

# Wahrscheinlichkeitstheorie

## Wahrscheinlichkeitsmaße

Bezeichne  $\Omega$  die Menge aller möglichen Ergebnisse eines Zufallsexperiments und  $A \in \Omega$  ein Ereignis. Ein Wahrscheinlichkeitsmaß ist eine Abbildung, die jedem  $A$  eine Zahl  $P(A) \in [0, 1]$  zuordnet. Es gilt:

- $P(\Omega) = 1$  (Normierung)
- $P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$  falls  $(A_i \cap A_j) = \emptyset$  für alle  $i \neq j$  ( $\sigma$ -Additivität)

Aus den Axiomen folgt

- $P(A^C) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(B \setminus A) = P(B) - P(A \cap B)$

Ist  $A$  eine echte Teilmenge von  $B$  ( $A \subset B$ ) gilt

- $P(B \setminus A) = P(B) - P(A)$
- $P(A) \subseteq P(B)$

Für Ereignisse  $A_1, \dots, A_n$  gilt

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

(a) Sind  $A, A_1, A_2, \dots \subset \Omega$  eine Folge von nicht disjunkten Ereignissen mit  $A_1 \subset A_2 \subset \dots \subset A$  und  $\bigcup_{n=1}^{\infty} A_n = A$ , dann ist

$$\lim_{n \rightarrow \infty} P(A_n) = P(A) = P\left(\bigcup_{n=1}^{\infty} A_n\right)$$

(b) Sind  $A, A_1, A_2, \dots \subset \Omega$  eine Folge von nicht disjunkten Ereignissen mit  $A_1 \supset A_2 \supset \dots \supset A$  und  $\bigcap_{n=1}^{\infty} A_n = A$ , dann gilt

$$\lim_{n \rightarrow \infty} P(A_n) = P(A) = P\left(\bigcap_{n=1}^{\infty} A_n\right)$$

(a) lässt sich folgendermaßen zeigen:

Seien

- $B_1 := A_1$
- $B_2 := A_2 \setminus A_1$
- ...
- $B_n := A_n \setminus A_{n-1}$

Mit der  $\sigma$ -Additivität ist dann

$$P(A) = P\left(\bigcup_{k=1}^{\infty} B_k\right) = \sum_{k=1}^{\infty} P(B_k) = \lim_{n \rightarrow \infty} \sum_{k=1}^n P(B_k) = \lim_{n \rightarrow \infty} P\left(\bigcup_{k=1}^n B_k\right) = \lim_{n \rightarrow \infty} P(A_n)$$

(b) lässt sich folgendermaßen zeigen:

Aus  $A_1 \subset A_2 \subset \dots \subset A_n$  folgt  $A_1^C \supset A_2^C \supset \dots \supset A_n^C$ . Sei nun

- $B_1 := A_1^C$
- $B_2 := A_2^C \setminus A_1^C$
- ...
- $A_n^C \setminus A_{n-1}^C$

Mit der  $\sigma$ -Additivität ist dann

$$P\left(\bigcup_{n=1}^{\infty} A_n^C\right) = P(A^C) = P\left(\bigcup_{k=1}^{\infty} B_k\right) = \sum_{k=1}^{\infty} P(B_k) = \lim_{n \rightarrow \infty} \sum_{i=1}^n P(B_i) = \lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n B_i\right) = \lim_{n \rightarrow \infty} P(A_n^C)$$

und folglich

$$\lim_{n \rightarrow \infty} P(A_n) = 1 - \lim_{n \rightarrow \infty} P(A_n^C) = 1 - P\left(\bigcup_{n=1}^{\infty} A_n^C\right) = 1 - P\left(\left[\bigcap_{n=1}^{\infty} A_n\right]^C\right) = P\left(\bigcap_{n=1}^{\infty} A_n\right) = P(A)$$

## Zufallsvariablen

Eine Zufallsvariable ist eine Abbildung  $X : \Omega \rightarrow \mathbb{R}$ . Für  $A \subset \mathbb{R}$  gilt

$$\{X \in A\} = \{\omega \in \Omega : X(\omega) \in A\}$$

und

$$P(X \in A) = P(\{\omega \in \Omega : X(\omega) \in A\}).$$

Für  $a \in \mathbb{R}$  ist

$$\{X = a\} = \{\omega \in \Omega : X(\omega) = a\}$$

und

$$P(X = a) = P(\{\omega \in \Omega : X(\omega) = a\}).$$

Die Verteilungsfunktion  $F = F_X$  der Zufallsvariablen  $X$  ist definiert durch  $F(x) = P(X \leq x)$  mit  $x \in \mathbb{R}$  und

$$P(X \in (a, b]) = P(a < X \leq b) = F(b) - F(a) \quad \text{für alle } a < b.$$

Zudem gilt

$$P(Y = a) = F(a) - F(b_-) = F(a) - \lim_{\epsilon \rightarrow 0} F(a - \epsilon).$$

Ist  $\{\epsilon_n\}_{n=1}^{\infty}$  eine fallende Folge mit  $\epsilon_n \subset \epsilon_{n+1}$  und  $\lim_{n \rightarrow \infty} \epsilon_n = 0$ , dann gilt für die Ereignisse  $A_n := \{a - \epsilon_n < X \leq a\}$  und  $\bigcap_{n=1}^{\infty} A_n = \{X = a\} = A$  gemäß Lemma 1b)

$$P(Y = a) = P\left(\bigcap_{n=1}^{\infty} A_n\right) = P(A) = \lim_{n \rightarrow \infty} P(A_n) = \lim_{n \rightarrow \infty} P(a - \epsilon_n < X \leq a) = F(a) - \lim_{n \rightarrow \infty} F(a - \epsilon_n)$$

Eine Zufallsvariable heißt diskret, falls ...

- ... sie (un)endlich viele abzählbare Werte hat.
- ...  $f(x) = P(X = x)$  gilt (Wahrscheinlichkeitsfunktion).
- ... für  $A \in \mathbb{R}$   $P(X \in A) = \sum_{x \in A} f(x)$  gilt.

Eine Zufallsvariable heißt stetig, falls es eine nicht negative Funktion  $f$  gibt, sodass

$$P(a \leq X \leq b) = \int_a^b f(x) dx \quad \text{für alle } a < b$$

$f$  heißt Dichte von  $X$ . Die Verteilungsfunktion ist dann

$$F(x) = \int_{-\infty}^x f(t) dt$$

Die gemeinsame Verteilungsfunktion  $F = F_{X_1, \dots, X_n}$  von  $n$  Zufallsvariablen  $X_1, \dots, X_n$  ist definiert durch

$$F(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n) \quad \text{für alle } x_1, \dots, x_n \in \mathbb{R}$$

Für die Randverteilungsfunktion  $F_{X_1}, \dots, F_{X_n}$  gilt

$$F_{X_i}(x_i) = \lim_{x_j \rightarrow \infty} F(x_1, \dots, x_n) \quad \text{für } i \neq j$$

Sind  $X_1, \dots, X_n$  diskrete Zufallsvariablen, so ist ihre gemeinsame Wahrscheinlichkeitsfunktion gegeben durch

$$f(x_1, \dots, x_n) = P(X_1 = x_1, \dots, X_n = x_n).$$

$X_1, \dots, X_n$  sind gemeinsam stetig verteilt, falls es eine gemeinsame Dichte  $f(x_1, \dots, x_n) \geq 0$  gibt, sodass

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(t_1, \dots, t_n) dt_n \dots dt_1 \quad \text{für alle } x_1, \dots, x_n \in \mathbb{R}$$

Für  $B_1, \dots, B_n \in \mathbb{R}$  gilt dann

$$P(X_1 \in B_1, \dots, X_n \in B_n) = \int_{B_1} \dots \int_{B_n} f(t_1, \dots, t_n) dt_n \dots dt_1$$

Für  $B \in \mathbb{R}^n$  ist

$$P(\{X_1, \dots, X_n\} \in B) = \int \dots \int_B f(t_1, \dots, t_n) dt_n \dots dt_1$$

und die Randdichte ist

$$f_{X_i} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n$$

Ist zum Beispiel  $f$  eine gemeinsame Dichte von  $X$  und  $Y$ , dann gilt

$$P(X < Y) = \int_{-\infty}^{\infty} \int_x^{\infty} f(x, y) dy dx$$

und

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx.$$

## Erwartungswert und Varianz

Ist  $X$  eine diskrete Zufallsvariable und  $g : \mathbb{R} \rightarrow \mathbb{R}$ , so ist

$$E[g(X)] = \sum_X g(x)P(X = x)$$

wobei die Summe über alle  $X$  mit  $P(X = x) > 0$  läuft.

Nimmt  $X$  nur Werte in  $\{0, 1, 2, \dots\}$  an, dann gilt

$$E[X] = \sum_{n=0}^{\infty} P(X > n),$$

denn

$$\begin{aligned} E[X] &= \sum_{n=0}^{\infty} nP(X = n) = \\ &P(X = 1) + \\ &P(X = 2) + P(X = 2) + \\ &P(X = 3) + P(X = 3) + P(X = 3) + \\ &\dots \\ &P(X = n) + P(X = n) + P(X = n) + \dots \\ &= P(X > 0) + P(X > 1) + P(X > 2) + \dots + P(X > n) \end{aligned}$$

Hat  $X$  eine Dichte  $f$ , so ist

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x) dx$$

Für jede nichtnegative Zufallsvariable  $X$  gilt

$$E[X] = \int_0^{\infty} P(X > x) dx = \int_0^{\infty} P(X \geq x) dx$$

Allgemein gilt

$$E[aX + bY] = aE[X] + bE[Y] \quad \text{für } a, b \in \mathbb{R}$$

Die Varianz der Zufallsvariable  $X$  ist

$$\text{Var}[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2$$

Die Kovarianz von Zufallsvariablen  $X$  und  $Y$  ist

$$\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$$

Es gilt

1.  $\text{Var}[aX + b] = a^2 \text{Var}[X]$  für  $a, b \in \mathbb{R}$
2.  $\text{Var}[aX + bY] = a^2 \text{Var}[X] + b^2 \text{Var}[Y] + 2ab \text{Cov}[X, Y]$  oder allgemein  $\text{Var}[\sum X_i] = \sum \text{Var}[X_i] + 2 \sum_{i < j} \text{Cov}[X_i, X_j]$

$$3. \text{Cov}[aX + b, cY + d] = ac\text{Cov}[X, Y] \text{ für } a, b, c, d \in \mathbb{R}$$

$$4. \text{Cov}\left[\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j\right] = \sum_{i=1}^n \sum_{j=1}^m \text{Cov}[X_i, Y_j]$$

## Unabhängigkeit

Ereignisse  $A_1, A_2, \dots$  heißen unabhängig, falls für jede endliche Auswahl von verschiedenen Indizes  $i_1, \dots, i_n$  gilt

$$P(A_{i_1} \cap \dots \cap A_{i_n}) = P(A_{i_1}) \dots P(A_{i_n})$$

Sind  $A_1, A_2, \dots$  unabhängige Ereignisse und ist für jedes  $i$   $B_i = A_i$  oder  $B_i = A_i^C$ , dann sind auch  $B_1, B_2, \dots$  unabhängig.

Zufallsvariablen  $X_1, \dots, X_n$  heißen unabhängig, falls für alle  $x_1, \dots, x_n \in \mathbb{R}$

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = F_{X_1}(x_1) F_{X_2}(x_2) \dots F_{X_n}(x_n)$$

Für eine unendliche Folge von Zufallsvariablen  $X_1, X_2, \dots$  bedeutet Unabhängigkeit, dass für jedes  $n$   $X_1, \dots, X_n$  unabhängig sind.

Für unabhängige Zufallsvariablen  $X_1, \dots, X_n$  gilt

$$1. \{X_1 \in B_1\}, \dots, \{X_n \in B_n\} \text{ sind unabhängige Ereignisse für alle } B_1, \dots, B_n \in \mathbb{R}$$

$$2. E\left[\prod_{i=1}^n X_i\right] = \prod_{i=1}^n E[X_i]$$

$$3. \text{Cov}[X_i, X_j] = 0 \text{ für alle } i \neq j \rightarrow \text{Var}[X_1, \dots, X_n] = \sum_{i=1}^n \text{Var}[X_i]$$

$$4. h_1(X_1, \dots, X_{n_1}), h_2(X_{n_1+1}, \dots, X_{n_2}), \dots, h_n(X_{n_{k-1}+1}, \dots, X_{n_k}) \text{ sind unabhängige Zufallsvariablen, wobei } 0 = n_0 < n_1 < \dots < n_k \leq n, \text{ wobei } h_i : \mathbb{R}^{n_i - n_{i-1}} \rightarrow \mathbb{R}$$

$$\left\{ \begin{aligned} &\{X_1, \dots, X_{n_1}\} \in B_1, \\ &\{X_{n_1+1}, \dots, X_{n_2}\} \in B_2, \end{aligned} \right.$$

...

$$\left\{ \begin{aligned} &\{X_{n_{k-1}+1}, \dots, X_{n_k}\} \in B_k, \end{aligned} \right.$$

sind somit unabhängige Ereignisse für alle  $B_1, \dots, B_k$ .

Für den Spezialfall diskreter Zufallsvariablen  $X_1, \dots, X_n$  gilt

$$X_1, \dots, X_n \text{ unabhängig} \Leftrightarrow P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) \dots P(X_n = x_n) \text{ für alle } x_1, \dots, x_n \in \mathbb{R}$$

Sind  $X_1, \dots, X_n$  stetige Zufallsvariablen mit gemeinsamer Dichte  $f$  und

$$f(x_1, \dots, x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n) \text{ für alle } x_1, \dots, x_n \in \mathbb{R},$$

dann sind  $X_1, \dots, X_n$  unabhängig. In diesem Fall ist

$$f_{X_1}(x_1) \dots f_{X_n}(x_n)$$

eine gemeinsame Dichte von  $X_1, \dots, X_n$ .

## Bedingte Wahrscheinlichkeiten und Erwartungswerte

Für  $A, B \in \Omega$  mit  $P(B) > 0$  ist die bedingte Wahrscheinlichkeit von  $A$  gegeben  $B$

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

- Bei festem  $B \in \Omega$  mit  $P(B) > 0$  ist  $P(\cdot | B)$  ein Wahrscheinlichkeitsmaß:  $P(\Omega | B) = 1$  und für disjunkte Ereignisse  $A_1, A_2, \dots$  ist  $P(A_1 \cup A_2 \cup \dots | B) = \sum_i P(A_i | B)$

- Sind A und B unabhängig, dann ist  $P(A | B) = P(A)$  (falls  $P(B) > 0$ )
- Bilden  $A_1, A_2, \dots$  eine Zerlegung von  $\Omega$ , also  $A_i \cap A_j = \emptyset$  für alle  $i \neq j$  und  $\bigcup_i A_i = \Omega$ , dann gilt für jedes  $A \in \Omega$

$$P(A) = \sum_i P(A \cap A_i) = \sum_{i: P(A_i) > 0} P(A | A_i) P(A_i) \quad (\text{Satz der totalen Wahrscheinlichkeit})$$

- Für beliebige Ereignisse  $A_1, \dots, A_n$  gilt

$$P(A_1 \cap A_2 \cap \dots) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 \cap A_2) \dots P(A_n | A_1 \cap \dots \cap A_{n-1}),$$

falls

$$P(A_1 \cap \dots \cap A_{n-1}) > 0 \quad (\text{Multiplikationssatz für bedingte Wahrscheinlichkeiten})$$

Für diskrete Zufallsvariablen  $X$  und  $Y$  ist die bedingte Wahrscheinlichkeitsfunktion  $f_{X|Y}(x | y)$  ( $X$ , gegeben  $Y = y$ ) definiert durch

$$f_{X|Y} = P(X = x | Y = y), \text{ falls } P(Y = y) > 0$$

Nach dem Satz der totalen Wahrscheinlichkeit ist dann

$$P(X = x) = \sum_{Y: P(Y=y) > 0} f_{X|Y}(x | y) P(Y = y)$$

$X$  und  $Y$  seien diskrete Zufallsvariablen mit Wahrscheinlichkeitsfunktionen  $f_X$  und  $f_Y$ . Für die Wahrscheinlichkeitsfunktion  $f_{X+Y}$  von  $X + Y$  gilt

$$\begin{aligned} f_{X+Y}(z) &= P(X + Y = z) = \sum_{Y: f_Y(y) > 0} P(X + Y = z | Y = y) f_Y(y) \\ &= \sum_Y f_{X|Y}(z - y | y) f_Y(y) \end{aligned}$$

Seien  $X$  und  $Y$  unabhängige diskrete Zufallsvariablen, dann gilt  $f_{X|Y}(x | y) = f_X(x)$  und

$$f_{X+Y}(z) = \sum_Y f_X(z - y) f_Y(y).$$

Diese Funktion heißt Faltung von  $f_X$  und  $f_Y$ .

Analog ist für unabhängige stetige Zufallsvariablen  $X$  und  $Y$  mit zugehörigen Dichten  $f_X(x)$  und  $f_Y(y)$  die Dichte von  $X + Y$  gegeben durch die Faltung

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) dy,$$

denn für alle  $t \in \mathbb{R}$  ist

$$\begin{aligned} P(X + Y + t) &= \int_{-\infty}^{\infty} \int_{-y}^{t-y} f_X(x) f_Y(y) dx dy = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^t f_X(z - y) f_Y(y) dz dy = \int_{-\infty}^t \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) dy dz \end{aligned}$$

Seien  $X$  und  $Y$  diskrete Zufallsvariablen und gilt  $g : \mathbb{R} \rightarrow \mathbb{R}$ . Der bedingte Erwartungswert von  $g(X)$  gegeben  $Y = y$  ist

$$E[g(X) \mid Y = y] = \sum_X g(x) f_{X \mid Y}(x \mid y) \quad \text{falls } P(Y = y) > 0$$

Zudem gilt

$$\begin{aligned} E[g(X)] &= \sum_{Y: P(Y=y)>0} E[g(X) \mid Y = y] P(Y = y) \quad , \text{ denn} \\ E[g(X)] &= \sum_X g(x) P(X = x) = \sum_X g(x) \sum_Y f_{X \mid Y}(x \mid y) P(Y = y) \\ &= \sum_Y P(Y = y) \sum_X g(x) f_{X \mid Y}(x \mid y) \end{aligned}$$

Im Falle diskreter Zufallsvariablen  $X, X_1, X_2, Y$  und  $y \in \mathbb{R}$  gelten folgende Rechenregeln für bedingte Erwartungswerte:

1.  $E[aX_1 + bX_2 \mid Y = y] = aE[X_1 \mid Y = y] + bE[X_2 \mid Y = y]$
2.  $E[h(X, Y) \mid Y = y] = E[h(X, y) \mid Y = y]$  für  $h : \mathbb{R}^2 \rightarrow \mathbb{R}$
3.  $E[g(X) \mid Y = y] = E[g(X)]$  falls  $X \perp Y$
4.  $E[g(X)h(Y) \mid Y = y] = h(y)E[g(X) \mid Y = y]$  für  $h : \mathbb{R} \rightarrow \mathbb{R}$

Ist  $Y$  diskret und  $X$  stetig mit bedingter Dichte  $f_{X \mid Y}(x \mid y)$ , also

$$P(a < X < b \mid Y = y) = \int_a^b f_{X \mid Y}(x \mid y) \quad \text{für alle } a < b \quad \text{und } P(Y = y) > 0,$$

dann gilt für die Randdichte  $f_X(x)$  von  $X$

$$f_X(x) = \sum_{Y: P(Y=y)>0} f_{X \mid Y}(x \mid y) P(Y = y)$$

und für  $g : \mathbb{R} \rightarrow \mathbb{R}$

$$E[g(X) \mid Y = y] := \int g(x) f_{X \mid Y}(x \mid y) dx \quad \text{falls } P(Y = y) > 0$$

Die Rechenregeln für bedingte Erwartungswerte gelten wie im diskreten Fall, wie zum Beispiel

$$E[g(X)] = \sum_{Y: P(Y=y)>0} E[g(X) \mid Y = y] P(Y = y).$$

## Markovketten

### Markov-Eigenschaft

Ein stochastischer Prozess ist eine Familie von Zufallsvariablen  $X_t$ , wobei der Parameter  $t$  eine Indexmenge  $T$  durchläuft. Oft ist  $T \in [0, \infty)$  oder  $T = \{0, 1, 2, \dots\}$  und  $t \in T$  wird als Zeitpunkt interpretiert.

Der Zustandsraum eines stochastischen Prozesses  $\{X_t : t \in T\}$  ist die Menge aller möglichen Werte der  $X_t$ . Man sagt der Prozess ist zur Zeit  $t$  in Zustand  $x$ , falls  $X_t = x$ .

Eine diskrete Markov-Kette ist ein stochastischer Prozess mit diskreter Zeit und diskretem Zustandsraum, sodass zu jeder Zeit die Verteilung des nächsten Zustands nur vom aktuellen Zustand abhängt aber nicht von den vorherigen. Formal bedeutet dies:

Sei  $S \neq \emptyset$  eine endliche oder abzählbar unendliche Menge. Sei  $(P_{ij})_{i,j \in S}$  eine stochastische Matrix, also  $p_{ij} \geq 0$  für alle  $i, j \in S$  und  $\sum_{j \in S} p_{ij} = 1$  für alle  $i \in S$ .

Eine Folge  $\{X_n : n \in \mathbb{N}\}$  von Zufallsvariablen mit Werten in  $S$  heißt Markovkette mit Zustandsraum  $S$  und Übergangsmatrix  $(P_{ij})_{i,j \in S}$ , falls für alle  $n \geq 0$  und alle möglichen  $i_0, i_1, \dots, i_{n+1} \in S$  gilt

$$(\star) P(X_{n+1} = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) = p_{i_n i_{n+1}},$$

sofern  $P(X_0 = i_0, \dots, X_n = i_n) > 0$ . Die Verteilung von  $X_0$  heißt Anfangsverteilung der Markov-Kette.

( $\star$ ) beinhaltet zwei Aussagen:

1. Die bedingte Verteilung von  $X_{n+1}$  für eine gegebene Vorgeschichte  $i_0, i_1, \dots, i_n$  hängt nur von der Gegenwart ab, aber nicht von der Vergangenheit. Dies ist die Markov-Eigenschaft.
2. Die bedingte Verteilung hängt nicht vom Zeitpunkt  $n$  ab. Die Übergangswahrscheinlichkeiten sind stationär.

Die  $i$ -te Zeile der Übergangsmatrix beschreibt die bedingte Verteilung des nächsten Zustands  $X_{n+1}$  gegeben  $X_n = i$  (und  $X_0 = i_0, \dots$ ).

Seien  $X_0, X_1, \dots$  unabhängige und identisch verteilte  $\mathbb{N}_0$ -wertige Zufallsvariablen mit  $P(X_0 = i) = a_i, i = 0, 1, \dots$ . Dann ist  $\{X_n : n \in \mathbb{N}\}$  eine Markov-Kette mit Zustandsraum  $\mathbb{N}_0$  und Übergangsmatrix  $(P_{ij})_{i,j \in \mathbb{N}_0}$  mit  $p_{ij} = a_j$  für alle  $i, j \in \mathbb{N}_0$ .

$$(P_{ij})_{i,j=0}^a = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots \\ a_0 & a_1 & a_2 & \cdots \\ a_0 & a_1 & a_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Denn für alle  $n \geq 0$  und alle  $i_0, \dots, i_{n+1} \in \mathbb{N}_0$  mit  $P(X_0 = i_0, \dots, X_n = i_n) > 0$  gilt:

$$P(X_{n+1} = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) = P(X_{n+1} = i_{n+1}) = a_{i_{n+1}} = p_{i_n i_{n+1}}$$

In Worten: Die Wahrscheinlichkeit von Zustand  $i$  in Zustand  $j$  zu gehen wird nur durch die Wahrscheinlichkeit in Zustand  $j$  zu sein bestimmt, oder anders gesagt  $X_{n+1}$  und  $X_n$  sind unabhängig.

### Irrfahrt auf $\mathbb{Z}$

Seien  $Y_1, Y_2, \dots$  unabhängige, identisch verteilte  $\mathbb{Z}$ -wertige Zufallsvariablen mit  $P(Y_1 = i), i \in \mathbb{Z}$ . Sei  $X_{0i} := 0$  und  $X_n := Y_1 + \dots + Y_n, n \in \mathbb{N}$ . In Worten:  $X_n$  springt zufällig von  $X_{n-1} = i_{n-1} = i_{n-2} + Y_{n-1}$  zu  $i_n = i_{n-1} + Y_n$  auf den ganzen Zahlen  $\mathbb{Z}$ . Damit gilt offensichtlich für jedes  $n \geq 1$ , dass  $X_n = i_{n-1} + Y_n \Leftrightarrow Y_n = i_n - i_{n-1}$ .

Für alle  $n \geq 0$  und alle  $i_0, \dots, i_{n+1} \in \mathbb{Z}$  mit  $P(X_0 = i_0, \dots, X_n = i_n) > 0$  ist

$$\begin{aligned} & P(X_{n+1} = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) \\ &= P(X_n + Y_{n+1} = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) \\ &= P(i_n + Y_{n+1} = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) \\ &= P(Y_{n+1} = i_{n+1} - i_n \mid X_0 = i_0, \dots, X_n = i_n) \\ &= P(Y_{n+1} = i_{n+1} - i_n \mid Y_1 = i_1, Y_2 = i_2 - i_1, \dots, Y_n = i_n - i_{n-1}) \\ &= P(Y_{n+1} = i_{n+1} - i_n) = a_{i_{n+1} - i_n}. \end{aligned}$$



$\Rightarrow \{X_n : n \in \mathbb{N}\}$  ist eine Markov-Kette mit Zustandsraum  $\mathbb{Z}$  und Übergangswahrscheinlichkeit  $p_{ij} = a_{j-1}, i, j \in \mathbb{Z}$ .

$$(P_{ij}) = \begin{pmatrix} \dots & a_{-2} & a_{-1} & a_0 & a_1 & a_2 & \dots \\ \dots & a_{-3} & a_{-2} & a_{-1} & a_0 & a_1 & \dots \\ \dots & a_{-4} & a_{-3} & a_{-2} & a_{-1} & a_0 & \dots \\ \dots & a_{-5} & a_{-4} & a_{-3} & a_{-2} & a_{-1} & \dots \\ \dots & a_{-6} & a_{-5} & a_{-4} & a_{-3} & a_{-2} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Ein Spezialfall wäre die einfache Irrfahrt auf  $\mathbb{Z}$  bei der  $P(Y_n = 1) = p$  und  $P(Y_n = -1) = 1 - p$  für alle  $n \in \mathbb{N}$ . Interpretation: Die Markov-Kette beschreibt die Position eines Teilchens, das sich auf  $\mathbb{Z}$  bewegt. Es startet im Ursprung  $X_0 = 0$  und springt zu jedem Zeitpunkt 1, 2, ... zufällig eine Einheit nach links oder rechts.

### Sukzessive Maxima

Seien  $Y_1, Y_2, \dots$  unabhängig und identisch verteilte  $\mathbb{N}_0$ -wertige Zufallsvariablen mit  $P(Y_1 = c) = a_c$ ,  $i \in \mathbb{N}_0$ ,  $X_0 := 0$  und  $X_n := \max(Y_1, \dots, Y_n)$ . In Worten: Die Kette nimmt zu jedem Zeitpunkt  $n$  den größten zufälligen Wert aus  $\{0, Y_1, \dots, Y_n\}$ , bzw. wegen der sukzessiven Natur des Prozesses aus  $\{i_{n-1}, Y_n\}$ , an.

Für alle  $n \geq 0$  und alle  $i_0, \dots, i_{n+1} \in \mathbb{N}_0$  mit  $P(X_0 = i_0, \dots, X_n = i_n) > 0$  ist

$$\begin{aligned} &P(X_{n+1} = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) \\ &= P(\max(Y_1, \dots, Y_{n+1}) = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) \\ &= P(\max(X_n, Y_{n+1}) = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) \\ &= P(\max(i_n, Y_{n+1}) = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) \\ &= P(\max(i_n, Y_{n+1}) = i_{n+1} \mid Y_1 = i_1, \max(Y_1, Y_2) = i_2, \dots, \max(Y_1, \dots, Y_n) = i_n) \\ &= P(\max(i_n, Y_{n+1}) = i_{n+1}) \end{aligned}$$

Diese Wahrscheinlichkeit hängt nun vom Verhältnis von  $i_n$  zu  $i_{n+1}$  ab. Per Konstruktion kann  $i_{n+1}$  nicht kleiner sein als  $i_n$ . Ist  $i_{n+1} < i_n$ , dann ist  $Y_{n+1}$  das Maximum und folglich auch  $P(\max(i_n, Y_{n+1}) = i_{n+1}) = P(Y_{n+1} = i_{n+1}) = a_{i_{n+1}}$ . Für den Fall, dass  $i_{n+1} = i_n$  ist, betrachten wir sowohl die Fälle, für die  $Y_{n+1} < i_n$  ist als auch den Fall  $Y_{n+1} = i_n$ . Daraus folgt  $P(\max(i_n, Y_{n+1}) = i_{n+1}) = P(Y_{n+1} \leq i_n) = \sum_{k=0}^n a_k$ .

$$P(\max(i_n, Y_{n+1}) = i_{n+1}) = \begin{cases} P(Y_{n+1} = i_{n+1}) = a_{i_{n+1}} & , \text{ falls } i_{n+1} > i_n \\ P(Y_{n+1} \leq i_n) = \sum_{k=0}^n a_k & , \text{ falls } i_{n+1} = i_n \\ 0 & , \text{ falls } i_{n+1} < i_n \end{cases}$$

$\Rightarrow \{X_n : n \in \mathbb{N}_0\}$  ist eine Markovkette mit Zustandsraum  $\mathbb{N}_0$  und Übergangsmatrix

$$p_{ij} = \begin{cases} a_j & , \text{ falls } j > i \\ \sum_{k=0}^i a_k & , \text{ falls } j = i \\ 0 & , \text{ falls } j < i \end{cases}$$

**Rekursive Darstellung einer Markovkette mittels zufälliger Funktionen** Sei  $X_0$  eine  $S$ -wertige Zufallsvariable und  $S$  endlich oder abzählbar unendlich. Setze für  $n \geq 0$  rekursiv

$$X_{n+1} = f(X_n, Y_{n+1})$$

Dabei sei  $f : S \times \mathbb{R} \rightarrow S$ ,  $Y_1, \dots, Y_n$  seien identisch verteilte Zufallsvariablen und  $Y_0, Y_1, \dots$  seien unabhängig.

$\Rightarrow \{X_n : n \in \mathbb{N}_0\}$  ist eine Markovkette mit Übergangswahrscheinlichkeiten  $p_{ij} = P(f(i, Y_1) = j)$ , denn für  $n \geq 1$  und  $i_0, \dots, i_{n+1} \in S$  mit  $P(X_0 = i_0, \dots, X_n = i_n) > 0$  gilt

$$\begin{aligned} & P(X_{n+1} = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) \\ &= P(f(X_n, Y_{n+1}) = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) \\ &= P(f(i_n, Y_{n+1}) = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) \\ &\stackrel{*}{=} P(f(i_n, Y_{n+1}) = i_{n+1}) \\ &\stackrel{**}{=} P(f(i_n, Y_1) = i_{n+1}) \\ &= p_{i_n i_{n+1}} \end{aligned}$$

★ Hinweis: Das Ereignis  $\{X_{n+1} = f(i_n, Y_{n+1}) \mid i_{n+1}\}$  hängt von einem festen Wert  $i_n$  und der Zufallsvariablen  $Y_{n+1}$  ab. Das gleiche gilt natürlich auch für  $X_0, \dots, X_n$ . Deswegen leigt hier Unabhängigkeit vor.

★ ★ Hinweis:  $Y_1, Y_2, \dots$  i.i.d

Bemerkung zum Beispiel: In Beispiel 3 war  $f(x, y) = x + y$  und in Beispiel 4 war  $f(x, y) = \max(x, y)$ .

## Mehrschritt-Übergangswahrscheinlichkeiten

Sei  $\{X_n : n \in \mathbb{N}_0\}$  eine Markovkette mit Zustandsraum  $S$ , Übergangsmatrix  $(P_{ij})_{i,j \in S}$  und Anfangsverteilung  $p_i = P(X_0 = i)$  ( $i \in S$ ). Bestimme die gemeinsame Verteilung von  $X_0, \dots, X_n$ : Für  $n \geq 1$  und  $i_0, \dots, i_n \in S$  gilt

$$\begin{aligned} & P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) \\ &= P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}) p_{i_{n-1} i_n} \\ &= P(X_0 = i_0, \dots, X_{n-2} = i_{n-2}) p_{i_{n-2} i_{n-1}} p_{i_{n-1} i_n} \\ &= \dots \\ &= P(X_0 = i_0) p_{i_0 i_1} \dots p_{i_{n-2} i_{n-1}} p_{i_{n-1} i_n} \end{aligned}$$

Für jede Folge von Zuständen  $i_0, \dots, i_n \in S$  gilt:

$$P(X_0 = i_0, \dots, X_n = i_n) = p_{i_0 i_1} \dots p_{i_{n-1} i_n}$$

Insbesondere ist für jedes  $n$  die Verteilung von  $X_0, \dots, X_n$  durch die Anfangsverteilung  $(p_i)$  und die Übergangsmatrix  $(P_{ij})_{i,j \in S}$  eindeutig festgelegt.

Betrachte eine Markovkette  $\{X_n : n \in \mathbb{N}_0\}$  mit Zustandsraum  $S = \{0, 1, 2, 3, 4\}$ ,  $X_0 := 0$ , also  $p_0 = P(X_0 = 0) = 1$ .

Sei  $X_n = \max(Y_1, \dots, Y_n)$ ,  $n \in \mathbb{N}$ , wobei  $Y_1, \dots, Y_n$  unabhängig und gleichverteilt ( $P(Y_n = j) = \frac{1}{5}$  für alle  $j \in S$  und  $n \in \mathbb{N}_0$ ).

In Worten: Die Markovkette nimmt zu jedem Zeitpunkt  $n$  das Maximum aus dem bisherigen Maximum  $i_{n-1}$  und dem zufälligen Wert  $Y_n$  an, welcher mit gleicher Wahrscheinlichkeit Werte zwischen 0 und 4 annimmt. Aus Beispiel 4 folgt diese Übergangsmatrix:

$$(P_{ij})_{i,j=0}^n = \begin{pmatrix} \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ 0 & \frac{2}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ 0 & 0 & \frac{3}{5} & \frac{1}{5} & \frac{1}{5} \\ 0 & 0 & 0 & \frac{4}{5} & \frac{1}{5} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Was wäre die Wahrscheinlichkeit, dass  $X_0 = 0$ ,  $X_1 = 2$  und  $X_3 = 4$ ?

$$\begin{aligned} P(X_0 = 0, X_1 = 2, X_2 = 2, X_3 = 4) \\ &= p_{00}p_{02}p_{22}p_{24} \\ &= 1 \cdot \frac{1}{5} \cdot \frac{3}{5} \cdot \frac{1}{5} = \frac{3}{125} \end{aligned}$$

Was wäre die Wahrscheinlichkeit, dass  $X_1 \geq 2$  und  $X_2 \leq X_1$ ?

Da es nicht möglich ist, dass  $i_{n+1} < i_n$ , können wir auf 3 mögliche Pfade mit  $X_1 = X_2 = 2$ ,  $X_1 = X_2 = 3$  und  $X_1 = X_2 = 4$  schließen.

$$\begin{aligned} P(X_1 \geq 2, X_2 \leq X_1) \\ &= P(X_1 \geq 2, X_1 = X_2) \\ &= P(X_0 = 0, X_1 = 2, X_2 = 2) + P(X_0 = 0, X_1 = 3, X_2 = 3) + \dots \\ &= \frac{12}{25} \end{aligned}$$

Die Markov-Eigenschaft lässt sich auf allgemeinere Ereignisse in Zukunft und Vergangenheit erweitern:

Für alle  $n, m \geq 1$ ,  $Z \subset S^m$ ,  $V \subset S^n$ ,  $i \in S$  mit  $P((X_0, \dots, X_{n-1}) \in V, X_n = i) > 0$  gilt

$$P((X_{n+1}, \dots, X_{n+m}) \in Z \mid (X_0, \dots, X_{n-1}) \in V, X_n = i) = \sum_{j_1, \dots, j_m \in Z} p_{ij_1} p_{j_1 j_2} \dots p_{j_{m-1} j_m}$$

In Worten: Die Wahrscheinlichkeit, dass  $\{X_n + 1 = i_{n+1}, \dots, X_{n+m} = i_{n+m}\} \in Z$ , hängt nur von  $X_n = i$  und nicht von  $\{X_0, \dots, X_{n-1}\} \in V$ . Die Wahrscheinlichkeit  $P(\{X_n + 1 = i_{n+1}, \dots, X_{n+m} = i_{n+m}\} \in Z \mid X_n = i)$  ist dann die Summe über alle (disjunkten) Pfad-Wahrscheinlichkeiten in  $Z$  (mit Start in  $X_n = i$  und daraus resultierender Anfangsverteilung  $p_i$ ). Sofern  $P(X_0 = i) > 0$  können wir die aufgrund der Zeithomogenität der Markovkette auch folgen:

$$P((X_{n+1}, \dots, X_{n+m}) \in Z \mid (X_0, \dots, X_{n-1}) \in V, X_n = i) = P((X_1, \dots, X_m) \in Z \mid X_0 = i)$$

### Beweis von Satz 7

Für alle  $(j_1, \dots, j_m) \in Z$  und  $(i_0, \dots, i_{n-1}) \in V$  mit  $P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i) > 0$  gilt

$$\begin{aligned} &P((X_{n+1}, \dots, X_{n+m}) = (j_1, \dots, j_m) \mid (X_0, \dots, X_{n-1}) = (i_0, \dots, i_{n-1}), X_n = i) \\ &= \frac{P((X_{n+1}, \dots, X_{n+m}) = (j_1, \dots, j_m), (X_0, \dots, X_{n-1}) = (i_0, \dots, i_{n-1}), X_n = i)}{P((X_0, \dots, X_{n-1}) = (i_0, \dots, i_{n-1}), X_n = i)} \\ &= \frac{p_{i_0} p_{i_0 i_1} \dots p_{i_{n-1} i} p_{ij_1} \dots p_{j_{m-1} j_m}}{p_{i_0} p_{i_0 i_1} \dots p_{i_{n-1} i}} \\ &= p_{ij_1} p_{j_1 j_2} \dots p_{j_{m-1} j_m} \end{aligned}$$

Oft werden für eine Markovkette  $\{X_n\}$  bei fester Übergangsmatrix  $(P_{ij})$  verschiedene Anfangszustände betrachtet. Bezeichne mit  $p_i(\cdot)$  die bedingte Verteilung  $P(X_n = j \mid X_0 = i)$  und mit  $E_i(\cdot)$  den Erwartungswert für den Anfangszustand  $i$ , das heißt  $P_i(X_0 = i) = 1$ . Dann gilt

$$\begin{aligned} & P((X_{n+1}, \dots, X_{n+m}) \in Z \mid (X_0, \dots, X_{n-1}) \in V, X_n = i) \\ &= P((X_1, \dots, X_m) \in Z \mid X_0 = i) \\ &= P_i((X_1, \dots, X_m) \in Z) \end{aligned}$$

und für  $f : f^m \rightarrow \mathbb{R}$

$$\begin{aligned} & E[f(X_{n+1}, \dots, X_{n+m}) \mid (X_0, \dots, X_{n-1}) \in V, X_n = i] \\ &= E[f(X_{n+1}, \dots, X_{n+m}) \mid X_0 = i] \\ &= E_i[f(X_1, \dots, X_m)] \end{aligned}$$

Beide Größen hängen vom gegenwärtigen Zustand  $i$  ab und weder von  $V$  noch von  $n$ .

Die Aussagen lassen sich mit einem unendlichen Zeithorizont erweitern:

Für die Menge  $Z \subset S^\infty$  und  $f : S^\infty \rightarrow \mathbb{R}$  gilt

$$P((X_{n+1}, X_{n+2}, \dots) \in Z \mid (X_0, \dots, X_{n-1}) \in V, X_n = i) = P_i((X_1, \dots) \in Z)$$

und

$$E[f(X_{n+1}, X_{n+2}, \dots) \mid (X_0, \dots, X_{n-1}) \in V, X_n = i] = E_i[f(X_{n+1}, X_{n+2}, \dots)]$$

Die  $n$ -Schritt Übergangswahrscheinlichkeit ist

$$p_{ij}^n := P(X_{n+m} = j \mid X_m = i)$$

mit  $i, j \in S, n \geq 0, P(X_m = i) > 0$ . Diese Wahrscheinlichkeit hängt nicht von  $m$  ab (Zeithomogenität).

Es gilt

$$p_{ij}^0 = P(X_m = j \mid X_m = i) = \begin{cases} 1 & , \text{ falls } i = j \\ 0 & , \text{ falls } i \neq j \end{cases}$$

und

$$P_{ij}^1 = P(X_{m+1} = j \mid X_m = i) = p_{ij}$$

Satz von Chapman-Kolmogorow: Für alle  $i, j \in S$  und  $m, n \in \mathbb{N}_0$  ist die Wahrscheinlichkeit, von Zustand  $i$  in  $m + n$  Schritten zu Zustand  $j$  zu gelangen gegeben durch:

$$p_{ij}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}$$

**Beweis von Satz 8:** Die Behauptung ist für  $m = 0, n = 0$  oder  $n = m = 0$  klar, da

$$p_{ij}^{(n)} = \sum_{k \in S} p_{ik}^0 p_{kj}^{(n)} = p_{ij}^{(n)}$$

$$p_{ij}^{(m)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^0 = p_{ij}^{(m)}$$

$$p_{ij}^0 = \sum_{k \in S} p_{ik}^0 p_{kj}^0 = \begin{cases} 1 & , \text{ falls } i = j \\ 0 & , \text{ falls } i \neq j \end{cases}$$

Für  $m, n \geq 1$  und  $P(X_0 = 1) > 0$  ist

$$\begin{aligned} p_{ij}^{m+n} &= P(X_{m+n} = j \mid X_0 = i) \\ &= \sum_{k \in S} P(X_{m+n} = j, X_m = k \mid X_0 = i) \\ &= \sum_{k \in S} \frac{P(X_{m+n} = j, X_m = k, X_0 = i)}{P(X_0 = i)} \times \frac{P(X_m = k, X_0 = i)}{P(X_m = k, X_0 = i)} \\ &= \sum_{k \in S} P(X_{m+n} = j \mid X_m = k, X_0 = i) P(X_m = k \mid X_0 = i) \\ &= \sum_{k \in S} P(X_m = k \mid X_0 = i) P(X_{m+n} = j \mid X_m = k) \\ &= \sum_{k \in S} p_{ik}^m p_{kj}^n \end{aligned}$$

Für einen Übergang von  $i$  nach  $j$  in  $m + n$  Schritten muss die Markovkette in  $m$  Schritten von  $i$  zu einem beliebigen Zustand  $k$  gehen und dann von  $k$  in  $n$  Schritten zu  $j$ .

Die  $n$ -Schritt Übergangsmatrix ist definiert als

$$\Pi^{(n)} := (P_{ij})^{(n)}$$

mit

$$\Pi = \Pi^{(1)} := (P_{ij})_{i,j \in S}^{(1)} = (P_{ij})_{i,j \in S}$$

Die  $n$ -Schritt Übergangsmatrix  $\Pi^{(n)}$  ist die  $n$ -te Potenz der Übergangsmatrix  $\Pi^{(1)}$ .

Somit besagt der Satz von Chapman-Kolmogorow, dass

$$\Pi^{m+n} = \Pi^m \Pi^n$$

Alternativ kann auch für einen Start bei  $X_n = i$  mit korrespondierender Verteilung, welche dem Zeilenvektor

$$(p_i^{(n)})_{i \in S}$$

der Übergangsmatrix mit  $p_i^{(n)} = P(X_n = i)$  und  $n \in \mathbb{N}_0$  entspricht, die Wahrscheinlichkeit in weiteren  $m \in \mathbb{N}_0$  Schritten ein beliebiges  $j \in S$  zu erreichen folgendermaßen erfasst werden:

$$p_j^{(n+m)} = P(X_{n+m} = j) = \sum_{i \in S} P(X_{n+m} = j \mid X_n = i) P(X_n = i) = \sum_{i \in S} p_i^{(n)} p_{ij}^{(m)}$$

Daraus folgt wiederum:

$$p^{(n+m)} = p^{(n)} \Pi^m$$

Letzteres kann genutzt werden, um die Verteilung  $X_n$  zu jeden Zeitpunkt zu berechnen:

Die Verteilung von  $X_n$  zu einem beliebigen Zeitpunkt  $n \geq 0$  ergibt sich aus dem Produkt der Anfangsverteilung und der  $n$ -ten Potenz der Übergangsmatrix.

$$p^{(n)} = p^{(0)} \Pi^n$$

Betrachte eine Markovkette  $X_n : n \in \mathbb{N}_0$  mit  $S = \{1, 2, 3\}$ , Übergangsmatrix

$$\Pi = \frac{1}{3} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

und Anfangsverteilung  $p^{(0)} = P(X_0 = i) = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ .

Berechne die bedingte Verteilung von  $X_2$  gegeben  $X_0 = 3$ .

$$\Pi^2 = \Pi \Pi = \frac{1}{9} \begin{pmatrix} 3 & 2 & 4 \\ 1 & 5 & 3 \\ 2 & 3 & 4 \end{pmatrix}$$

$$\Pi^4 = \Pi^2 \Pi^2 = \frac{1}{81} \begin{pmatrix} 19 & 28 & 34 \\ 14 & 36 & 31 \\ 17 & 31 & 33 \end{pmatrix}$$

- $P(X_2 = 1 \mid X_0 = 3) = p_{31}^{(2)} = \frac{2}{9}$
- $P(X_2 = 2 \mid X_0 = 3) = p_{32}^{(2)} = \frac{1}{3}$
- $P(X_2 = 3 \mid X_0 = 3) = p_{33}^{(2)} = \frac{4}{9}$

Berechne zudem die Verteilung von  $X_4$ .

$$p^{(4)} = p^{(0)} \Pi^{(4)} = (P(X_4 = 1), P(X_4 = 2), P(X_4 = 3)) = \left( \frac{16}{81}, \frac{131}{324}, \frac{43}{108} \right)$$

## Absorptionswahrscheinlichkeiten und -zeiten

Sei  $\{X_n : n \in \mathbb{N}_0\}$  eine Markovkette mit Zustandsraum  $S$  und Übergangsmatrix  $(P_{ij})$ .

Sei  $A \subset S$ ,  $A \neq \emptyset$ . Setze  $T_i = \inf\{n \in \mathbb{N}_0 : X_n \in A\}$ , wobei  $\inf \emptyset = \infty$ .  $T$  ist die Eintrittszeit in  $A$ , also der zufällige Zeitpunkt des ersten Besuchs der Menge  $A$ , falls es einen gibt.

Ziel: Berechne für jeden Anfangszustand  $i$  die Wahrscheinlichkeit, dass  $A$  in endlicher Zeit erreicht wird.

$$P(T < \infty \mid X_0 = i) = P\left(\bigcup_{n=0}^{\infty} \{X_n \in A\} \mid X_0 = i\right) = p_i(T < \infty)$$

Ein Zustand  $z \in S$  heißt absorbierend, falls  $p_{zz} = 1$ . In dem Spezialfall, dass  $A$  nur aus absorbierenden Zuständen besteht, heißt  $p_i(T < \infty)$  Absorptionswahrscheinlichkeit bei Start in  $i \in S$  und  $T$  heißt Absorptionszeit.

Folgender Satz gilt für beliebige  $A$ :

Sei  $\{X_n : n \in \mathbb{N}_0\}$  eine Markovkette mit Zustandsraum  $S$  und Übergangsmatrix  $(P_{ij})$ . Sei  $A \subset S$ ,  $A \neq \emptyset$  und

$$T_i = \inf\{n \in \mathbb{N}_0 : X_n \in A\}$$

$$h_i = P(T < \infty \mid X_0 = i)$$

$$\Rightarrow h_i = 1 \text{ für alle } i \in A \text{ und } h_i = \sum_{j \in S} p_{ij} h_j \text{ für alle } i \in S \setminus A$$

### Beweis Satz 10

Für  $i \in A$  gilt  $P(T = 0 \mid X_0 = i) = 1$ , also  $h_i = 1$ .

Sei nun  $i \in S \setminus A$ . Der Beweisansatz ist die Einschnitt-Analyse: Zerlege die gesuchte

Wahrscheinlichkeit  $h_i$  mit dem Satz der totalen Wahrscheinlichkeit danach, was im ersten Schritt der Markov-Kette passiert ist.

$$\begin{aligned} h_i &= P(T < \infty \mid X_0 = i) \\ &= \sum_{j \in S} P(T < \infty, X_1 = j \mid X_0 = i) \\ &= \sum_{j \in S} \overbrace{P(X_1 = j \mid X_0 = i)}^{p_{ij}} P(T < \infty \mid X_1 = j, X_0 = i) \end{aligned}$$

Für  $j \in A$  ist

$$P(T < \infty \mid X_1 = j, X_0 = i) = 1,$$

und für  $j \in S \setminus A$

$$\begin{aligned} P(T < \infty \mid X_1 = j, X_0 = i) &= P\left(\bigcup_{n=2}^{\infty} \{X_n \in A\} \mid X_1 = j, X_0 = i\right) \\ &= \lim_{N \rightarrow \infty} P\left(\bigcup_{n=2}^N \{X_n \in A\} \mid X_1 = j, X_0 = i\right) \\ &= \lim_{N \rightarrow \infty} P\left(\bigcup_{n=1}^{N-1} \{X_n \in A\} \mid X_0 = j\right) \\ &= P\left(\bigcup_{n=1}^{\infty} \{X_n \in A\} \mid X_0 = j\right) \\ &= P(T < \infty \mid X_0 = j) \\ &= h_j \end{aligned}$$

Damit folgt:

$$h_i = \sum_{j \in S} p_{ij} h_j$$

### Ruinproblem

In jeder Runde eines Spiels gewinnt eine Spieler 1€ mit Wahrscheinlichkeit  $p \in (0, 1)$  und verliert 1€ mit Wahrscheinlichkeit  $(1 - p)$ . Sein Anfangskapital sei  $i$ €. Er spielt so lange, bis er  $M$ € hat ( $M \geq i$ ) oder ruiniert ist (0€).

Berechne die Ruinwahrscheinlichkeit.

Betrachte hierfür die Markovkette  $\{X_n : n \in \mathbb{N}_0\}$  mit Zustandsraum  $S = \{0, 1, 2, \dots, M\}$  und Übergangswahrscheinlichkeiten

$$p_{ij} = \begin{cases} 1 & , \text{ falls } i = 0 \text{ und } j = 0 \\ p & , \text{ falls } i \in \{1, \dots, M-1\} \text{ und } j = i+1 \\ 1-p & , \text{ falls } i \in \{1, \dots, M-1\} \text{ und } j = i-1 \\ 1 & , \text{ falls } i = M \text{ und } j = M \\ 0 & , \text{ sonst} \end{cases}$$

Gesucht sind die Absorbtionswahrscheinlichkeiten

$$h_i = P(T < \infty \mid X_0 = i)$$

mit  $i = 0, \dots, M$  und  $T = \inf\{n \in \mathbb{N}_0 : X_n = 0\}$ .

Es ist  $h_0 = 1$  (sofort ruiniert),  $h_M = 0$  (sofort Ende des Spiels wegen maximalem Gewinn) und nach Satz 10 ist die Wahrscheinlichkeit jemals ruiniert zu sein beim Start in Periode  $i$  eine Gewichtung der Ruinwahrscheinlichkeit nach einer Aufwärtsbewegung (also ab  $i + 1$ ) und einer Abwärtsbewegung (also ab  $i - 1$ ).

$$h_i = ph_{i+1} + (1-p)h_{i-1} \quad \text{für } i = 1, \dots, M-1$$

$$\Leftrightarrow h_{i+1} - h_i = \frac{1-p}{p}(h_i - h_{i-1})$$

Falls es sich um ein faires Spiel handelt, also  $p = \frac{1}{2}$ , dann ist  $h_{i+1} - h_i = h_i - h_{i-1}$  für  $i = 1, \dots, M-1$ . Die Änderung in der Ruinwahrscheinlichkeit ist also konstant für jedes der  $i$ , man kann  $h_i$  also als lineare Funktion betrachten:

$$h_i = h_0 + a \times i$$

Da  $h_0 = 1$

$$h_i = 1 + a \times i$$

und  $h_M = 0$  folgt für das  $a$  aus der Gleichung

$$0 = 1 + a \times M$$

$$\Leftrightarrow a = -\frac{1}{M}$$

und somit für die Ruinwahrscheinlichkeit  $h_i$ :

$$h_i = 1 - \frac{i}{M}$$

Sei nun  $p \neq \frac{1}{2}$  und setze  $\Theta = \frac{1-p}{p}$ . Durch rekursives Einsetzen erhält man

$$h_2 - h_1 = \Theta(h_1 - h_0) = \Theta(h_1 - 1)$$

$$h_3 - h_2 = \Theta(h_2 - h_1) = \Theta^2(h_1 - 1)$$

$$\vdots$$

$$h_{i+1} - h_i = \Theta^i(h_1 - 1) \quad \text{für alle } i = 0, \dots, M-1$$

Zunächst bildet man die Summe über  $i = 0, \dots, j-1$  (über alle  $j \in S$ ) und nutzt die Eigenschaft der Teleskopsumme und der geometrischen Reihe aus:



$$\begin{aligned}
\sum_{i=0}^{j-1} h_{i+1} - h_i &= \sum_{i=0}^{j-1} \Theta^i (h_1 - 1) \\
\Leftrightarrow \sum_{i=0}^{j-1} h_i - h_{i+1} &= \sum_{i=0}^{j-1} \Theta^i (1 - h_1) \\
\Leftrightarrow h_0 - h_j &= \sum_{i=0}^{j-1} \Theta^i (1 - h_1) \\
\Leftrightarrow 1 - h_j &= (1 - h_1) \sum_{i=0}^{j-1} \Theta^i \\
\Leftrightarrow 1 - h_j &= (1 - h_1) \frac{1 - \Theta^j}{1 - \Theta} \\
\Leftrightarrow h_j &= (h_1 - 1) \frac{1 - \Theta^j}{1 - \Theta} + 1 \quad \text{für } j = 1, \dots, M
\end{aligned}$$

Da diese Gleichung für alle  $j$  gilt, gilt sie auch für  $j = M$  und da  $h_M = 0$  ist, ist

$$\begin{aligned}
0 &= (h_1 - 1) \frac{1 - \Theta^M}{1 - \Theta} + 1 \\
\Leftrightarrow 1 - h_1 &= \frac{1 - \Theta}{1 - \Theta^M}
\end{aligned}$$

Damit ist

$$1 - h_j = \frac{1 - \Theta}{1 - \Theta^M} \times \frac{1 - \Theta^j}{1 - \Theta}$$

und folglich

$$\Rightarrow h_j = 1 - \frac{1 - \Theta^j}{1 - \Theta^M} \quad \text{für } j = 0, \dots, M$$

### Ruinproblem bei unendlich reichem Gegner/Einfache Irrfahrt mit absorbierender Schranke

Betrachte die Markovkette  $\{X_n : n \in \mathbb{N}_0\}$  mit Zustandsraum  $S = \mathbb{N}_0$  und

$$p_{ij} = \begin{cases} 1 & , \text{ falls } i = 0 \text{ und } j = 0 \\ p & , \text{ falls } i \in \mathbb{N}_0 \text{ und } j = i + 1 \\ 1 - p & , \text{ falls } i \in \mathbb{N}_0 \text{ und } j = i - 1 \\ 0 & , \text{ sonst} \end{cases}$$

Berechne die Absorbtionswahrscheinlichkeiten.

$h_i = P(T < \infty \mid X_0 = i)$  mit  $i \in S$  und  $T = \inf\{n \in \mathbb{N}_0 : X_n = 0\}$ .

Es gilt  $h_0 = 1$  und Nach Satz 10

$$h_i = ph_{i+1} + (1 - p)h_{i-1} \quad \text{für } i = 1, 2, \dots$$

Mit  $\Theta = \frac{1-p}{p}$  gilt somit (siehe Bsp. 9)

$$h_{i+1} - h_i = \Theta(h_i - h_{i-1}) \quad \text{für } i = 1, 2, \dots$$

und daher

$$h_{i+1} - h_i = \Theta^i(h_1 - h_0) = \Theta^i(h_1 - 1) \quad \text{für } i = 0, 1, \dots \quad (\star)$$

Sei nun  $p \leq \frac{1}{2}$ , also  $\Theta \geq 1$ .

Dann gilt

$$h_i - h_{i+1} \stackrel{(\star)}{=} \Theta^i(1 - h_1) \geq 1 - h_1$$

also

$$h_{i+1} \leq h_i - (1 - h_1) \quad \text{für } i = 0, 1, \dots$$

Angenommen  $h_1 < 1$ , dann würde daraus folgen, dass  $h_{i+1} - h_i \leq 0$  und damit wäre  $\lim_{i \rightarrow \infty} h_i = -\infty$ . Dies wäre offensichtlich ein Widerspruch!

Also muss  $h_1 = 1$  sein.

$$\stackrel{(\star)}{\Rightarrow} h_{i+1} - h_i = 0 \quad \text{für } i = 0, 1, \dots$$

$$\Rightarrow h_i = 1 \quad \text{für alle } i = 0, 1, \dots, \text{ falls } p \leq \frac{1}{2}$$

Insbesondere ist auch bei einem fairen Spiel ( $p = \frac{1}{2}$ ) die Ruinwahrscheinlichkeit  $h_i = 1$  für jedes Anfangskapital  $i$ .

Sei nun  $p > \frac{1}{2}$  (nicht-faires Spiel mit höherer Wahrscheinlichkeit Geld zu verlieren), also  $\Theta < 1$ . Aus  $(\star)$  folgt durch Summation (wie in Bsp. 9)

$$\begin{aligned} \sum_{i=0}^{j-1} (h_{i+1} - h_i) &= h_j - h_0 \\ &= h_j - 1 \\ &= (h_1 - 1) \sum_{i=0}^{j-1} \Theta^i \\ &= (h_1 - 1) \frac{1 - \Theta^j}{1 - \Theta} \\ \Rightarrow h_i &= 1 + (h_1 - 1) \frac{1 - \Theta^j}{1 - \Theta} \quad \text{für } j = 0, 1, \dots \end{aligned}$$

Für jedes  $h_1 \in [0, 1]$  erfüllt diese Gleichung die Gleichung  $(\star)$ .

Wegen  $0 \leq \lim_{j \rightarrow \infty} h_i = 1 + \frac{(h_1 - 1)}{1 - \Theta}$

$$\begin{aligned} 0 &\leq 1 + \frac{h_1 - 1}{1 - \Theta} \\ \Theta - 1 &\leq h_1 - 1 \\ h_1 &\geq \Theta \end{aligned}$$

Welches  $h_1 \in [\Theta, 1]$  liefert die gesuchten Wahrscheinlichkeiten?

Hinweis: Hier gibt es keine zweite Randbedingung, wie in Bsp. 9!

Sei  $\{X_n : n \in \mathbb{N}_0\}$  eine Markovkette mit Zustandsraum  $S$  und Übergangsmatrix  $(P_{ij})$ . Sei zudem  $A \subset S$  mit  $A \neq \emptyset$ . Sei  $T_i = \inf\{n \in \mathbb{N} : X_n \in A\}$  und  $h_i = P(T_i < \infty \mid X_0 = i)$ . Dann ist  $(h_i)_{i \in S}$  die **komponentenweise kleinste nicht-negative Lösung** des Gleichungssystems

$$x_i = 1 \text{ für alle } i \in A$$

$$x_i = \sum_{j \in S} p_{ij} x_j \text{ für alle } i \in S \setminus A$$

Das heißt

1.  $(x_i) = (h_i)$  ist eine nicht-negative Lösung.
2. Für jede Lösung  $(x_i)$  mit  $x_i \geq 0, i \in S$  gilt  $h_i \leq x_i$ .

### Beweis von Satz 11

Die Behauptung ist klar, wenn  $A = S$ . Deswegen sei  $A \neq S$ :

1.  $(h_i)$  ist eine nichtnegative Lösung nach Satz 10.
2. Sei  $(x_i)$  eine Lösung mit  $x_i \geq 0, i \in S$ .

$$\Rightarrow x_i = \overbrace{\sum_{j \in A} p_{ij}}^{:=r_i} \times 1 + \overbrace{\sum_{j \in A^c} p_{ij} x_j}^{:=Q}, i \in A^c$$

Für die Spaltenvektoren  $x = (x_i)_{i \in A^c}$ ,  $r = (r_i)_{i \in A^c}$  und die Teilmatrix  $Q = (p_{ij})_{i,j \in A^c}$  gilt durch iteratives Einsetzen

$$\begin{aligned} x &= r + Qx \\ &= r + Q(r + Qx) \\ &= r + Qr + Q^2x \\ &= r + Qr + Q^2(r + Qx) \\ &\vdots \\ x &= r + \sum_{k=1}^n Q^k r + Q^{n+1}x, n \geq 1 \end{aligned}$$

Daraus folgt also für alle  $i \in A^c, j \in A$  und  $n \geq 1$  durch Einsetzen, wobei  $Q^k r = p_{ij_k}^k p_{j_k j}$  und  $Q^{n+1}x$  als strikt positiver Wert für die untere Schranke ausgelassen werden kann, folgendes:

$$\begin{aligned} x_i &\geq p_{ij} + \sum_{k=1}^n \sum_{j_1, \dots, j_k \in A^c, j \in A} p_{ij_1} p_{j_1 j_2} \dots p_{j_k j} \\ &\Leftrightarrow x_i \geq P_i(X_1 \in A) + \sum_{k=1}^n P(X_1 \in A^c, \dots, X_k \in A^c, X_{k+1} \in A) \\ &\Leftrightarrow x_i \geq P_i(T \leq n+1) \\ &\Rightarrow x_i \geq \lim_{n \rightarrow \infty} P_i(T \leq n+1) = \lim_{n \rightarrow \infty} P_i(T \leq n) = h_i, i \in A^c \end{aligned}$$

Im Fall von  $p > \frac{1}{2}$  wurde gezeigt, dass die nichtnegativen Lösungen des Gleichungssystems (wie in Satz 11) charakterisiert sind durch

$$x_i = 1 + (x_1 - 1) \frac{1 - \Theta^i}{1 - \Theta}$$

mit  $x_1 \in [\Theta, 1]$ .

Aus Satz 11 lässt sich nun folgern, dass die Komponentenweise kleinste Lösung gegeben ist bei  $x_1 = \Theta$  und damit

$$h_i = 1 + (\Theta - 1) \frac{1 - \Theta^i}{1 - \Theta} = \Theta^i = \left( \frac{1-p}{p} \right)^i$$

Für die folgenden Schritte werden folgende Rechenregeln für  $\mathbb{R}^- = \mathbb{R} \cup \{-\infty, \infty\}$  benötigt:

- $-\infty < a < \infty \quad \forall a \in \mathbb{R}$
- $a + \infty = \infty + a = \infty \quad \forall a \in \mathbb{R} \cup \{\infty\}$
- $a - \infty = -\infty + a = -\infty \quad \forall a \in \mathbb{R} \cup \{-\infty\}$
- $a \times \infty = \infty \times a = \infty \quad \forall a \in \mathbb{R}^{(\cdot)+}$
- $a \times \infty = \infty \times a = -\infty \quad \forall a \in \mathbb{R}^{(\cdot)-}$
- $a \times (-\infty) = (-\infty) \times a = -\infty \quad \forall a \in \mathbb{R}^{(\cdot)+}$
- $a \times (-\infty) = (-\infty) \times a = \infty \quad \forall a \in \mathbb{R}^{(\cdot)-}$
- $0 \times \infty = \infty \times 0 = 0$  (spezielle Konnotation in diesem Kontext)
- Nicht definiert sind:  $\infty - \infty$  und  $-\infty + \infty$ .

Ist  $X$  eine Zufallsvariable mit Werten in  $\mathbb{N}_0 \cup \{\infty\}$ , dann ist

$$\begin{aligned} E[X] &= \sum_{n=0}^{\infty} nP(X = n) \\ &= \sum_{n=0}^{\infty} P(X > n) \\ &= \sum_{n=1}^{\infty} P(X \geq n) \end{aligned}$$

Insbesondere ist also falls  $P(X = \infty) > 0$  auch  $E[X] = \infty$ .

Sei  $\{X_n : n \in \mathbb{N}_0\}$  eine Markovkette mit Zustandsraum  $S$  und Übergangsmatrix  $(P_{ij})$ . Sei  $A \in S$  und  $A \neq \emptyset$ . Für  $i \in S$  sei

$$\begin{aligned} T &:= \inf\{n \in \mathbb{N}_0 : X_n \in A\} \\ k_i &:= E[T \mid X_0 = i] \end{aligned}$$

$(k_i)_{i \in S}$  ist die **komponentenweise kleinst enichtnegative Lösung** des Gleichungssystems  $(\star)$

$$\begin{aligned} k_i &= 0, \quad i \in A \\ k_i &= 1 + \sum_{j \in S \setminus A} p_{ij} k_j, \quad \text{für alle } i \in S \setminus A \end{aligned}$$

mit  $k_i \in [0, \infty]$ .

### Beweis von Satz 12

Die Behauptung ist klar für  $A = S$ . Sei also  $A \neq S$ .

- Für  $(k_i)_{i \in S}$  gilt

$$\begin{aligned} k_i &\in [0, \infty] \quad \forall i \in S \\ k_i &= 0 \quad \forall i \in A \end{aligned}$$

Für  $i \in S \setminus A$  gilt

$$\begin{aligned}
k_i &= \sum_{j \in S} E[T \mid X_0 = i, X_1 = j] p_{ij} \\
&= \sum_{j \in S} (1 + E[T \mid Y_0 = j]) p_{ij} \quad (\text{siehe Aufgaben}) \\
&= 1 + \sum_{j \in S \setminus A} p_{ij} k_j
\end{aligned}$$

$\Rightarrow (k_i)$  löst  $(\star)$ .

- Sie nun  $(x_i)_{i \in S}$  eine Lösung von  $(\star)$  mit  $x_i \in [0, \infty]$ . Zu zeigen ist, dass  $x_i \geq k_i \quad \forall i$ . Für  $i \in A$  gilt  $x_i = k_i = 0$ . Für die Spaltenvektoren

$$x = (x_i)_{i \in A^c} \quad \text{und} \quad e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

und

$$Q = (p_{ij})_{i,j \in A^c}$$

gilt (siehe auch Beweis Satz 11)

$$\begin{aligned}
x &= e + Qx \\
&\vdots \\
&= e + \sum_{k=1}^n Q^k e + Q^{n+1} x, \quad \text{für } n \geq 1
\end{aligned}$$

und somit für alle  $i \in A^c$  und  $n \geq 2$ , wobei ... und  $Q^{n+1}x$  als strikt positiver Wert für die untere Schranke ausgelassen werden kann,

$$\begin{aligned}
x_i &\geq 1 + \sum_{j \in A^c} p_{ij} + \sum_{k=2}^n \sum_{j_1, \dots, j_k \in A^c} p_{ij_1} \cdots p_{j_{k-1}j_k} \\
&\Leftrightarrow x_i \geq P_i(T > 0) + P_i(X_1 \in A^c) + \sum_{k=2}^n P_i(X_1 \in A^c, X_2 \in A^c, \dots, X_k \in A^c) \\
&\Leftrightarrow x_i \geq P_i(T > 0) + P_i(T > 1) + \sum_{k=2}^n P_i(T > k) \\
&\Leftrightarrow x_i \geq \sum_{k=0}^n P_i(T > k) \\
&\Rightarrow x_i \geq \sum_{k=0}^{\infty} P_i(T > k) = E_i[T] = k_i
\end{aligned}$$

Nun ist die erwartete Arbsorbtionszeit  $k_i = E_i[T]$  mit gesucht.

Offensichtlich ist  $k_0 = 0$  und  $k_M = 0$ . Aus Satz 12 folgt hier

$$(\star) \quad k_i = 1 + (1-p)k_{i-1} + pk_{i+1}$$

Sei zunächst  $p = \frac{1}{2}$ .

$$\begin{aligned}
k_i &= 1 + \frac{k_{i-1}}{2} + \frac{k_{i+1}}{2} \\
&\Leftrightarrow k_{i+1} - k_i = k_i - k_{i-1} - 2
\end{aligned}$$

Für steigende  $i$  fällt das Wachstum von  $k_i$  um den Faktor 2.

$$\Rightarrow k_{i+1} - k_i = k_1 - k_0 - 2i = k_1 - 2i \quad i = 0, \dots, M-1$$

Für  $j = 1, \dots, M$  ist

$$\begin{aligned} k_j - k_0 &= k_j = \sum_{i=0}^{j-1} (k_{i+1} - k_i) \\ &= jk_1 - 2 \sum_{i=0}^{j-1} i \\ &= jk_1 - (j-1)j \end{aligned}$$

Wegen  $k_M = 0$  folgt aus dieser Gleichung (durch einsetzen)  $k_1 = M - 1$  und somit

$$E[T \mid X_0 = j] = k_j = j(M - j), \quad j = 0, \dots, M$$

Sei nun  $p \neq \frac{1}{2}$ .

In diesem Fall lässt sich zeigen, dass  $k_i$  als Lösung von  $(\star)$  die Form

$$k_i = \frac{i}{1-2p} + \alpha + \beta \times \left( \frac{1-p}{p} \right)^i \quad \text{für } i = 0, \dots, M$$

haben muss für gewisse  $\alpha, \beta \in \mathbb{R}$ . Mit  $k_0 = k_M = 0$  folgt

$$\begin{aligned} 0 &= \alpha + \beta \\ 0 &= \frac{M}{1-2p} + \alpha + \beta \times \left( \frac{1-p}{p} \right)^M \\ \Rightarrow \alpha &= -\beta = \frac{-\frac{M}{1-2p}}{1 - \left( \frac{1-p}{p} \right)^M} \\ \Rightarrow E[T \mid X_0 = i] &= \frac{i}{1-2p} - \frac{M}{1-2p} \times \frac{1 - \left( \frac{1-p}{p} \right)^i}{1 - \left( \frac{1-p}{p} \right)^M}, \quad i = 0, \dots, M \end{aligned}$$

## Konvergenzsätze

Sei  $\{X_n : n \in \mathbb{N}_0\}$  eine Markovkette mit Zustandsraum  $S$  und Übergangsmatrix  $(P_{ij})$ .

Die stationäre Verteilung der Markovkette  $(\pi_i)$  mit

- $\pi_i \geq 0 \quad \forall i \in S$
- $\sum_{i \in S} \pi_i = 1$

ist definiert durch

$$\pi_j = \sum_{i \in S} \pi_i p_{ij}$$

oder ausgedrückt in Matrix Schreibweise

$$\begin{aligned} \pi &= \pi(P_{ij}) \\ \Leftrightarrow \pi((P_{ij}) - I) &= 0 \end{aligned}$$

## Bemerkung zu Satz 13

- Für jede Markovkette mit endlichem Zustandsraum existiert eine stationäre Verteilung (s. Aufgaben).
- Bezeichnet  $p^{(n)} = \left( P_i^{(n)} \right)_{i \in S}$ ,  $P_i^{(n)} = P(X_n = i)$  die Anfangsverteilung von  $X_n$  und ist die Anfangsverteilung stationär, dann gilt

$$P^{(n)} = P^{(0)} \Pi^n = p^{(0)} \Pi \Pi^{n-1} = \dots = P^{(0)},$$

d.h.  $P(X_n = i) = P(X_0 = i) \quad \forall n \geq 0 \quad \text{und} \quad i \in S$ .

- Ist der Zustandsraum endlich und existiert für jedes  $j \in S$  der Grenzwert

$$\pi_j = \lim_{n \rightarrow \infty} P(X_n = j),$$

dann muss  $(\pi_i)$  eine stationäre Verteilung sein, denn

$$\begin{aligned} \pi_i &\geq 0 \quad \forall i \in S \\ \sum_{i \in S} \pi_i &= \lim_{n \rightarrow \infty} \sum_{i \in S} P(X_n = i) = 1 \end{aligned}$$

und

$$\begin{aligned} \pi_j &= \lim_{n \rightarrow \infty} P(X_{n+1} = j) \\ &= \lim_{n \rightarrow \infty} \sum_{i \in S} P(X_n = i) p_{ij} \\ &= \sum_{i \in S} \lim_{n \rightarrow \infty} P(X_n = i) p_{ij} \\ &= \sum_{i \in S} \pi_i p_{ij} \quad \forall j \in S \end{aligned}$$

Gegeben sei eine Markovkette  $\{X_n : n \in \mathbb{N}_0\}$  mit Zustandsraum  $S = \{1, 2\}$  und Übergangsmatrix  $\Pi = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$  ( $a, b \in (0, 1)$ ) Es gilt

$$\Pi^{(n)} = \frac{1}{a+b} \begin{pmatrix} b + (1-a-b)^n a & a - (1-a-b)^n a \\ b - (1-a-b)^n b & b + (1-a-b)^n b \end{pmatrix}$$

und da  $\lim_{n \rightarrow \infty} (1-a-b)^n = 0$  folgt

$$\lim_{n \rightarrow \infty} \Pi^{(n)} = \frac{1}{a+b} \begin{pmatrix} b & a \\ b & a \end{pmatrix}$$

Damit ist für  $i \in S$

$$\lim_{n \rightarrow \infty} P_{i1}^{(n)} = \frac{b}{a+b}$$

und

$$\lim_{n \rightarrow \infty} P_{i2}^{(n)} = \frac{a}{a+b}$$

Für jede Anfangsverteilung  $(p_1 \ p_2) = (P(X_0 = 1), P(X_0 = 2))$  gilt

$$\begin{aligned}
\lim_{n \rightarrow \infty} (P(X_n = 1), P(X_n = 2)) &= \lim_{n \rightarrow \infty} (p_1 \ p_2) \Pi^n \\
&= \frac{1}{a+b} (p_1 \ 1-p_1) \begin{pmatrix} b & a \\ b & a \end{pmatrix} \\
&= \left( \frac{b}{a+b} \mid \frac{a}{a+b} \right)
\end{aligned}$$

Die Grenzverteilung ist also unabhängig von der Anfangsverteilung.

Ist die Anfangsverteilung gleich der Grenzverteilung, also  $p_1 = \frac{b}{a+b}$  und  $p_2 = \frac{a}{a+b}$ , dann gilt

$$(p_1 \ p_2) \Pi = \frac{1}{a+b} (b \ a) \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} = \frac{1}{a+b} (b \ a) = (p_1 \ p_2)$$

und daher  $(p_1 \ p_2) \Pi^n = (p_1 \ p_2)$ .

Sei  $\Pi = (P_{ij})$  die Übergangsmatrix einer Markovkette  $\{X_n : n \in \mathbb{N}_0\}$  mit stationäre Verteilung  $\pi = (\pi_1 \ \dots \ \pi_S)$ .

$\exists \ 0 < \epsilon < 1$ , sodass

$$\begin{aligned}
p_{ij} &\geq \epsilon \pi_j \quad \forall \ i, j = 1, \dots, S \\
\Rightarrow \sum_{j=1}^S |p_{ij}^n - \pi_j| &\leq 2(1-\epsilon)^n
\end{aligned}$$

Aus dem Satz folgt also, dass eine Markovkette für steigende  $n$  exponentiell zur stationären Verteilung konvergiert.

#### **Beweis Satz 14**

Setze  $M = \begin{pmatrix} \pi \\ \vdots \\ p_i \end{pmatrix} \in \mathbb{R}^{S \times S}$  und  $Q := \frac{1}{1-\epsilon}(\pi - \epsilon M)$ .  $M$  und  $Q$  sind positive stochastische Matrizen und

$$\pi = (1-\epsilon)Q + \epsilon M$$

Mit  $e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^S$  ist  $M = e\pi$  und

$$\begin{aligned}
M^2 &= e \overbrace{\pi e}^1 \pi = M \\
\Pi M &= \overbrace{\Pi e}^e \pi = e\pi = M \\
M\Pi &= e \overbrace{\pi \Pi}^\pi = M,
\end{aligned}$$

da  $\pi$  stationär ist. Daraus folgt also  $QM = M$  und  $MQ = M$ .

Das Produkt einer endlichen Folge von  $Q$ s und  $M$ s, die mindestens ein  $M$  enthält, ist gleich  $M$ .

Für jeden  $n \in \mathbb{N}$  gilt:

$$\begin{aligned}
\pi^n &= [(1-\epsilon)Q + \epsilon M]^n \\
&= (1-\epsilon)^n Q^n + \sum_{l=0}^{n-1} \binom{n}{l} (1-\epsilon)^l \epsilon^{n-l} M \\
&= (1-\epsilon)^n Q^n + (1 - (1-\epsilon)^n) M \\
\Rightarrow \pi^n - M &= (1-\epsilon)^n (Q^n - M) \quad \forall \ n \in \mathbb{N}_0
\end{aligned}$$

Für jede Zeile  $i = 1, \dots, S$  folgt damit (unter Verwendung der Dreiecksungleichung):



$$\begin{aligned}\sum_{j=1}^S |p_{ij}^n - \pi_j| &= (1 - \epsilon)^n \sum_{j=1}^S |\{Q_{ij}^n\} - \pi_j| \\ &\leq (1 - \epsilon)^n \sum_{j=1}^S (\{Q_{ij}^n\} + \pi_j) = 2(1 - \epsilon)^n\end{aligned}$$

Sei  $\Pi = (P_{ij})$  die Übergangsmatrix einer Markovkette mit Zustandsraum  $\{1, \dots, S\}$  und stationärer Verteilung  $\pi = (\pi_1 \dots \pi_S)$ .

Es existiere  $k \in \mathbb{N}$ ,  $\epsilon \in (0, 1)$ , sodass

$$\begin{aligned}p_{ij}^k &\geq \epsilon \pi_j \quad \forall \quad i, j = 1, \dots, S \\ \Rightarrow \sum_{j=1}^S |p_{ij}^n - \pi_j| &\leq 2(1 - \epsilon)^{\lfloor \frac{n}{k} \rfloor}\end{aligned}$$

### Beweis Satz 15

Wende hier Satz 14 auf  $\Pi^k$  an.

$\pi$  ist auch eine stationäre Verteilung für  $\Pi^k$  und

$$(\Pi^k)^m = \Pi^{km} = (P_{ij}^{km})$$

Damit liefert Satz 14:

$$\sum_{j=1}^S |P_{ij}^{km} - \pi_j| \leq 2(1 - \epsilon)^m \quad \forall \quad i = 1, \dots, S \quad \text{und} \quad m = 1, 2, 3, \dots$$

Sei  $n \in \mathbb{N}_0$ ,  $m = \lfloor \frac{n}{k} \rfloor$ ,  $r = n - km \geq 0$ .

$$\begin{aligned}\Rightarrow \sum_{j=1}^S |p_{ij}^n - \pi_j| &= \sum_{j=1}^S \left| \sum_{l=1}^S p_{il}^r (p_{lj}^{km} - \pi_j) \right| \\ &\leq \sum_{l=1}^S p_{il}^r \sum_{j=1}^S |p_{lj}^{km} - \pi_j| \\ &\leq 2(1 - \epsilon)^m = 2(1 - \epsilon)^{\lfloor \frac{n}{k} \rfloor}\end{aligned}$$

Unter den Voraussetzungen von Satz 15 gilt:

$\lim_{n \rightarrow \infty} P_{ij}^n$

- existiert für alle  $i, j \in S$ ,
- ist unabhängig von  $i$ ,
- und ist gegeben durch die stationäre Verteilung.

Ist  $\{X_n : n \in \mathbb{N}_0\}$  eine Markovkette mit Übergangsmatrix  $(P_{ij})$ , dann

$$\begin{aligned}
\sum_{j=1}^S |P(x_n = j) - \pi_j| &= \sum_{j=1}^S \left| \left( \sum_{i=1}^S P(X_0 = i) \right) p_{ij}^n - \pi_j \right| \\
&= \sum_{j=1}^S \left| \sum_{i=1}^S P(X_0 = i) (p_{ij}^n - \pi_j) \right| \\
&\leq \sum_{i=1}^S P(X_0 = i) \sum_{j=1}^S |p_{ij}^n - \pi_j| \\
&\leq 2(1 - \epsilon)^{\lfloor \frac{n}{k} \rfloor} \\
&\Rightarrow \lim_{n \rightarrow \infty} P(X_n = j) = \pi_j
\end{aligned}$$

Sind alle Einträge von  $\Pi^n$  positiv, dann ist die Voraussetzung  $p_{ij}^n \geq \epsilon \pi_j$  ( $i, j = 1, \dots, S$ ) mit  $\epsilon := \min_{i,j} p_{ij}^n$  erfüllt.

Sei nun  $a = b = 1$ , also

$$\Pi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

dann gilt

$$\Pi^n = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{für } n \text{ gerade} \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \text{für } n \text{ ungerade} \end{cases}$$

$\lim_{n \rightarrow \infty} P_{ij}^n$  existiert also nicht!

Die Markovkette zeigt periodisches Verhalten.

Die Periode  $d_i$  eines Zustands  $i$  einer Markovkette ist die größte ganze Zahl, die alle  $n \in \mathbb{N}$  mit  $P_{ii}^n > 0$  teilt:

$$d_i = \text{ggT}(\{n \in \mathbb{N} : P_{ii}^n > 0\})$$

$$d_i = \infty, \text{ falls } P_{ii}^n = 0 \quad \forall \quad n \in \mathbb{N}$$

Ein Zustandsraum mit Periode 1 heißt **aperiodisch**. Die Markovkette heißt aperiodisch, falls alle ihre Zustände aperiodisch sind.

- Für  $S = \{1, 2\}$  und  $\Pi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  gilt

$$d_1 = \text{ggT}(\{2, 4, 6, \dots\}) = 2$$

$$d_2 = 2$$

- Jeder Zustand  $i$  mit  $p_{ii} > 0$  ist aperiodisch.
- Für die einfache Irrfahrt auf  $\mathbb{Z}$  mit  $p \in (0, 1)$  gilt

$$d_i = 2 \quad \forall \quad i \in \mathbb{Z}$$

Für jeden aperiodischen Zustand  $i$  existiert  $n_0(i) \in \mathbb{N}$ , sodass

$$p_{ii}^n > 0 \quad \forall \quad n \geq n_0(i)$$

Sei

$$\Pi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

$P_{11}^1 = 0, P_{11}^2 = 0$ . Für  $n \geq 3$  ist aber

$$p_{11}^n \geq p_{12}p_{23}p_{33}^{n-3}p_{31} = \left(\frac{1}{2}\right)^{n-2} > 0$$

$\Rightarrow$  Zustand 1 ist aperiodisch. Es gilt nicht, dass  $P_{11}^n > 0$  für alle  $n \in \mathbb{N}$ , aber für  $n \geq 3 = n_0(1)$ .

### Beweis Satz 16

Zunächst ist für den Beweis folgendes Lemma notwendig:

Ist  $N \subset \mathbb{N}$ ,  $N \neq \emptyset$  abgeschlossen unter Addition, d.h.  $n + m \in N \quad \forall \quad n, m \in N$  und ist  $\text{ggT}(N) = 1$ , dann existiert  $n_0 \in \mathbb{N}$ , sodass  $n \in N \quad \forall n \geq n_0$ .

Um auch dies zu beweisen, muss zunächst gezeigt werden, dass  $N$  zwei aufeinander folgende Zahlen  $n_1, n_1 + 1$  enthält. Setze dazu

$$M := \{n - n' : n, n' \in N, n' < n\}$$

$$m := \min(M)$$

und sei  $n_1, n_2 \in N$ , sodass  $m = n_2 - n_1$ . Zeige  $m = 1$ . Nehme dafür zunächst an, dass  $m \neq 1$ .

$\Rightarrow m > 1$  und somit  $m > \text{ggT}(N)$ .  $m$  teilt nicht alle Elemente von  $N$ .

$\Rightarrow$  Es gibt  $n \in N$  und  $k \in \mathbb{N}_0$  mit

$$km < n < (k+1)m \quad ,$$

also

$$0 < n - km < m$$

und

$$\begin{aligned} n - km &= n - k(n_2 - n_1) \\ &= \underbrace{n + kn_1}_{\in N} - \underbrace{(k+1)n_2}_{\in N} \in M \end{aligned}$$

Widerpruch zur Definition von  $m$ ! Also muss  $m = 1$  und  $n_1, n_1 + 1 \in N$ .

Jede natürliche Zahl  $n \in \mathbb{N}$  lässt sich schreiben als

$$\begin{aligned} n &= qn_1 + r \quad \text{mit } q \in \mathbb{N}_0 \quad \text{und } r < n_1 \\ \Rightarrow n &= qn_1 + r(n_1 + 1 - n_1) \\ &= r(n_1 + 1) + (q - r)n_1 \end{aligned}$$

Für  $m \geq n_1^2$  ist  $q \geq n_1$ , also  $q - r > 0$  und es folgt  $n \in N$ . Es folgt also die Behauptung mit  $n_0 = n_1^2$ .

Nun zum Beweis von Satz 16: Sei  $i$  ein aperiodischer Zustand, für

$$N_i := \{n \in \mathbb{N} : P_{ii}^n > 0\}$$

gelte also  $\text{ggT}(N_i) = 1$ , insbesondere  $N_i \neq \emptyset$ .

$N_i$  ist abgeschlossen unter Addition, denn für  $n, m \in N$  gilt

$$P_{ii}^{n+m} = \sum_j p_{ij}^n p_{ji}^m \geq p_{ii}^n p_{ii}^m > 0$$

$\Rightarrow$  Es gibt  $n_0(i)$  mit  $P_{ii}^n > 0 \forall n \geq n_0(i)$ .

Eine Markovkette und ihre Übergangsmatrix heißen **irreduzibel**, falls es für alle  $i, j \in S$  ein  $n \geq 0$  gibt mit  $p_{ij}^n > 0$ .

- $(P_{ij})$  ist genau dann irreduzibel, wenn es für alle  $i, j \in S$  mit  $i \neq j$  eine Folge von Zuständen  $i_0, \dots, i_n \in S$  gibt, mit  $i_0 = i, \dots, i_n = j$  und  $p_{i_0 i_1} > 0, \dots, p_{i_{n-1} i_n} > 0$ .
- Jede irreduzibel Markovkette mit endlichem Zustandsraum  $S$  hat eine eindeutige stationäre Verteilung  $(\pi_i)_{i \in S}$ .

Positivität:

Es existiert ein  $i_0 \in S$  mit  $\pi_{i_0} > 0$ . Für jedes  $j \in S$  existiert  $n \in \mathbb{N}_0$  mit  $P_{i_0 j}^n$  und daher

$$\pi_j = \sum_{i \in S} \pi_i p_{ij}^n \geq \pi_{i_0} p_{i_0 j}^n > 0$$

Endeutigkeit:

Sei  $(\tilde{\pi}_i)_{i \in S}$  eine weitere stationäre Verteilung. Sei  $k \in S$  so, dass

$$\frac{\tilde{\pi}_k}{\pi_k} \leq \frac{\tilde{\pi}_i}{\pi_i} \quad \forall i \in S$$

$$\Rightarrow \tilde{\pi}_k = \sum_{j \in S} \tilde{\pi}_j p_{jk}^n \geq \frac{\tilde{\pi}_k}{\pi_k} \sum_{i \in S} \pi_i p_{ik}^n = \tilde{\pi}_k \quad \forall n \geq 0$$

$\Rightarrow$  Für  $n \geq 0$  und  $i \in S$  gilt

$$\tilde{\pi}_i p_{ik}^n = \frac{\tilde{\pi}_k}{\pi_k} \pi_i p_{ik}^n$$

Wähle  $n$  mit  $P_{ik}^n > 0$

$$\Rightarrow \tilde{\pi}_i = \frac{\tilde{\pi}_k}{\pi_k} \pi_i \quad ,$$

d.h.  $(\tilde{\pi}_i)$  ist proportional zu  $(\pi_i)$ . Und da  $\sum \tilde{\pi}_i = \sum \pi_i$

$$\Rightarrow \tilde{\pi}_i = \pi_i \quad \forall i \in S$$

Sei  $(P_{ij})$  die Übergangsmatrix einer irreduziblen Markovkette mit Zustandsraum  $S = \{1, \dots, s\}$  und mindestens einem aperiodischen Zustand.

$\Rightarrow$  Es existiert eine eindeutige stationäre Verteilung  $(\pi_i)_{i \in S}$  und

$$\lim_{n \rightarrow \infty} P_{ij}^n = \pi_j > 0 \quad \forall i, j \in S$$

### Beweis Satz 17

Zeige: Es gibt ein  $N \in \mathbb{N}$  (unabhängig von  $i, j$ ), sodass

$$P_{ij}^N > \beta \quad \forall i, j \in S$$

Sei  $i_0 \in S$  aperiodisch.

$\Rightarrow$  Es gibt ein  $n_0 \in \mathbb{N}$ , sodass

$$P_{i_0 i_0}^n > 0 \quad \forall n \geq n_0$$

Da  $(P_{ij})$  irreduzibel ist, gibt es für alle  $i, j \in S$  ein  $n_{ij} \in \mathbb{N}_0$  mit  $P_{ij}^{n_{ij}} > 0$ .  
Setze  $N := 2 \max\{n_{ij} : i, j \in S\} + n_0$ .

$$\Rightarrow p_{ij}^N \geq p_{ii_0}^{n_{ii_0}} \overbrace{p_{i_0 i_0}^{N - n_{ii_0} - n_{i_0 j}}^{\geq n_0}} p_{i_0 j}^{n_{i_0 j}} > 0 \quad \forall i, j \in S$$

Bemerkung: Unter den Voraussetzungen von Satz 17 gibt es für jeden Zustand  $i$  ein  $n_0(i) \in \mathbb{N}$ , sodass  $P_{ii}^n > 0 \quad \forall n \geq n_0(i)$ . Deher ist dann jeder Zustand aperiodisch.

### Interpretation der Grenzverteilung

Es gelte

$$\pi_j = \lim_{n \rightarrow \infty} P_{ij}^n = \lim_{n \rightarrow \infty} P(X_n = j \mid X_0 = i)$$

Ist  $(a_n)_{n=1}^\infty$  eine Folge reeller Zahlen mit  $\lim_{n \rightarrow \infty} a_n = a$ , dann gilt auch

$$\frac{1}{n} \sum_{k=0}^{n-1} a_k = a \quad ,$$

also hier

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} P_{ij}^k = \pi_j$$

und dahier ist

$$\begin{aligned} \frac{1}{n} \sum_{k=0}^{n-1} P_{ij}^k &= \frac{1}{n} \sum_{k=0}^{n-1} P(X_k = j \mid X_0 = i) \\ &= \frac{1}{n} \sum_{k=0}^{n-1} E[\mathbf{1}_{\{X_k=j\}} \mid X_0 = i] \\ &= E\left[\frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}_{\{X_k=j\}} \mid X_0 = i\right] \end{aligned}$$

Das heißt  $\pi_j$  ist der Grenzwert der erwarteten Zeitanteile, die die Markovkette in Zustand  $j$  verbringt.

HIER FEHLT EINE ERGÄNZUNG ZU AUFGABE 29

### Rekurrenz und Transienz

Sei  $\{X_n : n \in \mathbb{N}_0\}$  eine Markovkette mit Zustandsraum  $S$ .

Für jedes  $j \in S$  sei

$$\tau_j := \inf\{n \geq 1 : X_n = j\} \quad (\inf \emptyset = \infty)$$

In Worten:  $\tau_j$  ist der erste Zeitpunkt  $n \geq 1$ , zu dem die Markovkette  $j$  besucht, falls es einen gibt.

Für  $i, j \in S$  sei

$$f_{ij} = P(\tau_j < \infty \mid X_0 = i)$$

In Worten:  $f_{ij}$  ist die Wahrscheinlichkeit, dass  $j$  in endlicher Zeit ( $\geq 1$ ) erreicht wird bei Start in  $i$ .  
Insbesondere ist  $f_{jj}$  die Wahrscheinlichkeit einer Rückkehr.

Ein Zustand  $j \in S$  heißt **rekurrent**, falls  $f_{jj} = 1$ . Er heißt **transient**, falls  $f_{jj} < 1$ . Sind alle Zustände rekurrent oder transient, dann heißt auch die Markovkette rekurrent oder transient.

Sei  $S = \{1, 2, 3\}$  und

$$(P_{ij}) = \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

Dann ist

$$\begin{aligned} f_{11} &= \sum_{n=1}^{\infty} P(\tau_1 = n \mid X_0 = 1) \\ &= \sum_{k=1}^{\infty} P(X_n = 1, X_k = 2 \text{ mit } 1 \leq k < n \mid X_0 = 1) \\ &= \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n = 1 \end{aligned}$$

Folglich ist 1 rekurrent. Ebenso auch 2.

$$f_{33} = P(\tau < \infty \mid X_0 = 3) = P(X_1 = 3 \mid X_0 = 3) = \frac{1}{3} < 1$$

3 ist also transient und somit die Markovkette weder transient noch rekurrent.

Für  $j \in S$  sei

$$N_j = \sum_{n=1}^{\infty} \mathbf{1}_{\{X_n = j\}}$$

die Anzahl der Zeitpunkte  $\geq 1$  zu denen Zustand  $j$  besucht wird.

- Ist  $j$  rekurrent, dann gilt für alle  $i \in S$

$$P(N_j = \infty \mid X_0 = i) = f_{ij} \quad , \quad P(N_j = 0 \mid X_0 = i) = 1 - f_{ij}$$

und insbesondere

$$P(N_j = \infty \mid X_0 = j) = f_{jj} = 1 \quad \text{und} \quad E[N_j \mid X_0 = j] = \infty$$

- Ist  $j$  transient, dann gilt für alle  $i \in S$

$$P(N_j < \infty \mid X_0 = i) = 1 \quad \text{und} \quad E[N_j \mid X_0 = i] = \frac{f_{ij}}{1 - f_{jj}} < \infty$$

### Beweis Satz 18

Setze  $\sigma_j = \sup\{n \geq 1 : X_n = j\}$ , wobei  $\sup \emptyset = \infty$ . Ist  $\sigma_j$  endlich, dann ist  $\sigma_j$  der Zeitpunkt des letzten Aufenthalts in  $j$ . Für jedes  $n \in \mathbb{N}$  gilt

$$\begin{aligned}
P(\sigma_j = n \mid X_0 = i) &= P(X_n = j, X_m \neq j \text{ für alle } m > n \mid X_0 = i) \\
&= P(X_M \neq j \text{ für alle } m > n \mid X_n = j, X_0 = i) P(X_n = j \mid X_0 = i) \\
&= \lim_{n \rightarrow \infty} P(X_{n+1} \neq j, \dots, X_{n+m} \neq j \mid X_n = j, X_0 = i) P_{ij}^n \\
&= \lim_{m \rightarrow \infty} P(X_1 \neq j, \dots, X_m \neq j \mid X_0 = j) P_{ij}^n \\
&= (1 - f_{jj}) P_{ij}^n
\end{aligned}$$

$$\begin{aligned}
P(1 \leq \sigma_j < \infty \mid X_0 = i) &= \sum_{n=1}^{\infty} P(\sigma_j = n \mid X_0 = i) \\
&= (1 - f_{jj}) \sum_{n=1}^{\infty} P_{ij}^n \\
&= (1 - f_{jj}) \sum_{n=1}^{\infty} E[\mathbf{1}_{X_n=j} \mid X_0 = i] \\
&= (1 - f_{jj}) E[N_j \mid X_0 = i]
\end{aligned}$$

- Sei  $j$  rekurrent, also  $f_{jj} = 1$ .  
 $\Rightarrow P(\sigma_j = n \mid X_0 = i) = 0$  für alle  $n \in \mathbb{N}$ .  
 $\Rightarrow P(\sigma_j \in \{-\infty, \infty\}) = 1$ , also  $P(N_j \in \{0, \infty\} \mid X_0 = i) = 1$ .  
Es ist  $P(N_j = 0 \mid X_0 = i) = 1 - f_{ij}$  und es folgt  $P(N_j = \infty \mid X_0 = i) = f_{ij}$ .
- Sei  $j$  transient, also  $f_{jj} < 1$ .

$$\begin{aligned}
(1 - f_{jj}) E[N_j \mid X_0 = i] &= P(1 \leq \sigma_j < \infty \mid X_0 = i) \leq 1 \\
\Rightarrow E[N_j \mid X_0 = i] &< \infty \text{ und daher ist } P(N_j < \infty \mid X_0 = i) = 1. \\
\Rightarrow P(1 \leq \sigma_j < \infty \mid X_0 = i) &= P(1 \leq \sigma_j) \\
&= P(\tau_j < \infty \mid X_0 = i) \\
&= f_{ij}
\end{aligned}$$

$$\Rightarrow E[N_j \mid X_0 = i] = \frac{f_{ij}}{1 - f_{jj}}.$$

### Bemerkung zu Satz 18

- Wegen  $E[N_j \mid X_0 = j] = \sum_{n=1}^{\infty} P_{jj}^n$  liefert der Satz ein einfacheres Rekurrenzkriterium

$$j \text{ ist rekurrent} \Leftrightarrow \sum_{n=1}^{\infty} P_{jj}^n = \infty$$

- Im Allgemeinen kann eine Markovkette sowohl rekurrente als auch transiente Zustände haben.  
Ist  $\{X_n\}$  irreduzibel, dann ist entweder alle Zustände rekurrent oder alle sind transient.  
Denn: Falls es einen rekurrenten Zustand  $j$  gibt, dann gibt es für jedes  $i \in S$   $n_1, n_2 \geq 0$  mit  $P_{ij}^{n_1} > 0$ ,  $P_{ji}^{n_2} > 0$  und daher

$$\sum_{n=1}^{\infty} P_{ii}^n \geq \sum_{n=1}^{\infty} P_{ij}^{n_1} P_{jj}^n P_{ji}^{n_2} = \overbrace{P_{ij}^{n_1} P_{ji}^{n_2}}^{>0} \sum_{n=1}^{\infty} P_{jj}^n = \infty$$

$\Rightarrow i$  ist ebenfalls rekurrent.

- Ist  $j$  transient, dann gilt  $\lim_{n \rightarrow \infty} P_{ij}^n = 0 \quad \forall \quad i \in S$ , denn

$$\sum_{n=1}^{\infty} P_{ij}^n = E[N_j \mid X_0 = i] = \frac{f_{ij}}{1 - f_{jj}} < \infty$$

- Jede Markovkette mit endlichem Zustandsraum  $S = \{1, \dots, s\}$  hat mindestens einen rekurrenten Zustand, denn sonst wäre

$$1 = \lim_{n \rightarrow \infty} \sum_{j=1}^s P_{ij}^n = \sum_{j=1}^s \lim_{n \rightarrow \infty} P_{ij}^n = 0 \quad \text{Widerspruch!}$$

- Jede irreduzible Markovkette mit endlichem Zustandsraum ist rekurrent. Außerdem gilt für solche Markovketten

$$E[\tau_j \mid X_0 = i] < \infty$$

und daher ist

$$P(N_j = \infty \mid X_0 = i) = f_{ij} = 1 \quad \forall \quad i, j \in S$$

Wie betrachten wieder die einfache Irrfahrt auf  $\mathbb{Z}$ .

Da eine Rückkehr zum Zustand  $i$  immer nur in 2 Schritten möglich ist, ist die Periode  $d_i = 2$ .

$$P_{ii}^{2n} = \binom{2n}{n} p^n (1-p)^{2n-n}$$

Wegen  $\frac{\binom{2n}{n}}{4^n} \leq 1$  und  $\binom{2}{1} = 2$  gilt

$$\binom{2n}{n} \leq 4^n$$

Für  $p \neq \frac{1}{2}$  ist  $p(1-p) < \frac{1}{4}$  und daher (unter Nutzung geom. Reihe)

$$\sum_{n=1}^{\infty} P_{ii}^n = \sum_{n=1}^{\infty} P_{ii}^{2n} = \sum_{n=1}^{\infty} \binom{2n}{n} [p(1-p)]^n \leq \sum_{n=1}^{\infty} [4p(1-p)]^n < \infty$$

$\Rightarrow$  Für  $p \neq \frac{1}{2}$  ist die einfache Irrfahrt transient.

Andererseits gilt  $\binom{2n}{n} \geq \frac{4^n}{2\sqrt{n}}$  (Beweis durch vollst. Induktion hier ausgelassen).

Für  $p = \frac{1}{2}$  gilt daher

$$\sum_{n=1}^{\infty} P_{ii}^n = \sum_{n=1}^{\infty} \binom{2n}{n} [p(1-p)]^n \geq \sum_{n=1}^{\infty} \frac{4^n}{2\sqrt{n}} \left(\frac{1}{4}\right)^n = \frac{1}{2} \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} = \infty$$

$\Rightarrow$  für  $p = \frac{1}{2}$  ist die einfache Irrfahrt rekurrent.

Ein Zustand  $j$  heißt positiv rekurrent, falls  $E_j[T_j] < \infty$  und er heißt nullrekurrent, falls  $E_j[T_j] = \infty$ . Ist der Zustandsraum der Markovkette endlich, folgt aus Irreduzibilität bereits positive Rekurrenz und die Existenz einer eindeutigen stationären Verteilung. Die Konvergenz zu einer eindeutigen stationären Verteilung (unabhängig von der Anfangsverteilung) hingegen basiert (s. Satz 17) darauf, dass mindestens ein Zustand aperiodisch ist.

## Erneuerungsprozesse

Eine Erneuerung ist ein Ereignis, das zu zufälligen Zeitpunkten auftritt. Ein Erneuerungsprozess ist ein stochastischer Prozess, der die Anzahl der Erneuerungen bis zu einem Zeitpunkt  $t$  beschreibt. Seien  $Y_1, Y_2, \dots$  i.i.d die Wartezeiten, also die Zeiten zwischen aufeinanderfolgenden Erneuerungen, mit  $E[Y_i] = \mu \in (0, \infty)$  und  $\text{Var}[Y_i] = \sigma^2$ . Die Zeitpunkte der Erneuerungen seien definiert als



$$T_n = Y_1 + Y_2 + \dots + Y_n$$

Sei  $N(t)$  die Anzahl der Erneuerungen bis zu einem Zeitpunkt  $t$ :

$$N(t) = \max\{n \in \mathbb{N}_0 : T_n \leq t\}$$

In Worten:  $N(t)$  ist die maximale Anzahl an Erneuerungen, deren Gesamtdauer  $T_n$  nicht über  $t$  hinausgeht.

Für jeden Zeitpunkt  $t$  gilt

$$T_{N(t)} \leq t < T_{N(t)+1} \quad (\star)$$

Nach dem starken Gesetz der großen Zahlen gilt mit Wahrscheinlichkeit 1, dass

$$\lim_{n \rightarrow \infty} \frac{T_n}{n} = \lim_{n \rightarrow \infty} \frac{Y_1 + \dots + Y_n}{n} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n Y_i = \mu \quad (\star \star)$$

Es gilt zudem

$$\lim_{t \rightarrow \infty} N(t) = \infty$$

Um zu zeigen, dass dies gilt, sei anzunehmen, dass  $N(t) \leq M \quad \forall t$  mit  $N(t^*) = M$ . Das wiederum impliziert  $T_{M+1} > t \quad \forall t > t^*$ . Das wiederum impliziert

$$T_{M+1} = \overbrace{Y_1 + Y_2 + \dots + Y_M}^{=t^* < \infty} + Y_{M+1} = \infty$$

und damit  $Y_{M+1} = \infty$ . Das wäre ein Widerspruch, da dies  $P(Y = \infty) > 0$  voraussetzen würde, was gemäß  $E[Y] < \infty$  nicht der Fall ist.

Sei  $0 < \mu = E[Y_1] < \infty$ . Mit Wahrscheinlichkeit 1 gilt

$$\lim_{t \rightarrow \infty} \frac{N(t)}{t} = \frac{1}{\mu}$$

### Beweis Satz 19

Mit  $(\star)$  und  $(\star \star)$  folgt

$$\Rightarrow \frac{T_{N(t)}}{N(t)} \leq \frac{t}{N(t)} < \frac{T_{N(t)+1}}{N(t)+1} \frac{N(t)+1}{N(t)}$$

und

$$\begin{aligned} \Rightarrow \lim_{t \rightarrow \infty} \frac{T_{N(t)}}{N(t)} &= \mu \\ \lim_{t \rightarrow \infty} \frac{T_{N(t)+1}}{N(t)+1} &= \mu \\ \lim_{t \rightarrow \infty} \frac{N(t)+1}{N(t)} &= 1 \\ \Rightarrow \lim_{t \rightarrow \infty} \frac{t}{N(t)} &= \mu \end{aligned}$$

Zu betrachten sei ein Erneuerungsprozess  $\{N(t) : t \in \mathbb{N}_0\}$ , bei dem  $Y_1, Y_2, \dots$  i.i.d die Wartezeit bis zum Ausfall einer Glühbirne in einer Lampe mit einer Glühbirne beschreibt. Folglich ist  $T_n$  die Zeit, bis die  $n$ -te Glühbirne ausgefallen ist, bzw. ausgetauscht werden musste.

Nehmen wir an, wir beobachten in Tagen  $Y_1 = 100, Y_2 = 150, Y_3 = 125, Y_4 = 110$  und wir interessieren uns für  $N(270)$ .

$$N(270) = \max\{n \in \mathbb{N}_0 : T_n \leq 270\} = 2$$

Angenommen, wir beobachten an Tag 1000, dass die 7. Glühbirne ausgetauscht wird, was ist dann die erwartete Lebensdauer einer Glühbirne? Nach Satz 19 ist für einen unendlichen Zeithorizont

$$\frac{\text{Anzahl ausgetauschter Glühbirnen bis Zeitpunkt } t}{t} \rightarrow \frac{1}{\text{Erwartete Lebensdauer einer Glühbirne}}$$

$$\text{Also ist } \mu \approx \frac{1000}{7} = 143$$

Zu betrachten sei ein Erneuerungsprozess  $\{N(t) : t \in \mathbb{N}_0\}$ , bei dem  $Y_1, Y_2, \dots$  i.i.d. die Zeitabstände zwischen Schadensfällen ist. Folglich ist  $T_n$  die Zeit, bis die n-te Schadensmeldung eingegangen ist. Wenn die Frage nach dem erwarteten Zeitabends zwischen zwei Meldungen ist, könnte nach Satz 19 wieder

$$\mu \approx \frac{t}{\text{Anzahl Meldungen bis Zeit } t}$$

berechnet werden.

Nun erhalte die Versicherung zu den Zeiten  $T_1, T_2, \dots$  Schadensforderungen in Höhe von  $X_1, X_2, \dots$  i.i.d mit  $EX_i < \infty$ . Über die Gesamte Zeit häufen sich also

$$X(t) = \sum_{i=1}^{N(t)} X_i$$

an. Berechne  $\lim_{t \rightarrow \infty} \frac{X(t)}{t}$ .  
Da gilt

$$\lim_{t \rightarrow \infty} \frac{N(t)}{t} = \frac{1}{E[Y_1]}$$

und

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = E[X_i]$$

folgt mit  $\lim_{t \rightarrow \infty} N(t) = \infty$

$$\Rightarrow \lim_{n \rightarrow \infty} \frac{1}{N(t)} \sum_{i=1}^{N(t)} X_i = E[X_i]$$

Mittels Satz 19 erhalten wir

$$\lim_{t \rightarrow \infty} \frac{X(t)}{t} = \lim_{t \rightarrow \infty} \left( \frac{1}{N(t)} \sum_{i=1}^{N(t)} X_i \right) \frac{N(t)}{t} = \frac{E[X_1]}{E[Y_1]}$$

Sei  $\mu := E[Y_1], \sigma^2 := \text{Var}[Y_1]$  mit  $0 < \mu < \infty$  und  $0 < \sigma < \infty$ .

$\Rightarrow$  Für  $t \rightarrow \infty$  ist  $N(t)$  asymptotisch normalverteilt mit Parametern

$$\frac{t}{\mu} \quad \text{und} \quad \frac{t\sigma^2}{\mu^3},$$

das heißt

$$\lim_{t \rightarrow \infty} P\left(\frac{N(t) - \frac{t}{\mu}}{\sqrt{\frac{t\sigma^2}{\mu^3}}} < y\right) = \Phi(y) \quad \forall y \in \mathbb{R}$$

### Beweis Satz 20

Der Beweis nutzt  $P(N(t) < n) = P(T_n > t)$ , den zentralen Grenzwertsatz und folgendes Lemma:

Seien  $X_1, X_2, \dots$  Zufallsvariablen, sodass  $\lim_{n \rightarrow \infty} P(X_n > x) = G(x) \quad \forall x \in \mathbb{R}$ , wobei  $G$  stetig ist.

$\Rightarrow$  Für alle  $x, x_1, x_2, \dots \in \mathbb{R}$  mit  $\lim_{n \rightarrow \infty} X_n = x$  gilt

$$\lim_{n \rightarrow \infty} P(X_n > x_n) = G(x)$$

Sei nun  $y \in \mathbb{R}$ . Für alle  $t > 0$  gilt

$$\begin{aligned} P\left(\frac{N(t) - \frac{t}{\mu}}{\sqrt{\frac{t\sigma^2}{\mu^3}}} < y\right) &= P\left(N(t) < \frac{t}{\mu} + y\sigma\sqrt{\frac{t}{\mu^3}}\right) \\ &= P(N(t) < n_t) \end{aligned}$$

mit  $n_t = \left\lceil \frac{t}{\mu} + y\sigma\sqrt{\frac{t}{\mu^3}} \right\rceil$ .

Sei  $t > 0$  so groß, dass  $n_t \geq 1$ . Dann gilt

$$\begin{aligned} P(N(t) < n_t) &= P(T_{n_t} > t) \\ &= P\left(\frac{T_{n_t} - n_t\mu}{\sigma\sqrt{n_t}} > \frac{t - n_t\mu}{\sigma\sqrt{n_t}}\right) \end{aligned}$$

Und mit  $\lim_{t \rightarrow \infty} \frac{n_t}{t} = \frac{1}{\mu}$  ist

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{t - n_t\mu}{\sigma\sqrt{n_t}} &= \lim_{t \rightarrow \infty} \frac{t - \left(\frac{t}{\mu} + y\sigma\sqrt{\frac{t}{\mu^3}}\right)\mu}{\sigma\sqrt{n_t}} \\ &= \lim_{t \rightarrow \infty} \frac{-y\sqrt{\frac{t}{\mu}}}{\sqrt{n_t}} \\ &= -y \end{aligned}$$

Mit dem zentralen Grenzwertsatz und dem Lemma folgt also

$$\lim_{t \rightarrow \infty} P(N(t) < n_t) = 1 - \Phi(-y) = \Phi(y)$$

## Poisson-Prozesse

Eine Zufallsvariable  $X$  heißt exponentialverteilt mit Parameter  $\lambda > 0$  ( $X \sim \text{EXP}(\lambda)$ ), falls  $X$  die Dichte

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{falls } x > 0 \\ 0, & \text{sonst} \end{cases}$$

hat.

Sei  $\lambda$  eine Konstante  $> 0$ . Seien  $Y_1, Y_2, \dots$  i.i.d. exponentialverteilte Zufallsvariablen für alle  $n = 1, 2, \dots$

Setze

$$\begin{aligned} T_0 &:= 0 \\ T_n &:= Y_1 + \dots + Y_n \\ N(t) &:= \max\{n \in \mathbb{N}_0 : T_n \leq t\}, \quad t \geq 0 \end{aligned}$$

Dann heißt  $\{N(t) : t \geq 0\}$  Poisson-Prozess mit Intensität  $\lambda$ .

Ein stochastischer Prozess  $\{N(t) : t \geq 0\}$  heißt ein Prozess mit unabhängigen Zuwächsen, falls für je endlich viele Zeitpunkte  $0 = t_0 < t_1 < \dots < t_k$  die Zufallsvariablen

$$\begin{aligned} &N(t_0), \\ &N(t_1) - N(t_0), \\ &\vdots \\ &N(t_k) - N(t_{k-1}) \end{aligned}$$

unabhängig sind.

Sei  $\{N(t) : t \geq 0\}$  ein Poisson-Prozess mit Intensität  $\lambda > 0$ .

$\Rightarrow \{N(t) : t \geq 0\}$  hat unabhängige Zuwächse und für  $0 \leq s < t$  gilt

$$N(t) - N(s) \sim POI(\lambda(t-s)) \quad ,$$

$$P(N(t) - N(s) = n) = e^{-\lambda(t-s)} \frac{[\lambda(t-s)]^n}{n!} \quad \text{für } n = 0, 1, \dots$$

### Beweis Satz 21

Seien  $Y_i, T_n, N(t)$  wie in der Definition eines Poisson-Prozesses.

Sei  $k \geq 2, 0 = t_0 < t_1 < \dots, t_k, n_1, \dots, n_k \in \mathbb{N}$  beliebig fest.

$$\Rightarrow P(N(t_i) - N(t_{i-1}) = n_i, i = 1, \dots, k) = P((T_1, \dots, T_{m+1}) \in B \times (t_k, \infty)),$$

wobei  $m = n_1 + \dots + n_k$  und  $B = \{(s_1, \dots, s_m) \in \mathbb{R}^m : 0 < s_1 < \dots < s_m, |\{j : t_{i-1} < s_j \leq t_i\}| = n_i, i = 1, \dots, k\}$ .  $(T_1, \dots, T_{m+1})$  hat die Dichte

$$f(s_1, \dots, s_{m+1}) = \begin{cases} \lambda^{m+1} e^{-\lambda s_{m+1}} & , \text{ falls } 0 < s_1 < \dots < s_{m+1} \\ 0 & , \text{ sonst} \end{cases}$$

$$\begin{aligned} P(N(t_i) - N(t_{i-1}) = n_i, i = 1, \dots, k) &= \int_{B \times (t_k, \infty)} f(s_1, \dots, s_{m+1}) d(s_1, \dots, s_{m+1}) \\ &= \lambda^{m+1} \left( \int_B 1 d(s_1, \dots, s_m) \right) \int_{t_k}^{\infty} e^{-\lambda s_{m+1}} ds_{m+1} \end{aligned}$$

An dieser Stelle ist zu nutzen, dass

$$\begin{aligned} \int_B 1 d(s_1, \dots, s_m) &= \int_{B_1} 1 d(x_1, \dots, x_{n_1}) \int_{B_2} 1 d(x_1, \dots, x_{n_2}) \dots \int_{B_k} 1 d(x_1, \dots, x_{n_k}) \\ &= \prod_{i=1}^k \frac{(t_i - t_{i-1})^{n_i}}{n_i!} \end{aligned}$$

und

$$\int_{t_k}^{\infty} e^{-\lambda s_{m+1}} ds_{m+1} = \frac{e^{-\lambda t_k}}{\lambda}$$

Deswegen gilt

$$\begin{aligned} P(N(t_i) - N(t_{i-1}) = n_i, i = 1, \dots, k) &= \lambda^{1+\sum_{i=1}^k n_i} \left( \prod_{i=1}^k \frac{(t_i - t_{i-1})^{n_i}}{n_i!} \right) \frac{e^{-\lambda t_k}}{\lambda} \\ &= \prod_{i=1}^k \frac{(t_i - t_{i-1})^{n_i}}{n_i!} e^{t_i - t_{i-1}} \end{aligned}$$

Sei  $\{N(t) : t \geq 0\}$  ein Poisson-Prozess mit Intensität  $\lambda > 0$ . Berechne  $P(N(1) = 1, N(3) = 5)$ .

$$\begin{aligned} P(N(1) = 1, N(3) = 5) &= P(N(1) = 1, N(3) - N(1) = 4) \\ &= P\left(\overbrace{N(1) = 1}^{\sim \text{POI}(\lambda)}\right) P\left(\overbrace{N(3) - N(1) = 4}^{\sim \text{POI}(2\lambda)}\right) \\ &= \frac{e^{-\lambda} \lambda}{1!} \frac{e^{-2\lambda} (2\lambda)^4}{4!} \\ &= \frac{2}{3} e^{-3\lambda} \lambda^5 \end{aligned}$$

Ab 09:00 erfolgen Anrufe gemäß eines Poisson-Prozesses. Gegeben, dass bis 09:30 genau ein Anruf eingeht, wie groß ist die bedingte Wahrscheinlichkeit, dass er vor 09:20 eingeht, wenn  $t$  in Stunden gemessen wird?

$$\begin{aligned} P\left(T_1 < \frac{1}{3} \mid N\left(\frac{1}{2}\right) = 1\right) &= \frac{P(T_1 < \frac{1}{3}, N(\frac{1}{2}) = 1)}{P(N(\frac{1}{2}) = 1)} \\ &= \frac{P(T_1 < \frac{1}{3}, N(\frac{1}{2}) - N(\frac{1}{3}) = 0)}{P(N(\frac{1}{2}) = 1)} \\ &= \frac{P(T_1 < \frac{1}{3}) P(N(\frac{1}{2}) - N(\frac{1}{3}) = 0)}{P(N(\frac{1}{2}) = 1)} \\ &= \frac{\frac{e^{-\lambda \frac{1}{3}}}{1!} \left(\frac{\lambda}{3}\right)^1 \frac{e^{-\lambda(\frac{1}{2}-\frac{1}{3})}}{0!} \left(\lambda(\frac{1}{2} - \frac{1}{3})\right)^0}{\frac{e^{-\lambda \frac{1}{2}}}{1!} \left(\frac{\lambda}{2}\right)^1} \\ &= \frac{2}{3} \end{aligned}$$

Das entspricht genau  $\frac{2}{3}$  der Zeit!