INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Class: M.Sc. SEM 3

Subject: Statistical and Risk Modelling 3

Chapter: Unit 2 Chapter 1

Chapter Name: Markov Chains

Today’s Agenda

1. Introduction to Markov chains
   1.1 Transition graph and transition matrix

2. Chapman-Kolmogorov equations

3. Time homogenous and Time in-homogenous

4. Long Term Stationary Distribution of a Markov Chain

5. Modelling using Markov Chains

1 Markov Chains

The Markov property means that the future value of a process is independent of the past history and only depends on the current value. Any process satisfying the Markov property is a Markov process.

The term Markov chain refers to Markov processes in discrete time and with a discrete state space.

1 Example

A simple model of a No Claims Discount (NCD) policy.

The No Claims Discount system (NCD) in motor insurance, whereby the premium charged depends on the driver's claim record, is a prime application of Markov chains.

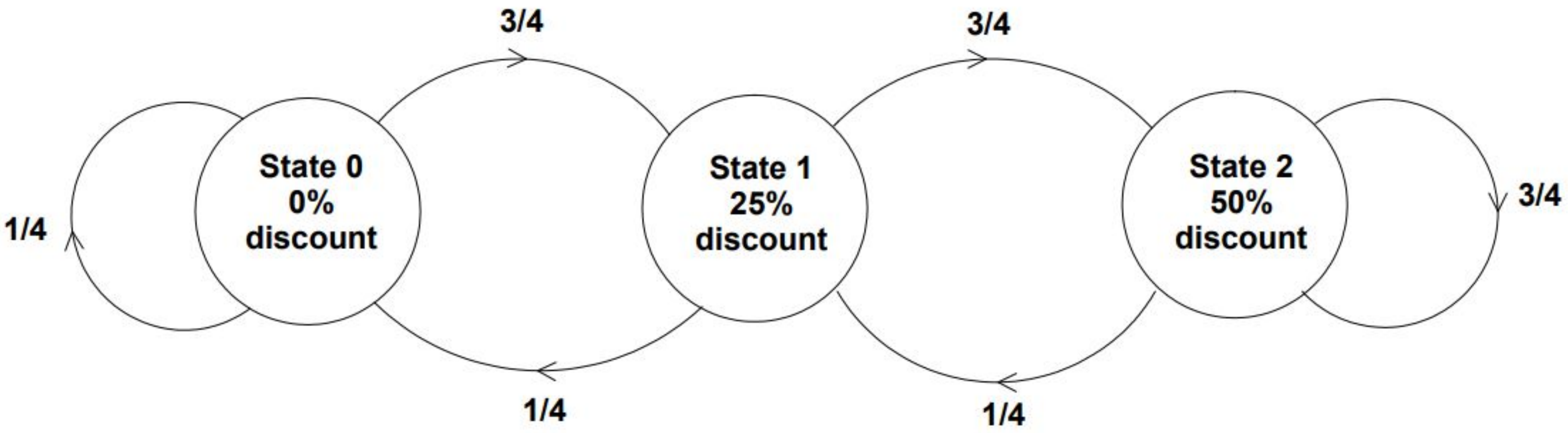
A motor insurance company grants its customers either no discount (state 0) or 25% discount (state 1) or 50% discount (state 2).

A claim-free year results in a transition to the next higher state the following year (or in the retention of the maximum discount); similarly, a year with one claim or more causes a transition to the next lower state (or the retention of the zero-discount status). This is an example of a discrete-time process that satisfies the Markov property (because the future only depends on the current level, and not on the past history). Hence this is a Markov chain.

Under these rules, the discount status of a policyholder forms a Markov chain with state space $S = \{0,1,2\}$: consider that the probability of a claim-free year is ¾.

1.1 Transition Graph

One way of representing a Markov chain is by its transition graph. In case of the example stated above, the transition graph will be given as:



The states are represented by the circles, and each arrow represents a possible transition. Staying in a state for one time period is also a possible transition, so we need an arrow to show this. Next to each arrow is written the corresponding transition probability.

1.2 Transition Matrix

The transition probability between two states in unit time is therefore given explicitly. These can be written in the form of a transition matrix:

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

The (i, j)th entry in the matrix, ie: i-th row and j-th column, gives the probability of moving in one step from state i to state j .

The entries in each row sum to 1. This can be interpreted as saying that at the end of each year, some transition into another state must occur or the process stays where it is. The sum of the probabilities of the mutually exclusive and exhaustive events that can happen when the process is in state i is 1.

It turns out that to calculate transition probabilities over 2 steps, we use the matrix P^2 , and for three steps we use P^3 , and so on.

2 Chapman-Kolmogorov equation

A Markov chain is a sequence of random variables $X_0, X_1, \ldots, X_N, \ldots$ with the following property:

$$P\left[X_n = j \middle| X_0 = i_0, X_1 = i_1, \ldots, X_{m-1} = i_{m-1}, X_m = i\right] = P\left[X_n = j \middle| X_m = i\right]$$

for all integer times n > m and states $i_0, i_1, \ldots, i_{m-1}$, i, j in S.

The Markov property has the following interpretation - given the present state of the process X_m = i, the additional knowledge of the past is irrelevant for the calculation of the probability distribution of future values of the process.

The conditional probabilities on the right-hand side of the property above are the key objects for the description of a Markov chain; we call them transition probabilities, and we denote them by:

$$\mathbf{P}\ [X_n = j \mid X_m = i] = p_{ij}^{(m,n)}$$

So $p_{ij}^{(m,n)}$ is the probability of being in state j at time n having been in state i at time m .

2 Chapman-Kolmogorov equation

The transition probabilities of a discrete-time Markov chain obey the Chapman-Kolmogorov equations:

$$\boldsymbol{p}_{\boldsymbol{i}\,\boldsymbol{j}}^{(\boldsymbol{m},\boldsymbol{n})} = \sum_{k \in S} \boldsymbol{p}_{\boldsymbol{i}\,\boldsymbol{k}}^{(\boldsymbol{m},\boldsymbol{l})}\ \boldsymbol{p}_{\boldsymbol{k}\,\boldsymbol{j}}^{(\boldsymbol{l},\boldsymbol{n})}$$

for all states i, j in S and all integer times m < l < n.

3 Time homogenous and Time inhomogenous

Time homogenous – Transition probabilities are independent of time, i.e. only depend upon duration and not on starting time.

The transition probabilities are given as: $\boldsymbol{p_{i\,j}^{(n,\,n+1)}} = \boldsymbol{p_{i\,j}}$

Eg: [$p_{ij}^{(5,7)} = p_{ij}^{(0,2)} = p_{ij}^{(10,12)} = p_{ij}^{(k,k+2)}$]

Time in-homogenous – Transition probabilities are dependent on starting time as well as duration.

The transition probabilities are given as: $\boldsymbol{p_{i\,j}^{(n,\,n+1)}}$

Eg: [$p_{ij}^{(5,7)} \neq p_{ij}^{(0,2)}$]

4 The long-term distribution of a Markov Chain

The stationary probability distribution

We say that $\pi_j, j \in S$ is a stationary probability distribution for a Markov chain with transition matrix P if the following conditions hold for all j in S :

$$\pi_j = \sum_{i \in S} \pi_i \, p_{ij}$$
$$\pi_j \geq 0$$
$$\sum_{j \in S} \pi_j = 1$$

First condition can be stated in the compact form $\pi = \pi P$ where π is viewed as a row vector.
The interpretation is that, if we take π as our initial probability distribution, that is to say $P[X_0 = i] = \pi_i$, then the distribution at time 1 is again given by π :

$$P[X_1 = j] = \sum_{i \in S} P[X_1 = j \mid X_0 = i] P[X_0 = i] = \sum_{i \in S} p_{ij} \pi_i = \pi_j$$

Stationary distribution result (1) - A Markov chain with a finite state space has at least one stationary probability distribution.

4 The long-term distribution of a Markov Chain

The method and steps to compute the stationary probability for a Markov model are:

Step 1: Apply $\pi = \pi P$ and get the system of equations.

Step 2: Discard one of the equations. [Preferably complicated one]

Step 3: Apply normalization theory. ie. $\sum_{j \in S} \pi_j = 1$

Step 4: Select one of the π_j 's as working variable.

Step 5: Rewrite remaining equations in terms of the working variable. (It is good practice to solve for the components of π in terms of one of them which we will refer to as the working variable).

Step 6: Solve the equations in terms of the working variable.

Step 7: Solve for the working variable.

Step 8: Find the stationary distribution by substituting the working variable.

4.1 Irreducible chains

Irreducibility

A Markov chain is said to be irreducible if any state ‘j’ can be reached from any other state i in any number of steps.

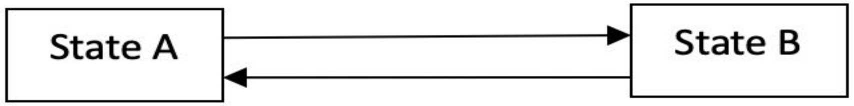
This is a property that can normally be judged from the transition graph alone.

Stationary distribution result (2) – An irreducible Markov chain with a finite state space has a unique stationary probability distribution.

Question

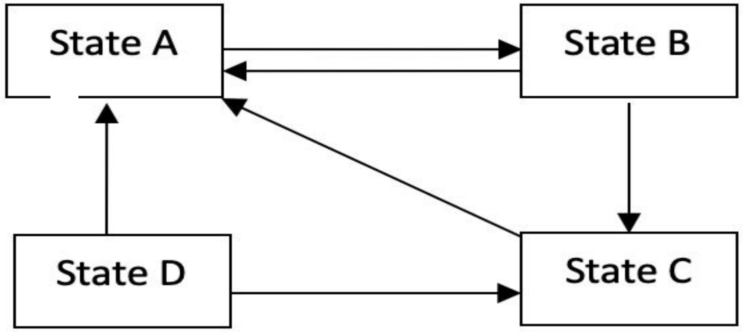
Identify whether these chains are irreducible or not

1)



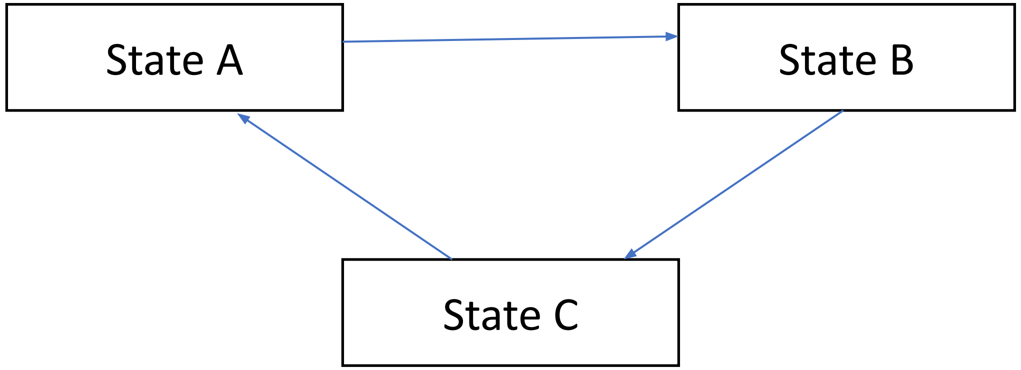
2)

3)



Question

4)



5)

4 The long-term distribution of a Markov Chain

It is natural to expect the distribution of a Markov chain to tend to the invariant distribution π for large times. This is why the stationary distribution is so important.

If the above convergence holds, $p_{i\,j}^{(n)}$ will be close to π_j for an overwhelming fraction of the time in the long run.

The period of a state
A state i is said to be periodic with period $d > 1$ if a return to i is possible only in a number of steps that is a multiple of d (ie $p_{ii}^{(n)} = 0$ unless $n = md$ for some integer m).
A state is said to be aperiodic if it is not periodic.

Periodicity result - If a Markov chain is irreducible all its states have the same period (or all are aperiodic).

Aperiodic Markov chains - A Markov chain is said to be aperiodic if all its states are aperiodic.

4 The long-term distribution of a Markov Chain

Stationary distribution result (3)

Let $p_{ij}^{(n)}$ be the n-step transition probability of an irreducible aperiodic Markov chain on a finite state space. Then for every i and j :

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

where π is the stationary probability distribution.

Summary

- A Markov chain with a finite state space has at least one stationary distribution.
- An irreducible Markov chain with a finite state space has a unique stationary distribution.
- An irreducible, aperiodic Markov chain with a finite state space will settle down to its unique stationary distribution in the long run.

Question

A Markov Chain with state space {A, B, C} has the following properties:
- it is irreducible
- it is periodic
- the probability of moving from A to B equals the probability of moving from A to C

(i) Show that these properties uniquely define the process. [4]

(ii) Sketch a transition diagram for the process. [1]
[Total 5]

Solution

(i) As periodic and irreducible then all states are periodic, hence probability of staying in any state is zero.

By law of total probability, $P_{AA} + P_{AB} + P_{AC} = 1$.

But $P_{AB} = P_{AC}$ and $P_{AA} = 0$ so $P_{AB} = P_{AC} = 0.5$.

To be irreducible at least one of P_{BA} or P_{CA} must be greater than zero.

If $P_{BA} > 0$ then to be periodic must have $P_{CB} = 0$,

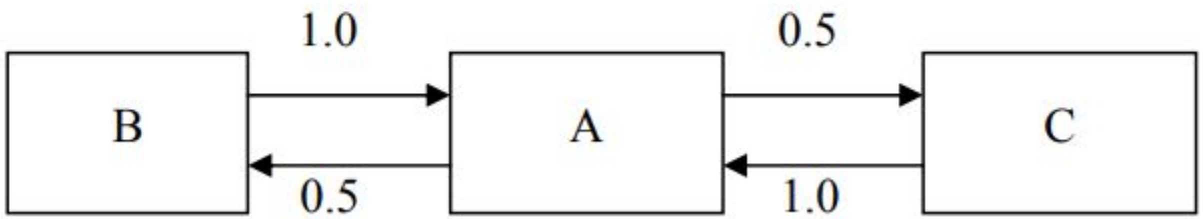
and to be irreducible $P_{CA} > 0$,

and if $P_{CA} > 0$ then to be periodic must have $P_{BC} = 0$, and to be irreducible $P_{BA} > 0$.

So must have $P_{BC} = P_{CB} = 0$ and $P_{BA} = P_{CA} = 1$.

Solution

(ii)



5 Modelling using Markov Chains

Estimating transition probabilities

The transition probability p_{ij} is estimated by:

$$\hat{p}_{ij} = \frac{n_{ij}}{n_i} = \frac{\text{number of transitions from state } i \text{ to state } j}{\text{number of transitions out of state } i}$$

Testing the Markov assumption

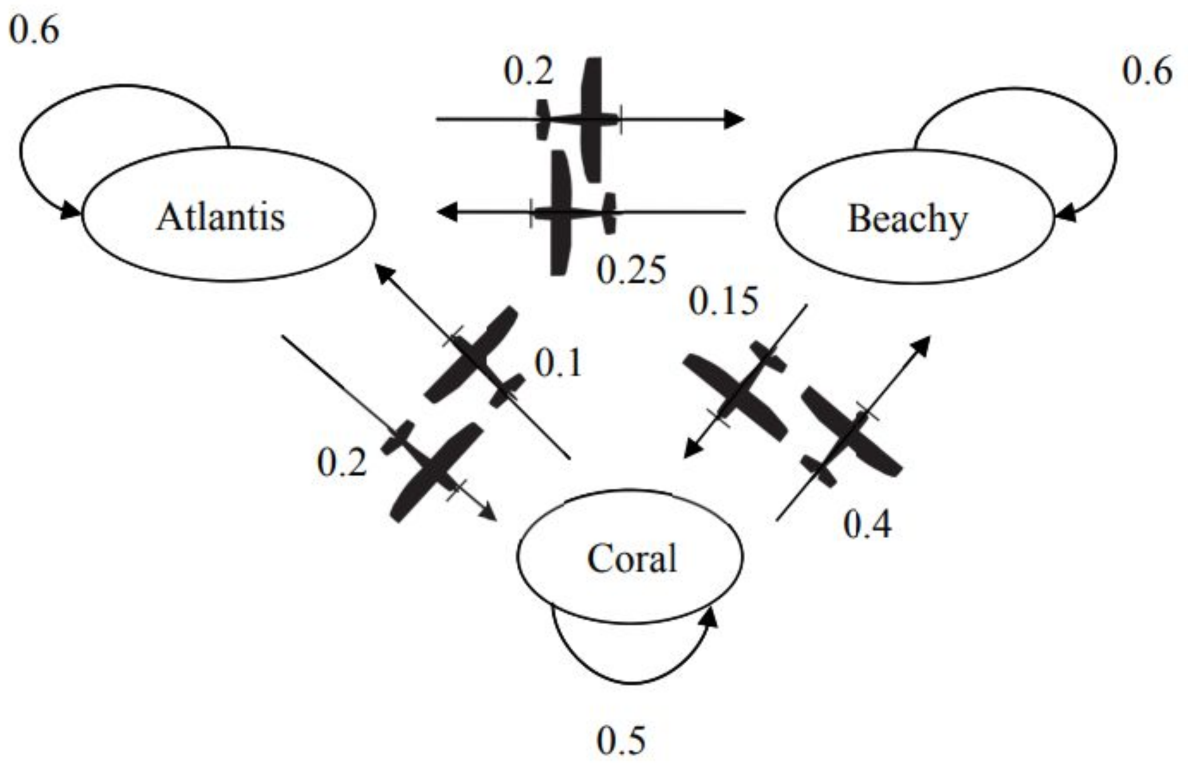
We can test the Markov assumption using a triplets test. The formula for the test statistic is:

$$\sum_{i,j,k} \frac{\left(n_{ijk} - n_{ij}\hat{p}_{jk}\right)^2}{n_{ij}\hat{p}_{jk}}$$

Question

CT4 A2017 Q9

A journalist has just been appointed as a foreign correspondent covering news stories in three island countries Atlantis, Beachy and Coral.
He will spend each day in the country likely to have the most exciting news, taking the flights available between each country which go once per day at the end of the day.
The previous foreign correspondent tells him "If you want to know how many flights you are likely to take, I estimate my movements have been like this" and she drew this diagram showing the transition probabilities:



Question

(i) Give the transition matrix for this process. [1]

On his first day in the job the new foreign correspondent will be in Atlantis.
(ii) Calculate the probability that the foreign correspondent will be in each country in his third day in the job. [2]

The previous correspondent also reported that Beachy must be the most interesting of the islands in terms of news because she spent 41.9% of her time there compared with 32.6% on Atlantis and 25.6% on Coral.

(iii) Sketch a graph showing the probability that the journalist is in each country over time, using the information above. [3]

(iv) Calculate the proportion of days on which the foreign correspondent will take a flight. [1]

Question

The first time the foreign correspondent visits each of the countries he takes a photograph to mark the occasion.

(v) Identify a suitable state space for modelling as a Markov chain which countries he has visited so far. [2]

(vi) Draw a transition diagram for the possible transitions between these states. [3]
[Total 12]

Solution

(i) The matrix is

$$\begin{array}{c} A \\ B \\ C \end{array}\begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.25 & 0.6 & 0.15 \\ 0.1 & 0.4 & 0.5 \end{pmatrix}$$

(ii) We need second order transition probabilities.

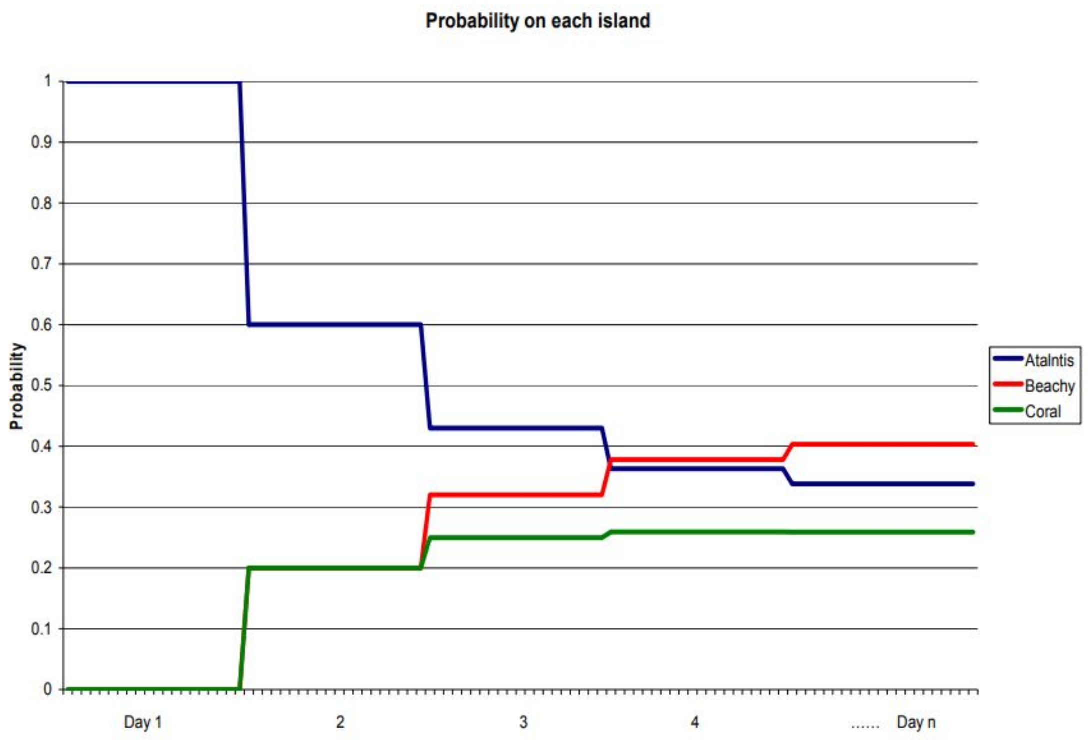
The second order transition matrix is:

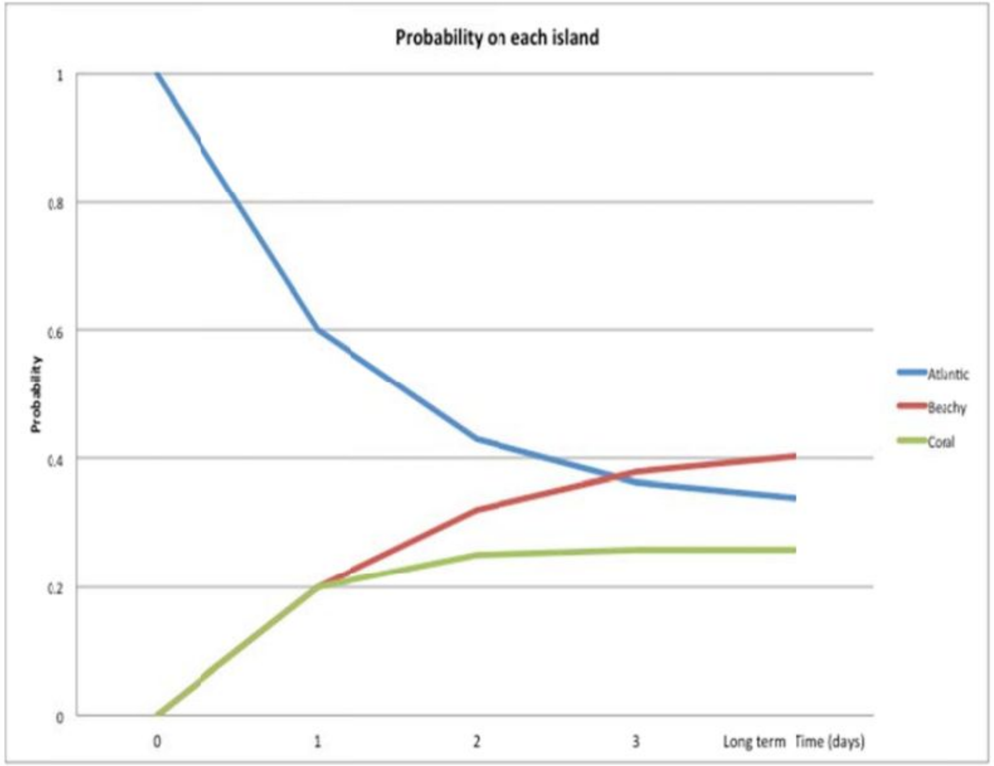
$$\begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.25 & 0.6 & 0.15 \\ 0.1 & 0.4 & 0.5 \end{pmatrix} * \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.25 & 0.6 & 0.15 \\ 0.1 & 0.4 & 0.5 \end{pmatrix} = \begin{pmatrix} 0.43 & 0.32 & 0.25 \\ 0.315 & 0.47 & 0.215 \\ 0.21 & 0.46 & 0.33 \end{pmatrix}$$

So the probabilities are 0.43 in Atlantis, 0.32 in Beachy and 0.25 in Coral.

Solution

(iii) *Either of the following two graphs was acceptable.*





Solution

(iv) This is the probability he is in each state multiplied by the chance of taking a flight if he is in that state.

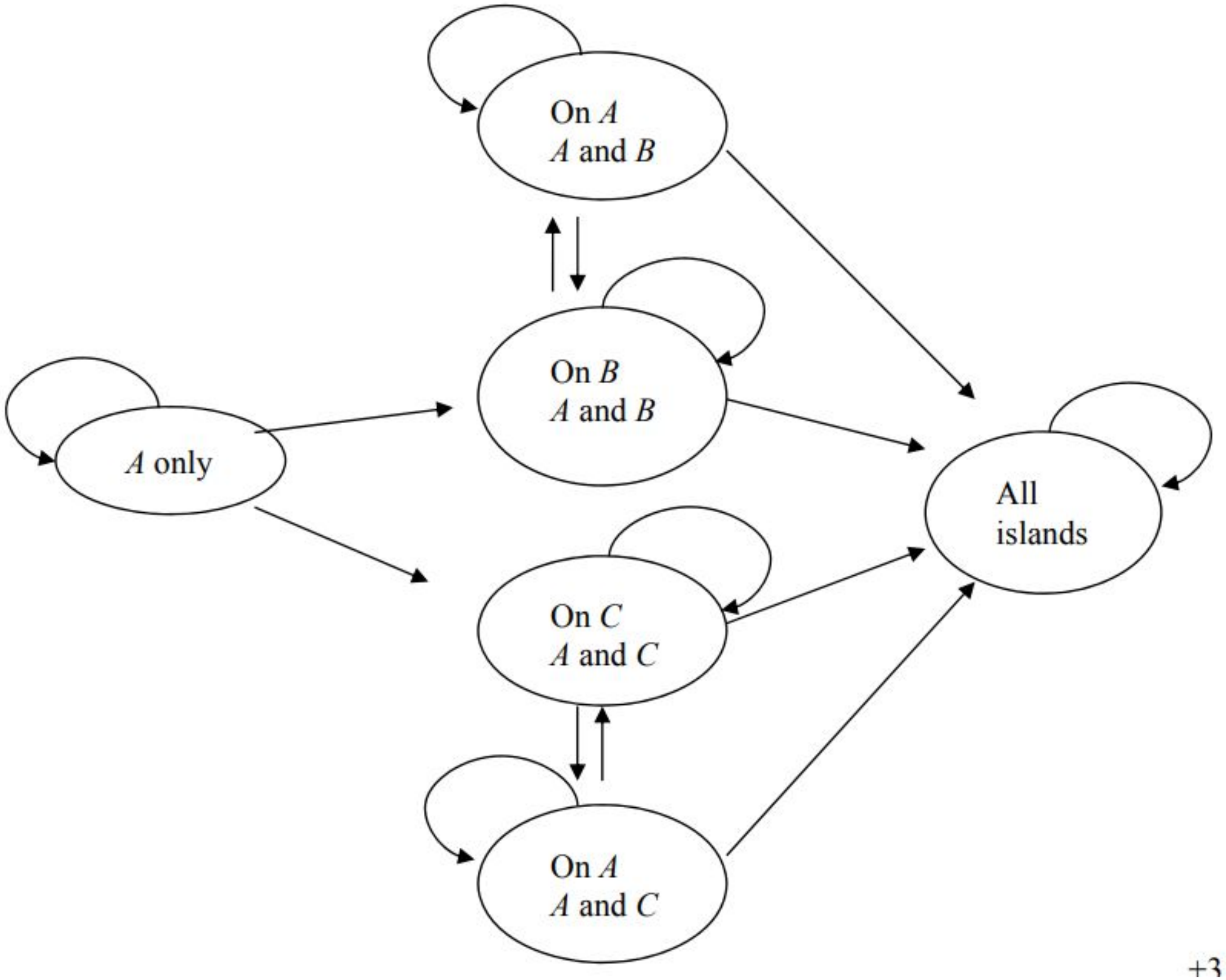
$$(0.326 \times 0.4) + (0.419 \times 0.4) + (0.256 \times 0.5) = 0.426.$$

(v) This needs six states:

{A visited only, On B (A and B visited), On A (A and B visited),
On C (A and C visited), On A (A and C visited), All islands visited}

Solution

(vi)



+3