# 2. allgemeine Wahrscheinlichkeiten

## Diskrete Stochastik

Prof. Dr. Andreas Vogt

Fachhochschule Nordwestschweiz

# Allgemeine Wahrscheinlichkeiten

Kolmogoroff hat erkannt, dass man für eine allgemeine Theorie der Wahrscheinlichkeit die Wahrscheinlichkeiten am besten durch eine beliebige Funktion $P$ vorgibt, welche jedem Ereignis $E \subseteq \Omega$ eine Zahl $P(E)$ zuordnet, sodass bestimmte Bedingungen erfüllt sind, die sogenannten Axiome.

Andrei Nikolajewitsch Kolmogoroff
12.04.1903 – 20.10.1987

## Definition:

Eine Wahrscheinlichkeit (oder: Wahrscheinlichkeitsmass, Wahrscheinlichkeitsverteilung) $P : \underbrace{2^\Omega}_{\text{Potenzmenge von } \Omega \text{, d.h. die Menge aller Teilmengen von } \Omega} \to [0,1]$ erfüllt:

(1) $P(\Omega) = 1$

Je zwei Ereignisse haben kein gemeinsames Element

(2) Für endlich oder abzählbar viele paarweise disjunkte Ereignisse $E_1, E_2, E_3, \ldots$ gilt

$$P(E_1 \cup E_2 \cup E_3 \cup \ldots) = P(E_1) + P(E_2) + P(E_3) + \ldots .$$

## Bemerkungen:

- Diese beiden Bedingungen sind offensichtlich im Laplace-Raum erfüllt.
- Es gibt Situationen, wo nicht jedem Ereignis, sondern nur Ereignissen einer sogenannten $\sigma$-Algebra, eine sinnvolle Wahrscheinlichkeit zugewiesen werden kann. Diese Problematik ist für unsere Zwecke ohne Bedeutung.

Satz: Es sei $P$ eine Wahrscheinlichkeit auf $\Omega$. Dann gilt:

(1) $\forall E \subseteq \Omega : \; P(E^c) = 1 - P(E)$

(2) $P(\emptyset) = 0$

(3) $\forall E_1, E_2 \subseteq \Omega : \; P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

(4) Für eine endliche oder abzählbare Menge $E = \{e_1, e_2, \ldots\}$ gilt

$$P(E) = P(\{e_1\}) + P(\{e_2\}) + \ldots$$

Beweis (von (1)):

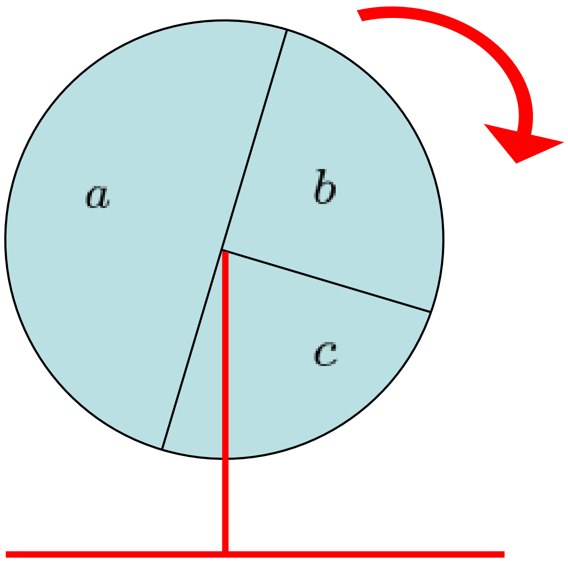
Es gilt $1 \underset{\substack{\uparrow \\ \text{Axiom (1)}}}{=} P(\Omega) = P(E \cup E^c) \underset{\substack{\uparrow \\ \text{Axiom (2)}}}{=} P(E) + P(E^c)$, woraus die Behauptung folgt.

Aus (4) folgt, dass die gesamte Wahrscheinlichkeitsverteilung $P$ durch die Angabe der Wahrscheinlichkeiten für die Elementarereignisse festgelegt ist, falls $\Omega$ endlich oder abzählbar ist.

Dann heisst die Funktion $f_P : \Omega \to [0, 1]$ mit $f_P(\omega) = P(\{\omega\})$ Zähldichte (kurz Z-Dichte) von $P$.

In diesem Fall gilt also $P(E) = \sum_{e \in E} f_P(e)$, insbesondere $\sum_{e \in \Omega} f_P(e) = 1$.

# Allgemeine Wahrscheinlichkeiten




$\Omega = \{a, b, c\}$

Modellierung als Laplace-Experiment scheint nicht sinnvoll.

Bessere Definition von $P : 2^{\Omega} \to [0, 1]$:

$2^{\Omega} = \{\emptyset,$
{a},{b},{c},
{a,b},{a,c},{b,c},
{a,b,c}}

Es reicht die Angabe der Zähldichte, also die Angabe der Wahrscheinlichkeiten für die Elementarereignisse: $f_P(a) = \frac{1}{2}$ $\quad f_P(b) = \frac{1}{4}$ $\quad f_P(c) = \frac{1}{4}$

Wir erhalten so: $P(\emptyset) = 0$

$P(\{a\}) = \frac{1}{2} \qquad P(\{b\}) = \frac{1}{4} \qquad P(\{c\}) = \frac{1}{4}$

$P(\{a, b\}) = f_P(a) + f_P(b) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4} \quad P(\{b, c\}) = f_P(b) + f_P(c) = \frac{1}{2}$

$P(\{a, c\}) = f_P(a) + f_P(c) = \frac{3}{4}$

$P(\{a, b, c\}) = 1$

# bedingte Wahrscheinlichkeit

Würfeln mit einem normalen Würfel: $\Omega = \{1, 2, 3, 4, 5, 6\}$, $P =$ Gleichverteilung

Wie hoch ist die Wahrscheinlichkeit für $A = \{2, 3\}$, wenn ich weiss, dass eine ungerade Zahl gewürfelt worden ist?

Es verbleibt noch die eingeschränkte Ergebnismenge $B = \{1, 3, 5\}$.

Es sind also nur noch $|B| = 3$ Ergebnisse möglich.

Davon ist $|A \cap B| = 1$ Ergebnis günstig.

Die gesuchte Wahrscheinlichkeit ist somit $\dfrac{|A \cap B|}{|B|} = \dfrac{1}{3}$.

Dies führt zu folgender Definition:

**Definition:**

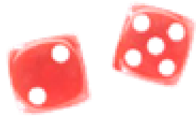
Es sei $\Omega$ eine nichtleere endliche Menge und $P$ eine Wahrscheinlichkeitsverteilung auf $\Omega$. Ferner sei $B \subseteq \Omega$ mit $P(B) > 0$.

Dann heisst

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

(elementare) bedingte Wahrscheinlichkeit von $A$ unter $B$.

Beispiel: Zweimaliges Würfeln



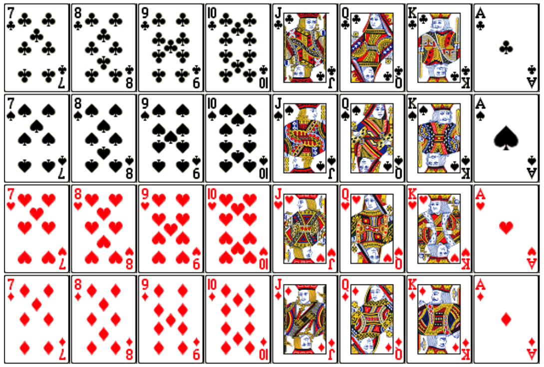
Wie hoch ist Wahrscheinlichkeit, dass die Augensumme 7 ist, wenn das Produkt der Augen 12 ist?

Modellierung: $\Omega = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$

$P =$ Gleichverteilung

$A = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$, $B = \{(2,6),(3,4),(4,3),(6,2)\}$

$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{|A\cap B|}{|B|} = \frac{2}{4} = \frac{1}{2}$



Beispiel: Ziehen aus einem Kartenspiel (32 Karten)

Wie gross ist die Wahrscheinlichkeit, dass ich einen König ziehe, wenn ich weiss, dass ich eine Herzkarte gezogen habe?

Modellierung: $\Omega = \{7_{Karo}, 8_{Karo}, \ldots, Ass_{Kreuz}\}$

$P =$ Gleichverteilung

$A = \{K_{Karo}, K_{Herz}, K_{Pik}, K_{Kreuz}\}$, $B = \{7_{Herz}, 8_{Herz}, \ldots, Ass_{Herz}\}$

$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{|A\cap B|}{|B|} = \frac{1}{8}$

Hier gilt $P(A) = \frac{1}{8} = P(A|B)$.

Die Wahrscheinlichkeit von $A$ wird also nicht durch $B$ beeinflusst.

## Formel von Bayes

Es sei $\Omega$ eine nichtleere endliche Menge und $P$ eine Wahrscheinlichkeitsverteilung auf $\Omega$. Ferner seien $A, B \subseteq \Omega$ mit $P(A) > 0$ und $P(B) > 0$.

Dann gilt $$P(A \mid B) = \frac{P(A)}{P(B)} \cdot P(B \mid A).$$

**Beweis:**

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

Die Formel von Bayes hat sehr wichtige Anwendungen, auch in der Informatik.

Der sogenannte Bayes-Spamfilter basiert etwa darauf.

# bedingte Wahrscheinlichkeit

Von einer bestimmen Population sind 40% weiblich und 60% männlich.
Von den weiblichen Personen wiegen 10% mehr als 70 kg, von den männlichen Personen 80%.
Wieviel Prozent der Gesamtbevölkerung wiegt mehr als 70 kg?

männlich und mehr als 70 kg

Insgesamt wiegen $\boxed{0.1 \cdot 0.4} + \boxed{0.8 \cdot 0.6}$ mehr als 70 kg.

weiblich und mehr als 70 kg

Wahrscheinlickeit für weiblich

Wahrscheinlichkeit für männlich

Andere Interpretation: $\boxed{0.1} \cdot \boxed{0.4} + \boxed{0.8} \cdot \boxed{0.6}$

Wahrscheinlickeit für > 70 kg
unter der Bedingung weiblich

Wahrscheinlickeit für > 70 kg
unter der Bedingung männlich

Allgemein gilt:

Formel von der totalen Wahrscheinlichkeit

Es sei $\Omega$ eine nichtleere endliche Menge und $P$ eine Wahrscheinlichkeitsverteilung auf $\Omega$. Ferner sei $B_i\,(i \in I)$ eine Zerlegung von $\Omega$ (d.h. die $B_i$ sind paarweise disjunkt und $\Omega = \bigcup_{i\in I} B_i$) mit $P(B_i) > 0$.

Dann gilt $P(A) = \sum_{i\in I} P(A \mid B_i) \cdot P(B_i)$

Für $0 < P(A) < 1$ gilt mit $\Omega = A \cup A^c$ insbesondere:

$$P(A \mid B) = \frac{P(A) \cdot P(B \mid A)}{P(B \mid A) \cdot P(A) + P(B \mid A^c) \cdot P(A^c)}.$$

$\rightarrow$ Übung

# bedingte Wahrscheinlichkeit

Beispiel:

Eine bestimmte Krankheit tritt bei ca. zwei von 10.000 Personen auf. Der Sachverhalt $K$, dass ein Mensch diese Krankheit in sich trägt, hat also die Wahrscheinlichkeit $P(K) = 0.0002$.

In einem Screening-Test soll ermittelt werden, welche Personen Träger dieser Krankheit sind. $T$ bezeichne die Tatsache, dass der Test bei einer Person positiv ausgefallen ist. Der Hersteller des Tests versichert, dass der Test die Krankheit zu 99 Prozent erkennt (Sensitivität $P(T \mid K) = 0.99$) und in 99 Prozent der Fälle richtig liegt (Spezifität $P(T^c \mid K^c) = 0.99$), also nur in einem Prozent der Fälle fälschlicherweise anschlägt, obwohl die Krankheit gar nicht vorliegt ($P(T \mid K^c) = 0.01$).

Wir wissen also, mit welcher Wahrscheinlichkeit der Test positiv ausfällt, wenn die Krankheit vorliegt (nämlich mit 99-prozentiger Wahrscheinlichkeit). Die Frage ist: Wie wahrscheinlich ist das Vorliegen der Krankheit, wenn der Test positiv ausfällt?

Gesucht ist also der positive prädiktive Wert $P(K \mid T)$.

$$P(K \mid T) = \frac{P(K) \cdot P(T \mid K)}{P(T \mid K) \cdot P(K) + P(T \mid K^c) \cdot P(K^c)} = \frac{0.0002 \cdot 0.99}{0.99 \cdot 0.0002 + 0.01 \cdot 0.9998} \approx 0.019.$$

Die Wahrscheinlichkeit, dass ein positiv Getesteter tatsächlich krank ist, beträgt in diesem Beispiel also nur rund 1.9 Prozent, d.h. ein positiv Getesteter hat immer noch eine Chance von über 98 Prozent, gesund zu sein, obwohl der Test ihn als krank eingestuft hat!

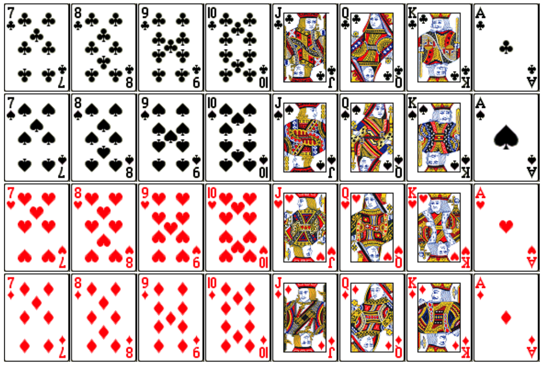
## stochastische Unabhängigkeit

**Definition:** Es sei $P$ eine Wahrscheinlichkeitsverteilung auf $\Omega$.

Zwei Ereignisse $A, B \subseteq \Omega$ heissen stochastisch unabhängig, falls

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

Im Falle $P(B) \neq 0$ ist dies äquivalent zu $\underbrace{P(A|B)}_{\frac{P(A\cap B)}{P(B)}} = P(A)$.



**Beispiel:** Ziehen aus einem Kartenspiel (32 Karten)

Modellierung: $\Omega = \{7_{Karo}, 8_{Karo}, \ldots, Ass_{Kreuz}\}$

$P =$ Gleichverteilung

Das Ereignis "König" ist unabhängig vom Ereignis "Herz".

König$= \{K_{Karo}, K_{Herz}, K_{Pik}, K_{Kreuz}\}$, Herz$= \{7_{Herz}, 8_{Herz}, \ldots, Ass_{Herz}\}$

$P(\text{König} \cap \text{Herz}) = P(\{K_{Herz}\}) = \frac{1}{32}$

$P(\text{König}) \cdot P(\text{Herz}) = \frac{4}{32} \cdot \frac{8}{32} = \frac{1}{8} \cdot \frac{1}{4} = \frac{1}{32}$

Serie 2, A.1–A.9

# Mehrstufige Zufallsexperimente

Beispiel: Urne mit 3 roten und 4 grünen Kugeln

(1) Ziehen einer Kugel

Modellierung: $\Omega_{\text{1-mal}} = \{\text{rot, grün}\}$

$f_{P_{\text{1-mal}}}(\text{rot}) = \frac{3}{7} \quad f_{P_{\text{1-mal}}}(\text{grün}) = \frac{4}{7}$ ⟵ Zähldichte

(2) Ziehen von zwei Kugeln nacheinander (ohne Zurücklegen)

$P(\text{1. Kugel rot} \cap \text{2. Kugel grün}) = P(\text{1. Kugel rot}) \cdot P(\text{2. Kugel grün} \mid \text{1. Kugel rot}) =$

$= \frac{3}{7} \cdot \frac{4}{6}$

Ziehen einer Kugel aus Urne mit 2 roten und 4 grünen Kugeln

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

Modellierung:

$\Omega_{\text{2-mal}} = \{\text{rot, grün}\} \times \{\text{rot, grün}\}$

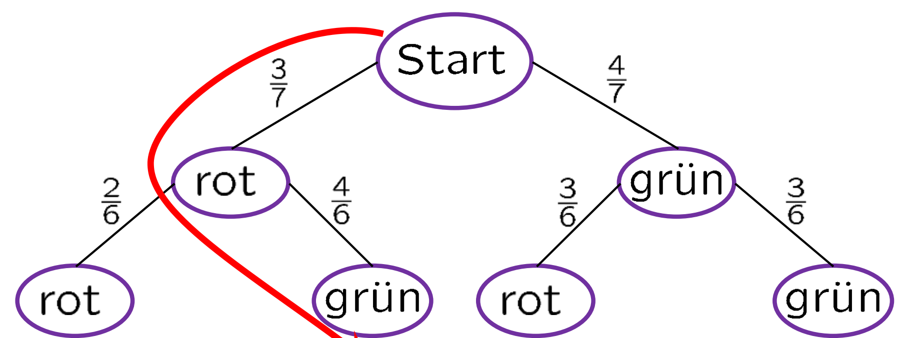
$f_{P_{\text{2-mal}}}(\text{rot, rot}) = \frac{3}{7} \cdot \frac{2}{6}$

$f_{P_{\text{2-mal}}}(\text{rot, grün}) = \frac{3}{7} \cdot \frac{4}{6}$

$f_{P_{\text{2-mal}}}(\text{grün, rot}) = \frac{4}{7} \cdot \frac{3}{6}$

$f_{P_{\text{2-mal}}}(\text{grün, grün}) = \frac{4}{7} \cdot \frac{3}{6}$

Diese Modellierung kann übersichtlich mit Hilfe eines Baumdiagrammes dargestellt werden:




$f_{P_{\text{2-mal}}}(\text{rot,grün}) = \frac{3}{7} \cdot \frac{4}{6}$

Multiplikation der Wahrscheinlichkeiten entlang eines Pfades (manchmal als Pfadregel bezeichnet)

# Mehrstufige Zufallsexperimente

Ein $n$-stufiger Versuch mit Ergebnismenge $\Omega_i$ für den $i$-ten Versuch wird meist wie folgt modelliert:

(1) Man legt die Dichte $f_1(\omega_1)$ auf $\Omega_1$ für den ersten Versuch fest.

(2) Man legt die Dichte $f_2(\omega_2 \mid \omega_1)$ auf $\Omega_2$ für den zweiten Versuch in Abhängigkeit vom Ergebnis $\omega_1$ des ersten Versuchs fest.

(3) Man legt die Dichte $f_3(\omega_3 \mid \omega_1, \omega_2)$ auf $\Omega_3$ für den dritten Versuch in Abhängigkeit der Ergebnisse $(\omega_1, \omega_2)$ der ersten beiden Versuche fest.

$\vdots$

($n$) Man legt die Dichte $f_n(\omega_n \mid \omega_1, \ldots, \omega_{n-1})$ auf $\Omega_n$ für den $n$-ten Versuch in Abhängigkeit der Ergebnisse $(\omega_1, \ldots, \omega_{n-1})$ der ersten $n-1$ Versuche fest.

($n+1$) Die resultierende Dichte auf $\Omega_1 \times \Omega_2 \times \ldots \times \Omega_n$ ist dann
$$f(\omega_1, \ldots, \omega_n) = f_1(\omega_1) \cdot f_2(\omega_2 \mid \omega_1) \cdot \ldots \cdot f_n(\omega_n \mid \omega_1, \ldots, \omega_{n-1}).$$

Wenn die Versuche nicht voneinander abhängen, dann modelliert man die Versuche einzeln, mit Dichten $f_i(\omega_i)$ auf $\Omega_i$, und erhält als resultierende Dichte

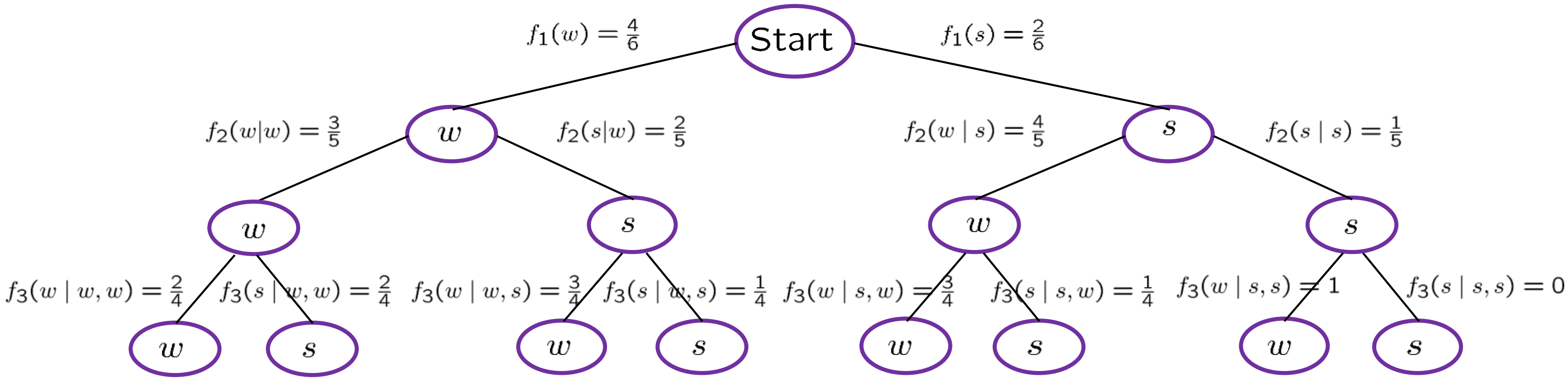
$$f(\omega_1, \ldots, \omega_n) = f_1(\omega_1) \cdot \ldots \cdot f_n(\omega_n).$$

# Mehrstufige Zufallsexperimente

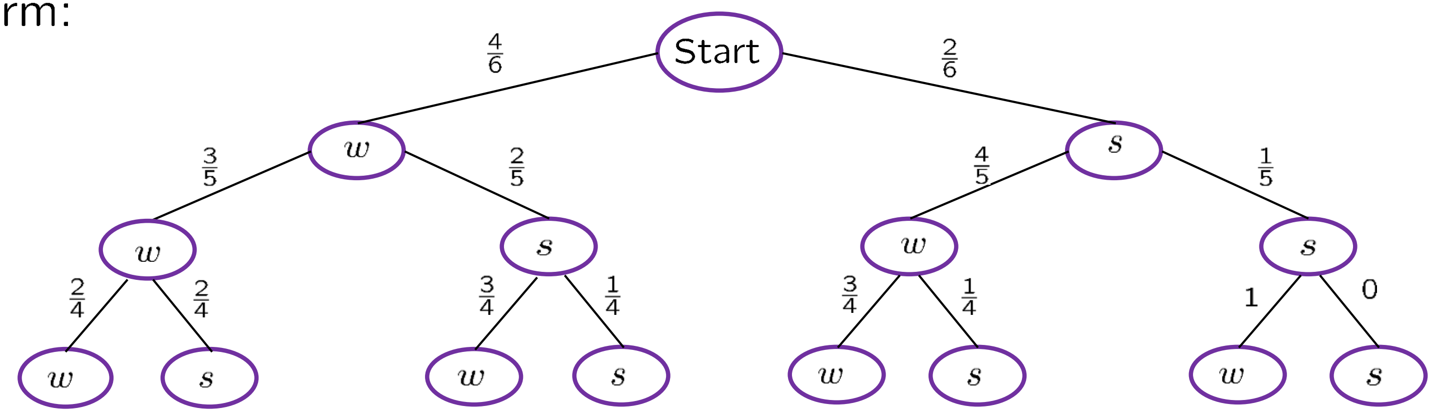
**Beispiel:** Gegeben ist eine Urne mit 4 weissen und 2 schwarzen Kugeln.

Wir ziehen dreimal ohne Zurücklegen.

Modellierung: $\Omega = \{w,s\} \times \{w,s\} \times \{w,s\}$



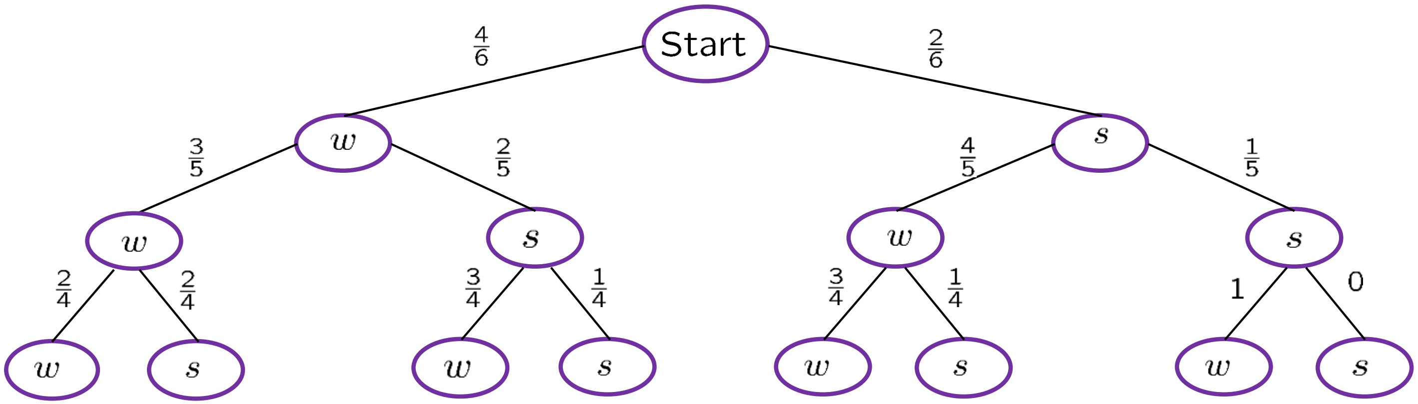

Kurzform:



# Mehrstufige Zufallsexperimente

Beispiel: Gegeben ist eine Urne mit 4 weissen und 2 schwarzen Kugeln.

Wir ziehen dreimal ohne Zurücklegen.




Was ist die Wahrscheinlichkeit für $A$: "Dritte Kugel weiss"?

Als Teilmenge von $\Omega = \{w,s\}\times\{w,s\}\times\{w,s\}$ ist $A = \{(w,w,w),(w,s,w),(s,w,w),(s,s,w)\}$.

Es gilt also:

$P(A) = f(w,w,w)+f(w,s,w)+f(s,w,w)+f(s,s,w)$, wobei $f$ die gemeinsame Dichte ist, die der Multiplikation entlang des Pfades entspricht.

Also: $P(A) = \frac{4}{6}\cdot\frac{3}{5}\cdot\frac{2}{4} + \frac{4}{6}\cdot\frac{2}{5}\cdot\frac{3}{4} + \frac{2}{6}\cdot\frac{4}{5}\cdot\frac{3}{4} + \frac{2}{6}\cdot\frac{1}{5}\cdot 1$

Was ist die Wahrscheinlichkeit für $B$: "Erste Kugel weiss"?

Als Teilmenge von $\Omega = \{w,s\}\times\{w,s\}\times\{w,s\}$ ist $B = \{(w,w,w),(w,s,w),(w,w,s),(w,s,s)\}$.

Also: $P(B) = \frac{4}{6}\cdot\frac{3}{5}\cdot\frac{2}{4} + \frac{4}{6}\cdot\frac{3}{5}\cdot\frac{2}{4} + \frac{4}{6}\cdot\frac{2}{5}\cdot\frac{3}{4} + \frac{4}{6}\cdot\frac{2}{5}\cdot\frac{1}{4} = \frac{4}{6}(\frac{3}{5}\cdot\frac{2}{4} + \frac{3}{5}\cdot\frac{2}{4} + \frac{2}{5}\cdot\frac{3}{4} + \frac{2}{5}\cdot\frac{1}{4}) = \frac{4}{6}.$

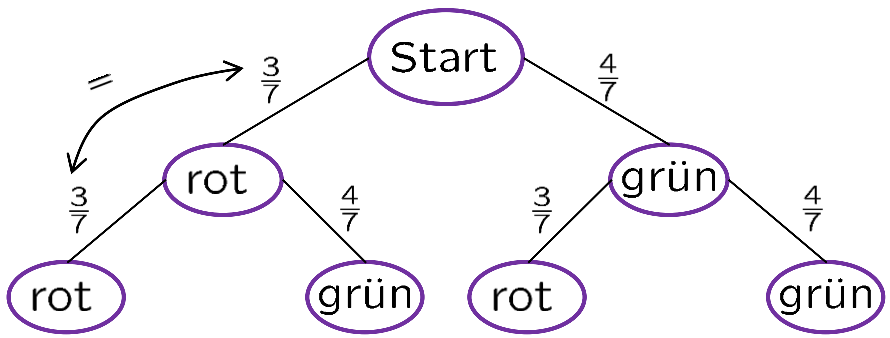
(Das hätte man auch direkt sagen können.)

# Mehrstufige Zufallsexperimente

Beispiel: Urne mit 3 roten und 4 grünen Kugeln

Ziehen von zwei Kugeln nacheinander (**mit Zurücklegen**)

Modellierung als unabhängige Versuche




Man kann leicht nachrechnen, dass das Ereignis $A$ =1. Kugel rot= {(rot,rot), (rot,grün)} unabhängig vom Ereignis $B$ =2. Kugel rot= {(rot,rot), (grün,rot)} ist:

$$P(A \cap B) = P(\{(\text{rot,rot})\}) = \tfrac{3}{7} \cdot \tfrac{3}{7}$$

$$P(A) \cdot P(B) = \underbrace{\left(\tfrac{3}{7} \cdot \tfrac{3}{7} + \tfrac{3}{7} \cdot \tfrac{4}{7}\right)}_{= \frac{3}{7} \cdot \left(\frac{3}{7} + \frac{4}{7}\right)} \cdot \underbrace{\left(\tfrac{3}{7} \cdot \tfrac{3}{7} + \tfrac{4}{7} \cdot \tfrac{3}{7}\right)}_{= \frac{3}{7} \cdot \left(\frac{3}{7} + \frac{4}{7}\right)} = \tfrac{3}{7} \cdot \tfrac{3}{7}$$

Serie 2, A.10–A.14