

# Algorithmische Mathematik II

Maximilian Keßler

26. April 2021

## Inhaltsverzeichnis

<b>1</b>	<b>Diskrete Stochastik</b>	<b>2</b>
1.1	Einleitung . . . . .	2
1.2	Ereignisse und Wahrscheinlichkeiten . . . . .	4
1.3	Diskrete Verteilungen . . . . .	9
1.4	Die Gleichverteilung . . . . .	12
1.5	Die empirische Verteilung . . . . .	14
1.6	Zufallsvariablen . . . . .	14
1.6.1	Die Bernoulli-Verteilung . . . . .	16
1.6.2	Die Binomial-Verteilung . . . . .	17
1.6.3	Die Poisson-Verteilung . . . . .	20
1.6.4	Die geometrische Verteilung . . . . .	20
1.7	Simulation von Gleichverteilung . . . . .	21
1.7.1	Lineare Kongruenzgeneratoren (LCG) . . . . .	21
1.7.2	Zufallsvariablen aus $[0, 1)$ . . . . .	21
1.7.3	Zufallspermutationen . . . . .	22
1.7.4	Geometrische Verteilung . . . . .	22
1.8	Erwartungswert und Varianz . . . . .	22

## Lecture 1: Grundbegriffe

Mo 12 Apr 2021 10:16

- Es gibt ein Helpdesk, auch explizit für Studentinnen
- die Vorlesung wird aufgenommen, und zwar ohne Videos der Teilnehmenden sowie des Dozenten, die Aufzeichnung werden anschließend in Sciebo hochgeladen.
- Es gibt ein Diskussionsforum für Fragen (auf eCampus).
- Ab heute Abend, 18 Uhr (Mo 12 Apr 2021 18:00), kann man sich auf eCampus für die Übungsgruppen registrieren und endet am Dienstag Abend um 24 Uhr (Di 12 Apr 2021 24:00), es wird versucht, die Studenten gleichmäßig zu verteilen.
- Falls ihr in der Warteliste landet und gewünscht ist, in der Gruppe abzugeben, schreibt eine Mail mit den gewünschten Abgabepartner, dann kann eine gemeinsame Einteilung erfolgen.

- Es gibt auch das Modul **AlmaITb**. Registriert euch noch nicht, dies ist für den 2. Teil der Vorlesung notwendig.
- Die Abgabe der Übungsblätter erfolgt einheitlich jeden Freitag um 12 Uhr.
- Gruppenabgaben sind erlaubt, bis zu einer Größe von maximal 4 StudentInnen.
- Das 1. Blatt ist freiwillig und gibt Bonuspunkte.
- Für die Klausurzulassung werden 50% der Punkte benötigt. Von den Programmieraufgaben müssen mindestens 4 von 6 zufriedenstellend bearbeitet werden.
- Programmieraufgaben gibt es ab dem 2. Übungsblatt auf jedem 2. Blatt. Die Bearbeitungszeit beträgt dann 2 Wochen.

## Einleitung

In der Vorlesung werden wir sehen:

**Teil 1: Diskrete Stochastik** • Zufallsvariablen

- Bedingte Wahrscheinlichkeiten
- Unabhängigkeit von Variablen
- Monte-Carlo Methoden

**Teil 2: Numerische Analysis** • Iterative Verfahren

- Interpolation von Daten (durch Polynome, trigonometrische Funktionen, ...)
- Numerische Verfahren für die Integration

## 1 Diskrete Stochastik

### 1.1 Einleitung

**Ziel.** Beschreibung von Systemen, die einen Anteil an **Zufall** haben, d.h. nicht 100% deterministisch sind.

- Example.** • Spiele: Kartenspiele, Glücksspiele, ...
- Statistik: Umfragen, Versicherung
  - Komplexe Systeme: Wettermodelle, Finanzmärkte

◇

Was sind Quellen von Zufall?

- Zu komplexe Systeme. Dann sieht der Gesamteffekt zufällig aus.
- Fehlende Informationen (z.B. bei einem Kartenspiel)

- Chaotische Systeme (Wetter)
- Intrinsisch unvorhersagbare Systeme (z.B. radioaktiver Zerfall)

**Question.** (1) Wie modelliert man ein System mit Zufall?

- (2) Wie simuliert man ein System mit Zufall? (anwendungstechnischer)
- (3) Welche Voraussagen kann man machen?

**Example.** Die **Brown'sche Bewegung**. Das System ist implizit ein Pollen mit vielen Wassermolekülen ( $\sim 10^{23}$ ), die sich im Prinzip deterministisch bewegen.

$\Rightarrow$  Wir erhalten ein Gleichungssystem mit  $(N+1) \cdot 6$  (3 Positionen, 3 Geschwindigkeit) Variablen. Dieses ist de facto unlösbar.

Was wollen wir hier eigentlich untersuchen?  $\rightarrow$  Die Bewegung des Pollens, jedoch nicht die der einzelnen Wassermoleküle.

In einer **Modellierung** ersetzt man die Stöße, die durch die Wassermoleküle entstehen durch **zufällige Stöße**.  $\diamond$

Diskretes Modell: Die Zeit bewegt sich in  $n \in \{0, 1, 2, \dots\}$ . Sei

$$Z(n) := (\text{Position des Pollens zur Zeit } n) \in \mathbb{Z}^3.$$

OBdA setzen wir  $Z(0) = 0$ .

Dynamik:  $Z(n+1) = Z(n) + \xi_n$ , wobei wir  $\xi_n$  aus dem Ergebnis eines Würfelwurfs bestimmen werden:

$$\xi_n = \begin{cases} (1, 0, 0) & \text{wenn Würfel} = 1 \\ (-1, 0, 0) & \text{wenn Würfel} = 2 \\ (0, 1, 0) & \text{wenn Würfel} = 3 \\ (0, -1, 0) & \text{wenn Würfel} = 4 \\ (0, 0, 1) & \text{wenn Würfel} = 5 \\ (0, 0, -1) & \text{wenn Würfel} = 6 \end{cases}.$$

**Question.** Welche Fragen können wir mit solch einem System nun beantworten? Was passiert, wenn  $n \gg 1$ .

- (a) Typischerweise erhalten wir  $|Z(n)| = O(\sqrt{n})$
- (b) Wenn wir die Frequenz von  $[Z(n)]_i$  betrachten, sehen wir typischerweise: (Füge Graph mit grober Binomialverteilung ein)  
Für  $n \gg 1$  sieht diese Verteilung dann ungefähr wie die Gaussglocke aus.

Skalierung: Wir setzen nun

$$B(t) = \lim_{n \rightarrow \infty} \frac{Z(\lfloor nt \rfloor)}{\sqrt{n}}.$$

und dies ist dann die Brownsche Bewegung.

- Question.**
- Ist  $Z(n)$  in einer gegebenen Menge  $A$ ?  
-> Im Allgemeinen kann man das nicht mit 'Ja' oder 'Nein' beantworten.
  - Wenn man  $Z(n)$  beobachtet, wie häufig wird  $Z(n)$  in  $A$  sein?  
-> Diese Frage lässt sich mit einer Zahl  $\in [0, 1]$  beantworten.

## 1.2 Ereignisse und Wahrscheinlichkeiten

Wir benötigen 3 Grundelemente:

- (1) Die Menge  $\Omega$  von möglichen **Ergebnissen**. die Elemente von  $\Omega$  heißen auch **Elementarereignisse**.
- (2) Die Menge  $\mathcal{F}$  der **Ereignisse**. Ein Ereignis  $E$  ist eine Eigenschaft, die an einer Teilmenge von  $G \subset \Omega$  assoziiert ist:  $\omega \in G \Leftrightarrow$  Eigenschaft  $E$  ist erfüllt.
- (3) Eine **Wahrscheinlichkeitsverteilung (auch W-maß)**:

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

**Remark.** Wir werden noch sehen, dass gewisse Dinge für unsere Begriffe erfüllt sein müssen, dazu aber später mehr.

**Example.** Eine Urne hat 12 nummerierte Kugeln (von 1 bis 12).

- (1) Das Zufallsexperiment besteht daraus, dass wir eine Kugel aus der Urne ziehen und die Zahl notieren, die wir sehen. D.h.

$$\Omega = \{1, \dots, 12\}.$$

Ein Elementarereignis ist nun z.B. gegeben durch  $\omega = \{5\} \equiv 5$  (wir vereinfachen die Notation).

- (2) Mögliche Ereignisse sind z.B:

$$\begin{aligned} A &= \text{"Die Zahl ist gerade"} \\ B &= \text{"Die Zahl ist } \leq 5 \\ C &= \text{"Die Zahl ist 8"} \end{aligned} \tag{1}$$

Die assoziierten Mengen sind dann

$$\begin{aligned} A &= \{2, 4, 6, 8, 10, 12\} \\ B &= \{1, 2, 3, 4, 5\} \\ C &= \{8\} \end{aligned} \tag{2}$$

- (3) Für die Wahrscheinlichkeiten nehmen wir an, dass jede Kugel die

gleiche Chance hat, gezogen zu werden, d.h.

$$\forall G \in \mathcal{F} : \mathbb{P}(G) = \frac{|G|}{|\Omega|}.$$

Wir erhalten als

$$\mathbb{P}(A) = \frac{6}{12} = \frac{1}{2} \quad \mathbb{P}(B) = \frac{5}{12} \quad \mathbb{P}(C) = \frac{1}{12}.$$

◇

**Notation.**  $A \equiv \{\omega \in \Omega \mid \omega \in A\} \equiv \{\omega \in A\} \equiv \{A \text{ tritt ein}\}$

## Lecture 2: Wahrscheinlichkeitsräume

Mi 14 Apr 2021 10:17

Wir kennen nun die Grundbegriffe  $\Omega, \mathcal{F}, \mathbb{P}$  zur Beschreibung von Zufallsexperimenten, die wir uns nun genauer ansehen wollen:

**Question.** Welche Struktur muss  $\mathcal{F}$  besitzen.

Seien  $A, B \in \mathcal{F}$ , dann können wir das Ereignis  $A \cap B$  betrachten, d.h. beide der Eigenschaften treten ein. Genauso sollte

$$A^c := \Omega \setminus A.$$

, das **Komplement von  $A$** , bzw. das **Gegenereignis** von  $A$  ebenfalls in  $\mathcal{F}$  sein. Aus den beiden vorherigen Eigenschaften folgt bereits, dass

$$A \cup B = (A^c \cap B^c)^c.$$

ebenfalls in  $\mathcal{F}$  sein wird. Eine Menge  $\mathcal{F}$  mit solchen Eigenschaften heißt **Algebra**, d.h. wir fordern von  $\mathcal{F}$ , dass es sich um eine solche Algebra handelt.

Seien nun  $A, B, (A_i)_{i \in I}$  Ereignisse, wobei  $I$  endlich oder abzählbar sei. Dann notieren wir folgendermaßen

- (a)  $A \cup B$ :  $\omega \in A \cup B \Leftrightarrow \omega \in A \vee \omega \in B$ , d.h.  $A \cup B$  tritt ein, genau dann, wenn  $A$  eintritt oder  $B$  eintritt
- (b)  $\bigcup_{i \in I} A_i$ :  $\omega \in \bigcup_{i \in I} A_i$ , wenn es ein  $i \in I$  gibt, sodass  $\omega \in A_i$
- (c)  $A \cap B$ :  $\omega \in A \cap B \Leftrightarrow A$  und  $B$  treten ein.
- (d)  $\bigcap_{i \in I} A_i$ :  $\omega \in \bigcap_{i \in I} A_i \Leftrightarrow \forall i \in I$ :  $A_i$  tritt ein.
- (e)  $A = \emptyset$  ist das Ereignis, das nie eintritt.  
 $A = \Omega$  ist das Ereignis, dass immer eintritt.

**Definition 1.1.** Sei  $\mathcal{F}$  eine nicht leere Menge von Teilmengen von  $\Omega$  mit den Eigenschaften:

- (a)  $\Omega \in \mathcal{F}$

- ⓑ  $\forall A \in \mathcal{F}: A^c \in \mathcal{F}$ .
- ⓒ Falls  $(A_i)_{i \in I} \in \mathcal{F}$ , dann auch  $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

Dann nennen wir  $\mathcal{F}$  eine  **$\sigma$ -Algebra** und das Paar  $(\Omega, \mathcal{F})$  einen **Messraum**.

**Lemma 1.2.** Sei  $\mathcal{F}$  eine  $\sigma$ -Algebra, dann ist:

- ⓐ  $\emptyset \in \mathcal{F}$
- ⓑ  $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$  und  $A \cap B \in \mathcal{F}$ .
- ⓒ  $(A_i)_{i \in I} \in \mathcal{F} \Rightarrow \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$ .

*Beweis.* ⓐ  $\emptyset = \Omega^c \in \mathcal{F}$  nach Eigenschaften ⓐ und ⓑ aus der Definition.

ⓑ  $A \cup B = A \cup B \cup \emptyset \cup \emptyset \dots \in \mathcal{F}$  nach Eigenschaften ⓑ und ⓒ.  $A \cap B = (A^c \cup B^c)^c \in \mathcal{F}$

ⓒ  $\bigcap_{i=1}^{\infty} A_i = (\bigcup_{i=1}^{\infty} A_i^c)^c \in \mathcal{F}$  nach ⓑ und ⓒ.

□

Wir haben nun  $(\Omega, \mathcal{F})$  näher untersucht, es fehlt nun noch  $\mathbb{P}$ .

Seien  $A, B \in \mathcal{F}$  mit  $A \cap B = \emptyset$ , d.h.  $A$  und  $B$  können nicht gleichzeitig eintreten. Dann fordern wir

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) \quad (\text{endliche Additivität}).$$

Dazu wollen wir, dass  $\Omega \in \mathcal{F}$  immer eintritt, d.h.  $\mathbb{P}(\Omega) = 1 \equiv 100\%$  (Normierung).

**Definition 1.3.** Sei  $(\Omega, \mathcal{F})$  ein Messraum. Eine Abbildung  $\mathbb{P} : \mathcal{F} \rightarrow \mathbb{R}_+$  ist eine **Wahrscheinlichkeitsverteilung auf  $(\Omega, \mathcal{F})$** , falls

- (1)  $\mathbb{P}(\Omega) = 1$
- (2)  $(A_i)_{i \in I} \in \mathcal{F}$  mit  $A_i \cap A_j = \emptyset$  für  $i \neq j$ , dann ist:

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) \quad (\sigma\text{-Additivität}).$$

**Definition 1.4.** Ein **Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$**  besteht aus einer Menge  $\Omega$ , einer  $\sigma$ -Algebra  $\mathcal{F} \subset \mathcal{P}(\Omega)$  und einem Wahrscheinlichkeitsmass  $\mathbb{P}$  auf  $(\Omega, \mathcal{F})$

**Lemma 1.5.** Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  ein Wahrscheinlichkeitsraum. Dann ist

(a)  $\mathbb{P}(\emptyset) = 0$

(b)  $\forall A, B \in \mathcal{F}$  mit  $A \cap B = \emptyset$  ist

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

(c)  $\forall A, B \in \mathcal{F}$  mit  $A \subset B$  ist

$$\begin{aligned}\mathbb{P}(B) &= \mathbb{P}(A) + \mathbb{P}(B \setminus A) \\ \mathbb{P}(A^c) &= 1 - \mathbb{P}(A) \\ \mathbb{P}(A) &\leq \mathbb{P}(B) \leq 1\end{aligned}\tag{3}$$

(d)  $\forall A, B \in \mathcal{F}$  ist

$$\begin{aligned}\mathbb{P}(A \cup B) &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \\ &\leq \mathbb{P}(A) + \mathbb{P}(B)\end{aligned}\tag{4}$$

(e) Wenn  $A_n$

*Beweis.* (a) Wir wissen:

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(\Omega \cup \emptyset \cup \emptyset \cup \emptyset \dots) = \mathbb{P}(\Omega) + \mathbb{P}(\emptyset) + \mathbb{P}(\emptyset) + \dots$$

subtrahieren von  $\mathbb{P}(\Omega) = 1$  liefert dann  $\mathbb{P}(\emptyset) = 0$ .

(b) Sei  $A \cap B = \emptyset$ , dann ist:

$$\begin{aligned}\mathbb{P}(A \cup B) &= \mathbb{P}(A \cup B \cup \emptyset \cup \emptyset \dots) \\ &\stackrel{\sigma\text{-Additivität}}{=} \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(\emptyset) + \mathbb{P}(\emptyset) + \dots \\ &= \mathbb{P}(A) + \mathbb{P}(B)\end{aligned}\tag{5}$$

(c) Sei  $A \subset B$ . Dann ist  $B = A \cup (B \setminus A)$  eine disjunkte Vereinigung, also erhalten wir

$$\mathbb{P}(B) = \mathbb{P}(A) + \underbrace{\mathbb{P}(B \setminus A)}_{\geq 0} \geq \mathbb{P}(A).$$

Mit  $B = \Omega$  ergibt sich  $1 = \mathbb{P}(A) + \mathbb{P}(A^c)$

(d) Es ist

$$\begin{aligned}\mathbb{P}(A \cup B) &= \mathbb{P}(A) + \mathbb{P}((A \cup B) \setminus A) \\ &= \mathbb{P}(A) + \mathbb{P}(B \setminus (A \cap B)) \\ &= \mathbb{P}(A) + \mathbb{P}(B) - \underbrace{\mathbb{P}(A \cap B)}_{\geq 0} \\ &\geq \mathbb{P}(A) + \mathbb{P}(B)\end{aligned}\tag{6}$$

(e) Übung

□

**Corollary 1.6** (Einschluss-Ausschluss-Prinzip). Seien  $A_1, \dots, A_n \in \mathcal{F}$ . Dann gilt

$$\mathbb{P}(A_1 \cup \dots \cup A_n) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}).$$

*Beweis.* Per Induktion, der Induktionsanfang lautet  $\mathbb{P}(A_1) = \mathbb{P}(A_1)$  und ist offensichtlich wahr.

Die Aussage gelte nun für ein  $n \in \mathbb{N}$ , dann erhalten wir

$$\begin{aligned} \mathbb{P}\left(\bigcup_{i=1}^{n+1} A_i\right) &= \mathbb{P}\left(\left(\bigcup_{i=1}^n A_i\right) \cup A_{n+1}\right) \\ &= \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) + \mathbb{P}(A_{n+1}) - \mathbb{P}\left(\left(\bigcup_{i=1}^n A_i\right) \cap A_{n+1}\right) \\ &= \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) + \mathbb{P}(A_{n+1}) - \mathbb{P}\left(\bigcup_{i=1}^n \underbrace{(A_i \cap A_{n+1})}_{=: \tilde{A}_i}\right) \\ &= \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{P}(A_{i_1} \cap \dots \cap A_{i_k}) + \mathbb{P}(A_{n+1}) \\ &\quad - \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{P}\left(\underbrace{\tilde{A}_{i_1} \cap \dots \cap \tilde{A}_{i_k}}_{A_{i_1} \cap \dots \cap A_{i_k} \cap A_{n+1}}\right) \end{aligned} \quad (7)$$

Andererseits ist aber auch:

$$\begin{aligned} &\sum_{k=1}^{n+1} (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n+1} \mathbb{P}(A_{i_1} \cap \dots \cap A_{i_k}) \\ &= \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq \textcolor{red}{n}} \mathbb{P}(A_{i_1} \cap \dots \cap A_{i_k}) \quad \left. \vphantom{\sum_{k=1}^n} \right\} \text{Terme mit } i_k \leq n \\ &+ \underbrace{\sum_{k=2}^{n+1} (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_{k-1} \leq n} \mathbb{P}(A_{i_1} \cap \dots \cap A_{i_{k-1}} \cap A_{n+1})}_{\stackrel{l:=k-1}{=} \sum_{l=1}^n (-1)^{l-1} \sum_{1 \leq i_1 < \dots < i_l \leq n} \mathbb{P}(A_{i_1} \cap \dots \cap A_{i_l} \cap A_{n+1})} \quad \left. \vphantom{\sum_{k=2}^{n+1}} \right\} \text{Terme mit } i_k = n+1 \text{ und } k \geq 2 \\ &+ \mathbb{P}(A_{n+1}) \quad \left. \vphantom{\mathbb{P}(A_{n+1})} \right\} \text{Terme mit } i_k = n+1 \text{ und } k=1 \end{aligned}$$

und damit sehen wir, dass die beiden Ausdrücke übereinstimmen, also ist der Induktionsschritt erbracht. □



### 1.3 Diskrete Verteilungen

- Sei nun  $\Omega$  endlich oder abzählbar.
- Falls wir  $\mathcal{F}$  nicht explizit angeben, dann wird  $\mathcal{F} = \mathcal{P}(\Omega)$  gewählt, d.h.

$$\text{Card}(\mathcal{P}(\Omega)) \equiv |\mathcal{P}(\Omega)| = 2^{|\Omega|}.$$

**Example** (Münzwurf). Es sei  $\Omega = \{K, Z\}$ , wobei  $K$  für Kopf stehe und  $Z$  für Zahl. Dann ist

$$\mathcal{F} = \{\{K\}, \{Z\}, \{Z, K\}, \emptyset\}.$$

Sei  $p \in [0, 1]$  die Wahrscheinlichkeit, dass man Kopf erhält. Da  $\mathbb{P}$  für alle Element aus  $\mathcal{F}$  definiert sein muss, erhalten wir

$$\mathbb{P}(\emptyset) = 0 \quad \mathbb{P}(K) = p, \quad \mathbb{P}(Z) = \mathbb{P}(K^c) = 1-p \quad \mathbb{P}(\{Z, K\}) = \mathbb{P}(\Omega) = 1.$$

◇

Was müssen wir fordern, sodass es ein wohldefiniertes  $\mathbb{P}$  auf  $\mathcal{P}(\Omega)$  gibt?.

**Example.**  $\Omega = \{1, 2, \dots, 10\}$  würde genügen, da dann  $|\mathcal{P}(\Omega)| = 2^{|\Omega|} = 2^{10} = 1024$  endlich (diskret) ist. ◇

## Lecture 3: Gleichverteilung, empirische Verteilung

Mo 19 Apr 2021 10:23

Wir stellen fest, dass es im letzten Beispiel auch genügt hätte,  $\mathbb{P}(\{k\})$  für  $k = 1, \dots, 10$  anzugeben, das motiviert Folgendes:

**Theorem 1.7.** (a) Sei  $p(\omega) \in [0, 1], \omega \in \Omega$ , sodass

$$\sum_{\omega \in \Omega} p(\omega) = 1.$$

Dann ist  $\mathbb{P}$  definiert durch:

$$\mathbb{P} : \begin{cases} \mathcal{P}(\Omega) & \longrightarrow [0, 1] \\ A & \longmapsto \sum_{\omega \in A} p(\omega) \end{cases}$$

eine Wahrscheinlichkeitsverteilung auf  $(\Omega, \mathcal{P}(\Omega))$ .

- (b) Jede Wahrscheinlichkeitsverteilung  $\mathbb{P}$  auf  $(\Omega, \mathcal{P}(\Omega))$  hat obige Form, wobei  $p(\omega) = \mathbb{P}(\{\omega\})$ .

**Remark.**  $p : \Omega \rightarrow [0, 1]$  heißt Massenfunktion der Wahrscheinlichkeitsverteilung  $\mathbb{P}$ .

**Warning.** Der Satz gilt nicht für  $\Omega$  überabzählbar.

**Remark.** Sei  $A$  abzählbar und  $p(\omega) \geq 0$  für  $\omega \in A$ . Dann definiert

$$\sum_{\omega \in A} p(\omega) := \sum_{k \geq 1} p(\omega_k).$$

mit einer beliebigen Abzählung  $\omega_1, \omega_2, \dots, \omega_k$  von  $A$  eine wohldefinierte Summe der  $p(\omega)$ . Es ist wichtig, dass hier  $p(\omega) \geq 0$ , sonst ist obiges nicht wohldefiniert.

**Lemma 1.8.** (a) Sei  $p(\omega) \in [0, 1]$  für alle  $\omega$ . Dann ist

$$\sum_{\omega \in A} p(\omega) \in [0, \infty].$$

wohldefiniert. Setzen wir

$$\mathbb{P}(A) := \sum_{\omega \in A} p(\omega).$$

so gilt

$$P(A) = \sup_{\substack{F \subset A \\ |F| < \infty}} P(F).$$

und  $P(A) \leq P(B)$  für  $A \subset B$ .

(b) Ist  $A = \bigsqcup_{k=1}^{\infty} A_k$  eine disjunkte Vereinigung, so ist

$$P(A) = \sum_{k=1}^{\infty} P(A_k).$$

*Beweis.* (a) Sei  $\omega_1, \omega_2, \dots$  eine beliebige Abzählung von  $A$ . Dann ist die Funktion

$$n \mapsto \sum_{k=1}^n p(\omega_k).$$

monoton wachsend. Also ist

$$\sum_{k=1}^{\infty} p(\omega_k) := \lim_{n \rightarrow \infty} \sum_{k=1}^n p(\omega_k) = \sup_{n \in \mathbb{N}} \sum_{k=1}^n p(\omega_k) \in [0, \infty].$$

wohldefiniert.

Wir wollen nun noch zeigen, dass

$$P(A) := \sum_{k=1}^{\infty} p(\omega_k) \stackrel{!}{=} \sup_{\substack{F \subset A \\ |F| < \infty}} P(F).$$

Die Ungleichung ' $\leq$ ' folgt sofort, da wir mit  $F_n := \{\omega_1, \dots, \omega_n\}$  feststellen, dass

$$\sum_{k=1}^n p(\omega_k) = \sum_{\omega \in F_n} p(\omega) = P(F_n) \leq \sup_{\substack{F \subset A \\ |F| < \infty}} P(F).$$

Also ergibt sich im Limes genau wie gewünscht

$$P(A) = \lim_{n \rightarrow \infty} \sum_{k=1}^n p(\omega_k) \leq \lim_{n \rightarrow \infty} \sup_{\substack{F \subset A \\ |F| < \infty}} P(F) = \sup_{\substack{F \subset A \\ |F| < \infty}} P(F).$$

Für ' $\geq$ ' stellen wir fest, dass es für jedes  $F \subset A$  endlich ein  $n \in \mathbb{N}$  gibt, sodass  $F \subset \{\omega_1, \dots, \omega_n\}$ , und somit ist

$$P(F) = \sum_{\omega \in F} P(\omega) \leq \sum_{k=1}^n p(\omega_k) \leq \sum_{k=1}^{\infty} p(\omega_k) = P(A).$$

und somit ist das Supremum der  $P(F)$  für  $F \subset A, |F| < \infty$  durch  $P(A)$  beschränkt.

Für die letzte Behauptung sehen wir mit  $A \subset B$  leicht, dass

$$P(A) = \sup_{\substack{F \subset A \\ |F| < \infty}} P(F) \leq \sup_{F \subset B, |F| < \infty} P(F) = P(B).$$

(b) ( $\sigma$ -Additivität) Wir unterscheiden zwei Fälle:

1) Falls  $|A| < \infty$ , so ist  $A = \bigsqcup_{k=1}^n A_k$  für ein  $n$ , und somit ist

$$\begin{aligned} P(A) &= \sum_{l=1}^{|A|} p(\omega_l) = \sum_{l=1}^{|A|} \sum_{k=1}^n p(\omega_l) \mathbb{1}_{A_k}(\omega_l) \\ &= \sum_{k=1}^n \sum_{l=1}^{|A|} p(\omega_l) \mathbb{1}_{A_k}(\omega_l) = \sum_{k=1}^n P(A_k) \end{aligned} \tag{8}$$

2) Sei nun  $|A| = \infty$ . Wir zeigen zunächst ' $\leq$ '. Für ein endliches  $F \subset A$  ist

$$F = \bigcup_{k=1}^{\infty} (F \cap A_k).$$

eine disjunkte Vereinigung mit endlich vielen Termen, also ist

$$P(F) = \sum_{k=1}^{\infty} P(F \cap A_k) \leq \sum_{k=1}^{\infty} P(A_k).$$

und somit liefert das Supremum über beide Seiten, dass

$$P(A) = \sup_{\substack{F \subset A \\ |F| < \infty}} P(F) \leq \sum_{k=1}^{\infty} P(A_k).$$

Wir zeigen nun ' $\geq$ '.

**Idea.** Wir können  $P(A_k) = \sup_{\substack{F_k \subset A_k \\ |F_k| \leq \infty}} P(F_k)$  schreiben und 'optimieren' nun jedes einzelne  $F_k$ .

Seien also  $F_k \subset A_k$  jeweils endlich. Dann ist  $F_k \cap F_l \subset A_k \cap A_l = \emptyset$ , also sind auch die  $F_k$  paarweise disjunkt, und wir lernen

$$\sum_{k=1}^n P(A_k) = \sum_{k=1}^n \sup_{\substack{F_k \subset A_k \\ |F_k| < \infty}} P(F_k) = \sup_{\substack{F_1 \subset A_1 \\ \text{abs } F_1 < \infty}} \dots \sup_{\substack{F_k \subset A_k \\ |F_k| < \infty}} \sum_{k=1}^n P(F_k) \quad (9)$$

Also ist

$$\sum_{k=1}^n P(F_k) = P\left(\bigcup_{k=1}^n F_k\right) \leq P\left(\bigcup_{k=1}^{\infty} F_k\right) \leq P\left(\bigcup_{k=1}^{\infty} A_k\right) \stackrel{\text{def}}{=} P(A).$$

setzen wir dies nun in die rechte Seite von (1) ein, so ergibt sich

$$\sum_{k=1}^n P(A_k) \leq P(A) \quad \Rightarrow \quad \sum_{k=1}^{\infty} P(A_k) \leq P(A).$$

□

*Beweis von Satz 1.7.* (a) Es gilt

$$\sum_{\omega \in \Omega} p(\omega) = P(\Omega) = 1.$$

nach Voraussetzung. Die  $\sigma$ -Additivität folgt nun aus Lemma 1.8. Deswegen ist  $P(A)$  eine Wahrscheinlichkeitsverteilung.

(b) Da  $P$   $\sigma$ -additiv ist, ist  $\forall A \subset \Omega$ :

$$P(A) = P\left(\bigcup_{\omega \in A} \{\omega\}\right) = \sum_{\omega \in A} P(\{\omega\}).$$

und dies hat genau die angegebene Form mit  $p(\omega) := P(\{\omega\})$

□

## 1.4 Die Gleichverteilung

Sei  $\Omega$  endlich ( $\neq \emptyset$ ) und betrachte  $\sigma$ -Algebra  $\mathcal{F} = \mathcal{P}(\Omega)$ .

Die **Gleichverteilung** ist die Wahrscheinlichkeitsverteilung, die ein uniformes "Gewicht" (Massenfunktion) auf die Elementarereignisse verteilt:

$$\forall \omega \in \Omega : p(\omega) = \mathbb{P}(\{\omega\}) = \frac{1}{|\Omega|}.$$

Aus Satz 7 folgt dann bereits, dass  $\forall A \subset \Omega$ :

$$\mathbb{P}(A) = \sum_{\omega \in A} p(\omega) = \frac{|A|}{|\Omega|}.$$

**Example.** (a) Betrachte  $n$  Würfe eines fairen Würfels. In diesem Fall ist  $\Omega = \{1, 2, 3, 4, 5, 6\}^n = \{\omega = (\omega_1, \dots, \omega_n) \mid \omega_k \in \{1, \dots, 6\}\}$  und somit  $|\Omega| = 6^n$  und die Gleichverteilung ist gegeben durch

$$\mathbb{P}(\omega) = \frac{1}{6^n}.$$

(b) (Zufällige Permutationen).

- Eine Permutation  $\sigma \in \mathfrak{S}_n$  von  $\{1, \dots, n\}$  ist eine Abbildung von  $\{1, \dots, n\}$  nach  $\{1, \dots, n\}$ , die bijektiv ist. Oft schreiben wir

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}.$$

und meinen damit  $\sigma(1) = 4, \sigma(2) = 3, \sigma(3) = 1, \sigma(4) = 2$ . Manchmal schreiben wir dann auch

$$\sigma = (4, 3, 1, 2) = (\sigma_1, \sigma_2, \sigma_3, \sigma_4).$$

- Sei  $\Omega = \mathfrak{S}_n$  die Menge aller Permutation von  $\{1, \dots, n\}$ . Dann ergibt sich

$$|\mathfrak{S}_n| = n!.$$

Also ergibt sich für die Gleichverteilung eine Wahrscheinlichkeit von

$$\mathbb{P}(\sigma) = \frac{1}{n!} \quad \forall \sigma \in \mathfrak{S}_n.$$

◇

**Exercise.** Sei  $N$  die Anzahl von Karten eines Kartenspiels, die gut gemischt sind, d.h. jede Reihenfolge ist gleich wahrscheinlich.

- (1) Was ist die Wahrscheinlichkeit, dass die  $k$ -te Karte auf der  $l$ . Stelle ist?  
D.h., was ist:

$$\mathbb{P}(\{\omega \in \mathfrak{S}_n \mid \omega(k) = l\}).$$

Es ergibt sich

$$\mathbb{P}(\{\omega \in \mathfrak{S}_n \mid \omega(k) = l\}) = \frac{|\{\omega \in \mathfrak{S}_n \mid \omega(k) = l\}|}{|\Omega|} = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

- (2) Was ist die Wahrscheinlichkeit, dass eine Karte 'auf ihrer Stelle' ist, dh.

$$\mathbb{P}(\{\omega \mid \exists k: \omega(k) = k\}).$$

Definiere die Ereignisse  $A_k := \{\omega(k) = k\}$ . Diese sind nicht disjunkt für

verschiedene  $k$ . Es ergibt sich:

$$\begin{aligned}
 \mathbb{P}(\exists k: \omega(k) = k) &= \mathbb{P}\left(\bigcup_{k=1}^n A_k\right) \\
 &= \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \underbrace{\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k})}_{= \frac{(n-k)!}{n!}} \\
 &= \sum_{k=1}^n (-1)^{k-1} \frac{(n-k)!}{n!} \underbrace{\sum_{1 \leq i_1 < \dots < i_k \leq n} 1}_{= \binom{n}{k}} \\
 &= \sum_{k=1}^n (-1)^{k-1} \frac{(n-k)!}{n!} \cdot \frac{n!}{(n-k)!k!} \\
 &= - \sum_{k=1}^n \frac{(-1)^k}{k!} \\
 &= 1 - \frac{1}{e} + \sum_{k=n+1}^{\infty} \frac{(-1)^k}{k!}
 \end{aligned} \tag{10}$$

Für  $n \rightarrow \infty$  geht das gegen  $1 - \frac{1}{e} \in (0, 1)$ .

## 1.5 Die empirische Verteilung

Diese wird aus den Beobachtungen definiert. Seien  $x_1, x_2, \dots, x_n \in \Omega$   $n$  Beobachtungen. Setze

$$N(A) := |\{k \in \{1, \dots, n\} \mid x_k \in A\}|.$$

Dazu setzen wir

$$\mathbb{P}(A) = \frac{N(A)}{n}.$$

, die **empirische Häufigkeit** von  $A$ .  $\mathbb{P}$  ist die **empirische Verteilung**.

$$p(\omega) = \frac{N(\{\omega\})}{n}.$$

ist die **relative Häufigkeit** von  $\omega \in \Omega$ .

**Example.** Die empirische Verteilung von  $n$  Zufallswürfeln eines Würfels wird gegeben durch  $x_1, \dots, x_n \in \{1, \dots, 6\}$ . Die Plots für  $p_k := \frac{N(k)}{n}$  für verschieden  $n$  sehen wie folgt aus:  $\diamond$

## Lecture 4

Mi 21 Apr 2021 10:15

## 1.6 Zufallsvariablen

Wir werden Funktionen der Ergebnisse betrachten:

**Definition 1.9.** Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  ein Wahrscheinlichkeitsraum. Eine **diskrete Zufallsvariable** ist eine messbare Abbildung

$$X : \Omega \longrightarrow \mathcal{S}.$$

mit  $\mathcal{S}$  abzählbar (denke: 'diskret').

Messbar bedeutet hierbei, dass

$$\forall s \in \mathcal{S} : X^{-1}(s) = \{\omega \in \Omega \mid X(\omega) = s\} \in \mathcal{F}.$$

**Notation.** Wir schreiben auch kurz:

$$X^{-1}(s) = \{X(\omega) = s\} = \{X = s\}.$$

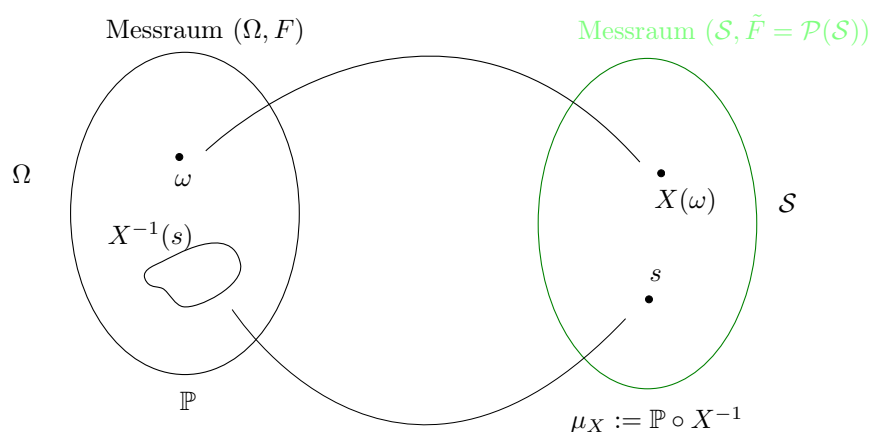


Abbildung 1: Diskrete Zufallsvariable

**Definition 1.10.** Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  ein Wahrscheinlichkeitsraum, und  $X : \Omega \rightarrow \mathcal{S}$  eine diskrete Zufallsvariable.

Die **Verteilung von  $X$**  ist die Wahrscheinlichkeitsverteilung  $\mu_X$  auf  $\mathcal{S}$ ,  $\mathcal{P}(\mathcal{S}) = \tilde{\mathcal{F}}$ , s.d.  $\forall B \in \tilde{\mathcal{F}} : \mu_X(B) := \mathbb{P}(X^{-1}(B))$ .

$\mu_X$  hat eine **Massenfunktion**

$$p_X(s) := \mathbb{P}.$$

**Example** (Werfen von  $n$  Münzen). Betrachte folgende Situation:

- Sei  $\Omega = \{\omega = (\omega_1, \dots, \omega_n \mid \omega_i \in \{0, 1\} \text{ für } 1 \leq i \leq n\}$  wobei

$$\omega_k = \begin{cases} 0 & \text{falls } k\text{-ter Wurf ist Zahl} \\ 1 & \text{falls } k\text{-ter Wurf ist Kopf} \end{cases}.$$

- $\mathcal{F} = \mathcal{P}(\Omega)$  und  $\mathbb{P}$  die Gleichverteilung

(1) Setze

$$X_k : \begin{cases} \Omega & \longrightarrow \mathcal{S} = \{0, 1\} \\ \omega & \longmapsto \omega_k \end{cases}$$

für  $k = 1, \dots, n$ . Dies ist eine diskrete Zufallsvariable mit Verteilung  $\mu_{X_k}$  mit

$$p_{X_k}(s) = \mathbb{P}(X_k = s) = \frac{2^{n-1}}{2^n} = \frac{1}{2}.$$

Wir sehen also, dass  $X_k$  gleichverteilt ist.

(2) Definiere

$$Y : \begin{cases} \Omega & \longrightarrow \mathcal{S} := \{0, 1, \dots, n\} \\ \omega & \longmapsto \omega_1 + \dots + \omega_n \end{cases}$$

d.h.

$$Y(\omega) = \# \{\text{geworfene Köpfe}\}.$$

Es hat nun  $\mu_Y$  die Massenfunktion:

$$p_Y(k) = \frac{1}{2^n} |\{\omega \mid \omega_1 + \dots + \omega_n = k\}| = \frac{\binom{n}{k}}{2^n}.$$

Diese Verteilung sieht wie folgt aus:

TODO: Binomialverteilung

Diese sind Sonderfälle der **Bernoulli-Verteilung** und der **Binomialverteilung**

◇

### 1.6.1 Die Bernoulli-Verteilung

**Definition 1.11.** Sei  $p \in [0, 1]$  gegeben. Die Wahrscheinlichkeitsverteilung auf  $\{0, 1\}$  mit Massenfunktion

$$p(k) = \begin{cases} p & k = 1 \\ 1 - p & k = 0 \end{cases}.$$

heißt **Bernoulli-Verteilung mit Parameter  $p$** .

**Notation.** Wir notieren auch  $\text{Ber}(p)$  für die Bernoulli-Verteilung mit Parameter  $p$ .



**Example.** (a) Eine Münze, die mit Wahrscheinlichkeit  $p$  Kopf zeigt.  
Hier ist

$$\Omega = \{\text{Zahl}, \text{Kopf}\} \quad \mathbb{P}(\text{Kopf}) = p = 1 - \mathbb{P}(\text{Zahl}).$$

Sei

$$X(\omega) = \begin{cases} 1 & \omega = \text{Kopf} \\ 0 & \omega = \text{Zahl} \end{cases}.$$

Dann ist  $\mathbb{P}(X = 1) = p$  und  $X \sim \text{Ber}(p)$ .

**Notation.** Wir schreiben  $X \sim \text{Ber}(p)$ , wenn  $X$  die Verteilung  $\text{Ber}(p)$  hat.

(b) In einer Urne befinden sich  $n$  blaue Kugeln und  $m$  rote Kugeln. Wir ziehen eine Kugel aus der Urne (Annahme: Gleichverteilung). Dann ist

$$\Omega = \{\omega = (\omega_1, \dots, \omega_{n+m}) \mid \omega_i \in \{\text{blau}, \text{rot}\} \text{ mit } n \text{ mal blau}\}.$$

Setze  $\mathcal{F} = \mathcal{P}(\Omega)$  und wähle  $\mathbb{P}$  als die Gleichverteilung. Betrachte

$$X(\omega) = \begin{cases} 1 & \text{falls } \omega_i \text{ ist blau} \\ 0 & \text{sonst} \end{cases}.$$

Diese hat also die Verteilung

$$\mathbb{P}(X = 1) = \frac{\binom{m+n-1}{n-1}}{\binom{m+n}{n}} = \frac{(m+n-1)!}{(n-1)!m!} \frac{n!m!}{(m+n)!} = \frac{n}{m+n}.$$

Also ist  $X \sim \text{Ber}\left(\frac{n}{n+m}\right)$

◇

### 1.6.2 Die Binomial-Verteilung

**Definition 1.12.** Seien  $n \in \mathbb{N}$  und  $p \in [0, 1]$  gegeben. Die Wahrscheinlichkeitsverteilung auf  $\{0, 1, \dots, n\}$  mit Massenfunktion

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

für  $k = 0, \dots, n$  heißt **Binomialverteilung mit Parametern  $n$  und  $p$** .

**Notation.** Wir notieren  $\text{Bin}(n, p)$  für die Binomialverteilung mit Parametern  $n$  und  $p$ .

**Example** (Ziehen mit Zurücklegen). • Seien  $m$  Kugeln in einer Urne, davon  $p \cdot m \in \mathbb{N}$  weiße Kugeln und  $(1 - p)m$  schwarze Kugeln.

- Wir ziehen eine Kugel, notieren uns die Farbe und legen sie wieder zurück.
- Wir mischen die übrigen Kugeln wieder gut.
- Wir wiederholen die vorherigen Schritte, bis wir  $n$  Ziehungen durchgeführt haben.
- Dies modellieren wir durch

$$\Omega = \{0, 1\}^n.$$

wobei  $\omega = (\omega_1, \dots, \omega_n) \in \Omega$  gegeben ist durch

$$\omega_i = \begin{cases} 1 & \text{falls Farbe der } i\text{-ten Kugel weiß ist} \\ 0 & \text{sonst} \end{cases}.$$

- Sei nun  $X(\omega) = \sum_{k=1}^n \omega_k = \#\{\text{weiße Kugeln}\}$

Dann behaupten wir, dass  $X \sim \text{Bin}(n, p)$ . In der Tat:

$$\begin{aligned} \frac{|\{\omega \in \Omega \mid X(\omega) = l\}|}{|\Omega|} &= \frac{\binom{n}{l} \cdot (pm)^l ((1-p)m)^{n-l}}{m^n} \\ &= \frac{\binom{n}{l} p^l (1-p)^{n-l} \cdot m^n}{m^n} = \binom{n}{l} p^l (1-p)^{n-l} \end{aligned} \quad (11)$$

◇

**Remark.** Wir haben hier den Begriff der **Unabhängigkeit** genutzt, den wir nun genauer kennenlernen wollen.

**Definition 1.13** (Unabhängige Ereignisse). Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  ein Wahrscheinlichkeitsraum. Die Ereignisse  $E_1, E_2, \dots, E_n$  heißen **unabhängig**, falls

$$\mathbb{P}(E_{i_1} \cap E_{i_2} \cap \dots \cap E_{i_k}) = \prod_{l=1}^k \mathbb{P}(E_{i_l}).$$

für alle  $2 \leq k \leq n$  und  $1 \leq i_1 < i_2 < \dots < i_k \leq n$ .

**Example.** • Betrachte zwei Würfelwürfe, d.h.  $\Omega = \{1, \dots, 6\}^2$  und notiere  $\omega = (\omega_1, \omega_2)$ . Dann können wir

$$E_1 = \{\omega_1 = 3\} \quad E_2 = \{\omega_2 \geq 4\}.$$

betrachten. Wir rechnen nach, dass

$$\mathbb{P}(\omega_1 = 3 \cap \omega_2 \geq 4) = \frac{3}{36} = \frac{1}{6} \cdot \frac{3}{6} = \mathbb{P}(\omega_1 = 3) \cdot \mathbb{P}(\omega_2 \geq 4).$$

also sind die beiden Ereignisse unabhängig voneinander. Das macht auch semantisch Sinn, weil wir durch das Ergebnis des einen Würfelwurfs keine Informationen über das Ergebnis des zweiten Würfelwurfs erhalten.

- Falls  $E_1, E_2, \dots, E_n$  unabhängige Ereignisse sind, mit  $\mathbb{P}(E_i) = p$  für  $1 \leq i \leq n$ , dann ist

$$\mathbb{P}(\text{genau } k \text{ der Ereignisse treten ein}) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Dies rechnen wir nach. Setze hierzu

$$A_{(i_1, \dots, i_k)} = \{\omega \in \Omega \mid E_{i_1}, \dots, E_{i_k} \text{ treten ein, die anderen nicht}\}$$

Dann ist

$$\tilde{A} = \bigsqcup_{1 \leq i_1 < \dots < i_k \leq n} A_{i_1, \dots, i_k}.$$

eine disjunkte Vereinigung, also erhalten wir

$$\begin{aligned} \mathbb{P}(\tilde{A}) &= \sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{P}(A_{(i_1, \dots, i_k)}) \\ &\stackrel{\text{unabhängig}}{=} \sum_{1 \leq i_1 < \dots < i_k \leq n} \prod_{j \in \{i_1, \dots, i_k\}} \mathbb{P}(E_j) \cdot \prod_{l \notin \{i_1, \dots, i_k\}} \mathbb{P}(E_l^c) \\ &= \sum_{1 \leq i_1 < \dots < i_k \leq n} p^k (1-p)^{n-k} \\ &= \binom{n}{k} p^k (1-p)^{n-k} \end{aligned} \tag{12}$$

◇

**Remark.** Strenggenommen haben wir in der letzten Rechnung verwendet, dass mit  $E_1, \dots, E_n$  unabhängig auch  $F_1, \dots, F_n$  für  $F_i = E_i$  oder  $F_i = E_i^c$  unabhängig voneinander sind. Dies müssten wir noch einmal nachrechnen, dazu für den Fall  $n = 2$  ist z.B:

$$\mathbb{P}(E_1 \cap E_2^c) + \mathbb{P}(E_1 \cap E_2) = \mathbb{P}(E_1 \cap (E_2 \cup E_2^c)) = \mathbb{P}(E_1).$$

Also ergibt sich

$$\mathbb{P}(E_1 \cap E_2^c) = \mathbb{P}(E_1) - \mathbb{P}(E_1) \mathbb{P}(E_2) = \mathbb{P}(E_1)(1 - \mathbb{P}(E_2)) = \mathbb{P}(E_1) \mathbb{P}(E_2^c).$$

wie zu zeigen war.

### 1.6.3 Die Poisson-Verteilung

Betrachte Ereignisse  $E_1, \dots, E_n$ , die unabhängig sind und jeweils Wahrscheinlichkeit  $p$  haben, einzutreten.

**Question.** Was passiert wenn  $n \gg 1$ .

Typischerweise haben wir dann  $\mathcal{O}(pn)$  Erfolge in  $E_1, \dots, E_n$ .

- Sei  $p = p(n)$  sodass  $\lim_{n \rightarrow \infty} pn = \lambda \in (0, \infty)$
- Wähle Zeiteinheit  $\delta = \frac{1}{n}$

Wir fragen uns nun: Was ist

$$\lim_{n \rightarrow \infty} \mathbb{P}(\exists k \text{ Erfolge in } A).$$

für ein Intervall  $A \subset [0, 1]$

**Theorem 1.14.** Sei  $\lambda \in (0, \infty)$ . Dann ist

$$\lim_{n \rightarrow \infty} \text{Bin}\left(n, \frac{\lambda}{n}\right)(k) = \frac{e^{-\lambda} \lambda^k}{k!}.$$

für  $k = 0, 1, 2, \dots$

*Beweis.* Sei  $k$  fest. Dann ist

$$\text{Bin}\left(n, \frac{\lambda}{n}\right)(k) = \frac{n!}{(n-k)!k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{n(n-1) \cdots (n-k+1)}{n^k} \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}.$$

□

**Definition 1.15.** Sei  $\lambda \in (0, \infty)$  fest. Die Wahrscheinlichkeitsverteilung auf  $\{0, 1, 2, \dots\}$  mit Massenfunktion

$$p(k) = \frac{e^{-\lambda} \lambda^k}{k!}.$$

heißt **Poisson-Verteilung mit Parameter  $\lambda$** .

**Notation.** Wir schreiben auch  $\text{Poi}(\lambda)$  für die Poisson-Verteilung zum Parameter  $\lambda$ .

## Lecture 5

Mo 26 Apr 2021 10:17

### 1.6.4 Die geometrische Verteilung

- Seien  $E_1, E_2, \dots$  unabhängige Ereignisse mit

## 1.7 Simulation von Gleichverteilung

Typischerweise benutzen wir folgende Situation:

**Input** Zahl(en), z.B. Redinerzeit

**Output** 'Zufällige Zahl' in  $\{0, \dots, n\}$

### 1.7.1 Lineare Kongruenzgeneratoren (LCG)

**Startwert**  $x_0 \in \mathbb{N}$  gegeben.

**Parameter**  $a, c, m \in \mathbb{N}$

**Schritt** Setze  $x_{n+1} := (a \cdot x_n + c) \bmod m$ .

Dieses Vorgehen produziert eine scheinbar zufällige Folge.

**Example.**

◇

**Example** (Eine schlechte Wahl). Wenn wir  $a = 4, c = 1, m = 31$  wählen sowie  $x_0 = 3$ , so erreichen wir Periode 9, und somit werden nicht alle Zahlen erreichen / generieren.

◇

**Lemma 1.16** (Knuth). Die Periode eines LCG ist gleich  $m$ , genau dann, wenn

- (a)  $c$  und  $m$  haben keine gemeinsamen Primfaktoren
- (b) Jeder Primfaktor von  $m$  ist ein Teiler von  $a - 1$
- (c) Falls  $4 \mid m$ , dann  $4 \mid a - 1$ .

**Example.**

◇

### 1.7.2 Zufallsvariablen aus $[0, 1)$

- Sei  $(x_n)_{n \geq 1}$  eine Folge von (Pseudo)zufallszahlen aus  $\{0, 1, \dots, m - 1\}$ . Dann ist

$$u_n := \left( \frac{x_n}{m} \right)_{n \geq 1}.$$

eine Folge von Pseudozahlen in  $[0, 1)$ . Gut ist aber nur der Fall, wenn  $m \approx 10^N$ , wobei  $N = \text{Rechnergenauigkeit, d.h. \#Ziffern}$ .

### 1.7.3 Zufallspermutationen

Wie erzeugt man eine gleichverteilte Permutation von  $\{1, \dots, N\}$ ?

---

**Algorithmus 1 :** Zufallspermutationen

---

**Eingabe :** Möglichkeit, aus endlicher Menge gleichverteilt zufällige Zahlen zu ziehen

**Ausgabe :** Eine zufällige Permutation von  $\{1, \dots, N\}$

Setze  $\sigma_0 := \{1, \dots, N\}$

**for**  $i = 1$  **to**  $n - 1$  **do**

| wähle  $k \in \{i, \dots, N\}$  gleichverteilt  
 | Setze  $\sigma_k := \sigma_{k-1} \circ \tau_{i,k}$

---

**Lemma 1.17.** Der Algorithmus erzeugt eine zufällige gleichverteilte Permutation.

*Beweis.* Der Algorithmus benutzt eine Gleichverteilung auf

$$\Omega_n := \{1, \dots, N\} \times \{2, \dots, n\} \times \{n - 1, n\}.$$

Für  $\omega = (w_1, \dots, w_{N-1}) \in \Omega_N$  ist

$$\sigma(\omega) = \tau_{N-1, w_{N-1}} \circ \dots \circ \tau_{1, w} \circ \underbrace{(1, \dots, N)}_{\sigma_0}.$$

Es genügt also zu zeigen, dass  $\sigma : \Omega_N \rightarrow \mathcal{S}_N$  eine Bijektion ist. Wir sehen:

- (a)  $|\Omega_N| = |\mathcal{S}_N| = N!$
- (b) Sei  $w \neq \tilde{w}$  und setze  $k = \min \{j \mid \omega_j \neq \tilde{\omega}_j\}$ . Dann ist  $\sigma(\omega)_k \neq \sigma(\tilde{\omega})_k$  und somit ist die Funktion injektiv

Damit ist die Abbildung sogar bijektiv und wir sind fertig. □

### 1.7.4 Geometrische Verteilung

- Sei  $X \sim \text{Geo}(q)$ , d.h.

$$\mathbb{P}(X = k) = (1 - q)q^k.$$

Wie simuliert man nun  $X$ ?

- (a) Erzeuge  $n \sim U[0, 1)$  als Gleichverteilte Zufallsvariable auf  $[0, 1)$ .
- (b) Sei  $T_k := \mathbb{P}(X < k)$ . Falls  $n \in [T_k, T_{k+1})$ , dann setze  $X = k$ .

## 1.8 Erwartungswert und Varianz

- Sei  $X$  eine reellwertige diskrete Zufallsverteilung. Sei

$$X : \Omega \rightarrow \mathcal{S} \subset \mathbb{R}.$$

eine diskrete Zufallsvariable, d.h.  $\mathcal{S}$  abzählbar.

**Definition 1.18.** Seien  $x_1, \dots, x_n \in \mathcal{S}$   $n$  Beobachtungen einer Zufallsvariable  $X$ . Der **empirische Mittelwert** ist durch

$$\frac{1}{n} \sum_{i=1}^n x_i.$$

definiert.

- Wir wollen eine Sorte von Mittelwert definieren, der nur von  $X$  abhängig ist, und nicht von den Beobachtungen.
- Folgende Forderungen ergeben sich an solch einen Mittelwert:
  - Falls  $X(\omega) = x$  für jedes  $\omega$ , dann muss der Mittelwert von  $X$  gleich  $x$  sein.
  - Jeder Wert  $x \in \mathcal{S}$  muss bezüglich der Massenfunktion  $p_X(x)$  gewichtet sein.

**Definition 1.19.** Der **Erwartungswert** von  $X$  bzgl.  $\mathbb{P}$  ist durch

$$\mathbb{E}(X) = \sum_{s \in \mathcal{S}} s \cdot \mathbb{P}(X = s) = \sum_{s \in \mathcal{S}} s \cdot p_X(s).$$

definiert. Dies ist wohldefiniert, falls die Reihe absolut gegen einen Wert  $< \infty$  konvergiert.

**Remark.** Nicht alle Wahrscheinlichkeitsverteilungen besitzen einen endlichen Mittelwert, das zeigt folgendes

**Example.** Sei  $X$  auf  $\{1, 2, \dots\}$  verteilt mit

$$\mathbb{P}_X(s) = \frac{6}{\pi^2 s^2}.$$

dann ergibt sich für den Erwartungswert:

$$\mathbb{E}(X) = \sum_{s \geq 1} s \cdot \frac{6}{\pi^2 s^2} = \frac{6}{\pi^2} \cdot \sum_{s \geq 1} \frac{1}{s} \rightarrow \infty.$$

◇