

# **Stochastik SS 2019**

Dozent: Prof. Dr. ANITA BEHME

7. Mai 2019

# *Inhaltsverzeichnis*

<b>I</b>	<b>Grundbegriffe der Wahrscheinlichkeitstheorie</b>	<b>3</b>
1	Wahrscheinlichkeitsräume . . . . .	3
2	Zufallsvariablen . . . . .	7
<b>II</b>	<b>Erste Standardmodelle der Wahrscheinlichkeitstheorie</b>	<b>12</b>
1	Diskrete Gleichverteilungen . . . . .	12
2	Urnenmodelle . . . . .	12
2.1	Urnenmodell mit Zurücklegen: Multinomial-Verteilung . . . . .	13
2.2	Urnenmodell ohne Zurücklegen: Hypergeometrische Verteilung . . . . .	15
3	Poisson-Approximation und POISSON-Verteilung . . . . .	15
<b>III</b>	<b>Bedingte Wkeiten und (Un-)abhängigkeit</b>	<b>17</b>
1	Bedingte Wahrscheinlichkeiten . . . . .	17
2	(Un)abhängigkeit . . . . .	22
<b>IV</b>	<b>Weitere Standardmodelle der Wahrscheinlichkeitstheorie</b>	<b>32</b>
1	Stetige Gleichverteilung . . . . .	32
2	Wartezeitverteilungen . . . . .	32
2.1	Exponential- und Gammaverteilung . . . . .	33
<b>V</b>	<b>Erwartungswerte &amp; Varianz</b>	<b>37</b>
1	Der Erwartungswert . . . . .	37
	<b>Anhang</b>	<b>40</b>
	<b>Index</b>	<b>41</b>

# *Vorwort*

# *Was ist Stochastik?*

Altgriechisch Stochastikos ( $\sigma\tau\omicron\chi\alpha\sigma\tau\iota\kappa\omicron\varsigma$ ) 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 Sprache!  
Beispiel: “Beim Würfeln erscheint jedes sechste Mal (im Schnitt) eine 6.”  $\rightarrow$  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 I.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 Ergebnis des Zufallsexperiments, sondern nur für das Eintreten gewisser Ereignisse.

#### ■ Beispiel

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

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

#### **Definition I.2 (Ereignisraum, messbarer Raum)**

Sei  $\Omega \neq \emptyset$  ein Ergebnisraum und  $\mathcal{F}$  eine  $\sigma$ -Algebra auf  $\Omega$ , d.h. eine Familie von Teilmenge von  $\Omega$ , sodass

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

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

### Wahrscheinlichkeiten

Ordne Ereignissen Wahrscheinlichkeiten zu mittels der Abbildung

$$\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$$

sodass

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

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

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

**Definition I.3 (Wahrscheinlichkeitsmaß, Wahrscheinlichkeitsverteilung)**

Sei  $(\Omega, \mathcal{F})$  ein Ereignisraum und  $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$  eine Abbildung mit Eigenschaften (N) und (A).

Dann heißt  $\mathbb{P}$  Wahrscheinlichkeitsmaß oder auch Wahrscheinlichkeitsverteilung.

Aus der Definition folgen direkt:

**Satz I.4 (Rechenregeln für W-Maße)**

Sei  $\mathbb{P}$  ein W-Maß, Ereignisse  $(\Omega, \mathcal{F}), A, B, A_1, A_2, \dots \in \mathcal{F}$ . Dann gelten:

1.  $\mathbb{P}(\emptyset) = 0$
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.  $A_1 \subseteq A_2 \subseteq \dots$  und  $A = \bigcup_{i=1}^{\infty} (A_i)$ ) oder  $A_n \downarrow A$ , so gilt:

$$\mathbb{P}(A_n) \longrightarrow \mathbb{P}(A), n \rightarrow \infty$$

*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:  $A \cap B = \emptyset \implies \mathbb{P}(A \uplus B) = \mathbb{P}(A) + \mathbb{P}(B)$ .

Es kann  $\sigma$ -Additivität verwendet werden, indem "fehlende" Mengen durch  $\emptyset$  ergänzt werden:

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

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

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

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  $F_1 := A_1, F_2 := A_2 \setminus A_1, \dots, F_{i+1} := A_{i+1} \setminus A_n$ . Die  $F_i$  Mengen sind paarweise disjunkt und damit folgt für  $m \rightarrow \infty$

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

und

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

### ■ Beispiel I.5

Für ein beliebiges Ereignisraum  $(\Omega, \mathcal{F})$  ( $\Omega \neq \emptyset$ ) 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  $(\Omega, \mathcal{F})$ , welches wir als DIRAC-Maß oder DIRAC-Verteilung bezeichnen.

### ■ Beispiel I.6

Würfel mit fairem, 6-(gleich)seitigem Würfel mit Ergebnismenge  $\Omega = \{1, \dots, 6\}$  und Ereignisraum  $\mathcal{F} = \mathcal{P}(\Omega)$  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 I.7

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

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

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 I.8 (Konstruktion von Wahrscheinlichkeitsmaßen durch Dichten)**

Sei  $(\Omega, \mathcal{F})$  ein Ereignisraum.

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

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

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

- Umgekehrt definiert jedes Wahrscheinlichkeitsmaß  $\mathbb{P}$  auf  $(\Omega, \mathcal{F})$  definiert Folge  $\rho(\omega) = \mathbb{P}(\{\omega\})$ ,  $\omega \in \Omega$  eine Folge  $\rho$  mit den obigen Eigenschaften.
- $\Omega \subset \mathbb{R}^n$ ,  $\mathcal{F} = \mathcal{B}(\Omega)$ : Sei  $\rho : \Omega \rightarrow [0, \infty)$  eine Funktion, sodass

$$1. \int_{\Omega} \rho(x) dx = 1$$

$$2. \{x \in \Omega : f(x) \leq c\} \in \mathcal{B}(\Omega) \text{ für alle } c > 0$$

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

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

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 Behauptung aus den bekannten Eigenschaften des Lebesgue-Integrals ( $\nearrow$  Schilling MINT, Lemma 8.9)  $\square$

► **Bemerkung**

- Die eindeutige Beziehung zwischen Dichte und Wahrscheinlichkeitsmaß überträgt sich nicht auf den stetigen Fall.
  - Nicht jedes Wahrscheinlichkeitsmaß auf  $(\Omega, \mathcal{B}(\Omega))$  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  $\Omega \subset \mathbb{R}^n$  definierte Dichtefunktion  $\rho$  lässt sich auf ganz  $\mathbb{R}^n$  fortsetzen durch  $\rho(x) = 0$  mit  $x \notin \Omega$ . Das erzeugte Wahrscheinlichkeitsmaß auf  $(\mathbb{R}^n, \mathcal{B}(\Omega))$  lässt mit dem Wahrscheinlichkeitsmaß auf  $(\Omega, \mathcal{B}(\Omega))$  identifizieren.
- Mittels Dirac-Maß  $\delta_x$  können auch jedes diskrete Wahrscheinlichkeitsmaß auf  $\Omega \subset \mathbb{R}^n$  als



Wahrscheinlichkeitsmaß auf  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  interpretieren

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

stetige und diskrete Wahrscheinlichkeitsmaße lassen sich kombinieren z.B.

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

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

Abschließend erinnern wir uns an:

**Satz I.9 (Eindeutigkeitssatz für Wahrscheinlichkeitsmaße)**

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

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

Insbesondere definiert z.B.

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

bereits die Exponentialverteilung aus Beispiel I.7.

**Definition I.10 (Gleichverteilung)**

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

## Wahrscheinlichkeitsräume

**Definition I.11 (Wahrscheinlichkeitsraum)**

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

## 2. Zufallsvariablen

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

$$A' \in \mathcal{F}' \Rightarrow X^{-1}A' \in \mathcal{F}$$

d.h.  $X$  sollte messbar sein.

**Definition I.12 (Zufallsvariable)**

Seien  $(\Omega, \mathcal{F})$  und  $(\Omega', \mathcal{F}')$  Ereignisräume. Dann heißt jede messbare Abbildung

$$X : \Omega \rightarrow \Omega'$$

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

**■ Beispiel I.13**

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

**Satz I.14**

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

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

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

*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  $A'_1, A'_2, \dots \in \mathcal{F}'$  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  $\mathbb{P}(X \in A) = \mathbb{P}(\{\omega : X(\omega) \in A\})$
- 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  $(\Omega, \mathcal{F}, \mathbb{P})$  mit  $\mathcal{F} = \mathcal{P}(\Omega)/\mathcal{B}(\Omega)$ .
- 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 I.15 (identisch verteilt, reellen Zufallsvariablen)

Zwei Zufallsvariablen sind identisch verteilt, wenn sie dieselbe Verteilung haben. Von besonderem Interesse sind für uns die Zufallsvariablen, die nach  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  abbilden, sogenannte reelle Zufallsvariablen.

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

### Definition I.16 ((kumulative) Verteilungsfunktion von $\mathbb{P}$ )

Sei  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$  Wahrscheinlichkeitsraum, so heißt

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

(kumulative) Verteilungsfunktion von  $\mathbb{P}$ .

Ist  $X$  eine reelle Zufallsvariable auf beliebigem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$ , so heißt

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

die (kumulative) Verteilungsfunktion von  $X$ .

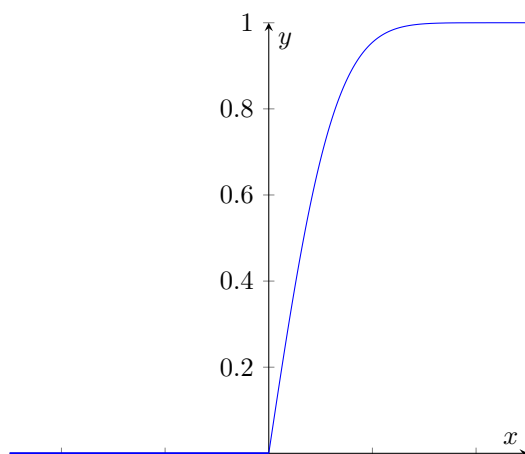
### ■ Beispiel I.17

Sei  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$  mit  $\mathbb{P}$  Exponentialverteilung mit Parameter  $\lambda > 0$

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

Dann ist

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



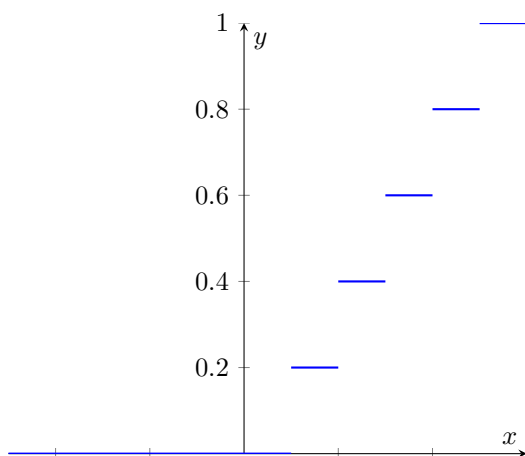
### ■ Beispiel I.18

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

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

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 I.19**

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

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

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

*Beweis.* Ist  $F$  Verteilungsfunktion, so folgt mit Satz I.4

$$x \leq y \Rightarrow F(x) = \mathbb{P}((-\infty, x]) \stackrel{\text{I.4.3}}{\leq} \mathbb{P}((-\infty, y]) = F(y)$$

und

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

sowie

$$\begin{aligned} \lim_{x \rightarrow -\infty} F(x) &\stackrel{\text{I.4.5}}{=} \mathbb{P}(\emptyset) \stackrel{\text{I.4.1}}{=} 0 \\ \lim_{x \rightarrow \infty} F(x) &\stackrel{\text{I.4.5}}{=} \mathbb{P}(\mathbb{R}) = 1. \end{aligned}$$

Umgekehrt wähle

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

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

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

und zudem

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

Da diese halboffene Mengen ein Erzeugendensystem von  $\mathcal{B}(\mathbb{R})$  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  $F$ .  $\square$

**Folgerung I.20**

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

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

*Beweis.* Folgt aus Satz I.8, der Definition I.16 der Verteilungsfunktion und dem Eindeigkeitssatz ???.  $\square$

## Kapitel II

# *Erste Standardmodelle der Wahrscheinlichkeitstheorie*

## Diskrete Verteilungen

### 1. Diskrete Gleichverteilungen

Erinnerung:

► **Erinnerung (Definition I.I.10)**

Ist  $\Omega$  endlich, so heißt Wahrscheinlichkeitsmaß mit Zähldichte

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

(diskrete) Gleichverteilung auf  $\Omega \rightarrow U(\Omega)$

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

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

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

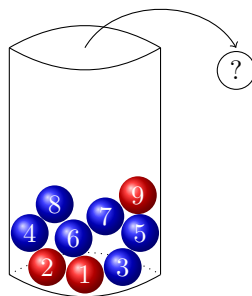
### 2. Urnenmodelle

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

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

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

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

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

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

Der Zufallsvektor

$$X = (X_1, \dots, X_n) : \bar{\Omega} \rightarrow \Omega$$

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

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

und damit

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

Zähldichten, die sich als Produkte von Zähl-dichten schreiben lassen, werden auch als Produktdichten bezeichnet ( $\nearrow$  Abschnitt 2). 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

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

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

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

und

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

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

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

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 II.1

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

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

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

### ■ Beispiel II.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 II.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  $\text{Bin}(n, p)$ . 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  $\text{Bernoulli}(p)$ .

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 II.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 II.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  $\text{Hyper}(N, W, n)$ .

## 3. Poisson-Approximation und Poisson-Verteilung

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

### Satz II.6 (Poisson-Approximation)

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

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

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

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

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

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

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

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

□

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 II.7

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  $\text{Poisson}(\lambda)$ .

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)-abhängigkeit*

### 1. Bedingte Wahrscheinlichkeiten

#### ■ Beispiel III.1

Das Würfeln mit zwei fairen, sechsseitigen Würfeln können wir mit

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

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

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

Betrachte das Ereignis

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

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  $A \subseteq \mathcal{F}$  mit  $A \subseteq B$  gilt  $\mathbb{P}_B(A) = c_B \mathbb{P}(A)$  mit einer Konstante  $c_B$ . (P)

**Lemma III.2**

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

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

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

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

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 III.3**

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

$$\mathbb{P}(A | B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad \text{mit } A \in \mathcal{F}$$

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

$$\mathbb{P}(A | B) = 0 \quad \forall A \in \mathcal{F}$$

**■ Beispiel III.4**

In der Situation Beispiel III.1 gilt

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

und damit

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

Aus Definition III.3 ergibt sich

**Lemma III.5 (Multiplikationsformel)**

Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  ein Wahrscheinlichkeitsraum und  $A_1, \dots, A_n \in \mathcal{F}$ . Dann gilt

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

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

Seite ungleich Null und

$$\begin{aligned}
\mathbb{P}(A_1)\mathbb{P}(A_2 | A_1) \dots \mathbb{P}(A_n | \bigcap_{i=1}^{n-1} A_i) \\
&= \mathbb{P}(A_1) \cdot \frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_1)} \dots \frac{\mathbb{P}(\bigcap_{i=1}^n A_i)}{\mathbb{P}(\bigcap_{i=1}^{n-1} A_i)} \\
&= \mathbb{P}\left(\bigcap_{i=1}^n A_i\right)
\end{aligned}$$

□

Stehen die  $A_i$  in Lemma III.5 in einer (zeitlichen) Abfolge, so liefert Formel einen Hinweis wie Wahrscheinlichkeitsmaße für Stufenexperimente konstruiert werden können. Ein Stufenexperiment aus  $n$  nacheinander ausgeführten Telexperimenten lässt sich als Baumdiagramm darstellen.

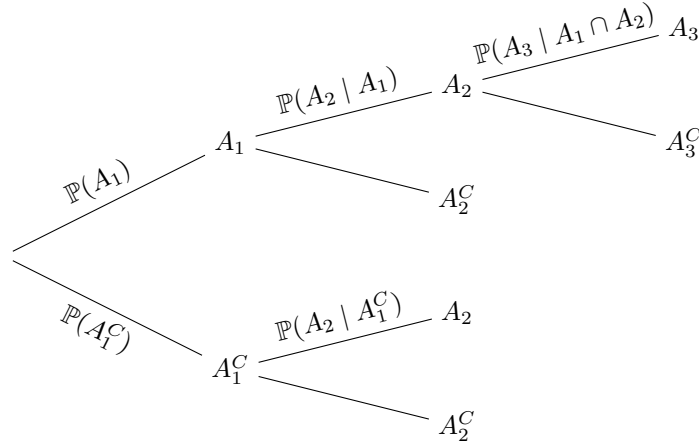


Abbildung III.1: Lemma III.5

### Satz III.6 (Konstruktion des Wahrscheinlichkeitsmaßes eines Stufenexperimentes)

Gegeben seien  $n$  Ergebnisräume  $\Omega_i = \{\omega_i(1), \dots, \omega_i(k)\}$ ,  $k \in \mathbb{N} \cup \{\infty\}$  und es sei  $\Omega = \times_{i=1}^n \Omega_i$  der zugehörige Produktraum. Weiter seien  $\mathcal{F}_i$   $\sigma$ -Algebren auf  $\Omega_i$  und  $\mathcal{F} = \bigotimes_{i=1}^n \mathcal{F}_i$  die Produkt- $\sigma$ -Algebra auf  $\Omega$ . Setze  $\omega = (\omega_1, \dots, \omega_n)$  und

$$\begin{aligned}
[\omega_1, \dots, \omega_m] &:= \{\omega_1\} \times \dots \times \{\omega_m\} \times \Omega_{m+1} \times \dots \times \Omega_n, \quad m \leq n \\
&\mathbb{P}(\{\omega_m\}[\omega_1, \dots, \omega_{m-1}])
\end{aligned}$$

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

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

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

*Beweis.* Nachrechnen!

□

### ■ Beispiel III.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

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

Mit Satz III.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

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

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

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

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

■ **Beispiel III.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

$$W = \{\text{richtige Antwort gewusst}\}$$

$$R = \{\text{Richtige Antwort gewählt}\}$$

Dann gilt

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

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

**Satz III.9**

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

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

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

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

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

*Beweis.* 1. Es gilt:

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

2.

$$\mathbb{P}(B_k | A) \stackrel{\text{Def}}{=} \frac{\mathbb{P}(A \cap B_k)}{\mathbb{P}(A)} \stackrel{\text{Def}}{=} \frac{\mathbb{P}(A | B_k) \mathbb{P}(B_k)}{\mathbb{P}(A)}$$

also folgt (b) aus (a). □

■ **Beispiel III.10**

In der Situation von Definition III.3 folgt mit Satz III.9 ([totale Wahrscheinlichkeit](#))

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

und mit Satz III.9 (Bayes)

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

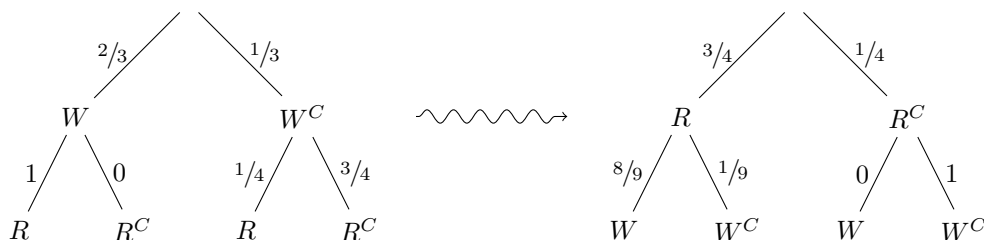


Abbildung III.2: Satz III.3

## 2. (Un)abhängigkeit

In vielen Fällen besagt die Intuition über verschiedene Zufallsexperimente / Ereignisse, dass diese sich nicht gegenseitig beeinflussen. Für solche  $A, B \in \mathcal{F}$  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 III.11 ((stochastisch) unabhängig)

Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  Wahrscheinlichkeitsraum. Zwei Ereignisse  $A, B \in \mathcal{F}$  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 III.12

Würfeln mit 2 fairen, sechsseitigen Würfeln:

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

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

$$C := \{(i, j) \in \Omega \mid i + j = 7\}$$

$$D := \{(i, j) \in \Omega \mid i = 6\}$$

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 III.13 (unabhängig bezüglich $\mathbb{P}$ )

Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  Wahrscheinlichkeitsraum und  $I \neq \emptyset$  endliche Indexmenge. Dann heißt die Familie  $(A_i)_{i \in I}$  von Ereignissen in  $\mathcal{F}$  unabhängig bezüglich  $\mathbb{P}$ , falls für alle  $J \subseteq I, J \neq \emptyset$  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 III.11. Umgekehrt gilt dies nicht!

### ■ Beispiel III.14 (Abhängigkeit trotz paarweiser Unabhängigkeit)

Betrachte zweifaches Bernoulliexperiment mit Erfolgswahrscheinlichkeit  $1/2$ , d.h.

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

sowie

$$A = \{1\} \times \{0, 1\} \quad (\text{Münzwurf: erster Wurf ist Zahl})$$

$$B = \{0, 1\} \times \{1\} \quad (\text{Münzwurf: zweiter Wurf ist Zahl})$$

$$C = \{(0, 0), (1, 1)\} \quad (\text{beide Würfe haben selbes Ergebnis})$$

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 III.15 (Unabhängige $\sigma$ -Algebren)

Seien  $(\Omega, \mathcal{F}, \mathbb{P})$  Wahrscheinlichkeitsraum,  $I \neq \emptyset$  Indexmenge und  $(E_i, \mathcal{E}_i)$  Messräume

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

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

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

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

unabhängig sind.

### Lemma III.16 (Zusammenhang der Definitionen)

Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  Wahrscheinlichkeitsraum,  $I \neq \emptyset, A \in \mathcal{F}, i \in I$ . 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 Teilmengen definiert ist, können wir oBdA  $I = \{1, \dots, n\}$  annehmen.

- Da  $\sigma(\mathbb{1}_{A_i}) = \sigma(A_i)$  folgt die Äquivalenz von 2. und 3. direkt aus Definition III.15.
- Zudem ist 2.  $\rightarrow$  1. klar!

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

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

Rekursiv folgt dies bereits aus

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

Für  $B_1 \in \{\emptyset, A_1, \Omega\}$  ist dies klar.

Sei also  $B_1 = A_1^C$  und  $J \subseteq I, J \neq \emptyset$ . 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

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

Insbesondere zeigt Lemma III.16, dass wir in einer Familie unabhängiger Ereignisse beliebig viele Ereignisse durch ihr Komplement,  $\emptyset$  oder  $\Omega$  ersetzen können, ohne die Unabhängigkeit zu verlieren.

### Satz III.17

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

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

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

- $\Leftarrow$ : trivial, da  $\mathcal{F}_i \subseteq \sigma(\mathcal{F}_i)$  und das Weglassen von Mengen erlaubt ist.
- $\Rightarrow$ : zeigen wir rekursiv

1. Wähle  $F_i \in \mathcal{F}_i, i = 2, \dots, n$  und definiere für  $F \in \sigma(\mathcal{F}_1)$  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  $\mathcal{F}_i$  unabhängig sind, gilt  $\mu|_{\mathcal{F}_1} = \nu|_{\mathcal{F}_1}$ . Nach dem Eindeutigkeitssatz für Maße (Satz I.1.9) folgt  $\mu|_{\sigma(\mathcal{F}_1)} = \nu|_{\sigma(\mathcal{F}_1)}$  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  $F \in \sigma(\mathcal{F}_1)$  und  $F_i \in \mathcal{F}_i, i = 2, \dots, n$ . Da  $\Omega \in \mathcal{F}_i$  für alle  $i$  gilt die erhaltene Produktformel für alle Teilmengen  $J \subseteq I$ .

Also sind

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

3. Wiederholtes Anwenden von 1 und 2 liefert den Satz.  $\square$

Mit Satz III.17 folgen:

**Folgerung III.18**

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

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

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

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

unabhängig.

**Folgerung III.19**

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

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

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

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

unabhängig.

■ **Beispiel III.20**

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

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

unabhängig.

*Beweis* (Folgerung III.18). OBdA sei  $\Omega \in \mathcal{F}_{i,j} \forall i, j$ . Dann sind die Familien:

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

$\cap$ -stabil, unabhängig und es gilt:  $\mathcal{F}_{i,1}, \dots, \mathcal{F}_{i,m(i)} \subseteq \mathcal{F}_i^\cap$  ( $\nearrow$  HA)! Nach Satz III.17 sind auch  $\sigma(\mathcal{F}_i^\cap)$  unabhängig. Damit folgt die Behauptung, da  $\sigma(\mathcal{F}_i^\cap) = \mathcal{G}_i$ :

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

*Beweis* (Folgerung III.19). Setze  $\mathcal{F}_{i,j} = \sigma(X_{i,j})$  und  $\mathcal{G}_i = \sigma(\mathcal{F}_{i,1}, \dots, \mathcal{F}_{i,m(i)})$ , dann sind nach Folgerung III.18

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

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

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

**Satz III.21 (Unabhängigkeit von Zufallsvariablen)**

$(\Omega, \mathcal{F}, \mathbb{P})$  Wahrscheinlichkeitsraum und  $X_1, \dots, X_n : (\Omega, \mathcal{F}) \rightarrow (E, \mathcal{E})$  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  $\Rightarrow$  2: Seien  $A_1, \dots, A_n \in E$  beliebig, dann gilt per Definition

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

2  $\Rightarrow$  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 Satz 1.1.9.

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

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

Dann

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

### ■ Beispiel III.22

Im Urnenmodell mit Zurücklegen hat der Vektor  $X = (X_1, \dots, X_n)$  mit  $X_i = \text{Farbe im } i\text{-ten Zug}$  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  $(\Omega, \mathcal{F}, \mathbb{P})$  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  $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i), i = 1, 2$  und Zufallsvariablen  $X_1, X_2$  mit  $X_i \sim \mathbb{P}_{X_i}$ . Auf dem Produktraum

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

definiere

$$\begin{aligned} X'_1 : \Omega_1 \times \Omega_2 &\rightarrow E : (\omega_1, \omega_2) \mapsto X_1(\omega_1) \\ X'_2 : \Omega_1 \times \Omega_2 &\rightarrow E : (\omega_1, \omega_2) \mapsto X_2(\omega_2) \end{aligned}$$

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

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

und damit folgt die Messbarkeit der Abbildungen  $X'_1, X'_2$ , d.h.  $X'_1, X'_2$  sind Zufallsvariablen auf  $(\Omega, \mathcal{F})$ . 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 III.23  $X'_1 \perp\!\!\!\perp X'_2$ .

Wenn  $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2) = (\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ , 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 III.23 (Satz von Kolmogorov)**

Sei  $I$  beliebige Indexmenge und  $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i), i \in I$  Wahrscheinlichkeitsräume. Setze

$$\Omega_I := \prod_{i \in I} \Omega_i = \left\{ \omega : I \rightarrow \bigcup_{i \in I} \Omega_i, \omega_i \in \Omega_i, i \in I \right\}$$

$$\mathcal{F}_I := \sigma(\pi^{-1}(\mathcal{F}_i), i \in I)$$

wobei  $\pi_i : \Omega_I \rightarrow \Omega_i$  mit  $\omega \mapsto \omega_i$  die Projektionsabbildung. Dann existiert auf  $(\Omega_I, \mathcal{F}_I)$  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  $\pi_H : \Omega_I \rightarrow \Omega_H$  wiederum die Projektionsabbildung.

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

Sind auf den Wahrscheinlichkeitsräumen  $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i), i \in I$ , nun Zufallsvariablen  $X_i : \Omega_i \rightarrow E$  gegeben, so definieren wir wie im Satz von Kolmogorov (Satz III.23)

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

und wie im endlichen Fall

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

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

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

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

**Definition III.24 (Faltung)**

Seien  $\mathbb{P}_1, \mathbb{P}_2$  Wahrscheinlichkeitsmaße auf  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ . Das durch

$$\mathbb{P}_1 \star \mathbb{P}_2(F) = \iint \mathbb{1}_F(x+y) \mathbb{P}_1(dx) \mathbb{P}_2(dy)$$

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

**Satz III.25**

Seien  $X, Y : \Omega \rightarrow \mathbb{R}^n$  unabhängige Zufallsvariablen mit Verteilungen  $\mathbb{P}_X, \mathbb{P}_Y$ . Dann ist

$$\mathbb{P}_{X+Y} = \mathbb{P}_X + \mathbb{P}_Y,$$

die Verteilung von  $X + Y$ .

*Beweis.* Siehe Herleitung Faltung. □

Faltung von Wahrscheinlichkeitsmaßen und Dichten besitzen wieder eine Dichte.

**Satz III.26**

Seien  $\mathbb{P}_1, \mathbb{P}_2$  Wahrscheinlichkeitsmaße auf  $(\mathbb{R}, \mathcal{B}(\mathbb{R}^n))$

1. Diskreter Fall: Sind  $\mathbb{P}_1, \mathbb{P}_2$  de facto Wahrscheinlichkeitsmaße auf  $(\mathbb{Z}, \mathcal{P}(\mathbb{Z}))$  mit Zähldichte  $\rho_1, \rho_2$ . Dann ist die Faltung  $\mathbb{P}_1 \star \mathbb{P}_2$  Wahrscheinlichkeitsmaß auf  $(\mathbb{Z}, \mathcal{B}(\mathbb{Z}))$  mit Zähldichte

$$\rho_1 \star \rho_2(k) = \sum_{l \in \mathbb{Z}} \rho_1(l) \rho_2(k-l).$$

2. Stetiger Fall: Besitzt  $\mathbb{P}_1, \mathbb{P}_2$  Dichtefunktionen  $\rho_1, \rho_2$ , so besitzt die Faltung  $\mathbb{P}_1 \star \mathbb{P}_2$  die Dichtefunktion

$$\rho_1 \star \rho_2(x) = \int_{\mathbb{R}} \rho_1(y) \rho_2(x-y) dy \quad x \in \mathbb{R}$$

*Beweis.* 1. Diskrete Fall: Sei  $k \in \mathbb{Z}$

$$\begin{aligned} (\mathbb{P}_1 \otimes \mathbb{P}_2)(A = k) &= \sum_{\substack{l_1, l_2 \in \mathbb{Z} \\ l_1 + l_2 = k}} \rho_1(l_1) \rho_2(l_2) \\ &= \rho_1 \star \rho_2(k) \end{aligned}$$



2. Stetiger Fall: Sei  $c \in \mathbb{R}$

$$\begin{aligned}
 \mathbb{P}_1 + \mathbb{P}_2((-\infty, c]) &= (\mathbb{P}_1 \otimes \mathbb{P}_2)(A \leq c) \\
 &= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_{(-\infty, c]}(x+y) \rho_1(x) \rho_2(y) \, dx \, dy \\
 &\stackrel{y=z-x}{=} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_{(-\infty, c]}(z) \rho_1(x) \rho_2(z-x) \, dx \, dz \\
 &= \int_{-\infty}^c \underbrace{\int_{\mathbb{R}} \rho_1(x) \rho_2(z-x) \, dx}_{\rho_1 \star \rho_2(z)} \, dz. \quad \square
 \end{aligned}$$

### ■ Beispiel III.27

Seien  $X \sim \text{Poisson}(\lambda), Y \sim \text{Poisson}(\mu)$  zwei unabhängigen reellen Zufallsvariablen (mit Werten in  $\mathbb{N}_0$ ). Dann ist  $X + Y$  eine Zufallsvariable mit Werten in  $\mathbb{N}_0$  und Zähldichte

$$\begin{aligned}
 \mathbb{P}(X + Y = k) &= \sum_{l \in \mathbb{Z}} \mathbb{P}(X = l) \mathbb{P}(Y = k - l) \\
 &= \sum_{l \in \mathbb{Z}} \frac{\lambda^l}{l!} e^{-\lambda} \frac{\mu^{k-l}}{(k-l)!} e^{-\mu} \\
 &= e^{-(\lambda+\mu)} \frac{1}{k!} \sum_{l=0}^k \binom{k}{l} \lambda^l \mu^{k-l} \\
 &= e^{-(\lambda+\mu)} \frac{1}{k!} (\lambda + \mu)^k \quad \forall k \in \mathbb{N}_0,
 \end{aligned}$$

so dass

$$X + Y \sim \text{Poisson}(\lambda + \mu).$$

D.h. der Typ der Verteilung ist bei der Faltung erhalten geblieben.

### ► Bemerkung III.28

Das ist aber nicht immer der Fall!

### ■ Beispiel III.29

Seien  $X, Y \sim U([0, 1])$  zwei unabhängige Zufallsvariablen mit Dichten  $\rho(x) = \mathbb{1}_{[0,1]}(x)$ . Dann ist  $X + Y$  eine Zufallsvariable mit Werten in  $[0, 2]$  und Dichte

$$\begin{aligned}
 \rho \star \rho(x) &= \int_{\mathbb{R}} \rho(y) \rho(x-y) \, dx \, dy \\
 &= \int_{\mathbb{R}} \mathbb{1}_{[0,1]}(y) \mathbb{1}_{[0,1]}(x-y) \, dy \\
 &= \int_{0 \vee (x-1)}^{1 \wedge x} = \begin{cases} x & 0 \leq x \leq 1 \\ 2-x & 1 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}
 \end{aligned}$$

## Kapitel IV

# Weitere Standardmodelle der Wahrscheinlichkeitstheorie

## 1. Stetige Gleichverteilung

### ► Erinnerung

$\Omega \subset \mathbb{R}^n$  Borel-messbar mit LEBESGUE-Volumen  $0 < \lambda(\Omega) < \infty$ . Wahrscheinlichkeitsmaß ist  $(\Omega, \mathcal{B}(\Omega))$  mit Dichte

$$q\rho(x) = \frac{1}{\lambda(\Omega)}$$

heißt stetige Gleichverteilung auf  $\Omega$ :  $U(\Omega)$ .

Für alle  $A \in \mathcal{B}(\Omega)$  gilt:

$$\mathbb{P}(A) = \int_A \rho(x) \, dx = \frac{\lambda(A)}{\lambda(\Omega)}.$$

Meist verwenden wir  $U([a, b])$ ,  $a < b$  (Gleichverteilung auf Intervall) mit  $\rho(x) = 1/(b-a)$ ,  $a \leq x \leq b$  und Verteilungsfunktion

$$F(x) = \begin{cases} 0 & x < a \\ \int_a^x \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b. \end{cases}$$

## 2. Wartezeitverteilungen

Negative Binomialverteilung:

Wir wiederholen ein Bernoulliexperiment mit Erfolgswahrscheinlichkeit  $p \in [0, 1]$  unendlich oft. Gesucht ist die Anzahl der Misserfolge bis zum  $r$ -ten Erfolg,  $r \in \mathbb{N}$ . Ein passender Ergebnisraum ist  $\Omega = \mathbb{N}_0$ . Für Modellierung ist es jedoch leichter in jedem Versuch erfolgt ("1") oder Misserfolg ("0") festzuhalten und  $i$  mit dem unendlichen Produktmaß des Bernoullimaßes auf  $\{0, 1\}^{\mathbb{N}}$  zu arbeiten.

Als Zufallsvariable

$$X_r : \{0, 1\}^{\mathbb{N}} \rightarrow \Omega$$

welche die Anzahl der Misserfolge bis zum  $r$ -ten Erfolg darstellt, setze

$$X_r(\omega) = \min \left\{ \sum_{i=1}^k \omega_i = r \right\} = r.$$

Dann

$$\mathbb{P}(X_r = k) = \sum_{\substack{\omega \in \{0,1\}^{\mathbb{N}} \\ X_r(\omega) = k}} \prod_{i=1}^{\infty} \rho(\omega_i)$$

mit  $\rho(0) = 1 - p, \rho(1) = p$  (Zähldichte der Bernoulliverteilung), also

$$\mathbb{P}(X_r = k) = \binom{r+k-1}{k} (1-p)^k p^r \quad r \in \mathbb{N}_0.$$

### Definition IV.1 (negative Binomialverteilung, geometrische Verteilung)

Sei  $p \in [0, 1]$  und  $r \in \mathbb{N}$ , dann heißt die Verteilung auf  $\mathbb{N}_0$  mit Zähldichte

$$\rho(k) = \binom{r+k-1}{k} p^r (1-p)^k$$

die negative Binomialverteilung mit Parametern  $(r, p)$ . Schreibe  $\text{negBin}(r, p)$ . Im Fall  $r = 1$  nennen wir die Verteilung mit Zähldichte

$$\rho(k) = p(1-p)^k \quad k \in \mathbb{N}_0$$

die geometrische Verteilung mit Parametern  $p$ . Schreibe  $\text{Geom}(p)$ .

## 2.1. Exponential- und Gammaverteilung

1. Ziel: Modelliere die Wartezeit auf  $r$  Ereignisse in kontinuierlicher Zeit.
2. Wähle:  $(\Omega, \mathcal{F}) = (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$
3. Annahmen:

- Jedes Ereignis geschieht zu einer zufälligen Zeit
- Die Anzahl der Ereignisse bis zur Zeit  $t$  sei  $\text{Poisson}(\lambda t)$  verteilt.

Die zweite Annahme macht Sinn, denn

- Poissonverteilung ist Modell für Anzahl seltener Ereignisse
- Nach Beispiel 3.26:

$$\text{Poisson}(\lambda t) \star \text{Poisson}(\lambda s) = \text{Poisson}(\lambda(t+s))$$

Die Linearität des Parameters entspricht also einer Stationaritätsvoraussetzung:

Modelliert

$X \sim \text{Poisson}(\lambda t)$  die Ereignisse in  $(0, t]$ ,

$Y \sim \text{Poisson}(\lambda s)$  die Ereignisse in  $(t, t + s]$

so modelliert

$X + Y \sim \text{Poisson}(\lambda(t + s))$  die Ereignisse in  $(0, t + s]$

Unter diesen Annahmen folgt für die Wahrscheinlichkeit in  $(0, t]$  mindestens  $r$  Ereignisse zu beobachten

$$\begin{aligned} \mathbb{P}((0, t)) &= 1 - \sum_{k=0}^{r-1} \underbrace{\exp(-\lambda t) \frac{(\lambda t)^k}{k!}}_{\text{Zähldichte Poisson}(\lambda t) \text{ in } t} \\ &= \frac{d}{dt} \left( 1 - \sum_{k=0}^{r-1} \exp(-\lambda t) \frac{(\lambda t)^k}{k!} \right) \\ &= -(-\lambda) e^{-\lambda t} \sum_{k=0}^{r-1} \frac{(\lambda t)^k}{k!} - e^{-\lambda t} \sum_{k=0}^{r-1} \frac{\lambda^k t^{k-1}}{(k-1)!} \\ &= e^{-\lambda t} \left( \sum_{k=0}^{r-1} \frac{\lambda^k t^{k-1}}{k!} - \sum_{l=0}^{r-1} \frac{\lambda^k t^{l-1}}{(l-1)!} \right) \end{aligned}$$

gilt

$$\mathbb{P}((0, t)) = \int_0^t e^{-\lambda x} \frac{\lambda^r x^{r-1}}{(r-1)!} dx$$

Wir definieren allgemeiner:

### Definition IV.2

Seien  $\lambda > 0, r > 0$ , dann heißt die Verteilung auf  $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$  mit Dichte

$$f(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x},$$

wobei

$$\Gamma(r) = \int_0^\infty y^{r-1} e^{-y} dy, r > 0$$

Gammefunktion, Gammaverteilung mit Parametern  $\lambda, r$ . Schreibe  $\text{Gamma}(\lambda, r)$ . Insbesondere ist  $\text{Geom}(\lambda, 1)$  gerade die Exponentialverteilung (vgl. Bsp 17?).

Die Gammaverteilung ist reproduktiv: Die Wartezeit auf  $r + s$  Ereignisse entspricht der Wartezeit auf  $r$  Ereignisse +  $s$  (weitere) Ereignisse:

**Lemma IV.3**

Seien  $X \sim \text{Gamma}(\lambda, r)$ ,  $Y \sim \text{Gamma}(\lambda, s)$  unabhängig, dann impliziert das

$$X + Y \sim \text{Gamma}(\lambda, r + s)$$

*Beweis.* Hier nur für  $r, s \in \mathbb{N}$ , allgemein später mit momenterzeugenden Funktionen.

Seien  $\rho(x), \rho(y)$  Dichten von  $X, Y$ . Nach ??? folgt

$$\begin{aligned} \rho_{(X+Y)} &= \rho_X \star \rho_Y(x) = \int_{\mathbb{R}} \rho_X(y) \rho_Y(x-y) dy \\ &= \int_{\mathbb{R}} \frac{\lambda^r}{\Gamma(r)} y^{r-1} e^{-\lambda y} \frac{\lambda^s}{\Gamma(s)} (x-y)^{s-1} e^{-\lambda(x-y)} dy \\ &= \frac{\lambda^{r+s}}{\Gamma(r)\Gamma(s)} e^{-\lambda x} \int_0^x y^{r-1} (x-y)^{s-1} dy \\ &\stackrel{\text{part Int}}{=} \frac{\lambda^{r+s}}{\Gamma(r)\Gamma(s)} e^{-\lambda x} \left( \underbrace{\left( \frac{1}{r} y^r (x-y)^{s-1} \right)_{y=0}}_{=0} + \frac{s-1}{r} \int_0^x y^r (x-y)^{s-2} dy \right) \\ &\stackrel{\text{Ind}}{=} \frac{\lambda^{r+s}}{\Gamma(r)\Gamma(r)} e^{-\lambda x} \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)} = \frac{\lambda^{r+s}}{\Gamma(r+s)} e^{-\lambda x} \end{aligned} \quad \square$$

Exponentialverteilung sind zudem gedächtnislos

**Lemma IV.4**

Sei  $X \sim \text{Exp}(\lambda)$ , dann gilt

$$\mathbb{P}(X > t) = \mathbb{P}(X > t + s \mid X > s) \quad t, s \geq 0 \quad (*)$$

**■ Beispiel IV.5**

Eine Studentin wartet morgens eine  $\text{Exp}(\frac{1}{s})$  verteilte Zeit  $X$  auf den Bus zur Uni. Die Wahrscheinlichkeit einer Wartezeit  $\geq 5$  Minuten

$$\mathbb{P}(X \geq 5) = e^{-\frac{1}{s}} = e^{-1} \approx 0,37.$$

An einen kalten, stürmischen Frühlingstag hat die Studentin bereits 10 Minuten gewartet. Die Wahrscheinlichkeit mindestens 5 weitere Minuten zu warten ist

$$\mathbb{P}(X \geq 15 \mid X > 10) = \mathbb{P}(X \geq 5) = e^{-1} \approx 0,37.$$

*Beweis* (Lemma 4.4).

$$\begin{aligned} \mathbb{P}(X > t + s \mid X \geq s) &= \frac{\mathbb{P}(X > t + s, X \geq s)}{\mathbb{P}(X \geq s)} \\ &= \frac{\mathbb{P}(X > t + s)}{\mathbb{P}(X \geq s)} \\ &= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda t} = \mathbb{P}(X > t). \end{aligned} \quad \square$$

Hinweis: Man kann sogar zeigen, dass die Exponentialverteilung die einzige absolutstetige Verteilung mit (\*) ist.

## Kapitel V

# Erwartungswerte & Varianz

1. Frage: Beispiel ??? Durchschnittliche Wartezeit? → Erwartungswert  
Wie stark ist die Streuung um den Durchschnitt? → Varianz

## 1. Der Erwartungswert

### Definition V.1 (title)

Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  Wahrscheinlichkeitsraum und  $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  Zufallsvariable. Dann ist

$$\mathbb{E}[X] = \int_{\Omega} X(\omega) \mathbb{P}(\mathrm{d}\omega) = \int_{\mathbb{R}} x \mathbb{P}(X \in \mathrm{d}x)$$

der Erwartungswert von  $X$ .

Hinweis: Der Erwartungswert von  $X$  existiert, genau dann wenn

$$\int_{\Omega} |X(\omega)| \mathbb{P}(\mathrm{d}\omega) < \infty \text{ bzw. } \mathbb{E}[|X|] < \infty$$

d.h. genau dann wenn  $X \in \mathcal{L}^1(\mathbb{P})$ .

Für nichtnegative Zufallsvariablen ist der Erwartungswert immer definiert, wenn wir  $+\infty$  als zulässigen Wert annehmen, was wir in der Folge auch tun.

### ■ Beispiel V.2

$(\Omega, \mathcal{F}, \mathbb{P})$  Wahrscheinlichkeitsraum,  $A \in \mathcal{F}$  und sei  $Y : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  die Indikatorvariable

$$X(\omega) = \mathbb{1}_A(\omega)$$

Dann gilt:  $X \in \mathcal{L}^1(\mathbb{P})$  und

$$\mathbb{E}[X] = \int_{\Omega} \mathbb{1}_A(\omega) \mathbb{P}(\mathrm{d}\omega) = \int_A \mathbb{P}(\mathrm{d}\omega) = \mathbb{P}(A).$$

### Satz V.3

Sei  $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  Zufallsvariable und

$f : \mathbb{R}^n \rightarrow \mathbb{R}$  Borel-messbar.

Dann

$$\mathbb{E}[f(X)] = \int f(X) \mathrm{d}\mathbb{P} = \int_{\Omega} \mathbb{E}(X(\omega)) \mathrm{d}\mathbb{P}(\omega) = \int_{\mathbb{R}^n} f(X) \mathbb{P}(X \in \mathrm{d}x).$$

*Beweis.* Sei  $f(X)$  eine reelle Zufallsvariable. Die Formel folgt direkt auf dem Transformationssatz für Bildmaße

(↗ Schilling MINT 18.1). □

**Satz V.4 (Erwartungswerte bei Existenz einer (Zähl-)dichte)**

Sei  $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  Zufallsvariable und

$f : \mathbb{R}^n \rightarrow \mathbb{R}$  Borel-messbar.

1. diskrete Fall: Ist  $\mathbb{P}_X$  ein Wahrscheinlichkeitsmaß auf  $(\mathbb{Z}, \mathbb{P}(\mathbb{Z}))$  und der Zähldichte  $\rho$ , so

$$\mathbb{E}[f(X)] = \sum_{x \in \mathbb{Z}} f(x) \rho(x).$$

2. stetiger Fall: Besitzt  $\mathbb{P}_X$  eine Dichte  $\rho$  (bzgl Lebesguemaß), so

$$\mathbb{E}[f(X)] = \int_{\mathbb{R}} f(x) \rho(x) \, dx$$

*Beweis.* Klar aus Def. 5.1 und Prop. 5.3. □

■ **Beispiel V.5**

Sei  $X \sim \text{Bin}(n, p)$ . dann gilt

$$\begin{aligned} \mathbb{E}[f(X)] &= \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n \frac{n!}{(n-k)!(k-1)!} p^k (1-p)^{n-k} \\ &= np \sum_{k=1}^n \underbrace{\binom{n-1}{k-1} p^{k-1} (1-p)^{n-1-(k-1)}}_{\text{Zähldichte Bin}(n-1, p) \text{ in } k-1} \\ &= np \end{aligned}$$



# 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] GEORGI, 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

- (absolut)stetig (bzgl. denn Lebesgue-Maß), [6](#)
- (diskrete) Gleichverteilung, [7](#), [12](#)
- (kumulative) Verteilungsfunktion von  $\mathbb{P}$ , [9](#)
- (stetige) Gleichverteilung, [7](#)
- (stochastisch) unabhängig bezüglich  $\mathbb{P}$ , [22](#)
- (wiederholtes) Bernoulliexperiment, [14](#)
- DIRAC-Maß, [5](#)
- DIRAC-Verteilung, [5](#)
- KOLMOGOROVsche Axiome, [4](#)
- POLYA-Verteilung, [20](#)
  
- Baumdiagramm, [19](#)
- bedingte Wahrscheinlichkeit von  $A$  gegeben  $B$ ,  
[18](#)
- Bernoulliverteilung, [15](#)
- Binomialverteilung, [15](#)
  
- Dichte, [6](#)
- Dichtefunktion, [6](#)
  
- Ereignisraum, [3](#)
- Erfolgswahrscheinlichkeit, [15](#)
- Ergebnisraum, [3](#)
- Erwartungswert von  $X$ , [37](#)
- Exponentialverteilung, [6](#), [34](#)
  
- Faltung, [30](#)
  
- Gammaverteilung, [34](#)
- Gammefunktion, [34](#)
- gedächtnislos, [35](#)
- geometrische Verteilung, [33](#)
  
- Hypergeometrische Verteilung, [15](#)
  
- identisch verteilt, [9](#)
  
- messbar, [3](#)
- messbarer Raum, [3](#)
- momenterzeugenden Funktionen, [35](#)
- Multinomialkoeffizient, [14](#)
- Multinomialverteilung, [14](#)
  
- negative Binomialverteilung, [33](#)
  
- Poissonverteilung, [16](#)
- Produktdichten, [13](#)
  
- Quantilfunktion, [11](#)
  
- reelle Zufallsvariablen, [9](#)
  
- Stufenexperimente, [19](#)
  
- unabhängig, [24](#)
- unabhängig bezüglich  $\mathbb{P}$ , [23](#)
  
- verallgemeinerte Inverse, [11](#)
  
- Wahrscheinlichkeitsdichte, [6](#)
- Wahrscheinlichkeitsmaß, [4](#)
- Wahrscheinlichkeitsraum, [7](#)
- Wahrscheinlichkeitsverteilung, [4](#)
- Wahrscheinlichkeitsverteilung von  $X$  unter  $\mathbb{P}$ , [8](#)
  
- Zähldichte, [6](#)
- Zufallselement, [8](#)