2 Bedingte Wahrscheinlichkeit und Unabhängigkeit

2.1 Bedingte Wahrscheinlichkeit

Ist über den Ausgang eines Zufallsexperiments bereits eine Teilinformation verfügbar, ändern sich entsprechend die Wahrscheinlichkeiten der Ereignisse.

Beispiel

Zweimaliges Würfeln eines fairen Würfels

$$P(\texttt{Augensumme} \ > 10) = P(\{(5,6),(6,5),(6,6)\}) = \frac{1}{12}\,.$$

Wie ändert sich diese Wahrscheinlichkeit, wenn bereits bekannt ist, dass beim ersten Würfeln eine 6 gewürfelt wurde? Unter dieser Annahme bleiben nur noch sechs gleichwahrscheinliche Möglichkeiten für die zweite Augenzahl übrig, von denen die Augenzahlen 5 und 6 insgesamt zu einer Augensumme größer als 10 führen. Für die Wahrscheinlichkeit des Ereignisses `Augenzahl` > 10 unter der Bedingung `1.Augenzahl` 6 ergibt sich somit

$$\begin{aligned} &P(\texttt{Augensumme} \ > 10 \mid \texttt{1.Augenzahl} \ 6) \\ &\qquad = \frac{P(\texttt{Augensumme > 10, 1.Augenzahl 6})}{P(\texttt{1.Augenzahl 6})} = \frac{\frac{2}{36}}{\frac{6}{36}} = \frac{1}{3}\,. \end{aligned}$$

Die bedingte Wahrscheinlichkeit ist also viermal höher als die ursprüngliche “a priori” Wahrscheinlichkeit.

Definition 2.1. Es sei $(\Omega, \mathcal{A}, P)$ ein beliebiger Wahrscheinlichkeitsraum und A, $B \in \mathcal{A}$ Ereignisse mit $P(B) > 0$. Dann heißt

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

die **bedingte Wahrscheinlichkeit von A unter der Bedingung B** (oder auch: **die bedingte Wahrscheinlichkeit von A gegeben B)**. Im Falle $P(B) = 0$ setzen wir einfach $P(A \mid B) := 0$.

Satz 2.2. Es sei $(\Omega, \mathcal{A}, P)$ ein beliebiger Wahrscheinlichkeitsraum und A, $B \in \mathcal{A}$. Dann gelten:

(i) Für $P(B) > 0$ und $A \supseteq B$ gilt $P(A|B) = 1$.

(ii) Für $A \cap B = \varnothing$ gilt $P(A \mid B) = 0$.

(iii) Für $P(B) > 0$ ist $P(\cdot \mid B)$ ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{A})$.

Beweis. (i) $A \supseteq B$ impliziert $A \cap B = B$ und damit

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1\,.$$

(ii) $A \cap B = \varnothing$ impliziert $P(A \cap B) = 0$, also $P(A \mid B) = 0$.

(iii) Offensichtlich ist die Abbildung

$$A \mapsto P(A \mid B)\,, A \in \mathcal{A}$$

wohldefiniert, nichtnegativ und, wegen $P(B) > 0$, auch normiert $P(\Omega \mid B) = 1$. Um die σ-Additivität zu zeigen, sei $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ eine Folge paarweise disjunkter Ereignisse. Dann ist auch $(A_n \cap B)_{n\in\mathbb{N}}$ eine Folge paarweiser disjunkter Ereignisse in $\mathcal{A}$, also

$$P\left(\left(\bigcup_{n\geq 1} A_n\right) \cap B\right) = P\left(\bigcup_{n\geq 1} (A_n \cap B)\right) = \sum_{n\geq 1} P(A_n \cap B) \tag{2.1}$$

aufgrund der σ-Additivität von P. Folglich

$$\begin{aligned} P\left(\bigcup_{n\in\mathbb{N}} A_n \mid B\right) &= \frac{P\left(\left(\bigcup_{n\in\mathbb{N}} A_n\right) \cap B\right)}{P(B)} = \frac{\sum_{n\geq 1} P(A_n \cap B)}{P(B)} \\ &= \sum_{n\geq 1} \frac{P(A_n \cap B)}{P(B)} = \sum_{n\geq 1} P(A_n \mid B)\,. \end{aligned}$$

□

Beispiel 2.3. (Laplacescher Wahrscheinlichkeitsraum)

Es sei Ω endlich und $P(A) = \frac{|A|}{|\Omega|}$ die Gleichverteilung auf Ω. Dann folgt für $B \neq \emptyset$

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

Insbesondere: Die bedingte Wahrscheinlichkeitsverteilung ist im Falle des Laplaceschen Wahrscheinlichkeitsraumes gerade die Gleichverteilung auf B.

Beispiel 2.4. Bedingte Wahrscheinlichkeiten bilden die Grundlage für das Tarifsystem von Versicherungen. Verunglücken etwa mehr Männer als Frauen, sollten entsprechende Prämien einer Versicherung gegen Arbeitsunfälle für Männer höher als für Frauen sein, etwa:

$$\begin{aligned} P(\texttt{Unfall} \mid \texttt{V weiblich}) &= 0.002 \\ P(\texttt{Unfall} \mid \texttt{V männlich}) &= 0.005\,. \end{aligned}$$

Kennt man noch den Anteil der männlichen und weiblichen Versicherungsnehmer, etwa

$$P(\texttt{V weiblich}) = \frac{2}{5} = 1 - P(\texttt{V männlich})\,,$$

so kann man hieraus die totale Wahrscheinlichkeit für einen Arbeitsunfall eines Versicherungsnehmers errechnen:

$$\begin{aligned} P(\texttt{Unfall}) &= P(\texttt{Unfall und V weiblich}) + P(\texttt{Unfall und V männlich}) \\ &= P(\texttt{Unfall} \mid \texttt{V weiblich})P(\texttt{V weiblich}) \\ &\quad + P(\texttt{Unfall} \mid \texttt{V männlich})P(\texttt{V männlich}) \\ &= 0.002\frac{2}{5} + 0.005\frac{3}{5} = 0.0038\,. \end{aligned}$$

Die Berechnung dieser "totalen" Wahrscheinlichkeit für einen Arbeitsunfall ist ein Spezialfall des folgenden Satzes.

Week 03
26.4
Fr.

Satz 2.5. (Satz von der totalen Wahrscheinlichkeit) Es sei $(\Omega, \mathcal{A}, P)$ ein beliebiger Wahrscheinlichkeitsraum und $A \in \mathcal{A}$. Weiter sei $(B_n) \subset \mathcal{A}$ eine (endliche oder unendliche) Folge paarweiser disjunkter Ereignisse in $\mathcal{A}$ mit $A \subseteq \bigcup_n B_n$. Dann gilt: abz.

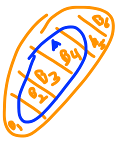
$$P(A) = \sum_n P(A \mid B_n) P(B_n)\,. \tag{2.2}$$

Überdeckung

Beweis. Offensichtlich gilt

$P(B_n) > 0$

$$P(A \mid B_n) P(B_n) = \frac{P(A \cap B_n)}{P(B_n)} P(B_n) = P(A \cap B_n) \tag{2.3}$$



für alle B_n mit $P(B_n) > 0$. Ist $P(B_n) = 0$, so gilt $P(A \mid B_n)P(B_n) = P(A \cap B_n)$ ebenfalls, da beide Seiten nach Definition 0 sind, denn aufgrund der Monotonie ist $P(A \cap B_n) \leq P(B_n) = 0$. Summation über n ergibt

disjunkt

$$\begin{aligned} \sum_n P(A \mid B_n) P(B_n) &= \sum_n P(A \cap B_n) = P\left(\dot{\bigcup_n} A \cap B_n \right) \underset{A \subseteq \bigcup_n B_n}{=} P\left(A \cap \left(\dot{\bigcup_n} B_n \right) \right) \\ &= P(A) \end{aligned}$$

wobei wir in der zweiten Gleichheit benutzt haben, dass $(A \cap B_n)_n$ paarweise disjunkte Ereignisse sind, und in der letzten Gleichheit, dass $A \subseteq \bigcup_n B_n$. □

Beispiel 2.6. In obigem Beispiel 2.4 kennt man bereits die totale Wahrscheinlichkeit für einen Arbeitsunfall eines Versicherungsnehmers

$$P(\texttt{Arbeitsunfall}) = 0.0038\,.$$

P(Unf | w), P(Unf | m)

Wie lässt sich hieraus die "umgekehrte" Wahrscheinlichkeit

$$P(\texttt{V männlich} \mid \texttt{Arbeitsunfall})$$

dafür berechnen, dass ein verunglückter Versicherungsnehmer männlich ist? Offensichtlich gilt

$$\begin{aligned} P(\texttt{V männlich} \mid \texttt{Arbeitsunfall}) P(\texttt{Arbeitsunfall}) &= P(\texttt{V männlich, Arbeitsunfall}) \\ &= P(\texttt{Arbeitsunfall} \mid \texttt{V männlich}) P(\texttt{V männlich}) \end{aligned}$$

also

$$\begin{aligned} P(\texttt{V männlich} \mid \texttt{Arbeitsunfall}) &= \frac{P(\texttt{Arbeitsunfall} \mid \texttt{V männlich}) \overbrace{P(\texttt{V männlich})}^{=\frac{3}{5}}}{P(\texttt{Arbeitsunfall})} \\ &= \frac{0.003}{0.0038} = 0.789\,. \end{aligned}$$

Wie zu erwarten, handelt es sich bei einer verunglückten Person in fast 80% aller Fälle um Männer. Dieses Verhältnis kann sich aber ins Gegenteil verkehren, wenn entweder der Anteil der weiblichen Versicherungsnehmer den Anteil der männlichen Versicherungsnehmer weit übersteigt oder die bedingte Wahrscheinlichkeit $P(\texttt{Arbeitsunfall} \mid \texttt{V weiblich})$ für einen Arbeitsunfall eines weiblichen Versicherungsnehmers die entsprechende Wahrscheinlichkeit eines Arbeitsunfalles eines männlichen Versicherungsnehmers weit übersteigt.

Der folgende Satz formuliert die Rechnung aus dem letzten Beispiel allgemein.

Satz 2.7. (Formel von Bayes) Unter den Voraussetzungen von Satz 2.5 und der zusätzlichen Annahme $P(A) > 0$ gilt:

$$P(B_n \mid A) = \frac{P(A \mid B_n)P(B_n)}{\sum_k P(A \mid B_k)P(B_k)} .$$

totale W.

Beweis. Nach Satz 2.5 gilt $\sum_k P(A \mid B_k)P(B_k) = P(A)$. Also

$$\begin{aligned} P(B_n \mid A) &\overset{\text{def}}{=} \frac{P(B_n \cap A)}{P(A)} = \frac{P(A \cap B_n)}{P(B_n)} P(B_n) \frac{1}{P(A)} \\ &= \frac{P(A \mid B_n)P(B_n)}{\sum_k P(A \mid B_k)P(B_k)} . \end{aligned}$$

$= P(A|B_N)$

□

Beispiel 2.8. Die Formel von Bayes liefert mitunter scheinbar überraschende Aussagen wie im Falle des folgenden Tests auf eine seltene Krankheit.

Angenommen, 5 Promille der Bevölkerung haben eine seltene Krankheit K, d.h.

$$P(K) = 0.005 .$$

Ein medizinischer Test zeigt bei 99% der Erkrankten eine positive Reaktion, d.h.

$$P(\texttt{Test positiv} \mid K) = 0.99 .$$

Allerdings zeigt besagter Test auch bei 2% der Gesunden eine positive Reaktion, d.h.

$$P(\texttt{Test positiv} \mid K^c) = 0.02 .$$

Von besonderem Interesse ist nun offenbar folgende

Frage: Angenommen, der Test ist positiv. Wie groß ist die Wahrscheinlichkeit, dass die getestete Person tatsächlich an K erkrankt ist? Wie groß ist also die bedingte Wahrscheinlichkeit

$$P(K \mid \texttt{Test positiv})?$$

Die Formel von Bayes liefert

$$\begin{aligned} P(K \mid \texttt{Test positiv}) &= \frac{P(\texttt{Test positiv} \mid K)P(K)}{P(\text{Test positiv} \mid K) \cdot P(K) + P(\texttt{Test positiv} \mid K^c)P(K^c)} \\ &= \frac{0.99 \cdot 0.005}{0.99 \cdot 0.005 + 0.02 \cdot 0.995} = \frac{495}{2485} \sim 0.2 . \end{aligned}$$

Also: Nur in 2 von 10 Fällen mit postivem Testergebnis ist die getestete Person auch wirklich an K erkrankt.

2.2 Unabhängigkeit

Ist $P(A) = P(A \mid B)$, d.h. die Wahrscheinlichkeit von A **unabhängig** davon, ob das Ereignis B eingetreten ist oder nicht, so folgt:

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

und damit

$$P(A \cap B) = P(A) \cdot P(B)\,. \tag{2.4}$$

Zwei Ereignisse A und B mit (2.4) heißen **(stochastisch) unabhängig**.

Allgemeiner definieren wir:

Definition 2.9. Eine Familie von Ereignissen $(A_i)_{i \in I}$ heißt **(stochastisch) unabhängig**, falls für jede nichtleere endliche Teilmenge $J \subseteq I$ die Produktformel

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j)$$

gilt, d.h. also, falls für je endlich viele paarweise verschiedene $j_1, \ldots, j_n \in I$ gilt

$$P(A_{j_1} \cap \ldots \cap A_{j_n}) = P(A_{j_1}) \cdot \ldots \cdot P(A_{j_n})\,.$$

Bemerkung 2.10. (i) Aus der Produktformel $P(A_1 \cap \ldots \cap A_n) = P(A_1) \cdot \ldots \cdot P(A_n)$ für Ereignisse $A_1, \ldots, A_n \in \mathcal{A}$ $(n \geq 3)$ allein folgt noch nicht deren Unabhängigkeit. Als Gegenbeispiel betrachte $A_1 = A_2 = \ldots = A_{n-1} = A$ für $A \in \mathcal{A}$ mit $0 < P(A) < 1$ und $A_n = \varnothing$. Dann folgt

$$P(A_1 \cap \ldots \cap A_n) \overset{\text{„unab."}}{=} P(A_1) \cdot \ldots \cdot P(A_n) = 0$$

aber

$$P(\underbrace{A_1 \cap A_2}_{=A}) = P(A) \overset{0<P(A)<1}{\neq} P(A)^2 = P(A_1)P(A_2)\,.$$

(ii) Aus der paarweisen Unabhängigkeit von Ereignissen folgt noch nicht deren (vollständige) Unabhängigkeit. Als Beispiel betrachten wir beim zweimaligen Werfen einer fairen Münze die Ereignisse

$$A = \texttt{1.Wurf Zahl}$$
$$B = \texttt{2.Wurf Zahl}$$
$$C = \texttt{1.und 2.Wurf gleich}\,.$$

Diese sind paarweise unabhängig aber nicht unabhängig, denn $P(A) = P(B) = P(C) = \frac{1}{2}$, $P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4}$, aber

$$P(A \cap B \cap C) = \frac{1}{4} \neq P(A)P(B)P(C)\,.$$

Der folgende Satz liefert ein einfach zu handhabendes Kriterium für die Unabhängigkeit endlich vieler Ereignisse:

Satz 2.11. Es sei $(\Omega, \mathcal{A}, P)$ ein beliebiger Wahrscheinlichkeitsraum und $I = \{1, \ldots, n\}$. Dann sind $(A_i)_{i\in I}$ genau dann unabhängig, wenn für jede Wahl von $B_i \in \{A_i, A_i^c\}$ die Produktformel

$$P\left(\bigcap_{i\in I} B_i\right) = \prod_{i\in I} P(B_i)$$

gilt.

Beweis. (I) Es seien (A_i) zunächst unabhängig. Wir zeigen im folgenden mit einer Induktion über m: Ist $|\{i \in I : B_i \neq A_i\}| = m$, so sind die Ereignisse $(B_i)_{i\in I}$ unabhängig. Insbesondere gilt die Produktformel

$$P\left(\bigcap_{j\in J} B_j\right) = \prod_{j\in J} P(B_j)$$

für alle $\varnothing \neq J \subseteq I$.

I.A. Für $m = 0$ folgt $B_i = A_i$ für alle $i \in I$ und damit ergibt sich die Unabhängigkeit der Ereignisse $(B_i)_{i\in I}$ direkt aus der Annahme.

I.S. (von m nach $m+1$): Ist $|\{i \in I : B_i \neq A_i\}| = m+1$, so können wir durch Umnummerierung $B_1 = A_1^c$ annehmen. Nach Annahme sind die Ereignisse $A_1, B_2, \ldots, B_n$ unabhängig, also für $\varnothing \neq J \subseteq \{2, \ldots, n\}$ gelten

$$P\left(\bigcap_{j\in J} B_j\right) = \prod_{j\in J} P(B_j) \quad \text{sowie} \quad P\left(A_1 \cap \left(\bigcap_{j\in J} B_j\right)\right) = P(A_1)\prod_{j\in J} P(B_j)\,,$$

und damit auch

$$\begin{aligned} P\left(\bigcap_{j\in J\cup\{1\}} B_j\right) &= P\left(\bigcap_{j\in J} B_j\right) - P\left(A_1 \cap \left(\bigcap_{j\in J} B_j\right)\right) \\ &= \prod_{j\in J} P(B_j) - P(A_1)\prod_{j\in J} P(B_j) = (1 - P(A_1))\prod_{j\in J} P(B_j) \\ &= P(B_1)\prod_{j\in J} P(B_j)\,. \end{aligned}$$

Hieraus folgt die Unabhängigkeit der $(B_i)_{i\in I}$.

(II) Gilt umgekehrt die Produktformel für jede Wahl von $B_i \in \{A_i, A_i^c\}$, so folgt durch Addition der Produktformeln für $B_1, B_2, \ldots, B_n$ und $B_1^c, B_2, \ldots, B_n$ dass

$$P(B_2 \cap \ldots \cap B_n) = \prod_{j=2}^{n} P(B_j)\,.$$

Somit erhählt man die Produktformel für Durchschnitte von $n-1$ Mengen, dann für Durchschnitte aus $n-2$ Mengen, usw. □

Beispiel 2.12. Betrachte beim n-maligen Wurf einer fairen Münze die Ereignisse

$$A_k = k\texttt{-ter Wurf Kopf}\,, k = 1, \ldots, n\,.$$

Dann sind $(A_k)_{k\in\{1,\ldots,n\}}$ unabhängig, denn für beliebige Wahl $B_k \in \{A_n, A_k^c\}$ gilt

$$P\left(\bigcap_{k=1}^{n} B_k\right) = 2^{-n} = \prod_{k=1}^{n} P(B_k)\,.$$

Der n-malige Münzwurf ist ein Spezialfall für die unabhängige Hintereinanderausführung von Teilexperimenten (1-maliger Münzwurf). Dies führt auf Produktexperimente, die im folgenden Abschnitt betrachtet werden sollen.

2.3 Produkte von Wahrscheinlichkeitsräumen

Es seien (Ω_i, P_i) für $1 \le i \le n$ diskrete Wahrscheinlichkeitsräume. Hieraus konstruieren wir einen neuen diskreten Wahrscheinlichkeitsraum (Ω, P) durch

$$\Omega := \bigtimes_{i=1}^{n} \Omega_i = \{\omega = (\omega_1, \dots, \omega_n) \mid \omega_i \in \Omega_i \,, 1 \le i \le n\}$$

Zähldichte

$$P(\{\omega\}) := \prod_{i=1}^{n} P_i(\{\omega_i\}) \quad \text{für} \quad \omega = (\omega_1, \dots, \omega_n) \in \Omega$$

und schließlich

$$P(A) = \sum_{\omega \in A} P(\{\omega\}), A \in \mathcal{P}(\Omega)\,.$$

Dann gilt für eine Zylindermenge $A = A_1 \times \ldots \times A_n \in \mathcal{P}(\Omega)$

$$P(A) = \prod_{i=1}^{n} P_i(A_i)\,, \tag{2.5}$$

denn

$$\begin{aligned} P(A) &= \sum_{\omega \in A} P(\{\omega\}) = \sum_{\omega_1 \in A_1} \cdots \sum_{\omega_n \in A_n} \prod_{i=1}^{n} P_i(\{\omega_i\}) \\ &= \sum_{\omega_1 \in A_1} \cdots \sum_{\omega_{n-1} \in A_{n-1}} \left(\prod_{j=1}^{n-1} P_i(\{\omega_i\}) \right) P_n(A_n) \\ &= \ldots = \prod_{i=1}^{n} P_i(A_i)\,. \end{aligned}$$

Definition 2.13. Der diskrete Wahrscheinlichkeitsraum (Ω, P) heißt der **Produktraum** der Wahrscheinlichkeitsräume (Ω_i, P_i), $1 \le i \le n$. Man schreibt $\Omega = \bigtimes_{i=1}^{n} \Omega_i$, $P = \bigotimes_{i=1}^{n} P_i$ und bezeichnet P auch als **Produktmaß** der Wahrscheinlichkeitsmaße P_i, $1 \le i \le n$.

Beispiel 2.14. (i) Der n-malige Münzwurf (Ω, P) ist der Produktraum der Wahrscheinlichkeitsräume $\Omega_i = \{0,1\}$, $P_i(\{0\}) = P_i(\{1\}) = \frac{1}{2}$, $1 \le i \le n$, also der Produktraum der Wahrscheinlichkeitsräume zum 1-maligen Münzwurf: $\Omega = \bigtimes_{i=1}^{n} \Omega_i$, $P = \bigotimes_{i=1}^{n} P_i$. Betrachtet man allgemeiner für $p \in [0,1]$ das Produkt der Wahrscheinlichkeitsmaße $P_i(\{1\}) = p = 1 - P_i(\{0\})$, $1 \le i \le n$, so heißt das durch den Produktraum beschriebene Zufallsexperiment **Bernoulli-Experiment der Länge** n mit **Erfolgswahrscheinlichkeit** p.

(ii) Das Produkt Laplacescher Wahrscheinlichkeitsräume ist wieder eine Laplacescher Wahrscheinlichkeitsraum. Um dies einzusehen, seien $\Omega_1, \dots, \Omega_n$ nichtleere endliche Mengen und P_i die Gleichverteilung auf Ω_i, $1 \le i \le n$, mit zugehöriger Zähldichte $p_i(\omega_i) = \frac{1}{|\Omega_i|}$. Dann gilt für das Produktmaß

$$P(\{(\omega_1, \dots, \omega_n)\}) = P_1(\{\omega_1\}) \cdot \ldots \cdot P_n(\{\omega_n\}) = \frac{1}{|\Omega_1|} \cdot \ldots \cdot \frac{1}{|\Omega_n|} = \frac{1}{|\Omega_1 \times \ldots \times \Omega_n|}\,.$$

Also ist das Produktmaß $P = \bigotimes_{i=1}^{n} P_i$ die Gleichverteilung auf $\Omega_1 \times \ldots \times \Omega_n$.

Der folgende Satz entspricht unserer Vorstellung, dass der Produktraum $\Omega = \bigtimes_{i=1}^{n} \Omega_i$, $P = \bigotimes_{i=1}^{n} P_i$ die **unabhängige** Hintereinanderausführung der Zufallsexperimente (Ω_i, P_i) modelliert.

Satz 2.15. Für Ereignisse $A_i \in \mathcal{P}(\Omega_i)$, $1 \le i \le n$, seien

$$A_i^{(i)} = \{\omega = (\omega_1, \dots, \omega_n) \in \Omega \mid \omega_i \in A_i\} = \Omega_1 \times \dots \times \Omega_{i-1} \times A_i \times \Omega_{i+1} \times \dots \times \Omega_n$$

die Ereignisse im Produktraum, die nur vom Ereignis A_i im i-ten Zufallsexperiment abhängen. Dann sind die Ereignisse $\left(A_i^{(i)}\right)_{1 \le i \le n}$ unabhängig.

nach Satz 2.11

Beweis. Zunächst beachte, dass $\left(A_i^{(i)}\right)^c = \Omega_1 \times \dots \times \Omega_{i-1} \times A_i^c \times \Omega_{i+1} \times \dots \times \Omega_n$. Um die Produktformel $P\left(\bigcap_{i=1}^n B_i\right) = \prod_{i=1}^n P(B_i)$ für beliebige Wahl $B_i \in \{A_i^{(i)}, (A_i^{(i)})^c\}$ zu zeigen, reicht es also aus, die Produktformel

$$P\left(\bigcap_{i=1}^n A_i^{(i)}\right) = \prod_{i=1}^n P(A_i^{(i)})$$

für **beliebige** Ereignisse $A_i \in \mathcal{P}(\Omega_i)$ zu zeigen. Dies folgt jedoch aus (2.5) wegen

$$\begin{aligned}
P\left(\bigcap_{i=1}^n A_i^{(i)}\right) &= P\left(\bigcap_{i=1}^n (\Omega_1 \times \dots \times \Omega_{i-1} \times A_i \times \Omega_{i+1} \times \dots \times \Omega_n)\right) \\
&= P\left(A_1 \times \dots \times A_n\right) = \prod_{i=1}^n P_i(A_i) \\
&= \prod_{i=1}^n P\left(\Omega_1 \times \dots \times \Omega_{i-1} \times A_i \times \Omega_{i+1} \times \dots \times \Omega_n\right) \\
&= \prod_{i=1}^n P\left(A_i^{(i)}\right).
\end{aligned}$$

□

Bemerkung 2.16. Analog zum diskreten Fall kan man auch für beliebige Wahrscheinlichkeitsräume $(\Omega_i, \mathcal{A}_i, P_i)$, $1 \le i \le n$, das Produkt $(\Omega, \mathcal{A}, P)$ folgendermaßen definieren:

(a) $\Omega = \Omega_1 \times \dots \times \Omega_n$

(b) $\mathcal{A} = \mathcal{A}_1 \otimes \dots \otimes \mathcal{A}_n$ - kleinste σ-Algebra, die von allen Zylindermengen $A_1 \times \dots \times A_n$ mit $A_1 \in \mathcal{A}_1, \dots, A_n \in \mathcal{A}_n$ erzeugt wird,

(c) $P = P_1 \otimes \dots\dots \otimes P_n$ - dasjenige eindeutig bestimmte Wahrscheinlichkeitsmaß auf $\mathcal{A}$, für das $P(A_1 \times \dots \times A_n) = P_1(A_1) \cdot \dots \cdot P_n(A_n)$ $(A_1 \in \mathcal{A}_1, \dots, A_n \in \mathcal{A}_n)$.

Weiteres hierzu findet man in der Literatur zur Maß- und Integrationstheorie (siehe [Ba91]).

Beispiel 2.17. (i) Sind $f_1, \dots, f_n : \mathbb{R} \to [0, \infty)$ Dichten auf $\mathbb{R}$ und $P_i(A) = \int_A f_i(x)\,dx$, $A \in \mathcal{B}(\mathbb{R})$ die zugehörigen Wahrscheinlichkeitsmaße auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, so ist

$$f(x_1, \dots, x_n) = f_1(x_1) \cdot \dots \cdot f_n(x_n), \quad (x_1, \dots, x_n) \in \mathbb{R}^n$$

eine Dichte auf $\mathbb{R}^n$ und das zugehörige Wahrscheinlichkeitsmaß

$$P(A) = \int_A f(x)\,dx\,, A \in \mathcal{B}(\mathbb{R}^n) \tag{2.6}$$

genau das Produktmaß der $P_1, \dots, P_n$. Aus (2.6) folgt nämlich für Zylindermengen $A = A_1 \times \ldots \times A_n$

$$P(A_1 \times \ldots \times A_n) = \int_{A_1 \times \ldots \times A_n} f(x)\,dx = \int_{A_1} f_1(x_1)\,dx_1 \ldots \int_{A_n} f_n(x_n)\,dx_n$$
$$= P_1(A_1) \cdot \ldots \cdot P_n(A_n)\,.$$

(ii) Aus (i) folgt analog zum diskreten Fall auch hier: Sind $\Omega_1, \dots, \Omega_n \subseteq \mathbb{R}$ Borelmengen endlicher Länge $|\Omega_i|$, und ist $P_i(A_i) = \frac{|A_i|}{|\Omega_i|}$, die Gleichverteilung auf Ω_i, so ist das zugehörige Produktmaß $P = P_1 \otimes \ldots \otimes P_n$ die Gleichverteilung auf $\Omega = \Omega_1 \times \ldots \times \Omega_n$.