

## Definition

Gegeben sei ein Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, \mathbb{P})$  und Ereignisse  $A_1, \dots, A_n \in \mathcal{A}$  für ein  $n \in \mathbb{N}$ . Die Ereignisse heißen *unabhängig*, falls für jedes  $1 \leq m \leq n$  und jede Teilmenge  $\{i_1, \dots, i_m\} \subseteq \{1, \dots, n\}$  gilt

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

Natürlich impliziert die Unabhängigkeit von  $A_1, \dots, A_n$  auch die **paarweise Unabhängigkeit**, d.h. die Unabhängigkeit je zweier Ereignisse:

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$$

für  $i \neq j$

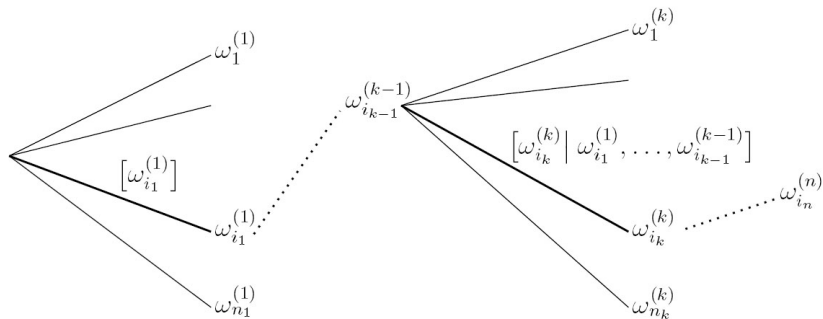
- Ein  $n$ -stufiges Zufallsexperiment ist ein Experiment, dass aus  $n$  Einzelversuchen besteht, wobei wir annehmen, dass die Versuchsausgänge der späteren Versuche die früheren nicht beeinflussen, wohl aber möglicherweise umgekehrt.
- Als Grundmenge  $\Omega$  nehmen wir

$$\Omega := \Omega_1 \times \Omega_2 \times \dots \times \Omega_n := \{\omega = (\omega_1, \dots, \omega_n) \mid \omega_i \in \Omega_i, 1 \leq i \leq n\}.$$

- Dabei enthält für jedes  $i = 1, \dots, n$  die Menge  $\Omega_i$  die möglichen Versuchsausgänge des  $i$ -ten Experimentes, und zwar *ohne* Berücksichtigung aller anderen Experimente.
- Die Menge  $\Omega_i$  beschreibt den  $i$ -ten Versuch so, als sei er unbeeinflusst von den übrigen Versuchen. Die Abhängigkeit wird also nicht mit der Menge  $\Omega$  selbst modelliert, sondern mit  $\mathbb{P}$ .
- Wenn alle  $\Omega_i$  endliche Mengen sind, dann ist auch  $\Omega$  endlich und wir wählen, wie immer,  $\mathcal{P}(\Omega)$  als  $\sigma$ -Algebra über  $\Omega$ .

- Konstruieren wir ein Baumdiagramm, d.h. einen Baum, der sich von der Wurzel aus in  $|\Omega_1| =: n_1$  viele Zweige spaltet.
- Die so entstehenden Knoten der *ersten Generation* beschriften wir mit den Elementen  $\omega_1^{(1)}$  bis  $\omega_{n_1}^{(1)}$  aus  $\Omega_1$ . Von jedem Knoten  $\omega_{i_1}^{(1)}$  der ersten Generation verzweigt sich der Baum um den Faktor  $|\Omega_2| =: n_2$ .
- Die Knoten dieser zweiten Generation werden jeweils mit  $\omega_{i_2}^{(2)}$ ,  $1 \leq i_2 \leq n_2$  indiziert.
- Ganz allgemein verzweigt sich der Baum von jedem Knoten  $\omega_{i_{k-1}}^{(k-1)}$ ,  $1 \leq i_{k-1} \leq n_{k-1}$ , der  $(k-1)$ -ten Generation um den Faktor  $|\Omega_k| =: n_k$  und die Knoten der  $k$ -ten Generation werden jeweils mit  $\omega_{i_k}^{(k)}$ ,  $1 \leq i_k \leq n_k$ , indiziert.

# Mehrstufige Experimente/Baumdiagramm



- Zu einem Knoten  $\omega_{i_k}^{(k)}$  der  $k$ -ten Generation für  $2 \leq k \leq n$  gibt es genau einen Pfad von diesem Knoten zurück zur Wurzel.
- Verläuft dieser Pfad über die Knoten  $\omega_{i_{k-1}}^{(k-1)}$  bis  $\omega_{i_1}^{(1)}$ , so beschriften wir den Zweig führt, mit der bedingten Wahrscheinlichkeit

$$\mathbb{P}\left(\{\omega \in \Omega \mid \omega_k = \omega_{i_k}^{(k)}\} \mid \{\omega \in \Omega \mid \omega_j = \omega_{i_j}^{(j)} \text{ für alle } 1 \leq j \leq k-1\}\right).$$

- Hierfür schreiben wir abkürzend auch  $[\omega_{i_k}^{(k)} \mid \omega_{i_1}^{(1)}, \dots, \omega_{i_{k-1}}^{(k-1)}]$ .
- Einen Zweig, der zu dem Knoten  $\omega_{i_1}^{(1)}$  der ersten Generation führt, beschriften wir einfach mit

$$\mathbb{P}\left(\{\omega \in \Omega \mid \omega_1 = \omega_{i_1}^{(1)}\}\right) =: [\omega_{i_1}^{(1)}].$$

- Mithilfe dieses Baumdiagramms haben wir nun für unser mehrstufiges Zufallsexperiment eine leichte Regel, mit der man die Wahrscheinlichkeiten einzelner  $\omega \in \Omega$  berechnet.

## Regel (Pfadregel)

*In einem mehrstufigen Zufallsexperiment ist die Wahrscheinlichkeit eines  $\omega$  das Produkt der Wahrscheinlichkeiten entlang seines Pfades im zugehörigen Baumdiagramm.*

- Wann immer der zugrunde liegende Wahrscheinlichkeitsraum ein Laplace-Experiment beschreibt – egal, ob ein ein- oder ein mehrstufiges – ist es wichtig, dass wir die Mächtigkeit von Ereignismengen bestimmen können.
- Denn wir beschreiben ein Ereignis häufig durch einfaches Aufzählen all seiner Elementarereignisse; diese abzuzählen ist dann letztlich gleich der Berechnung der gesuchten Wahrscheinlichkeit.
- Die mathematische Disziplin, die sich mit dem Abzählen von Mengen beschäftigt, ist die *Kombinatorik*.
- Wer sich jemals ein wenig mit Kombinatorik beschäftigt hat, weiß, dass Zählen in der Tat ein Problem sein kann und es kombinatorische Probleme gibt, die extrem schwierig und bis zum heutigen Tage ungelöst sind

## Definition

Eine Familie  $A_1, A_2, \dots$  von Teilmengen einer Menge  $\Omega$  heißt *Partition von  $\Omega$* , falls

- $A_1, A_2, \dots$  paarweise disjunkt sind und
- $\Omega = A_1 \cup A_2 \cup \dots$

Die Partition heißt *endlich*, wenn die Familie  $A_1, A_2, \dots$  endlich ist.

## Regel (Summenregel)

Sei  $\Omega$  eine endliche Menge und  $A_1, \dots, A_r$  eine endliche Partition von  $\Omega$ , dann gilt

$$|\Omega| = |A_1| + \dots + |A_r| = \sum_{i=1}^r |A_i|.$$

## Regel (spezielle Produktregel)

Seien  $\Omega_1, \dots, \Omega_r$  endliche Mengen, dann gilt

$$\begin{aligned} |\Omega_1 \times \dots \times \Omega_r| &= |\{(\omega_1, \dots, \omega_r) \mid \omega_i \in \Omega_i \text{ für alle } 1 \leq i \leq r\}| \\ &= |\Omega_1| \cdot \dots \cdot |\Omega_r|. \end{aligned}$$

## Regel (allgemeine Produktregel)

*Ein Versuch bestehe aus  $r$  nacheinander durchgeführten Experimenten, wobei der Ausgang eines Experiments zwar alle nachfolgenden beeinflussen kann, aber nur derart, dass die Anzahl der möglichen Ausgänge der nachfolgenden Experimente von diesem Ausgang unabhängig sind. Sind  $n_1, n_2, \dots, n_r$  die erwähnten Anzahlen, dann hat der Gesamtversuch*

$$n_1 \cdot n_2 \cdot \dots \cdot n_r$$

*mögliche Ausgänge.*

- Wir werden nun die Regeln benutzen, um die Größe einer aus einer Grundmenge gezogenen Stichprobe zu berechnen.
- Wichtige Faktoren: Die Art, wie wir die Stichprobe auswählen, und die Frage, ob wir die Anordnung der Elemente der Stichprobe berücksichtigen wollen.
- Bei der Auswahl der Stichprobe können wir entweder bereits gezogene Elemente der Grundmenge für weitere Ziehungen zulassen oder nicht.
- Je nachdem spricht man vom *Ziehen mit Zurücklegen* oder *Ziehen ohne Zurücklegen*.
- Außerdem nennen wir die Stichprobe *geordnet* bzw. *ungeordnet*, je nachdem, ob wir der Reihenfolge der gezogenen Elemente eine Bedeutung beimessen oder nicht.
- Ausgangssituation: Wir haben eine Grundmenge der Größe  $n$ , aus der wir  $s$  Elemente wählen. Dazu stellen wir uns eine Urne mit  $n$  Kugeln vor, die von 1 bis  $n$  nummeriert sind.

## Geordnete Stichproben mit Zurücklegen

- Wir ziehen hier eine von  $n$  Kugeln zufällig, notieren ihre Nummer und legen sie wieder zurück in die Urne.
- Dies wird  $s$ -mal wiederholt.
- Offenbar besitzt dieser Versuch  $s$  Schritte oder Stufen, wobei keine Stufe die anderen Stufen beeinflusst.
- Jeder Schritt hat  $n$  mögliche Versuchsausgänge und daher gibt es aufgrund der Produktregel  $n^s$  mögliche Versuchsausgänge des Gesamtversuchs.

## Duale Darstellung:

- Statt  $s$  Mal aus  $n$  Kugeln zu ziehen, legen wir  $s$  Kugeln in  $n$  Urnen.
- Genauer seien  $s$  unterscheidbare (z.B. nummerierte) Kugeln gegeben und  $n$  unterscheidbare leere Urnen.
- Wir legen nun die  $s$  Kugeln in die  $n$  Urnen; sowohl für die erste Kugel, wie auch für alle nachfolgenden haben wir  $n$  Möglichkeiten zur Auswahl.
- Insgesamt haben wir also wieder  $n^s$  mögliche Versuchsausgänge.

## Geordnete Stichproben ohne Zurücklegen:

- Wieder ziehen wir  $s$ -mal aus einer Urne mit  $n$  nummerierten Kugeln, diesmal jedoch, *ohne* die Kugeln zurückzulegen.
- Natürlich erfordert dies  $s \leq n$ .
- Wir notieren die Nummern in der Reihenfolge, in der sie erscheinen. Es gibt also  $s$  Stufen, die jedoch keine Kopien voneinander sind.
- Wir haben also für den ersten Schritt  $n$  mögliche Ausgänge, für den zweiten noch  $n - 1$  usw. Für den  $s$ -ten Schritt bleiben  $n - s + 1$  Möglichkeiten.
- Nach der allgemeinen Produktregel hat das Gesamtexperiment

$$(n)_s := n \cdot (n - 1) \cdot \dots \cdot (n - s + 1)$$

mögliche Ausgänge.

- Ist  $s = n$ , d.h. wird die Urne vollständig leer gezogen, so gibt es

$$n! := (n)_n = n \cdot (n - 1) \cdot \dots \cdot 3 \cdot 2 \cdot 1$$

mögliche Versuchsausgänge.

- Für größere  $n$  benutzt man häufig die *Stirlingformel*

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}.$$

zur Approximation, die nach James Stirling (1696–1770) benannt wurde.

- Hierbei schreiben wir  $a_n \sim b_n$  für zwei Folgen  $(a_n)_n$  und  $(b_n)_n$ , falls

$$\frac{a_n}{b_n} \longrightarrow 1 \text{ für } n \rightarrow \infty$$

gilt.

- Dies bedeutet, dass der *relative Fehler*  $(a_n - b_n)/b_n$ , den wir durch die Approximation in Kauf nehmen, gegen Null geht.
- Für den *absoluten Fehler*  $a_n - b_n$  muss dies keineswegs gelten!

## Duale Darstellung:

- Es sollen  $s$  unterscheidbare Kugeln auf  $n$  unterscheidbare Urnen verteilt werden.
- Es gelte aber das folgende *Ausschließungsprinzip*: In jeder Urne darf höchstens eine Kugel liegen.
- Für die erste Kugel stehen dann  $n$  Urnen zur Verfügung, für die zweite noch  $n - 1$  und so weiter; für die  $s$ -te Kugel kann noch zwischen  $n - s + 1$  Urnen gewählt werden.
- Auch hier ergeben sich also

$$(n)_s = n \cdot (n - 1) \dots (n - s + 1)$$

verschiedene Versuchsausgänge.