Applied Probability
Concepts & Examples

Johar M. Ashfaque

1   *What is a Random Variable?*

Let X be a function from the sample space Ω into the set of reals $\mathbf{R}$. In order for X to be RV, some extra requirements are needed. What are they?

The necessary and sufficient condition is: for each $x \in \mathbf{R}$ and $\omega \in \Omega$ (an element in the set of all the possible outcomes for the named event),

$$\mathcal{P}(x) := \{\omega \in \Omega : X(\omega) \leq x\}$$

must be an **event** such that the probability $P \in \mathbf{R}_{\geq 0}$ that is the probability of the event taking place is well-defined on $\mathcal{P}(x)$.

The point here is that for any function $X : \Omega \to \mathbf{R}$, the set $\mathcal{P}(x)$ is not necessarily an event (in this case, such an X is not a RV). Only for that function X which guarantees $\mathcal{P}(x)$ as an event for any $x \in \mathbf{R}$, then X is a RV.

There are two types of random variables: discrete and continuous. Where discrete random variables can only take particular values or specified values in $\mathbf{R}$, continuous random variables can take any value in $\mathbf{R}$.

A random variable has either an associated probability distribution (discrete random variable) or probability density function (continuous random variable).

1.1   *Summary*

Ω: The Sample Space (the set of all outcomes for the named event).
ω: An element of the sample space.
$\mathbf{R}$: The set of real numbers.
X: A random variable if and only if for each $x \in \mathbf{R}$ and $\omega \in \Omega$

$$\mathcal{P}(x) := \{\omega \in \Omega : X(\omega) \leq x\}$$

is an event with probability $P \in \mathbf{R}_{\geq 0}$ defined on $\mathcal{P}(x)$.

2   *What is Conditional Probability?*

Given any two events A and B which are not independent, the probability of event A taking place given that event B has already taken place is defined by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0.$$

That is the probability of event A given that event B has already taken place is defined as the probability of both event A and event B taking place divided by the probability of event B taking place.

If event A and event B are two independent events then

$$P(A \cap B) = P(A)P(B)$$

and
$$P(A|B) = P(A).$$
This can be interpreted as the definition of independence.

3   The Law of Total Probability

Let $\cup_{i=1}^{\infty} B_i$ be i mutually exclusive events. If $\cup_{i=1}^{\infty} B_i = \Omega$ where Ω is the sample space and for any $i \neq j$
$$B_i \cap B_j = \emptyset$$
that is B_i and B_j are disjoint then
$$P(A) = \sum_{i=1}^{\infty} P(B_i)P(A|B_i).$$

4   Conditional Probability: The Discrete Case

Suppose that X and Y are discrete RVs. For any x, $y \in \mathbb{R}$ consider the events $\{X = x\}$ and $\{Y = y\}$ with
$$P(Y = y) > 0.$$
Define the conditional probability mass function of X given that $Y = y$ by
$$p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)}.$$

4.1   Example

Suppose that the joint mass function $p(x,y)$ of X and Y is
$$p(1,1) = 0.5, \quad p(1,2) = 0.1, \quad p(2,1) = 0.1, \quad p(2,2) = 0.3.$$
Let us calculate $p_{X|Y}(x|Y = 1)$.

To begin with let us compute $p_Y(1)$ which is
$$p_Y(1) = \sum_x p(x,1) = p(1,1) + p(2,1) = 0.6.$$

Now we can compute $p_{X|Y}(1|1)$ and $p_{X|Y}(2|1)$ very easily which can be found to be
$$p_{X|Y}(1|1) = \frac{5}{6}$$
and
$$p_{X|Y}(2|1) = \frac{1}{6}.$$

5   Conditional Probability: The Continuous Case

Suppose that X and Y are continuous RVs. For any x, $y \in \mathbb{R}$ consider the events $\{X = x\}$ and $\{Y = y\}$ with
$$P(Y = y) > 0.$$

Define the conditional probability mass function of X given that $Y = y$ by

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)dx}$$

and the associated conditional expectation is given by

$$E(X|Y = y) = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx.$$

5.1   *Example 1*

Suppose the joint density of X and Y is given by

$$f(x,y) = \begin{cases} 6xy(2 - x - y), & 0 < x < 1, \quad 0 < y < 1 \\ 0, & otherwise. \end{cases}$$

Let us compute the conditional expectation of X given $Y = y$ where $0 < y < 1$.

If $y \in (0,1)$ we have for $x \in (0,1)$ that

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)dx} = \frac{6xy(2 - x - y)}{\int_0^1 6xy(2 - x - y)dx} = \frac{6xy(2 - x - y)}{4 - 3y}$$

Then

$$E(X|Y = y) = \frac{5 - 4y}{8 - 6y}$$

where $y \in (0,1)$.

5.2   *Example 2*

Consider the triangle in the plane $\mathbb{R}$ whose vertices are $(0,0)$, $(0,1)$ and $(1,0)$.

Let X and Y be continuous random variables for which the joint density is given by

$$f(x,y) = \begin{cases} 2, & 0 \leq x \leq 1, \quad 0 \leq y \leq 1, \quad x + y \leq 1 \\ 0, & otherwise. \end{cases}$$

Let us compute the conditional expectation of X given $Y = y$ where $0 \leq y \leq 1$. $0 \leq y \leq 1$ tells us that $x \in [0, 1 - y]$ and so

$$f_{X|Y}(x|y) = \frac{2}{\int_0^{1-y} 2dx} = \frac{1}{1 - y}$$

. Then

$$E(X|Y = y) = \int_0^{1-y} x(1 - y)^{-1}dx = \frac{1 - y}{2}.$$

6   *Discrete-Time Markov Chains & Transition Matrix*

Let X_1, X_2,..., be independent RVs each taking values -1 and 1 with probabilities $q = 1 - p$ and p where $p \in (0,1)$.

Define the random walk

$$S_n = a + \sum_{i=1}^{n} X_i$$

by the position of the movement after n steps having started at $S_0 = a$ (in most cases $a = 0$).

For $p = \frac{1}{2}$, $q = 1 - p = \frac{1}{2}$ and so we would have a symmetric random walk.

6.1 Example: The Gambler's Ruin

A man is saving to buy a BMW at a cost of N units of money. He stars with a ($0 \leq a \leq N$) units and tries to win the remainder by a gamble.

He flips a fair coin repeatedly; if its heads his friend pays him one unit; for tails, he pays his friend one unit. He keeps playing this game until either he runs out of money or he wins enough to buy.

This is an example of a symmetric random walk.

6.2 Random Stochastic Process

A random stochastic process is a group of random variables

$$\{X_t : t \in T\}$$

indexed by some set $T \subset \mathbb{R}$ where T can be be thought of as time.

If $T \subset \mathbb{R}$ is a discrete set, we speak of discrete time process $\{X_t\}$, $t \in T$.

If T happens to be continuous we speak of a continuous time process $\{X_t\}$, $t \in T$.

If the state space S of this process is finite then $\{X_t\}$ is called a chain.

6.3 Discrete Time Markov Chain (DTMC)

$\{X_n\}$ is a DTMC if it satisfies for all n

$$\begin{aligned} P\{X_{n+1} &= i_{n+1} | X_0 = i_0, ..., X_{n-1} = i_{n-1}, X_n = i_n\} \\ &= P\{X_{n+1} = i_{n+1} | X_n = i_n\} \end{aligned}$$

holds for arbitrary $i_k \in S$ for $k = 0, 1, ..., n+1$, $n \geq 1$. This means that the probability of any future behaviour of the process when its present state is known is not altered by the additional knowledge of its past behaviour. This is the Markovian property.

6.4 Example: Random Walk on the Cube

Let $S = \{(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1)\}$ where each state represents one corner of a cube in the obvious way in $\mathbb{R}^3$:
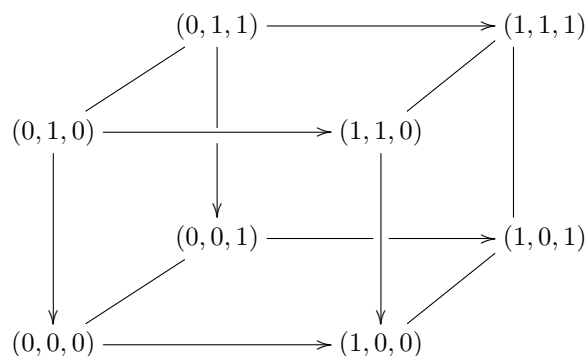
We wish to write down the transition matrix of the Markov chain in which from any corner independently of the past the next transition is equally likely to be to any of the three adjacent corners with probability

$\frac{1}{3}$.

	$(0,0,0)$	$(0,0,1)$	$(0,1,0)$	$(0,1,1)$	$(1,0,0)$	$(1,0,1)$	$(1,1,0)$	$(1,1,1)$
$(0,0,0)$	0	1/3	1/3	0	1/3	0	0	0
$(0,0,1)$	1/3	0	0	1/3	0	1/3	0	0
$(0,1,0)$	1/3	0	0	1/3	0	0	1/3	0
$(0,1,1)$	0	1/3	1/3	0	0	0	0	1/3
$(1,0,0)$	1/3	0	0	0	0	1/3	1/3	0
$(1,0,1)$	0	1/3	0	0	1/3	0	0	1/3
$(1,1,0)$	0	0	1/3	0	1/3	0	0	1/3
$(1,1,1)$	0	0	0	1/3	0	1/3	1/3	0

6.5 *Another Example*

Consider the Markov chain with three states $S = \{1, 2, 3\}$ with

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

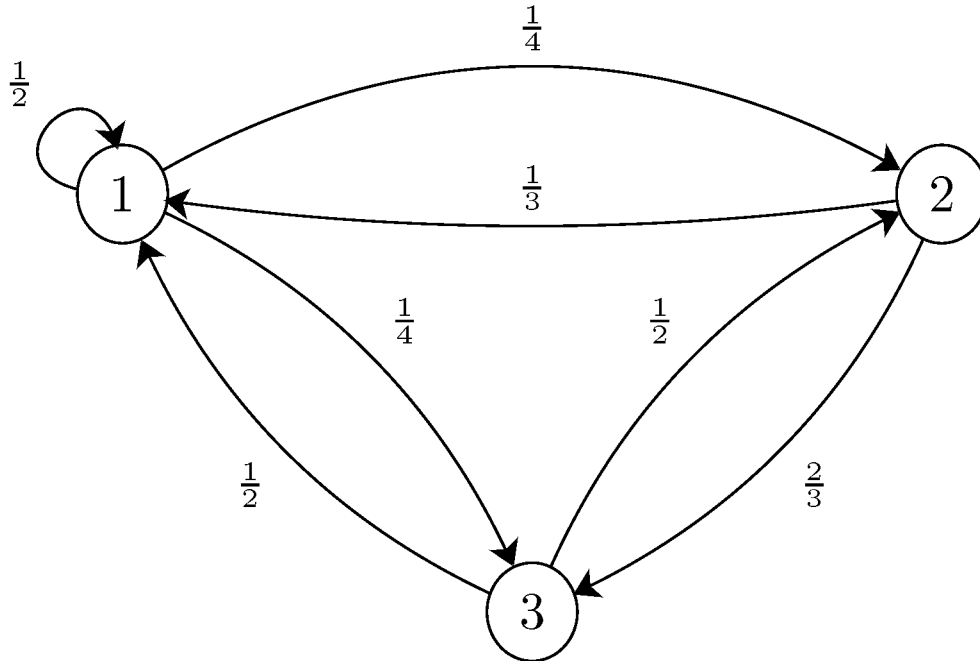then the state transition diagram for this chain is as shown below:

Figure 1: The State Transition Diagram

Furthermore, given

$$P(X_1 = 1) = P(X_1 = 2) = \frac{1}{4}$$

we can find $P(X_1 = 3, X_2 = 2, X_3 = 1)$. Firstly, we will have to compute $P(X_1 = 3)$. We proceed as follows

$$P(X_1 = 3) = 1 - P(X_1 = 1) - P(X_1 = 2) = 1 - \frac{1}{4} - \frac{1}{4} = \frac{1}{2}.$$

5

Then
$$P(X_1 = 3, X_2 = 2, X_3 = 1) = P(X_1 = 3) \cdot p_{3,2} \cdot p_{2,1} = \frac{1}{12}.$$

7 *Regular Transition Matrices*

A transition matrix is regular if some power of the transition matrix contains all positive entries.

For example
$$A = \begin{pmatrix} 0.75 & 0.25 & 0 \\ 0 & 0.5 & 0.5 \\ 0.6 & 0.4 & 0 \end{pmatrix}$$

which squares to
$$A^2 = \begin{pmatrix} 0.5625 & 0.3125 & 0.125 \\ 0.3 & 0.45 & 0.25 \\ 0.45 & 0.35 & 0.2 \end{pmatrix}$$

since all the entries in A^2 are positive, matrix A is regular. But
$$B = \begin{pmatrix} 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is not regular since
$$B^2 = \begin{pmatrix} 0.25 & 0 & 0.75 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and using the fact that if zeros occur in the identical places in both B and B^2 than for higher power they will appear in the same places. So no power of matrix B contains all positive entries.

8 *Absorbing Markov Chains*

State i of a Markov chain is an absorbing state if $p_{ii} = 1$.

A Markov chain is an absorbing chain if

- the chain has at least one absorbing state
- it is possible to go from any non-absorbing state to an absorbing state.

   For example
$$\begin{pmatrix} 1 & 0 & 0 \\ 0.3 & 0.5 & 0.2 \\ 0 & 0 & 1 \end{pmatrix}$$

where $p_{11} = 3 = p_{33}$ and we deduce that state 1 and state 3 are absorbing states. This is an example of a Markov chain that is an absorbing chain since state 2 is non-absorbing and it is possible to go from state 2 to state 1 and state 3.

   However,
$$\begin{pmatrix} 0.6 & 0 & 0.4 & 0 \\ 0 & 1 & 0 & 0 \\ 0.9 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is a Markov chain that is non-absorbing.

9   Steady-State Distribution

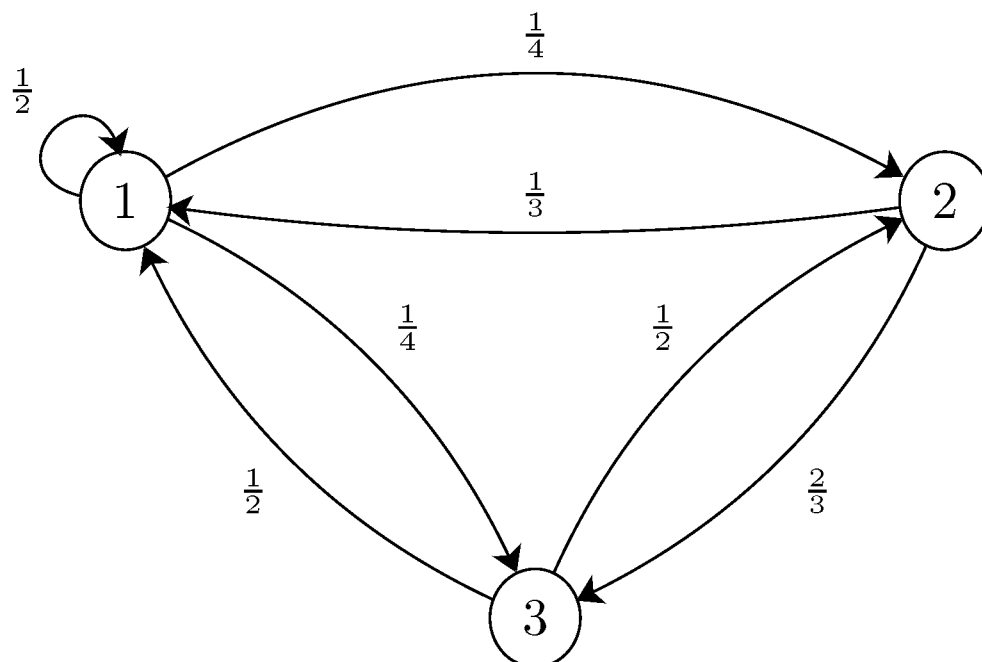We wish to find the steady state distribution for the following chain

Figure 2: The State Transition Diagram

we find

$$\frac{1}{2}\pi_1 + \frac{1}{3}\pi_2 + \frac{1}{2}\pi_3 = \pi_1$$
$$\frac{1}{4}\pi_1 + \frac{1}{2}\pi_3 = \pi_2$$
$$\frac{1}{4}\pi_1 + \frac{2}{3}\pi_2 = \pi_3$$
$$\pi_1 + \pi_2 + \pi_3 = 1$$

from which we deduce that

$$\pi_3 = \frac{10}{9}\pi_2 = \frac{5}{8}\pi_1$$

and

$$\pi_1 = \frac{16}{9}\pi_2$$

giving

$$\pi_1 = \frac{16}{35}$$
$$\pi_2 = \frac{9}{35}$$
$$\pi_3 = \frac{10}{35}.$$

10 *More Examples of DTMC*

A mobile robot randomly moves along a circular path divided into three sectors. At every sampling instant the robot with probability p to move clockwise and with probability $1 - p$ to move counter-clockwise. During the sampling interval the robot accomplishes the length of a sector.
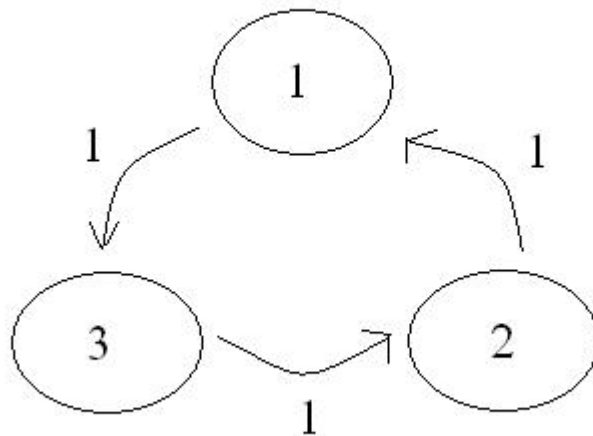
- Define a discrete Markov chain for the random walk of the robot.

- Study the stationary probability that the robot is localized in each sector for $p \in [0, 1]$.
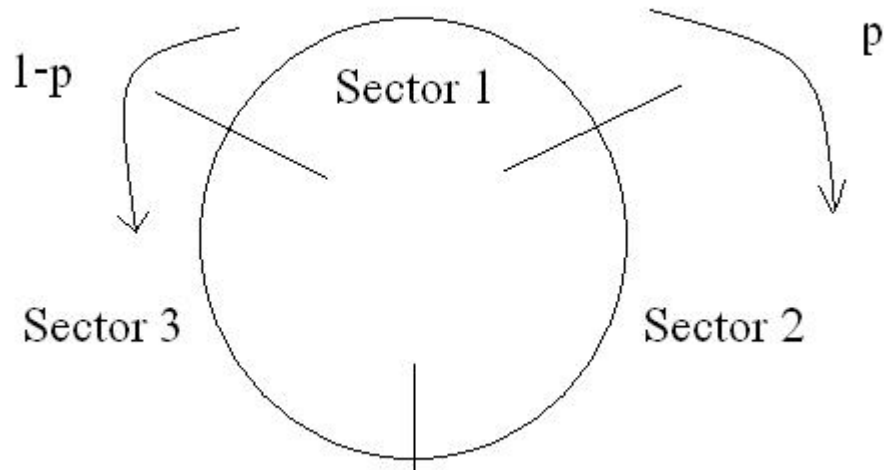
State X = sector occupied by the robot $\in \{1, 2, 3\}$. The transition matrix reads

$$\begin{pmatrix} 0 & p & 1 - p \\ 1 - p & 0 & p \\ p & 1 - p & 0 \end{pmatrix}$$

For $p = 0$ we have

Each state of the Markov chain is periodic with period 3. Therefore, stationary probabilities do not exist.

$p = 1$ case is similar to the $p = 0$ case.

If $p \in (0, 1)$ the Markov chain is irreducible aperiodic and finite. Therefore, the stationary-state probabilities exist.

The unique steady state probability vector is given by

$$\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$$

which is the solution of the system of equations

$$\begin{aligned} \pi &= \pi P \\ \sum_{i=1}^{3} \pi_i &= 1 \end{aligned}$$

11   Reversible Markov Chains

A Markov chain is with stationary distribution π is reversible if and only if

$$\pi_i p_{ij} = \pi_j p_{ji}$$

for all states i and j.

11.1   Random Walk on a Graph

Recall that a graph is a collection of vertices where the ith vertex v_i is directly joined by at most one edge each to d_i other vertices (called its neighbour).

If

$$\sigma = \sum_i d_i < \infty$$

then a random walk on the vertices of the graph simply goes to any one of the neighbours of v_i with equal probability $\frac{1}{d_i}$. Such a walk is reversible. To see this note that the detailed balance equations take the form

$$\frac{\pi_i}{d_i} = \frac{\pi_j}{d_j}$$

for neighbours i and j and the solution such that

$$\sum_i \pi_i = 1$$

is obviously

$$\pi_i = \frac{d_i}{\sigma}.$$

11.2   The Ehrenfest Model

M balls distributed in two urns. Each time, pick a ball at random, move it from one urn to the other one. Let X_n be the number of balls in urn 1. Prove that the chain has a reversible distribution.

The non-zero transition probabilities are

$$\begin{aligned} p_{i,i-1} &= \frac{i}{M} \\ p_{i,i+1} &= \frac{M-i}{M} \end{aligned}$$

We are seeking solutions to

$$\pi_i p_{i,i+1} = \pi_{i+1} p_{i+1,i} \Rightarrow \pi_i \frac{M-i}{M} = \pi_{i+1} \frac{i+1}{M}$$

leading to

$$\frac{\pi_i}{\pi_0} = \frac{(M-i-1)(M-i)...M}{i(i-1)...1} = \binom{M}{i}.$$

Note.

$$\sum_i \pi_i = 1 = \pi_0 \sum_i \binom{M}{i} \Rightarrow \pi_0 = \frac{1}{2^m}.$$

Note. Using

$$p_{i,i-1} = \frac{i}{M}$$

and $i \to i+1$ arrive at

$$p_{i+1,i} = \frac{i+1}{M}.$$

The chain is irreducible and periodic.

12 *Continuous-Time Markov Chains*

We will consider a random process

$$\{X(t), t \in [0, \infty)\}.$$

We assume that we have a countable state space

$$S \subset \{0, 1, 2, ...\}.$$

If $X(0) = i$, then $X(t)$ stays in state i for a random amount of time say T_1 where T_1 is a continuous random variable. At time T_1, the process jumps to a new state j and will spend a random amount of time T_2 in that state and so on. The random variables $T_1, T_2, ...$ follow an exponential distribution. The probability of going from state i to state j is denoted p_{ij}.

In particular, suppose that at time t, we know that $X(t) = i$. To make any prediction about the future, it should not matter how long the process spends in each state and therefore must have a "memoryless" property. The exponential distribution is the only continuous distribution with the "memoryless" property. Thus, the time that a continuous-time Markov chain spends in state i known as the holding time will have an exponential distribution: e^{λ_i} where λ_i is a non-negative real number. We further assume that the λ_is are bounded that is there exists a real number $M < \infty$ such that $\lambda_i < M$ for all $i \in S$.

A continuous-time Markov chain has two components. First is the discrete-time Markov chain known as the jump chain or an embedded Markov chain which give the transition probabilities p_{ij}. The second is, for each state, there is a holding time parameter λ_i that controls the time spent in each state.

12.1 *An Example*

Consider a Markov chain with two states $S = \{0, 1\}$. Assume the holding time parameters are given by $\lambda_0 = \lambda_1 = \lambda > 0$. That is, the time that the chain spends in each state before going to the other state has e^λ distribution.

- Draw the state diagram of the embedded (jump) chain.

- Find the transition matrix $P(t)$. You may find the following identities useful

$$\sinh(x) = \frac{e^x - e^{-x}}{2} = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}$$

$$\cosh(x) = \frac{e^x + e^{-x}}{2} = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}$$

12.2 The Solution: Part 1

There are two states in the chain and none of them are absorbing as $\lambda_0 = \lambda_1 = \lambda > 0$. The jump chain must therefore have the following transition matrix

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

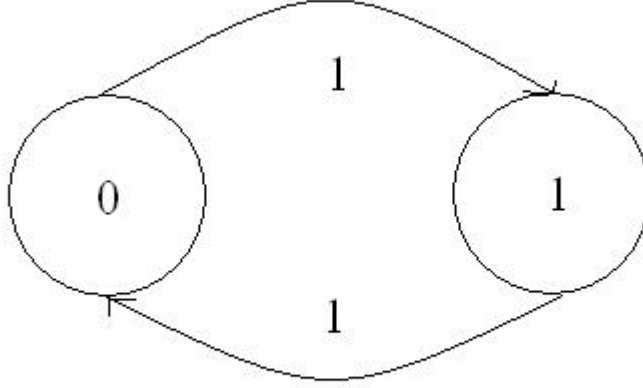where the state-transition diagram of the embedded (jump) chain is

Figure 3: The State Transition Diagram

12.3 The Solution: Part 2

The Markov chain has a simple structure. Lets find $P_{00}(t)$. By definition

$$P_{00}(t) = P(X(t) = 0 | X(0) = 0), \quad \forall t \in [0, \infty).$$

Assuming that $X(0) = 0$, $X(t)$ will be 0 if and only if we have an even number of transitions in the time interval $[0, t]$. The time between each transition is an e^λ random variable. Therefore, the transitions occur according to a Poisson process with parameter λ.

We find

$$
\begin{aligned}
P_{00}(t) &= \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^{2n}}{(2n)!} \\
&= e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^{2n}}{(2n)!} \\
&= e^{-\lambda t} \left[\frac{e^{\lambda t} + e^{-\lambda t}}{2} \right] \\
&= \frac{1}{2} + \frac{1}{2} e^{-2\lambda t}
\end{aligned}
$$

$$P_{01}(t) = 1 - P_{00}(t) = \frac{1}{2} - \frac{1}{2} e^{-2\lambda t}.$$

11

$$P_{11}(t) = P_{00}(t).$$
$$P_{10}(t) = P_{01}(t).$$

Therefore, the transition matrix for any $t \geq 0$ is given by

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

Letting t go to infinity, the transition matrix reads

$$\lim_{t \to \infty} P(t) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

12.4 *CTMC Transition Matrix: Properties*

The transition matrix satisfies the following properties:

1. $P(0)$ is equal to the identity matrix, $P(0) = I$;

2. The rows of the transition matrix must sum to 1

$$\sum_{j \in S} p_{ij}(t) = 1, \quad \forall t \geq 0;$$

3. For all $s, t \geq 0$,
$$P(s + t) = P(s)P(t).$$

13 *CTMC: Limiting Distribution Examples*

13.1 *Example 1*

Consider a continuous-time Markov chain $X(t)$ that has the jump chain
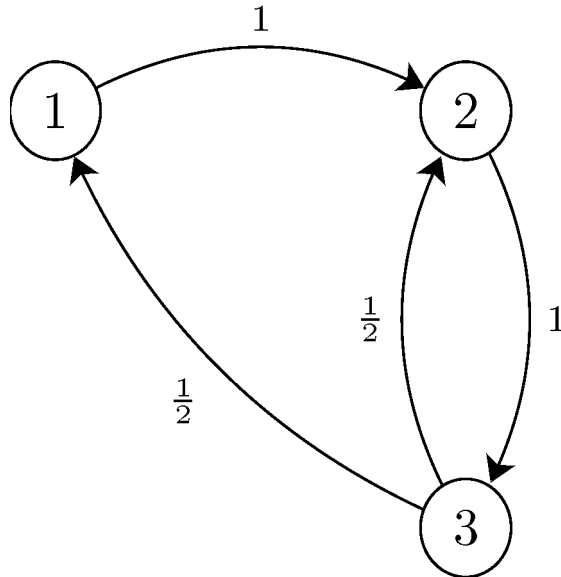

Figure 4: The State Transition Diagram

Assume the holding parameters are given by

$$\lambda_1 = 2, \qquad \lambda_2 = 1, \qquad \lambda_3 = 3.$$

We wish to find the limiting distribution of $X(t)$.

First notice that the chain is irreducible. The transition matrix of the jump chain reads

$$P(t) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

the stationary distribution for the jump chain being

$$\tilde{\pi} = \frac{1}{5}(1, 2, 2).$$

This is obtained by solving the following system of equations

$$
\begin{aligned}
\pi_3 &= 2\pi_1 \\
\pi_2 &= 2\pi_1 \\
\pi_2 &= \pi_3 \\
\sum_{i=1}^{3} \pi_i &= 1
\end{aligned}
$$

The limiting distribution of $X(t)$ can then be obtained by using

$$\pi_j = \frac{\frac{\tilde{\pi}_j}{\lambda_j}}{\sum_{k \in S} \frac{\tilde{\pi}_k}{\lambda_k}}.$$

Hence, we conclude that

$$\pi = \frac{1}{19}(3, 12, 4)$$

is the limiting distribution of $X(t)$.

13.2 *Example 2*

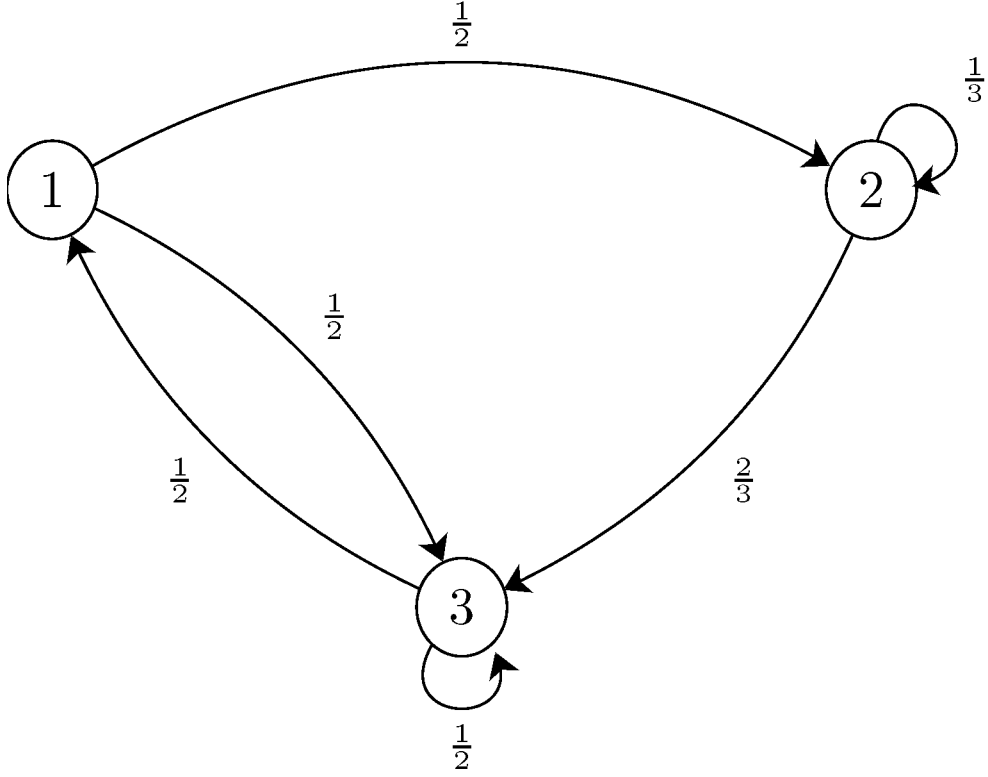Consider a continuous-time Markov chain $X(t)$ that has the jump chain

Figure 5: The State Transition Diagram

Assume the holding parameters are given by

$$\lambda_1 = 2, \qquad \lambda_2 = 3, \qquad \lambda_3 = 4.$$

We wish to find the limiting distribution of $X(t)$.

First notice that the chain is irreducible. The transition matrix of the jump chain reads

$$P(t) = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

the stationary distribution for the jump chain being

$$\tilde{\pi} = \frac{1}{15}(4, 3, 8).$$

The limiting distribution of $X(t)$ can then be obtained by using

$$\pi_j = \frac{\frac{\tilde{\pi}_j}{\lambda_j}}{\sum_{k \in S} \frac{\tilde{\pi}_k}{\lambda_k}}$$

which gives

$$\pi = \frac{1}{5}(2, 1, 2)$$

as the limiting distribution of $X(t)$.

14

14 Continuous-Time Markov Chains: The Generator Matrix

For a continuous-time Markov chain, define the generator matrix G as a transition matrix whose (i, j)-th entry is given by

$$g_{ij} = \begin{cases} \lambda_i p_{ij} & \text{if } i \neq j \\ -\lambda_i & \text{if } i = j \end{cases}$$

14.1 Example

A chain with two states $S = \{0, 1\}$ and $\lambda_0 = \lambda_1 = \lambda > 0$. We found that the transition matrix for any $t \geq 0$ is given by

$$P(t) = \begin{pmatrix} \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} \\ \frac{1}{2} - \frac{1}{2}e^{-2\lambda t} & \frac{1}{2} + \frac{1}{2}e^{-2\lambda t} \end{pmatrix}.$$

We wish

1. To find the generator matrix G

2. Show that for any $t \geq 0$

$$P'(t) = P(t)G = GP(t).$$

By definition

$$\begin{aligned} g_{00} &= -\lambda_0 = -\lambda \\ g_{11} &= -\lambda_1 = -\lambda \\ g_{01} &= \lambda_0 p_{01} \\ &= \lambda \\ g_{10} &= \lambda_1 p_{10} \\ &= \lambda \end{aligned}$$

Thus, the generator matrix is given by

$$G = \begin{pmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{pmatrix}$$

To show that for any $t \geq 0$

$$P'(t) = P(t)G = GP(t)$$

we note that

$$P'(t) = \begin{pmatrix} -\lambda e^{-2\lambda t} & \lambda e^{-2\lambda t} \\ \lambda e^{-2\lambda t} & -\lambda e^{-2\lambda t} \end{pmatrix}$$

We also have that

$$P(t)G = \begin{pmatrix} -\lambda e^{-2\lambda t} & \lambda e^{-2\lambda t} \\ \lambda e^{-2\lambda t} & -\lambda e^{-2\lambda t} \end{pmatrix} = GP(t).$$

14.2 The Forward and Backward Equations

- The forward equations state that

$$P'(t) = P(t)G$$

- The backward equations state that

$$P'(t) = GP(t)$$

15

14.3   *The Limiting Distribution Via Generator Matrix*

Consider a continuous-time Markov chain $X(t)$ with the state space S and the generator matrix G. The probability distribution π on S is stationary distribution for $X(t)$ if and only if it satisfies

$$\pi G = 0.$$

14.4   *Example 1*

The generator matrix for the continuous-time Markov chain is given by

$$G = \begin{pmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{pmatrix}$$

We wish to find the stationary distribution for this chain by solving

$$\pi G = 0.$$

We have that $\pi_0 = \pi_1$ and by imposing that $\sum_i \pi_i = 1$, we obtain

$$\pi = \frac{1}{2}(1,1).$$

14.5   *Example 2*

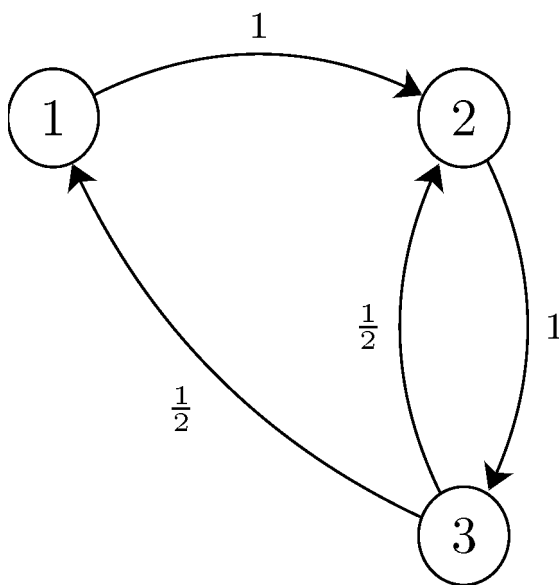Consider the Markov chain $X(t)$ that has the following jump chain

Figure 6: The State Transition Diagram

and assume $\lambda_1 = 2$, $\lambda_2 = 1$ and $\lambda_3 = 3$.

1. We wish to find the generator matrix for this chain.

2. We wish to find the limiting distribution for $X(t)$ by solving $\pi G = 0$.

The jump chain is irreducible and the transition matrix of the jump chain is given by

$$P(t) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

and by definition, we have the generator matrix G with entries

$$G = \begin{pmatrix} -2 & 2 & 0 \\ 0 & -1 & 1 \\ \frac{3}{2} & \frac{3}{2} & -3 \end{pmatrix}.$$

Now by solving

$$\pi G = 0$$

we will be able to find the limiting distribution. We obtain the following system of equations

$$
\begin{aligned}
-2\pi_1 + \frac{3}{2}\pi_3 &= 0 \\
2\pi_1 - \pi_2 + \frac{3}{2}\pi_3 &= 0 \\
\pi_2 - 3\pi_3 &= 0
\end{aligned}
$$

Then imposing $\sum_i \pi_i = 1$, we find that

$$\pi = \frac{1}{19}(3, 12, 4)$$

is the limiting distribution of $X(t)$.

A   Axioms of Probability

Given

- $P(\Omega) = 1$
- $0 \le P(A) \le 1$ for any event $A \subseteq \Omega$
- if $A_1, ..., A_n$ are events and $A_i \cap A_j = \emptyset$ for all $i \ne j$ then

$$PP(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{i=1}^{n} P(A_i).$$

A.1 $P(A^c) = 1 - P(A)$

A and A^c are disjoint that is $A \cap A^c = \emptyset$.

$$A \cup A^c = \Omega \Rightarrow P(A) + P(A^c) = P(\Sigma) = 1$$

from which follows that

$$P(A^c) = 1 - P(A).$$

A.2 $P(\emptyset) = 0$

$$P(\emptyset) = P(\Omega^c) = 1 - P(\Omega) = 1 - 1 = 0.$$

A.3 $P(B) \geq P(A)$ *given* $A \subseteq B$

If $A \subseteq B$ then
$$B = A \cup (B \cap A^c)$$
and
$$A \cap (B \cap A^c) = \emptyset.$$
Then
$$P(B) = P(A) + \underbrace{P(B \cap A^c)}_{\geq 0} \Rightarrow P(B) \geq P(A).$$

A.4 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$A = (A \cap B^c) \cup (A \cap B)$$
$$B = (B \cap A^c) \cup (A \cap B)$$
$$\Rightarrow A \cup B = (A \cap B^c) \cup (A \cap B) \cup (B \cap A^c) \Rightarrow P(A \cup B) = P(A \cap B^c) + P(A \cap B) + P(B \cap A^c)$$

But
$$P(A \cap B^c) = P(A) - P(A \cap B)$$
and
$$P(B \cap A^c) = P(B) - P(A \cap B)$$
which leads to
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

B   *Poisson Distribution Through Examples*

A Poisson random variable is the number of successes that result from a Poisson experiment. The probability distribution of a Poisson random variable is called a Poisson distribution.

Given the mean number of successes λ we have the **PoissonFormula** which states that if we were to conduct a Poisson experiment in which the average number of successes within a given region is λ then the Poisson probability is
$$P(n; \lambda) = \frac{e^{-\lambda} \lambda^n}{n!}, \quad \lambda > 0$$
where n is the actual number of successes.

A cumulative Poisson distribution refers to the probability that the Poisson random variable is greater than some specified lower limit and less than some specified upper limit.

B.1   *Example 1*

Each customer who enters Rebecca's clothing store will purchase a suit with probability $p > 0$. If the number of customers entering the store is Poisson distributed with parameter $\lambda > 0$

- we want to know the probability that Rebecca does not sell any suits:

Let X be the numbers of suits that Rebecca sells and N be the number of customers who enter the store. Then

$$
\begin{aligned}
P\{X = 0\} &= \sum_{n=0}^{\infty} P\{X = 0|N = n\}P\{N = n\} \\
&= \sum_{n=0}^{\infty} P\{X = 0|N = n\}\frac{e^{-\lambda}\lambda^n}{n!} \\
&= \sum_{n=0}^{\infty} (1-p)^n \frac{e^{-\lambda}\lambda^n}{n!} \\
&= \sum_{n=0}^{\infty} e^{-\lambda}\frac{(\lambda(1-p))^n}{n!} \\
&= e^{-\lambda}e^{\lambda(1-p)} \\
&= e^{-\lambda p}
\end{aligned}
$$

- we want to know the probability that Rebecca sells k suits:

  First note that given $N = n$ X has a binomial distribution with parameter n and p

  $$
  P\{X = k|N = n\} = \binom{n}{k}p^k(1-p)^{n-k} \text{ if } n \geq k
  $$

  and

  $$
  P\{X = k|N = n\} = 0 \text{ if } n < k.
  $$

  Then

$$
\begin{aligned}
P\{X = 0\} &= \sum_{n=0}^{\infty} P\{X = k|N = n\}P\{N = n\} \\
&= \sum_{n=0}^{\infty} P\{X = k|N = n\}\frac{e^{-\lambda}\lambda^n}{n!} \\
&= \sum_{n=0}^{\infty} \binom{n}{k}p^k(1-p)^{n-k}\frac{e^{-\lambda}\lambda^n}{n!} \\
&= \frac{e^{-\lambda}(\lambda p)^k}{k!} \sum_{n=0}^{\infty} \frac{(\lambda(1-p))^n}{n!} \\
&= \frac{e^{-\lambda}(\lambda p)^k}{k!} e^{\lambda(1-p)} \\
&= \frac{e^{-\lambda p}(\lambda p)^k}{k!}
\end{aligned}
$$

B.2   *Example 2*

Suppose the average number of lions seen on a one-day safari is 5. What is the probability that tourists will see fewer than four lions on the next one-day safari?

This is a Poisson experiment in which $\lambda = 5$ and $n = 0, 1, 2, 3$ and we wish to calculate the sum of the probabilities $P(0; 5)$, $P(1; 5)$, $P(2; 5)$, $P(3; 5)$.

This leads to

$$
P(n \leq 3; 5) = \sum_{n=0}^{3} P(n; 5) = 0.2650
$$

B.3 *Example 3*

The number of baseball games rained out is a Poisson process with arrival rate of 5 per 30 days.

- We wish to find the probability that there are more than 5 rained out games in 15 days.

- We wish to find the probability that there are no rained out games in 7 days.

First note that

$$\lambda = \frac{5}{30} = \frac{1}{6}$$

and the total number of rained out games in t days is Poisson($\frac{t}{6}$). To find the probability that there are more than 5 rained out games in 15 days reduces to

$$1 - P(n \leq 5; \frac{5}{2}) = 0.042.$$

Similarly, the probability that there are no rained out games in 7 days is

$$e^{-7/6} = 0.311$$