MATHEMATIK 2
WAHRSCHEINLICHKEITSRECHNUNG UND STATISTIK
VERSION 29. Juni 2021

LISIBACH ANDRÉ

1. Kombinatorische Grundlagen

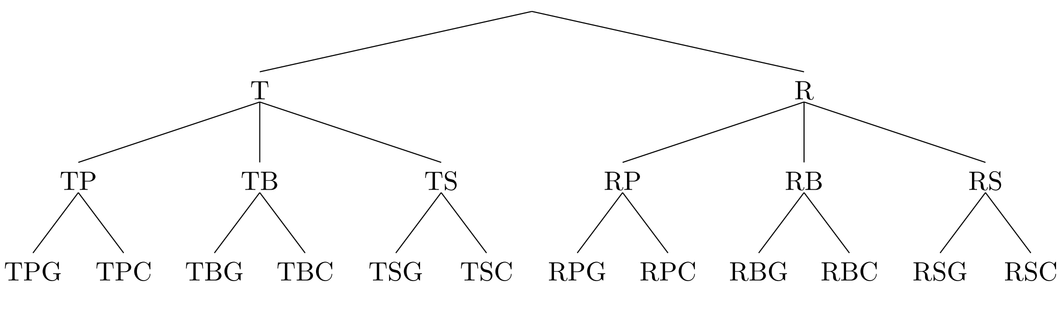
Die Kombinatorik ist ein Teilgebiet der Mathematik welches sich mit dem Anordnen beziehungsweise Auswählen endlicher Mengen von Objekten und dem Abzählen der verschiedenen Anordnungs- beziehungsweise Auswahlmöglichkeiten beschäftigt. Beispielsweise beantwortet die Kombinatorik die Frage nach der Anzahl Möglichkeiten beim Lotto aus 45 Zahlen 6 auszuwählen. Bei Fragen dieser Natur ist es prinzipiell möglich die Anzahl Möglichkeiten einzeln aufzuschreiben und abzuzählen. Die Kombinatorik liefert aber Resultate zur systematischen Herangehensweise an solche Probleme.

1.1. **Hauptsatz der Kombinatorik.** Wir illustrieren an einem Beispiel. Die folgende Abbildung zeigt eine Speisekarte in einem Restaurant:

Speisekarte

Vorspeisen	**Hauptgerichte**	**Dessert**
Tomatensuppe	Poulet mit Reis	Glace
Rindfleischsuppe	Bratwurst mit Pommes	Cremeschnitte
	Schnitzel mit Nudeln	

Ein Essen besteht aus einer Vorspeise, einem Haupgericht und einem Dessert. Die Frage nach der Anzahl Möglichkeiten sich ein Essen zusammenzustellen kann mit Hilfe eines *Baumdiagrammes* beantwortet werden:



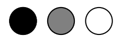
Es müssen drei Entscheidungen getroffen werden. Die Wahl einer Vorspeise, die Wahl eines Haupgerichts und die Wahl eines Dessert. Jede Wahl entspricht einer Verästelung im Baum. Wir haben einen dreistufigen Prozess, mit zwei, dann drei und dann wieder zwei Wahlmöglichkeiten. Insgesamt gibt es $2 \cdot 3 \cdot 2 = 12$ verschiedene Essen. Die am Baumdiagramm illustrierte Zählstrategie heisst in der Kombinatorik *Fundamentalprinzip des Zählens* (dies ist der Hauptsatz der Kombinatorik).

Allgemein halten wir fest: Ein Vorgang habe k Stufen. Die Anzahl Möglichkeiten in der i-ten Stufe sei n_i. Dann hat dieser k-stufige Vorgang insgesamt

$$a = n_1 \cdot n_2 \cdot n_3 \cdot \ldots \cdot n_k$$

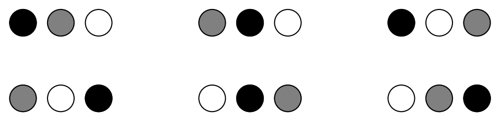
verschiedene Möglichkeiten (Ausgänge). Der dazugehörige Baum hat genau so viele Pfade.

1.2. **Permutationen.** Wir betrachten drei Kugeln in den Farben schwarz, grau und weiss:



und interessieren uns für die Anzahl unterschiedlicher Anordnungen (in einer Reihe) dieser drei Kugeln. Wir lösen das Problem auf zwei verschiedenen Wegen:

(i) Wir stellen die möglichen Anordnungen auf und zählen sie ab. Die möglichen Anordnungen sind:



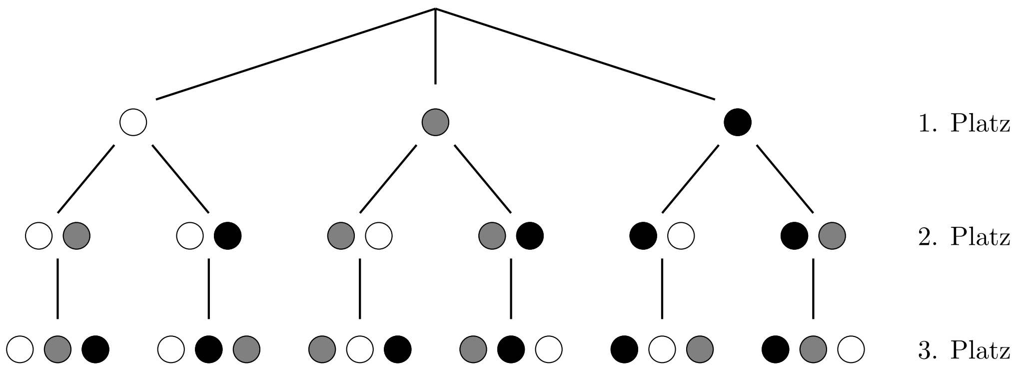
Somit ist die Anzahl möglicher Anordnungen der drei Kugeln gleich 6.

(ii) Wir machen die folgenden (kombinatorischen) Überlegungen:
  (a) Es sind drei Plätze in der Anordnung vorhanden die es zu besetzen gilt.
  (b) Wir besetzen den ersten Platz. Dazu können wir jede der drei Kugeln verwenden. Somit gibt es drei Möglichkeiten den ersten Platz zu besetzen, i.e. $n_1 = 3$.
  (c) Wir besetzen den zweiten Platz. Dafür sind noch zwei Kugeln übrig, somit gibt es zwei Möglichkeiten den zweiten Platz zu besetzen, i.e. $n_2 = 2$.
  (d) Wir besetzen den dritten Platz. Dazu ist noch eine Kugel übrig, somit gibt es nur eine Möglichkeit den dritten Platz zu besetzen $n_3 = 1$.
  (e) Es handelt sich somit um einen dreistufigen Prozess mit drei, dann 2 und dann einer Wahlmöglichkeit. Für die Besetzung aller drei Plätze gibt es somit

$$n_1 \cdot n_2 \cdot n_3 = 3 \cdot 2 \cdot 1 = 3! = 6$$

  Möglichkeiten.

Das dazugehörige Baumdiagramm ist:



Allgemein ist eine *Permutation ohne Wiederholung* eine Anordnung von n unterschiedlichen Objekten in einer bestimmten Reihenfolge. Die Anzahl möglicher Permutationen bezeichnen wir mit P_n. Es gilt

$$P_n = n!$$

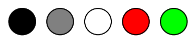
wobei

$$n! = \begin{cases} 1 & n = 0, \\ 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (n-1) \cdot n & n > 0, \end{cases}$$

gelesen "n Fakultät". Die Fakultät wächst sehr schnell, es gilt beispielsweise $10! > 10^6$. Für grosse Werte von n kann die Näherungsformel von Stirling verwendet werden:

$$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}.$$

1.3. **Variationen.** Wir betrachten fünf Kugeln in den Farben schwarz, grau, weiss, rot und grün:



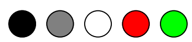
und interessieren uns für die Anzahl unterschiedlicher Anordnungen von drei aus diesen fünf Kugeln. Das Aufstellen und abzählen der möglichen Anordnungen lassen wir weg und gehen direkt über zu den folgenden kombinatorischen Überlegungen:

(i) Es sind drei Plätze in der Anordnung vorhanden die es zu besetzen gilt.
(ii) Wir besetzen den ersten Platz, dazu stehen fünf Kugeln zur Auswahl, i.e. es gibt fünf Möglichkeiten.
(iii) Zur Besetzung des zweiten Platzes sind vier Kugeln übrig, i.e. es gibt vier Möglichkeiten.
(iv) Zur Besetzung des dritten Platzes sind drei Kugeln übrig, i.e. es gibt drei Möglichkeiten.
(v) Für die Besetzung der drei Plätze gibt es somit $5 \cdot 4 \cdot 3 = 60$ Möglichkeiten.

Allgemein ist eine *Variation ohne Wiederholung* eine Anordnung von k Objekten aus einer Menge von n unterschiedlichen Objekten, wobei $n > k$ (der Fall $n = k$ ist durch die Permutationen abgedeckt). Die Anzahl möglicher Variationen von k aus n Objekten bezeichnen wir mit V_n^k. Es gilt

$$\begin{aligned} V_n^k &= n(n-1)(n-2)\cdots(n-k+2)(n-k+1) \\ &= \frac{n(n-1)(n-2)\cdots(n-k+2)(n-k+1)(n-k)(n-k-1)\cdots 1}{(n-k)(n-k-1)\cdots 1} \\ &= \frac{n!}{(n-k)!}. \end{aligned}$$

1.4. **Kombinationen.** Wir betrachten wieder die fünf Kugeln:

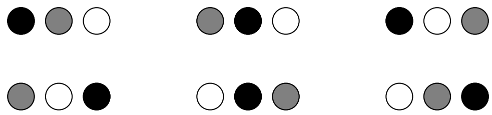


interessieren uns jetzt aber für die Anzahl Möglichkeiten aus diesen fünf Kugeln drei auszuwählen. Im Gegensatz zu einer Anordnung soll bei einer Auswahl die Reihenfolge keine Rolle spielen.

Für die folgende kombinatorische Betrachtung gehen wir zuerst einen Schritt zurück und nehmen an dass die Reihenfolge doch relevant sei. Es ergeben sich dann (siehe Variationen)

$$\frac{5!}{(5-3)!} = 60$$

mögliche Anordnungen. Unter diesen 60 Anordnungen sind beispielsweise die folgenden sechs zu finden:



i.e. die Anordnungen bei welchen nur die schwarze, weisse und graue Kugel vorkommt.

Im jetzigen Problem soll aber die Reihenfolge keine Rolle spielen. Es folgt dass diese sechs Anordnungen ein und der selben Auswahl enstprechen. Die Zahl sechs folgt aus der Tatsache dass sich die drei Kugeln (schwarz, weiss und grau) auf $3! = 6$ mögliche Arten anordnen lassen.

Der Zusammenhang zwischen Auswahl und Anordnungen gilt für jede beliebige Dreiergruppe von Kugeln aus den ursprünglichen fünf. I.e. jeweils $3! = 6$ mögliche Anordnungen entsprechen einer möglichen Auswahl von drei Kugeln. Somit müssen wir die total mögliche Anzahl Anordnungen durch $3! = 6$ dividieren um die Anzahl der möglichen Auswahlen zu bekommen.

Es ergeben sich

$$\frac{\frac{5!}{(5-3)!}}{3!} = \frac{60}{6} = 10$$

Möglichkeiten aus den fünf Kugeln drei auszuwählen.

Allgemein heisst eine Auswahl von k Objekten aus einer Menge von n unterschiedlichen Objekten eine *Kombination* $(n \geq k)$. Die Anzahl Möglichkeiten aus n Objekten k auszuwählen, i.e. die Anzahl Kombinationen, ist

$$C_n^k = \frac{n!}{(n-k)!k!} = \binom{n}{k}.$$

Dieser Ausdruck heisst *Binomialkoeffizient* und wird "n tief k" gelesen.

1.5. **Binomischer Lehrsatz.** Seien $x, y \in \mathbb{R}$, $n \in \mathbb{N}$. Es gilt

$$\begin{aligned}(x+y)^n &= \binom{n}{0} x^n + \binom{n}{1} x^{n-1}y + \binom{n}{2} x^{n-2}y^2 + \ldots + \binom{n}{n-1} xy^{n-1} + \binom{n}{n} y^n \\ &= \sum_{k=0}^{n} \binom{n}{k} x^{n-k}y^k.\end{aligned}$$

Wir leiten diesen Satz mit Hilfe von kombinatorischen Überlegungen für das Beispiel $n = 5$ her. Wir betrachten das Produkt

$$(x+y)^5 = (x+y)(x+y)(x+y)(x+y)(x+y).$$

Beim ausmultiplizieren erhalten wir einen Ausdruck der Form

$$(x+y)^5 = \ldots x^5 + \ldots x^4y + \ldots x^3y^2 + \ldots x^2y^3 + \ldots xy^4 + \ldots y^5,$$

wobei $\ldots$ für die (noch) unbekannten Koeffizienten steht. Wir betrachten den Term $\ldots x^3y^2$. Dieser kommt aus dem obigen Produkt durch die Multiplikation von drei Faktoren mit x und zwei Faktoren mit y zustande. Die Anzahl Möglichkeiten die zu diesen Potenzen führen ist die Anzahl Möglichkeiten aus den fünf Faktoren zwei auszuwählen (die zwei von welchen das y verwendet wird). Somit ist der Koeffizient vor x^3y^2 gegeben durch den Binomialkoeffizienten

$$\binom{5}{2}.$$

Allgemein ist der Koeffizient vor $x^{5-k}x^k$ gegeben durch

$$\binom{5}{k}$$

und wir erhalten

$$(x+y)^5 = \binom{5}{0} x^5 + \binom{5}{1} x^4y + \binom{5}{2} x^3y^2 + \binom{5}{3} x^2y^3 + \binom{5}{4} xy^4 + \binom{5}{5} y^5$$

$$= \sum_{k=0}^{5} \binom{5}{k} x^{5-k} y^k.$$

1.6. **Variationen mit Wiederholung.** Wir betrachten Kugeln in den Farben schwarz, grau und weiss und es soll von jeder Farbe ein beliebig grosser Vorrat an Kugeln vorhanden sein. Wir interessieren uns für die Anzahl unterschiedlicher Anordnungen von fünf dieser Kugeln. Die kombinatorische Überlegung ist in diesem Fall die folgende:

(i) Es sind fünf Plätze vorhanden die es zu besetzen gilt.

(ii) Wir besetzen den ersten Platz, dazu stehen drei verschiedene Farben von Kugeln zur Verfügung, i.e. wir haben drei Möglichkeiten den ersten Platz zu besetzen.

(iii) Zur Besetzung des zweiten Platzes können wir wie bei der Besetzung des ersten Platzes aus den drei Farben eine Kugel auswählen, i.e. wir haben auch bei der Besetzung des zweiten Platzes drei Möglichkeiten.

(iv) Die Anzahl Möglichkeiten ist für jeden Platz gleich, für die Besetzung der fünf Plätze gibt es somit insgesamt $3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 = 3^5 = 243$ Möglichkeiten.

Allgemein heisst eine Anordnung von k Objekten aus n verschiedenen Objekten, wobei die Objekte mehrmals verwendet werden dürfen, *Variation mit Wiederholung*. Die Anzahl solcher Möglichkeiten ist

$$n^k$$

Man sagt in diesem Fall auch dass k Objekte aus n *Kategorien* ausgewählt werden (im Beispiel der farbigen Kugeln sind die Kategorien die Farben der Kugeln).

2. Wahrscheinlichkeitsrechnung

2.1. **Enführung, was ist Wahrscheinlichkeit?** Wir illustrieren das Vorkommen von Wahrscheinlichkeit an Hand von drei Beispielen:

(i) Ein idealer Würfel wird geworfen. Die Wahrscheinlichkeit dass eine 'drei' geworfen wird ist $\frac{1}{6}$. Man kommt zu diesem Ergebnis, indem man das Verhältnis der Anzahl günstiger Ergebnisse zur Anzahl möglicher Ergebnisse berechnet. Durch die gleiche Überlegung kommt man auf die Wahrscheinlichkeit von $\frac{1}{2}$ dass eine gerade Zahl geworfen wird. In diesem Beispiel handelt es sich um die *klassische* Definition von Wahrscheinlichkeit.

(ii) Ein nicht-idealer Würfel wird geworfen. Gesucht ist die Wahrscheinlichkeit dass eine 'drei' geworfen wird. Eine Möglichkeit diese Wahrscheinlichkeit zu bestimmen ist den Würfel durch die Durchführung von Würfen 'auszumessen'. Die Anzahl Würfe sowie die Anzahl der aufgetretenen 'drei' wird festgehalten:

Anzahl Würfe	Anzahl 'drei'	Relative Häufigkeit
100	15	0.15
500	80	0.16
1000	151	0.151
5000	783	0.157
9000	1461	0.162
12000	1940	0.162
20000	3244	0.162

Wir sehen aus dieser Tabelle dass sich die relative Häufigkeit für eine 'drei' bei 0.162 stabilisiert. Somit ist davon auszugehn dass die Wahrscheinlichkeit für eine 'drei' durch diese relative Häufigkeit gegeben ist, i.e. die Wahrscheinlicheit ist 0.162. Die Stabilisierung der relativen Häufigkeit bei einem Wert ist rein hypothetisch und kann auch durch noch so viele Würfe nicht mit Sicherheit etabliert werden. In diesem Beispiel ist die Wahrscheinlichkeit somit durch einen *hypothetischen Grenzwert* gegeben.

(iii) Wir fragen nach der Wahrscheinlichkeit dass eine Person, welche im Jar 2009 in der Schweiz geboren ist männlich ist. Eine Antwort darauf liefern die folgenden Zahlen zu den Geburten im Jahr 2009 in der Schweiz:

Mädchen	37879
Knaben	40407
Total	78286

Die Wahrscheinlichkeit ist die relative Häufigkeit für eine Knabengeburt:

$$\text{Wahrscheinlichkeit} = \frac{40407}{78286} = 0.516.$$

Im Gegensatz zum vorhergehenden Beispiel ist es nicht möglich die Daten zu erweitern. Es handelt sich um *empirische Wahrscheinlichkeit.*

2.2. **Zufallsexperiment, Ergebnis, Stichprobenraum, Ereignis.** Die Grundlage der Wahrscheinlichkeitsrechung ist das *Zufallsexperiment.* Hierbei handelt es sich um einen Vorgang, welcher unter genau festgelegten Bedingungen durchgeführt wird und ein zufälliges *Ergebnis* hat, i.e. das Ergebnis kann nicht vorhergesagt werden.

Die Menge aller möglichen Ergebnisse eines Zufallsexperiments heisst *Stichprobenraum.* Bezeichnung: Ω. Beispiele von Zufallsexperimenten mit zugehörigem Stichprobenraum Ω:

(i) Wurf mit einem Würfel:

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

(ii) Geburt:

$$\Omega = \{\text{Knabe}, \text{Mädchen}\}.$$

(iii) Wurf einer Münze:

$$\Omega = \{\text{Kopf}, \text{Zahl}\}.$$

(iv) Wurf von zwei Münzen:

$$\Omega = \Big\{(\text{Kopf}, \text{Kopf}), (\text{Kopf}, \text{Zahl}), (\text{Zahl}, \text{Kopf}), (\text{Zahl}, \text{Zahl})\Big\}.$$

(v) Wurf mit einem roten und einem schwarzen Würfel:

$$\Omega = \{(i, j) : i, j = 1, 2, 3, 4, 5, 6\},$$

wobei (i, j) das Ergebnis bezeichnet, in welchem der rote Würfel i anzeigt und der schwarze j.

Reale Fragestellungen beziehen sich oft nicht auf ein einzelnes Ergebnis. Wir definieren: Eine Teilmenge E von Ω heisst ein *Ereignis*. Bezeichnung: $E \subset \Omega$.

Beispiele von Ereignissen beim Wurf mit einem Würfel:

(i) Die geworfene Zahl ist gerade:

$$E_{(\text{i})} = \{2, 4, 6\}.$$

(ii) Die geworfene Zahl ist ungerade:

$$E_{(\text{ii})} = \{1, 3, 5\}.$$

(iii) Die geworfene Zahl ist gleich zwei:

$$E_{(\text{iii})} = \{2\}.$$

(iv) Die geworfene Zahl ist kleiner oder gleich sieben:

$$E_{(\text{iv})} = \{1, 2, 3, 4, 5, 6\}.$$

(v) Die geworfene Zahl ist negativ:

$$E_{(\text{v})} = \{\}.$$

Wir notieren die folgenden speziellen Ereignisse:

(i) Das *sichere Ereignis*:

$$E = \Omega.$$

(ii) Das *unmögliche Ereignis*:

$$E = \{\}$$

(iii) Das *Gegenereignis*:

$$E^c, \qquad \text{wobei } E \subset \Omega.$$

(iv) Ein *Elementarereignis*:

$$E = \{\omega\}.$$

(v) *Unvereinbare Ereignisse*: Zwei Ereignisse E, F heissen *unvereinbar*, falls

$$E \cap F = \{\}.$$

2.3. **Wahrscheinlichkeit.** Sei Ω ein Stichprobenraum. Wir bezeichnen mit $\mathcal{P}(\Omega)$ die Menge aller Ereignisse, i.e. die Menge der Teilmengen von Ω. Eine Funktion

$$P : \mathcal{P}(\Omega) \to \mathbb{R}$$

heisst *Wahrscheinlichkeit*, falls sie die folgenden Axiome erfüllt:

(i) $0 \leq P(E) \leq 1, \quad$ für alle $E \subset \Omega$,
(ii) $P(\Omega) = 1$,
(iii) Wenn E, F unvereinbar sind ($E \cap F = \{\}$), dann gilt

$$P(E \cup F) = P(E) + P(F).$$

Ein Stichprobenraum Ω, zusammen mit einer Wahrscheinlichkeit P heisst *Wahrscheinlichkeitsraum*.

Bemerkung: Die obige Definition von Wahrscheinlichkeit charakterisiert P durch Eigenschaften, liefert aber keine Methode zur Berechnung von P. Die konkrete Berechnung von P wird durch Modellbildung, Versuche oder statistische Erhebungen durchgeführt (siehe Beispiele in der Einführung zu diesem Abschnitt).

Als Beispiel betrachten wir den Wurf einer Münze. Wir haben den Stichprobenraum $\Omega = \{\text{Kopf}, \text{Zahl}\}$ und somit ist die Menge aller Ereignisse:

$$\mathcal{P}(\Omega) = \Big\{\{\}, \{\text{Kopf}\}, \{\text{Zahl}\}, \Omega\Big\}.$$

Wir gehen nicht davon aus dass es eine faire Münze ist. Wir erwarten aber dass die Wahrscheinlichkeit für Kopf gleich Eins minus die Wahrscheinlichkeit für Zahl ist, i.e. wir erwarten

$$P(\{\text{Zahl}\}) = 1 - P(\{\text{Kopf}\}).$$

Dies lässt sich aus den obigen Axiomen herleiten: Wegen $\{\text{Kopf}\} \cap \{\text{Zahl}\} = \{\}$ folgt aus dem dritten Axiom

$$P(\{\text{Kopf}\} \cup \{\text{Zahl}\}) = P(\{\text{Kopf}\}) + P(\{\text{Zahl}\}).$$

Die linke Seite dieser Gleichung ist $P(\Omega) = 1$. Daraus folgt

$$P(\{\text{Zahl}\}) = 1 - P(\{\text{Kopf}\}).$$

Was wir auch erwarten ist

$$P(\{\}) = 0.$$

Herleitung aus den Axiomen: Durch das zweite Axiom haben wir $P(\Omega) = 1$. Da die Ereignisse Ω und $\{\}$ unvereinbar sind, i.e. $\Omega \cap \{\} = \{\}$, folgt aus dem dritten Axiom dass

$$P(\Omega \cup \{\}) = P(\Omega) + P(\{\}).$$

Die linke Seite dieser Gleichung ist $P(\Omega) = 1$, die rechte Seite dieser Gleichung ist $1 + P(\{\})$. Somit folgt

$$P(\{\}) = 0.$$

Wie in der obigen Bemerkung bereits festgehalten ist es jedoch nicht möglich aus den Axiomen die Wahrscheinlichkeit für Zahl oder Kopf herzuleiten. Diese Information muss durch Modellbildung, Versuche oder statistische Daten gegeben sein. Nehmen wir beispielsweise an dass Versuche ergeben haben dass die Münze zweimal so oft Kopf wie Zahl anzeigt. Dann folgt

$$P(\{\text{Kopf}\}) = \frac{2}{3}, \qquad P(\{\text{Zahl}\}) = \frac{1}{3}.$$

Seien $E, F \subset \Omega$. Die folgenden Aussagen lassen sich direkt aus den obigen Axiomen herleiten:

(i) $P(E^c) = 1 - P(E)$,
(ii) $P(\{\}) = 0$,
(iii) Für $E \subset F$ folgt $P(E) \leq P(F)$,
(iv) $P(E \backslash F) = P(E) - P(E \cap F)$,
(v) $P(E \cup F) = P(E) + P(F) - P(E \cap F)$.
(vi) Für $\Omega = \{E_1, E_2, \ldots, E_n\}$ und $E_i \cap E_j = \{\}$ für $i \neq j$ folgt

$$P\Big(\bigcup_{i=1}^{n} E_i\Big) = \sum_{i=1}^{n} P(E_i).$$

2.4. **Beispiel.** Wir betrachten den Wurf eines Würfels mit den folgenden Ereignissen und zugehörigen Wahrscheinlichkeiten:

A : 'Zahl ist Primzahl', $\quad P(A) = 0.36$,
B : 'Zahl ist eine Zweierpotenz', $\quad P(B) = 0.50$,
C : 'Zahl ist gleich zwei', $\quad P(C) = 0.15$

und fragen nach den Wahrscheinlichkeiten der Ereignisse

N : 'Zahl ist Primzahl und Zweierpotenz',
W : 'Zahl ist Primzahl oder Zweierpotenz'.

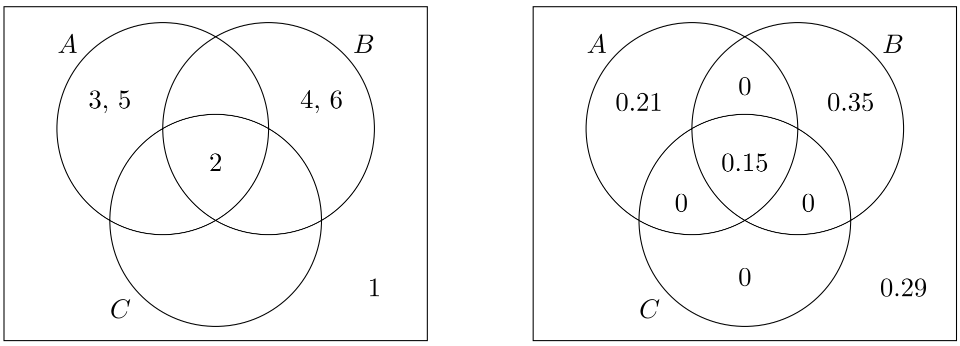
Wir haben $N = A \cap B = C$. Somit

$$P(N) = P(C) = 0.15.$$

Wir haben $W = A \cup B$. Somit

$$\begin{aligned} P(W) = P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0.36 + 0.50 - 0.15 = 0.71. \end{aligned}$$

Einfacher als die Rechnungen ist oft die Verwendung eines Venndiagramms:



Hier haben wir auf der linken Seite die Mengen mit ihren Elementen gezeichnet und auf der rechten Seite die Mengen mit den dazugehörigen Wahrscheinlichkeiten.

2.5. **Gleichwahrscheinlichkeit.** Der Stichprobenraum $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ zusammen mit der Wahrscheinlichkeit P heisst *Laplaceraum*, falls gilt

$$P(\{\omega_1\}) = P(\{\omega_2\}) = \ldots = P(\{\omega_n\}),$$

i.e. alle Ergebnisse sind gleichwahrscheinlich.

Beispiele:

(i) Münzwurf mit Münze.[1]

[1] Wird bei einer Münze oder einem Würfel nichts anderes vermerkt, so handelt es sich um eine ideale Münze, resp. um einen idealen Würfel. Ideal bedeutet bei einer Münze dass $P(\text{Kopf}) = P(\text{Zahl})$ und bei einem Würfel dass $P(\text{'eins'}) = P(\text{'zwei'}) = \ldots = P(\text{'sechs'})$

(ii) Würfeln mit einem Würfel.

(iii) Eine Person aus 100 Personen zufällig auswählen. Die zufällige Auswahl sorgt in diesem Beispiel dafür dass jede Person mit der selben Wahrscheinlichkeit ausgewählt wird.

Es ergeben sich die folgenden Konsequenzen:

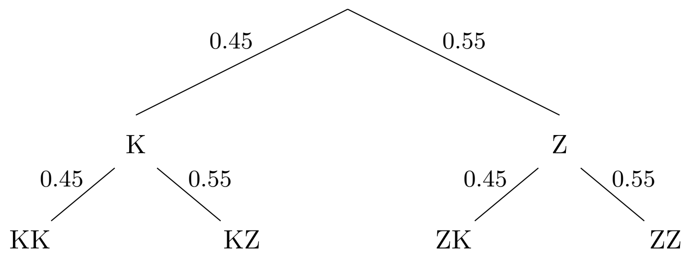
(i) $P(\{\omega_i\}) = \frac{1}{n}$, wobei[2] $n = |\Omega|$

(ii) Sei $E \subset \Omega$. Es folgt dass

$$P(E) = \frac{|E|}{|\Omega|} = \frac{|E|}{n} = \frac{\text{Anzahl günstige Fälle}}{\text{Anzahl mögliche Fälle}}.$$

Die Wahrscheinlichkeit P für einen Laplaceraum ist also durch die klassische Definition gegeben (siehe erstes Beispiel in der Einführung zu diesem Abschnitt).

2.6. **Mehrstufiges Zufallsexperiment.** Wir illustrieren an einem Beispiel: In einem Experiment wird eine Münze zweimal geworfen. Die Münze sei nicht fair aber man weiss dass sie bei jedem Wurf mit einer Wahrscheinlichkeit von 45% Kopf und mit einer Wahrscheinlichkeit von 55% Zahl zeigt. Wir verwenden ein Baumdiagramm und zeichnen auf den Ästen die Wahrscheinlichkeiten ein:



Das Baumdiagramm liefert den Stichprobenraum

$$\Omega = \{\text{KK}, \text{KZ}, \text{ZK}, \text{ZZ}\}.$$

Nun wird das Experiment 1000-mal wiederholt. Von 1000 ersten Würfen werden ca. $1000 \cdot 0.45 = 450$ Kopf anzeigen. Von diesen 450 ersten Würfen, welche Kopf angezeigt haben, werden im zweiten Wurf ca. $450 \cdot 0.45 = 202.5$ wieder Kopf anzeigen. Die Wahrscheinlichkeit für das Ergebnis KK ist somit gegeben durch $0.45 \cdot 0.45 = 0.45^2 = 0.2025$. Analog ist die Wahrscheinlichkeit für das Ergebnis KZ gegeben durch $0.45 \cdot 0.55$. Wir sehen dass die Wahrscheinlichkeiten auf den Ästen entlang eines Pfades zum Ergebnis multipliziert werden müssen um die Wahrscheinlichkeit für dieses Ergebnis zu erhalten.

Nun bestimmen wir die Wahrscheinlichkeit für das Ereignis E: 'Es wird in den beiden Würfen mindestens einmal Zahl geworfen'. Wir haben

$$E = \{\text{KZ}, \text{ZK}, \text{ZZ}\}.$$

und somit

$$\begin{aligned} P(E) &= P(\text{KZ}) + P(\text{ZK}) + P(\text{ZZ}) \\ &= 0.45 \cdot 0.55 + 0.55 \cdot 0.45 + 0.55 \cdot 0.55. \end{aligned}$$

[2]Bei einer endlichen Menge A bezeichnen wir mit $|A|$ die Anzahl Elemente dieser Menge.

3. Diskrete Zufallsvariablen

Bei vielen Zufallsexperimenten sind die Ergebnisse keine Zahlen, sondern beispielsweise eine Jasskarte, das angezeigte Bild einer Münze,... In diesem Abschnitt führen wir den Begriff einer Zufallsvariablen ein. Die Zufallsvariable wird dazu verwendet einem Ergebnis eines Zufallsexperiments eine Zahl zuzuordnen.

3.1. **Definition.** Eine *Zufallsvariable* auf einem Stichprobenraum Ω ist eine Abbildung:

$$X : \Omega \to \mathbb{R}$$
$$\omega \mapsto X(\omega).$$

Eine Zufallsvariable, welche nur endlich viele (oder abzählbar unendlich viele) Werte annehmen kann, heisst *diskrete Zufallsvariable*.

Als Anwendungsbeispiel betrachten wir ein Spiel in welchem eine Münze dreimal geworfen wird und der Gewinn sei gleich der Anzahl Köpfe die geworfen werden. Der Stichprobenraum ist

$$\Omega = \Big\{\text{KKK}, \text{KKZ}, \text{KZK}, \text{ZKK}, \text{KZZ}, \text{ZKZ}, \text{ZZK}, \text{ZZZ}\Big\}$$

und die Zufallsvariable X sei die Abbildung welche den Elementen in Ω den Gewinn zuordnet. Tabelle mit den Ergebnissen und den dazugehörigen Werten der Zufallsvariable X:

Ergebnis: ω	KKK	KKZ	KZK	ZKK	KZZ	ZKZ	ZZK	ZZZ
Gewinn: $X(\omega)$	3	2	2	2	1	1	1	0

Es ist wichtig zu sehen dass die Elemente in Ω nicht Zahlen sein müssen, die Zufallsvariable diesen Elementen jedoch eine (reelle) Zahl zuordnet. Somit werden die Ergebnisse eines Zufallsexperiments durch eine Zufallsvariable quantifiziert!

Bemerkungen:

(i) Für die Bezeichnung von Zufallsvariablen verwenden wir wann immer möglich Grossbuchstaben: $X, Y, Z, \ldots$

(ii) Es handelt sich bei einer Zufallsvariable weder um eine Variable (es ist eine Abbildung), noch ist sie zufällig (wie jede Abbildung oder Funktion ist die Zuordnung eindeutig). Nur das zu Grunde liegende Zufallsexperiment ist zufällig.

(iii) Die Interpretation des Wertes einer Zufallsvariable als Gewinn (oder falls negativ als Verlust) eines Glücksspiels ist immer möglich.

(iv) Aus einer Zufallsvariable lassen sich neue Zufallsvariablen generieren. Als Beispiel betrachten wir ein Glücksspiel in welchem ein Würfel geworfen wird und der Gewinn (Zufallsvariable X) sei gegeben durch

$$\text{Gewinn} = \begin{cases} 2 \cdot \text{Augenzahl}, & \text{falls Augenzahl gerade,} \\ -3 \cdot \text{Augenzahl}, & \text{falls Augenzahl ungerade.} \end{cases}$$

Wir haben

Ergebnis: ω	1	2	3	4	5	6
Gewinn: $X(\omega)$	-3	4	-9	8	-15	12

Nun definieren wir eine neue Zufallsvariable $Y(\omega) = 2X(\omega) + 1$. Wir haben

Ergebnis: ω	1	2	3	4	5	6
$Y(\omega)$	-5	9	-17	17	-29	25

(v) Meistens wird der Stichprobenraum Ω nicht näher spezifiziert. Als Beispiel betrachten wir die Aussage: "X bezeichnet die Anzahl Mädchen in Familien mit vier Kindern." Somit ist Ω die Menge der Familien mit vier Kindern, ω ist eine solche Familie ($\omega \in \Omega$) und $X(\omega) =$ Anzahl Mädchen in der Familie ω.

3.2. **Wahrscheinlichkeitsverteilung.** Wir haben gesehen wie sich durch eine Zufallsvariable X die Ergebnisse eines Zufallsexperiments durch die Zuordnung einer reellen Zahl zu jedem Ergebnis quantifizieren lassen. Auch haben wir gesehen dass eine Wahrscheinlichkeit P jedem Ereignis und somit auch jedem Ergebnis eine Zahl zwischen Null und Eins zuordnet (siehe vorhergehendes Kapitel). Die Wahrscheinlichkeitsverteilung setzt diese beiden Zahlen folgendermassen in Relation: Die *Wahrscheinlichkeitsverteilung* ist die Zuordnung:

$$\text{Wert von } X \mapsto \text{Wahrscheinlichkeit } P \text{ für diesen Wert.}$$

Im folgenden bezeichnen wir die Werte von X mit x_i und die dazugehörigen Werte von P, i.e. die Wahrscheinlichkeiten, mit p_i. Die Wahrscheinlichkeitesverteilung ist dann die Zuordnung:

$$x_i \mapsto p_i = P(X = x_i).$$

Die p_i besitzen die Eigenschaften:

(i) $0 \leq p_i \leq 1$,
(ii) $\sum_{i=1}^{n} p_i = 1$.

Die Eigenschaften ergeben sich direkt aus dem ersten und zweiten Axiom für P.

3.3. **Kumulative Verteilungsfunktion.** Die Funktion

$$F_X(x) = P(X \leq x)$$

heisst *kumulative Verteilungsfunktion* der Zufallsvariablen X. Diese Funktion ist somit die Wahrscheinlichkeit dass die Zufallsvariable X höchstens gleich x ist. $F_X(x)$ ist für alle $x \in \mathbb{R}$ definiert.

3.4. **Beispiel.** Als Beispiel betrachten wir das Glücksspiel in welchem eine Münze dreimal geworfen wird und der Gewinn (= Zufallsvariable X) sei die Anzahl Köpfe. Wir haben

ω	KKK	KKZ	KZK	ZKK	KZZ	ZKZ	ZZK	ZZZ
$P(\{\omega\})$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$X(\omega)$	3	2	2	2	1	1	1	0

Die Anzahl möglicher Werte der diskreten Zufallsvariablen X ist vier. Wir betrachten zum Beispiel den Werte $X(\omega) = 1$. Das dazugehörige Ereignis ist

$$E = \Big\{\text{KZZ}, \text{ZKZ}, \text{ZZK}\Big\}.$$

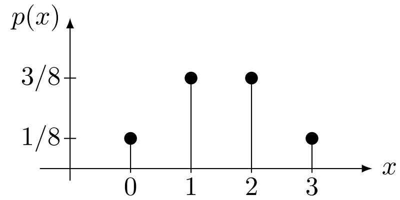
Somit haben wir

$$P(E) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}.$$

Die Wahrscheinlichkeitsverteilung ist

x_i	0	1	2	3
p_i	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

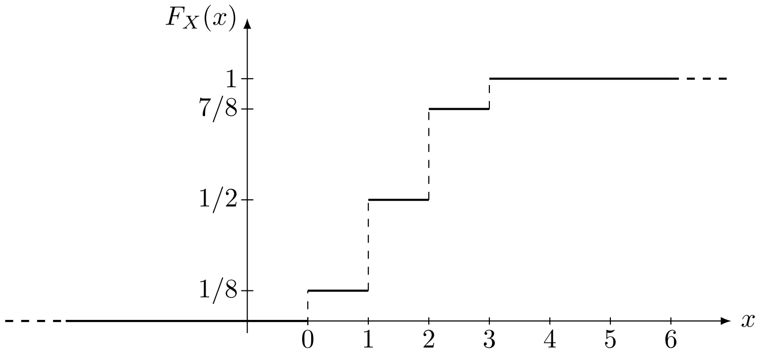
Die grafische Darstellung dieser Verteilung ist:



Nun bestimmen wir die kumulative Verteilungsfunktion. Wir haben $P(X < 0) = 0$. Somit folgt $F_X(x) = 0$ für $x < 0$. Weiter haben wir

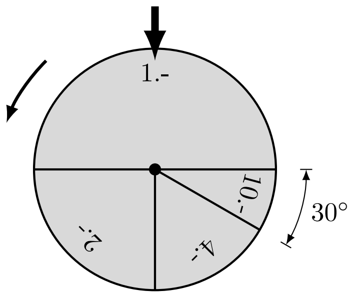
$$F_X(0) = P(X \leq 0) = P(X = 0) = \frac{1}{8},$$
$$F_X(1) = P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2},$$
$$F_X(2) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8},$$
$$F_X(3) = P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1.$$

Die grafische Darstellung der kumulativen Verteilungsfunktion ist:



Dies sieht aus wie eine Treppe und die Stufenhöhe bei x_i entspricht dem Wert der Verteilung $p(x_i)$.

3.5. **Erwartungswert.** Als einführendes Beispiel betrachten wir das Glücksrad:



Die eingetragenen Zahlen sind der Gewinn, wenn das Glücksrad im entsprechenden Bereich stehenbleibt. Wir fragen nach dem durchschnittlichen Gewinn bei 12 Spielen. Der Gewinn (Zufallsvariable X) besitzt die Verteilung:

x_i	1	2	4	10
p_i	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{12}$

In 12 Spielen erwarten wir somit:

$$6 \text{ mal } 1.- = 6.-$$
$$3 \text{ mal } 2.- = 6.-$$
$$2 \text{ mal } 4.- = 8.-$$
$$1 \text{ mal } 10.- = 10.-$$

Dies ergibt total 30.- und somit ergibt sich ein durchschnittlicher Gewinn pro Spiel von $\frac{30.-}{12} = 2.50$. Dieser mittlere hypothetische Gewinn heisst *Erwartungswert* der Zufallsvariablen X. Man erwartet bei nur 12 Spielen natürlich nicht gerade diese Gewinnverteilung, bei sehr vielen Spielen jedoch schon.

Allgemein erwarten wir bei einem Glücksspiel mit einer Verteilung des Gewinns

x_i	x_1	x_2	$\ldots$	x_n
p_i	p_1	p_2	$\ldots$	p_n

bei n Spielen

$$\begin{gathered} np_1 \text{ mal den Wert } x_1, \\ np_2 \text{ mal den Wert } x_2, \\ \vdots \\ np_n \text{ mal den Wert } x_n, \end{gathered}$$

und somit einen totalen Gewinn von

$$np_1x_1 + np_2x_2 + \ldots + np_nx_n = n\left(p_1x_1 + p_2x_2 + \ldots + p_nx_n\right).$$

Daraus ergibt sich der durchschnittliche Gewinn von

$$p_1x_1 + p_2x_2 + \ldots + p_nx_n.$$

Aufgrund der gemachten Beobachtungen definieren wir für eine diskrete Zufallsvariable X mit Werten x_i $(1 \le i \le n)$ und zugehörigen Wahrscheinlichkeiten $P(X = x_i) = p_i$ den *Erwartungswert* durch

$$E(X) = \sum_{i=1}^{n} p_i x_i.$$

3.6. **Varianz, Standardabweichung.** Als einführendes Beispiel betrachten wir die Notenergebnisse (Skala: $1, 2, \ldots, 6$) von vier Schülern X, Y, Z, W nach vier Prüfungen:

$$\begin{aligned} X &: 4, 4, 4, 4, \\ Y &: 1, 5, 5, 5, \\ Z &: 2, 4, 4, 6, \\ W &: 4, 5, 5, 6. \end{aligned}$$

Die Verteilungen sind

x_i	4
p_i	1

y_i	1	5
p_i	$\frac{1}{4}$	$\frac{3}{4}$

z_i	2	4	6
p_i	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

w_i	4	5	6
p_i	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Wir haben somit die Erwartungswerte

$$\begin{aligned} E(X) &= 1 \cdot 4 = 4, \\ E(Y) &= \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot 5 = 4, \\ E(Z) &= \frac{1}{4} \cdot 2 + \frac{1}{2} \cdot 4 + \frac{1}{4} \cdot 6 = 4, \\ E(W) &= \frac{1}{4} \cdot 4 + \frac{1}{2} \cdot 5 + \frac{1}{4} \cdot 6 = 5. \end{aligned}$$

Wir sehen dass

$$E(X) = E(Y) = E(Z),$$

jedoch ist es offensichtlich dass die Streuung der Noten des Schülers Y viel grösser ist als die Streuung der Noten des Schülers X. Um die Streuung zu quantifizieren definieren wir die *Varianz* einer Zufallsvariable als durchschnittliches Abweichungsquadrat zum Mittelwert. Für Y beträgt dieses Mass zum Beispiel:

$$\begin{aligned}\text{Varianz}(Y) &= \frac{1}{4}\Big((4-1)^2 + (4-5)^2 + (4-5)^2 + (4-5)^2\Big)\\ &= \frac{1}{4}(4-1)^2 + \frac{3}{4}(4-5)^2 = 3.\end{aligned}$$

Analog finden wir

$$\text{Varianz}(X) = 0, \qquad \text{Varianz}(Z) = 2, \qquad \text{Varianz}(W) = 2.$$

Allgemein definieren wir für eine diskrete Zufallsvariable X mit Werten x_i $(1 \leq i \leq n)$ und zugehörigen Wahrscheinlichkeiten $P(X = x_i) = p_i$ die *Varianz* durch

$$\begin{aligned}V(X) &= E\left((X - E(X))^2\right)\\ &= \sum_{i=1}^{n} p_i(x_i - E(X))^2.\end{aligned}$$

Die *Standardabweichung* ist definiert als

$$\sigma(X) = \sqrt{V(X)}.$$

Bemerkung: σ besitzt die selbe Einheit wie X.

Eine wichtige Formel um die Varianz zu berechnen ist

$$V(X) = E(X^2) - (E(X))^2.$$

Herleitung:

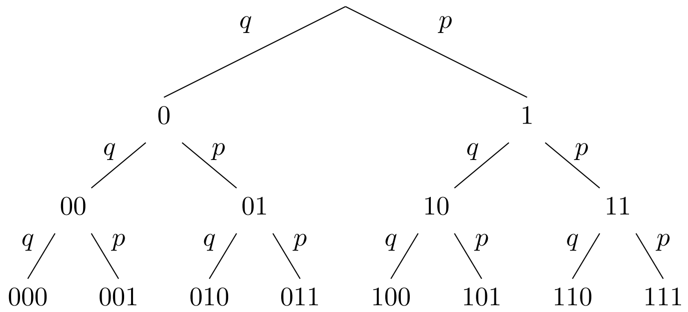
$$\begin{aligned}V(X) &= \sum_i p_i(x_i - E(X))^2\\ &= \sum_i p_i\left(x_i^2 - 2x_iE(X) + E(X)^2\right)\\ &= \sum_i p_i x_i^2 - 2E(X)\sum_i p_i x_i + E(X)^2 \sum_i p_i\\ &= E(X^2) - 2E(X)^2 + E(X)^2\\ &= E(X^2) - E(X)^2.\end{aligned}$$

3.7. **Binomialverteilung.** Wir betrachten ein Zufallsexperiment mit zwei möglichen Ergebnissen (auch *Bernoulliexperiment* genannt). Wir nennen die Ergebnisse *Erfolg* und *Misserfolg*. Sei p die Wahrscheinlichkeit für Erfolg. Weiter defnieren wir eine Zufallsvariable X als gleich Eins für Erfolg und gleich Null für Misserfolg:

ω	Erfolg	Misserfolg
x_i	1	0
p_i	p	$q = 1 - p$

Nun betrachten wir n Wiederholungen dieses Zufallsexperiments. Wir nehmen an dass die Wiederholungen unabhängig voneinander sind, i.e. das Ergebnis einer Durchführung hängt nicht von den vorhergehenden Ergebnissen ab. Sei die Zufallsvariable S gleich der Anzahl Erfolge bei diesen n Wiederholungen. S heisst *Binomialvariable.*

Wir betrachten die Situation für $n = 3$ an Hand eines Baumdiagramms:



Daraus lesen wir ab:

$$
\begin{aligned}
&P(S = 0) = P(000) = q^3, \\
&P(S = 1) = P(001) + P(010) + P(100) = q^2p + qpq + pq^2 = 3pq^2, \\
&P(S = 2) = P(011) + P(101) + P(110) = qp^2 + pqp + ppq = 3p^2q, \\
&P(S = 3) = P(111) = p^3.
\end{aligned}
$$

Wir sehen dass

$$P(S = k) = \binom{3}{k} p^k q^{3-k}.$$

Allgemein, bei n Wiederholungen, ergibt sich

$$P(S = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Die dazugehörige Wahrscheinlichkeitsverteilung heisst *Binomialverteilung* mit Parametern n, p. Für eine Zufallsvariable S, welche binomialverteilt ist mit Parametern n, p, schreiben wir $S \sim \mathcal{B}(n, p)$.

Die Binomialverteilung besitzt die folgenden Eigenschaften (ohne Beweis):

$$E(X) = np, \qquad V(X) = np(p-1).$$

Anwendungen:

(i) Wir betrachten eine Münze welche fünf mal geworfen wird. Sei X = Anzahl Kopf. X ist somit eine binomialverteilte Zufallsvariable mit $n = 5$, $p = 1/2$. Somit ist beispielsweise die Wahrscheinlichkeit dass dreimal Kopf geworfen wird gegeben durch

$$P(X = 3) = \binom{5}{3} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2 = \frac{10}{32} = 0.3125.$$

Falls die Münze manipuliert ist, so dass $P(\{\text{Kopf}\}) = 2/3$, dann folgt

$$P(X = 3) = \binom{5}{3} \left(\frac{2}{3}\right)^3 \left(\frac{1}{3}\right)^2 = \frac{80}{243} = 0.3292.$$

(ii) Wir betrachten die Herstellung von Schrauben. Diese werden so produziert, dass jede Schraube mit einer Wahrscheinlichkeit von $p = \frac{1}{100}$ Ausschuss ist. Die Schrauben werden in Paketen zu zehn Stück verkauft. Dabei wird ein Paket zurückgenommen, falls mehr als eine Schraube im Paket Ausschuss ist. Gesucht ist der Anteil an Paketen welche zurückgenommen werden.

Wir definieren die Zufallsvariable X als die Anzahl Schrauben eines Paketes die Ausschuss sind. Dann ist X binomialverteilt mit $n = 10$, $p = \frac{1}{100}$. Die

Wahrscheinlichkeit dass ein Paket zurückgenommen wird ist gleich der Wahrscheinlichkeit dass mehr als eine Schraube im Paket Ausschuss ist, i.e. gleich $P(X > 1)$. Wir haben

$$\begin{aligned} P(X > 1) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - \binom{10}{0}\left(\frac{1}{100}\right)^0\left(\frac{99}{100}\right)^{10} - \binom{10}{1}\left(\frac{1}{100}\right)^1\left(\frac{1}{99}\right)^9 \\ &= 0.00427, \end{aligned}$$

i.e. der Anteil an Paketen welche zurückgenommen werden ist 0.43%.

4. Stetige Zufallsvariablen

4.1. **Einführendes Beispiel.** Wir betrachten als Zufallsvariable X die Körperlänge von Neugeborenenen. Die folgenden Werte (in cm) wurden gemessen:

$$48.7, \quad 51.4, \quad 53.2, \quad 49, \quad 49.8, \quad 45.7, \quad 48.5, \dots$$

Die Variable X kann (in einem gewissen Bereich) beliebig viele Werte annehmen (überabzählbar viele). Man spricht von einer *stetigen Zufallsvariable.* Weitere Beispiele sind die Temperatur von Schmiedeteilen, die Geschwindigkeit von Fahrzeugen,...

Würde man nun jedem auftretenden Wert x_i eine Wahrscheinlichkeit p_i zuordnen, so würde die Summe aller Wahrscheinlichkeiten über alle möglichen Ergebnisse unendlich werden. Aus diesem Grund erfolgt die Beschreibung von stetigen Zufallsvariablen durch eine *Dichtefunktion.*

4.2. **Dichtefunktion.** Eine *Dichtefunktion* ist eine Funktion $f : \mathbb{R} \to \mathbb{R}$ mit den Eigenschaften:

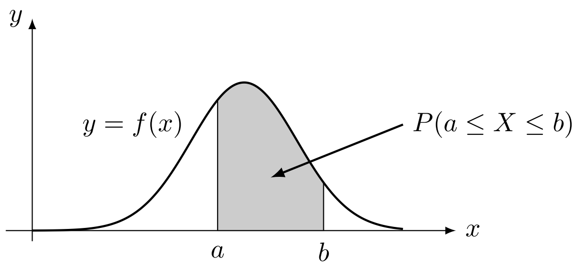
$$\text{(i)} \qquad f(x) \geq 0, \quad \text{für alle } x \in \mathbb{R},$$

$$\text{(ii)} \qquad \int_{-\infty}^{\infty} f(x)dx = 1.$$

Eine Zufallsvariable X heisst *stetig*, falls sie eine Dichte $f(x)$ besitzt, welche die Wahrscheinlichkeit für die Werte von X folgendermassen beschreibt: $f(x)dx$ ist gleich der Wahrscheinlichkeit dass die Zufallsvariable X einen Wert zwischen x und $x + dx$ annimmt. Somit ist $\int_a^b f(x)dx$ die Wahrscheinlichkeit dass die Zufallsvariable einen Wert zwischen a und b annimmt, i.e.

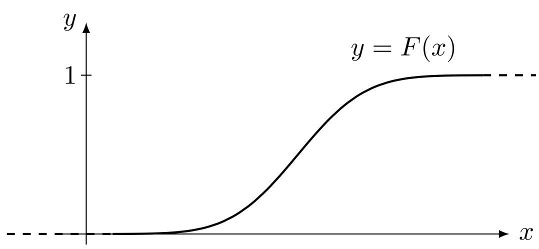
$$P(a \leq X \leq b) = \int_a^b f(x)dx.$$

Die Wahrscheinlichkeit entspricht somit dem Flächeninhalt unter dem Graphen der Dichtefunktion im Bereich $[a, b]$:



4.3. **Kumulative Verteilungsfunktion.** Die zu $f(x)$ dazugehörige *kumulative Verteilungsfunktion* $F(x)$ ist die Wahrscheinlichkeit dafür, dass die Zufallsvariable X einen Wert annimmt, der kleiner oder gleich einer vorgegebenen reellen Zahl x ist. Somit gilt

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du.$$



Wir haben die Eigenschaften:

(i)

$$P(a \le X \le b) = F(b) - F(a).$$

(ii) $F(x)$ ist monoton wachsend und es gilt $F'(x) = f(x)$ (wo die Ableitung existiert).

(iii) Es gilt

$$\lim_{x\to-\infty} F(x) = 0, \qquad \lim_{x\to\infty} F(x) = 1.$$

4.4. **Erwartungswert und Varianz.** Wir definieren den Erwartungswert und die Varianz als

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx,$$
$$V(X) = \int_{-\infty}^{\infty} \big(x - E(X)\big)^2 f(x) dx.$$

Diese Formeln sind analog zu den Formeln für diskrete Zufallsvariablen, wenn man die Summen mit Integralen ersetzt. Oft wird für den Erwartungswert das Symbol μ verwendet, i.e. $E(X) = \mu(X)$.

Die *Standardabweichung* ist definiert als

$$\sigma(X) = \sqrt{V(X)}.$$

Bemerkung: σ besitzt die selbe Einheit wie X.

4.5. **Gleichmässige Verteilung.** Eine Zufallsvariable X heisst *gleichmässig verteilt* im Intervall $[a, b]$, falls die Dichtefunktion $f(x)$ gegeben ist durch

$$f(x) = \begin{cases} C & a \le x \le b, \\ 0 & \text{sonst.} \end{cases}$$

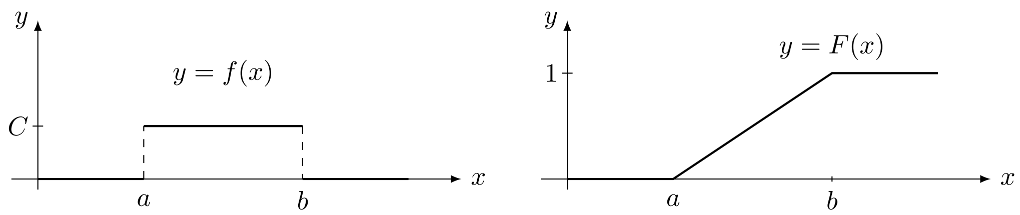
Die Konstante C ist gegeben durch

$$1 \overset{!}{=} \int_{-\infty}^{\infty} f(x) dx = \int_a^b C dx = C(b-a), \qquad \text{i.e.} \qquad C = \frac{1}{b-a}.$$

Die kumulative Verteilungsfunktion ist

$$F(x) = \int_{-\infty}^{x} f(u) du = \begin{cases} 0 & x \le a, \\ \frac{x-a}{b-a} & a < x < b, \\ 1 & b < x. \end{cases}$$

Grafische Darstellung:



Erwartungswert und Varianz sind (durch direkte Rechnung)

$$E(X) = \frac{a+b}{2}, \qquad V(X) = \frac{(b-a)^2}{12}.$$

4.6. **Normalverteilung.** Die wichtigste Wahrscheinlichkeitsverteilung ist die Normalverteilung.[3]

[3]Man nennt sie auch Gauss'sche Verteilung, Gauss'sche Fehlerkurve oder Gauss'sche Glockenkurve.

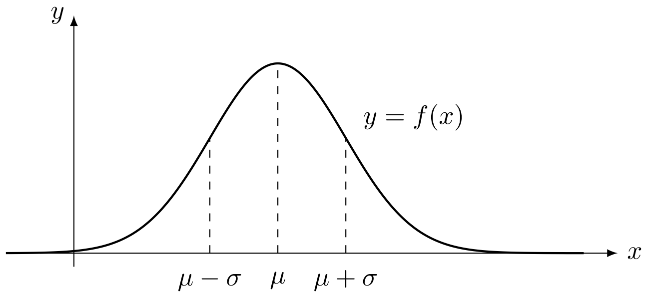
4.6.1. *Dichtefunktion und kumulative Verteilungsfunktion.* Die Dichtefunktion der *Normalverteilung* ist

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Hier sind μ, σ Konstanten mit $\sigma > 0$. Für eine normalverteilte Zufallsvariable X schreiben wir $X \sim \mathcal{N}(\mu, \sigma)$. Es gilt (ohne Beweis)

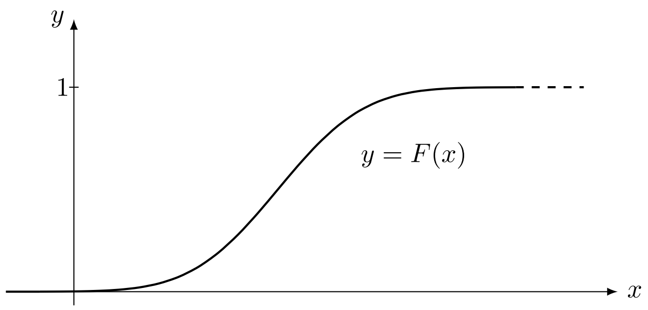
$$\int_{-\infty}^{\infty} f(x)dx = 1, \qquad E(X) = \mu, \qquad V(X) = \sigma^2,$$

i.e. die Konstanten μ und σ sind der Erwartungswert und die Standardabweichung der Normalverteilung. Der Graph $y = f(x)$ besitzt bei $x = \mu$ ein globales Maximum und bei $x = \mu \pm \sigma$ Wendepunkte. Der Graph ist symmetrisch bezüglich der Achse $x = \mu$.



Die dazugehörige kumulative Verteilungsfunktion ist

$$F(x) = \int_{-\infty}^{x} f(u)du = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du,$$



Die Wahrscheinlichkeit dass die normalverteilte Zufallsvariable X einen Wert im Intervall $[a, b]$ annimmt ist gegeben durch

$$P(a \leq X \leq b) = F(b) - F(a)$$

und entspricht der Fläche unter der Glockenkurve $y = f(x)$ im Intervall $[a, b]$.

4.6.2. *Standardnormalverteilung.* Die Normalverteilung mit Parametern $\mu = 0$, $\sigma = 1$ heisst *Standardnormalverteilung.* Ihre Dichtefunktion ist somit gegeben durch

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Die dazugehörige kumulative Verteilungsfunktion ist

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du.$$

Die Funktion im Integral besitzt keine Stammfunktion. Die Werte von $\Phi(x)$ sind im Anhang tabelliert.

4.6.3. *Berechnung von Wahrscheinlichkeiten.* Sei X eine normalverteilte Zufallsvariablen mit Parametern μ, σ. Die Wahrscheinlichkeit dass die Zufallsvariable X einen Wert im Intervall $[a,b]$ annimmt ist gegeben durch

$$P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

Die im Integral auftretende Funktion besitzt keine Stammfunktion. Wir führen die Berechnung auf die Standardnormalverteilung zurück:

$$\frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{t^2}{2}} dt = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

wobei wir die Substitution $t = \frac{x-\mu}{\sigma}$ verwendet haben. Wir finden also

$$P(a \leq X \leq b) = F(b) - F(a) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

Man nennt

$$X^* = \frac{X-\mu}{\sigma}$$

die *standardisierte Zufallsvariable* zu X. Es gilt $E(X^*) = 0$, $V(X^*) = 1$, i.e. X^* ist standardnormalverteilt.

Wir illustrieren die Berechnung von Wahrscheinlichkeiten an einem Beispiel: Sei $X \sim \mathcal{N}(0.8, 2)$, i.e. die Zufallsvariable X sei normalverteilt mit Erwartungswert $\mu = 0.8$ und Standardabweichung $\sigma = 2$. Gesucht sei $P(X \leq 2.44)$, i.e. gesucht sei die Wahrscheinlichkeit dass die Zufallsvariable X einen Wert kleiner oder gleich 2.44 annimmt. Wir führen die Berechnung auf die Standardnormalverteilung zurück:

$$P(X \leq 2.44) = F(2.44) = \Phi\left(\frac{2.44-\mu}{\sigma}\right) = \Phi\left(\frac{2.44-0.8}{2}\right) = \Phi(0.82) = 0.7939,$$

wobei wir den Wert $\Phi(0.82)$ in der Tabelle im Anhang, in der 0.8-Zeile und der 0.02-Spalte abgelesen haben.

4.6.4. *Beispiel.* Bei der Produktion von Achsen ist festgestellt worden, dass die Werte für den Durchmesser (Zufallsvariable X) normalverteilt sind mit einem Mittelwert $\mu = 30\,\text{mm}$ und einer Streuung $\sigma = 2\,\text{mm}$. Die vorgegebene Toleranzgrenze für die Achsen sei $30 \pm 5\,\text{mm}$. Wir bestimmen den Ausschussanteil. Dieser Anteil entspricht der Wahrscheinlichkeit dass die Achsen entweder zu klein ($\leq 25\,\text{mm}$) oder zu gross ($\geq 35\,\text{mm}$) sind:

$$\begin{aligned} P(X \leq 25) + P(X \geq 35) &= 1 - P(25 \leq X \leq 35) \\ &= 1 - (F(35) - F(25)) \\ &= 1 - \left(\Phi\left(\frac{35-\mu}{\sigma}\right) - \Phi\left(\frac{25-\mu}{\sigma}\right)\right) \\ &= 1 - \left(\Phi\left(\frac{35-30}{2}\right) - \Phi\left(\frac{25-30}{2}\right)\right) \\ &= 1 - (\Phi(2.5) - \Phi(-2.5)). \end{aligned}$$

Der Wert $\Phi(-2.5)$ kann aus der Tabelle nicht direkt abgelesen werden. Aus Symmetriegründen gilt jedoch $\Phi(-x) = 1 - \Phi(x)$, i.e. $\Phi(-2.5) = 1 - \Phi(2.5)$. Somit ist der Ausschussanteil

$$\begin{aligned} P(X \leq 25) + P(X \geq 35) &= 1 - (\Phi(2.5) - (1 - \Phi(2.5))) \\ &= 2(1 - \Phi(2.5)) \\ &= 2(1 - 0.9938) = 0.0124, \end{aligned}$$

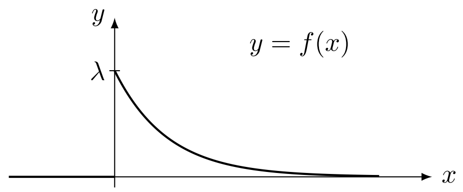
i.e. der Ausschussanteil beträgt 1.24 %.

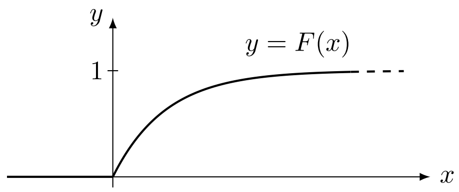
4.7. **Exponentialverteilung.** Eine Zufallsvariable X heisst *exponentialverteilt* mit Parameter $\lambda > 0$, falls die Dichtefunktion gegeben ist durch

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

Wir schreiben $X \sim \mathrm{Exp}(\lambda)$. Die kumulative Verteilungsfunktion ist

$$F(x) = \int_{-\infty}^{x} f(u)du = \begin{cases} 1 - e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$





Es gilt

$$E(X) = \frac{1}{\lambda}, \qquad V(X) = \frac{1}{\lambda^2}.$$

5. Zentraler Grenzwertsatz

Wir haben im vorhergehenden Kapitel die Normalverteilung behandelt und sie als die wichtigste Wahrscheinlichkeitesverteilung bezeichnet. Der zentrale Grenzwertsatz liefert die Begründung dafür. Grob gesprochen ist seine Aussage: Lässt sich eine Zufallsvariable als Summe von vielen Zufallsvariablen darstellen, welche unabhängig voneinander sind, so ist die Zufallsvariable näherungsweise normalverteilt.

In der Praxis hängen Grössen die untersucht werden oft von vielen Einflussfaktoren ab, welche meist nicht alle bekannt sind und gemessen werden können. In der Summer ergibt sich aber approximativ eine Normalverteilung.

Eine wichtige Anwendung sind Messfehler. Diese kommen normalerweise aufgrund von vielen voneinander unabhängigen Einflüssen zustande und sind somit in guter Näherung normalverteilt.

5.1. Aussage des Satzes.

Theorem 5.1. *Seien X_i, $i = 1, 2, \dots, n$ unabhängige, identisch verteilte Zufallsvariablen mit*

$$E(X_i) = \mu, \qquad V(X_i) = \sigma^2.$$

und sei

$$S_n = \sum_{i=1}^{n} X_i.$$

Dann haben wir

$$E(S_n) = nE(X_i) = n\mu, \qquad V(S_n) = nV(X_i) = n\sigma^2$$

i.e.

$$\mu(S_n) = n\mu, \qquad \sigma(S_n) = \sqrt{n}\sigma$$

und für die standardisierte Zufallsvariable

$$S_n^* = \frac{S_n - \mu(S_n)}{\sigma(S_n)} = \frac{S_n - n\mu}{\sqrt{n}\sigma}$$

gilt

$$\lim_{n\to\infty} P(a \le S_n^* \le b) = \Phi(b) - \Phi(a),$$

i.e. die Verteilung von S_n^* konvergiert gegen die Standardnormalverteilung.

5.2. Bemerkungen.

(i) Wir geben keinen Beweis dieses Satzes.

(ii) Der Satz wird für Approximationen verwendet, i.e. man verwendet

$$S_n^* \sim \mathcal{N}(0, 1)$$

für grosse n. In der Praxis wird die Approximation meist für $n \ge 50$ verwendet.

(iii) Wie zu Beginn des Kapitels erwähnt liefert der Satz den Grund für die Wichtigkeit der Normalverteilung.

(iv) Aus

$$\sigma(S_n) = \sqrt{n}\sigma$$

folgt

$$\sigma\left(\frac{S_n}{n}\right) = \frac{\sigma}{\sqrt{n}}.$$

I.e. die Grösse $\frac{S_n}{n}$ (dies ist der Mittelwert der X_i) streut um einen Faktor $\frac{1}{\sqrt{n}}$ weniger stark als X_i.

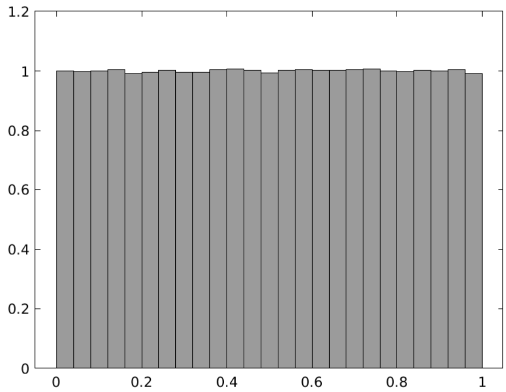
5.3. **Illustration mit Matlab.** An Stelle eines Beweises illustrieren wir den zentralen Grenzwertsatz numerisch (mit Matlab) mit Hilfe von zufällig generierten Zahlen.

Der Befehl `rand(m,n)` liefert eine $m \times n$ - Matrix mit Zufallszahlen zwischen 0 und 1. In einem ersten Schritt untersuchen wir qualitativ die Verteilung dieser Zufallszahlen indem wir einen Zeilenvektor mit 10^6 dieser Zufallszahlen erzeugen und ein Histogramm zeichnen:

Matlab

```
X=rand(1,10^6);
histogram(X,20,'Normalization','pdf')
```

Das zweite Argument im Histogrammbefehl sind die Anzahl Unterteilungen auf der horizontalen Achse. Das dritte und vierte Argument im Histogrammbefehl bewirken dass auf der vertikalen Achse relative Häufigkeiten pro Balkenbreite dargestellt werden. Somit kann der gezeigte Verlauf als Graph einer Dichtefunktion interpretiert werden, i.e. die Fläche zwischen zwei Werten a, b auf der horizontalen Achse entspricht der relativen Häufgkeit von Werten im Intervall $[a, b]$ und entpricht somit der empirischen Wahrscheinlichkeit für einen Wert im Intervall $[a, b]$.



Wir sehen dass der Befehl `rand` in guter Näherung gleichmässig verteilte Zufallszahlen zwischen 0 und 1 liefert. Dazugehöriger Mittelwert und Standardabweichung sind

$$\mu = \frac{1}{2}, \qquad \sigma = \frac{1}{\sqrt{12}}$$

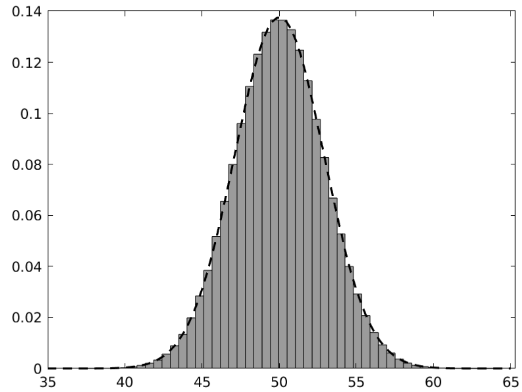
(siehe 4.5 auf Seite 19).

Nun generieren wir diese Verteilung für $n = 100$ Zufallsvariablen $X_1, X_2, \ldots, X_{100}$. Wir verwenden dazu `rand(100,10^6)`. Die i-te Zeile dieser Matrix enthält jeweils 10^6 Werte der Zufallsvariablen X_i. Mit dem Befehl `sum(y,1)` bilden wir die Summe jeder Spalte von y und erhalten somit 10^6 Werte von $S_{100} = \sum_{i=1}^{100} X_i$. Schliesslich erzeugen wir das Histogramm für S_{100}.

Matlab

```
X=rand(100,10^6);
S=sum(X,1);
histogram(S,100,'Normalization','pdf')
```

Wir erhalten:



Nach dem zentralen Grenzwertsatz erwarten wir eine Normalverteilung mit Erwartungswert und Standardabweichung

$$\mu(S_n) = n\mu = 100 \cdot \frac{1}{2} = 50, \qquad \sigma(S_n) = \sqrt{n}\sigma = 10 \cdot \frac{1}{\sqrt{12}} \approx 2.9,$$

i.e. die Wendepunkte liegen bei

$$\mu(S_n) - \sigma(S_n) \approx 47.1, \qquad \mu(S_n) + \sigma(S_n) \approx 52.9.$$

Die Zeilen

Matlab

```
x=linspace(35,65);
muSn = 50;
sigmaSn = 2.9;
f = exp(-(x-muSn).^2./(2*sigmaSn^2))./(sigmaSn*sqrt(2*pi));
plot(x,f,'--','LineWidth',1.5)
```

erzeugen den entsprechenden Graphen der Dichtefunktion (siehe gestrichelte Linie in der obigen Grafik).

5.4. **Anwendung.** Ein Arbeiter fertigt eine Serie von 50 Bauteilen. Für ein Bauteil benötigt er durchschnittlich 20 min bei einer Standardabweichung von 4 min. Im folgenden bestimmen wir die Wahrscheinlichkeit dass er in den ersten 450 min mindestens 25 Bauteile fertigt.

Sei X_i die Zeit welche der Arbeiter für das i-te Bauteil benötigt. X_i sind Zufallsvariablen mit $\mu = 20$, $\sigma = 4$. Die Zeit welche der Arbeiter für die ersten 25 Bauteile benötigt ist gegeben durch die Zufallsvariable

$$S_{25} = \sum_{i=1}^{25} X_i.$$

Gesucht ist

$$P(S_{25} \leq 450).$$

Erwartungswert und Standardabweichung von S_{25} sind

$$\mu(S_{25}) = 25 \cdot \mu = 25 \cdot 20 = 500, \qquad \sigma(S_{25}) = \sqrt{25} \cdot \sigma = 5 \cdot 4 = 20.$$

Die standardisierte Zufallsvariable zu S_{25} ist

$$S_{25}^* = \frac{S_{25} - \mu(S_{25})}{\sigma(S_{25})} = \frac{S_{25} - 500}{20}.$$

Nach dem zentralen Grenzwertsatz ist S_{25}^* approximativ standardnormalverteilt. Somit haben wir

$$\begin{aligned} P(S_{25} \leq 450) &= P\left(S_{25}^* \leq \frac{450 - 500}{20}\right) \\ &= P(S_{25}^* \leq -2.5) \\ &\approx \Phi(-2.5) \\ &= 1 - \Phi(2.5) \\ &= 1 - 0.9938 = 0.0062, \end{aligned}$$

wobei wir den Zahlenwert 0.9938 aus der Tabelle im Anhang rausgelesen haben. Das heisst dass der Arbeiter mit einer Wahrscheinlichkeit von (approximativ) 0.62% in den ersten 450 min 25 Bauteile fertigt.

Bemerkungen:

(i) Die Zeit welche der Arbeiter für das Fertigen eines Bauteils benötigt muss nicht normalverteilt sein. Diese Verteilung ist unbekannt und wird auch nicht benötigt. Was benötigt wird ist nur der Erwartungswert μ und die Standardabweichung σ dieser Verteilung.

(ii) Die Streuung der Zeit, welche der Arbeiter für ein Bauteil benötigt ist mit $\sigma = 4$ min angegeben. Beim Fertigen von 25 Bauteilen ist die Streuung der durchschnittlichen Zeit für ein Bauteil gegeben durch

$$\sigma\left(\frac{S_{25}}{n}\right) = \frac{\sigma}{\sqrt{n}} = \frac{4}{5} = 0.8,$$

i.e. die Streuung der durchschnittlichen Zeit für ein Bauteil beträgt 0.8 min.

6. Schliessende Statistik

Oft sind die Parameter einer Wahrscheinlichkeitsverteilung unbekannt und sollen mittels Messungen gefunden werden. Im folgenden ist der Erwartungswert μ dieser unbekannte Parameter und weitere Parameter sollen jeweils bekannt sein. Wir betrachten im folgenden drei Methoden um den unbekannten Parameter μ zu finden:

(i) Parameterschätzung: Es wird ein approximativer Wert für den Parameter gefunden.
(ii) Vertrauensintervall: Es wird ein Intervall gefunden, in welchem der Parameter mit einer vordefinierten Wahrscheinlichkeit liegt.
(iii) Statistischer Test: Es wird eine Annahme für den Parameter getroffen und untersucht ob diese Annahme haltbar ist oder verworfen werden sollte.

Zugrunde liegt jeweils die Auswertung einer Stichprobe aus einer Grundgesamtheit, welche gemessen wurde.

6.1. **Statistische Grundbegriffe.** Eine zentrale Aufgabe der Statistik besteht darin, Informationen über eine bestimmte Menge von Objekten zu gewinnen. Beispiele:

(i) Es soll die durchschnittliche Körperlänge aller Neugeborenen im Jahre 2020 ermittelt werden.
(ii) Es soll ermittelt werden, welcher Anteil der Knallfrösche, welche eine Maschine herstellt, nicht funktionieren.
(iii) Es soll ermittelt werden, zu wieviel Prozent eine unfaire Münze Kopf anzeigt.

Die Menge der Objekte heisst *Grundgesamtheit*. Oft ist es nicht möglich die ganze Grundgesamtheit zu untersuchen. In Beispiel (i) weil der Aufwand zu gross ist und in Beispiel (ii) weil dazu die Knallfrösche getestet und somit vernichtet werden müssen. Aus diesem Grund wird eine *Stichprobe* der Grösse n verwendet (gleich Anzahl Neugeborene welche gemessen werden in Beispiel (i) oder Anzahl Knallfrösche welche getestet werden in Beispiel (ii)). Eine Stichprobe ist somit eine Teilmenge der Grundgesamtheit.

Wir nehmen an, dass die Grundgesamtheit einer Verteilung folgt, welche durch eine Zufallsvariable X beschrieben wird. Wir verwenden eine Stichprobe der Grösse n. <u>Vor dem Ausmessen</u> der Stichprobe sind die potentiellen Werte von X die Zufallsvariablen $X_1, X_2, \ldots, X_n$. Man sagt auch dies seien nicht ausgewertete Kopien der Zufallsvariablen X. Der Mittelwert der Stichprobe ist die Zufallsvariable

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

<u>Nach dem Ausmessen</u> der Stichprobe haben wir die gemessenen Zahlenwerte

$$x_1, x_2, \ldots, x_n.$$

Man sagt diesen Werten auch *Realisierungen* der Zufallsvariablen $X_1, X_2, \ldots, X_n$. Der Wert

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

ist eine Realisierung der Zufallsvariablen $\overline{X}$.

6.2. **Parameterschätzung.** Das naheliegenste Vorgehen um Informationen über den unbekannten Parameter μ zu bekommen, ist den Mittelwert der Stichprobe dafür zu verwenden:

$$\hat{\mu} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Hier ist n die Grösse der Stichprobe (gleich Anzahl gemessener Werte) und x_i sind die gemessenen Werte der Stichprobe. Man nennt dieses Vorgehen *Punktschätzung* und den erhaltenen Wert $\hat{\mu}$ eine Schätzung für den unbekannten Parameter μ. Im allgemeinen stimmt die Schätzung nicht mit dem wahren Mittelwert überein, i.e. $\hat{\mu} \neq \mu$, aber die Schätzung wird verlässlicher, je grösser n gewählt wird.

Soll zum Beispiel die durchschnittliche Körperlänge von Neugeborenen im Jahr 2020 ermittelt werden, so kann eine Stichprobe von 1000 Neugeborenen ausgemessen werden. Der erhaltene Mittelwert dieser Stichrpobe ist eine Schätzung für den wahren Mittelwert aller Neugeborenen.

6.3. **Vertrauensintervall.** Es soll ein Intervall angegeben werden, in welchem der unbekannte Parameter μ mit einer vorgegebenen Wahrscheinlichkeit liegt. I.e. es sollen Grenzen θ_1, θ_2 gefunden werden, so dass

$$P(\theta_1 \leq \mu \leq \theta_2) = \gamma$$

für vorgegebenes γ (z.B. $\gamma = 0.95$ oder $\gamma = 0.99$). Dazu sollen Werte $x_1, x_2, \ldots, x_n$ einer Stichprobe verwendet werden. Die Werte $x_1, x_2, \ldots, x_n$ der Stichprobe entsprechen Realisierungen der Zufallsvariablen $X_1, X_2, \ldots, X_n$, welche Kopien der Zufallsvariablen X der Grundgesamtheit sind. Für grosses n gilt wegen dem zentralen Grenzwertsatz, dass

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

approximativ normalverteilt ist mit

$$\mu(\overline{X}) = \mu, \qquad \sigma(\overline{X}) = \frac{\sigma}{\sqrt{n}}.$$

Somit gilt

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1),$$

i.e. die zum Mittelwert der Stichprobe dazugehörige standardisierte Zufallsvariable ist (approximativ) standardnormalverteilt.

Die Forderung $P(\theta_1 \leq \mu \leq \theta_2) = \gamma$ ist äquivalent zu

$$P(-\theta_1 \geq -\mu \geq -\theta_2) = \gamma$$

und dies ist wiederum äquivalent zu

$$P\left(\frac{\overline{X} - \theta_1}{\sigma/\sqrt{n}} \geq \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \geq \frac{\overline{X} - \theta_2}{\sigma/\sqrt{n}}\right) = \gamma.$$

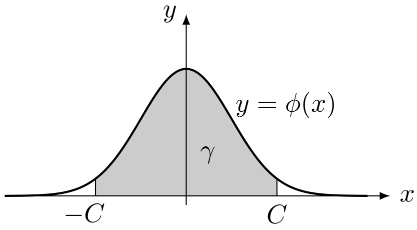
Da nun $(\overline{X} - \mu)/(\sigma/\sqrt{n})$ standardnormalverteilt ist (siehe oben), wird die Forderung zu

$$\Phi\left(\frac{\overline{X} - \theta_1}{\sigma/\sqrt{n}}\right) - \Phi\left(\frac{\overline{X} - \theta_2}{\sigma/\sqrt{n}}\right) = \gamma.$$

Wir nehmen an dass das Vertrauensintervall symmetrisch um den Wert μ liegt. Somit liegt das Vertrauensintervall der standardisierten Zufallsvariable symmetrisch um 0. I.e. wir haben

$$\frac{\overline{X} - \theta_1}{\sigma/\sqrt{n}} = C, \qquad \frac{\overline{X} - \theta_2}{\sigma/\sqrt{n}} = -C,$$

für eine Konstante C. Grafisch:



Die Konstante C folgt aus γ unter Verwendung der kumulativen Verteilungsfunktion der Standardnormalverteilung $\Phi(x)$. Die gesuchten Grenzen ergeben sich zu

$$\theta_1 = \overline{X} - \frac{\sigma}{\sqrt{n}} C, \qquad \theta_2 = \overline{X} + \frac{\sigma}{\sqrt{n}} C.$$

Wie angenommen soll σ bekannt sein, n entspricht der Stichprobengrösse und für $\overline{X}$ wird der Mittelwert der gemessenen Stichprobe $\overline{x}$ verwendet. Somit gilt für den unbekannten Paramter μ

$$\overline{x} - \frac{\sigma}{\sqrt{n}} C \leq \mu \leq \overline{x} + \frac{\sigma}{\sqrt{n}} C.$$

I.e. der unbekannte Parameter μ liegt mit einer Wahrscheinlichkeit von γ im Intervall

$$\left[\overline{x} - \frac{\sigma}{\sqrt{n}} C, \, \overline{x} + \frac{\sigma}{\sqrt{n}} C \right].$$

Dieses Intervall wird das γ-*Vertrauensintervall* genannt.

Beispiele für Werte von C: Für $\gamma = 0.95$ erhalten wir aus der Tabelle für $\Phi(x)$:

$$C = \Phi^{-1}(0.975) = 1.96$$

und für $\gamma = 0.99$ erhalten wir

$$C = \Phi^{-1}(0.995) = 2.575.$$

Anwendungsbeispiel: Eine Maschine stellt Bonbons her. Uns interessiert das mittlere Gewicht der Bonbons. Die Genauigkeit der Maschine ist mit einer Standardabweichung von 3 g angegeben. Von 100 Bonbons wird das Gewicht gemessen und der Durchschnitt davon beträgt 5 g. Es soll das 95%-Vertrauensintervall für das mittlere Gewicht der Bonbons bestimmt werden.

Wir haben $n = 100$, $\overline{x} = 5$, $\sigma = 3$, $C = 1.96$. Somit folgen die Intervallgrenzen:

$$\begin{aligned} \overline{x} - \frac{\sigma}{\sqrt{n}} C &= 5 - \frac{3}{10} 1.96 = 4.412, \\ \overline{x} + \frac{\sigma}{\sqrt{n}} C &= 5 + \frac{3}{10} 1.96 = 5.588. \end{aligned}$$

Das 95%-Vertrauensintervall ist

$$[4.412, 5.588],$$

i.e. mit einer Wahrscheinlichkeit von 95% liegt der Mittelwert der hergestellten Bonbons zwischen 4.412 g und 5.588 g.

6.4. **Statistischer Test.** Wir illustrieren das Vorgehen an einem Beispiel: Von einer Pulverabfüllanlage werden Behälter mit je 4 mg Pulver gefüllt. Die Standardabweichung sei 1 mg. Ein Inspektor inspiziert eine Stichprobe von neun Behältern (welche er zufällig auswählt) und findet in diesen Behältern durchschnittlich 4.6 mg Pulver. Es soll mit einem statistischen Test untersucht werden, ob die Anlage richtig funktioniert. Wir verwenden die Notation:

$$\mu_0 = 4, \qquad \sigma = 1, \qquad n = 9, \qquad \overline{x} = 4.6.$$

Wir stellen die folgenden Hypothesen auf:

(i) *Nullhypothese*: Der unbekannte Mittelwert μ ist gleich dem Referenzwert μ_0, i.e.

$$H_0 : \quad \mu = \mu_0.$$

(ii) *Alternativhypothese*: Der unbekannte Mittelwert ist nicht gleich dem dem Referenzwert, i.e.

$$H_1 : \quad \mu \neq \mu_0.$$

Unter der Hypothese H_0 ($\mu = \mu_0$) gilt nun, wie bereits im vorhergehenden Abschnitt zum Vertrauensintervall benutzt, dass die Zufallsvariable

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

approximativ standardnormalverteilt ist. Die Stichprobe gibt uns die folgende Realisierung dieser Variable:

$$z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{4.6 - 4}{1/\sqrt{9}} = 1.8.$$

Die Wahrscheinlichkeit, unter der Hypothese H_0, dass die Variable Z einen mindestens so extremen Wert annimmt ist (dies ist der p-Wert des Tests)

$$p = P(Z \leq -1.8) + P(Z \geq 1.8) = 2(1 - \Phi(1.8)) = 2(1 - 0.9641) = 0.0718.$$

I.e. die Wahrscheinlichkeit für mindestens so einen extremen Wert ist 7.2%. Wir wählen nun eine Wahrscheinlichkeit α, ab welcher die Nullhypothese verworfen wird:

$$\alpha = 0.05.$$

Dies nennt man das *Signifikanzniveau*. Wir haben $p > \alpha$ und somit verwerfen wir die Nullhypothese nicht. I.e. unter der Annahme dass der ubekannte Mittelwert μ gleich dem angegebenen Referenzwert μ_0 ist, ist es genügend wahrscheinlich, dass in einer Stichprobe ein Mittelwert entsteht, welcher gemessen wurde. Somit gibt es keinen Hinweis darauf dass die Maschine nicht korrekt läuft.

Dieser Test war *zweiseitig*, i.e. es wurde getestet ob die Maschine zuviel oder zuwenig abfüllt. Wenn uns nur interessiert, ob die Maschine zuviel abfüllt dann formulieren wir die Hypothesen folgendermassen:

$$H_0 : \quad \mu = \mu_0,$$
$$H_1 : \quad \mu > \mu_0.$$

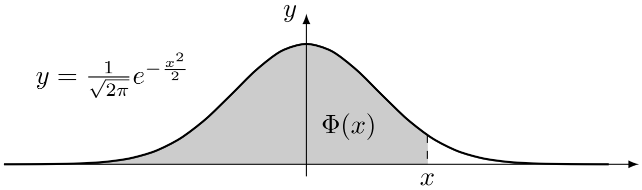
Man spricht von einem *einseitigen* Test. Der p-Wert ist dann

$$p = P(Z \geq 1.8) = 1 - \Phi(1.8) = 1 - 0.9641 = 0.0359.$$

I.e. die Wahrscheinlichkeit für mindestens so einen extremen Wert ist 3.6%. Wird das gleiche Signifikanzniveau $\alpha = 0.05$ verwendet, so haben wir nun $p < \alpha$, i.e. der gemessene Wert in der Stichprobe ist zu extrem und die Hypothese H_0 wird verworfen. I.e. es gibt einen Grund anzunehmen dass die Maschine nicht korrekt läuft (Achtung: Auch wenn $p > \alpha$ gilt, ist die Hypothese H_1 nicht bewiesen).

7. Anhang

7.1. Kumulative Verteilungsfunktion $\Phi(x)$ der Standardnormalverteilung.



$\Phi(x)$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7793	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8906	0.8925	0.8943	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9986	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

8. Aufgaben zu Wahrscheinlichkeitsrechnung

Aufgabe 1

Wir betrachten ein Rennen mit fünf Läufern. Wieviele verschiedene Ranglisten sind möglich?

Aufgabe 2

Man vereinfache die folgenden Ausdrücke ($n, i, r \in \mathbb{N}$):

$$\frac{10!}{7!}, \qquad \frac{(n+i)!}{n!}, \qquad \frac{n!}{(n-2)!}, \qquad \frac{(n-r+1)!}{(n-r-1)!}.$$

Aufgabe 3

Auf wieviele Arten können sich sieben Personen

(i) auf eine Stuhlreihe mit sieben Stühlen,
(ii) an einen runden Tisch mit sieben Stühlen setzen?

Aufgabe 4

(i) Wieviele vierstellige Zahlen lassen sich mit den Ziffern 1, 3, 5, 7 bilden?
(ii) Man bestimme die Summe dieser Zahlen.

Aufgabe 5

Man bestimme die Anzahl Möglichkeiten acht Türme auf einem Schachbrett aufzustellen, so dass sie sich nicht bedrohen, falls

(i) die Türme unterscheidbar sind (z.B. durch unterschiedliche Farben) und
(ii) falls die Türme ununterscheidbar sind.

Aufgabe 6

Anne besitzt 10 Bücher welche sie in ihr Bücherregal stellt. Davon sind vier Mathematikbücher, drei sind Chemiebücher, zwei sind Geschichtsbücher und eines ist ein Roman. Man bestimme die Anzahl Möglichkeiten die Bücher ins Regal zu stellen, wenn die Bücher eines Themas jeweils nebeneinander stehen sollen.

Aufgabe 7

Wir betrachten ein Rennen mit fünf Läufern. Wieviele verschiedene Klassierungen der ersten drei Läufer sind möglich?

Aufgabe 8

Wie viele ganze Zahlen n, mit $1000 \leq n \leq 9999$, besitzen vier unterschiedliche Ziffern?

Aufgabe 9

Wir betrachten die Menge

$$M = \{A, B, C, D, E\}$$

Wieviele Teilmengen mit drei Elementen gibt es?

Aufgabe 10

Auf wieviele Arten kann ein Gremium mit drei Männern und zwei Frauen aus sieben Männern und fünf Frauen gebildet werden?

Aufgabe 11

Man zeige

$$\binom{n}{k} = \binom{n}{n-k}.$$

Aufgabe 12

Ein Schüler muss in einer Prüfung acht von zehn Fragen richtig beantworten.

(i) Wieviele Möglichkeiten hat er?
(ii) Wieviele Möglichkeiten hat er wenn er die ersten drei Fragen richtig beantworten muss?
(iii) Wieviele Möglichkeiten hat er wenn er mindestens vier der ersten fünf Fragen richtig beantworten muss?

Aufgabe 13

Man bestimme den Term in der Entwicklung von $(2x^2 - \frac{1}{2}y^3)^8$, welcher x zur achten Potenz enthält.

Aufgabe 14

Auf wieviele Arten kann man 36 Spielkarten auf vier Spieler A, B, C, D verteilen, wenn jeder Spieler neun Karten erhält?

Aufgabe 15

Man bestimme den Term in der Entwichlung von $(xy - y^2 + 2z)^6$, welcher x zur dritten und y zur fünften Potenz enthält.

Aufgabe 16

Man zeige

(i)

$$\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \ldots + (-1)^n \binom{n}{n} = 0.$$

(ii)

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3} + \ldots + \binom{n}{n} = 2^n.$$

Hinweis: Man verwende den binomischen Lehrsatz und wähle die x, y entsprechend den obigen Gleichungen.

Aufgabe 17

Man bestimme die Anzahl der Teilmengen einer Menge mit n Elementen.

Aufgabe 18

Wir betrachten zehn Punkte A, B, C, $\ldots$, J in der Ebene, wobei keine drei Punkte auf einer Geraden liegen.

(i) Wieviele Geraden sind durch die Punkte bestimmt?
(ii) Wieviele dieser Geraden verlaufen durch A oder B?
(iii) Wieviele Dreiecke sind durch diese Punkte bestimmt?
(iv) Wieviele dieser Dreiecke enthalten die Seite AB?

Aufgabe 19

Man bestimme die Anzahl der binären Zeichenketten der Länge n.

Aufgabe 20

Wir betrachten das Tippspiel Sport Toto, bei welchem für 13 Fussballspiele ein Tipp über deren Ausgang abgegeben werden soll. Ein Tipp beinhaltet die Angabe der Ausgänge jedes Spiels. Die möglichen Ausgänge eines Spiels sind: Gewinn Heimmannschaft, Gewinn Gastmannschaft und unentschieden. Man bestimme die Anzahl möglicher Tippabgaben.

Aufgabe 21

Man bestimme die Anzahl symmetrischer $n \times n$-Matrizen mit Einträgen in $\{1, 2, \ldots, 9\}$.

Aufgabe 22

Wir betrachten die Gleichung

$$x_1 + x_2 + x_3 + x_4 = 15,$$

wobei $x_i \in \{1, 2, 3, \ldots\}$. Man bestimme die Anzahl Lösungen. Hinweis: Die Lösung $x_1 = 3$, $x_2 = 4$, $x_3 = 3$, $x_4 = 5$ lässt sich in der folgenden Form notieren:

$$1 + 1 + 1 | 1 + 1 + 1 + 1 | 1 + 1 + 1 | 1 + 1 + 1 + 1 + 1.$$

Aufgabe 23

Ein Würfel wird geworfen. Man formuliere für die folgenden Ereignisse das Gegenereignis und bestimme seine Wahrscheinlichkeit:

(i) E: 'Die Zahl ist gerade',
(ii) F: 'Die Zahl ist grösser als zwei',
(iii) G: 'Es wird mindestens eine zwei gewürfelt'.

Aufgabe 24

Anne geht in die Ferien und nimmt zwei Bücher mit. Mit einer Wahrscheinlichkeit von 0.5 mag sie das erste Buch. Mit einer Wahrscheinlichkeit von 0.4 mag sie das zweite Buch. Mit einer Wahrscheinlichkeit von 0.3 mag sie beide Bücher. Man bestimme die Wahrscheinlichkeit dass Anne keines der Bücher mag.

Aufgabe 25

Ein Kunde besucht einen Kleiderladen. Mit einer Wahrscheinlichkeit von 0.22 kauft er einen Anzug, mit Wahrscheinlichkeit von 0.30 kauft er ein Shirt, mit Wahrscheinlichkeit von 0.28 kauft er eine Kravatte. Der Kunde kauft mit einer Wahrscheinlichkeit von 0.11 einen Anzug und ein Shirt, mit einer Wahrscheinlichkeit von 0.14 einen Anzug und eine Kravatte und mit einer Warscheinlichkeit von 0.10 kauft er ein Shirt und eine Kravatte. Der Kunde kauft mit einer Wahrcheinlichkeit von 0.06 alle drei Dinge.

(i) Mit welcher Wahrscheinlichkeit kauft der Kunde nichts?
(ii) Mit welcher Wahrscheinlichkeit kauft der Kunde genau ein Ding?

Aufgabe 26

In einem Restaurant gibt es acht Tische mit je vier Plätzen. Ein Platz im Restaurant ist von einem Gast besetzt. Ein zweiter Gast betritt das Restaurant und entscheidet sich willkürlich für einen der freien Plätze. Man bestimme die Wahrscheinlichkeit dass die beiden Gäste am selben Tisch sitzen.

Aufgabe 27

Wir betrachten die Summe der Augenzahlen beim Wurf mit zwei Würfeln.

(i) Man bestimme Ω.
(ii) Handelt es sich um einen Laplaceraum?
(iii) Was muss betrachtet werden, damit es sich um einen Laplaceraum handelt?

Aufgabe 28

Ein Kommitee von fünf Personen wird aus einer Gruppe von sechs Männern und neun Frauen ausgewählt. Man bestimme die Wahrscheinlichkeit für ein Kommitee bestehend aus drei Männern und zwei Frauen.

Aufgabe 29

Wurf mit zwei Würfeln.

(i) Mit welcher Wahrscheinlichkeit werden zwei Sechsen gewürfelt?
(ii) Mit welcher Wahrscheinlichkeit werden nicht zwei Sechsen gewürfelt?
(iii) Mit welcher Wahrscheinlichkeit wird mindestens eine Sechs gewürfelt?
(iv) Mit welcher Wahrscheinlichkeit wird keine Sechs gewürfelt?

Aufgabe 30

Man wirft 6 Würfel. Was ist die Wahrscheinlichkeit, dass die 6 Würfel unterschiedliche Augenzahlen anzeigen?

Aufgabe 31

In einer Schublade liegen drei Paar Schuhe zerstreut herum. Man zieht zufällig drei Schuhe. Wie gross ist die Wahrscheinlichkeit, dass sich darunter ein Paar befindet?

Aufgabe 32

Drei Freunde gehen Kaffee trinken. Sie werfen jeder eine Münze und die Person, deren Resultat alleine vorkommt bezahlt. Sind die Resultate alle gleich, so wird nochmals geworfen.

(i) Was ist die Wahrscheinlichkeit dass nach einem Wurf der drei Münzen feststeht welche Person zahlen muss?
(ii) Was ist die Wahrscheinlichkeit bei einem Wurf der drei Münzen, dass nochmals geworfen werden muss?
(iii) Was ist die Wahrscheinlichkeit dass genau dreimal geworfen werden muss um die zahlende Person festzustellen?
(iv) Was ist die Wahrscheinlichkeit dass die drei Münzen mehr als viermal geworfen werden müssen um die zahlende Person festzustellen?

Aufgabe 33

Es wird mit zwei Würfeln geworfen. Man finde die Wahrscheinlichkeitsverteilung und kumulative Verteilungsfunktion für die Summe der Augenzahlen und stelle die Resultate grafisch dar.

Aufgabe 34

Man wirft zwei Würfel. Sei X die Zufallsvariable, die als Wert das Maximum der beiden gewürfelten Zahlen annimmt. Man bestimme die Verteilung von X.

Aufgabe 35

Wir betrachten das folgende Spiel: Eine Münze wird dreimal nacheinander geworfen. Die Wahrscheinlichkeit für Kopf ist bei jedem Wurf 3/5. Der Gewinn/Verlust ist gegeben durch zweimal die Anzahl Kopf minus einmal die Anzahl Zahl.

(i) Man beschreibe die möglichen Verläufe das Spiels mit einem Baumdiagramm und trage die Wahrscheinlichkeiten auf den Ästen ein.
(ii) Man bestimme den Stichprobenraum Ω und die Wahrscheinlichkeiten für alle Elementarereignisse.
(iii) Man bestimme die Zufallsvariable X welche den Gewinn/Verlust beschreibt.
(iv) Für X bestimme man die Wahrscheinlichkeits- und Verteilungsfunktion und stelle diese Funktionen grafisch dar.
(v) Man bestimme den Erwartungswert und die Standardabweichung von X.
(vi) In einer neuen Version des Spiels muss vor jedem Spiel ein Einsatz gezahlt werden. Wie tief muss dieser Einsatz sein, damit unter Berücksichtigung dieses Einsatzes, trotzdem ein Gewinn erwartet werden kann?

Aufgabe 36

Eine Münze wird sechs mal geworfen. Man bestimme die Wahrscheinlichkeit,

(i) dass genau zweimal Kopf auftritt,
(ii) dass mindestens viermal Kopf auftritt,
(iii) dass nie Kopf auftritt.

Aufgabe 37

Bei einem Flugzeug fällt jedes Triebwerk mit der gleichen Wahrscheinlichkeit p während des Fluges aus und zwar unabhängig vom Zustand der anderen Triebwerke. Wir nehmen an dass ein Flugzeug abstürzt wenn mehr als die Hälfte der Triebwerke ausfallen.

Für welche Werte von p ist ein Flugzeug mit zwei Triebwerken sicherer als ein Flugzeug mit vier Triebwerken?

Aufgabe 38

Wir nehmen an der Anteil der Linkshänder in der Bevölkerung sei 1%.

(i) Wie gross ist die Wahrscheinlichkeit, dass unter 300 zufällig ausgewählten Personen mindestens 5 Linkshänder sind?

(ii) Wie gross muss die Stichprobe sein, damit sich mit einer Wahrscheinlichkeit von 95% mindestens 1 Linkshänder darin befindet?

Aufgabe 39

Ein Spital hat 15 Zimmer, jedes hat eine Wahrscheinlichkeit von 30%, besetzt zu sein.

(i) Was ist die Wahrscheinlichkeit, dass es nicht mehr als 3 besetzte Zimmer gibt?

(ii) Was ist die mittlere Anzahl besetzter Räume?

Aufgabe 40

Wir betrachten die Funktion

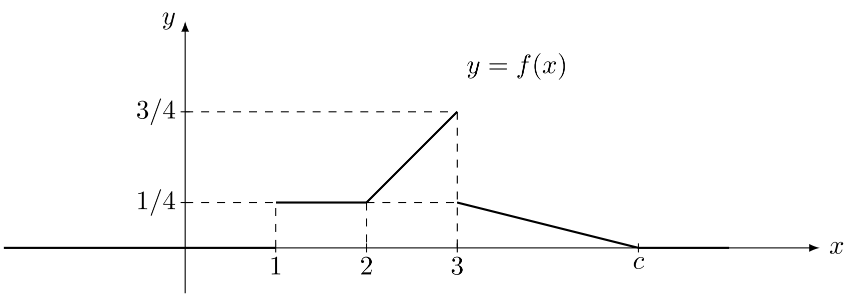
$$f(x) = \begin{cases} 0.02x & 0 \leq x \leq 10, \\ 0 & \text{sonst.} \end{cases}$$

(i) Man zeige dass es sich bei $f(x)$ um eine Dichtefunktion handelt.

(ii) Man berechne den Erwartungswert und die Varianz.

Aufgabe 41

Eine stetige Zufallsvariable X ist durch ihre Dichtefunktion $f(x)$ folgendermassen gegeben:



(i) Man bestimme die Konstante c (mit Begründung).

(ii) Man bestimme $P(X \leq \frac{3}{2})$, $P(X \geq 2)$, $P(X = \frac{5}{2})$.

(iii) Man skizziere die zugehörige Verteilungsfunktion.

Aufgabe 42

Eine Zufallvariable Y ist durch die folgende Wahrscheinlichkeitsdichte gegeben:

$$f(y) = \begin{cases} 4y^3 & \text{für } 0 \leq y \leq 1 \\ 0 & \text{sonst} \end{cases}$$

(i) Man finde y_0, so dass

$$P(Y \leq y_0) = 0.3.$$

(ii) Man berechne den Erwartungswert $E(Y)$.

(iii) Man berechne die Varianz $V(Y)$.
(iv) Man berechne $E(e^Y)$.

Aufgabe 43

Eine Zufallsvariable X ist gleichmässig verteilt zwischen a und b, i.e.

$$f(x) = \begin{cases} C & a \leq x \leq b, \\ 0 & \text{sonst.} \end{cases}$$

Die Zufallsvariable X besitzt den folgenden Erwartungswert $E(X)$ und die folgende Varianz $V(X)$:

$$E(X) = 6 \qquad V(X) = 3.$$

(i) Man bestimme die Konstante C.
(ii) Man bestimme die Wahrscheinlichkeit dass $P(X \geq 5)$.

Aufgabe 44

Auf einer Hauptstrasse, wo die Geschwindigkeit auf 80 km/h begrenzt ist, misst ein Radar die Geschwindigkeit aller Autos während eines Tages. Unter der Annahme dass die gemessenen Geschwindigkeiten normalverteilt sind, mit einem Mittelwert von 72 km/h und einer Standardabweichung von 8 km/h, bestimme man den Anteil der Autofahrer, welche zu schnell gefahren sind.

Aufgabe 45

Angenommen die Körpergrösse in cm von 25-jährigen Männern sei normalverteilt mit den Parametern $\mu = 175$ und $\sigma = 6$.

(i) Welcher prozentuale Anteil der 25-jährigen Männer besitzt eine Körpergrösse von mehr als 185 cm?
(ii) Welcher Prozentsatz der Männer, deren Körpergrösse mehr als 180 cm beträgt, besitzt eine Körpergrösse von mehr als 192 cm?

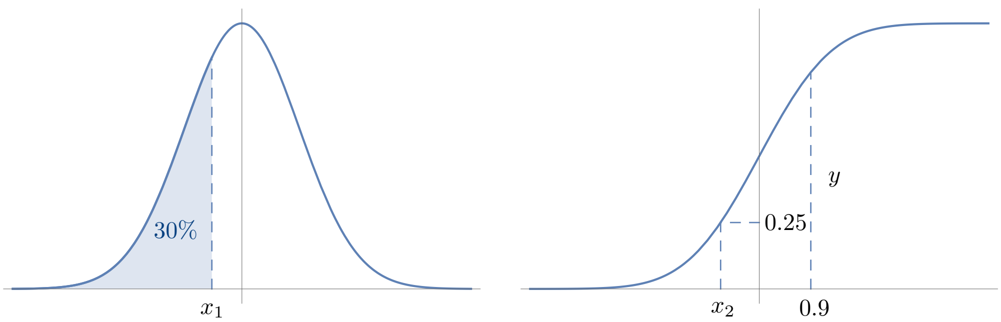
Aufgabe 46

Sei X eine standard-normalverteilte Zufallvariable. Man bestimme

(i) $P(X \leq 2.44)$
(ii) $P(X \leq -1.16)$
(iii) $P(X \leq 1.92)$
(iv) $P(X \geq 1)$
(v) $P(X \geq -2.9)$
(vi) $P(2 \leq X \leq 10)$

Aufgabe 47

Die nachfolgenden Graphiken zeigen Dichte und kumulative Verteilungsfunktion der Standardnormalverteilung:



(i) Bestimmen Sie x_1, x_2 und y.
(ii) Geben Sie eine Interpretation für y.

Aufgabe 48

Die Geschwindigkeit von Fahrzeugen in einem Autobahnabschnitt ist normalverteilt mit einem Mittelwert von 91 km/h und einer Standardabweichung von 12 km/h.

(i) Das Tempolimit liegt bei 100 km/h. Wie gross ist der Anteil der Fahrzeuge die mit einem Tempo unterhalb oder gleich dem Tempolimit unterwegs sind?
(ii) Wie gross ist der Anteil der Fahrzeuge die mit weniger als 49 km/h unterwegs sind?
(iii) Ein neues Tempolimit tritt in Kraft, so dass ca. 5 % der Fahrzeuge über dem Tempolimit unterwegs sind. Welches ist das neue Tempolimit aufgrund dieses Kriteriums? (Annahme: Die Geschwindigkeitsverteilung der Fahrzeuge ändert sich aufgrund des neuen Tempolimits nicht.)

Aufgabe 49

Sei X eine normalverteilte Zufallvariable mit Parametern $\mu = 0.8$ und $\sigma = 2$. Man bestimme

(i) $P(X \leq 2.44)$
(ii) $P(X \leq -1.16)$
(iii) $P(X \leq 1.923)$
(iv) $P(X \geq 1)$
(v) $P(X \geq -2.9)$
(vi) $P(2 \leq X \leq 10)$

Aufgabe 50

Eine Metallhobelmaschine stellt Platten her. Die Plattendicke X (in mm) lässt sich als Zufallsvariable auffassen, die von Platte zu Platte etwas andere Werte annimmt. X sei normalverteilt und habe bei einer bestimmten Maschineneinstellung den Mittelwert $\mu = 10$ mm und die Standardabweichung $\sigma = 0.02$ mm.

(i) Wieviel Prozent Ausschuss sind zu erwarten, wenn die Platten mindestens 9.97 mm dick sein dürfen?
(ii) Wieviel Prozent Ausschuss sind zu erwarten, wenn die Platten höchstens 10.05 mm dick sein dürfen?
(iii) Wieviel Prozent Ausschuss sind zu erwarten, wenn die Platten um maximal ± 0.03 mm vom Sollwert 10 mm abweichen dürfen?
(iv) Mit dieser Toleranz von ± 0.03 mm, wie stark verringert sich der Ausschussanteil, wenn man eine genauere Hobelmaschine mit $\sigma = 0.01$ mm einsetzt?
(v) Mit dieser Toleranz von ± 0.03 mm, wie gross müsste σ sein, damit man nur 1 % Ausschuss erhält?

(vi) Wie muss man die Toleranzgrenzen $10 - c$ und $10 + c$ wählen, damit man nicht mehr als 5 % Ausschuss erhält?

Aufgabe 51

Die Anzahl Kilometer, während der eine Autobatterie ohne Panne funktioniert, ist exponentiell verteilt mit Mittelwert 12000 km. Jemand möchte eine Reise von 6000 km unternehmen.

(i) Man bestimme die Wahrscheinlichkeit dass diese Reise ohne Batteriepanne durchgeführt werden kann.

(ii) Man bestimme eine Formel für die Wahrscheinlichkeit im Punkt (i), falls die Lebensdauer nicht exponentiell verteilt ist. Man bezeichne dazu die kumulative Verteilungsfunktion der Lebensdauer der Batterie mit F und die Anzahl gefahrener Kilometer seit ihrer Inbetriebnahme mit n.

Aufgabe 52

Sei X eine normalverteilte Zufallvariable mit Parametern $\mu = 0$ und $\sigma = 1$. Man bestimme den Wert der Konstanten c, so dass

(i) $P(X \geq c) = 0.1$

(ii) $P(X \leq c) = 0.05$

(iii) $P(0 \leq X \leq c) = 0.45$

(iv) $P(-c \leq X \leq c) = 0.99$

Aufgabe 53

Es sei $X \sim \mathrm{Exp}(\lambda)$. Man zeige dass

$$E(X) = \frac{1}{\lambda}, \qquad V(X) = \frac{1}{\lambda^2}.$$

Aufgabe 54

Sei X eine normalverteilte Zufallsvariable mit Parametern $\mu = -2$ und $\sigma = 0.5$. Man bestimme den Wert der Konstanten c, so dass

(i) $P(X \geq c) = 0.2$

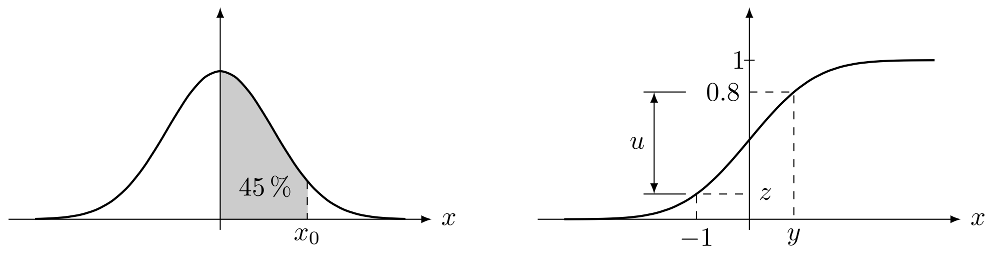
(ii) $P(-c \leq X \leq -1) = 0.5$

(iii) $P(-2 - c \leq X \leq -2 + c) = 0.9$

(iv) $P(-2 - c \leq X \leq -2 + c) = 0.996$

Aufgabe 55

Die nachfolgenden Graphiken zeigen Dichtefunktion und kumulative Verteilungsfunktion der Standardnormalverteilung:



(i) Man bestimme x_0, y, z.

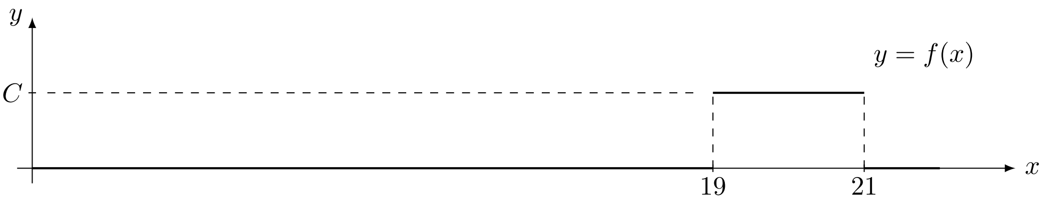
(ii) Man gebe eine Interpretation für u.

Aufgabe 56

Ein Warenlift besitzt eine maximale Nutzlast von genau 10000 kg. Mit dem Warenlift sollen 200 Säcke Zement transportiert werden. Aus Erfahrung weiss man dass das Gewicht eines Zementsackes einer Wahrscheinlichkeitsverteilung mit $\mu = 49.5\,\text{kg}$ und $\sigma = 5 \cdot \sqrt{2}\,\text{kg}$ folgt. Wie gross ist die Wahrscheinlichkeit dass beim Transport etwas schief geht?

Aufgabe 57

Eine Firma produziert Betonelemente. Die Länge eines Elements ist eine Zufallsvariable X. Aus dem Herstellprozess ergibt sich die folgende Dichtefunktion $f(x)$ für X (Angaben in Meter):



(i) Man bestimme die Konstante C.
(ii) Man skizziere die dazugehörige kumulative Verteilungsfunktion.
(iii) Man bestimme den Erwartungswert μ und die Standardabweichung σ dieser Wahrscheinlichkeitsverteilung. Hinweis: Die auftretenden Integrale für μ und σ können durch Substitution vereinfacht werden.
(iv) Über einen Fluss mit Breite 2005 m wird eine Brücke mit diesen Elementen gebaut. Man bestimme die Wahrscheinlichkeit, dass die Brücke mit 100 Elementen gebaut werden kann. Hinweis: $\sqrt{3}/2 = 0.866$.

Aufgabe 58

Eine Tabakfirma behauptet dass die Nikotinmenge in einer ihrer Zigaretten eine Zufallsvariable mit Mittelwert 2.2 mg und Standardabweichung 0.3 mg ist. Bei einer Messung wurde in 100 zufällig ausgewählten Zigaretten ein durchschnittlicher Nikotingehalt von 3.1 mg gemessen. Man finde (approximativ) die Wahrscheinlichkeit dass der durchschnittliche Nikotingehalt von 100 ausgewählten Zigaretten dem gemessenen oder einem höheren Wert entspricht, wenn davon ausgegangen wird dass die Behauptung der Tabakfirma wahr ist.

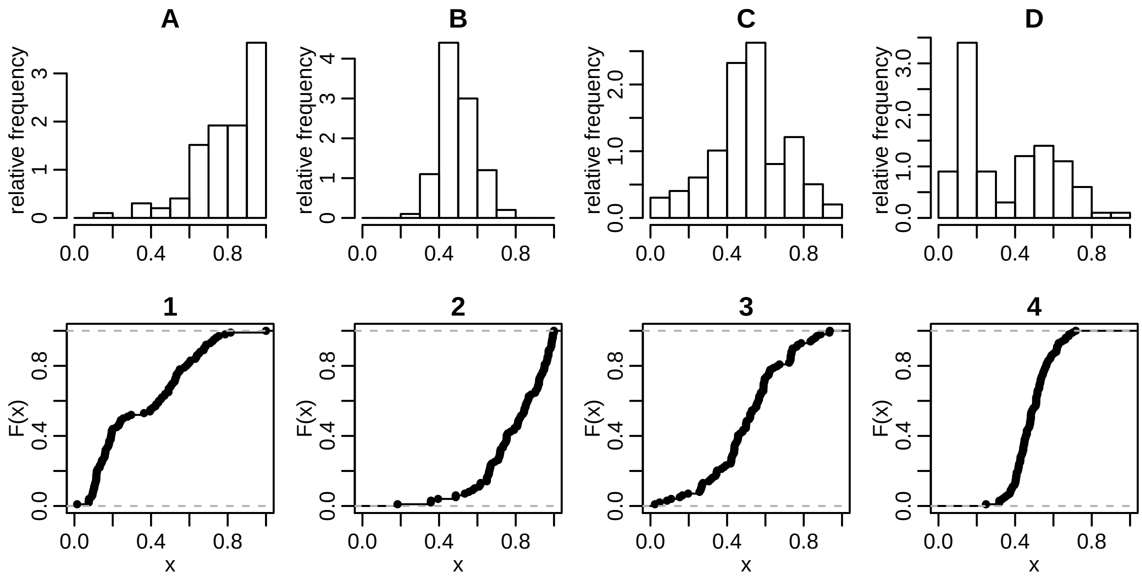
Aufgabe 59

Sie möchten die durchschnittliche Körperlänge von 30-jährigen Männern bestimmen. Sie nehmen an, dass die Körperlänge eine bestimmte Verteilung besitzt, die nicht zwingend die Normalverteilung sein muss, mit Standardabweichung $\sigma = 4\,\text{cm}$.

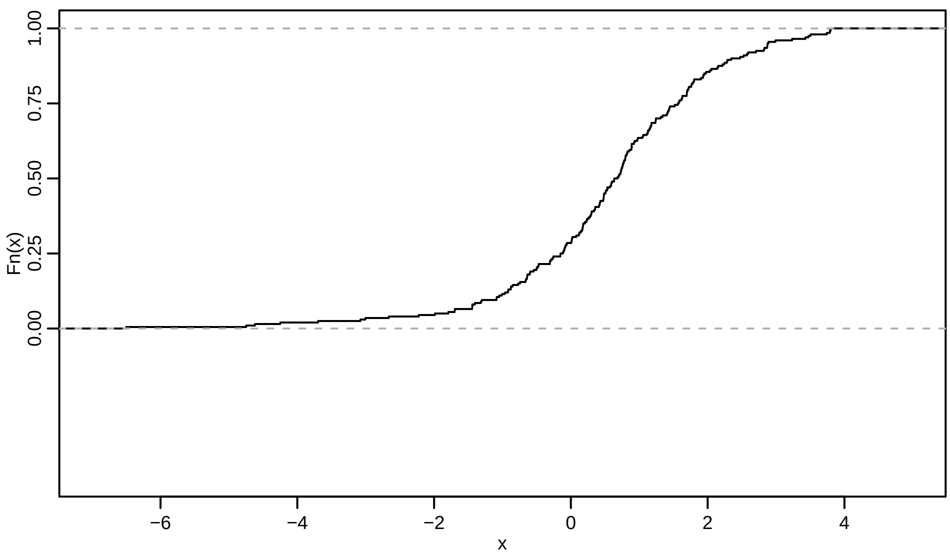
(i) Wie gross muss die Stichprobe sein (i.e. wie viele Männer müssen ausgemessen werden), wenn die Standardabweichung des Stichprobenmittels höchstens 0.25 cm betragen soll?
(ii) Sie haben eine Stichprobe der Grösse n. Sie möchten die Standardabweichung des Stichprobenmittels halbieren. Wie gross muss die neue Stichprobe sein?

Aufgabe 60

(i) Man ordne jedem der folgenden Histogramme die zugehörige empirische kumulative Verteilungsfunktion $F(x)$ zu:



(ii) Die folgende Darstellung zeigt die empirische kumulative Verteilungsfunktion eines Datensatzes:



Man skizziere einen Box-Plot des Datensatzes in die gegebene Grafik hinein, ohne Ausreisser zu berücksichtigen.

Aufgabe 61

Durch eine Pulverabfüllanlage werden Behälter mit je 4 mg Pulver abgefüllt. Die Genauigkeit der Anlage wird durch eine Standardabweichung von 1 mg angegeben. Ein Inspektor inspiziert 36 (zufällig ausgewählte) Behälter und findet eine durchschnittliche Pulvermenge von 4.3 mg.

Man untersuche mit einem statistischen Test ob die Anlage neu justiert werden muss. Man verwende dazu ein Signifikanzniveau von 5 %. Man gebe den p-Wert des Tests an. Man erkläre in einem Satz die Interpretation des p-Werts.

Aufgabe 62

Eine Maschine produziert Schrauben der Länge 10 cm. Die Schrauben sollen werder zu lang noch zu kurz sein. Aus der Maschinengenauigkeit ist eine Standardabweichung von 0.5 cm bekannt. Aus einer Messung von 36 Schrauben ergibt sich eine durchschnittliche Länge von 9.7 cm. Es soll mit einem statistischen Test festgestellt werden, ob die Maschine neu justiert werden muss.

Aufgabe 63

Es soll mit einem statistischen Test herausgefunden werden, ob die mittlere tägliche Kalorienmenge bei der erwachsenen, ländlichen Bevölkerung in einem Entwicklungsland weniger als 2000 beträgt. Eine Stichprobe von 500 Personen hat einen Mittelwert von 1985 und eine Standardabweichung von $\sigma = 210$ ergeben. Das Signifikanzniveau für den statistischen Test sei $\alpha = 0.05$.

Aufgabe 64

Es soll geprüft werden, ob der mittlere Cholesterinspiegel bei asiatischen Frauen zwischen 21 und 40 Jahren, die kürzlich in die USA immigirert sind, sich signifikant von jenem ihrer Altersgenossinnnen, die in den USA aufgewachsen sind, unterscheidet. Der Mittelwert des Cholesterinspiegels der Frauen zwischen 21 und 40 Jahren, die in den USA aufgewachsen sind, ist gleich 190 mg/dl. Es wird eine Stichprobe von 100 asiatischen Immigrantinnen zwischen 21 und 40 Jahren bestimmt. Es ergibt sich ein Mittelwert von $\overline{x} = 181.52\,\mathrm{mg/dl}$ und eine Standardabweichung von $\sigma = 40\,\mathrm{mg/dl}$. Liegt ein signifikanter Unterschied vor, wenn $\alpha = 0.05$ ist?

9. Lösungen zu den Aufgaben

Lösung zu Aufgabe 1

Es handelt sich um die Anzahl möglicher Anordnungen der fünf Läufer. Somit gilt: $P_5 = 5! = 120$.

Lösung zu Aufgabe 2

$10 \cdot 9 \cdot 8 = 720$,
$(n+1)(n+2)\cdots(n+i-1)(n+i)$
$(n-1)n$
$(n-r+1)(n-r)$

Lösung zu Aufgabe 3

(i) 7!, (ii) Eine Person kann sich an einen beliebigen Platz des runden Tisches setzen, die sechs anderen können sich danach auf 6! Arten an den Tisch setzen.

Lösung zu Aufgabe 4

(i) Es handelt sich um die Anzahl möglicher Anordnungen der vier Ziffern. Diese ist $4! = 24$.

(ii) Wir betrachten die möglichen Positionen der Ziffer 1. Steht die Ziffer 1 an erster Position, so sind drei weitere Positionen offen, auf welche die restlichen drei Ziffern verteilt werden können. Somit gibt es $3! = 6$ mögliche Zahlen mit der Ziffer 1 in der ersten Position. Der Beitrag der Ziffer 1 dieser Zahlen beträgt somit $6 \cdot 1000$. Steht die Ziffer 1 an zweiter Stelle so sind wiederum $3! = 6$ Zahlen möglich und der Beitrag der Ziffer 1 dieser Zahlen beträgt $6 \cdot 100$. Die Ziffer 1 an der dritten Position liefert einen Beitrag $6 \cdot 10$ und die Ziffer 1 an der vierten Position liefert einen Beitrag $6 \cdot 1$. Insgesamt liefert die Ziffer 1 somit den Beitrag $6 \cdot 1111$. Analog liefert die Ziffer 3 den Beitrag $6 \cdot 3333$, die Ziffer 5 den Beitrag $6 \cdot 5555$ und die Ziffer 7 den Beitrag $6 \cdot 7777$. Somit ist die Summe

$$6 \cdot (1111 + 3333 + 5555 + 7777).$$

Lösung zu Aufgabe 5

(i) Für den ersten Turm welcher platziert wird gibt es $8 \cdot 8$ Möglichkeiten. Um einer Bedrohung auszuweichen darf der zweite Turm nicht in der selben Zeile oder Spalte wie der erste Turm stehen. Somit ergeben sich für die Platzierung des zweiten Turms $7 \cdot 7$ Möglichkeiten. Folgen wir diesem Schema so ergeben sich insgesamt

$$8 \cdot 8 \cdot 7 \cdot 7 \cdots 1 \cdot 1 = (8!)^2$$

Möglichkeiten.

(ii) Es muss in jeder Zeile und in jeder Spalte genau ein Turm stehen. Wir betrachten die erste Spalte. Es gibt acht Möglichkeiten den Turm in der ersten Spalte zu platzieren. Der Turm in der ersten Spalte blockiert nun eine Zeile, somit gibt es in der zweiten Spalte nur noch sieben Möglichkeiten den Turm zu platzieren. Folgen wir diesem Schema, so erhalten wir insgesamt 8! Möglichkeiten.

Lösung zu Aufgabe 6

Es gibt 4!3!2!1! Möglichkeiten die Bücher ins Regal zu stellen, wenn die Mathematikbücher zuerst kommen, dann die Chemiebücher, dann die Geschichtsbücher und dann der Roman. Analog gibt es für jede mögliche Anordnung der Themen 4!3!2!1! Möglichkeiten die Bücher ins Regal zu stellen. Da es 4! mögliche Anordnungen der Themen gibt, gibt es insgesamt $4!4!3!2!1! = 6912$ Möglichkeiten die Bücher ins Regal zu stellen, wenn die Bücher eines Themas jeweils nebeneinander stehen sollen.

Lösung zu Aufgabe 7

Es handelt sich um die Anzahl möglicher Anordnungen von drei aus den fünf Läufern. Somit gilt: $V_5^3 = \frac{5!}{(5-3)!} = \frac{5!}{2!} = 60$.

Lösung zu Aufgabe 8

Für die Wahl der ersten Ziffer gibt es neun Möglichkeiten (diese muss ein Element der Menge $\{1, 2, \cdots, 9\}$ sein). Für die Wahl der zweiten, dritten und vierten Ziffer können wir nun aus den verbleibenden neun Ziffern drei auswählen (und anordnen), i.e. für diese Ziffern gibt es $9 \cdot 8 \cdot 7$ Möglichkeiten. Total gibt es somit $9 \cdot 9 \cdot 8 \cdot 7 = 4536$ mögliche Zahlen.

Lösung zu Aufgabe 9

Es handelt sich um die Anzahl Möglichkeiten aus den fünf Elementen drei Elemente auszuwählen, wobei die Reihenfolge keine Rolle spielt. I.e. es gibt

$$\binom{5}{3} = \frac{5!}{(5-3)!3!} = 10$$

mögliche Teilmengen mit drei Elementen.

Lösung zu Aufgabe 10

Die Anzahl Möglichkeiten die Frauen auszuwählen ist $\binom{5}{2}$, die Anzahl Möglichkeiten die Männer auszuwählen ist $\binom{7}{3}$. Somit ist die Anzahl Möglichkeiten das Gremium zu bilden $\binom{5}{2}\binom{7}{3}$.

Lösung zu Aufgabe 11

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

Lösung zu Aufgabe 12

(i) $\binom{10}{8} = 45$

(ii) $\binom{7}{5} = 21$

(iii) Wenn er die ersten fünf Fragen richtig beantwortet, gibt es für die restlichen noch $\binom{5}{3} = 10$ Möglichkeiten. Beantwortet er nur vier der ersten fünf Fragen richtig, wofür es $\binom{5}{4} = 5$ Möglichkeiten gibt, so kann er die restlichen vier richtigen Antworten auf die anderen fünf Fragen geben, wofür es $\binom{5}{4} = 5$ Möglichkeiten gibt. In diesem Fall hat er also $5 \cdot 5 = 25$ Möglichkeiten. Insgesamt gibt es somit 35 Möglicheiten.

Lösung zu Aufgabe 13

Wir haben

$$(2x^2 - \tfrac{1}{2}y^3)^8 = \sum_{k=0}^{8} \binom{8}{k} (2x^2)^{8-k}(-\tfrac{1}{2}y^3)^k.$$

Der Term welcher x zur achten Potenz enthält entspricht dem Term in der Summe mit $k = 4$. Dieser Term ist

$$\binom{8}{4} (2x^2)^{8-4}(-\tfrac{1}{2}y^3)^4 = \binom{8}{4} x^8y^{12} = 70x^8y^{12}.$$

Lösung zu Aufgabe 14

Wir können die Karten in einem vierstufigen Prozess verteilen:

(i) Spieler A erhält neun der 36 Karten, Anzahl Möglichkeiten:

$$n_1 = \binom{36}{9}$$

(ii) Spieler B erhält neun der verbleibenden 27 Karten, Anzahl Möglichkeiten:

$$n_2 = \binom{27}{9}$$

(iii) Spieler C erhält neun der verbleibenden 18 Karten, Anzahl Möglichkeiten:

$$n_3 = \binom{18}{9}$$

(iv) Spieler D erhält neun der verbleibenden 9 Karten, Anzahl Möglichkeiten:

$$n_4 = \binom{9}{9} = 1$$

Insgesamt erhalten wir

$$\binom{36}{9}\binom{27}{9}\binom{18}{9}\binom{9}{9} \approx 2.15 * 10^{19}.$$

Lösung zu Aufgabe 15

Aus den sechs Faktoren in $(xy - y^2 + 2z)^6$ müssen wir dreimal xy auswählen. Aus den verbleibenden drei Faktoren müssen wir einmal $-y^2$ auswählen. Aus den verbleibenden zwei Faktoren müssen wir zweimal $2z$ auswählen. Es ergibt sich somit der Term

$$\binom{6}{3} (xy)^3 \binom{3}{1} (-y^2) \binom{2}{2} (2z)^2 = -240x^3y^5z^2.$$

Lösung zu Aufgabe 16

(i) Binomischer Lehrsatz mit $x = 1$, $y = -1$. (ii) Binomischer Lehrsatz mit $x = y = 1$.

Lösung zu Aufgabe 17

Wir überlegen uns die Anzahl Möglichkeiten Teilmengen mit k Elementen zu bilden. Für $k = 0$ gibt es eine ($\binom{n}{0} = 1$) Möglichkeit (dies entspricht der leeren Menge), für $k = 1$ gibt es $\binom{n}{1}$ Möglichkeiten, für $k = 2$ gibt es $\binom{n}{2}$ Möglichkeiten, $\ldots$, für $k = n$ gibt es eine Möglichkeit ($\binom{n}{n} = 1$), dies entspricht der Menge selbst. Somit gibt es insgesamt

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3} + \ldots + \binom{n}{n} = 2^n$$

Teilmengen, wobei wir für diese Gleichung das Resultat der vorhergehenden Übung (ii) verwendet haben.

Lösung zu Aufgabe 18

(i) Eine Gerade wird durch zwei Punkte bestimmt, Anzahl Geraden ist somit $\binom{10}{2} = 45$.

(ii) Es verlaufen $\binom{8}{2} = 28$ Geraden weder durch A noch durch B. Somit verlaufen $45 - 28 = 17$ Geraden durch A oder B.

(iii) Ein Dreieck ist durch drei Punkte bestimmt, Anzahl Dreiecke ist somit $\binom{10}{3} = 120$.

(iv) Zwei Punkte sind fixiert (A, B), bleibt ein Punkt auszuwählen, Anzahl Möglichkeiten ist $\binom{8}{1} = 8$.

Lösung zu Aufgabe 19

Für jede der n Ziffern kann eine 1 oder eine 0 gewählt werden. Somit ergibt sich eine Anzahl von 2^n.

Lösung zu Aufgabe 20

Für jedes Spiel haben wir die drei Möglichkeiten Gewinn Heimmannschaft, Gewinn Gastmannschaft und unentschieden. Somit ergeben sich total $3^{13} = 1594323$ Mögliche Tippabgaben.

Lösung zu Aufgabe 21

Die symmetrischen $n \times n$-Matrizen besitzen $\frac{n(n+1)}{2}$ wählbare Einträge (die restlichen sind durch die Symmetrie bestimmt). Jeder Eintrag in $\{1, 2, \ldots, 9\}$ ergibt $9^{\frac{n(n+1)}{2}}$ solche Matrizen.

Lösung zu Aufgabe 22

Mit dem Hinweis sehen wir dass Anzahl Separatoren | und Anzahl '+'-Zeichen bei jeder Lösung gleich 14 sein muss. Somit ist die Anzahl Lösungen gleich der Anzahl

Möglichkeiten die drei Separatoren auf diesen 14 Positionen zu platzieren. Diese Anzahl ist $\binom{14}{3} = 364$.

Lösung zu Aufgabe 23

(i) E^c: 'Die Zahl ist ungerade'. $E^c = \{1, 3, 5\}$, $P(E^c) = 1 - P(G) = 0.5$.
(ii) F^c: 'Die Zahl ist kleiner oder gleich zwei'. $F^c = \{1, 2\}$, $P(F^c) = P(1)+P(2) = \frac{1}{3}$.
(iii) G^c: 'Es wird weniger als eine zwei gewürfelt'. $G^c = \{1\}$, $P(G^c) = P(1) = \frac{1}{6}$.

Lösung zu Aufgabe 24

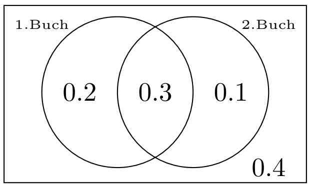
Sei E das Ereignis dass Anne das erste Buch mag und sei F das Ereignis dass Anne das zweite Buch mag. Wir haben $P(E) = 0.5$, $P(F) = 0.4$, $P(E \cap F) = 0.3$. Somit ist die Wahrscheinlichkeit dass Anne eines der Bücher mag

$$P(E \cup F) = P(E) + P(F) - P(E \cap F) = 0.5 + 0.4 - 0.3 = 0.6.$$

Und die Wahrscheinlichkeit dass Anne keines der Bücher mag ist

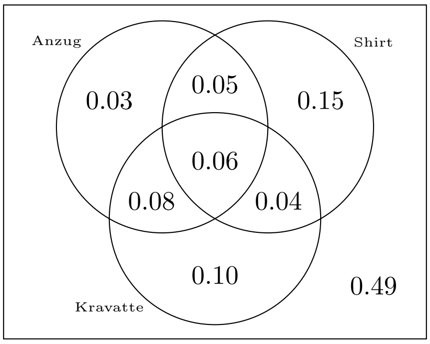
$$P((E \cup F)^c) = 1 - P(E \cup F) = 1 - 0.6 = 0.4.$$

Veranschaulichung in Venndiagramm mit eingetragenen Wahrscheinlichkeiten:



Lösung zu Aufgabe 25

Veranschaulichung in Venndiagramm mit eingetragenen Wahrscheinlichkeiten:



(i) 0.49, (ii) 0.28.

Lösung zu Aufgabe 26

Für den zweiten Gast gibt es $8 \cdot 4 - 1 = 31$ freie Plätze. Die willkürliche Wahl deutet darauf hin dass der zweite Gast jeden dieser freien Plätze mit der selben Wahrscheinlichkeit aussucht. Die Anzahl günstiger Fälle ist die Anzahl freier Plätze am Tisch des ersten Gastes, also drei. Somit ergibt sich $P(E) = \frac{3}{31}$, wobei wir mit E das Ereignis bezeichnen dass beide Gäste am selben Tisch sitzen.

Lösung zu Aufgabe 27

(i) Die möglichen Ergebnisse der Betrachtung sind die möglichen Summen der Augenzahlen, i.e.

$$\Omega = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}.$$

(ii) Nein. Es ist zum Beispiel wahrscheinlicher dass die Summe gleich sieben ist als dass sie gleich zwei ist (für eine konkrete Berechung dieser Wahrscheinlichkeiten siehe später).

(iii) Man betrachtet die möglichen Würfe mit zwei Würfen, i.e.

$$\Omega = \{(i, j) : i, j = 1, 2, 3, 4, 5, 6\}.$$

Bei dieser Betrachtung sind alle Ergebnisse gleichwahrscheinlich.

Lösung zu Aufgabe 28

Die Anzahl möglicher Kommitees ist $\binom{15}{5} = 3003$. Die Anzahl günstiger Kommitees, i.e. die Anzahl möglicher Kommitees bestehend aus drei Männern und zwei Frauen, ist gleich $\binom{6}{3}\binom{9}{2} = 720$. Die Wahrscheinlichkeit für ein Kommitee bestehend aus drei Männern und zwei Frauen ist somit $\frac{720}{3003} = 0.24$.

Lösung zu Aufgabe 29

(i) 1/36. (ii) 35/36, (iii) 11/36, (iv) 25/36.

Lösung zu Aufgabe 30

$6!/6^6$.

Lösung zu Aufgabe 31

$\frac{3}{5}$.

Lösung zu Aufgabe 32

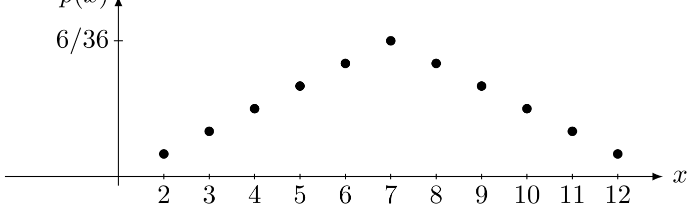
(i) $\frac{3}{4}$, (ii) $\frac{1}{4}$, (iii) $\frac{1}{4}\frac{1}{4}\frac{3}{4} = \frac{3}{64}$, (iv) $\left(\frac{1}{4}\right)^4 = \frac{1}{256}$.

Lösung zu Aufgabe 33

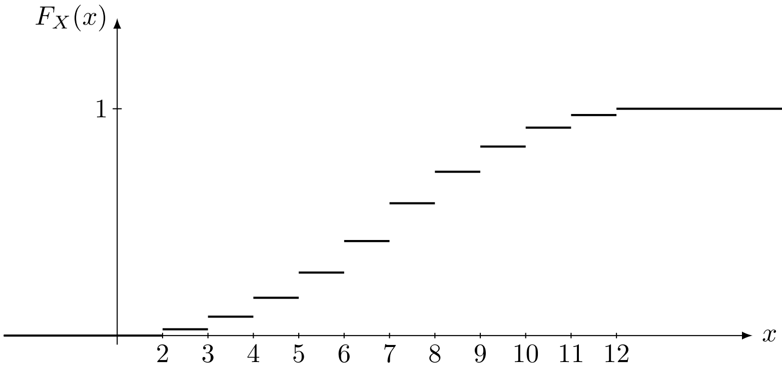
Sei $X =$ Summe der Augenzahlen. Wir finden

x_i	2	3	4	5	6	7	8	9	10	11	12
p_1	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Die grafische Darstellung der Wahrscheinlichkeitsverteilung ist:



Die grafische Darstellung der kumulativen Wahrscheinlichkeitsverteilung ist:

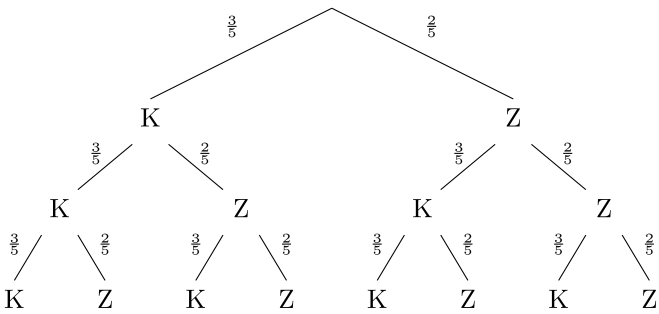


Lösung zu Aufgabe 34

x_i	1	2	3	4	5	6
p_i	1/36	3/36	5/36	7/36	9/36	11/36

Lösung zu Aufgabe 35

(i) Das Baumdiagramm ist:



(ii) Der Stichprobenraum ist

$$\Omega = \{\mathrm{KKK}, \mathrm{KKZ}, \mathrm{KZK}, \mathrm{ZKK}, \mathrm{KZZ}, \mathrm{ZKZ}, \mathrm{ZZK}, \mathrm{ZZZ}\}$$

und die Wahrscheinlichkeiten für die Elementarereignisse $\omega \in \Omega$ sind

ω	KKK	KKZ	KZK	ZKK	KZZ	ZKZ	ZZK	ZZZ
$P(\omega)$	$\frac{27}{125}$	$\frac{18}{125}$	$\frac{18}{125}$	$\frac{18}{125}$	$\frac{12}{125}$	$\frac{12}{125}$	$\frac{12}{125}$	$\frac{8}{125}$

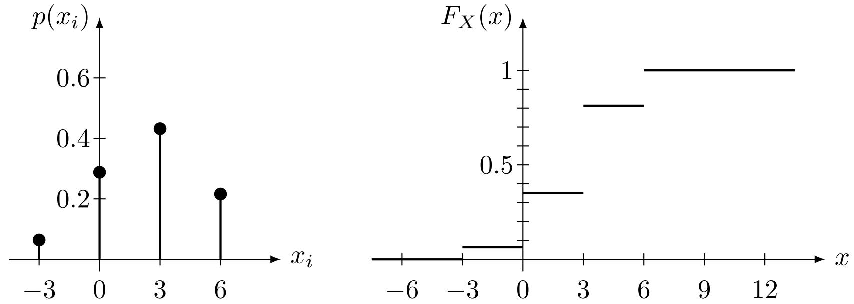
(iii) Die Werte der Zufallsvariablen X sind

ω	KKK	KKZ	KZK	ZKK	KZZ	ZKZ	ZZK	ZZZ
$X(\omega)$	6	3	3	3	0	0	0	-3

(iv) Die möglichen Werte x_i von X und die dazugehörigen Wahrscheinlichkeiten sind

x_i	-3	0	3	6
$p_i = P(X = x_i)$	$\frac{8}{125}$	$\frac{36}{125}$	$\frac{54}{125}$	$\frac{27}{125}$

Grafische Darstellung der Wahrscheinlichkeitsverteilung der zugehörigen kumulativen Verteilungsfunktion:



(v) Erwartungswert und Varianz sind

$$E(X) = \sum_i p_i x_i = \frac{8}{125} \cdot (-3) + \frac{36}{125} \cdot 0 + \frac{54}{125} \cdot 3 + \frac{27}{125} \cdot 6 = 2.4,$$

$$V(X) = \sum_i p_i (x_i - E(X))^2 = \frac{8}{125} \cdot 5.4^2 + \frac{36}{125} \cdot 2.4^2 + \frac{54}{125} \cdot 0.6^2 + \frac{27}{125} \cdot 3.6^2 = 6.48.$$

Die Standardabweichung ist somit

$$\sigma = \sqrt{V(X)} = \sqrt{6.48} \approx 2.546.$$

(vi) Der Einsatz muss kleiner oder gleich $E(X)$ sein, i.e. kleiner oder gleich 2.4.

Lösung zu Aufgabe 36

Sei X die Anzahl Kopf. X ist binomialverteilt.

(i) $P(X = 2) = \binom{6}{2} \left(\frac{1}{2}\right)^6 = \frac{15}{64} = 0.2344,$

(ii) $P(X \geq 4) = \left(\binom{6}{4} + \binom{6}{5} + \binom{6}{6}\right) \left(\frac{1}{2}\right)^6 = \frac{11}{32} = 0.34375,$

(iii) $P(X = 0) = \binom{6}{0} \left(\frac{1}{2}\right)^6 = \frac{1}{64} = 0.015625.$

Lösung zu Aufgabe 37

Die Wahrscheinlichkeit dass das Flugzeug mit zwei Triebwerken abstürzt ist

$$\binom{2}{2} p^2 = p^2.$$

Die Wahrscheinlichkeit dass das Flugzeug mit vier Triebwerken abstürzt ist

$$\binom{4}{3} p^3(1-p) + \binom{4}{4} p^4 = 4p^3(1-p) + p^4 = 4p^3 - 3p^4.$$

Somit ist das Flugzeug mit zwei Triebwerken sicherer, falls

$$4p^3 - 3p^4 > p^2,$$

i.e.

$$4p - 3p^2 > 1,$$

i.e.

$$0 > 3p^2 - 4p + 1.$$

Das quadratische Polynom auf der rechten Seite hat Nullstellen bei $p_1 = \frac{1}{3}$, $p_2 = 1$. Somit ist für $p > 1/3$ das Flugzeug mit zwei Triebwerken sicherer.

Lösung zu Aufgabe 38

Binomialverteilung mit $p = 0.01$.

(i)

$$1 - \sum_{i=0}^{4} \binom{300}{i} 0.01^i 0.99^{300-i} \approx 18.4\%$$

(ii) Wenn n die Grösse der Stichprobe ist, so muss die Ungleichung $1 - 0.99^n \geq 0.95$ gelten. Dies ergibt

$$n \geq \frac{\log(0.05)}{\log(0.99)} \approx 298.07$$

Lösung zu Aufgabe 39

Binomialverteilung mit $p = 0.3$.

(i)

$$\sum_{i=0}^{3} \binom{15}{i} 0.3^i 0.7^{15-i} \approx 29.7\%$$

(ii) Der Erwartungswert einer Binomialverteilung mit Parametern $p = 0.3$ und $n = 15$ ist $pn = 4.5$.

Lösung zu Aufgabe 40

(i) Die Funktionswerte sind grösser gleich Null und wir haben

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^{10} 0.02x dx = 0.01x^2 \Big|_0^{10} = 1.$$

(ii) Wir haben

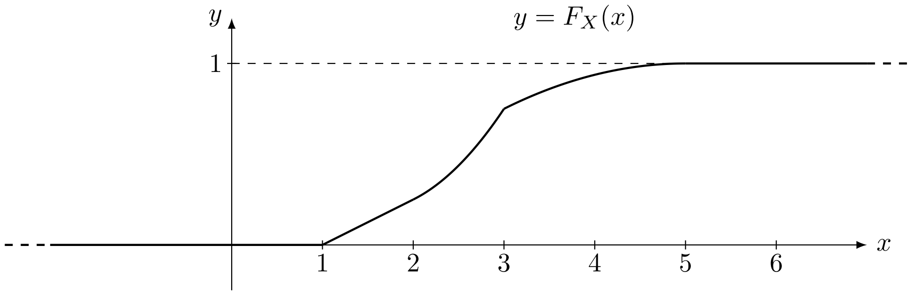
$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = \ldots = \frac{20}{3},$$
$$V(X) = \int_{-\infty}^{\infty} (x - E(X))^2 f(x) dx = \int_0^{10} \left(x - \frac{20}{3}\right)^2 0.02x dx = \ldots = \frac{50}{9}.$$

Lösung zu Aufgabe 41

(i) Wir fordern $\int_{-\infty}^{\infty} f(x)dx = 1$, i.e. die Fläche unter $y = f(x)$ muss eins sein. Die Fäche unter $y = f(x)$ bis $x = 3$ ist $3/4$. Somit bleiben $1/4$ für den Abschnitt zwischen $x = 3$ und $x = c$. Daraus ergibt sich $c = 5$.

(ii) $P(X \leq \frac{3}{2}) = 1/8$, $P(X \geq 2) = 1 - P(X \leq 2) = 3/4$, $P(X = \frac{5}{2}) = 0$.

(iii)



Lösung zu Aufgabe 42

(i) $\sqrt[4]{0.3}$
(ii) $4/5$
(iii) $2/75$
(iv) $24 - 8e$

Lösung zu Aufgabe 43

(i) Es gilt

$$E(X) = \frac{a+b}{2} = 6,$$
$$V(X) = \frac{1}{12}(b-a)^2 = 3 \quad \Rightarrow \quad b-a = 6.$$

Dies ist ein lineares Gleichungssystem für a und b. Die Lösung lautet

$$a = 3, \qquad b = 9.$$

Somit ergibt sich

$$C = \frac{1}{b-a} = \frac{1}{6}$$

und die Wahrscheinlichkeitsdichte ist

$$f(x) = \begin{cases} 1/6 & 3 \leq x \leq 9, \\ 0 & \text{sonst.} \end{cases}$$

(ii) Wir haben

$$P(X \geq 5) = \int_5^\infty f(x)dx = \int_5^9 \frac{1}{6}dx = \frac{2}{3}.$$

Lösung zu Aufgabe 44

$P(X \geq 80) = 1 - P(X \leq 80) = 1 - F(80) = 1 - \Phi\left(\frac{80-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{80-72}{8}\right) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587 = 15.9\,\%$

Lösung zu Aufgabe 45

(i) $\approx 4.8\,\%$, (ii) $\approx 1\,\%$

Lösung zu Aufgabe 46

(i) $P(X \leq 2.44) = 0.9927$
(ii) $P(X \leq -1.16) = 0.1230$
(iii) $P(X \leq 1.923) = 0.9726$
(iv) $P(X \geq 1) = 0.1587$
(v) $P(X \geq -2.9) = 0.9981$
(vi) $P(2 \leq X \leq 10) = 0.0228$

Lösung zu Aufgabe 47

(i) $x_1 = -0.525$, $x_2 = -0.675$, $y = 0.8159$.
(ii) y ist die Wahrscheinlichkeit für einen Wert kleiner gleich 0.9.

Lösung zu Aufgabe 48

(i) 77.34 %
(ii) 0.02 %
(iii) $1.645 \cdot 12 + 91\,\mathrm{km/h} = 110.74\,\mathrm{km/h}$

Lösung zu Aufgabe 49

(i) $P(X \leq 2.44) = 0.7939$
(ii) $P(X \leq -1.16) = 0.1635$
(iii) $P(X \leq 1.923) = 0.7123$
(iv) $P(X \geq 1) = 0.4602$
(v) $P(X \geq -2.9) = 0.9678$
(vi) $P(2 \leq X \leq 10) = 0.2743$

Lösung zu Aufgabe 50

(i) 6.68 %
(ii) 0.62 %
(iii) 13.4 %
(iv) 13.4 % $\rightarrow$ 0.28 %
(v) $\sigma = 0.0116$
(vi) $c = 1.96\,\sigma$

Lösung zu Aufgabe 51

(i) $e^{-1/2} = 0.60653$,
(ii) $1 - F(n)$.

Lösung zu Aufgabe 52

(i) $c = 1.285$
(ii) $c = -1.645$
(iii) $c = 1.645$
(iv) $c = 2.575$

Lösung zu Aufgabe 53

Beide Resultate folgen durch partielle Integration.

Lösung zu Aufgabe 54

(i) $c = -1.58$
(ii) $c = 2.03$
(iii) $c = 0.8225$
(iv) $c = 1.439$

Lösung zu Aufgabe 55

(i) $x_0 = 1.645$, $y = 0.845$, $z = 0.159$
(ii) u ist die Wahrscheinlichkeit dass eine standard-normalverteilte Zufallsvariable einen Wert zwischen -1 und y annimmt.

Lösung zu Aufgabe 56

Sei X_i die Zufallsvariable welche das Gewicht des i-ten Sackes beschreibt. Die Zufallsvariable welche das Gewicht der 200 Säcke beschreibt ist $S_n = \sum_{n=1}^{200} X_i$. Wir haben

$$\mu(S_n) = n\mu = 200 \cdot 49.5 = 9900, \qquad \sigma(S_n) = \sqrt{n}\sigma = \sqrt{200} \cdot 5\sqrt{2} = 100.$$

Die standardisierte Variable

$$S_n^* = \frac{S_n - \mu(S_n)}{\sigma(S_n)} = \frac{S_n - 9900}{100}$$

ist standardnormalverteilt. Somit haben wir

$$P(S_n \geq 10000) = P(S_n^* \geq \frac{10000 - 9900}{100}) = P(S_n^* \geq 1) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587.$$

I.e. mit 15.9 % Wahrscheinlichkeit geht etwas schief.

Lösung zu Aufgabe 57

(i) Fläche unter Dichtekurve muss gleich Eins sein, i.e. $2C = 1$, i.e. $C = 1/2$.

(ii) Linearer Verlauf von $y = 0$ bis $y = 1$ zwischen $x = 19$ bis $x = 21$, links davon gleich Null, rechts davon gleich Eins.

(iii) $\mu = 20$ und

$$\sigma^2 = \int_{-\infty}^{\infty} f(x)(x-20)^2 dx = \frac{1}{2}\int_{19}^{21} (x-20)^2 dx = \frac{1}{2}\int_{-1}^{1} x^2 dx = \frac{1}{3},$$

i.e. $\sigma = 1/\sqrt{3}$.

(iv) Sei S_n die Zufallsvariable welche die Verteilung der Summe von $n = 100$ Elementen beschreibt. S_n ist (approximativ) normalverteilt. Sei S_n^* die standardisierte Zufallsvariable zu S_n. Wir haben

$$\begin{aligned} P(S_n \geq 2005) &= P\left(S_n^* \geq \frac{2005 - n\mu}{\sqrt{n}\sigma}\right) \\ &= P\left(S_n^* \geq \frac{2005 - 100 \cdot 20}{\sqrt{100} \cdot \frac{1}{\sqrt{3}}}\right) \\ &= P\left(S_n^* \geq \frac{\sqrt{3}}{2}\right) = 1 - \Phi(0.866) = 1 - 0.8078 = 0.1922. \end{aligned}$$

I.e. die Wahrscheinlichkeit dass nur 100 Elemente genügen ist ca. 19.2 %.

Lösung zu Aufgabe 58

Falls die Behauptung wahr ist, dann würde aus dem zentralen Grenzwertsatz folgen, dass der durchschnittliche Nikotingehalt (wir bezeichnen ihn mit X) approximativ normalverteilt ist, mit Mittelwert 2.2 und Standardabweichung 0.03 (Einheiten: mg). Somit wäre die Wahrscheinlichkeit dass der Durchschnitt grösser oder gleich 3.1 ist gegeben durch

$$P(X \geq 3.1) = P\left(\frac{X - 2.2}{0.03} \geq \frac{3.1 - 2.2}{0.03}\right)$$

$$\approx P(Z \geq 30)$$
$$\approx 0,$$

wobei wir mit Z die standardisierte Zufallsvariable bezeichnen.

Lösung zu Aufgabe 59

(i) Aus $0.25 = \sigma/\sqrt{n}$ folgt $n = (4/0.25)^2 = 16^2 = 256$.
(ii) Man muss n vervierfachen.

Lösung zu Aufgabe 60

(i) A2, B4, C3, D1

Lösung zu Aufgabe 61

Nullhypothese: $\mu = \mu_0$, Alternativhypothese: $\mu \neq \mu_0$. Angaben: $\mu_0 = 4$, $\sigma = 1$, $n = 36$, $\bar{x} = 4.3$, $\alpha = 0.05$. Daraus folgt

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{4.3 - 4}{1/\sqrt{36}} = 1.8.$$

Und somit ist der p-Wert

$$p = P(Z \leq -1.8) + P(Z \geq 1.8) = 2(1 - \Phi(1.8)) = 2(1 - 0.9641) = 0.0718.$$

Somit haben wir $p > \alpha$ und verwerfen die Nullhypothese nicht, i.e. aufgrund dieses statistischen Tests gibt es keinen Hinweis darauf dass die Maschine neu justiert werden soll. Der p-Wert ist die Wahrscheinlichkeit unter der Nullhypothese, dass ein Ergebnis so extrem wie das gemessene oder noch extremer auftritt.

Lösung zu Aufgabe 62

Wir haben $\mu_0 = 10$, $\sigma = 0.5$, $n = 36$, $\overline{x} = 9.7$. Wir wählen ein Signifikanzniveau von 5%, i.e. $\alpha = 0.05$. Es wird zweiseitig getestet (da Schrauben werder zu lang noch zu kurz sein sollten). Nullhypothese H_0: $\mu = \mu_0$. Alternativhypothese: H_1: $\mu \neq \mu_0$. Die Realisierung der standardisierten Zufallsvariable Z für den Mittelwert ist

$$z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{9.7 - 10}{0.5/6} = -3.6.$$

Es ergibt sich ein p-Wert von

$$p = P(Z \geq 3.6) + P(Z \leq -3.6) = 2(1 - \Phi(3.6)) = 2(1 - 0.9998) = 0.0004.$$

Da $p < \alpha$, wird die Nullhypothese verworfen. Dies ist ein Hinweis darauf dass die Maschine neu justiert werden sollte. Jedoch ist n eher klein und vielleicht macht es Sinn die Situation genauer zu untersuchen.

Lösung zu Aufgabe 63

Wir haben $n = 500$, $\overline{x} = 1985$, $\mu_0 = 2000$, $\alpha = 0.05$. Für σ verwenden wir den Wert aus der Stichprobe, da nichts anderes bekannt, i.e. wir nehmen an $\sigma = 210$. Nullhypothese H_0: $\mu = \mu_0$, Alternativhypothese: H_1: $\mu < \mu_0$, i.e. wir testen einseitig (die Frage ist ob die Kalorienmenge kleiner weniger als 2000 beträgt). Wir haben

$$z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{1985 - 2000}{210/\sqrt{500}} = -1.6.$$

Der p-Wert ist

$$p = P(Z \leq -1.6) = 1 - \Phi(1.6) = 1 - 0.9452 = 0.0548.$$

Da $p > \alpha$, wird H_0 nicht verworfen, i.e. es liegt kein signifikanter Unterschied vor.

Lösung zu Aufgabe 64

Wir haben $\mu_0 = 190$, $n = 100$, $\overline{x} = 181.52$, $\alpha = 0.05$. Für σ verwenden wir die Standardabweichung der Stichprobe (da nichts anderes bekannt), i.e. $\sigma = 40$. Nullhypothese H_0: $\mu = \mu_0$, Alternativhypothese H_1: $\mu < \mu_0$. Wir testen einseitig, da wir aus dem Zusammenhang davon ausgehen dass wir nur einen Cholesterinspiegel erwarten der niedriger ist. Wir haben

$$z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{181.52 - 190}{40/\sqrt{100}} = -2.12.$$

Der p-Wert ist

$$p = P(Z < -2.12) = 1 - \Phi(2.12) = 1 - 0.9830 = 0.017.$$

Da $p < 0.05$, verwerfen wir die Hypothese H_0, i.e. es liegt ein signifikanter Unterschied zwischen den Cholesterinspiegeln vor.