Wahrscheinlichkeitsrechnung

Dr. rer. nat Dennis Müller

November 6, 2015

Table of Contents

Einführung in die Wahrscheinlichkeitsrechnung

Einführung

- Viele Ereignisse der Realität lassen sich nicht exakt vorhersagen.
- Theorie des *Determinismus* geht davon aus das alle - auch zukünftige - Ereignisse durch Vorbedingungen eindeutig festgelegt sind.
- z.B. Bewegung der Planeten exakt vorhersagbar wenn alle Daten bekannt.
- Aber: Bei chaotischen Systemen reicht minimale Änderung der Ausgangsbedingungen für verändertes Ergebniss.
- Heisenbergsche Unschärferelation (Quantenmechanik): Ort und Impuls nicht gleichzeitig beliebig genau meßbar.
- Daher: Es gibt nicht-determinischte Ereignisse, deren Eintreten nur mit einer bestimmten Wahrscheinlichkeit vorhergesagt werden kann.

Einführung

- Für die meisten reelen Probleme spielt Quantenmechanik keine Rolle.
- Trotzdem ist es oft einfacher oder eleganter ein Problem nicht-deterministisch zu betrachten (z.B. Würfelwurf)
- Wahrscheinlichkeitstheorie liefert oft sehr gute und belastbare Aussagen (z.B. mit vernachlässigbarer oder aktzeptierter Restwahrscheinlichkeit für Irrtümer)

Einführung

- Frequentische Interpretation von Wahrscheinlichkeiten:
- Wiederholt man ein Experiment N mal (N sehr groß) und tritt dabei ein bestimmtes Ereigniss k mal auf, so sagt man, dass Ereigniss habe die *Eintrittswahrscheinlichkeit* $p = k/N$.
- Umgekehrt: Sei p die Eintrittswahrscheinlichkeit für das Ereigniss. Wird das Experiment nun N mal wiederholt, so *erwartet* man dass das Ereigniss $k = p * N$ mal auftritt.
- Beispiel: Experiment: Fairer Würfelwurf. Ereigniss: Gerade Augenzahl. Eintrittswahrscheinlichkeit: p = 0.5. Unter 1000 Würfen erwarten wir etwa 500 mal eine gerade Augenzahl zu finden.

Mengenlehre

- Wie fast alle anderen Teilgebiete der Mathematik baut auch die Wahrscheinlichkeitsrechnung auf der Mengenlehre auf.
- Dabei ist eine Menge M eine *Zusammenfassung von bestimmten, wohlunterscheidbaren Objekten zu einem Ganzen* [Georg Cantor]
- Wir sagen m ist *Element* von M ($m \in M$), wenn m in M enthalten ist. Umgekehrt auch $m \notin M$, falls m nicht in M enthalten ist.
- Die s.g. leere Menge $\emptyset$ enthält keine Element.
- Wir sagen A ist *Teilmenge* von M ($A \subset M$), wenn jedes Element aus A auch in M ist. Insbesondere ist die leere Menge Teilmenge jeder anderen Menge.

Mengenlehre

- Die s.g. *Schnittmenge* $(A \cap B)$ von A und B enthält alle Elemente die in A und B enthalten sind.
- Die s.g. *Vereinigung* $(A \cup B)$ von A und B enthält alle Elemente die in A oder B enthalten sind.
- Zwei Mengen heissen *disjunkt*, wenn ihre Schnittmenge leer ist, sie also keine gemeinsamen Elemente enthalten.
- Die s.g. *Differenzmenge* $(A \setminus B)$ enthält alle Elemente von A, die nicht auch in B enthalten sind.
- Die s.g. *Potenzmenge* $\mathcal{P}(A)$ einer Menge M enthält alle Teilmenge von A. Insbesondere ist stets $\emptyset \in \mathcal{P}(A)$ und $A \in \mathcal{P}(A)$.

Axiomatische Wahrscheinlichkeitstheorie

Axiomatische Einführung

Was Sie heute mitnehmen sollen

- Sie kennen die wichtigsten Begriffe der Wahrscheinlichkeitstheorie
- Sie kennen und verstehen die Kolmogorowschen Axiome
- Sie kennen und verstehen den Begriff der Bedingte Wahrscheinlichkeiten und Unabhängigkeit
- Sie kennen und verstehen den Satz von Bayes und den Satz der totalen Wahrscheinlichkeit

Axiomatische Einführung

- Wir brauchen einige Begriffe und Definitionen um die Wahrscheinlichkeitstheorie formal korrekt (axiomatisch) einführen zu können.

Ergebnissmenge, Ereignniss, Elementarereigniss

Die Menge aller möglichen Elementarereignisse eines Zufallsexperimentes wird *Ergebnissraum* oder *Ergebnissmenge* genannt und im allgemeinen mit Ω bezeichnet.

Sei $W \subset \Omega$, dann nennt man W *Ereigniss* (Teilmenge von Ω).

Sei $w \in \Omega$, dann nennt man w *Elementarereigniss* (Element von Ω).

Axiomatische Einführung

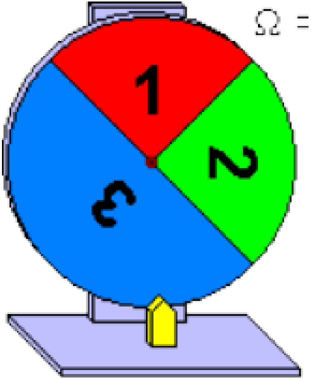
Ereignissraum, Wahrscheinlichkeitsmaß, Wahrscheinlichkeitsraum

Sei Ω die Ergebnissmenge eines Zufallsexperimentes und sei $\sum = \{W | W \subset \Omega\}$ die Menge aller Teilmengen von Ω (Potenzmenge), dann heisst $\sum$ *Ereignisraum*.

Sei P eine Abbildung $P : \sum \mapsto [0, 1]$ mit bestimmten Eigenschaften (siehe nächste Folien), dann nennt man P *Wahrscheinlichkeitsmaß*.

Das Tripel $(\Omega, \sum, P)$ wird als *Wahrscheinlichkeitsraum* bezeichnet.

Axiomatische Einführung



$$\Omega = \{1, 2, 3\}$$

$$\sum = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \Omega\}$$

$$P(\emptyset) = 0, P(\{1\}) = P(\{2\}) = 0.25, P(\{1,2\}) = P(\{3\}) = 0.5, \ldots$$

Axiomatische Einführung

Axiome von Kolmogorow

1. Für jedes Ereigniss $A \in \sum$ ist die Wahrscheinlichkeit von A eine reele Zahl zwischen 0 und 1: $0 \leq P(A) \leq 1$.
2. Das sichere Ereigniss $\Omega \in \sum$ hat die Wahrscheinlichkeit 1: $P(\Omega) = 1$.
3. Seien $A_1, \ldots, A_n$ paarweise disjunkt ($A_i \cap A_j = \emptyset$ für alle $i \neq j$), dann ist $P(A_1 \cup \cdots \cup A_n) = \sum P(A_i)$.

Axiomatische Einführung

- Sei wie im obigen Beispiel

$$\sum = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \Omega\}$$

und

A	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	Ω
P(A)	0	0.25	0.25	0.5	0.5	0.75	0.75	1

- Dann gilt z.B. das erste Axiom

$$0 \leq P(A) \leq 1 \quad \text{für alle } A \in \sum$$

- Es gilt auch das zweite Axiom

$$P(\Omega) = 1$$

Axiomatische Einführung

- Sei wie im obigen Beispiel

$$\sum = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \Omega\}$$

und

A	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	Ω
$P(A)$	0	0.25	0.25	0.5	0.5	0.75	0.75	1

- Auch das dritte Axiom gilt wegen

$$\begin{aligned} 0.5 &= P(\{1\} \cup \{2\}) = P(\{1\}) + P(\{2\}) = 0.25 + 0.25 \\ 0.75 &= P(\{1\} \cup \{3\}) = P(\{1\}) + P(\{3\}) = 0.25 + 0.5 \\ 0.75 &= P(\{2\} \cup \{3\}) = P(\{2\}) + P(\{3\}) = 0.25 + 0.25 \end{aligned}$$

- $(\Omega, \sum, P)$ bildet daher einen gültigen Wahrscheinlichkeitsraum.

Axiomatische Einführung

- Achtung: Die Kolmogorowschen Axiome sagen nichts darüber aus, ob das gewählte Wahrscheinlichkeitsmaß *sinnvoll* ist, nur ob es den formalen Anforderungen genügt.
- Sei z.B. $\sum$ wie oben, aber

A	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	Ω
P(A)	0	0.1	0.3	0.6	0.4	0.7	0.9	1

- Auch hier bildet $(\Omega, \sum, P)$ einen gültigen Wahrscheinlichkeitsraum, beschreibt jedoch ein anderes Zufallsexperiment.

Axiomatische Einführung

- Aus den Axiomen ergeben sich unmittelbar einige Folgerungen:
- Aus der Additivität der Wahrscheinlichkeiten disjkunter Ereignisse folgt, dass komplementäre Ereignisse komplementäre Wahrscheinlichkeiten haben: $P(\Omega \setminus A) = 1 - P(A)$.
- Beweis: Es ist $(\Omega \setminus A) \cup A = \Omega$ sowie $(\Omega \setminus A) \cap A = \emptyset$. Folglich nach Axiom (3): $P(\Omega \setminus A) + P(A) = P(\Omega)$ und dann nach Axiom (2): $P(\Omega \setminus A) + P(A) = 1$.

Axiomatische Einführung

- Aus den Axiomen ergeben sich unmittelbar einige Folgerungen:
- Das unmögliche Ereigniss (leere Menge) hat die Wahrscheinlichkeit 0.
- Beweis: $\emptyset$ ist das komplementäre Ereigniss zu Ω, also $P(\emptyset) = 1 - P(\Omega) = 0$.

Axiomatische Einführung

- Aus den Axiomen ergeben sich unmittelbar einige Folgerungen:

Vereinigung von nicht disjunkten Ereignissen

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Beweis: Die Menge $A \cup B$ kann als Vereinigung von drei disjunkten Mengen dargestellt werden: $A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A)$
- Nun ist $P(A) = P(A \setminus B) + P(A \cap B)$ und $P(B) = P(B \setminus A) + P(A \cap B)$.
- Addition liefert $P(A) + P(B) = P(A \setminus B) + P(A \cap B) + P(A \cap B) + P(B \setminus A) = P(A \cup B) + P(A \cap B)$
- Umstellen liefert dann die Behauptung.

Axiomatische Einführung

- Beispiel: Wir betrachten einen fairen Würfelwurf.
- Sei A das Ereigniss *gerade Augenzahl*, also $A = \{2, 4, 6\}$.
- Sei B das Ereigniss *4 oder mehr*, also $B = \{4, 5, 6\}$.
- Es ist $P(A) = P(\{2\}) + P(\{4\}) + P(\{6\}) = 3/6 = 1/2$.
- Es ist $P(B) = P(\{4\}) + P(\{5\}) + P(\{6\}) = 3/6 = 1/2$.

QUIZ

Was bedeutet hier $A \cup B$? Was ist $P(A \cup B)$?

Bedingte Wahrscheinlichkeiten

- Unter einer *bedingten Wahrscheinlichkeit* versteht man die Wahrscheinlichkeit für das Eintreten eines Ereignisses A unter der Vorraussetzung, dass das Ereigniss B bereits eingetreten ist. Dabei darf B natürlich nicht das unmögliche Ereigniss sein. Man sagt auch *A wird auf B konditioniert*.
- Man schreibt $P(A|B)$, seltener $P_B(A)$. Sprich: *P von A gegeben B*.
- Beispiel: Die Wahrscheinlichkeit aus einem Skatblatt eine Herz-Karte zu ziehen beträgt $1/4$. Also $P(Herz) = 1/4$. Wenn man aber schon weiß das die Karte rot ist, beträgt die Wahrscheinlichkeit für Herz $1/2$, also $P(Herz|Rot) = 1/2$.

Bedingte Wahrscheinlichkeiten

- Wenn B bereits eingetreten ist, schränkt dies die mögliche Ereignissmenge ein $(\Omega = B)$. Der neue Ereignissraum ist $\sum_B$.
- Es können dann nur noch solche Ereignisse eintreten die auch in B sind $(A \cap B \neq \emptyset)$.
- $P(A \cap B)$ beschreibt die Wahrscheinlichkeit das A und B gleichzeitig eintreten. Da B schon eingetreten ist, muß die Wahrscheinlichkeit neu normiert werden.

Bedingte Wahrscheinlichkeit

Es gilt $P(A|B) = \frac{P(A \cap B)}{P(B)}$

Bedingte Wahrscheinlichkeiten

- Ist $P(A|B)$ ein gültiges Wahrscheinlichkeitsmaß?
- $P(A|B)$ erfüllt die Kolmogorowschen Axiome, also ja.
- **Axiom 1**. Zu zeigen: $0 \leq P(A|B) \leq 1$ für jedes Ereigniss $A \in \sum_B$.
- Es ist $P(A \cap B) \geq 0$ und $P(B) > 0$ (weil B nicht unmöglich ist). Also ist $P(A|B) = P(A \cap B)/P(B) \geq 0$.
- Ferner sind $A \cap B$ und $B \setminus A$ disjunkt, ihre Vereinigung ist B. Also ist nach Axiom (3) $P(A \cap B) + P(B \setminus A) = P(B)$. Weil $P(B \setminus A) \geq 0$ ist, folgt $P(A \cap B) \leq P(B)$ und daher $P(A|B) \leq 1$.

Bedingte Wahrscheinlichkeiten

- Ist $P(A|B)$ ein gültiges Wahrscheinlichkeitsmaß?
- $P(A|B)$ erfüllt die Kolmogorowschen Axiome, also ja.
- **Axiom 2**. Zu zeigen: $P(B|B) = 1$.
- Es ist $P(B|B) = \frac{P(B \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$.

Bedingte Wahrscheinlichkeiten

- Ist $P(A|B)$ ein gültiges Wahrscheinlichkeitsmaß?
- $P(A|B)$ erfüllt die Kolmogorowschen Axiome, also ja.
- **Axiom 3**. Zu zeigen: $P(A_1 \cup \cdots \cup A_n|B) = \sum P(A_i|B)$ für paarweise disjunkte A_i.
- Es ist

$$
\begin{aligned}
P(A_1 \cup \cdots \cup A_n|B) &= \frac{P((A_1 \cup \cdots \cup A_n) \cap B)}{P(B)} \\
&= \frac{P((A_1 \cap B) \cup \cdots \cup (A_n \cap B))}{P(B)} \\
&= \frac{P(A_1 \cap B) + \cdots + P(A_n \cap B)}{P(B)} \\
&= P(A_1|B) + \cdots + P(A_n|B)
\end{aligned}
$$

Bedingte Wahrscheinlichkeiten

- Wir haben gezeigt: Sei $(\Omega, \sum, P)$ ein Wahrscheinlichkeitsraum, d.h. P erfüllt die Kolmogorowschen Axiome auf $\sum$. Dann ist für ein Ereigniss $B \in \sum$ das Tripel $(B, \sum_B, P_B)$ mit $P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)}$ ebenfalls ein Wahrscheinlichkeitsraum.
- Das bedeutet: Alles was wir hergeleitet haben und auf den folgenden Folien herleiten werden gilt uneingeschränkt genauso, wenn der Wahrscheinlichkeitsraum auf ein (nicht unmögliches) Ereigniss aus $\sum$ konditioniert wird.
- Beispiel: Wir wissen schon

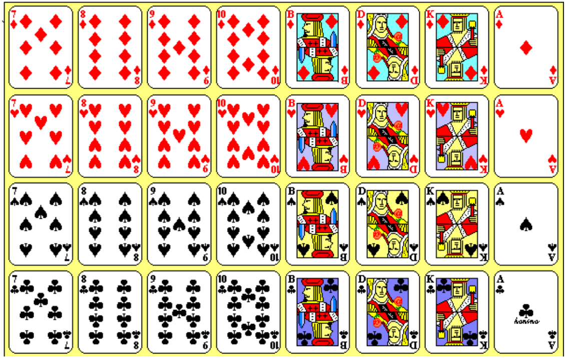
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

. Konditionieren wir auf $C \neq \emptyset$, wissen wir sofort auch

$$P(A \cup B|C) = P(A|C) + P(B|C) - P(A \cap B|C)$$

Bedingte Wahrscheinlichkeiten

- Beispiel: Sei A das Ereignis *Herz König* beim ziehen einer Karte aus einem Skatspiel. Sei B das Ereigniss *rote Karte* und C das Ereigniss *Bildkarte*.



QUIZ

Was ist $P(A|B)$, $P(B|A)$, $P(B|C)$, $P(B \cap C)$, $P(B \cup C)$, $P(A|B \cap C)$, $P(A|B \cup C)$

Verbundwahrscheinlichkeiten

- Es gilt

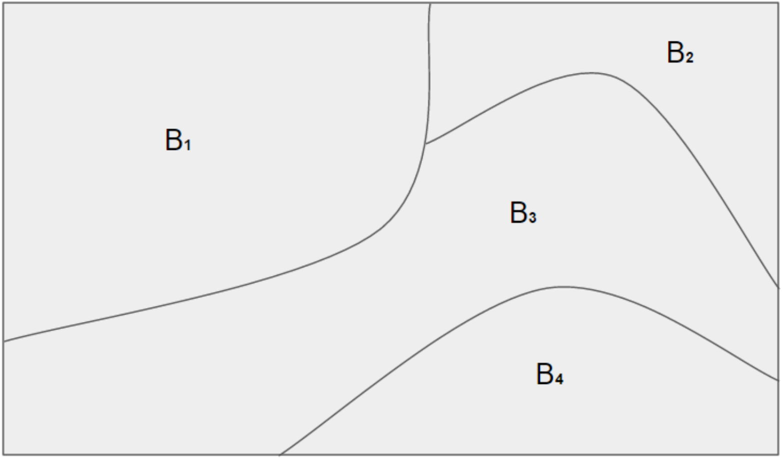
Verbundwahrscheinlichkeit

$$P(A \cap B) = P(B|A) \cdot P(A) = P(A|B) \cdot P(B)$$

- Beweis: Direkt aus der Definition der bedingten Wahrscheinlichkeit folgt $P(A|B) = \frac{P(A \cap B)}{P(B)}$ also $P(A \cap B) = P(A|B) \cdot P(B)$. Genau für die andere Seite.

Satz der totalen Wahrscheinlichkeit

- Eine Familie von Mengen $B_1, \ldots, B_n$ heisst *Partition* von Ω, wenn die B_i paarweise disjunkt sind ($B_i \cap B_j = \emptyset$ für $i \neq j$) und Ihre Vereingung wieder Ω ergibt, d.h. $\bigcup_i B_i = \Omega$.



Satz der totalen Wahrscheinlichkeit

Satz der totalen Wahrscheinlichkeit

Sei $B_1, \ldots, B_n$ eine *Partition* von Ω mit $P(B_i) > 0$ für alle i. Dann gilt

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

- Beweis: Es sind $(A \cap B_1), \ldots, (A \cap B_n)$ paarweise disjunkt (warum?) mit $\bigcup_i (A \cap B_i) = A$ (warum?).
- Nach Axiom 3 folgt $P(A) = \sum_{i=1}^{n} P(A \cap B_i)$ und mit $P(A \cap B_i) = P(A|B_i) \cdot P(B_i)$ die Behauptung.

Satz der totalen Wahrscheinlichkeit

Satz der totalen Wahrscheinlichkeit - Spezialfall

Sei $B \neq \emptyset$ und $\overline{B} = \Omega \setminus B$. Dann ist $\{B, \overline{B}\}$ eine Partition von Ω und es gilt

$$P(A) = P(A|B) \cdot P(B) + P(A|\overline{B}) \cdot P(\overline{B})$$

- B und $\overline{B}$ sind komplementäre Ereignisse.

Satz der totalen Wahrscheinlichkeit

Satz der totalen Wahrscheinlichkeit - Alternative Formulierung

Sei $B_1, \ldots, B_n$ eine *Partition* von Ω mit $P(B_i) > 0$ für alle i. Dann gilt

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i)$$

- Dies folgt trivial aus

$$P(A|B_i) \cdot P(B_i) = \frac{P(A \cap B_i)}{P(B_i)} \cdot P(B_i)$$

Satz der totalen Wahrscheinlichkeit

- Beispiel: Für ein Navigationsystem möchten Sie die Wahrscheinlichkeit berechnen, das der Verkehr sich auf einer bestimmten Strasse stauen wird (S).
- Sie wissen aus einer statistischer Erhebung in der Vergangenheit: Wenn es regnet (R) staut sich der Verker mit einer Wahrscheinlichkeit von 80%, also $P(S|R) = 0.8$. Wenn es nicht regnet staut sich der Verkehr nur mit einer Wahrscheinlichkeit von 30%, also $P(S|\overline{R}) = 0.3$. Die Wettervorhersage für den Reisetag sagt eine Regenwahrscheinlichkeit von $P(R) = 0.6$ vorraus.
- Damit ist die Stauwahrscheinlichkeit $P(S) = P(S|R) \cdot P(R) + P(S|\overline{R}) \cdot P(\overline{R}) = 0.8 \cdot 0.6 + 0.3 \cdot 0.4 = 0.6$

Satz von Bayes

- Es gilt

Satz von Bayes

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

- Beweis: Aus $P(A|B) = \frac{P(A \cap B)}{P(B)}$ folgt mit $P(A \cap B) = P(B|A) \cdot P(A)$ die Behauptung.

Satz von Bayes - Beispiel

- Sie entwickeln einen Notbremsassitenten für Fahrzeuge. Sie wissen das das System im Falle einer kritischen Situation (C) mit 99.9% Wahrscheinlichkeit eine Bremsung anfordert (B).
- Weiterhin wissen Sie das das System mit einer Wahrscheinlichkeit von 0.0005% eine Bremsung anfordert wenn keine kritische)Situation vorliegt.
- Aus Verkehrsstatistiken wissen sie, das nur 0.0001% aller Fahrsituationen kritisch sind.
- Ihr System fordert nun eine Bremsung an. Mit welcher Wahrscheinlichkeit ist die Fahrsituation kritisch?

Satz von Bayes - Beispiel

- Wir wissen

$$\begin{aligned} P(B|C) &= 0.999 \\ P(B|\overline{C}) &= 0.000005 \\ P(C) &= 0.000001 \end{aligned}$$

- Nun gilt nach Satz von Bayes $P(C|B) = \frac{P(B|C) \cdot P(C)}{P(B)}$
- Ferner gilt nach dem Satz der totalen Wahrscheinlichkeit $P(B) = P(B|C) \cdot P(C) + P(B|\overline{C}) \cdot P(\overline{C}) = 0.000005998995$
- Somit ist $P(C|B) = 16.653\%$

Satz von Bayes - Beispiel

- Sie spielen in einer Samstag-Abend Show mit, bei der Sie eines von drei Toren auswählen müssen. Hinter einem Tor befindet sich der Hauptgewinn, hinter den anderen beiden nur Nieten.
- Nachdem Sie gewählt haben öffnet der Moderator ein Tor mit Niete (das gibt es immer!) und bietet Ihnen an, Ihre Entscheidung nochmal zu überdenken und das andere Tor zu wählen.
- Sollten Sie wechseln oder bei Ihrer Wahl bleiben? Oder ist das egal?

Sie sollten wechseln, weil das Ihre Gewinnchancen verdoppelt!

Satz von Bayes - Beispiel

- O.B.D.A. haben Sie Tor 1 gewählt und der Moderator daraufhin Tor 3 geöffnet. Sei ferner G_i das Ereigniss "*Der Gewinn ist hinter Tor i*" und M_i das Ereigniss "*Der Moderator öffnet Tor i*".
- Dann gilt

$$\begin{aligned} P(G_1) = P(G_2) = P(G_3) &= 1/3 \\ P(M_3|G_1) &= 1/2 \\ P(M_3|G_2) &= 1 \\ P(M_3|G_3) &= 0 \end{aligned}$$

- Nach dem Satz von Bayes und dem Satz der totalen Wahrscheinlichkeit folgt damit

$$\begin{aligned} & P(G_2|M_3) \\ = & \frac{P(M_3|G_2)P(G_2)}{P(M_3|G_1)P(G_1) + P(M_3|G_2)P(G_2) + P(M_3|G_3)P(G_3)} \\ = & 2/3 \end{aligned}$$

Unabhängigkeit

Unabhängigkeit von Ereignissen

Zwei Ereignisse A und B sind unabhängig, wenn gilt

$$P(A \cap B) = P(A) \cdot P(B)$$

- Dies bedeutet äquivalent: $P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A)$, d.h. die Wahrscheinlichkeit für das eintreten von A ändert sich nicht wenn man weiß das B eingetreten ist.

Unabhängigkeit

- Beispiel: Sei A das Ereigniss *Herz* beim ziehen einer Karte aus einem Skatspiel. Sei B das Ereigniss *Bild*.
- Es gilt $P(A) = 1/4$ und $P(B) = 3/8$ sowie $P(A \cap B) = 3/32$ und weil $P(A) \cdot P(B) = P(A \cap B)$ gilt, sind A und B unabhängig.
- Sei C das Ereigniss *Rot*. Dann ist $P(C) = 1/2$ und $P(A \cap C) = 1/4 \neq P(A) \cdot P(C) = 1/8$, also sind A und C abhängig.

Unabhängigkeit

- Beispiel: Beim fairen Münzwurf gilt $P(Kopf) = 1/2$. Wenn man annimmt das mehrere Münzwürfe unabhängig voneinander sind, dann gilt

$$P(5\ mal\ Kopf) = (1/2) \cdot (1/2) \cdot (1/2) \cdot (1/2) \cdot (1/2) = 1/32$$

$$\begin{aligned} P(2\ Kopf\ \ bei\ 3\ Wuerfen) &= P(ZKK) + P(KZK) + P(KKZ) \\ &= 1/8 + 1/8 + 1/8 = 3/8 \end{aligned}$$

$$P(Zwei\ gleiche\ Wuerfe) = P(KK) + P(ZZ) = 1/4 + 1/4 = 1/2$$

Unabhängigkeit

- Wir werfen eine Münze 10 mal und versuchen, nur Kopf zu bekommen. Wie hoch ist die Wahrscheinlichkeit dafür das wir dies mindestens einmal schaffen, wenn wir das Experiment 1000 mal wiederholen?
- Die Wahrscheinlichkeit für 10 mal Kopf hintereinander beträgt $(1/2)^{10} \approx 0.01\%$. Das Gegenereigniss (unter 10 Würfen mindestens einmal Zahl) hat also die Wahrscheinlichkeit $1 - (1/2)^{10} \approx 99,9\%$.
- Die Wahrscheinlichkeit dies 1000 mal in Folge zu schaffen beträgt also $(1 - (1/2)^{10})^{1000} \approx 37.64\%$.
- Das Gegenereigniss (Bei 1000 Wiederholungen mindestens einmal 10 Köpfe) hat also die Wahrscheinlichkeit $1 - 37.64\% = 62.36\%$.

Zusammenfassung

Was Sie heute gelernt haben

- Sie wissen, das die Wahrscheinlichkeiten paarweiser disjunkter Ereignisse additiv sind. (Axiom 3)
- Sie wissen, wie man die Wahrscheinlichkeit der Vereinigung zweier Ereignisse berechnen kann.
- Sie wissen, wie sich Wahrscheinlichkeiten verändern wenn konditionale Ereignisse eingetreten sind.
- Sie wissen, was eine Partition ist und wie dies helfen kann, Wahrscheinlichkeiten zu berechnen.
- Sie wissen, wie bedingte Wahrscheinlichkeiten "vertauscht" werden können um Schlussfolgerungen zu ziehen.
- Sie wissen, was es heisst wenn zwei Ereignisse unabhängig sind.

Zusammenfassung

Wichtige Formeln / Sätze

$$
\begin{aligned}
P(A_1 \cup \dots \cup A_n) &= \sum P(A_i) \quad \text{für paarweise disjunkte } A_i \\
P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\
P(A|B) &= \frac{P(A \cap B)}{P(B)} \\
P(A \cap B) &= P(A|B) \cdot P(B) = P(B|A) \cdot P(A) \\
P(A) &= \sum P(A|B_i) \cdot P(B_i) \quad \text{falls } B_1, \dots, B_n \text{ Partition} \\
P(A|B) &= \frac{P(B|A) \cdot P(A)}{P(B)} \\
P(A \cap B) &= P(A) \cdot P(B) \quad \text{wenn } A \text{ und } B \text{ unabhängig}
\end{aligned}
$$

Diskrete Wahrscheinlichkeitsräume

Was Sie heute mitnehmen sollen

- Sie kennen und verstehen den Begriff der Zufallsvariable und Massefunktion
- Sie kennen und verstehen den Begriff Erwartungswert und Varianz
- Sie kennen und die Binomialverteilung

Diskrete Wahrscheinlichkeitsräume

Diskreter Wahrscheinlichkeitsraum

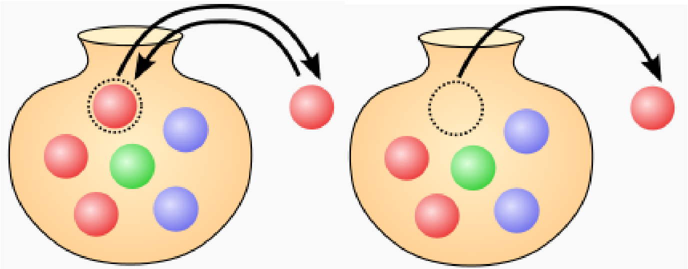
Ein Wahrscheinlichkeitsraum $(\Omega, \sum, P)$ heisst *diskreter Wahrscheinlichkeitsraum*, wenn die Ergebnismenge Ω endlich oder abzählbar unendlich ist.

Diskrete Wahrscheinlichkeitsräume

- Beispiel: Wie hoch ist bei einer Gruppe von 23 Kindern die Wahrscheinlichkeit dafür, das mindestens zwei am selben Tag Geburtstag haben?
- Wir betrachten das Gegenereigniss: "Alle Kinder haben an verschiedenen Tagen Geburtstag".
- Es gibt $365^{23} = 8.56 \cdot 10^{58}$ mögliche Geburstagsvarianten.
- Wann das erste Kind Geburtstag hat ist egal. Für das zweite gibt es dann noch 364 "günstige" Tage, für das dritte 363, und so weiter...
- Also gibt es $365 \cdot 364 \cdot \dots \cdot 344 \cdot 343 = 4.22 \cdot 10^{58}$ günstige Kombinationen. Die Wahrscheinlichkeit dafür beträgt also $p = 49.27\%$.
- Damit beträgt die gesuchte Wahrscheinlichkeit dafür, dass mindestens zwei Kinder haben am selben Tag Geburtstag haben, $p = 51.73\%$.

Kombinatorik - Urnenmodelle

- Man unterscheidet zwei verschiedene Modelle.
- Beim *Ziehen mit zurücklegen* wird die gezogene Kugel gedanklich zurückgelegt und kann daher erneut gezogen werden.
- Beim *Ziehen ohne zurücklegen* wird die Kugel nicht zurückgelegt. Die Zahl der Kugeln verändert sich mit jedem Zug.



Kombinatorik - Urnenmodelle

- Ziehen ohne Reihenfolge: In einer Urne sind 10 Bälle mit den Ziffern von 0 bis 9.

QUIZ

Wieviele Möglichkeiten gibt es, aus der Urne 5 Kugeln **ohne** zurücklegen zu ziehen, wenn die Reihenfolge keine Rolle spielt.

- Mit Reihenfolge gäbe es $10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 = 30.240$ Möglichkeiten.
- Um die erste Zahl anzuordnen gibt es 5 Plätze. Für die zweite Zahl gibt es noch 4, für die dritte 3 und so weiter. Also gibt es $5! = 120$ verschiedene Reihenfolgen.
- Da alle diese Reihenfolgen äquivalent sein sollen, gibt es nur noch $30.240/120 = 252$ Möglichkeiten.

Kombinatorik - Urnenmodelle

- Allgemein: In einer Urne seien n (unterscheidbare) Kugeln. Wir ziehen k ohne zurücklegen und ohne Berücksichtigung der Reihenfolge.
- Dann gibt es

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

  Möglichkeiten.
- Dabei heisst $\binom{n}{k}$ (lies: n über k) *binomialkoeffizient*.

Kombinatorik - Urnenmodelle

- Beispiel: Aus einem Pokerspiel (52 Karten) werden fünf Karten (ohne zurücklegen) gezogen. Wie hoch ist die Wahrscheinlichkeit fünf Karten der gleichen Farbe zu bekommen?
- Es sind $\binom{52}{5} = 2.598.960$ *Hände* möglich (ziehen ohne zurücklegen).
- Davon sind $4 \cdot \binom{13}{5} = 5.148$ günstig. Also ist die gesuchte Wahrscheinlichkeit $p = 0.2\%$.

Kombinatorik - Urnenmodelle

- Beispiel: Aus einem Pokerspiel (52 Karten) werden sieben Karten (ohne zurücklegen) gezogen. Wie hoch ist die Wahrscheinlichkeit dafür mindestens fünf Karten gleicher Farbe zu haben?
- Es sind $\begin{pmatrix} 52 \\ 7 \end{pmatrix} = 133.784.560$ *Hände* möglich (ziehen ohne zurücklegen).
- Es gibt $\begin{pmatrix} 13 \\ 5 \end{pmatrix} \cdot 4 \cdot 39 \cdot 38$ Kombinationen mit genau fünf Karten gleicher Farbe.
- Es gibt $\begin{pmatrix} 13 \\ 6 \end{pmatrix} \cdot 4 \cdot 39$ Kombinationen mit genau sechs Karten gleicher Farbe.
- Es gibt $\begin{pmatrix} 13 \\ 7 \end{pmatrix} \cdot 4$ Kombinationen mit genau sieben Karten gleicher Farbe.
- Damit gibt es insgesamt 7.903.896 günstige Möglichkeiten. Die Wahrscheinlichkeit beträgt damit $p = 5.91\%$.

Zufallsvariablen

Zufallsvariablen

Zufallsvariablen

Sei $(\Omega, \sum, P)$ ein Wahrscheinlichkeitsraum. Sei ferner $X : \Omega \mapsto \Omega^*$ mit $\Omega^* \subset \mathbb{R}$ eine Abbildung von der Ereignismenge in die reelen Zahlen. Dann heisst X *Zufallsvariable* wenn X die Eigenschaft

$$\forall x \in \Omega^* : \{\omega | X(\omega) \leq x\} \in \sum$$

erfüllt.

- X bildet Elementarereignisse auf reelen Zahlen ab. Damit ist der Wert, den X annimmt zufällig.
- Die Bedingung besagt, das die Menge aller Elementarereignisse, für welche X einen Wert kleiner oder gleich x annimmt, stets ein gültiges Ereigniss sein muß, und zwar für alle x.
- Wenn Ω endlich ist und für $\sum$ die Potenzmenge gewählt wird, ist dies immer erfüllt.

Zufallsvariablen

- Beispiel: Wir werfen zwei faire Würfel. Als Ereignissmenge wählen wir $\Omega = \{(1,1);(1,2);\ldots;(6,6)\}$ die Menge aller möglichen Tupel und für $\sum$ wählen wir die Potenzmenge. Dann sind alle Elementarereignisse gleich wahrscheinlich und es gilt $P(\omega) = 1/36$ für alle $\omega \in \Omega$.
- Wir definieren die Zufallsvariable X als Summe der Augenzahlen. Dann ist $\Omega^* = \{2,3,4,5,6,7,8,9,10,11,12\} \subset \mathbb{R}$.
- Ferner ist $P(X < 4) = P(\{(1,1);(1,2);(2,1)\}) = 3/36 = 1/12$
- X kann die Werte zwischen 2 und 12 annehmen. Dabei gilt die folgende Wahrscheinlichkeitstabelle

x	2	3	4	5	6	7	8	9	10	11	12
P(X=x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Zufallsvariablen

Verteilung einer Zufallsvariable

Gegeben sei eine Zufallsvariable X auf dem Wahrscheinlichkeitsraum $(\Omega, \sum, P)$, welche nach $\Omega^* \subset \mathbb{R}$ abbildet. Dann heisst

$$P_X(A^*) = P(X^{-1}(A^*)) \text{ für alle } A^*$$

Verteilung der Zufallsvariable.

- $X^{-1}(A)$ ist das *Urbild* von A^* unter X, also das Ereigniss

$$\{\omega \in \Omega | X(\omega) \in A^*\}$$

- P_X definiert ein Wahrscheinlichkeitsmaß auf $(\Omega^*, \sum^*, P_X)$, wobei hier nicht näher erläutert werden soll was $\sum *$ eigentlich ist.

Zufallsvariablen

- Wir wollen zeigen, dass P_X tatsächlich ein Wahrscheinlichkeitsmaß ist. Dazu müssen wir zeigen, das P_X die drei Kolmogorowschen Axiome erfüllt.
- zu (1): Wir wollen zeigen das $0 \leq P_X(A^*) \leq 1$ für alle A^*.
- Nach Definition folgt $0 \leq P(X^{-1}(A^*)) \leq 1$.
- Da P Wahrscheinlichkeitsmaß auf Ω ist, folgt die Behauptung.

Zufallsvariablen

- Wir wollen zeigen, dass P_X tatsächlich ein Wahrscheinlichkeitsmaß ist. Dazu müssen wir zeigen, das P_X die drei Kolmogorowschen Axiome erfüllt.
- zu (2): Wir wollen zeigen das $P_X(\Omega^*) = 1$ gilt.
- Das Urbild von Ω^* ist $\{\omega \in \Omega | X(\omega) \in \Omega^*\} = \Omega$
- Damit ist $P_X(\Omega^*) = P(\Omega) = 1$ nach Vorraussetung.

Zufallsvariablen

- Wir wollen zeigen, dass P_X tatsächlich ein Wahrscheinlichkeitsmaß ist. Dazu müssen wir zeigen, das P_X die drei Kolmogorowschen Axiome erfüllt.
- zu (3): Wir wollen zeigen das für paarweise disjunkte $A_1^*, \dots, A_n^*$ gilt
$$P_X(A_1^* \cup \dots \cup A_n^*) = \sum P_X(A_i^*)$$
- Das Urbild zu $A_1^* \cup \dots A_n^*$ ist

$$\{\omega \in \Omega | X(\omega) \in A_1^* \vee \dots \vee X(\omega \in A_n^*\}$$
$$=$$
$$\{\omega \in \Omega | X(\omega) \in A_1^*\} \cup \dots \cup \{\omega \in \Omega | X(\omega) \in A_n^*\}$$

Zufallsvariablen

- Wir wollen zeigen, dass P_X tatsächlich ein Wahrscheinlichkeitsmaß ist. Dazu müssen wir zeigen, das P_X die drei Kolmogorowschen Axiome erfüllt.
- zu (3): Wir wollen zeigen das für paarweise disjunkte $A_1^*, \ldots, A_n^*$ gilt
$$P_X(A_1^* \cup \cdots \cup A_n^*) = \sum P_X(A_i^*)$$
- Die einzelnen $X^{-1}(A_i^*) = \{\omega \in \Omega | X(\omega) \in A_i^*\}$ sind paarweise disjunkt.
- Beweis: Sei $\omega \in X^{-1}(A_i^*)$, d.h. $X(\omega) \in A_i^*$, also X bildet ω auf ein Element in A_i^* ab. Da die A_i^* nach Vorraussetzung paarweise disjunkt sind, ist dieses Element $X(\omega)$ in keiner anderen Menge A_j^* $(i \neq j)$ enthalten. Also ist ω in keiner anderen Menge $X^{-1}(A_j^*)$ $(i \neq j)$ enthalten.

Zufallsvariablen

- Wir wollen zeigen, dass P_X tatsächlich ein Wahrscheinlichkeitsmaß ist. Dazu müssen wir zeigen, das P_X die drei Kolmogorowschen Axiome erfüllt.
- zu (3): Wir wollen zeigen das für paarweise disjunkte $A_1^*, \dots, A_n^*$ gilt

$$P_X(A_1^* \cup \dots \cup A_n^*) = \sum P_X(A_i^*)$$

- Damit gilt weiter

$$\begin{aligned} P_X(A_1^* \cup \dots \cup A_n^*) &= \sum P\left(\{\omega \in \Omega | X(\omega) \in A_i^*\}\right) \\ &= \sum P_X(A_i^*) \end{aligned}$$

Zufallsvariablen

QUIZ

Sei wieder X die Summe der Augenzahlen bei zwei fairen Würfelwürfen, also wie oben

$$\Omega = \{(1,1);(1,2);\dots;(6,6)\}$$

und

$$\Omega^* = \{2,3,\dots,11,12\}$$

.

- Was ist $X(\{(1,1)\})$ und $X(\{(3,4)\})$?
- $X(\{(1,1)\}) = 2$, $X(\{(3,4)\}) = 7$
- Was ist $X^{-1}(\{3\})$, $X^{-1}(\{11,12\})$
- $X^{-1}(\{3\}) = \{(1,2);(2,1)\}$
- $X^{-1}(\{5,6,7\}) = \{(5,6);(6,5);(6,6)\}$

Zufallsvariablen

QUIZ

Sei nun Y die kleinere Augenzahl bei zwei fairen Würfelwürfen.

- Was ist nun $Y(\{(1,1)\})$ und $Y(\{(3,4)\})$?
- $Y(\{(1,1)\}) = 1,\ Y(\{(3,4)\}) = 3$
- Was ist $Y^{-1}(\{3\})$
- $Y^{-1}(\{3\}) = \{(3,3); (3,4); (3,5); (3,6); (4,3); (5,3); (6,3)\}$

Zufallsvariablen

- Das Wahrscheinlichkeitsmaß P_X ordnet jedem Ereignis A^* aus dem Ereignisraum $\sum^*$ die gleiche Wahrscheinlichkeit zu wie das Wahrscheinlichmaß P demjenigen Ereigniss A aus $\sum$ zuordnet, welches von X auf A^* abgebildet wird.
- Da $(\Omega^*, \sum^*, P_X)$ einen Wahrscheinlichkeitsraum bilden, gelten alle Sätze und Eigenschaften die wir bereits bewiesen haben uneingeschränkt weiter.
- Beispielsweise gilt mit dem Satz der totalen Wahrscheinlichkeit:

$$\begin{aligned} P(Y \geq 3) &= P(Y \geq 3 | X \leq 8) \cdot P(X \leq 8) \\ &+ P(Y \geq 3 | X > 8) \cdot P(X > 8) \end{aligned}$$

- Es gilt natürlich auch der Satz von Bayes

$$P(X = 5 | Y = 1) = \frac{P(Y = 1 | X = 5) \cdot P(X = 5)}{P(Y = 1)}$$

Zufallsvariablen

Massenfunktion

Sei X eine Zufallsvariable. Dann heisst

$$p_X(x) = P(X = x)$$

Massenfunktion von X.

- Die Massenfunktion ordnet jeder reelen Zahl x die Wahrscheinlichkeit dafür zu, dass X den Wert x annimmt.
- Es gilt (warum?)
$$\sum_{x \in \Omega^*} p_X(x) = 1$$
- Mit einem Interval $I = [a, b]$ mit $a, b \in \mathbb{R}$ und $a \leq b$ gilt ferner
$$P(X \in I) = \sum_{x \in I} p(x)$$

Zufallsvariablen

Verteilungsfunktion

Sei X eine Zufallsvariable. Dann heisst

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

Verteilungsfunktion von X.

- Die Verteilungsfunktion ordnet jeder reelen Zahl x die Wahrscheinlichkeit dafür zu, dass X einen Wert kleiner oder gleich x annimmt.

Zufallsvariablen

- Beispiel: Wir werfen eine fairen Münze 8 mal. Die Zufallsvariable X beschreibe die Anzahl "Kopf" unter diesen 8 Würfen. Dann gilt

$$
\begin{aligned}
P(X=0) &= \binom{8}{0} 0.5^8 = 0.4\% \\
P(X=1) &= \binom{8}{1} 0.5^8 = 3.1\% \\
P(X=2) &= \binom{8}{2} 0.5^8 = 10.9\%
\end{aligned}
$$

und so weiter. Damit ist die Massen- und Verteilungsfunktion gegeben durch

X	0	1	2	3	4	5	6	7	8
p_X	0.4%	3.1%	10.9%	21.9%	27.3%	21.9%	10.9%	3.1%	0.4%
F_X	0.4%	3.5%	14.4%	36,3%	63.6%	85.5%	96.5%	99.6%	100%

Zufallsvariablen

- Beispiel: Sei Y wieder die kleinere der beiden Augenzahlen bei zwei fairen Würfelwürfen. Es ist

X	1	2	3	4	5	6
p_X	30.5%	25%	19.4%	13.8%	8.3%	2.7%
F_X	30.5%	55.5%	75.0%	88.8%	97.2%	100.0%

Zufallsvariablen

Erwartungswert

Sei X eine Zufallsvariable. Dann heisst

$$E[X] = \sum_{x \in \Omega^*} x \cdot P(X = x)$$

Erwartungswert von X.

- Der Erwartungswert gibt an, welches "Ergebnis" man im Mittel erwartet.
- Dabei kann es durchaus vorkommen das der Erwartungswert nicht Element von Ω^* ist, das Zufallsexperiment diesen also nie direkt erzeugen kann.
- Es gibt auch diskrete Verteilungen für die der Erwartungswert nicht definiert ist.

Zufallsvariablen

- Beispiel: Sei X die Augenzahl bei einem fairen Würfelwurf. Dann ist $p_X(i) = 1/6$ für $i = 1, \ldots, 6$ und es gilt

$$E[X] = \sum_{i=1}^{6} i \cdot p_X(i) = 3.5$$

- Dabei kann 3.5 nie als Augenzahl geworfen werden.
- St. Petersburger Spiel: Man werfe eine Münze. Zeigt sie Kopf, erhält man 2 Euro und das Spiel ist beendet, zeigt sie Zahl, darf man nochmals werfen. Wirft man nun Kopf, erhält man 4 Euro und das Spiel ist beendet. Wirft man Zahl, darf man wieder werfen und erhält die Chance auf 8 Euro, und so weiter.
- Der Erwartungswert des St. Petersburger Spiels ist

$$E[X] = 2 \cdot (1/2) + 4 \cdot (1/4) + 8 \cdot (1/8) + \cdots = 1 + 1 + 1 + \cdots = \infty$$

Zufallsvariablen

- Sei X eine Zufallsvariable mit folgender Massefunktion

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- X beschreibt die Anzahl *Erfolge* bei n unabhängigen Wiederholungen eines Zufallsexperimentes mit Erfolgswahrscheinlichkeit p.
- Nun ist $E[X] = n \cdot p$.
- Beweis: Wir betrachten zunächst die Gleichung

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

- Nun leiten wir auf beiden Seiten nach a ab:

$$n(a+b)^{n-1} = \sum_{k=0}^{n} k \binom{n}{k} a^{k-1} b^{n-k}$$

Zufallsvariablen

- Wir haben gezeigt:

$$n(a+b)^{n-1} = \sum_{k=0}^{n} k \begin{pmatrix} n \\ k \end{pmatrix} a^{k-1} b^{n-k}$$

- Multiplikation mit a liefert dann

$$n \cdot a(a+b)^{n-1} = \sum_{k=0}^{n} k \begin{pmatrix} n \\ k \end{pmatrix} a^{k} b^{n-k}$$

- Substituieren wir nun $a = p$ und $b = 1 - p$ erhalten wir die Behauptung

$$n \cdot p = \sum_{k=0}^{n} k p_X(k) = E[X]$$

Zufallsvariablen

- Der Erwartungswert ist linear, es gilt also für beliebige, nicht notwendigerweise unabhängig Zufallsvariablen X_1, X_2 sowie zwei Konstanten $a, b \in \mathbb{R}$

$$E[aX_1 + bX_2] = aE[X_1] + bE[X_2]$$

- Beweis: Übung!
- Falls X_1, X_2 **unabhängig** (und nur dann!) sind, gilt darüber hinaus auch
$$E[X_1 \cdot X_2] = E[X_1] \cdot E[X_2]$$
- Beweis: Übung!

Zufallsvariablen

- Beispiel: Sei X_1 die Augenzahl beim ersten Wurf und X_2 die Augenzahl beim zweiten Wurf eines fairen Würfels.
- Dann ist $X = X_1 + X_2$ wieder die Summe beider Augenzahlen und es gilt mit Hilfe der Linearität

$$E[X] = E[X_1 + X_2] = E[X_1] + E[X_2] = 3.5 + 3.5 = 7$$

- Da die Würfe auch unabhängig sind gilt ferner für das Produkt der beiden Augenzahlen

$$E[X_1 \cdot X_2] = E[X_1] \cdot E[X_2] = 12.25$$

Zufallsvariablen

- Beispiel: Wir werfen eine *unfaire* Münze, welche mit Wahrscheinlichkeit p Kopf und mit Wahrscheinlichkeit $1-p$ Zahl zeigt. Die Zufallsvariable X sei 1 g.d.w. Kopf fällt und 0 g.d.w. Zahl fällt.
- Damit ist $p_X(1) = p$ und $p_X(0) = 0$. Ferner ist $E[X] = p$.
- Wir wiederholen das Experiment nun n mal. X_i seien die einzelnen Ausgänge während X die Summe (Anzahl *Kopf*) angibt, also $X = \sum_{i=1}^{n} X_i$.
- Dann folgt aus der Linearität sofort

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = n \cdot p$$

- X entspricht hier wie oben der Anzahl "Erfolge" bei n unabhängigen Wiederholungen eines Zufallsexperimentes mit Erfolgswahrscheinlichkeit p.

Zufallsvariablen

Varianz

Sei X eine Zufallsvariable. Dann heisst

$$V[X] = \sum_{x \in \Omega^*} (x - E[X])^2 \cdot P(X = x)$$

Varianz von X.

- Die Varianz gibt an, wie stark das "Ergebnis" um den Mittelwert streut.
- Es gilt sowohl
  $$V[X] = E\left[(X - E[X])^2\right]$$
  als auch
  $$V[X] = E[X^2] - E[X]^2$$
- Dieser Zusammenhang heisst auch *Verschiebungssatz*

Zufallsvariablen

- Beweis: Direkt nach Definition des Erwartungswertes folgt

$$
\begin{aligned}
E\left[(X-E[X])^2\right] &= \sum_{x\in\Omega^*}(x-E[X])^2\cdot P(X=x)\\
&= \sum_{x\in\Omega^*} x^2\cdot P(X=x)\\
&- 2E[X]\sum_{x\in\Omega^*} x\cdot P(X=x)\\
&+ E[X]^2\sum_{x\in\Omega^*} P(X=x)\\
&= E[X^2]-2E[X]E[X]+E[X]^2\\
&= E[X^2]-E[X]^2\\
&= V[X]
\end{aligned}
$$

Zufallsvariablen

- Sei X eine Zufallsvariale. Für reelen Zahlen a und b gilt dann

$$V[aX+b] = a^2 V[X]$$

- Beweis: Direkt aus dem Verschiebungssatz folgt

$$\begin{aligned} V[aX+b] &= E\left[(aX+b) - E[aX+b])^2\right] \\ &= E\left[(aX+b-b-aE[X])^2\right] \\ &= E\left[a^2(X-E[X])^2\right] \\ &= a^2 V[X] \end{aligned}$$

- Insbesondere gilt mit $a=-1$ und $b=0$

$$V[-X] = V[X]$$

Zufallsvariablen

- Beispiel: Sei X die Augenzahl beim fairen Würfelwurf
- Wir wissen bereits $P(X = i) = 1/6$ für $i = 1, \ldots, 6$ und

$$E[X] = 3.5$$

- Ferner gilt

$$E[X^2] = \sum_{i=1}^{6} i^2 \cdot (1/6) = 15.16$$

- und damit gilt

$$V[X] = E[X^2] - E[X]^2 = 2.92$$

Zufallsvariablen

- Beispiel: Sei X eine Zufallvariable mit der Massenfunktion

X	0	2	5
p_X	0.5	0.3	0.2

- Dann gilt

$$E[X] = 0 \cdot 0.5 + 2 \cdot 0.3 + 5 \cdot 0.2 = 1.6$$

- Ferner gilt

$$V[X] = (0 - 1.6)^2 \cdot 0.5 + (2 - 1.6)^2 \cdot 0.3 + (5 - 1.6)^2 \cdot 0.2 = 3.64$$

Zufallsvariablen

- Für *unabhängige* Zufallsvariablen X und Y gilt ferner

$$V[X+Y] = V[X] + V[Y]$$

- Beweis: Übung!
- Beispiel: Sei X die Summe $X = X_1 + X_2$ zweier fairer Würfelwürfe. Wir wissen bereits das $V[X_1] = V[X_2] = 2.92$. Da die Würfe unabhängig sind, ist damit auch

$$V[X] = V[X_1] + V[X_2] = 2 \cdot 2.92 = 5.84$$

Zufallsvariablen

- Beispiel: Sei wieder X_i eine Zufallsvariable mit $P(X_i = 1) = p$ und $P(X_i = 0) = 1 - p$. Wir wissen bereits das gilt

$$E[X_i] = p$$

und ferner gilt ebenfalls

$$E[X_i^2] = p$$

also ist nach dem Verschiebungssatz

$$V[X_i] = E[X_i^2] - E[X_i]^2 = p - p2 = p \cdot (1 - p)$$

- Sei nun $X = \sum_{i=1}^{n} X_i$ wieder die Summe von n solcher, unabhängiger, Zufallsvariablen.
- Dann ist

$$V[X] = V\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} V[X_i] = n \cdot p \cdot (1 - p)$$

Binomialverteilung

Binomialverteilung

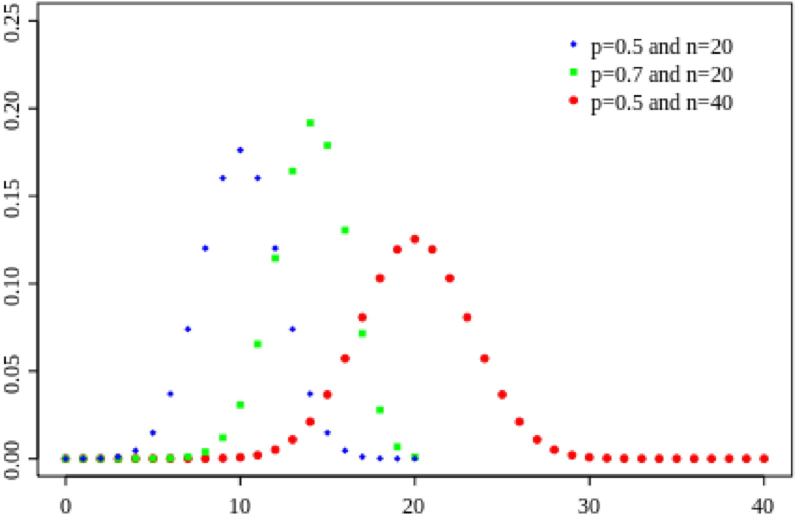
Sei $p \in [0,1]$ und $n \in \mathbb{N}^+$. Dann heisst $X \sim \mathcal{B}_{n,p}$ *Binomialverteilt*, wenn X die Massenfunktion

$$P(k) = \begin{pmatrix} n \\ k \end{pmatrix} p^k (1-p)^{n-k} \text{ für } k = 0, \ldots, n$$

hat.

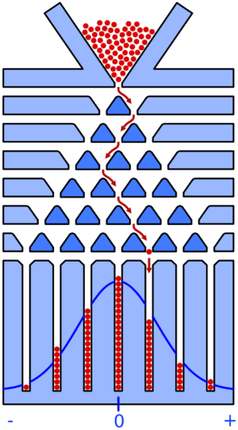
- X beschreibt die Anzahl *Erfolge* bei n unabhängigen Wiederholungen eines Zufallsexperimentes mit jeweils p als Erfolgswahrscheinlichkeit.
- X kann, wie oben, interpretiert werden als Summe von n unabhängigen Zufallsvariablen $X_1, \ldots, X_n$, wobei jeweils $P(X_i = 0) = 1 - p$ und $P(X_i = 1) = p$ gilt.

Binomialverteilung



Binomialverteilung

- Die Binomialverteilung wird z.B. am so genannten *Galton* Brett angenommen



- Sei $X \sim \mathcal{B}_{n,p}$, dann gilt $E[X] = n \cdot p$ und $V[x] = n \cdot p \cdot q$. Beweis siehe oben!

Experiment

Zufallsvariablen

Kovarianz

Seien X und Y Zufallsvariablen. Dann heisst

$$Cov[X, Y] = E[(X - E[X]) \cdot (Y - E[Y])]$$

Kovarianz von X und Y.

- Die Kovarianz ist positiv, wenn X und Y linear *korreliert* sind, d.h. hohe Werte von X (tendenziel) mit hohen Werten von Y einhergehen.
- Die Kovarianz ist negativ, wenn hohe Werte von X mit niedrigen Werten von Y einhergehen.
- Ist die Kovarianz 0, so besteht kein linearer Zusammenhang. Andere Zusammenhänge (z.B. quadratisch) sind jedoch möglich.

Zufallsvariablen

Verschiebungssatz für Kovarianzen

Seien X und Y Zufallsvariablen. Dann gilt für die Kovarianz

$$Cov[X,Y] = E[XY] - E[X] \cdot E[Y]$$

Kovarianz von X und Y.

- Beweis: Folgt trivial aus der Linearität des Erwartungswertes:

$$\begin{aligned} Cov[X,Y] &= E[(X - E[X]) \cdot (Y - E[Y])] \\ &= E[XY] - XE[Y] - YE[X] + E[X]E[Y] \\ &= E[XY] - E[X]E[Y] - E[Y]E[X] + E[X]E[Y] \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

Zufallsvariablen

- Die Kovarianz is bilinear, das heisst es gilt

$$Cov[aX + b, cY + d] = abCov[X, Y]$$

und

$$Cov[X, Y + Z] = Cov[X, Y] + Cov[X, Z]$$

- Beweis: Auf Grund der Linearität des Erwartungswertes gilt $E[aX + b] = aE[X] + b$ und damit

$$\begin{aligned} & Cov[aX + b, cY + d] \\ = \quad & E[(aX + b - E[aX + b])(cY + d - E[cY + d])] \\ = \quad & E[ac(X - E[X])(Y - E[Y])] \\ = \quad & ac \cdot Cov[X, Y] \end{aligned}$$

Zufallsvariablen

- Die Kovarianz is bilinear, das heisst es gilt

$$Cov[aX+b, cY+d] = abCov[X,Y]$$

und

$$Cov[X, Y+Z] = Cov[X,Y] + Cov[X,Z]$$

- Beweis: Ferner gilt mit $E[Y+Z] = E[Y] + E[Z]$

$$\begin{aligned} & Cov[X, Y+Z] \\ = \; & E[(X-E[X])(Y+Z-E[Y+Z])] \\ = \; & E[(X-E[X])(Y-E[Y]) + (X-E[X])(Z-E[Z])] \\ = \; & Cov[X,Y] + Cov[X,Z] \end{aligned}$$

Zufallsvariablen

- Beispiel: Seien X und Y zwei Zufallsvariabeln mit der folgenden gemeinsamen Massenfunktion

	X=-1	X=0	X=1	
Y=1	0.1	0.2	0.1	0.4
Y=2	0.2	0.0	0.1	0.3
Y=4	0.0	0.1	0.2	0.3
	0.3	0.3	0.4	

- Dann ist

$$\begin{aligned} E[X] &= -0.3 + 0.4 = 0.1 \\ E[Y] &= 0.4 + 2 \cdot 0.3 + 4 \cdot 0.3 = 2.2 \end{aligned}$$

und

$$E[XY] = -1 \cdot 0.1 - 2 \cdot 0.2 + 1 \cdot 0.1 + 2 \cdot 0.2 + 4 \cdot 0.2 = 0.8$$

und damit auch

$$Cov[X, Y] = E[XY] - E[X]E[Y] = 0.8 - 0.1 \cdot 2.2 = 0.58$$

Zufallsvariablen

- Beispiel: Seien X und Y wie oben, dann gilt auch

$$Cov[2X+1, 1-2Y] = -4Cov[X, Y] = -2.32$$

- Übung: Manuell nachrechnen!

Zufallsvariablen

- Beispiel: Seien X_1 und X_2 unabhängige Zufallsvariable jeweils mit Erwartungswert 0 und Varianz 1. Seien ferner $Y = X_1 + X_2$ und $Z = X_1 - X_2$. Dann gilt

$$E[Y] = E[X_1 + X_2] = E[X_1] + E[X_2] = 0$$

und auch

$$E[Z] = E[X_1 - X_2] = E[X_1] - E[X_2] = 0$$

- Damit gilt auch

$$\begin{aligned} Cov[Y, Z] &= E[(Y - E[Y])(Z - E[Z])] \\ &= E[(X_1 + X_2)(X_1 - X_2)] \\ &= E[X_1^2 - X_2^2] \\ &= E[X_1^2] - E[X_2^2] \end{aligned}$$

- Nun ist wegen $V[X_i] = E[X_i^2] - E[X_i]^2 = 1$ (Verschiebungssatz) auch $E[X_i^2] = 1$ und damit $Cov[Y, Z] = 0$

Zufallsvariablen

- Für unabhängige Zufallsvariablen X und Y gilt

$$Cov[X, Y] = 0$$

- Beweis: Da X und Y unabhängig seien sollen gilt $E[XY] = E[X]E[Y]$ und damit auch nach dem Verschiebungssatz

$$Cov[X, Y] = E[XY] - E[X] \cdot E[Y] = 0$$

- Haben zwei Zufallsvariablen eine Kovarianz von 0 heissen die beiden Variablen *unkorreliert*.
- Unabhängige Zufallsvariablen sind immer unkorrelierit. Umgekehrt gilt dies nicht notwendigerweise.

Zufallsvariablen

- Beispiel: Seien X und Y Zufallsvariablen mit

$$\begin{aligned} P(X = 0 \wedge Y = 1) &= 1/2 \\ P(X = 2 \wedge Y = 0) &= 1/4 \\ P(X = 2 \wedge Y = 2) &= 1/4 \end{aligned}$$

- Dann ist $P(X = 0) = P(X = 2) = 1/2$ und $P(Y = 0) = P(Y = 2) = 1/4$ sowie $P(Y = 1) = 1/2$
- Damit ist $E[X] = E[Y] = 1$ und ebenfalls $E[XY] = 1$, also $Cov[X, Y] = 0$, damit sind X und Y unkorreliert.
- Andererseits sind X und Y wegen $P(X = 0, Y = 1) = 1/2 \neq (1/2) \cdot (1/2) = P(X = 0)P(Y = 1)$ nicht unabhängig.

Zufallsvariablen

Kovarianzmatrix

Sei $X = (X_1, \dots, X_n)^T$ ein *Zufallsvektor*, also ein Vektor von Zufallsvariablen. Dann heisst

$$Cov[X] = \begin{pmatrix} Cov[X_1, X_1] & \dots & Cov[X_1, X_n] \\ \vdots & \ddots & \vdots \\ Cov[X_n, X_1] & \dots & Cov[X_n, X_n] \end{pmatrix} \in \mathbb{R}^{nxn}$$

Kovarianzmatrix.

- Es gilt für alle Matrizen $A \in \mathbb{R}^{mxn}$ (ohne Beweis):

$$Cov[AX] = A \cdot Cov[X] \cdot A^T$$

Die Tschebyscheffsche Ungleichung

Tschebyscheffsche Ungleichung

Tschebyscheffsche Ungleichung

Sei X ein Zufallsvariable mit beliebiger Verteilung aber endlichem Erwartungswert μ und endlicher Varianz $\sigma^2 > 0$. Dann gilt für alle $k > 0$

$$P\left(|X - \mu| \geq k\sigma\right) \leq \frac{1}{k^2}$$

Tschebyscheffsche Ungleichung

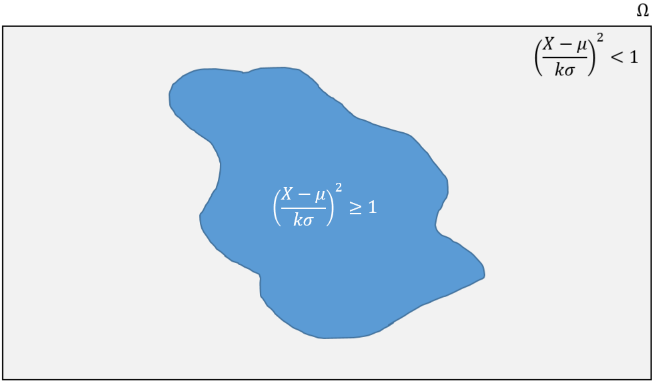
- **Beweis:** Zunächst gilt

$$P\left(|X-\mu| \geq k\sigma\right) = P\left(\left(\frac{X-\mu}{k\sigma}\right)^2 \geq 1\right)$$

- Sei nun I_A eine neue Zufallsvariable, welche genau dann 1 ist, falls das Ereigniss A Eintritt und 0, falls nicht. I_A heisst *Indikatorvariable* zu dem Ereigniss A.
- Dann gilt weiter

$$P\left(\left(\frac{X-\mu}{k\sigma}\right)^2 \geq 1\right) = E\left[I_{\left(\frac{X-\mu}{k\sigma}\right)^2}\right]$$

Tschebyscheffsche Ungleichung



- Die blaue Fläche repräsentiert das Ereigniss $A : \left(\frac{X-\mu}{k\sigma}\right)^2 \geq 1$.
- Der Erwartungswert integriert nun über ganz Ω, wobei der graue Teile 0 ist (weil A nicht eingetreten ist) und der blaue Teil 1 ist (weil A eingetreten ist).

Tschebyscheffsche Ungleichung

QUIZ

Was passiert, wenn wir auf Ω statt über $I_{\left(\frac{X-\mu}{k\sigma}\right)^2}$ einfach über $\left(\frac{X-\mu}{k\sigma}\right)^2$ integrieren würden?

- Die grauen Flächenanteile waren vorher 0 und sind jetzt

$$\left(\frac{X-\mu}{k\sigma}\right)^2 \geq 0$$

.

- Die blauen Flächenanteile waren vorher 1 und sind jetzt

$$\left(\frac{X-\mu}{k\sigma}\right)^2 \geq 1$$

.

Tschebyscheffsche Ungleichung

- Es gilt also

$$P\left(\left(\frac{X-\mu}{k\sigma}\right)^2 \geq 1\right) = E\left[I_{\left(\frac{X-\mu}{k\sigma}\right)^2}\right] \leq E\left[\left(\frac{X-\mu}{k\sigma}\right)^2\right]$$

- Nun folgt wegen die Linearität des Erwartungswertes

$$E\left[\left(\frac{X-\mu}{k\sigma}\right)^2\right] = \frac{1}{k^2\sigma^2}E\left[(X-\mu)^2\right] = \frac{1}{k^2\sigma^2}V[X] = \frac{1}{k^2}$$

Zusammenfassung

Was Sie heute gelernt haben

- Sie wissen, was eine Zufallsvariable ist.
- Sie wissen, was die Massenfunktion und die Verteilungsfunktion einer Zufallsvariable angibt.
- Sie kennen den Begriff des Erwartungswertes und seine Eigenschaften.
- Sie kennen den Begriff der Varianz und Kovarianz und ihre jeweiligen Eigenschaften.
- Sie kennen die Binomialverteilung und ihre Eigenschaften.

Kontinuierliche Wahrscheinlichkeitsräume

Kontinuierliche Wahrscheinlichkeitsräume

- Viele reale Systeme lassen sich besser durch kontinuierliche Größen beschreiben als durch diskrete.
- Bisher waren alle betrachteten Ereignissmengen höchstens abzählbar unendlich, also diskretisierbar.
- Im folgenden betrachten wir, was passiert wenn wir überabzählbar unendliche Ereignissmengen zulassen.
- Speziell interessieren uns reelwertige Zufallsvariablen, also Zufallsvariablen die in einer (Teil-)Menge der reelen Zahlen überabzählbar viele Werte annehmen können.
- Wir werden sehen das sehr viele Eigenschaften, die wir schon kennen erhalten bleiben, aber formal anders geschrieben werden müssen. Sie brauchen also nur die Unterschiede lernen!

Kontinuierliche Wahrscheinlichkeitsräume

Kontinuierlicher Wahrscheinlichkeitsraum

Ein Wahrscheinlichkeitsraum $(\Omega, \sum, P)$ heisst *kontinuierlicher Wahrscheinlichkeitsraum*, wenn die Ergebnismenge Ω eine Teilmenge der reelen Zahlen ist und $\sum$ bestimmte Eigenschaften erfüllt (siehe unten).

Um später sinnvolle Wahrscheinlichkeitsmaße auf $\sum$ definieren zu können muß $\sum$ eine so genannten *Borelsche* Menge sein. Das bedeutet

1. Alle abgeschlossenen Intervalle $[a, b] \subset \mathbb{R}$ sind in $\sum$ enthalten.
2. Wenn $A_1, A_2, A_3, \ldots$ in $\sum$ enthalten sind, dann auch $A_1 \cup A_2 \cup A_3 \cup \ldots$
3. Wenn A in $\sum$ enthalten ist, dann auch $\overline{A} = \mathbb{R} \setminus A$.

Kontinuierliche Wahrscheinlichkeitsräume

- Wie in Übung 1 macht es Sinn, das Wahrscheinlichkeitsmaß über Ω als *Flächenverhältniss* zu definieren.
- D.h. $P(A)$ soll proportional zur von A eingenommen *Fläche* bezogen auf ganz Ω sein.
- Wir wollen an dieser Stelle nicht näher erläutern was genau wir mit *Fläche* eigentlich meinen, das intuitive Verständniss reicht hier.

Dichtefunktion

- Wenn, wie oben beschrieben, das Wahrscheinlichkeitsmaß von A proportional zur *Fläche* von A sein soll, kann man diese Fläche über ein Integral ausdrücken
- Dann ist nämlich
$$P(X \in A) = \int_A 1 dx$$
wobei hier über alle Elemente in A integriert wird.
- Dies entspricht intuitiv der Annahme, das alle Elementarereignisse in Ω gleich wahrscheinlich sind (kontinuerlicher Laplaceraum).
- Wir wollen dieses Prinzip im folgenden verallgemeinern...

Dichtefunktion

Dichtefunktion

Sei $f : \mathbb{R} \mapsto \mathbb{R}_0^+$ eine *stetige* Funktion mit der Bedingung

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

dann heisst f *Dichtefunktion*

- Die Dichtefunktion nimmt nur positive Werte inklusive 0 an
- Die Dichtefunktion gibt das **relative** Verhältnis der Wahrscheinlichkeiten bestimmter Elementarereignisse an.
- Die Dichtefunktion gibt insbesondere **nicht** die absolute Wahrscheinlichkeit für ein Elementarereignis an.
- Anschauliche Interpretation: Die Dichtefunktion gibt an, wie *dicht* Ausgänge des Zufallsexperimentes in einem (infinitesimal kleinem) Gebiet gestreut sind.

Dichtefunktion

Beispiel

- Es sei $\Omega = [0,1] \subset \mathbb{R}$ und

$$f_1(x) = (3/2)\sqrt{x}$$

sowie

$$f_2(x) = 3x^2$$

- Dann sind wegen

$$\int_0^1 (3/2)\sqrt{x}dx = x^{3/2}\Big|_0^1 = 1$$

und

$$\int_0^1 3x^2 dx = x^3\Big|_0^1 = 1$$

sowohl f_1 als auch f_2 Dichtefunktionen über Ω.

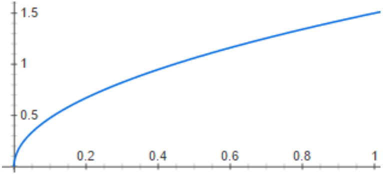
Dichtefunktion

Beispiel

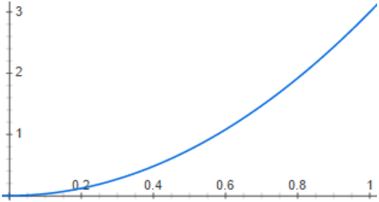
- Es sei $\Omega = [0,1] \subset \mathbb{R}$ und
$$f_1(x) = (3/2)\sqrt{x}$$
sowie
$$f_2(x) = 3x^2$$
- Nun ist aber $f_1(1) = 1.5$ und $f_2(1) = 3$.
- Also kann die Dichtefunktion allein kein sinnvolles Wahrscheinlichkeitsmaß angeben (Axiom 1 verletzt).

Dichtefunktion

Beispiel



$f_1(x)$



$f_2(x)$

Verteilungsfunktion

- Um sinnvolle Wahrscheinlichkeitsmaße über Ω definieren zu können brauchen wir auch noch die so genannte *Verteilungsfunktion*

Verteilungsfunktion

Sei f eine Dichtefunktion über Ω, dann heißt

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

Verteilungsfunktion

- Mit der Verteilungsfunktion ist

$$P(a \leq X \leq b) = F(b) - F(a) = \int_{a}^{b} f(t)dt$$

ein Wahrscheinlichkeitsmaß über Ω.

Verteilungsfunktion

- Beweis (Axiom 1): F ist monoton steigend (da stets positive Flächenstücke auf-integriert werden), d.h. $F(b) \geq F(a) \Leftrightarrow b \geq a$.
- Damit ist $P(a \leq X \leq b) = F(b) - F(a) \geq 0$
- Ferner gilt nach Voraussetzung (f ist eine Dichtefunktion)

$$F(\infty) = \lim_{x \mapsto \infty} F(x) = \int_{-\infty}^{\infty} f(t)dt = 1$$

und

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

also auch

$$P(a \leq X \leq b) = F(b) - F(a) \leq F(\infty) - F(-\infty) = 1$$

.

Verteilungsfunktion

- Beweis (Axiom 2): Es gilt wieder nach Voraussetzung (f ist eine Dichtefunktion)

$$P(\Omega) = P(-\infty \leq X \leq \infty) = 1$$

- Axiom 3: Seien nun $I_1 = [a_1, b_1]$ und $I_2 = [a_2, b_2]$ disjunkte Intervalle, d.h. O.B.D.A. $b_1 < a_2$, dann gilt

$$\begin{aligned} P(I_1 \cup I_2) &= \int_{I_1 \cup I_2} f(t)dt \\ &= \int_{a_1}^{b_1} f(t)dt + \int_{a_2}^{b_2} f(t)dt \\ &= P(I_1) + P(I_2) \end{aligned}$$

Wahrscheinlichkeiten in kontinuierlichen Wahrscheinlichkeitsräumen

- Die Wahrscheinlichkeit dafür, das die (kontinuierliche) Zufallsvariable X im Intervall $[a, b]$ liegt ist somit also Fläche unterhalb der Dichtefunktion f gegeben.
- Das bedeutet insbesondere, das die Wahrscheinlichkeit für ein (singuläres) Ereignis $X = x$ stets 0 ist
- Es gilt nämlich stets

$$P(X = x) = P(x \leq X \leq x) = F(x) - F(x) = 0$$

- Dies zeigt erneut, das die Dichtefunktion **nicht** die Wahrscheinlichkeit für das auftreten eines Elementarereignisses angibt.

Erwartungswert und Varianz

- Obwohl die Dichtefunktion und die Massenfunktion so unterschiedliche Interpretationen bedürfen, nehmen beide bei den meisten Überlegungen trotzdem die selbe Rolle ein.
- Wir werden dies anhand des Erwartungswertes und der Varianz sehen

Erwartungswert, Varianz

Sei X eine kontiniuerliche Zufallsvariable mit Dichte f, dann heißt

$$E[X] = \int_{-\infty}^{\infty} tf(t)dt$$

Erwartungswert von X. Ferner heißt

$$V[X] = \int_{-\infty}^{\infty} (t - E[X])^2 f(t)dt$$

Varianz von X.

Erwartungswert und Varianz

- Die bekannten Rechenregeln für Erwartungswert und Varianz im diskreten gelten sinngemäß auch im kontinuierlichen, d.h. es gilt z.B.
$$E[aX + bY] = aE[X] + bE[Y]$$
für zwei kontinuierliche Zufallsvariablen X und Y.
- Ferner gilt weiterhin der wichtige *Verschiebungssatz*
$$V[X] = E[X^2] - E[X]^2$$

Dichtefunktion

Beispiel

- Sei X eine Zufallsvariable über $\Omega = [a, b] \subset \mathbb{R}$ mit $b > a$ und der Dichte
$$f(x) = 1/(b-a)$$
- X heißt *gleichverteilt*, da sie überall die selbe Dichte hat.
- Nun ist
$$E[X] = \int_a^b \frac{1}{b-a} \cdot t \; dt = \frac{1}{2(b-a)} t^2 \Big|_a^b = \frac{a+b}{2}$$
- Es gilt ferner
$$E[X^2] = \int_a^b \frac{1}{b-a} \cdot t^2 \; dt = \frac{1}{3(b-a)} t^3 \Big|_a^b = \frac{b^3 - a^3}{3(b-a)}$$
und damit auch (selber nachrechnen!)
$$V[X] = E[X^2] - E[X]^2 = \cdots = \frac{1}{12}(b-a)^2$$

Normalverteilung

- Die mit Abstand wichtigste kontinuierliche Wahrscheinlichkeitsverteilung ist die sogenannte *Normalverteilung* $\mathcal{N}$.
- Die Wichtigkeit wird erst aus dem (hier nicht behandelten) *zentralen Grenzwertsatz* klar.
- Das Gesetz besagt, flapsig formuliert, das, falls $X_1, \ldots, X_n$ hinreichend viele Zufallsvariablen mit **beliebiger** (also insbesondere auch unbekannter) Verteilung sind, trotzdem stets gilt

  $$Z = \sum_i X_i \sim \mathcal{N}$$

  also die Summe vieler solcher Zufallsvariablen stets Normalverteilt ist.

Normalverteilung

- Praktische Bedeutung: Oft hat man ein System mit unbekannten Störeinflüssen.
- Beispiel: Fahrzeug während Kurvenfahrt. Mögliche Einflüsse: Wind, Reibung, Straßenneigung, etc...
- Jeder einzelne Störeinfluss ist eher minimal, aber die Summe aus vielen Störeinflüssen ist nicht mehr zu vernachlässigen. Anstatt nun jeden Einfluss individuell modellieren zu müssen, kann man mit dem zentralen Grenzwertsatz einfach annehmen, das die Summe dieser Einflüsse Normalverteilt ist, ohne die einzelnen Einflüsse genau zu kennen.
- Damit findet die Normalverteilung sehr oft Anwendung in sehr vielen praktischen und theoretischen Problemen.

Normalverteilung

Normalverteilung

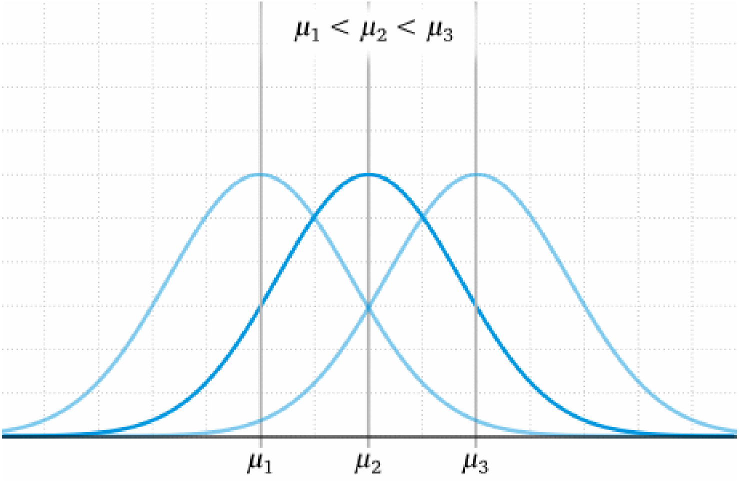
Sei $\mu \in \mathbb{R}$ und $\sigma \in \mathbb{R}$. Dann heißt $X \sim \mathcal{N}(\mu, \sigma^2)$ *normalverteilt* wenn X die Dichtefunktion

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

hat.

Einfluss von μ

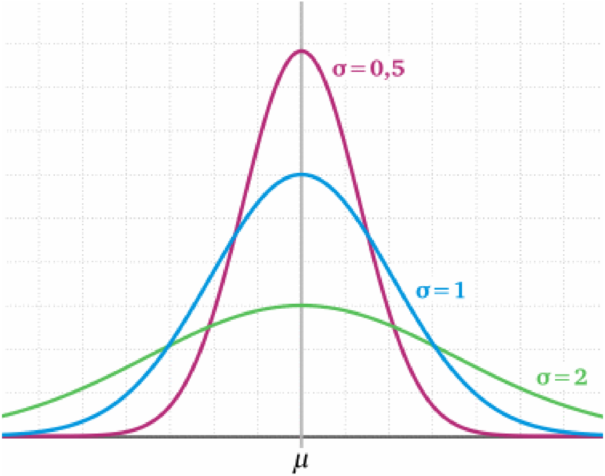
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$



Normalverteilung

Einfluss von σ

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$



Normalverteilung

- Sei $X \sim \mathcal{N}(\mu, \sigma^2)$. Dann ist μ der Erwartungswert von X.
- Beweis: Wir betrachten zunächst $Z \sim \mathcal{N}(0,1)$. Dann ist

$$f_Z(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right)$$

und damit

$$E[Z] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) t \; dt$$

- Da f_Z symmetrisch um 0 ist gilt also $E[Z] = 0$.
- Ferner ist $X = \sigma Z + \mu$, also $E[X] = E[\sigma Z + \mu] = \sigma E[Z] + \mu = \mu$.
- Ohne Beweis: Es gilt ferner $V[X] = \sigma^2$.