# Chapter 9: Equilibrium
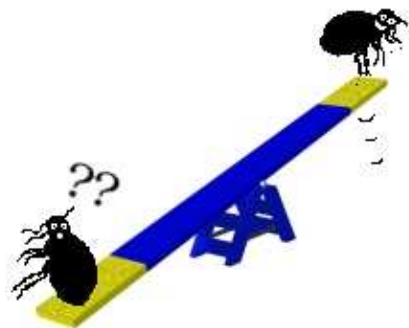
In Chapter 8, we saw that if $\{X_0, X_1, X_2, \ldots\}$ is a Markov chain with transition matrix $P$, then
$$X_t \sim \boldsymbol{\pi}^T \quad \Rightarrow \quad X_{t+1} \sim \boldsymbol{\pi}^T P.$$

This raises the question: is there any distribution $\boldsymbol{\pi}$ such that

If $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$, then

In other words, if $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$, and $X_t \sim \boldsymbol{\pi}^T$, then

Thus, once a Markov chain has reached a distribution $\boldsymbol{\pi}^T$ such that $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$,

If $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$, we say that the distribution $\boldsymbol{\pi}^T$ is an

***Equilibrium*** means a ***level position:*** there is *no more change* in the distribution of $X_t$ as we wander through the Markov chain.

***Note:*** Equilibrium does not mean that the **value** of $X_{t+1}$ equals the **value** of $X_t$. It means that the **distribution** of $X_{t+1}$ is the same as the **distribution** of $X_t$:

In this chapter, we will first see how to *calculate* the equilibrium distribution $\boldsymbol{\pi}^T$. We will then see the remarkable result that many Markov chains automatically *find their own way* to an equilibrium distribution as the chain wanders through time. This happens for many Markov chains, but not all. We will see the conditions required for the chain to find its way to an equilibrium distribution.

## 9.1   Equilibrium distribution in pictures
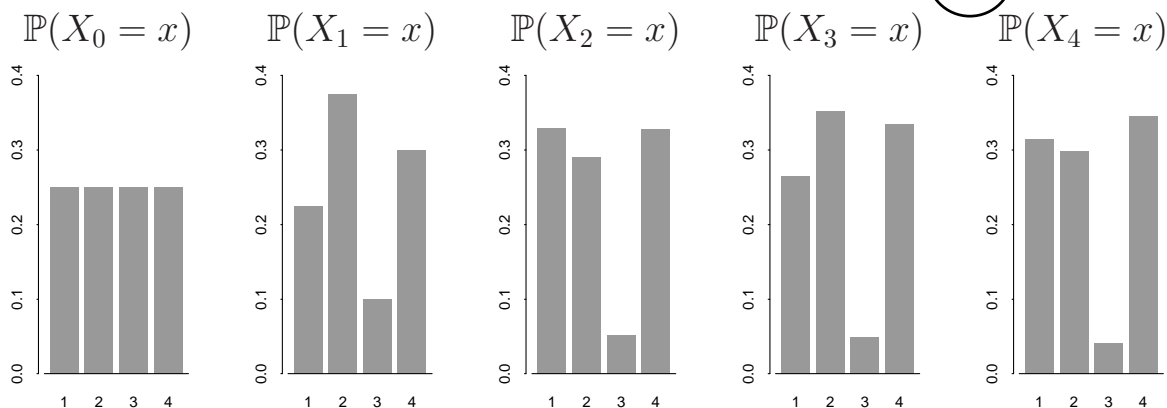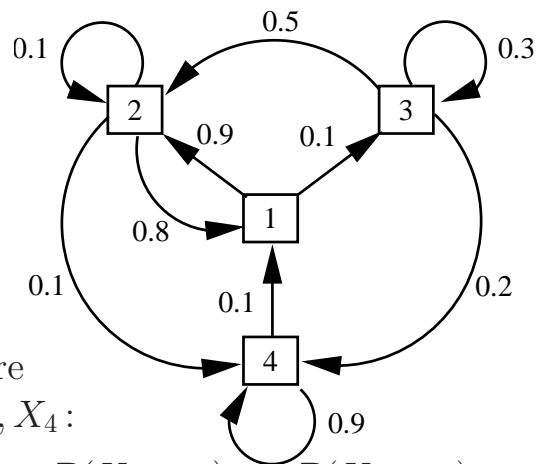
Consider the following 4-state Markov chain:

$$P = \begin{pmatrix} 0.0 & 0.9 & 0.1 & 0.0 \\ 0.8 & 0.1 & 0.0 & 0.1 \\ 0.0 & 0.5 & 0.3 & 0.2 \\ 0.1 & 0.0 & 0.0 & 0.9 \end{pmatrix}$$

Suppose we start at time 0 with
$X_0 \sim \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$: so the chain is equally
likely to start from any of the four states. Here
are pictures of the distributions of $X_0, X_1, \ldots, X_4$:



The distribution starts off level, but quickly changes: for example the chain is
least likely to be found in state 3. The distribution of $X_t$ changes between each
$t = 0, 1, 2, 3, 4$. Now look at the distribution of $X_t$ 500 steps into the future:
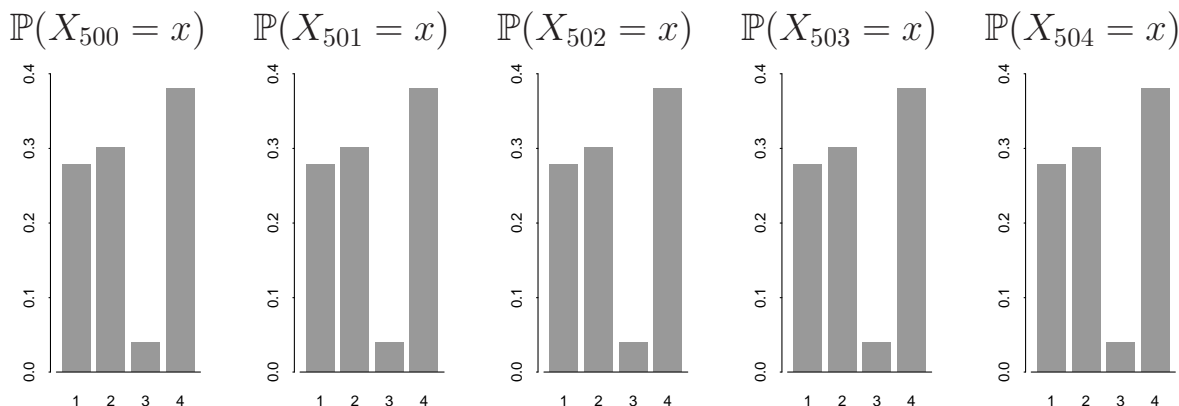


The distribution has reached a steady state: it ***does not change*** between
$t = 500, 501, \ldots, 504$.

## 9.2 Calculating equilibrium distributions

*Definition:* Let $\{X_0, X_1, \ldots\}$ be a Markov chain with transition matrix $P$ and state space $S$, where $|S| = N$ (possibly infinite). Let $\boldsymbol{\pi}^T$ be a row vector denoting a probability distribution on $S$: so each element $\pi_i$ denotes the probability of being in state $i$, and $\sum_{i=1}^{N} \pi_i = 1$, where $\pi_i \geq 0$ for all $i = 1, \ldots, N$. The probability distribution $\boldsymbol{\pi}^T$ is an **equilibrium** distribution for the Markov chain if

That is, $\boldsymbol{\pi}^T$ is an equilibrium distribution if

By the argument given on page 164, we have the following Theorem:

**Theorem 9.2:** Let $\{X_0, X_1, \ldots\}$ be a Markov chain with transition matrix $P$. Suppose that $\boldsymbol{\pi}^T$ is an equilibrium distribution for the chain. If $X_t \sim \boldsymbol{\pi}^T$ for any $t$, then $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Once a chain has hit an equilibrium distribution,

***Note:*** There are several other names for an equilibrium distribution. If $\boldsymbol{\pi}^T$ is an equilibrium distribution, it is also called:

- **invariant:**
- **stationary:**

## Stationarity: the Chain Station



a B U S station is where a B U S stops

a t r a i n station is where a t r a i n stops

a **workstation** is where . . . ? ? ?



a stationary distribution is where a Markov chain stops

## 9.3  Finding an equilibrium distribution

Vector $\boldsymbol{\pi}^T$ is an equilibrium distribution for $P$ if:

   1.

   2.

   3.

Conditions 2 and 3 ensure that

Condition 1 means that $\boldsymbol{\pi}$ is a row eigenvector of $P$.

Solving $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$ by itself will just specify $\boldsymbol{\pi}$ up to
We need to include Condition 2 to scale $\boldsymbol{\pi}$ to a genuine probability distribution,
and then check with Condition 3 that the scaled distribution is valid.

***Example:***  Find an equilibrium distribution for the Markov chain below.

$$
P = \begin{pmatrix}
0.0 & 0.9 & 0.1 & 0.0 \\
0.8 & 0.1 & 0.0 & 0.1 \\
0.0 & 0.5 & 0.3 & 0.2 \\
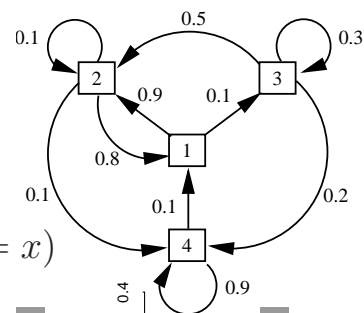0.1 & 0.0 & 0.0 & 0.9
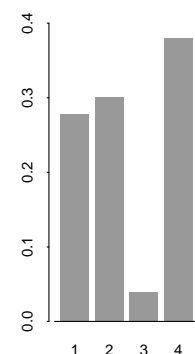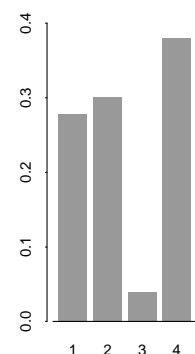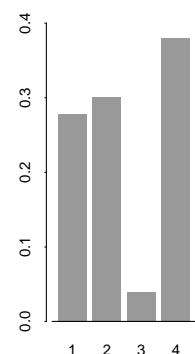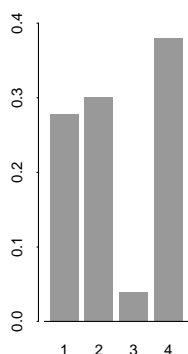\end{pmatrix}
$$

***Solution:***
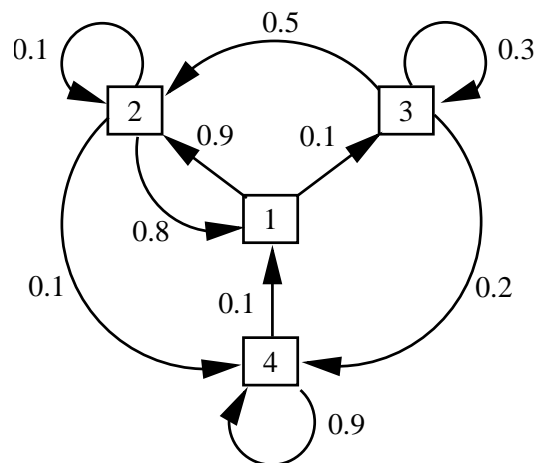
This is the distribution the chain converged to in Section 9.1.

## 9.4  Long-term behaviour

In Section 9.1, we saw an example where the Markov chain wandered of its own accord into its equilibrium distribution:

$$\mathbb{P}(X_{500} = x) \quad \mathbb{P}(X_{501} = x) \quad \mathbb{P}(X_{502} = x) \quad \mathbb{P}(X_{503} = x)$$

This will always happen for this Markov chain. In fact, the distribution it converges to (found above) does not depend upon the starting conditions:

What is happening here is that

$$P = \begin{pmatrix} 0.0 & 0.9 & 0.1 & 0.0 \\ 0.8 & 0.1 & 0.0 & 0.1 \\ 0.0 & 0.5 & 0.3 & 0.2 \\ 0.1 & 0.0 & 0.0 & 0.9 \end{pmatrix} \quad \Rightarrow \quad P^t \rightarrow \begin{pmatrix} 0.28 & 0.30 & 0.04 & 0.38 \\ 0.28 & 0.30 & 0.04 & 0.38 \\ 0.28 & 0.30 & 0.04 & 0.38 \\ 0.28 & 0.30 & 0.04 & 0.38 \end{pmatrix} \quad \text{as } t \to \infty.$$

(If you have a calculator that can handle matrices, try finding $P^t$ for $t = 20$ and $t = 30$: you will find the matrix is already converging as above.)
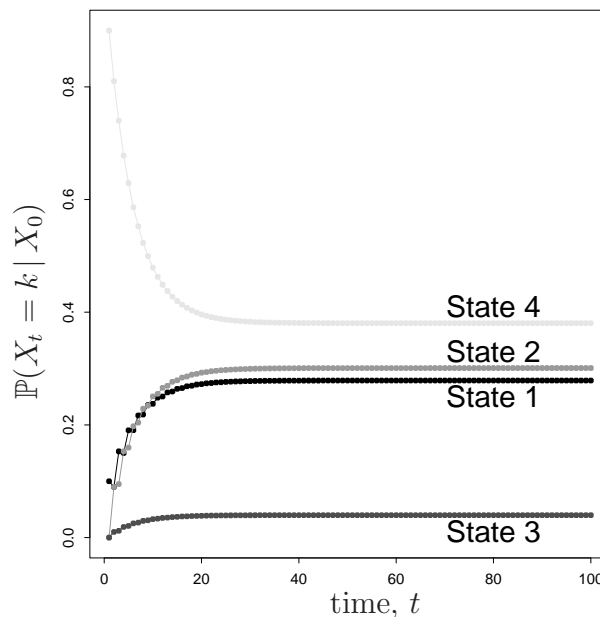
This convergence of $P^t$ means that

- about          of being in State    after $t$ steps;

- about          of being in State    after $t$ steps;

- about          of being in State    after $t$ steps;

- about          of being in State    after $t$ steps.

## Start at $X_0 = 2$
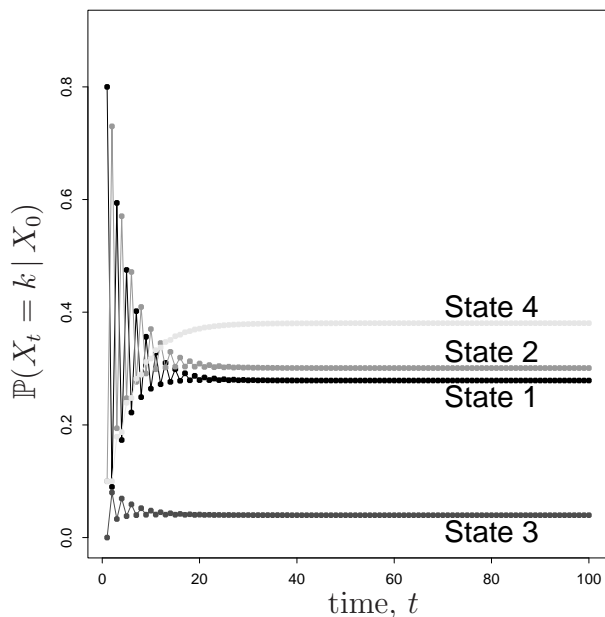


## Start at $X_0 = 4$



The **left graph** shows the probability of getting from state 2 to state $k$ in $t$ steps, as $t$ changes: $(P^t)_{2,k}$ for $k = 1, 2, 3, 4$.

The **right graph** shows the probability of getting from state 4 to state $k$ in $t$ steps, as $t$ changes: $(P^t)_{4,k}$ for $k = 1, 2, 3, 4$.

The ***initial behaviour*** differs greatly for the different start states.

The ***long-term behaviour*** (large $t$) is the same for both start states.

However, this does not always happen. Consider the two-state chain below:

$$ \begin{array}{cc} \text{1} \xrightleftharpoons[1]{1} \text{2} & \qquad P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array} $$

As $t$ gets large,

$$ P^{500} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad P^{501} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad P^{502} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad P^{503} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdots $$

For this Markov chain,

## General formula for $P^t$

We have seen that we are interested in whether $P^t$ converges to

If it does, then the Markov chain will

The equilibrium distribution is then given by

It can be shown that a general formula is available for $P^t$ for any $t$, based on the eigenvalues of $P$. Producing this formula is beyond the scope of this course, but if you are given the formula, you should be able to recognise whether $P^t$ is going to converge to a fixed matrix with all rows the same.

*Example 1:*



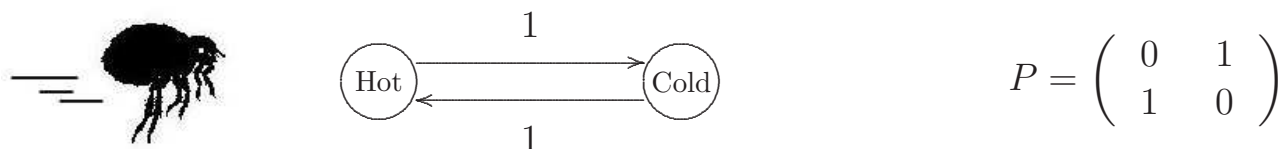$$P = \begin{pmatrix} 0.2 & 0.8 \\ 0.6 & 0.4 \end{pmatrix}$$
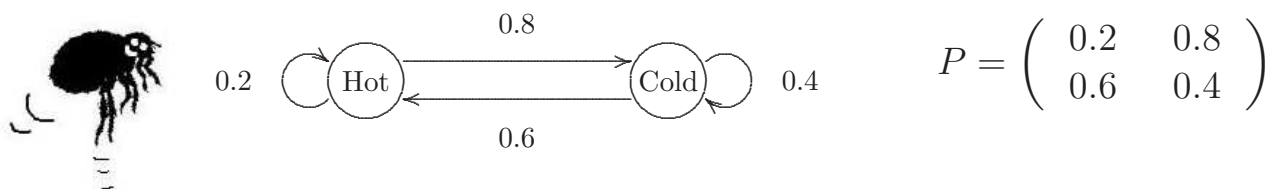
We can show that the general solution for $P^t$ is:

$$P^t = \frac{1}{7}\left\{\begin{pmatrix} 3 & 4 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 4 & -4 \\ -3 & 3 \end{pmatrix}(-0.4)^t\right\}$$

*Exercise:* Verify that $\boldsymbol{\pi}^T = \left(\frac{3}{7}, \frac{4}{7}\right)$ is the same as the result you obtain from solving the equilibrium equations: $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$ and $\pi_1 + \pi_2 = 1$.

**Example 2:** Purposeflea knows exactly what he is doing, so his probabilities are
all 1:



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

We can show that the general solution for $P^t$ is:

$$P^t = \frac{1}{2}\left\{ \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} (-1)^t \right\}$$

**Exercise:** Verify that this Markov chain *does* have an equilibrium distribution,
$\boldsymbol{\pi}^T = \left(\frac{1}{2}, \frac{1}{2}\right)$. However, the chain does not *converge* to this distribution as
$t \to \infty$.

These examples show that some Markov chains forget their starting conditions
in the long term, and ensure that $X_t$ will have the same distribution as $t \to \infty$
regardless of where we started at $X_0$. However, for other Markov chains, the
initial conditions are never forgotten. In the next sections we look for general
criteria that will ensure the chain converges.

**Target Result:**

- If a Markov chain is ***irreducible*** and ***aperiodic***, and if an equilibrium distribution $\boldsymbol{\pi}^T$ ***exists***, then the chain converges to this distribution as $t \to \infty$, regardless of the initial starting states.

To make sense of this, we need to revise the concept of ***irreducibility***, and introduce the idea of ***aperiodicity.***

---

## 9.5  Irreducibility

Recall from Chapter 8:

*Definition:*  A Markov chain or transition matrix $P$ is said to be **irreducible** if



Irreducibility of a Markov chain is important for convergence to equilibrium as $t \to \infty$, because

This can happen if the chain is irreducible. When the chain is not irreducible, different start states might cause the chain to get stuck in different closed classes. In the example above, a start state of $X_0 = 1$ means that the chain is restricted to states 1 and 2 as $t \to \infty$, whereas a start state of $X_0 = 4$ means that the chain is restricted to states 4 and 5 as $t \to \infty$. A single convergence that 'forgets' the initial state is therefore not possible.

---

## 9.6 Periodicity

Consider the Markov chain with transition matrix $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
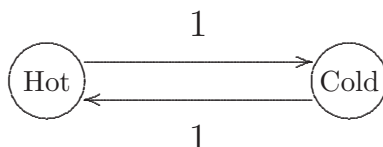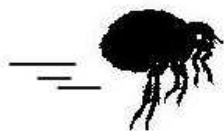


Suppose that $X_0 = 1$.

Then

This sort of behaviour is called **periodicity:**

Clearly, periodicity of the chain will interfere with convergence to an equilibrium distribution as $t \to \infty$. For example,

$$\mathbb{P}(X_t = 1 \,|\, X_0 = 1) = \begin{cases} 1 \text{ for even values of } t, \\ 0 \text{ for odd values of } t. \end{cases}$$

Therefore, the probability can not converge to any <u>single</u> value as $t \to \infty$.

---

### Period of state $i$

To formalize the notion of periodicity, we define the **period** of a state $i$. Intuitively,

In the example above, the chain can return to state 1 after

The period of state 1 is therefore

In general, the chain can return from state $i$ back to state $i$ again in $t$ steps if
        This prompts the following definition.

*Definition:* The **period** $d(i)$ of a state $i$ is

*Definition:* The state $i$ is said to be **periodic** if

> For a periodic state $i$, $(P^t)_{ii} = 0$ if $t$ is not a multiple of $d(i)$.

*Definition:* The state $i$ is said to be **aperiodic** if

If state $i$ is aperiodic, it means that

For convergence to equilibrium as $t \to \infty$, we will be interested only in

The following examples show how to calculate the period for both aperiodic and periodic states.
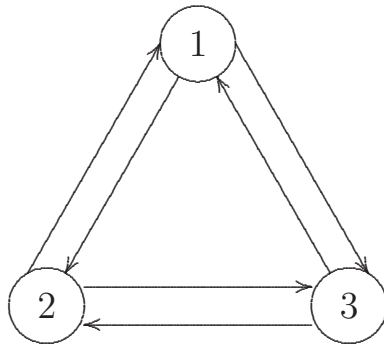
***Examples:*** Find the periods of the given states in the following Markov chains, and state whether or not the chain is irreducible.
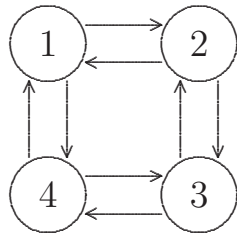
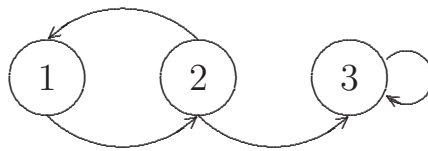1. The simple random walk.
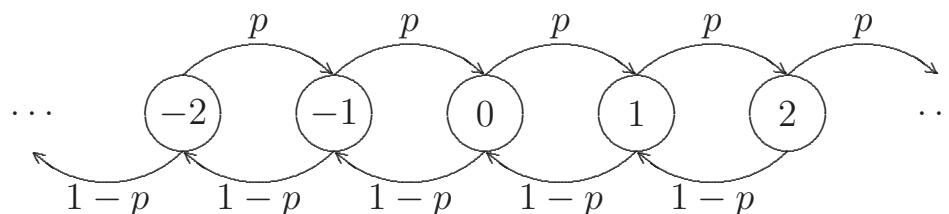


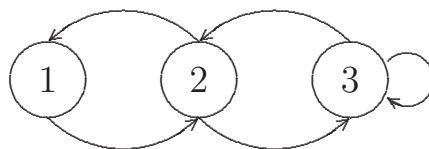$$d(0) =$$

2.

$$d(1) =$$

3.

$$d(1) =$$

4.

$$d(1) =$$

5.

$$d(1) =$$

## 9.7 Convergence to Equilibrium

We now draw together the threads of the previous sections with the following results.

**Fact:** If $i \leftrightarrow j$, then $i$ and $j$ have the same period. (Proof omitted.)

This leads immediately to the following result:

> If a Markov chain is ***irreducible*** and has ***one*** aperiodic state,
> then ***all*** states are aperiodic.

We can therefore talk about an **irreducible, aperiodic chain,**

**Theorem 9.7:** Let $\{X_0, X_1, \ldots\}$ be an **irreducible and aperiodic** Markov chain with transition matrix $P$. Suppose that there ***exists*** an equilibrium distribution $\boldsymbol{\pi}^T$. Then, from ***any*** starting state $i$, and for any end state $j$,

In particular,

**Note:** If the state space is infinite, it is not guaranteed that an equilibrium distribution $\boldsymbol{\pi}^T$ exists. See Example 3 below.
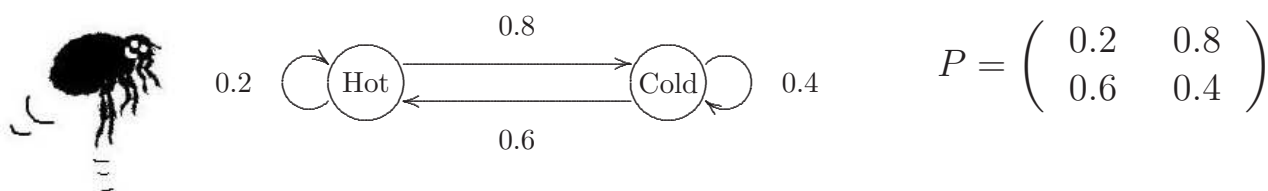
**Note:** If the chain converges to an equilibrium distribution $\boldsymbol{\pi}^T$ as $t \to \infty$, then

---

## 9.8 Examples

A typical exam question gives you a Markov chain on a finite state space and asks if it converges to an equilibrium distribution as $t \to \infty$. An equilibrium distribution will always exist for a finite state space. You need to check whether the chain is irreducible and aperiodic. If so, it will converge to equilibrium. If the chain is *periodic*, it cannot converge to an equilibrium distribution that is independent of start state. If the chain is *reducible*, it may or may not converge.

The first two examples are the same as the ones given in Section 9.4.

**Example 1:** State whether the Markov chain below converges to an equilibrium distribution as $t \to \infty$.



$$P = \left( \begin{array}{cc} 0.2 & 0.8 \\ 0.6 & 0.4 \end{array} \right)$$

***Example 2:*** State whether the Markov chain below converges to an equilibrium distribution as $t \to \infty$.

Hot →(1)→ Cold, Cold →(1)→ Hot
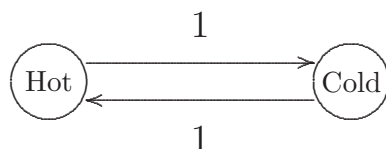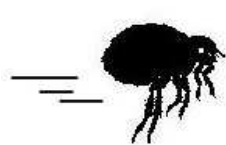
$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

It is important to check for aperiodicity, because the existence of an equilibrium distribution does NOT ensure convergence to this distribution if the matrix is not aperiodic.

---

***Example 3:*** Random walk with retaining barrier at 0.
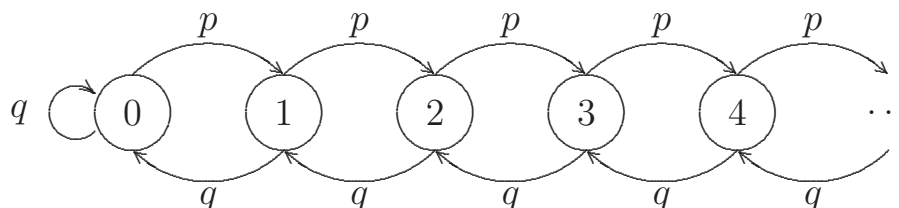
Find whether the chain converges to equilibrium as $t \to \infty$, and if so, find the equilibrium distribution.

***Note:*** Equilibrium results also exist for chains that are ***not*** aperiodic. Also, states can be classified as **transient** (return to the state is not certain), **null recurrent** (return to the state is certain, but the expected return time is infinite), and **positive recurrent** (return to the state is certain, and the expected return time is finite). For each type of state, the long-term behaviour is known:

- If the state $k$ is **transient** or **null-recurrent**,

$$\mathbb{P}(X_t = k \mid X_0 = k) = \left(P^t\right)_{kk} \to 0 \text{ as } t \to \infty.$$

- If the state is **positive recurrent**, then

$$\mathbb{P}(X_t = k \mid X_0 = k) = \left(P^t\right)_{kk} \to \pi_k \text{ as } t \to \infty, \text{ where } \pi_k > 0.$$

The expected return time for the state is $1/\pi_k$.

A detailed treatment is available at
http://www.statslab.cam.ac.uk/~james/Markov/.