ST. PÖLTEN UNIVERSITY
OF APPLIED SCIENCES

Informatik & Security

Wahrscheinlichkeitsrechnung

Dipl.-Ing. Hubert Schölnast, BSc
Stand: 21. Juni 2023

Inhaltsverzeichnis

1 Grundbegriffe

1.1 Zufallsexperiment

Ein Zufallsexperiment ist ein Vorgang, den man

- beliebig oft
- unter gleichbleibenden Bedingungen

wiederholen kann, und

- dessen Ergebnis man nicht mit Sicherheit vorhersagen kann.

Klassische Beispiele dafür sind:

- Der Wurf einer Münze
- Der Wurf eines Würfels
- Das Ziehen einer Spielkarte

Aber auch das sind Beispiele für Zufallsexperimente bzw. für deren Ergebnisse:

- Die IP-Adresse des nächsten Besuchers einer bestimmten öffentlichen Website
- Das Kennzeichen des Autos, das bei der nächsten Rot-Phase als erstes an einer Ampel hält
- Die Anzahl der Menschen, die aus dem Aufzug kommen, wenn seine Tür das nächste Mal aufgeht
- Das Geschlecht des dritten Kindes eines Elternpaares
- usw.

1.2 Elementarereignis, Ereignismenge

Alle gleichen Ergebnisse eines Zufallsexperiments sind zusammengenommen ein Elementarereignis. Beispielsweise sind alle Würfe eines gewöhnlichen sechsseitigen Spielwürfels, bei denen die Augenzahl ⚂ erscheint, erschienen ist, oder noch erscheinen wird, zusammengenommen das Elementarereignis {⚂}. Elementarereignisse sind Mengen. Sie enthalten alle Zufallsexperimente mit demselben Ausgang.

Elementarereignisse sind atomar. Das bedeutet, dass man sie nicht in weitere Sub-Ereignisse unterteilen kann.

Elementarereignisse schließen sich gegenseitig aus. Das heißt beispielsweise: Es ist ausgeschlossen, dass jemand bei nur einem Wurf mit einem normalen Spielwürfel zugleich die Augenzahlen ⚀ und ⚄ wirft. Mit anderen Worten: Die Elementarereignisse eines Zufallsexperiments sind paarweise disjunkt:

Es seien $E_1, E_2, \cdots, E_n$ die Elementarereignisse eines Zufallsexperiments, dann gilt:

$$i \neq j \implies E_i \cap E_j = \emptyset$$

Die Vereinigungsmenge aller Elementarereignisse eines Zufallsexperiments wird die Ereignismenge genannt. Als Formelsymbol für die Ereignismenge wird häufig der große griechische Buchstabe Omega verwendet: Ω. Dies ist aber nur eine Konvention und man kann, wenn man dafür einen guten Grund hat, stattdessen auch jedes beliebige andere Symbol verwenden. Wenn in diesem Skriptum das Symbol Ω verwendet wird, und nichts anderes gesagt wird, ist damit immer eine Ereignismenge gemeint.

$$\Omega := \bigcup_{i=1}^{n} E_i$$

Nachdem die Elementarereignisse paarweise disjunkt sind und ihre Vereinigung die Ereignismenge ist, bilden die Elementarereignisse eine Partition der Ergebnismenge.

1.3 Ereignis

Ein Ereignis E ist eine Teilmenge einer Ereignismenge.

$$E \subseteq \Omega$$

Beispiel:

Wenn man einen gewöhnlichen sechsseitigen Spielwürfel wirft, können die Augenzahlen ⚀, ⚁, ⚂, ⚃, ⚄, und ⚅ auftreten. Also:

$$\Omega = \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\}$$

Das Werfen einer geraden Zahl wäre dann ein Ereignis

$$\mathrm{G} = \{\text{⚁}, \text{⚃}, \text{⚅}\}$$

Denn es gilt ja:

$$\{\text{⚁}, \text{⚃}, \text{⚅}\} \subseteq \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\} \quad \Longleftrightarrow \quad G \subseteq \Omega$$

Ebenso wäre das Werfen einer Zahl, die kleiner als 4 ist, ein Ereignis:

$$\mathrm{K} = \{\text{⚀}, \text{⚁}, \text{⚂}\}$$
$$\{\text{⚀}, \text{⚁}, \text{⚂}\} \subseteq \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\} \quad \Longleftrightarrow \quad K \subseteq \Omega$$

Natürlich ist auch das Werfen einer bestimmten (einzelnen) Zahl ein Ereignis. Es ist sogar ein Elementarereignis:

$$\mathrm{V} = \{\text{⚃}\}$$
$$\{\text{⚃}\} \subseteq \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\} \quad \Longleftrightarrow \quad V \subseteq \Omega$$

1.4 Das sichere und das unmögliche Ereignis

Zu jeder Menge gehören zwei ganz besondere Teilmengen[1]:

1. Die leere Menge $\emptyset \subseteq M$
2. Die ursprüngliche Menge selbst. $M \subseteq M$

Da jede Teilmenge der Ereignismenge ein Ereignis ist, trifft das auch auf diese beiden speziellen Teilmengen zu. Diese beiden Ereignisse, die es bei jeder Ereignismenge gibt, haben spezielle Namen:

- **Das unmögliche Ereignis**
  Das ist die leere Teilmenge der Ereignismenge
  Dieses Ereignis enthält kein einziges Elementarereignis und kann daher niemals eintreten.
- **Das sichere Ereignis**
  Das ist jene Teilmenge der Ereignismenge, die alle Elementarereignisse enthält.
  Dieses Ereignis tritt bei jedem Zufallsexperiment mit Sicherheit immer ein.

Beispiel:

Das Zufallsexperiment ist ein Münzwurf. Es gibt zwei Elementarereignisse, nämlich »Kopf« und »Zahl«.

Das unmögliche Ereignis ist jenes Ereignis, bei dem die Münze weder »Kopf« noch »Zahl« zeigt (sondern z.B. »gestern«, »O Tannenbaum« oder »π«). Das kann niemals passieren, daher ist das ein unmögliches Ereignis.

Das Eintreten des sicheren Ereignisses bedeutet nichts anderes, als dass die Münze nach dem Wurf ein Element der Menge $\{\text{Kopf}, \text{Zahl}\}$ zeigt; also entweder »Kopf« oder »Zahl«. Eines dieser Elementarereignisse muss jedes Mal eintreten, daher nennt man dieses Ereignis das »sichere Ereignis«.

Beachte, dass das nur gilt, weil davon ausgegangen wird, dass die Münze niemals auf der Kante stehenbleibt. Will man auch diesen Fall berücksichtigen, muss man die Ereignismenge nur um dieses Element erweitern. Sie enthält dann diese drei Elemente:

$$\{\text{Kopf}, \text{Zahl}, \text{steht auf der Kante}\}$$

Dieselbe Menge ist dann auch das sichere Ereignis.

[1] Davon ausgenommen ist nur die leere Menge, denn bei ihr sind die Teilmenge, die keine Elemente enthält und die Teilmenge, die alle Elemente enthält, identisch. Von der leeren Menge kann man nur eine einzige Teilmenge bilden.

1.5 Gegenereignis

In der Mengenlehre gibt es, bezogen auf eine bestimmte Grundmenge G, zu jeder Menge M immer eine Komplementärmenge M^C. Die Komplementärmenge M^C enthält alle Elemente der Grundmenge G, die nicht in der Menge M vorkommen:

$$M \subseteq G \ \wedge \ x \in G \ \wedge \ x \notin M \ \Longrightarrow \ x \in M^C$$

$$M \subseteq G \ \wedge \ x \in G \ \wedge \ x \in M \ \Longrightarrow \ x \notin M^C$$

In der Wahrscheinlichkeitsrechnung ist auf diese Weise das Gegenereignis eines Ereignisses definiert. Das Gegenereignis ist ganz einfach die Komplementärmenge des Ereignisses.

Beispiel:

Bleiben wir beim Werfen eines gewöhnlichen Spielwürfels. Die Ereignismenge Ω ist $\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}$ und wir betrachten das Ereignis, bei dem eine gerade Augenzahl geworfen wird:

$$\mathrm{G} = \{⚁, ⚃, ⚅\}$$

Dann ist das Gegenereignis jenes Ereignis, bei dem eine ungerade Augenzahl geworfen wird:

$$\mathrm{G}^C = \mathrm{U} = \{⚀, ⚂, ⚄\}$$

Schlussfolgerungen:

- Das sichere Ereignis ist das Gegenereignis des unmöglichen Ereignisses.
- Das unmögliche Ereignis ist das Gegenereignis des sicheren Ereignisses.

1.6 Vereinbare und unvereinbare Ereignisse

Zwei Ereignisse sind vereinbar, wenn die Schnittmenge der beiden Ereignisse nicht leer ist. Anders gesagt: Wenn es Elementarereignisse gibt, die in zwei verschiedenen Ereignissen vorkommen, dann sind diese Ereignisse vereinbar. Daraus folgt auch, dass es möglich ist, dass Ereignisse, die miteinander vereinbar sind, gleichzeitig auftreten.

Beispiel 1:

Das Ereignis G sei wieder das Werfen einer geraden Augenzahl

$$\mathrm{G} = \{⚁, ⚃, ⚅\}$$

und K sei das Ereignis, bei dem eine Augenzahl kleiner als 4 geworfen wird:

$$\mathrm{K} = \{⚀, ⚁, ⚂\}$$

Die Schnittmenge aus beiden Mengen ist jene Menge, die alle Elemente enthält, die sowohl in G als auch in K vorkommen:

$$S = G \cap K = \{\text{⚁}, \text{⚃}, \text{⚅}\} \cap \{\text{⚀}, \text{⚁}, \text{⚂}\} = \{\text{⚁}\}$$

Diese Schnittmenge S ist nicht leer, daher sind die Ereignisse G und K vereinbar. Es ist also möglich, mit nur einem Wurf zugleich eine ungerade Zahl als auch eine Zahl kleiner als 4 zu werfen.

Beispiel 2:

Das Ereignis U sei das Werfen einer ungeraden Augenzahl, und X sei das Werfen der Augenzahl ⚅.

$$\mathrm{U} = \{\text{⚀}, \text{⚂}, \text{⚄}\}$$
$$X = \{\text{⚅}\}$$

Hier gilt:

$$S = U \cap X = \{\text{⚀}, \text{⚂}, \text{⚄}\} \cap \{\text{⚅}\} = \emptyset$$

Die beiden Mengen haben kein gemeinsames Element. Es ist unmöglich, mit nur einem Wurf sowohl eine ungerade Zahl als auch die Zahl 6 zu werfen. Daher nennt man diese beiden Ereignisse »unvereinbar«.

Schlussfolgerungen:

- Ein Ereignis ist niemals mit seinem Gegenereignis vereinbar.
- Elementarereignisse sind niemals mit anderen Elementarereignissen vereinbar. (Das folgt ohnehin bereits aus der Definition von Elementarereignissen.)
- Das unmögliche Ereignis ist mit keinem Ereignis vereinbar. (Nicht mal mit sich selbst.)
- Jedes Ereignis ist mit sich selbst vereinbar. (Davon ausgenommen ist nur das unmögliche Ereignis.)
- Das sichere Ereignis ist mit allen Ereignissen vereinbar (mit Ausnahme des unmöglichen Ereignisses).

1.7 Laplace-Experiment

Viele (aber nicht alle) der bisherigen Ereignismengen hatten folgende Eigenschaften:

- Die Ereignismenge ist endlich (d.h. es gibt nicht unendlich viele mögliche Elementarereignisse)
- Alle Elementarereignisse treten mit gleicher Wahrscheinlichkeit auf. Langfristig ist also nicht damit zu rechnen, dass bestimmte Elementarereignisse häufiger auftreten als andere.

Ein Zufallsexperiment, das eine Ereignismenge mit solchen Eigenschaften hervorbringt, nennt man ein Laplace-Experiment[2].

Viele Glücksspiele beruhen auf Laplace-Experimenten, zum Beispiel:

- Der Wurf einer Münze (die nie auf der Kante stehen bleibt)
- Der Wurf eines Spiel-Würfels
- Das Ziehen einer Karte aus einem gemischten Kartenspiel
- Roulette (Die Kugel fällt in eines von 37 Fächern)
- Lotto (Aus mehreren Millionen Tipps wird ein Sechser gezogen)

Keine Laplace-Experimente sind Zufallsexperimente, bei denen es unendlich viele Elementarereignisse gibt.

- Der Drehwinkel eines Windrades, das auf einem Foto zu sehen ist.
  Es kann zwar jeder Drehwinkel gleich häufig auftreten, aber es gibt unendlich viele mögliche Drehwinkel.

Zufallsexperimente, bei denen die Elementarereignisse unterschiedlich häufig auftreten, sind ebenfalls keine Laplace-Experimente.

- Die Anzahl der Personen, die an einer Haltestelle aus einem Bus aussteigen.
  Die Anzahl der Elementarereignisse ist zwar endlich (eine natürliche Zahl, die nach oben durch das Fassungsvermögen des Busses beschränkt ist), aber dass genau so viele Menschen aussteigen, wie im Bus überhaupt Platz haben, passiert seltener als dass z.B. nur 2 oder 3 Personen aussteigen. Die Elementarereignisse treten also langfristig nicht gleich häufig auf.

2 Wahrscheinlichkeit

Eine Wahrscheinlichkeit eines Ereignisses ist ein Zahlenwert, der angibt, wie sehr damit zu rechnen ist, dass dieses Ereignis eintritt. Für die Wahrscheinlichkeit verwendet man sehr häufig das Symbol P.[3] Um anzuzeigen, wessen Wahrscheinlichkeit gemeint ist, wird das entsprechende Ereignis in runden Klammern hinter das P geschrieben:

$$P(E) \text{ = Die Wahrscheinlichkeit des Ereignisses } E.$$

---

[2] Benannt nach dem Marquis Pierre-Simon de Laplace (1749-1827). Laplace war ein französischer Mathematiker, Physiker und Astronom. Neben seinen Beiträgen zur Wahrscheinlichkeitsrechnung gehen auch wichtige Erkenntnisse aus den Bereichen Spieltheorie, Differentialgleichungen und der Mechanik von Himmelskörpern auf ihn zurück.

[3] von Lateinisch »probabilis« = glaubhaft, es erscheint wahr, wahrscheinlich bzw. »probābilitās« = Glaubwürdigkeit, Wahrscheinlichkeit über Mittelfranzösisch »probabilité« = Wahrscheinlichkeit zu Englisch »probability« = Wahrscheinlichkeit

2.1 Kolmogorow-Axiome

Andrei Kolmogorow[4] formulierte im Jahr 1933 erstmals drei Axiome[5], die die Grundlage der Wahrscheinlichkeitstheorie sind. Implizit wurden diese Axiome zwar schon lange davor verwendet, ab erst durch Kolmogorow wurden sie explizit als solche formuliert.

2.1.1 Erstes Axiom

Für jedes Ereignis $E \subseteq \Omega$ ist die Wahrscheinlichkeit von E (also $P(E)$) eine reelle Zahl zwischen 0 und 1:

$$P(E) \in \mathbb{R}; \quad 0 \leq P(E) \leq 1$$

2.1.2 Zweites Axiom

Das sichere Ereignis Ω hat die Wahrscheinlichkeit 1:

$$P(\Omega) = 1$$

2.1.3 Drittes Axiom

Die Wahrscheinlichkeit einer Vereinigung abzählbar[6] vieler unvereinbarer Ereignisse ist gleich der Summe der Wahrscheinlichkeiten der einzelnen Ereignisse:

Für $E_1, E_2, \cdots, E_n \subseteq \Omega$ mit $E_i \cap E_j = \emptyset$ wenn $i \neq j$ gilt:

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$$

Das dritte Kolmogorow-Axiom heißt auch »Additionsregel für unvereinbare Ereignisse«.

---

[4] Андре́й Никола́евич Колмого́ров (Andrei Nikolajewitsch Kolmogorow) (1903-1987) war ein sowjetischer Mathematiker. Bekannt wurde er vor allem als Mitbegründer der algorithmischen Komplexitätstheorie (»Kolmogorow-Komplexität«) und durch die Axiomatisierung der Wahrscheinlichkeitstheorie.

[5] Ein Axiom ist ein mathematischer Satz, der nicht per Beweis wahr ist, sondern per Definition. Mehrere Axiome bilden gemeinsam die Basis eines axiomatischen Systems. Jede wahre Aussage innerhalb eines solchen Systems kann durch logische Operationen aus den Axiomen abgeleitet werden (so eine Ableitung heißt »Beweis«). Aber kein Axiom kann durch logische Operationen aus anderen Axiomen hergeleitet werden. Ein Axiom in der Zahlentheorie ist z.B. »0 ist eine Zahl«. Das ist ein Satz, der per Definition wahr ist. Man kann nicht beweisen, dass 0 eine Zahl ist. Dass das so ist, ist die Folge einer Definition. Auf dieser Definition (und ein paar anderen) beruht das ganze System der Zahlentheorie.

[6] Der Ausdruck »abzählbar« bedeutet: »entweder von endlicher Größe (also nicht unendlich groß) oder so groß wie die Menge der natürlichen Zahlen«. Unendlich große Mengen, die größer als die Menge der natürlichen Zahlen sind, nennt man »überabzählbar«. »Abzählbar« und »überabzählbar« sind also Größenangaben von verschieden großen unendlichen Mengen. Die abzählbaren Mengen sind die kleinsten unendlich großen Mengen.

2.2 Wahrscheinlichkeiten von Laplace-Experimenten

2.2.1 Wahrscheinlichkeiten von Laplace-Elementarereignissen

In Übereinstimmung mit diesen Axiomen können wir nun die Wahrscheinlichkeiten von Laplace-Experimenten definieren:

- Das dritte Axiom sagt uns, dass wir die Wahrscheinlichkeiten von Elementarereignissen immer addieren können, denn Elementarereignisse sind per Definition ja immer paarweise unvereinbar.
- Das zweite Axiom sagt uns, dass die Summe der Wahrscheinlichkeiten aller Elementarereignisse immer genau 1 ist, denn die Vereinigung aller Elementarereignisse ist ja das sichere Ereignis.
- Und weil wir Laplace-Experimente untersuchen, wissen wir, dass alle Elementarereignisse langfristig gleich häufig sind, was bedeutet, dass sie gleich wahrscheinlich sind.

Damit hat jedes Elementarereignis eines Laplace-Experiments die Wahrscheinlichkeit

$$P(E) = \frac{1}{n}$$

wobei n die Mächtigkeit der Ereignismenge Ω ist.

$$n = |\Omega|$$

Beispiel:

Die Wahrscheinlichkeit, mit einem gewöhnlichen 6-seitigen Spielwürfel die Augenzahl ⚁ zu würfeln, ist $\frac{1}{6}$.

$$P(⚁) = \frac{1}{6}$$

Im Nenner steht die Zahl 6, weil das die Mächtigkeit der Ereignismenge eines sechsseitigen Würfels ist:

$$\Omega = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\} \quad \Longrightarrow \quad |\Omega| = 6$$

2.2.2 Wahrscheinlichkeiten von allgemeinen Laplace-Ereignissen

Jedes Ereignis ist per Definition eine Teilmenge der Ereignismenge, besteht also aus einer bestimmten Anzahl von Elementarereignissen, die ja bekanntermaßen paarweise unvereinbar sind. Beispielsweise besteht das Ereignis »es wird eine gerade Augenzahl gewürfelt« aus der Vereinigung der Elementarereignisse $\{⚁\}$, $\{⚃\}$ und $\{⚅\}$

$$G = \{⚁, ⚃, ⚅\} = \{⚁\} \cup \{⚃\} \cup \{⚅\}$$

Die Wahrscheinlichkeiten der Elementarereignisse sind inzwischen bekannt (siehe 2.2.1)

$$P(⚁) = P(⚃) = P(⚅) = \frac{1}{6}$$

Das dritte Kolmogorow-Axiom sagt uns nun, wie wir nun die Wahrscheinlichkeit von G berechnen können:

$$P(G) \;=\; P(\{\text{⚁}\} \cup \{\text{⚃}\} \cup \{\text{⚅}\}) \;=\; P(\text{⚁}) + P(\text{⚃}) + P(\text{⚅}) \;= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \;=\; \frac{3}{6}$$

Das kann man weiter zu $\frac{1}{2}$ kürzen. Die Wahrscheinlichkeit, mit einem gewöhnlichen Würfel eine gerade Augenzahl zu würfeln, ist also $\frac{1}{2}$.

Aber bleiben wir noch beim ungekürzten Ergebnis. Hier steht unten, im Nenner, die Mächtigkeit der Ergebnismenge, das ist gleich die Anzahl aller möglichen Fälle. Und oben, im Zähler steht die Mächtigkeit des Ereignisses, das wir gerade untersuchen, also die Anzahl der Elementarereignisse, aus denen sich dieses Ereignis zusammensetzt. Das bezeichnet man häufig auch als »die Anzahl der günstigen Fälle«. Damit ist eigentlich gemeint: »die Anzahl der zum Ereignis gehörenden Elementarereignisse«.

$$P(G) \;=\; \frac{|G|}{|\Omega|} \;=\; \frac{\text{Anzahl »günstige« Fälle}}{\text{Anzahl alle Fälle}}$$

2.3 Rechenregeln und Gesetzmäßigkeiten für Wahrscheinlichkeiten

2.3.1 Wahrscheinlichkeit eines Gegenereignisses

Laut Definition sind Ereignis E und Gegenereignis E^{C} unvereinbar, das heißt

$$E \cap E^{\mathsf{C}} = \emptyset$$

und ihre Vereinigungsmenge ist die Ereignismenge Ω:

$$E \cup E^{\mathsf{C}} = \Omega$$

Aus dem zweiten und dritten Axiom folgt daher:

$$P(E) + P\left(E^{\mathsf{C}}\right) \;=\; P(\Omega) \;=\; 1$$

und daher

$$P\left(E^{\mathsf{C}}\right) \;=\; 1 - P(E)$$

Die Wahrscheinlichkeit eines Gegenereignisses ist daher 1 minus die Wahrscheinlichkeit des Ereignisses.

2.3.2 Wahrscheinlichkeit des unmöglichen Ereignisses

Das unmögliche Ereignis ist die leere Menge (siehe 1.4 auf Seite 5), und die Mächtigkeit der leeren Menge ist 0:

$$|\emptyset| = 0$$

Daher ist die Wahrscheinlichkeit dafür, dass etwas unmögliches geschieht, gleich 0:

$$P(\emptyset) = \frac{|\emptyset|}{|\Omega|} = \frac{0}{\text{Anzahl alle Fälle}} = 0$$

2.3.3 Wahrscheinlichkeit eines Teilereignisses (Monotonie)

Ereignisse sind Mengen, und Mengen können Teilmengen haben. Im Fall von Ereignissen nennt man diese Teilmengen »Teilereignisse«. Dem entsprechend sind »Oberereignisse« Obermengen von Ereignissen.

Betrachten wir 2 Ereignisse, die beim Ziehen von Spielkarten aus einem Französischen Blatt[7] auftreten können:

- R: Eine Karte mit roter Kartenfarbe wird gezogen (♥ oder ♦)
- K: Eine Karte der Kartenfarbe Karo gezogen (♦)

Ganz offensichtlich ist K ein Teilereignis von R. Das heißt, dass R alle Elementarereignisse enthält, die in K vorkommen, und eventuell gibt es in R sogar noch weitere zusätzliche Elementarereignisse (in diesem Fall sogar ganz sicher, in anderen Fällen muss das nicht zwingend so sein). Da keines dieser zusätzlichen Elementarereignisse eine negative Wahrscheinlichkeit haben kann (siehe Axiom 1), ist es unmöglich, dass die Wahrscheinlichkeit des Teilereignisses größer als die Wahrscheinlichkeit des Oberereignisses ist:

$$K \subseteq R \quad \Longrightarrow \quad P(K) \leq P(R)$$

Eine fachsprachliche Formulierung dieses Umstandes ist:
»Wahrscheinlichkeiten sind monoton«.

2.4 Additionsregeln

2.4.1 Additionsregel für unvereinbare Ereignisse

Diese Regel ist identisch mit dem dritten Kolmogorow-Axiom. Hier ist eine andere, etwas vereinfachte Formulierung dieses Axioms:

Die Wahrscheinlichkeit dafür, dass eines von mehreren unvereinbaren Ereignissen eintritt, ist die Summe der Wahrscheinlichkeiten dieser unvereinbaren Ereignisse.

---

[7] Kartenspiel mit 52 Karten in den Farben Kreuz (=Treff) ♣, Pik ♠, Herz ♥ und Karo ♦. Zu jeder Kartenfarbe gehören 13 Karten.

Beispiel 1:

Es seien A und B zwei unvereinbare Ereignisse. Die Wahrscheinlichkeiten für das Eintreten dieser Ereignisse seien $P(A)$ und $P(B)$. Dann ist die Wahrscheinlichkeit dafür, dass entweder A oder B eintritt die Summe dieser beiden Wahrscheinlichkeiten:

$$A \cap B = \emptyset \quad \Longrightarrow \quad P(A \cup B) = P(A) + P(B)$$

Beispiel 2:

Wie zuvor, nur sollen es nun drei unvereinbare Ereignisse A, B und C seien, mit den dazugehörenden Wahrscheinlichkeiten $P(A)$, $P(B)$ und $P(C)$:

$$A \cap B = \emptyset \quad \wedge \quad A \cap C = \emptyset \quad \wedge \quad B \cap C = \emptyset \quad \Longrightarrow \quad P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

2.4.2 **Additionsregel für vereinbare Ereignisse**

Wenn zwei Ereignisse A und B vereinbar sind, bedeutet das nichts anderes, als dass ihre Schnittmenge nicht leer ist:

$$A \cap B \neq \emptyset$$

Nennen wir diese Schnittmenge S:

$$S := A \cap B$$

Wenn wir diese Schnittmenge aus einer der beiden Mengen entfernen, erhalten wir eine kleinere Menge, die nun mit der anderen unvereinbar ist. Bilden wir zuerst diese Restmenge, z.B. von der Menge B:

$$B' := B \setminus S$$

Weil B' nur mehr Elemente enthält, die in A nicht vorkommen, sind A und B' unvereinbar, und man kann nun leicht die gewünschte Wahrscheinlichkeit nach Axiom 3 berechnen:

$$P(A \cup B) = P(A \cup B') = P(A) + P(B')$$

Dummerweise kennen wir aber die Wahrscheinlichkeit von B' (also $P(B')$) nicht; besser gesagt: noch nicht. Aber wir wissen zwei Dinge:

1. B' und S enthalten keine gleichen Elemente, d.h. B' und S sind unvereinbar:
$$B' \cap S = \emptyset$$
2. Die Vereinigungsmenge von B' und S ist B:
$$B' \cup S = B$$

Falls wir nun die Wahrscheinlichkeit für das Eintreten des Ereignisses S kennen, können wir daraus $P(B')$ berechnen. Denn, wegen 1 und 2 gilt ja:

$$P(B') + P(S) = P(B)$$

daher

$$P(B') = P(B) - P(S)$$

bzw. (weil $S = A \cap B$)

$$P(B') = P(B) - P(A \cap B)$$

Das kann man nun in die Formel einsetzen, die wir gerade eben schon hergeleitet haben, und erhalten nun endlich die ...

Additionsregel für vereinbare Ereignisse:

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

Das heißt: Man addiert einfach die Wahrscheinlichkeiten der beiden Ereignisse und zieht davon die Wahrscheinlichkeit für die Schnittmenge der beiden Ereignisse ab.

Man erkennt auch leicht, dass die Additionsregel für unvereinbare Ereignisse nur ein Spezialfall der soeben hergeleiteten Formel ist, mit $P(A \cap B) = 0$.

Beispiel 1:

Wir spielen Roulette und gehen der Einfachheit halber davon aus, dass es die Zahl 0 nicht gibt. Es stehen also nur die Zahlen von 1 bis 36 zur Auswahl. Die Kugel landet mit gleicher Wahrscheinlichkeit bei diesen 36 Zahlen, d.h. ein Roulett-Spiel ist ein Laplace-Experiment.

Wie groß ist die Wahrscheinlichkeit, dass eine ungerade Zahl (»impair« = »I«) oder eine Zahl im mittleren Dutzend (»medium douzaine« = »12^M« = Zahlen 13-24) gezogen wird?

Unter den 36 Zahlen, die gezogen werden können, sind 18 ungerade

$$P(I) = \frac{18}{36} = \frac{1}{2}$$

und 12 Zahlen liegen im mittleren Dutzend

$$P(12^M) = \frac{12}{36} = \frac{1}{3}$$

6 Zahlen sind sowohl ungerade als auch im mittleren Dutzend, nämlich $\{13, 15, 17, 19, 21, 23\}$

$$P(I \cap 12^M) = \frac{6}{36} = \frac{1}{6}$$

Die Wahrscheinlichkeit, dass eine Zahl gezogen wird, die entweder ungerade ist oder im mittleren Dutzend liegt (oder auf die beides zutrifft) ist daher

$$P(I \cup 12^M) = P(I) + P(12^M) - P(I \cap 12^M)$$

$$P(I \cup 12^M) = \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{3+2-1}{6} = \frac{4}{6} = \frac{2}{3}$$

Beispiel 2:

Im Protokoll eines Botanikers stehen diese Zeilen:

»Die untersuchte Wiese ist 360 m² groß, sie wurde in 36 Quadrate von je 10 m² unterteilt. Die einzigen Blütenpflanzen auf dieser Wiese waren Löwenzahn und Gänseblümchen. Mindestens jeweils 1 Löwenzahn wurde in 18 der 36 Quadrate gefunden, bei den Gänseblümchen waren es 12 Quadrate, in denen jeweils ein Exemplar blühte.«

Können Sie anhand dieser Angaben ermitteln, auf wie vielen der 36 Testfelder überhaupt irgendetwas blüht?

Antwort:
Nein, das ist nicht möglich.

Warum? Es sind doch dieselben Zahlen wie im vorherigen Beispiel (Roulette). Was ist jetzt anders?

Wir haben in diesem Beispiel keinerlei Information darüber, ob Gänseblümchen und Löwenzahn gerne nahe beieinander oder lieber weit voneinander entfernt wachsen, oder ob es beiden Pflanzen egal ist, ob die jeweils andere in ihrer Nachbarschaft wächst.

Es könnte sein, dass Gänseblümchen nur dort wachsen, wo es auch Löwenzahn gibt, dann würde nur dort etwas blühen, wo der Löwenzahl blüht, und die Antwort wäre $\frac{1}{2}$. Die andere Möglichkeit wäre, dass die beiden Pflanzen nie in demselben Quadrat zu finden sind, dann würde auf $\frac{5}{6}$ aller Quadrate etwas blühen.

Die wahre Antwort muss also irgendwo zwischen diesen beiden Extremwerten liegen, aber wo genau lässt sich nicht sagen.

Im Roulette-Beispiel war $P(I \cap 12^M)$ bekannt. Bei der Blumenwiese fehlt diese Information, und ohne diese Information ist die Berechnung nicht möglich.

3 Bedingte Wahrscheinlichkeiten

Oft ist man an der Wahrscheinlichkeit für ein bestimmtes Ereignis interessiert, unter der Voraussetzung, dass davor bereits ein anders Ereignis eingetreten ist.

Beispiele:

- Wie groß ist die Wahrscheinlichkeit, dass auf jenen Feldern, auf denen Löwenzahn wächst, auch Gänseblümchen wachsen?
- Wie groß ist die Wahrscheinlichkeit, dass man mit einem Virus infiziert ist, wenn ein Virus-Test eine Infektion anzeigt?

Bleiben wir bei der Blumenwiese und den Zahlen aus dem obigen Beispiel:

- 36 Felder gibt es insgesamt
  $|\Omega| = 36$
- auf 18 dieser Felder blüht mindestens 1 Löwenzahn
  $\Big|$  $\Big| = 18$
- auf 12 Feldern blüht mindestens 1 Gänseblümchen
  $\Big|$  $\Big| = 12$

Nehmen wir nun folgende Information hinzu, die vorher noch unbekannt war:

- auf 9 Feldern blüht von beiden Spezies mindestens 1 Exemplar
  $\Big|$  $\cap$  $\Big| = 9$

Die Wahrscheinlichkeit, auf irgendeinem der 36 Felder blühende Gänseblümchen zu finden, ist gemäß der Formel in Abschnitt 2.2.2 auf Seite 10:

$P\Big($  $\Big) = \dfrac{\text{Anzahl »günstige« Fälle}}{\text{Anzahl alle Fälle}} = \dfrac{\Big|\text{ }$  $\Big|}{|\Omega|} = \dfrac{12}{36} = \dfrac{1}{3}$

Nun wollen wir aber nur jene Fälle als neue Ereignismenge ansehen, bei denen Löwenzahn wächst. Dazu brauchen wir zunächst einmal eine neue Notation:

3.1 Notation

$$P(A|B)$$

ist die Wahrscheinlichkeit dafür, dass das Ereignis A eintritt, wenn das Ereignis B bereits eingetreten ist, bzw. unter der Voraussetzung, dass das Ereignis B eintritt.

Fortsetzung des Beispiels:

Die Wahrscheinlichkeit, ein Gänseblümchen auf einem Feld zu finden, auf dem bereits Löwenzahn wächst, schreibt man mit der neuen Notation so:

$$P\left(\text{🌼}|\text{🌻}\right)$$

und wieder gilt

$$\frac{\text{Anzahl »günstige« Fälle}}{\text{Anzahl alle Fälle}}$$

aber die Anzahl der günstigen Fälle ist nun die Mächtigkeit der Menge aller Felder, auf denen beide Arten wachsen, also $\left|\text{🌻} \cap \text{🌼}\right|$.

Und mit »alle Fälle« meint man nicht mehr alle 36 Felder, sondern nur mehr jene Felder, auf denen Löwenzahn wächst. Also

$$P\left(\text{🌼}|\text{🌻}\right) = \frac{\left|\text{🌻} \cap \text{🌼}\right|}{\left|\text{🌻}\right|} = \frac{9}{18} = \frac{1}{2}$$

Schreiben wir mal ein paar Zusammenhänge auf, die wir bei diesem Beispiel bereits jetzt leicht aufschreiben können:

Die Wahrscheinlichkeit dafür, dass auf irgendeinem Feld beide Arten blühen:

$$P\left(\text{🌼} \cap \text{🌻}\right) = \frac{\left|\text{🌻} \cap \text{🌼}\right|}{|\Omega|} = \frac{9}{36} = \frac{1}{4}$$

Die Wahrscheinlichkeit dafür, dass auf irgendeinem Feld Löwenzahn blüht:

$$P\left(\text{🌻}\right) = \frac{\left|\text{🌻}\right|}{|\Omega|} = \frac{18}{36} = \frac{1}{2}$$

3.2 Multiplikationssatz

Wenn man sich die letzten drei Formeln ansieht, erkennt man folgendes:

$$P\left(\text{🌼}|\text{🌻}\right) \cdot P\left(\text{🌻}\right) = \frac{\left|\text{🌻} \cap \text{🌼}\right|}{\left|\text{🌻}\right|} \cdot \frac{\left|\text{🌻}\right|}{|\Omega|} = \frac{\left|\text{🌻} \cap \text{🌼}\right| \cdot \left|\text{🌻}\right|}{\left|\text{🌻}\right| \cdot |\Omega|} = \frac{\left|\text{🌻} \cap \text{🌼}\right|}{|\Omega|} = P\left(\text{🌼} \cap \text{🌻}\right)$$

oder, auf das Wesentliche reduziert:

$$P\left(\text{🌼}|\text{🌻}\right) \cdot P\left(\text{🌻}\right) = P\left(\text{🌼} \cap \text{🌻}\right)$$

Mit vertauschen Seiten

$$P(🌼 \cap 🌻) \quad = \quad P(🌼 | 🌻) \cdot P(🌻)$$

Und weil das nicht nur für Blumen gilt, schreibt man das gerne auch allgemeiner so:

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

Die Wahrscheinlichkeit dafür, dass die Ereignisse A und B zugleich stattfinden, ist die Wahrscheinlichkeit dafür, dass A unter der Voraussetzung von B stattfindet, multipliziert mit der Wahrscheinlichkeit dafür, dass B stattfindet.

Diesen Satz nennt man den **Multiplikationssatz für Wahrscheinlichkeiten**.

Aus der Kommutativität der Schnittmengenbildung ($A \cap B = B \cap A$) folgt natürlich (durch simples Vertauschen der Buchstaben A und B) auch diese gleichwerte Variante des Multiplikationssatzes:

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

Die Wahrscheinlichkeit dafür, dass die Ereignisse A und B zugleich stattfinden, ist die Wahrscheinlichkeit dafür, dass B unter der Voraussetzung von A stattfindet, multipliziert mit der Wahrscheinlichkeit dafür, dass A stattfindet.

Man hätte diese alternative Form des Multiplikationssatzes auch herleiten können, indem man nach der Wahrscheinlich dafür gesucht hätte, dass Löwenzahn dort wächst, wo bereits Gänseblümchen wachsen. Wer mag, kann das gerne selbst versuchen.

3.3 Satz von Bayes

Der Multiplikationssatz ist für sich allein schon von sehr großer Wichtigkeit. Wenn man aber die rechten Seiten beiden Varianten dieses Satzes gleichsetzt und ein wenig umformt, erhält man einen Satz, den Thomas Bayes[8] entdeckt hat, der aber erst 1764, also 3 Jahre nach Bayes Tod, publiziert wurde. Dieser Satz von Bayes übertrifft den Multiplikationssatz vermutlich sogar noch an Wichtigkeit. Denn der Satz von Bayes gibt einem die Möglichkeit, die Wahrscheinlichkeit eines bedingten Ereignisses zu berechnen, bei dem die Rollen der beiden Ereignisse (welches hängt von wem ab?) vertauscht sind.

---

[8] Thomas Bayes (1701-1761) war eigentlich ein presbyterianischer Pfarrer in Tunbridge Wells, südöstlich von London. Er hat in seinem ganzen Leben nur 2 Schriftstücke veröffentlicht. Das eine war ein religiöses Werk, in dem er versucht zu beweisen, dass der Sinn des Lebens Glückseligkeit ist. Das andere Werk war eine mathematische Streitschrift gegen einen Bischof, der wiederum die Lehren Newtons anzweifelte. Zu Lebzeiten hat Bayes nie etwas über Statistik oder Wahrscheinlichkeiten veröffentlicht. Die Werke, für die Bayes heute berühmt ist, wurden alle erst aus seinem Nachlass veröffentlicht, einige Jahre nach seinem Tod.

Gleichsetzen der beiden rechten Seiten der beiden Varianten des Multiplikationssatzes:

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

Durch die Wahrscheinlichkeit des Eintretens von B dividieren:

Satz von Bayes

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

Man erhält die umgekehrte bedingte Wahrscheinlichkeit, wenn man eine bedingte Wahrscheinlichkeit mit der Wahrscheinlichkeit des bedingenden Ereignisses multipliziert und durch die Wahrscheinlichkeit des bedingten Ereignisses dividiert.

Andere Schreibweise:

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

Zu ergänzen ist lediglich noch, dass dieser Satz nur gilt, wenn $P(B) \neq 0$.

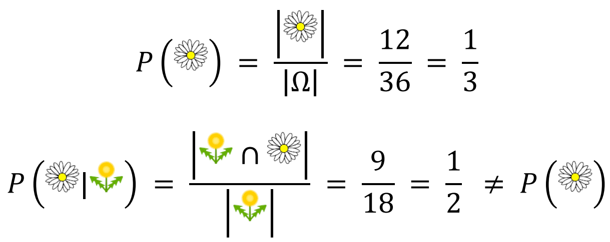
Falls $P(B) = 0$, ist automatisch auch $P(B|A) = 0$. Wenn man nun die Formel in dieser Variante schreibt

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

erkennt man, dass sich dann mit $\frac{P(B|A)}{P(B)}$ der Ausdruck $\frac{0}{0}$ ergibt, dessen Wert unbestimmt ist. $P(A|B)$ kann also jeden beliebigen Wert haben, wenn $P(B) = 0$. In Wahrheit steckt in $P(A|B)$ aber ohnehin keine sinnvolle Information, wenn $P(B) = 0$.

3.4 Abhängige und unabhängige Ereignisse

Bei der Blumenwiese hatten wir diese Wahrscheinlichkeiten:



Oder etwas allgemeiner:

$$P(A \mid B) \;\neq\; P(A)$$

Die Wahrscheinlichkeit dafür, dass man auf einem Feld, auf dem schon Löwenzahn wächst, auch ein Gänseblümchen findet, ist nicht identisch mit der Wahrscheinlichkeit dafür, überhaupt auf irgendeinem Feld ein Gänseblümchen zu finden. Die Blüh-Wahrscheinlichkeit der Gänseblümchen wird also vom Löwenzahn beeinflusst (zumindest auf der untersuchten Blumenwiese). Hier gibt es also eine Abhängigkeit.

Sehen wir uns hingegen noch einmal das Roulette-Spiel an. Wir hatten diese Zahlen:

Ungerade (»impair«)

$$P(I) = \frac{18}{36} = \frac{1}{2}$$

Mittleres Dutzend

$$P(12^M) = \frac{12}{36} = \frac{1}{3}$$

Schnittmenge (ungerade Zahlen im mittleren Dutzend)

$$P(I \cap 12^M) = \frac{6}{36} = \frac{1}{6}$$

Mit dem nun verfügbaren Wissen können wir auch berechnen, wie groß die Wahrscheinlichkeit des einen Ereignisses unter der Voraussetzung des anderen ist.

Ungerade Zahl, wenn man weiß, dass die gezogene Zahl im mittleren Dutzend liegt:

$$P(I \mid 12^M) = \frac{|I \cap 12^M|}{|12^M|} = \frac{6}{12} = \frac{1}{2} = P(I)$$

Umgekehrt:

$$P(12^M \mid I) = \frac{|12^M \cap I|}{|I|} = \frac{6}{18} = \frac{1}{3} = P(12^M)$$

In beiden Fällen gilt also

$$P(A \mid B) = P(A)$$

Es ist also egal, ob ich mir ansehe, mit welcher Wahrscheinlichkeit das Ereignis A überhaupt eintritt, oder ob ich voraussetze, dass B eingetreten ist. In beiden Fällen kommt genau dasselbe heraus. Mit anderen Worten: A ist nicht von B abhängig.

Zwei Ereignisse A und B nennt man »unabhängig«, wenn einer diese drei Bedingungen zutrifft:

$P(A \mid B) = P(A)$ wobei $P(B) > 0$ sein muss

$P(B \mid A) = P(B)$ wobei $P(A) > 0$ sein muss

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

Wenn keine dieser Bedingungen zutrifft, sind A und B abhängig voneinander.

Mit dem Multiplikationssatz (3.2 auf Seite 18) kann man leicht zeigen, dass alle drei Bedingungen äquivalent sind. Man kann aus je der der drei Bedingungen die beiden anderen ableiten. Mit anderen Worten: Sobald eine dieser drei Bedingungen erfüllt ist, sind das die beiden anderen automatisch auch. Auch die Umkehrung gilt: Wenn eine dieser Bedingungen verletzt ist, trifft keine der drei zu.

3.4.1 **Multiplikationssatz für unabhängige Ereignisse**

Die letzte der drei Formeln im obenstehenden Block hat einen eigenen Namen. Aus naheliegenden Gründen heißt sie:

Multiplikationssatz für unabhängige Ereignisse

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

Man kann diese Formel anstelle des »normalen« Multiplikationssatzes anwenden, wenn man bereits weiß, dass A und B unabhängig voneinander sind.

4 Totale Wahrscheinlichkeit

Gleich zu Beginn ein Beispiel:

In einem bestimmten Studiengang sind 20% der Studierenden weiblich, alle anderen sind männlich. Von den weiblichen Studierenden erhalten 30% ein Leistungsstipendium, von den männlichen sind es nur 10%. Die Richtlinien des Stipendiums sehen vor, dass mindestens 15%, höchstens aber 20% aller Studierenden ein Leistungsstipendium erhalten. Werden diese Richtlinien eingehalten?

Wir haben also zwei bedingte Wahrscheinlichkeiten gegeben, wobei dasselbe abhängige Ereignis von zwei verschiedenen vorgegebenen Ereignissen abhängt:

$$P(€ \mid ♀) = \frac{30}{100} = \frac{3}{10}$$

$$P(€ \mid ♂) = \frac{10}{100} = \frac{1}{10}$$

Außerdem kennen wir die Wahrscheinlichkeit $P(♀)$ und damit die Wahrscheinlichkeit des Gegenereignisses $♀^C = ♂$, also $P(♂)$

$$P(♀) = \frac{20}{100} = \frac{1}{5}$$

$$P(♂) = P\left(♀^C\right) = 1 - P(♀) = 1 - \frac{1}{5} = \frac{4}{5}$$

Und wir wollen daraus $P(€)$ berechnen, um diesem Wert mit den Vorgaben vergleichen zu können.

Mit dem Multiplikationssatz können wir $P(€ \cap ♀)$ und $P(€ \cap ♂)$ berechnen:

$$P(€ \cap ♀) = P(€ \mid ♀) \cdot P(♀) = \frac{3}{10} \cdot \frac{1}{5} = \frac{3}{50}$$

$$P(€ \cap ♂) = P(€ \mid ♂) \cdot P(♂) = \frac{1}{10} \cdot \frac{4}{5} = \frac{4}{50}$$

Diese beiden Ereignisse können wir vereinigen und auf den entstehenden Ausdruck das Distributivgesetz für Vereinigungs- und Schnittmengenbildung anwenden

$$(€ \cap ♀) \cup (€ \cap ♂) = € \cap (♀ \cup ♂)$$

Wegen $♀^C = ♂$ steht in der letzten Klammer aber eigentlich $♀ \cup ♀^C$ und das ist nichts weiter als die gesamte Ereignismenge Ω. Rechts vom Gleichheitszeichen steht also $€ \cap \Omega$. Die Menge der Stipendienbezieher € ist aber eine Teilmenge von Ω: $€ \subseteq \Omega$. Und daher gilt $€ \cap \Omega = €$. Wir haben also:

$$(€ \cap ♀) \cup (€ \cap ♂) = €$$

und daher auch

$$P\big((€ \cap ♀) \cup (€ \cap ♂)\big) = P(€)$$

Der Ausdruck rechts vom Gleichheitszeichen ist also bereits die gesuchte Wahrscheinlichkeit, daher steht auch links dieser Wert. Auf der linken Seite steht etwas in der Form $P(A \cup B)$ mit $A = € \cap ♀$ und $B = € \cap ♂$. Weil aber ♀ und ♂ disjunkt sind, sind auch A und B disjunkt, oder, in der Sprache der Wahrscheinlichkeiten: A und B sind unvereinbar: $A \cap B = \emptyset$. Und damit haben wir die Möglichkeit, die Additionsregel für unvereinbare Ereignisse anzuwenden:

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

Wir setzen ein $A = € \cap ♀$ und $B = € \cap ♂$:

$$P\big((€ \cap ♀) \cup (€ \cap ♂)\big) = P(€ \cap ♀) + P(€ \cap ♂)$$

Die linke Seite ersetzen wir durch das, was wir eigentlich suchen

$$P(€) = P(€ \cap ♀) + P(€ \cap ♂)$$

Und für beiden Summanden auf der rechten Seite haben wir weiter oben auch schon etwas hergeleitet, das wir hier einsetzen können:

$$P(€) = P(€ \mid ♀) \cdot P(♀) + P(€ \mid ♂) \cdot P(♂)$$

4.1 Formel für die totale Wahrscheinlichkeit

Bevor wir am Beispiel weiterrechnen, wollen wir diese Formel etwas allgemeiner machen, indem wir einfach andere Symbole einsetzen:

$$P(A) = P(A \mid E_1) \cdot P(E_1) + P(A \mid E_2) \cdot P(E_2); \qquad E_1 \cap E_2 = \emptyset; \qquad E_1 \cup E_2 = \Omega$$

Das, was hier links vom Gleichheitszeichen steht, ist eine totale Wahrscheinlichkeit (»total«, weil sie nicht bedingt ist), und diese Formel heißt daher »Formel für die totale Wahrscheinlichkeit«. So, wie sie hier steht, gilt sie aber nur für eine Partition der Ereignismenge, die aus nur zwei Teilereignissen E_1 und E_2 besteht. Man kann diese Formel auch noch allgemeiner, für alle möglichen Partitionen formulieren:

Formel für die totale Wahrscheinlichkeit

Für eine beliebige Partition $E_1, \cdots, E_n$ mit

$$E_i \cap E_j = \emptyset \text{ für alle } i \neq j$$

und

$$\bigcup_{i=1}^{n} E_i = \Omega$$

und für ein beliebiges Ereignis $A \subseteq \Omega$ kann man die totale Wahrscheinlichkeit von A nach dieser Formel berechnen:

$$P(A) = \sum_{i=1}^{n} P(A \cap E_i) = \sum_{i=1}^{n} P(A \mid E_i) \cdot P(E_i)$$

Fortsetzung des Beispiels:

Wir sind stehengeblieben bei

$$P(€) = P(€ \mid ♀) \cdot P(♀) + P(€ \mid ♂) \cdot P(♂)$$

Hier können wir die schon bekannten Werte einsetzen:

$$P(€) = \frac{3}{10} \cdot \frac{1}{5} + \frac{1}{10} \cdot \frac{4}{5}$$

$$P(€) = \frac{3}{50} + \frac{4}{50} = \frac{7}{50} = \frac{14}{100}$$

Die totale Wahrscheinlichkeit für die Vergabe eines Stipendiums liegt bei 14%. Dieser Wert liegt unter der vorgegebenen Untergrenze von 15%. Die Richtlinien zur Vergabe der Stipendien wurden daher nicht eingehalten.

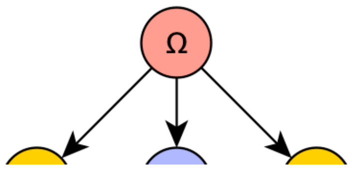
5 Wahrscheinlichkeitsbäume

Die Formeln für bedingte und totale Wahrscheinlichkeiten, der Multiplikationssatz, der Satz von Bayes und weitere Formeln sind nicht immer leicht zu merken. Alles, was man mit den genannten Formeln berechnen kann, kann man fast immer auch mit einem Wahrscheinlichkeitsbaum berechnen. Die entsprechenden Formeln ergeben sich dann fast wie von selbst.

5.1 Wahrscheinlichkeitsbaum

Ein Wahrscheinlichkeitsbaum ist ein Baum im Sinn der Graphentheorie. Wenn Ihnen die Graphentheorie nicht geläufig ist, lesen Sie bitte das Dokument *Graphentheorie*. Machen sie das bitte auch, wenn Ihnen nicht vollständig klar ist, was die Begriffe *Knoten, Kante, Pfad, Wurzel, Blatt* und *Teilbaum* bedeuten. In dem genannten Dokument wird alles über Graphen und Bäume erklärt, was Sie wissen müssen, um Wahrscheinlichkeitsbäume zu verstehen.

In einem Wahrscheinlichkeitsbaum repräsentiert jeder Knoten ein Ereignis, und an der Wurzel steht Ω, die Ereignismenge.



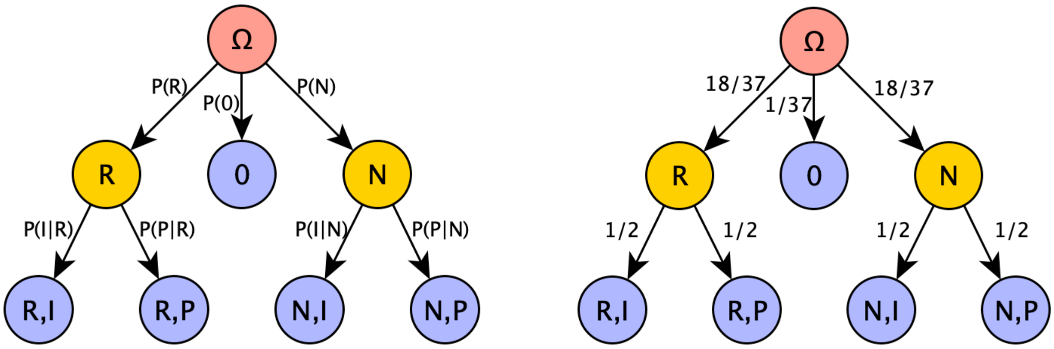
Die Wurzel eines Wahrscheinlichkeitsbaumes ist immer Ω.

Alle Knoten, die direkt auf einen bestimmten Knoten folgen, sind unvereinbare Ereignisse, und ihre Vereinigungsmenge ist genau jenes Ereignis, aus dem sie hervorgegangen sind. Mit anderen Worten: Alle unmittelbaren Nachfolgeereignisse eines Ereignisses in einem Wahrscheinlichkeitsbaum bilden gemeinsam eine Partition dieses Ereignisses.

Beispiel:

Beim Roulette kann die Zahl 0 »gezogen« werden, oder eine rote (R = rouge) oder eine schwarze (N = noir) Zahl. Die roten und schwarzen Zahlen können gerade (P = pair) oder ungerade (I = impair) sein.

Die Kanten werden mit den jeweiligen Wahrscheinlichkeiten beschriftet. Wenn der Pfeil nicht direkt von Ω ausgeht, muss das eine bedingte Wahrscheinlichkeit sein.



Wahrscheinlichkeitsbaum für Kombinationen von 0/R/N und I/P,
links sind die Kanten mit Formelzeichen beschriftet, rechts mit konkreten Werten.
R,I steht für das gemeinsame Auftreten von R und I, soll also $R \cap I$ bedeuten.
Das gilt sinngemäß auch für die anderen Blätter.

Die Werte, die bei den Pfeilen stehen, die von einem bestimmten Knoten ausgehen, müssen zusammen immer die Summe 1 ergeben.

Jeder Knoten in diesem Baum steht bekanntermaßen für ein Ereignis, und die Wahrscheinlichkeit jedes Ereignisses kann man errechnen, indem man alle Wahrscheinlichkeiten, die entlang des Pfades von der Wurzel bis zum jeweiligen Knoten stehen, miteinander multipliziert. Das ist auch genau das, was auch der Multiplikationssatz aussagt (siehe 3.2 auf Seite 17):

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

Zur Wurzel führt kein Pfad. Da dort aber das sichere Ereignis steht, ist dessen Wahrscheinlichkeit auch immer klar, sie beträgt 1.

Die im Beispiel vorkommenden absoluten Wahrscheinlichkeiten sind daher:

Ω	1	$= 1$	$R \cap I$	$\frac{18}{37} \cdot \frac{1}{2}$	$= \frac{9}{37}$
0	$\frac{1}{37}$	$= \frac{1}{37}$	$R \cap P$	$\frac{18}{37} \cdot \frac{1}{2}$	$= \frac{9}{37}$
R	$\frac{18}{37}$	$= \frac{18}{37}$	$N \cap I$	$\frac{18}{37} \cdot \frac{1}{2}$	$= \frac{9}{37}$
N	$\frac{18}{37}$	$= \frac{18}{37}$	$N \cap P$	$\frac{18}{37} \cdot \frac{1}{2}$	$= \frac{9}{37}$

Eine weitere wichtige Regel von Wahrscheinlichkeitsbäumen ist, dass die Summe der Wahrscheinlichkeiten aller Blätter 1 ergeben muss. Im Beispiel sind das die blauen Knoten:

$$P(0) + P(R \cap I) + P(R \cap P) + P(N \cap I) + P(N \cap P) \; = \; \frac{1}{37} + \frac{9}{37} + \frac{9}{37} + \frac{9}{37} + \frac{9}{37} \; = \; \frac{37}{37} \; = \; 1$$

Das ist so, weil alle Blätter eines Wahrscheinlichkeitsbaumes gemeinsam eine Partition der Ereignismenge Ω sind.

5.2 Totale Wahrscheinlichkeit im Wahrscheinlichkeitsbaum

Bleiben wir gleich beim oben dargestellten Roulette-Beispiel: Wie groß ist die totale Wahrscheinlichkeit dafür, dass die Kugel bei einer ungeraden Zahl landet? Wir suchen also die totale Wahrscheinlichkeit für das Ereignis I = impair.

Nach der Formel auf Seite 23 (Abschnitt 4.1) geht das so:

Allgemeine Formel:

$$P(A) = \sum_{i=1}^{n} P(A \cap E_i) = \sum_{i=1}^{n} P(A \mid E_i) \cdot P(E_i)$$

Aufs Beispiel angewandt:

$$P(P) = P(R \cap I) + P(N \cap I) = P(I \mid R) \cdot P(R) + P(I \mid N) \cdot P(N)$$

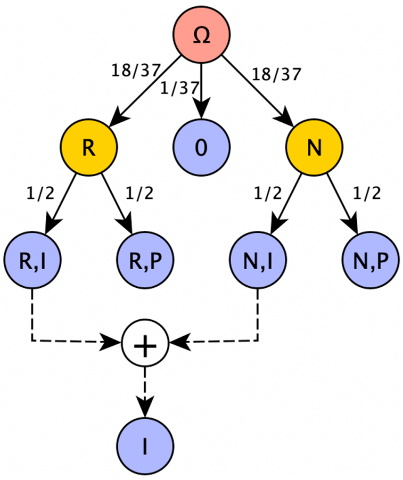
Zahlen aus der Angabe einsetzen:

$$P(I) = \frac{1}{2} \cdot \frac{18}{37} + \frac{1}{2} \cdot \frac{18}{37} = \frac{9}{37} + \frac{9}{37} = \frac{18}{37}$$

Wie geht das nun im Wahrscheinlichkeitsbaum?

Dabei muss man ausnutzen, dass alle Blätter untereinander unvereinbar sind. Damit kann man nach Axiom 3 die Wahrscheinlichkeit von I als Summe der Wahrscheinlichkeit zweier Blätter ausrechnen.

Mehr muss man gar nicht berücksichtigen. Man muss einfach nur die (bedingten) Wahrscheinlichkeiten entlang der Pfade zu den beiden relevanten Blättern multiplizieren und dann addieren, und kommt so ganz einfach zum gewünschten Ergebnis, ohne sich viele Formeln merken zu müssen:

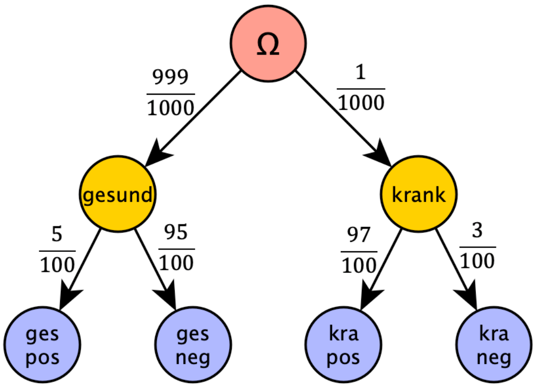


$$P(I) = \frac{18}{37} \cdot \frac{1}{2} + \frac{18}{37} \cdot \frac{1}{2} = \frac{9}{37} + \frac{9}{37} = \frac{18}{37}$$

Gesucht ist die Wahrscheinlichkeit dafür, tatsächlich an einer bestimmten Krankheit erkrankt zu sein, wenn ein medizinischer Test eine Krankheit anzeigt, wenn man also ein positives Testergebnis hat und gleichzeitig anzunehmen ist, dass der Test nicht zu 100% zuverlässig ist. Wir suchen also

$$P(\text{krank} \mid \text{pos})$$

Dabei gehen wir in diesem Beispiel von folgenden Wahrscheinlichkeiten aus:



Dieser Baum ist wie folgt zu interpretieren:

- **Erste Ebene**
  Die Wahrscheinlichkeit, dass eine zufällig ausgewählte Person an der betrachteten Krankheit erkrankt ist, beträgt $\frac{1}{1000}$. Die Gegenwahrscheinlichkeit (Wahrscheinlichkeit gesund zu sein) beträgt $\frac{1}{1000}$. Das heißt: Im statistischen Mittel ist eine von 1000 Personen ist tatsächlich krank. Im Fall von Österreich (ziemlich genau 9 Millionen Einwohner) entspricht das insgesamt 9.000 Menschen, die die betrachtete Krankheit haben.

- **Zweite Ebene**
  Links: Die Wahrscheinlichkeit, dass der Test bei einer gesunden Person positiv ist, also fälschlicherweise anzeigt, dass sie krank ist, beträgt 5%. Da hier ein falsches positives Ergebnis angezeigt wird, nennt man diese Wahrscheinlichkeit auch die Falsch-Positiv-Rate.
  Links: Die Wahrscheinlich, dass der Test bei einer kranken Person ein negatives Ergebnis hat, also fälschlicherweise anzeigt, dass sie gesund ist, beträgt 3%. Da hier ein falsches negatives Ergebnis angezeigt wird, nennt man diese Wahrscheinlichkeit auch die Falsch-Negativ-Rate.

Bei medizinischen Tests, aber auch bei fast allen anderen Prüfverfahren, ist es oftmals aus vielerlei Gründen nicht möglich, diese beiden Raten ganz auf 0 zu senken.

Wir erinnern uns an den Multiplikationssatz:

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

Daher

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

$$P(\text{krank}|\text{pos}) = \frac{P(\text{krank} \cap \text{pos})}{P(\text{pos})}$$

Wir müssen aus dem Baum also $P(\text{krank} \cap \text{pos})$ (eines der Blätter) und $P(\text{pos})$ (eine totale Wahrscheinlichkeit) ermitteln, und dann den Quotienten bilden.

Zuerst $P(\text{krank} \cap \text{pos})$:

$$P(\text{krank} \cap \text{pos}) = \frac{1}{1000} \cdot \frac{97}{100} = \frac{97}{100.000}$$

Dann die totale Wahrscheinlichkeit von $P(\text{pos})$. Wie wir diesen Wert zu berechnen haben, können wir am Baum ablesen

$$P(\text{pos}) = P(\text{krank} \cap \text{pos}) \; + \; P(\text{gesund} \cap \text{pos})$$

Den linken Summanden haben wir gerade vorhin ausgerechnet, und wie wir den rechten erhalten, sieht man auch am Baum

$$P(\text{pos}) \; = \frac{97}{100.000} + \frac{999}{1000} \cdot \frac{5}{100} = \frac{97}{100.000} + \frac{4995}{100.000} = \frac{5092}{100.000}$$

Nun dividieren:

$$P(\text{krank}|\text{pos}) = \frac{P(\text{krank} \cap \text{pos})}{P(\text{pos})} = \frac{\frac{97}{100.000}}{\frac{5092}{100.000}} = \frac{97}{100.000} \cdot \frac{100.000}{5092} = \frac{97}{5092}$$

$$P(\text{krank}|\text{pos}) = \frac{97}{5092} \cong 1{,}9\%$$

Die Wahrscheinlichkeit wirklich krank zu sein, wenn der Test ein positives Ergebnis anzeigt, liegt also bei nur 1,9%.

Wie ginge das mit der Formel von Bayes? Also ohne Baum?

Das ist die Formel:

$$P(\text{krank}|\text{pos}) = P(\text{pos}|\text{krank}) \cdot \frac{P(\text{krank})}{P(\text{pos})}$$

$P(\text{pos}|\text{krank})$ und $P(\text{krank})$ sind Werte, die direkt in der Angabe stehen:

- $P(\text{pos}|\text{krank}) = \frac{97}{100}$
- $P(\text{krank}) = \frac{1}{1000}$

Für $P(\text{pos})$ muss man auswendig die Formel für totale Wahrscheinlichkeiten kennen:

$$P(\text{pos}) = P(\text{pos}|\text{krank}) \cdot P(\text{krank}) \; + \; P(\text{pos}|\text{gesund}) \cdot P(\text{gesund})$$

$$P(\text{pos}) = \frac{97}{100} \cdot \frac{1}{1000} + \frac{5}{100} \cdot \frac{9999}{1000} = \frac{5092}{100.000}$$

Dann in obige Formel einsetzen:

$$P(\text{krank}|\text{pos}) = \frac{97}{100} \cdot \frac{\frac{1}{1000}}{\frac{5092}{100.000}} = \frac{97}{5092} \cong 1{,}9\%$$

Die Rechnerei ist natürlich dieselbe, und wer auch ohne Baum den Überblick behält, wie man was ausrechnet, kommt auch ohne den Wahrscheinlichkeitsbaum aus. Ihn aufzuzeichnen kann aber helfen, den Überblick über die Rechnung zu bewahren.

6 Wahrscheinlichkeiten mit Mengen rechnen

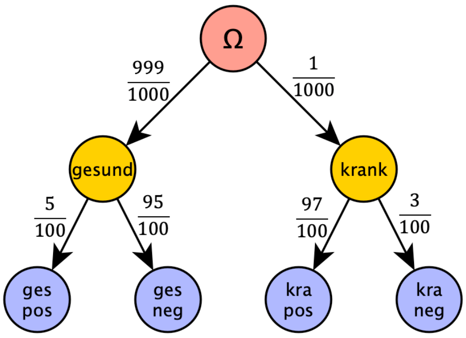
Viele Menschen haben Schwierigkeiten, die Bedeutung von Wahrscheinlichkeiten zu erfassen, die rationale oder sogar reelle Zahlen zwischen 0 und 1 sind. Im »wirklichen Leben« hat man es mit ganzen Zahlen zu tun: Der Satz *»Von 1000 Menschen ist derzeit eine Person an Corona erkrankt«* ist wesentlich leichter intuitiv zu erfassen als *»0,1% aller Menschen sind derzeit an Corona erkrankt«*. Vor allem dann, wenn man sich klar macht, dass z.B. ein voll besetzter U-Bahn-Zug der Wiener Linien zu den Stoßzeiten ungefähr 800 Personen transportiert (U6: 776 Personen pro Zug, andere Linien: getrennte Wagons: 840, Zug zum Durchgehen: 882)

Andere Größen von Menschenmengen, die man sich merken kann:

- ca. 100 = 1 voller U-Bahn-Wagon
- ca. 2000 = Fassungsvermögen des Großen Saals im Wiener Konzerthaus, Fassungsvermögen der »kleinen« Wiener Stadthalle (Halle F)
- ca. 4000 = Fassungsvermögen der Gasometer-Halle
- ca. 15.000 = Fassungsvermögen der Wiener Stadthalle (Halle D).
- ca. 50.000 = Fassungsvermögen des Ernst-Happel-Stadions.
- ca. 300.000 = Bewohner von Graz. Bewohner des Burgenlandes
- ca. 2.000.000 = Bewohner von Wien oder Hamburg
- ca. 9.000.000 = Bewohner Österreichs
- ca. 25.000.000 = Einwohner Australiens
- ca. 50.000.000 = Bewohner von Kenia oder Südkorea
- ca. 80.000.000 = Bewohner Deutschlands
- ca. 100.000.000 = Bewohner der Philippinen oder von Äthiopien
- ca. 330.000.000 = Einwohner der USA
- ca. 450.000.000 = Einwohner der EU. Einwohner Südamerikas
- ca. 600.000.000 = Einwohner Nordamerikas
- ca. 750.000.000 = Einwohner Europas
- ca. 1,200.000.000 = Einwohner Afrikas
- ca. 1.400.000.000 = Bewohner von China oder Indien
- ca. 4.500.000.000 = Einwohner Asiens
- knapp 8.000.000.000 = Aktuelle Weltbevölkerung

6.1 Satz von Bayes mit Mengen rechnen

Bleiben wir bei dem Beispiel von vorhin und orientieren wir uns am Wahrscheinlichkeitsbaum (das ist nicht zwingend notwendig, erhöht aber die Übersichtlichkeit)



Zunächst suchen wir eine passend große Zahl für die Anzahl der Menschen, die wir betrachten wollen. Das Produkt der Nenner entlang eines Pfades ist dafür eine gute Wahl. In unserem Beispiel also $1000 \cdot 100 = 100.000$, das ist zweimal das vollbesetzte Ernst-Happel-Stadion. Das ist die Anzahl der Menschen, die mit Ω gleichzusetzen sind.

Diese Zahl multipliziert man nun mit $\frac{999}{1000}$ bzw. $\frac{1}{1000}$ um die Anzahl der gesunden bzw. kranken Menschen zu erhalten. (Um genau zu sein: Diese Zahlen entsprechen dem Erwartungswert für die jeweilige Anzahl.)

- gesund: 99.900
- krank: 100

Dann folgt man den Pfaden im Baum und kommt zu folgenden Zahlen:

- gesund und positiv: $99.900 \cdot \frac{5}{100} = 4995$
- gesund und negativ: $99.900 \cdot \frac{95}{100} = 94905$
- krank und positiv: $100 \cdot \frac{97}{100} = 97$
- krank und negativ: $100 \cdot \frac{3}{100} = 3$

Die Anzahl der Menschen, bei denen der Test ein positives Ergebnis meldet, beträgt

$$4995 + 97 = 5092$$

Wirklich krank sind davon aber nur 97 Menschen. Also 97 von 5092. Die exakte Wahrscheinlichkeit ist dann

$$\frac{97}{5092} = 0{,}01904949 \cong 1{,}9\%$$

Wenn man ein wenig rundet, sind das ungefähr 100 von 5000 oder 1 von 50. (Entspricht 2%, was als Schätzung ausreichend genau ist.)

Wenn man einen U-Bahn-Wagon ausschließlich mit Menschen füllt, die ein positives Testergebnis bekommen haben, sind nur 2 kranke Menschen im Wagon.