einen Pasch. Allerdings ist die Wahrscheinlichkeit nicht $\frac{6}{21}$, sondern $\frac{6}{36}$. Ähnlich ist die W.K., dass ein Würfel eine 1, ein anderer eine 2 zeigt, nicht $\frac{1}{21}$ sondern $\frac{2}{36}$. Es gibt also 15 Elementarereignisse, die jeweils mit Wahrscheinlichkeit $\frac{1}{18}$ eintreten, und 6 Elementarereignisse, die jeweils mit W.K. $\frac{1}{36}$ eintreten (die Paschs), in der Summe gibt das 1. Es handelt sich hierbei also nicht um ein Laplace-Experiment.

1.4 Einschub: Kombinatorik

1.5 Beispiel eines unendlichen Ereignisraumes

Hier sei $\Omega = \{0, 1, \ldots\}$ und $\lambda > 0$. Die Wahrscheinlichkeit für das Eintreten eines Elementarereignisses $\{i\}$ sei

$$P(i) = \frac{\lambda^i}{i!} e^{-\lambda}$$

$A = \{1, 4, 7\}$

$P(A) = P(1) + P(4) + P(7)^9$

und

$$P(A) = \sum_{i \in A} P(i).$$

Offenbar gilt (K1) und (K3). Nicht ganz klar ist (K2): Benutze den bekannten Grenzwert

$$\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{\lambda},$$

also

$$P(\Omega) = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} = 1.$$

Wir werden diesem Wahrscheinlichkeitsmaß noch unter dem Stichwort "Poisson-Verteilung" begegnen.

1.6 Elementare Eigenschaften eines Wahrscheinlichkeitsmaßes

Satz 1.9. *Sei (Ω, P) ein Wahrscheinlichkeitsraum (d.h. Ω ist höchstens abzählbar unendlich). Dann gilt:*

1. $P(\Omega \setminus A) = 1 - P(A)$.
2. $P(\{\ \}) = 0$.
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
4. *Wenn $A \subseteq B$, dann gilt $P(A) \leq P(B)$.*

Es ist wichtig zu bemerken, dass diese Eigenschaften aus den Axiomen in Definition 1.6 folgen, also beweisbar sind. Sie sind nicht Bestandteil des Axiomensystems.

1.7 Geburtstagsparadoxon: Anwendung in der Kryptographie

Angenommen, jedes Dokument wird mit einem sogenannten **Hashwert**, der eine Zahl h mit $1 \leq h \leq m$ ist, "unterschrieben". Diese Zahl berechnet sich aus dem Dokument (ist also so etwas wie eine Prüfziffer). Wie groß ist die W.K., dass bei n unterzeichneten Dokumenten alle verwendeten Hashwerte verschieden sind? Warum ist das wichtig? Hashwerte werden zur digitalen Signatur benutzt, d.h. eine Person (nennen wir sie Alice) unterzeichnet ein Dokument so, dass sie erst den Hashwert berechnet und dann den Hashwert digital unterschreibt. Der Empfänger, sagen wir Bob, kann überprüfen, ob die Signatur korrekt ist (dazu sagen wir hier nichts), und wenn ja, überprüft er, ob der Hashwert zum Dokument passt (denn sonst hätte unterwegs ja jemand das Dokument austauschen können!). Nun könnte aber ein "bad guy", sagen wir Eve, in betrügerischer Absicht ganz viele harmlose Varianten desjenigen Dokumentes erzeugen, das Al-

ice unterschreiben soll. Gleichzeitig erzeugt Eve viele Varianten eines gefälschten Dokumentes, zusammen mit den Hashwerten. Wenn sie ein Paar (S, T) mit identischen Hashwerten gefunden hat (eine sogenannte Kollision), wobei S das richtige und T das gefälschte Dokument ist, legt sie S und den Hashwert von S Alice zur Unterschrift vor, und dann tauscht sie beim Senden an Bob den Text S durch T aus. Weil der Hashwert von T mit dem von S übereinstimmt, erkennt Bob die betrügerische Absicht von Eve nicht.

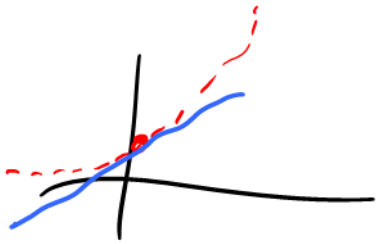
Wie groß ist also die W.K. für eine Kollision, wenn man m Hashwerte hat, d.h. wie leicht ist es für Eve, zwei Dokumente mit gleichem Hashwert zu finden?

Es werden n Dokumente zufällig gewählt, und dann sind auch die Hashwerte zufällig. Wir können uns das als ein Urnenexperiment vorstellen, wobei in der Urne m verschiedene Kugeln liegen (die Hashwerte) und es werden, mit Zurücklegen, n gezogen. Dann ist die An-

$m = 365$

$n = 50$

Normal



zahl "günstiger" Ausgänge ohne Kollision

$$m(m-1)(m-2)\cdots(m-n+1) = \frac{m!}{(m-n)!}.$$

Das ist die Anzahl injektiver Abbildungen einer n-elementigen auf eine m-elementige Menge. Die Anzahl aller möglichen Ziehungen ist m^n. Das liefert die gesuchte W.K.

$$\begin{aligned} P &= \frac{m!}{(m-n)!m^n} = \frac{m}{m}\cdot\frac{m-1}{m}\cdots\frac{m-n+1}{m} \\ &= \left(1-\frac{0}{m}\right)\left(1-\frac{1}{m}\right)\left(1-\frac{2}{m}\right)\cdots\left(1-\frac{n-1}{m}\right). \end{aligned}$$

Ist n/m klein, benutzen wir $e^x \approx 1+x$ und erhalten

$$P \approx e^{[-0-1-2-\ldots-(n-1)]/m} = e^{-\frac{n(n-1)}{2m}}.$$

Wann ist dieser Wert z.B. $1/2$? Er ist $1/2$ wenn

$$-\frac{n(n-1)}{2m} = \ln(\frac{1}{2}) = -\ln(2),$$

also

$$n^2 - n = 2m \ln(2)$$

gilt, d.h.

$$n = \frac{1}{2} + \sqrt{\frac{1}{4} + 2m \ln(2)} \approx \sqrt{2 \ln(2) m}$$

für große m. Für $m = 365$ erhalten wir $n \approx 23$, also in einer Gruppe von 23 Menschen gibt es mit W.K. etwa $\frac{1}{2}$ zwei, die am selben Tag Geburtstag feiern (daher der Name Geburtstagsparadoxon).

1.8 Bedingte Wahrscheinlichkeiten

Definition 1.10. Zwei Ereignisse A und B heißen **unabhängig**, wenn

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

gilt.

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

$P(A \cup B)$

$= P(A) + P(B)$

falls $A \cap B = \{\}$

Beachten Sie den Unterschied zu unvereinbaren Ereignissen A und B (da gilt $P(A \cap B) = 0$.
Anschaulich soll bedingte Wahrscheinlichkeit folgendes bedeuten: Wir wollen $P(B)$ bestimmen, wenn wir schon wissen, dass A eingetreten ist.

Definition 1.11. Es seien A und B Ereignisse eines Ereignisraums Ω, auf dem ein Wahrscheinlichkeitsmaß P definiert ist. Ferner sei $P(B) \neq 0$. Dann definieren wir

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

Gesprochen: P von A gegeben B. Wir nennen $P(A|B)$ die **bedingte** Wahrscheinlichkeit.

Bemerkung 1.12. Es gilt

$$\begin{aligned} P(A \cap B) &= P(B) \cdot P(A|B) \\ &= P(A) \cdot P(B|A). \end{aligned}$$

Satz 1.13 (Satz von Bayes)**.**

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

sofern $P(A), P(B) \neq 0$.

Wenn zwei Ereignisse A und B unabhängig voneinander sind, gilt

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

Manchmal nennt man $P(A)$ die "a-priori" Wahrscheinlichkeit und $P(A|B)$ die "a-posteriori"-Wahrscheinlichkeit. Das Problem, bedingte Wahrscheinlichkeiten vernünftig zu interpretieren, ist folgendes: Von Wahrscheinlichkeiten reden wir, wenn ein Zufallsexperiment durchgeführt wird, und die W.K. sagt dann etwas aus über das Ergebnis eines Experimentes, das in der Zukunft liegt, also erst ausgeführt wird. Wenn uns jemand nach Durchführung des Zufallsexperimentes ein wenig Informationen gibt (also sagt, das Ereignis B sei eingetreten), haben

$$A_1 \cap A_2 = \{\}$$

$$P(A_1 \mid B) = \frac{P(A_1 \cap B)}{P(B)}$$

$$P(A_2 \mid B) = \frac{P(A_2 \cap B)}{P(B)}$$

Beachte: $(A_1 \cap B) \cap (A_2 \cap B) = \{\}$

$P(A_1 \mid B) + P(A_2 \mid B) =$

$$\frac{P(A_1 \cap B) + P(A_2 \cap B)}{P(B)} = \frac{P((A_1 \cap B) \cup (A_2 \cap B))}{P(B)}$$

$$= \frac{P((A_1 \cup A_2) \cap B)}{P(B)} = P(A_1 \cup A_2 \mid B)$$

falls $A_1 \cap A_2 = \{\}$

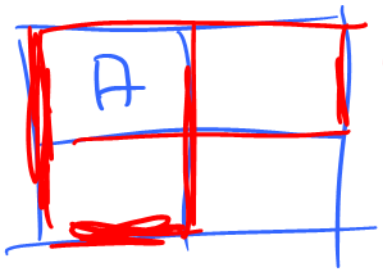
wir ja gar kein Zufallsexperiment mehr! Man muss vielmehr **vor** der Durchführung des Experimentes vereinbaren, ob die Information B weitergegeben wird. Wir schauen uns also all die Ausgänge eines Zufallsexperimentes an, die eintreten, wenn auch B eingetreten ist und auch wirklich nur dann! Wir erhalten so einen neuen Wahrscheinlichkeitsraum:

Satz 1.14. *Sei (Ω, P) ein Wahrscheinlichkeitsraum und sei B ein Ereignis mit $P(B) \neq 0$. Dann definiert $P(A|B)$ für $A \subseteq \Omega$ ein Wahrscheinlichkeitsmaß auf Ω. Dieses neue Maß wird auch $P_B(A)$ bezeichnet.*

Beispiel 1.15. Angenommen wir würfeln mit einem Würfel und B

$P(A_1 \cup A_2 | B) = P(A_1 | B) + P(A_2 | B)$

falls $A_1 \cap A_2 = \{\}$.



$P(A|B) = \frac{1}{3}$

$\frac{P(\{2\} \cap B)}{P(B)} = \frac{2}{6} = \frac{1}{3}$

sei das Ereignis "es wird eine gerade Zahl gewürfelt". Dann ist

$P(B) = \frac{1}{2}$

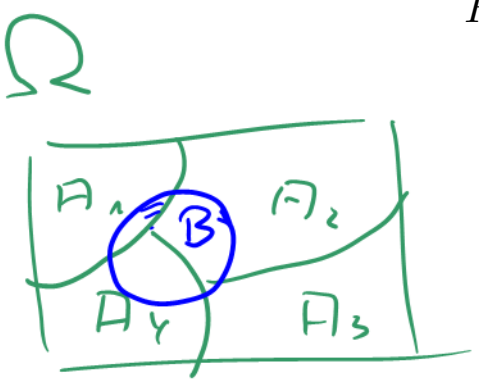
$$\begin{aligned} P(\{1\}|B) &= 0 \\ P(\{2\}|B) &= 1/3 \\ P(\{3\}|B) &= 0 \\ P(\{4\}|B) &= 1/3 \\ P(\{5\}|B) &= 0 \\ P(\{6\}|B) &= 1/3. \end{aligned}$$

$\frac{P(\{1\} \cap B)}{P(B)} = 0$

Satz 1.16 (Satz von der totalen Wahrscheinlichkeit)**.** *(Ω, P) sei ein Wahrscheinlichkeitsraum, $A_i \subseteq \Omega$ mit $1 \leq i \leq m$ paarweise disjunkte Ereignisse mit $\Omega = \dot{\bigcup}_{i=1}^{m} A_i$. Ferner sei $B \subseteq \Omega$ sowie $P(B), P(A_i) \neq 0$ für $i = 1, \ldots, m$. Dann gilt*

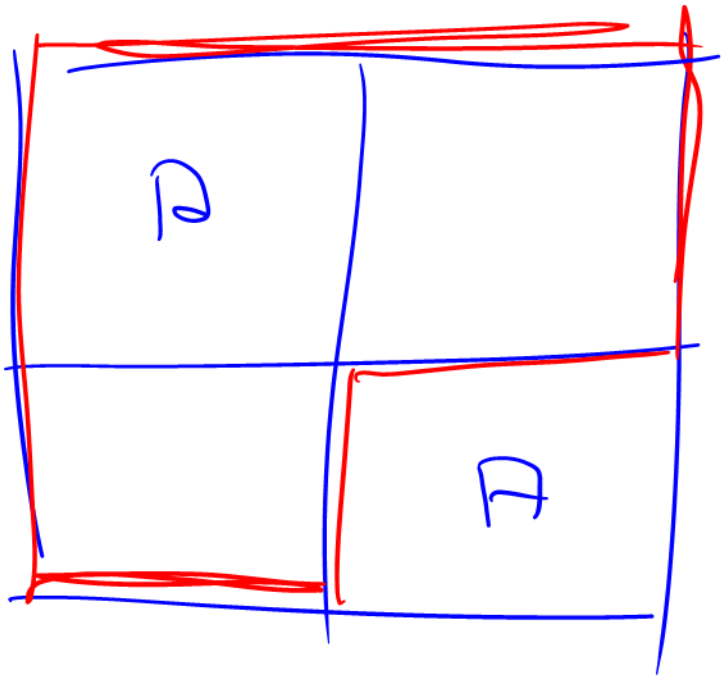
$$P(B) = \sum_{i=1}^{m} P(A_i) \cdot P(B|A_i) = \sum_{i=1}^{m} P(A_i \cap B)$$

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



disjunkt

$B = \bigcup_{i=1}^{m} (A_i \cap B)$



$P(D) = \frac{1}{2}$

$P(B) = \frac{3}{4}$

$\frac{P(D)}{P(B)} = \frac{2}{3}$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{3/4} = \frac{1}{3}$$

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

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

$= \frac{1}{3}$

sowie (Satz von BAYES*)*

$$P(A_k|B) = \frac{P(A_k)}{P(B)} \cdot P(B|A_k) = \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{m} P(A_i) \cdot P(B|A_i)}.$$

Beispiel 1.17. Nehmen wir an, wir leben in einem Land, in dem jede Familie genau zwei Kinder hat, jeweils 1/4 der Familien haben zwei Jungen und zwei Mädchen, und bei jeweils 1/4 der Familien ist die Verteilung mj und jm, wobei im ersten Fall das Mädchen (m) das erstgeborene Kind ist, im zweiten Fall der Junge (j). Es gilt hier also

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

und wir haben, wenn eine Familie zufällig ausgewählt wird, ein Laplace-Experiment. Nun nehmen wir an, dass in dem Land einem Besucher stets (falls möglich) zuerst die Tochter vorgestellt wird. Wenn man also eine Familie besucht und es wird eine Tochter vorgestellt, erhält man Teilinformationen, nämlich man weiß, dass das Ereignis

$$B = \{mj, jm, mm\}$$

eingetreten ist. Das liefert

$$\frac{P(mm \cap B)}{P(B)} = \frac{1/4}{3/4}$$

$$\begin{aligned} P_B(mm) &= 1/3 \\ P_B(mj) &= 1/3 \\ P_B(jm) &= 1/3 \\ P_B(jj) &= 0. \end{aligned}$$

Das bedeutet, mit Wahrscheinlichkeit 2/3 handelt es sich bei der Familie um eine mit einem Mädchen und einem Jungen.
Nun ändern sich die Traditionen in dem Land, und einem Besucher wird stets der/die Erstgeborene vorgestellt. Angenommen, dann wird uns eine Tochter vorgestellt. Das Ereignis C wäre dann

$$C = \{mm, mj\}$$

und wir erhalten

$$\begin{aligned} P_C(mm) &= 1/2 \\ P_C(mj) &= 1/2 \\ P_C(jm) &= 0 \\ P_C(jj) &= 0. \end{aligned}$$

Damit ist nun die Wahrscheinlichkeit für eine Familie mit einem Mädchen und einem Jungen genau 1/2.
Nun ändern sich die Traditionen noch mehr und die Eltern schnappen sich irgendein Kind, das sie dem Besucher als erstes vorstellen. Angenommen, das ist ein Mädchen. Wie groß ist jetzt die W.K., dass das zweite Kind ein Junge ist? Oft wird hier so argumentiert wie im ersten Fall. Andererseits hat man intuitiv das Gefühl, die W.K. für das Geschlecht des zweiten Kindes sollte unabhängig sein vom Geschlecht eines zufällig beobachteten Kindes. Wie können wir hier vorgehen?

Gehen wir zu einem gößeren Ereignisraum über:

$$\Omega_e := \Omega \times \Omega_v,$$

wobei $\Omega_v = \{j, m\}$ das Geschlecht des Kindes ist, das vorgestellt wird. Wir erhalten

$$\begin{aligned} P(mm, m) &= 1/4 \\ P(mm, j) &= 0 \\ P(mj, m) &= 1/8 \\ P(mj, j) &= 1/8 \\ P(jm, m) &= 1/8 \\ P(jm, j) &= 1/8 \\ P(jj, m) &= 0 \\ P(jj, j) &= 1/4. \end{aligned}$$

Die Wahrscheinlichkeit, dass uns ein Mädchen vorgestellt wird (nennen wir das Ereignis D) ist $1/2$. Wenn A das Ereignis bezeichnet "Ein

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

Kind ist ein Junge", so ist die Wahrscheinlichkeit $P(A \cap D) = 1/4$, somit $P(A|D) = 1/2$.

Beispiel 1.18. In einer Bevölkerungsgruppe seien 0.1% der Bevölkerung mit einem Virus infiziert, der Rest ist nicht infiziert. Ein Test habe eine zuverlässigkeit von 99%, d.h. er liefert in 1% der Fälle ein falsches Ergebnis. Mit welcher W.K. ist ein positiv auf das Virus getesteter Mensch in Wirklichkeit gesund (also nicht infiziert). Wir können das mit dem Satz von der totalen W.K. machen: Jemand kann gesund oder krank (infiziert) sein (G, K) und jemand kann positiv (auf Virus) p oder negativ (gesund) n getestet werden. $P(\cdot)$ bezeichne die W.K. für diese Ereignisse. Dann gilt

$$P(G|p) = \frac{P(G) \cdot P(p|G)}{P(G) \cdot P(p|G) + P(K) \cdot P(p|K)} = \frac{0.999 \cdot 0.01}{0.999 \cdot 0.01 + 0.001 \cdot 0.99} \approx 0.9098\ldots$$

Also: Die meisten positiv getesteten Menschen sind gesund!