Stochastic Processes and Markov Chains

Notes by Holly Hirst

Adapted from Chapter 5 of Discrete Mathematical Models

By Fred Roberts

Introduction

Stochastic Process: A sequence of events in which the outcome of the n^{th} event (also called a stage or trial) depends on some chance element and perhaps also outcomes from some set of the previous stages. If the number of possible outcomes at each stage is finite, the process is considered to be finite.

Memory-less Processes: Stochastic processes in which no information from previous stages is needed for the next stage. Example: coin tossing.

Markov Chains: Processes in which the outcomes at any stage depend upon the previous stage (and no further back).

Markov Chain Example 1: *Weather* – A study of the weather in Tel Aviv showed that the sequence of wet and dry days could be predicted quite accurately as follows.

If the current day is dry then there is a

.250 probability of having a wet day the next day

.750 probability of having a dry day the next day

If the current day is a wet day there this is a

.662 probability of having a wet day the next day

.338 probability of having a dry day the next day

Markov Chain Example 2: *Russian Roulette* – There is a gun with six cylinders, one of which has a bullet in it. The barrel is spun and then the gun is fired at a person's head. After each firing, the person is either dead or alive. If the person survives, the barrel is spun again and fired again. This is repeated until the person is dead.

Representing Markov Chains

Here is a formal definition:

A Markov Chain is a sequence of events for which

(1) *There is a finite set of outcomes, which includes all possible outcomes – more commonly called "states" – for all possible stages:* $U = \{u_1, u_2, \ldots, u_n\}$.
(2) *The probability that outcome* u_i *will occur at trial* $t+1$ *is known if we know the outcome of trial* t*, but is independent of* t*, i.e., the same for trial 2 as for trial 10. So we can consider the set of probabilities* $p_{i,j}$ *as the probability that if the outcome on any trial is* u_j *then the outcome on the next trial is* u_i.

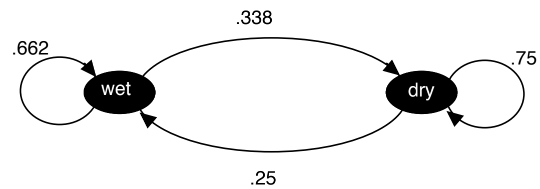
In other words, information about trials other than the current one or the trial number itself doesn't affect the probabilities of future events.

Weather Example (1): The set of outcomes is U = {wet, dry}, and the probabilities can be put in a *transition* matrix:

$$\begin{array}{ll} & \begin{array}{cc} \textit{wet now} & \textit{dry now} \end{array} \\ \begin{array}{l} \textit{wet later} \\ \textit{dry later} \end{array} & \begin{bmatrix} .662 & .250 \\ .338 & .750 \end{bmatrix} \end{array}$$

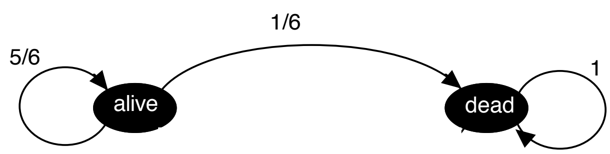
This matrix is a really nice way to characterize Markov chains. There is no ambiguity in the definition, provided one remembers that the *i,j* entry is the probability of going from state u_j to state u_i. Note also that the columns sum to 1, since each column lists probabilities for all possible future outcomes for a given current state. Be aware that the fact that the probabilities are listed column-wise is not always used; other references may list the entries in rows, which would change the matrix calculations that are described in what follows.

We can also draw a graph of the process:



Russian Roulette Example (2): The set of outcomes is U = {alive, dead}, and the probabilities are:

$$\begin{array}{ll} & \begin{array}{cc} \textit{alive now} & \textit{dead now} \end{array} \\ \begin{array}{l} \textit{alive later} \\ \textit{dead later} \end{array} & \begin{bmatrix} 5/6 & 0 \\ 1/6 & 1 \end{bmatrix} \end{array}$$



Making Predictions with Markov Chains

The main question we want to answer eventually is, "what happens over the long term?" In the weather example, what is the long-term probability of a dry or wet day? To do this we need to be able to calculate the probability of a wet or dry day two days, three days, four days, etc., after a wet or dry day. These are called *higher order transition probabilities* and are denoted $P_{i,j}^{(k)}$.

<u>Theorem 1</u>: If P is the transition matrix for a Markov chain, then $P_{i,j}^{(k)}$ is the *i, j* entry of P^k, i.e., the k^{th} power of the matrix.

Weather Example (1): The probability that a dry day will occur 3 days after a wet day is .53463 from entry 2,1 in the matrix:

$$\begin{bmatrix} .662 & .250 \\ .338 & .750 \end{bmatrix}^3 = \begin{bmatrix} .46537 & .39544 \\ .53463 & .60456 \end{bmatrix}$$

Raising this matrix to different powers can answer many questions about the situation being modeled.

Using Initial Conditions

Sometimes we know the initial state or the initial probability distribution, and we want to learn what happens in the future based upon that starting information. We can use matrix calculations to find the probability of being in a certain state several stages later.

Theorem 2: If P is the transition matrix for a Markov chain and v^0 is a vector of initial probabilities for being in the states (in the same order as in the matrix), then the matrix multiplication $v^k = P^k v^0$ gives the probabilities of being in the states after k stages.

Russian Roulette Example (2): Initially (assuming the person is alive), the probability vector is (alive, dead) = $(1,0)^T$. After 20 rounds, the probability vector becomes:

$$\begin{bmatrix} 5/6 & 0 \\ 1/6 & 1 \end{bmatrix}^{20} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} .02608 \\ .97392 \end{bmatrix}$$

So, we have .02608 probability of being alive after 20 rounds.

With the information at our disposal we can answer quite a few questions about what happens in the future based upon current events. However, in certain instances we can say a lot more. We will look at two kinds of Markov Chains with interesting properties.

Regular Markov Chains

Chains that have the property that there is an integer k such that every state can be reached from every other state in exactly k steps are called *regular chains*. These chains have two interesting properties:

Theorem 3: In a regular chain, some power of the transition matrix has all of its entries positive.

Theorem 4: The powers of the transition matrix approach a matrix with all columns the same. More over, this column vector – called the *fixed-point probability vector* – contains the long-term probabilities of being in each state.

In addition to the usual question, "if we start in state i, what is the probability we get to j in k steps?" we will ask the following question about regular chains:

What is the long term probability of being in state i?

Weather Example (1): The weather example is a regular chain. Let's look at some high powers of the transition matrix:

$$\begin{bmatrix} .662 & .250 \\ .338 & .750 \end{bmatrix}^{10} = \begin{bmatrix} .425251 & .425251 \\ .574749 & .574749 \end{bmatrix}$$

$$\begin{bmatrix} .662 & .250 \\ .338 & .750 \end{bmatrix}^{100} = \begin{bmatrix} .42517 & .42517 \\ .57483 & .57483 \end{bmatrix}$$

$$\begin{bmatrix} .662 & .250 \\ .338 & .750 \end{bmatrix}^{1000} = \begin{bmatrix} .42517 & .42517 \\ .57483 & .57483 \end{bmatrix}$$

We appear to be converging to a probability of having a wet day of .425 and a probability of having a dry day of .575.

Absorbing Markov Chains

A state that cannot be left is called an *absorbing state*. Chains that have at least one absorbing state and from every non-absorbing state it is possible to reach an absorbing state are called *absorbing chains*.

<u>Theorem 5</u>: As the number of stages approaches infinity in an absorbing chain, the probability of being in a non-absorbing state approaches 0.

Russian Roulette (2): Russian Roulette is an example of an absorbing chain. Note that very high powers of the matrix give:

$$\begin{bmatrix} 5/6 & 0 \\ 1/6 & 1 \end{bmatrix}^{20} = \begin{bmatrix} .02608 & 0 \\ .97392 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 5/6 & 0 \\ 1/6 & 1 \end{bmatrix}^{100} = \begin{bmatrix} .00000002 & 0 \\ .99999998 & 1 \end{bmatrix}$$

We will be interested in the following questions related to absorbing chains:

(1) What is the probability of ending up in absorbing state i when starting in state j?
(2) What is the expected number of times we will be in state j before being absorbed?
(3) What is the expected number of stages before being absorbed?

These can all be answered by studying the transition matrix.

For any absorbing Markov chain, the transition matrix can be rewritten in the following form:

$$\begin{array}{ll} & \quad \textit{absorbing} \quad \textit{nonabsorbing} \\ \textit{absorbing} & \begin{bmatrix} I & \qquad\qquad R \\ 0 & \qquad\qquad Q \end{bmatrix} \\ \textit{nonabsorbing} & \end{array}$$

where I is the identity matrix, 0 is the matrix of zeros, R and Q are submatrices that give the probabilities of going from nonabsorbing to absorbing states and the probabilities of staying in nonabsorbing states respectively.

The key ideas are listed in the following theorem:

Theorem 6: In an absorbing Markov chain, the matrix $N = (I - Q)^{-1}$ exists, and

(1) the expected number of times we are in state i given that we start in state j is given by the i,j entry of N;
(2) the expected number of steps before absorption given that the process starts in row i is the sum of column i of N;
(3) the probability of absorption in state i given that the process starts in state j is given by the i,j entry of the matrix RN.

Russian Roulette (2): The Russian roulette problem gives matrices:

$$R = \left[\tfrac{1}{6}\right], Q = \left[\tfrac{5}{6}\right]$$

This overly simple example gives:

$$(I - Q) = \left[1 - \tfrac{5}{6}\right] = \left[\tfrac{1}{6}\right]$$
$$N = (I - Q)^{-1} = [6]$$
$$RN = [1]$$

Interpreting these numbers gives us:

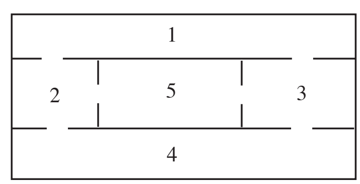
(1) We expect to be alive 6 stages given that we start out alive.
(2) We expect to be alive 6 stages before dying.
(3) The probability of dying given that we start out alive is 1.

Other Kinds of Markov Chains

These two special cases are not the only situations we could encounter. Chains can cycle between two sets of states, chains can have parts that act like regular chains and parts that act like absorbing chains. There is a rich theory behind the behaviors of Markov chains, and they remain a topic of interest to researchers today.

Problems

1. A rat is dropped into chamber 1 in the maze, and wanders through the chambers at random until it finds the cheese in chamber 5. Assuming that at any stage the rat chooses a door out of any chamber at random, what is the expected number of stages before the rat finds the cheese? What is the expected number of times that each chamber is entered? If the rat is particularly stupid and only sees the cheese half the time when entering chamber 5, how do these answers change?



2. Consider the maze from problem (1) again, this time assuming that there is no cheese, i.e., the mouse will wander at random indefinitely. Build the transition matrix, and investigate the process for this situation. Is it an absorbing chain? Is it regular? Answer all of the questions one can pose about this sort of chain.

3. A particle moves on a circle through points that have been marked 0, 1, 2, 3, 4 (in a clockwise order). The particle starts at 0. At each stage the particle moves one step clockwise (0 follows 4)

with probability q or one step counterclockwise with probability $1-q$. Let the position of the particle be the states of a Markov chain. Find the transition matrix and the equilibrium probabilities.

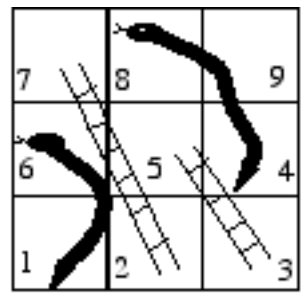
4. Consider the game of ping pong with the following states and transition matrix as pictured on the right:

1. Player A hits the ball toward B's side.
2. Player B hits the ball towards A's side.
3. Play is dead because A hit the ball out or hit the net.
4. Play is dead because B hit the ball out or hit the net.

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0 & .9 & 0 & 0 \\ .4 & 0 & 0 & 0 \\ .6 & 0 & 1 & 0 \\ 0 & .1 & 0 & 1 \end{bmatrix} \end{array}$$

The game starts with player 1 hitting the ball. What is the expected number of times the ball will be hit without error before play is dead? What is the probability that A will win the point (i.e., B will be the one to make an error)?

5. Play the games snakes and ladders on the board to the right as follows: Start on square 1, toss a fair coin and move 1 if heads and 2 if tails. On squares where a ladder is based, slide immediately up to the square at the top of the ladder. On squares where a snake rests its head, slide immediately down to the square at the tail end of the snake. Find the expected number of tosses to completion of the game. What is the probability that a person on the middle square completes the game without sliding back to 1?



6. Consider the following gambling game. A coin is tossed, with a bet placed on the outcome. Each person must start with \$2, and the first one with \$10 wins. The allowed bets are: Stake all if current cash is \$5 or less, stake enough to get to \$10 otherwise. What is the probability that a player gets to \$10? What is the expected number of tosses to losing (\$0) or winning (\$10)?

7. Three tanks are engaged in a battle. Tank A, when it fires, hits its target with probability 1/2. B hits its target with probability 1/3, and C will hit with probability 1/6. Initially, B and C fire at A and A fires at B. Once one tank it hit, the remaining tanks aim at each other. The battle ends when there is one or no tanks left. Analyze the outcomes of this battle. (Hint: Let the states be the possible subsets of tanks in action at any one time.)

8. Suppose an urn contains 2 unpainted balls to start. We choose a ball at random and flip a coin. If the ball is unpainted and the coin is heads we paint the ball red. If the ball is unpainted and the coin is tails, we paint the ball black. If the ball is already painted, we change the color of the ball (regardless of the coin toss) from red to black or black to red. Analyze this process. After 2 balls are painted, what is the probability that there will be 2 red balls? 2 black balls? 1 red and 1 black?

9. Consider an inventory system in which the sequence of events during each period is as follows: (1) We observe the inventory level i at the beginning of each period. (2) If $i \leq 1$, $(4-i)$ units are ordered. If $i \geq 2$, 0 units are ordered. The probability of demand in any period has been observed over time to be equal for 0, 1, or 2 units. Analyze the ordering process.

10. A company has 2 machines. During any given day, each machine that is working has a 1/3 chance of breaking down. If a machine breaks down during the day, it is sent to a repair facility and will be working 2 days after it breaks down. (Thus if a machine breaks during day 3, it will be

working again during day 5.) Analyze this system from the point of view of number of machines working at the beginning of any given day.

11. Each American family is classified as living in an urban, rural or suburban location. During a given year, 15% of all urban families move to a suburban location, and 5% move to a rural location. 6% of all suburban families move to an urban location, and 4% move to a rural location. 4% of all rural families move to an urban area and 6% move to a suburban area. If a family now lives in an urban location, what is the probability that it will live in an urban area 2 years from now? In a rural area? In a suburban area?

12. A college admissions office has modeled the path of a student through his or her college career as a Markov chain. Each state is observed as the start of the fall semester. If a student enters as a freshman how many years can he or she expect to spend as a student? What is the probability that a freshman graduates? If the college works really hard on freshman retention, reducing the number who quit down to 5%, how much does the probability that a freshman graduates increase?

Estimated percentages:
Freshmen: 10% stay freshmen and 10% quit
Sophomores: 10% stay sophomores and 5% quit
Juniors: 15% stay juniors and 5% quit
Seniors: 10% stay seniors and 5% quit

13. A baseball team consists of 2 stars, 13 starters, and 10 substitutes. For tax purposes the owner must "value" the players. The value of a player is defined as the total salary that a player will earn while on the team. Determine the value of the teams players. At the beginning of each season the non-retired players are classified into one of three categories:

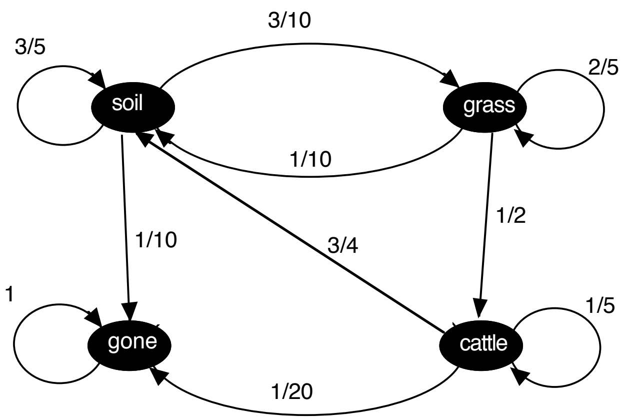
1: Star -- earns $1 million per year
2: Starter -- earns $400,000 per year
3: Substitute -- earns $100,000 per year

Here is the transition matrix as estimated from past data:

	1	**2**	**3**	**ret**
1	.5	.3	.15	.05
2	.2	.5	.2	.1
3	.05	.15	.5	.3

14. The manager of a factory claims that of the wastes from his plant, which are emptied into a nearby river, the majority are carried out to sea. Specifically, for a given molecule of mercury found the wastes, its probability of being swept out to sea within a day is .999. If that molecule is still around after a certain number of days, its probability of being swept out to sea on the next day is still .999. Once it is swept out to sea, we assume it cannot return. Suppose a particular molecule of mercury can be tagged and on the “t” th day we could record its location. Would this sequence of observances form a Markov chain? If so, build the transition matrix. Is it an absorbing chain? Is it regular? Answer all of the questions one can pose about this sort of chain.

15. Here is a graphical representation of a simple model for passage of a phosphorus molecule through a pasture ecosystem. Build the transition matrix, and investigate the process. Is it an absorbing chain? Is it regular? Answer all of the questions one can pose about this sort of chain.



16. Sociologists are concerned with the movement between different occupational classes from one generation to the next. Below is the transition matrix for such a situation modeled as a Markov chain, where the occupations have been classified as U (upper), M (middle) and L (lower) socioeconomic. The entries are the probabilities that the son of a man in occupation level *j* has occupation *i*. Build the graph, and investigate the process. Is it an absorbing chain? Is it regular? Answer all of the questions one can pose about this sort of chain.

$$\begin{array}{c c} & \begin{array}{ccc} U & M & L \end{array} \\ \begin{array}{c} U \\ M \\ L \end{array} & \begin{bmatrix} .448 & .054 & .011 \\ .484 & .699 & .503 \\ .068 & .247 & .486 \end{bmatrix} \end{array}$$

17. Three armies are engaged in a battle. The British will win a battle with probability 1/3. The Russians win with probability 1/4, and the French will win with probability 1/2. Initially, there is an alliance between the British and the Russians, i.e., the British and Russians will not engage each other in battle. Once one army is defeated the other two engage each other. The war ends when there is one or no army left. Analyze the outcomes of this war.

18. A college's faculty consists of assistant, associate, and full professors. The provost wants to get a feel for the overall salary structure of the university. At the beginning of each year the faculty who haven't retired or left are classified into:

1. assistant professor year 1 – earns \$43000 per year on average
2. assistant professor year 2 – earns \$44000 per year on average
3. assistant professor year 3 – earns \$45000 per year on average
4. assistant professor year 4 – earns \$46000 per year on average
5. assistant professor year 5 – earns \$47000 per year on average
6. assistant professor year 6 – earns \$48000 per year on average
7. assistant professor post tenure – earns \$50000 per year on average
8. associate professor – earns \$54000 per year on average
9. full professor – earns \$68000 per year on average

Here is the transition matrix as estimated from past data:

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**leave**	**retire**
1		.9								.1	
2			.9							.1	
3				.8				.01		.19	
4					.8			.02		.18	
5						.8		.03		.17	
6							.05	.65		.30	
7							.04	.93		.01	.02
8								.61	.31	.04	.04
9									.81	.02	.17

What is the expected number of years at each salary level for an entry-level assistant professor? What is the expected total earnings potential for an entry level assistant professor if 12% of the salary is set aside for savings?

19. A major problem for a hospital is managing the database containing patient records. The General Hospital has considered two different policies and needs help with its decision:

- policy 1: dispose of a patient's records if he or she has not reentered the hospital in the last 5 years
- policy 2: dispose of a patient's records is he or she has not reentered the hospital in the last 10 years.

The following information is available: If a patient has been hospitalized in the last year, there is a 30% chance that he or she will reenter the hospital during the next year. For each passing since hospitalization, chances go down:

number of years since hospitalization	1	2	3	4	5	6 or more
percent chance of hospitalization	20	10	5	3	2	1

Assume that the hospital admits an average of 10000 patients to the hospital in any given year. Estimate the number of patient records that will be in the system under each policy.

20. At the beginning of each day a patient in a hospital is classified into one of three conditions: good, fair or critical. At the beginning of the next day, a patient will be discharges in one of three conditions: improved, unimproved or dead. The transition probabilities are:

	Good	Fair	Critical			Improved	Unimproved	Dead
Good	.65	.20	.05		Good	.06	.03	.01
Fair	.5	.3	.12		Fair	.03	.02	.03
Critical	.51	.25	.20		Critical	.01	.01	.02

Consider a patient who enters the hospital in good condition. On average, how many days does this person spend in the hospital? What is the probability that he or she will die? What fraction of the patients who enter the hospital in critical condition leave the hospital in improved condition? On average, the hospital's daily admissions are 20 patients in good condition, 10 in fair condition and 10 in critical condition. How many patients of each type would you expect to see in the hospital?

21. Freezco, Inc., sells refrigerators. The company has issued a warranty on all refrigerators that requires free replacement of any refrigerator that fails before it is three years old. We know the following:

- 3% of all new refrigerators fail during the first year of operation
- 5% of all one year old refrigerators fail during the second year of operation
- 7% of all two year old refrigerators fail during the third year of operation
- A replacement refrigerator is not covered by the warrantee.

Predict the fraction of all refrigerators that Freezco will have to replace. Suppose that it costs Freezco $500 to replace a refrigerator that fails and that Freezco sells approximately 10000 refrigerators per year. If the company reduced the warranty period to 2 years, how much money in replacement costs will it save?

22. A taxi company has divided the city into 4 zones: Northside, Downtown, Uptown and Southside. Their data indicate that:

- of Northside fares, 10% stay in Northside, 40% go Downtown, 40% go Uptown and the rest go to Southside.
- of Downtown fares, 60% stay in Downtown, 10% go to Uptown, and the rest are equally likely to go to either of the other two zones.
- of the Southside fares, 40% stay in Southside, 15% go to Uptown, 15% go to Downtown, and the rest go to Northside
- of the Uptown fares, 55% stay in Uptown, and the rest are evenly divided between the other three zones.

a. If a taxi starts Downtown, what is the probability that it will be Downtown after 5 fares?
b. Where should a taxi start to maximize the chances that it will be Uptown after 4 fares?
c. If the taxi company starts with 25% of the taxis in each zone, what happens by the end of the day (assuming approximately 20 fares per day)?
d. What is the long term distribution of taxis?

23. The 1980 census yielded the following information about occupations of fathers and sons:

- Of the fathers who were in professional jobs, 63% of their sons were in professional jobs and 25% were in service jobs.
- Of the fathers who were in service jobs, 30% of their sons were in professional jobs and 45% were in service jobs.
- Of the fathers who were in manufacturing jobs, only 18% of their sons followed their profession and 41% went into service jobs.

a. What is the probability that a man's great grandson will be in the same profession as he is?
b. If the population was distributed as 32% in manufacturing and 27% in service in 1980 and these trends hold, what is the long-term distribution of professions for men?

24. After farmland has been abandoned, various species of grass arrive to reclaim the land. Based on a study of 30 old fields in Oklahoma, a succession of five dominant or co-dominant grass species

were identified. *Helianthus annus* (sunflower) and *Digitaria sanguinalis* identify the first stage, *Aristida oligantha* (three-awn grass) the second stage, *Aristada basiramea* (red three-awn) the third stage, *Eragrostis secundiflora* (lovegrass) the fourth stage and *Andropogen scoparius* (little bluestem) and *Bouteloua curtipendula* (sidecoats grama) the final, climax stage. We wish to model the percent of area dominated by each type of grass for a period of 10 years. Initially, 85% of the field is in the first successional stage, 10% in the second stage, and 5% in the third stage. For each time step, the flow rate is as follows:

	From 1	From 2	From 3	From 4
To 2	0.8	--	--	--
To 3	0.2	1.0	--	--
To 4	--	--	0.7	--
To 5	--	--	--	0.8

25. (A slightly different kind of problem: Leslie population models) Consider the table below, giving the birth and death rates by age group for a population of small woodland animals, where the age ranges are given in months.

	0-3	3-6	6-9	9-12	12-15	15-18
Birth rate	0	.3	.8	.7	.4	0
Death rate	.4	.1	.1	.2	.4	1

What is the average life expectancy for one of these critters? If the current population is distributed as below, what will the distribution be in 1 month? 6 months? 1 year? 20 years?

	0-3	3-6	6-9	9-12	12-15	15-18
Population	14	8	12	4	2	1

What is the long-term population growth rate the population? What is the long-term age distribution (as a percent of the total population) of the animals? Determine a uniform harvesting rate that will ensure a stable population distribution.

26. (Investigate the background theory) Expand upon this treatment of Markov Chains by writing up proofs of the six theorems stated in this document.