Stochastic processes 1

Author:
O.Univ.-Prof. Dr.
Jürgen Pilz
Institute of Statistics

Prepared by:
Dženana Alagić
Vedo Alagić

Table of Contents

1 Introduction

1.1 Definition: what is a stochastic process?

Random systems are of complex nature. Usually, such systems cannot be described by a single random variable: they evolve in space and time. Description can be based on a family of random variables $\{X_t = X(t,\cdot) : t \in T\}$ where T is the index set (parameter set), e.g. T is the set of points in time and/or space.

Examples:

- Energy consumption over time.
- Growth of population over time (demography!).
- Capital growth over time (stocks, portfolios, ...).
- Dynamics of queuing systems (in banks, shops, computer networks, access to URL's, ...).
- Spatial structure of materials, substances, ...
- Spatial structure of aquifers, oil and gas deposits.
- Spatio - temporal spread of pollution sources (oil spills, radioactive contamination, ...).
- Neurobiology: spatio - temporal brain activities ... .

Definition 1.1 *A stochastic process means a family $\varkappa = \{X_t = X(t,\cdot) : t \in T\}$ of random variables X_t which are defined on a joint probability space $[\Omega, \mathscr{L}, P]$*

$$X_t : \Omega \longrightarrow E \quad \forall t \in T$$

where T denotes the index set (parameter set) and E denotes the state space of the process $\varkappa$.

Remark:

a) $\varkappa$ is called a discrete process if $T = \mathbb{N}$ (or $\mathbb{Z}$).

b) For $T = \mathbb{R}^+$ (or $T = \mathbb{R}^1$), $\varkappa$ is called a continuous process.

c) For $dim(T) \geq 2$, $\varkappa$ is called a random field (subject of Master course on "Spatial Statistics").

d) For $dim(E) \geq 2$, the elements of X_t of $\varkappa$ form a vector-valued random process (field).

Definition 1.2 *Let $\varkappa = \{X(t,\cdot) : t \in T\}$ be a stochastic process defined on $[\Omega, \mathscr{L}, P]$ with index set T. Then, for each $\omega \in \Omega$, the mapping*

$$X(\cdot,\omega) : T \longrightarrow E \quad with \quad X(\cdot,\omega)(t) := X(t,\omega)$$

is called the path (trajectory) of $\varkappa$ with respect to ω.

Remark: The mapping $X(\cdot,\cdot)$ defined on $T \times \Omega$ is called the random function associated with $\varkappa$. Its sample space is formed by the set of all possible paths.

Geometric visualization:

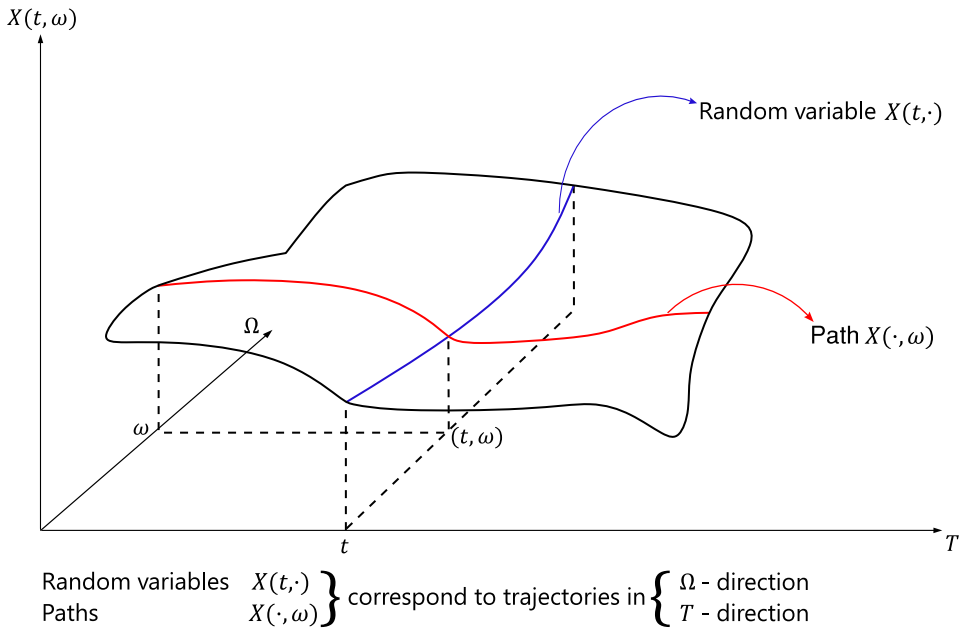


Figure 1.1: $\varkappa$ can be interpreted as "surface" defined over the $T \times \Omega$ - plane.

Well known fact: Random variables X are completely characterized by their probability law P_X (distribution function F_X). A "similar" assertion holds for random processes:

Theorem 1.1 *The random process $\varkappa = \{X(t,\cdot) : t \in T\}$ is completely determined by the family*

$$\mathscr{P}_\varkappa := \{P_{(X_{t_1},X_{t_2},\ldots,X_{t_n})} : n \in \mathbb{N}, \underbrace{t_1, t_2, \ldots, t_n}_{\text{pairwise different}} \in T\}$$

of all of its finite-dimensional distributions.

Problems:

- For increasing n these distributions are hard to determine.
- The random variables $X_{t_1}, \ldots, X_{t_n}$ are not independent, i.e. $P_{(X_{t_1},\ldots,X_{t_n})} \neq P_{X_{t_1}} \otimes \cdots \otimes P_{X_{t_n}}$ and thus, in general, $F_{(X_{t_1},\ldots,X_{t_n})} \neq \prod_{i=1}^{n} F_{X_{t_i}}$.
- Dependence structure (correlation structure) is difficult to describe.

1.2 Projective families, Kolmogorov's existence theorem

Clearly, there are some natural relationships between the finite dimensional distributions within $\mathscr{P}_{\mathscr{X}}$. This, in turn, means that not every possible family of probability measures

$$\mathscr{P} = \{P_{(t_1,t_2,\ldots,t_n)} : n \in \mathbb{N}; t_1, t_2, \ldots, t_n \in T \text{ p.w. different}\}$$

can serve as family of finite-dimensional distributions of a random process. The analysis of the requirements on such a family leads to the concept of projective families, whereby we limit ourselves to $E \subseteq \mathbb{R}^1$.

Definition 1.3 *A family $\mathscr{P} = \{P_{(t_1,t_2,\ldots,t_n)} : n \in \mathbb{N}; t_1, t_2, \ldots, t_n \in T$ p.w. different$\}$ of probability measures is called a projective family if it holds:*

a) Symmetry: For all $n \in \mathbb{N}$, all p.w. different $t_1, t_2, \ldots, t_n \in T$, all $B_1, B_2, \ldots, B_n \in \mathscr{L}^1(\sigma - algebra$ of the Borel sets of $\mathbb{R}^1)$, and all permutations π of $\{1, 2, \ldots, n\}$

$$P_{(t_{\pi(1)},t_{\pi(2)},\ldots,t_{\pi(n)})}(B_{\pi(1)} \times B_{\pi(2)} \times \ldots \times B_{\pi(n)}) = P_{(t_1,t_2,\ldots,t_n)}(B_1 \times B_2 \times \ldots \times B_n).$$

b) Consistency: For all $n \in \mathbb{N}$, all p.w. different $t_1, t_2, \ldots t_n, t_{n+1} \in T$, and all $B_1, B_2, \ldots, B_n \in \mathscr{L}^1$

$$P_{(t_1,t_2,\ldots,t_n,t_{n+1})}(B_1 \times B_2 \times \ldots \times B_n \times \mathbb{R}^1) = P_{(t_1,t_2,\ldots,t_n)}(B_1 \times B_2 \times \ldots \times B_n).$$

Remark:

a) Symmetry thus means permutation invariance of the distributions.

b) Consistency means compatibility w.r.t. marginalization, i.e. all marginal distributions are proper.

Illustration of consistency: $n = 1, B_1 = (-\infty, a]$ for some $a \in \mathbb{R}^1$

$$\underset{b)}{\Rightarrow} \quad P_{(t_1,t_2)}(B_1 \times \mathbb{R}^1) = \int_{-\infty}^{a} \int_{-\infty}^{+\infty} dP_{(t_1,t_2)}(x, y) = \int_{-\infty}^{a} dP_{t_1}(x) = P_{t_1}(B_1).$$

In more familiar terms: b) requires the existence of a random vector (X_1, X_2) with cdf $F_{(X_1,X_2)}$ such that

$$P_{(t_1,t_2)}(B_1 \times \mathbb{R}^1) = F_{(X_1,X_2)}(a, \infty) = \int_{-\infty}^{a} \int_{-\infty}^{+\infty} dF_{(X_1,X_2)}(x_1, x_2) = \int_{-\infty}^{a} dF_{X_1}(x_1) = F_{X_1}(a)$$

where F_{X_1} is the marginal distribution of X_1.

Corollary 1.1 *The symmetry property implies that a projective family $\mathscr{P}$ with $T \subseteq \mathbb{R}^1$ is completely determined by the subfamily $\mathscr{P}' = \{P_{(t_1,t_2,\ldots,t_n)} : n \in \mathbb{N}; t_1 < t_2 < \cdots < t_n \in T\}$ (in the sequel we therefore identify $\mathscr{P}$ and $\mathscr{P}'$).*

The following theorem plays a central role in the theory of stochastic processes. Its proof is beyond the scope of our course, it can be found e.g. in P. Billingsley: Probability and Measure, 3rd ed., Wiley, New York 1995.

Theorem 1.2 (Kolmogorov's existence theorem) *Let*

$$\mathscr{P} = \{P_{(t_1,t_2,\ldots,t_n)} : n \in \mathbb{N}; t_1, t_2, \ldots, t_n \in T \text{ p.w. different}\}$$

be a projective family. Then there exists a probability space $[\Omega, \mathscr{L}, P]$ *and a stochastic process* $\varkappa = \{X(t,\cdot) : t \in T\}$ *(with state space* $E \subseteq \mathbb{R}^1$*) defined on this space such that, for all* $n \in \mathbb{N}$ *and p.w. different* $t_1, t_2, \ldots, t_n \in T$ *it holds*

$$P_{(X_{t_1},X_{t_2},\ldots,X_{t_n})} = P_{(t_1,t_2,\ldots,t_n)}.$$

Roughly spoken: every projective family of measures forms the family of finite-dimensional distribution of some suitable stochastic process.

Remark: More generally, Kolmogorov's existence theorem holds for projective families and stochastic processes for which the state space E is a Banach space (endowed with the Borel-σ-algebra of open subsets of E).

Example 1.1 (Projective family associated with the Poisson process) *Let* $\lambda > 0$. *The family*

$$\mathscr{P} = \{P_{(t_1,t_2,\ldots,t_n)} : n \in \mathbb{N}; t_1 < t_2 < \cdots < t_n \in \mathbb{R}^+\}$$

of discrete probability measures with densities (with respect to the counting measure)

$$f_{(t_1,t_2,\ldots,t_n)}(x_1, x_2, \ldots, x_n) := \begin{cases} \prod_{i=1}^{n} \frac{[\lambda(t_i - t_{i-1})]^{x_i - x_{i-1}}}{(x_i - x_{i-1})!} e^{-\lambda(t_i - t_{i-1})} & \text{if } x_1 \le x_2 \le \cdots \le x_n \in \mathbb{N} \\ 0, & \text{else} \end{cases}$$

where $t_0 := 0, x_0 := 0$, *is symmetric and consistent.*
Every stochastic process having this family as family of finite-dimensional distributions is called Poisson process with parameter λ. ♢

Interpretation: $PP(\lambda)$ is a random point configuration obtained when rolling a dice and, for each infinitesimal interval $[t, t+dt)$, a decision is made, independently of the preceding results, whether to place a "point" in to this interval. Identify "points" with random events.
Define $N(T)$ = Random number of points (events) in the interval $(0, t]$, e.g. # radioactive decays, # of costumers, vehicles etc.
Then it holds:

$$P(\{N(t_1) = x_1\} \cap \ldots \cap \{N(t_n) = x_n\}) = \prod_{i=1}^{n} \frac{[\lambda(t_i - t_{i-1})]^{x_i - x_{i-1}}}{(x_i - x_{i-1})!} e^{-\lambda(t_i - t_{i-1})}$$

for all configurations $0 =: t_0 < t_1 < \cdots < t_n \in \mathbb{R}^+$ and $0 := x_0 \le x_1 \le \cdots \le x_n \in \mathbb{N}$.

Example 1.2 (Projective family associated with the Wiener Process (Brownian motion)) *Let $\sigma > 0$. The family $\mathscr{P} = \{P_{(t_1,t_2,\ldots,t_n)} : n \in \mathbb{N}; t_1 < t_2 < \cdots < t_n \in \mathbb{R}^+\}$ of continuous probability measures with densities (w.r.t. the Lebesgue measure)*

$$f_{(t_1,t_2,\ldots,t_n)}(x_1, x_2, \ldots, x_n) := \prod_{i=1}^{n} [2\pi\sigma^2(t_i - t_{i-1})]^{-\frac{1}{2}} \exp\left(-\frac{(x_i - x_{i-1})^2}{2\sigma^2(t_i - t_{i-1})}\right)$$

where $t_0 := 0$, $x_0 := 0$, is symmetric and consistent.
Every stochastic process having this family as family of finite-dimensional distributions is called Brownian motion[1] (process) or Wiener process[2]. ♢

Interpretation: Brownian motion is a stochastic process $\{X(t,\cdot) : t \geq 0\}$ with the properties

a) $X(t + \Delta t) - X(t) \sim N(0, \sigma^2 \Delta t); t > 0, \Delta t > 0.$

b) For any pair of disjoint intervals $(t_1, t_2]$ and $(t_3, t_4]$ with $0 \leq t_1 < t_2 \leq t_3 < t_4$ the increments $[X(t_4,\cdot) - X(x_3,\cdot)]$ and $[X(t_2,\cdot) - X(x_1,\cdot)]$ are (stochastically) independent. Correspondingly, for all $n \geq 2$

$$\begin{pmatrix} X_{t_1} - X_{t_0} \\ X_{t_2} - X_{t_1} \\ \vdots \\ X_{t_n} - X_{t_{n-1}} \end{pmatrix} \quad \sim \quad N_n(\underline{0}, \sigma^2 D)$$

where $t_0 = 0$ and $D = diag(t_1 - t_0, t_2 - t_1, \ldots, t_n - t_{n-1})$.

c) The paths of the process are continuous functions of the time.

Roughly spoken: Brownian motion is a stochastic process with independent and normally distributed increments and given initial state $x_0 \in \mathbb{R}^1$.
Many applications in stochastic finance: modeling the dynamics of stokes, currencies, interest rates etc. Basic building blocks of stochastic differential equations, stochastic integrals in physics, chemistry, biology, etc.

1.3 Classification of stochastic process

Let $\varkappa = \{X_t = X(t,\cdot) : t \in T\}$, T- index set, $X_t : [\Omega, \mathscr{L}] \to [E, \varepsilon]$, E- state space, ε- Borel σ-algebra of open subsets of E.
Here we will consider the (most important) special cases where $E \subseteq \mathbb{R}^1 (\varepsilon \subseteq \mathscr{L}^1)$ and

- T is discrete ($T = \mathbb{Z}$ or $T = \mathbb{N}$) : $\varkappa =$ discrete stochastic process.
- T is continuous ($T = \mathbb{R}^1$ or $T = \mathbb{R}^+$) : $\varkappa =$ continuous stochastic process.

[1] Motion of suspended particles (Robert Brown 1827).
[2] Norbert Wiener 1923 established mathematical foundations.

- If E is discrete ($E \subseteq \mathbb{Z}$) : $\varkappa$ = stochastic chain.

Classification:

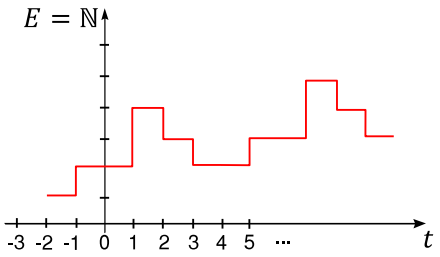
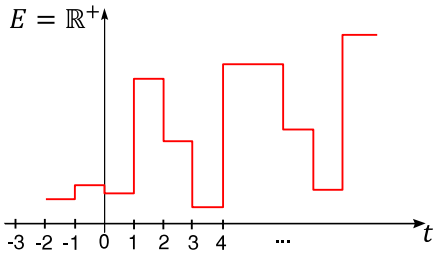
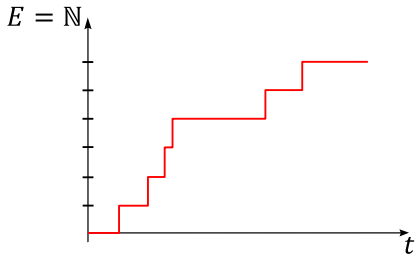
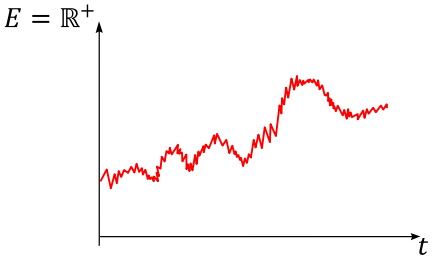
Index set	E discrete	E continuous
T discrete	Discrete stochastic chain  # accidents (crimes,...) per hour (day,...)	Discrete stochastic process  Currency rates per day, commodity prices/10 min
T continuous	Continuous stochastic chain  Radioactive decays	Continuous stochastic process  EEG, energy consumption,...

Figure 1.2: Classification of stochastic processes

Most important and most useful class of stochastic processes in applications are Markov processes[3].
Markov property (roughly spoken): Future only depends on present states, not on past records.
Examples:

a) Gains/losses in a casino.

b) Sequence of cards after shuffling.

c) Sequence of generation.

d) More general: branching processes.

[3] Date back to A. N. Markov 1906 - a student of Chebyshev.

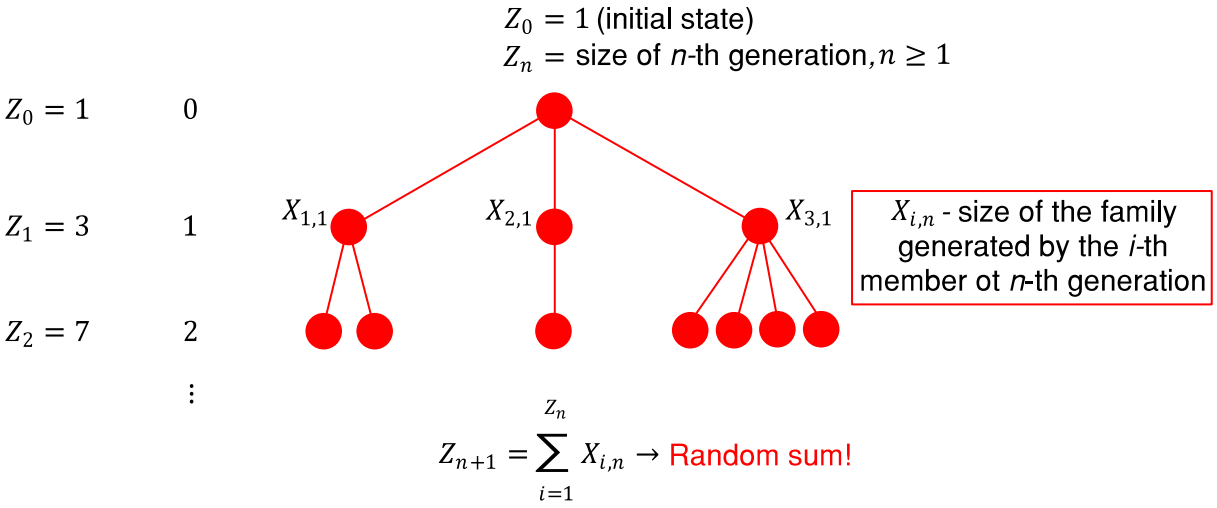


Figure 1.3: Generation tree

In general: $\varkappa = \{X_t : t \in T\}$ with $X_t : [\Omega, \mathscr{L}] \to [E, \varepsilon]$ is called a **Markov process** if it holds

$$P(X_{t_{n+1}} \in E_{n+1} | X_{t_n} \in E_n, X_{t_{n-1}} \in E_{n-1}, \ldots, X_{t_1} \in E_1) = P(X_{t_{n+1}} \in E_{n+1} | X_{t_n} \in E_n)$$

$$\forall n \geq 1, \overset{n+1}{\underset{i=1}{\forall}} E_i \in \varepsilon, \overset{n+1}{\underset{i=1}{\forall}} t_i \in T : t_1 < t_2 < \ldots < t_n < t_{n+1}.$$

Special cases: For $E = \mathbb{R}^1 (\varepsilon \in \mathscr{L}^1) : \varkappa$ - real valued Markov process, for $E = \mathbb{Z} : \varkappa$ - Markov chain.

2 Discrete Markov chain

In this chapter: both E (state space) and T (index set) are discrete. Without loss of generality: $T \subseteq \mathbb{N}$, $E \subseteq \mathbb{Z} \Rightarrow \varkappa = \{X_k\}_{k=0,1,2,\ldots}$ sequence of discrete random variables. X_k then describes the random state at time k.

2.1 Introductory example

Let us consider the following version of:
Gambler's ruin: Sequence of (independent) games or lotteries with:

(i) $X_0 = x_0, 0 < x_0 < K$; starting capital at time $k = 0$,

(ii) in the $k-$th game ($k \geq 1$) we gain a unit with probability p, $0 < p < 1$ and lose a unit with probability $q = 1 - p$,

(iii) we stop whenever our capital is zero or has reached a predetermined amount $K \in \mathbb{N}$.

Obviously: $\{X_k\}_{k=0,1,2,\ldots}$ is a Markov Chain (MC) with $E = \{0, 1, 2, \ldots, K\}, T = \mathbb{N}$.
1st problem: one - step probabilities for a transition from $i \in E$ to $j \in E$ in the n - th game (step):

$$\begin{aligned}
p_{j,j+1}^{(m-1,m)} &:= P(X_m = j+1 | X_{m-1} = j) = p \text{ for } m = 1, 2, \ldots; 1 \leq j \leq K-1 \\
p_{0j}^{(m-1,m)} &:= 0, \quad \forall j \in \{1, \ldots, K\}, \forall m \geq 1 \\
p_{00}^{(m-1,m)} &:= 1, \quad p_{KK}^{(m-1,m)} = 1, \forall m \geq 1 \\
p_{Kj}^{(m-1,m)} &:= 0, \quad \forall j \in \{0, \ldots, K-1\}, \forall m \geq 1 \\
p_{jj}^{(m-1,m)} &:= 0, \quad \forall j \in \{1, \ldots, K-1\}, \forall m \geq 1 \\
p_{j,j-1}^{(m-1,m)} &:= P(X_m = j-1 | X_{m-1} = j) = 1 - p \text{ for } 1 \leq j \leq K-1, \forall m \geq 1 \\
p_{i,j}^{(m-1,m)} &:= P(X_m = j | X_{m-1} = i) = 0 \text{ for } |i - j| \geq 2.
\end{aligned}$$

Conclusions:

a) The sequence $\{X_k\}_{k=0,1,2,\ldots}$ of gains is <u>not</u> independent.

b) The one - step transition probabilities $p_{ij}^{(m-1,m)}$ do not depend on m, we write

$$p_{ij}^{(m-1,m)} = p_{ij}^{(1)}, \forall m \geq 1, \forall i, j \in E.$$

c) The matrix of one - step transition probabilities

$$P^{(1)} = (p_{ij}^{(1)})_{i,j=0,\ldots,K} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ K-3 \\ K-2 \\ K-1 \\ K \end{array} \begin{array}{c} \begin{array}{ccccccccc} 0 & 1 & 2 & 3 & \cdots & K-2 & K-1 & K \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1-p & 0 & p & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1-p & 0 & p & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & 0 \\ 0 & \cdots & \cdots & 1-p & 0 & p & 0 & 0 \\ 0 & \cdots & \cdots & \cdots & 1-p & 0 & p & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 1-p & 0 & p \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 & 1 \end{bmatrix} \end{array}$$

is a stochastic matrix, i.e.

$$\begin{aligned} p_{ij}^{(1)} &\geq 0, \quad \forall i,j \in E \\ \sum_{j \in E} p_{ij}^{(1)} &= 1, \quad \forall i \in E \end{aligned}$$

(nonnegative entries, all rows sum to one).

2nd problem: Are there higher order dependencies (higher then of first order)?
Illustration:

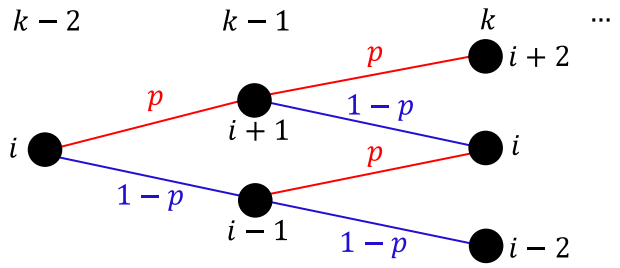


Figure 2.1: For two intermediate steps.

$$P(X_k = j | X_{k-1} = i_{k-1}, X_{k-2} = i_{k-2}, \ldots, X_0 = i_0) = P(X_k = j | X_{k-1} = i_{k-1})$$

There are only 1st order dependencies - we then say that the chain $\varkappa = \{X_k\}_{k=0,1,2,\ldots}$ has a 1st order Markov property.
Note, however, that, in general:

$$P(X_k = j | X_{k-1} = i) \neq P(X_k = j), \forall i,j \in E, \forall k \geq 0$$

(Equality would imply stochastic independence of the r.v.s. X_k! Observe that the games are independent, but not the gains (capital growth) X_k).

2.2 Basic definitions

Definition 2.1 *A stochastic chain $\varkappa = \{X_k\}_{k=0,1,2,\ldots}$ defined on a probability space $[\Omega, \mathscr{L}, P]$ is called a* **discrete Markov chain** *(DMC) with state space E if and only if it holds:*

(i) E is countable and for all $k \in \mathbb{N}$: $P(X_k \in E) = 1$,

(ii) for all $k \in \mathbb{N}$ and $i_0, i_1, \ldots, i_k, i_{k+1} \in E$:

$$P(X_{k+1} = i_{k+1} | X_k = i_k, X_{k-1} = i_{k-1}, \ldots, X_0 = i_0) = P(X_{k+1} = i_{k+1} | X_k = i_k).$$

Note: More precisely, (ii) describes the 1st order Markov property.

Denotations: Let $\varkappa = \{X_k\}_{k=0,1,2,\ldots}$ be a DMC.

a) For all $k \in \mathbb{N}$ we denote

$$\underline{p}(k) = (P(X_k = j))_{j \in E}$$

the state vector of (marginal) probabilities at time k.

b) For all $k, l \in \mathbb{N}, l > k$, we denote

$$P(k,l) := (p_{i,j}^{(k,l)})_{i,j \in E} = (P(X_l = j | X_k = i))_{i,j \in E}$$

the transition matrix of $\varkappa$ from (time) k to l.

In particular, for $k = 0$ we call

$$\underline{p}(0) = (P(X_0 = j))_{j \in E} = (p_j(0))_{j \in E}$$

the initial state distribution of $\varkappa$. Further, $P(k, k+1)$ is called the one - step transition matrix at the $(k+1)$st transition.

Theorem 2.1 *Let $\varkappa$ be a DMC with state space E. Then it holds for all $k \in \mathbb{N}$ and $i_0, i_1, \ldots, i_k \in E$:*

$$P(X_0 = i_0, X_1 = i_1, \ldots, X_k = i_k) = p_{i_{k-1}, i_k}^{(k-1,k)} \cdot \ldots \cdot p_{i_0, i_1}^{(0,1)} \cdot p_{i_0}^{(0)}.$$

Remark upon the proof: this follows immediately from the multiplication theorem:

$$P(A_0 \cap A_1 \cap \cdots \cap A_k) = P(A_k | A_{k-1} \cap \cdots \cap A_0) \cdot P(A_{k-1} | A_{k-2} \cap \cdots \cap A_0) \cdot \ldots \cdot P(A_1 | A_0) \cdot P(A_0)$$

for events $A_j (= \{X_j = i_j\}); j = 0, \ldots, k$; and observing the (first order) Markov property:

$$P(A_0 \cap A_1 \cap \cdots \cap A_k) = P(A_k | A_{k-1}) \cdot P(A_{k-1} | A_{k-2}) \cdot \ldots \cdot P(A_1 | A_0) \cdot P(A_0).$$

Corollary 2.1 *All finite - dimensional distributions $P_{(X_0, X_1, \ldots, X_k)}$ and thus the stochastic behaviour of the whole DMC $\varkappa$ are completely determined by the initial state distribution $\underline{p}(0)$ and the corresponding one - step transition matrices*

$$P(0,1), P(1,2), \ldots, P(k-1, k).$$

This characterization is often utilized in practical applications: First one finds $\underline{p}(0)$ and determines the one - step transition matrices and then, assuming the Markov property, one is able to model the stochastic system under consideration as a Markov Chain. The validity of the Markov property must be checked, however, carefully.
Very often, transition probabilities $p_{i,j}^{(k,l)}$ do not depend on the particular instances $k, l \in T$, but only on the number of steps $n = l - k$ to be taken. In this case, we call

$$p_{i,j}^{(n)} := p_{i,j}^{(k,k+n)} = P(X_{k+n} = j | X_k = i)$$

the n-step transition probability from i to $j; n \geq 1$.

Definition 2.2 *A (first order) DMC is said to be homogeneous, if it holds:*

$$p_{i,j}^{(k,k+n)} =: p_{i,j}^{(n)}, \forall i, j \in E, \forall k, n \in \mathbb{N}.$$

Thus, for a homogeneous DMC (HDMC for short), the starting time k is irrelevant, only the number of transition steps matters.
Denotations: $P^{(n)} = (p_{ij}^{(n)})$, n - step transition matrix; in particular: $P^{(1)} =: P$ - one step transition matrix.
Lemma 2.1 *a)* $P^{(0)} = (p_{ij}^{(0)}) = (\delta_{ij})_{i,j \in E} = I_s$ *where* $s = |E|$.

b) $p_{ij}^{(n+1)} = \sum_{k \in E} p_{ik}^{(n)} p_{kj}^{(1)}, \forall n \geq 1, \forall i, j \in E$ *(special case of Chapman - Kolmogorov equation, see Theorem 2.2).*

Proof:

a) Is trivial.

b) We make use of the following identity for events B, C and complete systems $\{A_k\}$ of events:

$$\begin{aligned} P(B|C) = \sum_k P(A_k \cap B|C) &= \sum_k \frac{P(A_k \cap B \cap C)}{P(C)} = \sum_k \frac{P(A_k \cap B \cap C)}{P(C)} \cdot \frac{P(A_k \cap C)}{P(A_k \cap C)} \\ &= \sum_k \frac{P(A_k \cap B \cap C)}{P(A_k \cap C)} \cdot \frac{P(A_k \cap C)}{P(C)} \\ &= \sum_k P(B|A_k \cap C) \cdot P(A_k|C) \qquad (*) \end{aligned}$$

Now, consider $p_{ij}^{(n+1)} = P(\underbrace{X_{l+n+1} = j}_{B} | \underbrace{X_l = i}_{C})$ with an arbitrary $l \geq 0$.
Obviously, $\{A_k\} = \{X_{l+n} = k\}_{k \in E}$ defines a complete system of events (at time $l + n$). Applying the

above identity (*) we get:

$$\begin{aligned} p_{ij}^{(n+1)} &= P(X_{l+n+1} \overset{B}{=} j \,|\, X_l \overset{C}{=} i) \\ &= \sum_{k\in E} P(X_{l+n+1} \overset{B}{=} j \,|\, X_{l+n} \overset{A_k}{=} k, X_l \overset{C}{=} i) \cdot P(X_{l+n} \overset{A_k}{=} k \,|\, X_l \overset{C}{=} i) \\ \text{[Markov property]} &= \sum_{k\in E} P(X_{l+n+1} = j | X_{l+n} = k) \cdot P(X_{l+n} = k | X_l = i) \\ \text{[Homogeneity]} &= \sum_{k\in E} p_{ik}^{(n)} p_{kj}^{(1)}. \end{aligned}$$

■

In matrix form Lemma 2.1 b) reads:

$$P^{(n+1)} = P^{(n)} \cdot P = P^{(n-1)} \cdot P \cdot P = \dots$$

Corollary 2.2 $P^{(n)} = P^n, \forall n \geq 0.$

Observing that $\underline{p}(k)^T = \underline{p}(k-1)^T P$ and using an induction argument, we may summarize our findings as follows.

Theorem 2.2 (Chapman - Kolmogorov Equation) *Let $\varkappa = \{X_k\}_{k=0,1,2,\dots}$ be a homogeneous DMC with state space E. Then it holds:*
a) $\underline{p}(k+n)^T = \underline{p}(k)^T P^n, \forall k, n \in \mathbb{N}$
b) $P^{(k+n)} = P^{(k)} P^{(n)} = P^{k+n}, \forall k, n \in \mathbb{N}.$

Remark: More generally, for a DMC not necessarily being homogeneous, we have:

a) $\underline{p}(k+n)^T = \underline{p}(k)^T P(k, k+n)$

b) $P(k, k+n) = P(k, k+l) P(k+l, k+n), \forall k, l, n \in \mathbb{N}$ with $0 \leq l < n$.

Note: All powers $P^{(n)} = P^n$ are stochastic matrices, $n = 0, 1, 2, \dots$.

Example 2.1 (Introductory example 2.1 continued) *Let $K = 4$, i.e. $E = \{0, 1, 2, 3, 4\}$*

$$P^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = P \text{ implies}$$

$$P^{(2)} = P^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ q & pq & 0 & p^2 & 0 \\ q^2 & 0 & 2pq & 0 & p^2 \\ 0 & q^2 & 0 & pq & p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

in particular, for $p = 0.5$ *we obtain*

$$P^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.5 & 0.25 & 0 & 0.25 & 0 \\ 0.25 & 0 & 0.5 & 0 & 0.25 \\ 0 & 0.25 & 0 & 0.25 & 0.5 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and with an initial state distribution

$$\begin{aligned} \underline{p}(0) &= (P(X_0 = 0), P(X_0 = 1), P(X_0 = 2), P(X_0 = 3), P(X_0 = 4))^T \\ &= (0, 0.2, 0.5, 0.3, 0)^T, \text{ say,} \\ \underline{p}(1)^T &= \underline{p}(0)^T \cdot P = (0.1, 0.25, 0.25, 0.25, 0.15). \end{aligned}$$

Thus, the initial state distribution has changed after the first transition, $\underline{p}(1) \neq \underline{p}(0)$.
Proceeding further, we obtain e.g.

$$P^{(9)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.734 & 0 & 0.031 & 0 & 0 \\ 0.469 & 0.031 & 0 & 0.031 & 0.25 \\ 0.234 & 0 & 0 .031 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \underline{p}(9) = \begin{pmatrix} 0.451 \\ 0.016 \\ 0.016 \\ 0.016 \\ 0.501 \end{pmatrix}$$

♢

Are there limits $\lim_{n\to\infty} P^{(n)}$ and $\lim_{n\to\infty} \underline{p}(n)$, respectively?

Definition 2.3 *The homogeneous DMC* $\varkappa$ *is said to have a stationary initial distribution if it holds*

$$\underline{p}(k) = \underline{p}(0), \forall k \geq 1.$$

(In this case $\underline{p}(0)$ is also called a equilibrium state).
If a stationary initial distribution $\underline{p}(0)$ exists then it follows from Theorem 2.2 a) that it can be determined as the solution to the linear equation system

$$\begin{aligned} \underline{p}(0)^T P &= \underline{p}(0)^T \\ \underline{p}(0)^T \underline{1}_s &= 1 \text{ (normalization)} \end{aligned}$$

where $\underline{1}_s$ is a vector of s ones, $s = |E|$. We will come back to this issue in Section 3.4, where we will also make clear how $\underline{p}(0)$ relates to $\lim_{n\to\infty} \underline{p}(n)$.

2.3 Typical examples of discrete MCs

2.3.1 One dimensional random walk

Model: A particle is moving on the set of integers according to the following rules

(i) The particle starts at the origin, i.e. $P(X_0 = 0) = 1$.

(ii) If, at time $k \in \mathbb{N}$, the particle is in the state $z \in \mathbb{Z}$ then it jumps during the next time interval $(k, k+1)$ with probability p, $0 < p < 1$, to $z+1$, and with probability $q = 1-p$ to $z-1$.

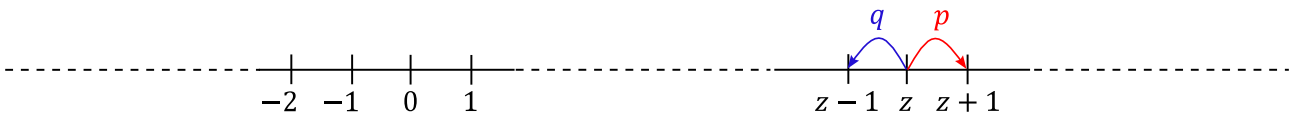


Figure 2.2: State diagram.

Obviously, this one-dimensional random walk constitutes a homogeneous DMC with state space $E = \mathbb{Z}$, initial state distribution $\underline{p}(0) = (\ldots, 0, \underbrace{1}_{X_0=0}, \underbrace{0}_{X_0=1}, \ldots)^T$ and transition matrix

$$P = \begin{array}{c} \\ \vdots \\ i=-1 \\ i=0 \\ i=1 \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccccccc} \ldots & j=-2 & j=-1 & j=0 & j=1 & j=2 & \ldots & \ldots \end{array} \\ \begin{bmatrix} \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & q & 0 & p & 0 & 0 & 0 & \ldots \\ \ldots & 0 & q & 0 & p & 0 & 0 & \ldots \\ \ldots & 0 & 0 & q & 0 & p & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{bmatrix} \end{array}$$

Interesting questions:

(i) What is the probability that the particle ever return to the origin and how long will it take to do so, on the average? **(recurrence time)**

(ii) How long will it take, on the average, until the particle reaches the state $z \in \mathbb{Z}$ for the first time? **(1st passage time)**

2.3.2 First modification: random walk with reflecting barriers

Model: As before, the particle is moving on $E = \mathbb{Z}$ according to a random walk, where, however, there are now reflecting barriers at the states $a \in \mathbb{Z}$, $a < 0$, and/or $b \in \mathbb{Z}$,$b > 0$, respectively, at which the particle

is reflected (like a pinball).

Illustration:

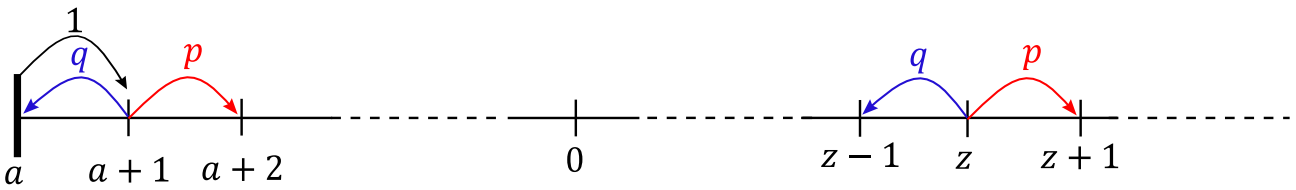


Figure 2.3: Just one reflecting barrier at $a \in \mathbb{Z}$, $a < 0$.

We have a homogeneous DMC with state space $E = \{a, a+1, \ldots, 0, 1, \ldots\}$ and the initial state distribution

$$\underline{p}(0) = (\underbrace{0}_{X_0=a}, \ldots, 0, \underbrace{1}_{X_0=0}, 0, \ldots)^T$$

$$P = \begin{array}{c} \\ i = a \\ i = a+1 \\ i = a+2 \\ i = a+3 \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccccc} j = a & j = a+1 & j = a+2 & j = a+3 & \cdots & \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ q & 0 & p & 0 & 0 & \cdots \\ 0 & q & 0 & p & 0 & \cdots \\ 0 & 0 & q & 0 & p & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \end{array}$$

Questions:

(i) Do the state distribution vectors $\underline{p}(k)$ converge as $k \to \infty$, and if so, is it independent od the initial state $\underline{p}(0)$?

(ii) Is there an equilibrium state $\underline{p} = (p_j)_{j \in E}$ for which $\sum_{j \in E} p_j p_{jk} = p_k, \forall k \in E$?

2.3.3 Second modification: random walk with absorbing barriers

Model: random walk as in Section 2.3.1, but now there are barriers at the states $a \in \mathbb{Z}, a < 0$, and $b \in \mathbb{Z}, b > 0$, where the particle gets absorbed and remains there, forever.

Illustration:

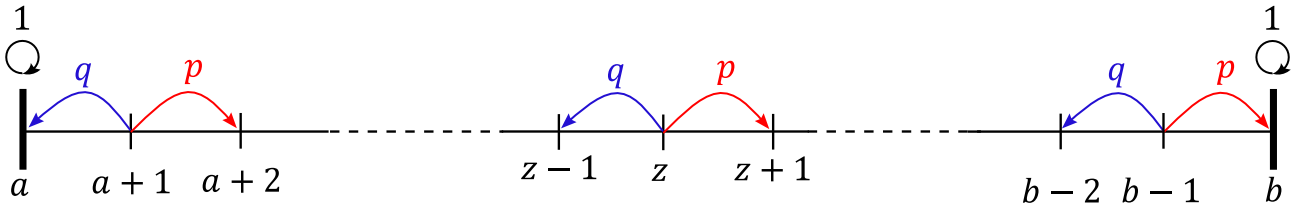


Figure 2.4: State diagram.

This forms a homogeneous DMC with finite state space $E = \{a, a+1, \ldots, 0, 1, \ldots, b-1, b\}$, initial state

$\underline{p}(0) = (\underbrace{0}_{a}, \ldots, 0, \underbrace{1}_{0}, 0, \ldots, \underbrace{0}_{b})$ has finite length.

$$P = \begin{array}{c} \\ i=a \\ i=a+1 \\ \vdots \\ i=b-1 \\ i=b \end{array} \begin{array}{c} \begin{array}{ccccccc} j=a & j=a+1 & j=a+2 & \cdots & j=b-2 & j=b-1 & j=b \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ q & 0 & p & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & q & 0 & p \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Questions:

a) What is the probability of absorption at state $a \in \mathbb{Z} (b \in \mathbb{Z})$?

b) How long will it take, on the average, until the particle gets absorbed at all?

We will give a partial answer in Section 2.4.

Remark: The random walk and its modifications as considered before can be easily generalized to higher dimensions. Consider:

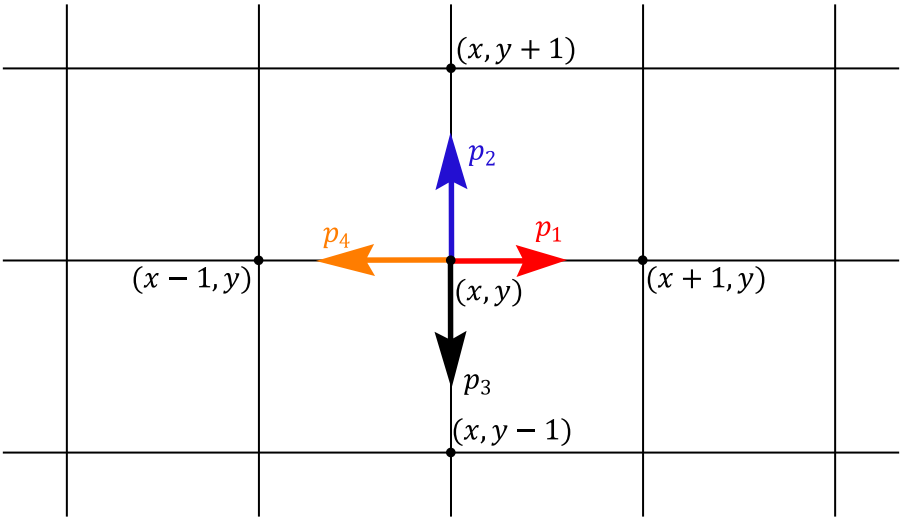


Figure 2.5: *Two - dimensional random walk* on a regular grid.

$p_i \geq 0$ with $p_1 + p_2 + p_3 + p_4 = 1$. Particle starts at $(x, y) = (0, 0)$. This is a 2D - homogeneous DMC.

Remark: Only in the symmetric case $p_1 = p_2 = p_3 = p_4 = \frac{1}{4}$ we can guarantee recurrence to the origin $(0, 0)$, up Section 3.3.

2.3.4 Galton's branching process

Model: Growth of a population in which the individuals of the k-th generation $(k = 0, 1, 2, \ldots)$ will have zero, one, two, ... offspring with probabilities $\alpha_0, \alpha_1, \alpha_2, \ldots$. The total number of offspring of the

individuals of the k-th generation form the $(k+1)$th generation. Originally, (generation $k=0$) there is only one individual:

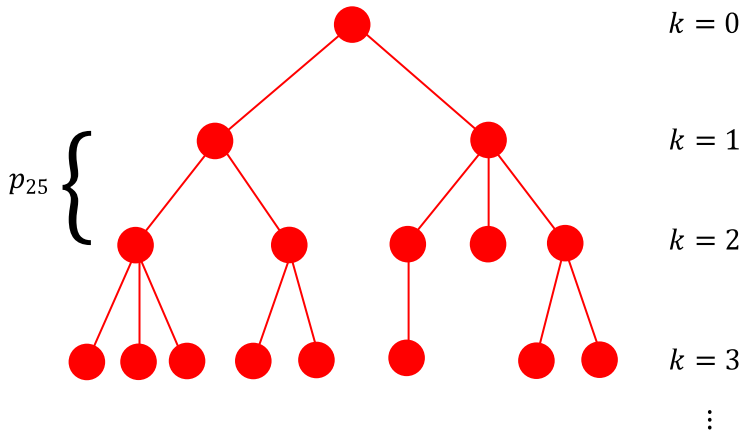


Figure 2.6: Branching process.

$$p_{25} = \alpha_0\alpha_5 + \alpha_1\alpha_4 + \alpha_2\alpha_3 + \alpha_3\alpha_2 + \alpha_4\alpha_1 + \alpha_5\alpha_0.$$

This results in a homogeneous DMC $\{X_k\}_{k=0,1,\ldots}$ with state space $E = \mathbb{N}$, where $X_k = i \in \mathbb{N}$ means that there are exactly i individuals at time k (k-th generation). Initial state distribution:

$$\underline{p}(0) = (P(X_0), P(X_1), \ldots)^T = (0, 1, 0, \ldots)^T$$

Transition matrix $P = (p_{ij})_{i,j=0,1,\ldots}$ has elements

$$p_{ij} = \begin{cases} 1 & \text{for } i = 0, j = 0 \\ 0 & \text{for } i = 0, j \neq 0 \\ \sum\limits_{\substack{k_1,k_2,\ldots,k_i \in E \\ k_1+k_2+\ldots+k_i=j}} \alpha_{k_1} \cdot \alpha_{k_2} \cdot \ldots \cdot \alpha_{k_i} & \text{else} \end{cases}$$

Question: What is the probability of extinction of the population?
We'll come back to this issue in Chapter 3.

2.4 Hitting times, absorption probabilities

Let $\varkappa = \{X_k\}_{k\geq 0}$ be a HDMC with transition matrix $P = P^{(1)}$.

Definition 2.4 *For $A \subset E$, the random variable*

$$H^A : \Omega \to \{0, 1, 2, \ldots\} \cup \{\infty\}$$

given by

$$H^A(\omega) := \inf\{n \geq 0 : X_n(\omega) \in A\}$$

is called the hitting time of the subset A of E. (Agreement: $\inf(\emptyset) = \infty$.)

Consider the probability starting from $i \in E$ that the MC $\varkappa$ ever hits A:

$$h_i^A = P(H^A < \infty | X_0 = i).$$

When A consists of absorbing states, h_i^A is called the absorption probability. The mean time taken for $\varkappa$ to reach A is given by:

$$t_i^A = E(H^A | X_0 = i) = \sum_{k=0}^{\infty} k P(H^A = k | X_0 = i).$$

Less formally:

$$h_i^A = P(\text{hit } A | X_0 = i),$$

$$t_i^A = E(\text{time to hit } A | X_0 = i).$$

Remarkably, these quantities can be calculated explicitly by means of certain linear equations associated with the transition matrix P.

Example 2.2 *Let $E = \{1, 2, 3, 4\}$ and consider the chain $\varkappa$ with state diagram:*

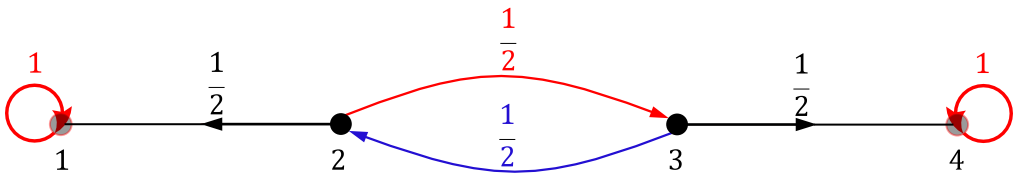


Figure 2.7: State diagram.

Starting from $X_0 = 2$, what is the probability of absorption in 4? How long does it take until the chain is absorbed in 1 or 4?
Define:

$$h_i(4) = P(\textit{hit } 4 | X_0 = i),$$
$$t_i^A = E(\textit{time to hit } A = \{1, 4\} | X_0 = i).$$

Clearly,

$$h_1(4) = P(\textit{hit } 4 | X_0 = 1) = 0,$$
$$h_4(4) = 1,$$
$$t_1^A = t_4^A = 0.$$

Now,

$$h_2(4) = \frac{1}{2} h_1(4) + \frac{1}{2} h_3(4),$$
$$t_2^A = 1 + \frac{1}{2} t_1^A + \frac{1}{2} t_3^A.$$

Similarly,

$$\begin{aligned} h_3 &= \frac{1}{2}h_2 + \frac{1}{2}h_4, \\ t_3^A &= 1 + \frac{1}{2}t_2^A + \frac{1}{2}t_4^A \\ \Rightarrow h_2 = \frac{1}{2}h_3 &= \frac{1}{2}\left(\frac{1}{2}h_2 + \frac{1}{2}h_4\right) \\ t_2^A &= 1 + \frac{1}{2}t_3^A = 1 + \frac{1}{2}\left(1 + \frac{1}{2}t_2^A\right) \\ \Rightarrow h_2 = \frac{1}{3}, t_2^A &= 2 = t_3^A, h_3 = \frac{2}{3}. \end{aligned}$$

♢

The general result for hitting probabilities reads:

Theorem 2.3 *The vector of hitting probabilities $\underline{h}^A = (h_i^A)_{i\in E}$ is the minimal non - negative solution to the system of linear equations:*

$$\begin{aligned} h_i^A &= 1 \text{ for } i \in A \\ h_i^A &= \sum_{j\in E} p_{ij} h_j^A \text{ for } i \notin A. \end{aligned}$$

Remark: Minimality means that if $\underline{x} = (x_i)_{i\in E}$ is another solution with $x_i \geq 0$ for all i, then $x_i \geq h_i$ for all i.

Example 2.2 continued: $h_4 = 1, h_2 = \frac{1}{2}h_1 + \frac{1}{2}h_3, h_3 = \frac{1}{2}h_2 + \frac{1}{2}h_4, h_1$ *was not determined, minimality implies* $h_1 = 0$. ♢

In cases where E is infinite, the minimality condition is essential.

Example 2.3 *Gambler's ruin, no capital limit.*

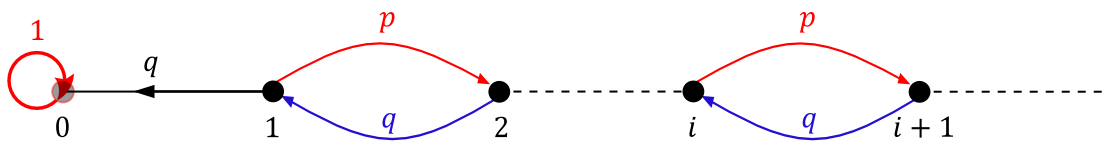


Figure 2.8: State diagram.

$$0 < p = 1 - q < 1$$

$$p_{00} = 1, p_{i,i-1} = q, p_{i,i+1} = p \text{ for } i = 1, 2, \ldots.$$

Resources of the casino are regarded as infinite. What is the probability that you leave broke?

$$\begin{aligned} h_i &= P(\text{ hit } 0 | X_0 = i) \\ h_0 &= 1 \\ h_i &= ph_{i+1} + qh_{i-1}, i = 1, 2, \ldots \end{aligned}$$

If $p \neq q$ this recurrence relation has a general solution $h_i = A + B(\frac{q}{p})^i$ with constants A and B.

a) $p < q$ *(which is the case in most casinos)*
$\Rightarrow B = 0$ *because of the restriction* $0 \leq h_i \leq 1$
$\Rightarrow h_i = 1$ *for all* i.

b) $p > q \Rightarrow$ *observing that* $h_0 = 1$
$h_i = A + (1 - A)(\frac{q}{p})^i$ *nonnegativity of solution implies* $A \geq 0$. $h_i = (\frac{q}{p})^i$ *provides the minimal solution.*

c) $p = q \Rightarrow$ *general solution reads*
$h_i = A + B(i)$
Restriction $0 \leq h_i \leq 1$ *forces* $B = 0$, *so* $h_i = 1$ *for all* i. ♢

Conclusion: For $p \leq q$ you are certain to end up broke (even in the "fair" case $p = q = 0.5$). This apparent paradox is called gambler's ruin.

Remark: More generally, Theorem 2.3 holds for any closed set $A \subset E$ (see Section 3.1, Def 3.3). Absorption classes are special cases of closed sets.

3 State classification of homogeneous DMCs

In this chapter, we assume, again:

$$E \subseteq \mathbb{Z}, \text{ (discrete state space)}$$
$$T \subseteq \mathbb{Z}, \text{ (discrete index set, "points" in time),}$$

and, additionally, $\varkappa = \{X_k\}_{k=0,1,...}$ to be homogeneous.

3.1 Communicating states, essential classes

Definition 3.1 *a) The state $j \in E$ is said to be reachable from the state $i \in E$ (we write $i \rightsquigarrow j$, for short) if $p_{ij}^{(n)} > 0$ for some $n \geq 1$.*
b) The states $i, j \in E$ are said to be communicating (we write $i \leftrightsquigarrow j$, for short) if it holds: $i \rightsquigarrow j$ and $j \rightsquigarrow i$.

Agreement: $i \leftrightsquigarrow i$ if and only if $i \rightsquigarrow i$, i.e. $\sum_{n=1}^{\infty} p_{ii}^{(n)} > 0$.
This way, $\leftrightsquigarrow$ defines an equivalence relation and implies a decomposition of the state space E into disjoint equivalence classes.

Definition 3.2 *a) The state $i \in E$ is said to be essential, if it holds: $\forall j \in E : i \rightsquigarrow j \Rightarrow j \rightsquigarrow i$.*
b) The state $i \in E$ is said to be non - essential if it is not essential.

Denotation:
$M_i = \{j \in E : i \rightsquigarrow j\}$ = set of all states which can be reached from i,
$C_i = \{j \in E : i \leftrightsquigarrow j\}$ = equivalence class of all states which are communicating with i.
Clearly, $C_i \subseteq M_i$ for all $i \in E$.
Corollary 3.1 *Let $i \in E$ be an essential state. Then it holds: $C_i = M_i$.*

Therefore, if $i \in E$ is essential, we call C_i an essential class.
Note: If $i \in E$ is non - essential and $i \rightsquigarrow j$, this does not necessarily imply that j is non - essential as well. However, it holds:
Corollary 3.2 *If $i \in E$ is non - essential, then each $j \in C_i$ is non - essential.*

Proof: Assume, there exists $j \in C_i$ such that j is essential. Then, by Corollary 3.1, all $k \in C_i$ are essential states. This, in turn, implies that i is also essential, contradictory to the assumption of our corollary. ■
Consequently, "non - essential" defines a class property, too. Therefore, if $i \in E$ is non - essential, we call C_i a non - essential class.
Corollary 3.3 *If $i \in E$ is non - essential, then it holds:*

$$C_i \subset M_i,$$

where C_i consists of all non - essential states which communicate with i and M_i consists of all essential

and non - essential states which are reachable from i.

Definition 3.3 *$K \subseteq E$ is called a closed set of states if it holds:*

$$p_{ij} = 0, \forall i \in K, \forall j \notin K,$$

or, equivalently,

$$\sum_{j \in K} p_{ij} = 1, \forall i \in K.$$

Corollary 3.4 *If K is closed, then we have:*

$$p_{ij}^{(n)} = 0, \forall i \in K, \forall j \notin K, \forall n \geq 1.$$

Remark upon the proof: For $n = 2$, it follows from Chapman Kolmogorov's equation

$$p_{ij}^{(2)} = \sum_{m \in E} p_{im} p_{mj} = \sum_{m \in K} p_{im} \underbrace{p_{mj}}_{=0} + \sum_{m \notin K} \underbrace{p_{im}}_{=0} p_{mj}$$

i.e. $p_{ij}^{(2)} = 0$ for all $i \in K, j \notin K$. The proof then follows by complete induction.

Example 3.1 *Find the communicating classes associated to the transition matrix:*

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

The solution is obvious from the state diagram:

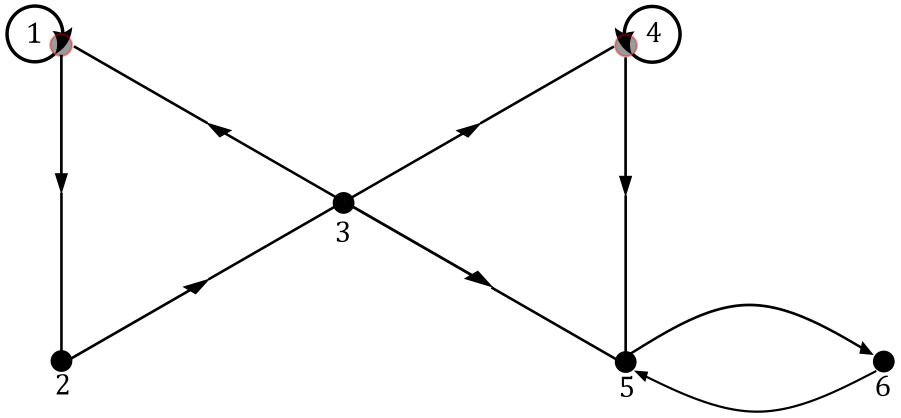


Figure 3.1: State diagram.

$M_1 = \{1, 2, 3, 4, 5, 6\}, M_4 = \{4, 5, 6\}, M_5 = \{5, 6\}$, *but,* $C_1 = \{1, 2, 3\}, C_4 = \{4\}, C_5 = \{5, 6\}$, *i.e. 3 communicating classes with only* $\{5, 6\}$ *being closed (and essential). The states* $1, 2, 3, 4$ *are non - essential.*

Corollary 3.4 says: there is no escape from a closed set!
Special case: absorbing state $i \in E$ (for which $p_{ii}^{(n)} = 1, \forall n \geq 1$)
$\Rightarrow C_i = \{i\}$ are single - element classes being closed and essential.

Theorem 3.1 *(i) Essential classes are closed.*
(ii) Non - essential classes are not closed.

Proof:

(i) For essential classes, $C_i = M_i$ (see Corollary 3.1). M_i is closed, however.

(ii) If $i \in E$ is non essential, then there exists $j \in E$ such that: $i \rightsquigarrow j$ and $j \not\rightsquigarrow i$, i.e. $j \notin C_i$. Now $i \rightsquigarrow j$ implies that $p_{ij}^{(n)} > 0$ for some $n \geq 1$. Thus, by Corollary 3.4, C_i is not closed. ■

Summary:

(1) E can be decomposed into disjoint equivalence classes of communicating states.

(2) Equivalence classes are either essential or non - essential.

(3) Essential classes are closed (and minimal).

(4) Non - essential classes are not closed. From such classes transitions can be made into other non - essential classes or into essential classes, from which then there is no escape.

3.2 Irreducible chains

Definition 3.4 *A homogeneous discrete Markov chain is said to be irreducible, if the state space E consists of a single class of (communicating) essential states.*

Equivalent formulations: The HDMC is irreducible

$\Leftrightarrow \forall i \in E : C_i = M_i = E$, i. e. each state $i \in E$ can be reached from any other state $j \in E$.

$\Leftrightarrow$ For each pair of states (i, j) there exists a finite sequence of intermediate states $i_1, i_2, \ldots, i_n \in E$ such that: $p_{ii_1} > 0, p_{i_1 i_2} > 0, \ldots, p_{i_n j} > 0$, i.e. $i \to i_1 \to i_2 \to \ldots \to i_n \to j$.

Conversely, if the state space E consists of at least two equivalent classes, then the HDMC is called reducible.
We can decide on reducibility/irreducibility of a HDMC solely on the basis of the one - step transition matrix $P^{(1)} = P$. We "simply" have to generate a (generalized) lower block - triangular structure, after an appropriate reordering of rows and columns of P.

Theorem 3.2 (Decomposition of the state space E) *The state space E of a HDMC can be uniquely decomposed into $k \geq 0$ (closed) essential classes $B_1, \ldots, B_k$ and $m \geq 0$ non - essential classes $B_{k+1}, \ldots, B_{k+m}$ with $k + m \geq 1$ such that transitions from states of a non - essential class can be made only to other states within this class or to non - essential classes with smaller ordering number, or to essential classes.*

This decomposition is equivalent to representing the transition matrix in the following form:

$$P = \begin{bmatrix} Q_{11} & & & & & & \\ O & Q_{22} & & & & & \\ \vdots & & \ddots & & & \mathcal{O} & \\ O & \dots & & Q_{kk} & & & \\ Q_{k+1,1} & Q_{k+1,2} & \dots & \dots & Q_{k+1,k+1} & & \\ \vdots & & & & & & \\ \vdots & & & & & \ddots & \\ Q_{k+m,1} & Q_{k+m,2} & \dots & \dots & \dots & & Q_{k+m,k+m} \\ \uparrow & \uparrow & & \uparrow & \uparrow & & \uparrow \\ B_1 & B_2 & \dots & B_k & B_{k+1} & \dots & B_{k+m} \end{bmatrix}$$

where $Q_{ii} \neq O$ are quadratic $(i = 1, \dots, k+m)$ and $Q_{ij}(i > j)$ are rectangular submatrices (which may be null matrices).

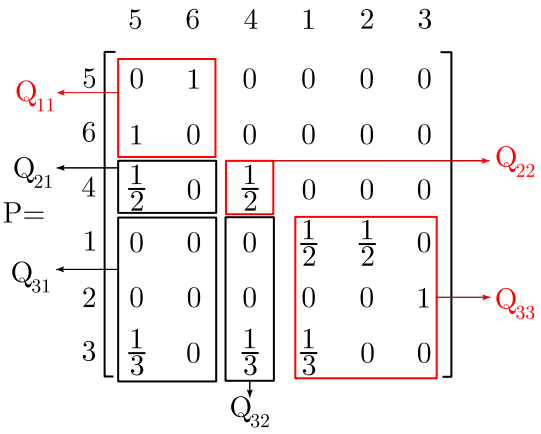
Theorem 3.2 tells us that the classification of E into essential and non - essential classes can be achieved by reordering of P.

Matrices of the above type are called decomposable whenever $k + m \geq 2$.

Corollary 3.5 *The HDMC is irreducible if and only if its transition matrix is not decomposable (i. e. if and only if $k + m = 1$).*

Remark: If E is finite state space then it holds $k \geq 1$, i. e. than there exists at least one essential state/class.

Example 3.2 (Previous example 3.1 from Section 3.1 cont'd) *We may reorder the columns and rows of P as follows:*



$k = 1$ *essential class* $B_1 := \{5, 6\}$ *(closed),*

$m = 2$ *non - essential classes* $B_2 := \{4\}, B_3 := \{1, 2, 3\}$ *(not closed).* ♢

Example 3.3

$$E = \{0, 1, 2, 3, 4, 5, 6, 7\}$$

$$P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0.4 & 0 & 0.2 & 0.1 & 0.1 & 0.2 & 0 & 0 \\ 0.1 & 0.2 & 0.1 & 0.2 & 0.3 & 0.1 & 0 & 0 \\ 0.1 & 0 & 0 & 0 & 0 & 0.2 & 0.1 & 0.6 \\ 0.1 & 0.1 & 0 & 0 & 0.1 & 0.2 & 0.3 & 0.2 \end{bmatrix}$$

P is already in a generalized lower triangular form. We conclude:
$B_1 = \{0, 1\}, B_2 = \{2, 3\}; k = 2$ *essential classes;*
$B_3 = \{4, 5\}, B_4 = \{6, 7\}; m = 2$ *non - essential classes.*
$p_{ij}^{(n)} = 0$ *for* $i \in B_1 \cup B_2, j \in B_3 \cup B_4; n \geq 1$, *but:*
$p_{ij} > 0$ *for* $i \in B_3 \cup B_4, j \in B_1 \cup B_2$ *(transitions from non - essential to essential classes)*
B_1 *and* B_2 *are closed,* B_3 *and* B_4 *are not closed.*

$$\begin{aligned} C_0 &= C_1 = M_0 = M_1 = B_1 \\ C_6 &= \{6, 7\} = B_4 \\ M_6 &= \{0, 5, 6, 7\}, C_6 \subset M_6. \end{aligned}$$

♢

Remarks on irreducible matrices

Matrix $A \in \mathbb{R}^{n \times n}$ is said to be reducible if there exists a permutation matrix Π such that:

$$\Pi^T A \Pi = \begin{bmatrix} A_{1_{(s,s)}} & O_{(s,n-s)} \\ A_{2_{(n-s,s)}} & A_{3_{(n-s,n-s)}} \end{bmatrix}$$

with $1 \leq s < n$. Otherwise, A is irreducible[4].
Equivalently: A is irreducible if $B = A + A^2 + \cdots + A^n$ has non - zero entries, i.e. $b_{ij} \neq 0$ elements of $B = (b_{ij})$ for all $i, j = 1, \ldots, n$.

Theorem 3.3 *Let* $\varkappa = \{X_k\}_{k=0,1,2,\ldots}$ *be a HDMC with finite state space* E, $|E| = n$. *The chain is irreducible if*

$$B = \sum_{k=1}^{n} P^k > O$$

i.e. all $b_{ij} > 0$.

[4] R.Brualdi, H.Ryser: Combinatorial Matrix Theory, Cambridge University Press, New York 1991.

3.3 Periodicity of HDMC's

We are investigate whether there is a periodic behaviour for transitions $i \rightsquigarrow i$. Clearly, if $p_{ii}^{(n)} > 0$ for some $n \geq 1$, then it holds: $p_{ii}^{(k \cdot n)} > 0$ for all $k \geq 1$. This follows from Chapman - Kolmogorov's equation $p_{ii}^{(k \cdot n)} \geq (p_{ii}^{(n)})^k > 0$.

Definition 3.5 *Let $i \in E$ be such that $i \leftrightsquigarrow i$, i.e. there is some $n \geq 1$ such that $p_{ii}^{(n)} > 0$. The greatest common divisor (gcd) of the set $N_i = \{n \in N : p_{ii}^{(n)} > 0\}$ is called the period d_i of the state $i \in E : d_i = gcd(N_i)$.*

Agreement: The period is not defined for states $i \in E$ with $i \not\rightsquigarrow i$, i. e. for states i with $N_i = \emptyset$.
The state $i \in E$ with $N_i \neq \emptyset$ is said to be periodic if $d_i > 1$; it is said to be aperiodic if $d_i = 1$.

Example 3.4 *The simple random walk as defined in Section 2.3.1 forms an irreducible homogeneous discrete Markov chain with $d_i = 2$ for all $i \in E = \mathbb{Z}$.* ♢

Obviously, if $p_{ii}^{(n)} > 0$ for some $n \geq 1$, then it must hold $n = k \cdot d_i$ for some $k \geq 1$, i. e. n must be a multiple of the period. Moreover, it is easily seen that periodicity is a class property.

Corollary 3.6 *Let $i \in E$ be a state of a HDMC with period $d_i \geq 1$. Then all the states in the equivalence class C_i have the same period: $d_j = d_i, \forall j \in C_i$.*

For $d_i \geq 2$ we say: C_i has the period d_i.
For $d_i = 1$ we say: C_i is aperiodic.
In particular, we define:

a) An irreducible HDMC is said to be periodic (or cyclic, respectively) if E consists of a single essential class with period $d > 1$.

b) For $d = 1$ we call this chain aperiodic.

Example 3.5 (Example 3.2 from Section 3.2 cont'd) *Computing the $n-$step transition matrices $(n = 2, 3, 4)$ we obtain:*

$$\begin{aligned}
p_{55}^{(2)} &= p_{55}^{(4)} = 1 = p_{66}^{(2)} = p_{66}^{(4)}, p_{55}^{(3)} = p_{66}^{(3)} = 0 \\
p_{44}^{(2)} &= \frac{1}{4}, p_{44}^{(3)} = \frac{1}{8}, p_{44}^{(4)} = \frac{1}{16} \\
p_{11}^{(2)} &= \frac{1}{4}, p_{11}^{(3)} = 0.2917, p_{11}^{(4)} = 0.2292 \\
p_{22}^{(2)} &= 0, p_{22}^{(3)} = \frac{1}{6}, p_{22}^{(4)} = \frac{1}{12} \\
p_{33}^{(2)} &= 0, p_{33}^{(3)} = \frac{1}{6}, p_{33}^{(4)} = \frac{1}{12}.
\end{aligned}$$

Thus, $B_1 = \{5, 6\}$ has period $d = 2$, whereas $B_2 = \{4\}$ and $B_3 = \{1, 2, 3\}$ are both aperiodic, in accordance with Corollary 3.5. ♢

3.4 Recurrence and transience of states

Assumption: At state i we start with probability 1, i.e. $P(X_0 = i) = 1$.

Definition 3.6 (First passage probabilities)

$$f_{ij}^{(n)} := P(X_n = j, X_k \neq j \text{ for } k = 1, \ldots, n-1 | X_0 = i)$$

probability that starting from state i, we are reaching after n steps the state j, for the first time.

Denotation:

η_{ij} = random number of steps for reaching state j for the first time after starting at state i **(first passage time)**.

η_{ii} = random number of steps for the first return into state i **(recurrence time)**.

Corollary 3.7

$$f_{ij}^{(n)} = P(\eta_{ij} = n).$$

Relationship between $f_{ij}^{(n)}$ and $p_{ij}^{(n)}$:

$$\begin{aligned} f_{ij}^{(1)} &= p_{ij}^{(1)} = p_{ij} \\ f_{ij}^{(0)} &= \begin{cases} 1 & \text{for } i = j \\ 0 & \text{else} \end{cases} \\ &= \delta_{ij} = p_{ij}^{(0)} \end{aligned}$$

Theorem 3.4

$$p_{ij}^{(n)} = \sum_{k=0}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)}, \forall n \geq 1.$$

Denotation:

$f_{ij}^{*} := \sum_{n=1}^{\infty} f_{ij}^{(n)}$ = probability that starting from i, we will ever reach j.

f_{ii}^{*} := probability that starting from i, we ever return to i.

Definition 3.7 *The state $i \in E$ of a HDMC is said to be:*

a) recurrent $\Leftrightarrow f_{ii}^{*} = 1$, i. e. we return to i with probability 1,

b) transient $\Leftrightarrow f_{ii}^{*} < 1$.

Essential probabilistic tools for discussing recurrence/transience:

a) PGF (probability generating function)
Let X be discrete random variable taking values $0, 1, 2, \ldots$ with probabilities

$p_k = P(X = k), k = 0, 1, \ldots$

$$G_x(s) := E(s^x) = \sum_{k=0}^{\infty} s^k \cdot p_k, s \in [0, 1]$$

.

b) Abel's limit theorem
Let $\sum_{n=0}^{\infty} a_n$ be a convergent series of real number a_n. Then we define $f(x) := \sum_{n=0}^{\infty} a_n x^n, x \in [0, 1]$ that converges and $f(\cdot)$ is continuous everywhere on $[0, 1]$ with $f(1) = \sum_{n=0}^{\infty} a_n < \infty$.

Theorem 3.5 *a) The state i is recurrent if and only if $\sum_{n=0}^{\infty} p_{ii}^{(n)} = +\infty$.*

b) The state i is transient if and only if $\sum_{n=0}^{\infty} p_{ii}^{(n)} < +\infty$.

Remark upon the proof. Recall (Theorem 3.4): $p_{ij}^{(n)} = \sum_{k=0}^{n} f_{ij}^{(k)} p_{ij}^{(n-k)}, \forall n \geq 1$. Define the following PGF's:

$$\begin{aligned}
F_{ij}(s) &:= \sum_{n=0}^{\infty} f_{ij}^{(n)} \cdot s^n \\
G_{ij}(s) &:= \sum_{n=0}^{\infty} p_{ij}^{(n)} \cdot s^n \\
\overset{Th.3.4}{\Rightarrow} G_{ij}(s) &:= \delta_{ij} + F_{ij}(s) G_{jj}(s) \\
f_{ii}^* &:= \sum_{n=0}^{\infty} f_{ii}^{(n)} = F_{ii}(1); \\
G_{ii}(1) &:= 1 + F_{ii}(1) G_{ii}(1) \Leftrightarrow F_{ii}(1) = 1 - \frac{1}{G_{ii}(1)} \\
&= 1 - \frac{1}{\sum_{n=0}^{\infty} p_{ii}^{(n)}} = f_{ii}^*. \\
f_{ii}^* &= 1 \Leftrightarrow \sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty \\
f_{ii}^* &< 1 \Leftrightarrow \sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty.
\end{aligned}$$

Corollary 3.8 *Transience and recurrence are class properties, i. e. i recurrent (transient) $\Rightarrow j$ is recurrent (transient) for all $j \in C_i$.*

Relationship between essential/non - essential classes:

a) i non - essential state $\Rightarrow i$ is transient

b) i recurrent $\Rightarrow i$ is essential state.

If E is a finite state space, then at least one state is recurrent.
Obviously: if j is transient $\Rightarrow \lim_{n\to\infty} p_{ij}^{(n)} = 0, \forall i \in E$.
Subdivision of recurrent states:

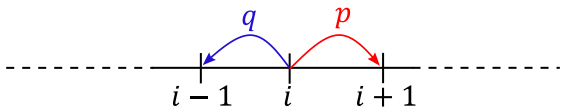
$$\begin{aligned} i \text{ recurrent} \quad &\Rightarrow \quad f_{ii}^{*} = \sum_{n=1}^{\infty} f_{ii}^{(n)} = 1 \\ \eta_{ii} \quad &\sim \quad \left\{ \begin{array}{cccccc} 1 & 2 & \cdots & k & \cdots \\ f_{ii}^{(1)} & f_{ii}^{(2)} & \cdots & f_{ii}^{(k)} & \cdots \end{array} \right\} \\ \mu_i \quad &= \quad \sum_{k=1}^{\infty} k \cdot f_{ii}^{(k)} = E(\eta_{ii}) = \text{ mean recurrence time.} \end{aligned}$$

Definition 3.8 *A recurrent state i is said to be:*
a) zero - recurrent (weakly - ergodic) if $\mu_i = +\infty$
b) positive recurrent (strongly - ergodic) if $\mu_i < \infty$.

Again: zero - recurrence and positive recurrence are also class properties.

Theorem 3.6 (Limit Theorem for DMC's) *Asume that $\varkappa = \{X_k\}_{k=0,1,\ldots}$ is recurrent, irreducible and aperiodic with $P(X_0 = i) = 1$ (i.e. initial state $= i$). Then it holds:*
a) $\lim_{n\to\infty} p_{ii}^{(n)} = \frac{1}{\mu_i}$
b) $\lim_{n\to\infty} p_{ji}^{(n)} = \lim_{n\to\infty} p_{ii}^{(n)}, \forall j \in E.$

Example 3.6 (Recurrence/transience for simple random walk/s) *Consider*



Then

$$\begin{aligned} p_{ii}^{(2k)} \quad &> \quad 0, \forall k \geq 1 \\ p_{ii}^{(2k+1)} \quad &= \quad 0, \forall k \geq 1 \end{aligned}$$

Simple random walk is irreducible with period $d = 2$.

$$\begin{aligned} p_{00}^{(2n)} \quad &= \quad \binom{2n}{n} p^n q^n = \frac{(2n)!}{n! \cdot n!} p^n q^n = (\textit{Stirling's formula: } n! \sim n^{n+\frac{1}{2}} \cdot e^{-n}\sqrt{2\pi}) = \\ &= \quad \frac{(2n)^{2n+\frac{1}{2}} \cdot e^{-2n} \cdot \sqrt{2\pi}}{n^{n+\frac{1}{2}} \cdot e^{-n} \cdot \sqrt{2\pi} \cdot n^{n+\frac{1}{2}} \cdot e^{-n} \cdot \sqrt{2\pi}} p^n q^n = \frac{(4pq)^n}{\sqrt{\pi n}} \leq \frac{1}{\sqrt{\pi n}} \end{aligned}$$

Hence,

$$\sum_{n=1}^{\infty} p_{00}^{(2n)} \quad = \quad \sum_{n=1}^{\infty} \frac{(4pq)^n}{\sqrt{\pi n}} = \begin{cases} \sum_{n=1}^{\infty} \frac{1}{\sqrt{\pi n}} = \infty & \textit{if } p = q = \frac{1}{2} \\ < \infty & \textit{if } p \neq q \end{cases}$$

That means: simple random walk is recurrent if and only if $p = q = \frac{1}{2}$ and transient if and only if $p \neq q$.$\diamondsuit$

Two dimensional random walk:

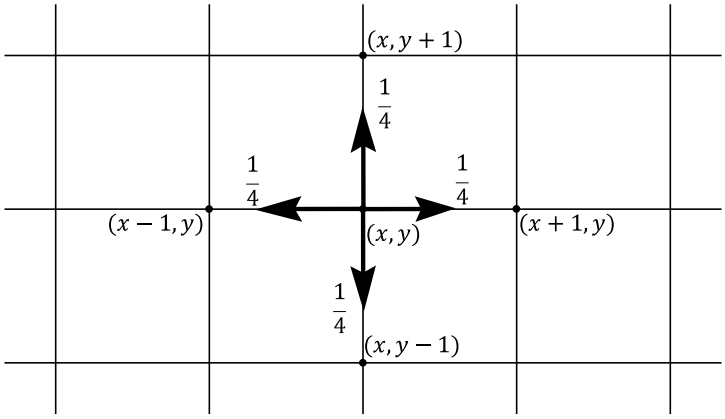
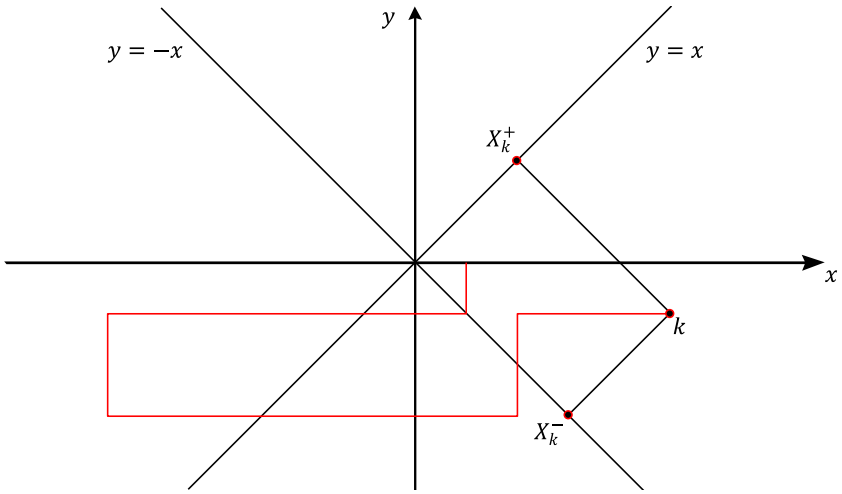


Figure 3.2: 2D Random walk.

$$p_{ij} = \begin{cases} \frac{1}{4} & \text{if } |i-j| = 1 \\ 0 & \text{otherwise} \end{cases}$$

We start at $(x, y) = (0, 0)$.



X_k^+ = orthogonal projection on $y = +x$
X_k^- = orthogonal projection on $y = -x$

$\Rightarrow X_k^+$ and X_k^- are independent simple symmetric random walks on $\frac{1}{\sqrt{2}}\mathbb{Z}$.

$$\begin{aligned} X_k &= (0,0) \Leftrightarrow X_k^+ = 0 = X_k^- \\ \Rightarrow p_{00}^{(2n)} &= \left[\binom{2n}{n}\left[\frac{1}{4}\right]^n\right]^2 \sim \frac{1}{\pi n} \\ \Rightarrow \sum_{n=1}^{\infty} p_{00}^{(n)} &= \infty \end{aligned}$$

$\Rightarrow$ recurrent random walk (transient for the non - symmetric case).
Remark: for the 3D symmetric random walk with:

$$p_{ij} = \begin{cases} \frac{1}{6} & \text{if } |i-j| = 1 \\ 0 & \text{otherwise} \end{cases}$$

we can show

$$\begin{aligned} p_{00}^{(2n)} &= \binom{2n}{n}\left[\frac{1}{4}\right]^n \binom{n}{mmm}\left[\frac{1}{3}\right]^n \\ &\sim \frac{1}{2a^3}\left[\frac{6}{n}\right]^{\frac{3}{2}} \sim \frac{C}{(\pi n)^{\frac{3}{2}}} \end{aligned}$$

where $a = \sqrt{2\pi}, n = 3m$, i. e. transience even in the symmetric case.

3.5 Ergodic states, long - term behaviour of DMC'c

The steady state (ergodic state) is defined by:

$$\underline{p}_\infty := \lim_{n\to\infty} \underline{p}(n) = \lim_{n\to\infty} [P(X_n = i)]_{i\in E}\,.$$

If $\underline{p}_\infty$ exists then we call $\underline{p}_\infty$ the vector of probabilities of ergodic states. It exists trivially, if we have a stationary initial distribution: $\underline{p}(0) = \underline{p}(1) = \underline{p}(2) = \ldots$.
Recall: $\underline{p}(1)^T = \underline{p}(0)^T P$. In case of initial stationarity, this means: $\underline{p}(0)^T = \underline{p}(0)^T P$ (linear equation system): $\underline{1}\underline{p}(0) = 1, \underline{p}(0) \geq 0$. Clearly, if initial distribution $\underline{p}(0)$ exists, then it is also an ergodic state:

$$\lim_{n\to\infty} \underline{p}(n) = \underline{p}(0).$$

Question: Are the limiting probabilities (stable system) independent of the initial distribution $\underline{p}(0)$? Yes!
After $n+1$ steps:

$$\|\underline{p}(n+1) - \underline{p}(n)\| < \varepsilon.$$

That means: if $\underline{p}_\infty$ exists, then it must hold:

$$\begin{cases} \underline{p}_\infty^T = \underline{p}_\infty^T P & \text{we don't have changes after } n+1 \text{ steps} \\ \underline{1}_s^T \underline{p}_\infty = 1, \quad s = |E| & \\ \underline{p}_\infty \geq \underline{0} & \end{cases} \tag{**}$$

Theorem 3.7 *Let be given a HDMC with transition matrix $(P^{(1)})$. Let further denote E_{pr} the set of all positive recurrent states.*

a) If $E_{pr} = \emptyset \Rightarrow$ there exists no ergodic states.

b) If $E_{pr} \neq \emptyset$ and E_{pr} consists of at least 2 equivalent classes, then there exist infinitely many solutions of (**).

c) If $E_{pr} \neq \emptyset$ and E_{pr} consists of only one class, then there exists a unique solution:

$$\underline{p}_\infty = (I_s - P^T + \underline{1}_s \underline{1}_s^T)^{-1} \underline{1}_s.$$

Example 3.7 *Consider the following diagram:*

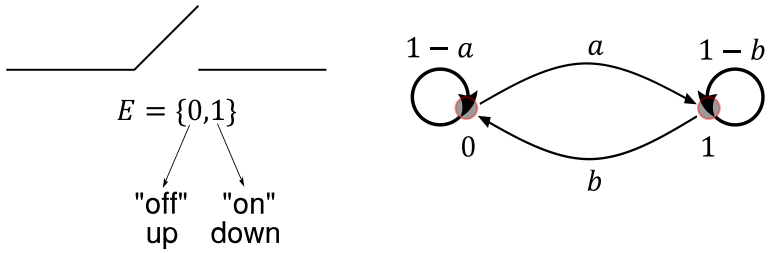


Figure 3.3: Flip-flop circuit.

Probability transition matrix:

$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}$$

We have two cases:

a) $a = b = 1$

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, P^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, P + P^2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

irreducible, $d = 2$ period.

$$P^{2n+1} = P, P^{2n} = I_2 \Rightarrow \lim_{n\to\infty} P^{(n)} \text{ does not exist.}$$

b) $0 < a + b < 2; ab \neq 0 \Rightarrow P$ *indecomposable, i. e. Markov chain is irreducible, one essential class with period* $d = 1$.

$P = U\Lambda U^{-1}$ *mode decomposition (U - matrix of eigen vectors, Λ - matrix of eigen values).*
$P^n = U\Lambda^n U^{-1}, \Lambda =$ *diag* $(\lambda_1, \lambda_2, \ldots, \lambda_s,), s = |E|$.

$$\begin{aligned} |P - \lambda I_2| = 0 &\Leftrightarrow \begin{vmatrix} 1-a-\lambda & a \\ b & 1-b-\lambda \end{vmatrix} = 0 \\ &\Leftrightarrow (1-a-\lambda)(1-b-\lambda) - ab = 0 \\ &\Rightarrow \lambda_1 = 1, \lambda_2 = 1-a-b. \end{aligned}$$

$$\begin{aligned} (P - \lambda_1 I) \cdot \underline{u}_1 &= \begin{bmatrix} -a & a \\ b & -b \end{bmatrix} \underline{u}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ (P - \lambda_2 I_2) \cdot \underline{u}_2 &= \begin{bmatrix} b & a \\ b & a \end{bmatrix} \underline{u}_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow u_2 = \begin{pmatrix} 1 \\ -\frac{b}{a} \end{pmatrix} \\ \Lambda^n &= \begin{bmatrix} 1 & 0 \\ 0 & (1-a-b)^n \end{bmatrix} \\ U &= \begin{bmatrix} 1 & 1 \\ 1 & -\frac{b}{a} \end{bmatrix}, U^{-1} = \frac{1}{1+\frac{b}{a}} \begin{bmatrix} \frac{b}{a} & 1 \\ 1 & -1 \end{bmatrix} \\ \Rightarrow P^n &= U\Lambda^n U^{-1} \\ &= \frac{1}{1+\frac{b}{a}} \begin{bmatrix} 1 & 1 \\ 1 & -\frac{b}{a} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & (1-a-b)^n \end{bmatrix} \begin{bmatrix} \frac{b}{a} & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{a+b} \begin{bmatrix} b & a \\ b & a \end{bmatrix} + \frac{(1-a-b)^n}{a+b} \begin{bmatrix} a & -a \\ -b & b \end{bmatrix}, |1-a-b| < 1 \\ \Rightarrow \lim_{n\to\infty} P^n &= \frac{1}{a+b} \begin{bmatrix} b & a \\ b & a \end{bmatrix} \\ \lim_{n\to\infty} p_{00}^{(n)} &= \frac{b}{a+b}, \\ \lim_{n\to\infty} p_{11}^{(n)} &= \frac{a}{a+b}. \end{aligned}$$

Mean recurrence times:

$$
\begin{aligned}
f_{00}^{(1)} &= P(\eta_{00}=1) = 1-a \\
\mu_{00} &= E(\eta_{00}) \\
f_{00}^{(2)} &= P(\eta_{00}=2) = ab \\
f_{00}^{(n)} &= P(\eta_{00}=n) = ab(1-b)^{n-2}, n \geq 2 \\
\mu_0 = E(\eta_{00}) &= \sum_{n=1}^{\infty} nP(\eta_{00}=n) = \sum_{n=1}^{\infty} nf_{00}^{(n)} = (1-a) + \sum_{n=2}^{\infty} n \cdot ab(1-b)^{n-2} \\
&\overset{[n=k+1]}{=} 1-a+ab\sum_{k=1}^{\infty}(k+1)(1-b)^{k-1} = 1-a+ab\sum_{k=1}^{\infty}(1-b)^{k-1} + ab\sum_{k=1}^{\infty}k(1-b)^{k-1} \\
&= 1-a+ab\frac{1}{1-(1-b)} + a \underbrace{\sum_{k=1}^{\infty} kb(1-b)^{k-1}}_{=\frac{1}{b}, EX \text{ with } X\sim Geo(b)} \\
&= 1-a+a+\frac{a}{b} = 1+\frac{a}{b} = \frac{a+b}{b} \\
P(\eta_{11}=1) &= f_{11}^{(1)} = 1-b \\
P(\eta_{11}=n) &= f_{11}^{(n)} = ab(1-a)^{n-2}, n \geq 2 \\
\Rightarrow \mu_1 = E(\eta_{11}) &= \sum_{n=1}^{\infty} nf_{11}^{(n)} = \cdots = \frac{a+b}{a}
\end{aligned}
$$

$\Rightarrow$ *both state are positive recurrent,*

$$E_{pr} = E = \{0,1\}$$

$\Rightarrow$ *exists only one (strong) ergodic state,* $\underline{p} = \underline{p}(0) = (p_0^*, p_1^*)^T$

$$
\begin{aligned}
(p_0^*, p_1^*) &= (p_0^*, p_1^*) \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \\
p_0^* &= (1-a)p_0^* + bp_1^* \Leftrightarrow ap_0^* = bp_1^* \\
p_1^* &= ap_0^* + (1-b)p_1^* \Leftrightarrow ap_0^* = bp_1^* \\
\Rightarrow p_0^* &= \frac{b}{a+b} = \frac{1}{\mu_0} = \lim_{n\to\infty} p_{00}^{(n)} \\
p_1^* &= \frac{a}{a+b} = \frac{1}{\mu_1} = \lim_{n\to\infty} p_{11}^{(n)}.
\end{aligned}
$$

The ergodic state $(p_0^*, p_1^*)^T$ *is independent from the stationary state distribution:*

$$\underline{p}_{(0)}^T = p(0)^T P = p(0)^T \cdot \lim_{n\to\infty} P^n = \underline{p}_{(0)}^T \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{bmatrix} = \left[\frac{b}{a+b}, \frac{a}{a+b}\right]$$

for arbitrary $\underline{p}(0) = (p_0, p_1)$ *such that* $p_0, p_1 \geq 0, p_0 + p_1 = 1$. ♢

Remark: The preceding example can be generalized for every finite irreducible HDMC. The following holds: $E_{pr} = E$, i.e. in finite irreducible HDMC all states are recurrent.
If the finite HDMC is reducible, then there are transient and positive recurrent states, but no zero-recurrent states.

Example 3.8 *Dynamics of species after introduction of new or alien species.*
At the end of the 19th century gray squirrels were first introduced in Great Britain. They quickly took over areas occupied by the native red squirrel. Data from different regions of GB were collected. Lets denote with:

- R *regions with red squirrels only,*
- G *regions with gray squirrels only,*
- B *regions with both squirrels,*
- A *regions with absence of squirrels.*

For example, let for a 1 Year period, the transition between states be (estimated):
$p_{RR} = 0.8797, p_{RG} = 0.0382, \ldots$. *The transition matrix is given by:*

$$P = \begin{array}{c} \begin{array}{cccc} R\hat{=}1 & G\hat{=}2 & B\hat{=}3 & A\hat{=}4 \end{array} \\ \begin{bmatrix} 0.8797 & 0.0382 & 0.0527 & 0.0008 \\ 0.0212 & 0.8002 & 0.0041 & 0.0143 \\ 0.0981 & 0.0273 & 0.8802 & 0.0527 \\ 0.0010 & 0.1343 & 0.0630 & 0.9322 \end{bmatrix} \end{array}$$

$P > \underline{0}$, *i.e.* P *is regular, hence* $\varkappa$ *is irreducible and aperiodic. The eigenvector for the eigenvalue* $\lambda(P) = 1$ *is:*

$$\underline{p}_\infty = \begin{array}{c} (\quad 0.1705, \quad 0.0560, \quad 0.3421, \quad 0.4314)^T \\ \quad Red \qquad Gray \qquad Both \qquad Neither \end{array}$$

The mean recurrent times $\mu_i = \frac{1}{p_i^*}, i = 1, 2, 3, 4$ *are given by:*

$$\mu = \begin{array}{c} (\quad 5.865, \quad 17.857, \quad 2.923, \quad 2.318)^T \\ \quad R \qquad G \qquad B \qquad A \end{array}$$

Interpretation: The area with red squirrels may change to other states (G,B,A), but on average, will be populated with red squirrels again after about six years. ♢

Remark: For $\varkappa$ with finite E and $P > 0$, i.e. $p_{ij} > 0, \forall i, j \in E$, exists a ergodic state (steady state) $\underline{p}_\infty$ which is unique (for $P > 0$ is P indecomposable). This holds, in general, for so called regular Markov chains. We call $\varkappa = \{X_k\}_{k=0,1,\ldots}$ regular if and only if $\exists n \geq 1 : p_{ij}^{(n)} > 0, \forall i, j \in E$. Clearly, if a HDMC has a absorbing state, then it can't be regular. Long - term behaviour of systems can be completely described

by the following theorem:

Theorem 3.8 *Let $\varkappa = \{X_k\}_{k=1,2,\dots}$ be an irreducible and recurrent DMC with state space $E \subseteq \mathbb{Z}$, and further, let $p_i^*(i \in \mathbb{Z})$ be (the) ergodic state probabilities, i.e. the components of $\underline{p}_\infty$. Moreover, let $g(\cdot)$ be a bounded function on $\mathbb{Z}$. Then it holds:*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} g(X_k) = \sum_{i\in\mathbb{Z}} p_i^* g(i).$$

Example: Production system with 3 states:

$$\begin{aligned} 1 &\hat{=} \text{ full performance,} \\ 2 &\hat{=} \text{ reduced productivity,} \\ 3 &\hat{=} \text{ system failure (repairs).} \end{aligned}$$

State changes might occur at each hour (discretization step).
X_k = State of the system at time $k = 0, 1, 2, \dots \Rightarrow \varkappa = \{X_k\}_{k=0,1,2,\dots}$ is a HDMC with probability transition matrix:

$$\begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0 & 0.6 & 0.4 \\ 0.8 & 0 & 0.2 \end{bmatrix}.$$

Gain/loss function g:

$$\begin{aligned} g(1) &= 1200€/h \\ g(2) &= 750€/h \\ g(3) &= -150€/h \end{aligned}$$

Do we have a unique solution?

$$P^2 = \begin{bmatrix} 0.72 & 0.14 & 0.14 \\ 0.32 & 0.36 & 0.32 \\ 0.80 & 0.08 & 0.12 \end{bmatrix} > 0$$

Clearly, $P + P^2 + P^3 > 0 \to \{X_k\}_{k=0,1,2,\dots}$ is irreducible, so that we have a unique solution p^*.
$\underline{p}_\infty = (p_1^*, p_2^*, p_3^*)$ is uniquely determined.

$$(I_3 - P^T + \mathscr{J})\underline{p}_\infty = \underline{1}_3.$$

Solution in R:

```
> P=matrix(c(0.8,0,0.8,0.1,0.6,0,0.1,0.4,0.2),3,3)
> P%*%P # all entries are positive, i.e. the chain is irreducible
[,1] [,2] [,3]
[1,] 0.72 0.14 0.14
[2,] 0.32 0.36 0.32
[3,] 0.80 0.08 0.12
> I3=diag(c(1,1,1))
> J=matrix(rep(1,9),3,3)
> b=c(1,1,1)
> pstar=solve(I3-t(P)+J,b)
> pstar # unique ergodic probabilities
[1] 0.6666667 0.1666667 0.1666667
> g=c(1200,750,-150)
> Gain=sum(pstar*g)
> Gain
[1] 900
```

Thus, the gain is 900 €/h.

4 Continuous Markov Chains (CMC)

4.1 Definition

Recall: a stochastic process $\varkappa = \{X_t : t \in T\}$ is a family of random variables defined on a joint measurable space $[E, \mathscr{M}]$, where T is the index set and E the state space.
In this chapter the state space will be $E \subseteq \mathbb{Z}$ (as before), i.e. Markov Chain but now let $T \subseteq \mathbb{R}^1$, (usually $T \subseteq [0, \infty)$).

Definition 4.1 *The stochastic process $\varkappa = \{X_t : t \in T\}$ is called a continuous Markov Chain (CMC) if it holds:*

$$P(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}, \ldots, X_{t_1} = i_1, X_{t_0} = i_0) = P(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}),$$
$$\forall n \geq 1, \quad \forall i_0, i_1, \ldots, i_n \in E; \quad \forall 0 \leq t_0 < t_1 < \cdots < t_{n-1} < t_n \in T.$$

Denotation: $P(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}) = p_{i_{n-1} i_n}(t_{n-1}, t_n)$ is called the transition (probability) function. Recall the analogy with $p_{ij}(n-1, n)$ in the discrete case: one-step transition probability. However,the notion of one-step and multi-step transition makes no sense in the continuous case where $t_{n-1}, t_n \in T = [0, \infty)$. Instead, we are now interested in analytical properties of $p_{i_{n-1} i_n}(t_{n-1}, t_n)$, e.g. in limiting and differential properties (differential equations). In analogy with the discrete case, homogeneity is an essential structural property.

Definition 4.2 *The CMC $\varkappa = \{X_t : t \in T\}$ is said to be homogeneous if it holds:*

$$p_{ij}(t_{n-1}, t_n) = p_{ij}(t_n - t_{n-1}), \forall i, j \in E, \forall 0 \leq t_{n-1} < t_n \in T.$$

This means, for a homogeneous CMC (HCMC for short), the starting and end points are irrelevant, only the time difference $t = t_n - t_{n-1}$ matters.
Denotation: in the homogeneous case: $p_{ij}(s, s+t) = p_{ij}(t), \forall i, j, \in E, \forall 0 \leq s, t.$
Again, we are collecting all this functions in a matrix $P(t) = (p_{ij}(t))_{i,j \in E}$.
Clearly, for each fixed $t \in [0, \infty)$, $P(t)$ is a stochastic matrix. In particular, we have:

$$p_{ij}(0) = \lim_{t \downarrow 0} p_{ij}(t) = \delta_{ij}, \forall i, j, \in E \text{ i.e. } P(0) = (p_{i,j}(0))_{i,j \in E} = I_s, s = |E|.$$

4.2 The Poisson process

This is the most important type of a CMC, with many applications, e.g. in modeling:

- arrivals of emails and information packets,
- customer arrival (and queueing) processes in shops, banks and in tourism (incoming, outgoing),
- faults in materials (cables, textiles, yarns, iron and steel casts, tissues ...),

- epidemiological/toxicological processes (spread of infection diseases/viruses threatening humans, animals,plants).

Modeling then proceeds on the basis of counting processes.

Definition 4.3 *$\mathcal{N} = \{N(t) : t \geq 0\}$ is called a counting process if it holds:*

a)	$N(t) \geq 0, \forall t \geq 0$	*non-negativity*
b)	$N(t) : \Omega \to \mathbb{N}, \forall t \geq 0$	*integer valued*
c)	$N(t_2) \geq N(t_1), \forall 0 \leq t_1 < t_2 \in T$	*monotone non-decreasing.*

The most prominent example of a counting process is the Poisson Process defined as follows:

Definition 4.4 *The counting process $\{N(t) : t \geq 0\}$ is called a Poisson process with intensity $\lambda > 0$ if it holds:*

a) $N(0) = 0$ *starting condition,*

b) $N(t+s) - N(s) \sim Po(\lambda t), \forall s, t > 0$,

c) For all $n \geq 1$, *and* $0 \leq t_0 < t_1 < \cdots < t_n < \infty$

$$N(t_0) - N(0), N(t_1) - N(t_0), N(t_2) - N(t_1), \ldots, N(t_n) - N(t_{n-1})$$

are stochastically independent random variables.

(Property c) is referred to as the property of independent increments).

Briefly, the Poisson process is defined as a counting process with independent Poisson distributed increments. It is the basis for modeling "completely" random events such as the sequence of cosmic particles hitting a given region on the Earth.

We will now give two equivalent characterizations of the Poisson process.

4.2.1 Modeling by means of exponentially distributed interarrival times

Let $\{X_n\}_{n=1,2,\ldots}$ be a sequence of r.v.s with $X_i \underset{i.i.d}{\sim} \text{Ex}(\lambda), i = 1, 2, \ldots$ and define:

$$S_n = \sum_{i=1}^{n} X_i, S_0 = 0, n = 1, 2, \ldots$$

(Think of S_n as the arrival time of the n-th cosmic particle).

Consider the counting process $\mathcal{N} = \{N(t) : t \geq 0\}$ defined by $N(t) = \max\{n \in \mathbb{N} : S_n \leq t\}$. (In our example, $N(t)$ counts the numbers of particles which have hitted the region in the interval $(0, t]$).

Theorem 4.1 *The counting process $\mathcal{N} = \{N(t) : t \geq 0\}$ defined above forms a Poisson process with intensity λ.*

Proof: By definition of $\mathcal{N}$ we have

$$P(N(t) = n) = P(S_n \leq t, S_{n+1} > t) = P(t - X_{n+1} < S_n \leq t) = P(S_n \leq t) - P(S_{n+1} \leq t)$$

observing that $S_{n+1} = S_n + X_{n+1}$. Since the sum of i.i.d. exponentially distributed r.v.s. follows a Gamma distribution, $S_n \sim Ga(n, \lambda)$, we further obtain

$$P(N(t) = n) = \int_0^t \left[\underbrace{\frac{\lambda^n}{(n-1)!} x^{n-1}}_{u'(x)} \underbrace{e^{-\lambda x}}_{v(x)} - \underbrace{\frac{\lambda^{n+1}}{n!} x^n}_{u(x)} \underbrace{e^{-\lambda x}}_{v'(x)} \right] dx = u(x)v(x) \Big|_0^t = \frac{\lambda^n}{n!} t^n e^{-\lambda t} = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

Therefore, $N(t) \sim Po(\lambda t)$.
It remains to show that $\bar{N}(t) = N(t+s) - N(s)$ is independent of $N(u), \forall u \in [0, s]$ and $\bar{N}(t) \sim Po(\lambda t)$. This, however, follows from the well-known property of "memoryless" - property of the exponential distribution. ■
Conversely, it is easy to show that for a Poisson process $\mathscr{N} = \{N(t) : t \geq 0\}$ with intensity λ, the times between the jumps of $\mathscr{N}$ are i.i.d. distributed as Ex(λ).
Therefore, we have the following characterization:
$\mathscr{N} = \{N(t) : t \geq 0\}$ is a Poisson process PP(λ) if and only if the interarrival times X_i between the jumps of $\mathscr{N}$ are i.i.d. Ex(λ).

4.2.2 Alternative modeling of Poisson processes

We are now studying an infinitesimal characterization (local behaviour) of the Poisson process:

$$N(t+h) - N(t) \text{ as } h \to 0.$$

Let $\mathscr{N} = \{N(t) : t \geq 0\}$ be a counting process with properties:

(1) Independence: The number of events (arrivals) in $(t, t+h)$ is independent of that in $(0, t]$.

(2) Constant intensity: for any interval $(t, t+h)$ of length h

$$P(N(t+h) - N(t) = 1) = \lambda h + o(h)$$

for some $\lambda > 0$.

(3) Rareness: Multiple events within short time intervals are improbable, i.e.

$$P(N(t+h) - N(t) \geq 2) = o(h).$$

Theorem 4.2 *The counting process $\mathscr{N} = \{N(t) : t \geq 0\}$ defined by (1)-(3) forms a Poisson process with parameter λ.*

Proof: Acording to (1), $\mathscr{N}$ has independent increments. Therefore we just need to show that $N(t) \sim Po(\lambda t)$. Let $p_n(t) := P(N(t) = n)$. Then, by (2) and (3),

$$p_n(t+h) = (1 - \lambda h - o(h))p_n(t) + (\lambda h + o(h))p_{n-1}(t) + o(h)$$

for all $n \geq 1$, and $p_0(t+h) = (1 - \lambda h - o(h))p_0(t)$. Then it follows

$$\frac{p_n(t+h) - p_n(t)}{h} = -\frac{\lambda h + o(h)}{h} p_n(t) + \frac{\lambda h + o(h)}{h} p_{n-1}(t) + \frac{o(h)}{h}$$

and

$$\frac{p_0(t+h) - p_0(t)}{h} = -\frac{\lambda h + o(h)}{h} p_0(t).$$

Now, letting $h \to 0$, we obtain

$$\begin{aligned} p_n'(t) &= -\lambda p_n(t) + \lambda p_{n-1}(t) \quad n \geq 1 \\ p_0'(t) &= -\lambda p_0(t). \end{aligned}$$

This, in turn, implies,

$$\begin{aligned} p_0(t) &= e^{-\lambda t} \\ p_1'(t) &= -\lambda p_1(t) + \lambda e^{-\lambda t} \leftrightarrow p_1(t) = (\lambda t) e^{-\lambda t}. \end{aligned}$$

By induction over n we finally obtain:

$$p_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \quad \forall n \geq 0.$$

■

4.2.3 Modifications of the Poisson process

4.2.3.1 Compound Poisson process (CPP(λ))

Let $\mathscr{N} = \{N(t) : t \geq 0\}$ be a PP(λ) and $\{X_i\}_{i=1,2,\ldots}$ a family of i.i.d. random variables $X_i \sim X$. Then $Y(t) = \sum_{i=1}^{N(t)} X_i = X_1 + \ldots + X_{N(t)}$ is called a compound PP(λ) (short CPP(λ)).
$Y(t)$ is also called the Poisson sum of the random variable X_i. It has the mean

$$\begin{aligned} EY(t) &= E_{N(t)}\{E[Y(t)|N(t)]\} \\ &= E_{N(t)}\{E[X_1 + \ldots + X_{N(t)}|N(t)]\} \\ &= E_{N(t)}\{N(t) \cdot E(X)\} \\ &= [EN(t)] \cdot E(X) = (\lambda t) EX \end{aligned}$$

and variance

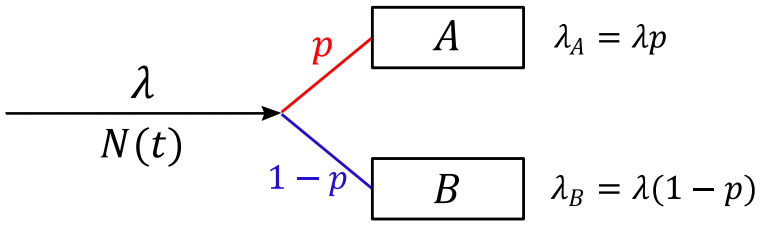
$$\begin{aligned} VarY(t) &= Var_N\{E[Y|N(t)]\} + E_{N(t)}\{Var[Y|N]\} \\ &= Var_N\{N(t)\cdot EX\} + E_{N(t)}\{N(t)\cdot Var(X)\} \\ &= (EX)^2\cdot\lambda t + \lambda t\cdot Var(X) \\ &= \lambda t\cdot EX^2. \end{aligned}$$

Example: Suppose that, on the average, 25 customers arrive at a grocery shop (per hour). Also, suppose that a costumer spends an amount of money which is uniformly distributed between 8€ and 20€. What is the mean and the standard deviation of the total amount of money spent within 4 hours?
CPP(λ) where $\lambda = 25h^{-1}$ and $X \sim U[8,20]$, $EX = 14$, $Var(X) = \frac{(20-8)^2}{12} = 12$

$$\begin{aligned} \Rightarrow EY(t) &= (\lambda t)(EX) = (25\cdot 4)\cdot 14 = 1400€ \\ VarY(t) &= (\lambda t)(EX^2) = 100\cdot(14^2+12) = 20800€^2 \\ SDY(t) &= 144.22€. \end{aligned}$$

♢

4.2.3.2 Filtered Poisson process FPP(λp)



A Poisson process $\mathcal{N} = \{N(t) : t \geq 0\}$ with intensity λ where all events are "filtered" with the same probability $p > 0$, independently of each other, is called a filtered Poisson process (FPP). This is a PP with intensity λp.

Example: On the average, $38\frac{\text{cars}}{h}$ stop at a highway restaurant. Twenty per cent of the drivers refill their cars at the neighbouring gas station before continuing to travel. Compute the probability that at least 25 cars come to the gas station, within a period of 3 hours.

$$\lambda_R = 38h^{-1}, \lambda_G = 38\cdot 0.2 = 7.6$$

$$P(N_G(3) \geq 25) = \sum_{k=25}^{\infty} \frac{(3\lambda_G)^k}{k!} e^{-3\lambda_G} = 1 - \text{ppois}(24, 3\cdot 7.6) = 0.3496.$$

Additional remarks: In analogy with the PGF (probability generating function) for random variables we can define the PGF of a counting process $\{N(t) : t \geq 0\}$ as follows:

$$G_{N(t)}(s) = E\{s^{N(t)}\} = \sum_{k=0}^{\infty} s^k P(N(t) = k), |s| \leq 1.$$

In particular, the PGF of the Poisson process reads:

$$G_{N(t)}(s) = \sum_{k=0}^{\infty} s^k \frac{(\lambda t)^k}{k!} e^{-\lambda t} = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda s t)^k}{k!} = e^{-\lambda t} e^{\lambda s t} = e^{-\lambda t(1-s)}.$$

The PGF of the compound PP(λ) is then defined by:

$$\begin{aligned} G_{Y(t)}(s) &= \sum_{n=0}^{\infty} (E s^{X_1 + \ldots + X_n}) P(N(t) = n) \\ &= \sum_{n=0}^{\infty} (E s^X)^n \frac{(\lambda t)^n}{n!} e^{-\lambda t} \\ &= e^{-\lambda t} e^{\lambda t G_X(s)} = e^{-\lambda t[1 - G_X(s)]}. \end{aligned}$$

4.3 Further properties of the Poisson process PP(λ)

By definition of the Poisson process, the realizations (paths) are integer-valued and monotonically non-decreasing:

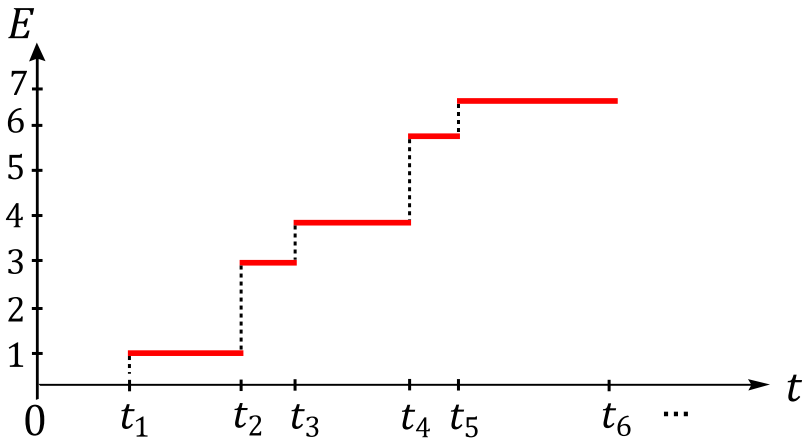


Figure 4.1: Realization (path) X_t of a PP(λ).

Interpretation: $\varkappa = \{X_t : t \geq 0\}$ records the number of events (of the same type) occuring in the interval $(0, t]$, e.g. the number of customers arriving, the number of radioactive decays etc.
From this we immediately infer

P1) $P(X_{t_1} = x_1, X_{t_2} = x_2, \ldots, X_{t_n} = x_n) = \prod_{i=1}^{n} \frac{[\lambda(t_i - t_{i-1})]^{x_i - x_{i-1}}}{(x_i - x_{i-1})!} e^{-\lambda(t_i - t_{i-1})}$
for all $x_0 := 0 \leq x_1 \leq x_2 \leq \cdots \leq x_n, t_0 := 0 < t_1 < t_2 < \cdots < t_n < \infty$ and $n = 1, 2, \ldots$.

Compare this result with the definition of the projective family of the PP(λ) given in Chapter 1!
For $n = 1$ we have the one-dimensional distribution: $P(X_t = x) = \frac{(\lambda t)^x}{x!}e^{-\lambda t}, \forall x \in \mathbb{N}$ i.e. $X_t \sim Po(\lambda t)$ and $E(X_t) = Var(X_t) = \lambda t$.
(For normed time periods, $t = 1$, this yields the well-known characterization $EX = VarX = \lambda$ of $X \sim Po(\lambda)$).

P2) The Poisson process has the Markov property, i.e.

$$P(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}, \ldots, X_{t_1} = x_1) = P(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}).$$

Proof: By definition of the conditional probability the l.h.s. expression reads

$$\frac{P(X_{t_1} = x_1, \ldots, X_{t_n} = x_n)}{P(X_{t_1} = x_1, \ldots, X_{t_{n-1}} = x_{n-1})} = \frac{[\lambda(t_n - t_{n-1})]^{x_n - x_{n-1}}}{(x_n - x_{n-1})!}e^{-\lambda(t_n - t_{n-1})} \text{ (using P1)}$$

The r.h.s. expression evaluates to

$$\frac{P(X_{t_{n-1}} = x_{n-1}, X_{t_n} = x_n)}{P(X_{t_{n-1}} = x_{n-1})}.$$

By virtue of the independent increments property we have

$$P(X_{t_{n-1}} = x_{n-1}, X_{t_n} = x_n) = \frac{[\lambda(t_n - t_{n-1})]^{x_n - x_{n-1}}}{(x_n - x_{n-1})!}e^{-\lambda(t_n - t_{n-1})} \cdot \underbrace{\frac{(\lambda t_{n-1})^{x_{n-1}}}{x_{n-1}!}e^{-\lambda t_{n-1}}}_{=P(X_{t_{n-1}} = x_{n-1})}$$

which leads to the equality of the l.h.s. and r.h.s. expression. ■

Note: The proof shows that the Markov property follows immediately from the (stronger) assumption of the independence of increments $(X_{t_n} - X_{t_{n-1}}), (X_{t_{n-1}} - X_{t_{n-2}}), \ldots, (X_{t_2} - X_{t_1}), X_{t_1}$.

P3) The Poisson process is homogeneous, i.e. $p_{ij}(s, s+t) = P(X_{t+s} = j | X_s = i) = p_{ij}(t), \forall s, t \geq 0$.
Proof:

$$\begin{aligned} p_{ij}(s, s+t) &= P(X_{s+t} = j | X_s = i) = \frac{P(X_s = i, X_{s+t} = j)}{P(X_s = i)} = \\ &= \frac{\frac{(\lambda s)^i}{i!}e^{-\lambda s}\frac{(\lambda t)^{j-i}}{(j-i)!}e^{-\lambda t}}{\frac{(\lambda s)^i}{i!}e^{-\lambda s}} = \begin{cases} \frac{(\lambda t)^{j-i}}{(j-i)!}e^{-\lambda t} & \text{for } j \geq i \\ 0 & \text{else.} \end{cases} \end{aligned}$$

We observe that $p_{ij}(s, s+t)$ does not depend on the starting point $s \geq 0$, but solely on the time difference $(s+t) - s = t$. ■
Again, this property is immediately clear from the independent increments property of the PP(λ): the number of events occuring in $(s, s+t]$ is independent from the number of events occurring in (0,s].
In particular, for $i = 0$:$p_{0j}(t) = \frac{(\lambda t)^j}{j!}e^{-\lambda t} = P(X_t = j)$. $\forall j \in E : p_{0j}(t) = P(X_t = j)$ are the one dimensional distributions.

P4) Let ξ denote the (random) pause time (sojourn time, interarrival time) associated with a PP(λ). It is the time between two jumps of the process. Then it holds $\xi \sim Ex(\lambda)$.
Proof: $F_\xi(t) = P(\xi < t) = 1 - P(\xi \geq t) = 1 - P(X_t = 0)$. Thus, $F_\xi(t) = 1 - \frac{(\lambda t)^0}{0!}e^{-\lambda t}$, which is the c.d.f. of $Ex(\lambda)$. ■
Note: $\underbrace{E\xi = \frac{1}{\lambda}}_{\text{mean sojourn time}}$ and $Var(\xi) = \frac{1}{\lambda^2}$.

P5) The Poisson process is a pure birth process:

$$\begin{aligned} p_{i,i+k}(t) &= \frac{(\lambda t)^k}{k!}e^{-\lambda t}, \forall i \in E, \forall k \geq 0 \\ p_{ij}(t) &= 0, \text{for } j < i \end{aligned}$$

(See proof of P3).
Letting $t = \Delta t \to 0$, we obtain

$$\begin{aligned} p_{ii}(\Delta t) &= e^{-\lambda \Delta t} = 1 - \lambda \Delta t + \frac{1}{2}(\lambda \Delta t)^2 - \cdots = 1 - \lambda \Delta t + o(\lambda \Delta t) \\ p_{i,i+1}(\Delta t) &= (\lambda \Delta t)e^{-\lambda \Delta t} = \lambda \Delta t - (\lambda \Delta t)^2 + \frac{1}{2}(\lambda \Delta t)^3 - \cdots = \lambda \Delta t + o(\Delta t) \\ p_{i,i+k}(\Delta t) &= \frac{1}{k!}(\lambda \Delta t)^k e^{-\lambda \Delta t} = o(\Delta t) \text{ for } k > 1. \end{aligned}$$

The matrix of infinitesimal transition functions thus reads:

$$P(\Delta t) = (p_{ij}(\Delta t))_{i,j \in E} = \begin{array}{c} \begin{array}{ccccc} \cdots & j = i & j = i+1 & j > i+1 & \cdots \end{array} \\ \left[\begin{array}{ccccc} & \ddots & \ddots & & \\ & 1 - \lambda \Delta t + o(\Delta t) & \lambda \Delta t + o(\Delta t) & & o(\Delta t) \\ & & \ddots & \ddots & \\ & \mathcal{O} & & & \end{array}\right] \end{array}$$

4.4 Continuity and differentiability properties of the transition functions $p_{ij}(t)$

Let $\varkappa = \{X_t : t \geq 0\}$ be a homogeneous and continuous Markov chain (HCMC).

4.4.1 Chapman-Kolmogorov equation

Let denote $\underline{p}(t) = (P(X_t = j))_{j \in E}$ the vector of the one-dimensional state probabilities of $\varkappa$ (marginal distribution of X_t). In particular, $\underline{p}(0) = (P(X_0 = j))_{j \in E} =: (p_j(0))_{j \in E}$ is the initial state distribution.

Theorem 4.3 *For any HCMC it holds:*
a) $P(X_t = j) = \sum_{i \in E} p_i(0) p_{ij}(t), \forall t \geq 0.$
b) $p_{ij}(s+t) = \sum_{k \in E} p_{ik}(s) p_{kj}(t), \forall s, t \geq 0.$

Proof:

a) $P(X_0 = i, X_t = j) = P(X_t = j | X_0 = i) \cdot P(X_0 = i)$
$\Rightarrow \underbrace{\sum_{i \in E} P(X_0 = i, X_t = j)}_{= P(X_t = j)} = \sum_{i \in E} p_{ij}(t) \cdot p_i(0)$ (marginalization!).

b) This can be proven in analogy with the discrete case (see Section 3.)

In matrix form the theorem reads:

a) $\underline{p}(t)^T = \underline{p}(0)^T P(t)$

b) $P(s+t) = P(s)P(t)$

(Recall the discrete case: $\underline{p}(k)^T = \underline{p}(0)^T P^{(k)}$ and $P^{(n+k)} = P^{(n)} P^{(k)} = P^n P^k$.)
Recall: Similar result hold true for inhomogeneous CMC's, e.g. in generalization of b):

$$P(s,t) = P(s,u)P(u,t) \text{ for all } 0 \leq u < t < \infty.$$

From Theorem 4.3 b) we get the following:
Corollary 4.1 *For any fixed pair $(i,j) \in E \times E$,the transition function $p_{ij}(\cdot)$ is uniformly continuous, i.e.* $p_{ij}(t \pm s) \underset{s \to 0}{\longrightarrow} p_{ij}(t), \forall t > 0.$

4.4.2 Differentiability of $p_{ij}(t)$

Theorem 4.4 (Differentiability at $t = 0$) *a) For every $i \in E$ the limiting value*

$$\lim_{t \downarrow 0} \frac{p_{ii}(0) - p_{ii}(t)}{t} = -p'_{ii}(0) =: q_i$$

exists and it holds: $0 \leq q_i \leq \infty$ ($q_i < \infty$ for finite E).
b) For all pairs $(i,j) \in E \times E, i \neq j, \lim_{t \downarrow 0} \frac{p_{ij}(t) - p_{ij}(0)}{t} = \lim_{t \downarrow 0} \frac{p_{ij}(t)}{t} = p'_{ij}(0) =: q_{ij}$ exists and it holds: $0 \leq q_{ij} < \infty$.

Remarks upon the proof: for brevity, we indicate only the main steps for proving a), the proof of b) proceeds in a similar way.

1) Using the fact that $\lim_{t \downarrow 0} p_{ii}(t) = 1$ and observing that $p_{ii}(t)$ is continuous for all $t \geq 0$ it follows: $p_{ii}(t) > 0, \forall i \in E$ and $t > 0$.

2) Further, with $0 < p_{ii}(t) \leq 1$ for $t \geq 0$ we have $\varphi_i(t) := -\ln p_{ii}(t) \geq 0 \wedge \lim_{t \downarrow 0} \varphi_i(t) = 0.$

3) It can be shown that $\lim\limits_{t\downarrow 0}\frac{\varphi_i(t)}{t} = \sup\limits_{t>0}\frac{\varphi_i(t)}{t} \underset{def.}{=} q_i$ where $0 \le q_i = \sup\limits_{t>0}\frac{\varphi_i(t)}{t} \le \infty$.

4) $\lim\limits_{t\downarrow 0}\frac{p_{ii}(0)-p_{ii}(t)}{t} = \lim\limits_{t\downarrow 0}\frac{1-p_{ii}(t)}{t} = \lim\limits_{t\downarrow 0}\underbrace{\frac{1-e^{\varphi_i(t)}}{\varphi_i(t)}}_{\to 1(\text{for } t\to 0^+)}\cdot\frac{\varphi_i(t)}{t} = \lim\limits_{t\downarrow 0}\frac{\varphi_i(t)}{t} = q_i$ observing that

$\frac{1-\exp(-\varphi_i(t))}{\varphi_i(t)} = \frac{1-(1-\varphi_i(t)+\frac{\varphi_i(t)^2}{2}-\cdots)}{\varphi_i(t)} = 1-\frac{\varphi_i(t)}{2}+\frac{\varphi_i(t)^2}{3!}-\cdots \underset{t\downarrow 0}{\longrightarrow} 1.$

4.4.3 Transition intensities

We are now interpreting the infinitesimal quantities $q_{ij}(j \neq i)$ and q_i:

a) For q_{ij} we have $q_{ij} := p'_{ij}(0) = \lim\limits_{\Delta t\downarrow 0}\frac{p_{ij}(\Delta t)}{\Delta t} < \infty$ and thus $p_{ij}(\Delta t) = q_{ij}\Delta t + o(\Delta t)$, i.e., apart from remainder terms $o(\Delta t)$, the probability of transition from i to j within the infinitesimal period Δt is given by $q_{ij}\Delta t$. Therefore, q_{ij} is called the transition intensity from i to $j \neq i$.

b) For $i \in E$ with $q_i := -p'_{ii}(0) < \infty$ it follows from Theorem 4.4 a) that $\frac{1-p_{ii}(\Delta t)}{\Delta t} \underset{\Delta t\to 0^+}{\longrightarrow} -p'_{ii}(0) = q_i$ and thus $\underbrace{1-p_{ii}(\Delta t)}_{P(X_{t+\Delta t}\in E\setminus\{i\}|X_t=i)} = q_i\Delta t + o(\Delta t)$ as $\Delta t \downarrow 0$. Therefore, q_i is called the transition intensity out of state i, conversely, $p_{ii}(\Delta t) = 1 - q_i\Delta t + o(\Delta t)$ represents the probability of staying in the state $i \in E$ for at least a further period of the length Δt.

As an example, we look at the transition intensities of the Poisson process.
Example: PP(λ) , $E = \{0, 1, 2, \ldots\}$. From E5, Section 4.3, we recall:

$$p_{ij}(\Delta t) = \begin{cases} \lambda\Delta t + o(\Delta t) & \text{, for } j = i+1 \\ 1 - \lambda\Delta t + o(\Delta t) & \text{, for } j = i \\ o(\Delta t) & \text{, for } j \ge i+2 \\ 0 & \text{, for } j < i \end{cases}$$

Therefore, $q_i = \lambda$ for all $i \in E$, $q_{i,i+1} = \lambda$ for all $i \in E$, $q_{ij} = 0$ for all $j < i$ or $j \ge i+2$. ♢

Corollary 4.2 (Relationship between q_{ij} and q_i)

$$\sum_{j\neq i} q_{ij} \le q_i \text{ for all } i \in E$$

with equality holding in case that E is finite.

Proof: $\sum_{j\neq i} p_{ij}(\Delta t) = 1 - p_{ii}(\Delta t)$. Clearly, then, for all $n < \infty$,

$$\begin{aligned} \sum_{j=1, j\neq i}^{n} p_{ij}(\Delta t) &\leq 1 - p_{ii}(\Delta t) \\ &(=) \quad \text{if } |E| = n \\ \Rightarrow \sum_{j=1, j\neq i}^{n} \frac{p_{ij}(\Delta t)}{\Delta t} &\leq \frac{1 - p_{ii}(\Delta t)}{\Delta t} \end{aligned}$$

and letting $\Delta t \to 0$: $\sum_{j=1, j\neq i}^{n} q_{ij} \leq q_i$. ■

Definition 4.5 *A homogeneous CMC is said to be conservative if it holds $\sum_{j\neq i} q_{ij} = q_i < \infty$ for all $i \in E$.*

Corollary 4.3 *Each homogeneous CMC with finite state space E is conservative.*

4.4.4 Infinitesimal matrix

Denotation: $q_{ii} := p'_{ii}(0) = -q_i, i \in E$.

Definition 4.6 *The matrix of transition intensities $Q = (q_{ij})_{i,j\in E} = (p'_{ij}(0))_{i,j\in E}$ is called the infinitesimal matrix or generator, respectively.*

In matrix form: $Q = P'(0) = \frac{d}{dt}P(t)|_{t=0}$.
The generator Q has the following properties:

i) $q_{ij} \geq 0$ for $j \neq i, i \in E$

ii) $q_{ii} \leq 0$ for all $i \in E$

iii) $\sum_{j\neq i} q_{ij} \leq q_i$, i.e. $\sum_{j\in E} q_{ij} \leq 0$ for all $i \in E$ ($\sum_{j\in E} q_{ij} = 0$ for conservative CMC).

Remark: Clearly, Q is not a stochastic matrix, neither is $Q + I$, although for a conservative CMC the rows of $Q + I$ sum up to one, the main diagonal elements $q_{ii} + 1 = 1 - q_i$ are not necessarily non-negative.
Example: Generator (matrix) of PP(λ)

$$Q = \begin{bmatrix} -\lambda & \lambda & & & \\ & -\lambda & \lambda & & \mathcal{O} \\ & & \ddots & \ddots & \\ \mathcal{O} & & & & \end{bmatrix}$$

band matrix (pure birth process) with $\sum_{j\in E} q_{ij} = q_{ii} + q_{i,i+1} = -\lambda + \lambda = 0$. Thus, any PP($\lambda$) is a conservative CMC.

4.4.5 Random sojourn time in a given state

Let $i \in E$ be such that $0 < q_i < \infty$ and $p_i(0) = P(X_0 = i) = 1$. Denote by ξ_i the random sojourn time in state $i \in E$.

Theorem 4.5

$$P(\xi_i > t | X_0 = i) := P(X_s = i, 0 < s \leq t | X_0 = i) = e^{-q_i t}$$

for all $t > 0$, i.e. $\xi_i \sim Ex(q_i)$ for any hCMC.

Proof:

$$P(X_s = i, 0 < s \leq t | X_0 = i) = \lim_{n \to \infty} P(X_0 = i, X_{\frac{t}{n}} = i, X_{\frac{2t}{n}} = i, \cdots, X_{\frac{nt}{n}} = i). \qquad (\star)$$

Now, for any integer $n \in \mathbb{N}$ we have

$$\begin{aligned} P(X_0 = i, X_{\frac{t}{n}} = i, X_{\frac{2t}{n}} = i, \cdots, X_{\frac{nt}{n}} = i) &= p_i(0) \cdot p_{ii}\left(\frac{t}{n}\right) \cdot \ldots \cdot p_{ii}\left(\frac{t}{n}\right) \\ &= \left[p_{ii}(\frac{t}{n})\right]^n = \left[1 - \frac{t}{n} q_i + o\frac{t}{n}\right]^n \underset{n \to \infty}{\longrightarrow} e^{-q_i t}. \end{aligned}$$

■

Remark: ($\star$) actually requires that the CMC $\{X_t : t \geq 0\}$ is separable, i.e. $\{X_s : s \in (0, t]\}$ is already determined (with probability one) by $\{X_s : s \in S\}$ with $S = \{0, \frac{t}{n}, \frac{2t}{n}, \ldots, \frac{nt}{n}\}$, for all $n \in \mathbb{N}$.
Thus, Theorem 4.5 tells us that

$$F_{\xi_i}(t) = 1 - e^{-q_i t}, E(\xi_i) = \frac{1}{q_i}, i \in E.$$

Particularly, for a Poisson process PP(λ) we have:

$$\xi_i \underset{i.i.d}{\sim} Ex(\lambda) \quad \forall i \in E.$$

Note: For a general hCMC, the ξ_i are not necessarily i.i.d. .
On the basis of the mean sojourn time $E(\xi_i) = \frac{1}{q_i}$ we distinguish three basic categories of states as follows.

Definition 4.7 *The state $i \in E$ of the CMC $\{X_t : t > 0\}$ is called stable (instantaneous) if and only if $0 \leq q_i < \infty (q_i = \infty)$. In particular, the state $i \in E$ is said to be absorbing iff $q_i = 0$.*

Interpretation:

(i) If $i \in E$ is an absorbing state $(q_i = 0)$, then it holds:

$$P\left(X_s = i, 0 < s \leq t | X_0 = i\right) = 1, \forall t > 0,$$

and $E(\xi_i) = \frac{1}{q_i} = \infty$, i.e. the process remains in state i forever.

(ii) If $i \in E$ is instantaneous ($q_i = \infty$), then it holds:

$$P\left(X_s = i, 0 < s \leq t | X_0 = i\right) = 0, \forall t > 0,$$

and $E(\xi_i) = \frac{1}{q_i} = 0$, i.e. having reached $i \in E$, the process leaves the state $i \in E$ instantaneously.

(iii) For a stable state $i \in E$ with $0 < q_i < \infty$ it holds:

$$0 < P\left(X_s = i, 0 < s \leq t | X_0 = i\right) < 1.$$

Remark: For a conservative homogeneous CMC all states are stable.

4.4.6 First transition probabilities

Let $\varkappa = \{X_t : t \geq 0\}$ be a conservative hCMC, i.e. any state $i \in E$ is stable. Further, assume $P\left(X_0 = i\right) = 1$ for some $i \in E$. What is the probability of transition from state $i \in E$ to state $j \in E$ after the end of the sojourn time ξ_i?
Define (first transition times):

- $\tau_i := \sup\limits_{X_t = i} t$ (time of first transition of X_t from i to some $k \in E \backslash \{i\}$) and
- $\tau_{ij} := \inf\limits_{X_t = j} t$ (time of first transition from $i \in E$ to $j \in E$).

Definition 4.8 *First transition probabilities:*

$$r_{ij} := P\left(\tau_{ij} = \tau_i | X_0 = i\right); i \neq j \in E.$$

Theorem 4.6 *Let $\varkappa = \{X_t : t \geq 0\}$ be a conservative hCMC. Then it holds:*

$$r_{ij} = \frac{q_{ij}}{q_i}; j \neq i$$

for all non-absorbing states $i \in E$.

Proof: Obviously,

$$\begin{aligned} r_{ij} &= \lim_{\Delta t \to 0^+} P\left(X_{\Delta t} = j | X_0 = i, X_{\Delta t} \neq i\right) \\ &= \lim_{\Delta t \to 0^+} \frac{p_{ij}(\Delta t)}{1 - p_{ii}(\Delta t)} = \lim_{\Delta t \to 0^+} \frac{\frac{p_{ij}(\Delta t)}{\Delta t}}{\frac{1 - p_{ii}(\Delta t)}{\Delta t}} \\ &= \frac{p_{ij}^{'}(0)}{-p_{ii}^{'}(0)} = \frac{q_{ij}}{q_i}. \end{aligned}$$

■

Illustration:

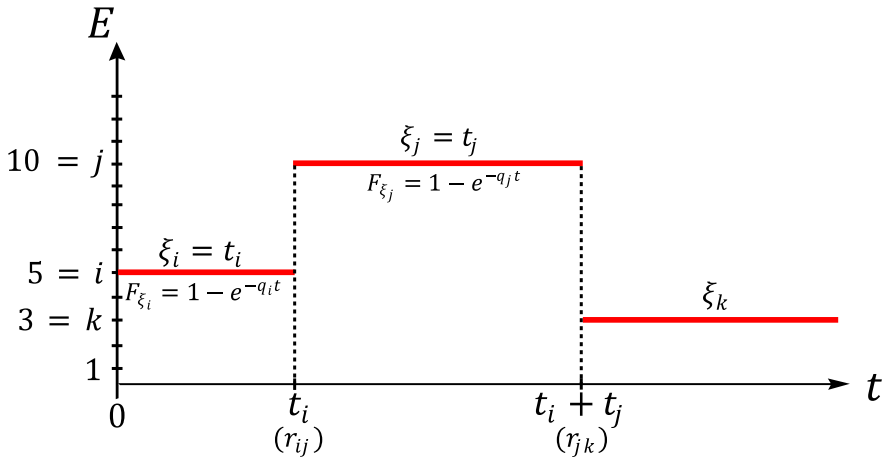


Figure 4.2: Path of a conservative hCMC.

Note: The distribution function of the random sojourn time ξ_i in state $i \in E$ only depends on i (via q_i), it does not depend on the state $j \in E$ of the next transition. In contrast with this situation, for a Semi - Markov process, F_{ξ_i} also depends on the state j of the next transition and, moreover, the cdf is not necessarily an exponential distribution.

Remark: The sojourn time ξ_i is characterized by the "memoryless" - property of the exponential distribution, i.e.

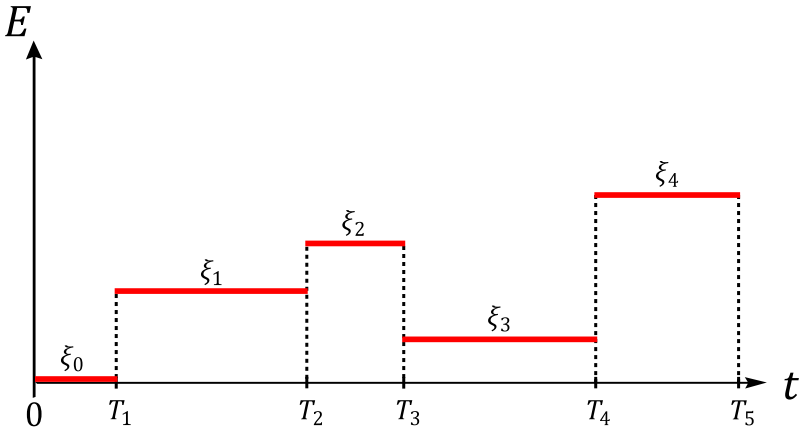
$$P(\xi_i > t + t_0 | \xi_i > t_0) = \frac{P(\xi_i > t + t_0)}{P(\xi_i > t_0)} = \frac{e^{-q_i(t+t_0)}}{e^{-q_i t_0}} = e^{-q_i t}$$

for all $t_0, t > 0$, i.e. ξ_i does not depend on the actual "age" $t_0 > 0$.

4.5 Embedded Markov Chain, State Classification

Let $\varkappa = \{X_t : t \geq 0\}$ be a hCMC.

Denote: $T_0 = 0, T_1, T_2, \ldots$ - (random) times at which the i - th jump of $\varkappa$ occurs.



Definition 4.9 *$y = \{Y_k : k \geq 0\}$ is called the embedded Markov chain (EMC) associated with $\varkappa$.*

For one - step transition probability of EMC y holds:

$$P_y = \left(p_{ij}^{(y)}\right) = \begin{cases} p_{ii}^{(y)} = \begin{cases} 0 & \text{if } q_i > 0 \\ 1 & \text{if } q_i = 0 \end{cases} \\ p_{ij}^{(y)} = \begin{cases} r_{ij} = \frac{q_{ij}}{q_i} & \text{if } i \neq j \wedge q_i > 0 \\ 0 & \text{if } i \neq j \wedge q_i = 0. \end{cases} \end{cases}$$

State classification is then based on P_y (essential, non - essential, transient, recurrent,...)

Definition 4.10 *The state $i \in E$ is said to be positive recurrent (null-recurrent) in $\varkappa$ iff*

$$\mu_i = E\left(\tau_{ii} | X_0 = i\right) < \infty (= \infty).$$

Theorem 4.7 (Limit theorem for hCMC) *If $\varkappa$ is an irreducible hCMC then it holds:*

$$\lim_{t \to \infty} p_{ji}(t) = \frac{1}{q_i \mu_i}, \forall i, j \in E.$$

4.6 Kolmogorov's Differential Equation System

The following differential equation system allows us to determine the transition functions $p_{ij}(t)$ in analytical form.

Theorem 4.8 (Kolmogorov's forward differential equation system, fKDE system) *Assume that $q_i < \infty$ for all $i \in E$. Then it holds:*

$$\begin{aligned} p_{ij}^{'}(t) &= \sum_{k \in E} q_{ik} p_{kj}(t) \quad \forall t \geq 0 \\ p_{ij}(0) &= \delta_{ij}, \forall i, j \in E \text{ (initial conditions).} \end{aligned}$$

In matrix form we have:

$$\begin{aligned} P^{'}(t) &= QP(t); \qquad P\left(t\right) = \left(p_{ij}(t)\right)_{i,j \in E}, Q = \left(q_{ij}\right)_{i,j \in E} \\ P(0) &= I_s; \qquad s = |E|. \end{aligned}$$

Proof: Starting from the Chapman - Kolmogorov equation system

$$P\left(s+t\right) = P\left(s\right) \cdot P\left(t\right), \quad s, t \geq 0$$

and setting $s = \Delta t$, we obtain

$$P\left(\Delta t + t\right) = P\left(\Delta t\right) \cdot P\left(t\right) \qquad / - P\left(t\right)$$

Thus,

$$\begin{aligned} P(t+\Delta t) - P(t) &= [P(\Delta t) - I_s]\, P(t) \\ \Rightarrow \frac{P(t+\Delta t) - P(t)}{\Delta t} &= \left[\frac{P(\Delta t) - I_s}{\Delta t}\right] P(t). \end{aligned}$$

Now, letting $\Delta t \to 0^+$, we arrive at

$$\begin{aligned} \lim_{\Delta t \to 0^+} \frac{P(t+\Delta t) - P(t)}{\Delta t} &= \lim_{\Delta t \to 0^+} \left[\frac{P(\Delta t) - P(0)}{\Delta t}\right] P(t) \\ \Leftrightarrow P'(t) &= QP(t) \end{aligned}$$

observing that $P(0) = (p_{ij}(0)) = I_s$ and $P'(0) = Q$. ■

Remark: The above equations form a system of s^2 homogeneous ordinary linear differential equations of first order with constant coefficients (such a system is called a d'Alambert system). The above differential equations are also referred to as "differential" form of the Chapman - Kolmogorov equations.
The converse of Theorem 4.8 also holds true, i.e. if the transition functions $p_{ij}(\cdot)$ of a homogeneous CMC satisfy the KDE system then the hCMC is conservative. Thus, in particular, for a homogeneous CMC with finite state space $E(s < \infty)$ the KDE system is automatically satisfied.
Interchanging the roles of s and t in the above proof we can analogously arrive at Kolmogorov's backward differential equation system (bKDE system for short):

$$\begin{aligned} P'(t) &= P(t) \cdot Q; \quad P(0) = I_s \\ \Leftrightarrow p'_{ij}(t) &= \sum_{k \in E} p_{ik}(t) q_{kj}; \quad \forall i, j \in E \\ p_{ij}(0) &= \delta_{ij}. \end{aligned}$$

Solution in case of finite $E : |E| = s < \infty$

$$P(t) = e^{Qt} := \underbrace{\sum_{k=0}^{\infty} Q^k \frac{t^k}{k!}}_{\text{(matrix exponential)}} = I_s + Qt + Q^2 \frac{t^2}{2} + \cdots$$

forms a solution of the fKDE system since

$$\frac{d}{dt} P(t) = \sum_{k=1}^{\infty} Q^k \frac{t^{k-1}}{(k-1)!} = Q \sum_{k=0}^{\infty} Q^k \frac{t^k}{k!} = Q e^{Qt} = Q \cdot P(t).$$

Example 4.1 *Let* $s = |E| = 2$ *and*

$$Q = \begin{pmatrix} -a & +a \\ b & -b \end{pmatrix} \text{ for some } a > 0, b > 0.$$

It is easily seen that

$$Q^2 = \begin{pmatrix} a^2+ab & ab-a^2 \\ -b^2-ab & b^2+ab \end{pmatrix} = -(a+b)\,Q$$

and, by induction,

$$Q^n = [-(a+b)]^{n-1}\,Q; \quad n \geq 1.$$

Thus, we obtain

$$\begin{aligned} P(t) = e^{Qt} &= I_2 + \sum_{k=1}^{\infty} Q^k \frac{t^k}{k!} \\ &= I_2 - \frac{Q}{a+b} \sum_{k=1}^{\infty} [-(a+b)]^k \frac{t^k}{k!} \\ &= I_2 - \frac{Q}{a+b}\left(e^{-(a+b)t} - 1\right). \end{aligned}$$

Moreover,

$$\begin{aligned} \lim_{t\to\infty} P(t) &= I_2 + \frac{Q}{a+b} = \frac{1}{a+b}\left[(a+b)I_2 + Q\right] \\ &= \frac{1}{a+b}\begin{pmatrix} b & a \\ b & a \end{pmatrix}. \end{aligned}$$

♢

4.7 Steady - state distribution

We now investigate the long - term behaviour of $\underline{p}(t) = (P(X_t = i))_{i\in E}$ as $t \to \infty$.
Formally, if the limiting distribution exists, then we call

$$\underline{p}_\infty := \lim_{t\to\infty} \underline{p}(t)$$

the steady - state (or: ergodic) distribution of the associated hCMC.
Heuristics: If $\underline{p}_\infty$ exists then, clearly,

$$\underline{p}(t+\Delta t) = \underline{p}(t)$$

for sufficiently large $t > t_0$. Recalling the Chapman - Kolmogorov equations,

$$\underline{p}(t+\Delta t)^T = \underline{p}(t)^T P(\Delta t)$$

we then have for $t > t_0$ and $\Delta t > 0$:

$$\begin{aligned} \underline{p}(t)^T &= \underline{p}(t)^T P(\Delta t) \\ \longleftrightarrow \underline{p}(t)^T \left[\frac{P(\Delta t) - I}{\Delta t}\right] &= \underline{0}^T \\ \underset{\Delta t \to 0}{\longleftrightarrow} \underline{p}(t)^T \cdot P'(0) &= \underline{0}^T \\ \underset{P'(0)=Q}{\longleftrightarrow} Q^T \underline{p}(t) &= \underline{0}. \end{aligned}$$

Definition 4.11 *Let $\varkappa = \{X_t : t \in T\}$ be a hCMC with generator matrix Q. Then $\underline{p}_\infty := (p_0, p_1, \ldots)^T$ is called the steady - state distribution (ergodic distribution) of $\varkappa$ if it holds*

$$Q^T \underline{p}_\infty = \underline{0}, \quad \underline{p}_\infty \geq 0, \quad \underline{1}^T \underline{p}_\infty = 1.$$

Theorem 4.9 *Let $\varkappa = \{X_t : t \in T\}$ be a positive recurrent and irreducible hCMC with transition matrix $P(t) = (p_{ij}(t))_{i,j \in E}$ and generator $Q = (q_{ij})$. Then there exists a unique steady - state distribution $\underline{p}_\infty = (p_i)_{i \in E}$ and it holds*

$$\lim_{t \to \infty} p_{ji}(t) = p_i = \frac{1}{q_i \mu_i} \quad \forall i, j \in E.$$

For a proof see [5].
Comparing this result with the limit result stated in Theorem 4.7, we note that the additional assumption of positive recurrence leads to a simple alternative for computing the mean recurrence times:

$$\mu_i = \frac{1}{p_i q_i}; \quad i \in E.$$

Further, note that in case that $\varkappa$ is finite the additional assumption of positive recurrence is not needed.
Example 4.1 cont'd: For the hCMC $\varkappa$ with $E = \{1, 2\}$ and

$$Q = \begin{pmatrix} -a & +a \\ b & -b \end{pmatrix}; \quad a > 0, b > 0;$$

considered in Section 4.6 we have

$$P^{(y)} = \begin{pmatrix} 0 & \frac{q_{12}}{q_1} \\ \frac{q_{21}}{q_2} & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Therefore, $\varkappa$ is irreducible. Moreover, it is positive recurrent, since E is finite. This implies the following unique steady - state distribution:

$$\underline{p}_\infty = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \lim_{t \to \infty} \begin{pmatrix} P(X_t = 1) \\ P(X_t = 2) \end{pmatrix}:$$

$$\left.\begin{array}{r} Q^T \underline{p}_\infty = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ p_1, p_2 \geq 0, \quad p_1 + p_2 = 1 \end{array}\right\} \quad \begin{array}{rl} \longleftrightarrow & -ap_1 + bp_2 = 0 \\ & p_1 + p_2 = 1 \\ \longleftrightarrow & p_1 = \dfrac{b}{a+b}, p_2 = \dfrac{a}{a+b}. \end{array}$$

Finally, the mean recurrence times read:

$$\mu_1 = \frac{1}{q_1 p_1} = \frac{a+b}{ab} = \frac{1}{q_2 p_2} = \mu_2.$$

To conclude this section, let us look at a simple but real - life application.

Example 4.2 *Consider a reserve system with two identical machines. Ideally, one machine is working while the other one is held in reserve. When a failure occurs, i.e. the machine stops working, it is going to be repaired and the other machine begins to work immediately. Assume that both the lifetime L and the repair time R of the machines follow an exponential distribution:*

$$\begin{aligned} L(t) &= 1 - e^{-\lambda t}; \quad t > 0, \quad \lambda > 0 \\ R(t) &= 1 - e^{-\mu t}; \quad t > 0, \quad \mu > 0. \end{aligned}$$

Define X_t = number of machines which may be used at time $t > 0$, and denote

$$E = \left\{ \begin{array}{ccc} 0, & 1, & 2 \\ \textit{no machine working} & \textit{1 machine working} & \textit{1 machine working} \\ \textit{(both in repair)} & \textit{(other one in repair)} & \textit{(second one in reserve)} \end{array} \right\}.$$

Determine the steady - state probabilities $p_i = P(X_\infty = i); i = 0, 1, 2;$ and the (overall) reliability of the system.

First, we have to determine the transition matrix:

$$P(\Delta t) = \begin{pmatrix} p_{00}(\Delta t) & p_{01}(\Delta t) & p_{02}(\Delta t) \\ p_{10}(\Delta t) & p_{11}(\Delta t) & p_{12}(\Delta t) \\ p_{20}(\Delta t) & p_{21}(\Delta t) & p_{22}(\Delta t) \end{pmatrix}$$

where

$$\begin{aligned} p_{00}(\Delta t) &= e^{-\mu \Delta t} e^{-\mu \Delta t} = e^{-2\mu \Delta t} = 1 - 2\mu \Delta t + o(\Delta t) \\ p_{02}(\Delta t) &= \left(1 - e^{-\mu \Delta t}\right)\left(1 - e^{-\mu \Delta t}\right) = o(\Delta t) \\ p_{10}(\Delta t) &= \left(1 - e^{-\lambda \Delta t}\right) e^{-\mu \Delta t} = \lambda \Delta t + o(\Delta t) \\ p_{12}(\Delta t) &= \left(1 - e^{-\mu \Delta t}\right) e^{-\lambda \Delta t} = \mu \Delta t + o(\Delta t) \\ p_{20}(\Delta t) &= \left(1 - e^{-\lambda \Delta t}\right)\left(1 - e^{-\lambda \Delta t}\right) = o(\Delta t) \\ p_{21}(\Delta t) &= \left(1 - e^{-\lambda \Delta t}\right) e^{-\lambda \Delta t} = \lambda \Delta t + o(\Delta t). \end{aligned}$$

Thus,

$$P(\Delta t) = \begin{bmatrix} 1-2\mu\Delta t + o(\Delta t) & 2\mu\Delta t + o(\Delta t) & o(\Delta t) \\ \lambda\Delta t + o(\Delta t) & 1-(\lambda+\mu)\Delta t + o(\Delta t) & \mu\Delta t + o(\Delta t) \\ o(\Delta t) & \lambda\Delta t + o(\Delta t) & 1-\lambda\Delta t + o(\Delta t) \end{bmatrix}$$

which implies

$$Q = P^{'}(0) = \begin{pmatrix} -2\mu & -2\mu & 0 \\ \lambda & -(\lambda+\mu) & \mu \\ 0 & \lambda & -\lambda \end{pmatrix}.$$

Finally, the steady - state equations $Q^T \underline{p}_\infty = \underline{0}_3$ *with* $\underline{p}_\infty = (p_0, p_1, p_2)^T$ *read:*

$$\begin{aligned} -2\mu p_0 \quad + \quad \lambda p_1 \qquad\qquad &= 0 \\ -2\mu p_0 - (\lambda+\mu)p_1 + \lambda p_2 &= 0 \\ \mu p_1 - \lambda p_2 &= 0, \end{aligned}$$

leading to $p_1 = \frac{\lambda}{\mu} p_2 = \frac{2\mu}{\lambda} p_0$. *Using*

$$1 = p_0 + p_1 + p_2 = p_0 \left(1 + \frac{2\mu}{\lambda} + \frac{2\mu^2}{\lambda^2}\right)$$

we therefore obtain

$$p_0 = \frac{\lambda^2}{(\lambda+\mu)^2+\mu^2}, \quad p_1 = \frac{2\lambda\mu}{(\lambda+\mu)^2+\mu^2}, \quad p_2 = \frac{2\mu^2}{(\lambda+\mu)^2+\mu^2}.$$

The reliability of the system is

$$p = p_1 + p_2 = \frac{2\lambda\mu + 2\mu^2}{\lambda^2 + 2\lambda\mu + 2\mu^2} = \frac{q+q^2}{\frac{1}{2}+q+q^2}$$

where

$$q = \frac{\frac{1}{\lambda}}{\frac{1}{\mu}} = \frac{\textit{mean lifetime}}{\textit{mean repairtime}}.$$

The following table gives numerical values for the reliability for some chosen values of q*:*

q	0.5	1	2	4	10
p	0.6	0.8	0.923	0.976	0.995

5 Basics of continuous stochastic processes

5.1 Basic concepts

Let $\varkappa = \{X_t = X(,) : t \in T\}$, where $T \subseteq \mathbb{R}^1(\mathbb{R}^+)$ is a continuous index set, and the state space E is also continuous, i.e. $X_t = \Omega \to E$, be a stochastic process.
In this chapter, we assume:

$$E \subseteq \mathbb{R}^1, \text{ (continuous state space)}$$
$$T \subseteq [0, \infty), \text{ (continuous index set).}$$

Typical applications are:

- Banks-insurance companies (market and interest rates, premium)
- Medicine (monitoring of brain-, hart-, organ activities)
- Meteorology
- Electrical engineering (measure of current intensity, voltage)

Definition 5.1 (Continuous Markov process (CMP)) *$\varkappa = \{X_t : t \in T\}$ is said to be a Markov process (Markovian) if it holds:*

$$P(X_t \in M | X_{t_n} = x_n, X_{t_{n-1}} = x_{n-1}, \ldots, X_{t_1} = x_1) = P(X_t \in M | X_{t_n} = x_n)$$

$\forall M \in \mathscr{L}_1$(Borel σ algebra in $\mathbb{R}^1$), $\forall n \in \mathbb{N}, \forall t_1 < t_2 < \cdots < t_n < t \in T, \forall x_i \in \mathbb{R}^1$.

Recall: $\varkappa$ is uniquely determined by the set of all n-dimensional distributions

$$F_{t_1,\ldots,t_n}(x_1, x_2, \ldots, x_n) = P(X_{t_1} \leq x_1, X_{t_2} \leq x_2, \ldots, X_{t_n} \leq x_n)$$

provided that the family of distributions is projective (Definition 1.3).
For $n = 1$ we get $F_t(x) = P(X(t) \leq x)$ the univariate (marginal) distribution of X_t.

Definition 5.2 *Let $\varkappa = \{X_t : t \in T\}$ be a continuous stochastic process. Then:*

$$m_X(t) := EX(t) \quad \forall t \in T = [0, \infty)$$

is called a trend function (mean or expectation function), if $EX(t)$ exists.

The trend function describes the average development of a stochastic process in time. If the densities $f_t(x) = \frac{dF_t(x)}{dx}, \forall t \in T$ exist, then it holds $m_X(t) = \int_{-\infty}^{+\infty} x f_t(x) dx$.

Definition 5.3 *Let $\varkappa = \{X_t : t \in T\}$ be a continuous stochastic process. Then:*

$$K_X(s,t) := Cov(X(s), X(t)) \quad \forall s,t \in T$$

is called the covariance function and

$$\rho_X(s,t) := \frac{Cov(X(s), X(t))}{\sqrt{Var(X_s)}\sqrt{Var(X_t)}}$$

is called the correlation function.

Remark: The covariance (correlation) function of a stochastic process is also called autocovariance (autocorrelation) function respectively.
By the definition of the covariance the following holds:

$$K_X(s,t) = E\{[X(s) - m_X(s)][X(t) - m_X(t)]\} = E\{[X(s)X(t)]\} - m_X(s)m_X(t).$$

In particular $K_X(t,t) = E\{[X(t) - m_X(t)]^2\} = Var\{X(t)\}$ is called the variance function. So we can write the correlation function as $\rho_X(s,t) := \frac{K_X(s,t)}{\sqrt{K_X(s,s)K_X(t,t)}}$.
Example: Let $X(t) = A\cos(\omega t + \Phi)$ be a oscillator of cosine waves whit a random amplitude and random phase where

- A : nonnegative random variable with finite expectation and variance,
- Φ : uniformly distributed random variable in $[0, 2\pi]$.

A and Φ are stochastically independent. The trend function is:

$$m_X(t) = E(A)E\{\cos(\omega t + \Phi)\} = E(A)\frac{1}{2\pi}\int_0^{2\pi} \cos(\omega t + \varphi)d\varphi = E(A)\frac{1}{2\pi}[\sin(\omega t + \varphi)]_0^{2\pi} = E(A) \cdot 0 \equiv 0.$$

The covariance function is:

$$\begin{aligned} K(s,t) &= E\{[A\cos(\omega s + \Phi)][A\cos(\omega t + \Phi)]\} = E(A^2)\frac{1}{2\pi}\int_0^{2\pi} \cos(\omega s + \varphi)\cos(\omega t + \varphi)d\varphi \\ &= E(A^2)\frac{1}{2\pi}\int_0^{2\pi} \frac{1}{2}\{\cos[\omega(t-s)]\cos[\omega(s+t) + 2\varphi]\}d\varphi = \frac{1}{2}E(A^2)\cos(\omega(t-s)). \end{aligned}$$

♢

Often, instead of considering $\varkappa = \{X_t : t \in T\}$, we consider the increments $X(t_2) - X(t_1)$ of $\varkappa$ in the interval $[t_1, t_2], t_1, t_2 \in T, t_1 < t_2$.
For practical purposes, we often assume independence of increments in disjoint intervals (for example finance statistics).

Definition 5.4 *$\varkappa = \{X_t : t \in T\}$ is said to be a continuous stochastic process with independent increments, if for all n-tuples $(t_1, t_2, \ldots, t_n)$ where $t_1 < t_2 < \ldots < t_n$, the increments*

$$X(t_2) - X(t_1), X(t_3) - X(t_2), \ldots, X(t_n) - X(t_{n-1})$$

are stochastically independent.

An important subclass of this type are Lévy processes.

5.2 Wiener process (Brownian motion)

In 1828, R. Brown, a Scottish botanist, summarized his observations on erratic movements of pollen particles in liquids. The physical explanation was given by A. Einstein in 1905 (molecular "'bombardment"' due to water molecules). The first mathematical approach was to model the Paris stock exchange by L. Bachelier (1900). Mathematical foundation was given by N. Wiener (1923).

5.2.1 Definition

Definition 5.5 *A stochastic process $\varkappa = \{X_t : t \in T\} = \{W_t\}_{t \geq 0}$ with continuous paths (trajectories) is said to be a Wiener process (Brownian motion), if it holds:*

i) All increments $W_t - W_s$ and $W_v - W_u$ are pairwise stochastically independent where $0 \leq s < t \leq u < v < \infty \in T$.

ii) The increments $W_t - W_s$ are normally distributed with parameter $\sigma^2 > 0$, i.e.

$$W_t - W_s \sim N(0, \sigma^2(t-s)), \quad \forall 0 \leq s < t.$$

iii) $P(W_0 = 0) = 1$ is the initial condition.

Briefly: A Wiener process (Brownian motion) is a stochastic process with independent normally distributed increments.

Remark: For $\sigma^2 = 1$ we call $\{W_t\}_{t \geq 0}$ a standard Wiener process. A non standard wiener process $\{W_t\}_{t \geq 0}$ can be standardized by taking $\tilde{W}_t = \frac{W_t}{\sigma}$.

5.2.2 Properties of the Wiener process

W1) **Covariance function**

Theorem 5.1 *The covariance function of the Wiener process $\{W_t\}_{t \geq 0}$ is*

$$K_W(s,t) = \sigma^2 \min(s,t).$$

Proof: Let $0 \leq s < t$:

$$\begin{aligned} K_W(s,t) &= Cov(W(s), W(t)) \\ &= E\{W(s)W(t)\} - \underbrace{m_W(s)}_{0}\underbrace{m_W(t)}_{0} \\ &= E\{W(s)[W(t) - W(s) + W(s)]\} \\ &= E\{W(s)[W(t) - W(s)]\} + E\{W(s)^2\} \\ \text{[independent increments]} &= \underbrace{E\{W(s)\}}_{0} E\{W(t) - W(s)\} + Var\{W(s)\} \\ &= \sigma^2 s. \end{aligned}$$

Analogous: $0 \leq t < s : Cov(W(s), W(t)) = \sigma^2 t$. ■

Corollary 5.1 *The variance function of the Wiener process is*

$$K_W(t,t) = Var\{W(t)\} = \sigma^2 t$$

and the correlation function is:

$$\rho_W(s,t) = \frac{\sigma^2 \min(s,t)}{\sqrt{\sigma^2 s}\sqrt{\sigma^2 t}} = \sqrt{\frac{s}{t}}$$

for $0 \leq s < t$.

W2) **Homogeneity:** $W_{t+h} - W_{s+h} \sim N(0, \sigma^2(t-s)), \forall h > 0$, i.e. the distribution of increments does not depend on the position of the time interval $[s,t]$, it just depends on the interval length $t - s$.

W3) **Non-differentiability**

Theorem 5.2 *The paths (trajectories) of the Wiener process are, with probability* 1*, nowhere differentiable.*

Proof: Consider the difference quotient $\frac{W_{t+h} - W_t}{h} \sim N(0, \frac{1}{h})$ where, w.l.o.g., we have set $\sigma^2 = 1$. Let $I = (a,b]$ with $-\infty < a < b < +\infty$ be an arbitrary interval.

$$P\left(\frac{W_{t+h} - W_t}{h} \in I\right) = \int_a^b \frac{\sqrt{h}}{\sqrt{2\pi}} e^{-\frac{hx^2}{2}} dx \underset{[\sqrt{h}x = t]}{=} \int_{\sqrt{h}a}^{\sqrt{h}b} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \underset{h \to 0}{\longrightarrow} 0$$

thus, the differential of a trajectory is, with probability 0, contained in a finite interval. ■

W4) **Finite dimensional distribution of the Wiener process:** Using the independence of the increments,

we can easily compute the following n dimensional probability distribution:

$$\underline{W} = \begin{pmatrix} W_{t_1} - W_{t_0} \\ W_{t_2} - W_{t_1} \\ \vdots \\ W_{t_n} - W_{t_{n-1}} \end{pmatrix} \sim N(\underline{0}, \sigma^2 D) \text{ where } D = \begin{pmatrix} t_1 - t_0 & & & \mathcal{O} \\ & t_2 - t_1 & & \\ & & \ddots & \\ \mathcal{O} & & & t_n - t_{n-1} \end{pmatrix}.$$

W5) **Markov property:** The Wiener process is a Markov process, since we have independence of increments.

Theorem 5.3 *For all $n \geq 2$ and n tuples $(t_1, t_2, \ldots, t_n), (x_1, x_2, \ldots, x_n)$ it holds:*

$$P(W(t_n) \leq x_n | W(t_{n-1}) \leq x_{n-1}, \ldots, W(t_1) \leq x_1) = P(W(t_n) \leq x_n | W(t_{n-1}) \leq x_{n-1}).$$

Proof: $P(W(t_n) \leq x_n | W(t_{n-1}) \leq x_{n-1}, \ldots, W(t_1) \leq x_1) = \frac{\prod_{i=1}^{n} P(W(t_i) - W(t_{i-1}) \leq x_i - x_{i-1})}{\prod_{i=1}^{n-1} P(W(t_i) - W(t_{i-1}) \leq x_i - x_{i-1})} =$
$P(W(t_n) - W(t_{n-1}) \leq x_n - x_{n-1}) = P(W(t_n) \leq x_n | W(t_{n-1}) \leq x_{n-1})$. ■

A stochastic process with independent increments is a Markov process, the converse does not have to be true.

5.2.3 Transition probability densities

Now we look at the probability transition function:

$$P(W(t) \leq y | W(s) = x) = P(\underbrace{W(t) - W(s)}_{\sim N(0, \sigma^2(t-s))} \leq y - x) = \Phi\left(\frac{y-x}{\sqrt{\sigma^2(t-s)}}\right) \text{ for } 0 \leq s < t. \qquad (\diamondsuit)$$

Definition 5.6 *The function $p_{x,y}(s,t) = \frac{\partial}{\partial y} P(W(t) \leq y | W(s) = x) = \frac{\partial}{\partial y} \Phi\left(\frac{y-x}{\sqrt{\sigma^2(t-s)}}\right)$ is called the probability (density) transition function of the Wiener process.*

Given ($\diamondsuit$) the following holds: the probability (density) transition function of the Wiener process is:

$$p_{x,y}(s,t) = \frac{1}{\sqrt{2\pi\sigma^2(t-s)}} \exp\left(-\frac{1}{2}\frac{(y-x)^2}{\sigma^2(t-s)}\right) \quad \forall 0 \leq s < t, \forall y, x \in \mathbb{R}^1.$$

This function satisfies the following equations:

$$\frac{\partial p}{\partial t} = \frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial y^2} \qquad \text{(Fokker-Planck equation)}$$

$$\frac{\partial p}{\partial s} = -\frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial x^2} \qquad \text{(Backward equation)}$$

$$p_{x,y}(s,t) = \int_{-\infty}^{+\infty} p_{x,z}(s,u)p_{z,y}(u,t)dz \text{ for arbitrary } s < u < t. \qquad \text{(Chapman-Kolmogorov equation)}$$

5.2.4 First passage time

Consider the event $W(t) > m, m > 0, t$ is fixed. We now look at the threshold exceedances of m. Since $W(0) = 0$ and $W(s)$ is continuous in $0 \leq s \leq t$, $W(s) = m$ for at least one $s \in [0, t]$.

Definition 5.7 *The first passage time of the threshold $m > 0$ is $T_m = \inf\{s > 0 : W(s) = m\}$.*

The reflection path of $W(s)$ is:

$$R(s) = \begin{cases} W(s) & , \text{ for } s < T_m \\ 2m - W(s) & , \text{ for } s \geq T_m. \end{cases}$$

Reflection principle: Both $W(\cdot)$ and $R(\cdot)$ are equally probable, i.e. $P(T_m \leq t, W(s) > m) = P(T_m \leq t, W(s) < m)$.

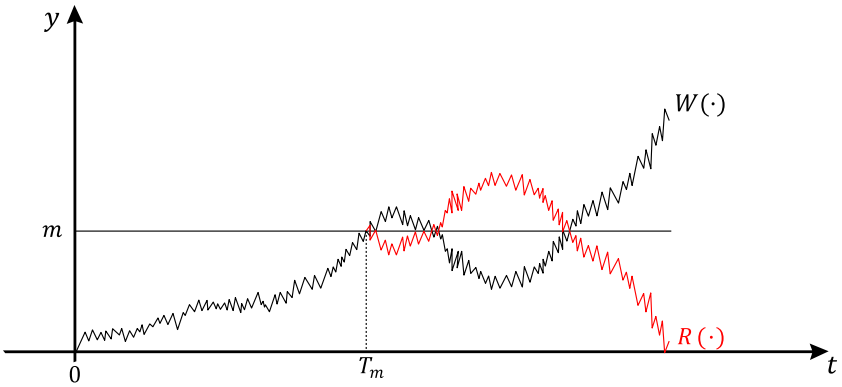


Figure 5.1: First passage time and reflection path.

Theorem 5.4 *For the first passage time it holds: $F_{T_m}(t) = 2\left(1 - \Phi\left(\frac{m}{\sigma\sqrt{t}}\right)\right), \forall t \geq 0$, the inverse Gaussian distribution with parameters m and σ. The density function is:*

$$f_{T_m}(t) = \frac{\partial}{\partial t} F_{T_m}(t) = \frac{1}{\sqrt{2\pi t^3}} \frac{m}{\sigma} \exp\left(-\frac{m^2}{2\sigma^2 t}\right), \forall t \geq 0.$$

5.2.5 Maximum of the Wiener process

Let $M_t = \sup\{W(s) : 0 \leq s \leq t\}$ be the maximum of the Wiener process in $[0, t]$.

Theorem 5.5 *The maximum M_t has the following probability density function:*

$$f_{M_t} = \sqrt{\frac{2}{\pi t\sigma^2}} \exp\left(-\frac{m^2}{2t\sigma^2}\right), 0 \leq m < \infty, M_t \sim N_{trunc}(0, \sigma^2 t).$$

Proof: For arbitrary $m > 0$ it holds $M_t \sim N_{[0,\infty)}(0, \sigma^2 t)$, thus $M_t \geq m \Leftrightarrow T_m \leq t$. We get:

$$P(M_t \geq m) = P(T_m \leq t) = 2P(W(t) > m) = \int_m^\infty \sqrt{\frac{2}{\pi t\sigma^2}} \exp\left(-\frac{x^2}{2t\sigma^2}\right) dx.$$

Differentiation by m yields the density:

$$f_{M_t} = \sqrt{\frac{2}{\pi t\sigma^2}} \exp\left(-\frac{m^2}{2t\sigma^2}\right), m \geq 0.$$

■

Consequence: $M_t \sim N_{[0,\infty)}(0, \sigma^2 t)$.

Remark: Truncated normal distribution $X \sim N(\mu, \sigma^2) \Rightarrow P(X \leq x | X \in [a, b])$. $N_{[a,b]}(\mu, \sigma^2)$ has the following density function:

$$f(x|\mu, \sigma, a, b) = \frac{\frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}.$$

Specially for $\mu = 0, a = 0, b = \infty$ we have:

$$f(x|0, \sigma, 0, \infty) = \frac{\frac{1}{\sigma}\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right)}{1 - \frac{1}{2}} = \sqrt{\frac{2}{\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

Remark: Simulation of the Wiener process. The Gaussian random walk: $\Delta t > 0$ sufficiently small, $t_n = n\Delta t; n = 1, 2, \ldots$. Generate $X_i \underset{i.i.d.}{\sim} N(0, \sigma^2), i = 1, 2, \ldots$. It holds $\sum_{i=1}^n X_i \sim N(0, n\sigma^2)$. It follows

$$\begin{aligned} W_{t_n} &= \sqrt{\Delta t} \sum_{i=1}^n X_i \sim N(0, n\Delta t\sigma^2) = N(0, \sigma^2 t_n) \\ W_{t_i} - W_{t_{i-1}} &\sim N(0, \sigma^2(t_i - t_{i-1}) = N(0, \sigma^2 \Delta t). \end{aligned}$$

R code:

```
> set.seed(139)
> N= 1000; T= 10; delta= T/N
> W=c(0, cumsum(sqrt(delta)*rnorm(N)))
> t= seq(0, T, by= delta)
> plot(t,W, type="l", main="Simulated Standard Wiener Process")
```

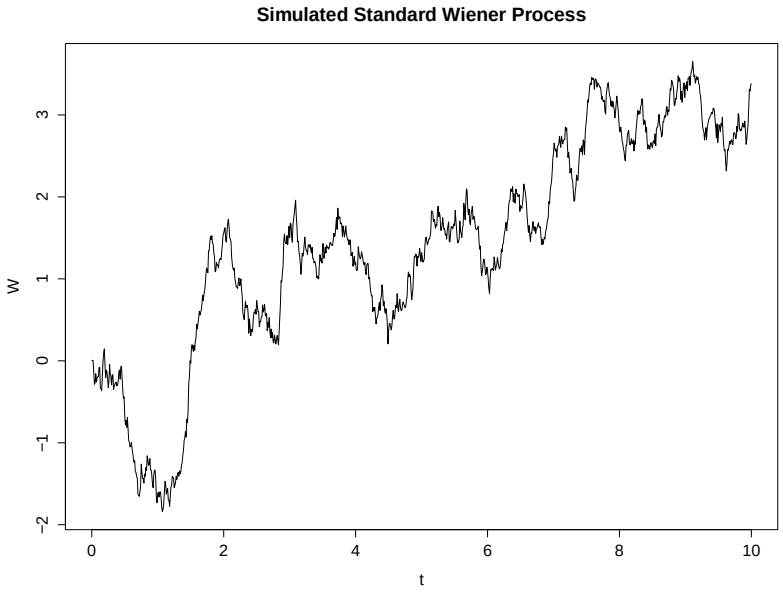


Figure 5.2: Output in R.

5.3 Generalization of the Wiener process

5.3.1 Wiener process with linear drift

Definition 5.8 *$X(t)$ is called a Wiener process with linear drift if*

$$X(t) = \mu t + W(t), W(t) - \textit{Wiener process}$$

where $\mu \in \mathbb{R}^1$ is called the drift parameter.

Corollary 5.2

$$m_X(t) = EX(t) = \mu t.$$

Interpretation: We get $X(t)$ by overlaying the Wiener process with linear increasing (decreasing) terms. From the definition of $W(t)$ we get:

a) $X(t)$ has independent, homogeneous increments.

b) For every increment, the following holds:

$$X(t) - X(s) = \mu(t-s) + W_{t-s} \sim N(\mu(t-s), \sigma^2(t-s)) \text{ for } 0 \leq s < t.$$

c) The first passage time T_m for the threshold m, of the Wiener process with drift parameter μ, is

$$f_{T_m}(t) = \frac{|m|}{\sqrt{2\pi\sigma^2 t^3}} \exp\left(-\frac{(|m| - |\mu|t)^2}{2\sigma^2 t}\right), 0 \leq t < \infty$$

the density function of the generalized inverse Gaussian distribution with parameters m, μ and σ^2. It holds:

$$E(T_m) = \frac{m}{\mu}, Var(T_m) = \frac{|m|\sigma^2}{|\mu|^3}.$$

Thus, for $\mu = 0$ (Wiener process without drift) the mathematical expectation and variance doesn't exist.

5.3.2 Geometric Wiener process with drift

Definition 5.9 *Let $X(t)$ be a Wiener process with drift. We call $Z(t) = e^{X(t)}$ the geometric Wiener process with drift.*

Application: In Finance statistics, the process is used to model share prices (interest rates) over time $Z(t) = z_0 e^{Xt}$, where $Z(0) = z_0$ is the stock price.

Example [American call-option]: An investor acquires at $t = 0$ the right to buy the option, in any subsequent time period $t > 0$ for z_0 Euro, independent of the market value. Most interesting are the threshold exceedances $z > z_0$, and the gain $z - z_0 : Z(t) \geq z \Leftrightarrow X(t) \geq \ln\left(\frac{z}{z_0}\right) =: m$.

Conversely: If $Z(t)$ is a geometric Wiener process with drift, then $X(t) = \ln[Z(t)]$ is a Wiener process.

5.3.3 Brownian bridge

Consider the conditional stochastic process $B_t := (W_t | W_1 = 0), t \in [0, 1]$. This process is known as the Brownian bridge, with the following properties:

- $EB_t = 0$,
- $Var(B_t) = t(1 - t)$,
- $Cov(B_s, B_t) = s(1 - t)$ for $0 \leq s < t \leq 1$.

Remark: the increments in the stochastic process $\{B_t : t \in [0, 1]\}$ are **not** independent.

Definition 5.10 (Alternative) *If $W_t, t \geq 0$ is a standard Wiener process, then $\{B_t : t \in [0, 1]\}$ is a Brownian bridge when $B_t = W_t - W(1)$ for $t \in [0, 1]$.*

Remark: The Brownian bridge can be represented as the series:

$$B_t = \sum_{k=1}^{\infty} X_k \frac{\sqrt{2}\sin(k\pi t)}{k\pi} \text{ where } X_k \underset{i.i.d}{\sim} N(0, 1)$$

known as the Karhunen–Loève decomposition.

5.4 Stochastic integral

Definition 5.11 *A Markov process $\varkappa = \{X_t\}_{t\in T}$ is called a diffusion process if it holds:*

i)
$$P(|X_{t+h} - X_t| \leq \varepsilon | X_t) = 1 - o(h), \forall \varepsilon > 0$$
where $\lim\limits_{h\to 0^+} \frac{o(h)}{h} = 0$, i.e. big changes in a small amount of time are improbable.

ii) $E\{X_{t+h} - X_t\} = \mu(t, X_t)h + o(h)$, *i.e. the local expectation of the (course) increment is proportional to the length of the time interval $[t, t+h]$. $\mu(t, X_t)$ is called the proportion factor, which depends on the time t and the current value of X_t, but doesn't dependent on the previous courses.*

iii) $Var\{X_{t+h} - X_t\} = \sigma(t, X_t)^2 h + o(h)$, *i.e. the local variance of the (course) increment is also proportional to the length of the time interval $[t, t+h]$ and also depends on the time t and the current value of X_t, but not on the previous courses.*

Remarks to Definition 5.11:

a) The factor $\mu(t, x)$ in definition 5.11 ii) is called drift of the stochastic process $\varkappa$.

b) The factor $\sigma(t, x)$ in definition 5.11 iii) is called the stochastic fluctuation (volatility) of the stochastic process $\varkappa$.

Now, considering the probability (density) transition function

$$p_{x,y}(s, t) = \frac{\partial}{\partial y} P(X_t \leq y | X_s = x), \forall t \geq s.$$

Under some regularity assumptions, we could uniquely determine the functions by $\mu(t, x)$ and $\sigma(t, x)$.

Regularity assumptions:

R1) The factors $\mu(t, x)$ and $\sigma(t, x)$ are measurable and real valued.

R2) Lipschitz condition:

$$\exists K_1 : |\mu(t, x) - \mu(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq K_1(n)|x - y|, \forall t \geq 0, \forall x, y : |x| \leq n, |y| \leq n.$$

R3) The growth bound:

$$\exists K_2 : |\mu(t, x)|^2 + |\sigma(t, x)|^2 \leq K_2(1 + x^2), \forall t \geq 0, \forall x \in \mathbb{R}^1.$$

Simple approach: Let $\mu(\cdot,\cdot)$ and $\sigma(\cdot,\cdot)$ be $\mu(t,X_t)=\mu X_t, \sigma(t,X_t)=\sigma|X_t|$. We get the following for this special case:

$$\begin{aligned} E[X_{t+h}-X_t|X_t] &= \mu X_t h + o(h) \\ E\left[\frac{X_{t+h}-X_t}{X_t}|X_t\right] &= \mu h + \frac{o(h)}{X_t} \\ \Rightarrow \lim_{h\to 0^+}\frac{1}{h}E\left[\frac{X_{t+h}-X_t}{X_t}|X_t\right] &= \mu \end{aligned}$$

I.e., infinitesimal (instantaneous) expected local percentual increments of the share price are constant. Analogous, we get for the local variance of the percentual increments:

$$\begin{aligned} & E\left\{\left[\frac{X_{t+h}-X_t}{X_t}-E\left(\frac{X_{t+h}-X_t}{X_t}|X_t\right)\right]^2|X_t\right\} \\ = \; & E\left[\left(\frac{X_{t+h}-X_t}{X_t}\right)^2|X_t\right]-\left[E\left(\frac{X_{t+h}-X_t}{X_t}|X_t\right)\right]^2 \\ = \; & \frac{\sigma^2X_t^2h}{X_t^2}+\frac{o(h)}{X_t^2}-\mu^2h^2-\frac{o^2(h)}{X_t^2}-2\frac{\mu h o(h)}{X_t} \\ \Rightarrow \; & \lim_{h\to 0^+}\frac{1}{h}Var\left[\frac{X_{t+h}-X_t}{X_t}|X_t\right]=\sigma^2 \end{aligned}$$

i.e., infinitesimal (moments) of local variance are also constant.
We would now like to describe the increment of a diffusion process over some time interval $\mathscr{I}=[t_0,t_e]$. Let $\mathscr{T}=\{t_0<t_1<\cdots<t_N=t_e\}$ be an equidistant partition of the time axis, i.e. $\Delta t=t_{i+1}-t_i=const.$ Choose the supporting points $\tau_i=[t_i,t_{i+1}]$, for $i=0,\ldots,N-1$ and model the increments, with given factors $\mu(\cdot,\cdot),\sigma(\cdot,\cdot)$ as follows:

$$X_{t_{i+1}}-X_{t_i}:=\mu(\tau_i,X_{\tau_i})(t_{i+1}-t_i)+\sigma(\tau_i,X_{\tau_i})(W_{t_{i+1}}-W_{t_i}). \qquad (\clubsuit)$$

Lemma 5.1 *If $\tau_i=t_i$, then the process defined by ($\clubsuit$), satisfies the requirements for the local expectation and variance of a diffusion process.*

Proof:

$$\begin{aligned} E\{X_{t_{i+1}}-X_{t_i}\} &= \mu(t_i,X_{t_i})(t_{i+1}-t_i)=\mu(t_i,X_{t_i})\Delta t. \\ Var\{X_{t_{i+1}}-X_{t_i}\} &= \sigma^2(t_i,X_{t_i})\Delta t. \end{aligned}$$

■

The increment to the time t_N is as follows:

$$X_{t_N} - X_{t_0} := \sum_{i=0}^{N-1} (X_{t_{i+1}} - X_{t_i}) = \underbrace{\sum_{i=0}^{N-1} \mu(\tau_i, X_{\tau_i})(t_{i+1} - t_i)}_{S_1} + \underbrace{\sum_{i=0}^{N-1} \sigma(\tau_i, X_{\tau_i})(W_{t_{i+1}} - W_{t_i})}_{S_2}. \quad \text{(L)}$$

Our goal: Determine the limit of this process for $N \to \infty$, i.e. $\Delta t \to 0$.
For the sum S_1 considering the **line integral** is sufficient, i.e. for arbitrary supporting points $\tau_i = [t_i, t_{i+1}]$ we define for every $\omega \in \Omega$ a line integral.

Definition 5.12 *Let $\mu : [0,T] \times \mathbb{R}^1 \to \mathbb{R}^1$ be a integrable function regarding time, then the following random variable*

$$Z_\mu(\omega) = \int_{t_0}^{t_e} \mu(t, X_t(\omega))dt := \lim_{\substack{\Delta t \to 0 \\ (N \to \infty)}} \sum_{i=0}^{N-1} \mu(\tau_i, X_{\tau_i})(t_{i+1} - t_i), \forall \omega \in \Omega$$

defines the line integral of $\varkappa$ with local expectation function μ.

Remark: The line integral is the generalisation of the Riemann integral from real analysis.
The limit of the sum S_2 does not exist considering line integrals, given that the trajectories of the Wiener process $\{W_t\}_{t \geq 0}$ are nowhere differentiable. Thus, for S_2 we need a different limiting approach, which leads us to the stochastic integral.
Recall: Convergence of sequence of random variables $\{X_i\}_{i=1,2,\ldots}$

- weak and strong convergence to X

$$P(\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\} \underset{n \to \infty}{\longrightarrow} 0 \quad (\Leftrightarrow) \quad X_n \underset{\substack{p \\ (a.s.)}}{\to} X,$$

- convergence in distribution

$$\lim_{n \to \infty} F_{X_n}(x) = F_X(x) \quad (\Leftrightarrow) \quad X_n \underset{d}{\to} X,$$

- $\{X_i\}_{i=1,2,\ldots}$ is convergent in quadratic mean over $[\Omega, \mathscr{L}, P]$ to X if

$$\lim_{n \to \infty} E_p\{X_n - X\}^2 = 0 \quad (\Leftrightarrow) \quad qm \lim_{n \to \infty} X_n = X \quad (\Leftrightarrow) \quad X_n \underset{q.m}{\to} X.$$

We can apply the convergent in quadratic mean to the summands of the sum S_2 through the sequence:

$$X_N := \sum_{i=0}^{N-1} \sigma(\tau_i, X_{\tau_i})(W_{t_{i+1}} - W_{t_i}), N = 1, 2, \ldots$$

where $\mathscr{T} = \mathscr{T}(N) = \{t_0 < t_1 < \ldots < t_N t_e\}$ and $\Delta t = \frac{t_e - t_0}{N} = t_{i+1} - t_i$.

Definition 5.13 *If the sequence $\{X_i\}_{i=1,2,\ldots}$ is convergent in quadratic mean, then*

$$\int_{t_0}^{T} \sigma(t, X_t) dW_t := \lim_{n\to\infty} \sum_{i=0}^{N-1} \sigma(\tau_i, X_{\tau_i})(W_{t_{i+1}} - W_{t_i})$$

is called the **stochastic integral** *of the factor $\sigma(\cdot,\cdot)$ regarding the Brownian motion.*

Thus, its possible to transform and do a discretization of the capital growth of a course into a continuous model.
If $\mu(\cdot,\cdot)$ and $\sigma(\cdot,\cdot)$ are given drift and fluctuation functions, such that the line and stochastic integral exist, then the process $\{X_t\}_{t\in\mathcal{T}}$ is called the **solution** of the stochastic integral equation:

$$\int_{t_0}^{t_e} dX_t = \int_{t_0}^{t_e} \mu(t, X_t)dt + \int_{t_0}^{t_e} \sigma(t, X_t)dW_t$$

with the initial condition $X_{t_0} = x_0$, or the **solution** of the stochastic differential equation:

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad t \in [t_0, t_e]$$

with the initial condition $X_{t_0} = x_0$.
Remark: The stochastic integral is defined as a real valued integral as a limit of a sequence, but the concepts of taking the limit are completely different.

5.4.1 Properties of stochastic integrals

Let W_t be a Wiener process, and, w.l.o.g., assume that $\sigma^2 = 1$.

S1) If $c \in \mathbb{R}^1$ is a constant, it holds:

$$\int_{t_0}^{t_e} cdW_t = c(W_{t_e} - W_{t_0}) \sim N(0, c^2(t_e - t_0))$$

which is a special case of the following property:

S2) If $\sigma(\cdot,\cdot) : [t_0, t_e] \to \mathbb{R}$ is square integrable, then it holds

$$\int_{t_0}^{t_e} \sigma(t)dW_t = X \sim N\left(0, \int_{t_0}^{t_e} \sigma(t)^2 dt\right)$$

Proof: For an arbitrary discretization $\{t_0, t_1, \ldots, t_N = t_e\}$, s.t. $t_0 < t_1 < \ldots < t_N$, and arbitrary

supporting points $\tau_i \in [t_i, t_{i+1}]$, it holds

$$\begin{aligned} & E \left\{ \sum_{i=0}^{N-1} \sigma(\tau_i)(W_{t_{i+1}} - W_{t_i}) \right\} = 0 \\ \Rightarrow \quad & E \left\{ \int_{t_0}^{t_e} \sigma(t) dW_t \right\} = 0 \end{aligned}$$

and given that the increments of the Brownian motion are independent, we get

$$\begin{aligned} & Var \left\{ \sum_{i=0}^{N-1} \sigma(\tau_i)(W_{t_{i+1}} - W_{t_i}) \right\} = \sum_{i=0}^{N-1} \sigma(\tau_i)(t_{i+1} - t_i) \underset{\Delta t \to 0}{\longrightarrow} \int_{t_0}^{t_e} \sigma(t)^2 dt \\ \Rightarrow \quad & Var \left\{ \int_{t_0}^{t_e} \sigma(t) dW_t \right\} = \int_{t_0}^{t_e} \sigma(t)^2 dt. \end{aligned}$$

Given that the sum of independent normally distributed random variables, is again a normal distributed random variable, we have that $\int_{t_0}^{t_e} \sigma(t) dW_t$ is normally distributed. ■

S3) Let $\sigma(\cdot)$ (deterministict factor) be a non deterministict factor. For the simplest stochastic integral of this form the following holds:

$$\begin{aligned} \int_{t_0}^{t_e} W_t dW_t &= qm \lim_{N \to \infty} \sum_{i=0}^{N-1} W_{\tau_i}(W_{t_{i+1}} - W_{t_i}) \\ &= \frac{1}{2}(W_{t_e}^2 - W_{t_0}^2) + (\alpha - \frac{1}{2})(t_e - t_0) \end{aligned}$$

where $\tau_i = (1-\alpha)t_i + \alpha t_{i+1}$ with $\alpha \in [0,1]$, are supporting points for $i = 1, \ldots, N-1$.
Remark upon the proof: define the sequence $\{X_i\}_{i=1,2,\ldots}$ with

$$X_n := \sum_{i=0}^{N-1} W_{\tau_i}(W_{t_{i+1}} - W_{t_i})$$

and shows that

$$qm \lim_{N \to \infty} = \frac{1}{2}(W_{t_e}^2 - W_{t_0}^2) + (\alpha - \frac{1}{2})(t_e - t_0)$$

using the relationship: $X \sim N(0, \sigma^2) \Rightarrow EX^4 = 3\sigma^2 EX^2 = 3\sigma^4$.
Remark: $\int_{t_0}^{T} W_t dW_t$ is dependent on the selection of supporting points (over α). We get a special integral for $\alpha = 0$, i.e. $\tau_i = t_i$ (left interval limit) , $i = 1, 2, \ldots, N-1$.

Definition 5.14 *For $\alpha = 0$ we get*

$$\int_{t_0}^{t_e} W_t dW_t = \frac{1}{2}(W_{t_e}^2 - W_{t_0}^2) - \frac{1}{2}(t_e - t_0)$$

which is called the Itô integral.

The $Itô$ integral has important applications in stochastic calculus of financial markets. $\alpha = 0$ means a non anticipatory trading strategy.

S4) Some properties that make it easier to work with the $Itô$ integral

i) $\int_{t_0}^{t_e} (dW_t)^2 = \int_{t_0}^{t_e} dt$, thus $(dW_t)^2 = dt$,

ii) $\int_{t_0}^{t_e} dt dW_t = 0$, thus $dt dW_t = 0$,

iii) $\int_{t_0}^{t_e} (dt)^2 = 0$, thus $(dt)^2 = 0$.

References

[1] H.M. Taylor, S. Karlin (1998) *An Introduction to Stochastic Modeling 3. ed.,* Academic Press

[2] R. Durrett (2012) *Essential of Stochastic Processes 2. ed.,* Springer

[3] S.M. Ross (2010) *Introduction to Probability Models 10. ed.,* Academic Press

[4] F. Beichelt (1997) *Stochastische Prozesse für Ingenieure,* Teubner - Verlag, Stuttgart

[5] J.R.Norris (1998) *Markov Chains,* Cambridge University Press

[6] R.N. Bhattacherrya, E.C. Waymire (2009) *Stochastic Processes with Applications,* SIAM Philadelphia

[7] Yu.V. Prokhorov (1997) *Probability Theory: Vol. 3: Stochastic Calculus*, Springer, Berlin

[8] P.Billingsley (1995) *Probability and MEasure 3. ed.,* Wiley, New York

[9] F.C. Klebaner (2012) *Introduction to Stochastic Calculus with applications 3. ed.,* Imperial College Press

[10] W. Paul (2000) *Stochastic Processes. From Physics to Finance*, Springer, Berlin

[11] L.J.S. Allen (2011) *An Introduction to Stochastic Processes with applications to Biology 2. ed.,* Chapman and Hall

[12] C. Gaetan, X. Guyon (2010) *Spatial Statistics and Modeling,* Springer

[13] F. Beichelt (2006) *Stochastic processes in science, engineering and finance,* Chapman and Hall

Remark:

a) introductory texts [1] - [6],

b) advanced texts [7] - [9],

c) specific applications [10] - [13].