

SS 2016

Diskrete Wahrscheinlichkeitstheorie

Susanne Albers

Fakultät für Informatik
TU München

<http://wwwalbers.in.tum.de/lehre/2016SS/dwt/index.html.de>

Sommersemester 2016

Kapitel 0 Organisatorisches

- Vorlesungen:
 - Fr 12:00–14:00 und Fr 14:00–15:00 (Interims Hörsaal 1)
Pflichtvorlesung Bachelor IN, Bioinformatik
Modulnr.: IN0018
- Übung:
 - 2SWS Tutorübung: siehe Webseite zur Übung
 - Übungsleitung: Marinus Gottschau, Dennis Kraft, Sebastian Schraink, Richard Stotz
- Umfang:
 - 3V+2TÜ, 6 ECTS-Punkte
- Sprechstunde:
 - nach Vereinbarung

- Übungsaufgaben:
 - Ausgabe jeweils am Freitag auf der Webseite der Vorlesung, ab 18:00 Uhr
 - Abgabe eine Woche später, jeweils Montag bis 10:00 Uhr, Briefkästen Westseite Untergeschoss FMI Magistrale
 - Vorbereitung in der Tutorübung
 - vorauss. 12 Übungsblätter, das letzte am 08. Juli 2016, jedes 20 Punkte
 - **Bonusregelung:** Werden bei den **ersten sechs und zweiten sechs Übungsblättern** jeweils mindestens 50% der insgesamt erreichbaren Punkte erzielt, so verbessert sich die Note einer *bestanden* Klausur um 1/3 Notenstufe.
- Klausur:
 - Klausur am **03. August 2016, 10:30–12:30 Uhr**
 - Wiederholungsklausur am **11. Oktober 2016, 13:30–15:30 Uhr**
 - bei den Klausuren sind *keine* Hilfsmittel außer einem handbeschriebenen DIN-A4-Blatt zugelassen

- Vorkenntnisse:
 - Einführung in die Informatik I/II
 - Diskrete Strukturen
- Weiterführende Vorlesungen:
 - Effiziente Algorithmen und Datenstrukturen
 - Randomisierte Algorithmen
 - Online- und Approximationsalgorithmen
 - Komplexitätstheorie
 - ...
- Webseite:

<http://www.albers.in.tum.de/lehre/2016SS/dwt/index.html.de>

1. Vorlesungsinhalt

- Diskrete Wahrscheinlichkeitsräume
 - Wahrscheinlichkeitsraum, Ereignis, Zufallsvariable
 - spezielle Verteilungen
 - Ungleichungen von Markov und Chebyshev
- Kontinuierliche Wahrscheinlichkeitsräume
 - Normalverteilung, Exponentialverteilung
 - Zentraler Grenzwertsatz
- Statistik
 - Schätzvariablen
 - Konfidenzintervalle
 - Testen von Hypothesen
- Stochastische Prozesse
 - Markovketten
 - Warteschlangen

2. Literatur

-  T. Schickinger, A. Steger:
Diskrete Strukturen - Band 2,
Springer Verlag, 2001
-  M. Greiner, G. Tinhofer:
Stochastik für Informatiker,
Carl Hanser Verlag, 1996
-  H. Gordon:
Discrete Probability,
Springer-Verlag, 1997
-  M. Mitzenmacher, E. Upfal:
Probability and Computing: Randomized Algorithms and Probabilistic Analysis,
Cambridge University Press, 2005

-  R. Motwani, P. Raghavan:
Randomized Algorithms,
Cambridge University Press, 1995
-  M. Hofri:
Probabilistic Analysis of Algorithms,
Springer Verlag, 1987
-  L. Fahrmeir, R. Künstler, I. Pigeot, G. Tutz:
Statistik - Der Weg zur Datenanalyse,
Springer-Verlag, 1997

3. Einleitung

Was bedeutet Zufall?

- Unkenntnis über den Ausgang eines durchgeführten Experiments
- Ein Experiment wird vielfach mit eventuell sich änderndem Ergebnis ausgeführt
- Ereignisse stehen in keinem kausalen Zusammenhang
- physikalischer Zufall (Rauschen, Kernzerfall)

Zufall in der *diskreten* Informatik

- Die Eingabe für einen bestimmten Algorithmus wird aus einer großen Menge möglicher Eingaben **zufällig** gewählt:

average case

- Kombination von Worst-Case- und Average-Case-Analyse, in der Eingaben gemäß einer Verteilung leicht *pertubiert* werden:

smoothed analysis

- Der Algorithmus verwendet Zufallsbits, um mit großer Wahrscheinlichkeit gewisse **Problemsituationen** zu vermeiden:

Randomisierung

Kapitel I Diskrete Wahrscheinlichkeitsräume

1. Grundlagen

Definition 1

- 1 Ein **diskreter Wahrscheinlichkeitsraum** ist durch eine **Ergebnismenge** $\Omega = \{\omega_1, \omega_2, \dots\}$ von **Elementarereignissen** gegeben.
- 2 Jedem Elementarereignis ω_i ist eine **(Elementar-)Wahrscheinlichkeit** $\Pr[\omega_i]$ zugeordnet, wobei wir fordern, dass $0 \leq \Pr[\omega_i] \leq 1$ und

$$\sum_{\omega \in \Omega} \Pr[\omega] = 1.$$

- 3 Eine Menge $E \subseteq \Omega$ heißt **Ereignis**. Die Wahrscheinlichkeit $\Pr[E]$ eines Ereignisses ist durch

$$\Pr[E] := \sum_{\omega \in E} \Pr[\omega]$$

definiert.

Beispiel 2

Zwei faire Würfel (einer weiß, einer schwarz) werden geworfen. Wir sind an der Gesamtzahl der angezeigten Augen interessiert:

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

- 1 Die Wahrscheinlichkeit $\Pr((i, j))$ eines jeden Elementarereignisses (i, j) ist $\frac{1}{36}$.
- 2 Die Wahrscheinlichkeit $\Pr(E)$ des Ereignisses

$$E = \{\text{Die Gesamtzahl der Augen ist } 10\}$$

ist $\frac{1}{12}$.

Wir hätten aber auch sagen können:

$$\Omega = \{2, 3, 4, \dots, 10, 11, 12\}$$

Die Wahrscheinlichkeiten der Elementarereignisse sind dann aber nicht mehr gleich. Es ist z.B.

- ① $\Pr(2) = \frac{1}{36}$;
- ② $\Pr(4) = \frac{1}{12}$;
- ③ $\Pr(7) = \frac{1}{6}$.

Beispiel 3

Eine faire Münze wird so lange geworfen, bis die gleiche Seite zweimal hintereinander fällt. Dann ist

$$\Omega = \{hh, tt, htt, thh, thtt, hthh, httht, ththh, \dots\}$$

Frage: Was sind die Wahrscheinlichkeiten der einzelnen Elementarereignisse?

\bar{E} heißt **komplementäres Ereignis** zu E .

Allgemein verwenden wir bei der Definition von Ereignissen alle bekannten Operatoren aus der Mengenlehre. Wenn also A und B Ereignisse sind, dann sind auch $A \cup B$, $A \cap B$, $A \setminus B$ etc. Ereignisse.

Zwei Ereignisse A und B heißen **disjunkt** oder auch **unvereinbar**, wenn $A \cap B = \emptyset$ gilt.

Definition 4

$$\begin{aligned} \text{relative H\u00e4ufigkeit von } E & := \frac{\text{absolute H\u00e4ufigkeit von } E}{\text{Anzahl aller Beobachtungen}} \\ & = \frac{\text{Anzahl Eintreten von } E}{\text{Anzahl aller Beobachtungen}}. \end{aligned}$$

Definition 5

Ein Wahrscheinlichkeitsraum mit $\Omega = \{\omega_1, \dots, \omega_n\}$ heißt **endlicher Wahrscheinlichkeitsraum**.

Bei unendlichen Wahrscheinlichkeitsräumen werden wir gewöhnlich nur den Fall $\Omega = \mathbb{N}_0$ betrachten. Dies stellt keine große Einschränkung dar, da wir statt einer Ergebnismenge $\Omega = \{\omega_1, \omega_2, \dots\}$ auch \mathbb{N}_0 als Ergebnismenge verwenden können, indem wir ω_i mit $i - 1$ identifizieren. Wir sagen, dass durch die Angabe der Elementarwahrscheinlichkeiten ein **Wahrscheinlichkeitsraum auf Ω** definiert ist.

Beispiel 6

Wir beobachten die an einer Straße in Bayern vorbeifahrenden Autos. Dabei gelte:

- 1 Es fahren doppelt so viele Autos von links nach rechts wie von rechts nach links.
- 2 Von zehn Autos haben zwei die Farbe hellelfenbein, die übrigen eine andere Lackierung.

- Das Ereignis “*Wir beobachten ein von links nach rechts fahrendes Auto*” hat die Wahrscheinlichkeit $\frac{2}{3}$.
- Das Ereignis “*Das nächste Auto ist ein Taxi von rechts*” passiert mit Wahrscheinlichkeit

$$\frac{1}{3} \cdot \frac{1}{5} \cdot$$

Beispiel 7 (Unendlicher Wahrscheinlichkeitsraum)

Wir betrachten eine Münze, die mit Wahrscheinlichkeit p Kopf zeigt und mit Wahrscheinlichkeit $q := 1 - p$ Zahl.

Wir führen Versuche aus, indem wir die Münze wiederholt solange werfen, bis *Zahl* fällt. Das *Ergebnis* eines solchen Versuchs ist die Anzahl der durchgeführten Münzwürfe. Damit ergibt sich hier als Ergebnismenge

$$\Omega = \mathbb{N} = \{1, 2, 3, \dots\} .$$

Beispiel 7 (Forts.)

Sei, für $i \in \mathbb{N}$, ω_i das Elementarereignis

$\omega_i \hat{=} \text{Die Münze wird } i\text{-mal geworfen.}$

Dann gilt:

$$\Pr[\omega_i] = p^{i-1}q,$$

und

$$\sum_{\omega \in \Omega} \Pr[\omega] = \sum_{i=1}^{\infty} p^{i-1}q = q \cdot \sum_{i=0}^{\infty} p^i = \frac{q}{1-p} = 1.$$

(wie es sein soll!)

Lemma 8

Für Ereignisse A, B, A_1, A_2, \dots gilt:

- 1 $\Pr[\emptyset] = 0, \Pr[\Omega] = 1.$
- 2 $0 \leq \Pr[A] \leq 1.$
- 3 $\Pr[\bar{A}] = 1 - \Pr[A].$
- 4 Wenn $A \subseteq B$, so folgt $\Pr[A] \leq \Pr[B].$

Lemma 8 (Forts.)

- ⑤ (Additionssatz) Wenn die Ereignisse A_1, \dots, A_n paarweise disjunkt sind (also wenn für alle Paare $i \neq j$ gilt, dass $A_i \cap A_j = \emptyset$), so folgt

$$\Pr \left[\bigcup_{i=1}^n A_i \right] = \sum_{i=1}^n \Pr[A_i].$$

Für disjunkte Ereignisse A, B erhalten wir insbesondere

$$\Pr[A \cup B] = \Pr[A] + \Pr[B].$$

Für eine unendliche Menge von disjunkten Ereignissen A_1, A_2, \dots gilt analog

$$\Pr \left[\bigcup_{i=1}^{\infty} A_i \right] = \sum_{i=1}^{\infty} \Pr[A_i].$$

Beweis:

Die Aussagen folgen unmittelbar aus Definition 1, den Eigenschaften der Addition und der Definition der Summe. □

Eigenschaft 5 in Lemma 8 gilt nur für disjunkte Ereignisse. Für den allgemeinen Fall erhalten wir folgenden

Satz 9 (Siebformel, Prinzip der Inklusion/Exklusion)

Für Ereignisse A_1, \dots, A_n ($n \geq 2$) gilt:

$$\begin{aligned} \Pr \left[\bigcup_{i=1}^n A_i \right] &= \sum_{i=1}^n \Pr[A_i] - \sum_{1 \leq i_1 < i_2 \leq n} \Pr[A_{i_1} \cap A_{i_2}] + \dots \\ &+ (-1)^{l-1} \sum_{1 \leq i_1 < \dots < i_l \leq n} \Pr[A_{i_1} \cap \dots \cap A_{i_l}] + \dots \\ &+ (-1)^{n-1} \cdot \Pr[A_1 \cap \dots \cap A_n]. \end{aligned}$$

Satz 9 (Forts.)

Insbesondere gilt für zwei Ereignisse A und B

$$\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B] .$$

Für drei Ereignisse A_1 , A_2 und A_3 erhalten wir

$$\begin{aligned} \Pr[A_1 \cup A_2 \cup A_3] &= \Pr[A_1] + \Pr[A_2] + \Pr[A_3] \\ &\quad - \Pr[A_1 \cap A_2] - \Pr[A_1 \cap A_3] \\ &\quad - \Pr[A_2 \cap A_3] \\ &\quad + \Pr[A_1 \cap A_2 \cap A_3] . \end{aligned}$$

Beweis:

Wir betrachten zunächst den Fall $n = 2$. Dazu setzen wir $C := A \setminus B = A \setminus (A \cap B)$. Gemäß dieser Definition gilt, dass C und $A \cap B$ sowie C und B disjunkt sind. Deshalb können wir Eigenschaft 5 von Lemma 8 anwenden:

$$\Pr[A] = \Pr[C \cup (A \cap B)] = \Pr[C] + \Pr[A \cap B].$$

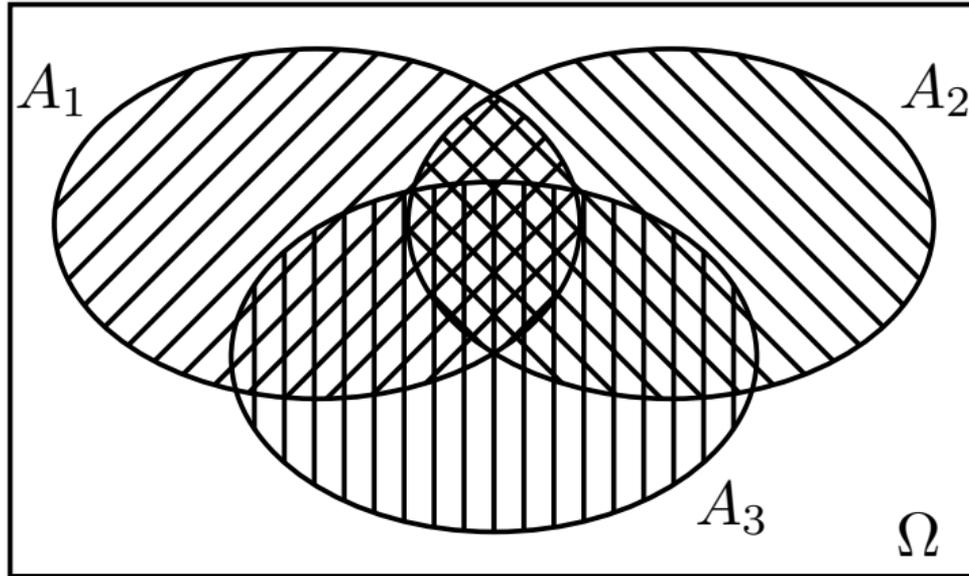
Wegen $A \cup B = C \cup B$ folgt daraus

$$\begin{aligned}\Pr[A \cup B] &= \Pr[C \cup B] = \Pr[C] + \Pr[B] = \\ &\Pr[A] - \Pr[A \cap B] + \Pr[B]\end{aligned}$$

und wir haben die Behauptung für $n = 2$ gezeigt.

Beweis (Forts.):

Der Fall $n = 3$:



Man beachte, dass durch die im Satz angegebene Summe jedes Flächenstück insgesamt genau einmal gezählt wird.

Beweis (Forts.):

Der allgemeine Fall kann nun durch Induktion über n gezeigt werden (was wir aber hier nicht ausführen!).

Satz 9 findet man manchmal auch unter der Bezeichnung *Satz von Poincaré-Sylvester*, nach dem Franzosen

Jules Henri Poincaré (1854–1912)

und dem Engländer

James Joseph Sylvester (1814–1897)

benannt.

Boolesche Ungleichung:

Die folgende Abschätzung ist nach **George Boole** (1815–1864) benannt:

Korollar 10

Für Ereignisse A_1, \dots, A_n gilt

$$\Pr \left[\bigcup_{i=1}^n A_i \right] \leq \sum_{i=1}^n \Pr[A_i] .$$

Analog gilt für eine unendliche Folge von Ereignissen A_1, A_2, \dots , dass

$$\Pr \left[\bigcup_{i=1}^{\infty} A_i \right] \leq \sum_{i=1}^{\infty} \Pr[A_i] .$$

Beweis:

Zunächst betrachten wir die linke Seite der Ungleichung für den endlichen Fall und erhalten

$$\Pr \left[\bigcup_{i=1}^n A_i \right] = \sum_{\omega \in \bigcup_{i=1}^n A_i} \Pr[\omega] .$$

Für die rechte Seite gilt

$$\sum_{i=1}^n \Pr[A_i] = \sum_{i=1}^n \sum_{\omega \in A_i} \Pr[\omega] .$$

Jedes Elementarereignis kommt links also genau einmal und rechts mindestens einmal vor. □

1.1 Wahl der Wahrscheinlichkeiten

Frage: Wie können Wahrscheinlichkeiten sinnvoll festgelegt werden?

Prinzip von Laplace (Pierre Simon Laplace (1749–1827)): *Wenn nichts dagegen spricht, gehen wir davon aus, dass alle Elementarereignisse gleich wahrscheinlich sind.*

Also:

$$\Pr[E] = \frac{|E|}{|\Omega|}$$

1.2 Historische Anfänge der Wahrscheinlichkeitstheorie

Die ersten Hinweise auf mathematische Untersuchungen zu Problemen der Wahrscheinlichkeitstheorie finden sich in einem Briefwechsel zwischen den französischen Mathematikern

Pierre Fermat (1601–1665)

und

Blaise Pascal (1623–1662).

Pascal beschäftigte sich neben der Mathematik auch mit Fragestellungen aus dem Bereich der Physik und auch aus der Informatik! Sein Vater hatte als Steuerinspektor in Rouen umfangreiche Rechnungen durchzuführen und so wurde Pascal zum Bau einer mechanischen Rechenmaschine, der so genannten *Pascaline*, motiviert.

In dem Briefwechsel taucht bereits der Ansatz $\Pr[E] = |E|/|\Omega|$ zur Berechnung der Wahrscheinlichkeit von E auf. Auch den Begriff des Erwartungswerts kann man dort schon finden. Weder Fermat noch Pascal publizierten ihre Überlegungen zur Wahrscheinlichkeitstheorie. Der Niederländer

Christiaan Huygens (1629–1695)

entwickelte ebenfalls Methoden zum Arbeiten mit Wahrscheinlichkeiten aus. Er publizierte im Jahre 1657 auch eine kleine Arbeit mit dem Titel „De ratiociniis in ludo aleae“ (Über die Gesetzmäßigkeiten beim Würfelspiel).

2. Bedingte Wahrscheinlichkeiten

Beispiel 11

A und B spielen Poker (52 Karten, 5 Karten pro Spieler, keine getauschten Karten). A hält vier Asse und eine Herz Zwei in der Hand. B kann dieses Blatt nur überbieten, wenn er einen Straight Flush (fünf Karten *einer* Farbe in aufsteigender Reihenfolge) hat. Die Wahrscheinlichkeit für das Ereignis $F :=$ „B hat einen Straight Flush“ beträgt

$$\Pr[F] = \frac{|F|}{|\Omega|} = \frac{3 \cdot 8 + 7}{\binom{52-5}{5}} = \frac{31}{1533939} = 2,02.. \cdot 10^{-5}.$$

Beispiel 11 (Forts.)

A hat die Karten allerdings gezinkt und weiß, dass B nur Kreuz in der Hand hält. Bezeichne nun Ω' den Wahrscheinlichkeitsraum aller Möglichkeiten für B und F' das Ereignis, dass B einen Straight Flush der Farbe Kreuz hat:

$$\Pr[F'] = \frac{|F'|}{|\Omega'|} = \frac{8}{\binom{12}{5}} = \frac{8}{792} \approx 0,01 !!$$

Für $\Pr[A|B]$ erforderliche Eigenschaften:

- 1 $\Pr[B|B] = 1$;
- 2 $\Pr[A|\Omega] = \Pr[A]$;
- 3 für festes B ist $\Pr[A|B]$ proportional zu $\Pr[A \cap B]$.

Definition 12

A und B seien Ereignisse mit $\Pr[B] > 0$. Die **bedingte Wahrscheinlichkeit** $\Pr[A|B]$ von A gegeben B ist definiert als

$$\Pr[A|B] := \frac{\Pr[A \cap B]}{\Pr[B]} .$$

Die bedingten Wahrscheinlichkeiten $\Pr[\cdot|B]$ bilden für ein beliebiges Ereignis $B \subseteq \Omega$ mit $\Pr[B] > 0$ einen neuen Wahrscheinlichkeitsraum über Ω .

Es ist leicht nachzurechnen, dass dadurch die Definition eines *diskreten Wahrscheinlichkeitsraums* erfüllt ist:

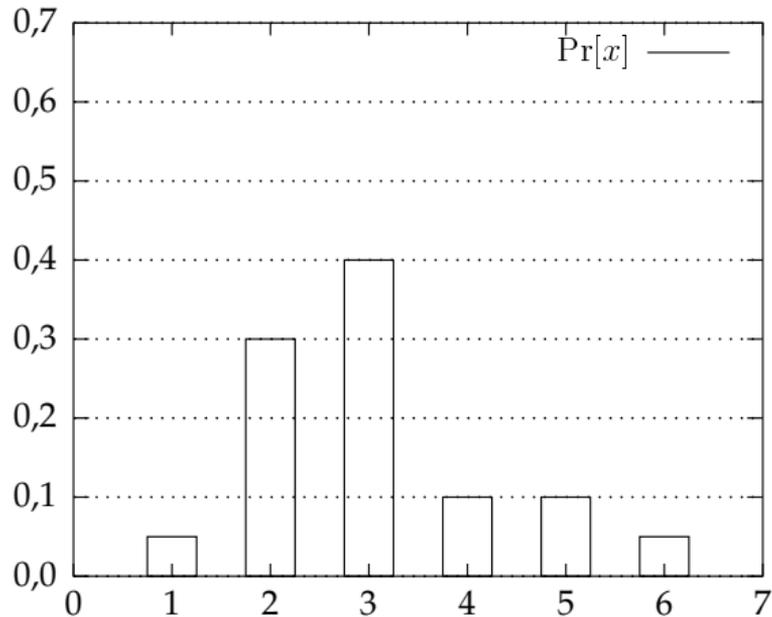
$$\sum_{\omega \in \Omega} \Pr[\omega|B] = \sum_{\omega \in \Omega} \frac{\Pr[\omega \cap B]}{\Pr[B]} = \sum_{\omega \in B} \frac{\Pr[\omega]}{\Pr[B]} = \frac{\Pr[B]}{\Pr[B]} = 1.$$

Damit gelten alle Rechenregeln für Wahrscheinlichkeiten auch für bedingte Wahrscheinlichkeiten. Beispielsweise:

$$\Pr[\emptyset|B] = 0 \text{ sowie } \Pr[\bar{A}|B] = 1 - \Pr[A|B].$$

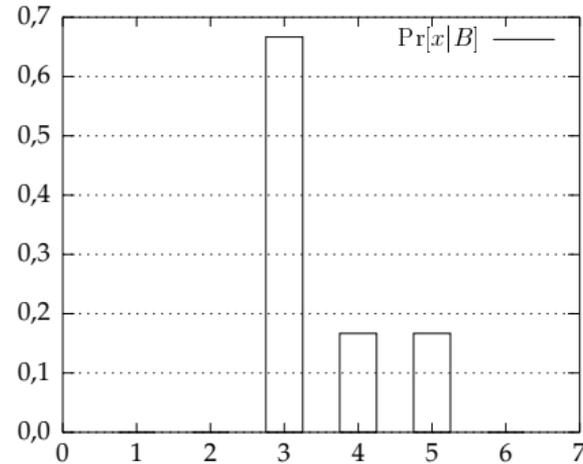
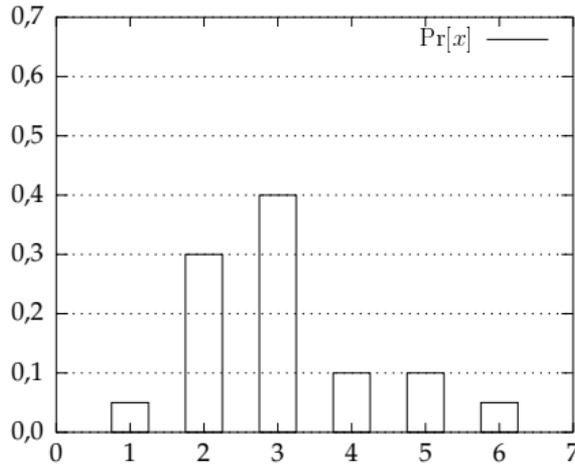
Beispiel 13 (Reskalierung bei bedingten Wahrscheinlichkeiten)

Betrachte folgenden **gezinkten** Würfel:



Beispiel 13 (Forts.)

Wir betrachten nun den durch $B := \{3, 4, 5\}$ gegebenen bedingten Wahrscheinlichkeitsraum:



Was genau war die Bedingung?

Beispiel 14 (Zweikinderproblem)

Wir nehmen an, dass bei der Geburt eines Kindes beide Geschlechter gleich wahrscheinlich sind. Wir wissen, dass eine bestimmte Familie zwei Kinder hat und eines davon ein Mädchen ist. Wie groß ist die Wahrscheinlichkeit, dass beide Kinder der Familie Mädchen sind?

Natürlich $\frac{1}{2}$.

Wirklich?

Beispiel 14 (Forts.)

Eigentlich gilt:

$$\Omega := \{mm, mj, jm, jj\}$$

und

$$M := \{mm, mj, jm\}.$$

Wir bedingen auf M , und damit gilt für $A := \{mm\}$:

$$\Pr[A|M] = \frac{\Pr[A \cap M]}{\Pr[M]} = \frac{1/4}{3/4} = \frac{1}{3}.$$

Beispiel 15 (Ziegenproblem)

Sie nehmen an einer Spielshow im Fernsehen teil, bei der Sie eine von drei verschlossenen Türen auswählen sollen. Hinter einer Tür wartet der Preis, ein Auto, hinter den beiden anderen stehen Ziegen. Sie zeigen auf eine Tür, sagen wir Nummer eins. Sie bleibt vorerst geschlossen. Der Moderator weiß, hinter welcher Tür sich das Auto befindet; mit den Worten “Ich gebe Ihnen mal einen kleinen Hinweis” öffnet er eine andere Tür, zum Beispiel Nummer drei, und eine Ziege schaut heraus und meckert. Er fragt: “Bleiben Sie bei Nummer eins, oder wählen sie Nummer zwei? ”

Frage: Welche Strategie ist günstiger:

- S1 Der Spieler bleibt immer bei seiner ursprünglichen Wahl.
- S2 Der Spieler wechselt stets die ausgewählte Tür.

Beispiel (Forts.)

Wir betrachten hier eine Diskussion des Ziegenproblems mit Hilfe von bedingten Wahrscheinlichkeiten. Wir betrachten bei jeder Variante den Fall, dass der Spieler

- a) die "richtige",
- b) eine falsche Tür gewählt hat.

Ersteres geschieht mit Wahrscheinlichkeit $\frac{1}{3}$, Letzteres mit Wahrscheinlichkeit $\frac{2}{3}$. Mit der vom Moderator gegebenen Information ergeben sich für die beiden Strategien die folgenden Gewinnwahrscheinlichkeiten:

	S1	S2
a)	?	?
b)	?	?