

Stochastic processes

1. | Preliminaries

Proposition 1. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and (X_n) be sequence of random variables such that:

$$\sum_{n=1}^{\infty} \mathbb{E}(|X_n|) < \infty$$

Then, $\sum_{n=1}^{\infty} X_n$ is a random variable and:

$$\mathbb{E}\left(\sum_{n=1}^{\infty} X_n\right) = \sum_{n=1}^{\infty} \mathbb{E}(X_n)$$

Proof. By the ?? ?? we have that

$$\mathbb{E}\left(\sum_{n=1}^{\infty} |X_n|\right) = \sum_{n=1}^{\infty} \mathbb{E}(|X_n|) < \infty$$

so the random variable $Y := \sum_{n=1}^{\infty} |X_n|$ is integrable and $\forall N \in \mathbb{N}$ satisfies:

$$\left|\sum_{n=1}^N X_n\right| \leq Y$$

Now use the ?? ??.

Proposition 2 (Law of total probability). Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $\{A_n : 1 \leq n \leq N\} \subset \mathcal{A}$, $N \in \mathbb{N} \cup \{\infty\}$, be such that $\bigcup_{n=1}^N A_n = \Omega'$ with $\mathbb{P}(\Omega') = 1$. Then, $\forall A \in \mathcal{A}$:

$$\mathbb{P}(A) = \sum_{n=1}^N \mathbb{P}(A_n) \mathbb{P}(A | A_n)$$

Sketch of the proof. See the proof of ?? ??.

Remark. Note that if there is some A_n for which $\mathbb{P}(A_n) = 0$, the conditional probability is not well-defined. But note that:

$$0 \leq \mathbb{P}(A_n) \mathbb{P}(A | A_n) = \mathbb{P}(A \cap A_n) \leq \mathbb{P}(A_n) = 0$$

So we can omit this term in the sum.

Conditional expectation

Proposition 3 (Substitution principle). Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\mathbf{X}$ be a discrete random vector with support $S_{\mathbf{X}}$, Y be a discrete random variable with support S_Y , $h : S_{\mathbf{X}} \times S_Y \rightarrow \mathbb{R}$ be a function and $y \in S_Y$. Then:

$$\mathbb{E}(h(\mathbf{X}, Y) | Y = y) = \mathbb{E}(h(\mathbf{X}, y) | Y = y)$$

Proposition 4 (Law of total expectation). Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\mathbf{X}$ is a discrete random vector with support $S_{\mathbf{X}}$, $\{A_n : 1 \leq n \leq N\} \subset \mathcal{A}$, $N \in \mathbb{N} \cup \{\infty\}$, be such that $\bigcup_{n=1}^N A_n = \Omega'$ with $\mathbb{P}(\Omega') = 1$ and $h : S_{\mathbf{X}} \rightarrow \mathbb{R}$ be a function. If $h(\mathbf{X})$ has finite expectation or $h \geq 0$, then:

$$\mathbb{E}(h(\mathbf{X})) = \sum_{n=1}^N \mathbb{E}(h(\mathbf{X}) | A_n) \mathbb{P}(A_n)$$

Sketch of the proof. See the proof of ?? ??.

Definition 5. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\mathbf{X}$ be a discrete random vector with support $S_{\mathbf{X}}$, Y be a discrete random variable with support S_Y and $h : S_{\mathbf{X}} \rightarrow \mathbb{R}$ be a function. For all $\omega \in \Omega$ we define the random variable $\mathbb{E}(h(\mathbf{X}) | Y)$ as

$$\mathbb{E}(h(\mathbf{X}) | Y)(\omega) = \mathbb{E}(h(\mathbf{X}) | Y = y)$$

provided that $Y(\omega) = y$. Note that it can also be written as:

$$\mathbb{E}(h(\mathbf{X}) | Y) = \sum_{y \in S_Y} \mathbb{E}(h(\mathbf{X}) | Y = y) \mathbf{1}_{\{Y=y\}}$$

Proposition 6. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\mathbf{X}$ be a discrete random vector with support $S_{\mathbf{X}}$, Y be a discrete random variable with support S_Y , $h, h_1, h_2 : S_{\mathbf{X}} \rightarrow \mathbb{R}$ be functions and $a, b \in \mathbb{R}$. Then:

1. $\mathbb{E}(ah_1(\mathbf{X}) + bh_2(\mathbf{X}) | Y) = a\mathbb{E}(h_1(\mathbf{X}) | Y) + b\mathbb{E}(h_2(\mathbf{X}) | Y)$
2. If $\mathbf{X}$ and Y are independent, then $\mathbb{E}(h(\mathbf{X}) | Y) = \mathbb{E}(h(\mathbf{X}))$.
3. $\mathbb{E}(\mathbb{E}(h(\mathbf{X}) | Y)) = \mathbb{E}(h(\mathbf{X}))$

Sketch of the proof. The first two properties are consequence of the fact that the conditional expectation is an expectation. For the last one note that:

$$\begin{aligned} \mathbb{E}(\mathbb{E}(h(\mathbf{X}) | Y)) &= \sum_{y \in S_Y} \mathbb{E}(h(\mathbf{X}) | Y = y) \mathbb{P}(Y = y) \\ &= \mathbb{E}(h(\mathbf{X})) \end{aligned}$$

where the last equality is due to the 4 Law of total expectation.

Theorem 7 (Wald theorem). Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, (Z_n) be a sequence of random variables, all of them with expectation $\mu \in \mathbb{R}$, such that $\sup_{n \geq 1} \mathbb{E}(|Z_n|) = A < \infty$. If N is an integrable random variable with support $\mathbb{N}$ independent of $Z_n \forall n \in \mathbb{N}$, we have that:

$$\mathbb{E}\left(\sum_{n=1}^N Z_n\right) = \mu \mathbb{E}(N)$$

Proof. Note that $\mathbb{E}\left(\sum_{n=1}^N Z_n\right) = \mathbb{E}\left(\sum_{n=1}^{\infty} Z_n \mathbf{1}_{N \geq n}\right)$, and it is integrable because:

$$\begin{aligned} \mathbb{E}\left(\sum_{n=1}^{\infty} |Z_n| \mathbf{1}_{N \geq n}\right) &= \sum_{n=1}^{\infty} \mathbb{E}(|Z_n|) \mathbb{E}(\mathbf{1}_{N \geq n}) \leq \\ &\leq A \sum_{n=1}^{\infty} \mathbb{P}(N \geq n) = A \mathbb{E}(N) < \infty \end{aligned}$$

where we have used the Independence of (Z_n) and N in the first equality. And so:

$$\mathbb{E}\left(\sum_{n=1}^N Z_n\right) = \sum_{n=1}^{\infty} \mathbb{E}(Z_n) \mathbb{E}(\mathbf{1}_{N \geq n}) = \mu \mathbb{E}(N)$$

□

Remark. Note that the equality remains true if $Z_n \geq 0 \forall n \in \mathbb{N}$ even if $\mathbb{E}(N) = \infty$ because the equality $\sum_{n=1}^{\infty} \mathbb{P}(N \geq n) = \mathbb{E}(N)$ remains true.

Probability-generating function

Definition 8. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X be a random variable with support $\mathbb{N} \cup \{0\}$. The *probability-generating function* (or *pgf*) of X is the function $g_X : \mathcal{D}_X \rightarrow \mathbb{R}$ defined as:

$$g_X(s) = \sum_{k=0}^{\infty} s^k \mathbb{P}(X = k) = \mathbb{P}(X = 0) + \sum_{k=1}^{\infty} s^k \mathbb{P}(X = k) \quad (1)$$

The set $\mathcal{D}_X$ is defined as all the points for which the series of Eq. (1) converges absolutely.

Lemma 9. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X be a random variable with support $\mathbb{N} \cup \{0\}$. Then, $[-1, 1] \subseteq \mathcal{D}_X$ and $g_X(s) = \mathbb{E}(s^X)$ (with the convention that $0^0 = 1$).

Proof. Clearly $g_X(s) = \mathbb{E}(s^X)$ and furthermore $\forall s \in [-1, 1]$ we have $s \in \mathcal{D}_X$ because:

$$\sum_{k=0}^{\infty} |s|^k \mathbb{P}(X = k) \leq \sum_{k=0}^{\infty} \mathbb{P}(X = k) = 1 < \infty$$

Theorem 10. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X, Y be random variables with support $\mathbb{N} \cup \{0\}$. Then:

$$X \stackrel{d}{=} Y \iff g_X = g_Y$$

Moreover:

$$\mathbb{P}(X = k) = \frac{g_Y^{(k)}(0)}{k!} \quad \forall k \geq 0$$

Sketch of the proof. Note that $g_X(s)$ is a power series defined in a neighbourhood of $s = 0$ (recall ??). □

Theorem 11. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X, Y be independent random variables with support $\mathbb{N} \cup \{0\}$. Then, $\forall s \in \mathcal{D}_X \cap \mathcal{D}_Y$ we have:

$$g_{X+Y}(s) = g_X(s)g_Y(s)$$

Proof. Note that if $s \in \mathcal{D}_X \cap \mathcal{D}_Y$, then $s \in \mathcal{D}_{X+Y}$ because

$$g_{X+Y}(|s|) = \mathbb{E}(|s|^{X+Y}) = \mathbb{E}(|s|^X) \mathbb{E}(|s|^Y) < \infty$$

due to the independence of X and Y . To show the equality if $s = 0$, we have:

$$\mathbb{P}(X + Y = 0) = \mathbb{P}(X = 0, Y = 0) = \mathbb{P}(X = 0)\mathbb{P}(Y = 0)$$

where the first equality is due to $X, Y \in \mathbb{N} \cup \{0\}$ and the second one is because of the independence. If $s \neq 0$, as before:

$$g_{X+Y}(s) = \mathbb{E}(s^{X+Y}) = \mathbb{E}(s^X) \mathbb{E}(s^Y) = g_X(s)g_Y(s)$$

□

Theorem 12. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X be a random variable with support $\mathbb{N} \cup \{0\}$. Then, $\forall k \geq 1$ we have:

$$\lim_{s \rightarrow 1^-} g_X^{(k)}(s) = \mathbb{E}(X(X-1) \cdots (X-k+1))$$

Sketch of the proof. Take $(s_n) \in \mathbb{R}$ such that $s_n \nearrow 1$. Differentiating term by term we have that:

$$g_X^{(k)}(s_n) = \mathbb{E}(X(X-1) \cdots (X-k+1)(s_n)^{X-k})$$

for all $n \in \mathbb{N}$. Moreover, note that $X(X-1) \cdots (X-k+1)(s_n)^{X-k} \nearrow X(X-1) \cdots (X-k+1)$. Now use the ?? □

X	$g_X(s)$	$\mathcal{D}_X$
$k \in \mathbb{N}$	s^k	$\mathbb{R}$
$U(\{k_1, \dots, k_n\})$	$\frac{1}{n} \sum_{i=1}^n s^{k_i}$	$\mathbb{R}$
$B(n, p)$	$(ps + 1 - p)^n$	$\mathbb{R}$
$\text{Pois}(\lambda)$	$e^{\lambda(s-1)}$	$\mathbb{R}$
$\text{Geo}(p)$	$\frac{ps}{1 - (1-p)s}$	$\left(\frac{-1}{1-p}, \frac{1}{1-p}\right)$

□ Table 1: Probability-generating functions of common distributions.

2. | Discrete-time Markov chains

Stochastic processes

Definition 13 (Stochastic process). Let $T \subseteq \mathbb{R}^n$ be a set, $(E, \mathcal{E})$ be a measurable space and $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. A *stochastic process* on $(\Omega, \mathcal{A}, \mathbb{P})$ with *parameter set* T and *state space* $(E, \mathcal{E})$ is a family of random variables $\{X_t\}_{t \in T}$ from $(\Omega, \mathcal{A})$ to $(E, \mathcal{E})$. That is, $X_t : \Omega \rightarrow E$ satisfies $X_t^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{E}$ and all $t \in T$.

Remark. In general, we will consider stochastic processes with parameter sets $T = \mathbb{N}, \mathbb{N} \cup \{0\}, \mathbb{Z}, \mathbb{R}, \mathbb{R}_{\geq 0}$ and state spaces $(\mathbb{N} \cup \{0\}, \mathcal{P}(\mathbb{N} \cup \{0\}))$ or $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Definition 14. Let $(X_t)_{t \in T}, (Y_t)_{t \in T}$ be two stochastic processes defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We say that $(X_t)_{t \in T}$ and $(Y_t)_{t \in T}$ are *independent* if $\forall n, k \in \mathbb{N}$ and all $t_1, \dots, t_n, s_1, \dots, s_k \in T$ we have that the random vectors $(X_{t_1}, \dots, X_{t_n})$ and $(Y_{s_1}, \dots, Y_{s_k})$ are independent.

Galton-Watson process

Model 15. Let (X_n) , $n \in \mathbb{N} \cup \{0\}$ be a sequence of discrete random variables representing the number of new individuals of a certain population at the n -th generation. Suppose they are defined as

$$X_{n+1} = \sum_{k=1}^{X_n} Z_{n+1}^{(k)}$$

and $X_0 = 1$. Here $Z_{n+1}^{(k)}$ has support $\mathbb{N} \cup \{0\} \forall n, k$ and represent the number of descendants (to the next generation) of the k -th individual of the n -th generation. Suppose that $Z_{n+1}^{(k)} \sim Z$ are i.i.d. and independent of (X_n) . We would like to study the probability ρ of extinction of this population:

$$\rho = \mathbb{P}(\{X_n = 0 : \text{for some } n \in \mathbb{N}\}) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} \{X_n = 0\}\right)$$

Lemma 16. Let (Z_n) be a sequence of i.i.d. random variables distributed as Z with support $\mathbb{N} \cup \{0\}$, and N be a random variable also with support $\mathbb{N} \cup \{0\}$ and independent to (Z_n) . Let $X = \sum_{k=1}^N Z_k$. Then, $\forall s \in [-1, 1]$ we have:

$$g_X(s) = g_N(g_Z(s))$$

Proof. First suppose $N \leq M$ with $M \in \mathbb{N}$ fixed. Then using the independence, the 4 Law of total expectation and the 3 Substitution principle:

$$\begin{aligned} g_X(s) &= \mathbb{E}(s^X) = \sum_{k=1}^M \mathbb{E}(s^X \mid N = k) \mathbb{P}(N = k) \\ &= \sum_{k=1}^M \mathbb{E}\left(s^{\sum_{i=1}^N Z_i} \mid N = k\right) \mathbb{P}(N = k) \\ &= \sum_{k=1}^M \mathbb{E}\left(s^{\sum_{i=1}^k Z_i}\right) \mathbb{P}(N = k) \\ &= \sum_{k=1}^M g_Z(s)^k \mathbb{P}(N = k) \\ &= g_N(g_Z(s)) \end{aligned}$$

Now if N can take any value of $\mathbb{N} \cup \{0\}$ we have that:

$$\begin{aligned} g_X(s) &= \mathbb{E}\left(s^{\sum_{i=1}^N Z_i}\right) \\ &= \mathbb{E}\left(\lim_{M \rightarrow \infty} s^{\sum_{i=1}^{\min(N, M)} Z_i}\right) \\ &= \lim_{M \rightarrow \infty} \mathbb{E}\left(s^{\sum_{i=1}^{\min(N, M)} Z_i}\right) \\ &= \lim_{M \rightarrow \infty} g_{\min(N, M)}(g_Z(s)) \\ &= \lim_{M \rightarrow \infty} \mathbb{E}\left((g_Z(s))^{\min(N, M)}\right) \\ &= \mathbb{E}\left(\lim_{M \rightarrow \infty} (g_Z(s))^{\min(N, M)}\right) \\ &= g_N(g_Z(s)) \end{aligned}$$

where both limit exchangings are due to the ?? ?? using the integrable random variable 1. $\square$

Theorem 17. In the hypothesis of Theorem 15, we have that:

$$\rho = g_Z(\rho)$$

Proof. Note that $\{X_n = 0\} \subseteq \{X_{n+1} = 0\}$. Hence:

$$\rho = \mathbb{P}\left(\bigcup_{n=1}^{\infty} \{X_n = 0\}\right) = \lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0) = \lim_{n \rightarrow \infty} g_{X_n}(0)$$

Now, using Theorem 16 we have:

$$g_{X_n}(s) = g_{X_{n-1}}(g_Z(s)) = \cdots = g_{X_1}(g_Z^n(s))$$

But $X_1 = 1$ and so $g_{X_1}(s) = s$. So $g_{X_n}(s) = g_Z^n(s)$ and therefore $g_{X_{n+1}}(0) = g_Z(g_{X_n}(0))$. Taking the limit as $n \rightarrow \infty$ and using the continuity of the pgf we get the result. $\square$

Theorem 18. In the hypothesis of Theorem 15 and the additional assumption that $0 < \mathbb{P}(Z = 0) < 1$ we have:

1. If $\mathbb{E}(Z) \leq 1$, g_Z has only 1 fixed point (the trivial one, $s = 1$). Hence, the population will extinct with probability 1.
2. If $\mathbb{E}(Z) > 1$, g_Z has a unique non-trivial fixed point on $(0, 1)$.

Proof. First suppose $\mathbb{P}(Z = 0) + \mathbb{P}(Z = 1) = 1$. Thus, $Z \stackrel{\text{a.e.}}{\leq} 1$ and so $\mathbb{E}(Z) \leq 1$. Moreover, $g_Z(s) = \mathbb{P}(Z = 0) + s\mathbb{P}(Z = 1)$, which is a line with slope $\mathbb{P}(Z = 1) < 1$. Hence, it has a unique fixed point, which is $s = 1$.

Now assume $\mathbb{P}(Z = 0) + \mathbb{P}(Z = 1) < 1$. Then, $\exists k \geq 2$ with $\mathbb{P}(Z = k) > 0$. Hence, $g_Z'(s) > 0$ and $g_Z''(s) > 0 \forall s \in (0, 1)$. Now consider $f(s) = g(s) - s$. Note that f is strictly convex in $(0, 1)$ and $f(0) = g(0) = \mathbb{P}(Z = 0) > 0$. Finally, note that

$$\lim_{t \rightarrow 1^-} f'(s) = \lim_{t \rightarrow 1^-} g(s) - 1 = \mathbb{E}(Z) - 1$$

and so $\lim_{t \rightarrow 1^-} f'(s)$ is negative in the first case and positive in the second case. This implies that f has no zeros on $(0, 1)$ in the first case and exactly 1 zero in $(0, 1)$ in the second case.

It's missing to see that in the second case the probability of extinction ρ is given by the fixed point in $(0, 1)$, rather than 1. We have that:

$$\rho = \lim_{n \rightarrow \infty} g_{X_n}(0) = \lim_{n \rightarrow \infty} g_Z^n(0)$$

Since $g_Z' > 0$, we have that g_Z is increasing and so it is $g_Z^n \forall n \in \mathbb{N}$. Moreover, if $g_Z(x_0) = x_0$, we have that $g_Z^n(x_0) = x_0 \forall n \in \mathbb{N}$. Therefore,

$$0 < g_Z(0) < g_Z^2(0) < \cdots < g_Z^n(0) < \cdots < x_0 < \cdots < 1$$

And so the limit has to be x_0 (note that the limit does exist because $(g_Z^n(0))$ is an increasing bounded sequence). $\square$

Gambler's ruin

Definition 19 (Gambler's ruin problem). Consider a gambler with an initial capital $z \in \mathbb{Z}$ and suppose that he plays a game in which wins 1 unit of capital with probability p and loses 1 unit of capital with probability $q := 1 - p$. The game ends whenever the player is ruined or if he arrives to a capital of $a \in \mathbb{Z}$. All the plays are independent. We denote by (X_k) the variables that measure the k -th play. That is:

$$\mathbb{P}(X_k = 1) = p \quad \mathbb{P}(X_k = -1) = q$$

We define q_z as the probability of ruining himself starting with a capital of z , p_z as the probability of winning the game starting with a capital of z and D_z as the duration of the game starting with a capital of z .

Proposition 20. Consider the Gambler's ruin problem. Then:

$$q_z = \begin{cases} \frac{-\left(\frac{q}{p}\right)^a + \left(\frac{q}{p}\right)^z}{1 - \left(\frac{q}{p}\right)^a} & \text{if } p \neq 1/2 \\ 1 - \frac{z}{a} & \text{if } p = 1/2 \end{cases}$$

Sketch of the proof. We have that q_k solves the difference equation

$$q_k = \mathbb{P}(\text{ruin} \mid X_k = 1)\mathbb{P}(X_1 = 1) + \mathbb{P}(\text{ruin} \mid X_k = -1)\mathbb{P}(X_1 = -1) = q_{k+1}p + q_{k-1}q$$

with $q_0 = 1$ and $q_a = 0$, whose solution is straightforward. $\square$

Proposition 21. Consider the Gambler's ruin problem. Suppose that we play against another player (and so when we lose, he wins and vice versa). Let p_z^* , q_z^* be the respective probabilities for the other player. Then:

$$q_z + q_z^* = 1$$

Hence, $D_z \stackrel{\text{a.e.}}{<} \infty$.

Sketch of the proof. Note that

$$q_z^* = \begin{cases} \frac{-\left(\frac{p}{q}\right)^a + \left(\frac{p}{q}\right)^{a-z}}{1 - \left(\frac{p}{q}\right)^a} & \text{if } p \neq 1/2 \\ 1 - \frac{a-z}{a} & \text{if } p = 1/2 \end{cases}$$

Proposition 22. Let $d_z = \mathbb{E}(D_z)$ and suppose that this expectation is finite. Then:

$$d_z = \begin{cases} \frac{z}{q-p} - \frac{a}{q-p} \frac{1 - \left(\frac{q}{p}\right)^z}{1 - \left(\frac{q}{p}\right)^a} & \text{if } p \neq 1/2 \\ z(a-z) & \text{if } p = 1/2 \end{cases}$$

Proof. We have that q_k solves the difference equation:

$$\begin{aligned} d_k &= \mathbb{E}(D_k \mid X_k = 1)\mathbb{P}(X_1 = 1) + \mathbb{E}(D_k \mid X_k = -1)\mathbb{P}(X_1 = -1) = \\ &= \mathbb{E}(D_{k+1} + 1)p + \mathbb{E}(D_{k-1} + 1)q = d_{k+1}p + d_{k-1}q + 1 \end{aligned}$$

with $d_0 = 0$ and $d_a = 0$, whose solution is straightforward ($d_k = \frac{k}{p-q}$ and $d_k = -k^2$ are particular solutions for the case $p \neq q$ and $p = q$, respectively). $\square$

Markov chains

Definition 23. A *Markov chain* is a sequence of discrete random variables (X_n) with support I such that:

$$\begin{aligned} \mathbb{P}(X_{n+1} = j \mid X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i) = \\ = \mathbb{P}(X_{n+1} = j \mid X_n = i) \end{aligned}$$

for all $n \geq 0$ and all $i_0, \dots, i_{n-1}, i, j \in I$. This property is usually called *Markov property*. If moreover $\mathbb{P}(X_{n+1} = j \mid X_n = i)$ does not depend on n , that is

$$\mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(X_1 = j \mid X_0 = i)$$

then we say that the Markov chain is a *time-homogeneous Markov chain*. The set I is called *state space* and its elements are called *states* of the Markov chain.

Definition 24 (Stochastic matrix). Let I be an index set. A matrix $\mathbf{P} = (p_{ij})_{i,j \in I}$ is called a *stochastic matrix* if $p_{ij} \geq 0 \forall i, j \in I$ and:

$$\sum_{j \in I} p_{ij} = 1$$

Definition 25. Let (X_n) be a time-homogeneous Markov chain. We define the *transition probabilities* p_{ij} as the probability of going from state i to state j . That is:

$$p_{ij} = \mathbb{P}(X_1 = j \mid X_0 = i)$$

The matrix $\mathbf{P} = (p_{ij})_{i,j \in I}$ is called the *transition matrix* of the Markov chain. Finally, we define the probabilities π_i as $\pi_i = \mathbb{P}(X_0 = i)$. We define the vector $\boldsymbol{\pi} = (\pi_i)_{i \in I}$ as the *initial distribution* of the Markov chain.

Proposition 26. Let (X_n) be a time-homogeneous Markov chain. Then:

1. $\mathbf{P}$ is a stochastic matrix.

2. $\sum_{i \in I} \pi_i = 1$.

$\square$

Lemma 27. Let I, F be finite or countable set, (Z_n) be a sequence of random variables with support F , X_0 be a random variable with support I and $f : I \times F \rightarrow I$ be a function. Consider the sequence (X_n) defined by:

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

If $\forall i_0, \dots, i_{n-1}, i \in I$ and $\forall k \in F$ we have:

$$\begin{aligned} \mathbb{P}(Z_{n+1} = k \mid X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i) = \\ = \mathbb{P}(Z_{n+1} = k \mid X_n = i) = \mathbb{P}(Z_1 = k \mid X_0 = i) \end{aligned}$$

then (X_n) is a time-homogeneous Markov chain with transition matrix $\mathbf{P} = (p_{ij})_{i,j \in I}$ given by:

$$p_{ij} = \mathbb{P}(f(i, Z_1) = j \mid X_0 = i)$$

Proof. Let $C = \{X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i\}$ and let $A_{i,j} := \{k \in F : f(i, k) = j\}$. We have:

$$\begin{aligned} \mathbb{P}(X_{n+1} = j \mid C) &= \mathbb{P}(f(i, Z_{n+1}) = j \mid C) \\ &= \mathbb{P}(Z_{n+1} \in A_{i,j} \mid C) \\ &= \sum_{k \in A_{i,j}} \mathbb{P}(Z_{n+1} = k \mid C) \\ &= \sum_{k \in A_{i,j}} \mathbb{P}(Z_1 = k \mid X_0 = i) \\ &= \mathbb{P}(f(i, Z_1) = j \mid X_0 = i) \end{aligned}$$

Definition 28 (Random walk). A random walk is a sequence $(\mathbf{S}_n)$ with $\mathbf{S}_0 = \mathbf{X}_0$ and $\mathbf{S}_n = \sum_{k=0}^n \mathbf{X}_k$, where $(\mathbf{X}_k)_{k \geq 1}$ is a sequence of i.i.d. random vectors and $\mathbf{X}_0$ is a random vector independent of $(\mathbf{X}_k)$.

Definition 29. A simple random walk is a random walk in which in one step we can only pass from one state to its neighbours. That is, if the random walk is in $\mathbb{Z}$, X_k are random variables such that:

$$X_k = \begin{cases} 1 & \text{with probability } p \\ -1 & \text{with probability } 1 - p \end{cases}$$



Figure 1: A simple random walk of 10000 steps in $\mathbb{Z}^2$. The green and blue dots are the respective initial and final positions of the random walk.

Lemma 30. A sequence of i.i.d. random variables, a random walk and a Galton-Watson process are all time-homogeneous Markov chains.

Definition 31. Let (X_n) be a time-homogeneous Markov chain. We define the n -step transition probabilities $p_{ij}^{(n)}$ as the probability of going from state i to state j in n steps. That is:

$$p_{ij}^{(n)} = \mathbb{P}(X_n = j \mid X_0 = i)$$

The matrix $\mathbf{P}^{(n)} = (p_{ij}^{(n)})_{i,j \in I}$ is called n -step transition matrix of the Markov chain.

Definition 32. Let (X_n) be a time-homogeneous Markov chain. We define the probabilities $\pi_i^{(n)}$ as the probability of being in state i after n steps. That is:

$$\pi_i^{(n)} = \mathbb{P}(X_n = i)$$

We define the vector $\boldsymbol{\pi}^{(n)} = (\pi_i^{(n)})_{i \in I}$ as n -step distribution of the Markov chain.

Lemma 33. Let A, B, C be events in a probability space such that $\mathbb{P}(B \cap C) > 0$. Then:

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

Proof.

$$\begin{aligned} \mathbb{P}(A \cap B \mid C) &= \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(B \cap C)} \mathbb{P}(A \mid B \cap C) \\ &= \mathbb{P}(B \mid C) \mathbb{P}(A \mid B \cap C) \end{aligned}$$

Lemma 34. Let I be a finite or countable set and A and D_i for $i \in I$ be events in a probability space such that $\mathbb{P}(A \mid D_i) = p$ for all $i \in I$ and such that the D_i are pairwise disjoint. Then:

$$\mathbb{P}\left(A \mid \bigcup_{i \in I} D_i\right) = p$$

Proof.

$$\begin{aligned} \mathbb{P}\left(A \mid \bigcup_{i \in I} D_i\right) &= \frac{\mathbb{P}(A \cap \bigcup_{i \in I} D_i)}{\mathbb{P}(\bigcup_{i \in I} D_i)} \\ &= \frac{\sum_{i \in I} \mathbb{P}(A \cap D_i)}{\sum_{i \in I} \mathbb{P}(D_i)} \\ &= \frac{\sum_{i \in I} \mathbb{P}(A \mid D_i) \mathbb{P}(D_i)}{\sum_{i \in I} \mathbb{P}(D_i)} \\ &= \frac{\sum_{i \in I} p \mathbb{P}(D_i)}{\sum_{i \in I} \mathbb{P}(D_i)} \\ &= p \end{aligned}$$

Theorem 35. Let (X_n) be a time-homogeneous Markov chain. Then, $\mathbf{P}^{(n)} = \mathbf{P}^n$.

Proof. By induction on n . The case $n = 1$ is clear. For $n \geq 2$ we have:

$$\begin{aligned} p_{ij}^{(n+1)} &= \mathbb{P}(X_{n+1} = j \mid X_0 = i) \\ &= \sum_{k \in I} \mathbb{P}(X_{n+1} = j, X_n = k \mid X_0 = i) \\ &= \sum_{k \in I} \mathbb{P}(X_n = k \mid X_0 = i) \cdot \mathbb{P}(X_{n+1} = j \mid X_n = k, X_0 = i) \\ &= \sum_{k \in I} p_{ik}^{(n)} p_{kj}^{(1)} \end{aligned}$$

where the penultimate equality follows from [Theorem 33](#) and the last equality follows from [Theorem 34](#) and the Markov property because if $D = \{X_n = k, X_0 = i\}$ we have that:

$$D = \bigsqcup_{i_1, \dots, i_{n-1} \in I} \{X_n = k, X_0 = i, X_1 = i_1, \dots, X_{n-1} = i_{n-1}\}$$

and so:

$$\begin{aligned} \mathbb{P}(X_{n+1} = j \mid X_n = k, X_0 = i) &= \mathbb{P}(X_{n+1} = j \mid X_n = k, \\ &\quad X_0 = i, X_1 = i_1, \dots, X_{n-1} = i_{n-1}) = \\ &= \mathbb{P}(X_{n+1} = j \mid X_n = k) \end{aligned}$$

Therefore, $\mathbf{P}^{(n+1)} = \mathbf{P}^n \mathbf{P}$, by induction hypothesis. $\square$

Theorem 36 (Chapman-Kolmogorov equation). Let (X_n) be a time-homogeneous Markov chain and $i, j \in I$. Then:

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

Proposition 37. Let (X_n) be a time-homogeneous Markov chain. Then:

1. $\mathbf{P}^{(0)} = \mathbf{I}_I$
2. $\pi^{(n)} = \pi^{(0)} \mathbf{P}^n$
3. $\mathbb{P}(X_0 = i_0, \dots, X_n = i_n) = \pi_{i_0}^{(0)} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}$

Sketch of the proof.

1. $p_{ij}^{(0)} = \mathbb{P}(X_1 = j \mid X_0 = i) = \delta_{ij}$.

2.

$$\begin{aligned} \pi_i^{(n)} &= \mathbb{P}(X_n = i) = \sum_{k \in I} \mathbb{P}(X_n = i \mid X_0 = k) \cdot \\ &\quad \cdot \mathbb{P}(X_0 = k) = \sum_{k \in I} \pi_k^{(0)} p_{ki}^{(n)} \end{aligned}$$

3. Use the ?? ?? and the Markov property.

$\square$

Classification of states

Definition 38. Let (X_n) be a time-homogeneous Markov chain. We say that a state $j \in I$ is *reachable* from $i \in I$ if $\exists n \in \mathbb{N} \cup \{0\}$ such that $p_{ij}^{(n)} > 0$. In this case we will write $i \rightarrow j$.

Definition 39. Let (X_n) be a time-homogeneous Markov chain. We say that two states $i, j \in I$ *communicate* if $i \rightarrow j$ and $j \rightarrow i$. In this case we will write $i \leftrightarrow j$.

Lemma 40. Let (X_n) be a time-homogeneous Markov chain. Then, the relation $\leftrightarrow$ is an equivalence relation.

Proof. The reflexivity and symmetry are clear. For the transitivity, suppose $i \leftrightarrow j$ and $j \leftrightarrow k$. Then, $\exists n, m \in \mathbb{N}$ such that $p_{ij}^{(n)} > 0$ and $p_{jk}^{(m)} > 0$. Then by [36 Chapman-Kolmogorov equation](#):

$$p_{ik}^{(n+m)} = \sum_{\ell \in I} p_{i\ell}^{(n)} p_{\ell k}^{(m)} \geq p_{ij}^{(n)} p_{jk}^{(m)} > 0 \quad (2)$$

Similarly, we have $p_{ki}^{(r+s)} > 0$ for some $r, s \in \mathbb{N}$. Therefore, $i \leftrightarrow k$. $\square$

Definition 41. Let (X_n) be a time-homogeneous Markov chain. A subset $C \subseteq I$ is called *irreducible class* if for any $i, j \in C$ we have $i \leftrightarrow j$. That is, if C is an equivalence class of $\leftrightarrow$. If all the states are in the same equivalence class, then the Markov chain is called an *irreducible chain*.

Definition 42. Let (X_n) be a time-homogeneous Markov chain and $i \in I$. We define the *period* of i as:

$$d(i) := \gcd\{n \in \mathbb{N} : p_{ii}^{(n)} > 0\}$$

with the convention that if $\{n \in \mathbb{N} : p_{ii}^{(n)} > 0\} = \emptyset$, then $d(i) = \infty$. If $d(i) = 1$ we say that i is *aperiodic*.

Proposition 43. Let (X_n) be a time-homogeneous Markov chain and $i, j \in I$. Then:

$$i \leftrightarrow j \implies d(i) = d(j)$$

Proof. Suppose $i \neq j$. We will see that if $p_{jj}^{(n)} > 0$, then $d(i) \mid n$. Since $i \leftrightarrow j$, then $\exists r, s \in \mathbb{N}$ such that $p_{ij}^{(r)} > 0$ and $p_{ji}^{(s)} > 0$. So as in [Eq. \(2\)](#), we have $p_{ii}^{(r+s)} > 0$. Thus, $d(i) \mid r + s$. Moreover, if $p_{jj}^{(n)} > 0$, then:

$$p_{ii}^{(r+n+s)} \geq p_{ij}^{(r)} p_{jj}^{(n)} p_{ji}^{(s)} > 0$$

So $d(i) \mid r + n + s$. Thus, $d(i) \mid n$ and so $d(j) \geq d(i)$ because $d(j)$ is the greatest common divisor of all such n . Repeating the argument exchanging i and j we get $d(j) = d(i)$. $\square$

Definition 44. Let (X_n) be a time-homogeneous Markov chain. If the chain is irreducible, we will denote the common period by d . If $d = 1$ we say that the chain is *aperiodic*.

Proposition 45. Let (X_n) be a time-homogeneous Markov chain. Suppose we have an irreducible chain of period $d > 1$. Then, there exist subsets $C_0, \dots, C_{d-1} \subseteq I$ such that $I = C_0 \sqcup \dots \sqcup C_{d-1}$ and such that if $j \in C_\alpha$, then:

$$p_{jk} > 0 \implies k \in C_{[\alpha+1]_d}$$

for all $k \in I$. Here $[\alpha+1]_d$ denotes $\alpha+1 \pmod d$.

Proof. Let $i \in I$ and define

$$C_\alpha := \{j \in I : \exists n \in \mathbb{N} \cup \{0\} \text{ with } p_{ij}^{(nd+\alpha)} > 0\}$$

Clearly $C_0 \cup \dots \cup C_{d-1} = I$. Let's see that $C_\alpha \cap C_\beta = \emptyset$ if $\alpha \neq \beta$. Suppose $k \in C_\alpha \cap C_\beta$. Note that since the chain is irreducible, $\exists m \in \mathbb{N} \cup \{0\}$ such that $p_{ki}^{(m)} > 0$. And so, as in [Eq. \(2\)](#) we have $p_{kk}^{(nd+\alpha+m)} > 0$ because $k \in C_\alpha$. Thus, $d \mid \alpha + m$. The same argument with β implies $d \mid \beta + m$. So $d \mid \beta - \alpha$ and $\beta = \alpha$ because $\alpha, \beta \in \{0, \dots, d-1\}$. Finally, if $j \in C_\alpha$ is such that $p_{jk} > 0$ for $k \in I$, then as in [Eq. \(2\)](#) we have $p_{ik}^{(nd+\alpha+1)} > 0$. So, if $\alpha + 1 \leq d-1$, then $k \in C_{\alpha+1}$. Otherwise, $k \in C_0 = C_{[\alpha+1]_d}$. $\square$

Stopping time and strong Markov property

Proposition 46. Let (X_n) be a time-homogeneous Markov chain, $k \in \mathbb{N}$ and $A \subseteq I^k$ and $B \subseteq I^n$. Then:

$$\begin{aligned} \mathbb{P}((X_{n+1}, \dots, X_{n+k}) \in A \mid (X_0, \dots, X_{n-1}) \in B, X_n = i) &= \\ &= \mathbb{P}((X_{n+1}, \dots, X_{n+k}) \in A \mid X_n = i) = \\ &= \mathbb{P}((X_1, \dots, X_k) \in A \mid X_0 = i) \end{aligned}$$

for all $n \geq 0$.

Proof. By Theorem 34 it suffices to prove the statement for $B = \{i_0\} \times \dots \times \{i_{n-1}\}$. Moreover, since A is countable we can suppose $A = \{j_1\} \times \dots \times \{j_k\}$. We will prove it by induction on k the homogeneous equality (the other one is even easier). The case $k = 1$ is by definition. Now suppose $k \geq 2$. Then, denoting $C := \{X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i\}$ we have:

$$\begin{aligned} \mathbb{P}(X_{n+1} = j_1, \dots, X_{n+k+1} = j_{k+1} \mid C) &= \\ &= \mathbb{P}(X_{n+k+1} = j_{k+1} \mid C, X_{n+1} = j_1, \dots, X_{n+k} = j_k) \cdot \\ &\quad \cdot \mathbb{P}(X_{n+1} = j_1, \dots, X_{n+k} = j_k \mid C) \\ &= \mathbb{P}(X_{k+1} = j_{k+1} \mid X_0 = i, X_1 = j_1, \dots, X_k = j_k) \cdot \\ &\quad \cdot \mathbb{P}(X_1 = j_1, \dots, X_k = j_k \mid X_0 = i) \\ &= \mathbb{P}(X_1 = j_1, \dots, X_{k+1} = j_{k+1} \mid X_0 = i) \end{aligned}$$

where in the second equality we have used the Markov property, the homogeneous property and induction hypothesis. $\square$

Definition 47. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let I be a finite or countable set. For each $i \in I$, let $\mathcal{F}_i$ be a sub σ -algebra of $\mathcal{A}$, that is a subset of $\mathcal{A}$ which also σ -algebra. We say that $(\mathcal{F}_i)_{i \in I}$ is *filtration* if for all $i \in I$ we have $\mathcal{F}_i \subseteq \mathcal{F}_{i+1}$. The tuple $(\Omega, \mathcal{A}, (\mathcal{F}_i)_{i \in I}, \mathbb{P})$ is called a *filtration space*.

Definition 48. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $\mathbf{X}$ be a random vector. The σ -algebra generated by $\mathbf{X}$ is:

$$\sigma(\mathbf{X}) := \{\mathbf{X}^{-1}(B) : B \in \mathcal{B}(\mathbb{R}^n)\}$$

Proposition 49. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and (X_n) be a time-homogeneous Markov chain. Then, if $\mathcal{F}_n := \sigma(X_0, \dots, X_n)$, the sequence $(\mathcal{F}_n)_{n \geq 0}$ is a filtration.

Proof. Take $F \in \mathcal{F}_n$. Then:

$$\begin{aligned} F &= \{(X_0, \dots, X_n) \in B \subseteq I^{n+1}\} = \\ &= \{(X_0, \dots, X_n, X_{n+1}) \in B \times I \subseteq I^{n+2}\} \in \mathcal{F}_{n+1} \end{aligned}$$

$\square$

Definition 50. Let $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ be a filtration space and τ a random variable on it with support $\mathbb{N} \cup \{0\}$. We say that τ is a *stopping time* if $\forall n \geq 0$ we have:

$$\{\tau = n\} \in \mathcal{F}_n$$

Remark. Intuitively, this condition means that the “decision” of whether to stop at time n must be based only on the information present at time n , not on any future information.

Lemma 51. Let $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ be a filtration space and τ be a random variable. Then:

$$\tau \text{ is a stopping time} \iff \{\tau \leq n\} \in \mathcal{F}_n$$

Proof.

$$\implies \{\tau \leq n\} = \bigcup_{m=1}^n \{\tau = m\} \in \mathcal{F}_n \text{ because } \{\tau = m\} \in \mathcal{F}_m \subseteq \mathcal{F}_n \forall m \leq n.$$

$$\impliedby \{\tau = n\} = \{\tau \leq n\} \setminus \{\tau \leq n-1\} \in \mathcal{F}_n \text{ because } \{\tau \leq n-1\} \in \mathcal{F}_{n-1} \subseteq \mathcal{F}_n.$$

$\square$

Proposition 52. Let (X_n) be a time-homogeneous Markov chain, $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ be a filtration space defined with (X_n) , $i \in I$ and τ_i be the random variable with support $\mathbb{N} \cup \{0, \infty\}$ defined by:

$$\tau_i(\omega) = \inf\{n \geq 1 : X_n(\omega) = i\} \quad (3)$$

with the convention that $\inf \emptyset = +\infty$. Then, τ_i is a stopping time.

Proof. If $n = 0$, then $\{\tau_i = 0\} = \emptyset \in \mathcal{F}_0$. If $n = 1$, then $\{\tau_i = 1\} = \{X_0 \in I, X_1 = i\} \in \mathcal{F}_1$. If $n \geq 2$, then:

$$\{\tau_i = n\} = \{X_0 \in I, X_1, \dots, X_{n-1} \in \{i\}^c, X_n = i\} \in \mathcal{F}_n$$

$\square$

Theorem 53 (Strong Markov property). Let (X_n) be a time-homogeneous Markov chain, $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ be a filtration space defined with (X_n) and τ be a stopping time. Suppose that $\mathbb{P}(\tau < \infty) > 0$. Then:

$$\begin{aligned} \mathbb{P}(X_{\tau+n+1} = j \mid X_\tau = i_0, \dots, X_{\tau+n-1} = i_{n-1}, X_{\tau+n} = i, \\ \tau < \infty) &= \mathbb{P}(X_{\tau+n+1} = j \mid X_{\tau+n} = i, \tau < \infty) = \\ &= \mathbb{P}(X_1 = j \mid X_0 = i) \end{aligned}$$

on account that $\mathbb{P}(A) > 0$, where $A := \{X_\tau = i_0, \dots, X_{\tau+n-1} = i_{n-1}, X_{\tau+n} = i, \tau < \infty\}$.

Proof.

$$\begin{aligned} \mathbb{P}(X_{\tau+n+1} = j \mid A) &= \sum_{m=0}^{\infty} \mathbb{P}(X_{\tau+n+1} = j, \tau = m \mid A) = \\ &= \sum_{\substack{m=0 \\ \mathbb{P}(\tau=m, A) > 0}}^{\infty} \mathbb{P}(X_{m+n+1} = j \mid A, \tau = m) \mathbb{P}(\tau = m \mid A) \end{aligned}$$

Now note that since $\{\tau = m\} \in \mathcal{F}_m$, we can write:

$$\{\tau = m\} = \bigcup_{j_0, \dots, j_m} \{X_0 = j_0, \dots, X_m = j_m\}$$

for some $j_0, \dots, j_m \in I$. But since $\mathbb{P}(\tau = m, A) > 0$, we have that in this last expression $j_m = i_0$ and so using Theorem 46 we get:

$$\begin{aligned} \mathbb{P}(X_{\tau+n+1} = j \mid A) &= \sum_{\substack{m=0 \\ \mathbb{P}(\tau=m, A) > 0}}^{\infty} \mathbb{P}(X_1 = j \mid X_0 = i) \cdot \\ &\quad \cdot \mathbb{P}(\tau = m \mid A) = \mathbb{P}(X_1 = j \mid X_0 = i) \end{aligned}$$

$\square$

Corollary 54 (Strong Markov property). Let (X_n) be a time-homogeneous Markov chain, $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ be a filtration space defined with (X_n) , τ be a stopping time, $k \in \mathbb{N}$ and $A \subseteq I^k$ and $B \subseteq I^n$. Suppose that $\mathbb{P}(\tau < \infty) > 0$. Then:

$$\mathbb{P}((X_{\tau+n+1}, \dots, X_{\tau+n+k}) \in A \mid (X_\tau, \dots, X_{\tau+n-1}) \in B, \\ X_{\tau+n} = i, \tau < \infty) = \mathbb{P}((X_1, \dots, X_k) \in A \mid X_0 = i)$$

for all $n \geq 0$.

Recurrence and transience

From now on we will omit saying that a stopping time τ is defined in a filtration space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$. Moreover, given a Markov chain (X_n) , we will denote by $\mathbb{P}_i(A) := \mathbb{P}(A \mid X_0 = i)$ and $\mathbb{E}_i(A) := \mathbb{E}(A \mid X_0 = i)$, for any event A .

Definition 55. Let (X_n) be a time-homogeneous Markov chain, $i, j \in I$ and consider the stopping time τ_j of Eq. (3). We define $f_{ij} := \mathbb{P}_i(\tau_j < \infty)$. We say that i is *transient* if $f_{ii} < 1$ and *recurrent* if $f_{ii} = 1$. Finally, we define N_i as:

$$N_i := |\{n \in \mathbb{N} : X_n = i\}| = \sum_{n=1}^{\infty} \mathbf{1}_{\{X_n=i\}}$$

Remark. Roughly speaking, if i is recurrent it means that the chain will return at least once to i . On the other hand, if i is transient, it means that the chain may never return to i .

Definition 56. Let (X_n) be a time-homogeneous Markov chain, $i, j \in I$ and consider the stopping time τ_j of Eq. (3). For $k \geq 2$, we define the k -th hitting time of i by:

$$\tau_i^k := \inf\{n > \tau_i^{k-1} : X_n = i\}$$

with the convention that $\tau_i^1 = \tau_i$ and $\tau_i^0 = 0$. Moreover, we define the time difference $T_i^k := \tau_i^k - \tau_i^{k-1}$.

Lemma 57. Let (X_n) be a time-homogeneous Markov chain. Then, τ_i^k is a stopping time $\forall k \in \mathbb{N}$ and moreover T_i^k are i.i.d. random variables distributed as τ_i with respect to the probability $\mathbb{P}_i$.

Proof. We need to check that $\forall m_1, \dots, m_k \in \mathbb{N}$:

$$\mathbb{P}_i(T_i^1 = m_1, \dots, T_i^k = m_k) = \mathbb{P}_i(\tau_i = m_1) \cdots \mathbb{P}_i(\tau_i = m_k)$$

We expand the left-hand side using the ???. Now we examine each term of the product, which have the form:

$$p_\ell := \mathbb{P}_i(T_i^\ell = m_\ell \mid T_i^1 = m_1, \dots, T_i^{\ell-1} = m_{\ell-1})$$

We have that:

$$p_\ell = \mathbb{P}_i(\tau_i^\ell - \tau_i^{\ell-1} = m_\ell \mid A)$$

where $A = \{X_0 = i, X_1 \neq i, \dots, X_{m_1-1} \neq i, X_{m_1} = i, X_{m_1+1} \neq i, \dots, X_{m_1+\dots+m_{\ell-1}} = i\}$. So, by the 3 Substitution principle we have:

$$p_\ell = \mathbb{P}_i(X_{m_1+\dots+m_\ell} = i, X_{m_1+\dots+m_{\ell-1}} \neq i, \dots, \\ X_{m_1+\dots+m_{\ell-1}+1} \neq i \mid A) \\ = \mathbb{P}_i(X_{m_\ell} = i, X_{m_\ell-1} \neq i, \dots, X_1 \mid X_0 = i) \\ = \mathbb{P}(\tau_i = m_\ell)$$

□

Proposition 58. Let (X_n) be a time-homogeneous Markov chain and $i \in I$. Then:

$$\mathbb{P}_i(N_i \geq k) = (f_{ii})^k$$

Proof. First suppose $f_{ii} = 0$. Then:

$$\mathbb{P}_i(N_i \geq k) \leq \mathbb{P}_i(N_i \geq 1) = \mathbb{P}_i(\tau_i < \infty) = f_{ii} = 0$$

Now assume $f_{ii} = \mathbb{P}_i(\tau_i < \infty) > 0$. We will prove the statement by induction on k . The case $k = 1$ is clear. Note that $\{\tau_i^k < \infty\} \subseteq \{\tau_i^{k-1} < \infty\}$. Thus:

$$\mathbb{P}_i(N_i \geq k) = \mathbb{P}_i(\tau_i^k < \infty) = \mathbb{P}_i(\tau_i^k < \infty, \tau_i^{k-1} < \infty) \\ = \mathbb{P}_i(\tau_i^{k-1} < \infty) \mathbb{P}_i(\tau_i^k < \infty \mid \tau_i^{k-1} < \infty) = \\ = (f_{ii})^{k-1} \mathbb{P}_i(\tau_i^k < \infty \mid \tau_i^{k-1} < \infty)$$

So it's missing to prove that $\mathbb{P}_i(\tau_i^k < \infty \mid \tau_i^{k-1} < \infty) = f_{ii}$. But:

$$\mathbb{P}_i(\tau_i^k < \infty \mid \tau_i^{k-1} < \infty) = \\ = \sum_{m=1}^{\infty} \mathbb{P}_i(\tau_i^k = m + \tau_i^{k-1} \mid \tau_i^{k-1} < \infty) \\ = \sum_{m=1}^{\infty} \mathbb{P}_i(\tau_i^k = m + \tau_i^{k-1} \mid X_{\tau_i^{k-1}} = i, \tau_i^{k-1} < \infty) \\ = \sum_{m=1}^{\infty} \mathbb{P}_i(X_{\tau_i^{k-1}+1} \neq i, \dots, X_{\tau_i^{k-1}+m-1} \neq i, \\ X_{\tau_i^{k-1}+m} = i \mid X_{\tau_i^{k-1}} = i, \tau_i^{k-1} < \infty) \\ = \sum_{m=1}^{\infty} \mathbb{P}_i(X_1 \neq i, \dots, X_{m-1} \neq i, X_m = i) = \\ = \sum_{m=1}^{\infty} \mathbb{P}_i(\tau_i = m) = \mathbb{P}_i(\tau_i < \infty) = f_{ii}$$

where we have used the 53 Strong Markov property. □

Theorem 59. Let (X_n) be a time-homogeneous Markov chain and $i \in I$. Then:

- i is recurrent $\implies \mathbb{P}_i(N_i = \infty) = 1$
- i is transient $\implies \mathbb{P}_i(N_i < \infty) = 1$

Proof. Note that $\{N_i \geq k\} \searrow \{N_i = \infty\}$, so by Theorem 58 we get:

$$\mathbb{P}_i(N_i = \infty) = \lim_{k \rightarrow \infty} \mathbb{P}_i(N_i \geq k) = (f_{ii})^k = \\ = \begin{cases} 1 & \text{if } i \text{ is recurrent} \\ 0 & \text{if } i \text{ is transient} \end{cases}$$

□

Definition 60. Let (X_n) be a time-homogeneous Markov chain, $i, j \in I$ and $n \in \mathbb{N}$. We define $f_{ij}^{(n)}$ as:

$$f_{ij}^{(n)} := \mathbb{P}_i(\tau_j = n)$$

Note that in these conditions we have:

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$$

Proposition 61. Let (X_n) be a time-homogeneous Markov chain, $i, j \in I$ and $n \in \mathbb{N}$. Then:

$$p_{ij}^{(n)} = \sum_{m=1}^n f_{ij}^{(m)} p_{jj}^{(n-m)}$$

Proof. Note that $\{X_n = j\} \subseteq \{\tau_j \leq n\} = \bigsqcup_{m=1}^n \{\tau_j = m\}$. Hence, $\{X_n = j\} = \bigsqcup_{m=1}^n [\{X_n = j\} \cap \{\tau_j = m\}]$. Thus:

$$\begin{aligned} p_{ij}^{(n)} &= \mathbb{P}_i(X_n = j) \\ &= \sum_{m=1}^n \mathbb{P}_i(X_n = j, \tau_j = m) \\ &= \sum_{m=1}^n \mathbb{P}_i(X_n = j \mid \tau_j = m) \mathbb{P}_i(\tau_j = m) \\ &= \sum_{m=1}^n \mathbb{P}_i(X_n = j \mid X_m = j, X_{m-1} \neq j, \dots, X_1 \neq j) \cdot f_{ij}^{(m)} \\ &= \sum_{m=1}^n \mathbb{P}_j(X_{n-m} = j) f_{ij}^{(m)} \\ &= \sum_{m=1}^n f_{ij}^{(m)} p_{jj}^{(n-m)} \end{aligned}$$

□

Proposition 62. Let (X_n) be a time-homogeneous Markov chain, $i, j \in I$. Then:

$$i \rightarrow j \iff f_{ij} > 0$$

Proof. First note that $f_{ij} > 0 \iff \exists m \in \mathbb{N}$ such that $f_{ij}^{(m)} > 0$. So, we have:

$\implies$ If $i \rightarrow j$, $\exists n \in \mathbb{N}$ such that $p_{ij}^{(n)} > 0$. So by **Theorem 61** we have that $\exists m \in \mathbb{N}$ such that $f_{ij}^{(m)} > 0$.

$\impliedby$ Now suppose $f_{ij}^{(m)} > 0$ for some $m \in \mathbb{N}$. Then:

$$0 < f_{ij}^{(m)} = \mathbb{P}_i(\tau_j = m) \leq \mathbb{P}_i(X_m = j) = p_{ij}^{(m)}$$

Thus, $i \rightarrow j$.

□

Definition 63. Let (X_n) be a time-homogeneous Markov chain, $i, j \in I$. We define: $S_{ij} := \sum_{n=1}^{\infty} p_{ij}^{(n)}$.

Lemma 64. Let (X_n) be a time-homogeneous Markov chain and $j \in I$ be such that $S_{jj} < \infty$. Then, $\forall i \in I$ we have:

$$S_{ij} = f_{ij}(1 + S_{jj})$$

Proof.

$$\begin{aligned} \sum_{n=1}^{\infty} p_{ij}^{(n)} &= \sum_{n=1}^{\infty} \sum_{m=1}^n f_{ij}^{(m)} p_{jj}^{(n-m)} = \sum_{m=1}^{\infty} \sum_{n=m}^{\infty} f_{ij}^{(m)} p_{jj}^{(n-m)} = \\ &= \sum_{m=1}^{\infty} f_{ij}^{(m)} (1 + S_{jj}) = f_{ij}(1 + S_{jj}) \end{aligned}$$

□

Lemma 65. Let (X_n) be a time-homogeneous Markov chain and $j \in I$. Then, $\forall i \in I$ and all $N \in \mathbb{N}$ we have:

$$\sum_{n=1}^N p_{ij}^{(n)} \leq f_{ij} \left(1 + \sum_{n=1}^N p_{jj}^{(n)} \right)$$

Sketch of the proof. Same as in **Theorem 64**.

□

Theorem 66. Let (X_n) be a time-homogeneous Markov chain and $i, j \in I$. Then:

1. i is recurrent $\iff S_{ii} = \infty$.
2. If $i \leftrightarrow j$, then i is recurrent $\iff j$ is recurrent.
3. If j is recurrent and $i \rightarrow j$, then i is recurrent.
4. If j is transient, then $S_{ij} < \infty \forall i \in I$. In particular, $\forall i \in I$, we have $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$.

Proof.

1. $\implies$ Suppose $S_{ii} < \infty$. Then, since $f_{ii} = 1$, **Theorem 64** implies $S_{ii} = 1 + S_{ii}$, which is a contradiction.
- $\impliedby$ Using **Theorem 65** we have:

$$f_{ii} \geq \frac{\sum_{n=1}^N p_{ii}^{(n)}}{1 + \sum_{n=1}^N p_{ii}^{(n)}} \xrightarrow{N \rightarrow \infty} 1$$

2. If $i \leftrightarrow j$, then $\exists r, s \geq 1$ such that $p_{ij}^{(r)}, p_{ji}^{(s)} > 0$. So by **Eq. (2)** we have that $\forall n \geq 0$:

$$p_{jj}^{(n+r+s)} \geq p_{ji}^{(s)} p_{ii}^{(n)} p_{ij}^{(r)} =: C p_{ii}^{(n)}$$

And so, $\sum p_{ii}^{(n)} = \infty \implies \sum p_{jj}^{(n)} = \infty$ by **Item 66-1**.

3. Similarly, as before, since $i \rightarrow j$, $\exists r \geq 1$ such that $p_{ij}^{(r)} > 0$. Thus, $p_{ii}^{(n+r)} \geq p_{ij}^{(r)} p_{jj}^{(n)}$. So, $\sum p_{jj}^{(n)} = \infty \implies \sum p_{ii}^{(n)} = \infty$.
4. It follows from **Item 66-1** and **Theorem 64**.

□

Definition 67. Let (X_n) be a time-homogeneous Markov chain. We say that the chain is *recurrent* if i is recurrent for all $i \in I$. We say that the chain is *transient* if i is transient for all $i \in I$.

Theorem 68 (Polya's theorem on $\mathbb{Z}$). The simple random walk on $\mathbb{Z}$ is recurrent if and only if $p = q$.

Proof. By [Item 66-1](#) we need to study the convergence of $\sum p_{ii}^{(n)} = \sum_{ii}^{(2n)}$. Note that:

$$p_{ii}^{(2n)} = \binom{2n}{n} p^n q^n$$

because we choose n steps to the right from a total of $2n$ and the rest must be steps to the left. Finally, using ?? one can check that:

$$p_{ii}^{(2n)} \sim \frac{1}{\sqrt{n}} (4pq)^n \quad (4)$$

which lead to a convergent series if and only if $p \neq q$. $\square$

Lemma 69. Let $n \in \mathbb{N}$. Then:

$$\sum_{m=0}^n \binom{n}{m}^2 = \binom{2n}{n}$$

Sketch of the proof. Equate the coefficients of x^n of the two series

$$\sum_{j=0}^{2n} \binom{2n}{j} x^j = [(1+x)^n]^2 = \left(\sum_{j=0}^n \binom{n}{j} x^j \right)^2$$

and use the fact that $\binom{n}{m} = \binom{n}{n-m}$. $\square$

Theorem 70 (Polya's theorem on $\mathbb{Z}^2$). The simple random walk on $\mathbb{Z}^2$ is recurrent if and only if $\mathbb{P}(X_1 = (1, 0)) = \mathbb{P}(X_1 = (-1, 0)) = \mathbb{P}(X_1 = (0, 1)) = \mathbb{P}(X_1 = (0, -1)) = 1/4$.

Sketch of the proof. We will proof only the implication to the left, in order to keep the proof short. Note that we have:

$$p_{ii}^{(2n)} = \sum_{m=0}^n \frac{(2n)!}{(m!)^2 (n-m)!^2} \frac{1}{4^{2n}} = \frac{1}{4^{2n}} \binom{2n}{n} \sum_{m=0}^n \binom{n}{m}^2$$

In the formula m denotes the number of steps rightwards and leftwards, and $n - m$, the number of steps upwards and downwards. Now using [Theorem 69](#) and [Eq. \(4\)](#) we have:

$$p_{ii}^{(2n)} = \frac{1}{4^{2n}} \binom{2n}{n}^2 \sim \frac{1}{n}$$

$\square$

Theorem 71 (Polya's theorem on $\mathbb{Z}^3$). The simple random walk on $\mathbb{Z}^3$ is always transient.

Corollary 72. Let (X_n) be a time-homogeneous Markov chain and $i, j \in I$. Then, if j is transient, we have $\mathbb{E}_i(N_j) < \infty$.

Proof.

$$\mathbb{E}_i(N_j) = \mathbb{E}_i \left(\sum_{n=1}^{\infty} 1_{\{X_n=j\}} \right) = \sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$$

where the last inequality follows from [Item 66-4](#). $\square$

Limit distributions

Definition 73. Let (X_n) be a time-homogeneous Markov chain and $i \in I$ be recurrent. We denote $\mu_i := \mathbb{E}_i(\tau_i)$. We say that i is *positive recurrent* if $\mu_i < \infty$ and *null recurrent* if $\mu_i = \infty$.

Theorem 74 (Ergodic theorem). Let (X_n) be a time-homogeneous Markov chain and $i \in I$ be positively recurrent. Then:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n p_{ii}^{(m)} = \frac{1}{\mu_i}$$

Proof. By hypothesis T_i^k has finite expectation and so by the ?? we have:

$$\frac{1}{k} \sum_{m=1}^k T_i^m = \frac{\tau_i^k}{k} \xrightarrow{\text{a.e.}} \mu_i$$

Let $N_i^n = \sum_{m=1}^n 1_{\{X_m=i\}} \leq n$, which counts the number of visits of the state i in the first n steps. Note that if $N_i^n = k \leq n$, $\tau_i^k \leq n < \tau_i^{k+1}$ and so:

$$\frac{k}{k+1} \frac{k+1}{\tau_i^{k+1}} = \frac{k}{\tau_i^{k+1}} < \frac{N_i^n}{n} \leq \frac{k}{\tau_i^k}$$

Hence, taking the limit $k \rightarrow \infty$ we have:

$$\lim_{n \rightarrow \infty} \frac{N_i^n}{n} = \lim_{k \rightarrow \infty} \frac{k}{\tau_i^k} \xrightarrow{\text{a.e.}} \frac{1}{\mu_i}$$

Moreover note that $\frac{N_i^n}{n} \leq 1$. Thus, by the ?? we have that:

$$\mathbb{E}_i \left(\frac{N_i^n}{n} \right) = \frac{\sum_{m=1}^n \mathbb{P}_i(X_m = i)}{n} = \frac{\sum_{m=1}^n p_{ii}^{(m)}}{n} \xrightarrow{\text{a.e.}} \frac{1}{\mu_i}$$

$\square$

Corollary 75. Let (X_n) be a time-homogeneous Markov chain and $i \in I$ be positive recurrent such that $\exists \lim_{n \rightarrow \infty} p_{ii}^{(n)}$. Then:

$$\lim_{n \rightarrow \infty} p_{ii}^{(n)} = \frac{1}{\mu_i}$$

Sketch of the proof. Recall ??.

$\square$

Theorem 76 (Ergodic theorem). Let (X_n) be a time-homogeneous Markov chain and $i \in I$ be recurrent and aperiodic. Then, the limit $\lim_{n \rightarrow \infty} p_{ii}^{(n)}$ exists and:

$$\lim_{n \rightarrow \infty} p_{ii}^{(n)} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{n=1}^{\infty} p_{ii}^{(n)} = \frac{1}{\mu_i}$$

In particular, if i is positive recurrent, then $\lim_{n \rightarrow \infty} p_{ii}^{(n)} > 0$ and if i is null recurrent, then $\lim_{n \rightarrow \infty} p_{ii}^{(n)} = 0$.

Proposition 77. Let (X_n) be a time-homogeneous Markov chain, $i \in I$ be recurrent and aperiodic and $j \in I$ be such that $i \leftrightarrow j$. Then:

1. i positive recurrent $\implies j$ positive recurrent.
2. i null recurrent $\implies j$ null recurrent.

Proof. By Theorems 43 and 66 we have that j is recurrent and aperiodic. Thus, by 76 Ergodic theorem, the limits $\lim_{n \rightarrow \infty} p_{ii}^{(n)}$ and $\lim_{n \rightarrow \infty} p_{jj}^{(n)}$ exist. Moreover, since $i \leftrightarrow j$ $\exists r, s \in \mathbb{N}$ such that $p_{ij}^{(r)}, p_{ji}^{(s)} > 0$. By Eq. (2) we have that $p_{jj}^{(n+r+s)} \geq C p_{ii}^{(n)}$. If i is positive recurrent then:

$$\lim_{n \rightarrow \infty} p_{jj}^{(n+r+s)} \geq C \lim_{n \rightarrow \infty} p_{ii}^{(n)} > 0$$

If i is null and j was positive, then i would be positive by the previous argument, which is a contradiction. $\square$

Theorem 78. Let (X_n) be a time-homogeneous Markov chain and $i \in I$ be recurrent and periodic of period d . Then:

$$\lim_{n \rightarrow \infty} p_{ii}^{(nd)} = \frac{d}{\mu_i}$$

Proof. $(Y_n) := (X_{nd})$ is a time-homogeneous Markov chain and $i \in I$ is recurrent and aperiodic. Thus, by 76 Ergodic theorem we have that $\lim_{n \rightarrow \infty} p_{ii}^{(nd)} = \frac{1}{\mathbb{E}_i(\tau_i^Y)}$. But:

$$\tau_i^Y = \inf\{n \geq 1 : Y_n = i\} = \frac{1}{d} \inf\{n \geq 1 : X_n = i\} = \frac{\tau_i}{d}$$

$\square$

Theorem 79. Let (X_n) be a time-homogeneous irreducible and aperiodic Markov chain. Then, we have exactly one of the following results:

1. All the states are transient and $\forall i, j \in I$:

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \lim_{n \rightarrow \infty} \pi_j^{(n)} = 0$$

Moreover $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$.

2. All the states are null recurrent and $\forall i, j \in I$:

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \lim_{n \rightarrow \infty} \pi_j^{(n)} = 0$$

Moreover $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$.

3. All the states are positive recurrent and $\forall i, j \in I$:

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \lim_{n \rightarrow \infty} \pi_j^{(n)} = \frac{1}{\mu_j}$$

Proof. It can be seen that $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \lim_{n \rightarrow \infty} p_{jj}^{(n)} \forall i, j \in I$.

We will prove that $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \lim_{n \rightarrow \infty} \pi_j^{(n)} \forall i, j \in I$. We have that:

$$\begin{aligned} \lim_{n \rightarrow \infty} \pi_j^{(n)} &= \lim_{n \rightarrow \infty} \mathbb{P}(X_n = j) = \lim_{n \rightarrow \infty} \sum_{i \in I} p_{ij}^{(n)} \pi_i = \\ &= \sum_{i \in I} \lim_{n \rightarrow \infty} p_{ij}^{(n)} \pi_i = \sum_{i \in I} \frac{\pi_i}{\mu_j} = \frac{1}{\mu_j} \end{aligned}$$

where we have used the dominated convergence theorem for series. $\square$

Corollary 80. Let (X_n) be a time-homogeneous irreducible and aperiodic Markov chain such that I is finite. Then, all the states are positive recurrent.

Sketch of the proof. Note that we must have $\sum_{j \in I} \pi_j^{(n)} = 1 \forall n \in \mathbb{N}$. $\square$

Definition 81. Let (X_n) be a time-homogeneous Markov chain. A vector $\nu = (\nu_i)_{i \in I}$ is called a *stationary distribution* if:

$$\nu \geq 0 \quad \sum_{i \in I} \nu_i = 1 \quad \nu \mathbf{P} = \nu$$

Remark. In general, we cannot guarantee the existence or uniqueness of stationary distributions.

Lemma 82. Let (X_n) be a time-homogeneous Markov chain, ν be a stationary distribution and suppose $\pi^{(0)} = \nu$. Then, $\pi^{(n)} = \nu \forall n \in \mathbb{N}$.

Proof. $\pi^{(n)} = \nu \mathbf{P}^n = \nu \mathbf{P}^{n-1} = \dots = \nu$ $\square$

Theorem 83. Let (X_n) be a time-homogeneous irreducible and aperiodic Markov chain. Then, (X_n) is positive recurrent if and only if it admits a stationary distribution. Moreover, this distribution is unique, and it is given by $\nu_i = \frac{1}{\mu_i}$.

Proof. We will only proof the case when I is finite. By Theorem 80 we only need to prove the impication to the right. Since $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \frac{1}{\mu_j} \forall i, j \in I$ we have that $\nu = (1/\mu_i)_{i \in I} \geq 0$ satisfies:

$$\begin{aligned} \sum_{j \in I} \nu_j &= \sum_{j \in I} \lim_{n \rightarrow \infty} p_{ij}^{(n)} = \lim_{n \rightarrow \infty} \sum_{j \in I} p_{ij}^{(n)} = 1 \\ \sum_{i \in I} \nu_i p_{ij} &= \sum_{i \in I} \lim_{n \rightarrow \infty} p_{ki}^{(n)} p_{ij} = \lim_{n \rightarrow \infty} p_{kj}^{(n+1)} = \nu_j \end{aligned}$$

where we have used 36 Chapman-Kolmogorov equation. Hence, ν is a stationary distribution. Now, for the uniqueness, suppose ν is an arbitrary stationary distribution. Then, $\nu_j = \sum_{i \in I} \nu_i p_{ij}^{(n)} \forall n \in \mathbb{N}$. Thus, taking $n \rightarrow \infty$ we get that $\nu_j = \frac{1}{\mu_j} \forall j \in I$. $\square$

3. | Continuous-time Markov chains

Introduction

Definition 84. Let $(X_t)_{t \geq 0}$ be a stochastic process. We say that $(X_t)_{t \geq 0}$ is a *continuous-time Markov chain* with state space I (finite or countable) if $\forall n \in \mathbb{N}$ and all $0 \leq t_1 < \dots < t_n < t_{n+1}$ and all $i_1, \dots, i_{n-1}, i, j \in I$ we have that:

$$\begin{aligned} \mathbb{P}(X_{t_{n+1}} = j \mid X_{t_n} = i, X_{t_{n-1}} = i_{n-1}, \dots, X_{t_1} = i_1) &= \\ &= \mathbb{P}(X_{t_{n+1}} = j \mid X_{t_n} = i) \end{aligned}$$

The chain is called homogeneous if $\mathbb{P}(X_{t_{n+1}} = j \mid X_{t_n} = i)$ does only depend on the difference $t_{n+1} - t_n$. That is, if $\forall s \leq t$ we have:

$$\mathbb{P}(X_{t+s} = j \mid X_s = i) = \mathbb{P}(X_t = j \mid X_0 = i)$$

In order to simplify the lecture we will write CTHMC for continuous-time homogeneous Markov chains.

Definition 85. Let $(X_t)_{t \geq 0}$ be a CTHMC. We define the *transition probabilities* as:

$$p_{ij}(t) = \mathbb{P}(X_t = j \mid X_0 = i)$$

The transition matrix is $\mathbf{P}(t) = (p_{ij}(t))_{i,j \in I}$.

Proposition 86 (Chapman-Kolmogorov equation).

Let $(X_t)_{t \geq 0}$ be a CTHMC. Then, $\mathbf{P}(t)$ is a stochastic matrix $\forall t \geq 0$ and $\forall s, t \geq 0$:

$$\mathbf{P}(t+s) = \mathbf{P}(t)\mathbf{P}(s)$$

Proof.

$$\begin{aligned} p_{ij}(t+s) &= \sum_{k \in I} \mathbb{P}(X_{t+s} = j, X_s = k \mid X_0 = i) \\ &= \sum_{k \in I} \mathbb{P}(X_{t+s} = j \mid X_s = k, X_0 = i) \cdot \mathbb{P}(X_s = k \mid X_0 = i) \\ &= \sum_{k \in I} p_{ik}(t)p_{kj}(s) \end{aligned}$$

□

Proposition 87. Let $(X_t)_{t \geq 0}$ be a CTHMC. Then, for all $0 \leq t_1 < \dots < t_n$ and all $i_1, \dots, i_n \in I$ we have that:

$$\begin{aligned} \mathbb{P}(X_{t_n} = i_n, X_{t_{n-1}} = i_{n-1}, \dots, X_{t_1} = i_1) &= \\ &= p_{i_1}(t_1)p_{i_1 i_2}(t_2 - t_1) \dots p_{i_{n-1} i_n}(t_n - t_{n-1}) \end{aligned}$$

where $p_i(t) := \mathbb{P}(X_t = i)$.

Poisson process

Definition 88. Let $\lambda > 0$. A stochastic process $(N_t)_{t \geq 0}$ is called a *Poisson process* with parameter λ if:

1. $N_0 = 0$.
2. N_t has *independent increments*, that is $N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$ are independent random variables for all $0 \leq t_1 < \dots < t_n$ and all $n \in \mathbb{N}$.
3. N_t has *stationary increments*, that is $N_t - N_s \sim \text{Pois}(\lambda(t-s)) \forall 0 \leq s \leq t$.
4. For all $\omega \in \Omega$ the functions (called *trajectories*)

$$\begin{aligned} N(\omega) : [0, \infty) &\longrightarrow \mathbb{N} \cup \{0\} \\ t &\longmapsto N_t(\omega) \end{aligned}$$

are right-continuous (càd).

Proposition 89. Let $(X_t)_{t \geq 0}$ be a stochastic process with independent increments. Then, $(X_t)_{t \geq 0}$ is a continuous-time Markov chain.

Proof.

$$\begin{aligned} \mathbb{P}(X_{t_{n+1}} = j \mid X_{t_n} = i, X_{t_{n-1}} = i_{n-1}, \dots, X_{t_1} = i_1) &= \\ &= \mathbb{P}(X_{t_{n+1}} - X_{t_n} = j - i \mid X_{t_n} - X_{t_{n-1}} = i - i_{n-1}, \dots, X_{t_2} - X_{t_1} = i_2 - i_1, X_{t_1} = i_1) \\ &= \mathbb{P}(X_{t_{n+1}} - X_{t_n} = j - i) \\ &= \mathbb{P}(X_{t_{n+1}} = j \mid X_{t_n} = i) \end{aligned}$$

□

Corollary 90. Let $(N_t)_{t \geq 0}$ be a Poisson process with parameter λ . Then, $(N_t)_{t \geq 0}$ is a CTHMC with transition probabilities:

$$p_{ij}(t) = \mathbb{P}(N_t = j \mid N_0 = i) = \frac{(\lambda t)^j}{j!} e^{-\lambda t}$$

Proposition 91. Let $(N_t^1)_{t \geq 0}$ and $(N_t^2)_{t \geq 0}$ be two independent Poisson processes with parameters λ_1 and λ_2 respectively. Then, $(N_t^1 + N_t^2)_{t \geq 0}$ is a Poisson process with parameter $\lambda_1 + \lambda_2$.

Proof. Let $N_t := (N_t^1 + N_t^2)$. We only check the independent increment, the other properties are easier. We need to check that for all $0 \leq t_1 < \dots < t_n$ and all $n \in \mathbb{N}$ the random variables $X_{t_1} := N_{t_1}, X_{t_2} := N_{t_2} - N_{t_1}, \dots, X_{t_n} := N_{t_n} - N_{t_{n-1}}$ are independent. We have that

$$X_{t_\ell} = N_{t_\ell}^1 - N_{t_{\ell-1}}^1 + N_{t_\ell}^2 - N_{t_{\ell-1}}^2 =: Y_\ell^1 + Y_\ell^2$$

By hypothesis the variables Y_ℓ^1 and Y_ℓ^2 are independent. Moreover, since N_t^i are Poisson processes, we have that $\{(Y_k^i)\}_{k=1, \dots, n}$ pairwise independent, for $i = 1, 2$. Now using the characterization of independence with the characteristic function, we have:

$$\begin{aligned} \varphi_{X_{t_1}, \dots, X_{t_n}}(u_1, \dots, u_n) &= \mathbb{E} \left(e^{i \sum_{j=1}^n u_j X_j} \right) \\ &= \mathbb{E} \left(e^{i \sum_{j=1}^n u_j Y_j^1} e^{i \sum_{j=1}^n u_j Y_j^2} \right) \\ &= \prod_{j=1}^n \mathbb{E} \left(e^{i u_j Y_j^1} \right) \mathbb{E} \left(e^{i u_j Y_j^2} \right) \\ &= \prod_{j=1}^n \varphi_{X_j}(u_j) \end{aligned}$$

□

Lemma 92. Let $(N_t)_{t \geq 0}$ be a Poisson process with parameter λ . Then:

$$\mathbb{P}(N_h \geq 2) = o(h)$$

Proof. $\mathbb{P}(N_h \geq 2) = 1 - e^{-\lambda h} - \lambda h e^{-\lambda h} = o(h)$ □

Proposition 93. Let $(N_t)_{t \geq 0}$ be a Poisson process with parameter λ . Then, the trajectories are almost surely non-decreasing and have jumps of size at most 1.

Proof. We need to see that the event $\{N_s \leq N_t : \forall t, s \in \mathbb{R}, 0 \leq s < t\}$ has probability 1. We have:

$$\bigcap_{0 \leq s < t} \{N_s \leq N_t\} = \bigcap_{\substack{0 \leq s < t \\ s, t \in \mathbb{Q}}} \{N_s \leq N_t\}$$

because the trajectories are càd. Finally, since $\mathbb{P}(N_s \leq N_t) = \mathbb{P}(N_t - N_s \geq 0) = 1$, the intersection has probability 1. Now, let:

$$A := \{\omega \in \Omega : N(\omega) \text{ has jumps of size } \geq 2\}$$

$$A_R := \{\omega \in \Omega : N(\omega) \text{ has jumps of size } \geq 2 \text{ in } [0, R]\}$$

$$B_R^n := \{\exists k \in \{1, \dots, n\} : N_{\frac{kR}{n}} - N_{\frac{(k-1)R}{n}} \geq 2\}$$

□

Note that $A = \bigcup_{R=1}^{\infty} A_R$ and $A_R \subseteq B_R^n \forall n \geq 1$ because the trajectories are càd. Thus, $\forall R > 0$:

$$\begin{aligned} \mathbb{P}(A_R) &\leq \mathbb{P}(B_R^n) = \mathbb{P}\left(\bigcup_{k=1}^n \left\{N_{\frac{kR}{n}} - N_{\frac{(k-1)R}{n}} \geq 2\right\}\right) \leq \\ &\leq \sum_{k=1}^n \mathbb{P}\left(N_{\frac{kR}{n}} - N_{\frac{(k-1)R}{n}} \geq 2\right) = n\mathbb{P}(N_{\frac{R}{n}} \geq 2) = \\ &= n\mathbb{O}\left(\frac{R}{n}\right) \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

Hence, $\mathbb{P}(A) = 0$. $\square$

Definition 94. Let $(N_t)_{t \geq 0}$ be a Poisson process with parameter λ . We define the *holding times* as:

$$T_k := \inf\{t > T_{k-1} : N_t = k\}$$

with $T_0 := 0$. The *inter-arrival times* are:

$$S_k := T_k - T_{k-1}$$

Lemma 95. Let $(N_t)_{t \geq 0}$ be a Poisson process with parameter λ . Then, $\mathbb{P}(T_k < \infty) = 1 \forall k \in \mathbb{N} \cup \{0\}$.

Proof. Since the trajectories are càd:

$$\begin{aligned} \mathbb{P}(T_k = \infty) &= \mathbb{P}(\forall t \in \mathbb{R} : N_t \leq k-1) \leq \mathbb{P}(N_1 = 0, \\ N_2 - N_1 = 0, \dots, N_n - N_{n-1} = 0) &= e^{-\lambda n} \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

because the inequality is true for all $n \in \mathbb{N}$. $\square$

Theorem 96. Let $(N_t)_{t \geq 0}$ be a Poisson process with parameter λ . Then, the inter-arrival times (S_k) are i.i.d. random variables distributed as $\text{Exp}(\lambda)$.

Proof. Let $\mathbf{T} := (T_1, \dots, T_n)$. Recall that:

$$f_{\mathbf{T}}(\mathbf{t}) = \frac{\partial^n F_{\mathbf{X}}}{\partial x_1 \dots \partial x_n}(\mathbf{t}) = \lim_{\mathbf{h} \rightarrow \mathbf{0}^+} \frac{\mathbb{P}(\mathbf{T} \in \prod_{k=1}^n (t_k, t_k + h_k])}{h_1 \dots h_n}$$

Assume that $t_k + h_k < t_{k+1} \forall k \in \{1, \dots, n-1\}$. Then:

$$\begin{aligned} \mathbb{P}\left(\mathbf{T} \in \prod_{k=1}^n (t_k, t_k + h_k]\right) &= \mathbb{P}(N_{t_1} = 0, N_{t_1+h_1} - N_{t_1} = 1, \\ N_{t_2} - N_{t_1+h_1} = 0, \dots, N_{t_n} - N_{t_{n-1}+h_{n-1}} = 0, \\ N_{t_n+h_n} - N_{t_n} > 1) &= \lambda^{n-1} h_1 \dots h_{n-1} e^{-\lambda t_n} (1 - e^{-\lambda h_n}) \end{aligned}$$

Hence:

$$f_{\mathbf{T}}(\mathbf{t}) = \lambda^n e^{-\lambda t_n}$$

Now consider

$$\begin{aligned} \mathbf{g} : \{0 < t_1 < \dots < t_n\} &\longrightarrow (0, \infty)^n \\ (t_1, \dots, t_n) &\longmapsto (t_1, t_2 - t_1, \dots, t_n - t_{n-1}) \end{aligned}$$

which is a diffeomorphism such that $\det \mathbf{D}\mathbf{g}(\mathbf{t}) = 1$. The density of $\mathbf{S} := \mathbf{g}(\mathbf{T})$ is thus

$$f_{\mathbf{S}}(\mathbf{s}) = f_{\mathbf{T}}(\mathbf{g}^{-1}(\mathbf{s})) \mathbf{1}_{(0, \infty)^n}(\mathbf{s}) = \prod_{k=1}^n \lambda e^{-\lambda s_k} \mathbf{1}_{(0, \infty)}(s_k)$$

by ???. And this last expression is the joint pdf of n i.i.d. $\text{Exp}(\lambda)$ variables. $\square$

Theorem 97. Let (S_k) be a sequence of i.i.d. random variables distributed as $\text{Exp}(\lambda)$. Consider the sequence (T_n) defined as $T_0 := 0$ and $T_n = \sum_{k=1}^n S_k$. Let $N_t := \sup\{n \geq 1 : T_n \leq t\}$. Then, $(N_t)_{t \geq 0}$ is a Poisson process with parameter λ . In this case, we can also express N_t as:

$$N_t = \sum_{n=1}^{\infty} n \mathbf{1}_{\{T_n \leq t < T_{n+1}\}}$$

Sketch of the proof. We'll see only that $N_t \sim \text{Pois}(\lambda t)$. Let $k \in \mathbb{N} \cup \{0\}$. Then, $\mathbb{P}(N_t = k) = \mathbb{P}(T_k \leq t < T_{k+1})$. Since $T_k = \sum_{i=1}^k S_i$, we have that T_k and S_{k+1} are independent. Hence, $f_{T_k, S_{k+1}}(x, y) = f_{T_k}(x) f_{S_{k+1}}(y)$ and since $S_{k+1} \sim \text{Exp}(\lambda)$ and $T_k \sim \Gamma(k, \lambda)$ (because is a sum of exponentials), we have that:

$$f_{T_k, S_{k+1}}(x, y) = \frac{\lambda^k}{\Gamma(k)} t^{k-1} e^{-\lambda x} \lambda e^{-\lambda y} = \frac{\lambda^{k+1} t^{k-1}}{(k-1)!} e^{-\lambda(x+y)}$$

Finally:

$$\mathbb{P}(N_t = k) = \iint_A f_{T_k, S_{k+1}}(x, y) dx dy = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

where $A := \{(x, y) \in \mathbb{R}_{\geq 0}^2 : x \leq t < x + y\}$. $\square$

Kolmogorov's differential equations

From here on, we'll assume that the transition matrices $\mathbf{P}(t)$ satisfy that $\lim_{h \rightarrow 0} \mathbf{P}(h) = \mathbf{I}$. That is, we have right continuity at 0. This is equivalent to say that $\lim_{h \rightarrow 0} p_{ij}(h) = \delta_{ij}$.

Lemma 98. Let $(X_t)_{t \geq 0}$ be a CTHMC with transition matrix $\mathbf{P}(t)$. Then, $(p_{ij}(t))$ are continuous functions for all $i, j \in I$.

Proof. The result follows from the inequality:

$$|p_{ij}(t+h) - p_{ij}(t)| \leq 1 - p_{ii}(|h|)$$

and the right-continuity at 0. Let's prove the inequality. Suppose that $h > 0$. Then:

$$\begin{aligned} |p_{ij}(t+h) - p_{ij}(t)| &= \left| \sum_{k \in I} p_{ik}(h) p_{kj}(t) - p_{ij}(t) \right| \\ &= \left| \sum_{\substack{k \in I \\ k \neq i}} p_{ik}(h) p_{kj}(t) - p_{ij}(t) [1 - p_{ii}(h)] \right| \\ &= \left| \sum_{\substack{k \in I \\ k \neq i}} p_{ik}(h) p_{kj}(t) + p_{ij}(t) \sum_{\substack{k \in I \\ k \neq i}} p_{ik}(h) \right| \\ &\leq \sum_{\substack{k \in I \\ k \neq i}} p_{ik}(h) |p_{kj}(t) - p_{ij}(t)| \\ &\leq 1 - p_{ii}(h) \end{aligned}$$

The case $h < 0$ follows from considering $t' = t + h$. $\square$

Theorem 99. Let $\mathbf{P}(t)$ be a stochastic matrix such that:

$$1. \mathbf{P}(0) = \mathbf{I}.$$

2. $\mathbf{P}(t+s) = \mathbf{P}(t)\mathbf{P}(s)$ for all $t, s \geq 0$.

3. $\lim_{h \rightarrow 0} \mathbf{P}(h) = \mathbf{I}$.

Then, for any $i, j \in I$, the following limits exist:

$$\begin{cases} q_{ij} := \lim_{h \rightarrow 0} \frac{p_{ij}(h)}{h} \in [0, \infty) & \text{if } i \neq j \\ q_i := \lim_{h \rightarrow 0} \frac{1 - p_{ii}(h)}{h} \in [0, \infty] & \text{if } i = j \end{cases}$$

Note that if the limits are finite we have $q_{ij} = p'_{ij}(0)$ and $q_i = -p'_{ii}(0)$.

Definition 100. Let $(X_t)_{t \geq 0}$ be a CTHMC with transition matrix $\mathbf{P}(t)$. Then, the *infinitesimal generator* of $(X_t)_{t \geq 0}$ is the matrix $\mathbf{Q}$ defined as $\mathbf{Q} := (q_{ij})_{i,j \in I}$, where $q_{ii} := -q_i$. We define the *infinitesimal transition scheme* as:

$$\begin{cases} p_{ii}(h) = 1 - q_i h + o(h) & \text{if } i = j \\ p_{ij}(h) = q_{ij} h + o(h) & \text{if } i \neq j \end{cases}$$

Theorem 101. Let $(X_t)_{t \geq 0}$ be a CTHMC with infinitesimal generator $\mathbf{Q}$ and assume that I is finite. Then, $\mathbf{P}' = \mathbf{Q}\mathbf{P}$ and $\mathbf{P}' = \mathbf{P}\mathbf{Q}$. The first equation is called the *Kolmogorov's backward equation* and the second one the *Kolmogorov's forward equation*.

Proof. Note that since I is finite, $q_i < \infty \forall i \in I$. Indeed:

$$q_i = \lim_{h \rightarrow 0} \frac{1 - p_{ii}(h)}{h} = \sum_{\substack{k \in I \\ k \neq i}} \lim_{h \rightarrow 0} \frac{p_{ik}(h)}{h} = \sum_{\substack{k \in I \\ k \neq i}} q_{ik} < \infty$$

Now, let $t \geq 0$, $h > 0$ and $i, j \in I$. Then, using 86 Chapman-Kolmogorov equation

$$\begin{aligned} p_{ij}(t+h) - p_{ij}(t) &= \sum_{\substack{k \in I \\ k \neq i}} p_{ik}(h) p_{kj}(t) + p_{ii}(h) p_{ij}(t) - p_{ij}(t) \\ &= \sum_{\substack{k \in I \\ k \neq i}} (q_{ik} h + o(h)) p_{kj}(t) + (1 + q_{ii} h + o(h)) p_{ij}(t) - p_{ij}(t) \\ &= \sum_{k \in I} q_{ik} h p_{kj}(t) + o(h) \end{aligned}$$

Dividing by h and taking limits we get the result with the right derivative. Now take $t > 0$ and $h < 0$. Then, similarly:

$$p_{ij}(t) - p_{ij}(t+h) = - \sum_{k \in I} q_{ik} h p_{kj}(t+h) + o(h)$$

Using the continuity of the p_{ij} 's we get the result with the left derivative. The other equation follows analogously by exchanging the roles of h and t in the Chapman-Kolmogorov equations. $\square$

Theorem 102. Let $(X_t)_{t \geq 0}$ be a CTHMC with infinitesimal generator $\mathbf{Q}$. Assume that $q_i < \infty$ and $q_i = \sum_{\substack{k \in I \\ k \neq i}} q_{ik}$ for all $i \in I$. Then, $\mathbf{P}' = \mathbf{Q}\mathbf{P}$.

Theorem 103. Let $(X_t)_{t \geq 0}$ be a CTHMC with infinitesimal generator $\mathbf{Q}$. Assume that $q_i < \infty$ and $q_i = \sum_{\substack{k \in I \\ k \neq i}} q_{ik}$ for all $i \in I$ and that $\sum_{k \in I} p_{ik}(t) q_k < \infty$ for all $i \in I$ and $t \geq 0$. Then, $\mathbf{P}' = \mathbf{P}\mathbf{Q}$.

Remark. Note that in this latter theorem if $\sup_{k \in I} q_k < \infty$, then $\sum_{k \in I} p_{ik}(t) q_k < \infty$ for all $i \in I$ and $t \geq 0$.

Jump processes

Definition 104. Let $(X_t)_{t \geq 0}$ be a stochastic process with state space E (not necessarily countable) and $\tilde{\Omega}$ be such that $\mathbb{P}(\tilde{\Omega}) = 1$. We say that $(X_t)_{t \geq 0}$ is a *jump process* if $\forall \omega \in \tilde{\Omega}$ and $\forall t \geq 0$, $\exists \varepsilon > 0$ such that $X_t(\omega) = X_s(\omega)$ for all $s \in [t, t + \varepsilon)$.

Definition 105. Let $(X_t)_{t \geq 0}$ be a jump process. We say that the jump process is *regular* if $\forall C > 0$, the number of jumps of $X(\omega)$ in $[0, C]$ is finite for all $\omega \in \Omega$.

Theorem 106. Let $(X_t)_{t \geq 0}$ be a CTHMC and a regular jump process. Then, $\forall i \in I$, $q_i < \infty$ and $q_i = \sum_{\substack{k \in I \\ k \neq i}} q_{ik}$.

Definition 107. Let $(X_t)_{t \geq 0}$ be a CTHMC. Then, $(X_t)_{t \geq 0}$ is said to be *stable* if $\forall i \in I$, $q_i < \infty$, and is said to be *conservative* if $\forall i \in I$, $q_i = \sum_{\substack{k \in I \\ k \neq i}} q_{ik}$.

Theorem 108. Let $(X_t)_{t \geq 0}$ be a CTHMC and a regular jump process. Then, the two Kolmogorov ODEs are satisfied.

Limit and stationary distributions

Definition 109. Let $(X_t)_{t \geq 0}$ be a CTHMC. We say that $\bar{\mathbf{p}}$ is a *stationary distribution* for $(X_t)_{t \geq 0}$ if $\bar{p}_i \geq 0 \forall i \in I$, $\sum_{i \in I} \bar{p}_i = 1$ and $\bar{\mathbf{p}}\mathbf{P}(t) = \bar{\mathbf{p}} \forall t \geq 0$.

Lemma 110. Let $(X_t)_{t \geq 0}$ be a CTHMC and a regular jump process. Then:

$$p_j'(t) = \sum_{k \in I} q_{kj} p_k(t)$$

In particular if we have a stationary distribution $\bar{\mathbf{p}}$, then $\bar{\mathbf{p}}\mathbf{Q} = 0$, or equivalently:

$$\sum_{\substack{k \in I \\ k \neq j}} \bar{p}_k q_{kj} = \bar{p}_j q_j$$

This equation is called *balance equation*.

Proof.

$$p_j(t) = \mathbb{P}(X_t = j) = \sum_{i \in I} p_i(0) p_{ij}(t)$$

A result allows us to differentiate term by term and rearrange the following series (because $(X_t)_{t \geq 0}$ is a regular jump process):

$$\begin{aligned} p_j'(t) &= \sum_{i \in I} p_i(0) p'_{ij}(t) = \sum_{i \in I} p_i(0) \sum_{k \in I} p_{ik}(t) q_{kj} \\ &= \sum_{k \in I} q_{kj} \sum_{i \in I} p_i(0) p_{ik}(t) = \sum_{k \in I} q_{kj} p_k(t) \end{aligned}$$

If we have a stationary distribution, then $\bar{p}_j'(t) = 0$. $\square$

Remark. In the CTHMC there is no periodic behaviour as in the discrete case. Indeed, given $t > 0$ and $\delta > 0$ small enough, $\exists n \in \mathbb{N} \cup \{0\}$ such that $t = n\delta + h$ with $h \in [0, \delta)$. And so:

$$p_{ii}(t) \geq p_{ii}(n\delta)p_{ii}(h) \geq \dots \geq (p_{ii}(\delta))^n p_{ii}(h) > 0$$

where the last inequality is due to the fact that $\lim_{h \rightarrow 0} p_{ii}(h) = 1$.

Definition 111. Let $(X_t)_{t \geq 0}$ be a CTHMC. We say that the chain is *irreducible* if $\forall i, j \in I$, $\exists t_1, t_2 > 0$ such that $p_{ij}(t_1) > 0$ and $p_{ji}(t_2) > 0$.

Theorem 112. Let X be an irreducible CTHMC and a regular jump process. Then, we have exactly one of the following:

1. The balance equation has a unique solution $\bar{\mathbf{p}}$ (which must be the stationary distribution) and $\lim_{t \rightarrow \infty} p_{ij}(t) = \bar{p}_j \forall i, j \in I$. In that case, $\bar{\mathbf{p}}$ is called a *limit distribution*.
2. The balance equation has no solution and in that case $\lim_{t \rightarrow \infty} p_{ij}(t) = 0 \forall i, j \in I$.

Remark. Note that if I is finite, we are always in the first case since we always need to have $\sum_{j \in I} p_{ij}(t) = 1$.

Birth and death processes

Definition 113. Let $I = \mathbb{N} \cup \{0\}$. A *birth and death process* is a CTHMC and a regular jump process with the following infinitesimal transition scheme:

$$\begin{cases} p_{i,i+1}(h) = \lambda_i h + o(h) & i \geq 0 \\ p_{i,i-1}(h) = \mu_i h + o(h) & i \geq 1 \\ p_{ii}(h) = 1 - \lambda_i h + o(h) & i = 0 \\ p_{ii}(h) = 1 - (\lambda_i + \mu_i)h + o(h) & i \geq 1 \\ p_{ij} = o(h) & \text{otherwise} \end{cases}$$

This model describes a population of individuals, each of whom having $\lambda_i h + o(h)$ probability of giving birth to a new individual in the time interval $[t, t+h)$ and $\mu_i h + o(h)$ probability of dying in the same time interval. The probability of having more than one birth or death in that interval is $o(h)$. In this case the infinitesimal generator is:

$$\mathbf{Q} = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \dots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \dots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

If $\lambda_i = 0 \forall i \in I$, then we said that the process is a *pure death process*. If $\mu_i = 0 \forall i \in I$, then we said that the process is a *pure birth process*.

Proposition 114. The Poisson process is a birth and death process with $\lambda_i = \lambda$ and $\mu_i = 0 \forall i \in I$.

Theorem 115. Consider a birth and death process. Then, a limit distribution $\bar{\mathbf{p}}$ exists if $\lambda_i > 0$ and $\mu_i > 0 \forall i \in I$ and

$$\sum_{i \in I} \frac{\lambda_0 \lambda_1 \dots \lambda_{i-1}}{\mu_1 \mu_2 \dots \mu_i} < \infty$$

This distribution is given by:

$$\bar{p}_i = \frac{\lambda_0 \lambda_1 \dots \lambda_{i-1}}{\mu_1 \mu_2 \dots \mu_i} \bar{p}_0 \quad i \geq 1 \quad (5)$$

with $\bar{p}_0 = \left(1 + \sum_{i \in I} \frac{\lambda_0 \lambda_1 \dots \lambda_{i-1}}{\mu_1 \mu_2 \dots \mu_i}\right)^{-1}$.

Proof. First let's prove by induction that if $\bar{\mathbf{p}}$ is a stationary distribution, then the components are those of Eq. (5). Indeed, for $i = 1$, using the balance equation we have $\lambda_0 \bar{p}_0 = \mu_1 \bar{p}_1$. Now, suppose that the hypothesis holds $\forall i \leq j$. Then, we have:

$$\begin{aligned} (\lambda_j + \mu_j) \bar{p}_j &= \lambda_{j-1} \bar{p}_{j-1} + \mu_{j+1} \bar{p}_{j+1} \\ (\lambda_j + \mu_j) \frac{\lambda_0 \dots \lambda_{j-1}}{\mu_1 \dots \mu_j} \bar{p}_0 &= \lambda_{j-1} \frac{\lambda_0 \dots \lambda_{j-2}}{\mu_1 \dots \mu_{j-1}} \bar{p}_0 + \mu_{j+1} \bar{p}_{j+1} \\ \frac{\lambda_0 \dots \lambda_{j-1} \lambda_j}{\mu_1 \dots \mu_j} \bar{p}_0 &= \mu_{j+1} \bar{p}_{j+1} \end{aligned}$$

The first argument is determined from the condition $1 = \sum_{i \in I} \bar{p}_i = \bar{p}_0 + \bar{p}_0 \sum_{i \in I} \frac{\lambda_0 \lambda_1 \dots \lambda_{i-1}}{\mu_1 \mu_2 \dots \mu_i}$. Now if we see that for $\lambda_i > 0$ and $\mu_i > 0 \forall i \in I$ the chain is irreducible, then the theorem will be proved by Theorem 112. But this is clear because, for example if $i < j$ we have:

$$\begin{aligned} p_{ij}((j-i)h) &\geq p_{i,i+1}(h) p_{i+1,i+2}(h) \dots p_{j-1,j}(h) \\ &= \lambda_i \lambda_{i+1} \dots \lambda_{j-1} h^{j-i} + o(h^{j-i}) > 0 \end{aligned}$$

if $\lambda_i > 0 \forall i \in I$ and for some h small enough. The case $i > j$ is analogous. $\square$

Theorem 116 (Reuter criterion). Consider an infinitesimal generator for a birth and death process. Then, there is a CTHMC of regular jumps with this infinitesimal generator if and only if:

$$\sum_{n=1}^{\infty} \left[\frac{1}{\lambda_n} + \frac{\mu_n}{\lambda_n \lambda_{n-1}} + \dots + \frac{\mu_n \dots \mu_1}{\lambda_n \dots \lambda_0} \right] = \infty$$

4. | Brownian motion

Gaussian processes

Proposition 117. Let $\mathbf{x} \in \mathbb{R}^n$ be a random vector. Then, $\mathbf{x}$ is a *Gaussian vector*, that is it distributes as an n -dimensional normal, if and only if there exists $k \in \mathbb{N}$, $\mathbf{A} \in \mathcal{M}_{n \times k}(\mathbb{R})$, $\mathbf{z} \in \mathbb{R}^k$ with i.i.d. components distributed as $N(0, 1)$, and $\boldsymbol{\mu} \in \mathbb{R}^n$ such that:

$$\mathbf{x} = \mathbf{A}\mathbf{z} + \boldsymbol{\mu}$$

Definition 118. A stochastic process $(X_t)_{t \geq 0}$ is called a *Gaussian process* if for all $t_1, \dots, t_n \geq 0$ the random vector $(X_{t_1}, \dots, X_{t_n})$ is Gaussian.

Definition 119. Let $(X_t)_{t \geq 0}$ be a Gaussian process. Then, the *mean function* is defined as:

$$\begin{aligned} \mu : [0, \infty) &\longrightarrow \mathbb{R} \\ t &\longmapsto \mathbb{E}(X_t) =: \mu_t \end{aligned}$$

and the *covariance function* is defined as:

$$\begin{aligned} C : [0, \infty) \times [0, \infty) &\longrightarrow \mathbb{R} \\ (s, t) &\longmapsto \text{Cov}(X_s, X_t) = \mathbb{E}(X_s X_t) - \mu_s \mu_t \end{aligned}$$

Brownian motion

Definition 120. A stochastic process $(B_t)_{t \geq 0}$ is called a *Brownian motion* (or a *Wiener process*) with parameter λ if:

1. $B_0 = 0$.
2. B_t has independent increments.
3. B_t has stationary increments with distribution $B_t - B_s \sim N(0, \sigma^2(t-s)) \forall 0 \leq s \leq t$ with $\sigma > 0$.
4. The trajectories $t \rightarrow B_t$ are continuous.

The Brownian motion is said to be *standard* if $\sigma = 1$.

Proposition 121. Let $B := (B_t)_{t \geq 0}$ be a standard Brownian motion. Then, B is a Gaussian process with mean function $\mu_t = 0$ and covariance function $C(s, t) = \min(s, t)$.

Proof. Let $0 < t_1 < \dots < t_n$. We can write the vector $\mathbf{b} := (B_{t_1}, \dots, B_{t_n})^T$ as:

$$\mathbf{b} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & 0 \\ 1 & \dots & 1 & 1 \end{pmatrix} \begin{pmatrix} B_{t_1} \\ B_{t_2} - B_{t_1} \\ \vdots \\ B_{t_n} - B_{t_{n-1}} \end{pmatrix}$$

And so $\mathbf{b}$ is Gaussian because is a linear combination of Gaussian. Now in the general, let $s_1, \dots, s_n \geq 0$. We can write any vector $(B_{s_1}, \dots, B_{s_n})$ as a linear transformation of the vector $(B_{t_1}, \dots, B_{t_n})$ with $0 < t_1 < \dots < t_n$. On the other hand, it is clear that $\mu_t = \mathbb{E}(B_t) = \mathbb{E}(B_t - B_0) = 0$ and if $s \leq t$:

$$\mathbb{E}(B_s B_t) = \mathbb{E}(B_s(B_t - B_s)) + \mathbb{E}(B_s^2) = \mu_s \mu_{t-s} + s = s$$

□

Proposition 122. Let $B := (B_t)_{t \geq 0}$ be a Gaussian process with $B_0 = 0$, mean function $\mu_t = 0$ and covariance function $C(s, t) = \min(s, t)$. Then, B is a standard Brownian motion.

Proof. Since Gaussian uncorrelated variables are independent, it suffices to show that the covariance matrix of $(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}})$ is:

$$\begin{pmatrix} t_1 & 0 & \dots & 0 \\ 0 & t_2 - t_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & t_n - t_{n-1} \end{pmatrix}$$

for all $0 < t_1 < \dots < t_n$. On the one hand, if $t_i < t_j$, then $t_{i-1} < t_i \leq t_{j-1} < t_j$ and:

$$\mathbb{E}((B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})) = \mathbb{E}(B_{t_i} B_{t_j}) - \mathbb{E}(B_{t_i} B_{t_{j-1}}) - \mathbb{E}(B_{t_{i-1}} B_{t_j}) + \mathbb{E}(B_{t_{i-1}} B_{t_{j-1}}) = t_i - t_i - t_{i-1} + t_{i-1} = 0$$

On the other hand:

$$\begin{aligned} \mathbb{E}((B_{t_i} - B_{t_{i-1}})^2) &= \mathbb{E}(B_{t_i}^2) - 2\mathbb{E}(B_{t_i} B_{t_{i-1}}) + \mathbb{E}(B_{t_{i-1}}^2) = \\ &= t_i - t_{i-1} \end{aligned}$$

□

Proposition 123. Let $B := (B_t)_{t \geq 0}$ be a standard Brownian motion. Then, the following stochastic processes are also standard Brownian motions:

1. $\forall c \in \mathbb{R}^*, (cB_{t/c^2})_{t \geq 0}$
2. $(-B_t)_{t \geq 0}$
3. $(B_{t+s} - B_s)_{t \geq 0}$ for all $s \geq 0$.
- 4.

$$Y_t = \begin{cases} tB_{1/t} & \text{if } t > 0 \\ 0 & \text{if } t = 0 \end{cases}$$

Definition 124. Let $(X_t)_{t \in T}$ and $(Y_t)_{t \in T}$ be two stochastic processes. We say that $(X_t)_{t \in T}$ and $(Y_t)_{t \in T}$ are *stochastically equivalent* if $\forall t \in T$ we have:

$$\mathbb{P}(X_t = Y_t) = 1$$

In that case we also say that $(X_t)_{t \in T}$ is a *version* of $(Y_t)_{t \in T}$ (or vice versa). We say that $(X_t)_{t \in T}$ and $(Y_t)_{t \in T}$ are *indistinguishable* if:

$$\mathbb{P}(X_t = Y_t \forall t \in T) = 1$$

Remark. Note that if the set T is finite or countable, then the two notions are equivalent because:

$$\mathbb{P}(X_t = Y_t \forall t \in T) = \mathbb{P}\left(\bigcap_{t \in T} \{X_t = Y_t\}\right) = 1$$

Proposition 125. Let $(X_t)_{t \in T}$ and $(Y_t)_{t \in T}$ be equivalent stochastic processes. Then, $\forall n \in \mathbb{N}$ and all $t_1, \dots, t_n \in T$ we have:

$$\mathbf{X} := (X_{t_1}, \dots, X_{t_n}) \stackrel{d}{=} (Y_{t_1}, \dots, Y_{t_n}) =: \mathbf{Y}$$

Proof. Let $B \in \mathcal{B}(\mathbb{R}^n)$ and $A := \{X_{t_i} = Y_{t_i} \forall i\}$. Using that $\mathbb{P}(A) = 1$ we have:

$$\begin{aligned} \mathbb{P}(\mathbf{X} \in B) &= \mathbb{P}(\mathbf{X} \in B, A) + \mathbb{P}(\mathbf{X} \in B, A^c) \\ &= \mathbb{P}(\mathbf{Y} \in B, A) + 0 \\ &= \mathbb{P}(\mathbf{Y} \in B, A) + \mathbb{P}(\mathbf{Y} \in B, A^c) \\ &= \mathbb{P}(\mathbf{Y} \in B) \end{aligned}$$

□

Corollary 126. Let B be a standard Brownian motion and $\bar{B}$ be a version of B . Then, $\bar{B}$ is also a standard Brownian motion.

Theorem 127 (Kolmogorov's continuity theorem). Let $(X_t)_{t \geq 0}$ be a stochastic process such that $\exists \alpha, \beta, C > 0$ such that:

$$\mathbb{E}(|X_t - X_s|^\alpha) \leq C|t - s|^{1+\beta}$$

for all $t, s \geq 0$. Then, there exists a version of $(X_t)_{t \geq 0}$ with continuous trajectories.

Lemma 128. Let $X \sim N(0, \sigma^2)$. Then, $\mathbb{E}(|X|^n) = C_n \sigma^n$ where:

$$C_n = \mathbb{E}(|Z|^n) = \Gamma\left(\frac{n+1}{2}\right) \frac{2^{n/2}}{\sqrt{\pi}}$$

and $Z \sim N(0, 1)$.

Corollary 129. Let $B := (B_t)_{t \geq 0}$ be a standard Brownian motion. Then, there exists a version of B with continuous trajectories.

Proof. We use Theorems 127 and 128 with $\alpha = 3$ and $\beta = 1/2$. $\square$

Proposition 130. Let $B := (B_t)_{t \geq 0}$ be a standard Brownian motion. Then, for any interval $[a, b] \subset \mathbb{R}$:

$$\mathbb{P}(\omega \in \Omega : B(\omega) \text{ is monotone on } [a, b]) = 0$$

Proof. Let $A = \{\omega \in \Omega : B(\omega) \text{ is monotone on } [a, b]\}$. Using the density and continuity of B we have:

$$A = \{B_s \leq B_t : \forall s, t \in \mathbb{Q}, a \leq s < t \leq b\} + \{B_s \geq B_t : \forall s, t \in \mathbb{Q}, a \leq s < t \leq b\}$$

Hence, given $n \in \mathbb{N}$ and a partition $\{t_i\}_{0 \leq i \leq n}$ of $[a, b]$ we have:

$$A \subseteq \{B_{t_{i+1}} - B_{t_i} \geq 0 : i = 0, \dots, n-1\} + \{B_{t_{i+1}} - B_{t_i} \leq 0 : i = 0, \dots, n-1\}$$

Therefore, using the independence of the increments of B and the symmetry of the normal distribution we have:

$$\mathbb{P}(A) \leq 2 \prod_{i=0}^{n-1} \mathbb{P}(B_{t_{i+1}} - B_{t_i} \geq 0) = 2 \left(\frac{1}{2}\right)^n \xrightarrow{n \rightarrow \infty} 0$$

$\square$

Proposition 131. Let $B := (B_t)_{t \geq 0}$ be a standard Brownian motion. Then, $\forall t \geq 0$ the set

$$A := \left\{ \omega \in \Omega : \limsup_{h \rightarrow 0} \frac{|B_{t+h}(\omega) - B_t(\omega)|}{h} = +\infty \right\}$$

which may not belong in the σ -algebra, contains an event of probability 1.

Proof. Note that $A \supseteq \{\omega \in \Omega : \sup_{n \in \mathbb{N}} \frac{|B_{t+1/n}(\omega) - B_t(\omega)|}{1/n} = +\infty\} = \bigcap_{M \geq 1} A_M$ where:

$$A_M := \left\{ \omega \in \Omega : \sup_{n \in \mathbb{N}} \frac{|B_{t+1/n}(\omega) - B_t(\omega)|}{1/n} \geq M \right\}$$

If we see that $\mathbb{P}(A_M) = 1$ for all $M \geq 1$ we are done.

$$\begin{aligned} \mathbb{P}(A_M) &\geq \mathbb{P}\left(\frac{|B_{t+1/n}(\omega) - B_t(\omega)|}{1/n} \geq M\right) = \\ &= \mathbb{P}\left(|Z| \geq \frac{M}{\sqrt{n}}\right) = 2\left(1 - \Phi\left(\frac{M}{\sqrt{n}}\right)\right) \xrightarrow{n \rightarrow \infty} 1 \end{aligned}$$

where the first inequality holds $\forall n \in \mathbb{N}$, $Z \sim N(0, 1)$ and Φ is the cumulative distribution function of the standard normal distribution. $\square$

Theorem 132 (Paley-Wiener-Zygmund theorem).

The Brownian trajectories are almost surely nowhere differentiable. Namely, the set

$$\left\{ \omega \in \Omega : \forall t \geq 0, \limsup_{h \rightarrow 0^+} \frac{B_{t+h}(\omega) - B_t(\omega)}{h} = +\infty \text{ or } \liminf_{h \rightarrow 0^+} \frac{B_{t+h}(\omega) - B_t(\omega)}{h} = -\infty \right\}$$

contains an event of probability 1. And the same occurs for the left limit $h \rightarrow 0^-$ (in this case we need to exclude $t = 0$).

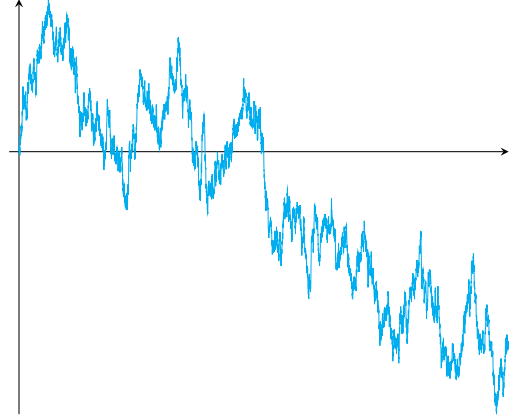


Figure 2: A Brownian motion simulated with 7500 increments. Observe the “non-differentiability” of the path.

Definition 133. Let B be a standard Brownian motion and $a > 0$. We define:

$$\tau_a := \inf\{t \geq 0 : B_t \geq a\}$$

If $a < 0$ we define:

$$\tau_a := \inf\{t \geq 0 : B_t \leq a\}$$

Remark. Note that if $\tau_a(\omega) < \infty$, then $B_{\tau_a(\omega)}(\omega) = a$ by the continuity of the trajectories.

Lemma 134. Let B be a standard Brownian motion. Then, $\forall t \geq 0$ and $a > 0$:

$$\mathbb{P}(\tau_a \leq t) = \mathbb{P}\left(\max_{0 \leq s \leq t} B_s \geq a\right) = 2\mathbb{P}(B_t \geq a)$$

If $a < 0$ we have:

$$\mathbb{P}(\tau_a \leq t) = \mathbb{P}\left(\min_{0 \leq s \leq t} B_s \leq a\right) = 2\mathbb{P}(B_t \leq a)$$

Corollary 135. Let B be a standard Brownian motion and $a \in \mathbb{R}^*$. Then, $\mathbb{P}(\tau_a < \infty) = 1$.

Proof. Assume $a > 0$, the other case is similar. Then:

$$\begin{aligned} \mathbb{P}(\tau_a < \infty) &= \mathbb{P}\left(\bigcup_{n=1}^{\infty} \{\tau_a \leq n\}\right) = \lim_{n \rightarrow \infty} \mathbb{P}(\tau_a \leq n) = \\ &= 2 \lim_{n \rightarrow \infty} \mathbb{P}(B_n \geq a) = 2 \lim_{n \rightarrow \infty} \left(1 - \Phi\left(\frac{a}{\sqrt{n}}\right)\right) = 1 \end{aligned}$$

$\square$

Proposition 136. Let B be a standard Brownian motion. Then:

$$\mathbb{P}\left(\sup_{t \geq 0} B_t = +\infty, \inf_{t \geq 0} B_t = -\infty\right) = 1$$

Proof. It suffices to prove only $\mathbb{P}(\sup_{t \geq 0} B_t = +\infty) = 1$. Note that:

$$\mathbb{P}\left(\sup_{t \geq 0} B_t = +\infty\right) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} \left\{\sup_{t \geq 0} B_t \geq n\right\}\right)$$

Let's see that all the events $A_n := \{\sup_{t \geq 0} B_t \geq n\}$ in the intersection have probability 1:

$$\begin{aligned} \mathbb{P}(A_n) &\geq \mathbb{P}\left(\max_{0 \leq t \leq s} B_t \geq n\right) = 2\mathbb{P}(B_s \geq n) = \\ &= 2\left(1 - \Phi\left(\frac{n}{\sqrt{s}}\right)\right) \xrightarrow{s \rightarrow \infty} 1 \end{aligned}$$

□

Corollary 137. The Brownian trajectories have infinite zeros almost surely, and they tend to infinity.

Proof. Let $B := \{B_t : t \geq 0\}$ be a standard Brownian motion and let $A = \{\omega \in \Omega : B(\omega) \text{ has finite zeros}\} \subseteq \bigcup_{n=1}^{\infty} \{\omega \in \Omega : B(\omega) \text{ doesn't vanish in } [n, \infty)\}$. Let's see that all the events $A_n := \{B \text{ doesn't vanish in } [n, \infty)\}$ in the union have probability 0.

$$\begin{aligned} \mathbb{P}(A_n) &= \mathbb{P}(B > 0 \text{ in } [n, \infty)) + \mathbb{P}(B < 0 \text{ in } [n, \infty)) \leq \\ &\leq \mathbb{P}\left(\inf_{t \geq n} B_t \neq -\infty\right) + \mathbb{P}\left(\sup_{t \geq n} B_t \neq +\infty\right) = 0 \end{aligned}$$

□

Corollary 138. The Brownian trajectories pass through every point $a \in \mathbb{R}$ infinitely many times almost surely.

Proposition 139. Let B be a standard Brownian motion. Then, $\forall h > 0$:

$$\mathbb{P}\left(\max_{0 \leq t \leq h} B_t > 0, \min_{0 \leq t \leq h} B_t < 0\right) = 1$$

Proof. It suffices to prove $\mathbb{P}(\max_{0 \leq t \leq h} B_t > 0) = 1$. Let $a > 0$. Then:

$$\begin{aligned} \mathbb{P}\left(\max_{0 \leq t \leq h} B_t > 0\right) &\geq \mathbb{P}\left(\max_{0 \leq t \leq h} B_t > a\right) = 2\mathbb{P}(B_h > a) = \\ &= 2\left(1 - \Phi\left(\frac{a}{\sqrt{h}}\right)\right) \xrightarrow{a \rightarrow 0} 1 \end{aligned}$$

□

Proposition 140. Let B be a standard Brownian motion and $a \in \mathbb{R}^*$. Then, $\mathbb{E}(\tau_a) = \infty$.

Proof. Using the symmetry of the Brownian motion, we have:

$$F_{\tau_a}(t) = \mathbb{P}(\tau_a \leq t) = 2\mathbb{P}(B_t < -a) = \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{a}{\sqrt{t}}} e^{-\frac{x^2}{2}} dx$$

and $F_{\tau_a}(0) = 0$. An easy check of the hypothesis of ?? shows that the density of τ_a is F_{τ_a}' and so:

$$\mathbb{E}(\tau_a) = \frac{a}{\sqrt{2\pi}} \int_0^{\infty} \frac{e^{-\frac{a^2}{2t}}}{\sqrt{t}} dt = \infty$$

because it diverges at ∞ . □

Definition 141. A d -dimensional standard Brownian motion is a d -dimensional stochastic process $\mathbf{B} = (B^1, \dots, B^d)$ such that $\forall i \in \{1, \dots, d\}$, B^i is a standard Brownian motion, and it is independent of the other components.

Theorem 142. Let $\mathbf{B}$ be a d -dimensional Brownian motion. Then:

1. If $d = 2$, then B is recurrent, that is $\forall \mathbf{x} \in \mathbb{R}^2$ and $\forall \delta > 0 \exists (\tau_n) \in \mathbb{R}$, with $\lim_{n \rightarrow \infty} \tau_n = +\infty$, such that:

$$\mathbb{P}(B_{\tau_n} \in B_{\delta}(\mathbf{x}) \forall n \in \mathbb{N}) = 1$$

Here $B_{\delta}(\mathbf{x})$ denotes the open ball of radius δ centered at $\mathbf{x}$.

2. If $d \geq 3$, then B is transient, that is $\forall M > 0, \exists T > 0$ such that:

$$\mathbb{P}(B_t \in B_M(\mathbf{0}) \forall t \geq T) = 0$$

Theorem 143 (Law of the iterated logarithm). Let B be a standard Brownian motion. Then:

$$\limsup_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \log \log t}} \stackrel{\text{a.e.}}{=} \liminf_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \log \log t}} \stackrel{\text{a.e.}}{=} 1$$

Corollary 144. Let B be a standard Brownian motion. Then:

$$\limsup_{h \rightarrow 0} \frac{B_h}{\sqrt{2h \log \log \frac{1}{h}}} \stackrel{\text{a.e.}}{=} \liminf_{h \rightarrow 0} \frac{B_h}{\sqrt{2h \log \log \frac{1}{h}}} \stackrel{\text{a.e.}}{=} 1$$

Sketch of the proof. Recall that $x B_{1/x}$ is a standard Brownian motion. □

Proposition 145. Let $S_n = \sum_{i=1}^n X_i$ be a simple random walk with $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$ and $B = (B_t)$ be a standard Brownian motion. We define the following sequence of stochastic processes:

$$Y_t^n := \frac{1}{\sqrt{n}} [S_{[nt]} + (nt - [nt])X_{[nt]+1}]$$

Then, $Y_t^n \xrightarrow{d} B_t$.

Existence of Brownian motion

Definition 146 (Finite-dimensional distributions).

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $X : I \times \Omega \rightarrow (E, \mathcal{E})$ be a stochastic process. The *finite-dimensional distributions* of X are the probability measures $\mathbb{P}_{t_1, \dots, t_n}$ defined on $(E^n, \mathcal{E}^n)$ by:

$$\mathbb{P}_{t_1, \dots, t_n}(B) := \mathbb{P}((X_{t_1}, \dots, X_{t_n}) \in B)$$

for all $B \in \mathcal{E}^n$.

Lemma 147. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $X : I \times \Omega \rightarrow (E, \mathcal{E})$ be a stochastic process. Then, the finite-dimensional distributions satisfy the following *consistency condition*:

1. For all $n \in \mathbb{N}$, $t_1, \dots, t_n \in I$, $B_1, \dots, B_n \in \mathcal{E}$ and $\sigma \in S_n$, we have:

$$\begin{aligned} \mathbb{P}_{t_1, \dots, t_n}(B_1 \times \dots \times B_n) &= \\ &= \mathbb{P}_{t_{\sigma(1)}, \dots, t_{\sigma(n)}}(B_{\sigma(1)} \times \dots \times B_{\sigma(n)}) \end{aligned}$$

2. For all $n \in \mathbb{N}$, $t_1, \dots, t_n \in I$ and $B_1, \dots, B_{n-1} \in \mathcal{E}$, we have:

$$\begin{aligned} \mathbb{P}_{t_1, \dots, t_n}(B_1 \times \dots \times B_{n-1} \times E) &= \\ &= \mathbb{P}_{t_1, \dots, t_{n-1}}(B_1 \times \dots \times B_{n-1}) \end{aligned}$$

Theorem 148 (Kolmogorov extension theorem).

Let I be a set and $\{\mathbb{P}_{t_1, \dots, t_n} : n \in \mathbb{N}, t_1, \dots, t_n \in I\}$ be a family of probabilities defined on $\mathcal{B}(\mathbb{R}^n)$ satisfying the consistency conditions. Then, there exists a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a stochastic process $X : I \times \Omega \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $\mathbb{P}_{t_1, \dots, t_n}$ is the finite-dimensional distribution of X for all $n \in \mathbb{N}$ and $t_1, \dots, t_n \in I$.