Introduction to Markov Chains

What is a Markov **Chain**?

X_1 X_2 X_3 X_n

Random Variables

Could model as i.i.d.
(independent and identically distributed)

Realistic? What if index is time?

What is a Markov **Chain**?

$$\boxed{X_1} \longrightarrow \boxed{X_2} \longrightarrow \boxed{X_3} \dashrightarrow \boxed{X_n}$$

Let’s add a dash of dependence
(but not too much!)

The Markov Property

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_{n+1} = j | X_n = i)$$

The Markov Property

Everything Before

Next Now Previous First Next Now

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_{n+1} = j | X_n = i)$$

Memoryless!

State Space

$$\boxed{X_1} \longrightarrow \boxed{X_2} \longrightarrow \boxed{X_3} \dashrightarrow \boxed{X_n}$$

$$X_i \in \{Rainy, Sunny\}$$

http://setosa.io/ev/markov-chains/

State Spaces

$$X_1 \rightarrow X_2 \rightarrow X_3 \dashrightarrow X_n$$

$$X_i \in \{Rainy, Sunny\}$$

$$X_i \in \{1, 2, 3, 4, 5, 6\}$$

$$X_i \in \{A, C, G, T\}$$

$$X_i \in \{AAA, AAC, AAG, \ldots, TTG, TTT\}$$

State Spaces (Discrete)

$$\boxed{X_1} \rightarrow \boxed{X_2} \rightarrow \boxed{X_3} \dashrightarrow \boxed{X_n}$$

$$X_i \in \{Rainy, Sunny\}$$

$$X_i \in \{1, 2, 3, 4, 5, 6\}$$

$$X_i \in \{A, C, G, T\}$$

$$X_i \in \{AAA, AAC, AAG, \ldots, TTG, TTT\}$$

State Spaces (Continuous)

$$X_1 \rightarrow X_2 \rightarrow X_3 \dashrightarrow X_n$$

$$X_i \in \mathbb{R}$$

$$X_i \in \mathbb{R}_{>0}$$

$$X_i \in [0, 1]$$

What sorts of continuous state spaces might we have in phylogenetics?

Transition Matrix

$$\begin{array}{c c} & \begin{array}{cc} R & S \end{array} \\ \begin{array}{c} R \\ S \end{array} & \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix} \end{array}$$

Transition Matrix

To

From

$$\begin{array}{c c} & \begin{array}{cc} R & S \end{array} \\ \begin{array}{c} R \\ S \end{array} & \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix} \end{array}$$

$$
\begin{array}{ccc}
 & & \text{To} \\
 & R & S \\
\text{From} \begin{array}{c} R \\ S \end{array} & \multicolumn{2}{c}{\begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix}}
\end{array}
$$

$$P(X_{n+1} = R | X_n = R) = 0.7$$

$$P(X_{n+1} = S | X_n = R) = 0.3$$

$$\begin{array}{c|cc} & \multicolumn{2}{c}{\text{To}} \\ & R & S \\ \hline \text{From} \quad R & 0.7 & 0.3 \\ S & 0.3 & 0.7 \end{array}$$

$$\begin{array}{cc} & \begin{array}{cc} R & S \end{array} \\ \begin{array}{c} R \\ S \end{array} & \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix} \end{array}$$

$$P(X_{n+1} = R | X_n = R) = 0.7$$

$$P(X_{n+1} = S | X_n = R) = 0.3$$

$$P(X_{n+1} = R | X_n = S) = 0.3$$

$$P(X_{n+1} = S | X_n = S) = 0.7$$

http://setosa.io/markov/index.html#%7B%22tm%22%3A%5B%5B0.7%2C0.3%5D%2C%5B0.3%2C0.7%5D%5D%7D

In-Class Exercise (pairs)

(1) Create a Markov chain class with these values:

- Number of steps (positive integer)
- State space (list)
- Transition matrix (list of lists of floats - or own class!)
- Sampled states (list)

and these methods:

- run (sample states for each step)
- clear (remove any sampled states)

(2) Create a list (or lists) to hold frequencies of states for different runs. For the {Rainy,Sunny} example, start each run in *S*. Now run 100 chains for 1 step. Record state frequencies across chains. Then run 100 chains for 2 steps. Record state frequencies. Then 5, then 10.

Transition Matrix

$$Q = \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix}$$

$$q_{ij} = P(X_{n+1} = j | X_n = i)$$

$$q_{11} = q_{RR} = 0.7$$

Q and ***q*** give us a sense for what will happen in the next step. But what about 2,3,4,…,100 steps in the future?

$$q_{ij}^{(100)} = ?$$

Transition Matrix

$$Q = \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix}$$

$$q_{ij} = P(X_{n+1} = j | X_n = i)$$

$$q_{11} = q_{RR} = 0.7$$

Q and ***q*** give us a sense for what will happen in the next step. But what about 2,3,4,…,100 steps in the future?

$$q_{ij}^{(100)} \neq \ (q_{ij})^{100}$$

Transition Matrix

$$P(X_{n+2} = R | X_n = R)$$

$$\boxed{R} \longrightarrow \boxed{?} \longrightarrow \boxed{R}$$

Transition Matrix

$$P(X_{n+2} = R | X_n = R)$$

$$\boxed{R} \longrightarrow \boxed{?} \longrightarrow \boxed{R}$$

$$\boxed{R} \longrightarrow \boxed{R} \longrightarrow \boxed{R}$$

$$\boxed{R} \longrightarrow \boxed{S} \longrightarrow \boxed{R}$$

Transition Matrix

$$P(X_{n+2} = R | X_n = R)$$

$$\boxed{R} \rightarrow \boxed{?} \rightarrow \boxed{R}$$

$$(q_{RR})^2 = \boxed{\begin{array}{c} \boxed{R} \rightarrow \boxed{R} \rightarrow \boxed{R} \\ \\ \boxed{R} \rightarrow \boxed{S} \rightarrow \boxed{R} \end{array}}$$

Transition Matrix

$$P(X_{n+3} = R | X_n = R)$$

$$R \rightarrow ? \rightarrow ? \rightarrow R$$

$$(q_{RR})^3 = \boxed{\begin{array}{c} R \rightarrow R \rightarrow R \rightarrow R \\ R \rightarrow R \rightarrow S \rightarrow R \\ R \rightarrow S \rightarrow R \rightarrow R \\ R \rightarrow S \rightarrow S \rightarrow R \end{array}}$$

Transition Matrix

$$P(X_{n+3} = R | X_n = R)$$

$$R \rightarrow ? \rightarrow ? \rightarrow R$$

$$q_{RR}^{(3)} = \boxed{\begin{array}{c} R \rightarrow R \rightarrow R \rightarrow R \\ R \rightarrow R \rightarrow S \rightarrow R \\ R \rightarrow S \rightarrow R \rightarrow R \\ R \rightarrow S \rightarrow S \rightarrow R \end{array}}$$

Transition Matrix

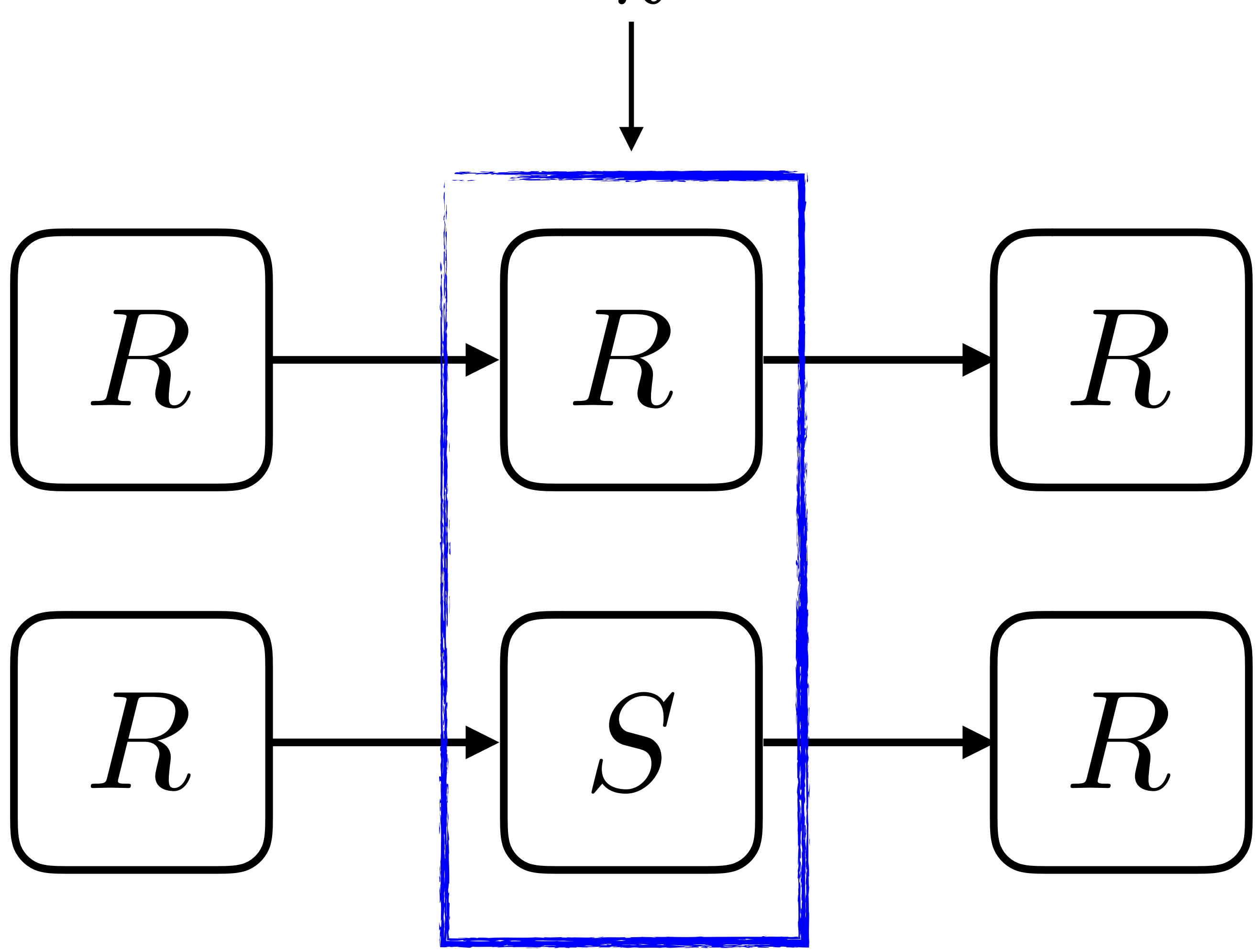
$q_{ij}^{(n)}$ is the (i, j) entry of Q^n

$$q_{ij}^{(2)} = \sum_k q_{ik} q_{kj}$$

Transition Matrix

$q_{ij}^{(n)}$ is the (i, j) entry of Q^n

$$q_{ij}^{(2)} = \sum_k q_{ik} q_{kj}$$



Stationary Distribution

A row vector $\mathbf{s} = (s_1, \ldots, s_M)$ such that $s_i \geq 0$ is a **stationary distribution** for a Markov chain with transition matrix Q if $\sum_i s_i q_{ij} = s_j$ for all j.

Equivalently, $\mathbf{s}Q = \mathbf{s}$.

Stationary Distribution

A row vector $\mathbf{s} = (s_1, \ldots, s_M)$ such that $s_i \geq 0$ is a **stationary distribution** for a Markov chain with transition matrix Q if $\sum_i s_i q_{ij} = s_j$ for all j.

Equivalently, $\mathbf{s}Q = \mathbf{s}$.

What does this mean??

Stationary Distribution

A row vector $\mathbf{s} = (s_1, \ldots, s_M)$ such that $s_i \geq 0$ is a **stationary distribution** for a Markov chain with transition matrix Q if $\sum_i s_i q_{ij} = s_j$ for all j.

Equivalently, $\mathbf{s}Q = \mathbf{s}$.

Simply put, if the states have certain probabilities of appearing in one iteration, they will have those same probabilities in the next iteration. The **probabilities do not change**!

Stationary Distribution

Using the Markov chain simulator you wrote above, try using this Q matrix:

$$\begin{array}{cc} & \begin{array}{cc} R & S \end{array} \\ \begin{array}{c} R \\ S \end{array} & \begin{pmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{pmatrix} \end{array}$$

Start 100 chains in *R*. Now calculate the frequencies of *R* and *S* at regular intervals for 100 generations.

What do you notice? How does this compare to the transition matrix you were using before? What are the stationary frequencies in each case?

Reversibility

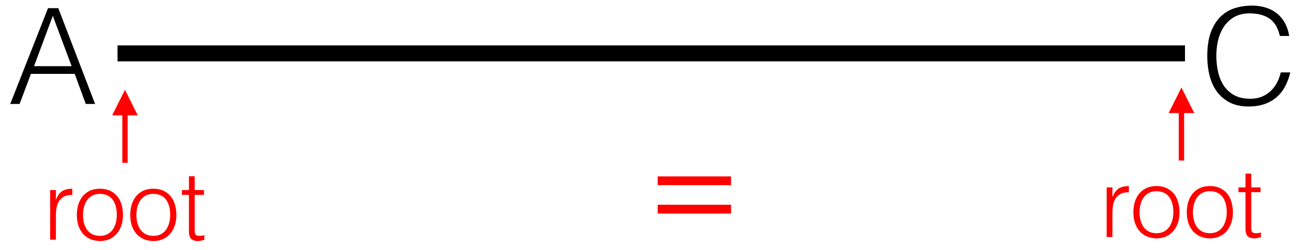
$$s_i q_{ij} = s_j q_{ji}$$

Detailed Balance Equation

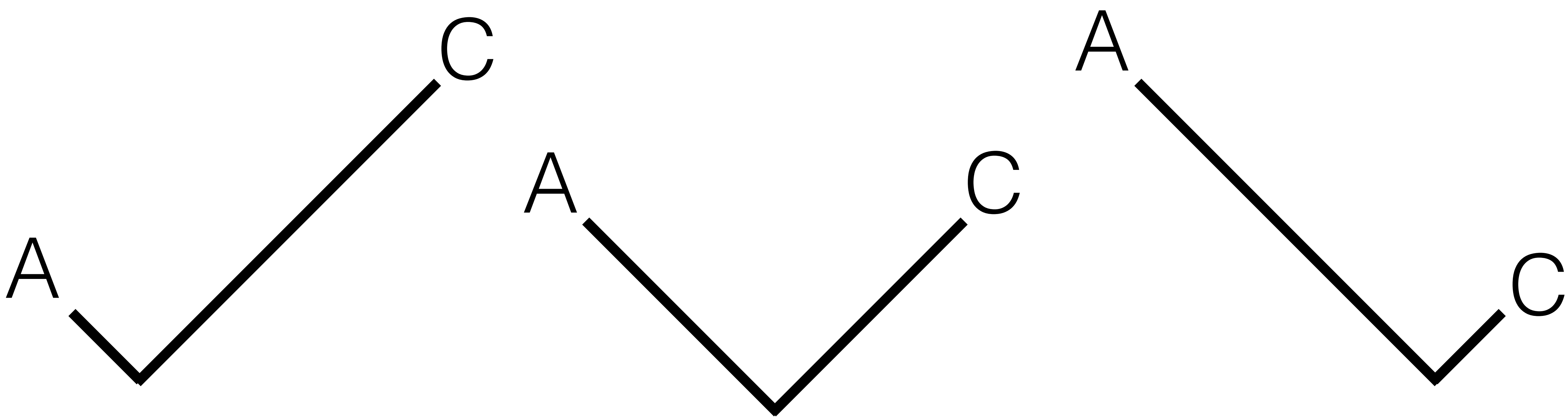
The probability of a series of states in the chain is the same forward, as it is in reverse? Why might this be important for phylogenetics?

The Pulley Principle

The Pulley Principle

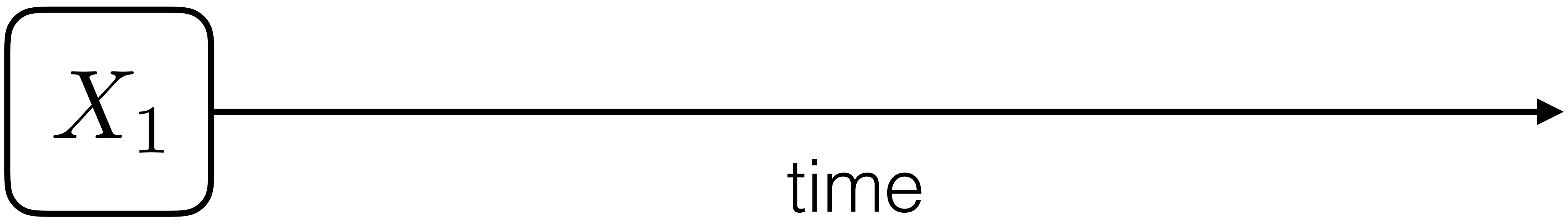


The Pulley Principle



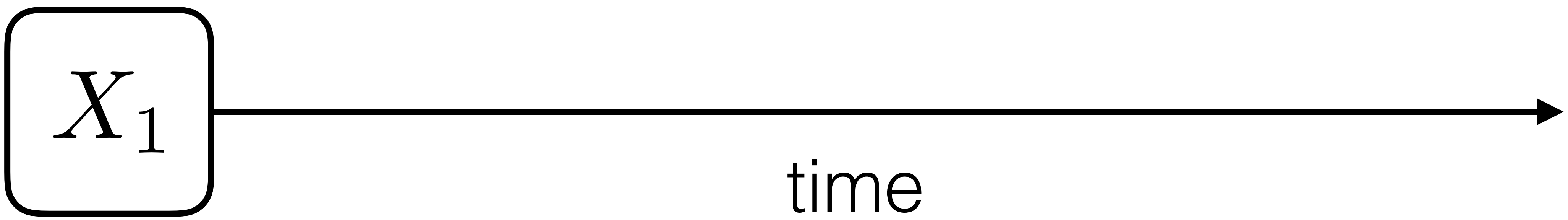
If the chain is reversible, the likelihood will be the same no matter where we put our "root". This is known as the pulley principle (Felsenstein 1981).

Continuous Time Markov Chains



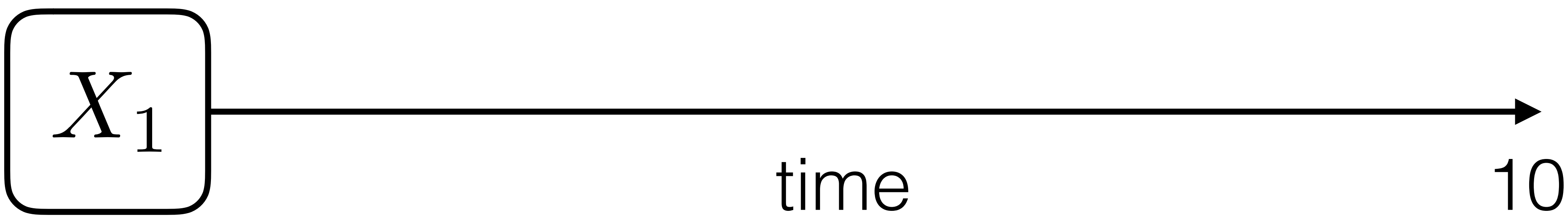
For a continuous-time Markov chain, time does not proceed in iterations. Rather, it is a continuous variable and state changes can occur at any point. Think of a discrete-time chain where the time between iterations is very short (too short to discern) and we go through many of them.

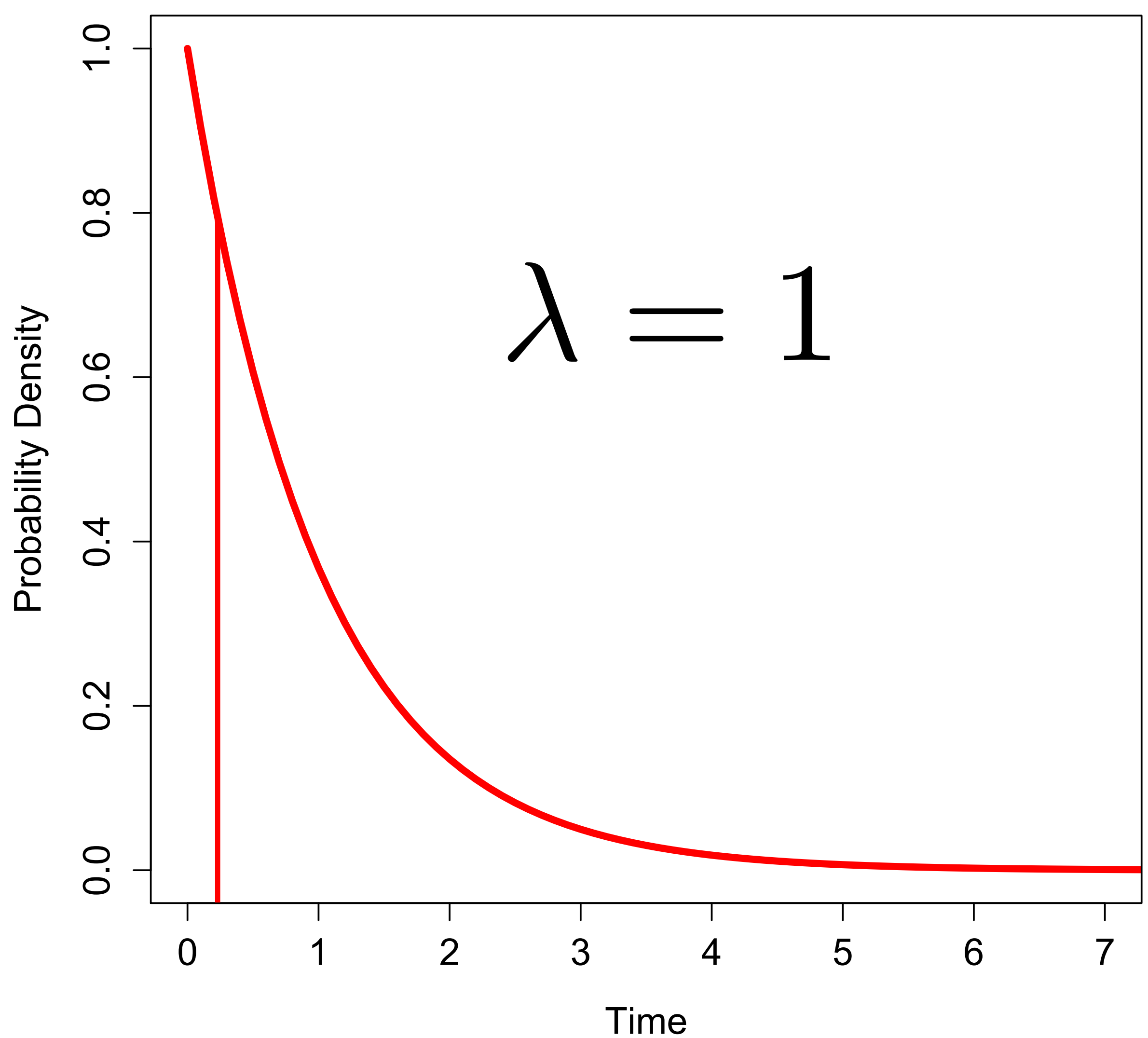
Continuous Time Markov Chains



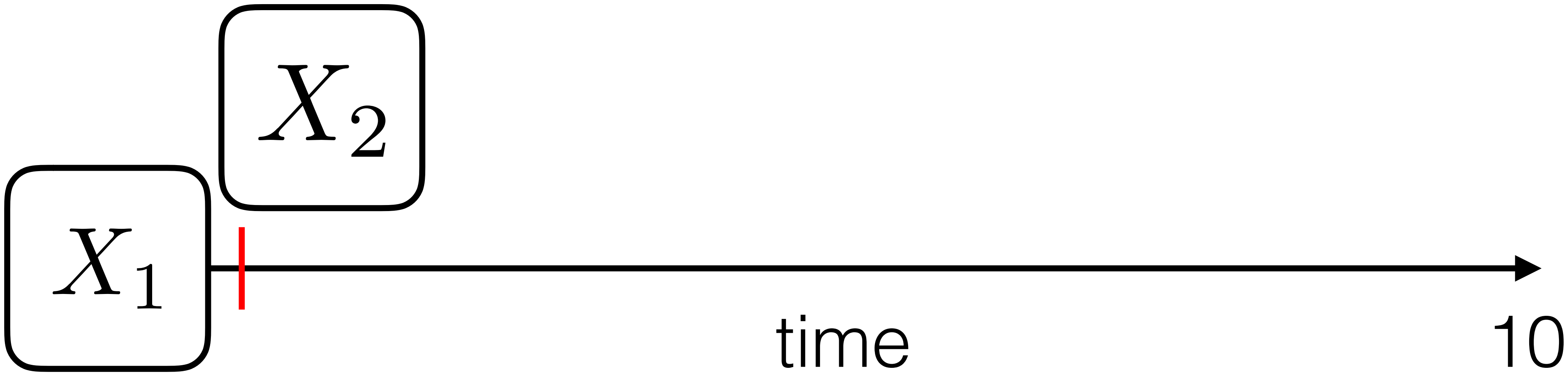
The waiting times between events (state changes) in a continuous-time Markov chain are exponentially distributed. The rate parameter (λ) determines how frequently those events occur.

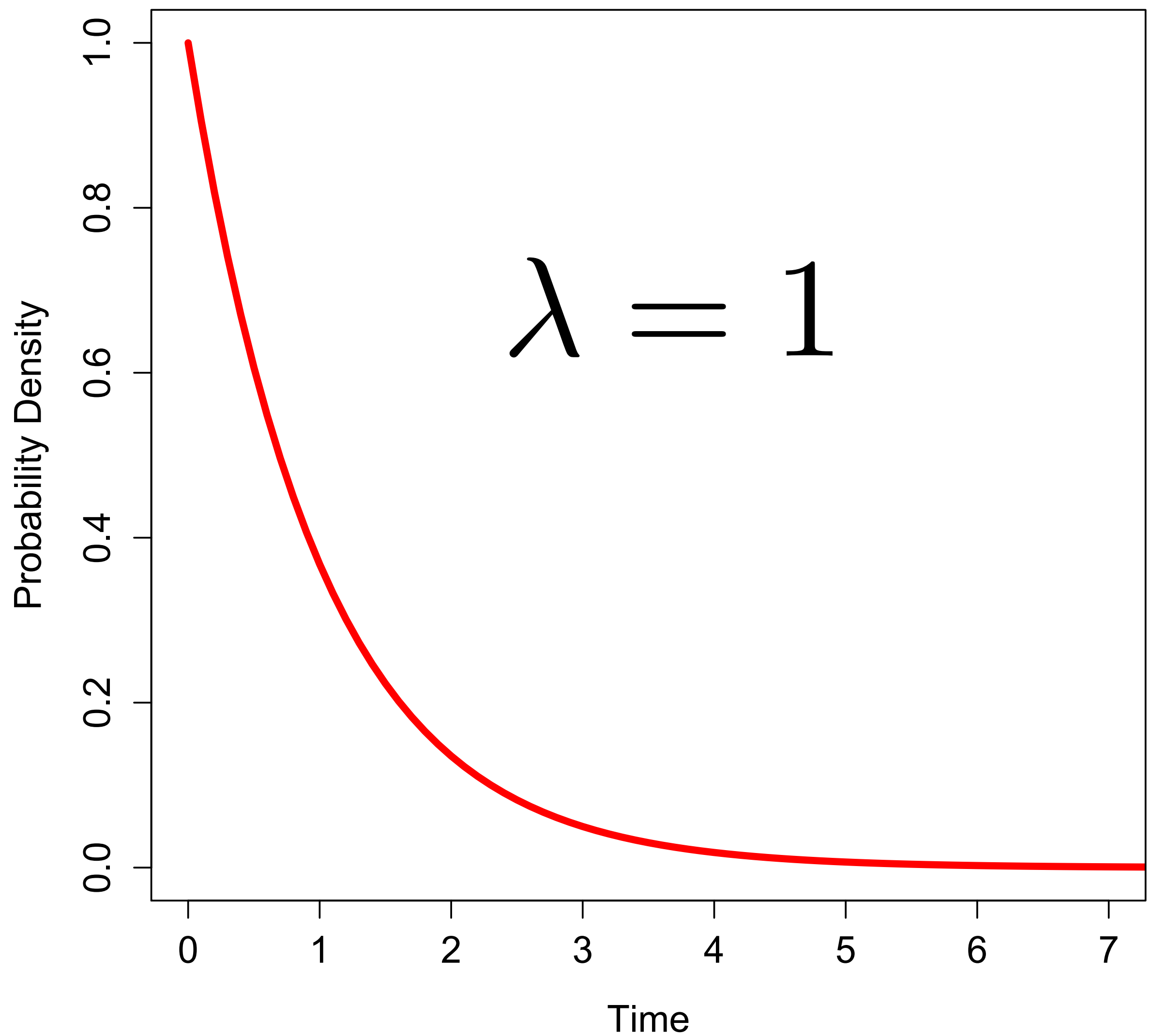
Continuous Time Markov Chains



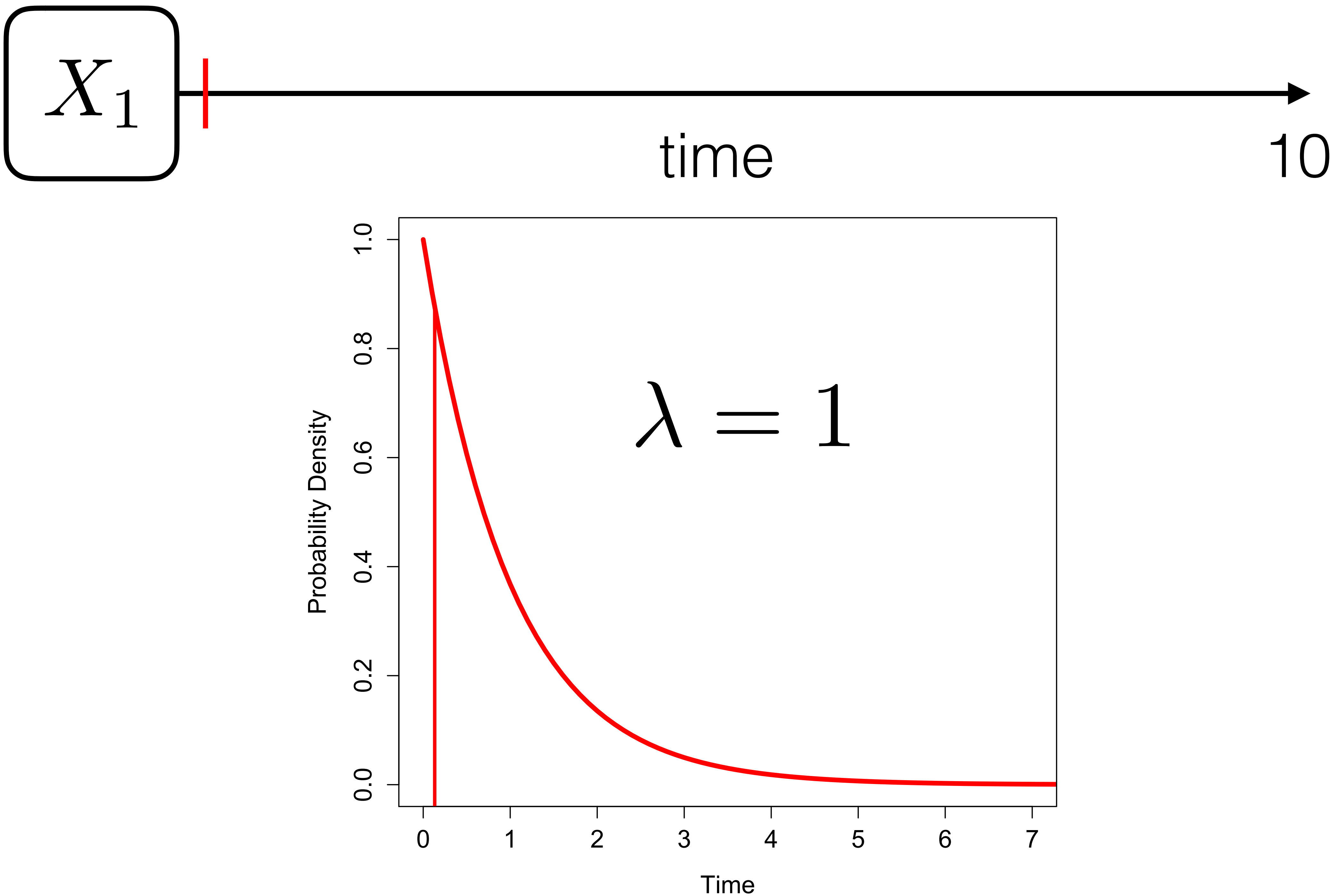


Continuous Time Markov Chains

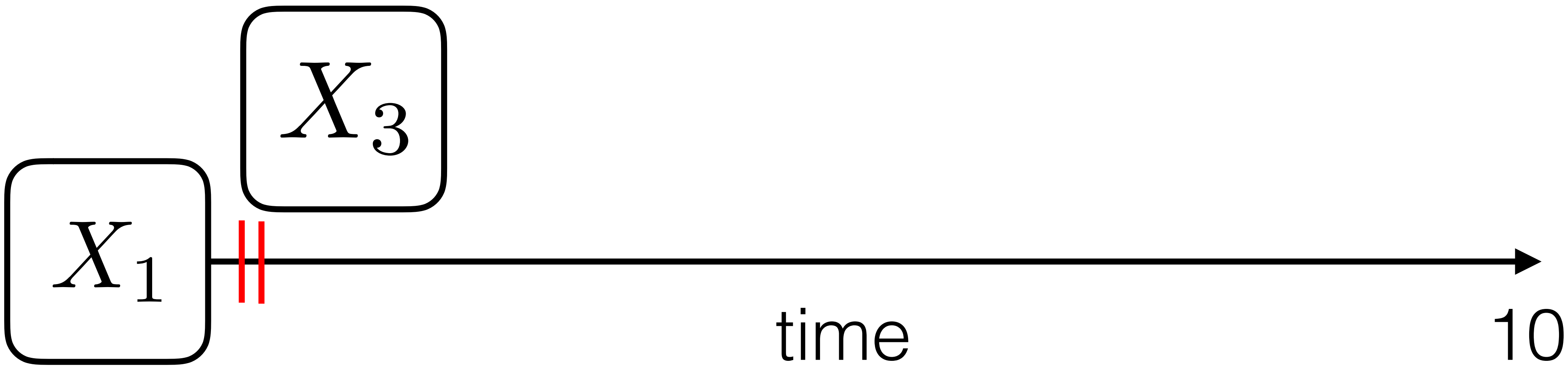


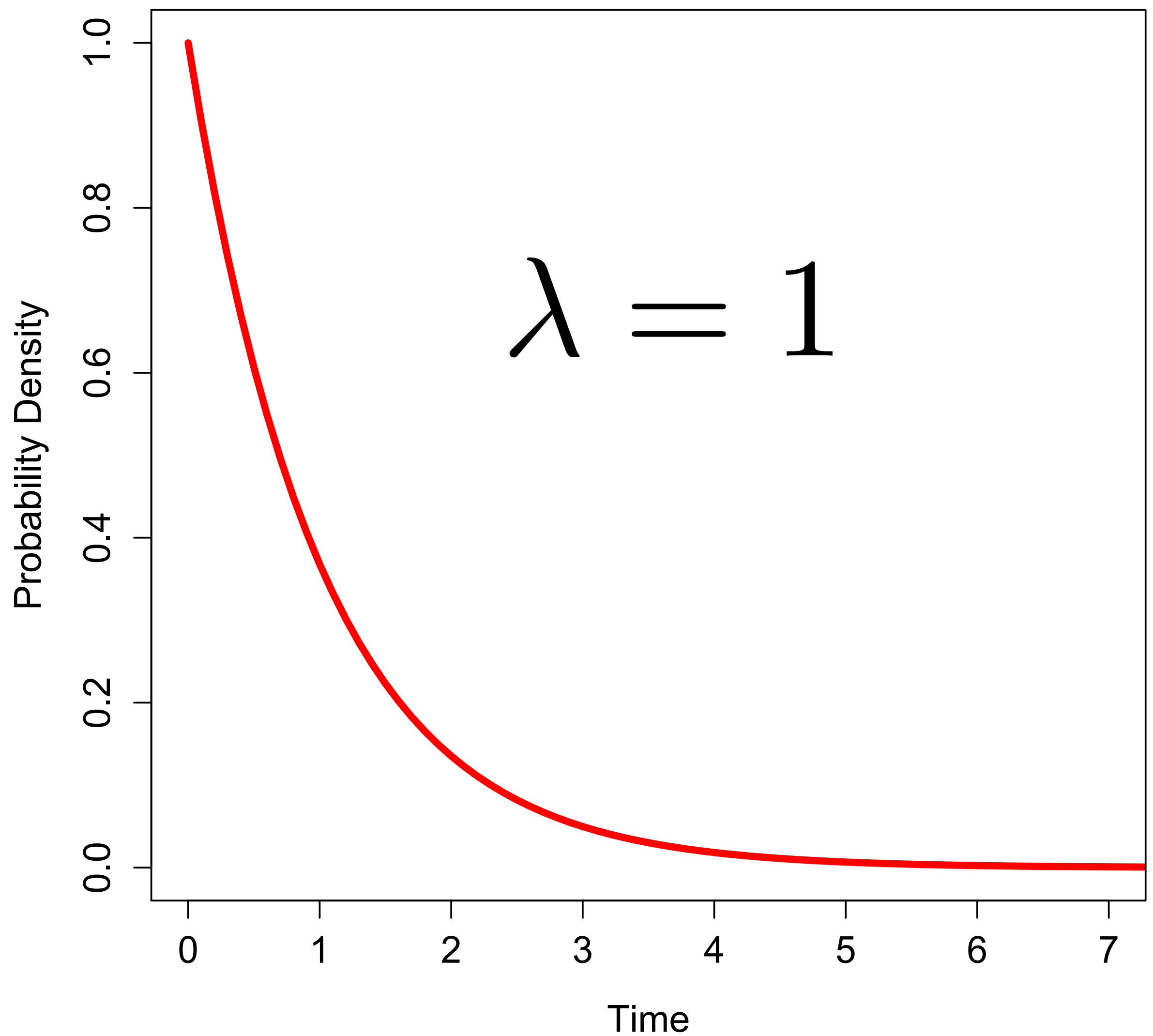


Continuous Time Markov Chains

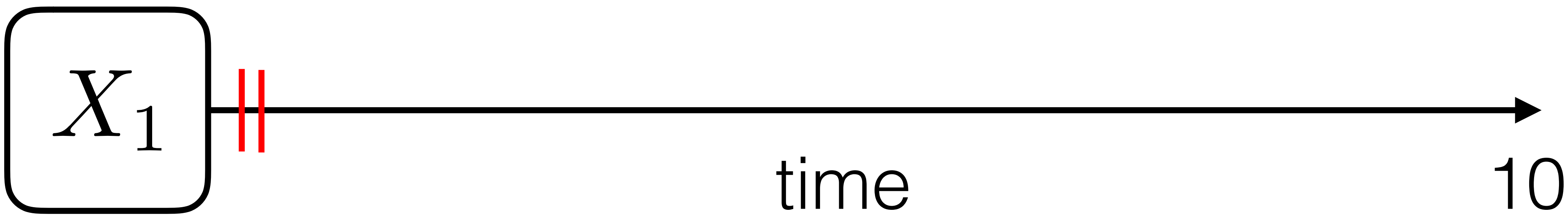


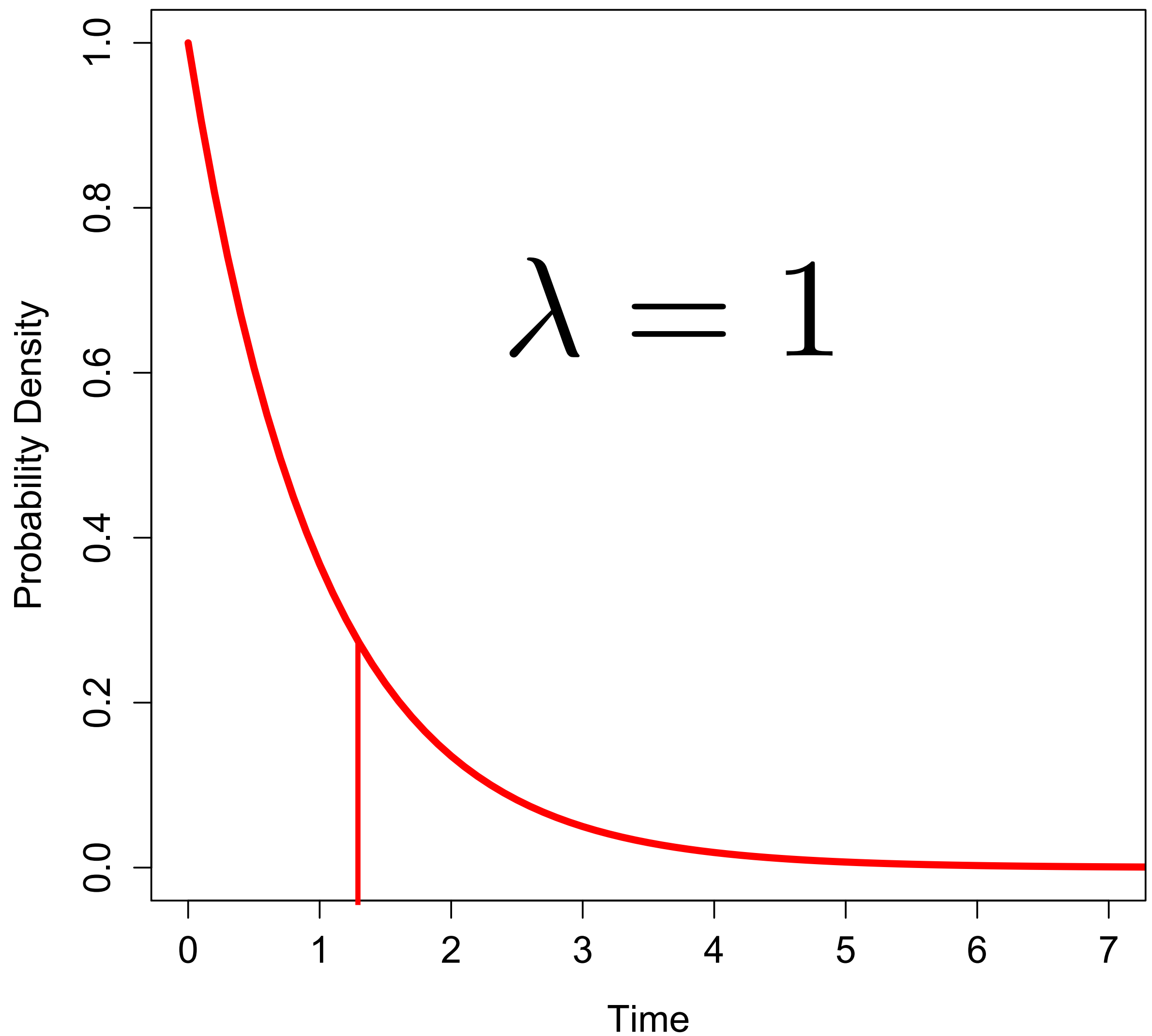
Continuous Time Markov Chains



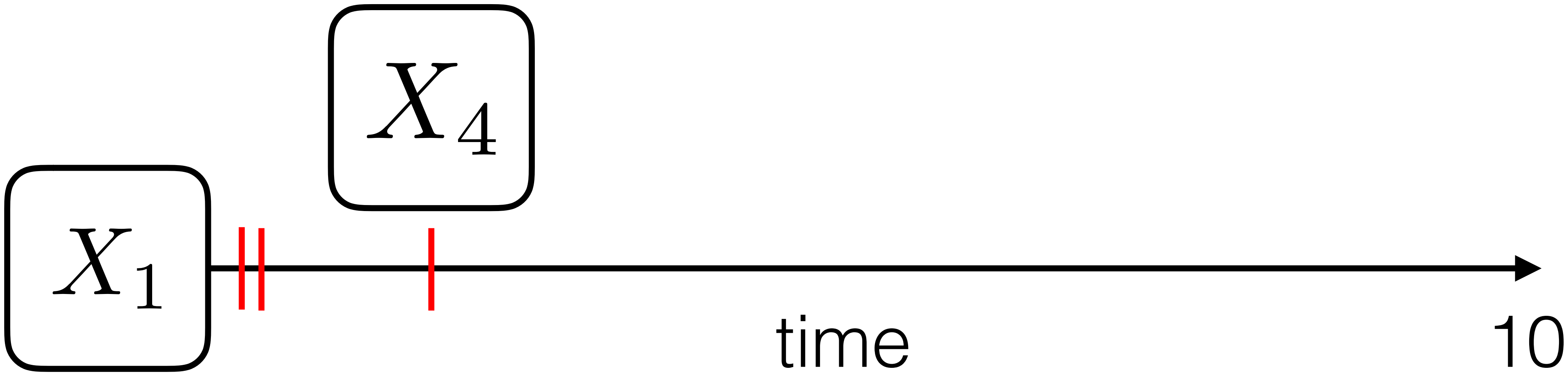


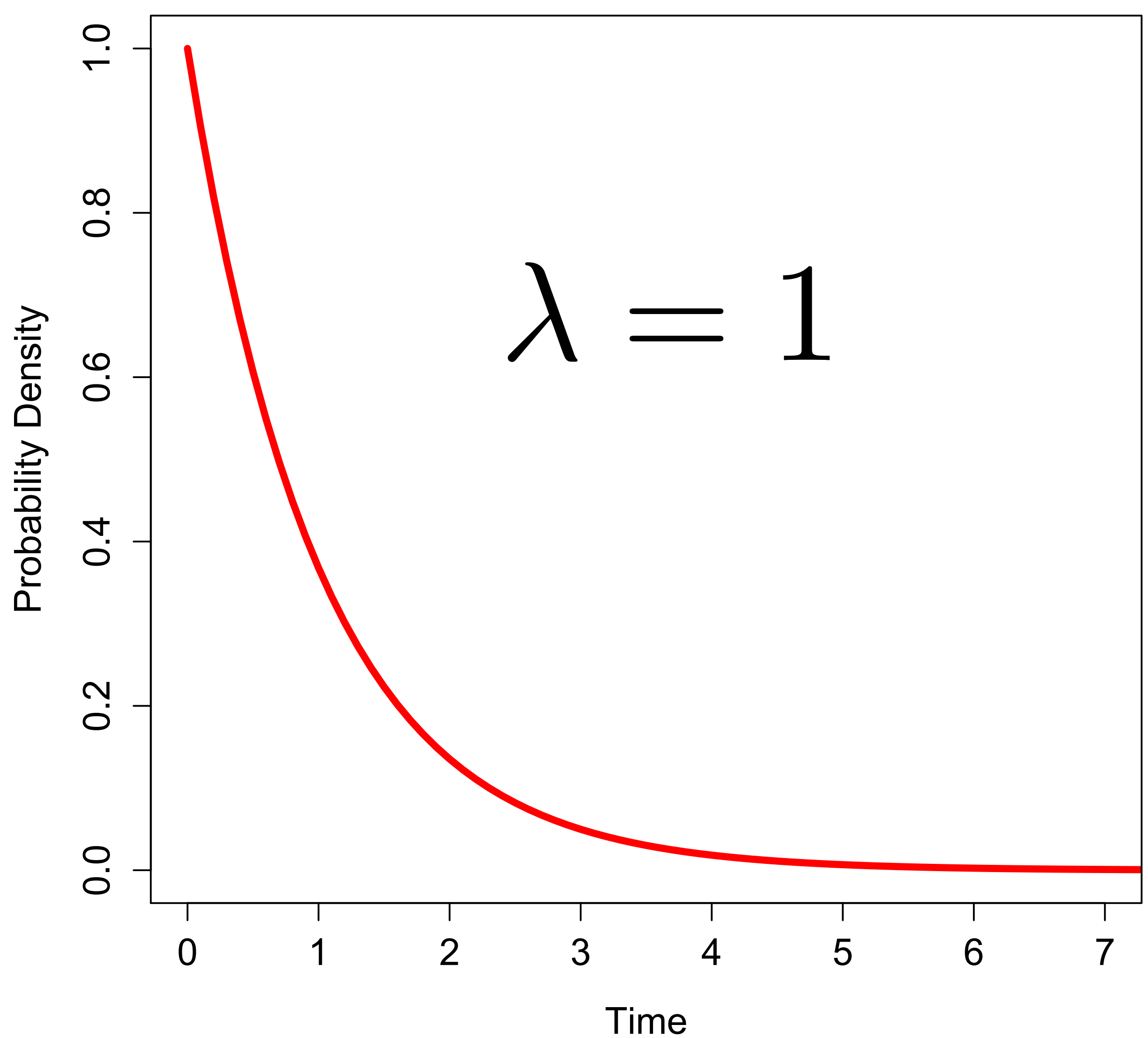
Continuous Time Markov Chains



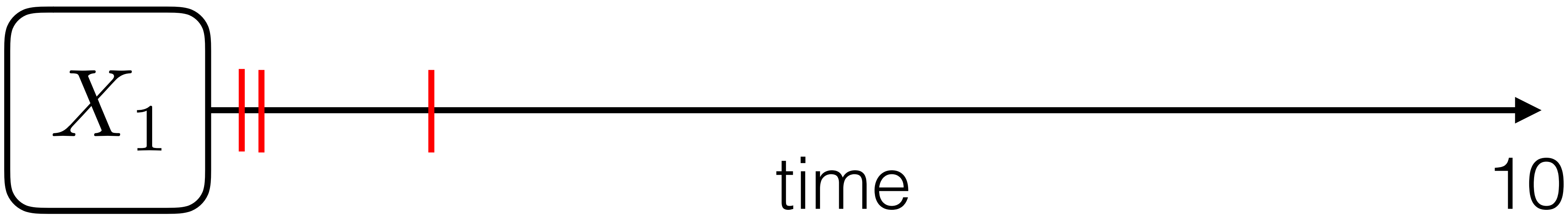


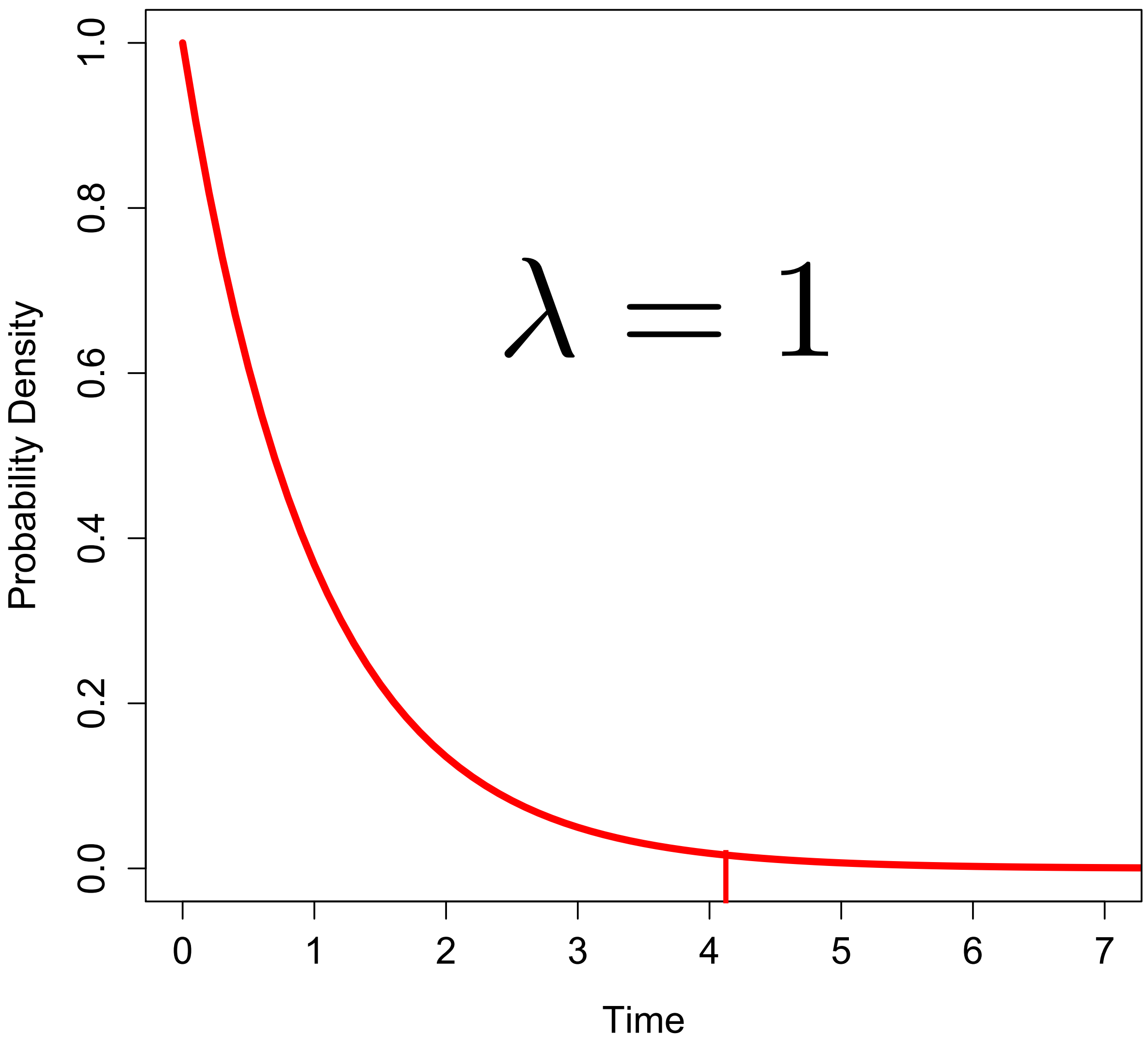
Continuous Time Markov Chains



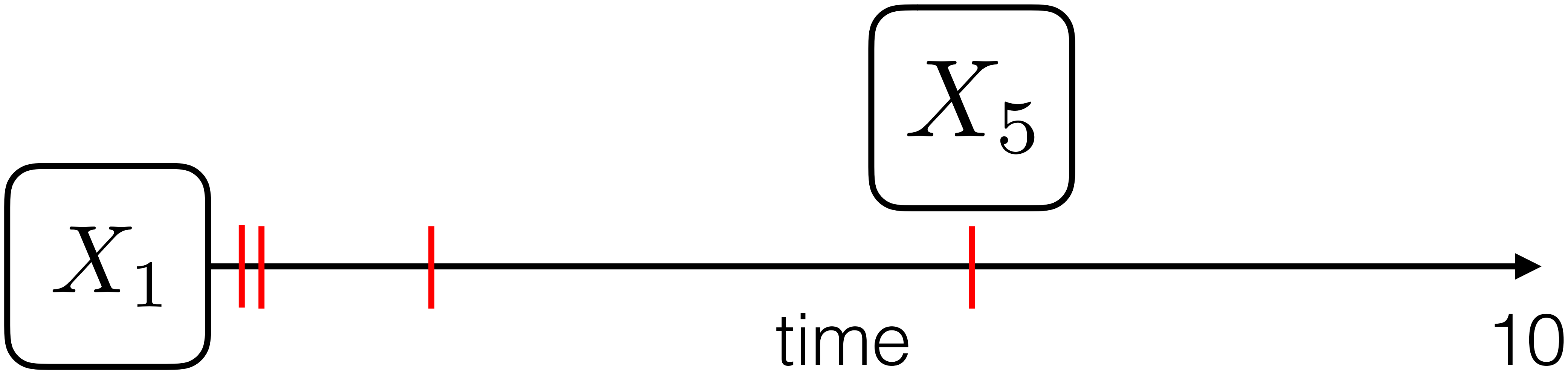


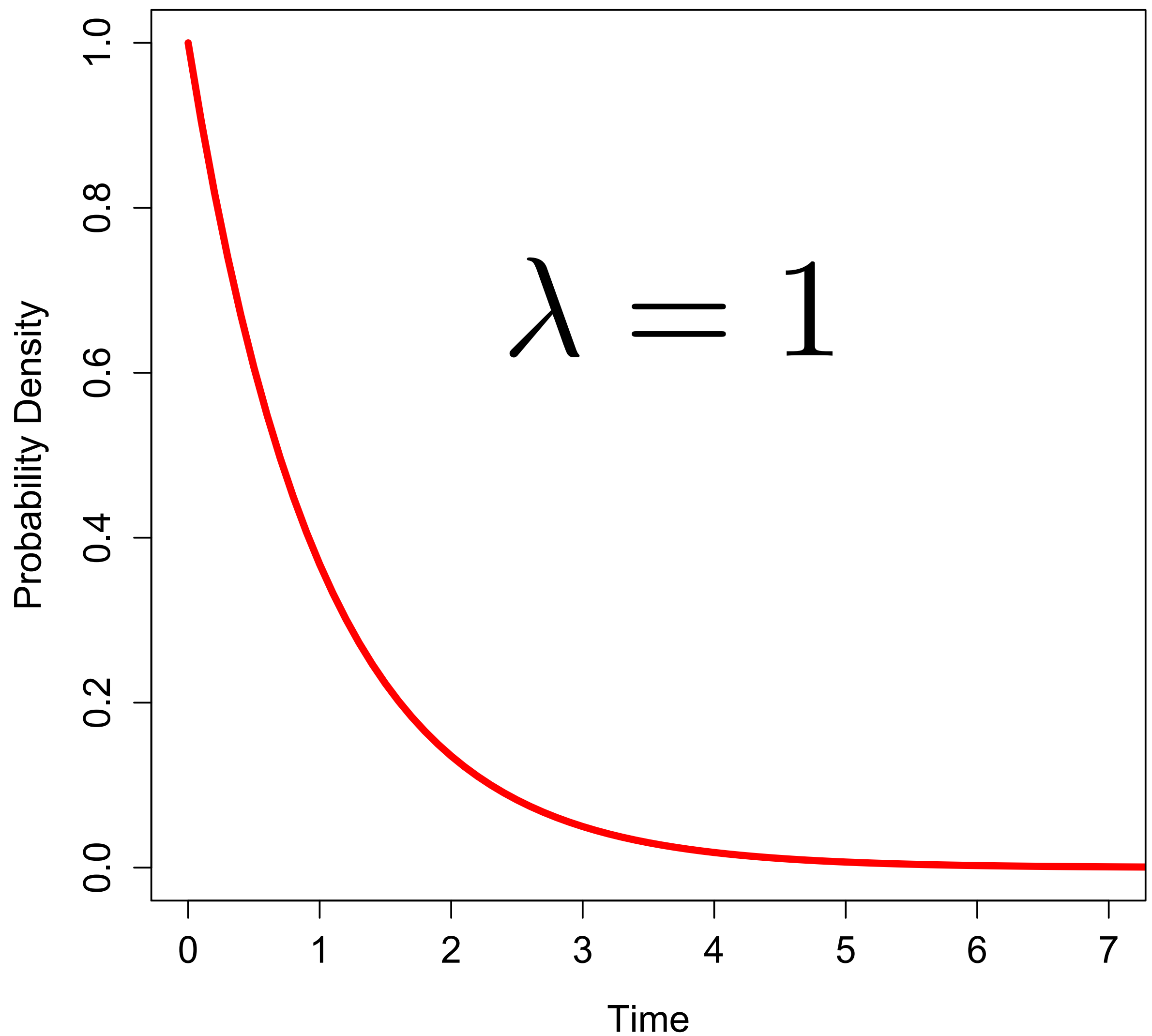
Continuous Time Markov Chains



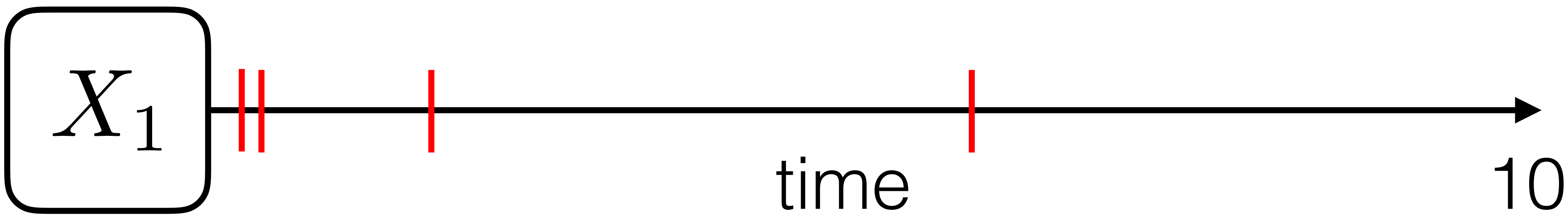


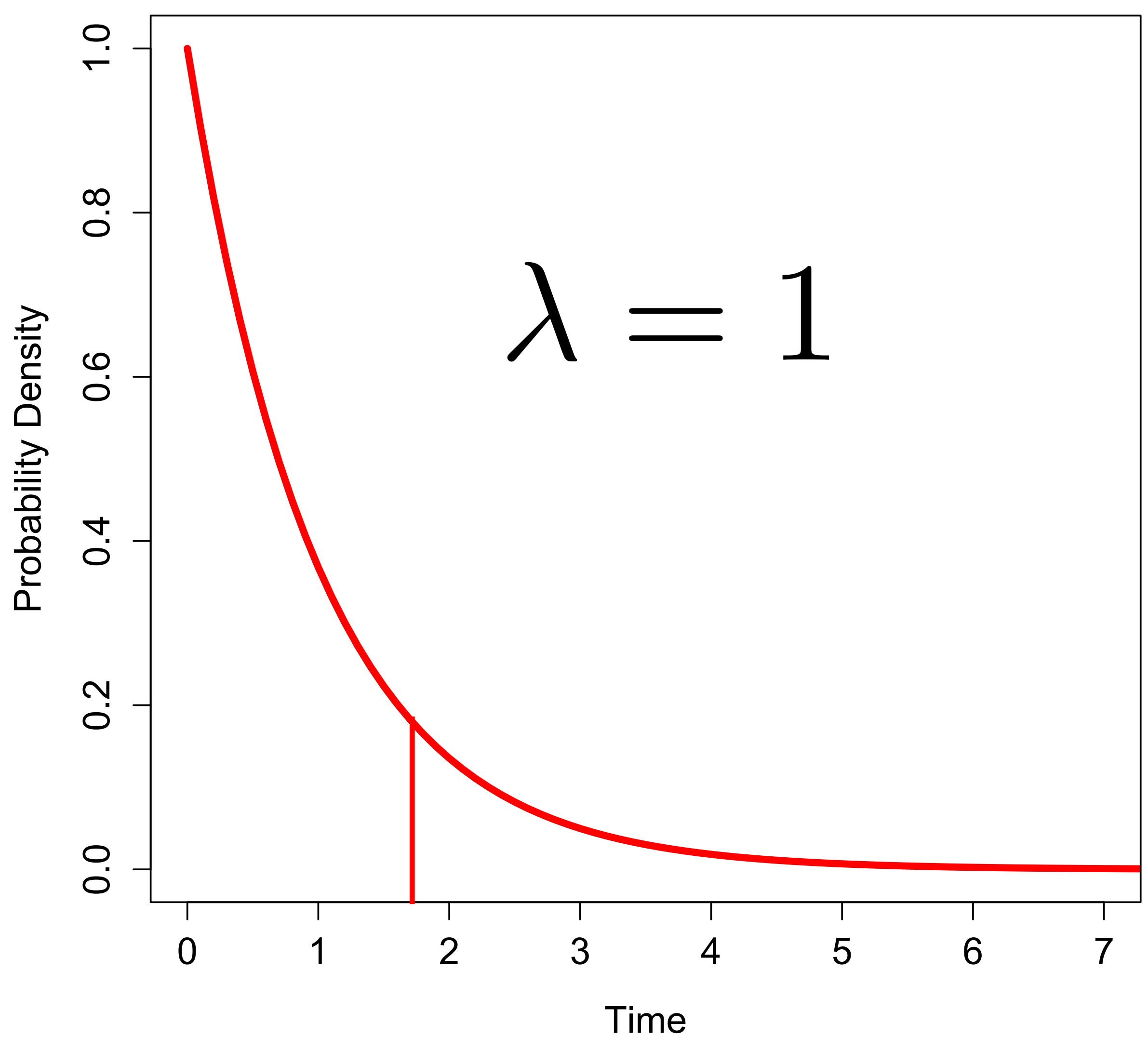
Continuous Time Markov Chains



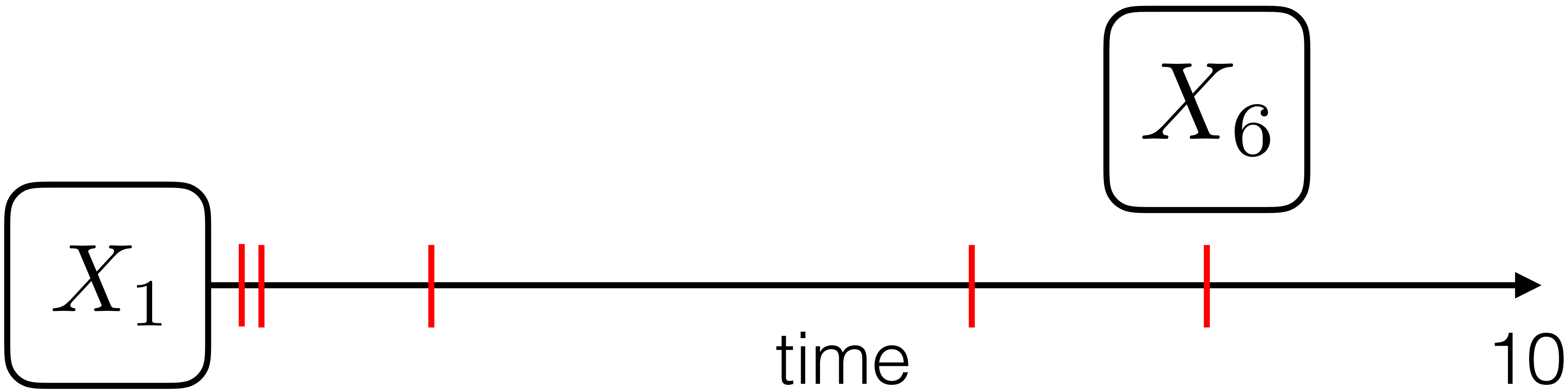


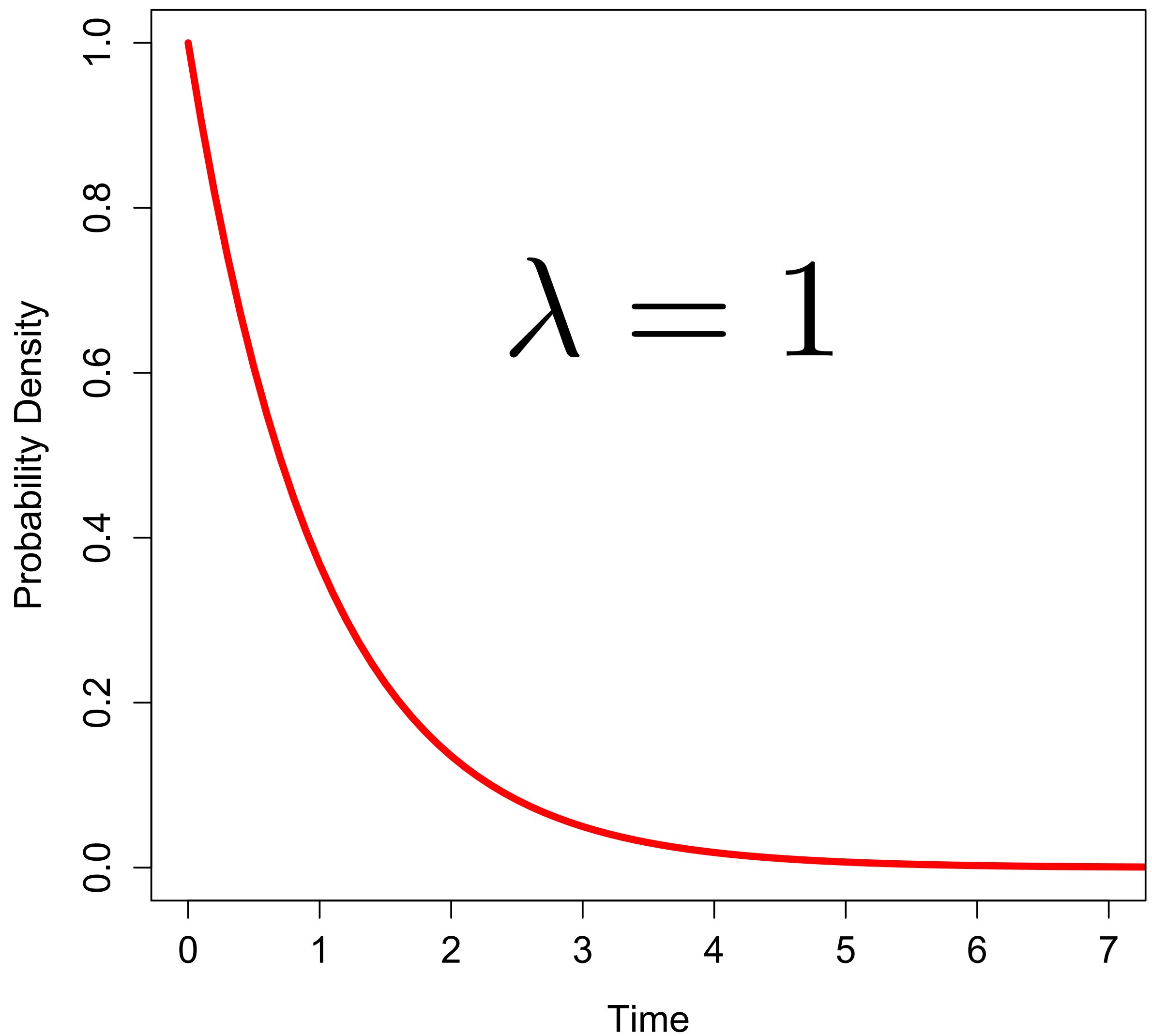
Continuous Time Markov Chains



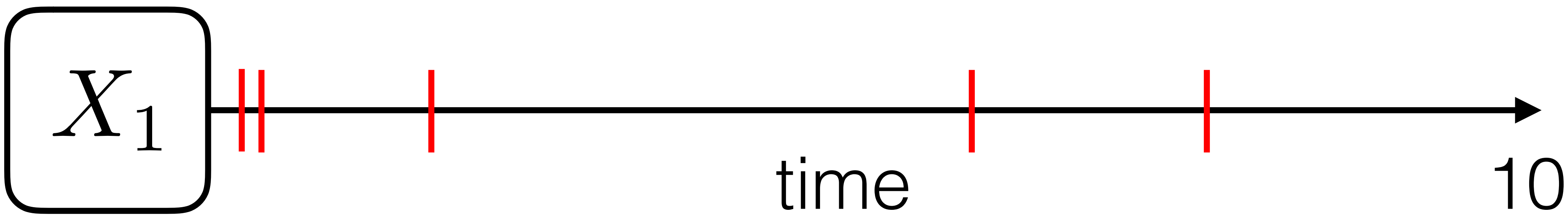


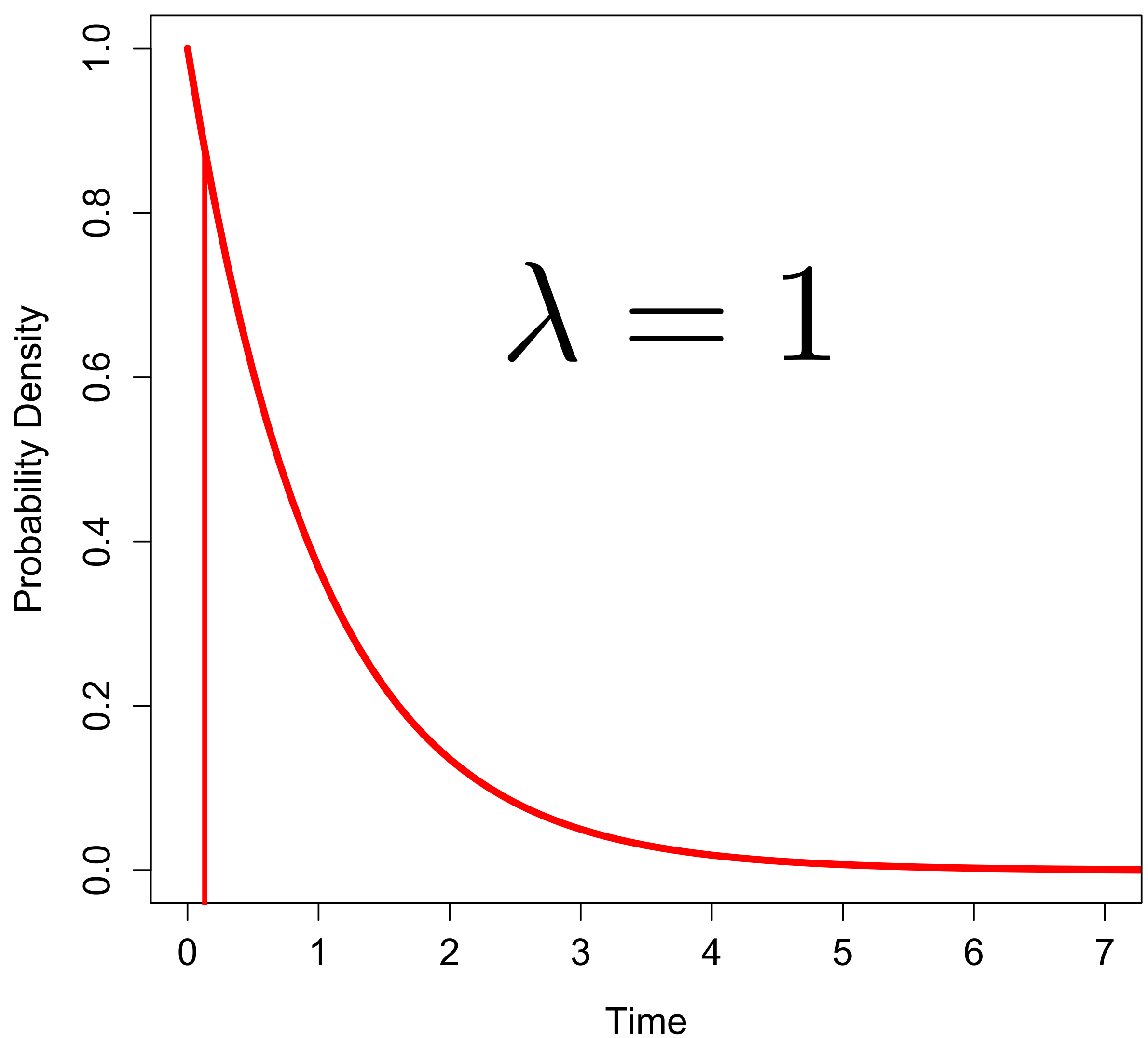
Continuous Time Markov Chains



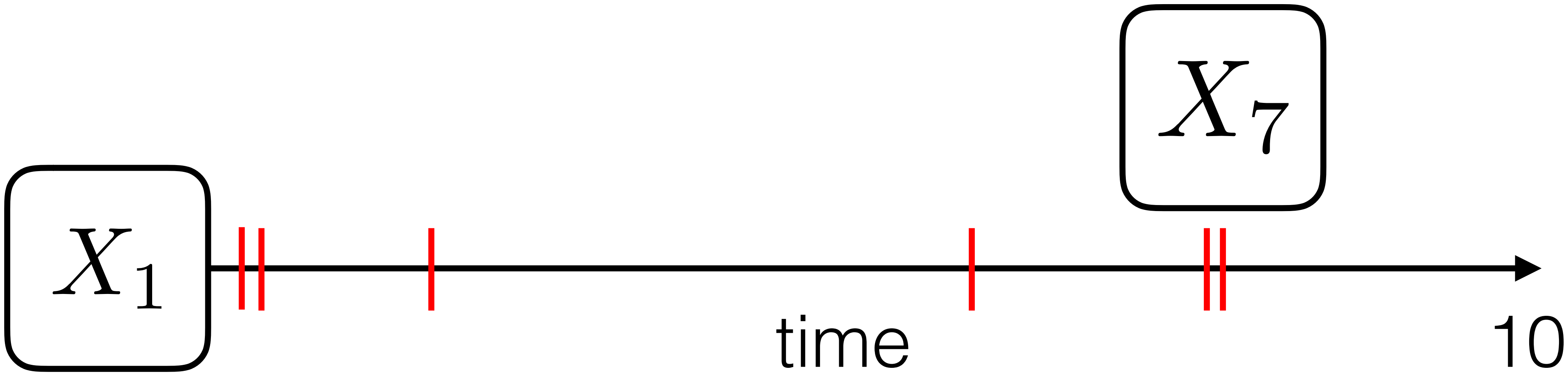


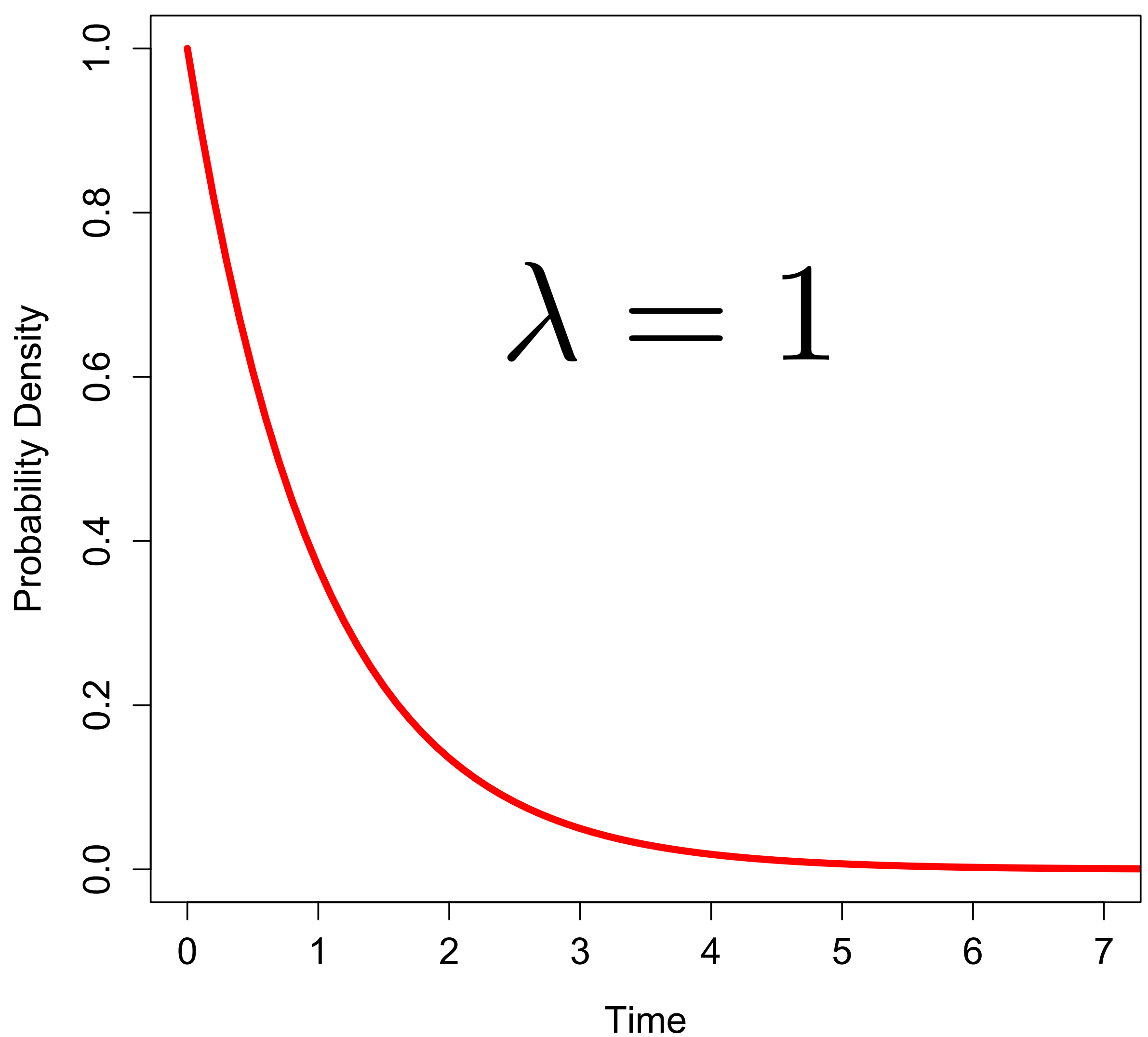
Continuous Time Markov Chains



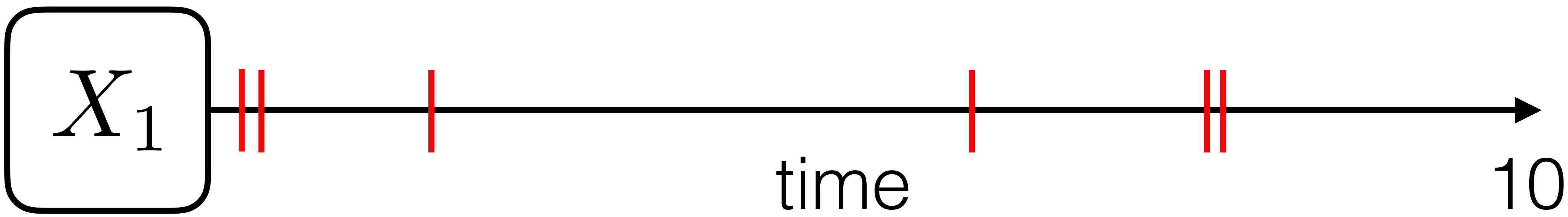


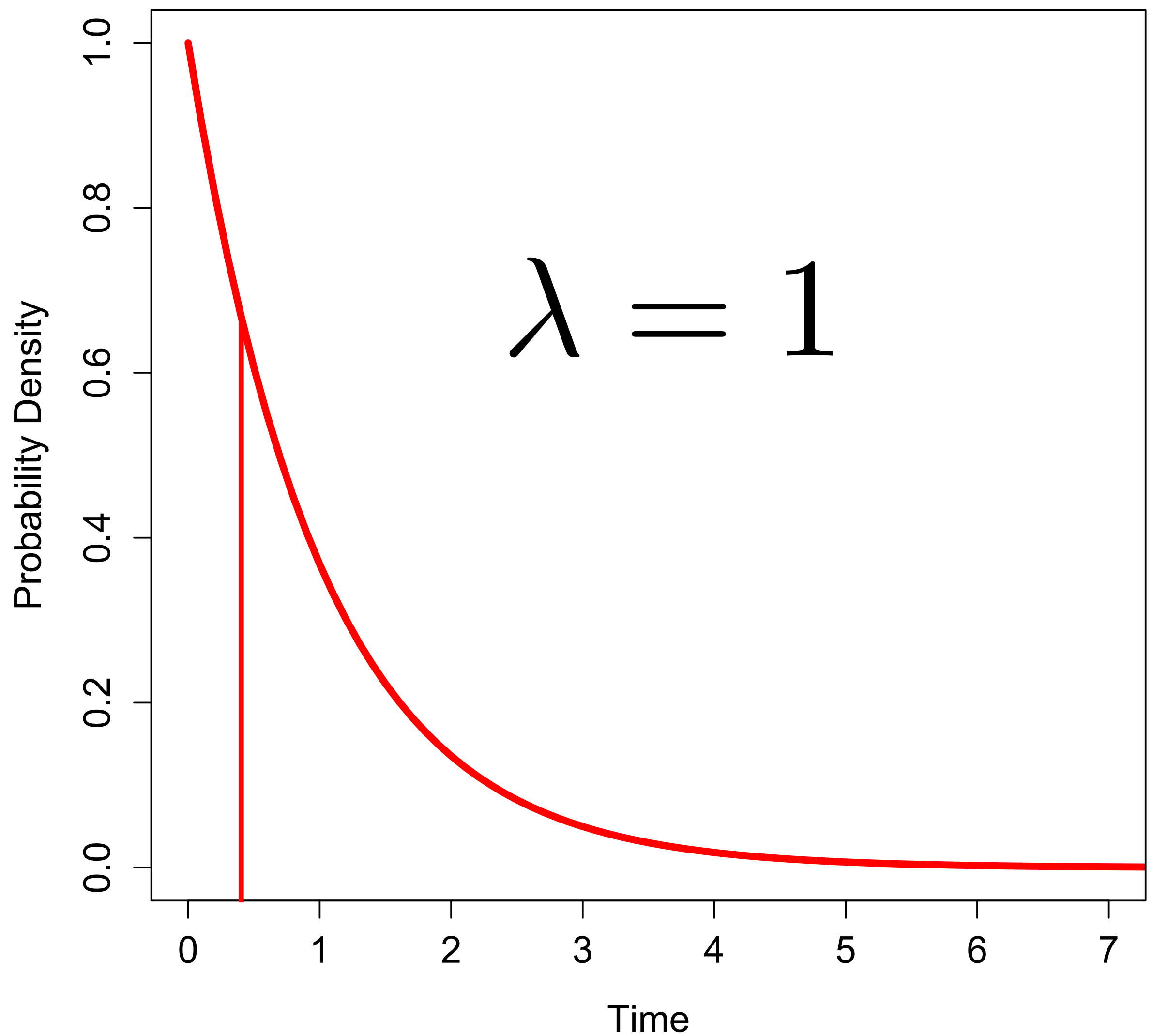
Continuous Time Markov Chains



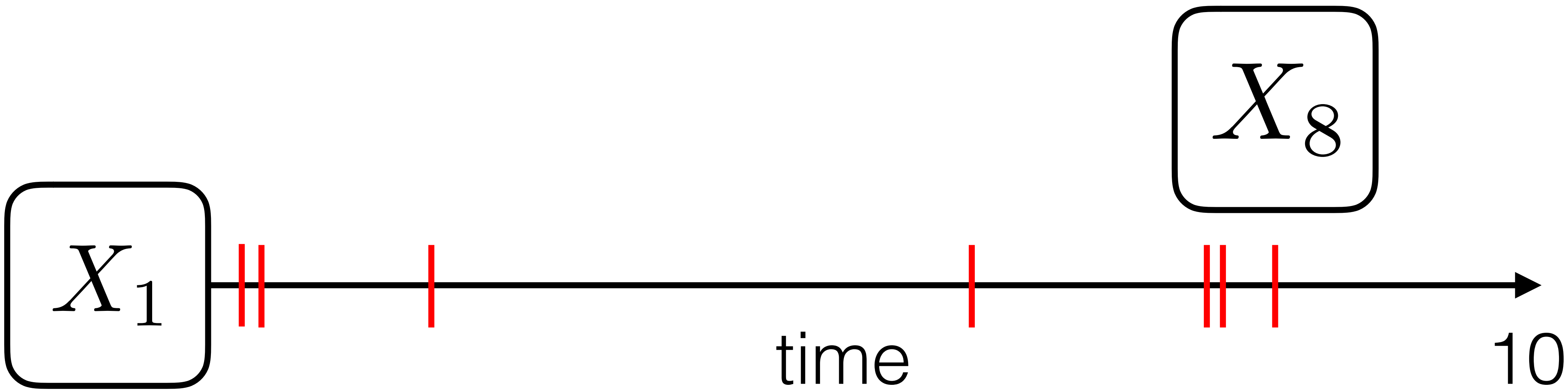


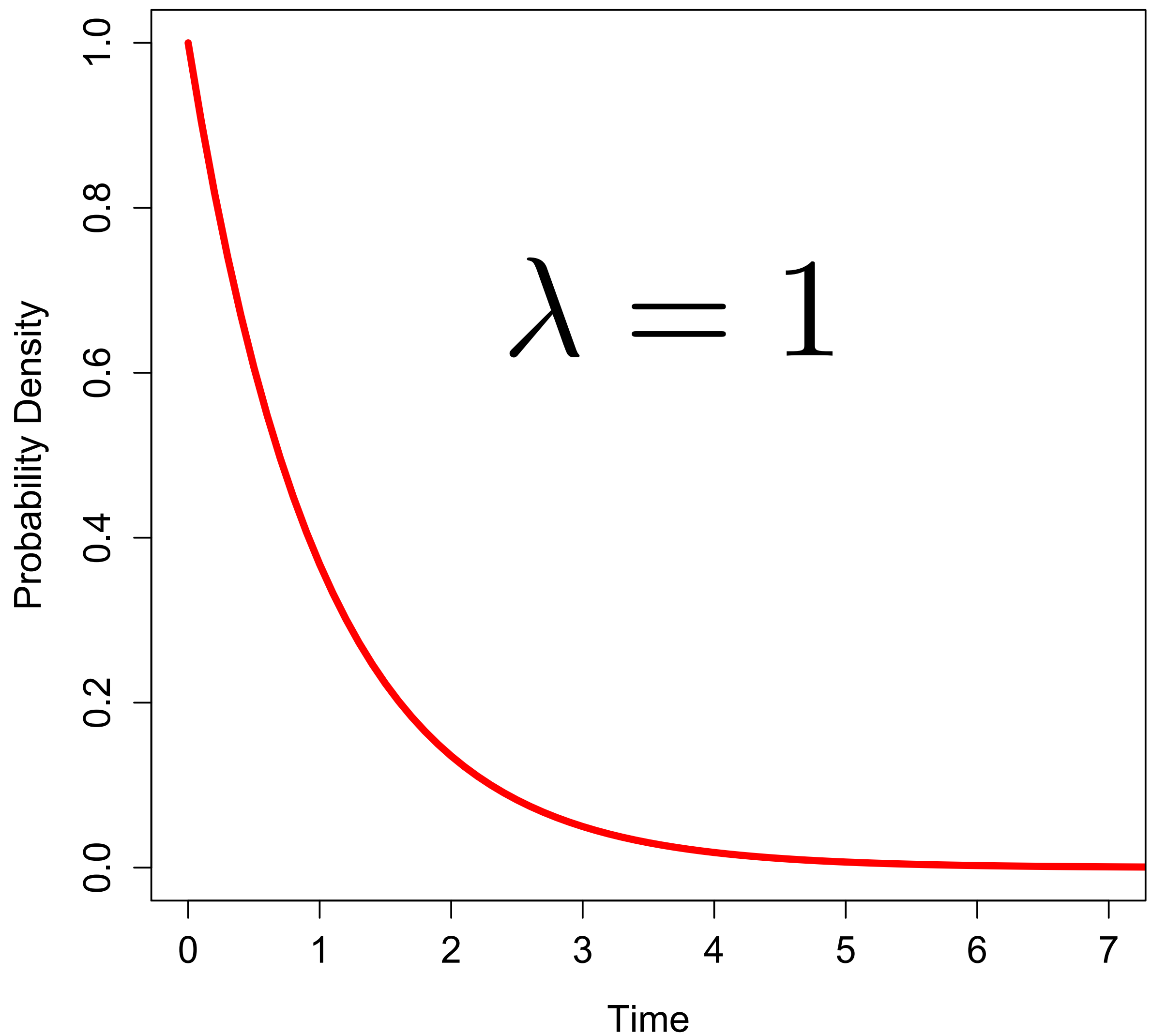
Continuous Time Markov Chains



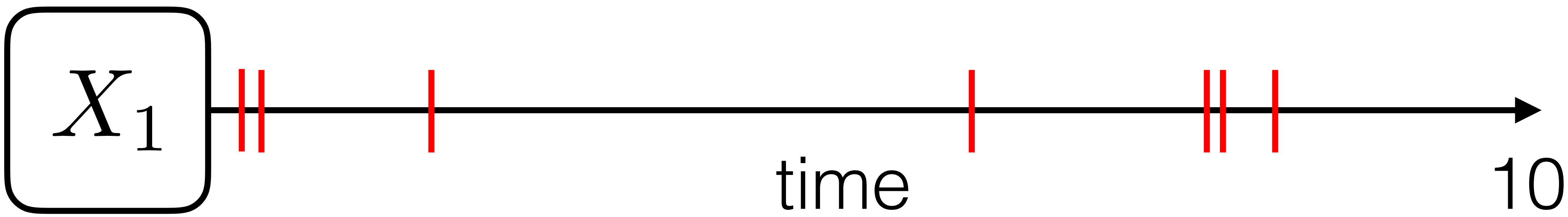


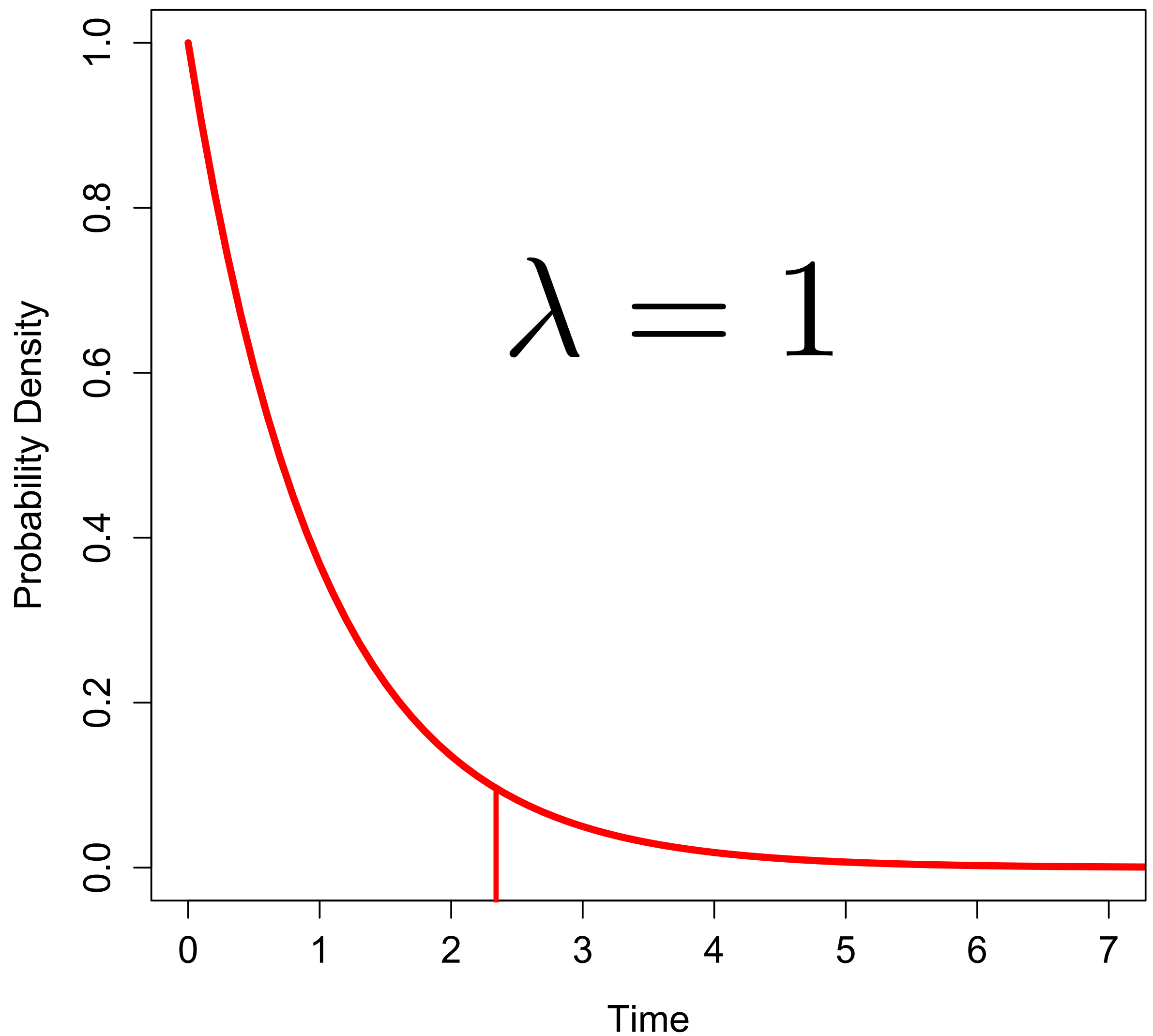
Continuous Time Markov Chains



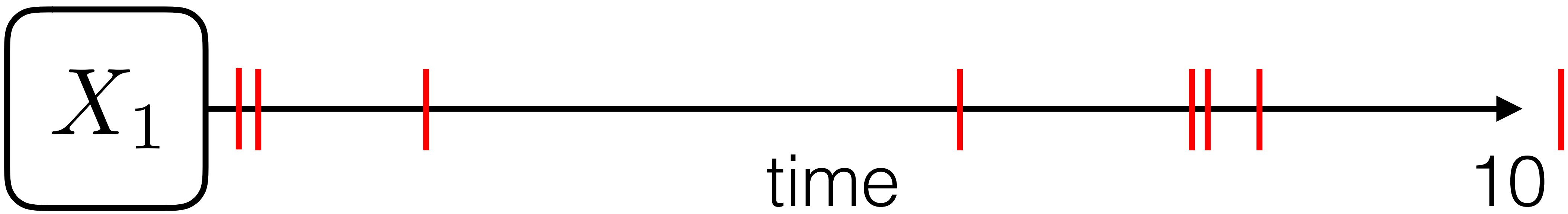


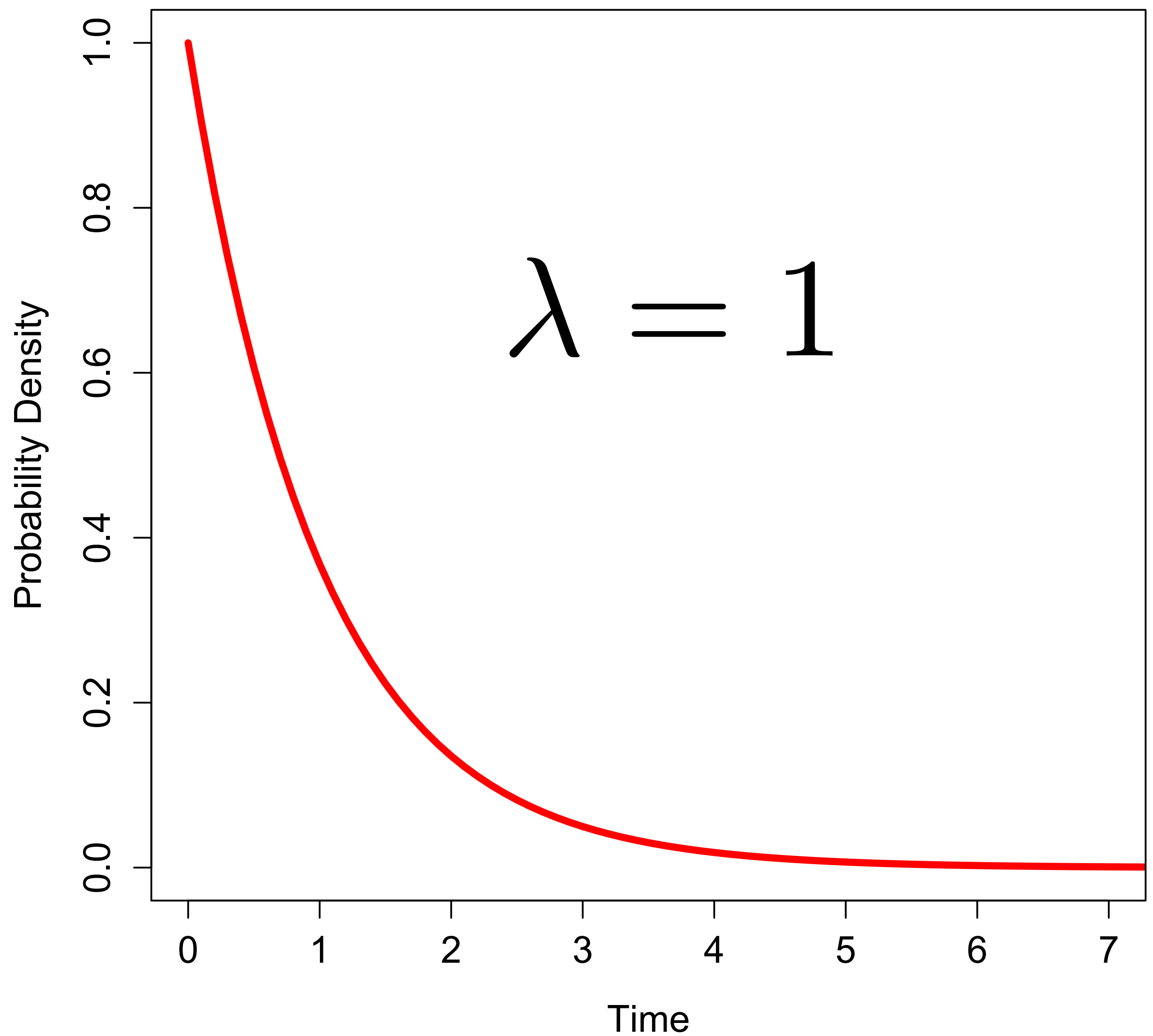
Continuous Time Markov Chains





Continuous Time Markov Chains





CTMC Transition Matrix

$$
\begin{array}{c c}
 & \begin{array}{cccc} A & C & G & T \end{array} \\
\begin{array}{c} A \\ C \\ G \\ T \end{array} & \begin{pmatrix} -1 & 1/3 & 1/3 & 1/3 \\ 1/3 & -1 & 1/3 & 1/3 \\ 1/3 & 1/3 & -1 & 1/3 \\ 1/3 & 1/3 & 1/3 & -1 \end{pmatrix}
\end{array}
$$

Jukes and Cantor (1969)

CTMC Transition Matrix

$$
\begin{array}{cc}
 & \begin{array}{cccc} A & C & G & T \end{array} \\
\begin{array}{c} A \\ C \\ G \\ T \end{array} &
\begin{pmatrix}
\boxed{-1} & 1/3 & 1/3 & 1/3 \\
1/3 & -1 & 1/3 & 1/3 \\
1/3 & 1/3 & -1 & 1/3 \\
1/3 & 1/3 & 1/3 & -1
\end{pmatrix}
\end{array}
$$

Jukes and Cantor (1969)

CTMC Transition Matrix

Exponential Rate
for State A

$$
\begin{array}{cc}
 & \begin{array}{cccc} A & C & G & T \end{array} \\
\begin{array}{c} A \\ C \\ G \\ T \end{array} & \begin{pmatrix} \boxed{-1} & 1/3 & 1/3 & 1/3 \\ 1/3 & -1 & 1/3 & 1/3 \\ 1/3 & 1/3 & -1 & 1/3 \\ 1/3 & 1/3 & 1/3 & -1 \end{pmatrix}
\end{array}
$$

Jukes and Cantor (1969)

CTMC Transition Matrix

Relative Probabilities of Transition

$$\begin{array}{c} \\ A \\ C \\ G \\ T \end{array} \begin{array}{c} \begin{array}{cccc} A & C & G & T \end{array} \\ \begin{pmatrix} -1 & \boxed{1/3 \quad 1/3 \quad 1/3} \\ 1/3 & -1 & 1/3 & 1/3 \\ 1/3 & 1/3 & -1 & 1/3 \\ 1/3 & 1/3 & 1/3 & -1 \end{pmatrix} \end{array}$$

Jukes and Cantor (1969)