

2014

## **The Application of Stochastic Processes in the Sciences: Using Markov Chains to Estimate Optimal Resource Allocation, Population Distribution, and Gene Inheritance**

Nate DeMaagd  
*Grand Valley State University*

Follow this and additional works at: <https://scholarworks.gvsu.edu/honorsprojects>

---

### **ScholarWorks Citation**

DeMaagd, Nate, "The Application of Stochastic Processes in the Sciences: Using Markov Chains to Estimate Optimal Resource Allocation, Population Distribution, and Gene Inheritance" (2014). *Honors Projects*. 290.

<https://scholarworks.gvsu.edu/honorsprojects/290>

This Open Access is brought to you for free and open access by the Undergraduate Research and Creative Practice at ScholarWorks@GVSU. It has been accepted for inclusion in Honors Projects by an authorized administrator of ScholarWorks@GVSU. For more information, please contact [scholarworks@gvsu.edu](mailto:scholarworks@gvsu.edu).

# The Application of Stochastic Processes in the Sciences

*Using Markov Chains to Estimate Optimal Resource Allocation,  
Population Distribution, and Gene Inheritance*

NATE DEMAGD

## ABSTRACT

The consideration of quantitative data is often required to perform research in both the physical and social sciences and, as a result, researchers must at times include mathematical models in their studies in order to use the data in a meaningful way. The use of stochastic, or probabilistic, methods is particularly useful for those who are attempting to predict future phenomena; for example, an economist may wish to advise a corporation by forecasting long-term prices of the corporation's assets, and an ecologist may wish to predict the migration of animal populations between regions and its effect on local ecosystems. To our dismay, such problems involving many uncontrollable factors and an aspect of uncertainty cannot be easily solved with the "simple" mathematical models we learn in our introductory algebra and calculus courses. Instead, we require the use of probabilities and the utilization of data that change over time. In this article we will present the method of stochastic processes (particularly Markov chains) in general, aiming to provide a working knowledge of the theory behind the method to be used. We will then demonstrate the wide applicability of Markov chains by introducing the method to several relevant fields of study, namely topics in economics and biology.

#### ACKNOWLEDGMENTS

I would like to thank Dr. Jiyeon Suh for editing this paper and answering my many questions. I also thank Dr. Laudo Ogura, Dr. Noor Ghiasvand, and Dr. Sonia Dalmia for their input pertaining to the applications of the topic discussed here.

In modern scientific experiments and analyses, mathematical methods are often employed in order to make predictions or explain how particular phenomena occur. Unfortunately this process is often difficult, as the collection of sufficient data may be tedious, expensive, or impossible. Thus, we must resort to statistical and probabilistic, or stochastic, techniques in order to infer generalized results and make accurate predictions. One such technique is a relatively simple process called Markov chains, named after the Russian mathematician Andrey Markov. In this process, we have a set of “states” that can be thought of as a set (or web) of possible positions in a Markov chain, where the chain “jumps” from position to position according to the process discussed below. The following definitions<sup>1</sup> provide a more formal description of the chain, and what then follows is a simple example to demonstrate the process.

**Definition.** The **state space**  $S$  of a Markov chain is a finite or countably infinite set of states with values of  $X_i$ .

$$S = \{X_1, X_2, \dots, X_n\} \text{ for some } n \in \mathbb{N}, \text{ or}$$

$$S = \{X_1, X_2, \dots\} \text{ when } S \text{ is countably infinite.}$$

To begin the process, our chain must begin at one of the states in the state space. Where we start is likewise a matter of probability; our certainty of which state we begin in can range from 0% to 100%. As such, there exists an initial distribution where each state is given a probability of being the initial state in our chain.

**Definition.** The **initial distribution**  $\pi_0$  is the probability distribution at time 0. For all  $i \in S$ , we denote  $\pi_0(i)$  to be the probability  $\mathbb{P}\{X_0 = i\}$  that the Markov chain starts in state  $i$ .

The initial distribution is a function taking  $S$  into the interval  $[0, 1]$  such that

$$\pi_0 \geq 0 \text{ for all } i \in S$$

and

$$\sum_{i \in S} \pi_0(i) = 1.$$

That is, a state’s probability of being our initial state cannot be less than 0, and the sum of the probabilities associated with all states in  $S$  is 1, or 100%. As an example, suppose that we have a state space  $S = \{1, 2, 3, 4, 5, 6\}$  with an initial distribution  $\pi_0 = \{\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\}$ . We note that the two conditions above are satisfied, and that each state has an equal probability of being the initial state. To determine the initial state  $X_0$ , we could roll a die and assign a roll of 1 to state space 1,

---

<sup>1</sup>Adapted from the book by P. Chang.

a roll of 2 to state space 2, and so on. We could also use a random number generator, which will be introduced in a different context later on.

Of course, the probabilities for each initial position need not be equal. For instance, consider once again our state space  $S = \{1, 2, 3, 4, 5, 6\}$ , and define the initial distribution as  $\pi_0 = \{\frac{1}{3}, 0, \frac{1}{6}, \frac{1}{4}, \frac{1}{6}, \frac{1}{12}\}$ . Notice that the two conditions for an initial distribution are still satisfied. Since each state does not have an equal probability of being the initial state of the Markov chain, we are clearly unable to use a die to randomly decide the initial state. This case would require the use of something like a random number generator which, again, will be introduced later since it is not of direct importance to the understanding of Markov chains. We next consider the subsequent steps of the Markov chain after the initial state.

Once we have our initial state determined, we must then determine which state the chain will “jump” to next. Between each state in  $S$  is a probability that determines the chances our chain will transition from its current position to the other states in  $S$ .

**Definition.** The **probability transition rule** is specified by giving a matrix  $P = (P_{ij})$ . If  $S$  contains  $N$  states, then  $P$  is an  $N \times N$  matrix.  $P_{ij}$  is the conditional probability: Given that the chain is in state  $i$  at time  $n$ , we denote the probability that the chain jumps to state  $j$  at time  $n+1$  as

$$P_{ij} = \mathbb{P}\{X_{n+1} = j | X_n = i\}.$$

For simplicity, we can alternatively express this as  $P(i, j)$ . This probability can be a function of  $i$  and  $j$ , or of  $n$ . If, however, the chain is assumed to be time homogenous (that is, if the probabilities expressed by  $P$  are constant with regard to changes in time), the probability of the chain being in a particular state is not dependent on time. The probabilities of moving from state to state can be expressed as a matrix whose rows sum to 1 (which is interpreted as all probabilities adding to 100%).

Given that these definitions are quite abstract, a simple example will help to solidify the concepts discussed above. Consider a frog jumping between 3 lily pads in Andrey Markov’s backyard pond, which is crudely depicted in Figure 1 on page 6. The state space is then given as  $S = \{1, 2, 3\}$ . We then define an initial distribution for the three lily pads as

$$\pi_0 = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right).$$

These numbers were decided arbitrarily, and they can be anything so long as their sum is 1. (The same is true for the following matrix, so long as the rows add to 1.) An example probability

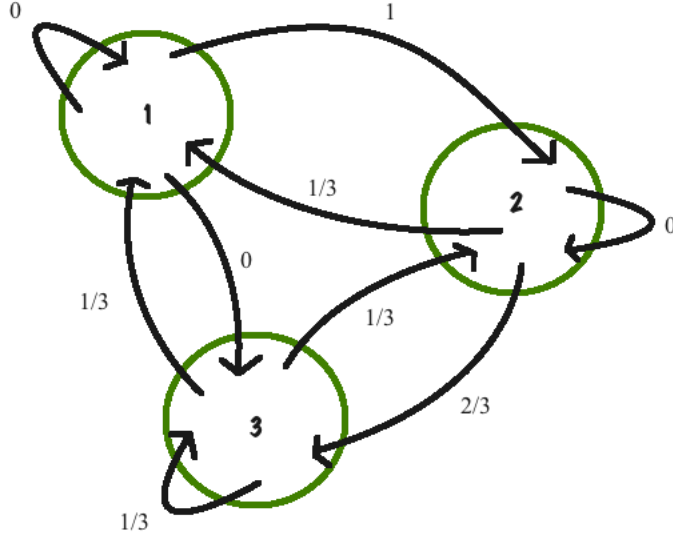


FIGURE 1. Markov's Lilies.

transition matrix for our lily pad example can be

$$\mathbb{P} = \begin{bmatrix} 0 & 1 & 0 \\ 1/3 & 0 & 2/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

The initial position  $X_0$  of the frog is determined by the initial distribution  $\pi_0$ . We can determine this by generating a random number  $U_0$  such that  $U_0 \in [0, 1]$  and using

$$X_0 = \begin{cases} 1, & \text{if } 0 \leq U_0 \leq \frac{1}{2} \\ 2, & \text{if } \frac{1}{2} < U_0 \leq \frac{3}{4} \\ 3, & \text{if } \frac{3}{4} < U_0 \leq 1, \end{cases}$$

where 1, 2, and 3 are the lilies. For example, suppose we have a random number generator that gives  $U_0 = 0.8539$ . This means the initial state of our frog will be lily (3). The next state of the frog,  $X_1$ , is then given by the probability matrix above. In our example where  $X_0 = 3$ ,  $X_1$  will be determined by the generation of a random number  $U_1 \in [0, 1]$  where

$$X_1 = \begin{cases} 1, & \text{if } 0 \leq U_0 \leq \frac{1}{3} \\ 2, & \text{if } \frac{1}{3} < U_0 \leq \frac{2}{3} \\ 3, & \text{if } \frac{2}{3} < U_0 \leq 1, \end{cases}$$

according to row (3) of the matrix. Suppose, then, that  $U_1 = 0.1247$  so that  $X_1 = 1$ . Then,  $X_2$  is determined by row (2) in the matrix, and so on.

Here, we notice something that is common to Markov chains in general. When our frog is situated on one lily, its next position is determined only by the probabilities associated with its current position and its possible next positions; the chain is “memoryless” in that as soon as the frog lands on a lily, its past positions on other lilies does not have any bearing on the frog’s future positions.

**Definition.** A process  $X_0, X_1, \dots$  satisfies the **Markov Property** if

$$\mathbb{P}\{X_{n+1} = i_{n+1} : X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\} = \mathbb{P}\{X_{n+1} = i_{n+1} : X_n = i_n\}$$

for all  $n \in \mathbb{N}$  and for all  $i_0, i_1, \dots, i_{n+1} \in S$ .

In other words, the state  $X_{n+1}$  is determined only by  $X_n$  and not  $X_{n-1}, X_{n-2}, \dots, X_0$ . Again, it is clear that the Markov property holds for our lily pad example. If the frog is situated at lily (3) for some state  $X_{i-1}$  in the Markov chain, we see that the state  $X_i$  is determined only by probabilities associated with lily (3). Insofar as the next move is concerned, we do not care where the frog was at states  $X_{i-2}, X_{i-3}, \dots, X_0$ .

**The Basic Limit Theorem.** Let  $X_0, X_1, \dots$  be an irreducible, aperiodic Markov chain having a stationary distribution  $\pi_0(\cdot)$ . Let  $X_0$  have the distribution  $\pi_0$ , an arbitrary initial distribution. Then,  $\lim_{n \rightarrow \infty} \pi_n(i) = \pi(i)$  for all states  $i$ . To make sense of the basic limit theorem, we must define some terms.

**Definition.** A **stationary distribution** is a distribution  $\pi$  on  $S$  such that if a Markov chain starts out with an initial distribution  $\pi_0 = \pi$ , then  $\pi_1 = \pi$ . Moreover, if the initial distribution is  $\pi$ , then  $\pi_n = \pi$ . Simply stated, the chain retains the distribution  $\pi$  indefinitely. We now describe some terms pertaining to the movement from one state space to another.

**Definition.** A state  $j$  is **accessible** from state  $i$  if  $\mathbb{P}(i, j) > 0$ . For example, refer once again to the lily pad example in Figure 1. We see that lily (1) is accessible from lily (3) since the probability that the frog moves from (3) to (1) is greater than 0. Conversely, lily (3) is not accessible from lily (1) since the probability that the frog moves from (1) to (3) is 0. If a pair of states are accessible from one another, they are said to communicate.

**Definition.** A state  $i$  **communicates** with a state  $j$  if  $j$  is accessible from  $i$  and  $i$  is accessible from  $j$ . Lilies (1) and (2) communicate since (1) is accessible from (2) and (2) is accessible from (1). Lilies (1) and (3) do *not* communicate since, while (1) is accessible from (3), (3) is not accessible from (1).



**Definition.** A Markov chain is **irreducible** if all pairs of states communicate. Our lily pads are not irreducible because all states (lilies) do not communicate. We can, however, define a probability transition matrix that is irreducible. An example of this would be the matrix

$$\mathbb{P} = \begin{bmatrix} 3/4 & 1/8 & 1/8 \\ 1/3 & 1/3 & 1/3 \\ 1/6 & 2/3 & 1/6 \end{bmatrix}.$$

Clearly, this chain is irreducible since any two given states communicate with one another. This can be quickly recognized by the fact that all of entries of this matrix are greater than 0.

**Definition.** The **period** of a state  $i$  is defined as

$$d_i = \gcd\{n : \mathbb{P}^n(i, i) > 0\}.$$

We can better understand the concept of a chain's periodicity if we consider a simple example. Let a Markov chain be given by the probability transition matrix

$$\mathbb{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

We can see by this matrix that if we are in state (1) at move  $n$ , then we will be in state (2) at move  $n + 1$  without exception. If we are in state (2) at move  $n$ , then we will be in state (1) at move  $n + 1$  without exception. We will alternate between state (1) and state (2) indefinitely with each subsequent move. Notice here the period of this chain is 2 since it takes exactly 2 moves to return once again to our initial position. In other words, if we are in state (1), then it will take exactly 2 moves to return to state (1). If we are in state (2), then it will take exactly 2 moves to return to state (2). Hence, a period of 2.

If the period of a Markov chain is 1, then it is said to be **aperiodic**.

**Definition.** The state  $i$  is **recurrent** if  $\mathbb{P}_i\{T_i < \infty\} = 1$ , where  $T_i$  is the first hitting time, which is defined as  $T_i = \inf\{n > 0 : X_n = i\}$ . In simpler terms, if we start at state  $i$ , then it is possible for the chain to eventually return to state  $i$ .

Another example: **Ehrenfest Chains**. Consider  $d$  balls in 2 urns and a Markov chain  $\{X_0, X_1, \dots\}$  with state space  $S = \{0, 1, \dots, d\}$ . We define the state  $X_n$  of the chain to be the number of balls contained in Urn 1 at time  $n$ . At each time, we randomly choose a ball from the two urns and switch it to the other urn. So, if the randomly selected ball is in Urn 1, it is taken out and placed in

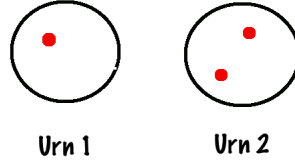


FIGURE 2. Ehrenfest Chain example with  $d = 3$  and  $X_0 = 1$ .

Urn 2. If it is Urn 2, then it is taken out and placed in Urn 1. Thus, the probability  $\mathbb{P}(i, i-1) = i/d$  and  $\mathbb{P}(i, i+1) = \frac{d-i}{d}$ .

As an example, let  $d = 3$ , and let  $X_0 = 1$ . This means we have 3 balls, 1 of which is located in Urn 1. This is shown in Figure 2 on page 9.

The chain can then be expressed as

$$\mathbb{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 2/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

These are the probabilities that, given we are in one state  $X_i$  (which can be described as having either 0, 1, 2, or 3 balls in Urn 1), what the next state  $X_{i+1}$  will be. The rows indicate the number of balls in Urn 1 (i.e. the first row is for 0 balls in Urn 1, the second row is for 1 ball in Urn 1, the third row is for 2 balls in Urn 1, and the fourth row is for 3 balls in Urn 1), while the columns indicate the number of balls in Urn 1 in the next step.

For example, if we have 1 ball in Urn 1, there is a  $\frac{1}{3}$  chance that the next state will have 0 balls in Urn 1 and a  $\frac{2}{3}$  chance that the next state will have 2 balls in Urn 1. We determine this by seeing that there is currently one ball in Urn 1, so we refer to row 2 of the matrix. We then see the respective probabilities for the number of balls in Urn 1 for the next turn by referring to the columns of the matrix.

We now have a theoretical background of Markov chains, and have seen a couple examples and variations of this process. What follows is an exploration of the applications of Markov chains in the sciences. To begin, we will consider the ways in which Markov chains might help us to efficiently govern the production of pollution in an economy.

There are several methods the government can use in order to control what companies or industries produce what levels of pollution. One of the simplest methods involves nothing more than dictating for each company the exact level of pollution it is allowed to produce. However, this method is inefficient in that it may inaccurately estimate what levels are cost-effective for these industries. This has led to the introduction to what are called “cap-and-trade” regulatory practices. With this method, a pollution goal or limit is imposed on all industries, and pollution allowance certificates are distributed to the industries. Unlike a strict cap, however, these certificates may be bought, sold, or retained by the companies, allowing each company to produce at the level most efficient for its own production needs under the restrictions imposed by a regulatory agency. That is, these allowances can be traded between companies according to their individual needs while still meeting the goals set forth by the government in terms of reducing pollution output. The Coase theorem states that, in the presence of well-defined property rights and negligible transaction costs, the<sup>2</sup> equilibrium allowance distribution will result regardless of the initial distribution of allowances among the firms involved. However, in reality, these assumptions are not realistic: it may be difficult for firms to locate other firms willing to buy or sell allowances, there are costs involved with negotiating prices, and so on. It would thus be beneficial to have the power to calculate the final equilibrium distribution and then allocate the allowances accordingly, without the need for firms to go through the often difficult and costly process of trading the permits. We can use the idea of stochastic processes to calculate these equilibria. We look first at the general case, then follow with an example.

Suppose we have  $n$  firms that desire pollution allowances. The value of buying, selling, or retaining the allowances can be represented in an  $n \times n$  matrix  $A$ , where each entry  $a_{ij}$  represents the value of company  $i$  selling an allowance to company  $j$ . (Note the diagonal entries  $a_{ii}$  then give the value of company  $i$  retaining an allowance for its own use.) In order to calculate the equilibrium allocation, we must transform  $A$  into a transition matrix  $T$ , where each entry  $t_{ij}$  is the *relative* value of an allowance being bought, sold, or retained. In other words, the entries of  $T$  will give the relative benefit of taking one action (buying, selling, or retaining) against the other actions a firm can take with the allowances it possesses. The entries of  $T$  are calculated by weighting the entry  $a_{ij}$  against all other entries in the column for firm  $j$ . Then,

$$t_{ij} = \frac{a_{ij}}{\sum_{i=1}^n a_{ij}}.$$

---

<sup>2</sup>With the assumption that the Coase theorem holds, a single, unique allocation of pollution permits will result. We will see later in our analysis that there is a possibility for many (infinite) efficient allocations.

Each column entry is divided by the sum of all the entries in the column to give the relative value. Each entry of  $T$  now takes a value  $0 \leq t_{ij} \leq 1$ , so we can think of the entries of  $T$  as the probability that the action represented by  $t_{ij}$  is taken given its corresponding value  $a_{ij}$  in  $A$ . Notice that, unlike with the general theory discussed above, our columns, as opposed to rows, will sum to 1. This is due to our desire to right multiply by a column vector as opposed to left multiply by a row vector; the process would work either way. If we right multiply this matrix by a vector with entries denoting the number of allowances given to each firm, we find the distribution of allowances after one “round” of trading in the market. We can then use the resulting vector and right multiply it once again to find the distribution of allowances after two rounds of trade. This process can be repeated until our input vector yields the same output vector, indicating that an equilibrium distribution has been achieved. This implies we wish to find a vector  $\mathbf{x}$  such that

$$T\mathbf{x} = \mathbf{x}. \quad (1)$$

We can choose any initial vector  $\mathbf{x}$  and find the equilibrium solution by writing a program that runs the vector through a loop using our transition matrix  $T$ . However, we can also solve for the equilibrium directly using the following process. Notice that equation 1 is equivalent to

$$\begin{aligned} T\mathbf{x} - \mathbf{x} &= \mathbf{0} \\ (T - I_n)\mathbf{x} &= \mathbf{0}, \end{aligned}$$

where  $I_n$  is the  $n \times n$  identity matrix. To solve this system, we row reduce the matrix  $(T - I_n)$ . Assuming that the Coase theorem holds, we know that our system is irreducible since all firms will be able to trade with one another. Durrett (2010) has proven that an irreducible Markov chain has one free variable, denoted here as  $x_i$ . Thus, we have

$$\begin{aligned} \mathbf{x} &= x_i \mathbf{c} \\ &= x_i \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \\ &= \begin{bmatrix} x_i c_1 & x_i c_2 & \dots & x_i c_n \end{bmatrix}^\top \end{aligned}$$

for some  $i \in \{1, 2, \dots, n\}$  and vector  $\mathbf{c}$ . The solution provided here is not useful to us, as it is simply a vector with constant entries multiplied by a free variable  $x_i$ , which can take any value and

thus provide us with infinite solutions. However, recall that this is ultimately the solution to the system  $A\mathbf{x} = \mathbf{x}$ . So, we can use the fact that  $x_i$  is free to find the proportion of allowances that should be allocated to each firm by the following process.

If we take the sum of the entries of  $\mathbf{x}$  and set it equal to 1, we can solve for the free variable  $x_i$ . This gives

$$x_i(c_1 + c_2 + \cdots + c_n) = 1,$$

which yields

$$x_i = \frac{1}{\sum_{k=1}^n c_k}.$$

We can then solve for each entry of

$$\frac{1}{\sum_{k=1}^n c_k} \mathbf{c},$$

which will give the equilibrium proportion of allowances that should be allocated to each of the firms.

**Example.** Suppose we have a market for pollution allocations with 3 firms. Let the values (in thousands of dollars) of buying, selling, and retaining allowances for each of the three firms be given by the matrix

$$A = \begin{bmatrix} 3 & 2 & 20 \\ 5 & 10 & 5 \\ 2 & 8 & 15 \end{bmatrix}.$$

The transition matrix  $T$  giving the relative values of trading or retaining the allowances is then

$$\begin{aligned} T &= \begin{bmatrix} 3/10 & 2/20 & 20/40 \\ 5/10 & 10/20 & 5/40 \\ 2/10 & 8/20 & 15/40 \end{bmatrix} \\ &= \begin{bmatrix} .3 & .1 & .5 \\ .5 & .5 & .125 \\ .2 & .4 & .375 \end{bmatrix}. \end{aligned}$$

We now wish to find the vector  $\mathbf{x}$  such that  $T\mathbf{x} = \mathbf{x}$ . Recall that this implies we wish to solve the system  $(T - I_3)\mathbf{x} = \mathbf{0}$ . We then have

$$\begin{bmatrix} -.7 & .1 & .5 \\ .5 & -.5 & .125 \\ .2 & .4 & -.625 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The reduced row echelon form of this system is

$$\begin{bmatrix} 1 & 0 & -.875 \\ 0 & 1 & -1.125 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

giving the solution

$$\mathbf{x} = x_3 \begin{bmatrix} 0.875 \\ 1.125 \\ 1 \end{bmatrix}.$$

We aim to find the proportion of allowances that should be allocated to each of the three firms, so we solve

$$0.875x_3 + 1.125x_3 + x_3 = 1$$

$$(0.875 + 1.125 + 1)x_3 = 1$$

$$3x_3 = 1$$

$$x_3 = \frac{1}{3}.$$

So, we find that our solution is

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7/24 \\ 3/8 \\ 1/3 \end{bmatrix}.$$

This means that firm 1 will likely hold approximately 29.2% of the allowances, firm 2 about 37.5%, and firm 3 about 33.3%. This sort of analysis can be applied to many other topics in business, finance, and economics in order to predict the outcomes of activities like stock trading. We will now explore an application of Markov chains to genetics.

The survival of a species is clearly dependent on the survival of the organisms of that species. In order for the species to survive, it is important that the organisms have some degree of variability in the adaptations they possess. In this way, a single virus or change in the environment will not likely kill every member of the species: various phenotypes, resulting from the existence of various genotypes, allow a species to withstand a variety of phenomena that would otherwise render it extinct. For simplicity, we will consider here only a single gene with several alleles. For instance, we could be considering a gene for color in a species that may result in any one of several different colors in the organisms of that species. We will show that a shift in the environment that increases the survivability of organisms with a particular allele will lead to a shift in the frequency of organisms with this allele.

First, suppose we are considering a gene  $A$  with alleles  $A_1$ ,  $A_2$ ,  $A_3$ , and  $A_4$ , and suppose our population consists of 500 organisms in a particular region. Further, we will assume that current environmental factors of the region favor allele  $A_1$  over other alleles. So, we will assume that most organisms in our population exhibit allele  $A_1$ . Also, assume that  $A_4$  does not yet exist; that is,  $A_4$

will arise as a mutation of gene  $A$ . We thus let

$$\mathbf{a} = \begin{bmatrix} 495 \\ 3 \\ 2 \\ 0 \end{bmatrix}$$

denote the number of organisms with each allele. As noted above, organisms with allele  $A_1$  are most well-suited for survival in this environment relative to those with alleles  $A_2$  and  $A_3$ , which is shown in their relatively large numbers. We may assume, then, that this distribution approximates the current equilibrium distribution of the alleles in our population.

Now, suppose the gene produces a mutation resulting in a new allele,  $A_4$ , and suppose organisms that inherit this allele are more suited for survival than those with other alleles. This would result in a change in the transition matrix denoting the inheritance (reflecting the change in relative survivability) of these alleles. Let the new transition matrix that accounts for the additional allele be given by

$$T = \begin{bmatrix} .4 & .2 & .21 & .04 \\ .001 & .12 & .01 & .006 \\ .002 & .005 & .09 & .004 \\ .597 & .675 & .69 & .95 \end{bmatrix}.$$

Notice that the last row of  $T$  indicates that many organisms who survive and reproduce inherit  $A_4$  and pass it on to their offspring. We will now find the distribution of alleles among the population in the manner we have seen before, and by applying a new method that involves matrix multiplication. First, we use the familiar method described above.

We solve for the equilibrium solution for allele distribution by first solving  $(T - I_4)\mathbf{x} = \mathbf{0}$  for  $\mathbf{x}$  as before. This gives us the result

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_4 \begin{bmatrix} 0.0706 \\ .0070 \\ .0046 \\ 1 \end{bmatrix}.$$

We wish for each  $x_i$  to give the proportion of the population with allele  $i$ , so we let

$$(0.0706 + 0.0070 + 0.0046 + 1)x_4 = 1,$$



which yields  $x_4 \approx 0.924$ . Solving for the remainder of the  $x_i$ s gives

$$\mathbf{x} \approx \begin{bmatrix} 0.065 \\ 0.006 \\ 0.004 \\ 0.924 \end{bmatrix}.$$

This solution for  $\mathbf{x}$  can be multiplied by our total population of 500 organisms to determine the approximate number of organisms with each of the four alleles once the equilibrium is reached. We find

$$\mathbf{a}^* \approx \begin{bmatrix} 33 \\ 3 \\ 2 \\ 462 \end{bmatrix}.$$

This solution makes sense since we know that organisms with the new, fourth allele are most fit for survival.

Now, we will see a new method to calculate this equilibrium. First, notice that if we find the product  $T\mathbf{a}$ , we will obtain the number of organisms with each allele after the passage of one generation. We find

$$\begin{bmatrix} .4 & .2 & .21 & .04 \\ .001 & .12 & .01 & .006 \\ .002 & .005 & .09 & .004 \\ .597 & .675 & .69 & .95 \end{bmatrix} \begin{bmatrix} 495 \\ 3 \\ 2 \\ 0 \end{bmatrix} \approx \begin{bmatrix} 200 \\ 0 \\ 1 \\ 299 \end{bmatrix} = \mathbf{b}.$$

This indicates that about 200 organisms exhibit allele  $A_1$  and 299 exhibit  $A_4$  after one generation. We may now left multiply this result by  $T$  once again to get the next generation's distribution,  $T\mathbf{b}$ . However, notice that since  $T\mathbf{a} = \mathbf{b}$ , we have

$$\begin{aligned} T\mathbf{b} &= T(T\mathbf{a}) \\ &= T^2\mathbf{a}. \end{aligned}$$

In general, then, we can find the distribution of alleles in the  $n$ th generation by calculating  $T^n\mathbf{a}$ . Performing this operation, we find that, for practical purposes, we reach an equilibrium at about

the 9th generation, since

$$T^9 \mathbf{a} \approx T^{10} \mathbf{a} \approx \begin{bmatrix} 33 \\ 3 \\ 2 \\ 462 \end{bmatrix}.$$

Comparing this result to what we calculated for  $\mathbf{a}^*$  above, we see these two methods are identical. However, the method discussed here may not be as straightforward or efficient as the previous method, as we are required here to do one of the following:

- (1) begin with the initial distribution and multiply by our transition matrix until an equilibrium is reached,
- (2) guess an initial power for our transition matrix and continue the process described above until an equilibrium is reached, or
- (3) write a computer program that has a loop to continue applying powers to the transition matrix, stopping only when an equilibrium is reached.

For these reasons, it will be simpler to use the previous method in many cases when solving for the equilibrium solution to a Markov chain.

Of course, there are many more applications of Markov chains. For instance, we could collect data on populations in major cities, and use estimates of population growth to estimate a hypothetical equilibrium population distribution. It is important to note that the applications covered in this paper are highly simplistic, and assume that no change occurs in the conditions of the situation, allowing for us to assume that the transition matrices do not change. This may or may not be a realistic assumption. For instance, assuming that there is no change in the environment during 10 generations of an organism in our last example may be appropriate for organisms that reproduce often and have short lifespans. It would be unrealistic, however, to make similar assumptions about organisms that live much longer, reproduce less often and later in life, and if we considered many more generations than we saw in our example. This applies to our application in economics as well. Both natural environments and economic markets can be highly volatile. We must be wary about the reality of the situations we research, as the “real world” does not always behave as well as we would like.

Regardless, mathematical models such as these can be extremely useful to researchers in many disciplines. Although our estimations may not be entirely accurate, they provide valuable insight into what we should expect under certain circumstances. It is the responsibility of the researchers

to make the necessary adjustments to account for the various possible outcomes of the phenomena they study. If this is done correctly, the models such as that considered in this paper may impart a vast array of information that allows for a more thorough, cogent study.

## REFERENCES

- [1] Campbell, N. A., & Reece, J. B. (2008). *Biology* (8th ed.). San Francisco, CA: Pearson.
- [2] Cummings, M. R. (2014). *Human Heredity: Principles and Issues* (10th ed.). Belmont, CA: Brooks/Cole.
- [3] Durrett, Rick (2010). *Probability : Theory and Examples*. Retrieved from <http://www.ebilib.com>
- [4] Tietenberg, T., & Lewis, L. (2012). *Environmental & Natural Resource Economics* (9th ed.). Boston, MA: Pearson.