# Stochastik SS 2019

Dozent: Prof. Dr. Anita Behme

25. April 2019

# *Inhaltsverzeichnis*

# *Vorwort*

# *Was ist Stochastik?*

Altgriechisch Stochastikos ($\sigma\tau o\chi\alpha\sigma\tau\iota\kappa\grave{o}\zeta$) und bedeutet sinngemäß "scharfsinnig in Vermuten". Fragestellung insbesondere aus Glücksspiel, Versicherungs-/Finanzmathematik, überall da wo Zufall/ Risiko / Chance auftauchen.
Was ist Stochastik?

- Beschreibt zufällige Phänomene in einer exakten Spache!
  Beispiel: "Beim Würfeln erscheint jedes sechste Mal (im Schnitt) eine 6." $\longrightarrow$ Gesetz der großen Zahlen ($\nearrow$ später)
- Lässt sich mathematische Stochastik in zwei Teilgebiete unterteilen
  Wahrscheinlichkeitstheorie (Wahrscheinlichkeitstheorie) & Statistik
  - *Wahrscheinlichkeitstheorie*: Beschreibt und untersucht konkret gegebene Zufallssituationen.
  - *Statistik*: Zieht Schlussfolgerungen aus Beobachtungen.

  Statistik benötigt Modelle der Wahrscheinlichkeitstheorie. Wahrscheinlichkeitstheorie benötigt die Bestätigung der Modelle durch Statistik.

In diesem Semester konzentrieren wir uns nur auf die Wahrscheinlichkeitstheorie!

## Kapitel I

# *Grundbegriffe der Wahrscheinlichkeitstheorie*

## 1. Wahrscheinlichkeitsräume

### Ergebnisraum

Welche der möglichen Ausgänge eines zufälligen Geschehens interessieren uns?
Würfeln? Augenzahl, nicht die Lage und die Fallhöhe

**Definition 1.1 (Ergebnisraum)**
Die Menge der relevanten Ergebnisse eines Zufallsgeschehens nennen wir Ergebnisraum und bezeichnen diesen mit $\Omega$.

■ **Beispiel**

- Würfeln: $\Omega = \{1, 2, \dots, 6\}$
- Wartezeiten: $\Omega = \mathbb{R}_+ = [0, \infty)$ (überabzählbar!)

### Ereignisse

Oft interessieren wir uns gar nicht für das konkrete Ergenis des Zufallsexperiments, sondern nur für das Eintreten gewisser Ereignisse.

■ **Beispiel**

- Würfeln: Zahl ist $\geq 3$
- Wartezeit: Wartezeit $\leq 5$ Minuten

$\longrightarrow$ Teilmenge des Ereignisraums, also Element der Potenzmenge $\mathscr{P}(\Omega)$, denen eine Wahrscheinlichkeit zugeordnet werden kann, d.h. welche messbar (mb) sind.

**Definition 1.2 (Ereignisraum, messbarer Raum)**
Sei $\Omega \neq \varnothing$ ein Ergebnisraum und $\mathscr{F}$ eine $\sigma$-Algebra auf $\Omega$, d.h. eine Familie von Teilmenge von $\Omega$, sodass

1. $\Omega \in \mathscr{F}$
2. $A \in \mathscr{F} \Rightarrow A^C \in \mathscr{F}$
3. $A_1, A_2, \dots \in \mathscr{F} \Rightarrow \bigcup_{i \geq 1} \in \mathscr{F}$

Dann heißt $(\Omega, \mathscr{F})$ Ereignisraum bzw. messbarer Raum.

### Wahrscheinlichkeiten

Ordne Ereignissen Wahrscheinlichkeiten zu mittels der Abbildung

$$\mathbb{P} : \mathscr{F} \to [0,1]$$

sodass

Normierung $\mathbb{P}(\Omega) = 1$ (N)

$\sigma$-Additivität für paarweise disjunkte Ereignisse$A_1, A_2, \dots \in \mathscr{F} \Rightarrow \mathbb{P}\left(\bigcup_{i\geq 1} A_i\right) = \sum_{i\geq 1} \mathbb{P}(A_i)$ (A)

(N), (A) und die Nichtnegativität von $\mathbb{P}$ werden als KOLMOGOROVsche Axiome bezeichnet (nach Kolomogorov: Grundbegriffe der Wahrscheinlichkeitstheorie, 1933)

**Definition 1.3 (Wahrscheinlichkeitsmaß, Wahrscheinlichkeitsverteilung)**
Sei $(\Omega, \mathscr{F})$ ein Ereignisraum und $\mathbb{P} : \mathscr{F} \to [0,1]$ eine Abbildung mit Eigenschaften (N) und (A). Dann heißt $\mathbb{P}$ Wahrscheinlichkeitsmaß oder auch Wahrscheinlichkeitsverteilung.

Aus der Definition folgen direkt:

**Satz 1.4 (Rechenregeln für W-Maße)**
Sei $\mathbb{P}$ ein W-Maß, Ereignisse $(\Omega, \mathscr{F}), A, B, A_1, A_2, \dots \in \mathscr{F}$. Dann gelten:

1. $\mathbb{P}(\varnothing) = 0$
2. Monotonie: $A \subseteq B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$
3. endliche $\sigma$-Additivität: $\mathbb{P}(A \cup B) + \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B)$ und insbesondere $\mathbb{P}(A) + \mathbb{P}(A^C) = 1$
4. $\sigma$-Subadditivität:
$$\mathbb{P}\left(\bigcup_{i\geq 1} A_i\right) \leq \sum_{i\geq 1} \mathbb{P}(A_i)$$
5. $\sigma$-Stetigkeit: Wenn $A_n \uparrow A$ (d.h. $A_1 \subseteq A_2 \subseteq \cdots$ und $A = \bigcup_{i=1}^{\infty}(A_i)$) oder $A_n \downarrow A$, so gilt:
$$\mathbb{P}(A_n) \longrightarrow \mathbb{P}(A), n \to \infty$$

*Beweis.* In der Vorlesung wurde auf Schilling MINT Satz 3.3 verwiesen. Ausserdem gab es dazu Präsenzübung 1.3. Der folgende Beweis wurde ergänzt.
Beweise erst Aussage: $A \cap B = \varnothing \Longrightarrow \mathbb{P}(A \uplus B) = \mathbb{P}(A) + \mathbb{P}(B)$.
Es kann $\sigma$-Additivität verwendet werden, indem "fehlende" Mengen durch $\varnothing$ ergänzt werden:

$$\mathbb{P}(A \uplus B) = \mathbb{P}(A \uplus B \uplus \varnothing \uplus \varnothing \dots) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(\varnothing) + \cdots = \mathbb{P}(A) + \mathbb{P}(B),$$

wobei Maßeigenschaften verwendet werden.

1. Definition des Maßes.

2. Da $A \subseteq B$ ist auch $B = A \uplus (B \setminus A) = A \uplus (B \setminus (A \cap B))$. Wende wieder Aussage von oben an, damit folgt

$$\mathbb{P}(B) = \mathbb{P}(A \uplus (B \setminus A)) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A) \tag{*}$$

3. Zerlege $A \cup B$ geschickt, dann sieht man mit oben gezeigter Aussage und (*)

$$\begin{aligned}\mathbb{P}(A \cup B) + \mathbb{P}(A \cap B) &= \mathbb{P}(A \uplus (B \setminus (A \cap B)) + \mathbb{P}(A \cap B) \\ &= \mathbb{P}(A) + \mathbb{P}(B \setminus (A \cap B)) + \mathbb{P}(A \cap B) \\ &= \mathbb{P}(A) + \mathbb{P}(B).\end{aligned}$$

   Im letzten Schritt wurde (*) verwendet.
4. Folgt aus endlicher $\sigma$-Additivität, da $\mathbb{P}\left(\bigcap_{i \geq 1} A_i\right) \geq 0$.
5. Definiere $F_1 := A_1, F_2 := A_2 \setminus A_1, \ldots, F_{i+1} := A_{i+1} \setminus A_n$. Die $F_i$ Mengen sind paarweise disjunkt und damit folgt für $m \to \infty$

$$A_m = \biguplus_{i=1}^{m} F_i \Rightarrow A = \biguplus_{i=1}^{\infty} F_i = \biguplus_{i=1}^{\infty} A_i$$

   und

$$\mathbb{P}(A) = \mathbb{P}\left(\biguplus_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(F_i) = \lim \lim_{m \to \infty} \mathbb{P}\left(\biguplus_{i=1}^{m} F_i\right) = \lim \lim_{m \to \infty} \mathbb{P}(A_m).$$

□

### ■ Beispiel 1.5

Für ein beliebigen Ereignisraum $(\Omega, \mathscr{F})$ $(\Omega \neq \varnothing)$ und eine beliebiges Element $\xi \in \Omega$ definiere

$$\delta_\xi(A) := \begin{cases} 1 & \xi \in A \\ 0 & \text{sonst} \end{cases}$$

eine (degeneriertes) W-Maß auf $(\Omega, \mathscr{F})$, welches wir als Dirac-Maß oder Dirac-Verteilung bezeichnen.

### ■ Beispiel 1.6

Würfeln mit fairem, 6-(gleich)seitigem Würfel mit Ergebnismenge $\Omega = \{1, \ldots, 6\}$ und Ereignisraum $\mathscr{F} = \mathscr{P}(\Omega)$ setzen wir als Symmetriegründen

$$\mathbb{P}(A) = \frac{\#A}{6}.$$

(Wobei $\#A$ oder auch $|A|$ die Kardinalität von $A$ ist.) Das definiert ein W-Maß.

### ■ Beispiel 1.7

Wartezeit an der Bushaltestelle mit Ergebnisraum $\Omega = \mathbb{R}_+$ und Ereignisraum Borelsche $\sigma$-Algebra $\mathscr{B}(\mathbb{R}_+) = \mathscr{F}$. Eine mögliches W-Maß können wir dann durch

$$\mathbb{P}(A) = \int_A \lambda e^{-\lambda x} \, \mathrm{d}x$$

für einen Parameter $\lambda > 0$ festlegen. (Offenbar gilt $\mathbb{P}(\Omega) = 1$ und die $\sigma$-Additivität aufgrund der

Additivität des Integrals.) Wir bezeichnen diese Maße als Exponentialverteilung. (Warum gerade dieses Maß für Wartezeiten gut geeignet ist $\nearrow$ später)

**Satz 1.8 (Konstruktion von Wahrscheinlichkeitsmaßen durch Dichten)**
Sei $(\Omega, \mathscr{F})$ ein Eriegnisraum.

- $\Omega$ abzählbar, $\mathscr{F} = \mathscr{P}(\Omega)$: Sei $\rho = (\rho(\omega))_{\omega\in\Omega}$ eine Folge in $[0,1]$ in $\sum_{\omega\in\Omega}\rho(\omega) = 1$, dann definiert

$$\mathbb{P}(A) = \sum_{\omega\in\Omega} \rho(\omega), A \in \mathscr{F}$$

  ein (diskretes) Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\Omega, \mathscr{F})$. $\rho$ wird als Zähldichte bezeichnet.

- Umgekehrt definiert jedes Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\Omega, \mathscr{F})$ definiert Folge $\rho(\omega) = \mathbb{P}(\{\omega\})$, $\omega \in \Omega$ eine Folge $\rho$ mit den obigen Eigenschaften.

- $\Omega \subset \mathbb{R}^n, \mathscr{F} = \mathscr{B}(\Omega)$: Sei $\rho : \Omega \to [0,\infty)$ eine Funktion, sodass

  1. $\int_\Omega \rho(x)\,\mathrm{d}x = 1$
  2. $\{x \in \Omega\colon f(x) \le c\} \in \mathscr{B}(\Omega)$ für alle $c > 0$

  dann definiert $\rho$ ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\Omega, \mathscr{F})$ durch

$$\mathbb{P}(A) = \int_A \rho(x)\,\mathrm{d}x = \int_A \rho\,\mathrm{d}\lambda, \quad A \in \mathscr{B}(\Omega).$$

  Das Integral interpretieren wir stets als Lebesgue-Integral bzw. Lebesgue-Maß $\lambda$. $\rho$ bezeichnet wir als Dichte, Dichtefunktion/Wahrscheinlichkeitsdichte von $\mathbb{P}$ und nennen ein solches $\mathbb{P}$ (absolut)stetig (bzgl. denn Lebesgue-Maß).

*Beweis.*

- Der diskrete Fall ist klar.
- Im stetigen Fall folgt die Bahuptung aus den bekannten Eigenschaften des Lebesgue-Integrals ($\nearrow$ Schilling MINT, Lemma 8.9) □

**▶ Bemerkung**

- Die eineindeutige Beziehung zwischen Dichte und Wahrscheinlichkeitsmaß überträgt sich nicht auf den stetigen Fall.
  - Nicht jedes Wahrscheinlichkeitsmaß auf $(\Omega, \mathscr{B}(\Omega))$ mit $\Omega \subset \mathbb{R}^n$ besitzt eine Dichte.
  - Zwei Dichtefunktionen definieren dasselbe Wahrscheinlichkeitsmaß, wenn sie sich nur auf einer Menge vom Lebesgue-Maß 0 unterscheiden.
- Jede auf $\Omega \subset \mathbb{R}^n$ definiert Dichtefunktion $\rho$ lässt sich auf ganz $\mathbb{R}^n$ fortsetzen durch $\rho(x) = 0$ mit $x \notin \Omega$. Das erzeugte Wahrscheinlichkeitsmaß auf $(\mathbb{R}^n, \mathscr{B}(\Omega))$ lässt mit dem Wahrscheinlichkeitsmaß auf $(\Omega, \mathscr{B}(\Omega)$ identifizieren.
- Mittels Dirac-Maß $\delta_x$ können auch jedes diskrete Wahrscheinlichkeitsmaß auf $\Omega \subset \mathbb{R}^n$ als

Wahrscheinlichkeitsmaß auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ intepretieren

$$\mathbb{P}(A) = \sum_{\omega \in A} \rho(\omega) = \int_A \mathrm{d}\left(\sum_{\omega \in \Omega} \rho(\omega)\delta_\omega\right)$$

stetige und diskrete Wahrscheinlichkeitsmaße lassen sich kombinieren z.B.

$$\mathbb{P}(A) = \frac{1}{2}\delta_0 + \frac{1}{2}\int_A \mathbb{1}_{[0,1]}(x)\,\mathrm{d}x, A \in \mathscr{B}(\mathbb{R})$$

ist ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

Abschließend erinnern wir uns an:

**Satz 1.9 (Eindeutigkeitssatz für Wahrscheinlichkeitsmaße)**
Sei $(\Omega, \mathscr{F})$ Ereignisraum und $\mathbb{P}$ ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathscr{F})$. Sei $\mathscr{F} = \omega(\mathscr{G})$ für ein $\cap$-stabiles Erzeugendensystem $\mathscr{G} \subset \mathscr{P}(\Omega)$. Dann ist $\mathbb{P}$ bereits durch seine Einschränkung $\mathbb{P}_{|\mathscr{G}}$ eindeutig bestimmt.

*Beweis.* ↗ Schilling MINT, Satz 4.5. □

Insbesondere definiert z.B.

$$\mathbb{P}([0,a]) = \int_0^a \lambda e^{-\lambda x}\,\mathrm{d}x = 1 - e^{-\lambda a}, a > 0$$

bereits die Exponentialverteilung aus Beispiel 1.7.

**Definition 1.10 (Gleichverteilung)**
Ist $\Omega$ endlich, so heißt das Wahrscheinlichkeitsmaß mit konstanter Zähldichte $\rho(\omega) = {}^1/_{|\Omega|}$ die (diskrete) Gleichverteilung auf $\Omega$ und wird mit $\mathrm{U}(\Omega)$ notiert (U = Uniform). Ist $\Omega \subset \mathbb{R}^n$ eine Borelmenge mit Lebesgue-Maß $0 < \lambda^n(\Omega) < \infty$, so heißt das Wahrscheinlichkeitsmaß auf $(\Omega, \mathscr{B}(\Omega))$ mit konstanter Dichtefunktion $\rho(x) = {}^1/_{\lambda^n(x)}$, die (stetige) Gleichverteilung auf $\Omega$. Sie wird ebenso mit $\mathrm{U}(\Omega)$ notiert.

### Wahrscheinlichkeitsräume

**Definition 1.11 (Wahrscheinlichkeitsraum)**
Ein Tripel $(\Omega, \mathscr{F}, \mathbb{P})$ mit $\Omega, \mathscr{F}$ Ereignisraum und $\mathbb{P}$ Wahrscheinlichkeitsmaß auf $(\Omega, \mathscr{F})$, nennen wir Wahrscheinlichkeitsraum.

## 2. Zufallsvariablen

Zufallsvariablen dienen dazu von einen gegebenen Ereignisraum $(\Omega, \mathscr{F})$ zu einem Modellausschnitt $\Omega', \mathscr{F}'$ überzugehen. Es handelt sich also um Abbildungen $X : \Omega \to \Omega'$. Damit wir auch jedem Ereignis in $\mathscr{F}'$ eine Wahrscheinlichkeit zuordnen können, benötigen wir

$$A' \in \mathscr{F}' \Rightarrow X' A' \in \mathscr{F}$$

d.h. $X$ sollte messbar sein.

**Definition 2.1 (Zufallsvariable)**
Seien $(\Omega, \mathscr{F})$ und $(\Omega', \mathscr{F}')$ Ereignisräume. Dann heißt jede messbare Abbildung

$$X : \Omega \to \Omega'$$

Zufallsvariable (von $(\Omega, \mathscr{F})$) nach $(\Omega', \mathscr{F}')$/ auf $(\Omega', \mathscr{F}')$ oder Zufallselement.

■ **Beispiel 2.2**

1. Ist $\Omega$ abzählbar und $\mathscr{F} = \mathcal{P}(\Omega)$, so ist jede Abbildung $X : \Omega \to \Omega'$ messbar und damit eine Zufallsvariable.
2. Ist $\Omega \subset \mathbb{R}^n$ und $\mathscr{F} = \mathscr{B}(\Omega)$, so ist jede stetige Funktion $X : \Omega \to \mathbb{R}$ messbar und damit eine Zufallsvariable.

**Satz 2.3**
Sei $(\Omega, \mathscr{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $X$ eine Zufallsvariable von $(\Omega, \mathscr{F})$ nach $(\Omega', \mathscr{F}')$. Dann definiert

$$\mathbb{P}'(A') := \mathbb{P}\left(X^{-1}(A')\right) = \mathbb{P}\left(\{X \in A'\}\right), \quad A' \in \mathscr{F}'$$

ein Wahrscheinlichkeitsmaß auf $(\Omega', \mathscr{F}')$ auf $(\Omega', \mathscr{F}')$, welches wir als Wahrscheinlichkeitsverteilung von X unter $\mathbb{P}$ bezeichnet.

*Beweis.* Aufgrund der Messbarkeit von $X$ ist die Definition sinnvoll. Zudem gelten

$$\mathbb{P}'(\Omega') = \mathbb{P}(X^{-1}(\Omega')) = \mathbb{P}(\Omega) = 1$$

und für $A'_1, A'_2, \dots \in \mathscr{F}'$ paarweise disjunkt.

$$\begin{aligned}\mathbb{P}'\left(\bigcup_{i\geq 1} A'_i\right) &= \mathbb{P}\left(X^{-1}\left(\bigcup_{i\geq 1} A'_i\right)\right) \\ &= \mathbb{P}\left(\bigcup_{i\geq 1} X^{-1}(A'_i)\right) \\ &= \sum_{i\geq 1} \mathbb{P}(X^{-1}A'_i)\end{aligned}$$

da auch $X^{-1}A'_1, X^{-1}A'_2, \dots$ paarweise disjunkt

$$\mathbb{P}'\left(\bigcup_{i\geq 1} A'_i\right) = \sum_{i\geq 1} \mathbb{P}'(A'_i).$$

Also ist $\mathbb{P}'$ ein Wahrscheinlichkeitsmaß. □

▶ **Bemerkung**

- Aus Gründen der Lesbarkeit schreiben wir in der Folge $\mathbb{P}(X \in A) = \mathbb{P}(\{\omega\colon X(\omega) \in A\})$
- Ist $X$ die Identität, so fallen die Begriffe Wahrscheinlichkeitsmaß und Wahrscheinlichkeitsver-

teilung zusammen.

- In der (weiterführenden) Literatur zu Wahrscheinlichkeitstheorie wird oft auf die Angabe eines zugrundeliegenden Wahrscheinlichkeitsraumes verzichtet und stattdessen eine “Zufallsvariable mit Verteilung $\mathbb{P}$ auf $\Omega$” eingeführt. Gemeint ist (fast) immer $X$ als Identität auf $(\Omega, \mathscr{F}, \mathbb{P})$ mit $\mathscr{F} = \mathcal{P}(\Omega)/\mathscr{B}(\Omega)$.
- Für die Verteilung von $X$ unter $\mathbb{P}$ schreibe $\mathbb{P}_X$ und $X \sim \mathbb{P}_X$ für die Tatsache, dass $X$ gemäß $\mathbb{P}_X$ verteilt ist.

**Definition 2.4 (identisch verteilt, reellen Zufallsvariablen)**
Zwei Zufallsvariablen sind identisch verteilt, wenn sie dieselbe Verteilung haben. Von besonderen Interesse sind für uns die Zufallsvariablen, die nach $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ abbilden, sogenannte reelle Zufallsvariablen.

Da die halboffenen Intervalle $\mathscr{B}(\mathbb{R})$ erzeugen, ist die Verteilung eine reelle Zufallsvariable durch die Werte $(-\infty, c], c \in \mathbb{R}$ eindeutig festgelegt.

**Definition 2.5 ((kumulative) Verteilungsfunktion von $\mathbb{P}$)**
Sei $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \mathbb{P})$ Wahrscheinlichkeitsraum, so heißt

$$F : \mathbb{R} \to [0,1] \text{ mit } x \mapsto \mathbb{P}((-\infty, x])$$

(kumulative) Verteilungsfunktion von $\mathbb{P}$.
Ist $X$ eine reelle Zufallsvariable auf beliebigen Wahrscheinlichkeitsraum $(\Omega, \mathscr{F}, \mathbb{P})$, so heißt

$$F : \mathbb{R} \to [0,1] \text{ mit } x \mapsto \mathbb{P}(X \leq x) = \mathbb{P}(X \in (-\infty, x])$$

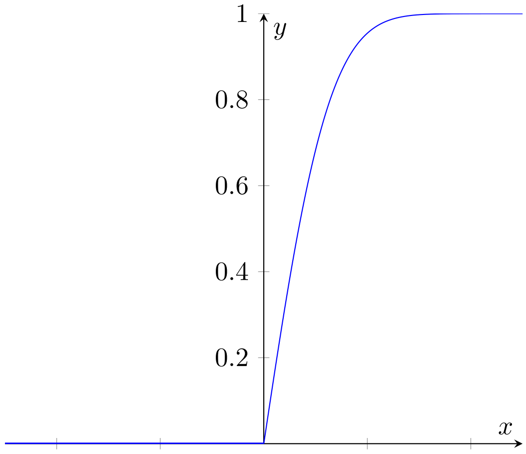
die (kumulative) Verteilungsfunktion von $X$.

■ **Beispiel 2.6**
Sei $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \mathbb{P})$ mit $\mathbb{P}$ Exponentialverteilung mit Parameter $\lambda > 0$

$$\mathbb{P}(A) = \int_{A \cap [0,\infty)} \lambda e^{-\lambda x} \,\mathrm{d}x \quad A \in \mathscr{B}(\mathbb{R}).$$

Dann ist

$$F(x) = \mathbb{P}((-\infty, x)) = \begin{cases} 0 & x \leq 0 \\ \int_0^x \lambda e^{-\lambda y} \,\mathrm{d}y = 1 - e^{-\lambda x} & x > 0 \end{cases}.$$



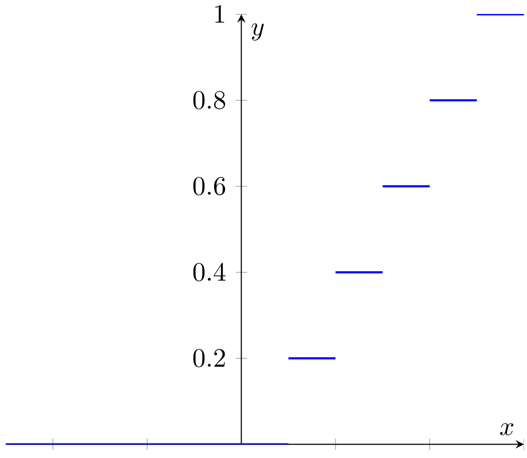

■ **Beispiel 2.7**

Das Würfeln mit einem fairen, sechsseitigen Würfel kann mittels einer reellen Zufallsvariablen

$$X : \{1, 2, \ldots, 6\} \to \mathbb{R} \text{ mit } x \mapsto x$$

modelliert werden. Es folgt als Verteilungsfunktion

$$\begin{aligned} F(x) &= \mathbb{P}'(X \leq x) = \mathbb{P}(X^{-1}(-\infty, x]) = \mathbb{P}((-\infty, x]) \\ &= \frac{1}{6} \sum_{i=1}^{6} \mathbb{1}_{i \leq x}. \end{aligned}$$




Allgemein:

**Satz 2.8**
Ist $\mathbb{P}$ ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ und $F$ die zugehörige Verteilungsfunktion, so gelten

1. $F$ ist monoton wachsend
2. $F$ ist rechtsseitig stetig
3. $\lim\limits_{x \to -\infty} F(x) = 0, \ \lim\limits_{x \to \infty} F(x) = 1$

Umgekehrt existiert zu jeder Funktion $F : \mathbb{R} \to [0,1]$ mit Eigenschaften 1-3 eine reelle Zufallsvariable auf $((0,1), \mathscr{B}((0,1)), \mathrm{U}((0,1))$ mit Verteilungsfunktion $F$.

*Beweis.* Ist $F$ Verteilungsfunktion, so folgt mit Satz 1.4

$$x \leq y \Rightarrow F(x) = \mathbb{P}((-\infty, x]) \overset{1.4.3}{\leq} \mathbb{P}((-\infty, y]) = F(y)$$

und

$$\lim_{x \searrow c} F(x) = \lim_{x \searrow c} \mathbb{P}((-\infty, x]) \overset{\sigma\text{-Stetigkeit}}{=} \mathbb{P}((-\infty, c]) = F(c)$$

sowie

$$\lim_{x \to -\infty} F(x) \overset{1.4.5}{=} \mathbb{P}(\varnothing) \overset{1.4.1}{=} 0$$
$$\lim_{x \to \infty} F(x) \overset{1.4.5}{=} \mathbb{P}(\mathbb{R}) = 1.$$

Umgekehrt wähle

$$X(u) := \inf\{x \in \mathbb{R} \colon F(x) \geq u\}, \quad u \in (0,1)$$

Dann ist $X$ eine "linksseitige Inverse" von $F$ (auch Quantilfunktion / verallgemeinerte Inverse). Wegen 3 gilt:

$$-\infty < X(u) < \infty$$

und zudem

$$\{X \leq x\} = (0, F(x)) \cap (0,1) \in \mathscr{B}((0,1)).$$

Da diese halboffene Mengen ein Erzeugendensystem von $\mathscr{B}(\mathbb{R})$ bilden, folgt bereits die Messbarkeit von $X$, also ist $X$ eine ZV. Insbesondere hat die Menge $\{X \leq x\}$ gerade Lebesgue-Maß $F(x)$ und damit hat $X$ die Verteilungsfunktion $F$. □

**Folgerung 2.9**
Ist $\mathbb{P}$ Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ und $F$ die zugehörige Verteilungsfunktion. Dann besitzt $\mathbb{P}$ genau eine Dichtefunktion $\rho$, wenn $F$ stetig differenzierbar ist, denn dann gelten

$$F(x) = \int_{-\infty}^{x} \rho(x) \,\mathrm{d}x, \ \text{bzw.} \ \rho(x) = F'(x)$$

*Beweis.* Folgt aus Satz 1.8, der Definition 2.5 der Verteilungsfunktion und dem Eindeutigkeitssatz Satz 1.9. □

## Kapitel II

# *Erste Standardmodelle der Wahrscheinlichkeitstheorie*

## Diskrete Verteilungen

### 1. Diskrete Gleichverteilungen

Erinnerung:

▶ **Erinnerung (Definition I.1.10)**
Ist $\Omega$ endlich, so heißt Wahrscheinlichkeitsmaß mit Zähldichte

$$\rho(\omega) = \frac{1}{\omega} \quad , \omega \in \Omega$$

<u>(diskrete) Gleichverteilung</u> auf $\Omega \to U(\Omega)$

Es gilt das für jedes $A \in \mathcal{P}(\Omega)$

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}$$

Anwendungsbeispiele sind faires Würfeln, fairer Münzwurf, Zahlenlotto, ...

### 2. Urnenmodelle

Ein "Urnenmodell" ist eine abstrakte Darstellung von Zufallsexperimenten, bei denen zufällig Stichproben aus einer gegebenen Menge "gezogen" werden.

**Definition (Urne)**
Eine Urne ist ein Behältnis in welchem sich farbige/nummerierte Kugeln befinden, die ansonsten ununterscheidbar sind.

Aus der Urne ziehe man blind/zufällig eine oder mehrere Kugeln und notiere Farbe/Zahl.

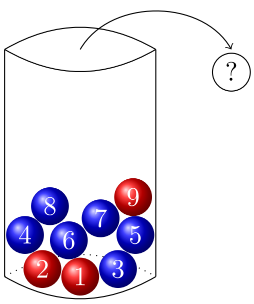


**Abbildung II.1:** Urnenmodell

## 2.1. Urnenmodell mit Zurücklegen: Multinomial-Verteilung

Gegeben: Urne mit $N$ Kugeln, verschiedenfarbig mit Farben aus $E$, $|E| \geq 2$

Ziehe: $n$ Stichproben/Kugeln, wobei nach jedem Zug die Kugel wieder zurückgelegt wird. Uns interessiert die Farbe in jedem Zug, setze also

$$\Omega = E^n \text{ und } \mathscr{F} = \mathcal{P}(\Omega)$$

Zur Bestimmung eines geeigneten Wahrscheinlichkeitsmaßes, nummerieren wir die Kugeln mit $1, \ldots, N$, so dass alle Kugeln der Farbe $a \in E$ eine Nummer aus $F_a \subset \{1, \ldots, N\}$ tragen. Würden wir die Nummern notieren, so wäre

$$\bar{\Omega} = \{1, \ldots, N\}^n \text{ und } \overline{\mathscr{F}} = \mathcal{P}(\overline{\Omega})$$

und wir könnten die Gleichverteilung $\overline{\mathbb{P}} = \mathrm{U}(\overline{\Omega})$ als Wahrscheinlichkeitsmaß für einem einzelnen Zug verwenden. Für den Übergang zu $\Omega$ konstruieren wir Zufallsvariablen. Die Farbe im $i$-ten Zug wird beschrieben durch

$$X_i : \overline{\Omega} \to E \text{ mit } \overline{\omega} = (\overline{\omega}_1, \ldots, \overline{\omega}_n) \mapsto a \text{ falls } \overline{\omega}_i \in F_a.$$

Der Zufallsvektor

$$X = (X_1, \ldots, X_n) : \overline{\Omega} \to \Omega$$

beschreibt dann die Abfolge der Farben. Für jedes $\omega \in \Omega$ gilt dann

$$\{X = \omega\} = F_{\omega_1} \times \cdots \times F_{\omega_n} = \bigtimes_{i=1}^{n} F_{\omega_i}$$

und damit

$$\begin{aligned} \mathbb{P}(\{\omega\}) &= \overline{\mathbb{P}}(X^{-1}(\{\omega\})) = \mathbb{P}(X = \omega) \\ &= \frac{|F_{\omega_1}| \cdots |F_{\omega_n}|}{\left|\overline{\Omega}\right|} \\ &= \prod_{i=1}^{n} \frac{|F_{\omega_i}|}{N} =: \prod_{i=1}^{n} \rho(\omega_i) \end{aligned}$$

Zähldichten, die sich als Produkte von Zähldichten schreiben lassen, werden auch als Produktdichten bezeichnet (↗ **??**). Sehr oft interessiert bei einem Urnenexperiment nicht die Reihenfolge der gezogenen Farben, sondern nur die Anzahl der Kugeln in Farbe $a \in E$ nach $n$ Zügen. Dies entspricht

$$\hat{\Omega} = \left\{ k = (k_a)_{a \in E} \in \mathbb{N}_0^{|E|} : \sum_{a \in E} k_a = n \right\} \text{ und } \hat{\mathscr{F}} = \mathcal{P}\left(\hat{\Omega}\right)$$

Den Übergang $\Omega \to \hat{\Omega}$ beschreiben wir durch die Zufallsvariablen

$$Y_a(\omega) : \Omega \to \mathbb{N}_0 \text{ mit } \omega = (\omega_1, \ldots, \omega_n) \mapsto \sum_{a \in E} \mathbb{1}_{\{a\}}(\omega_i)$$

und

$$Y = (Y_a)_{a \in E} : \Omega \to \hat{\Omega} = \left\{ k = (k_a)_{a \in E} : \sum_{a \in E} k_a = n \right\}$$

Wir erhalten

$$\begin{aligned} \mathbb{P}(Y = k) &= \mathbb{P}(Y_a = k_a, \ \ a \in E) \\ &= \sum_{\omega \in \Omega : Y(\omega) = k} \prod_{i=1}^{n} \rho(\omega_i) \\ &= \sum_{\omega \in \Omega : Y(\omega) = k} \prod_{a \in E} \rho(a) \\ &= \binom{n}{(k_a)_{a \in E}} \prod_{a \in E} \rho(a)^{k_a}, \end{aligned}$$

wobei

$$\binom{n}{(k_1, \ldots, k_l)} = \begin{cases} \frac{n!}{k_1! \, k_2! \cdots k_l!} & \sum_{i=1}^{l} k_i = n \\ 0 & \text{sonst} \end{cases}$$

der Multinomialkoeffizient ist, welcher die Anzahl der Möglichkeiten beschreibt, $n$ Objekte in $l$ Gruppen aufzuteilen, so dass Gruppe $i$ gerade $k_i$ Objekte beinhaltet.

**Definition 2.1**
Sei $l > 2, p = (p_1, \ldots, p_l)$ eine Zähldichte und $n \in \mathbb{N}$, dann heißt die Verteilung auf $\left\{ k = (k_i)_{i=1,\ldots,l} \in \mathbb{N}_0^l : \sum_{i=1}^{l} k_i = n \right\}$ mit Zähldichte

$$m((k_1, \ldots, k_l)) = \binom{n}{k_1, \ldots, k_l} \prod_{i=1}^{l} p_i^{k_i}$$

Multinomialverteilung mit Parametern $n$ und $p$. Wir schreiben auch $\text{Multi}(n, p)$.

■ **Beispiel 2.2**
Eine Urne enthalte nur schwarze "1" und weiße "0" Kugeln, d.h. $E = \{0, 1\}$, und es sei $\rho(1) = p$ gerade die Proportion der schwarzen Kugeln (= Wahrscheinlichkeit bei einem Zug schwarz zu ziehen), dann ist Wahrscheinlichkeit in $n$ Zügen $k$-mal schwarz zu ziehen:

$$\binom{n}{k} \prod_{i=0,1} \rho(i)^{k_i} = \binom{n}{k} p^k (1-p)^{n-k}.$$

Ein solches (wiederholtes) Experiment mit nur zwei möglichen Ereignissen und fester Wahrscheinlichkeit $p \in [0, 1]$ für eines der Ergebnisse nennen wir auch (wiederholtes) Bernoulliexperiment.

**Definition 2.3**
Sei $p \in [0,1]$ und $n \in \mathbb{N}$, dann heißt die Verteilung mit Zähldichte

$$\rho(k) = \binom{n}{k} p^k (1-p)^{n-k} \text{ mit } k \in \{0, 1, \dots, n\}.$$

Binomialverteilung auf $\{0, \dots, n\}$ mit Parameter $p$ (auch Erfolgswahrscheinlichkeit). Wir schreiben auch $\mathrm{Bin}(n,p)$. Im Fall $n = 1$ nennen wir die Verteilung mit Zähldichte

$$\rho(0) = 1 - p \text{ und } \rho(1) = p$$

auch Bernoulliverteilung mit Parameter $p$ und schreiben $\mathrm{Bernoulli}(p)$.

Urnenmodell ohne Zurücklegen: Hypergeometrische Verteilung
Gegeben: Urne mit $N$ Kugeln verschiedener Farben aus $E$,

$$|E| \geq 2.$$

Es werden $n \leq N$ Stichproben entnommen, wobei die gezogenen Kugeln nicht in die Urne zurückgelegt werden.

## 2.2. Urnenmodell ohne Zurücklegen: Hypergeometrische Verteilung

Gegeben: Urne mit $N$ Kugeln verschiedener Farben aus $E$, $|E| \geq 2$. Es werden $n \leq N$ Stichproben entnommen, wobei die gezogenen Kugeln nicht in die Urne zurückgelegt werden.

■ **Beispiel 2.4**
Eine Urne enthalte $S$ schwarze "1" und $W$ weiße Kugeln "0" Kugeln, $(E = \{0,1\}, S + W = N)$. Dann ist die Wahrscheinlichkeit in $n$ Zügen ohne Zurücklegen gerade $s$ schwarze und $w$ weiße Kugeln zu ziehen

$$\rho(w) = \frac{\binom{W}{w}\binom{S}{s}}{\binom{N}{n}}, \quad 0 \leq s \leq S, 0 \leq w \leq W, s + w = n, S + W = N.$$

*Beweis.* Hausaufgabe 2.3! □

**Definition 2.5**
Seien $N \in \mathbb{N}, W \leq N, n \leq N$, dann heißt die Verteilung auf $\{0, \dots, n\}$ mit Zähldichte

$$\rho(w) = \frac{\binom{W}{w}\binom{N-W}{n-w}}{\binom{N}{n}}, \quad w = \max\{0, n = N + W\}, \dots, \min\{W, n\},$$

die Hypergeometrische Verteilung mit Parametern $N, W, n$. Wir schreiben $\mathrm{Hyper}(N, W, n)$.

# 3. Poisson-Approximation und Poisson-Verteilung

$\mathrm{Bin}(n,p)$ ist zwar explizit und elementar definiert, jedoch für große $n$ mühsam auszuwerten. Für seltene Ereignisse ($n$ groß, $p$ klein) verwende daher:

**Satz 3.1 (Poisson-Approximation)**
Sei $\lambda > 0$ und $(p_n)_{n\in\mathbb{N}}$ eine Folge in $[0,1]$ mit

$$np_n \to \lambda, \quad n \to \infty.$$

Dann gilt $\forall k \in \mathbb{N}_0$ für die Zähldichte der $\mathrm{Bin}(n, p_n)$-Verteilung

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1-p)^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}.$$

*Beweis.* Sei $k \in \mathbb{N}_0$ fix, dann

$$\begin{aligned}\binom{n}{k} = \frac{n!}{k!(n-k)!} &= \frac{n^k}{k!}\frac{n(n-1)\cdots(n-k+1)}{n^k}\\ &= \frac{n^k}{k!}\cdot 1\cdot(1-\frac{1}{n}\cdots\frac{k-1}{n})\\ &\overset{n\to\infty}{\sim} \frac{n^k}{k!},\end{aligned}$$

wobei $a(l) \overset{n\to\infty}{\sim} b(l) \Leftrightarrow \frac{a(l)}{b(l)} \xrightarrow{n\to\infty} 1$. Damit

$$\begin{aligned}\binom{n}{k} p^k(1-p)^{n-k} &\overset{n\to\infty}{\sim} \frac{n^k}{k!}p_n^k(1-p_n)^{n-k}\\ &\overset{n\to\infty}{\sim} \frac{\lambda^k}{k!}(1-p_n)^n\\ &= \frac{\lambda^n}{k!}\left(1-\frac{np_n}{n}\right)^n\\ &\xrightarrow{n\to\infty} \frac{\lambda^n}{k!}e^{-\lambda}.\end{aligned}$$

□

Der erhaltene Grenzwert liefert die Zähldichte auf $\mathbb{N}_0$, denn

$$\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1$$

**Definition 3.2**
Sei $\lambda > 0$. Dann heißt das auf $(\mathbb{N}_0, \mathbb{P}(\mathbb{N}_0))$ definierte Wahrscheinlichkeitsmaß mit

$$\mathbb{P}(\{k\}) = \frac{\lambda^k}{k!}e^{-\lambda} \quad k \in \mathbb{N}_0,$$

Poissonverteilung mit Parameter $\lambda$. Schreibe $\mathrm{Poisson}(\lambda)$.

Die Poissonverteilung ist ein natürliches Modell für die Anzahl von zufälligen, seltenen Ereignissen (z.B. Tore im Fußballspiel, Schadensfälle einer Versicherung, ...).

## Kapitel III

# *Bedingte Wahrscheinlichkeiten und (Un)-abbhängigk*

## 1. Bedingte Wahrscheinlichkeiten

■ **Beispiel 1.1**

Das Würfeln mit zwei fairen, sechsseitigen Würfeln können wir mit

$$\Omega = \{(i,j) \colon i,j \in \{1,\dots,6\}\}$$

und $\mathbb{P} = \mathrm{U}(\Omega)$. Da $|\Omega| = 36$ gilt also

$$\mathbb{P}(\{\omega\}) = \frac{1}{36} \quad \forall \omega \in \Omega.$$

Betrachte das Ereignis

$$A = \{(i,j) \in \Omega \colon i + j = 8\},$$

dann folgt

$$\mathbb{P}(A) = \frac{5}{36}.$$

Werden die beiden Würfe nacheinander ausgeführt, so kann nach dem ersten Wurf eine Neubewertung der Wahrscheinlichkeit von $A$ erfolgen.
Ist z.B.

$$B = \{(i,j) \in \Omega, i = 4\}$$

eingetreten, so kann die Summe 8 nur durch eine weitere 4 realisiert werden, also mit Wahrscheinlichkeit

$$\frac{1}{6} = \frac{|A \cap B|}{|B|}.$$

Das Eintreten von $B$ führt also dazu, dass das Wahrscheinlichkeitsmaß $\mathbb{P}$ durch ein neues Wahrscheinlichkeitsmaß $\mathbb{P}_B$ ersetzt werden muss. Hierbei sollte gelten:

Renormierung: $\mathbb{P}_B = 1$ (R)

Proportionalität: Für alle$A \subseteq \mathscr{F}$ mit $A \subseteq B$ gilt $\mathbb{P}_B(A) = c_B \mathbb{P}(A)$ mit einer Konstante $c_B$. (P)

**Lemma 1.2**
Sei $(\Omega, \mathscr{F}, \mathbb{P})$ Wahrscheinlichkeitsraum und $B \in \mathscr{F}$ mit $\mathbb{P}(B) > 0$. Dann gibt es genau ein Wahrscheinlichkeitsmaß $\mathbb{P}_B$ auf $(\Omega, \mathscr{F})$ mit den Eigenschaften (R) und (P). Dieses ist gegeben durch

$$\mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad \forall A \in \mathscr{F}.$$

*Beweis.* Offenbar erfüllt $\mathbb{P}_B$ wie definiert (R) und (P). Umgekehrt erfüllt $\mathbb{P}_B$ (R) und (P). Dann folgt für $A \in \mathscr{F}$:

$$\mathbb{P}_B(A) = \mathbb{P}_B(A \cap B) + \underbrace{\mathbb{P}_B(A \setminus B)}_{=0,\text{ wegen (R)}} \stackrel{(P)}{=} c_B \mathbb{P}(A \cap B).$$

Für $A = B$ folgt zudem aus (R)

$$1 = \mathbb{P}_B(B) = c_B \mathbb{P}(B)$$

also $c_B = \mathbb{P}(B)^{-1}$. □

**Definition 1.3**
Sei $(\Omega, \mathscr{F}, \mathbb{P})$ Wahrscheinlichkeitsraum und $B \in \mathscr{F}$ mit $\mathbb{P}(B) > 0$. Dann heißt

$$\mathbb{P}(A \mid B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \text{ mit } A \in \mathscr{F}$$

die <u>bedingte Wahrscheinlichkeit von $A$ gegeben $B$</u>. Falls $\mathbb{P}(B) = 0$, setze

$$\mathbb{P}(A \mid B) = 0 \qquad \forall A \in \mathscr{F}$$

■ **Beispiel 1.4**
In der Situation Beispiel 1.1 gilt

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

und damit

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6}$$

Aus Definition 1.3 ergibt sich

**Lemma 1.5 (Multiplikationsformel)**
Sei $(\Omega, \mathscr{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $A_1, \dots, A_n \in \mathscr{F}$. Dann gilt

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2 \mid A_1) \dots \mathbb{P}(A_n \mid A_1 \cap \dots \cap A_{n-1})$$

*Beweis.* Ist $\mathbb{P}(A_1 \cap \dots \cap A_n) = 0$, so gilt auch $\mathbb{P}(A_n \mid \bigcap_{i=1}^{n-1} A_i) = 0$. Andernfalls sind alle Faktoren der rechten

Seite ungleich Null und

$$\begin{aligned}
&\mathbb{P}(A_1)\mathbb{P}(A_2 \mid A_1)\ldots\mathbb{P}(A_n \mid \bigcap_{i=1}^{n-1} A_i) \\
=&\mathbb{P}(A_1)\cdot\frac{\mathbb{P}(A_1\cap A_2)}{\mathbb{P}(A_1)}\cdots\frac{\mathbb{P}(\bigcap_{i=1}^{n} A_i)}{\mathbb{P}(\bigcap_{i=1}^{n-1} A_i)} \\
=&\mathbb{P}(\bigcap_{i=1}^{n} A_i)
\end{aligned}$$

□

Stehen die $A_i$ in Lemma 1.5 in einer (zeitlichen) Abfolge, so liefert Formel einen Hinweis wie Wahrscheinlichkeitsmaße für Stufenexperimente konstruiert werden können. Ein Stufenexperiment aus $n$ nacheinander ausgeführten Teilexperimenten lässt sich als Baumdiagramm darstellen.

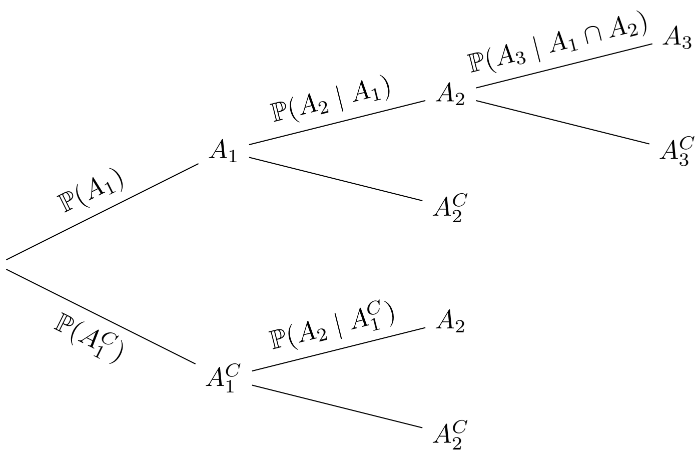



**Abbildung III.1:** Lemma 1.5

**Satz 1.6 (Konstruktion des Wahrscheinlichkeitsmaßes eines Stufenexperiments)** Gegeben seinen $n$ Ergebnisräume $\Omega_i = \{\omega_i(1),\ldots,\omega_i(k)\}, k\in\mathbb{N}\cup\{\infty\}$ und es sei $\Omega = \bigtimes_{i=1}^n \Omega_i$ der zugehörige Produktraum. Weiter seinen $\mathscr{F}_i$ $\sigma$-Algebren auf $\Omega_i$ und $\mathscr{F} = \bigotimes_{i=1}^n \mathscr{F}_i$ die Produkt-$\sigma$-Algebra auf $\Omega$. Setze $\omega = (\omega_1,\ldots,\omega_n)$ und

$$[\omega_1,\ldots,\omega_m] := \{\omega_1\}\times\cdots\times\{\omega_m\}\times\Omega_{m+1}\times\cdots\times\Omega_n,\quad m\leq n$$
$$\mathbb{P}(\{\omega_m\}[\omega_1,\ldots,\omega_{m-1}])$$

für die Wahrscheinlichkeit in der $m$-ten Stufe des Experiments $\omega_m$ zu beobachten, falls in den vorausgehenden Stufen $\omega_1,\ldots,\omega_{m-1}$ beobachten wurden. Dann definiert

$$\mathbb{P}(\{\omega\}) := \mathbb{P}(\{\omega_1\})\prod_{m=2}^{n}\mathbb{P}(\{\omega_m\}\mid[\omega_1,\ldots,\omega_{m-1}])$$

ein Wahrscheinlichkeitsmaß auf $(\Omega,\mathscr{F},\mathbb{P})$.

*Beweis.* Nachrechnen! □

■ **Beispiel 1.7 (Polya-Urne)**

Gegeben sei eine Urne mit $s$ schwarzen und $w$ weißen Kugeln. Bei jedem Zug wird die gezogene Kugel zusammen mit $c \in \mathbb{N}_0 \cup \{-1\}$ weiteren Kugeln derselben Farbe zurückgelegt.

- $c = 0$: Urnenmodell mit Zurücklegen
- $c = -1$: Urnenmodell ohne Zurücklegen

Beide haben wir schon in Kapitel 2.2 gesehen.
Sei deshalb $c \in \mathbb{N}$. (Modell für zwei konkurrierende Populationen) Ziehen wir $n$-mal, so haben wir ein $n$-Stufenexperiment mit

$$\Omega = \{0,1\}^n \text{ mit } 0 = \text{“weiß”}, 1 = \text{“schwarz”} \quad (\Omega_i = \{0,1\})$$

Zudem gelten im ersten Schritt

$$\mathbb{P}(\{0\}) = \frac{w}{s+w} \text{ und } \mathbb{P}(\{1\}) = \frac{s}{s+w}$$

sowie

$$\mathbb{P}(\{\omega_m\} \mid [\omega_1, \dots \omega_{m-1}]) = \begin{cases} \frac{w+c\left(m-1-\sum_{i=1}^{m-1}\omega_i\right)}{s+w+c(m-1)} & \omega_m = 0 \\ \frac{s+c\sum_{i=1}^{m-1}\omega_i}{s+w+c(m-1)} & \omega_m = 1 \end{cases}$$

Mit Satz 1.6 folgt als Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{P}(\Omega))$

$$\begin{aligned} \mathbb{P}(\{(\omega_1, \dots, \omega_n)\}) &= \mathbb{P}(\{\omega_1\}) \prod_{m=2}^{n} \mathbb{P}(\{\omega_m\} \mid [\omega_1, \dots, \omega_{m-1}]) \\ &= \frac{\prod_{i=0}^{l-1}(s + c\cdot i)\prod_{i=0}^{n-l-1}(w + c\cdot j)}{\prod_{i=0}^{n}(s + w + c\cdot i)} \text{ mit } l = \sum_{i=1}^{n}\omega_i. \end{aligned}$$

Definiere wir nun die Zufallsvariable

$$S_n : \Omega \to \mathbb{N}_0 \text{ mit } (\omega_1, \dots, \omega_n) \mapsto \sum_{i=1}^{n}\omega_i$$

welche die Anzahl der gezogenen schwarzen Kugeln modelliert, so folgt

$$\mathbb{P}(S_n = l) = \binom{n}{l}\frac{\prod_{i=0}^{l-1}(s + c\cdot i)\prod_{j=0}^{n-l-1}(w + c\cdot j)}{\prod_{i=0}^{n}(s + w + c\cdot i)}$$

Mittels $a := {}^{s}/_{c}, b := {}^{w}/_{c}$ folgt

$$\mathbb{P}(S_n = l) = \binom{n}{l}\frac{\prod_{i=0}^{l-1}(-a - i)\prod_{j=0}^{n-l-1}(-b - j)}{\prod_{i=0}^{n}(-a - b - i)} = \frac{\binom{-a}{l}\binom{-b}{n\cdot l}}{\binom{-a-b}{n}}$$
$$\text{mit } l \in \{0, \dots, n\}$$

Dies ist die <u>Polya-Verteilung</u> auf $\{0, \dots, n\}, n \in \mathbb{N}$ mit Parametern $a, b > 0$.

■ **Beispiel 1.8**

Ein Student beantwortet eine Multiple-Choice-Frage mit 4 Antwortmöglichkeiten, eine davon ist richtig. Er kennt die richtige Antwort mit Wahrscheinlichkeit $2/3$. Wenn er diese kennt, so wählt er diese aus. Andernfalls wählt er zufällig (gleichverteilt) eine Antwort. Betrachte

$$\begin{aligned} W &= \{\text{richtige Antwort gewusst}\} \\ R &= \{\text{Richtige Antwort gewählt}\} \end{aligned}$$

Dann gilt

$$\mathbb{P}(W) = \frac{2}{3}, \mathbb{P}(R \mid W) = 1, \mathbb{P}(R \mid W^C) = \frac{1}{4}$$

Angenommen, der Student gibt die richtige Antwort. Mit welcher Wahrscheinlichkeit hat er diese gewusst? $\longrightarrow \mathbb{P}(W \mid R) = \ ?$

**Satz 1.9**

Sei $(\Omega, \mathscr{F}, \mathbb{P})$ Wahrscheinlichkeitsraum und $\Omega = \bigcup_{i \in I} B_i$ eine höchstens abzählbare Zerlegung in paarweise disjunkte Ereignisse $B_i \in \mathscr{F}$.

1. Satz von der totalen Wahrscheinlichkeit: Für alle $A \in \mathscr{F}$ gilt

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A \mid B_i)\mathbb{P}(B_i) \qquad \text{(totale Wahrscheinlichkeit)}$$

2. Satz von Bayes: Für alle $A \in \mathscr{F}$ mit $\mathbb{P}(A) > 0$ und alle $k \in I$

$$\mathbb{P}(B_k \mid A) = \frac{\mathbb{P}(A \mid B_k)\mathbb{P}(B_k)}{\sum_{i \in I} \mathbb{P}(A \mid B_i)\mathbb{P}(B_i)} \qquad \text{(Bayes)}$$

*Beweis.* 1. Es gilt:

$$\sum_{i \in I} \mathbb{P}(A \mid B_i)\mathbb{P}(B_i) \overset{\text{Def}}{=} \sum_{i \in I} \frac{\mathbb{P}(A \cap B_i)}{\mathbb{P}(B_i)}\mathbb{P}(B_i) = \sum_{i \in I} \mathbb{P}(A \cap B_i) \overset{\sigma-Add.}{=} \mathbb{P}(A)$$

2.

$$\mathbb{P}(B_k \mid A) \overset{\text{Def}}{=} \frac{\mathbb{P}(A \cap B_k)}{\mathbb{P}(A)} \overset{\text{Def}}{=} \frac{\mathbb{P}(A \mid B_k)\mathbb{P}(B_k)}{\mathbb{P}(A)}$$

also folgt (b) aus (a). □

■ **Beispiel 1.10**

In der Situation von Definition 1.3 folgt mit Satz 1.9 (totale Wahrscheinlichkeit)

$$\begin{aligned} \mathbb{P}(R) &= \mathbb{P}(R \mid W)\mathbb{P}(W) + \mathbb{P}(R \mid W^C)\mathbb{P}(W^C) \\ &= 1 \cdot \frac{2}{3} + \frac{1}{4}\frac{1}{3} = \frac{3}{4} \end{aligned}$$

und mit Satz 1.9 (Bayes)

$$\mathbb{P}(W \mid R) = \frac{\mathbb{P}(R \mid W)\mathbb{P}(W)}{\mathbb{P}(R)} = \frac{1 \cdot \frac{2}{3}}{\frac{3}{4}} = \frac{8}{9} \text{ für die gesuchte Wahrscheinlichkeit.}$$

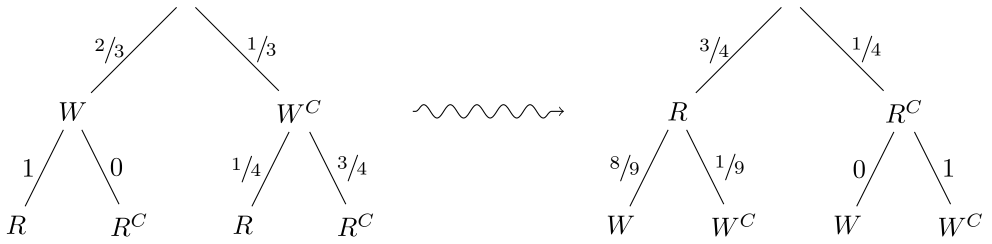



**Abbildung III.2:** Satz 1.3

## 2. (Un)abhängigkeit

In vielen Fällen besagt die Intuition über verschiedene Zufallsexperimente / Ereignisse, dass diese sich nicht gegenseitig beeinflussen. Für solche $A, B \in \mathscr{F}$ mit $\mathbb{P}(A) > 0, \mathbb{P}(B) > 0$ sollte gelten

$$\mathbb{P}(A \mid B) = \mathbb{P}(A), \quad \mathbb{P}(B \mid A) = \mathbb{P}(B).$$

**Definition 2.1 ((stochastisch) unabhängig)**
Sei $(\Omega, \mathscr{F}, \mathbb{P})$ Wahrscheinlichkeitsraum. Zwei Ereignisse $A, B \in \mathscr{F}$ heißt (stochastisch) unabhängig bezüglich $\mathbb{P}$, falls

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

Wir schreiben auch $A \perp\!\!\!\perp B$.

■ **Beispiel 2.2**
Würfeln mit 2 fairen, sechsseitigen Würfeln:

$$\Omega = \{(i,j) \mid i,j \in \{1,\dots,n\}\}, \quad \mathscr{F} = \mathcal{P}(\Omega), \quad \mathbb{P} = \mathrm{U}(\Omega)$$

Betrachte

$$A := \{(i,j) \in \Omega, i \text{ gerade}\}$$
$$B := \{(i,j) \in \Omega, j \leq 2\}$$

In diesem Fall, erwarten wir intuitiv Unabhängigkeit von $A$ und $B$.
In der Tat ist

$$\mathbb{P}(A) = \frac{1}{2}, \quad \mathbb{P}(B) = \frac{1}{3} \text{ und } \mathbb{P}(A \cap B) = \frac{1}{6}$$

was

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

erfüllt

Betrachte nun

$$\begin{aligned} C &:= \{(i,j) \in \Omega \mid i+j=7\} \\ D &:= \{(i,j) \in \Omega \mid i=6\} \end{aligned}$$

dann gilt

$$\mathbb{P}(C) = \frac{1}{6}, \quad \mathbb{P}(D) = \frac{1}{6}$$

und wegen $C \cap D = \{(6,1)\}$ folgt

$$\mathbb{P}(C \cap D) = \frac{1}{36} = \frac{1}{6}\frac{1}{6} = \mathbb{P}(C) \cdot \mathbb{P}(D)$$

$C$ und $D$ sind also stochastisch unabhängig, obwohl eine kausale Abhängigkeit vorliegt!

**Definition 2.3 (unabhängig bezüglich $\mathbb{P}$)**
Sei $(\Omega, \mathscr{F}, \mathbb{P})$ Wahrscheinlichkeitsraum und $I \neq \varnothing$ endliche Indexmenge. Dann heißt die Familie $(A_i)_{i \in I}$ von Ereignissen in $\mathscr{F}$ unabhängig bezüglich $\mathbb{P}$, falls für alle $J \subseteq I, J \neq \varnothing$ gilt:

$$\mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i)$$

Offensichtlich impliziert die Unabhängigkeit einer Familie die paarweise Unabhängigkeit je zweier Familienmitglieder nach Definition 2.1. Umgekehrt gilt dies nicht!

■ **Beispiel 2.4 (Abhängigkeit trotz paarweiser Unabhängigkeit)**
Betrachte zweifaches Bernoulliexperiment mit Erfolgswahrscheinlichkeit $1/2$, d.h.

$$\Omega = \{0,1\}^2, \quad \mathscr{F} = \mathcal{P}(\Omega), \quad \mathbb{P} = \mathrm{U}(\Omega)$$

sowie

$$\begin{aligned} A &= \{1\} \times \{0,1\} && \text{(Münzwurf: erster Wurf ist Zahl)} \\ B &= \{0,1\} \times \{1\} && \text{(Münzwurf: zweiter Wurf ist Zahl)} \\ C &= \{(0,0),(1,1)\} && \text{(beide Würfe haben selbes Ergebnis)} \end{aligned}$$

Dann gelten

$$\mathbb{P}(A) = \frac{1}{2} = \mathbb{P}(B) = \mathbb{P}(C)$$

und

$$\begin{aligned}\mathbb{P}(A \cap B) &= \mathbb{P}(\{(1,1)\}) = \frac{1}{4} = \mathbb{P}(A)\mathbb{P}(B)\\ \mathbb{P}(A \cap C) &= \mathbb{P}(\{(1,1)\}) = \frac{1}{4} = \mathbb{P}(A)\mathbb{P}(C)\\ \mathbb{P}(B \cap C) &= \mathbb{P}(\{(1,1)\}) = \frac{1}{4} = \mathbb{P}(B)\mathbb{P}(C)\end{aligned}$$

also paarweise Unabhängigkeit.
Aber

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

und $A, B, C$ sind nicht stochastisch unabhängig.

**Definition 2.5 (Unabhängige $\sigma$-Algebren)**
Seien $(\Omega, \mathscr{F}, \mathbb{P})$ Wahrscheinlichkeitsraum, $I \neq \varnothing$ Indexmenge und $(E_i, \mathcal{E}_i)$ Messräume

1. Die Familie $\mathscr{F}_i \subset \mathscr{F}, i \in I$, heißen unabhängig, wenn für die $J \subseteq I, J \neq \varnothing, |J| < \infty$ gilt

$$\mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i) \qquad \text{für beliebige } A_i \in \mathscr{F}_i, i \in J$$

2. Die Zufallsvariable $X_i : (\Omega, \mathscr{F}) \to (E_i, \mathcal{E}_i), i \in I$, heißen unabhängig, wenn die $\sigma$-Algebren

$$\sigma(X_i) = X^{-1}(\mathcal{E}_i) = \{\{X_i \in F\} \colon F \in \mathcal{E}_i\}, \quad i \in I$$

unabhängig sind.

**Lemma 2.6 (Zusammenhang der Definitionen)**
Sei $(\Omega, \mathscr{F}, \mathbb{P})$ Wahrscheinlichkeitsraum, $I \neq \varnothing, A \in \mathscr{F}, i \in I$. Die folgenden Aussagen sind äquivalent:

1. Die Ereignisse $A_i, i \in I$ sind unabhängig.
2. Die $\sigma$-Algebren $\sigma(A_i), i \in I$ sind unabhängig.
3. Die Zufallsvariablen $\mathbb{1}_{A_i}, i \in I$ sind unabhängig.

*Beweis.* Da die Unabhängigkeit über endliche Teilemengen definiert ist, können wir oBdA $I = \{1, \ldots, n\}$ annehmen.

- Da $\sigma(\mathbb{1}_{A_i}) = \sigma(A_i)$ folgt die Äquivalenz von 2. und 3. direkt aus Definition 2.5.
- Zudem ist 2. $\to$ 1. klar!

- Für $1 \to 2.$ genügt es zu zeigen, dass

$$A_1, \dots, A_n \text{ unabhängig } \Rightarrow B_1, \dots, B_n \text{ unabhängig mit } B_i \in \left\{\varnothing, A_i, A_i^C, \Omega\right\}.$$

Rekursiv folgt dies bereits aus

$$A_1, \dots, A_n \text{ unabhängig } \Rightarrow B_1, A_2, \dots, A_n \text{ unabhängig mit } B_1 \in \left\{\varnothing, A_1, A_1^C, \Omega\right\}.$$

Für $B_1 \in \{\varnothing, A_1, \Omega\}$ ist dies klar.
Sei also $B_1 = A_1^C$ und $J \subseteq I, J \neq \varnothing$. Falls $1 \notin J$, ist nichts zu zeigen. Sei $1 \in J$, dann gilt mit

$$A = \bigcap_{i \in J, i \neq 1} A_i$$

sicherlich

$$\begin{aligned}
\mathbb{P}\left(A_1^C \cap A\right) &= \mathbb{P}(A \setminus (A_1 \cap A)) \\
&= \mathbb{P}(A) - \mathbb{P}(A_1 \cap A) \\
&= \prod_{i \in J \setminus \{1\}} \mathbb{P}(A_i) - \prod_{i \in J} (A_i) \\
&= (1 - \mathbb{P}(A_1)) \prod_{i \in J \setminus \{1\}} \mathbb{P}(A_i) \\
&= \mathbb{P}\left(A_1^C\right)) \prod_{i \in J \setminus \{1\}} \mathbb{P}(A_i)
\end{aligned}$$

□

Insbesondere zeigt Lemma 2.6, dass wir in einer Familie unabhängiger Ereignisse beliebig viele Ereignisse durch ihr Komplement, $\varnothing$ oder $\Omega$ ersetzen können, ohne die Unabhängigkeit zu verlieren.

**Satz 2.7**
Sei $(\Omega, \mathscr{F}, \mathbb{P})$ Wahrscheinlichkeitsraum und $\mathscr{F}_i \subseteq \mathscr{F}, i \in I$, seien $\cap$-stabile Familien von Ereignissen. Dann gilt

$$\mathscr{F}_i, i \in I \text{ unabhängig } \Leftrightarrow \sigma(\mathscr{F}_i), i \in I \text{ unabhängig.}$$

*Beweis.* oBdA sei $I = \{1, \dots, n\}$ und $\Omega \in \mathscr{F}_i, i \in I$.

- $\Leftarrow$: trivial, da $\mathscr{F}_i \subseteq \sigma(\mathscr{F}_i)$ und das Weglassen von Mengen erlaubt ist.
- $\Rightarrow$: zeigen wir rekursiv
  1. Wähle $F_i \in \mathscr{F}_i, i = 2, \dots, n$ und defniere für $F \in \sigma(\mathscr{F}_i)$ die endlichen Maße

  $$\mu(F) = \mathbb{P}(F \cap F_2 \cap \dots \cap F_n) \text{ und } \nu(F) = \mathbb{P}(F)\,\mathbb{P}(F_2)\,\dots\,\mathbb{P}(F_n)$$

  2. Da die Familien $\mathscr{F}_i$ unabhängig sind, gilt $\mu \mid_{\mathscr{F}_1} = \nu \mid_{\mathscr{F}_1}$. Nach dem Eindeutigkeitssatz für Maße (**???**) folgt $\mu \mid_{\sigma(\mathscr{F}_1)} = \nu \mid_{\sigma(\mathscr{F}_1)}$ also

  $$\mathbb{P}(F \cap F_2 \cap \dots \cap F_n) = \mathbb{P}(F)\mathbb{P}(F_2)\dots\mathbb{P}(F_n)$$

  für alle $F \in \sigma(\mathscr{F}_i)$ und $F_i \in \mathscr{F}_i, i = 1, \dots, n$. Da $\Omega \in \mathscr{F}_i$ für alle $i$ gilt die erhaltene Produktformel für alle Teilemengen $J \subseteq I$.

Also sind

$$\sigma(\mathscr{F}_1), \mathscr{F}_2, \dots, \mathscr{F}_n \text{ unabhängig}$$

3. Wiederholtes Anwenden von 1 und 2 liefert den Satz. □

Mit Satz 2.7 folgen:

**Folgerung 2.8**

Sei $(\Omega, \mathscr{F}, \mathbb{P})$ Wahrscheinlichkeitsraum und

$$\mathscr{F}_{i,j} \subseteq \mathscr{F}, \quad 1 \leq i \leq n, 1 \leq j \leq m(i)$$

unabhängige, $\cap$-stabile Familien. Dann sind auch

$$\mathscr{G}_i = \sigma(\mathscr{F}_{i,1}, \dots, \mathscr{F}_{i,m(i)}), \quad 1 \leq i \leq n$$

unabhängig.

**Folgerung 2.9**

Sei $(\Omega, \mathscr{F}, \mathbb{P})$ Wahrscheinlichkeitsraum und

$$X_{ij} : \Omega \to E, \quad 1 \leq i \leq n, 1 \leq j \leq m(i)$$

unabhängige Zufallsvariablen. Zudem seinen $f_i : E^{m(i)} \to \mathbb{R}$ messbar. Dann sind auch die Zufallsvariablen

$$f_i(X_{i,1}, \dots, X_{i,m(i)}), \quad 1 \leq i \leq n$$

unabhängig.

■ **Beispiel 2.10**

$X_1, \dots, X_n$ unabhängige reelle Zufallsvariablen. Dann sind auch

$$Y_1 = X_1, Y_2 = X_2 + \dots + X_n$$

unabhängig.

*Beweis* (Folgerung 2.8). OBdA sei $\Omega \in \mathscr{F}_{i,j} \forall i, j$. Dann sind die Familien:

$$\mathscr{F}_i^{\cap} := \left\{ F_{i,1} \cap \dots \cap F_{i,m(i)} \mid F_{i,j} \in \mathscr{F}_{i,j}, 1 \leq j \leq m(i) \right\}, 1 \leq i \leq n$$

$\cap$-stabil, unabhängig und es gilt: $\mathscr{F}_{i,1}, \dots, \mathscr{F}_{i,m(i)} \subseteq \mathscr{F}_i^{\cap}$ ($\nearrow$ HA)! Nach Satz 2.7 sind auch $\sigma(\mathscr{F}_i^{\cap})$ unabhängig. Damit folgt die Behauptung, da $\sigma(\mathscr{F}_i^{\cap}) = \mathscr{G}_i$:

$$\begin{aligned} \mathscr{F}_{i,1}, \dots, \mathscr{F}_{i,m(i)} \subseteq \mathscr{F}_i^{\cap} \subseteq \sigma(\mathscr{F}_{i,1}, \dots, \mathscr{F}_{i,m(i)}) = \mathscr{G}_i \\ \Rightarrow \mathscr{G} = \sigma(\mathscr{F}_{i,1}, \dots, \mathscr{F}_{i,m(i)}) \subseteq \sigma(\mathscr{F}_i^{\cap}) \subseteq \mathscr{G}_i. \end{aligned}$$

□

*Beweis* (Folgerung 2.9). Setze $\mathscr{F}_{i,j} = \sigma(X_{i,j})$ und $\mathscr{G}_i = \sigma(\mathscr{F}_{i,1}, \dots, \mathscr{F}_{i,m(i)})$, dann sind nach Folgerung 2.8 die

$\mathscr{G}_i, i = 1, \dots, n$ unabhängig. Zudem ist

$$Y_i := f_i(X_{i,1}, \dots, X_{i,m(i)})$$

$\mathscr{G}_i$ messbar, also $\sigma(Y_i) \subseteq \mathscr{G}_i$. Damit erben die $Y_i$ die Unabhängigkeit der $\mathscr{G}_i$. □

**Satz 2.11 (Unabhängigkeit von Zufallsvariablen)**
$(\Omega, \mathscr{F}, \mathbb{P})$ Wahrscheinlichkeitsraum und $X_1, \dots, X_n : (\Omega, \mathscr{F}) \to (E, \mathcal{E})$ Zufallsvariablen. Dann sind

1. $X_1, \dots, X_n$ sind unabhängig
2. $\mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n) = \prod_{i=1}^n \mathbb{P}(X_i \in A_i) \quad \forall A_1, \dots, A_n \in \mathcal{E}$.
3. Die gemeinsame Verteilung der $X_i$ entspricht den Produktmaß der einzelnen Verteilungen

$$\mathbb{P}_{X_1,\dots,X_n} = \bigotimes_{i=1}^n \mathbb{P}_{X_i}$$

*Beweis.* Per Ringschluss:

1 ⇒ 2: Seien $A_1, \dots, A_n \in E$ beliebig, dann gilt per Definition

$$\begin{aligned}
\mathbb{P}_{X_1,\dots,X_n}(A_1 \times \dots \times A_n) &= \mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n) \\
&= \mathbb{P}\left(\bigcap_{i=1}^n \{X_i \in A_i\}\right) \\
&\overset{\text{unabh}}{=} \prod_{i=1}^n \mathbb{P}(X_i \in A_i) \\
&= \prod_{i=1}^n \mathbb{P}_{X_i}(A_i) = \left(\bigotimes_{i=1}^n \mathbb{P}_{X_i}\right)(A_1 \times \dots \times A_n)
\end{aligned}$$

2 ⇒ 3: Aus der obigen Rechnung sehen wir, dass 2 bereits 3 impliziert für alle Rechtecke: $\bigtimes_{i=1}^n A_i$. Da die Familie der Rechtecke $\cap$-stabil ist und $\mathcal{E}^{\otimes n}$ erzeugt, folgt die Aussage aus dem Eindeutigkeitssatz für Maße **???**.

3 ⇒ 1: Sei $J \subseteq \{1, \dots, n\}$ und setze

$$A_i := \begin{cases} \text{beliebig} & \in \mathcal{E}, i \in J \\ E & i \notin J. \end{cases}$$

Dann

$$\begin{aligned}
\mathbb{P}(X_i \in A_i, i \in J) &= \mathbb{P}(X_i \in A_i, i = 1, \dots, n) \\
&= \prod_{i=1}^n \mathbb{P}(X_i \in A_i) \\
&= \prod_{i \in J} \mathbb{P}(A_i \in A_i).
\end{aligned}$$ □

■ **Beispiel 2.12**
Im Urnenmodell mit Zurücklegen hat der Vektor $X = (X_1, \dots, X_n)$ mit $X_i =$ Farbe im $i$-ten Zug

als Zähldichte die Produktdichte der $X_i$. Die $X_1, \dots, X_n$ sind also unabhängig.

## Konstruktion unabhängiger Zufallsvariablen

Kapitel I: Zu beliebiger Wahrscheinlichkeitsverteilung $\mathbb{P}_X$ existiert Wahrscheinlichkeitsraum mit Zufallsvariable $X$ auf diesem Wahrscheinlichkeitsraum, so dass $X \sim \mathbb{P}_X$.

1. Seien $\mathbb{P}X_1, \dots, \mathbb{P}_{X_n}$ Wahrscheinlichkeitsverteilungen auf $(E, \mathcal{E})$. Gibt es einen Wahrscheinlichkeitsraum $(\Omega, \mathscr{F}, \mathbb{P})$ und Zufallsvariablen $X_1, X_2$ unabhängig, so dass $X_1 \sim \mathbb{P}_{X_1}$?
2. Wie kann ich beliebig (unendlich) viele unabhängige Zufallsvariablen konstruieren?

Wir beginnen mit 1:
Konstruiere zwei Wahrscheinlichkeitsräume $(\Omega_i, \mathscr{F}_i, \mathbb{P}_i), i = 1, 2$ und Zufallsvariablen $X_1, X_2$ mit $X_i \sim \mathbb{P}_{X_i}$. Auf dem Produktraum

$$\Omega = \Omega_1 \times \Omega_2, \quad \mathscr{F} := \mathscr{F}_1 \otimes \mathscr{F}_2 \text{ und } \mathbb{P} = \mathbb{P}_1 \otimes \mathbb{P}_2$$

definiere

$$\begin{aligned} X_1' &: \Omega_1 \times \Omega_2 \to E \colon (\omega_1, \omega_2) \mapsto X_1(\omega_1) \\ X_2' &: \Omega_1 \times \Omega_2 \to E \colon (\omega_1, \omega_2) \mapsto X_2(\omega_2) \end{aligned}$$

Dann gilt für beliebige Ereignisse: $F_1, F_2 \in \mathcal{E}$

$$\underbrace{\{X_1' \in F_1\} \cap \{X_2' \in F_2\}}_{\supseteq \Omega = \Omega_1 \times \Omega_2} = \underbrace{\{X_1 \in F_1\}}_{\supseteq \Omega_1} \times \underbrace{\{X_2 \in F_2\}}_{\supseteq \Omega_2} \in \mathscr{F}_1 \times \mathscr{F}_2$$

und damit folgt die Messbarkeit der Abbildungen $X_1', X_2'$, d.h. $X_1', X_2'$ sind Zufallsvariablen auf $(\Omega, \mathscr{F})$. Zudem gilt

$$\begin{aligned} \mathbb{P}(X_1' \in F_1, X_2' \in F_2) &= \mathbb{P}_1 \otimes \mathbb{P}_2(\{X_1 \in F_1\} \times \{X_2 \in F_2\}) \\ &= \mathbb{P}_1(X_1 \in F_1)\mathbb{P}_2(X_2 \in F_2), \end{aligned}$$

also

$$\mathbb{P}(X_i' \in F_i) = \mathbb{P}_i(X_i' \in F_i)$$

sowie nach Satz 2.13 $X_1' \perp\!\!\!\perp X_2'$.
Wenn $(\Omega_2, \mathscr{F}_1, \mathbb{P}_1) = (\Omega_2, \mathscr{F}_2, \mathbb{P}_2)$, so liefert die obige Konstruktion zwei unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum. Andernfalls können wir auf den Produktraum ausweichen und $X_i'$ anstelle von $X_i$ betrachten. Die obige Konstruktion lässt sich direkt auf endlich viele Zufallsvariablen übertragen.
Zu 2:

**Satz 2.13 (Satz von Kolmogorov)**
Sei $I$ beliebige Indexmenge und $(\Omega_i, \mathscr{F}_i, \mathbb{P}_i), i \in I$ Wahrscheinlichkeitsräume. Setze

$$\Omega_I := \bigtimes_{i\in I} \Omega_i = \left\{\omega : I \to \bigcup_{i\in I} \Omega_i, \omega_i \in \Omega_i, i \in I\right\}$$
$$\mathscr{F}_I := \sigma(\pi^{-1}(\mathscr{F}_i), i \in I)$$

wobei $\pi_i : \Omega_I \to \Omega_i$ mit $\omega \mapsto \omega_i$ die Projektionsabbildung. Dann existiert auf $(\Omega_I, \mathscr{F}_I)$ genau ein Maß $\mathbb{P}_I$, sodass für alle $H \subseteq I$ mit $0 < |H| < \infty$ gilt

$$\pi_H(\mathbb{P}_I) = \bigotimes_{i\in H} \mathbb{P}_i,$$

wobei $\pi_H : \Omega_I \to \Omega_H$ wiederum die Projektionsabbildung.

*Beweis.* $\nearrow$ Schilling Maß und Integral, Satz 17.4. □

Sind auf den Wahrscheinlichkeitsräumen $(\Omega_i, \mathscr{F}_i, \mathbb{P}_i), i \in I$, nun Zufallsvariablen $X_i : \Omega_i \to E$ gegeben, so definieren wir wie im Satz von Kolmogorov (Satz 2.13)

$$(\Omega, \mathscr{F}, \mathbb{P}) := (\Omega_I, \mathscr{F}_I, \mathbb{P}_I = \bigotimes_{i\in I}) \text{ mit } \omega = (\omega_i)_{i\in I}$$

und wie im endlichen Fall

$$X_i' : \Omega \to E \text{ mit } X_i'(\omega) = X_i(\omega_i).$$

Da die Unabhängigkeit der Zufallsvariablen über endliche Teilfamilien definiert ist, folgt diese wie im endlichen Fall.

## Faltungen

Seien $X, Y$ zwei reelle und unabhängige Zufallsvariablen mit

$$X \sim \mathbb{P}_X \text{ und } Y \sim \mathbb{P}_Y$$

Dann hat $(X, Y)$ die Verteilung $\mathbb{P}_X \otimes \mathbb{P}_Y$ auf $\mathbb{R}^2$. Andernfalls ist auch $X + Y$ eine reelle Zufallsvariable, dann

$$X + Y = A(X, Y) \text{ mit } A : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto x + y.$$

$A$ ist stetig, also messbar. Die Verteilung von $X + Y$ ist dann $(\mathbb{P}_X \otimes \mathbb{P}_Y) \circ A^{-1}$

**Definition 2.14 (Faltung)**
Seien $\mathbb{P}_1, \mathbb{P}_2$ Wahrscheinlichkeitsmaße auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$. Das durch

$$\mathbb{P}_1 * \mathbb{P}_2(F) = \iint \mathbb{1}_F(x+y)\mathbb{P}_1(\,\mathrm{d}x)\mathbb{P}_2(\,\mathrm{d}y)$$

definierte Wahrscheinlichkeitsmaß $\mathbb{P}_1 * \mathbb{P}_2 = (\mathbb{P}_1 \otimes \mathbb{P}_2) \circ A^{-1}$ auf $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ heißt Faltung von $\mathbb{P}_1$ und $\mathbb{P}_2$.

**Kapitel IV**

# *Test*

# Anhang

# Literaturverzeichnis

[1] Bauer, H. Wahrscheinlichkeitstheorie, 5 ed. De Gruyter, 2002.

[2] Dehling, H., and Haupt, B. Einführung in die Wahrscheinlichkeitstheorie und Statistik. Springer, 2003.

[3] Georgii, H.-O. Stochastik: Einführung in die Wahrscheinlichkeitstheorie und Statistik, 5 ed. De Gruyter, 2015.

[4] Krengel, U. Einführung in die Wahrscheinlichkeitstheorie und Statistik:. Vieweg, 2005.

[5] Schilling, R. L. Wahrscheinlichkeit: eine Einführung für Bachelor-Studenten, 1 ed. De Gruyter, 2017.

# Index