Häufig verwendet man die Definition der bedingten Wahrscheinlichkeit in der Form

$$\Pr[A \cap B] = \Pr[B|A] \cdot \Pr[A] = \Pr[A|B] \cdot \Pr[B]\,. \tag{1}$$

Damit:

Satz 16 (Multiplikationssatz)

Seien die Ereignisse $A_1, \ldots, A_n$ *gegeben. Falls* $\Pr[A_1 \cap \ldots \cap A_n] > 0$ *ist, gilt*

$$\begin{aligned}\Pr[A_1 \cap \ldots \cap A_n] = {} & \\ & \Pr[A_1] \cdot \Pr[A_2|A_1] \cdot \Pr[A_3|A_1 \cap A_2] \cdot \ldots \\ & \qquad \ldots \cdot \Pr[A_n|A_1 \cap \ldots \cap A_{n-1}]\,.\end{aligned}$$

Beweis:

Zunächst halten wir fest, dass alle bedingten Wahrscheinlichkeiten wohldefiniert sind, da $\Pr[A_1] \geq \Pr[A_1 \cap A_2] \geq \ldots \geq \Pr[A_1 \cap \ldots \cap A_n] > 0$.

Die rechte Seite der Aussage im Satz können wir umschreiben zu

$$\frac{\Pr[A_1]}{1} \cdot \frac{\Pr[A_1 \cap A_2]}{\Pr[A_1]} \cdot \frac{\Pr[A_1 \cap A_2 \cap A_3]}{\Pr[A_1 \cap A_2]} \cdot \ldots \cdot \frac{\Pr[A_1 \cap \ldots \cap A_n]}{\Pr[A_1 \cap \ldots \cap A_{n-1}]}.$$

Offensichtlich kürzen sich alle Terme bis auf $\Pr[A_1 \cap \ldots \cap A_n]$. $\square$

Beispiel 17 (Geburtstagsproblem)

Wie groß ist die Wahrscheinlichkeit, dass in einer m-köpfigen Gruppe zwei Personen am selben Tag Geburtstag haben?

Umformulierung:
Man werfe m Bälle zufällig und gleich wahrscheinlich in n Körbe. Wie groß ist die Wahrscheinlichkeit, dass nach dem Experiment jeder Ball allein in seinem Korb liegt?

Für das Geburtstagsproblem: $n = 365$

Offensichtlich muss $m \leq n$ sein, damit überhaupt jeder Ball allein in einem Korb liegen kann.
Wir nehmen an, dass die Bälle nacheinander geworfen werden. A_i bezeichne das Ereignis „Ball i landet in einem noch leeren Korb“. Das gesuchte Ereignis „Alle Bälle liegen allein in einem Korb“ bezeichnen wir mit A. Nach Satz 16 können wir $\Pr[A]$ berechnen durch

$$\begin{aligned}\Pr[A] &= \Pr\left[\cap_{i=1}^{m} A_i\right] \\ &= \Pr[A_1] \cdot \Pr[A_2|A_1] \cdot \ldots \cdot \Pr[A_m| \cap_{i=1}^{m-1} A_i].\end{aligned}$$

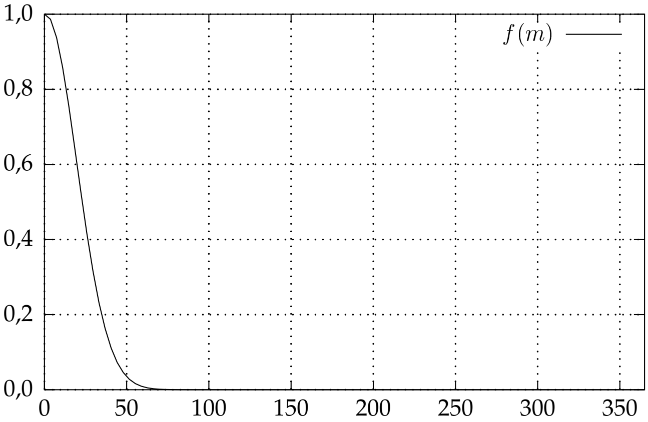
Unter der Bedingung, dass die ersten $j-1$ Bälle jeweils in einem leeren Korb gelandet sind, bedeutet A_j, dass der j-te Ball in eine der $n-(j-1)$ leeren Körbe fallen muss, die aus Symmetriegründen jeweils mit derselben Wahrscheinlichkeit gewählt werden.

Daraus folgt

$$\Pr[A_j | \cap_{i=1}^{j-1} A_i] = \frac{n-(j-1)}{n} = 1 - \frac{j-1}{n}.$$

Mit der Abschätzung $1 - x \le e^{-x}$ und wegen $\Pr[A_1] = 1$ erhalten wir

$$\begin{aligned}\Pr[A] &= \prod_{j=1}^{m}\left(1 - \frac{j-1}{n}\right)\\ &\le \prod_{j=2}^{m} e^{-(j-1)/n} = e^{-(1/n)\cdot\sum_{j=1}^{m-1} j}\\ &= e^{-m(m-1)/(2n)} =: f(m)\,.\end{aligned}$$



Verlauf von $f(m)$ für $n = 365$

Ausgehend von der Darstellung der bedingten Wahrscheinlichkeit in Gleichung 1 zeigen wir:

Satz 18 (Satz von der totalen Wahrscheinlichkeit)

Die Ereignisse $A_1, \ldots, A_n$ *seien paarweise disjunkt und es gelte* $B \subseteq A_1 \cup \ldots \cup A_n$*. Dann folgt*

$$\Pr[B] = \sum_{i=1}^{n} \Pr[B|A_i] \cdot \Pr[A_i] \,.$$

Analog gilt für paarweise disjunkte Ereignisse $A_1, A_2, \ldots$ *mit* $B \subseteq \bigcup_{i=1}^{\infty} A_i$*, dass*

$$\Pr[B] = \sum_{i=1}^{\infty} \Pr[B|A_i] \cdot \Pr[A_i] \,.$$

Beweis:

Wir zeigen zunächst den endlichen Fall. Wir halten fest, dass

$$B = (B \cap A_1) \cup \ldots \cup (B \cap A_n) \; .$$

Da für beliebige i, j mit $i \neq j$ gilt, dass $A_i \cap A_j = \emptyset$, sind auch die Ereignisse $B \cap A_i$ und $B \cap A_j$ disjunkt. Wegen (1) folgt $\Pr[B \cap A_i] = \Pr[B|A_i] \cdot \Pr[A_i]$ (auch für den Fall, dass $\Pr[A_i] = 0$!). Wir wenden nun den Additionssatz (Lemma 8, Teil 5) an

$$\begin{aligned} \Pr[B] = \Pr[B \cap A_1] + \ldots + \Pr[B \cap A_n] = \\ \Pr[B|A_1] \cdot \Pr[A_1] + \ldots + \Pr[B|A_n] \cdot \Pr[A_n] \end{aligned}$$

und haben damit die Behauptung gezeigt. Da der Additionssatz auch für unendlich viele Ereignisse $A_1, A_2, \ldots$ gilt, kann dieser Beweis direkt auf den unendlichen Fall übertragen werden. $\square$

Mit Hilfe von Satz 18 erhalten wir leicht einen weiteren nützlichen Satz:

Satz 19 (Satz von Bayes)

Die Ereignisse $A_1, \ldots, A_n$ *seien paarweise disjunkt, mit* $\Pr[A_j] > 0$ *für alle* j. *Ferner sei* $B \subseteq A_1 \cup \ldots \cup A_n$ *ein Ereignis mit* $\Pr[B] > 0$. *Dann gilt für ein beliebiges* $i = 1, \ldots, n$

$$\Pr[A_i|B] = \frac{\Pr[A_i \cap B]}{\Pr[B]} = \frac{\Pr[B|A_i] \cdot \Pr[A_i]}{\sum_{j=1}^{n} \Pr[B|A_j] \cdot \Pr[A_j]} .$$

Analog gilt für paarweise disjunkte Ereignisse $A_1, A_2, \ldots$ *mit* $B \subseteq \bigcup_{i=1}^{\infty} A_i$, *dass*

$$\Pr[A_i|B] = \frac{\Pr[A_i \cap B]}{\Pr[B]} = \frac{\Pr[B|A_i] \cdot \Pr[A_i]}{\sum_{j=1}^{\infty} \Pr[B|A_j] \cdot \Pr[A_j]} .$$

Mit dem Satz von Bayes dreht man gewissermaßen die Reihenfolge der Bedingung um. Gegeben die Wahrscheinlichkeit von B unter den Bedingungen A_i (sowie die Wahrscheinlichkeiten der A_i selbst), berechnet man die Wahrscheinlichkeit von A_i bedingt auf das Ereignis B.

Thomas Bayes (1702–1761) war ein bekannter Theologe und Mitglied der Royal Society. Als sein bedeutendstes Werk gilt sein Beitrag zur Wahrscheinlichkeitstheorie „Essay Towards Solving a Problem in the Doctrine of Chances". Diese Arbeit wurde erst 1763 publiziert.

3. Unabhängigkeit

Bei einer bedingten Wahrscheinlichkeit $\Pr[A|B]$ kann der Fall auftreten, dass die Bedingung auf B, also das Vorwissen, dass B eintritt, keinen Einfluss auf die Wahrscheinlichkeit hat, mit der wir das Eintreten von A erwarten. Es gilt also $\Pr[A|B] = \Pr[A]$, und wir nennen dann die Ereignisse A und B unabhängig.

Beispiel 20 (Zweimaliges Würfeln)

$$\Omega := \{(i,j) \mid 1 \leq i,j \leq 6\}\ .$$

Alle Elementarereignisse erhalten nach dem Prinzip von Laplace die Wahrscheinlichkeit $\frac{1}{36}$.
Wir definieren die Ereignisse

$$\begin{aligned} A &:= \text{Augenzahl im ersten Wurf ist gerade,}\\ B &:= \text{Augenzahl im zweiten Wurf ist gerade,}\\ C &:= \text{Summe der Augenzahlen beider Würfe beträgt 7.} \end{aligned}$$

Es gilt $\Pr[A] = \Pr[B] = \frac{1}{2}$ und $\Pr[C] = \frac{1}{6}$. Wie groß ist $\Pr[B|A]$?

Beispiel 20 (Forts.)

Nach unserer Intuition beeinflusst der Ausgang des ersten Wurfs den zweiten Wurf nicht. Daher gewinnen wir durch das Eintreten von A keine Information in Bezug auf das Ereignis B hinzu:

$$B \cap A = \{(2,2),(2,4),(2,6),(4,2),(4,4),(4,6),(6,2),(6,4),(6,6)\}.$$

Daraus folgt

$$\Pr[B|A] = \frac{\Pr[B \cap A]}{\Pr[A]} = \frac{\frac{9}{36}}{\frac{1}{2}} = \frac{1}{2} = \Pr[B]\,.$$

Das Eintreffen des Ereignisses B hat mit dem Ereignis A „nichts zu tun".

Definition 21

Die Ereignisse A und B heißen unabhängig, wenn gilt

$$\Pr[A \cap B] = \Pr[A] \cdot \Pr[B] \; .$$

Falls $\Pr[B] \neq 0$, so können wir diese Definition zu

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

umschreiben.

Beispiel 20 (Zweimaliges Würfeln, Forts.)

Zur Erinnerung:

$$
\begin{aligned}
A &:= \text{Augenzahl im ersten Wurf ist gerade,}\\
B &:= \text{Augenzahl im zweiten Wurf ist gerade,}\\
C &:= \text{Summe der Augenzahlen beider Würfe beträgt 7.}
\end{aligned}
$$

Bei den Ereignissen A und B ist die Unabhängigkeit klar, da offensichtlich kein kausaler Zusammenhang zwischen den Ereignissen besteht. Wie steht es mit A und C?

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

und damit

$$\Pr[A \cap C] = \frac{3}{36} = \frac{1}{2} \cdot \frac{1}{6} = \Pr[A] \cdot \Pr[C] \text{ bzw. } \Pr[C|A] = \Pr[C]\,.$$

Beispiel 20 (Forts.)

Also sind auch A und C (und analog B und C) unabhängig.

Bemerkung: Im Beispiel ist $A \cap C \neq \emptyset$.
Es gilt sogar allgemein für zwei unabhängige Ereignisse A und B mit $\Pr[A], \Pr[B] > 0$, dass sie gar nicht disjunkt sein können, da ansonsten

$$0 = \Pr[\emptyset] = \Pr[A \cap B] \neq \Pr[A] \cdot \Pr[B] \ .$$

Beispiel 20 (Zweimaliges Würfeln (Forts.))

Zur Erinnerung:

$$A := \text{Augenzahl im ersten Wurf ist gerade,}$$
$$B := \text{Augenzahl im zweiten Wurf ist gerade,}$$
$$C := \text{Summe der Augenzahlen beider Würfe beträgt 7.}$$

Wir betrachten das Ereignis $A \cap B \cap C$. Wenn $A \cap B$ eintritt, so sind beide gewürfelten Augenzahlen gerade und somit ergibt auch die Summe davon eine gerade Zahl. Daraus folgt $\Pr[A \cap B \cap C] = 0$ bzw. $\Pr[C|A \cap B] = 0 \neq \Pr[C]$. Das Ereignis $A \cap B$ liefert uns also Information über das Ereignis C.

Definition 22

Die paarweise verschiedenen Ereignisse $A_1, \ldots, A_n$ heißen unabhängig, wenn für alle Teilmengen $I = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$ mit $i_1 < i_2 < \ldots < i_k$ gilt, dass

$$\Pr[A_{i_1} \cap \ldots \cap A_{i_k}] = \Pr[A_{i_1}] \cdot \ldots \cdot \Pr[A_{i_k}]. \tag{2}$$

Eine unendliche Familie von paarweise verschiedenen Ereignissen A_i mit $i \in \mathbb{N}$ heißt unabhängig, wenn (2) für jede endliche Teilmenge $I \subseteq \mathbb{N}$ erfüllt ist.

Lemma 23

Die (paarweise verschiedenen) Ereignisse $A_1, \ldots, A_n$ sind genau dann unabhängig, wenn für alle $(s_1, \ldots, s_n) \in \{0,1\}^n$ gilt, dass

$$\Pr[A_1^{s_1} \cap \ldots \cap A_n^{s_n}] = \Pr[A_1^{s_1}] \cdot \ldots \cdot \Pr[A_n^{s_n}], \tag{3}$$

wobei $A_i^0 = \bar{A}_i$ und $A_i^1 = A_i$.

Beweis:

Zunächst zeigen wir, dass aus (2) die Bedingung (3) folgt. Wir beweisen dies durch Induktion über die Anzahl der Nullen in $s_1, \ldots, s_n$. Wenn $s_1 = \ldots = s_n = 1$ gilt, so ist nichts zu zeigen. Andernfalls gelte ohne Einschränkung $s_1 = 0$. Aus dem Additionssatz folgt dann

$$\begin{aligned}\Pr[\bar{A}_1 \cap A_2^{s_2} \cap \ldots \cap A_n^{s_n}] &= \Pr[A_2^{s_2} \cap \ldots \cap A_n^{s_n}] \\ &\quad -\Pr[A_1 \cap A_2^{s_2} \cap \ldots \cap A_n^{s_n}].\end{aligned}$$

Darauf können wir die Induktionsannahme anwenden und erhalten

$$\begin{aligned}&\Pr[\bar{A}_1 \cap A_2^{s_2} \cap \ldots \cap A_n^{s_n}] \\ &= \Pr[A_2^{s_2}] \cdot \ldots \cdot \Pr[A_n^{s_n}] - \Pr[A_1] \cdot \Pr[A_2^{s_2}] \cdot \ldots \cdot \Pr[A_n^{s_n}] \\ &= (1 - \Pr[A_1]) \cdot \Pr[A_2^{s_2}] \cdot \ldots \cdot \Pr[A_n^{s_n}],\end{aligned}$$

woraus die Behauptung wegen $1 - \Pr[A_1] = \Pr[\bar{A}_1]$ folgt.

Beweis (Forts.):

Für die Gegenrichtung zeigen wir nur, dass aus (3) $\Pr[A_1 \cap A_2] = \Pr[A_1] \cdot \Pr[A_2]$ folgt. Es gilt wegen des Satzes von der totalen Wahrscheinlichkeit, dass

$$\begin{aligned}
\Pr[A_1 \cap A_2] &= \sum_{s_3,\ldots,s_n \in \{0,1\}} \Pr[A_1 \cap A_2 \cap A_3^{s_3} \cap \ldots \cap A_n^{s_n}] \\
&= \sum_{s_3,\ldots,s_n \in \{0,1\}} \Pr[A_1] \cdot \Pr[A_2] \cdot \Pr[A_3^{s_3}] \cdot \ldots \cdot \Pr[A_n^{s_n}] \\
&= \Pr[A_1] \cdot \Pr[A_2] \cdot \sum_{s_3=0,1} \Pr[A_3^{s_3}] \cdot \ldots \cdot \sum_{s_n=0,1} \Pr[A_n^{s_n}] \\
&= \Pr[A_1] \cdot \Pr[A_2],
\end{aligned}$$

und es folgt die Behauptung. $\square$

Aus der Darstellung in Lemma 23 folgt die wichtige Beobachtung, dass für zwei unabhängige Ereignisse A und B auch die Ereignisse $\bar{A}$ und B (und analog auch A und $\bar{B}$ bzw. $\bar{A}$ und $\bar{B}$) unabhängig sind!

Ebenso folgt:

Lemma 24

Seien A, B und C unabhängige Ereignisse. Dann sind auch $A \cap B$ und C bzw. $A \cup B$ und C unabhängig.

Beweis:

Die Unabhängigkeit von $A \cap B$ und C folgt unmittelbar aus Definition 22.
Aus

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

folgt die Unabhängigkeit von $A \cup B$ und C. □

4. Zufallsvariablen

4.1 Grundlagen

Anstatt der Ereignisse selbst sind wir oft an „Auswirkungen“ oder „Merkmalen“ der (Elementar)Ereignisse interessiert.

Definition 25

Sei ein Wahrscheinlichkeitsraum auf der Ergebnismenge Ω gegeben. Eine Abbildung

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

heißt (numerische) Zufallsvariable.

Eine Zufallsvariable X über einer endlichen oder abzählbar unendlichen Ergebnismenge Ω heißt diskret.

Bei diskreten Zufallsvariablen ist der Wertebereich

$$W_X := X(\Omega) = \{x \in \mathbb{R};\ \exists\, \omega \in \Omega \text{ mit } X(\omega) = x\}$$

ebenfalls wieder endlich (bzw. abzählbar unendlich).

Beispiel 26

Wir werfen eine ideale Münze drei Mal. Als Ergebnismenge erhalten wir $\Omega := \{H, T\}^3$. Die Zufallsvariable Y bezeichne die Gesamtanzahl der Würfe mit Ergebnis „Head".

Beispielsweise gilt also $Y(HTH) = 2$ und $Y(HHH) = 3$. Y hat den Wertebereich $W_Y = \{0, 1, 2, 3\}$.

Für $W_X = \{x_1, \ldots, x_n\}$ bzw. $W_X = \{x_1, x_2, \ldots\}$ betrachten wir (für ein beliebiges $1 \leq i \leq n$ bzw. $x_i \in \mathbb{N}$) das Ereignis

$$A_i := \{\omega \in \Omega;\ X(\omega) = x_i\} = X^{-1}(x_i).$$

Bemerkung: Anstelle von $\Pr[X^{-1}(x_i)]$ verwendet man häufig auch die Schreibweise $\Pr[„X = x_i“]$. Analog setzt man

$$\begin{aligned} \Pr[„X \leq x_i“] &= \sum_{x \in W_X : x \leq x_i} \Pr[„X = x“] \\ &= \Pr[\{\omega \in \Omega;\ X(\omega) \leq x_i\}]\,. \end{aligned}$$

Oft lässt man auch die Anführungszeichen weg.

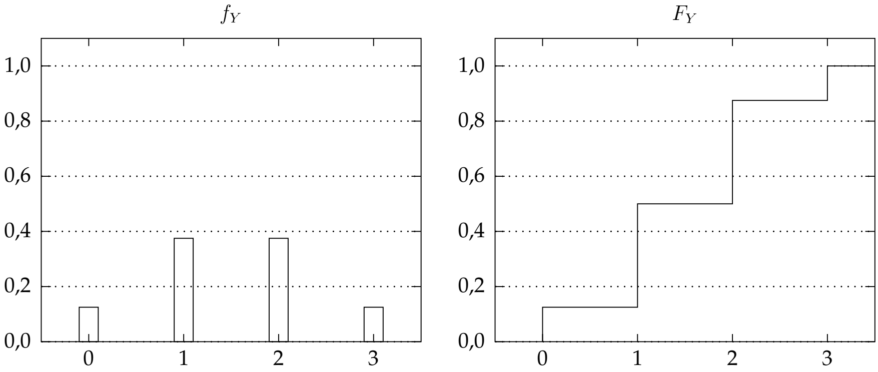
Definition 27

- Die Funktion
$$f_X : \mathbb{R} \ni x \mapsto \Pr[X = x] \in [0, 1] \tag{4}$$
nennt man (diskrete) Dichte(funktion) der Zufallsvariablen X.
- Die Funktion
$$F_X : \mathbb{R} \ni x \mapsto \Pr[X \leq x] = \sum_{x' \in W_X \,:\, x' \leq x} \Pr[X = x'] \in [0, 1] \tag{5}$$
heißt Verteilung(sfunktion) der Zufallsvariablen X.

Beispiel 28

Für die Zufallsvariable Y erhalten wir

$$\begin{aligned}
\Pr[Y=0] &= \Pr[TTT] = \frac{1}{8},\\
\Pr[Y=1] &= \Pr[HTT]+\Pr[THT]+\Pr[TTH] = \frac{3}{8},\\
\Pr[Y=2] &= \Pr[HHT]+\Pr[HTH]+\Pr[THH] = \frac{3}{8},\\
\Pr[Y=3] &= \Pr[HHH] = \frac{1}{8}.
\end{aligned}$$



Dichte und Verteilung von Y

Bemerkung: Man kann statt Ω auch den zugrunde liegenden Wahrscheinlichkeitsraum über W_X betrachten.