Die Wahrscheinlichkeit, dass Prof. L. die Wette verliert, lässt sich wie folgt berechnen:

$$\Omega = \{(i_1, \ldots, i_7) : i_j \in \{1 \ldots 7\},\ j = 1 \ldots 7\},$$

wobei i, j *für den Wochentag steht, an dem die Person* j *geboren ist. Die Wahrscheinlichkeit, dass Prof. L. die Wette verliert entspricht nun dem folgenden Pfad:*

Die Wahrscheinlichkeit für diesen Pfad ist

$$\frac{7 \cdot 6 \ldots 2 \cdot 1}{7^7} = \frac{7!}{7^7} = \frac{6!}{7^6} = 0,00612,$$

die Wette ist also günstig für den Prof.

Beispiel 3.15 *In M. gibt es nur zwei Arten Wetter: Nass (N) und Trocken (T). Ist es heute nass, so ist es mit Wahrscheinlichkeit* $\frac{5}{6}$ *auch morgen nass und mit Wahrscheilichkeit* $\frac{1}{6}$ *trocken. Ist es heute trocken, so ist es morgen mit Wahrscheinlichkeit* $\frac{3}{10}$ *auch trocken und mit Wahrscheinlichkeit* $\frac{7}{10}$ *nass. [Frage: Welche Stadt ist M.?] Heute ist es trocken: Mit welcher Wahrscheinlichkeit ist es übermorgen auch trocken?*

Offenbar gibt es zwei Pfade, dass es übermorgen trocken ist:

Die Wahrscheinlichkeit, dass es übermorgen trocken ist, berechnet sich also so:

$$\frac{3}{10} \cdot \frac{3}{10} + \frac{7}{10} \cdot \frac{1}{6} = \frac{9}{100} + \frac{7}{60} = \frac{27}{300} + \frac{36}{300} = \frac{52}{300} = \frac{13}{75}.$$

3.3 Unabhängigkeit und bedingte Wahrscheinlichkeit I

Zur Motivation begeben wir uns noch einmal zurück zu Beispiel 3.9 und 3.11, also den Beispielen, die in mehrstufige Zufallsexperimente eingeführt haben. Wir wollen erkennen, dass es sich bei der Multiplikationsregel um eine Form der *Unabhängigkeit* bzw. *bedingten*

Unabhängigkeit handelt, wobei wir den letzten Begriff nicht definieren, wohl aber kurz anreißen. In der Tat ist ja in Beispiel 3.9 durch das Zurücklegen der gezogenen Buchstaben die Grundsituation vor jedem Zug dieselbe. Da zudem die Buchstaben "gedächtnislos" sind, d. h. die Wahrscheinlichkeit ein a oder ein n zu ziehen nicht davon abhängt, ob ich zuvor ein a oder ein n gezogen habe, kann man guten Gewissens von der Unabhängigkeit der beiden Züge in Beispiel 3.9 sprechen. Die Situation in Beispiel 3.11 ist dagegen eine andere. Hier hängt die Wahrscheinlichkeit im zweiten Zug ein a oder ein n zu ziehen, sehr wohl vom ersten Zug ab, denn hier sind dann entsprechend weniger a's oder n's in der Urne, die beiden Züge sind also nicht unabhängig.

Wir wollen den Begriff der Unabhängigkeit nun mathematisch fassen. Dazu müssen wir zunächst den Begriff der "bedingten Wahrscheinlichkeit" definieren. Hierzu sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $A, B \in \mathcal{A}$ seien zwei Ereignisse. Nehmen wir nun an, jemand mit hellseherischen Fähigkeiten sagte uns, dass das Ereignis B eintritt. Was ist nun – bedingt darauf, dass wir wissen, dass B eintritt – die Wahrscheinlichkeit, dass auch A eintritt? Diese Wahrscheinlichkeit wollen wir mit $\mathbb{P}(A|B)$ bezeichnen. Da wir schon wissen, dass B eintritt, kommen für das Eintreten von A nur die $\omega \in \Omega$ in Frage, die sowohl in A ($\omega \in A$) als auch in B ($\omega \in B$) sind, also die ω mit

$$\omega \in A \cap B.$$

$\mathbb{P}(A|B)$ muss also proportional sein zu $\mathbb{P}(A \cap B)$. Allerdings ist $\mathbb{P}(A \cap B)$ in A keine Wahrscheinlichkeit mehr, denn für $A = \Omega$ gilt

$$\mathbb{P}(\Omega \cap B) = \mathbb{P}(B),$$

und dies ist in der Regel nicht 1. Um also eine Wahrscheinlichkeit zu erhalten, *teilen* wir $\mathbb{P}(A \cap B)$ durch $\mathbb{P}(B)$ (vorausgesetzt das ist nicht null, was wir aber getrost fordern dürfen, denn die Berechnung der Wahrscheinlichkeit eines Ereignisses A, vorausgesetzt ein Ereignis B tritt ein, dass aber sowieso nie eintritt, führt zu offensichtlichen logischen Problemen). Dies ergibt die folgende

Definition 3.16 *Es sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $A, B \in \mathcal{A}$ mit $\mathbb{P}(B) \neq 0$. Als die bedingte Wahrscheinlichkeit von A gegeben B definieren wir*

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

Beispiel 3.17 *Wir modellieren den einfachen fairen Würfelwurf mit $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{A} = \mathcal{P}\Omega$ und*

$$\mathbb{P}(\{\omega\}) = \frac{1}{6} \quad \textit{für alle} \quad \omega \in \Omega.$$

Es seien A und B die Ereignisse

$$\begin{aligned} A &= \textit{“Die Augenzahl ist durch 3 teilbar”}, \\ B &= \textit{“Die Augenzahl ist gerade”}, \end{aligned}$$

also

$$A = \{3,6\} \quad \textit{und} \quad B = \{2,4,6\}.$$

Dann ist

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{6\})}{\mathbb{P}(\{2,4,6\})} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}.$$

Mit Kenntnis von Definition 3.16 ist es nun einfach, den Begriff der Unabhängigkeit zu definieren. Es liegt nahe, zwei Ereignisse A und B, die den Voraussetzungen von Definition 3.16 genügen, unabhängig zu nennen, wenn die Kenntnis des Eintretens von B die Wahrscheinlichkeit des Eintretens von A nicht ändert, d. h. wenn

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

gilt. Mit anderen Worten soll gelten

$$\begin{aligned} \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} &= \mathbb{P}(A) \\ \Leftrightarrow \quad \mathbb{P}(A \cap B) &= \mathbb{P}(A) \cdot \mathbb{P}(B). \end{aligned}$$

Die letzte Zeile hat hierbei den Vorteil, dass sie symmetrisch in A und B ist, dass sie auch für $P(B) = 0$ sinnvoll und zweckmäßig ist, und dass deshalb auf die Voraussetzung $\mathbb{P}(B) > 0$ verzichtet werden kann.

Definition 3.18 *Es sei* $(\Omega, \mathcal{A}, \mathbb{P})$ *ein Wahrscheinlichkeitsraum,* $A, B \in \mathcal{A}$ *heißen unabhängig, falls*

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

gilt.

Beispiel 3.19 *Wir wollen nun zur Situation von Beispiel 3.9 zurückkehren und rechtfertigen, dass unsere Intuition dort nicht fehlgeleitet war und der erste und zweite Zug dort tatsächlich unabhängig sind. Hier war ja*

$$\begin{aligned} &\Omega = \{(a,a),(a,n),(n,a),(n,n)\} \\ &\mathcal{A} = \mathcal{P}\Omega \qquad \textit{und} \\ &\mathbb{P}(\{(a,a)\}) = \frac{3}{5} \cdot \frac{3}{5} = \frac{9}{25},\ \mathbb{P}(\{(a,n)\}) = \frac{3}{5} \cdot \frac{2}{5} = \frac{6}{25}, \\ &\mathbb{P}(\{(n,a)\}) = \frac{2}{5} \cdot \frac{3}{5} = \frac{6}{25},\ \mathbb{P}(\{(n,n)\}) = \frac{2}{5} \cdot \frac{2}{5} = \frac{4}{25}. \end{aligned}$$

Es sei nun

$$\begin{aligned} A &= \text{“}a \textit{ bei der ersten Ziehung”} \\ B &= \text{“}n \textit{ bei der zweiten Ziehung”}, \end{aligned}$$

also

$$A = \{(a,n),(a,a)\} \quad \textit{und} \quad B = \{(a,n),(n,n)\}.$$

Dann sind A und B unabhängig. In der Tat gilt ja

$$A \cap B = \{(a,n)\}$$

und weiter

$$\begin{aligned} \mathbb{P}(A) &= \frac{9}{25} + \frac{6}{25} = \frac{3}{5} \\ \mathbb{P}(B) &= \frac{6}{25} + \frac{4}{25} = \frac{2}{5}. \end{aligned}$$

Also

$$\mathbb{P}(A \cap B) = \frac{6}{25} = \frac{3}{5} \cdot \frac{2}{5} = \mathbb{P}(A)\mathbb{P}(B).$$

Dass Unabhängigkeit im stochastischen Sinn durchaus nicht immer etwas mit unserer Intuition, dass Unabhängigkeit von "unabhängigen" (d. h. verschiedenen, gegenseitig unbeeinflussten) Versuchen stammt, zu tun haben muss, zeigt ein weiterer Blick auf Beispiel 3.17.

Beispiel 3.20 *Die Ereignisse*

$$\begin{aligned} A &= \text{“Die Augenzahl ist durch 3 teilbar”} \qquad \textit{und} \\ B &= \text{“Die Augenzahl ist gerade”} \end{aligned}$$

sind beim einfachen fairen Würfeln unabhängig (obwohl beide Ereignisse sich auf denselben Wurf beziehen!). In der Tat haben wir ja in Beispiel 3.17 berechnet, dass

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

gilt. Außerdem gilt auch $\mathbb{P}(A) = \frac{1}{3}$*, also*

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

Man kann sich nun fragen, wie sich der Begriff der Unabhängigkeit auf mehrere Ereignisse überträgt. Dass man für Ereignisse $A_1, \ldots, A_n$ nicht einfach ihre Unabhängigkeit mit dem Bestehen der Identität

$$\mathbb{P}(A_1 \cap \ldots \cap A_n) = \mathbb{P}(A_1) \ldots \mathbb{P}(A_n)$$

gleichsetzt, hat den Grund darin, dass man Unabhängigkeit zu einer *vererbbaren* Eigenschaft machen möchte, d. h. man möchte, dass mit $A_1, \ldots, A_n$ auch jede Teilfamilie von $A_1, \ldots, A_n$ unabhängig ist. Dies wäre mit einer solchen Definition nicht gewährleistet. Daher definiert man direkt:

Definition 3.21 *Es sei* $(\Omega, \mathcal{A}, \mathbb{P})$ *ein Wahrscheinlichkeitsraum und* $A_1, \ldots, A_n \in \mathcal{A}$, $n \in \mathbb{N}$, $n \geq 2$. $A_1, \ldots, A_n$ *heißen unabhängig, falls für jedes* $1 \leq m \leq n$ *und alle* $i_1, \ldots, i_m$ *gilt*

$$\mathbb{P}(A_{i_1} \cap \ldots \cap A_{i_m}) = \mathbb{P}(A_{i_1}) \ldots \mathbb{P}(A_{i_m}).$$

Man sieht unmittelbar, dass diese Definition dem oben formulierten Vererbbarkeitsgrundsatz genügt. Der Zusammenhang zwischen der Unabhängigkeit zum einen und den im vorhergehenden Kapitel formulierten Pfadregeln zum anderen besteht nun darin, dass diese Pfadregeln für unabhängige Versuche unmittelbar evident sind. In der Tat wollen wir bei n Zufallsexperimenten, $n \in \mathbb{N}$, $n \geq 2$, davon sprechen, dass diese unabhängig sind, falls alle Ereignisse $A_1, \ldots, A_n$, wobei A_j, $j = 1, \ldots, n$, ein beliebiges Ereignis ist, das sich aber *nur auf das j'te Experiment beziehen darf*, unabhängig sind. Damit folgt

Proposition 3.22 *Sind die Experimente in einem mehrstufigen Zufallsversuch unabhängig, so gilt die Pfadregel, Definition 3.10.*

Beweis: Dies folgt in der Tat unmittelbar aus der Definition der unabhängigen Experimente bzw. der Definition der Unabhängigkeit von n Ereignissen, Definition 3.21. □

Allerdings haben wir in Definition 3.10 die Pfadregel nicht nur für unabhängige, sondern für beliebige Ereignisse formuliert. Wieso ist diese richtig? [Die Frage ist formal falsch, denn eine Definition kann natürlich nicht richtig oder falsch sein, die Frage ist vielmehr, wieso Definition 3.10 nicht im Widerspruch zur in diesem Kapitel entwickelten Theorie steht.] Die Antwort erhält man, indem man die Formel für die bedingte Wahrscheinlichkeit umstellt. Es ist ja

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

Analog erhält man für drei Ereignisse $A, B, C \in \mathcal{A}$ durch zweimaliges Anwenden dieser Regel:

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

(dies lässt sich auch beweisen, indem man einfach die Definitionen der entsprechenden bedingten Wahrscheinlichkeiten einsetzt).

Ebenso zeigt man die folgende Produktformel für bedingte Wahrscheinlichkeiten von Ereignissen $A_1, \ldots, A_n$, $n \in \mathbb{N}$, $n \geq 2$:

$$\mathbb{P}(A_1 \cap \ldots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_2 \cap A_1) \ldots \mathbb{P}(A_n|A_1 \cap \ldots \cap A_{n-1}). \tag{3.4}$$

Hat man nun ein mehrstufiges Zufallsexperiment mit n Stufen, und $A_1, \ldots, A_n$ sind Ereignisse, wobei sich das Ereignis A_j nur auf das j-te Experiment bezieht, so ist (3.4) nichts anderes als die Pfadregel Definition 3.10, die wir somit auf dem Hintergrund der bedingten Wahrscheinlichkeit gerechtfertigt haben.