ANALYSIS OF COLLEGE GRADUATION RATES USING MARKOV CHAINS

A Thesis
by
THOMAS KEENER

Submitted to the School of Graduate Studies
at Appalachian State University
in partial fulfillment of the requirement for the degree of
MASTER OF ARTS

May 2022
Department of Mathematical Sciences

ANALYSIS OF COLLEGE GRADUATION RATES USING MARKOV CHAINS

A Thesis
by
THOMAS KEENER
May 2022

APPROVED BY:

Holly P. Hirst, Ph.D.
Chairperson, Thesis Committee

Alan T. Arnholt, Ph.D.
Member, Thesis Committee

Richard E. Klima, Ph.D.
Member, Thesis Committee

Eric S. Marland, Ph.D.
Chairperson, Department of Mathematical Sciences

Marie Hoepfl, Ed.D.
Interim Dean, Cratis D. Williams School of Graduate Studies

Abstract

Thomas Keener
B.S., Appalachian State University
M.A., Appalachian State University

Chairperson: Holly P. Hirst

In this paper, I analyzed graduation rates of a large, comprehensive university in the southeast United States using Markov chains. This analysis was performed by applying Markov chains to different populations at the university including entering freshmen, sophomore transfers, and junior transfers. This was done to determine differences between the groups' long term behavior. Entering freshmen and sophomore transfers' progression were also analyzed within each of the individual colleges at the university. Some prescriptions based on the analysis include: increasing the freshman retention rate and encouraging sophomores to change majors rather than leave the university. Additional post-analysis recommendations and discussion are included, as well as limitations of the analysis.

Table of Contents

Chapter 1: Introduction to Absorbing Markov Chains

Often in mathematics we assume that systems are random, and we call these stochastic processes. Though these systems may be random, we can sometimes determine the states in which an individual in the system, or the system itself, could possibly be in. If we not only know these states, but the probabilities of the system or individual moving between states, we can establish a system of linear equations to predict the occupancy of each state at a given time in the future.

Example 1

Let us examine an example of these processes. A car has broken down on the side of the highway and been abandoned. There is a 30% chance the car will be removed from the side of the road each day. Once the car is removed from the highway there is no chance that it is seen in the same spot again.

In this system, we can determine that there are two possible outcomes at the end of each day; the car is either removed or the car remains on the side of the road. Thus, the set of possible outcomes is $U = \{remains\,,\, removed\}$. To represent these outcomes as variables, we will let u_1 be the state of the car remaining, and u_2 be the state of the car being removed. Then on some given day, t, we can model the state of the car on the next day, $t+1$, using the following equations:

$$u_1(t+1) = 0.7\, u_1(t)$$

$$u_2(t+1) = 0.3\, u_1(t) + 1\, u_2(t).$$

Since this is a system of linear equations, we can represent the system as a matrix in which the current, t, state is represented by the rows and the $t+1$ state is represented by the columns. The first row will represent the probabilities of transitioning from state $u_1(t)$ to $u_1(t+1)$ or $u_2(t+1)$ in the first and second columns respectively. The second row will represent the probabilities of moving from state $u_2(t)$ to $u_1(t+1)$ or $u_2(t+1)$ in the first and second columns respectively.

Expressing the system in terms of vectors and matrices allows us to see this more clearly. We can express the states in a vector $\overrightarrow{u(t)} = [u_1(t), u_2(t)]$. Then the vector $\overrightarrow{u(t+1)}$ can be expressed as $\overrightarrow{u(t+1)} = \overrightarrow{u(t)} \cdot \mathbf{P}$, where $\mathbf{P}$ is the coefficient matrix for the system and can be seen below.

$$\mathbf{P} = \begin{bmatrix} 0.7 & 0.3 \\ 0 & 1 \end{bmatrix}.$$

There are two things to note about this coefficient matrix. First, the rows of the matrix add to one; this guarantees that no matter which state that a given object is in, it will remain within the system in the next time step. Second, the matrix is square. This means that time steps can be iterated repeatedly by raising the matrix above to some positive power m. Because this matrix shows how the state of the car can transition from day to day, this matrix is known as a transition matrix.

If we wanted to know the probability of the car getting removed after 100 days, we would simply multiply $\mathbf{P}$ times itself repeatedly 100 times, or alternatively raise the transition matrix to the $m = 100$ power. In doing so, we would receive the following matrix:

$$\mathbf{P}^{100} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

This indicates that after 100 days, the probability of the car being removed, given that the car is remaining, is 100%. The reason we can deduce this information is because the probability of a transition from u_1 to u_2 is $P_{12} = 1$. Therefore, given enough time, the car will be removed from the side of the road.

Now that we know the long-term behavior of this system, we can now expand upon our previous observations. First, the probability of the car remaining in u_2, given that it is already in u_2, is 100%; if the car is removed, it stays removed. When there is no probability of an object leaving a state it has already entered, then the state is said to be absorbing. Second, the car is not guaranteed to return to the side of the road. When there is no guarantee that an object or system can return to a state–in this case u_1–that state is said to be transient.

The method used to model the car example is an example of a Markov chain. Markov chains are a method of probabilistic modeling used to model Markov Processes, a type of finite stochastic process in which the future state of a system is entirely contingent on the preceding state. If the probabilities of transitioning from one state to another in an arbitrary time step are known, the probabilities can be organized into a matrix known as a Markov chain. Subsequently, these chains can be used to determine the outcome of systems after multiple time steps [4].

With the addition of absorbing states, and the ability of an object or system to reach an absorbing state from a non-absorbing state, Markov chains become absorbing Markov chains . Given enough time, everything within a system will be absorbed into an absorbing state, as can be seen within the car example [3]. As a consequence, Markov chains with a single absorbing state do not provide interesting results, but Markov chains with multiple absorbing states do. A Markov chain with multiple absorbing states can be seen in the next

example.

Example 2

A neurology department of a hospital classifies patients under its care as either bedridden or ambulatory. Data for the past 10 years reveals that over a one-day time period, 20% of all ambulatory patients are discharged from the hospital, 75% remain ambulatory, and the remainder require complete bedrest. In contrast, 10% of all bedridden patients become ambulatory, 80% remain bedridden, and the other 10% die [4].

In this example, there are four states: ambulatory, bedridden, discharged or dead. Allow row one to be ambulatory, the second row to be bedridden, the third row to be discharged, and the fourth row to be dead. Also allow the columns to have the same states, but for the next time step. Then this situation can be modeled with the following transition matrix:

$$\mathbf{P} = \begin{bmatrix} 0.75 & 0.05 & 0.20 & 0 \\ 0.10 & 0.80 & 0 & 0.10 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Note that the bottom two rows of the matrix, discharged and dead respectively, are absorbing states, but all other states are transient. If we wish to know the probabilities of transition for 100 days, we can raise $\mathbf{P}$ to the one hundredth power as follows:

$$\mathbf{P}^{100} = \begin{bmatrix} 0 & 0 & 0.89 & 0.11 \\ 0 & 0 & 0.44 & 0.56 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Similarly to the car example, we can see that the probability of absorption is still 100%. However, in this case we can predict the percentage of patients that will die or be discharged. If we assume 100 days is sufficient time to determine long-term behavior–which we can because all transient probabilities have gone to zero–we can see that in the long-term 89% of ambulatory and 44% of bedridden patients are likely to be discharged from the neurology department. We can also see that 11% of ambulatory patients and 56% of bedridden patients are expected to die.

With the addition of multiple absorbing states, we become interested in the probability of being absorbed into each absorbing state. In the two previous examples, this was done by raising the transition matrices to exceptionally high powers. This method is possible and fast with the help of computer algorithms. However, there are alternative ways to calculate the long-term behaviors and expected time steps until absorption to be discussed in later sections.

Chapter 2: Mathematical Framework of Absorbing Markov Chains

Markov Processes

As previously mentioned, Markov Processes are processes in which the future state of a system is determined only by the immediately preceding state, and a state is a unique outcome of a trial. If we define chronological time steps $t_0, t_1, \ldots, t_n$ and random variables $\{X_{t_n}\} = \{x_1, x_2, \ldots, x_n\}$, this property can be formalized by the following mathematical statement [4]:

$$P\{X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}, \ldots, X_{t_0} = x_0\} = P\{X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}\}. \tag{1}$$

We will denote the probability of transitioning from one state to another in one time step, the one-step transition probability, as follows [4]:

$$p_{ij} = P\{X_t = j \,|\, X_{t-1} = i\},\ (i,j) = 1, 2, \ldots, n,\ \ t = 0, 1, \ldots, T \tag{2}$$

$$p_{ij} \geq 0,\ (i,j) = 1, 2, \ldots, n, \tag{3}$$

where n is the total number of states.

Additionally, the one-step transition probabilities exhibit the following property for n possible states [3]:

$$\sum_{j=1}^{n} p_{ij} = 1 \tag{4}$$

For the remainder of this paper, all Markov processes were assumed to be finite ($n < \infty$); this assumes that we can know all possible outcomes of all possible trials in a system [3].

Organizing Markov Chains

We will hereby denote each of the states in a Markov Process as u_i, $i = 1, \ldots, n$; the probabilities from equation 2 indicate the probability of transitioning from state u_i to state u_j in one time step. By arranging these probabilities into a matrix as follows, we can form a Markov Chain for which P is the transition matrix [3].

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & & \ddots & \\ p_{n1} & & & p_{nn} \end{bmatrix} \tag{5}$$

By applying the property from equation 4, we can note that the sum of the entries for each row is 1. Let it also be noted that each transition matrix is a square matrix with dimensions n as seen in the previous examples.

Transition Digraphs and Relevant Graph Theory

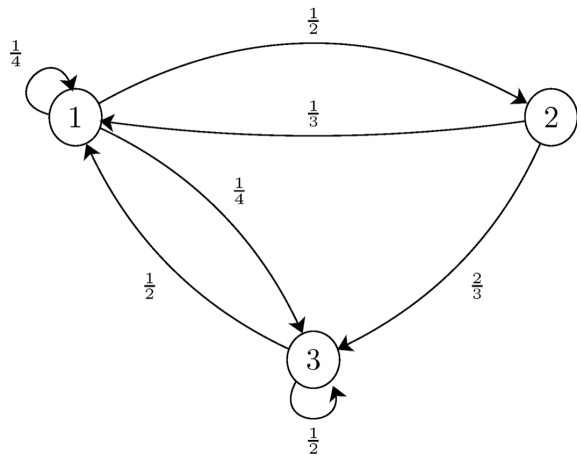
One method of building Markov Chains is to view all states as vertices, and all probabilities as weighted arcs connecting the vertices [3]. Not only is this a visualization method, but it will allow us to use graph theory to prove important theorems.

For example, we will construct a transition digraph for the following transition matrix.

$$P = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \tag{6}$$

From P, we can see that $p_{12} = \frac{1}{2}$. This indicates that a system in state one has a $\frac{1}{2}$ probability of transitioning to state two in a single time-step. Therefore, if we imagine states one and two as vertices, we could draw a directed arc from vertex one to vertex two with weight $\frac{1}{2}$. If we apply this process to each state and each p_{ij} within P, we receive the following transition digraph as seen in *Figure 1*.

Figure 1: **A Transition Digraph for Matrix P [2]**



Once a transition digraph is created, a vertex basis and vertex contrabasis can be determined. A vertex basis of a digraph, D, is the smallest set of points, B, such that any vertex in D can be reached from some vertex in B. If the direction of all arcs in D are reversed, any vertex basis in this reversed direction is known as a vertex contrabasis. Formally, a vertex contrabasis is a minimal set, B', such that any vertex in D can reach B' [3].

It must also be noted that we can divide a digraph into its strong components, which are vertices that can be reached by any other vertex. When we remove components that do not meet the criteria for a strong component, we will call this condensing [3].

The final required definition from the graph theory is the definition of an ergodic set. An ergodic set is a closed set in which no proper subset is closed. In a digraph, ergodic sets are strong components that form a vertex contrabasis in the condensation of D, D^* [3]. In the context of a Markov chains, ergodic Markov chains are those that contain states that objects or systems can return to, but in no certain time period. Sometimes ergodic chains

are described as recurrent, but aperiodic [4]. If an ergodic set contains only one element, that element is called an absorbing state. Once this vertex has been reached, there is no way to leave it [3].

Any vertex that cannot be returned to once it has been left is called a transient vertex, or a transient state in the language of Markov Chains.

Absorbing Markov Chains and Relevant Calculations

Absorbing Markov chains will be used for the upcoming analysis of graduation rates. These Markov chains contain at least one absorbing state, and a way to reach an absorbing state from every non-absorbing state. Once a system has entered an absorbing state, it cannot leave. In a transition matrix, these states are indicated by the presence of a 1 on the main diagonal, or more formally, $p_{jj} = 1$.

We can begin to address the properties of absorbing Markov Chains with the following theorem and corollary.

Theorem 1. *In any finite Markov Chain, independent of starting state, the probability after t steps that the process is in an ergodic state approaches 1 as t approaches ∞*

Corollary. *In an absorbing Markov chain, the probability of absorption is 1.*

Proof. Let D be an underlying digraph of the transition digraph T. Because ergodic sets form a vertex contrabasis for the condensation D^*, every transient vertex u_i can reach an ergodic vertex u_j. Since there are a finite number of vertices, there is a number r such that for each transient vertex u_i, there is some ergodic vertex u_j with $d(u_i, u_j) \leq r$, where $d(u_i, u_j)$ is the shortest path from u_i to u_j in D. Thus, independent of starting state, the probability of entering an ergodic vertex in at most r steps is a positive probability p.

Therefore, since the process is Markovian, the probability of not reaching an ergodic vertex in r time steps is $1-p$. This implies the probability of not reaching an ergodic vertex in kr time steps is $(1-p)^k$. Since $0 < p \leq 1$, the probability of not reaching an ergodic state

approaches 0 as $k \to \infty$ [3]. ▲

Before proceeding with additional theorems, we must declare a canonical form for the absorbing transition matrices. This is done by partitioning the transition matrix into recognizable parts, and in this paper, absorbing states will always come last. These parts are labeled $\mathbf{N}$, $\mathbf{A}$, $\mathbf{0}$, and $\mathbf{I}$. $\mathbf{N}$ denotes the matrix containing probabilities of transitioning from a non-absorbing state to another non-absorbing state, and the summation of these entries in each row is less than one. $\mathbf{A}$ is the matrix containing probabilities of transitioning from a non-absorbing state to an absorbing state. $\mathbf{0}$ is a zero matrix indicating a transition from absorbing state to non-absorbing states, which is impossible. Lastly, $\mathbf{I}$ is an identity matrix such that $p_{ii} = 1$ and all other $p_{ij} = 0$; this indicates that once an absorbing state has been entered, the system cannot transition to another state, not even another absorbing state. For the remainder of this paper, these matrices will be arranged as seen below

$$\mathbf{P} = \left[\begin{array}{c|c} \mathbf{N} & \mathbf{A} \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right]$$

If a matrix is organized according to the canonical form, the following theorem can be proven.

Theorem 2. *For a Markov chain in canonical form, the following statements hold.*

$$\mathbf{N}^t = 0 \tag{7}$$

$$(\mathbf{I} - \mathbf{N})^{-1} \textit{exists} \tag{8}$$

$$(\mathbf{I} - \mathbf{N})^{-1} = \sum_{s=0}^{\infty} \mathbf{N}^s \tag{9}$$

Proof. Equation (7) is a direct consequence of Theorem 1 as everything will eventually enter

the absorbing state. To prove Equations (8) and (9) we note the following:

$$(\mathbf{I}-\mathbf{N})(\mathbf{I}+\mathbf{N}+\ldots+\mathbf{N}^{t-1}) = (\mathbf{I}-\mathbf{N}) + (\mathbf{N}-\mathbf{N}^2) + \ldots + (\mathbf{N}^{t-1}-\mathbf{N}^{\mathbf{t}}).$$

Since this series is telescoping

$$(\mathbf{I}-\mathbf{N})(\mathbf{I}+\mathbf{N}+\ldots+\mathbf{N}^{t-1}) = \mathbf{I}-\mathbf{N}^t. \tag{10}$$

Then, as $t \to \infty$, $\mathbf{N}^t \to \mathbf{0}$ since all entries are positive numbers less than one or are otherwise zero. Thus, for sufficiently large t, $(\mathbf{I}-\mathbf{N}^t) \neq \mathbf{0}$. Furthermore, if we apply the fact that det $(\mathbf{AB})$= det $(\mathbf{A}) \cdot$ det$(\mathbf{B})$ and det $(\mathbf{I}) = 1$ to equation (10) we can see that det $(\mathbf{I}-\mathbf{N}) \neq 0$. Thus proving equation (8).

After proving equation (8), we know $(\mathbf{I}-\mathbf{N})^{-1}$ exists. If we multiply each side of equation (10) with this inverse, we can see

$$\mathbf{I}+\mathbf{N}+\mathbf{N}^2+\ldots+\mathbf{N}^{t-1} = (\mathbf{I}-\mathbf{N})^{-1}(\mathbf{I}-\mathbf{N}^t). \tag{11}$$

By letting $t \to \infty$, the right-hand side of equation (11) tends toward $(\mathbf{I}-\mathbf{N})^{-1}$. Thus, equation (9) has been proven [3]. ▲

It must now be noted that the matrix $(\mathbf{I}-\mathbf{N})^{-1}$ is called the fundamental matrix and will hereafter be denoted with the letter $\mathbf{Q}$. Now that we have established how to calculate the fundamental matrix, we can calculate: expected time in a given state, expected time before absorption, and probability of absorption. First we will prove the calculation of expected time in a given state, and then expected time before absorption will be proven by corollary.

Theorem 3. *The expected number of time steps before absorption that an absorbing chain is in non-absorbing state u_j, given that the chain starts in non-absorbing state u_i, is given*

by the i, j entry of the fundamental matrix.

Corollary. *The expected number of time steps before absorption, given that the process starts in non-absorbing state u_i, is given by the sum on the entries in the ith row of the fundamental matrix.*

Proof. Let e_{ij} be the expectation in question. Let $c_j^{(s)}$ be 1 if the process is in state u_j at time s and 0 otherwise. For any x, let $E_i(x)$ represent the expected value of x given that the process starts in state u_i. Then, given that the ith row corresponds to u_i

$$e_{ij} = E_i\left(\sum_{s=0}^{\infty} c_j^{(s)}\right) = \sum_{s=0}^{\infty} E_i C_j^{(s)} = \sum_{s=0}^{\infty}\left[(1 - p_{ij}^{(s)} \cdot 0 + p_{ij}^{(s)} \cdot 1)\right] = \sum_{s=0}^{\infty} p_{ij}^{(s)}$$

Now $p_{ij}^{(s)}$ is the i, j entry of N^s. Thus, e_{ij} is the i, j entry of $\sum_{s=0}^{\infty} \mathbf{Q^s}$, which is in the fundamental matrix $\mathbf{Q}$ by Theorem 2 [3]. ▲

Finally, we will prove how to calculate the probability of absorption using the fundamental matrix.

Theorem 4. *In an absorbing Markov chain with canonical transition matrix, let b_{ij} represent the probability of absorption in absorbing state u_j given that the process starts in non-absorbing state u_i. If $\mathbf{B}$ is the $(n-m) \times m$ matrix, then*

$$\mathbf{B} = \mathbf{QA}$$

where A represents the probabilities of transitioning from a non-absorbing state to an absorbing state, and $\mathbf{Q}$ is the fundamental matrix.

Proof. Begin by obtaining recursion for b_{ij}. Starting in state u_i, the chain can either be absorbed in absorbing state u_j, absorbed in another absorbing state u_k, such that $k \neq j$, or go to a non-absorbing state u_r. The respective probabilities will be denoted by p_{ij}, p_{ik}, p_{ir}. From u_j, u_k and u_r, the probabilities of being absorbed from state u_j are 1, 0 and b_{rj} respectively.

Therefore,

$$b_{ij} = p_{ij} + \sum_{r=m+1}^{n} p_{ir} b_{rj}.$$

Since u_i is non-absorbing and u_j is absorbing, p_{ij} is the i, j entry of the matrix $\mathbf{A}$. Since u_i is non-absorbing and u_r is non-absorbing, p_{ir} is the i, r entry of the matrix $\mathbf{N}$. Then,

$$\mathbf{B} = \mathbf{A} + \mathbf{NB}.$$

By subtracting $\mathbf{NB}$ and multiplying $\mathbf{Q}$, we can see

$$\mathbf{B} = (\mathbf{I} - \mathbf{N})^{-1}\mathbf{A} = \mathbf{QA}$$

Alternatively, we can examine raising matrices to large powers. If we let $\mathbf{P}$ be a transition matrix in canonical form:

$$\mathbf{P}^2 = \left[\begin{array}{c|c} \mathbf{N} & \mathbf{A} \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right] \left[\begin{array}{c|c} \mathbf{N} & \mathbf{A} \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{N}^2 & \mathbf{A} + \mathbf{NA} \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right].$$

Then,

$$\mathbf{P}^3 = \left[\begin{array}{c|c} \mathbf{N}^2 & \mathbf{A} + \mathbf{NA} \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right] \left[\begin{array}{c|c} \mathbf{N} & \mathbf{A} \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{N}^3 & \mathbf{A} + \mathbf{NA} + \mathbf{N^2A} \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right].$$

By Continuing to raise $\mathbf{P}$ to higher powers, we can recognize the pattern and generalize to the following:

$$\mathbf{P}^k = \left[\begin{array}{c|c} \mathbf{N}^k & \left(\sum_{\mathbf{i=0}}^{\mathbf{k}} \mathbf{N^i}\right) \mathbf{A} \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right]$$

Using Theorem 2, this matrix simplifies to the following:

$$\mathbf{P}^k = \left[\begin{array}{c|c} \mathbf{0} & \mathbf{QA} \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right]$$

Thus, raising matrices to high powers yields the same results as multiplying $\mathbf{Q}$ and $\mathbf{A}$ for sufficiently large k [3]. ▲

These theorems and corollaries will allow us to simplify calculations relevant to absorbing Markov chains.

Revisiting Example 1

Example 1 had the following transition matrix.

$$\mathbf{P} = \begin{bmatrix} 0.7 & .3 \\ 0 & 1 \end{bmatrix}.$$

In this case, the matrix $\mathbf{N}$ is not a matrix but a scalar. Thus the calculation of the fundamental matrix is:

$$\mathbf{Q} = (1 - .7)^{-1} = \frac{1}{0.3} = 3.\bar{3}.$$

We can now conclude the expected time until absorption is between three and four days. In addition, we can calculate the absorption probability as seen below:

$$\mathbf{QA} = \frac{1}{0.3} \cdot 0.3 = 1.$$

This calculation has shown that the results that we received from raising the transition matrix to high powers were true. Both results showed that the car would be removed with probability 1. Therefore, we can concluded that the car will be removed from the side of the

highway within 3 to 4 days based on our model.

Revisiting Example 2

Example 2 had the following transition matrix,

$$\mathbf{P} = \begin{bmatrix} 0.75 & 0.05 & 0.20 & 0 \\ 0.10 & 0.80 & 0 & 0.10 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

In this case, the matrix $\mathbf{N}$ is a 2×2 matrix containing the probabilities of transitioning from a non-absorbing state to another non-absorbing state. The calculation of the fundamental matrix utilizes an identity matrix the same size as $\mathbf{N}$ and appears as follows

$$\mathbf{Q} = (\mathbf{I} - \mathbf{N})^{-1} = \begin{bmatrix} 4.4 & 1.1 \\ 2.2 & 5.6 \end{bmatrix}$$

We can now see that the expected time until absorption for an ambulatory patient is 5.5 days by summation of the first row, and the expected time until absorption for a bedridden patient is 7.8 days by summation of the second row. This is something that cannot be calculated from raising the transition matrix to high powers. We can now calculate the absorption probabilities given that we now have the fundamental matrix. This calculation can be seen below, where $\mathbf{A}$ is the matrix of probabilities for non-absorbing state to absorbing state transitions.

$$\mathbf{QA} = \begin{bmatrix} .89 & .11 \\ .44 & .56 \end{bmatrix}$$

In this example, we can see that $\mathbf{QA}$ has four absorption probabilities. Each of these

is the probability of absorption given that a patient starts in either non-absorbing state. For example, the 0.89 represents the probability of a patient being discharged given that the patient entered the neurology unit ambulatory. The 0.11 represents the probability of an ambulatory patient dying. The 0.44 is the probability of a bedridden patient being discharged, and the 0.56 is the probability of a bedridden patient dying. Note that these are the exact probabilities we received earlier when we simply raised P to the one hundredth power. Thus, with the aid of a computer either method of calculating the long-term absorption probabilities is sufficient.

Now that we have a method of providing probabilities of objects transitioning from any non-absorbing state to any absorbing state, we can build matrices with many absorbing and non-absorbing states in order to analyze more complex systems.

Chapter 3: Analyzing Graduation Rates of New Students

When a student enters into a university they are assigned an academic standing based on the number of college classes, and their respective credit hours, taken prior to entry. These standings, arranged in order of fewest classes taken to most classes taken, are freshmen, sophomore, junior, and senior. Typically, enrolling students are either freshmen, sophomore transfers, or junior transfers.

Previous analyses of long-term student behavior in universities have often emphasized the use of advanced techniques to create accurate Markov chains for entire institutions [1]. I wanted to answer the question, what are the differences in the long-term behaviors for students entering a university in the three aforementioned groups? To address this problem I requested data from a large, comprehensive university in the southeastern United States, and data were provided, in varied format, by the university for each of the three groups requested– freshmen, sophomore transfers and junior transfers. From the raw data, probabilities of students transitioning through various academic states were calculated and organized into absorbing Markov chains. The probabilities of being absorbed into each absorbing state were then calculated, as well as the expected time until absorption using Theorem 4 and Theorem 3 respectively. A sample of this data is included in Appendix I.

The institution from which I received data considers a student to be counted toward the graduation rate if they graduate within six academic years of their degree program. So, when a freshman, sophomore, or junior enters this institution, they have six years to complete their degree to be considered a graduate for the purposes of this calculation. If a student finishes

their degree program in seven years, while they do receive a degree, they are not counted as a graduate when the graduation rate is calculated. Therefore, in the following analysis, at the end of a student's sixth year they were either counted for or against the graduation rate of the institution.

Analysis of Freshmen Cohorts from 2006-2015

The data regarding the academic progression of new freshmen classified students as either freshmen, sophomores, juniors, or seniors based on the number of credit hours accumulated over time. Students that left the university during this time period were either considered graduated or not enrolled, and these states were considered to be absorbing with not enrolled being equivalent to not graduated. The Markov chain corresponding to this data set contained six states. The rows of the transition matrix represent the following states at time t, from top to bottom:

1. freshman status
2. sophomore status
3. junior status
4. senior status
5. not enrolled
6. graduated

The same order or states describes the representation of each column, from left to right, at time $t + 1$. If t represents one academic year, the following transition matrix shows the transition from t to $t + 1$:

$$\mathbf{P} = \begin{bmatrix} .18 & .69 & 0 & 0 & .13 & 0 \\ 0 & .21 & .69 & 0 & .10 & 0 \\ 0 & 0 & .22 & .72 & .06 & 0 \\ 0 & 0 & 0 & .35 & .03 & 0.62 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{12}$$

Given that this transition matrix was expressed in canonical form, I utilized Theorem 3 to calculate the absorption times for each non-absorbing states. This calculation yields the following matrix:

$$\begin{bmatrix} 4.3 \\ 3.6 \\ 2.7 \\ 1.5 \end{bmatrix} \tag{13}$$

Based on the calculations above, it is expected that a freshman entering this institution will either graduate or become not enrolled between the fourth or fifth year. I then utilized Theorem 4 to calculate the probabilities of being absorbed into each absorbing state given current, non-absorbing state. This calculation renders the following matrix:

$$\mathbf{QA} = \begin{bmatrix} .35 & .65 \\ .23 & .77 \\ .12 & .88 \\ .05 & .95 \end{bmatrix} \tag{14}$$

.

In both equations 13 and 14 the rows represent the non-absorbing states in the same order as the transition matrix. In equation 14 the columns represent non-enrollment and gradua-

tion from left to right. Based on this information, an entering freshman has a 65% chance of graduating with an expected time until absorption of between four and five academic years.

Analysis of Sophomore Transfers from 2008 to 2014

Data regarding progression of sophomore transfers were provided in a format different than the data regarding entering freshmen. The data of sophomore transfers tracked the number of students that remained enrolled at the end of each year. At the end of the sixth year, a student would either be graduated or not counted toward the graduation rate, because of this, a third absorbing state was added. For this analysis, the rows of the transition matrix represent, at time t and from top to bottom:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year,
4. enrolled at end of fourth year
5. enrolled at end of fifth year
6. not enrolled at the university
7. not counted toward graduation
8. graduated.

The final three states are assumed to be absorbing. The same order of states describes the column states from left to right, but the column states represent the states at time $t+1$. The difference between time t and $t+1$ is one academic year. Note that at the end of a student's fifth year, they can only transition to a non-graduated, or graduated state. This is because in another academic year, any other status except graduated status is not counted toward the university's graduation rate. Below is the transition matrix for sophomore transfers:

$$\mathbf{P} = \begin{bmatrix} 0 & .76 & 0 & 0 & 0 & .15 & 0 & .09 \\ 0 & 0 & .37 & 0 & 0 & .11 & 0 & .52 \\ 0 & 0 & 0 & .20 & 0 & .16 & 0 & .65 \\ 0 & 0 & 0 & 0 & .20 & .31 & 0 & .49 \\ 0 & 0 & 0 & 0 & 0 & 0 & .59 & .41 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{15}$$

Because this transition matrix was expressed in canonical form, I utilized Theorem 3 to find the expected time until absorption and found a new sophomore to have an expected absorption time of 2 to 3 years. I also calculated the probability of absorption using Theorem 4, and the result of this calculation can be seen below:

$$\mathbf{QA} = \begin{bmatrix} .30 & .00 & .70 \\ .19 & .01 & .80 \\ .22 & .02 & .76 \\ .31 & .12 & .57 \\ 0 & .59 & .41 \end{bmatrix}. \tag{16}$$

Thus, a new sophomore transfer is expected to have an absorption time of 2 to 3 years at the end of their first year, and a graduation probability of 70%.

Analysis of Junior Transfers from 2014 to 2018

The data regarding junior transfers has the same formatting and states as the sophomore transfer data. This transition matrix can be found in Appendix II.

After calculating the expected absorbing time and absorption probabilities, I determined that a junior transfer has an expected absorption time of 1 to 2 years at the end of their

first year, and an 80% chance of graduating. The matrix of absorption probabilities can be seen below:

$$\mathbf{QA} = \begin{bmatrix} .20 & .00 & .80 \\ .16 & .01 & .83 \\ .22 & .04 & .74 \\ 0 & .19 & .81 \\ 0 & .42 & .58 \end{bmatrix}. \tag{17}$$

Chapter 4: Analysis of Individual Colleges

The university that provided the freshman, sophomore transfer, and junior transfer data houses seven different colleges. These colleges are: the College of Arts and Sciences, the College of Business, the College of Education, the College of Fine Arts, the College of Health Sciences, the College of Music, and the seventh college houses mostly undeclared majors. The seventh college will be discussed later in this paper. For now, I will present the results from the other six colleges for which I asked the question, is there any noticeable difference in the long-term behaviors of the individual colleges? The analysis of the long-term behavior was performed using freshman and sophomore transfer data, but not junior transfer data; I will discuss the reason for this decision later in this paper.

Analysis of the College of Arts and Sciences

The data regarding the individual colleges was in a very similar format to the data of sophomore and junior transfers but with an added absorbing state. The states represented by the rows of the transition matrix at time t, from top to bottom are:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year
4. enrolled at end of fourth year
5. enrolled at the end of fifth year
6. no longer in the given college

7. not enrolled at the university
8. not counted toward graduation
9. graduated

The last four states are absorbing states. This same order describes the states of the columns from left to right, except the column states are at time $t+1$, where one academic year separates t and $t+1$. Because of the continuity in the structure of the data, the transition matrix for the freshman data in the College of Arts and Sciences will be the only provided transition matrix in this discussion, but the remaining transition matrices can be found in Appendix II.

Freshmen Entering the College of Arts and Sciences

Freshmen entering the College of Arts and Sciences progress according to the following transition matrix:

$$\mathbf{P} \begin{bmatrix} 0 & .72 & 0 & 0 & 0 & .14 & .14 & 0 & 0 \\ 0 & 0 & .70 & 0 & 0 & .16 & .11 & 0 & .03 \\ 0 & 0 & 0 & .28 & 0 & .04 & .07 & 0 & .61 \\ 0 & 0 & 0 & 0 & .18 & .02 & .12 & 0 & .68 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .62 & .38 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{18}$$

Given this transition matrix, a freshman entering the College of Arts and Sciences has an expected time to absorption of between 2 and 3 years from the end of the student's first year. Additionally, the absorption probabilities can be calculated since the transition matrix is in canonical form. The result of this calculation can be seen below:

$$\mathbf{QA} = \begin{bmatrix} .27 & .27 & .02 & .44 \\ .18 & .19 & .02 & .61 \\ .04 & .10 & .03 & .83 \\ .02 & .12 & .11 & .75 \\ 0 & 0 & .62 & .38 \end{bmatrix}. \tag{19}$$

Sophomores Entering the College of Arts and Sciences

The individual college data for sophomore transfers had the structure as the data individual college data for freshman. Given this, only the results of sophomore analyses will be provided in this discussion, but all excluded transition matrices can be found in Appendix II.

For sophomores entering the College of Arts and Sciences, the expected time until absorption is between one and two academic years from the end of the student's first academic year. The absorption probabilities can be seen in the following matrix:

$$\mathbf{QA} = \begin{bmatrix} .15 & .29 & .00 & .56 \\ .05 & .22 & .00 & .73 \\ .02 & .29 & .02 & .67 \\ .02 & .41 & .09 & .48 \\ 0 & 0 & .60 & .40 \end{bmatrix}. \tag{20}$$

Therefore, a sophomore entering the College of Arts and Sciences has a 56% chance of graduating from the same college.

Analysis of the College of Business

Freshmen Entering the College of Business

The expected time until absorption of a freshman entering the College of Business is between two and three years at the end of the student's first year. The probabilities of absorption

can be seen in the matrix below:

$$\mathbf{QA} = \begin{bmatrix} .24 & .25 & .01 & .50 \\ .17 & .15 & .02 & .66 \\ .04 & .06 & .02 & .88 \\ .03 & .07 & .08 & .82 \\ 0 & 0 & .51 & .49 \end{bmatrix}. \tag{21}$$

Sophomores Entering the College of Business

The expected time until absorption of a sophomore entering the College of Business is between one and two academic years after the end of the student's first academic year. The probabilities of absorption can be seen in the matrix below:

$$\mathbf{QA} = \begin{bmatrix} .17 & .24 & .00 & .59 \\ .05 & .13 & .01 & .81 \\ .01 & .11 & .02 & .86 \\ .03 & .18 & .11 & .70 \\ 0 & 0 & .71 & .29 \end{bmatrix}. \tag{22}$$

Analysis of the College of Education

Freshmen Entering the College of Education

The expected time until absorption of a freshman entering the College of Education is between two and three academic years from the end of the students first academic year. The probabilities of absorption can be seen in the matrix below:

$$\mathbf{QA} = \begin{bmatrix} .33 & .16 & 0 & .51 \\ .23 & .10 & .01 & .66 \\ .05 & .06 & .01 & .88 \\ .03 & .08 & .04 & .85 \\ 0 & 0 & .44 & .56 \end{bmatrix}. \tag{23}$$

Sophomores Entering the College of Education

The expected time until absorption of a sophomore entering the College of Education is two academic years after the end of the student's first academic year. The probabilities of absorption can be seen in the matrix below:

$$\mathbf{QA} = \begin{bmatrix} .21 & .26 & 0 & .53 \\ .07 & .16 & .00 & .77 \\ .06 & .14 & .01 & .79 \\ 0 & .18 & .03 & .79 \\ 0 & 0 & .25 & .75 \end{bmatrix}. \tag{24}$$

Analysis of the College of Fine Arts

Freshmen Entering the College of Fine Arts

The expected time until absorption of a freshman entering the College of Fine Arts is between two and three academic years from the end of the student's first academic year. The probabilities of absorption can be seen in the matrix below:

$$\mathbf{QA} = \begin{bmatrix} .25 & .25 & 0 & .50 \\ .18 & .16 & .01 & .65 \\ .04 & .02 & .02 & .86 \\ .01 & .12 & .05 & .82 \\ 0 & 0 & .31 & .69 \end{bmatrix}. \tag{25}$$

Sophomores Entering the College of Fine Arts

The expected time until absorption of a sophomore entering the College of Fine Arts is two academic years after the student's first academic year. The probabilities of absorption can be seen in the matrix below:

$$\mathbf{QA} = \begin{bmatrix} .12 & .27 & .01 & .60 \\ .05 & .17 & .01 & .77 \\ .01 & .25 & .03 & .71 \\ .02 & .31 & .13 & .54 \\ 0 & 0 & .61 & .39 \end{bmatrix}. \tag{26}$$

Analysis of the College of Health Sciences

Freshmen Entering the College of Health Sciences

The expected time until absorption of a freshman entering the College of Health Sciences is between two and three academic years after the student's first academic year. The probabilities of absorption can be seen in the matrix below:

$$\mathbf{QA} = \begin{bmatrix} .29 & .23 & .01 & .47 \\ .20 & .14 & .01 & .65 \\ .04 & .06 & .02 & .88 \\ .01 & .08 & .06 & .85 \\ 0 & 0 & .49 & .51 \end{bmatrix} \tag{27}$$

Sophomores Entering the College of Health Sciences

The expected time until absorption of a sophomore entering the College of Health Sciences is between one and two academic years after the end of the student's first academic year. The probabilities of absorption can be seen in the matrix below:

$$\mathbf{QA} = \begin{bmatrix} .14 & .23 & .00 & .63 \\ .06 & .14 & .00 & .80 \\ .01 & .18 & .01 & .80 \\ .02 & .30 & .11 & .57 \\ 0 & 0 & .47 & .53 \end{bmatrix}. \tag{28}$$

Analysis of the College of Music

Freshmen Entering the College of Music

The expected time until absorption of a freshman entering the College of Music is between two and three academic years from the end of the student's first academic year. The probabilities of absorption can be seen in the matrix below:

$$\mathbf{QA} = \begin{bmatrix} .34 & .24 & .03 & .39 \\ .26 & .16 & .04 & .54 \\ .08 & .10 & .06 & .76 \\ .03 & .09 & .13 & .75 \\ 0 & 0 & .56 & .44 \end{bmatrix}. \tag{29}$$

Sophomores Entering the College of Music

The expected time until absorption of a sophomore entering the College of Music is two and three years at the end of the student's first academic year. The probabilities of absorption can be seen below:

$$\mathbf{QA} = \begin{bmatrix} .32 & .32 & 0 & 0 \\ .11 & .28 & 0 & .60 \\ .08 & .24 & 0 & .68 \\ .11 & .44 & 0 & .45 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{30}$$

Reasoning for Data Exclusion and Limitations

There are two data sets excluded from the analysis of the individual colleges housed within this university. First, the aforementioned seventh college has been excluded. Second, analysis of juniors throughout individual colleges have been excluded.

The junior transfer data set contained, by a large margin, the smallest population of any data set. The next smallest data set was the sophomore transfer data set. In the analysis of sophomores in the individual colleges, the smallest colleges had very few students in the the fourth, fifth, and sixth academic years. In some cases, results indicated that sophomore students in these years either had a 100% probability to graduate, or a 0% probability to graduate in the next year. This is evidenced in the probabilities of absorption of sophomores

entering the college of music. The absorption probabilities indicate that a student at the end of their fifth academic year has a 0% chance of graduation. This expectation is unrealistic. Because the population in the junior data set was much smaller than that of the sophomore data set, the analysis likely would have yielded more unrealistic results. Knowing these limitations, I did not perform an analysis of individual colleges using the junior transfer data.

The previously mentioned seventh college–I do not identify its official name since it could be an identifier of the university–mostly contains students of undeclared major. However, during some of the years that the data encapsulate there were some majors within this college, but not in others. This implied that in some years students could graduate from the seventh college, and in other years they could not. Because of this, the results of the individual college analysis results were impossible to interpret for the seventh college. Therefore, the results of this analysis were excluded from this paper entirely.

Chapter 5: Conclusions

From the analysis, I can conclude that there are differences in the long-term behaviors of entering freshmen, sophomore transfers, and junior transfers. Contrasting equation 14 to equation 16 and equation 17, we can see that entering freshmen have a significantly lower probability of graduating than their sophomore and junior counterparts. We can also see that junior transfers have the highest graduation probability of all the groups by a large margin. Equation 14 also indicates that freshmen also have the highest probability of leaving the university when contrasted with sophomore and junior transfers. These results can also be seen in Table 1 and Table 2.

Table 1: **Overview of Probability Results**

Entering Status	Graduation Probability	Not Graduating Probability
Freshman	0.65	0.35
Sophomore	0.70	0.30
Junior	0.80	0.20

Table 2: **Overview of Expected Time Until Absorption**

Entering Status	Expected Time Until Absorption (in years)
Freshman	4-5
Sophomore	2-3
Junior	1-2

Furthermore, when contrasting equation 16 and equation 17 we can see that juniors are less likely than sophomores to leave the university. But, the probability of a junior transfer not being counted toward the graduation rate rises at a faster rate than that of a sophomore transfer with each additional academic year. However, the maximum probability of a

sophomore not being counted toward graduation is lower than that of a junior transfer. Additionally, both junior and sophomore transfers gain probability of graduating with each year until the end of the second year, in which graduation rates become somewhat erratic.

Within the analysis of individual colleges, we can observe that the probabilities of a freshman changing colleges is higher than the probability of a sophomore changing college in every case. But, in contrast, sophomore transfers are more likely to leave the university than to change colleges, in general.

At this university, students are expected to receive a degree after their fourth year of college courses, but a fifth year of college is not uncommon. In general, someone in their first year is a freshman, someone in their second year is a sophomore, and someone in their third year is a junior. Given this information, we can see that all of the groups in the overall analysis are expected to leave after 4-5 years of college courses. These figures align with the timeline of expected graduation, meaning that most students are not leaving the university prior to their expected leaving time. The individual college analysis renders similar results, but sometimes these estimates are a bit below the 4-5 year mark. Based on these changes, the most likely reason for this is the fact that students can leave a given college, but graduate from another one.

This leads us to a second problem with the individual college analysis. Since the data keep track of a student's standing at the end of an academic year, the transition matrix shows probabilities of transitioning to a singular different state over the next academic year. If during the next year a student transfers colleges and also graduates from the new college, the student is considered to have graduated from the previous college from which they transferred. This is a direct result of the data measuring the results by the end of the years, because the next time a student in this situation is counted their status will be graduated without the transfer ever being recognized. Thus, the transition matrices cannot encapsulate both transitions in a single time step, and the provided graduation percentages do contain some error.

If I were to prescribe changes to increase the university's graduation rate, I would recommend increasing the freshman retention rate and encouraging sophomore transfers to change majors rather than leave the university. Junior transfers seem to have great success overall with high graduation rates within the expected time to receive a degree.

References

[1] Shahab Boumi and Adan Vela, *Improving Graduation Rate Estimates Using Regularly Updating Multi-Level Absorbing Markov Chains*, Education Sciences **10** (2020), 377-395.

[2] Hossein Pishro-Nik, *11.2.7 Solved Problems* (2020), available at `https://www.probabilitycourse.com/chapter11/11_2_7_solved_probs.php`.

[3] Fred Roberts, *Discrete Mathematical Models*, Practice-Hall, inc., 1976.

[4] Hamdy Taha, *Operations Research: An Introduction*, Pearson, 2007.

Appendix I: Original Junior Transfer Data Set

Seen below is a screenshot of the junior transfer data as received from the university. This particular data set was chosen as an example because it is the most concise and contains no unique identifiers to the university.

Figure 2: **Original Junior Transfer Data**

New Transfers entering as Juniors, fall 2014 through fall 2018

E = enrolled following year
G = graduated prior to or in the following year
X = no lonfer enrolled

			Year 2		Year 3		Year 4		Year 5		Year 6	
CLASS	Student_Type_New	N	Yr2	N	Yr3	N	Yr4	N	Yr5	N	Yr6	N
3 - Junior	New Transfer	2737	E	1525	E	397	E	78	E	40	E	17
			G	890	G	1866	G	2096	G	2146	G	2169
			X	322	X	474	X	563	X	551	X	551

Appendix II: Transition Matrices Excluded From Discussion

The transition matrices that were not presented in the discussion can be found in this appendix.

Transition Matrix for Sophomores Entering the College of Arts and Sciences

The states for this transition matrix are as follows:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year
4. enrolled at end of fourth year
5. enrolled at the end of fifth year
6. no longer in the given college
7. not enrolled at the university
8. not counted toward graduation
9. graduated

This order represents the row states from top to bottom, and the column states from left to right. The rows represent the state at time t and the columns represent the states at time

$t+1$. The difference ins t and $t+1$ is one academic year.

$$\mathbf{P} = \begin{bmatrix} 0 & .63 & 0 & 0 & 0 & .12 & .15 & 0 & .10 \\ 0 & 0 & .31 & 0 & 0 & .04 & .13 & 0 & .52 \\ 0 & 0 & 0 & .22 & 0 & .02 & .20 & 0 & .56 \\ 0 & 0 & 0 & 0 & .15 & .02 & .41 & 0 & .42 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .60 & .40 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Transition Matrix for Freshmen Entering the College of Business

The states for this transition matrix are as follows:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year
4. enrolled at end of fourth year
5. enrolled at the end of fifth year
6. no longer in the given college
7. not enrolled at the university
8. not counted toward graduation
9. graduated

This order represents the row states from top to bottom, and the column states from left to right. The rows represent the state at time t and the columns represent the states at time $t+1$. The difference ins t and $t+1$ is one academic year.

$$\mathbf{P} = \begin{bmatrix} 0 & .76 & 0 & 0 & 0 & .11 & .13 & 0 & 0 \\ 0 & 0 & .74 & 0 & 0 & .14 & .11 & 0 & .01 \\ 0 & 0 & 0 & .28 & 0 & .03 & .04 & 0 & .65 \\ 0 & 0 & 0 & 0 & .15 & .03 & .07 & 0 & .75 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .51 & .49 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Transition Matrix for Sophomores Entering the College of Business

The states for this transition matrix are as follows:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year
4. enrolled at end of fourth year
5. enrolled at the end of fifth year
6. no longer in the given college
7. not enrolled at the university
8. not counted toward graduation
9. graduated

This order represents the row states from top to bottom, and the column states from left to right. The rows represent the state at time t and the columns represent the states at time $t+1$. The difference ins t and $t+1$ is one academic year.

$$\mathbf{P} = \begin{bmatrix} 0 & .67 & 0 & 0 & 0 & .13 & .15 & 0 & .05 \\ 0 & 0 & .34 & 0 & 0 & .05 & .09 & 0 & .52 \\ 0 & 0 & 0 & .13 & 0 & .01 & .09 & 0 & .71 \\ 0 & 0 & 0 & 0 & .15 & .03 & .18 & 0 & .65 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .71 & .29 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Transition Matrix for Freshmen Entering the College of Education

The states for this transition matrix are as follows:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year
4. enrolled at end of fourth year
5. enrolled at the end of fifth year
6. no longer in the given college
7. not enrolled at the university
8. not counted toward graduation
9. graduated

This order represents the row states from top to bottom, and the column states from left to right. The rows represent the state at time t and the columns represent the states at time $t+1$. The difference ins t and $t+1$ is one academic year.

$$\mathbf{P} = \begin{bmatrix} 0 & .77 & 0 & 0 & 0 & .15 & .08 & 0 & 0 \\ 0 & 0 & .73 & 0 & 0 & .19 & .06 & 0 & .02 \\ 0 & 0 & 0 & .25 & 0 & .04 & .04 & 0 & .67 \\ 0 & 0 & 0 & 0 & .10 & .02 & .08 & 0 & .80 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .44 & .56 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Transition Matrix for Sophomores Entering the College of Education

The states for this transition matrix are as follows:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year
4. enrolled at end of fourth year
5. enrolled at the end of fifth year
6. no longer in the given college
7. not enrolled at the university
8. not counted toward graduation
9. graduated

This order represents the row states from top to bottom, and the column states from left to right. The rows represent the state at time t and the columns represent the states at time $t+1$. The difference ins t and $t+1$ is one academic year.

$$\mathbf{P} = \begin{bmatrix} 0 & .68 & 0 & 0 & 0 & .16 & .15 & 0 & .01 \\ 0 & 0 & .40 & 0 & 0 & .05 & .10 & 0 & .45 \\ 0 & 0 & 0 & .13 & 0 & .06 & .12 & 0 & .69 \\ 0 & 0 & 0 & 0 & .14 & 0 & .18 & 0 & .68 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .25 & .75 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Transition Matrix for Freshmen Entering the College of Fine Arts

The states for this transition matrix are as follows:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year
4. enrolled at end of fourth year
5. enrolled at the end of fifth year
6. no longer in the given college
7. not enrolled at the university
8. not counted toward graduation
9. graduated

This order represents the row states from top to bottom, and the column states from left to right. The rows represent the state at time t and the columns represent the states at time $t+1$. The difference ins t and $t+1$ is one academic year.

$$\mathbf{P} = \begin{bmatrix} 0 & .76 & 0 & 0 & 0 & .11 & .13 & 0 & 0 \\ 0 & 0 & .74 & 0 & 0 & .15 & .10 & 0 & .01 \\ 0 & 0 & 0 & .31 & 0 & .03 & .05 & 0 & .61 \\ 0 & 0 & 0 & 0 & .18 & .01 & .12 & 0 & .69 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .31 & .69 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Transition Matrix for Sophomores Entering the College of Fine Arts

The states for this transition matrix are as follows:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year
4. enrolled at end of fourth year
5. enrolled at the end of fifth year
6. no longer in the given college
7. not enrolled at the university
8. not counted toward graduation
9. graduated

This order represents the row states from top to bottom, and the column states from left to right. The rows represent the state at time t and the columns represent the states at time $t+1$. The difference ins t and $t+1$ is one academic year.

$$\mathbf{P} = \begin{bmatrix} 0 & .69 & 0 & 0 & 0 & .09 & .14 & 0 & .07 \\ 0 & 0 & .37 & 0 & 0 & .03 & .09 & 0 & .51 \\ 0 & 0 & 0 & .21 & 0 & .01 & .18 & 0 & .59 \\ 0 & 0 & 0 & 0 & .22 & .02 & .31 & 0 & .46 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .61 & .39 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Transition Matrix for Freshmen Entering the College of Health Sciences

The states for this transition matrix are as follows:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year
4. enrolled at end of fourth year
5. enrolled at the end of fifth year
6. no longer in the given college
7. not enrolled at the university
8. not counted toward graduation
9. graduated

This order represents the row states from top to bottom, and the column states from left to right. The rows represent the state at time t and the columns represent the states at time $t+1$. The difference ins t and $t+1$ is one academic year.

$$\mathbf{P} = \begin{bmatrix} 0 & .73 & 0 & 0 & 0 & .15 & .12 & 0 & 0 \\ 0 & 0 & .71 & 0 & 0 & .17 & .10 & 0 & .02 \\ 0 & 0 & 0 & .23 & 0 & .03 & .05 & 0 & .68 \\ 0 & 0 & 0 & 0 & .14 & .02 & .07 & 0 & .77 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .49 & .51 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Transition Matrix for Sophomores Entering the College of Health Sciences

The states for this transition matrix are as follows:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year
4. enrolled at end of fourth year
5. enrolled at the end of fifth year
6. no longer in the given college
7. not enrolled at the university
8. not counted toward graduation
9. graduated

This order represents the row states from top to bottom, and the column states from left to right. The rows represent the state at time t and the columns represent the states at time $t+1$. The difference ins t and $t+1$ is one academic year.

$$\mathbf{P} = \begin{bmatrix} 0 & .57 & 0 & 0 & 0 & .11 & .15 & 0 & .17 \\ 0 & 0 & .26 & 0 & 0 & .06 & .09 & 0 & .59 \\ 0 & 0 & 0 & .11 & 0 & 0 & .15 & 0 & .74 \\ 0 & 0 & 0 & 0 & .23 & .02 & .30 & 0 & .45 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .47 & .53 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Transition Matrix for Freshmen Entering the College of Music

The states for this transition matrix are as follows:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year
4. enrolled at end of fourth year
5. enrolled at the end of fifth year
6. no longer in the given college
7. not enrolled at the university
8. not counted toward graduation
9. graduated

This order represents the row states from top to bottom, and the column states from left to right. The rows represent the state at time t and the columns represent the states at time $t+1$. The difference ins t and $t+1$ is one academic year.

$$\mathbf{P} = \begin{bmatrix} 0 & .73 & 0 & 0 & 0 & .15 & .12 & 0 & 0 \\ 0 & 0 & .69 & 0 & 0 & .21 & .09 & 0 & .01 \\ 0 & 0 & 0 & .47 & 0 & .07 & .05 & 0 & .41 \\ 0 & 0 & 0 & 0 & .23 & .03 & .09 & 0 & .65 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .56 & .44 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Transition Matrix for Sophomores Entering the College of Music

The states for this transition matrix are as follows:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year
4. enrolled at end of fourth year
5. enrolled at the end of fifth year
6. no longer in the given college
7. not enrolled at the university
8. not counted toward graduation
9. graduated

This order represents the row states from top to bottom, and the column states from left to right. The rows represent the state at time t and the columns represent the states at time $t+1$. The difference ins t and $t+1$ is one academic year.

$$\mathbf{P} = \begin{bmatrix} 0 & .58 & 0 & 0 & 0 & .26 & .16 & 0 & 0 \\ 0 & 0 & .52 & 0 & 0 & .07 & .16 & 0 & .25 \\ 0 & 0 & 0 & .36 & 0 & .04 & .08 & 0 & .52 \\ 0 & 0 & 0 & 0 & 0 & .11 & .44 & 0 & .45 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Transition Matrix for Junior Transfers

The states for this transition matrix are as follows:

1. enrolled at end of first year
2. enrolled at end of second year
3. enrolled at the end of third year,
4. enrolled at end of fourth year
5. enrolled at end of fifth year
6. not enrolled at the university
7. not counted toward graduation
8. graduated.

This order represents the row states from top to bottom, and the column states from left to right. The rows represent the state at time t and the columns represent the states at time $t+1$. The difference ins t and $t+1$ is one academic year.

$$\mathbf{P} = \begin{bmatrix} 0 & .56 & 0 & 0 & 0 & .11 & 0 & .33 \\ 0 & 0 & .26 & 0 & 0 & .10 & 0 & .64 \\ 0 & 0 & 0 & .20 & 0 & .22 & 0 & .58 \\ 0 & 0 & 0 & 0 & .44 & 0 & 0 & .56 \\ 0 & 0 & 0 & 0 & 0 & 0 & .43 & .58 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Vita

Thomas Keener graduated with his BS in Mathematics from Appalachian State University in 2020. He continued his education at Appalachian State University, receiving his MA in Mathematics in 2022. He plans to attend the University of South Carolina at Columbia for a PhD in statistics. His interests include disease modeling, operations research, applied statistics and statistical modeling.