Kombinatorik und Wahrscheinlichkeitsrechnung

Übersicht:

1. **Teil 1: Elementare Kombinatorik**
2. Teil 2: Grundlagen der Wahrscheinlichkeitsrechnung

Motivation

- In der Vorlesung wurden bereits mehrfach kombinatorische Argumente verwendet (Anzahl von Teilmengen, Booleschen Funktionen, etc.). Solche Argumente spielen in vielen Überlegungen (und formalen Beweisen) eine wichtige Rolle.
- Die Kombinatorik bietet einen natürlichen Zugang zur diskreten Wahrscheinlichkeitsrechnung.
- Die Wahrscheinlichkeitsrechnung ist ein wichtiger Kalkül zur Repräsentation und Verarbeitung von Unsicherheit und spielt daher in der Modellierung eine zentrale Rolle.
- Beziehung zwischen Logik (Aussagenlogik) und Wahrscheinlichkeitsrechnung.

Ziehen von Elementen aus einer Menge

Anzahl der Möglichkeiten, k Objekte aus einer n-elementingen Menge zu ziehen:

	geordnet	ungeordnet
mit Zurücklegen	n^k	$\binom{n+k-1}{k}$
ohne Zurücklegen	$\frac{n!}{(n-k)!}$	$\binom{n}{k}$

Für $n \in \mathbb{N}_0$ ist dabei $n! \stackrel{\text{df}}{=} \prod_{i=1}^{n} i$ die **Fakultät** von n (wobei $0! \stackrel{\text{df}}{=} 1$).

Geordnetes Ziehen $\longrightarrow$ Sequenzen
Uneordnetes Ziehen $\longrightarrow$ Mengen bzw. Multimengen

Ziehen von Elementen aus einer Menge

Beispiel: $n = 3$ ($M = \{1, 2, 3\}$), $k = 2$

- geordnet, mit Zurücklegen:
  $A = \{(1,1), (1,2), \ldots, (2,1), \ldots, (3,3)\}$, $|A| = 9$
- geordnet, ohne Zurücklegen:
  $A = \{(1,2), (1,3), (2,1), (2,3), (3,1), (3,2)\}$, $|A| = 6$
- ungeordnet, mit Zurücklegen:
  $A = \{\{1,1\}, \{1,2\}, \{1,3\}, \{2,2\}, \{2,3\}, \{3,3\}\}$, $|A| = 6$
- ungeordnet, ohne Zurücklegen:
  $A = \{\{1,2\}, \{1,3\}, \{2,3\}, \{3,3\}\}$, $|A| = 3$

Ziehen von Elementen aus einer Menge

Erläuterung zum Fall "ungeordnet/mit Zurücklegen":

- Die abzuzählende Menge A kann offenbar repräsentiert werden durch die Menge aller n-Tupel $(a_1, \ldots, a_n)$ mit $a_i \in \mathbb{N}_0$ und $a_1 + \ldots + a_n = k$ (a_i ist die Häufigkeit des Elementes i in der Multimenge); vgl. Gleichheitsregel.
- Ein jedes Tupel dieser Art kann wiederum (eindeutig) repräsentiert werden durch eine Zeichenfolge mit $n-1$ Trennstrichen $|$ und a_i Kreisen $\bullet$ vor dem i-ten Strich (vgl. Gleichheitsregel). Beispiel: Für $n = 5$ und $k = 3$ wird $(0, 1, 0, 2, 0)$ repräsentiert durch

$$| \bullet || \bullet \bullet |$$

- Jede solche Zeichenfolge besteht aus insgesamt $n + k - 1$ Zeichen und ist eindeutig beschrieben durch die k Positionen der Kreise. Damit ist das Zählproblem auf den Fall "ungeordnet/ohne Zurücklegen" zurückgeführt (wähle k Positionen aus $n + k - 1$ möglichen).

Binomialkoeffizient

Für $k, n \in \mathbb{N}_0$, $k \leq n$, ist der **Binomialkoeffizient** $\binom{n}{k}$ definiert durch

$$\binom{n}{k} \stackrel{\text{df}}{=} \frac{n!}{k!(n-k)!}$$

(für $k > n$ setzt man den Wert auf 0). Es gilt die **binomische Formel**:

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

Man stellt leicht fest, dass für Binomialkoeffizienten die folgenden Identitäten gelten:

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

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

Kombinatorische Beweisprinzipien

In Beweisen und Herleitungen werden häufig die folgenden kombinatorischen Argumente verwendet:

- **Summenregel:** Ist $\{S_i \mid i \in I\}$ eine Zerlegung der Menge S (also $\bigcup_{i \in I} S_i = S$ und $S_i \cap S_j = \emptyset$ für $i \neq j$), so gilt
$$|S| = \sum_{i \in I} |S_i|.$$
- **Produktregel:** Ist S das kartesische Produkt der Mengen S_i, $i \in I$, so gilt
$$|S| = \prod_{i \in I} |S_i|.$$
- **Gleichheitsregel:** Ist $f : S \longrightarrow T$ eine Bijektion, so gilt $|S| = |T|$.
- **Doppeltes Abzählen:** Ist $R \subseteq S \times T$, so kann man die Kardinalität von R "zeilenweise" als auch "spaltenweise" bestimmen:
$$|R| = \sum_{s \in S} |\{t \in T \mid (s,t) \in R\}| = \sum_{t \in T} |\{s \in S \mid (s,t) \in R\}|$$

Kombinatorische Beweisprinzipien

- **Schubfachprinzip:** Verteilt man n Elemente auf $m < n$ Fächer, so gibt es mindestens ein Fach mit zwei Elementen (vgl. "pigeonhole principle"): Ist $f : X \longrightarrow Y$ ein Abbildung und gilt $|X| > |Y|$, so existiert ein $y \in Y$ mit $|f^{(-1)}(y)| \geq 2$.
  Die Aussage des Schubfachprinzips kann man offensichtlich wie folgt verfeinern: Ist $f : X \longrightarrow Y$ ein Abbildung, so existiert ein $y \in Y$ mit
  $$|f^{(-1)}(y)| \geq \left\lceil \frac{|X|}{|Y|} \right\rceil .$$

- **Prinzip der Inklusion und Exklusion:** Sind $A_1, \ldots, A_n$ endliche Mengen, so gilt
  $$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{r=1}^{n} (-1)^{r-1} \sum_{1 \leq i_1 < i_2 \leq \ldots < i_r \leq n} \left| \bigcap_{j=1}^{r} A_{i_j} \right| .$$

Zählprobleme: Teilmengen

Da eine n-elementigen Menge insgesamt 2^n Teilmengen besitzt, und diese sich aus der 0-elementigen, den 1-elementigen, 2-elementigen, etc. zusammensetzen, folgt aus der Summenregel

$$2^n = \sum_{k=0}^{n} \binom{n}{k}$$

Es gilt weiterhin (Summenregel und Produktregel) die **Vandermonde'sche Identität**:

$$\binom{n+m}{k} = \sum_{\ell=0}^{k} \binom{n}{\ell} \binom{m}{k-\ell}$$

(Um k Elemente aus einer Menge mit $n+m$ Elementen zu ziehen, kann man zunächst ℓ aus n ziehen und dann $k-\ell$ aus m.)

Zählprobleme: Mengenpartitionen

Die Zahl der k-Partitionen einer n-elementigen Menge wird durch die **Stirlingzahlen zweiter Art** beschrieben:

$$S_{n,k} = S_{n-1,k-1} + kS_{n-1,k}$$

für alle $k, n \in \mathbb{N}$ mit $n \geq k$ und $S_{n,0} = 0$ für alle n (sowie, per definitionem, $S_{0,0} = 1$ und $S_{n,k} = 0$ für $k > n$).

Zählprobleme: Permutationen

Eine Permutation einer Menge $A = \{a_1, \ldots, a_n\}$ ist eine bijektive Abbildung $\pi : A \longrightarrow A$. Bringt man die Elemente in eine feste Reihenfolge, so kann man o.B.d.A. $A = \{1, \ldots, n\}$ annehmen. Die Menge aller Permutationen von $\{1, \ldots, n\}$ bezeichnet man als **symmetrische Gruppe** $\mathfrak{S}_n$. Es gilt:

$$|\mathfrak{S}_n| = n!$$

Zählprobleme: Permutationen

Ein **Zyklus** (der Länge t) einer Permutation π ist eine Folge $(i_1, i_2, \ldots, i_t)$, sodass $i_j \neq i_k$ für $1 \leq j \neq k \leq t$, $\pi(i_j) = i_{j+1}$ für $1 \leq j < t$ und $\pi(i_t) = i_1$. Man beachte, dass ein Zyklus bis auf die Wahl des ersten Elementes eindeutig festgelegt ist ((3 6 1 2) ist also der selbe Zyklus wie (1 2 3 6)).

Die Anzahl der Permutationen $\pi \in \mathfrak{S}_n$ mit genau k Zyklen ist durch die **Stirlingzahlen erster Art** gegeben:

$$s_{n,k} = s_{n-1,k-1} + (n-1)s_{n-1,k}$$

für alle $k, n \in \mathbb{N}$ mit $n \geq k$ (sowie $s_{0,0} = 1$, $s_{n,0} = 0$ für alle $n > 0$, $S_{0,k} = 0$ für all $k > 0$).

Zählprobleme: Zahlpartitionen

Eine Darstellung einer natürlichen Zahl n als Summe

$$n = n_1 + n_2 + \ldots + n_k$$

von k natürlichen Zahlen n_i, $i = 1, \ldots, k$, nennt man Zahlpartition von n.
Die Anzahl **ungeordneter** k-Partitionen einer Zahl n ist gegeben durch

$$P_{n,k} = \sum_{j=0}^{k} P_{n-k,j}$$

für alle $k, n \in \mathbb{N}$ mit $n \geq k \geq 1$ (sowie $P_{0,0} = 1$, $P_{n,0} = 0$ für $n \geq 1$ und $P_{n,k} = 0$ für $k > n$).
Die Anzahl **geordneter** k-Partitionen einer Zahl n ist gegeben durch

$$\binom{n-1}{k-1}$$

für alle $k, n \in \mathbb{N}$ mit $n \geq k$.

Zählprobleme: Catalanzahlen

Die Catalanzahlen

$$C_n = \frac{1}{n+1}\binom{2n}{n}, \qquad (n \geq 0)$$

benannt nach Eugène Catalan (1814–1894), tauchen in diversen Zählproblemen auf. So ist z.B. C_n die Anzahl der Binärbäume auf $n+1$ Blättern.

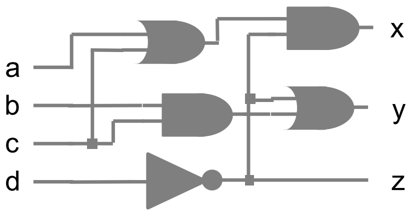
Folge der Catalanzahlen:
1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, 208012, ...

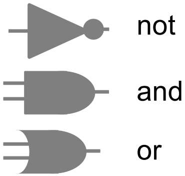
Kombinatorik und Wahrscheinlichkeitsrechnung

Übersicht:

1. Teil 1: Elementare Kombinatorik
2. **Teil 2: Grundlagen der Wahrscheinlichkeitsrechnung**

Motivation





Schaltnetz entspricht der folgenden booleschen Funktion:

$$x = (a \vee c) \wedge \neg d$$
$$y = \neg d \vee (b \wedge c)$$
$$z = \neg d$$

Wie modelliert man die Schaltung bzw. den Zusammenhang zwischen Ein- und Ausgaben, wenn davon auszugehen ist, dass die einzelnen Gatter fehlerhaft arbeiten können?

Bezeichnungen und Sprechweisen

- Ein **Zufallsexperiment** (Zufallsvorgang, stochastischer Vorgang) ist ein beliebig oft wiederholbarer, nach ganz bestimmten Vorschriften ablaufender Vorgang, dessen Ausgang "zufallsbedigt" ist.
- Die möglichen Ausgänge eines Zufallsexperiments fasst man zusammen zu einer **Ergebnismenge** $\Omega = \{\omega_1, \omega_2, \ldots\}$ sich gegenseitig ausschließender **Elementarereignisse**.
- Das tatsächlich eingetretene Elementarereignis heißt das **Ergebnis** des Zufallsvorgangs.
- Ein **Ereignis**, häufig durch eine logische oder nur umgangssprachliche Beschreibung charakterisiert, wird identifiziert mit einer Teilmenge $A \subseteq \Omega$. Man sagt, "A tritt ein", wenn $\omega \in A$ gilt, wobei ω das Ergebnis des Zufallsvorgangs ist. Man nennt $\emptyset$ das "unmögliche Ereignis" und Ω das "sichere Ereignis".

Diskreter Wahrscheinlichkeitsraum

Ein diskreter Wahrscheinlichkeitsraum ist ein Tupel $(\Omega, \mathbf{P})$, wobei Ω abzählbar ist und $\mathbf{P}$ ein (Wahrscheinlichkeits-)Maß auf Ω (eine reellwertige Abbildung mit Definitionsbereich $\mathrm{Pow}(\Omega)$) mit folgenden Eigenschaften:

A1 $\mathbf{P}(A) \geq 0$ für alle $A \subseteq \Omega$

A2 Normierung: $\mathbf{P}(\Omega) = 1$

A3 σ-Additivität:

$$\mathbf{P}(A_1 \cup A_2 \cup \ldots) = \mathbf{P}(A_1) + \mathbf{P}(A_2) + \ldots$$

für endlich oder abzählbar unendlich viele paarweise disjunkte Ereignisse $A_i \subseteq \Omega$.

Ist Ω endlich, so nennt man auch den Wahrscheinlichkeitsraum $(\Omega, \mathbf{P})$ endlich.

σ-Algebren

Allgemeiner definiert man Wahrscheinlichkeitsmaße auf einer σ-Algebra, ein Mengensystem auf Ω, das abgeschlossen gegenüber Komplementbildung und der Vereinigung abzählbar vieler Teilmengen ist, genauer:

Ein Mengensystem $\mathcal{A} \subseteq \mathrm{Pow}(\Omega)$ ist eine σ-Algebra, falls gilt:

- $\Omega \in \mathcal{A}$;
- ist $X \in \mathcal{A}$, so ist auch $X^c = (\Omega \setminus X) \in \mathcal{A}$;
- sind $X_1, X_2, \ldots \in \mathcal{A}$, so ist auch $\bigcup_{i=1}^{\infty} X_i \in \mathcal{A}$.

$\mathcal{A} = \mathrm{Pow}(\Omega)$ ist ein Spezialfall, der für diskretes Ω unproblematisch ist, jedoch nicht für überabzählbares Ω (in diesem Fall existiert i.A. kein Maß auf $\mathrm{Pow}(\Omega)$, das den Axiomen A1–A3 genügt).

Charakterisierung von Maßen

Aus den Axiomen A1–A3 folgt sofort, dass das Maß $\mathbf{P}$ bereits vollständig festgelegt ist durch seine Einschränkung auf Ω, also durch die Wahrscheinlichkeiten $\mathbf{P}(\omega) = \mathbf{P}(\{\omega\})$ der Elementarereignisse $\omega \in \Omega$. Es gilt nämlich für alle $A \subseteq \Omega$:

$$\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\{\omega\}) \tag{1}$$

Laplacescher Wahrscheinlichkeitsraum

Ist $(\Omega, \mathbf{P})$ ein endlicher Wahrscheinlichkeitsraum und ist jedes Elementarereignis gleich wahrscheinlich, so spricht man von einem Laplaceschen Wahrscheinlichkeitsraum.

Es gilt dann also $\mathbf{P}(\omega) = \frac{1}{|\Omega|}$ für jedes $\omega \in \Omega$, und aus (1) folgt

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|}$$

für alle $A \subseteq \Omega$.

$\longrightarrow$ Bezug zur Kombinatorik

Charakterisierung von Wahrscheinlichkeitsmaßen

Aus den Axiomen A1–A3 folgt, dass für Ereignisse $A, B \subseteq \Omega$ gilt:

a. $\mathbf{P}(A^c) = 1 - \mathbf{P}(A)$
b. $\mathbf{P}(\emptyset) = 0$
c. aus $A \subseteq B$ folgt $\mathbf{P}(A) \leq \mathbf{P}(B)$
d. $0 \leq \mathbf{P}(A) \leq 1$

Siebformel, Prinzip der Inklusion/Exklusion

Es sei $(\Omega, \mathbf{P})$ ein (diskreter) Wahrscheinlichkeitsraum. Für Ereignisse $A_1, A_2, \ldots, A_n \subseteq \Omega$, $n \geq 2$, gilt:

$$\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{r=1}^{n} (-1)^{r-1} \sum_{1 \leq i_1 < i_2 < \ldots < i_r \leq n} \mathbf{P}\left(\bigcap_{j=1}^{r} A_{i_j}\right).$$

Bedingte Wahrscheinlichkeit

Für Ereignisse A, B mit $\mathbf{P}(B) > 0$ ist die **bedingte Wahrscheinlichkeit** $\mathbf{P}(A \mid B)$ definiert durch

$$\mathbf{P}(A \mid B) \stackrel{\text{df}}{=} \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}.$$

Ist $(\Omega, \mathbf{P})$ ein (diskreter) Wahrscheinlichkeitsraum und $B \subseteq \Omega$ ein Ereignis mit $\mathbf{P}(B) > 0$, so ist $(\Omega, \mathbf{P}_B)$ mit

$$\mathbf{P}_B(A) = \mathbf{P}(A \mid B)$$

für alle $A \subseteq \Omega$ wiederum ein Wahrscheinlichkeitsraum.

Multiplikationssatz

Es seien $A_1, \ldots, A_n$ Ereignisse mit $\mathbf{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) > 0$. Dann gilt

$$\mathbf{P}(A_1 \cap A_2 \cap \ldots \cap A_n) = A_1 \cdot \prod_{i=1}^{n-1} \mathbf{P}(A_{i+1} \,|\, A_1 \cap A_2 \cap \ldots \cap A_i).$$

Satz der totalen Wahrscheinlichkeit

Es seien $A_1, \ldots, A_n \subseteq \Omega$ mit $A_i \cap A_j = \emptyset$ für $1 \leq i \neq j \leq n$ und $B \subseteq A_1 \cup A_2 \cup \ldots \cup A_n$. Dann gilt

$$\mathbf{P}(B) = \sum_{i=1}^{n} \mathbf{P}(B \mid A_i) \cdot \mathbf{P}(A_i).$$

Satz von Bayes

Es seien $A_1, \ldots, A_n \subseteq \Omega$ mit $A_i \cap A_j = \emptyset$ für $1 \leq i \neq j \leq n$ und $B \subseteq A_1 \cup A_2 \cup \ldots \cup A_n$ mit $\mathbf{P}(B) > 0$. Dann gilt

$$\mathbf{P}(A_i \,|\, B) = \frac{\mathbf{P}(A_i \cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}(B \,|\, A_i)\mathbf{P}(A_i)}{\sum_{j=1}^{n} \mathbf{P}(B \,|\, A_j)\mathbf{P}(A_j)}$$

für alle $i = 1, \ldots, n$.

Beispiel

- Im Rahmen eines Produktionsprozesses fällt 1% Ausschuss an.
- Mithilfe eines Prüfverfahrens kann getestet werden, ob ein produziertes Teil in Ordnung ist oder nicht.
- Das Verfahren ist allerdings nicht sicher, sondern liefert (unabhängig vom wahren Zustand des Teils) in 10% der Fälle eine falsche Antwort.
- Wie groß ist die Wahrscheinlichkeit, dass ein Teil wirklich nicht in Ordnung ist, wenn das Prüfverfahren dies anzeigt?

Unabhängigkeit

Zwei Ereignisse $A, B \subseteq \Omega$ heißen unabhängig, wenn gilt:

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$$

Allgemeiner heißen Ereignisse $A_1, A_2, \ldots, A_n$ unabhängig, wenn für alle $I = \{i_1, i_2, \ldots, i_k\} \subseteq \{1, 2, \ldots, n\}$ gilt:

$$\mathbf{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbf{P}(A_i).$$

Bemerkung: Konnektoren in der Logik sind **wahrheitsfunktional**, d.h. der Wahrheitswert einer Aussage hängt nur von den Wahrheitswerten der Teilaussagen ab.
Für die Wahrscheinlichkeitsrechnung gilt dies nicht: Um $\mathbf{P}(A \cap B)$ zu bestimmen, reicht es in der Regel nicht aus, $\mathbf{P}(A)$ und $\mathbf{P}(B)$ zu kennen.

Zufallsvariable

Sei Ω die Ergebnismenge eines Wahrscheinlichkeitsraumes. Eine Abbildung

$$X : \Omega \longrightarrow \mathbb{R}$$

heißt (reellwertige) Zufallsvariable (ZV).

Ist Ω diskret, so ist auch der Wertebereich von X,

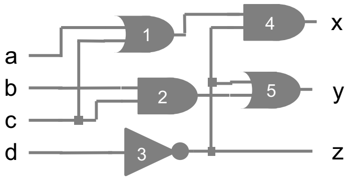
$$W_X = X(\Omega) = \{x \in \mathbb{R} \,|\, X(\omega) \text{ für ein } \omega \in \Omega\},$$

diskret, und X heißt diskrete Zufallsvariable.

Zufallsvariable

Beispiel: Schaltnetz

Wir betrachten die Ergebnismenge $\Omega = \{0,1\}^5$. Für $\omega = (\omega_1, \ldots, \omega_5) \in \Omega$ ist $\omega_i = 1$ genau dann, wenn das i-te Gatter fehlerhaft arbeitet.



Für eine feste Eingabe (z.B. $(a,b,c,d) = (0,0,1,1)$) ist die Ausgabe X dann eine ZV $X = X(\omega)$ mit Wertebereich $W_X = \{0,1\}$. Gleiches gilt für Y und Z.

Zufallsvariable

Beispiel: Schaltnetz

X hängt offenbar nur von $\omega_1, \omega_3, \omega_4$ ab.

ω_1	ω_3	ω_4	$X(\omega)$
0	0	0	0
0	0	1	1
0	1	0	1
1	0	0	0
0	1	1	0
1	0	1	1
1	1	0	0
1	1	1	1

Zufallsvariable

Sei $\mathbf{P}$ ein Wahrscheinlichkeitsmaß auf Ω und X eine Zufallsvariable.

Ist P ein (logisches) Prädikat auf den reellen Zahlen, so schreibt man vereinfacht $\mathbf{P}(P)$ statt

$$\mathbf{P}\Big(\{\omega \in \Omega \,|\, P(X(\omega))\}\Big).$$

Beispiel: Für ein festes $x \in \mathbb{R}$ ist

$$\begin{aligned}
\mathbf{P}(X \leq x) &= \mathbf{P}\big(\{\omega \in \Omega \,|\, X(\omega) \leq x\}\big), \\
\mathbf{P}(X = x) &= \mathbf{P}\big(\{\omega \in \Omega \,|\, X(\omega) = x\}\big).
\end{aligned}$$

Verteilung

Sei $X : \Omega \longrightarrow \mathbb{R}$ eine Zufallsvariable. Die Funktion

$$f_X : \mathbb{R} \longrightarrow [0,1],\, x \mapsto \mathbf{P}(X = x)$$

nennt man **Wahrscheinlichkeitsfunktion**, diskrete Dichtefunktion oder einfach Dichte von X.

Die Funktion

$$F_X : \mathbb{R} \longrightarrow [0,1],\, x \mapsto \mathbf{P}(X \leq x)$$

heißt **Verteilungsfunktion** oder einfach Verteilung von X.

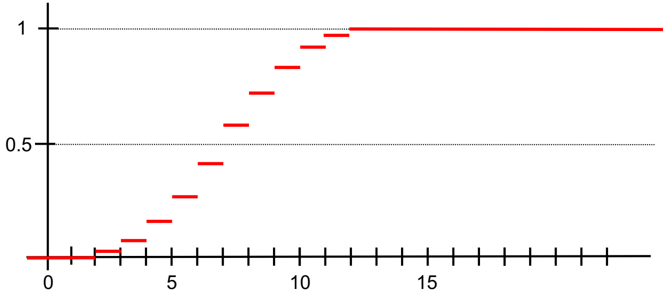
Verteilung

Beispiel: Augensumme beim zweifachen Würfelwurf
$X = X(\omega_1, \omega_2) = \omega_1 + \omega_2$ ist eine Zufallsvariable über $(\Omega, \mathbf{P})$, mit
$\Omega = \{1, \ldots, 6\}^2$ und $\mathbf{P}$ die Gleichverteilung auf Ω.

x	2	3	4	5	6	7	8	9	10	11	12
$f_X(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

(sowie $f_X(x) = 0$ für alle $x \notin \{2, \ldots, 11\}$)

Verteilungsfunktion F_X:



Verteilung

Bei der Untersuchung von Eigenschaften von Zufallsvariablen ist man häufig gar nicht mehr am zugrunde liegenden Wahrscheinlichkeitsraum $(\Omega, \mathbf{P})$ interessiert, sondern lediglich an dem “Bild” $(W_X, \mathbf{P}_X)$ dieses Raumes unter der Abbildung X, wobei $\mathbf{P}_X$ definiert ist durch

$$\mathbf{P}_X(A) = \mathbf{P}(X^{-1}(A)) = \mathbf{P}(\{\omega \in \Omega \,|\, X(\omega) \in A\})$$

für alle $A \subseteq W_X$.

Verteilung

Beispiel: Schaltnetz

Wir nehmen an, dass jedes Gatter mit Wahrscheinlichkeit $\epsilon \geq 0$ fehlerhaft arbeitet, und dass die Fehler unabhängig voneinander auftreten.

Für jedes w gilt dann: $\mathbf{P}(\omega) = \epsilon^{|\omega|} \cdot (1-\epsilon)^{5-|\omega|}$ mit $|w| = \omega_1 + \ldots + \omega_5$.

x	$\mathbf{P}_X(x)$
0	$1 - 2\epsilon + 3\epsilon^2 - 2\epsilon^3$
1	$2\epsilon - 3\epsilon^2 + 2\epsilon^3$

Erwartungswert

Der Erwartungswert einer (diskreten) ZV X ist definiert durch

$$E[X] \stackrel{\text{df}}{=} \sum_{x \in W_X} x \cdot \mathbf{P}(X = x) = \sum_{x \in W_X} x \cdot f_X(x),$$

sofern dieser Wert "existiert", also die Summe auf der rechten Seite konvergiert (was nur für den Fall $|W_X| = \infty$ von Belang ist).

Varianz

Für eine ZV X mit Erwartungswert $\mu = E[X]$ ist die Varianz $V[X]$ definiert durch

$$V[X] \stackrel{\text{df}}{=} E[(X-\mu)^2] = \sum_{x \in W_X} (x-\mu)^2 \cdot \mathbf{P}(X = x).$$

Weiterhin definiert man durch $\sigma \stackrel{\text{df}}{=} \sqrt{V[X]}$ die Standardabweichung von X.

Gemeinsame Verteilung

Seien $X_1, \ldots, X_n$ diskrete ZV über dem Wahrscheinlichkeitsraum $(\Omega, \mathbf{P})$. Man kann diese ZV zu einer Zufallsvariablen X mit Wertebereich

$$W_X \stackrel{\text{df}}{=} W_{X_1} \times \ldots \times W_{X_n}$$

zusammenfassen.

Die Verteilung von X nennt man die **gemeinsame Verteilung** der $X_1, \ldots, X_n$. Diese Verteilung ist bestimmt durch die Angabe der Wahrscheinlichkeiten $f(x_1, \ldots, x_n)$ mit $(x_1, \ldots, x_n) \in W_X$.

Gemeinsame Verteilung

Beispiel: Schaltnetz

$$\omega \mapsto (X(\omega), Y(\omega), Z(\omega))$$

ω_1	ω_2	ω_3	ω_4	ω_5	X	Y	Z
0	0	0	0	0	0	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
1	1	1	1	1	1	0	1

x	y	z	$f(x,y,z)$
0	0	0	0.664
0	0	1	0.002
0	1	0	0.146
1	0	0	0.016
0	1	1	0.074
1	0	1	0.008
1	1	0	0.016
1	1	1	0.074

Randverteilung

Es sei f die Wahrscheinlickeitsfunktion der (diskreten) ZV $X = (X_1, \ldots, X_n)$. Die Wahrscheinlickeitsfunktion f_1 von X_1 ist eine **Randverteilung** (Marginalverteilung) von X:

$$f_1(x_1) = \sum_{(x_2,\ldots,x_n) \in W_{X_2} \times \ldots \times W_{X_n}} f(x_1, x_2, \ldots, x_n).$$

Entsprechend sind die Randverteilungen für Variablen X_i $(2 \leq i \leq n)$ und für Variablen $(X_{i_1}, \ldots, X_{i_k})$ mit $\{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$ definiert.

Randverteilung

Beispiel: Schaltnetz

Mit welcher Wahrscheinlichkeit ist $X = 1$?

$$f_1(1) = f(1,0,0) + f(1,0,1) + f(1,1,0) + f(1,1,1)$$

Mit welcher Wahrscheinlichkeit ist $(X, Y) = (1, 0)$?

$$f_{1,2}(1,0) = f(1,0,0) + f(1,0,1)$$

Unabhängigkeit von Zufallsvariablen

Die Zufallsvariablen $X_1, \ldots, X_n$ heißen unabhängig, wenn für alle $A_i \subseteq W_{X_i}$, $1 \leq i \leq n$, die Ereignisse $(X_i \in A_i)$ unabhängig sind.

Es gilt: Die Zufallsvariablen $X_1, \ldots, X_n$ sind genau dann unabhängig, wenn

$$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_i(x_i)$$

für alle $(x_1, \ldots, x_n) \in W_{X_1} \times \ldots \times W_{X_n}$, wenn f also das (punktweise) Produkt der f_i ist.

Unabhängigkeit von Zufallsvariablen

Beispiel: Schaltnetz

x	y	z	$f(x,y,z)$	$f_1(x)f_2(y)f_3(z)$
0	0	0	0.664	0.557
0	0	1	0.002	0.062
0	1	0	0.146	0.188
1	0	0	0.016	0.021
0	1	1	0.074	0.116
1	0	1	0.008	0.013
1	1	0	0.016	0.039
1	1	1	0.074	0.004

Modellierung

- Modellierung mit Wahrscheinlichkeiten heißt in der Regel, zunächst die **(Zufalls-)Variablen** $X_1, \ldots, X_n$ über einem W-Raum $(\Omega, \mathbf{P})$ festzulegen, für die man sich interessiert (analog zur Logik).
- Ω sollte so gewählt werden, dass sich das Maß $\mathbf{P}$ hierauf einfach spezifizieren lässt.
- Ein vollständiges Modell ist dann gegeben durch die **gemeinsame Verteilung** der Variablen, die man oft auch direkt (ohne $(\Omega, \mathbf{P})$) spezifiziert.
- **Inferenz** besteht aus Operationen wie Marginalisierung, Berechnung bedingter Wahrscheinlichkeiten, etc.
- Das Festlegen der gemeinsamen Verteilung ist allerdings oft zu aufwändig, denn es erfordert die Spezifikation von

$$|W_{X_1}| \cdot |W_{X_2}| \cdot \ldots \cdot |W_{X_n}|$$

  vielen Wahrscheinlichkeiten.

Graphische Modelle

- Graphische Modelle wie z.B. Bayessche Netze vereinfachen die Modellierung, indem die (bedingte) **Unabhängigkeiten** zwischen Variablen ausnutzen.
- Für (evtl. mehrdimensionale) Zufallsvariablen X, Y, Z heißt X **bedingt unabhängig** von Y gegeben Z ($X \perp Y \,|\, Z$), wenn gilt:

$$\mathbf{P}(X, Y \,|\, Z) = \mathbf{P}(X \,|\, Z)\mathbf{P}(Y \,|\, Z)\,,$$

  also $\mathbf{P}(X = x, Y = y \,|\, Z = z) = \mathbf{P}(X = x \,|\, Z = z)\mathbf{P}(Y = y \,|\, Z = z)$ für alle Ausprägungen x, y, z.
- Beispiel: Sind X und Y unabhängig, und gilt $U \perp (X, Y) \,|\, Z$ sowie $V \perp (X, Y, U) \,|\, Z$, dann vereinfacht sich

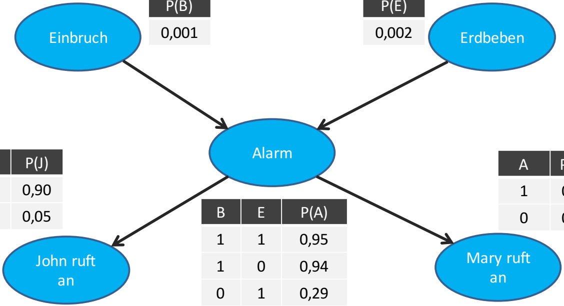
$$\mathbf{P}(X, Y, Z, U, V) = \mathbf{P}(X)\mathbf{P}(Y \,|\, X)\mathbf{P}(Z \,|\, X, Y)\mathbf{P}(U \,|\, X, Y, Z)\mathbf{P}(V \,|\, X, Y, Z, U)$$

  zu

$$\mathbf{P}(X, Y, Z, U, V) = \mathbf{P}(X)\mathbf{P}(Y)\mathbf{P}(Z \,|\, X, Y)\mathbf{P}(U \,|\, Z)\mathbf{P}(V \,|\, Z)$$

Graphische Modelle

Beispiel: 5 binäre (boolesche) Variable X, (bedingte) Wahrscheinlichkeiten nur für Ausprägung "wahr", also $\mathbf{P}(X = 1)$; man erhält $\mathbf{P}(X = 0)$ über Komplementärwahrscheinlichkeit.



P(B)
0,001

P(E)
0,002

B	E	P(A)
1	1	0,95
1	0	0,94
0	1	0,29
0	0	0,001

A	P(J)
1	0,90
0	0,05

A	P(M)
1	0,70
0	0,10

$$\mathbf{P}(B, E, A, J, M) = \mathbf{P}(B)\mathbf{P}(E)\mathbf{P}(A \mid B, E)\mathbf{P}(J \mid A)\mathbf{P}(M \mid A)$$