Albert-Ludwigs-Universität Freiburg

Vorlesungsskript

Stochastik I

Angelika Rohde

Wintersemester 2024/25

Inhaltsverzeichnis

1 Wahrscheinlichkeit

1.1 Axiome von Kolmogorov

Definition 1.1 (Diskreter Wahrscheinlichkeitsraum und diskreter Maßraum)**.** *Gegeben sei eine nicht-leere Menge Ω (genannt Grundraum oder Stichprobenraum). Eine Abbildung $\mu : \mathcal{P}(\Omega) \to [0, \infty]$ heißt diskretes Maß, falls folgende Eigenschaften erfüllt sind:*

(i) *Nicht-Negativität: $\forall A \subset \Omega : \mu(A) \geq 0$.*

(ii) *Nulltreue: $\mu(\varnothing) = 0$.*

(iii) *σ-Additivität: Für $A_i \subset \Omega \ \forall i \in \mathbb{N}$ mit $A_i \cap A_j = \varnothing \ \forall i, j \in \mathbb{N}, i \neq j$ gilt:*

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

(iv) *Diskretheit: Es gibt eine abzählbare Teilmenge $\Omega_0 \subset \Omega$ mit $\mu(\Omega_0^C) = 0$.*

Das Tripel $(\Omega, \mathcal{P}(\Omega), \mu)$ nennt man einen diskreten Maßraum. Die Abbildung μ heißt diskretes Wahrscheinlichkeitsmaß, falls $\mu(\Omega) = 1$ ist. In diesem Fall nennt man das Tripel $(\Omega, \mathcal{P}(\Omega), \mu)$ auch diskreten Wahrscheinlichkeitsraum. Für $A \subset \Omega$ heißt dann $\mu(A)$ die Wahrscheinlichkeit von A.

Bemerkung 1.2. *Ein (diskretes) Wahrscheinlichkeitsmaß wird typischerweise mit $\mathbb{P}$ bezeichnet (wobei der Buchstabe P für „Probability“ steht).*

Übungsaufgabe 1.3. *Beweisen Sie: Eine Abbildung $\mu : \mathcal{P}(\Omega) \to [0, \infty]$ mit (i), (iii) und $\mu(\Omega) = 1$ erfüllt auch die Eigenschaft (ii). N.b.: Für abzählbares $\Omega \neq \varnothing$ werden die Eigenschaften (i), (iii) und $\mu(\Omega) = 1$ auch Axiome von Kolmogorov genannt (1933).*

Bemerkung 1.4. *Die Menge Ω_0 aus Definition 1.1 ist nicht eindeutig. Jede abzählbare Menge $\Omega_0 \subset \Omega$ mit $\mu(\Omega_0^C) = 0$ heißt Träger von μ. Lässt man aus einem Träger alle Elemente ω weg, die $\mu(\{\omega\}) = 0$ erfüllen, so bleibt ein Träger mit $\mu(\{\omega\}) > 0 \ \forall \omega \in \Omega_0$ übrig. Dieses Ω_0 nennt man dann Träger im strengen Sinne.*

Konvention 1.5. *Ist in einer endlichen oder unendlichen Summe von Zahlen aus $[0, \infty]$ mindestens ein Summand ∞, so wird die Summe gleich ∞ gesetzt.*

Notation 1.6. *Sind $A_i \in \mathcal{P}(\Omega)$ für $i \in \mathbb{N}$ paarweise disjunkte Mengen (d.h. $A_i \cap A_j = \varnothing \ \forall i, j \in \mathbb{N}, i \neq j$), so schreibt man auch:*

$$\sum_{i=1}^{\infty} A_i \quad \text{für} \quad \bigcup_{i=1}^{\infty} A_i$$

Lemma 1.7 (Regel von De Morgan)**.** *Seien $\Omega \neq \varnothing$ eine beliebige Grundmenge und I eine beliebige Indexmenge. Für alle $i \in I$ sei $A_i \in \mathcal{P}(\Omega)$. Dann gilt:*

$$\left(\bigcap_{i\in I} A_i\right)^C = \bigcup_{i\in I} A_i^C \text{ und } \left(\bigcup_{i\in I} A_i\right)^C = \bigcap_{i\in I} A_i^C$$

Beweis. Übungsblatt 1, Aufgabe 2. □

Lemma 1.8. *Gegeben seien ein Grundraum $\Omega \neq \varnothing$, ein diskretes Maß $\mu : \mathcal{P}(\Omega) \to [0, \infty]$ und ein diskretes Wahrscheinlichkeitsmaß $\mathbb{P} : \mathcal{P}(\Omega) \to [0, \infty]$. Dann gelten folgende Aussagen:*

(*i*) *Endliche Additivität: Ist $n \in \mathbb{N}$ gegeben und sind $A_1, \ldots, A_n$ paarweise disjunkte Teilmengen von Ω (diese bezeichnet man gerne als „Ereignisse"), so gilt:*

$$\mu\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i)$$

(*ii*) *Isotonie: Ist $A \subset B$ für $A, B \in \mathcal{P}(\Omega)$, so ist $\mu(A) \leq \mu(B)$. Speziell folgt hieraus für das Wahrscheinlichkeitsmaß $\mathbb{P}$, dass $\forall A \in \mathcal{P}(\Omega)$ gilt, dass $0 \leq \mathbb{P}(A) \leq 1$ ist.*

(*iii*) *Subtraktivität: Ist $\mu(A) < \infty$ für ein $A \in \mathcal{P}(\Omega)$, so gilt die folgende Implikation $\forall B \in \mathcal{P}(\Omega)$:*

$$A \subset B \Rightarrow \mu(B \setminus A) = \mu(B) - \mu(A)$$

Wegen $\mathbb{P}(\Omega) = 1$ gilt diese Implikation für Wahrscheinlichkeitsmaße immer.

(*iv*) *Komplementarität: Ist $\mu(A) < \infty$ für ein $A \in \mathcal{P}(\Omega)$, so folgt, dass $\mu(A^C) = \mu(\Omega) - \mu(A)$ ist. Insbesondere ist $\mathbb{P}(A^C) = 1 - \mathbb{P}(A)$.*

(*v*) *Stetigkeit von unten: Ist $(A_i)_{i\in\mathbb{N}}$ eine aufsteigende Folge in $\mathcal{P}(\Omega)$ (d.h. $\forall i \in \mathbb{N}$ gilt $A_i \subset A_{i+1}$), so folgt:*

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} \mu(A_n)$$

(*vi*) *Stetigkeit von oben: Ist $(A_i)_{i\in\mathbb{N}}$ eine absteigende Folge in $\mathcal{P}(\Omega)$ (d.h. $\forall i \in \mathbb{N}$ gilt $A_i \supset A_{i+1}$), und ist $\mu(A_1) < \infty$, so folgt:*

$$\mu\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} \mu(A_n)$$

(*vii*) *Subadditivität: Für $n \in \mathbb{N}$ und $A_1, \ldots, A_n \in \mathcal{P}(\Omega)$ gilt immer:*

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mu(A_i)$$

(*viii*) *σ-Subadditivität: Ist $(A_i)_{i\in\mathbb{N}}$ eine Folge in $\mathcal{P}(\Omega)$, so gilt:*

$$\mu\left(\bigcup_{i=1}^{\infty}\right) \leq \sum_{i=1}^{\infty} \mu(A_i)$$

Beweis. Wir weisen die Eigenschaften $(i)-(iv)$ nach. Die Eigenschaften $(v)-(viii)$ sind Gegenstand von Aufgabe 3 auf Übungsblatt 1.

(*i*) Mit $A_i := \varnothing \ \forall i > n$ gilt

$$\mu\left(\sum_{i=1}^{n} A_i\right) = \mu\left(\sum_{i=1}^{\infty} A_i\right) \overset{\substack{\sigma\text{-Additivität}\\ \downarrow}}{=} \sum_{i=1}^{\infty} \mu(A_i) \overset{\substack{\mu(\varnothing)=0\\ \downarrow}}{=} \sum_{i=1}^{n} \mu(A_i).$$

(*ii*) Sei $B = A \cup (B \setminus A)$. Es folgt aus (i), dass $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$ ist (denn $\mu(B \setminus A) \geq 0$).

(*iii*) $\mu(B) \overset{(ii)}{=\!=} \mu(A) + \mu(B \setminus A) \xRightarrow{\mu(A)<\infty} \mu(B \setminus A) = \mu(B) - \mu(A)$

(*iv*) folgt aus (iii) mit $B = \Omega$.

□

Diskrete Maße lassen sich leicht mithilfe von Zähldichten konstruieren.

Definition 1.9 (Zähldichte)**.** *Sei $\Omega \neq \varnothing$ ein Grundraum.*

(*i*) *Eine Abbildung $f : \Omega \to [0, \infty]$ heißt Zähldichte, wenn $T = \{\omega \in \Omega \mid f(\omega) > 0\}$ abzählbar ist. Gemäß*

$$\mu(A) := \sum_{\omega \in A\cap T} f(\omega) \text{ mit } \sum_{\omega\in\varnothing} f(\omega) := 0 \tag{$*$}$$

wird offenbar ein diskretes Maß auf $(\Omega, \mathcal{P}(\Omega))$ definiert, wobei $\Omega_0 = T$ gesetzt werden kann. Gilt

$$\sum_{\omega\in T} f(\omega) = 1,$$

so ist das durch $(*)$ definierte Maß ein diskretes Wahrscheinlichkeitsmaß.

(ii) *Ist umgekehrt μ ein diskretes Maß auf $(\Omega, \mathcal{P}(\Omega))$, so ist die durch*

$$f : \Omega \to [0, \infty], \omega \mapsto \mu(\{\omega\})$$

definerte Funktion eine Zähldichte.

Beispiel 1.10. *(i)* *Seien $\Omega = \{1, 2, 3, 4, 5, 6\}$ und $f : \Omega \to [0, \infty], \omega \mapsto \frac{1}{|\Omega|} = \frac{1}{6}$. Das zu f assoziierte Wahrscheinlichkeitsmaß $\mathbb{P}$ heißt Laplace-Verteilung auf $(\Omega, \mathcal{P}(\Omega))$. $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ ist ein Modell für einen einmaligen Würfelwurf.*

(ii) *Betrachten wir nun einen einmaligen Münzwurf. Dafür verwenden wir die Grundmenge $\Omega = \{\text{Kopf}, \text{Zahl}\}$ mit der Zähldichte $f(\text{Kopf}) = \frac{1}{2}$, $f(\text{Zahl}) = \frac{1}{2}$. Das zugehörigeWahrscheinlichkeitsmaß ist ebenfalls eine Laplace-Verteilung auf $(\Omega, \mathcal{P}(\Omega))$.*

1.2 Urnenmodelle und abgeleitete Verteilungen

Wir betrachten n Ziehungen aus einer Urne mit N Kugeln, die mit den Zahlen $1, \ldots, N$ durchnummeriert seien.

Beispiel 1.11 (6 aus 49)**.** *Wieviele Möglichkeiten gibt es, aus 49 Zahlen 6 (unterschiedliche) Zahlen auszuwählen? Für die erste Zahl gibt es 49 Möglichkeiten; für die zweite dann nur noch 48, da eine Zahl dann ja bereits gewählt wurde; für die dritte gibt es nur noch 47 Möglichkeiten, usw.; wenn wir diese Anzahl noch durch die Anzahl der möglichen Permutationen eine 6-elementigen Menge teilen, ergibt sich schlussendlich folgende Formel:*

$$\frac{49 \cdot 48 \cdot \ldots \cdot 44}{6 \cdot \ldots \cdot 1} = \binom{49}{6} = 13.983.816$$

Dies ist genau die Anzahl aller 6-elementigen Teilmengen von $\{1, \ldots, 49\}$. Allgemeiner ist $\binom{n}{k}$ genau die Anzahl aller k-elementigen Teilmengen der Menge $\{1, \ldots, n\}$. Dieses Beispiel entspricht dem Ziehen ohne Reihenfolge und ohne Zurücklegen von 6 Kugeln aus einer Urne mit (ursprünglich) 49 Kugeln.

Satz 1.12 (Kombinatorik)**.** *Wir betrachten das Ziehen einer Stichprobe vom Umfang n aus einer N-elementigen Menge. Es gilt:*

	Mit Zurücklegen	Ohne Zurücklegen
Mit Reihenfolge	N^n	$N \cdot (N-1) \cdot \ldots \cdot (N-n+1) = \frac{N!}{(N-n)!}$
Ohne Reihenfolge	$\binom{N+n-1}{n}$	$\binom{N}{n}$

Ohne Beweis.

Beispiel 1.13 (Ziehen mit Zurücklegen, Binomialverteilung). *Wir betrachten eine Urne mit N Kugeln; R davon seien rot, und $N-R$ weiß. Aus dieser Urne ziehen wir n-mal hintereinander eine Kugel, wobei wir nach jedem Zug die jeweils gezogene Kugel zurücklegen. Die Kugeln seien von $1,\dots,N$ durchnummeriert – o.B.d.A. seien die ersten R Kugeln rot. Wir betrachten dazu den Stichprobenraum*

$$\Omega = \big\{(\omega_1,\dots,\omega_n) \mid 1 \le \omega_i \le N\big\}$$

(ω_1 bezeichnet dabei die Nummer der ersten gezogenen Kugel, ω_2 die der zweiten, usw.) mit Laplace-Verteilung, d.h. die Wahrscheinlichkeit jedes n-Tupels ist gleich.

Frage: *Wie groß ist die Wahrscheinlichkeit, dass genau r rote Kugeln in der Stichprobe sind? Dies entspricht dem Ereignis $E_r = \{(\omega_1,\dots,\omega_n) : |\{i : \omega_i \le R\}| = r\}$.*

Um die Mächtigkeit von E_r zu berechnen, schreiben wir E_r als Vereinigung disjunkter Ereignisse E_I, wobei $I \subseteq \{1,\dots,n\}$ die Nummern der Ziehungen enthält, bei denen genau eine rote Kugel gezogen wird. Die Idee ist, dass wir von diesen Ereignissen die Wahrscheinlichkeit leichter ermitteln können. Wir setzen also

$$E_I = \{(\omega_1,\dots,\omega_n) \mid \omega_i \in \{1,\dots,R\} \text{ für } i \in I \text{ und } \omega_i \in \{R+1,\dots,N\} \text{ für } i \notin I\}$$

und

$$E_r = \bigcup_{\substack{I \subseteq \{1,\dots,n\} \\ |I|=r}} E_I.$$

Da es genau $\binom{n}{r}$ Teilmengen $I \subseteq \{1,\dots,n\}$ mit Kardinalität r gibt, und da E_I für alle solchen I dieselbe Kardinalität hat, folgt:

$$|E_r| = \binom{n}{r} \cdot |E_{\{1,\dots,r\}}| = \binom{n}{r} \cdot R^r \cdot (N-R)^{n-r}$$

Mit der Laplace-Verteilung und der Tatsache, dass $|\Omega| = N^n$ ist, ergibt sich:

$$\mathbb{P}(E_r) = \frac{|E_r|}{|\Omega|} = \binom{n}{r} \cdot \left(\frac{R}{N}\right)^r \cdot \left(1 - \frac{R}{N}\right)^{n-r}$$

Da $E_0,\dots,E_n$ eine disjunkte Zerlegung des Ergebnisraumes Ω ist (denn irgendeine Anzahl roter Kugeln muss man ja gezogen haben, und diese Anzahl liegt bei n Kugeln zwischen 0 und n), wird durch

$$p(r) := \mathbb{P}(E_r) \ \forall r \in \{0,\dots,n\}$$

eine Zahldichte eines Wahrscheinlichkeitsmaßes auf $(\{0,\dots,n\}, \mathcal{P}(\{0,\dots,n\}))$ definiert, *denn es gilt*

$$1 = \mathbb{P}(\Omega) = \mathbb{P}\left(\sum_{i=0}^{n} E_i\right) = \sum_{i=0}^{n} \mathbb{P}(E_i) = \sum_{i=0}^{n} p(i).$$

Das Wahrscheinlichkeitsmaß heißt Binomialverteilung *(mit Parametern n und $\frac{R}{N}$).*

2 Unabhängigkeit, Produkträume, bedingte Wahrscheinlichkeit

2.1 Stochastische Unabhängigkeit

Definition 2.1 (Stochastische Unabhängigkeit)**.** *Sei $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ ein diskreter Wahrscheinlichkeitsraum. Dann heißen zwei Ereignisse A und B stochastisch unabhängig, wenn gilt:*

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

Beispiel 2.2. (i) *Wir werfen einen unverfälschten Würfel und definieren folgende Ereignisse:*

- *A: „Die Augenzahl ist gerade".*
- *B: „Die Augenzahl ist durch 3 teilbar".*

Nach obiger Definition sind diese Ereignisse stochastisch unabhängig, denn es gilt:

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{3}{6} = \frac{1}{2} \; ; \; \mathbb{P}(B) = \frac{|B|}{|\Omega|} = \frac{1}{3} \; ; \; \mathbb{P}(A \cap B) = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3}$$

(ii) *Wir ziehen zwei mal mit Zurücklegen aus einer Urne mit 3 roten und 5 weißen Kugeln, und betrachten folgende Ereignisse:*

- *A: „Die erste gezogene Kugel ist rot".*
- *B: „Die zweite gezogene Kugel ist weiß".*

Diese Ereignisse sind stochastisch unabhängig, denn es gilt:

$$\mathbb{P}(A) = \frac{3 \cdot 8}{8^2} = \frac{3}{8} \; ; \; \mathbb{P}(B) = \frac{8 \cdot 5}{8^2} = \frac{5}{8} \; ; \; \mathbb{P}(B) = \frac{15}{8^2} = \frac{3}{8} \cdot \frac{5}{8}$$

Definition 2.3. *Sei $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ ein diskreter Wahrscheinlichkeitsraum. Die Ereignisse $A_1, \ldots, A_n \in \mathcal{P}(\Omega)$ heißen stochastisch unabhängig, wenn $\forall\, k \in \{1, \ldots, n\}$ und jede Auswahl von Indizes $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ gilt:*

$$\mathbb{P}\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} \mathbb{P}(A_{i_j}).$$

Bemerkung. *Es folgt sofort, dass Teilfamilien von unabhängigen Ereignissen wieder unabhängig sind.*

Beispiel 2.4. *Wir werfen eine faire Münze zweimal. Wir betrachten also den folgenden Grundraum:*

$$\Omega = \{(K,K),(K,Z),(Z,K),(Z,Z)\}.$$

Setze:

- $A := \{(K,K),(K,Z)\} =$ *„Der erste Wurf ergibt Kopf"* $\rightsquigarrow \mathbb{P}(A) = 1/2$
- $B := \{(K,K),(Z,K)\} =$ *„Der zweite Wurf ergibt Kopf"* $\rightsquigarrow \mathbb{P}(B) = 1/2$
- $A := \{(K,Z),(Z,K)\} =$ *„Es wird genau einmal Kopf geworfen"* $\rightsquigarrow \mathbb{P}(C) = 1/2.$

Wir fragen uns, ob diese Ereignisse unabhängig sind. Dazu berechnen wir:

- $\mathbb{P}(A \cap B) = 1/4 = \mathbb{P}(A) \cdot \mathbb{P}(B) \rightsquigarrow A$ *und* B *sind unabhängig.*
- $\mathbb{P}(B \cap C) = 1/4 = \mathbb{P}(B) \cdot \mathbb{P}(C) \rightsquigarrow B$ *und* C *sind unabhängig.*
- $\mathbb{P}(A \cap C) = 1/4 = \mathbb{P}(A) \cdot \mathbb{P}(C) \rightsquigarrow A$ *und* C *sind unabhängig.*

Allerdings gilt: $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\varnothing) = 0 \neq \frac{1}{8} = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C)$*. Also sind zwar die Ereignisse* A, B *und* C *paarweise unabhängig, nicht aber das Gesamtsystem. Dies zeigt uns sofort, dass es nicht ausreicht, paarweise Unabhängigkeit zu überprüfen, um die Unabhängigkeit eines Systems nachzuweisen.*

Satz 2.5. *Sei* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *ein diskreter Wahrscheinlichkeitsraum. Seien die Ereignisse* $A_1, \ldots, A_n \in \mathcal{P}(\Omega)$ *stochastisch unabhängig. Dann sind auch die Ereignisse* $B_1, \ldots, B_n$ *mit* $B_i \in \{A_i, A_i^C\}\ \forall i \in \{1, \ldots, n\}$ *unabhängig.*

Beweis. Wir beweisen die Aussage mittels vollständiger Induktion nach n.

Induktionsanfang für $n = 2$*:* Es gilt:

$$A_1 = (A_1 \cap A_2) + (A_1 \cap A_2^C)$$

Dabei bezeichnet „+" eine disjunkte Vereinigung von Mengen.

$$\Rightarrow \mathbb{P}(A_1) = \underbrace{\mathbb{P}(A_1 \cap A_2)}_{=\mathbb{P}(A_1)\cdot\mathbb{P}(A_2)} + \mathbb{P}(A_1 \cap A_2^C)$$

$$\Rightarrow \mathbb{P}(A_1 \cap A_2^C) = \mathbb{P}(A_1) \cdot (1 - \mathbb{P}(A_2)) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2^C)$$

Damit sind die Ereignisse A_1 und A_2^C stochastisch unabhängig. Analog für A_1^C und A_2 bzw. A_1^C und A_2^C.

Induktionsschritt: Die Behauptung sei für $n-1$ richtig. Seien $A_1, \dots, A_n$ stochastisch unabhängig. Sei $\{i_1, \dots, i_m\} \subset \{1, \dots, n\}$ mit $m := n-1$. Nach Induktionsvoraussetzung sind dann die Ereignisse $B_{i_1}, \dots, B_{i_m}$ unabhängig (da dies $n-1$ Ereignisse sind). Es bleibt folgende Gleichung zu zeigen:

$$\mathbb{P}\left(\bigcap_{i=1}^{n} B_i\right) = \prod_{i=1}^{n} \mathbb{P}(B_i) \tag{$*$}$$

Es seien o.B.d.A. $B_1 = A_1^C, \dots, B_k = A_k^C, B_{k+1} = A_{k+1}, \dots, B_n = A_n$. Wir zeigen die Aussage $(*)$ per Induktion über k.

Induktionsanfang für $k=1$: Mit endlicher Additivität erhalten wir:

$$\mathbb{P}\left(A_1^C \cap \bigcap_{i=2}^{n} A_i\right) + \mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbb{P}\left(\bigcap_{i=2}^{n} A_i\right)$$

$$\Rightarrow \mathbb{P}\left(A_1^C \cap \bigcap_{i=2}^{n} A_i\right) = \left(\prod_{i=2}^{n} \mathbb{P}(A_i)\right) \cdot (1 - \mathbb{P}(A_1)) = \mathbb{P}(A_1^C) \cdot \prod_{i=2}^{n} \mathbb{P}(A_i)$$

Induktionsschritt von $k-1$ nach k: Es gilt:

$$\begin{aligned}\mathbb{P}\left(\left(\bigcap_{i=1}^{k} A_i^C\right) \cap \left(\bigcap_{i=k+1}^{n} A_i\right)\right) &+ \mathbb{P}\left(\left(\bigcap_{i=1}^{k-1} A_i^C\right) \cap \left(\bigcap_{i=k}^{n} A_i\right)\right) \\ &= \mathbb{P}\left(\left(\bigcap_{i=1}^{k-1} A_i^C\right) \cap \left(\bigcap_{i=k+1}^{n} A_i\right)\right) \overset{\text{I.V.}}{=\!=} \left(\prod_{i=1}^{k-1} \mathbb{P}(A_i^C)\right) \cdot \left(\prod_{i=k+1}^{n} \mathbb{P}(A_i)\right)\end{aligned}$$

Damit gilt analog zum Induktionsanfang:

$$\mathbb{P}\left(\left(\bigcap_{i=1}^{k} A_i^C\right) \cap \left(\bigcap_{i=k+1}^{n} A_i\right)\right) = \left(\prod_{i=1}^{k} \mathbb{P}(A_i^C)\right) \cdot \left(\prod_{i=k+1}^{n} \mathbb{P}(A_i)\right)$$

Insgesamt ergibt sich damit die Aussage $(*)$. □

2.2 Produkträume und Produktexperimente

Motivation 2.6. *Zwei Studenten – Student 1 in Freiburg, Student 2 in Konstanz – führen gleichzeitig ein Experiment aus: Student 1 wirft eine (faire) Münze, Student 2 wirft einen (fairen) Würfel. Beide haben ihre eigenen Wahrscheinlichkeitsräume,* $\Omega_1 = \{K, Z\}$ *mit Laplace-Verteilung* $\mathbb{P}_1$ *on* $\mathcal{P}(\Omega_1)$ *und* $\Omega_2 = \{1, \dots, 6\}$ *mit Laplace-Verteilung* $\mathbb{P}_2$ *on* $\mathcal{P}(\Omega_2)$*. Was die Wahrscheinlichkeit des Ereignisses* A *"Student 1 wirft Kopf, Student 2 erhält eine "6"" anbelangt, besteht das Problem , dass das Ereignis weder in* $\mathbb{P}(\Omega_1)$ *noch in* $\mathbb{P}(\Omega_2)$ *enthalten ist.*

Definition 2.7. *Es seien* $(\Omega_1, \mathcal{P}(\Omega_1), \mathbb{P}_1), \ldots, (\Omega_n, \mathcal{P}(\Omega_n), \mathbb{P}_n)$ *diskrete Wahrscheinlichkeitsräume. Der* *Produktraum* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *ist der diskrete Wahrscheinlichkeitsraum mit*

$$\Omega := \Omega_1 \times \ldots \times \Omega_n = \{(\omega_1, \ldots, \omega_n) \mid \omega_i \in \Omega_i \ \forall i \in \{1, \ldots, n\}\}$$

und der Zähldichte

$$p((\omega_1, \ldots, \omega_n)) := \underbrace{p_1(\omega_1)}_{=\mathbb{P}_1(\{\omega_1\})} \cdot \ldots \cdot \underbrace{p_n(\omega_n)}_{=\mathbb{P}_n(\{\omega_n\})}$$

Das zu p gehörende Wahrscheinlichkeitsmaß $\mathbb{P}$ *auf* $(\Omega, \mathcal{P}(\Omega))$ *heißt* *Produkt-Wahrscheinlichkeitsmaß* *und wird oft mit* $\mathbb{P} = \mathbb{P}_1 \otimes \cdots \otimes \mathbb{P}_n$ *bezzeichnet.*

Satz 2.8. *Es seien* $(\Omega_1, \mathcal{P}(\Omega_1), \mathbb{P}_1), \ldots, (\Omega_n, \mathcal{P}(\Omega_n), \mathbb{P}_n)$ *diskrete Wahrscheinlichkeitsräume und* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *der zugehörige Produktraum. Für* $i \in \{1, \ldots, n\}$ *seien* $A_i \subset \Omega$ *Ereignisse und* $A_i' := \{(\omega_1, \ldots, \omega_n) \in \Omega \mid \omega_i \in A_i\} = \Omega_1 \times \ldots \times \Omega_{i-1} \times A_i \times \Omega_{i+1} \times \ldots \times \Omega_n$. *Dann gilt* $\mathbb{P}(A_i') = \mathbb{P}_i(A_i)$ *für alle* $i \in \{1, \ldots, n\}$ *und die Ereignisse* $A_1', \ldots, A_n'$ *sind stochastisch unabhängig.*

Beweis. Es ist

$$\begin{aligned}
\mathbb{P}(A_i') = \sum_{\omega \in A_i'} p(\omega) &= \sum_{\substack{(\omega_1,\ldots,\omega_n) \in \Omega \\ \omega_i \in A_i}} \prod_{i=1}^{n} p_i(\omega_i) \\
&= \underbrace{\left(\sum_{\omega_1 \in \Omega_1} p_1(\omega_1) \right)}_{=1} \cdot \ldots \cdot \left(\sum_{\omega_i \in A_i} p_i(\omega_i) \right) \cdot \ldots \cdot \underbrace{\left(\sum_{\omega_n \in \Omega_n} p_n(\omega_n) \right)}_{=1} \\
&= \sum_{\omega_i \in A_i} p_i(\omega_i) = \mathbb{P}_i(A_i),
\end{aligned}$$

und ür die Indizes $1 \leq i_1 < i_2 < \ldots < i_n \leq n$ gilt

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{j=1}^{k} A_{i_j}' \right) &= \sum_{\substack{(\omega_1,\ldots,\omega_n) \in \Omega \\ \omega_{i_j} \in A_{i_j} \ \forall j \in \{1,\ldots,n\}}} \prod_{j=1}^{n} p_j(\omega_j) \\
&= \sum_{\substack{(\omega_1,\ldots,\omega_n) \in \Omega \\ \omega_{i_j} \in A_{i_j} \ \forall j \in \{1,\ldots,n\}}} \prod_{j=1}^{k} p_{i_j}(\omega_{i_j}) \\
&= \left(\sum_{\omega_{i_1} \in A_{i_1}} p_{i_1}(\omega_{i_1}) \right) \cdot \ldots \cdot \left(\sum_{\omega_{i_k} \in A_{i_k}} p_{i_k}(\omega_{i_k}) \right) = \prod_{j=1}^{k} \mathbb{P}_{i_j}(A_{i_j}) = \prod_{j=1}^{k} \mathbb{P}(A_{i_j}')
\end{aligned}$$

Damit sind die Ereignisse $A_1', \ldots, A_n'$ unabhängig. $\square$

Bemerkung. *Mit der Identität* $\bigcap_{i=1}^n A_i' = A_1 \times \ldots \times A_n$ *bedeutet Satz 2.8, dass auf dem Produkt-Wahrscheinlichkeitsraum gilt:*

$$\mathbb{P}(A_1 \times \ldots \times A_n) = \prod_{i=1}^n \mathbb{P}_i(A_i).$$

Diese Identität sit Ausgangspunkt einer alternativen Definition des Produkt-Wahrscheinlichkeitsmaßes $\to$ *Maßtheorie (Analysis III).*

Beispiel 2.9 (Noch einmal die Studenten aus Freiburg and Konstanz)**.** *Mit der Notation aus Motivation 2.6, definiere* $\Omega := \Omega_1 \times \Omega_2$ *sowie das Produktmaß* $\mathbb{P} := \mathbb{P}_1 \otimes \mathbb{P}_2$ *auf* $\mathcal{P}(\Omega)$ *mit Komponenten* $\mathbb{P}_1$ *und* $\mathbb{P}_2$*. Das Ereignis* $\{(K,6)\} =$ *"Student 1 wirft Kopf, Student 2 erhält eine "6""ist eine einelementige Teilmenge von* Ω *und*

$$\begin{aligned}
\mathbb{P}(\{K,6)\}) &= \mathbb{P}\Big(\big(\{K\} \times \Omega_2\big) \cap \big(\Omega_1 \times \{6\}\big)\Big) \\
&= \mathbb{P}(\{K\} \times \Omega_2) \cdot \mathbb{P}(\Omega_1 \times \{6\}) \\
&= \mathbb{P}_1(\{K\}) \cdot \mathbb{P}_2(\{6\}) = \frac{1}{2} \cdot \frac{1}{6}.
\end{aligned}$$

D.h., unter dem Produktmaß sind die Ereignisse in Experiment 1 *unabhängig von denen in Experiment* 2*.*

2.3 Bedingte Wahrscheinlichkeiten

Ziel: Beschreibung von Abhängigkeitseigenschaften von Ereignissen. Was ist die Wahrscheinlichkeit von einem Ereignis A, wenn wir wissen, dass ein anderes Ereignis B eingetreten ist?

Definition 2.10. *Sei* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *ein diskreter Wahrscheinlichkeitsraum. Weiter seien* $A, B \in \mathcal{P}(\Omega)$ *Ereignisse mit* $\mathbb{P}(A) > 0$*. Dann wird die bedingte Wahrscheinlichkeit von* B *gegeben* A *durch*

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

definiert.

Beispiel 2.11. *Wir betrachten den zweimaligen Würfelwurf sowie die Ereignisse*

- $\Omega := \{1, \ldots, 6\}^2$
- $A := \{(\omega_1, \omega_2) \in \Omega \mid \omega_1 + \omega_2 = 12\}$
- $B := \{(\omega_1, \omega_2) \in \Omega \mid \omega_1 = 5\}$

Damit gilt:

$$\mathbb{P}(A) = \frac{1}{36} \ ; \ \mathbb{P}(B) = \frac{1}{6} \ ; \ \mathbb{P}(A \mid B) = \frac{P(\overbrace{A \cap B}^{=\varnothing})}{P(B)} = 0$$

Bemerkung 2.12. *Für unabhängige Ereignisse $A, B \in \mathcal{P}(\Omega)$ mit $\mathbb{P}(A) > 0$ gilt*

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A) \cdot \mathbb{P}(B)}{\mathbb{P}(A)} = \mathbb{P}(B).$$

Anders gesagt: Sind die Ereignisse A und B unabhängig, so ändert das Wissen, dass das Ereignis A eingetreten ist, nichts an der Wahrscheinlichkeit, dass auch das Ereigniss B eintritt. Umgekehrt: Ist $\mathbb{P}(B \mid A) = \mathbb{P}(B)$, so folgt

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

d.h. A und B sind unabhängig.

Satz 2.13 (Satz von der totalen Wahrscheinlichkeit)**.** *Sei $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ ein diskreter Wahrscheinlichkeitsraum. Seien $B_1, \ldots, B_n \in \mathcal{P}(\Omega)$ mit folgenden Eigenschaften:*

- $\bigcup\limits_{i=1}^{n} B_i = \Omega$
- $B_i \cap B_j = \varnothing \ \forall i, j \in \{1, \ldots, n\}, i \neq j$
- $\mathbb{P}(B_j) > 0 \ \forall j \in \{1, \ldots, n\}.$

Eine solche Zerlegung des Raumes Ω wird auch Partition von Ω genannt. Dann gilt für $A \in \mathcal{P}(\Omega)$:

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \mid B_i) \cdot \mathbb{P}(B_i)$$

Beweis. Da auch die Mengen $A \cap B_i$, $i = 1, \ldots, n$, paarweise disjunkt sind, folgt

$$\begin{aligned} \mathbb{P}(A) = \mathbb{P}(A \cap \Omega) = \mathbb{P}\left(A \cap \bigcup_{j=1}^{n} B_j\right) &= \mathbb{P}\left(\bigcup_{j=1}^{n} (A \cap B_j)\right) \\ &= \sum_{j=1}^{n} \mathbb{P}(A \cap B_j) = \sum_{j=1}^{n} \mathbb{P}(A \mid B_j) \cdot \mathbb{P}(B_j) \end{aligned}$$

□

Satz 2.14 (Formel von Bayes)**.** *Sei $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ ein diskreter Wahrscheinlichkeitsraum. Sei $B_1, \ldots, B_n \in \mathcal{P}(\Omega)$ eine Partition von Ω wie oben, und sei $A \in \mathcal{P}(\Omega)$. Dann gilt $\forall \, i \in \{1, \ldots, n\}$:*

$$\mathbb{P}(B_i \mid A) = \frac{\mathbb{P}(A \mid B_i) \cdot \mathbb{P}(B_i)}{\sum_{j=1}^{n} \mathbb{P}(A \mid B_j) \cdot \mathbb{P}(B_j)}$$

Beweis. Es gilt $\forall\, i \in \{1, \dots, n\}$:

$$\mathbb{P}(B_i \mid A) = \frac{\mathbb{P}(A \cap B_i)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \mid B_i) \cdot \mathbb{P}(B_i)}{\mathbb{P}(A)} \overset{2.13}{=\!=} \frac{\mathbb{P}(A \mid B_i) \cdot \mathbb{P}(B_i)}{\sum_{j=1}^{n} \mathbb{P}(A \mid B_j) \cdot \mathbb{P}(B_j)}.$$

□

Beispiel 2.15. *Gegeben seien 6 Urnen, die jeweils 5 Kugeln enthalten. Die erste Urne enthält 5 weiße Kugeln, die zweite Urne enthält 4 weiße und eine schwarze Kugel, usw. bis zur 6. Urne, die 5 schwarze Kugeln enthält:*

Urne 1	○ ○ ○ ○ ○
Urne 2	● ○ ○ ○ ○
Urne 3	● ● ○ ○ ○
Urne 4	● ● ● ○ ○
Urne 5	● ● ● ● ○
Urne 6	● ● ● ● ●

Experiment: Wir werfen zunächst einen unverfälschten Würfel und wählen diejenige Urne aus, die der geworfenen Augenzahl entspricht. Anschließend ziehen wir aus der ausgewählten Urne mit Zurücklegen Kugeln. Wir definieren folgende Ereignisse:

- $B_i :=$ *„Es wird aus der i-ten Urne gezogen".*
- $A_1 :=$ *„Die erste gezogene Kugel ist schwarz".*
- $A_2 :=$ *„Die zweite gezogene Kugel ist schwarz".*

Ziel ist nun, die Wahrscheinlichkeiten $\mathbb{P}(A_1)$ und $\mathbb{P}(A_2 \mid A_1)$ zu ermitteln. Dafür halten wir zunächst fest, dass die Ereignisse $B_1, \dots, B_6$ eine disjunkte Zerlegung des Ergebnisraumes bilden. Ferner gilt $\forall\, i \in \{1, \dots, 6\}$:

$$\mathbb{P}(A_1 \mid B_i) = \frac{i-1}{5}$$

Damit gilt nach dem Satz der totalen Wahrscheinlichkeit:

$$\mathbb{P}(A_1) = \sum_{i=1}^{6} \underbrace{\mathbb{P}(A_1 \mid B_i)}_{=\frac{i-1}{5}} \cdot \underbrace{\mathbb{P}(B_i)}_{=\frac{1}{6}} = \frac{1}{30} \cdot \sum_{i=0}^{5} i = \frac{1}{2}$$

Bemerkung: Diese Aussage war bereits klar aus Symmetriegründen.
Außerdem gilt:

$$\mathbb{P}(A_2 \mid A_1) = \frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_1)} = 2\mathbb{P}(A_1 \cap A_2)$$

Nun ist

$$\mathbb{P}(A_1 \cap A_2 \mid B_i) = \left(\frac{i-1}{5}\right)^2$$

also gilt:

$$\mathbb{P}(A_1 \cap A_2) = \sum_{i=1}^{6} \underbrace{\mathbb{P}(A_1 \cap A_2 \mid B_i)}_{=\left(\frac{i-1}{5}\right)^2} \cdot \underbrace{\mathbb{P}(B_i)}_{=\frac{1}{6}} = \frac{1}{150} \cdot \sum_{i=0}^{5} i^2 = \frac{11}{30}$$

$$\Rightarrow \mathbb{P}(A_2 \mid A_1) = \frac{11}{15}.$$

3 Diskrete Zufallsvariablen

Beispiel 3.1. *Wir betrachten das dreimalige Werfen einer Münze mit dem Grundraum*

$$\Omega := \{(\omega_1, \omega_2, \omega_3) \mid \omega_i \in \{0,1\} \ \forall i \in \{1,2,3\}\}.$$

Wir nutzen die Laplace-Verteilung $\mathbb{P}$ *auf* $(\Omega, \mathcal{P}(\Omega))$, *d.h.* $\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{8}$ *für* $A \in \mathcal{P}(\Omega)$.
Frage: Wie groß ist die Wahrscheinlichkeit, dass k-mal Zahl geworfen wird?
Dafür gibt es zwei mögliche Herangehensweisen:

(*i*) *Wir definieren das Ereignis*

$$A_k := \left\{ (\omega_1, \omega_2, \omega_3) \in \Omega \;\middle|\; \sum_{i=1}^{3} \omega_i = k \right\}$$

und berechnen $\mathbb{P}(A_k)) = \frac{|A_k|}{8}$.

(*ii*) *Die Anzahl ist binomialverteilt, d.h. wir betrachten* $Bin(3, \frac{1}{2})$. *Diese Verteilung ist allerdings kein Wahrscheinlichkeitsmaß auf* $(\Omega, \mathcal{P}(\Omega))$. *Wir müssen also einen anderen Grundraum und ein anderes Wahrscheinlichkeitsmaß verwenden:*

$$\Omega^X = \{0,1,2,3\} \text{ und } \mathbb{P}^X(\{k\}) := \binom{3}{k} \cdot \left(\frac{1}{2}\right)^k \cdot \left(1 - \frac{1}{2}\right)^{3-k} = \binom{3}{k} \cdot \frac{1}{8}, \ k = 0,1,2,3.$$

Den Raum $(\Omega^X, \mathcal{P}(\Omega^X), \mathbb{P}^X)$ *erhält man aus* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *mit folgender Abbildung:*

$$X : \Omega \to \mathbb{R}, \quad \omega = (\omega_1, \omega_2, \omega_3) \mapsto \sum_{i=1}^{3} \omega_i.$$

Dann ist $\Omega^X = \mathrm{Bild}(X) = \{x \in \mathbb{R} \mid \exists \omega \in \Omega : x = X(\omega)\}$, *und es gilt* $\mathbb{P}^X(B) = \mathbb{P}(X^{-1}(B))$ *für jede Menge* $B \in \mathcal{P}(X)$.

Definition 3.2. *Sei* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *ein diskreter Wahrscheinlichkeitsraum. Eine Abbildung* $X : \Omega \to \mathbb{R}$ *heißt* <u>*diskrete Zufallsvariable*</u>. *Die durch* X *induzierte Verteilung*

$$\mathbb{P}^X(A) := \mathbb{P}(X^{-1}(A)) \ \forall A \in \mathcal{P}(\Omega^X)$$

ist ein diskretes Wahrscheinlichkeitsmaß auf $(\Omega^X, \mathcal{P}(\Omega^X))$. *Die Funktion*

$$F_X(x) = \mathbb{P}^X((-\infty, x] \cap \Omega^X) = \mathbb{P}(\underbrace{X^{-1}((-\infty, x] \cap \Omega^X)}_{=\{\omega | X(\omega) \leq x\}}) =: \mathbb{P}(X \leq x)$$

heißt <u>*Verteilungsfunktion von* X</u>.

$$\begin{tikzcd}
(\Omega, \mathcal{P}(\Omega), \mathbb{P}) \arrow[rr, bend left, "X"] & & (\Omega^X, \mathcal{P}(\Omega^X), \mathbb{P}^X)
\end{tikzcd}$$

Bemerkung 3.3. *(i) Man rechnet leicht nach, dass* $\mathbb{P}^X$ *tatsächlich (wie in obiger Definition behauptet) ein diskretes Wahrscheinlichkeitsmaß ist (da die Axiome von Kolmogorov gelten).*

(ii) Man rechnet ebenfalls leicht nach, dass F_X *monoton wachsend und rechtsseitig stetig ist, und folgende Limites hat:*

$$\lim_{x \to -\infty} F_X(x) = 0 \text{ und } \lim_{x \to \infty} F_X(x) = 1.$$

(iii) Da $\mathbb{P}$ *diskret ist, kann man* $\mathbb{P}^X$ *gleichermaßen auf* $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ *definieren.*

Die Poisson-Verteilung

In der Praxis stellt sich mitunter die Frage, explizite Zahlenwerte der Verteilungsfunktion F_X auszurechnen. Dies kann bspw. bei der Binomialverteilung für große Werte von n sehr aufwendig sein, da die Binomialkoeffizienten für große n sehr groß werden können.

Satz 3.4 (Poisson-Grenzwertsatz). *Es seien* $X_n \sim Bin(n, p_n)$ *für* $n \in \mathbb{N}$ *diskrete Wahrscheinlichkeitsräume. Falls ein* $\lambda \in (0, \infty)$ *existiert, sodass* $np_n \xrightarrow{n\to\infty} \lambda$ *gilt, so gilt*

$$\lim_{n\to\infty} \underbrace{\mathbb{P}(X_n = k)}_{=\binom{n}{k} p_n^k (1-p_n)^{n-k}} = e^{-\lambda} \cdot \frac{\lambda^k}{k!} \quad \forall k \in \mathbb{N}_0.$$

Die Abbildung $p_\lambda : \mathbb{N}_0 \to \mathbb{R}, \ k \mapsto e^{-\lambda} \frac{\lambda^k}{k!}$, *definiert die Zähldichte eines Wahrscheinlichkeitsmaßes auf* $(\mathbb{N}_0, \mathcal{P}(\mathbb{N}_0))$. *Dieses heißt* <u>*Poisson-Verteilung*</u>.

Beweis. Für festes $k \in \mathbb{N}_0$ ist

$$\begin{aligned}
\mathbb{P}(X_n = k) &= \binom{n}{k} \cdot p_n^k \cdot (1-p_n)^{n-k} \\
&= \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{k!} \cdot \frac{1}{n^k} \cdot (np_n)^k \cdot \left(1 - \frac{np_n}{n}\right)^n \cdot (1-p_n)^{-k} \\
&= \underbrace{\frac{n}{n} \cdot \frac{n-1}{n} \cdot \ldots \cdot \frac{n-k+1}{n}}_{\xrightarrow{n\to\infty} 1} \cdot \underbrace{\frac{1}{(1-p_n)^k}}_{\xrightarrow{n\to\infty} 1} \cdot \underbrace{\frac{(np_n)^k}{k!}}_{\xrightarrow{n\to\infty} \frac{\lambda^k}{k!}} \cdot \underbrace{\left(1 - \frac{np_n}{n}\right)^n}_{\xrightarrow{n\to\infty} e^{-\lambda}}
\end{aligned}$$

Nach Voraussetzung gilt $np_n \xrightarrow{n\to\infty} \lambda$, und damit $p_n \xrightarrow{n\to\infty} 0$. Wegen

$$\lim_{n\to\infty} \left(1 + \frac{x_n}{n}\right)^n = e^x$$

für $x_n \xrightarrow{n\to\infty} x$ folgt

$$\mathbb{P}(X_n = k) \xrightarrow{n\to\infty} e^{-\lambda} \cdot \frac{\lambda^k}{k!}.$$

Damit ist die erste Behauptung gezeigt. Wegen $p_\lambda(k) \geq 0$ und

$$\sum_{k=0}^{\infty} p_\lambda(k) = \sum_{k=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{k!} = e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{x^k} = e^{-\lambda} \cdot e^{\lambda} = e^0 = 1$$

ist p_λ tatsächlich Zähldichte eines diskreten Wahrscheinlichkeitsmaßes auf $(\mathbb{N}_0, \mathcal{P}(\mathbb{N}_0))$.

$\square$