \cap B^c) + \mathbb{P}(C | A^c \cap B) \cdot \mathbb{P}(A^c \cap B) \\
 &= \frac{3}{10} \cdot \frac{2}{5} + \frac{1}{2} \cdot \frac{1}{5} + \frac{1}{10} \cdot \frac{3}{10} + \frac{3}{5} \cdot \frac{1}{10} \\
 &= \frac{3}{25} + \frac{1}{10} + \frac{3}{100} + \frac{3}{50} \\
 &= \frac{31}{100} \\
 &= 31\%.
 \end{aligned}$$

Damit kommen im Mittel 62 Personen in die nächste Runde.

- b) Die in dieser Aufgabenstellung gesuchte Wahrscheinlichkeit lässt sich mathematisch durch $\mathbb{P}(B^c | C)$ ausdrücken. Diese möchten wir mit der Definition der bedingten Wahrscheinlichkeit berechnen, benötigen allerdings noch den Wert $\mathbb{P}(C \cap B^c)$, da wir $\mathbb{P}(C)$ bereits durch den ersten Teil kennen.

Die mathematische Herleitung ist an dieser Stelle anspruchsvoller. Wir nutzen zuerst aus, dass $B^c = B^c \cap \Omega$, da jedes Element von B^c gleichzeitig ein Element des gesamten Raums ist. Anschließend stellen wir Ω durch $\Omega = A \cup A^c$ dar, sodass wir die zweite Eigenschaft mit ins Spiel bringen und Wahrscheinlichkeiten erhalten, über die wir laut Aufgabentext etwas aussagen können. Zu guter Letzt benutzen wir zweimal den Multiplikationssatz:

$$\begin{aligned}\mathbb{P}(C \cap B^c \cap A) &= \mathbb{P}(C | B^c \cap A) \cdot \mathbb{P}(B^c \cap A) \\ \mathbb{P}(C \cap B^c \cap A^c) &= \mathbb{P}(C | B^c \cap A^c) \cdot \mathbb{P}(B^c \cap A^c).\end{aligned}$$

Damit erhalten wir nun die Wahrscheinlichkeit für den Schnitt durch

$$\begin{aligned}\mathbb{P}(C \cap B^c) &= \mathbb{P}(C \cap (B^c \cap \Omega)) = \mathbb{P}(C \cap (B^c \cap (A \cup A^c))) \\ &= \mathbb{P}((C \cap B^c \cap A) \cup (C \cap B^c \cap A^c)) \\ &= \mathbb{P}(C \cap B^c \cap A) + \mathbb{P}(C \cap B^c \cap A^c) \\ &= \mathbb{P}(C | B^c \cap A) \cdot \mathbb{P}(B^c \cap A) + \mathbb{P}(C | B^c \cap A^c) \cdot \mathbb{P}(B^c \cap A^c) \\ &= \frac{3}{10} \cdot \frac{2}{5} + \frac{1}{10} \cdot \frac{3}{10} = 15\%.\end{aligned}$$

Insgesamt beantworten wir die Aufgabenstellung mit der Definition der bedingten Wahrscheinlichkeit durch

$$\mathbb{P}(B^c | C) = \frac{\mathbb{P}(C \cap B^c)}{\mathbb{P}(C)} = \frac{\frac{15}{100}}{\frac{31}{100}} = \frac{15}{31} = \frac{15}{100} \cdot \frac{100}{31} = \frac{15}{31} \approx 48,39\%.$$

2.6 zu Stochastische Unabhängigkeit

zu A.2.12:

Bei jeder dieser Aufgaben berechnen wir die Unabhängigkeit per Formel direkt nach beziehungsweise widerlegen sie:

- a) Es gilt

$$\begin{aligned}\mathbb{P}(A \cap B) &= \mathbb{P}(\emptyset) = 0 \\ \mathbb{P}(A) &= \frac{|A|}{|\Omega|} = \frac{4}{8} = \frac{1}{2} \\ \mathbb{P}(B) &= \frac{|B|}{|\Omega|} = \frac{4}{8} = \frac{1}{2}\end{aligned}$$

und somit

$$\mathbb{P}(A \cap B) = 0 \neq \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Die Ereignisse sind demnach nicht stochastisch unabhängig.

b) Es gilt

$$\begin{aligned}\mathbb{P}(A \cap B) &= \mathbb{P}(\{3, 4\}) = \frac{|\{3, 4\}|}{|\Omega|} = \frac{2}{8} = \frac{1}{4} \\ \mathbb{P}(A) &= \frac{|A|}{|\Omega|} = \frac{4}{8} = \frac{1}{2} \\ \mathbb{P}(B) &= \frac{|B|}{|\Omega|} = \frac{4}{8} = \frac{1}{2}\end{aligned}$$

und somit

$$\mathbb{P}(A \cap B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Die Ereignisse sind demnach stochastisch unabhängig.

c) Es gilt

$$\begin{aligned}\mathbb{P}(A \cap B) &= \mathbb{P}(\{3\}) = \frac{|\{3\}|}{|\Omega|} = \frac{1}{8} \\ \mathbb{P}(A) &= \frac{|A|}{|\Omega|} = \frac{3}{8} \\ \mathbb{P}(B) &= \frac{|B|}{|\Omega|} = \frac{4}{8} = \frac{1}{2}\end{aligned}$$

und somit

$$\mathbb{P}(A \cap B) = \frac{1}{8} \neq \frac{3}{16} = \frac{3}{8} \cdot \frac{1}{2} = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Die Ereignisse sind demnach nicht stochastisch unabhängig.

d) Es gilt

$$\begin{aligned}\mathbb{P}(A \cap B) &= \mathbb{P}(\{3\}) = \frac{|\{3\}|}{|\Omega|} = \frac{1}{12} \\ \mathbb{P}(A) &= \frac{|A|}{|\Omega|} = \frac{3}{12} = \frac{1}{4} \\ \mathbb{P}(B) &= \frac{|B|}{|\Omega|} = \frac{4}{12} = \frac{1}{3}\end{aligned}$$

und somit

$$\mathbb{P}(A \cap B) = \frac{1}{12} = \frac{1}{3} \cdot \frac{1}{4} = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Die Ereignisse sind demnach stochastisch unabhängig.

zu A.2.13:

Um diese Frage zu beantworten, berechnen wir zunächst die einzelnen Wahrscheinlichkeiten sowie den Schnitt. Wir bezeichnen mit der Menge A das Ereignis „Kind 1 spielt im Fußballverein“ und mit der Menge B das Ereignis „Kind 2 spielt weder Fuß- noch Handball“. Für das erste Kind erhalten wir direkt

$$\mathbb{P}(A) = \frac{10}{30} = \frac{1}{3}.$$

Da in der Fragestellung explizit nach dem Kind an zweiter Stelle, also unter Beachtung der Reihenfolge, gefragt wurde, gehen wir für das zweite Kind insgesamt drei Fälle durch (das Ereignis A ist losgelöst von B):

1. Das erste Kind spielt Fußball.
2. Das erste Kind spielt Handball.
3. Das erste Kind spielt weder Fußball noch Handball.

Damit erhalten wir

$$\mathbb{P}(B) = \overbrace{\frac{10}{30}}^{\mathbb{P}(1.)} \cdot \frac{15}{29} + \overbrace{\frac{5}{30}}^{\mathbb{P}(2.)} \cdot \frac{15}{29} + \overbrace{\frac{15}{30}}^{\mathbb{P}(3.)} \cdot \frac{14}{29} = \frac{1}{2}.$$

Dabei verringern wir den Nenner bei allen drei Summanden um 1, da eines der 30 Kinder bereits gewählt wurde und noch 29 zur Auswahl stehen. Beim dritten Summanden müssen wir ebenfalls den Zähler um 1 reduzieren, da ein weder Fuß- noch Handball spielendes Kind bereits an erster Stelle ausgewählt wurde. Mathematisch lässt sich der ganze Vorgang mit dem Satz der totalen Wahrscheinlichkeit und dem allgemeinen Multiplikationssatz direkt erklären.

Für die Unabhängigkeit benötigen wir letztlich noch

$$\mathbb{P}(A \cap B) = \frac{10}{30} \cdot \frac{15}{29} = \frac{5}{29}.$$

Insgesamt erhalten wir

$$\mathbb{P}(A \cap B) = \frac{5}{29} \neq \frac{1}{6} = \frac{1}{3} \cdot \frac{1}{2} = \mathbb{P}(A) \cdot \mathbb{P}(B),$$

sodass die Ereignisse nicht stochastisch unabhängig sind.

zu A.2.14:

Da wir bei dieser Aufgabe gezielt die Unabhängigkeit zeigen beziehungsweise widerlegen möchten, definieren wir die Menge A als das Ereignis „Der zweite Wurf ist eine Primzahl“ und die Menge B als das Ereignis „Die Summe ist größer als 9“. Der zugrunde liegende Raum Ω ergibt sich durch den zweifachen Würfelwurf und damit $|\Omega| = 36$.

Für das Ereignis A zählen wir die günstigen Elemente

$$\begin{aligned} &(1, 2), (2, 2), (3, 2), (4, 2), (5, 2), (6, 2) \\ &(1, 3), (2, 3), (3, 3), (4, 3), (5, 3), (6, 3) \\ &(1, 5), (2, 5), (3, 5), (4, 5), (5, 5), (6, 5). \end{aligned}$$

Für das Ereignis B zählen wir die günstigen Elemente

$$\begin{aligned} &(4, 6) \\ &(5, 5), (5, 6) \\ &(6, 4), (6, 5), (6, 6). \end{aligned}$$

Für das Ereignis $B \cap A$, also dass die Summe größer als 9 ist und der zweite Wurf eine Primzahl ist, zählen wir

$$(5, 5) \text{ und } (6, 5).$$

Insgesamt erhalten wir

$$\mathbb{P}(B | A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{\frac{|A \cap B|}{|\Omega|}}{\frac{|A|}{|\Omega|}} = \frac{|A \cap B|}{|A|} = \frac{2}{18} = \frac{1}{9}.$$

Damit die beiden Ereignisse unabhängig sind, müsste die Wahrscheinlichkeit von B ohne Bedingung nach A ebenfalls $\frac{1}{9}$ ergeben, allerdings gilt

$$\mathbb{P}(B) = \frac{|B|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}.$$

Aus der Ungleichheit folgt, dass die beiden Ereignisse nicht stochastisch unabhängig sind.

2.7 zu Diskreten Zufallsvariablen

zu A.2.15:

a) Aufgrund von

$$f(0) + f(1) + f(2) = \frac{1}{4} + \frac{1}{5} + \frac{1}{2} = \frac{19}{20} \neq 1,$$

ist die zweite Bedingung einer Zähldichte nicht erfüllt.

b) Es sind sowohl alle Werte von f echt positiv als auch die Summe aller Werte aus $\Omega = \{0\}$ gleich 1 (wir haben nur einen Summanden). Daher handelt es sich in diesem Fall um eine Zähldichte.

c) Die Summe der einzelnen Funktionswerte ergibt zwar 1, jedoch ist die erste Bedingung der Positivität wegen $f(2) = -\frac{1}{4}$ nicht erfüllt, sodass es sich nicht um eine Zähldichte handelt.

zu A.2.16:

Damit es sich bei f um eine Zähldichte handelt, müssen wir nachweisen, dass jeder Funktionswert nicht negativ ist und die Summe über alle Funktionswerte 1 ergibt.

Da wir mit $\Omega = \{1, \dots, n\}$ und $n > 0$ nur positive Zahlen einsetzen, folgt direkt $f(k) = \frac{2 \cdot k}{n \cdot (n+1)} > 0$.

Für die Summe berechnen wir

$$\sum_{k=0}^n f(k) = \sum_{k=0}^n \frac{2 \cdot k}{n \cdot (n+1)} = \frac{2}{n \cdot (n+1)} \cdot \sum_{k=0}^n k \stackrel{\text{kleiner Gauß}}{=} \frac{2}{n \cdot (n+1)} \cdot \frac{n \cdot (n+1)}{2} = 1.$$

zu A.2.17:

a) Es gilt $f(1) = \frac{1^2}{30} = \frac{1}{30}$.

b) In diesem Fall erhalten wir $\mathbb{P}(X = 4) = f(4) = \frac{4^2}{30} = \frac{16}{30} = \frac{8}{15}$.

c) Bei dieser Aufgabenstellung berechnen wir

$$\mathbb{P}(X \in \{2, 3\}) = \mathbb{P}(X = 2) + \mathbb{P}(X = 3) = f(2) + f(3) = \frac{2^2}{30} + \frac{3^2}{30} = \frac{4}{30} + \frac{9}{30} = \frac{13}{30}.$$

Die 30 im Nenner kann nicht weggelassen werden, denn sonst würde die Summe der Funktionswerte nicht mehr 1 ergeben:

$$f(1) + f(2) + f(3) + f(4) = \frac{1^2 + 2^2 + 3^2 + 4^2}{30} = \frac{1 + 4 + 9 + 16}{30} = \frac{30}{30} = 1$$

zu A.2.18:

- a) Die angenommenen Werte von X können wir überschaubar aufschreiben (eine theoretische Überlegung/Begründung ist ebenfalls in Ordnung):

$$\begin{array}{ll} X((K, 1)) = 2 \cdot 1 = 2 & X((Z, 1)) = 1 + 1 = 2 \\ X((K, 2)) = 2 \cdot 2 = 4 & X((Z, 2)) = 2 + 1 = 3 \\ X((K, 3)) = 2 \cdot 3 = 6 & X((Z, 3)) = 3 + 1 = 4 \\ X((K, 4)) = 2 \cdot 4 = 8 & X((Z, 4)) = 4 + 1 = 5 \\ X((K, 5)) = 2 \cdot 5 = 10 & X((Z, 5)) = 5 + 1 = 6 \\ X((K, 6)) = 2 \cdot 6 = 12 & X((Z, 6)) = 6 + 1 = 7 \end{array}$$

Die Zufallsvariable nimmt also die Werte 2, 3, 4, 5, 6, 7, 8, 10 und 12 an.

- b) Für die Zähldichte geben wir die Wahrscheinlichkeiten der einzelnen Ereignisse an. Durch obige Aufzählung erhalten wir mit $\Omega = \{(\omega_1, \omega_2) : \omega_1 \in \{K, Z\} \text{ und } \omega_2 \in \{1, 2, 3, 4, 5, 6\}\}$

$$\begin{array}{llll} \mathbb{P}(X = 2) & = & \frac{|X^{-1}(2)|}{|\Omega|} & = \frac{|{(K, 1), (Z, 1)}|}{|\Omega|} = \frac{2}{12} = \frac{1}{6} \\ \mathbb{P}(X = 3) & = & \frac{|X^{-1}(3)|}{|\Omega|} & = \frac{|{(Z, 2)}|}{|\Omega|} = \frac{1}{12} \\ \mathbb{P}(X = 4) & = & \frac{|X^{-1}(4)|}{|\Omega|} & = \frac{|{(K, 2), (Z, 3)}|}{|\Omega|} = \frac{2}{12} = \frac{1}{6} \\ \mathbb{P}(X = 5) & = & \frac{|X^{-1}(5)|}{|\Omega|} & = \frac{|{(Z, 4)}|}{|\Omega|} = \frac{1}{12} \\ \mathbb{P}(X = 6) & = & \frac{|X^{-1}(6)|}{|\Omega|} & = \frac{|{(K, 3), (Z, 5)}|}{|\Omega|} = \frac{2}{12} = \frac{1}{6} \\ \mathbb{P}(X = 7) & = & \frac{|X^{-1}(7)|}{|\Omega|} & = \frac{|{(Z, 6)}|}{|\Omega|} = \frac{1}{12} \\ \mathbb{P}(X = 8) & = & \frac{|X^{-1}(8)|}{|\Omega|} & = \frac{|{(K, 4)}|}{|\Omega|} = \frac{1}{12} \\ \mathbb{P}(X = 10) & = & \frac{|X^{-1}(10)|}{|\Omega|} & = \frac{|{(K, 5)}|}{|\Omega|} = \frac{1}{12} \\ \mathbb{P}(X = 12) & = & \frac{|X^{-1}(12)|}{|\Omega|} & = \frac{|{(K, 6)}|}{|\Omega|} = \frac{1}{12}. \end{array}$$

zu A.2.19:

- a) Aufgrund von $\lim_{t \rightarrow -\infty} F(t) = \frac{1}{4} \neq 0$ handelt es sich nicht um eine Verteilungsfunktion.

- b) In diesem Fall ist die Monotonie verletzt, da beispielsweise

$$F(0) = \frac{2}{3} > \frac{1}{2} = F(1).$$

Daher handelt es sich nicht um eine Verteilungsfunktion.

- c) Aufgrund von $\lim_{t \rightarrow \infty} F(t) = 2 \neq 1$ handelt es sich nicht um eine Verteilungsfunktion.

- d) In diesem Fall sind sowohl die Grenzwerte als auch die Monotonie erfüllt, allerdings ist die rechtsseitige Stetigkeit verletzt. Dies erkennen wir an der Platzierung der $<$ und \leq Zeichen in der Definition von F . Es gilt

$$\lim_{t \downarrow 1} F(t) = \lim_{t \downarrow 1} \frac{2}{3} = \frac{2}{3} \neq \frac{3}{5} = F(1),$$

sodass es sich ebenfalls nicht um eine Verteilungsfunktion handelt.

zu A.2.20:

Damit es sich bei der angegebenen Funktion um eine Verteilungsfunktion handelt, müssen wir die vier Bedingungen nachweisen:

1. $\lim_{t \rightarrow \infty} F(t) = 1$,
2. $\lim_{t \rightarrow -\infty} F(t) = 0$,
3. $F(t)$ ist monoton wachsend und
4. $F(t)$ ist rechtsseitig stetig

Die Gültigkeit der Grenzwerte, also 1. und 2., erkennen wir direkt aus der Definition von F .

Da $0 < \frac{1}{4} < \frac{1}{2} < 1$, ist F zudem monoton wachsend.

Für die rechtsseitige Stetigkeit merken wir zunächst an, dass es sich bei den konstanten Abschnitten um stetige Funktionen handelt, die insbesondere rechtsseitig stetig sind. Es bleiben die Übergänge zu überprüfen, dazu berechnen wir

$$\lim_{t \downarrow -2} F(t) = \lim_{t \downarrow -2} \frac{1}{4} = \frac{1}{4} = F(-2)$$

$$\lim_{t \downarrow 2} F(t) = \lim_{t \downarrow 2} \frac{1}{2} = \frac{1}{2} = F(2)$$

$$\lim_{t \downarrow 4} F(t) = \lim_{t \downarrow 4} 1 = 1 = F(4).$$

Damit handelt es sich bei F um eine Verteilungsfunktion.

Die zugehörige Zähldichte können wir an den Sprüngen ablesen:

$$\mathbb{P}(X = -2) = \frac{1}{4}$$

$$\mathbb{P}(X = 2) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

$$\mathbb{P}(X = 4) = 1 - \frac{1}{2} = \frac{1}{2}.$$

zu A.2.21:

- a) Für jeden Sprung müssen wir eine und insgesamt eine zusätzliche Fallunterscheidung machen:

$$\begin{aligned} t < 0 & : F(t) = \sum_{k=0}^{\lfloor t \rfloor} f(k) = 0 \\ 0 \leq t < 1 & : F(t) = \sum_{k=0}^{\lfloor t \rfloor} f(k) = \sum_{k=0}^0 f(k) = f(0) = \frac{1}{3} \\ 1 \leq t < 2 & : F(t) = \sum_{k=0}^{\lfloor t \rfloor} f(k) = \sum_{k=0}^1 f(k) = f(0) + f(1) = \frac{2}{3} \\ 2 \leq t & : F(t) = \sum_{k=0}^{\lfloor t \rfloor} f(k) = \sum_{k=0}^2 f(k) = f(0) + f(1) + f(2) = 1 \end{aligned}$$

Damit erhalten wir die Verteilungsfunktion

$$F(t) = \begin{cases} 0 & \text{für } t < 0 \\ \frac{1}{3} & \text{für } 0 \leq t < 1 \\ \frac{2}{3} & \text{für } 1 \leq t < 2 \\ 1 & \text{für } 2 \leq t \end{cases}$$

b) Für jeden Sprung müssen wir erneut eine und insgesamt eine zusätzliche Fallunterscheidung machen:

$$\begin{aligned} t < 1 & : F(t) = \sum_{k=1}^{\lfloor t \rfloor} f(k) = 0 \\ 1 \leq t < 2 & : F(t) = \sum_{k=1}^{\lfloor t \rfloor} f(k) = \sum_{k=1}^1 f(k) = f(1) = \frac{1}{30} \\ 2 \leq t < 3 & : F(t) = \sum_{k=1}^{\lfloor t \rfloor} f(k) = \sum_{k=1}^2 f(k) = f(1) + f(2) = \frac{1}{6} \\ 3 \leq t < 4 & : F(t) = \sum_{k=1}^{\lfloor t \rfloor} f(k) = \sum_{k=1}^3 f(k) = f(1) + f(2) + f(3) = \frac{7}{15} \\ 4 \leq t & : F(t) = \sum_{k=1}^{\lfloor t \rfloor} f(k) = \sum_{k=1}^4 f(k) = f(1) + f(2) + f(3) + f(4) = 1 \end{aligned}$$

Damit erhalten wir die Verteilungsfunktion

$$F_X(t) = \begin{cases} 0 & \text{für } t < 1 \\ \frac{1}{30} & \text{für } 1 \leq t < 2 \\ \frac{1}{6} & \text{für } 2 \leq t < 3 \\ \frac{7}{15} & \text{für } 3 \leq t < 4 \\ 1 & \text{für } 4 \leq t \end{cases}$$

zu A.2.22:

Wir berechnen direkt

$$F(-5) = 0 \quad \text{aus erstem Bereich entnommen}$$

$$F(0,5) = \frac{1}{4} \quad \text{aus zweitem Bereich entnommen}$$

$$F(3) = \frac{1}{2} \quad \text{aus drittem Bereich entnommen.}$$

zu A.2.23:

Wir berechnen direkt

$$\begin{aligned}\mathbb{P}(X \leq 2,6) & \stackrel{\text{diskret}}{=} F(2) = \frac{1}{3} \\ \mathbb{P}(X > 2) & \stackrel{\text{diskret}}{=} 1 - F(2) = 1 - \frac{1}{3} = \frac{2}{3} \\ \mathbb{P}(-1 < X < 2) & \stackrel{\text{diskret}}{=} F(1) - F(-1) = \frac{1}{4} - 0 = \frac{1}{4}.\end{aligned}$$

zu A.2.24:

Damit es sich um eine Zähldichte handelt, muss zum Einen der Wert $f(3)$ positiv sein (alle anderen Werte sind positiv, sodass die Existenz einer Zähldichte möglich ist) und zum Anderen muss die Summe über alle Funktionswerte 1 ergeben. Es ergibt sich

$$\begin{aligned}\sum_{k=1}^5 f(k) &= f(1) + f(2) + f(3) + f(4) + f(5) \stackrel{!}{=} 1 \\ \Leftrightarrow f(3) &= 1 - f(1) - f(2) - f(4) - f(5) \\ \Leftrightarrow f(3) &= 1 - \frac{1}{6} - \frac{1}{6} - \frac{1}{4} - \frac{1}{4} \\ \Leftrightarrow f(3) &= \frac{1}{6}.\end{aligned}$$

Für $f(3) = \frac{1}{6}$ erhalten wir somit durch f eine Zähldichte.

Den Erwartungswert einer Zufallsvariable X mit Zähldichte f berechnen wir durch

$$\mathbb{E}[X] = \sum_{k=1}^5 k \cdot f(k) = 1 \cdot f(1) + 2 \cdot f(2) + 3 \cdot f(3) + 4 \cdot f(4) + 5 \cdot f(5) = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{4} + \frac{5}{4} = 3,25.$$

Damit erhalten wir die Varianz mit

$$\begin{aligned}\text{Var}(X) &= \sum_{k=1}^5 (k - \mathbb{E}[X])^2 \cdot f(k) = (1 - 3,25)^2 \cdot f(1) + (2 - 3,25)^2 \cdot f(2) + (3 - 3,25)^2 \cdot f(3) + (4 - 3,25)^2 \cdot f(4) \\ &\quad + (5 - 3,25)^2 \cdot f(5) + (6 - 3,25)^2 \cdot f(6) \approx 2,02.\end{aligned}$$

zu A.2.25:

Zunächst starten wir eine Vorüberlegung. Da die Zufallsvariable die Werte $-2, -1, 0, 1$ und 2 annehmen kann, folgt daraus, dass $|X + 1|$ die Werte

$$\begin{aligned}|-2+1| &= |-1| = 1 \\ |-1+1| &= |0| = 0 \\ |0+1| &= |1| = 1 \\ |1+1| &= |2| = 2 \\ |2+1| &= |3| = 3\end{aligned}$$

annehmen kann. Wir berechnen daher

$$\begin{aligned}
 f_Y(0) &= \mathbb{P}(Y=0) = \mathbb{P}(|X+1|=0) = \mathbb{P}(X=-1) &= \frac{1}{5} \\
 f_Y(1) &= \mathbb{P}(Y=1) = \mathbb{P}(|X+1|=1) = \mathbb{P}(X=-2)+\mathbb{P}(X=0) &= \frac{1}{6} + \frac{1}{4} = \frac{5}{12} \\
 f_Y(2) &= \mathbb{P}(Y=2) = \mathbb{P}(|X+1|=2) = \mathbb{P}(X=1) &= \frac{1}{3} \\
 f_Y(3) &= \mathbb{P}(Y=3) = \mathbb{P}(|X+1|=3) = \mathbb{P}(X=2) &= \frac{1}{20}
 \end{aligned}$$

Für den Erwartungswert berechnen wir

$$\begin{aligned}
 \mathbb{E}[Y] &= \sum_{k=0}^3 k \cdot f_Y(k) = 0 \cdot f_Y(0) + 1 \cdot f_Y(1) + 2 \cdot f_Y(2) + 3 \cdot f_Y(3) &\approx 1,23 \\
 \mathbb{E}[|X+1|] &= \sum_{k=-2}^2 |k+1| \cdot f_X(k) = |-2+1| \cdot f_X(-2) + |-1+1| \cdot f_X(-1) + |0+1| \cdot f_X(0) \\
 &\quad + |1+1| \cdot f_X(1) + |2+1| \cdot f_X(2) &\approx 1,23.
 \end{aligned}$$

2.8 zu Spezielle diskrete Verteilungen

zu A.2.26:

Mit $X \sim \mathcal{B}\left(10, \frac{1}{3}\right)$ folgt mit $f(k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$:

$$\mathbb{P}(X=3) = \binom{10}{3} \cdot \left(\frac{1}{3}\right)^3 \cdot \left(1 - \frac{1}{3}\right)^{10-3} = 120 \cdot \left(\frac{1}{3}\right)^3 \cdot \left(\frac{2}{3}\right)^7 \approx 0,2601 = 26,01\%$$

zu A.2.27:

Da X mit Parameter $p = \frac{1}{5}$ geometrisch verteilt ist, ergeben sich je nach Interpretation der geometrischen Verteilung zwei Erwartungswerte:

$$\begin{aligned}
 i) \quad \mathbb{E}[X] &= \frac{1-p}{p} = \frac{1-\frac{1}{5}}{\frac{1}{5}} = 4 \quad \text{sowie} \\
 ii) \quad \mathbb{E}[X] &= \frac{1}{p} = \frac{1}{\frac{1}{5}} = 5.
 \end{aligned}$$

Im ersten Fall handelt es sich um die erwartete Anzahl an Fehlversuchen bis zum ersten Erfolg bei einer Erfolgswahrscheinlichkeit von 20%. Der zweite Fall gibt die erwartete Anzahl an Versuchen (inklusive des Erfolgs) an.

zu A.2.28:

Dieses Experiment spiegelt eine Wiederholung eines gleichbleibenden Vorgangs wider, sodass wir es mit einer binomialverteilten Zufallsvariable X modellieren. Da es sich um eine faire Münze handelt, liegt die Erfolgswahrscheinlichkeit, also die Wahrscheinlichkeit, in einem Durchgang

„Kopf“ zu werfen, bei $\frac{1}{2}$. Insgesamt modellieren wir $X \sim \mathcal{B}\left(50, \frac{1}{2}\right)$. Wir interpretieren „zwischen“ als exklusive ($<$ und $>$) der Grenzen (inklusive, also \leq und \geq , ist bei dieser Aufgabenstellung ebenfalls vertretbar) und berechnen

$$\begin{aligned} \mathbb{P}(10 < X < 15) &= \mathbb{P}(X = 11) + \mathbb{P}(X = 12) + \mathbb{P}(X = 13) + \mathbb{P}(X = 14) \\ &= \binom{50}{11} \cdot \left(\frac{1}{2}\right)^{11} \cdot \left(1 - \frac{1}{2}\right)^{50-11} + \binom{50}{12} \cdot \left(\frac{1}{2}\right)^{12} \cdot \left(1 - \frac{1}{2}\right)^{50-12} \\ &\quad + \binom{50}{13} \cdot \left(\frac{1}{2}\right)^{13} \cdot \left(1 - \frac{1}{2}\right)^{50-13} + \binom{50}{14} \cdot \left(\frac{1}{2}\right)^{14} \cdot \left(1 - \frac{1}{2}\right)^{50-14} \\ &\approx 0,001289. \end{aligned}$$

Die **erwartete** Anzahl bei 500 Wiederholungen erhalten wir durch den Erwartungswert der Zufallsvariable Y mit $Y \sim \mathcal{B}\left(500, \frac{1}{2}\right)$ mit

$$\mathbb{E}[Y] = n \cdot p = 500 \cdot \frac{1}{2} = 250.$$

zu A.2.29:

Diese Aufgabe lässt sich mit zwei Verteilungen lösen. Entweder mit Hilfe der geometrischen Verteilung, da wir der Frage nachgehen, wie wahrscheinlich 6 Tage vergehen, bis Bernd diesen erstmals benutzen muss. Anders ausgedrückt: Wie wahrscheinlich sind 6 Versuche bis zum ersten Erfolg? Die zweite Möglichkeit besteht in der Modellierung mit der negativen Binomialverteilung. Diese entspricht bei $r = 1$, also einem Erfolg, gerade der geometrischen Verteilung.

Wir entscheiden uns daher für eine geometrisch verteilte Zufallsvariable X mit Erfolgswahrscheinlichkeit $p = \frac{1}{5}$, also $X \sim \mathcal{G}\left(\frac{1}{5}\right)$. Nun können wir mit der geometrischen Verteilung entweder die Fehlversuche oder die Gesamtversuche modellieren, auch hier entscheiden wir uns für die erste Variante. Damit erhalten wir $k = 5$, da Bernd erst am sechsten Tag erfolgreich seinen Schirm dabei hat und wir berechnen

$$\mathbb{P}(X = 5) = \frac{1}{5} \cdot \left(1 - \frac{1}{5}\right)^5 \approx 0,0655 = 6,55\%.$$

zu A.2.30:

Die Angabe der im Mittel geschossenen Tore in einem festen Zeitintervall (pro 90 Minuten) führt uns zur Modellierung mit der Poisson-Verteilung. Als Parameter wählen wir $\lambda = \frac{2}{5}$ und betrachten die Zufallsvariable $X \sim \mathcal{P}\left(\frac{2}{5}\right)$. Durch die Angabe **mindestens zwei** Tore betrachten wir den Fall

$X \geq 2$ und berechnen die gesuchte Wahrscheinlichkeit mit Hilfe der Zähldichte $f(k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$ durch

$$\mathbb{P}(X \geq 2) = 1 - \mathbb{P}(X < 2) = 1 - (\mathbb{P}(X = 0) + \mathbb{P}(X = 1)) = 1 - \left(\frac{\left(\frac{2}{5}\right)^0}{0!} \cdot e^{-\frac{2}{5}} + \frac{\left(\frac{2}{5}\right)^1}{1!} \cdot e^{-\frac{2}{5}} \right)$$

$$\approx 1 - 0,9384 \approx 0,0616 = 6,16\%.$$

zu A.2.31:

Wir gehen davon aus, dass ein Hacker immer 5 Versuche durchführt, auch wenn er bereits im ersten einen Glückstreffer generiert hat. Zusätzlich unterstellen wir, dass ein bereits benutzter Code nicht erneut verwendet. Somit modellieren das Ausprobieren der Codes mit der Hypergeometrischen-Verteilung („Ziehen ohne Zurücklegen“). Dazu bestimmen wir zunächst die vier Parameter aus dem Text:

- $M = 3$, da einem Kunden 3 Codes zur Verfügung stehen (Anzahl der günstigen Elemente).
- $n = 5$, da der Hacker 5 Versuche durchführt (Anzahl der Stichprobe).
- $k = 1$, da in der Aufgabenstellung nach einem Erfolg gefragt wird (Anzahl der Erfolge).
- $N = 10^5$, da die möglichen Codes einer Variation mit Wiederholung mit $n = 10$ (0,1,2,3,4,5,6,7,8,9,A) und $k = 5$ (5 Stellen) genügen.

Damit erhalten wir mit $X \sim H(10^5, 3, 5)$

$$\mathbb{P}(X = 1) = \frac{\binom{3}{1} \cdot \binom{10^5 - 3}{5 - 1}}{\binom{10^5}{5}} \approx 0,0015 = 0,15\%.$$

zu A.2.32:

Dieses Experiment modellieren wir mit Hilfe der Multi-Hypergeometrischen Verteilung. Wir finden mehrere Klassen vor (europäisch, asiatisch, afrikanisch und südamerikanisch), aus denen wir Elemente ohne Zurücklegen ziehen. Die Fragestellung bezieht sich dabei wiederum nur auf die Klassen. Nicht in der Aufgabenstellung als Text formuliert, jedoch trotzdem relevant ist, dass **kein südamerikanisches** Team in der Gruppe enthalten sein soll. Die Parameter erhalten wir durch:

- $M_1 = 7$ und $k_1 = 2$
- $M_2 = 4$ und $k_2 = 0$
- $M_3 = 5$ und $k_3 = 1$
- $M_4 = 6$ und $k_4 = 1$
- $n = k_1 + k_2 + k_3 + k_4 = 4$ Teams in der Gruppe
- $N = 7 + 4 + 5 + 6 = 22$ Teams insgesamt

Das Experiment modellieren wir mit dem Zufallsvektor $X = (X_1, X_2, X_3, X_4) \sim Hyp(7, 4, 5, 6, 4)$ und berechnen die gesuchte Wahrscheinlichkeit durch

$$\mathbb{P}(X_1 = 2, X_2 = 0, X_3 = 1, X_4 = 1) = \frac{\binom{7}{2} \cdot \binom{4}{0} \cdot \binom{5}{1} \cdot \binom{6}{1}}{\binom{22}{4}} = \frac{21 \cdot 1 \cdot 5 \cdot 6}{7315} \approx 0,0861 = 8,61\%.$$

zu A.2.33:

Da es sich im Gegensatz zur letzten Aufgabe um Ziehen mit Zurücklegen handelt, modellieren wir das Experiment mit Hilfe der Multinomialverteilung, da es erneut um zusammengefasste Klassen geht. Die Anzahl der Würfel im Becher sind dabei für die Wahrscheinlichkeiten p_i wichtig, während das Ziehen die n_k beschreibt. Da wir von jeder Farbe einen Würfel ziehen wollen, gilt $n_1 = n_2 = n_3 = 1$ und $n = n_1 + n_2 + n_3 = 3$. Für die Wahrscheinlichkeiten gilt

$$\bullet N = 3 + 3 + 4 = 10 \quad \bullet p_1 = \frac{3}{N} = \frac{3}{10} \quad \bullet p_2 = \frac{3}{N} = \frac{3}{10} \quad \bullet p_3 = \frac{4}{N} = \frac{4}{10}$$

Damit erhalten wir mit $X = (X_1, X_2, X_3) \sim M\left(3, \left(\frac{3}{10}, \frac{3}{10}, \frac{4}{10}\right)\right)$

$$\mathbb{P}(X_1 = 1, X_2 = 1, X_3 = 1) = \binom{3}{1, 1, 1} \cdot \frac{3}{10} \cdot \frac{3}{10} \cdot \frac{4}{10} = \frac{3!}{1! \cdot 1! \cdot 1!} \cdot \frac{36}{1000} = 0,216 = 21,6\%.$$

zu A.2.34:

Diese Aufgabe lösen wir mit der Binomialverteilung, da jeder Studierende die gleiche Wahrscheinlichkeit besitzt, die Klausur zu bestehen. Wir erhalten $n = 18$, $p = \frac{9}{20}$ und $k = 10$. Da wir die Wahrscheinlichkeit von mehr als bestimmen möchten, berechnen wir für die Zufallsvariable $X \sim \mathcal{B}\left(18, \frac{9}{20}\right)$ die Wahrscheinlichkeit

$$\begin{aligned} \mathbb{P}(X > 10) &= \mathbb{P}(X = 11) + \dots + \mathbb{P}(X = 18) \\ &= \binom{18}{11} \cdot \left(\frac{9}{20}\right)^{11} \cdot \left(1 - \frac{9}{20}\right)^{18-11} + \dots + \binom{18}{18} \cdot \left(\frac{9}{20}\right)^{18} \cdot \left(1 - \frac{9}{20}\right)^{18-18} \\ &\approx 0,07424 + \dots + 0,0000006 \approx 0,12796 = 12,796\%. \end{aligned}$$

zu A.2.35:

Bei dieser Aufgabe geht es um die Anzahl von Versuchen bis erstmals eine gewisse Anzahl von Erfolgen (hier 6) erreicht wird. Dies modellieren wir mit der Negativen Binomialverteilung. Als Parameter erhalten wir die gewünschten Erfolge $r = 6$ und die Erfolgswahrscheinlichkeit $p = \frac{2}{5}$, sodass wir mit der Zufallsvariable $X \sim NB\left(6, \frac{2}{5}\right)$ und erforderlichen $k = 4$ Misserfolgen (6 Erfolge + 4 Misserfolge = 10 Versuche)

$$\mathbb{P}(X = 4) = \binom{4+6-1}{4} \cdot \left(\frac{2}{5}\right)^6 \cdot \left(1 - \frac{2}{5}\right)^4 \approx 0,0669 = 6,69\%$$

erhalten.

zu A.2.36:

Für die Poisson-Approximation bestimmen wir zunächst den Parameter λ durch

$$\lambda = n \cdot p = 1000 \cdot 0,05 = 50.$$

Damit können wir für eine Zufallsvariable $X \sim \mathcal{P}(50)$ die Wahrscheinlichkeit

$$\mathbb{P}(X = 40) = \frac{50^{40}}{40!} \cdot e^{-50} \approx 0,0215 = 2,15\%$$

bestimmen. Das genaue Ergebnis mittels Binomialverteilung (war nicht in der Aufgabenstellung gefragt) liegt bei ca. 0,0208.

2.9 zu Stetige Zufallsvariablen

zu A.2.37:

Für eine Wahrscheinlichkeitsdichte muss zum Einen die Positivität $f(x) \geq 0$ und die Normierungseigenschaft $\int_{-\infty}^{\infty} f(x) dx = 1$ gelten. Für die Beispiele der Aufgabe gilt:

a) $f(-2) = \frac{-2+1}{4} \cdot \mathbb{1}_{[-2,2]}(-2) = -\frac{1}{4} \cdot 1 = -\frac{1}{4} < 0$ und somit ist f keine Dichte, da die Positivität verletzt ist.

b) Es ist $\mathbb{1}_{[0,1]} \geq 0$, da die Funktion nur die Werte 0 und 1 annehmen kann. Zudem ist

$$\int_{-\infty}^{\infty} \mathbb{1}_{[0,1]}(x) dx = \int_0^1 1 dx = [x]_0^1 = 1 - 0 = 1.$$

Damit handelt es sich bei der angegebenen Funktion um eine Zähldichte.

c) Sowohl die Funktion $\mathbb{1}_{[1,\infty)}$ als auch die Funktion $\frac{1}{x^2}$ nehmen nur positive Werte oder den Wert 0 an, sodass das Produkt stets nicht negativ ist. Für die Normierungseigenschaft berechnen wir das uneigentliche Integral

$$\int_{-\infty}^{\infty} \mathbb{1}_{[1,\infty)}(x) \cdot \frac{1}{x^2} dx = \int_1^{\infty} \frac{1}{x^2} dx = \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x^2} dx = \lim_{b \rightarrow \infty} \left[-\frac{1}{x} \right]_1^b = \lim_{b \rightarrow \infty} -\frac{1}{b} + 1 = 1.$$

Somit handelt es sich bei der angegebenen Funktion um eine Wahrscheinlichkeitsdichte.

d) Die Funktion $\mathbb{1}_{[1,3]}(x)$ nimmt wie bereits beobachtet nur die Werte 0 und 1 an. Die Funktion x nimmt zwar sowohl positive als auch negative Werte an, allerdings sind durch die Multiplikation mit $\mathbb{1}_{[1,3]}$ nur Werte im Bereich $[1, 3]$ interessant, sodass die Funktion stets nicht negativ ist. Für die Normierungseigenschaft berechnen wir

$$\int_{-\infty}^{\infty} x \cdot \mathbb{1}_{[1,3]}(x) dx = \int_1^3 x dx = \left[\frac{1}{2} \cdot x^2 \right]_1^3 = \frac{1}{2} \cdot 9 - \frac{1}{2} \cdot 1 = 4 \neq 1.$$

Die angegebene Funktion ist damit keine Wahrscheinlichkeitsdichte, da die Normierungseigenschaft verletzt ist.

zu A.2.38:

Damit es sich bei einer Funktion um eine Verteilungsfunktion handelt, müssen Grenzwerte, Monotonie und rechtsseitige Stetigkeit überprüft werden. Es gilt:

a) In diesem Fall ist F weder monoton wachsend, da $F(1) > F(2)$, aufgrund von

$$F(1) = \frac{1}{1} \cdot \mathbb{1}_{[1,\infty)}(1) = 1 \cdot 1 = 1 \quad \text{und}$$

$$F(2) = \frac{1}{2} \cdot \mathbb{1}_{[1,\infty)}(2) = \frac{1}{2} \cdot 1 = \frac{1}{2},$$

noch sind die Grenzwerte erfüllt, da

$$\lim_{t \rightarrow \infty} F(t) = \lim_{t \rightarrow \infty} \frac{1}{t} \cdot \mathbb{1}_{[1,\infty)}(t) = 0 \neq 1.$$

Somit handelt es bei $F(t)$ nicht um eine Verteilungsfunktion.

- b) Bei dieser Funktion können wir anbringen, dass es sich nicht um ein zugrunde liegendes Wahrscheinlichkeitsmaß handeln kann, denn

$$F(0) = (1 - 2 \cdot e^{-2 \cdot 0}) \cdot \mathbb{1}_{[0, \infty)}(0) = (-1) \cdot 1 = -1.$$

Dies steht im Widerspruch zu $F(t) = \mathbb{P}(X \leq t)$, da ein Wahrscheinlichkeitsmaß \mathbb{P} nur Werte zwischen 0 und 1 liefern darf, beziehungsweise zur Stetigkeit, da für jeden Wert < 0 aufgrund der charakteristischen Funktion $\mathbb{1}_{[0, \infty)}(t)$ direkt

$$\lim_{t \uparrow 0} F(t) = 0$$

folgt, sodass es sich nicht um eine Verteilungsfunktion einer stetigen Verteilung handeln kann.

zu A.2.39:

Da die Funktion $F_X(t)$ in jedem Bereich differenzierbar ist und nur 2 (also insbesondere abzählbar) viele Stellen existieren, an denen $F_X(t)$ eventuell (das können wir, müssen wir an dieser Stelle jedoch nicht überprüfen) nicht differenzierbar ist, erhalten wir die zugehörige Dichte durch komponentenweises Ableiten:

$$f_X(x) = \begin{cases} \frac{d}{dx} 0 & = 0 & \text{für } x < 0 \\ \frac{d}{dx} \left(x - \frac{x^3 - x^2}{3} \right) & = 1 - \frac{3x^2 - 2x}{3} & \text{für } 0 \leq x < 1 \\ \frac{d}{dx} 1 & = 0 & \text{für } 1 \leq x \end{cases}$$

Den Erwartungswert können wir nun mit Hilfe der Dichte berechnen:

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{\infty} x \cdot f_X(x) \, dx = \int_{-\infty}^0 x \cdot f_X(x) \, dx + \int_0^1 x \cdot f_X(x) \, dx + \int_1^{\infty} x \cdot f_X(x) \, dx \\ &= \int_{-\infty}^0 x \cdot 0 \, dx + \int_0^1 x \cdot \left(1 - \frac{3x^2 - 2x}{3} \right) \, dx + \int_1^{\infty} x \cdot 0 \, dx \\ &= \int_{-\infty}^0 0 \, dx + \int_0^1 x - \frac{3x^3 - 2x^2}{3} \, dx + \int_1^{\infty} 0 \, dx \\ &= 0 + \left[\frac{1}{2} \cdot x^2 - \frac{1}{3} \cdot \left(\frac{3}{4} \cdot x^4 - \frac{2}{3} \cdot x^3 \right) \right]_0^1 + 0 \\ &= \frac{1}{2} - \frac{1}{3} \cdot \left(\frac{3}{4} - \frac{2}{3} \right) \\ &= \frac{17}{36} \end{aligned}$$

zu A.2.40:

Wir berechnen direkt

- a) $F(1) = \frac{1}{4}$ ($t = 1$ liegt im zweiten Bereich).
- b) $\mathbb{P}(X > 3) = 1 - F(3) = 1 - \left(\frac{1}{2} + \frac{3^3}{128} \right) \approx 1 - 0,7109 = 0,2891$ ($t = 3$ liegt im dritten Bereich).
- c) $\mathbb{P}(X < 4) \stackrel{\text{stetig}}{=} F(4) = 1$ (Bestimmen wir $\lim_{t \uparrow 4} F(t)$ im dritten Bereich, erhalten wir ebenfalls 1).

zu A.2.41:

Damit es sich um eine Dichte handelt, muss f nicht negativ sein. Da sowohl x als auch $\frac{1}{x}$ auf den positiven Zahlen positiv sind, ist diese Bedingung für $c > 0$ erfüllt. Für die Normierungseigenschaft teilen wir das Integral an Übergängen der Funktionsvorschriften auf und setzen diese im entsprechenden Bereich ein:

$$\begin{aligned}
 \int_{-\infty}^{\infty} f(x) dx &= \int_{-\infty}^0 f(x) dx + \int_0^1 f(x) dx + \int_1^c f(x) dx + \int_c^{\infty} f(x) dx \\
 &= \int_{-\infty}^0 0 dx + \int_0^1 x dx + \int_1^c \frac{1}{x} dx + \int_c^{\infty} 0 dx \\
 &= 0 + \left[\frac{1}{2} \cdot x^2 \right]_0^1 + [\ln(x)]_1^c + 0 \\
 &= \left(\frac{1}{2} \cdot 1^2 - \frac{1}{2} \cdot 0^2 \right) + \left(\ln(c) - \underbrace{\ln(1)}_{=0} \right) \\
 &= \frac{1}{2} + \ln(c).
 \end{aligned}$$

Diesen Ausdruck setzen wir nun gleich 1 und berechnen den Parameter c :

$$\begin{aligned}
 \frac{1}{2} + \ln(c) &= 1 \quad | -\frac{1}{2} \\
 \Leftrightarrow \ln(c) &= \frac{1}{2} \quad | e^{\dots} \\
 \Leftrightarrow c &= e^{\frac{1}{2}}
 \end{aligned}$$

Für $c = e^{\frac{1}{2}}$ ist f damit eine Dichte. Die Verteilungsfunktion erhalten wir, indem wir eine Fallunterscheidung anhand der einzelnen Bereiche machen und das Integral entsprechend der Funktionsvorschriften aufteilen. Wir erhalten

für $t < 0$:

$$F(t) = \int_{-\infty}^t f(x) dx = \int_{-\infty}^t 0 dx = 0$$

für $0 \leq t < 1$:

$$\begin{aligned}
 F(t) = \int_{-\infty}^t f(x) dx &= \int_{-\infty}^0 f(x) dx + \int_0^t f(x) dx \\
 &= \int_{-\infty}^0 0 dx + \int_0^t x dx \\
 &= 0 + \left[\frac{1}{2} \cdot x^2 \right]_0^t \\
 &= \frac{1}{2} \cdot t^2
 \end{aligned}$$

für $1 \leq t < e^{\frac{1}{2}}$:

$$\begin{aligned}
 F(t) &= \int_{-\infty}^t f(x) dx = \int_{-\infty}^0 f(x) dx + \int_0^1 f(x) dx + \int_1^t f(x) dx \\
 &= \int_{-\infty}^0 0 dx + \int_0^1 1 dx + \int_1^t \frac{1}{x} dx \\
 &= 0 + \left[\frac{1}{2} \cdot x^2 \right]_0^1 + [\ln(x)]_1^t \\
 &= 0 + \frac{1}{2} + \ln(t) - \underbrace{\ln(1)}_{=0} \\
 &= \frac{1}{2} + \ln(t)
 \end{aligned}$$

für $e^{\frac{1}{2}} \leq t$:

$$\begin{aligned}
 F(t) &= \int_{-\infty}^t f(x) dx = \int_{-\infty}^0 f(x) dx + \int_0^1 f(x) dx + \int_1^{e^{\frac{1}{2}}} f(x) dx + \int_{e^{\frac{1}{2}}}^t f(x) dx \\
 &= \int_{-\infty}^0 0 dx + \int_0^1 1 dx + \int_1^{e^{\frac{1}{2}}} \frac{1}{x} dx + \int_{e^{\frac{1}{2}}}^t 0 dx \\
 &= 0 + \left[\frac{1}{2} \cdot x^2 \right]_0^1 + [\ln(x)]_1^{e^{\frac{1}{2}}} \\
 &= 0 + \frac{1}{2} + \ln(e^{\frac{1}{2}}) - \underbrace{\ln(1)}_{=0} + 0 \\
 &= \frac{1}{2} + \ln(e^{\frac{1}{2}}) \\
 &= \frac{1}{2} + \frac{1}{2} = 1
 \end{aligned}$$

Insgesamt erhalten wir somit die zugehörige Verteilungsfunktion

$$F(t) = \begin{cases} 0 & \text{für } t < 0 \\ \frac{1}{2} \cdot t^2 & \text{für } 0 \leq t < 1 \\ \frac{1}{2} + \ln(t) & \text{für } 1 \leq t < e^{\frac{1}{2}} \\ 1 & \text{für } e^{\frac{1}{2}} \leq t \end{cases} .$$

zu A.2.42:

Unabhängig von c erkennen wir, dass die Grenzwerte $\lim_{t \rightarrow -\infty} F(t) = 0$ und $\lim_{t \rightarrow \infty} F(t) = 1$ erfüllt sind. Wählen wir c nun so, dass die angegebene Funktion stetig und gleichzeitig monoton wachsend ist, so erhalten wir eine Verteilungsfunktion. Für die Stetigkeit reicht es aus,

$$\begin{aligned}
 0 &= \lim_{t \downarrow 0} F(t) = \lim_{t \downarrow 0} c \cdot t^2 = c \cdot 0 = 0 \quad \text{und} \\
 1 &= \lim_{t \uparrow 2} F(t) = \lim_{t \uparrow 2} c \cdot t^2 = c \cdot 2^2 = c \cdot 4
 \end{aligned}$$

zu berechnen, da die Funktionen der einzelnen Bereiche an sich stetig sind. Wir erhalten mit dieser Betrachtung $c = \frac{1}{4}$ und da $\frac{1}{4} \cdot t^2$ auf dem Bereich $0 \leq t < 2$ monoton wachsend ist, ist $F(t)$ mit dieser Wahl von c eine Verteilungsfunktion.

zu A.2.43:

- a) Die Wahrscheinlichkeit ermitteln wir per Integration, dabei spielt eine Unterscheidung zwischen $<$ und \leq beziehungsweise $>$ und \geq im stetigen Fall keine Rolle. Es gilt

$$\begin{aligned}\mathbb{P}(1 < X < 3) &= \int_1^3 f(x) \, dx = \int_1^3 \frac{1}{24} \cdot x^2 \cdot \mathbb{1}_{[-2,4]}(x) \, dx = \frac{1}{24} \cdot \int_1^3 x^2 \, dx \\ &= \frac{1}{24} \left[\frac{1}{3} \cdot x^3 \right]_1^3 = \frac{1}{72} \cdot (3^3 - 1^3) = \frac{1}{72} \cdot (27 - 1) = \frac{26}{72} = \frac{13}{36} \\ &\approx 0,3611 = 36,11\%.\end{aligned}$$

Dabei haben wir zunächst die Funktion eingesetzt und konnten anschließend die charakteristische Funktion weglassen, da alle Werte von x innerhalb der Grenzen 1 und 3 in der Menge $[-2, 4]$ enthalten sind.

- b) Für den Erwartungswert berechnen wir mit dem Transformationssatz

$$\begin{aligned}\mathbb{E}[X^2] &= \int_{-\infty}^{\infty} x^2 \cdot f(x) \, dx = \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{24} \cdot x^2 \cdot \mathbb{1}_{[-2,4]}(x) \, dx = \frac{1}{24} \cdot \int_{-2}^4 x^4 \, dx \\ &= \frac{1}{24} \cdot \left[\frac{1}{5} \cdot x^5 \right]_{-2}^4 = \frac{1}{120} \cdot (4^5 - (-2)^5) = \frac{1056}{120} = 8,8\end{aligned}$$

sowie wegen $\mathbb{E}[2 \cdot X + 1] = 2 \cdot \mathbb{E}[X] + 1$

$$\begin{aligned}\mathbb{E}[X] &= \int_{-\infty}^{\infty} x \cdot f(x) \, dx = \int_{-\infty}^{\infty} x \cdot \frac{1}{24} \cdot x^2 \cdot \mathbb{1}_{[-2,4]}(x) \, dx = \frac{1}{24} \cdot \int_{-2}^4 x^3 \, dx \\ &= \frac{1}{24} \cdot \left[\frac{1}{4} \cdot x^4 \right]_{-2}^4 = \frac{1}{96} \cdot (4^4 - (-2)^4) = \frac{240}{96} = 2,5.\end{aligned}$$

Insgesamt folgt somit $\mathbb{E}[2 \cdot X + 1] = 2 \cdot 2,5 + 1 = 6$.

- c) Die Varianz erhalten wir mit Hilfe des Verschiebungssatzes und der Rechnungen der vorigen Aufgabenteile durch

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 8,8 - 2,5^2 = 2,55.$$

2.10 zu Spezielle stetige Verteilungen

zu A.2.44:

- a) Die Dichte einer exponentialverteilten Zufallsvariable mit $\lambda = \frac{1}{2}$ ist von der Form $f(x) = \frac{1}{2} \cdot e^{-\frac{1}{2} \cdot x} \cdot \mathbb{1}_{[0, \infty)}(x)$ und die Verteilungsfunktion lautet in diesem Fall

$$F_X(t) = \begin{cases} 0 & \text{für } t < 0 \\ 1 - e^{-\frac{1}{2} \cdot t} & \text{für } 0 \leq t \end{cases}$$

- b) Die Wahrscheinlichkeit berechnen wir durch

$$\mathbb{P}(X > 3) = 1 - F_X(3) = 1 - (1 - e^{-\frac{1}{2} \cdot 3}) = e^{-\frac{3}{2}} \approx 0,2231 = 22,31\%.$$

c) Den Erwartungswert einer exponentialverteilten Zufallsvariable erhalten wir durch

$$\mathbb{E}[X] = \frac{1}{\lambda} = \frac{1}{\frac{1}{2}} = 2.$$

zu A.2.45:

Die Wahrscheinlichkeiten der Standardnormalverteilung erhalten wir mit der Tabelle durch

$$F(1) = \Phi(1) \stackrel{\text{aus Tab.}}{\approx} 0,84135$$

sowie mit der Formel $\Phi(-x) = 1 - \Phi(x)$

$$\mathbb{P}(-1 < X < 1) = \Phi(1) - \Phi(-1) = \Phi(1) - (1 - \Phi(1)) = 2 \cdot \Phi(1) - 1 \stackrel{\text{aus Tabelle}}{\approx} 2 \cdot 0,84135 - 1 = 0,6827.$$

zu A.2.46:

Sei X eine exponentialverteilte Zufallsvariable mit Parameter λ . Durch die Angabe der mittleren Dauer von 3 Jahren, erhalten wir

$$\mathbb{E}[X] = \frac{1}{\lambda} \stackrel{!}{=} 3 \Leftrightarrow \lambda = \frac{1}{3}.$$

Damit berechnen wir:

a) Die Wahrscheinlichkeit, dass das Smartphone innerhalb der nächsten beiden Jahre kaputt geht durch

$$\mathbb{P}(X < 2) = F_X(2) = 1 - e^{-\frac{1}{3} \cdot 2} = 1 - e^{-\frac{2}{3}} \approx 1 - 0,5134 = 0,4866 = 48,66\%.$$

b) Die Wahrscheinlichkeit, dass das Smartphone länger als 4 Jahre benutzt werden kann durch

$$\mathbb{P}(X > 4) = 1 - F_X(4) = 1 - (1 - e^{-\frac{1}{3} \cdot 4}) = e^{-\frac{4}{3}} \approx 0,2636 = 26,36\%.$$

c) Die Wahrscheinlichkeit, dass das Smartphone 6 Jahre lang benutzt werden kann, wenn es in drei Jahren noch funktioniert mit Hilfe der Gedächtnislosigkeit der Exponentialverteilung mit $6 = 3 + 3$ durch

$$\mathbb{P}(X < 6 \mid X > 3) = \mathbb{P}(X < 3) = F_X(3) = 1 - e^{-\frac{1}{3} \cdot 3} = 1 - \frac{1}{e} \approx 0,6321 = 63,21\%.$$

zu A.2.47:

Da X lognormalverteilt ist, berechnen wir die Wahrscheinlichkeiten durch

$$\mathbb{P}(X > 2) = 1 - F_X(2) = 1 - \Phi\left(\frac{\ln(2) - 1}{\sqrt{4}}\right) \approx 1 - \Phi(-0,15) = 1 - (1 - \Phi(0,15))$$

$$\stackrel{\text{aus Tab.}}{\approx} 1 - (1 - 0,55962) = 0,55962 = 55,962\%.$$

und

$$\mathbb{P}(1 < X < 3) = F_X(3) - F_X(1) = \Phi\left(\frac{\ln(3) - 1}{\sqrt{4}}\right) - \Phi\left(\frac{\ln(1) - 1}{\sqrt{4}}\right) \approx \Phi(0,05) - \Phi(-0,5)$$

$$= \Phi(0,05) - (1 - \Phi(0,5)) \stackrel{\text{aus Tab.}}{\approx} 0,51994 - (1 - 0,69146) = 0,2114 = 21,14\%.$$

Für die Verteilung von $Y = X^2$ erkennen wir zunächst, dass X^2 nur positive Werte erzeugen kann. Daher betrachten wir im Folgenden für $t \geq 0$ die Rechnung

$$\begin{aligned} F_Y(t) &= \mathbb{P}(Y \leq t) = \mathbb{P}(X^2 \leq t) = \mathbb{P}(|X| \leq \sqrt{t}) = \mathbb{P}(-\sqrt{t} \leq X \leq \sqrt{t}) = F_X(\sqrt{t}) - F_X(-\sqrt{t}) \\ &= \Phi\left(\frac{\ln(\sqrt{t}) - 1}{\sqrt{4}}\right) - 0 = \Phi\left(\frac{\ln(\sqrt{t}) - 1}{2}\right). \end{aligned}$$

Die Log-Normalverteilung nimmt nur positive Werte an, sodass **der Subtrahend gleich 0** ist. Mit $\sqrt{t} = t^{\frac{1}{2}}$ und $\ln(a^b) = b \cdot \ln(a)$ folgt nun

$$F_Y(t) = \Phi\left(\frac{\frac{1}{2} \cdot \ln(t) - 1}{2}\right) = \Phi\left(\frac{\ln(t) - 2}{4}\right).$$

Dies bedeutet also $Y \sim \mathcal{LN}(2, 16)$.

zu A.2.48:

Für die Zufallsvariable $X \sim \mathcal{N}(1,5, 0,16)$ berechnen wir:

- a) Die Wahrscheinlichkeit, dass der Preis pro Liter zwischen 1,40 € und 1,60 € liegt, ergibt sich mit Standardisieren durch

$$\begin{aligned} \mathbb{P}(1,4 < X < 1,6) &= \mathbb{P}\left(\frac{1,4 - 1,5}{\sqrt{0,16}} < \frac{X - 1,5}{\sqrt{0,16}} < \frac{1,6 - 1,5}{\sqrt{0,16}}\right) = \Phi\left(\frac{1,6 - 1,5}{0,4}\right) - \Phi\left(\frac{1,4 - 1,5}{0,4}\right) \\ &= \Phi\left(\frac{1}{4}\right) - \Phi\left(-\frac{1}{4}\right) = \Phi(0,25) - (1 - \Phi(0,25)) = 2 \cdot \Phi(0,25) - 1 \\ &\stackrel{\text{aus Tab.}}{\approx} 2 \cdot 0,59871 - 1 = 0,19742 = 19,742\%. \end{aligned}$$

- b) Die Wahrscheinlichkeit, dass der Preis pro Liter teurer als 2 € ist, ergibt sich durch

$$\begin{aligned} \mathbb{P}(X > 2) &= 1 - \mathbb{P}(X \leq 2) = 1 - \mathbb{P}\left(\frac{X - 1,5}{\sqrt{0,16}} \leq \frac{2 - 1,5}{\sqrt{0,16}}\right) = 1 - \Phi\left(\frac{2 - 1,5}{\sqrt{0,16}}\right) \\ &= 1 - \Phi(1,25) \stackrel{\text{aus Tab.}}{\approx} 1 - 0,89435 = 0,10565 = 10,565\%. \end{aligned}$$

- c) Der höchste Preis P , der mit einer Wahrscheinlichkeit von 90% unterschritten wird, ist das 90%-Quantil. Mit Standardisieren erhalten wir

$$\begin{aligned} 0,9 &= \mathbb{P}(X \leq P) = \mathbb{P}\left(\frac{X - 1,5}{\sqrt{0,16}} \leq \frac{P - 1,5}{\sqrt{0,16}}\right) = \Phi\left(\frac{P - 1,5}{\sqrt{0,16}}\right) \\ \Leftrightarrow \Phi^{-1}(0,9) &= \frac{P - 1,5}{0,4} \end{aligned}$$

Den Wert $\Phi^{-1}(0,9)$ lesen wir so ab, dass wir 0,9 innerhalb und nicht am Rand der Tabelle suchen. Dies ergibt $\Phi^{-1}(0,9) \stackrel{\text{aus Tab.}}{\approx} 1,28$. Nun formen wir nach P um und erhalten den maximalen Preis:

$$1,28 = \frac{P - 1,5}{0,4} \Leftrightarrow 0,512 = P - 1,5 \Leftrightarrow P = 2,012$$

zu A.2.49:

Damit im Mittel weniger in der Modellierung mit Y statt mit X produziert wird, vergleichen wir die Erwartungswerte:

$$\begin{aligned} \mathbb{E}[X] &\stackrel{!}{\geq} \mathbb{E}[Y] \\ \Leftrightarrow \frac{1}{\lambda} &\geq e^{\mu + \frac{1}{2} \cdot \sigma^2} \\ \Leftrightarrow \frac{1}{\frac{1}{4}} &\geq e^{1 + \frac{1}{2} \cdot \sigma^2} \\ \Leftrightarrow 4 &\geq e^1 \cdot e^{\frac{1}{2} \cdot \sigma^2} \\ \Leftrightarrow \frac{4}{e} &\geq e^{\frac{1}{2} \cdot \sigma^2} \\ \Leftrightarrow \ln\left(\frac{4}{e}\right) &\geq \frac{1}{2} \cdot \sigma^2 \\ \Leftrightarrow 2 \cdot \ln\left(\frac{4}{e}\right) &\geq \sigma^2 \end{aligned}$$

Formt man die linke Seite noch um (ist nicht notwendig), erhalten wir als obere Grenze

$$\sigma^2 \leq 2 \cdot (\ln(4) - \ln(e)) = 2 \cdot \ln(4) - 2 = \ln(16) - 2.$$

2.11 zu Mehrdimensionale Verteilungen

zu A.2.50:

Die fehlenden Werte der Tafel erhalten wir durch Addition der Wahrscheinlichkeiten einer Reihe oder einer Spalte und es ergibt sich

Kontingenztafel			
$X \setminus Y$	y_1	y_2	Summe
x_1	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$
x_2	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{5}{12}$
x_3	$\frac{1}{12}$	0	$\frac{1}{12}$
x_4	0	$\frac{1}{6}$	$\frac{1}{6}$
Summe	$\frac{1}{2}$	$\frac{1}{2}$	1

Anschließend berechnen wir die Wahrscheinlichkeiten:

a) $\mathbb{P}(X = 1, Y = 2) = \frac{1}{6}$, denn wir schauen in der **ersten** Zeile und **zweiten** Spalte.

b) Wir addieren die über Y summierten Werte in der dritten Spalte, die kleiner sind als $X = 4$:

$$\mathbb{P}(X < 4) = \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) = \frac{1}{3} + \frac{5}{12} + \frac{1}{12} = \frac{5}{6}$$

c) Den Wert $\mathbb{P}(Y > 1)$ können wir direkt ablesen:

$$\mathbb{P}(Y > 1) = \mathbb{P}(Y = 2) = \frac{1}{2}$$

d) Die Wahrscheinlichkeit berechnen wir durch

$$\mathbb{P}(X = 1 \mid Y = 2) = \frac{p_{12}}{p_{\bullet 2}} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

e) Die Wahrscheinlichkeit berechnen wir durch

$$\mathbb{P}(Y = 1 \mid X = 1) = \frac{p_{11}}{p_{1\bullet}} = \frac{\frac{1}{6}}{\frac{1}{3}} = \frac{1}{2}$$

Die Erwartungswerte erhalten wir durch

$$\mathbb{E}[X] = \sum_{i=1}^4 i \cdot p_{i\bullet} = 1 \cdot p_{1\bullet} + 2 \cdot p_{2\bullet} + 3 \cdot p_{3\bullet} + 4 \cdot p_{4\bullet} = 1 \cdot \frac{1}{3} + 2 \cdot \frac{5}{12} + 3 \cdot \frac{1}{12} + 4 \cdot \frac{1}{6} = \frac{25}{12}$$

und

$$\begin{aligned} \mathbb{E}[X \cdot Y] &= \sum_{i=1}^4 \sum_{j=1}^2 i \cdot j \cdot p_{ij} = 1 \cdot (1 \cdot p_{11} + 2 \cdot p_{12}) + 2 \cdot (1 \cdot p_{21} + 2 \cdot p_{22}) + 3 \cdot (1 \cdot p_{31} + 2 \cdot p_{32}) + 4 \cdot (1 \cdot p_{41} + 2 \cdot p_{42}) \\ &= 1 \cdot \left(1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6}\right) + 2 \cdot \left(1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{6}\right) + 3 \cdot \left(1 \cdot \frac{1}{12} + 2 \cdot 0\right) + 4 \cdot \left(1 \cdot 0 + 2 \cdot \frac{1}{6}\right) \\ &= 1 \cdot \frac{1}{2} + 2 \cdot \frac{7}{12} + 3 \cdot \frac{1}{12} + 4 \cdot \frac{1}{3} \\ &= \frac{39}{12} \end{aligned}$$

Für die Unabhängigkeit ist es oft interessant, eine Zeile mit Einträgen 0 zu suchen. Wir betrachten an dieser Stelle die dritte Zeile. Wären die Zufallsvariablen unabhängig, so müsste

$$p_{31} = p_{3\bullet} \cdot p_{\bullet 1} = \frac{1}{12} \cdot \frac{1}{2} = \frac{1}{24}$$

gelten. Da jedoch $p_{31} = \frac{1}{12}$, sind X und Y nicht unabhängig.

zu A.2.51:

Die zugehörige Kontingenztafel erhalten wir, indem wir die Wahrscheinlichkeiten $p_{ij} = \mathbb{P}(X = i, Y = j)$ bestimmen. Dabei können i und j jeweils 0, 1 oder 2 sein, da es sowohl von Kuschartieren als auch von kleinen Spielzeugen genug Objekte in der Tombola gibt. Das Experiment beinhaltet ein „Ziehen ohne Zurücklegen“ mit mehreren Klassen, sodass wir für die Wahrscheinlichkeit die Multi-Hypergeometrische Verteilung benutzen. Als Parameter erhalten wir:

- $M_1 = 3$ (die Anzahl der Kuschartiere)
- $M_2 = 5$ (die Anzahl der Spielzeuge)
- $M_3 = 7$ (die Anzahl der Seifenblasen)

und für die Wahrscheinlichkeiten somit

$$p_{ij} = \mathbb{P}(X = i, Y = j) = \begin{cases} \frac{\binom{3}{i} \cdot \binom{5}{j} \cdot \binom{7}{2-i-j}}{\binom{15}{2}} & \text{für } 0 \leq i+j \leq 2 \\ 0 & i+j > 2 \end{cases}.$$

Dabei entspricht $\binom{3}{i}$ der Anzahl, aus 3 Kuscheltieren i zu wählen, $\binom{5}{j}$ der Anzahl, aus 5 Spielzeugen j zu wählen und $\binom{7}{2-i-j}$ der Anzahl, aus 7 Seifenblasen die verbleibenden $2-i-j$ zu wählen. Die Wahrscheinlichkeit, mehr als zwei Gegenstände zu erhalten, ist gleich 0, da die Mutter nur zwei kauft.

Mit Hilfe der gemeinsamen Wahrscheinlichkeit können wir nun die Kontingenztafel erstellen (die Farben belegen wir für die folgenden Rechnungen ab hier neu):

Kontingenztafel				
$X \setminus Y$	0	1	2	Summe
0	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{2}{21}$	$\frac{22}{35}$
1	$\frac{1}{5}$	$\frac{1}{7}$	0	$\frac{12}{35}$
2	$\frac{1}{35}$	0	0	$\frac{1}{35}$
Summe	$\frac{3}{7}$	$\frac{10}{21}$	$\frac{2}{21}$	1

Im Folgenden berechnen wir mit der Tafel die Wahrscheinlichkeiten:

- a) Die gemeinsame Wahrscheinlichkeit können wir entweder erneut in die Funktion einsetzen oder direkt an der Tafel ablesen:

$$\mathbb{P}(X = 1, Y = 1) = \frac{1}{7}$$

- b) Die Wahrscheinlichkeit, die nur von der Zufallsvariable X abhängt, lesen wir am Rand in der entsprechenden Zeile (hier für $X = 0$) ab:

$$\mathbb{P}(X = 0) = \frac{22}{35}$$

- c) Die bedingte Wahrscheinlichkeit erhalten wir mit der Formel

$$\mathbb{P}(Y = 0 \mid X = 1) = \frac{p_{10}}{p_{1\bullet}} = \frac{\frac{1}{5}}{\frac{12}{35}} = \frac{7}{12}.$$

zu A.2.52:

a) Wir berechnen direkt

$$f(2,2) = \frac{12 \cdot e^{-3 \cdot 2}}{e^{4 \cdot 2}} = \frac{12}{e^{14}} \approx 0,000009.$$

b) Für die gemeinsame Verteilungsfunktion berechnen wir das Integral über die gemeinsame Dichte:

$$\begin{aligned} F_{XY}(t_x, t_y) &= \mathbb{P}(X \leq t_x, Y \leq t_y) = \int_{-\infty}^{t_x} \int_{-\infty}^{t_y} f(x,y) dy dx \\ &= \int_0^{t_x} \int_0^{t_y} \frac{12 \cdot e^{-3x}}{e^{4y}} dy dx = \int_0^{t_x} 12 \cdot e^{-3x} \int_0^{t_y} e^{-4y} dy dx \\ &= \int_0^{t_x} 12 \cdot e^{-3x} \left[-\frac{1}{4} \cdot e^{-4y} \right]_0^{t_y} dx = \int_0^{t_x} 12 \cdot e^{-3x} \left(-\frac{1}{4} \cdot e^{-4 \cdot t_y} + \frac{1}{4} \right) dx \\ &= 12 \cdot \left(\frac{1}{4} - \frac{1}{4} \cdot e^{-4 \cdot t_y} \right) \cdot \int_0^{t_x} e^{-3x} dx = \left(3 - 3 \cdot e^{-4 \cdot t_y} \right) \cdot \left[-\frac{1}{3} \cdot e^{-3x} \right]_0^{t_x} \\ &= 3 \cdot \left(1 - 1 \cdot e^{-4 \cdot t_y} \right) \cdot \left[-\frac{1}{3} \cdot e^{-3x} \right]_0^{t_x} = \left(1 - e^{-4 \cdot t_y} \right) \cdot \left(1 - e^{-3 \cdot t_x} \right) \end{aligned}$$

c) Die erste Wahrscheinlichkeit können wir direkt über die gemeinsame Verteilungsfunktion bestimmen

$$\mathbb{P}(X < 1, Y < 1) = F(1,1) = \left(1 - e^{-3} \right) \cdot \left(1 - e^{-4} \right) \approx 0,9328 = 93,28\%.$$

Die zweite Wahrscheinlichkeit erhalten wir mittels Integration durch

$$\begin{aligned} \mathbb{P}(X \leq 2, Y > 2) &= \int_{-\infty}^2 \int_2^{\infty} f(x,y) dy dx = \int_0^2 \int_2^{\infty} 12 \cdot e^{-3x} \cdot e^{-4y} dy dx \\ &= \int_0^2 12 \cdot e^{-3x} \cdot \left[-\frac{1}{4} \cdot e^{-4y} \right]_2^{\infty} dx = \int_0^2 3 \cdot e^{-3x} \cdot e^{-8} dx \\ &= e^{-8} \cdot 3 \cdot \left[-\frac{1}{3} \cdot e^{-3x} \right]_0^2 = e^{-8} \cdot \left(1 - e^{-6} \right) \\ &\approx 0,00033 = 0,033\%. \end{aligned}$$

d) Die einzelnen Dichten erhalten wir durch „Herausintegrieren“ der anderen Variable:

$$\begin{aligned} F_X(t) &= \int_{-\infty}^t \int_{-\infty}^{\infty} f(x,y) dy dx = \int_0^t \int_0^{\infty} 12 \cdot e^{-3x} \cdot e^{-4y} dy dx = \int_0^t 12 \cdot e^{-3x} \int_0^{\infty} e^{-4y} dy dx \\ &= \int_0^t 12 \cdot e^{-3x} \cdot \left[-\frac{1}{4} \cdot e^{-4y} \right]_0^{\infty} dx = \int_0^t 3 \cdot e^{-3x} dx = 3 \cdot \left[-\frac{1}{3} \cdot e^{-3x} \right]_0^t = 1 - e^{-3t}. \end{aligned}$$

Analog erhalten wir $F_Y(t) = 1 - e^{-4t}$.

e) Da $F_{XY}(t_x, t_y) = F_X(t_x) \cdot F_Y(t_y)$ gilt, sind die Zufallsvariablen stochastisch unabhängig. Damit ergibt sich

$$f(x | Y = y) = f_X(x) = \frac{d}{dx} F_X = 3 \cdot e^{-3x}.$$

Das gleiche Ergebnis erhalten wir anhand der Formel $f(x | Y = y) = \frac{f(x,y)}{f_Y(y)}$.

zu A.2.53:

Für die Dichte stellen wir zunächst die Kovarianzmatrix auf

$$\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$$

und bilden die Determinante $\det(\Sigma) = 1 \cdot 4 - 1 \cdot 1 = 3$.

Für den Exponenten berechnen wir

$$(x-0, y-0) \cdot \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} x-0 \\ y-0 \end{pmatrix} = (x, y) \cdot \begin{pmatrix} x+y \\ x+4y \end{pmatrix} = x^2 + xy + xy + 4y^2 = x^2 + 2xy + 4y^2.$$

Insgesamt erhalten wir damit die Dichte

$$f(x, y) = \frac{1}{2\pi \cdot \sqrt{3}} \cdot e^{-\frac{1}{2}x^2 - xy - 2y^2}.$$

2.12 zu Abschätzungen und Grenzwertsätze

zu A.2.54:

Da jedes Tor mit gleicher Wahrscheinlichkeit erzielt wird, liegt eine Wiederholung eines Experiments mit Erfolgswahrscheinlichkeit $p = \frac{1}{4}$ vor. Wir modellieren daher das Vorgehen mit der Zufallsvariable $X \sim \mathcal{B}\left(100, \frac{1}{4}\right)$. Mit der Markow-Ungleichung berechnen wir

$$\mathbb{P}(X \geq 30) \leq \frac{\mathbb{E}[X]}{30} = \frac{100 \cdot \frac{1}{4}}{30} = \frac{5}{6} \approx 83,33\%.$$

zu A.2.55:

Diese Aufgabe lösen wir mit Hilfe der Tschebyscheff-Ungleichung (T.U.). Das Vorgehen modellieren wir mit einer Zufallsvariable $X \sim \mathcal{B}\left(1800, \frac{1}{6}\right)$, da es sich um das Wiederholen eines gleichbleibenden Vorgangs handelt. Alternativ ist die Modellierung mit einer Summe von Bernoulli-Verteilungen denkbar. In unserem Fall gilt mit $\mathbb{E}[X] = 1800 \cdot \frac{1}{6} = 300$ und $\text{Var}(X) = 1800 \cdot \frac{1}{6} \cdot \left(1 - \frac{1}{6}\right) = 250$

$$\begin{aligned} \mathbb{P}(250 < X < 350) &= \mathbb{P}(250 - 300 < X - 300 < 350 - 300) = \mathbb{P}(-50 < X - 300 < 50) \\ &= \mathbb{P}(|X - 300| < 50) \stackrel{\text{T.U.}}{\geq} 1 - \frac{\text{Var}(X)}{k^2} = 1 - \frac{250}{50^2} = 1 - \frac{1}{10} = 90\%. \end{aligned}$$

zu A.2.56:

Bei diesem Beispiel modellieren wir den wiederholten Münzwurf durch die Zufallsvariable $X \sim \mathcal{B}\left(1500, \frac{1}{2}\right)$ und formen die Wahrscheinlichkeit auf die gleiche Weise wie in der letzten Aufgabe um und erhalten mit $\mathbb{E}[X] = 1500 \cdot \frac{1}{2} = 750$ und $\text{Var}(X) = 1500 \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right) = 375$

$$\begin{aligned} \mathbb{P}(650 < X < 850) &= \mathbb{P}(650 - 750 < X - 750 < 850 - 750) = \mathbb{P}(-100 < X - 750 < 100) = \mathbb{P}(|X - 750| < 100) \\ &\stackrel{\text{T.U.}}{\geq} 1 - \frac{\text{Var}(X)}{100^2} = 1 - \frac{375}{10000} = 0,9625 = 96,25\%. \end{aligned}$$

zu A.2.57:

Für das Standardisieren im zentralen Grenzwertsatz benötigen wir den Erwartungswert $\mathbb{E}[X_k] = 0$, da $X_k \sim t(6)$ und die Varianz $\text{Var}(X_k) = \frac{6}{6-4} = 1,5$. Damit erhalten wir

$$\begin{aligned} \mathbb{P}(X > 25) &= \mathbb{P}\left(\frac{X - 100 \cdot 0}{\sqrt{100 \cdot \sqrt{1,5}}} > \frac{25 - 100 \cdot 0}{\sqrt{100 \cdot \sqrt{1,5}}}\right) \approx \mathbb{P}\left(\frac{X}{10 \cdot \sqrt{1,5}} > 2,04\right) = 1 - \mathbb{P}\left(\frac{X}{10 \cdot \sqrt{1,5}} \leq 2,04\right) \\ &\approx 1 - \phi(2,04) \stackrel{\text{aus Tab.}}{\approx} 1 - 0,97933 = 0,02067 = 2,067\%. \end{aligned}$$

zu A.2.58:

Zunächst übersetzen wir den Text in einen mathematischen Ausdruck. Gesucht ist n , sodass

$$\mathbb{P}\left(0,95 \cdot \frac{1}{6} < \frac{1}{n} \cdot \sum_{k=1}^n X_k < 1,05 \cdot \frac{1}{6}\right)$$

gilt. Die Ausdrücke $0,95 \cdot \frac{1}{6}$ und $1,05 \cdot \frac{1}{6}$ beschreiben die Abweichung sowohl nach unten als auch nach oben zu $\frac{1}{6}$, die nach Aufgabenstellung jeweils maximal 5% betragen soll.

Zunächst bestimmen wir n mit Hilfe des zentralen Grenzwertsatzes. Da X_k nur den Ausgang eines Wurfs beschreibt, erhalten wir $\mathbb{E}[X_k] = 1 \cdot \frac{1}{6} = \frac{1}{6}$ und $\text{Var}(X_k) = 1 \cdot \frac{1}{6} \cdot \left(1 - \frac{1}{6}\right) = \frac{5}{36}$. Insgesamt folgt

mit abkürzender Schreibweise $S_n = \sum_{k=1}^n X_k$

$$\begin{aligned} \mathbb{P}\left(0,95 \cdot \frac{1}{6} < \frac{1}{n} \cdot \sum_{k=1}^n X_k < 1,05 \cdot \frac{1}{6}\right) &= \mathbb{P}\left(n \cdot 0,95 \cdot \frac{1}{6} < S_n < n \cdot 1,05 \cdot \frac{1}{6}\right) \\ &= \mathbb{P}\left(n \cdot 0,95 \cdot \frac{1}{6} - n \cdot \frac{1}{6} < S_n - n \cdot \frac{1}{6} < n \cdot 1,05 \cdot \frac{1}{6} - n \cdot \frac{1}{6}\right) \\ &= \mathbb{P}\left(-0,05 \cdot n \cdot \frac{1}{6} < S_n - n \cdot \frac{1}{6} < 0,05 \cdot n \cdot \frac{1}{6}\right) \\ &= \mathbb{P}\left(\frac{-0,05 \cdot n \cdot \frac{1}{6}}{\sqrt{n} \cdot \sqrt{\frac{5}{36}}} < \frac{S_n - n \cdot \frac{1}{6}}{\sqrt{n} \cdot \sqrt{\frac{5}{36}}} < \frac{0,05 \cdot n \cdot \frac{1}{6}}{\sqrt{n} \cdot \sqrt{\frac{5}{36}}}\right) \\ &\approx \phi\left(\frac{0,05 \cdot n \cdot \frac{1}{6}}{\sqrt{n} \cdot \sqrt{\frac{5}{36}}}\right) - \phi\left(\frac{-0,05 \cdot n \cdot \frac{1}{6}}{\sqrt{n} \cdot \sqrt{\frac{5}{36}}}\right) \\ &= \phi\left(\frac{0,05 \cdot n \cdot \frac{1}{6}}{\sqrt{n} \cdot \sqrt{\frac{5}{36}}}\right) - \left(1 - \phi\left(\frac{0,05 \cdot n \cdot \frac{1}{6}}{\sqrt{n} \cdot \sqrt{\frac{5}{36}}}\right)\right) \\ &= 2 \cdot \phi\left(\frac{0,05 \cdot n \cdot \frac{1}{6}}{\sqrt{n} \cdot \sqrt{\frac{5}{36}}}\right) - 1. \end{aligned}$$

Dieser Wert soll nun nach Aufgabenstellung 90% ergeben. Bevor wir n bestimmen, vereinfachen wir den Term innerhalb von ϕ durch

$$\frac{0,05 \cdot n \cdot \frac{1}{6}}{\sqrt{n} \cdot \sqrt{\frac{5}{36}}} = 0,05 \cdot \sqrt{n} \cdot \frac{1}{6} \cdot \sqrt{\frac{36}{5}} = 0,05 \cdot \sqrt{n} \cdot \frac{1}{\sqrt{5}}.$$

Damit berechnen wir

$$\begin{aligned}0,9 &= 2 \cdot \Phi\left(0,05 \cdot \sqrt{n} \cdot \frac{1}{\sqrt{5}}\right) - 1 \\ \Leftrightarrow 1,9 &= 2 \cdot \Phi\left(0,05 \cdot \sqrt{n} \cdot \frac{1}{\sqrt{5}}\right) \\ \Leftrightarrow 0,95 &= \Phi\left(0,05 \cdot \sqrt{n} \cdot \frac{1}{\sqrt{5}}\right) \\ \Leftrightarrow \Phi^{-1}(0,95) &= 0,05 \cdot \sqrt{n} \cdot \frac{1}{\sqrt{5}} \\ \Leftrightarrow \Phi^{-1}(0,95) \cdot 20 \cdot \sqrt{5} &= \sqrt{n} \\ \Leftrightarrow n &= 2000 \cdot (\Phi^{-1}(0,95))^2\end{aligned}$$

Den Wert $\Phi^{-1}(0,95) = 1,645$ lesen wir aus der Tabelle ab, sodass wir wegen $n \approx 5412,05$ mindestens $n = 5413$ Mal würfeln müssen.

3 Lösungen zu Inferenzstatistik

3.1 zu Grundbegriffe

keine Aufgaben vorhanden

3.2 zu Parameterschätzung

zu A.3.1:

Wir berechnen anhand der Definition

$$\begin{aligned}\bar{x} &= \frac{1}{n} \cdot \sum_{k=1}^n x_k &= \frac{1}{10} \cdot (1+2+3+1+5+3+2+5+6+7) &= 3,5 \\ s^2 &= \frac{1}{n} \cdot \sum_{k=1}^n (x_k - \bar{x})^2 &= \frac{1}{10} \cdot ((1-3,5)^2 + (2-3,5)^2 + \dots + (6-3,5)^2 + (7-3,5)^2) &= 4,05 \\ s^{*2} &= \frac{1}{n-1} \cdot \sum_{k=1}^n (x_k - \bar{x})^2 &= \frac{1}{9} \cdot ((1-3,5)^2 + (2-3,5)^2 + \dots + (6-3,5)^2 + (7-3,5)^2) &= 4,5.\end{aligned}$$

zu A.3.2:

Für die Erwartungstreue berechnen wir jeweils den Erwartungswert der Schätzer T_i :

$$\begin{aligned}\text{a) } \mathbb{E}[T_1] &= \mathbb{E}\left[\frac{1}{4} \cdot (X_1 + X_2 + X_4)\right] = \frac{1}{4} \cdot (\mathbb{E}[X_1] + \mathbb{E}[X_2] + \mathbb{E}[X_3]) = \frac{1}{4} \cdot (\mu + \mu + \mu) = \frac{3}{4} \cdot \mu \\ \text{b) } \mathbb{E}[T_2] &= \mathbb{E}\left[\frac{X_1}{2} + 2 \cdot X_2 - \frac{3}{4} \cdot (X_3 + X_4)\right] = \frac{1}{2} \cdot \mathbb{E}[X_1] + 2 \cdot \mathbb{E}[X_2] - \frac{3}{4} \cdot (\mathbb{E}[X_3] + \mathbb{E}[X_4]) \\ &= \frac{1}{2} \cdot \mu + 2 \cdot \mu - \frac{3}{4} \cdot (\mu + \mu) = \mu. \\ \text{c) } \mathbb{E}[T_3] &= \mathbb{E}\left[\frac{1}{10} \cdot (2 \cdot X_1 + 2 \cdot X_2 + 3 \cdot X_3 + 4 \cdot X_4)\right] = \frac{2}{10} \cdot \mathbb{E}[X_1] + \frac{2}{10} \cdot \mathbb{E}[X_2] + \frac{3}{10} \cdot \mathbb{E}[X_3] + \frac{4}{10} \cdot \\ &\mathbb{E}[X_4] = \frac{11}{10} \cdot \mu \\ \text{d) } \mathbb{E}[T_4] &= \mathbb{E}\left[2 \cdot \left(\frac{1}{8} \cdot X_1 + \frac{1}{10} \cdot X_2 + \frac{1}{4} \cdot X_3 + \frac{1}{40} \cdot X_4\right)\right] = \frac{1}{4} \cdot \mathbb{E}[X_1] + \frac{1}{5} \cdot \mathbb{E}[X_2] + \frac{1}{2} \cdot \mathbb{E}[X_3] + \frac{1}{20} \cdot \\ &\mathbb{E}[X_4] = \frac{1}{4} \cdot \mu + \frac{1}{5} \cdot \mu + \frac{1}{2} \cdot \mu + \frac{1}{20} \cdot \mu = \mu.\end{aligned}$$

Damit sind die Schätzer T_2 und T_4 erwartungstreue Schätzer für den Mittelwert μ .

Für die Effizienz können wir die Varianz der Summe durch die Summe der Varianzen berechnen, da die Stichprobenvariablen unabhängig sind. Anschließend ziehen wir die Faktoren quadratisch heraus. Es gilt

$$\begin{aligned} \text{Var}(T_2) &= \text{Var}\left(\frac{X_1}{2} + 2 \cdot X_2 - \frac{3}{4} \cdot (X_3 + X_4)\right) \\ &= \text{Var}\left(\frac{1}{2} \cdot X_1\right) + \text{Var}(2 \cdot X_2) + \text{Var}\left(-\frac{3}{4} \cdot X_3\right) + \text{Var}\left(-\frac{3}{4} \cdot X_4\right) \\ &= \frac{1}{4} \cdot \text{Var}(X_1) + 4 \cdot \text{Var}(X_2) + \frac{9}{16} \cdot \text{Var}(X_3) + \frac{9}{16} \cdot \text{Var}(X_4) \\ &= \frac{1}{4} \cdot \sigma^2 + 4 \cdot \sigma^2 + \frac{9}{16} \cdot \sigma^2 + \frac{9}{16} \cdot \sigma^2 \\ &= \frac{43}{8} \cdot \sigma^2 \end{aligned}$$

sowie

$$\begin{aligned} \text{Var}(T_4) &= \text{Var}\left(2 \cdot \left(\frac{1}{8} \cdot X_1 + \frac{1}{10} \cdot X_2 + \frac{1}{4} \cdot X_3 + \frac{1}{40} \cdot X_4\right)\right) \\ &= \text{Var}\left(\frac{1}{4} \cdot X_1\right) + \text{Var}\left(\frac{1}{5} \cdot X_2\right) + \text{Var}\left(\frac{1}{2} \cdot X_3\right) + \text{Var}\left(\frac{1}{20} \cdot X_4\right) \\ &= \frac{1}{16} \cdot \text{Var}(X_1) + \frac{1}{25} \cdot \text{Var}(X_2) + \frac{1}{4} \cdot \text{Var}(X_3) + \frac{1}{400} \cdot \text{Var}(X_4) \\ &= \frac{1}{16} \cdot \sigma^2 + \frac{1}{25} \cdot \sigma^2 + \frac{1}{4} \cdot \sigma^2 + \frac{1}{400} \cdot \sigma^2 \\ &= \frac{5}{4} \cdot \sigma^2. \end{aligned}$$

Da $\frac{5}{4} < \frac{43}{8}$, ist T_4 effizienter als T_2 .

zu A.3.3:

Da die Minuten, die bis zum Anfahren der Raststätte exponentialverteilt sind, gilt

$$\mathbb{E}[X] = \frac{1}{\lambda} \Leftrightarrow \lambda = \frac{1}{\mathbb{E}[X]}.$$

Anschließend berechnen wir das erste Moment, da wir nur einen Parameter schätzen durch

$$\bar{x} = \frac{1}{12} \cdot (15 + 22 + 15 + 17 + 25 + 31 + 13 + 39 + 22 + 14 + 17 + 20) = \frac{125}{6}$$

und erhalten mit der Momentenmethode den Schätzer

$$\hat{\lambda} = \frac{1}{\bar{x}} = \frac{1}{\frac{125}{6}} = 0,048.$$

zu A.3.4:

Für die Bestimmung des Schätzers $\hat{\lambda}$ stellen wir zunächst die Likelihood-Funktion über das Produkt der Dichten auf:

$$L(x_1, \dots, x_n, \lambda) = f_\lambda(x_1) \cdot \dots \cdot f_\lambda(x_n) = \frac{\lambda^{x_1}}{x_1!} \cdot e^{-\lambda} \cdot \dots \cdot \frac{\lambda^{x_n}}{x_n!} \cdot e^{-\lambda} = \frac{\lambda^{\sum_{k=1}^n x_k}}{x_1! \cdot \dots \cdot x_n!} \cdot e^{-n \cdot \lambda}$$

Anschließend bilden wir die Log-Likelihood-Funktion, damit wir das Maximum einfacher bestimmen können. Diese vereinfachen wir zudem mit Hilfe der Logarithmusgesetze

$$\begin{aligned}\ell(x_1, \dots, x_n, \lambda) &= \ell(\lambda) = \ln \left(\frac{\lambda^{\sum_{k=1}^n x_k}}{x_1! \cdot \dots \cdot x_n!} \cdot e^{-n \cdot \lambda} \right) = \ln \left(\frac{\lambda^{\sum_{k=1}^n x_k}}{x_1! \cdot \dots \cdot x_n!} \right) + \ln \left(e^{-n \cdot \lambda} \right) \\ &= \ln \left(\lambda^{\sum_{k=1}^n x_k} \right) - \ln(x_1! \cdot \dots \cdot x_n!) - n \cdot \lambda = \sum_{k=1}^n x_k \cdot \ln(\lambda) - \ln(x_1! \cdot \dots \cdot x_n!) - n \cdot \lambda.\end{aligned}$$

Nun bilden wir die ersten beiden Ableitungen und bestimmen das Maximum der Log-Likelihood-Funktion. Es ist

$$\frac{\partial}{\partial \lambda} \ell(\lambda) = \frac{\sum_{k=1}^n x_k}{\lambda} - n \quad \text{und} \quad \frac{\partial^2}{\partial \lambda^2} \ell(\lambda) = -\frac{\sum_{k=1}^n x_k}{\lambda^2}.$$

Für die möglichen Kandidaten der Extremstellen berechnen wir nun die Nullstellen der ersten Ableitung:

$$\begin{aligned}\frac{\partial}{\partial \lambda} \ell & \stackrel{!}{=} 0 \\ \Leftrightarrow \frac{\sum_{k=1}^n x_k}{\lambda} - n &= 0 \\ \Leftrightarrow \frac{\sum_{k=1}^n x_k}{\lambda} &= n \\ \Leftrightarrow \sum_{k=1}^n x_k &= n \cdot \lambda \\ \Leftrightarrow \frac{\sum_{k=1}^n x_k}{n} &= \lambda\end{aligned}$$

Da die Realisationen x_1, \dots, x_n von Poisson-verteilten Stichprobenvariablen stets positiv sind, ist die zweite Ableitung stets negativ, sodass wir mit $\hat{\lambda} = \bar{x}$ tatsächlich ein Maximum der Log-Likelihood und somit einen Schätzer für den Parameter λ mit der Maximum-Likelihood Methode gefunden haben.

zu A.3.5:

Bei dieser Aufgabe liegt die zusätzliche Schwierigkeit in den unterschiedlichen n_k . Das Vorgehen bleibt gleich, sodass wir zunächst die Likelihood-Funktion aufstellen:

$$L(x_1, \dots, x_n, p) = f_p(x_1) \cdot \dots \cdot f_p(x_m) = \binom{n_1}{x_1} \cdot p^{x_1} \cdot (1-p)^{n_1-x_1} \cdot \dots \cdot \binom{n_m}{x_m} \cdot p^{x_m} \cdot (1-p)^{n_m-x_m}$$

Diese schreiben wir kompakt mit Hilfe des Produktzeichens als

$$L(p) = L(x_1, \dots, x_n, p) = \prod_{k=1}^m \binom{n_k}{x_k} \cdot p^{x_k} \cdot (1-p)^{n_k-x_k}.$$

Anschließend wenden wir den Logarithmus auf die Funktion und diverse Logarithmengesetze an. Damit erhalten wir die Log-Likelihood-Funktion mit

$$\begin{aligned}\ell(p) = \ln(L(p)) &= \ln \left(\prod_{k=1}^m \binom{n_k}{x_k} \cdot p^{x_k} \cdot (1-p)^{n_k-x_k} \right) = \sum_{k=1}^m \ln \left(\binom{n_k}{x_k} \cdot p^{x_k} \cdot (1-p)^{n_k-x_k} \right) \\ &= \sum_{k=1}^m \left[\ln \left(\binom{n_k}{x_k} \right) + \ln(p^{x_k}) + \ln((1-p)^{n_k-x_k}) \right] \\ &= \sum_{k=1}^m \left[\ln \left(\binom{n_k}{x_k} \right) + x_k \cdot \ln(p) + (n_k - x_k) \cdot \ln(1-p) \right].\end{aligned}$$

Wir bestimmen nun die ersten beiden Ableitungen nach dem gesuchten Parameter p der Log-Likelihood-Funktion und nutzen aus, dass wir Summen einzeln ableiten dürfen, sodass wir alle Summanden innerhalb der äußeren Summe ableiten. Dabei ergibt die Ableitung des **ersten Summanden 0**, da dieser nicht von p abhängt. Beim **letzten Summanden** beachten wir neben der Ableitung des Logarithmus noch die innere Ableitung von $(1-p)' = -1$ und erhalten

$$\frac{\partial}{\partial p} \ell(p) = \sum_{k=1}^m \left[0 + \frac{x_k}{p} - \frac{n_k - x_k}{1-p} \right] = \frac{1}{p} \cdot \sum_{k=1}^m x_k - \frac{1}{1-p} \cdot \sum_{k=1}^m n_k - x_k.$$

Für die zweite Ableitung gilt erneut die innere Ableitung von $\frac{1}{1-p}$ zu beachten, nämlich $\left(\frac{\partial}{\partial p} \frac{1}{1-p} = \frac{1}{(1-p)^2} \right)$, sodass wir

$$\frac{\partial^2}{\partial p^2} \ell(p) = -\frac{1}{p^2} \cdot \sum_{k=1}^m x_k - \frac{1}{(1-p)^2} \cdot \sum_{k=1}^m n_k - x_k$$

erhalten. Wir erkennen, dass $\frac{\partial^2}{\partial p^2} \ell(p)$ stets negativ ist, da p , x_k und $n_k - x_k$ bei binomialverteilten Stichprobenvariablen stets positiv sind. Damit ist ein möglicher Kandidat, den wir im Folgenden berechnen, direkt ein Maximum. Es gilt

$$\begin{aligned}\frac{\partial}{\partial p} \ell(p) &\stackrel{!}{=} 0 \\ \Leftrightarrow \frac{1}{p} \cdot \sum_{k=1}^m x_k - \frac{1}{1-p} \cdot \sum_{k=1}^m n_k - x_k &= 0 \\ \Leftrightarrow \frac{1}{p} \cdot \sum_{k=1}^m x_k &= \frac{1}{1-p} \cdot \sum_{k=1}^m n_k - x_k \\ \Leftrightarrow (1-p) \cdot \sum_{k=1}^m x_k &= p \cdot \sum_{k=1}^m n_k - x_k \\ \Leftrightarrow 1 \cdot \sum_{k=1}^m x_k - p \cdot \sum_{k=1}^m x_k &= p \cdot \sum_{k=1}^m n_k - x_k \\ \Leftrightarrow \sum_{k=1}^m x_k &= p \cdot \sum_{k=1}^m n_k - x_k + p \cdot \sum_{k=1}^m x_k \\ \Leftrightarrow \sum_{k=1}^m x_k &= p \cdot \sum_{k=1}^m n_k \\ \Leftrightarrow \frac{\sum_{k=1}^m x_k}{\sum_{k=1}^m n_k} &= p.\end{aligned}$$

Wir erhalten mit der Maximum-Likelihood-Methode den Schätzer $\hat{\rho} = \frac{\sum_{k=1}^m x_k}{\sum_{k=1}^m n_k}$.

zu A.3.6:

Wir gehen wie in den beiden vorigen Aufgaben vor und stellen zunächst die Likelihood-Funktion auf. Diese vereinfachen wir mit Hilfe des Produktzeichens und der Potenzgesetze zu

$$\begin{aligned} L(\mu) &= L(x_1, \dots, x_n, \mu) = f_\mu(x_1) \cdot \dots \cdot f_\mu(x_n) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x_1 - \mu)^2}{2}} \cdot \dots \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x_n - \mu)^2}{2}} \\ &= \prod_{k=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_k - \mu)^2}{2}} = \frac{1}{(\sqrt{2\pi})^n} \cdot e^{-\frac{1}{2} \cdot \sum_{k=1}^n (x_k - \mu)^2} \end{aligned}$$

Im Anschluss bilden wir die Log-Likelihood-Funktion

$$\begin{aligned} \ell(\mu) &= \ln(L(\mu)) = \ln\left(\frac{1}{(\sqrt{2\pi})^n} \cdot e^{-\frac{1}{2} \cdot \sum_{k=1}^n (x_k - \mu)^2}\right) = \ln\left(\frac{1}{(\sqrt{2\pi})^n}\right) + \ln\left(e^{-\frac{1}{2} \cdot \sum_{k=1}^n (x_k - \mu)^2}\right) \\ &= \ln\left(\frac{1}{(\sqrt{2\pi})^n}\right) - \frac{1}{2} \cdot \sum_{k=1}^n (x_k - \mu)^2 \cdot \ln(e) = -n \cdot \ln(\sqrt{2\pi}) - \frac{1}{2} \cdot \sum_{k=1}^n (x_k - \mu)^2. \end{aligned}$$

Als weiteren Schritt berechnen wir die ersten beiden Ableitungen. Bei der Berechnung müssen wir bei der inneren Ableitung von $(x_k - \mu)^2$ aufpassen und erhalten

$$\begin{aligned} \frac{\partial}{\partial \mu} \ell(\mu) &= \sum_{k=1}^n x_k - \mu \\ \frac{\partial^2}{\partial \mu^2} \ell(\mu) &= \sum_{k=1}^n (-1). \end{aligned}$$

Wir erkennen, dass die zweite Ableitung in jedem Fall negativ ist und somit erhalten wir mit den Nullstellen der ersten Ableitung direkt ein Maximum. Dazu berechnen wir

$$\begin{aligned} \frac{\partial}{\partial \mu} \ell(\mu) &\stackrel{!}{=} 0 \\ \Leftrightarrow \sum_{k=1}^n x_k - \mu &= 0 \\ \Leftrightarrow \sum_{k=1}^n x_k - \sum_{k=1}^n \mu &= 0 \\ \Leftrightarrow \sum_{k=1}^n x_k &= \sum_{k=1}^n \mu \\ \Leftrightarrow \sum_{k=1}^n x_k &= n \cdot \mu \end{aligned}$$

Somit erhalten wir mit der Maximum-Likelihood-Methode den Schätzer $\hat{\mu} = \frac{1}{n} \cdot \sum_{k=1}^n x_k = \bar{x}$.

zu A.3.7:

Wir gehen nach dem gängigen Muster vor:

1. Zunächst entnehmen wir der Aufgabenstellung, dass wir eine Normalverteilung mit **bekanntem Varianz $\sigma^2 = 2,5$** vorliegen haben.

2. Bei einem $90\% = 1 - \alpha$ -Konfidenzniveau erhalten wir $\alpha = 10\%$.
3. Wir erhalten die Quantile im zweiseitigen Fall durch

$$Z_{1-\frac{\alpha}{2}} = Z_{0,95} \stackrel{\text{aus Tabelle}}{\approx} 1,645.$$

4. Aus dem Aufgabentext können wir $\bar{x} = 17,5$ und $n = 54$ ablesen.
5. Wir erhalten das Intervall durch

$$\begin{aligned} I_{90\%} &= \left[\bar{x} - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right] = \left[17,5 - 1,645 \cdot \frac{\sqrt{2,5}}{\sqrt{54}}, 17,5 + 1,645 \cdot \frac{\sqrt{2,5}}{\sqrt{54}} \right] \\ &= [17,15, 17,85]. \end{aligned}$$

Die Anzahl an Testfahrten, die Hilde für ein Intervall der Länge von 30 Sekunden = 0,5 Minuten erhalten wir mit der Formel

$$n = \left(\frac{2 \cdot Z_{1-\frac{\alpha}{2}} \cdot \sigma}{L} \right)^2 = \left(\frac{2 \cdot 1,645 \cdot \sqrt{2,5}}{0,5} \right)^2 = 108,241.$$

Hilde muss demnach 109 Testfahrten durchführen.

zu A.3.8:

Wir bestimmen das Intervall anhand der bekannten Schritte:

- Wir entnehmen dem Aufgabentext, dass es sich um eine Normalverteilung handelt, bei dem die Varianz unbekannt ist, da wir lediglich die Stichprobenvarianz gegeben haben.
- Aus einem Konfidenzniveau von 95% folgt $1 - \alpha = 0,95 \Leftrightarrow \alpha = 0,05$.
- Da der Stichprobenumfang mit 8 sehr klein ist, benötigen wir Werte der t -Verteilung und lesen

$$t_{n-1, 1-\alpha} = t_{7, 0,95} \stackrel{\text{aus Tab.}}{\approx} 1,895$$

ab.

- Wir lesen aus der Aufgabenstellung die Stichprobenvarianz $s^{*2} \approx 21,71$ ab und berechnen das Stichprobenmittel durch

$$\bar{x} = \frac{1}{n} \cdot \sum_{k=1}^n x_k = \frac{1}{8} \cdot (23 + 12 + 11 + 9 + 14 + 20 + 15 + 16) = 15.$$

- Wir erhalten die beiden einseitigen Konfidenzintervalle durch

$$I_{0,95}^b = \left[-\infty, \bar{x} + t_{n-1, 1-\alpha} \cdot \frac{s^*}{\sqrt{n}} \right] = \left[-\infty, 15 + 1,895 \cdot \frac{\sqrt{21,71}}{\sqrt{8}} \right] \approx [-\infty, 18,12]$$

und

$$I_{0,95}^a = \left[\bar{x} - t_{n-1, 1-\alpha} \cdot \frac{s^*}{\sqrt{n}}, \infty \right] = \left[15 - 1,895 \cdot \frac{\sqrt{21,71}}{\sqrt{8}}, \infty \right] \approx [11,88, \infty].$$

zu A.3.9:

Wir gehen nach dem bekannten Schema vor:

1. Der Aufgabenstellung entnehmen wir, dass wir das zweiseitige Konfidenzintervall für den Anteilswert bestimmen sollen.
2. Aus der Angabe des 85% Quantils errechnen wir $1 - \alpha = 0,85 \Leftrightarrow \alpha = 0,15$.
3. Da wir ein zweiseitiges Konfidenzintervall bestimmen, lesen wir das Quantil

$$z_{1-\frac{\alpha}{2}} = z_{0,925} \stackrel{\text{aus Tab.}}{\approx} 1,44$$

ab.

4. Als fehlende Größe berechnen wir das Stichprobenmittel durch

$$\bar{x} = \frac{12}{400} = 0,03.$$

5. Wir erhalten somit das zweiseitige Konfidenzintervall für den Anteilswert

$$\begin{aligned} I_{0,85} &= \left[\bar{x} - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{x} \cdot (1-\bar{x})}{n}}, \bar{x} + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{x} \cdot (1-\bar{x})}{n}} \right] \\ &= \left[0,03 - 1,44 \cdot \sqrt{\frac{0,03 \cdot (1-0,03)}{400}}, 0,03 + 1,44 \cdot \sqrt{\frac{0,03 \cdot (1-0,03)}{400}} \right] \\ &\approx [0,018, 0,042]. \end{aligned}$$

Die notwendige Stichprobengröße für eine Intervalllänge von $L = 0,05$ erhalten wir durch

$$n \geq \left(\frac{z_{1-\frac{\alpha}{2}}}{L} \right)^2 = \left(\frac{1,44}{0,05} \right)^2 = 829,44$$

und benötigen somit $n = 830$.

zu A.3.10:

Nach unserem Schema gehen wir wie folgt vor:

1. Es handelt sich um normalverteilte Stichprobenvariablen und gesucht ist ein Konfidenzintervall der Varianz bei unbekanntem Erwartungswert.
2. Aus der Angabe des Konfidenzniveaus 90% erhalten wir $1 - \alpha = 0,9 \Leftrightarrow \alpha = 0,1$.
3. Da ein zweiseitiges Konfidenzintervall für die Varianz gesucht ist, lesen wir zwei Werte der χ^2 -Verteilung

$$\chi_{n-1, 1-\frac{\alpha}{2}}^2 = \chi_{7, 0,95}^2 \stackrel{\text{aus Tab.}}{\approx} 14,067$$

sowie

$$\chi_{n-1, \frac{\alpha}{2}}^2 = \chi_{7, 0,05}^2 \stackrel{\text{aus Tab.}}{\approx} 2,167$$

ab.

4. Als fehlende Größe benötigen wir in diesem Fall die korrigierte Stichprobenvarianz und für diese ebenso das Stichprobenmittel. Wir berechnen

$$\bar{x} = \frac{1}{n} \cdot \sum_{k=1}^n x_k = \frac{1}{8} \cdot (135 + 132 + \dots + 133 + 130) = 130 \quad \text{und}$$

$$s^{*2} = \frac{1}{n-1} \cdot \sum_{k=1}^n (x_k - \bar{x})^2 = \frac{1}{7} \cdot ((135 - 130)^2 + \dots + (130 - 130)^2) \approx 20,57.$$

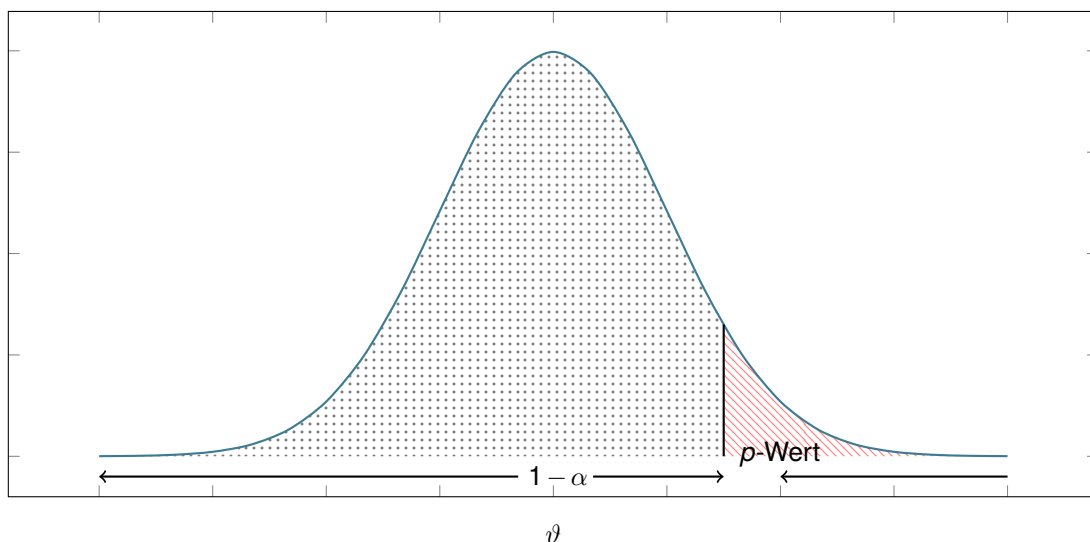
5. Wir erhalten das zweiseitige Konfidenzintervall für die Varianz durch

$$I_{0,9} = \left[\frac{(n-1) \cdot s^{*2}}{\chi_{n-1, 1-\frac{\alpha}{2}}^2}, \frac{(n-1) \cdot s^{*2}}{\chi_{n-1, \frac{\alpha}{2}}^2} \right] = \left[\frac{7 \cdot 20,57}{14,067}, \frac{7 \cdot 20,57}{2,167} \right] \approx [10,24, 66,45].$$

3.3 zu Hypothesentests

zu A.3.11:

Der p -Wert korrespondiert mit dem Fehler 1. Art, dem α -Fehler. Ist der p -Wert kleiner als die vorgegebene Signifikanz α , so wird die Nullhypothese verworfen. Je kleiner der p -Wert ist, desto unwahrscheinlicher wird die Nullhypothese. Ein Schaubild kann beispielhaft wie folgt aussehen. In diesem Schaubild ist der aus einer Stichprobe ermittelte p -Wert zu klein, da er im rot markierten Bereich liegt (also kleiner als α ist), sodass die Nullhypothese zum Niveau α verworfen wird:



zu A.3.12:

- a) Da wir **nachweisen** möchten, dass sich der **Preis erhöht** hat, definieren wir einen **rechtsseitigen Test**. Mit dem Zusatz „**im Mittel**“ wissen wir, dass die Hypothesen über den Mittelwert μ formuliert sind. Wir erhalten:

$$H_0 : \mu \leq \mu_0 \quad \text{und} \quad H_1 : \mu > \mu_0$$

- b) In diesem Beispiel möchten wir eine Aussage **widerlegen**, sodass wir die Aussage in der Nullhypothese festhalten. An dieser Stelle lässt sich debattieren, ob es sich um die durchschnittliche Anzahl von **100g Nüssen** handelt, die in einem Paket enthalten sein sollen, sodass wir eine Aussage über den Erwartungswert treffen. Da dieser Zusatz jedoch fehlt, ist es schlüssiger und genauer, hier den Fokus auf das Wort **immer** zu legen. Dies bedeutet, dass keine nennenswerte Streuung existiert, also diese gleich 0 ist. Damit formulieren wir die Thesen:

$$H_0 : \sigma^2 = 0 \quad \text{und} \quad H_1 : \sigma^2 \neq 0$$

- c) Erneut möchten wir eine **Behauptung widerlegen**, sodass wir diese Behauptung als Nullhypothese festhalten. Der Zusatz „**durchschnittlich**“ zeigt, dass die Aussagen über den Mittelwert zu formulieren sind. Da dieser mindestens eine gewisse Größe sein soll, lautet die Alternativhypothese, dass er kleiner ist und somit erhalten wir einen linksseitigen Test mit den Hypothesen:

$$H_0 : \mu \geq 1,9 \quad \text{und} \quad H_1 : \mu < 1,9$$

zu A.3.13:

Wir beginnen mit der Aufstellung der Hypothesen. Da wir nachweisen möchten, dass der Würfel nicht fair ist, vermuten wir, dass p , also die Wahrscheinlichkeit, dass eine 1 gewürfelt wird, nicht mehr $\frac{1}{6}$ entspricht. Damit erhalten wir:

$$H_0 : p = \frac{1}{6} \quad \text{und} \quad H_1 : p \neq \frac{1}{6}$$

Da wir einen zweiseitigen Test durchführen (wir möchten nachweisen, dass der Würfel nicht fair ist), berechnen wir entweder sowohl p_r als auch p_l nach den bekannten Formeln oder überführen den Test in einen rechtsseitigen Test, sodass nur die Berechnung des Wertes p_r notwendig ist. Dazu benutzen wir die Modellierung mit der Teststatistik $T = |A - 3|$, wobei A der Menge der gewürfelten 1en entspricht. Wir betrachten dabei gerade diese Statistik, da wir im Mittel bei einer Wahrscheinlichkeit von $p = \frac{1}{6}$ und 18 Würfeln drei 1en erwarten. Somit erhalten wir mit dem Betrag und der Differenz sowohl die zu viel als auch die zu wenig geworfenen 1en der Stichprobe und es genügt die Berechnung von p_r . Aus der Stichprobe erhalten wir, dass nur eine 1 gewürfelt wurde, sodass wir $\varphi(x_1, \dots, x_{18}) = |1 - 3| = 2$ erhalten. Damit ergibt sich

$$\begin{aligned} p_r &= \mathbb{P}(T \geq \varphi \mid H_0 \text{ ist wahr}) = \mathbb{P}\left(T \geq 2 \mid p = \frac{1}{6}\right) = 1 - \mathbb{P}\left(T < 2 \mid p = \frac{1}{6}\right) \\ &= 1 - \left(\mathbb{P}\left(T = 0 \mid p = \frac{1}{6}\right) + \mathbb{P}\left(T = 1 \mid p = \frac{1}{6}\right)\right) \\ &= 1 - \left(\mathbb{P}\left(|A - 3| = 0 \mid p = \frac{1}{6}\right) + \mathbb{P}\left(|A - 3| = 1 \mid p = \frac{1}{6}\right)\right) \\ &= 1 - \left(\mathbb{P}\left(A = 3 \mid p = \frac{1}{6}\right) + \mathbb{P}\left(A = 2 \mid p = \frac{1}{6}\right) + \mathbb{P}\left(A = 4 \mid p = \frac{1}{6}\right)\right) \\ &= 1 - \left(\binom{18}{3} \cdot \left(\frac{1}{6}\right)^3 \cdot \left(1 - \frac{1}{6}\right)^{18-3} + \binom{18}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(1 - \frac{1}{6}\right)^{18-2} + \binom{18}{4} \cdot \left(\frac{1}{6}\right)^4 \cdot \left(1 - \frac{1}{6}\right)^{18-4}\right) \\ &\approx 1 - (0,245198 + 0,229874 + 0,183899) \\ &\approx 0,3410 = 34,1\%. \end{aligned}$$

Die Nullhypothese wird demnach nicht verworfen, da $p_r > \alpha$.

zu A.3.14:

Die Vermutung des Fachmanns ist, dass das Gerät defekt ist und im Mittel nicht mehr die erwartete Anzahl an Meldungen pro Stunde überträgt. Gehen wir streng vor, so ist die Aussage des Fachmanns in keine Richtung gegeben, sondern nur durch zu viele Meldungen entstanden. Daher gehen wir von einem zweiseitigen Test des Erwartungswerts aus. Für eine poisson-verteilte Zufallsvariable X mit Parameter $\lambda = \frac{1}{4}$ gilt

$$\mathbb{E}[X] = \lambda = \frac{1}{4}.$$

Wir formulieren daher

$$H_0 : \mu = \lambda = \frac{1}{4} \quad \text{und} \quad H_1 : \mu = \lambda \neq \frac{1}{4}.$$

Für den p -Wert berechnen wir nun direkt

$$\begin{aligned} p_r &= \mathbb{P}(X > 8 \mid H_0 \text{ ist wahr}) = \mathbb{P}\left(X > 8 \mid \lambda = \frac{1}{4}\right) = 1 - \mathbb{P}\left(X \leq 8 \mid \lambda = \frac{1}{4}\right) = 1 - F_X(8) \\ &= \sum_{k=0}^8 \frac{\left(\frac{1}{4}\right)^k}{k!} \cdot e^{-\frac{1}{4}} = e^{-\frac{1}{4}} \cdot \left(\frac{\left(\frac{1}{4}\right)^0}{0!} + \frac{\left(\frac{1}{4}\right)^1}{1!} + \dots + \frac{\left(\frac{1}{4}\right)^8}{8!}\right) \approx 1 - 0,999999999999 \approx 0. \end{aligned}$$

Für unser gewähltes α ist die Rundung auf 0 nicht mehr ausschlaggebend. Somit verwerfen wir zu gegebenem $\alpha = 5\%$ die Nullhypothese, da p_r bereits hinreichend klein ist.

3.4 zu Parametrische Einstichproben-Tests

zu A.3.15:

Bei der Lösung dieser Aufgabe gehen wir nach dem bekannten Schema vor:

1. Für das Aufstellen der Hypothesen erkennen wir:

- Wir möchten **nachweisen**, dass die Golfbälle zu schwer ($>$) sind (spricht für H_1).
- Die Aussage, dass der Club sich an das Maximalgewicht (\leq) hält, wird **in Frage gestellt** und soll somit widerlegt werden (spricht für H_0).

Aus dem Text geht hervor, dass wir einen rechtsseitigen Test betrachten, da wir nachweisen möchten, dass die Bälle schwerer sind, also das Gewicht größer als erlaubt ist, und formulieren daher

$$H_0 : \mu \leq \mu_0 = 46 \quad \text{und} \quad H_1 : \mu > \mu_0 = 46.$$

2. Wir erhalten direkt aus der Aufgabenstellung $\alpha = 1\%$.

3. Durch die Angabe der **Stichprobenstandardabweichung** erkennen wir, dass die tatsächliche Standardabweichung und somit die Varianz nicht bekannt sind. Daher befinden wir uns im einseitigen Einstichproben t-Test und betrachten die Teststatistik

$$T = \sqrt{n} \cdot \frac{\bar{X} - \mu_0}{S^*}.$$

4. Für die Prüfgröße liegen bereits alle Größen vor, sodass wir

$$t = \sqrt{n} \cdot \frac{\bar{x} - \mu_0}{s^*} = \sqrt{25} \cdot \frac{50,5 - 46}{9} = 2,5$$

berechnen.

5. Den kritischen Wert erhalten wir für einen rechtsseitigen Test durch

$$t_{n-1, 1-\alpha} = t_{24, 0,99} \stackrel{\text{aus Tab.}}{\approx} 2,492.$$

6. Damit können wir die Nullhypothese zum Signifikanzniveau $\alpha = 1\%$ verwerfen, denn

$$t = 2,5 > 2,492 \approx t_{24, 0,99}.$$

zu A.3.16:

Wir gehen nach unserem bekannten Schema vor:

1. Für das Aufstellen der Hypothesen erkennen wir:

- Wir möchten **nachweisen**, dass der Vater durchschnittlich mehr ($>$) Müll verursacht (spricht für H_1).

Aus der Aufgabenstellung erkennen wir zudem, dass es sich um einen rechtsseitigen Test handelt, da wir nachweisen möchten, dass mehr Müll produziert wird und formulieren daher die Thesen

$$H_0 : \mu \leq \mu_0 = 38 \quad \text{und} \quad H_1 : \mu > \mu_0 = 38.$$

2. Wir erhalten direkt aus der Aufgabenstellung $\alpha = 10\%$.

3. Durch die Angabe der **Varianz** befinden wir uns im einseitigen Einstichproben Gauß-Test und betrachten die Teststatistik

$$T = \sqrt{n} \cdot \frac{\bar{X} - \mu_0}{\sigma}.$$

4. Für die Prüfgröße benötigen wir noch das Stichprobenmittel und berechnen dazu

$$\bar{x} = \frac{1}{n} \cdot \sum_{k=1}^n x_k = \frac{1}{6} \cdot (40 + 37 + 35 + 41 + 45 + 41) \approx 39,83.$$

Damit erhalten wir

$$t = \sqrt{n} \cdot \frac{\bar{x} - \mu_0}{\sigma} = \sqrt{6} \cdot \frac{39,83 - 38}{\sqrt{8}} \approx 1,58.$$

5. Den kritischen Wert erhalten wir für einen rechtsseitigen Test durch

$$t_{n-1, 1-\alpha} = t_{5, 0,90} \stackrel{\text{aus Tab.}}{\approx} 1,476.$$

6. Damit können wir die Nullhypothese zum Signifikanzniveau $\alpha = 10\%$ verwerfen, denn

$$t = 1,58 > 1,476 \approx t_{5, 0,90}.$$

zu A.3.17:

Wir gehen nach unserem bekannten Schema vor:

1. Aus dem Text lesen wir, dass

- wir **nachweisen** möchten, dass der Frauenanteil nicht (\neq) 50% beträgt (spricht für H_1).

Zusätzlich wird im Text keine Richtung der Abweichung vorgegeben und es wird eine Aussage über den Anteil der Frauen in der Führungsetage getätigt, daher formulieren wir

$$H_0 : p = p_0 = \frac{1}{2} \quad \text{und} \quad H_1 : p \neq p_0 = \frac{1}{2}.$$

2. Wir erhalten direkt aus der Aufgabenstellung $\alpha = 2\%$.

3. Durch die Frage nach dem Anteil befinden wir uns in einem zweiseitigen approximativen Einstichproben Binomialtest für den Anteilswert und benötigen die Teststatistik

$$T = \sqrt{n} \cdot \frac{\bar{X} - p_0}{\sqrt{\bar{X} \cdot (1 - \bar{X})}}.$$

4. Den Anteil der Stichprobe erhalten wir aus der Aussage, dass **31** von **82** Personen der Führungsetage Frauen sind, sodass $\bar{x} = \frac{31}{82} \approx 0,38$ gilt und wir somit die Prüfgröße

$$t = \sqrt{n} \cdot \frac{\bar{x} - p_0}{\sqrt{\bar{x} \cdot (1 - \bar{x})}} = \sqrt{82} \cdot \frac{0,38 - 0,5}{\sqrt{0,38 \cdot (1 - 0,38)}} \approx -2,24$$

berechnen.

5. Den kritischen Wert erhalten wir für einen zweiseitigen Test durch

$$z_{1 - \frac{\alpha}{2}} = z_{0,99} \stackrel{\text{aus Tab.}}{\approx} 2,33.$$

6. Damit können wir die Nullhypothese zum Signifikanzniveau $\alpha = 2\%$ nicht verwerfen, denn

$$|t| \approx 2,24 \not\geq 2,33 \approx z_{0,99}.$$

zu A.3.18:

Mit dem Schema gilt:

1. Aus dem Text lesen wir, dass

- wir **nachweisen** möchten, dass die Varianz **niedriger** ($<$) geworden ist (spricht für H_1).

Da in der Aufgabenstellung Thesen über die Varianz aufgestellt werden und wir nachweisen möchten, dass sie kleiner ist, formulieren wir einen linksseitigen Test durch

$$H_0 : \sigma^2 \geq \sigma_0^2 = 9 \quad \text{und} \quad H_1 : \sigma^2 < \sigma_0^2 = 9.$$

2. Wir erhalten direkt aus der Aufgabenstellung $\alpha = 10\%$.

3. Durch die Frage nach der Varianz führen wir einen linksseitigen Einstichproben χ^2 -Test für die Varianz durch und benötigen betrachten die Teststatistik

$$T = \frac{(n-1) \cdot S^{*2}}{\sigma_0^2}.$$

4. Für die weitere Berechnung fehlt und die korrigierte Stichprobenvarianz und somit ebenfalls das Stichprobenmittel. Diese berechnen wir durch

$$\bar{x} = \frac{1}{n} \cdot \sum_{k=1}^n x_k = \frac{1}{10} \cdot (62 + 63 + 65 + 61 + 68 + 63 + 64 + 62 + 61 + 60) = 62,9$$

und

$$s^{*2} = \frac{1}{n-1} \cdot \sum_{k=1}^n (x_k - \bar{x})^2 = \frac{1}{10-1} \cdot ((62 - 62,9)^2 + \dots + (60 - 62,9)^2) = 5,43.$$

Damit berechnen wir die Prüfgröße durch

$$t = \frac{(n-1) \cdot s^{*2}}{\sigma_0^2} = \frac{(10-1) \cdot 5,43}{9} \approx 5,43.$$

5. Den kritischen Wert erhalten wir für einen linksseitigen Test durch

$$\chi_{n-1, \alpha}^2 = \chi_{9, 0,10}^2 \stackrel{\text{aus Tab.}}{\approx} 4,168.$$

6. Damit können wir die Nullhypothese zum Signifikanzniveau $\alpha = 2\%$ nicht verwerfen, denn

$$t \approx 5,43 \not\leq 4,168 \approx \chi_{9, 0,10}^2.$$

3.5 zu Parametrische Zweistichproben-Tests

zu A.3.19:

Bei dieser Aufgabe gehen wir nach dem bekannten Schema vor:

1. Aus dem Text lesen wir, dass

- wir **nachweisen** möchten, dass der Blutdruck **gesenkt** ($<$) werden konnte (spricht für H_1).

Laut Aufgabenstellung soll nachgewiesen werden, dass der Blutdruck durch Hinzugabe des neuen Medikaments (dargestellt durch Y_k und für uns mit μ_2) signifikant mehr verringert werden konnte als durch das ältere Medikament (dargestellt durch X_k und für uns mit μ_1). Wir formulieren daher einen rechtsseitigen Test

$$H_0: \mu_1 \leq \mu_2$$

$$H_1: \mu_1 > \mu_2$$

2. Wir erhalten direkt aus der Aufgabenstellung $\alpha = 1\%$.
3. Wir befinden uns in einem Testfall mit zwei Stichproben, der ohne Kenntnis der Varianz eine Aussage über die Erwartungswerte treffen soll. Wir nehmen allerdings an, dass sich die Varianz nicht unterscheidet, sodass wir den linksseitigen Zweistichproben-t-Test durchführen und die Teststatistik

$$T = \frac{\bar{X} - \bar{Y}}{\bar{S} \cdot \sqrt{n+mn} \cdot m}$$

betrachten.

4. Bevor wir diese Prüfgröße berechnen können, benötigen wir diverse weitere Kennzahlen. Im Folgenden berechnen wir daher die Stichprobenmittel und die korrigierten Stichprobenvarianzen durch

$$\begin{aligned}\bar{x} &= \frac{1}{n} \cdot \sum_{k=1}^n x_k &= \frac{1}{9} \cdot (142 + 143 + 141 + 143 + 145 + 146 + 143 + 144 + 140) \\ &= 143\end{aligned}$$

$$\begin{aligned}\bar{y} &= \frac{1}{m} \cdot \sum_{k=1}^m y_k &= \frac{1}{9} \cdot (137 + 139 + 141 + 137 + 135 + 135 + 141 + 142 + 144) \\ &= 139\end{aligned}$$

$$\begin{aligned}s_x^{*2} &= \frac{1}{n-1} \cdot \sum_{k=1}^n (x_k - \bar{x})^2 &= \frac{1}{9-1} \cdot ((142-143)^2 + \dots + (140-143)^2) \\ &= 3,5\end{aligned}$$

$$\begin{aligned}s_y^{*2} &= \frac{1}{m-1} \cdot \sum_{k=1}^m (y_k - \bar{y})^2 &= \frac{1}{9-1} \cdot ((137-139)^2 + \dots + (144-139)^2) \\ &= 10,25.\end{aligned}$$

Anschließend können wir das gewichtete Mittel der Stichprobenvarianzen bestimmen:

$$\bar{s}^2 = \frac{(n-1) \cdot s_x^{*2} + (m-1) \cdot s_y^{*2}}{n+m-2} = \frac{8 \cdot 3,5 + 8 \cdot 10,25}{16} = 6,875$$

Insgesamt erhalten wir damit die Prüfgröße durch

$$t = \frac{\bar{x} - \bar{y}}{\bar{s} \cdot \sqrt{\frac{n+m}{n \cdot m}}} = \frac{143 - 139}{\sqrt{6,875} \cdot \sqrt{\frac{9+9}{9 \cdot 9}}} \approx 3,24.$$

5. Den kritischen Wert erhalten wir für einen linksseitigen Test durch

$$t_{n+m-2, 1-\alpha} = t_{16, 0,99} \stackrel{\text{aus Tab.}}{\approx} 1,337.$$

6. Damit können wir die Nullhypothese zum Signifikanzniveau $\alpha = 1\%$ verwerfen, denn

$$t \approx 3,24 > 1,337 \approx t_{16, 0,99}.$$

zu A.3.20:

1. Aus dem Text lesen wir, dass

- wir **widerlegen** möchten, dass das Medikament Einfluss auf die Träume der Probanden hat.

Laut Aufgabenstellung möchten wir also widerlegen, dass sich der Anteil der Personen mit schönen Träumen mit und ohne Medikament signifikant unterscheidet, daher formulieren wir

$$H_0: p_1 = p_2$$

$$H_1: p_1 \neq p_2.$$

- Wir erhalten direkt aus der Aufgabenstellung $\alpha = 5\%$.
- Wir möchten einen Test auf Gleichheit/Ungleichheit der Anteile in zwei Stichproben durchführen, sodass wir die Teststatistik des zweiseitigen approximativen Binomialtests betrachten

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\bar{P} \cdot (1 - \bar{P})}{n} + \frac{\bar{P} \cdot (1 - \bar{P})}{m}}}$$

- Da beide Stichproben den gleichen Umfang $n = m = 1000$ besitzen, erhalten wir

$$\bar{x} = \frac{600}{1000} = 0,60$$

$$\bar{y} = \frac{650}{1000} = 0,65.$$

Die Faustregeln $n \cdot p_x \cdot (1 - p_x) = n \cdot \bar{x} \cdot (1 - \bar{x}) > 9$ und $m \cdot p_y \cdot (1 - p_y) = m \cdot \bar{y} \cdot (1 - \bar{y}) > 9$ sind jeweils erfüllt. Für die Prüfgröße bestimmen wir noch den fehlenden Wert, den gewichteten Anteilswert der Gesamtstichprobe \bar{p} :

$$\bar{p} = \frac{n \cdot p_x + m \cdot p_y}{n + m} = \frac{n \cdot \bar{x} + m \cdot \bar{y}}{n + m} = \frac{1000 \cdot 0,6 + 1000 \cdot 0,65}{2000} = 0,625.$$

Insgesamt erhalten wir damit die Prüfgröße durch

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\bar{p} \cdot (1 - \bar{p})}{n} + \frac{\bar{p} \cdot (1 - \bar{p})}{m}}} = \frac{0,6 - 0,65}{\sqrt{\frac{0,625 \cdot (1 - 0,625)}{1000} + \frac{0,625 \cdot (1 - 0,625)}{1000}}} \approx -2,31.$$

- Den kritischen Wert erhalten wir für einen zweiseitigen Test durch

$$Z_{1 - \frac{\alpha}{2}} = Z_{0,975} \stackrel{\text{aus Tab.}}{=} 1,96.$$

- Damit können wir die Nullhypothese zum Signifikanzniveau $\alpha = 5\%$ verwerfen, denn

$$|t| \approx 2,31 > 1,96 = Z_{0,975}.$$

zu A.3.21:

Es gilt:

- Aus dem Text lesen wir, dass

- wir **nachweisen** möchten, dass die durchschnittliche Füllmenge vom Kalkanteil beeinflusst wird (spricht für H_1).

Anhand der Aufgabenstellung erkennen wir einen zweiseitigen Test des Erwartungswerts, da der Einfluss nicht in eine Richtung nachgewiesen werden soll, sodass wir

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

formulieren.

- Wir erhalten direkt aus der Aufgabenstellung $\alpha = 10\%$.

3. Laut des Textes ist die Varianz bekannt, sodass wir die Teststatistik eines zweiseitigen Zweistichproben Gauß-Test betrachten

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}}$$

4. Aus dem Text können wir bereits alle notwendigen Größen ablesen und berechnen die Prüfgröße

$$t = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}} = \frac{435 - 439}{\sqrt{\frac{81}{100} + \frac{64}{120}}} \approx -3,45.$$

5. Den kritischen Wert erhalten wir für einen zweiseitigen Test durch

$$Z_{1-\frac{\alpha}{2}} = Z_{0,95} \stackrel{\text{aus Tab.}}{=} 1,645.$$

6. Damit können wir die Nullhypothese zum Signifikanzniveau $\alpha = 10\%$ verwerfen, denn

$$|t| \approx 3,45 > 1,645 \approx Z_{0,95}.$$

zu A.3.22:

Wir gehen erneut nach dem bekannten Schema vor:

1. Aus dem Text lesen wir, dass

- wir **nachweisen** möchten, dass das Gericht durch die Umstellung **weniger** ($<$) Kalorien besitzt (spricht für H_1).

Wir erkennen also einen einseitigen Test. Diesen formulieren wir jedoch als rechtsseitigen Test, da unsere zweite Stichprobe die Kalorien nach Veränderung widerspiegelt und wir somit $\mu_x > \mu_y$ zeigen möchten. Somit gilt

$$H_0: \mu_1 \leq \mu_2$$

$$H_1: \mu_1 > \mu_2.$$

2. Wir erhalten direkt aus der Aufgabenstellung $\alpha = 5\%$.
3. Laut des Textes ist die Varianz unbekannt und die Stichprobenumfänge zu niedrig, sodass wir die Teststatistik eines zweiseitigen Test von Welch betrachten

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^{*2}}{n} + \frac{S_Y^{*2}}{m}}}$$

4. Bevor wir die Prüfgröße berechnen können, benötigen wir noch die Stichprobenmittel sowie die

korrigierten Stichprobenvarianzen. Diese berechnen wir durch

$$\begin{aligned}\bar{x} &= \frac{1}{n} \cdot \sum_{k=1}^n x_k &= \frac{1}{8} \cdot (1290 + 1295 + 1285 + 1301 + 1270 + 1305 + 1293 + 1297) \\ &= 1292\end{aligned}$$

$$\begin{aligned}\bar{y} &= \frac{1}{m} \cdot \sum_{k=1}^m y_k &= \frac{1}{7} \cdot (1298 + 1325 + 1275 + 1271 + 1263 + 1255 + 1336) \\ &= 1289\end{aligned}$$

$$\begin{aligned}s_x^{*2} &= \frac{1}{n-1} \cdot \sum_{k=1}^n (x_k - \bar{x})^2 &= \frac{1}{8-1} \cdot ((1290 - 1292)^2 + \dots + (1297 - 1292)^2) \\ &\approx 117,43\end{aligned}$$

$$\begin{aligned}s_y^{*2} &= \frac{1}{m-1} \cdot \sum_{k=1}^m (y_k - \bar{y})^2 &= \frac{1}{7-1} \cdot ((1298 - 1289)^2 + \dots + (1336 - 1289)^2) \\ &\approx 989,67.\end{aligned}$$

Insgesamt ergibt sich damit

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^{*2}}{n} + \frac{s_y^{*2}}{m}}} = \frac{1292 - 1289}{\sqrt{\frac{117,43}{8} + \frac{989,67}{7}}} \approx 0,24$$

5. Für den kritischen Wert benötigen wir beim Test von Welch zunächst die korrekte Anzahl an Freiheitsgraden. Diese bestimmen wir durch

$$\nu = \left\lfloor \frac{\left(\frac{s_x^{*2}}{n} + \frac{s_y^{*2}}{m}\right)^2}{\frac{1}{n-1} \cdot \left(\frac{s_x^{*2}}{n}\right)^2 + \frac{1}{m-1} \cdot \left(\frac{s_y^{*2}}{m}\right)^2} \right\rfloor = \left\lfloor \frac{\left(\frac{117,43}{8} + \frac{989,67}{7}\right)^2}{\frac{1}{8-1} \cdot \left(\frac{117,43}{8}\right)^2 + \frac{1}{7-1} \cdot \left(\frac{989,67}{7}\right)^2} \right\rfloor \approx \lfloor 7,24 \rfloor = 7$$

und erhalten den kritischen Wert

$$t_{\nu, 1-\alpha} = t_{7, 0,95} \stackrel{\text{aus Tab.}}{=} 1,895.$$

6. Damit können wir die Nullhypothese zum Signifikanzniveau $\alpha = 5\%$ nicht verwerfen, denn

$$t \approx 0,24 \not\geq 1,895 \approx t_{7, 0,95}.$$

zu A.3.23:

1. Aus dem Text lesen wir, dass

- wir **nachweisen** möchten, dass das Gewicht Einfluss auf die Abbaugeschwindigkeit von Alkohol im Blut nimmt und zu größeren ($>$) Schwankungen führt (spricht für H_1).

Wir erkennen, dass der Einfluss der Größe mit einer Richtung identifiziert wird, nämlich, dass die Abbaumenge bei größeren Personen auch deutlicher schwankt. Da die zweite Gruppe die größeren Personen umfasst, formulieren wir für die Schwankung die linksseitigen Hypothesen

$$H_0: \sigma_1^2 \geq \sigma_2^2$$

$$H_1: \sigma_1^2 < \sigma_2^2$$

2. Wir erhalten direkt aus der Aufgabenstellung $\alpha = 5\%$.
3. Laut des Textes möchten wir eine Aussage über die Schwankungen, also die Varianz treffen und benötigen daher für einen linksseitigen Zweistichproben F-Test die Teststatistik

$$T = \frac{S_X^{*2}}{S_Y^{*2}}$$

4. Aus den Stichproben berechnen wir zunächst die Stichprobenmittel und die korrigierten Stichprobenvarianzen

$$\begin{aligned}\bar{x} &= \frac{1}{n} \cdot \sum_{k=1}^n x_k &= \frac{1}{8} \cdot (0,11 + 0,13 + 0,09 + 0,11 + 0,12 + 0,1 + 0,08 + 0,09) \\ &= 0,10375\end{aligned}$$

$$\begin{aligned}\bar{y} &= \frac{1}{m} \cdot \sum_{k=1}^m y_k &= \frac{1}{9} \cdot (0,11 + 0,14 + 0,1 + 0,09 + 0,13 + 0,08 + 0,12 + 0,07 + 0,11) \\ &\approx 0,10556\end{aligned}$$

$$\begin{aligned}s_x^{*2} &= \frac{1}{n-1} \cdot \sum_{k=1}^n (x_k - \bar{x})^2 &= \frac{1}{8-1} \cdot \left((0,11 - 0,10375)^2 + \dots + (0,09 - 0,10375)^2 \right) \\ &\approx 0,00028\end{aligned}$$

$$\begin{aligned}s_y^{*2} &= \frac{1}{m-1} \cdot \sum_{k=1}^m (y_k - \bar{y})^2 &= \frac{1}{9-1} \cdot \left((0,11 - 0,10556)^2 + \dots + (0,11 - 0,10556)^2 \right) \\ &\approx 0,00053.\end{aligned}$$

Insgesamt ergibt sich damit

$$t = \frac{s_x^{*2}}{s_y^{*2}} \approx \frac{0,00028}{0,00053} \approx 0,5283.$$

5. Entweder liegt bei solch einer Aufgabe eine Tabelle vor, in welcher der Wert der F -Verteilung abgelesen werden kann oder er ist wie in diesem Fall in der Aufgabenstellung angegeben. Wir erhalten somit

$$f_{n-1, m-1, \alpha} = f_{7, 8, 0,05} \approx 0,268.$$

6. Damit können wir die Nullhypothese zum Signifikanzniveau $\alpha = 5\%$ nicht verwerfen, denn in die Prüfgröße nicht kleiner als der kritische Wert.

zu A.3.24:

Wir gehen nach dem bekannten Schema vor:

1. Wir erkennen, dass es sich um ein Merkmal „Haarwachstum“ und einen Faktor „Koffein im Shampoo“ mit insgesamt 3 Stufen handelt, sodass wir bei dieser Aufgabenstellung eine Varianzanalyse durchführen (one-way-ANOVA). Daher formulieren wir

$$H_0: \mu_1 = \mu_2 = \mu_3$$

$$H_1: \mu_i \neq \mu_j \text{ für } i \neq j.$$

2. Das Signifikanzniveau lesen wir direkt aus der Aufgabe ab: $\alpha = 5\%$.
3. Für die Prüfgröße müssen wir zunächst einige andere Werte berechnen:

Die Mittelwerte der Randverteilungen:

$$\begin{aligned}\bar{y}_{1\bullet} &= \frac{1}{n_1} \cdot \sum_{j=1}^{n_1} y_{1j} = \frac{1}{10} \cdot (1,5 + 1,3 + 1,4 + 1,2 + 1,6 + 1,5 + 1,5 + 1,3 + 1,4 + 1,6) = 1,43 \\ \bar{y}_{2\bullet} &= \frac{1}{n_2} \cdot \sum_{j=1}^{n_2} y_{2j} = \frac{1}{10} \cdot (1,4 + 1,6 + 1,5 + 1,3 + 1,7 + 1,7 + 1,2 + 1,5 + 1,8 + 1,2) = 1,49 \\ \bar{y}_{3\bullet} &= \frac{1}{n_3} \cdot \sum_{j=1}^{n_3} y_{3j} = \frac{1}{10} \cdot (1,8 + 1,3 + 1,2 + 1,1 + 1,9 + 1,8 + 1,5 + 1,4 + 1,3 + 1,4) = 1,47\end{aligned}$$

Das Gesamtmittel:

$$\bar{y}_{\bullet\bullet} = \frac{1}{N} \cdot \sum_{i=1}^{n_i} n_i \cdot y_{i\bullet} = \frac{1}{10+10+10} \cdot (10 \cdot 1,43 + 10 \cdot 1,49 + 10 \cdot 1,47) \approx 1,463.$$

Die korrigierten Faktorstufenvarianzen:

$$\begin{aligned}s_{1\bullet}^{*2} &= \frac{1}{n_1-1} \cdot \sum_{j=1}^{n_1} (y_{1j} - \bar{y}_{1\bullet})^2 = \frac{1}{10-1} \cdot \left((1,5 - 1,43)^2 + \dots + (1,6 - 1,43)^2 \right) \approx 0,0179 \\ s_{2\bullet}^{*2} &= \frac{1}{n_2-1} \cdot \sum_{j=1}^{n_2} (y_{2j} - \bar{y}_{2\bullet})^2 = \frac{1}{10-1} \cdot \left((1,4 - 1,49)^2 + \dots + (1,2 - 1,49)^2 \right) \approx 0,0454 \\ s_{3\bullet}^{*2} &= \frac{1}{n_3-1} \cdot \sum_{j=1}^{n_3} (y_{3j} - \bar{y}_{3\bullet})^2 = \frac{1}{10-1} \cdot \left((1,8 - 1,47)^2 + \dots + (1,4 - 1,47)^2 \right) \approx 0,0757\end{aligned}$$

Und die mittleren Abweichungsquadrate:

$$\begin{aligned}\text{MQA} &= \frac{1}{k-1} \cdot \sum_{i=1}^k n_i \cdot (\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet})^2 = \frac{1}{3-1} \cdot \left((1,43 - 1,463)^2 + (1,49 - 1,463)^2 \right. \\ &\quad \left. + (1,47 - 1,463)^2 \right) \approx 0,00093 \\ \text{MQR} &= \frac{1}{N-k} \cdot \sum_{i=1}^k (n_i - 1) \cdot s_{i\bullet}^{*2} = \frac{1}{10+10+10-3} \cdot \left((10-1) \cdot 0,0179 + (10-1) \cdot 0,0454 \right. \\ &\quad \left. + (10-1) \cdot 0,0757 \right) \approx 0,0463.\end{aligned}$$

Insgesamt erhalten wir die Prüfgröße durch

$$t = \frac{\text{MQA}}{\text{MQR}} \approx \frac{0,00093}{0,0463} \approx 0,02.$$

4. Den kritischen Wert entnehmen wir der Aufgabenstellung $f_{k-1, N-k, 1-\alpha} = f_{2, 27, 0,95} \approx 3,354$.
5. Wir lehnen die Nullhypothese zum Signifikanzniveau $\alpha = 5\%$ nicht ab, denn es gilt

$$t \approx 0,02 \not\geq 3,354 \approx f_{2, 27, 0,95}.$$

3.6 zu Nichtparametrische Tests

zu A.3.25:

Da die Unabhängigkeit gezeigt werden soll, führen wir den χ^2 -Unabhängigkeitstest durch.

- Zunächst stellen wir die Hypothesen auf. beim χ^2 -Unabhängigkeitstest wird die stochastische Unabhängigkeit immer in der Nullhypothese festgehalten:

H_0 : Die Zufallsvariablen „Alter“ und „Autoantrieb“ sind stochastisch unabhängig.

H_1 : Die Zufallsvariablen „Alter“ und „Autoantrieb“ sind nicht stochastisch unabhängig.

- Aus den Daten des Textes erstellen wir folgende Kontingenztabelle:

zu Autoantrieb			
	E-Auto	Benziner	Summe
Ü 40	35	30	65
U 40	25	15	40
Summe	60	45	105

- Das Signifikanzniveau beträgt $\alpha = 5\%$.
- Nun berechnen wir die erwarteten Häufigkeiten E_{ij} . Die Randhäufigkeiten $n_{i\bullet}$ entnehmen wir der dritten Spalte und der i -ten Zeile, die Randhäufigkeiten $n_{\bullet j}$ entnehmen wir der dritten Zeile und der j -ten Spalte sowie $N = 105$. Damit ergibt sich:

$$E_{11} = \frac{n_{1\bullet} \cdot n_{\bullet 1}}{N} = \frac{65 \cdot 60}{105} \approx 37,14, \quad E_{12} = \frac{n_{1\bullet} \cdot n_{\bullet 2}}{N} = \frac{65 \cdot 45}{105} \approx 27,86$$

$$E_{21} = \frac{n_{2\bullet} \cdot n_{\bullet 1}}{N} = \frac{40 \cdot 60}{105} \approx 22,86, \quad E_{22} = \frac{n_{2\bullet} \cdot n_{\bullet 2}}{N} = \frac{40 \cdot 45}{105} \approx 17,14$$

Damit erhalten wir die Prüfgröße mit den Werten n_{ij} der Kontingenztabelle durch

$$t = \sum_{j=1}^2 \sum_{i=1}^2 \frac{(n_{ij} - E_{ij})^2}{E_{ij}} = \frac{(n_{11} - E_{11})^2}{E_{11}} + \dots + \frac{(n_{22} - E_{22})^2}{E_{22}}$$

$$= \frac{(35 - 37,14)^2}{37,14} + \dots + \frac{(15 - 17,14)^2}{17,14}$$

$$\approx 0,75.$$

- Den kritischen Wert erhalten wir mit $m = \ell = 2$ durch

$$\chi_{(m-1) \cdot (\ell-1), 1-\alpha}^2 = \chi_{1, 0,95}^2 \approx 3,841.$$

- Somit verwerfen wir die Nullhypothese nicht, denn

$$t \approx 0,75 \not> 3,841 \approx \chi_{4, 0,95}^2.$$

zu A.3.26:

Wir möchten eine Aussage über die Verteilung der Anzahl der Gäste treffen und führen somit den χ^2 -Anpassungstest durch:

1. Zunächst stellen wir die Hypothesen auf:

H_0 : Das Merkmal „Anzahl der Gäste“ ist gleichverteilt.

H_1 : Das Merkmal „Anzahl der Gäste“ ist nicht gleichverteilt.

2. Das Signifikanzniveau beträgt $\alpha = 5\%$ und es liegen 6 Kategorien vor.

3. Die erwartete Anzahlen an Gästen unter der Bedingung „Gleichverteilung“ (n_j) lauten bei $N = 125 + 135 + 115 + 140 + 135 + 140 = 790$:

Dienstag: 131,67 Mittwoch: 131,67

Donnerstag: 131,67 Freitag: 131,67

Samstag: 131,67 Sonntag: 131,67

Wir berechnen die Prüfgröße und bezeichnen die tatsächlichen Besucheranzahlen mit N_j :

$$t = \sum_{k=1}^m \frac{(N_j - n_j)^2}{n_j} = \left(\frac{(125 - 131,67)^2}{131,67} + \dots + \frac{(140 - 131,67)^2}{131,67} \right) \approx 3,67$$

4. Den kritischen Wert erhalten wir mit

$$\chi_{m-1, 1-\alpha}^2 = \chi_{5, 0,95}^2 \approx 11,07.$$

5. Somit verwerfen wir die Nullhypothese nicht, denn

$$t \approx 3,67 \not\geq 11,07 \approx \chi_{5, 0,95}^2.$$

zu A.3.27:

Für die Prüfgröße des Wilcoxon-Vorzeichen-Rang-Tests bestimmen wir zunächst die Ränge anhand der Kontingenztafel:

Weingummis						
Supermarkt	mit Gelatine	ohne Gelatine	D_k	$ D_k $	R_k	
1	51	32	19	19	7	
2	38	61	-23	23	10	
3	29	51	-22	22	9	
4	55	35	20	20	8	
5	71	72	-1	1	1,5	
6	85	83	2	2	3	
7	112	127	-15	15	6	
8	14	13	1	1	1,5	
9	92	99	-7	7	4,5	
10	87	80	7	7	4,5	

Dabei haben wir D_k gebildet, indem wir die Differenz von Spalte 2 und Spalte 3 gebildet haben. Anschließend wenden wir den Betrag auf diese Differenzen an und erhalten $|D_k|$, mit denen wir die Ränge festlegen. Wir erkennen zwei Dopplungen in den absoluten Differenzen bei 1 und 7, sodass wir dort das arithmetische Mittel der Ränge bilden. Die Positionen mit $|D_k| = 1$ würden die Ränge 1 und 2 belegen, sodass wir $1,5 = \frac{1+2}{2}$ erhalten sowie $|D_k| = 7$ die Ränge 4 und 5, sodass wir $\frac{4+5}{2} = 4,5$ erhalten.

Für die Prüfgröße bestimmen wir mit $x_k =$ „Anzahl mit Gelatine“ und $y_k =$ „Anzahl ohne Gelatine“

$$w_+ = \sum_{k=1}^n \mathbb{1}_{\{x_k - y_k > 0\}} \cdot R_k = 1,5 + 3 + 4,5 + 7 + 8 = 24 \quad \text{sowie}$$

$$w_- = \sum_{k=1}^n \mathbb{1}_{\{x_k - y_k < 0\}} \cdot R_k = 1,5 + 4,5 + 6 + 9 + 10 = 31.$$

Insgesamt erhalten wir somit die Prüfgröße $w = \min(w_+, w_-) = w_+ = 24$.

zu A.3.28:

Nach unserem Schema gilt:

- Wir möchten nachweisen, dass ein Einfluss besteht, welcher nicht in eine spezielle Richtung vorgegeben ist, sodass wir die Hypothesen

$$H_0 : a = 0 \quad \text{und} \quad H_1 : a \neq 0$$

formulieren.

- Das Signifikanzniveau beträgt $\alpha = 5\%$.
- Die Teststatistik im Wilcoxon-Rangsummen-Test lautet

$$W = n \cdot m + \frac{m \cdot (m+1)}{2} - T_Y.$$

- Anhand der Anzahl der Werte in der Kontingenztafel, können wir ablesen, dass $n = 12$ und $m = 14$ gilt. Anschließend bestimmen wir die Ränge in der Gesamtstichprobe:

Gewicht und zuckerfreie Softgetränke				
Person	zuckerfreie Softgetränke	Rang	keine zuckerfreien Softgetränke	Rang
1	71	7	80	16
2	76	12	73	9
3	67	3	79	15
4	65	1	82	18
5	81	17	69	5
6	72	8	68	4
7	66	2	87	23
8	95	26	88	24
9	83	19	89	25
10	77	13	74	10
11	86	22	84	21
12	78	14	84	20
13	-	-	70	6
14	-	-	75	11

Nun bilden wir die Summe der Ränge in den einzelnen Kategorien:

$$t_x = 144 \text{ sowie}$$

$$t_y = 207$$

Somit erhalten wir die Prüfgröße

$$t = n \cdot m + \frac{m \cdot (m+1)}{2} - t_y = 12 \cdot 14 + \frac{14 \cdot 15}{2} - 207 = 66.$$

5. Den kritischen Wert erhalten wir mit

$$w_{n,m,1-\frac{\alpha}{2}} = w_{12,14,0,975} \stackrel{\text{aus Tab.}}{=} 45.$$

6. Damit lehnen wir die Nullhypothese ab, denn es gilt

$$t = 66 > 45 = w_{12,14,0,975}.$$