School Policy Evaluated with Time-Reversible Markov Chains

ABSTRACT

In this work we propose a reversible Markov chain scheme to model for the mobility of students affected by a grade school leveling policy. This model provides unified and mathematically tractable framework in which transition functions are sampled uniformly from the set of **reversible** transition functions. The results from the study appear to confirm the disadvantageous effects of this school policy, on par with the of a previous model on the same policy.

Keywords— Markov chain, stochastic process, ergodicity, numerical methods

School Policy Evaluated with Time-Reversible Markov Chains

Trajan Murphy

2019

Honors Thesis

School Policy Evaluated with Time-Reversible Markov Chains

Trajan Murphy

trajan.murphy@gmail.com

Advisor: Iddo Ben-Ari PhD, Department of Mathematics.

Date ______________________________

Advisor signature ______________________________

University of Connecticut

2019

ACKNOWLEDGMENTS

I would like to extend special thanks to Deborah Brandon, director of the Student Undergraduate Applied Mathematics Institute at Carnegie Mellon University who gave me the opportunity to perform the research project in 2018 leading to this one. I would also like to give thanks to Dr. Aris Winger, Assistant Professor of Mathematics at Georgia Gwinnett College who advised me during that research project. I would also like to thank my advisor for this project, Dr. Iddo Ben-Ari.

Contents

Chapter 1

Introduction

1.1 Background

In 2017, Pittsburgh Arlington PreK-8 implemented a disciplinary policy in which students are sorted into three different categories: red, yellow, and green based on their behaviors and performance. Examples of behaviors that would earn students the title of green include following instructions and attending class more than 95% of the time. Examples of red behaviors include insubordination and attending class less than 90% of the time. Students in the yellow category are in between the red and green categories. The faculty at Pittsburgh Arlington PreK-8 evaluate the behavior and performance of each individual every two weeks in order to place each individual in their respective category [Mur18].

The policy further stipulates that each student will be required to wear an armband that publicizes the color category that they are placed in. This prompted research into the potential long-term outcome of such a policy, in that we wanted to know how many students would end up in the red category as time approaches infinity [Mur18].

1.2 Mathematical Models

Previous work on this policy informed the construction of a mathematical model for the Pittsburgh Arlington PreK-8 leveling system, considering it as a 3-state discrete-time homogeneous Markov chain, with the three states being the red, yellow, and green categories. This model by definition assumes that the students move through the categories probabilistically, and the probability of going from one category to another category depends only on the student's placement in the previous two-week period.

The ergodic theorem for Markov chains is applicable in this case. In particular, as the length of the time interval observed tends to infinity, the proportion of times a student (or any given group of students)

- spends in a category A, tends to $\pi(A)$;
- moves from category A to category B, relative to the number of times the student is in category A, tends to $p_{A,B}$.

These facts are of high empirical significance as they allow to estimate the unknown π and p from (a large number of) observations [Mur18].

Because the policy was implemented so recently, the data on the average proportion of individuals moving from category to category was not available. Furthermore, it was likely not being recorded at all, and even if it were, it would not be up for disclosure to the public.

In order to counteract this obstruction, we made simplifying assumptions to reduce the number of unknown transition probabilities. We also assumed that the transition probabilities were themselves random variables. We assumed that the transition matrix of the Markov chain was regular, and thus converged to a unique stationary distribution independent of the initial conditions. In fact, assuming that the transition probabilities are represented by continuous random variables, the probability of obtaining a non-regular transition matrix is equal to zero. By treating the transition probabilities as random variables, we were able to find numerically the probability that the stationary distribution would have more than 10% of students in the red category in the long term [Mur18].

Chapter 2

Our work

In this work we operate under a similar assumption that the movement of students through the different color categories represents a three-state, discrete-time Markov chain. We furthermore maintain the assumption transition probabilities are random variables. We introduce a different assumption that the Markov Chain is reversible (generalization of symmetric, see Section 2.2). This latter assumption leads to relatively simple and explicit formulas for the stationary distribution and other quantities, allowing for a more detailed, rigorous analysis of time-evolution.

Many processes in physics and biology, such as the path of a pollen particle in a breeze can be modeled as a reversible Markov chain, so it is a reasonable assumption that this Markov chain might be reversible too.

2.1 Reversibility

Let $p = (p_{i,j})$ be a transition function of a Markov chain on the state-space $S = \{1, \dots, N\}$. That is, $p_{i,j}$ is the probability of transition from i to j. If $\mathbf{X} = (X_n : n \in \mathbb{Z}_+)$ is a Markov chain with transition function p, then

$$P(X_{n+1} = j | X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = p_{i,j}.$$

Definition 2.1.1. *The transition function p is reversible, if there exists a nonnegative function $\pi : S \to \mathbb{R}$, not identically zero, such that*

$$\pi(i) p_{i,j} = \pi(j) p_{j,i},$$

for all $i, j \in S$.

In what follows we will always assume the function π, if exits, is normalized to be a probability measure, i.e. :

$$\sum_i \pi(i) = 1.$$

Proposition 2.1.2. *If p is reversible with the function π. Then π is a stationary distribution for p.*

Proof. We need to show

$$\pi(j) = \sum_i \pi(i) p_{i,j}.$$

By the reversibility assumption, each summand on the righthand side is equal to $\pi(j) p_{j,i}$, therefore the righthand side is equal to

$$\sum_i \pi(j) p_{i,j} = \pi(j),$$

because p is a transition function and so $\sum_i p_{j,i} = 1$. □

For example, any symmetric transition function (i.e. $p_{i,j} = p_{j,i}$) is reversible with π being uniform. The converse is not true. Indeed, it is an easy exercise to see that any transition function on two states which is strictly positive off the diagonal (i.e. $p_{1,2}, p_{2,1} > 0$)

is reversible, but clearly not every such transition function is symmetric. On more than two states, reversibility is, of course, harder to achieve. As an example, consider the transition matrix $\begin{bmatrix} 0.90 & 0.05 & 0.05 \\ 0.20 & 0.55 & 0.25 \\ 0.15 & 0.75 & 0.10 \end{bmatrix}$ has stationary distribution $\pi = \begin{bmatrix} 0.616 & 0.081 & 0.302 \end{bmatrix}$. However, we observe that $0.05(0.081) \neq 0.20(0.616)$.

The following explains why we refer to chains that obey Definition 2.1 as "reversible":

Theorem 2.1.3. *Let* $\mathbf{X}$ *be a Markov chain with transition function* p *and stationary distribution* π*. Then* p *is reversible with* π *if and only if the distribution of the vectors* $(X_0, \dots, X_n)$ *and* $(X_n, 0, \dots, X_0)$ *are the same when* X_0 *is* π*-distributed.*

Proof. [Dur11]. Fix n and let $Y_m = X_{n-m}$ for $0 \leq m \leq n$. Then Y_m is a Markov chain with transition probability

$$\hat{p}_{i,j} = P(Y_{m+1} = j | Y_m = i) = \frac{\pi(j) p_{j,i}}{\pi(i)}$$

To show this, we compute the conditional probability.

$$P(Y_{m+1} = i_{m+1} | Y_m = i_m, Y_{m-1} = i_{m-1} ... Y_0 = i_0)$$

$$= \frac{P(X_{n-(m+1)} = i_{m+1}, X_{n-m} = i_m, X_{n-m+1} = i_{m-1} ... X_n = i_0)}{P(X_{n-m} = i_m, X_{n-m+1} = i_{m-1} ... X_n = i_0)}.$$

Using the Markov property, the numerator is equal to

$$\pi(i_{m+1}) p_{i_{m+1}, i_m} P(X_{n-m+1} = i_{m-1}, ... X_n = i_0 | X_{n-m} = i_m).$$

Similarly the denominator can be written as

$$\pi(i_m) P(X_{n-m+1} = i_{m1}, ... X_n = i_0 | X_{n-m} = i_m).$$

Dividing the last two formulas and noticing that the conditional probabilities cancel we have

$$P(Y_{m+1} = i_{m+1}|Y_m = i_m, ...Y_0 = i_0) = \pi(i_{m+1})p_{i_{m+1},i_m}(i_m).$$

This shows Y_m is a Markov chain with the indicated transition probability. □

Corollary 2.1.4. *An irreducible three-state Markov Chain is reversible iff*

$$p_{1,2}p_{2,3}p_{3,1} = p_{3,2}p_{2,1}p_{1,3} > 0. \tag{2.1.1}$$

Proof. Going in the forward direction, we have

$$\pi(1)p_{1,2} = \pi(2)p_{2,1}$$
$$\pi(2)p_{2,3} = \pi(3)p_{3,2}$$
$$\pi(3)p_{3,1} = \pi(1)p_{1,3},$$

with $\pi(i) > 0$ for all i. Multiplying the elements in right column and all elements in left column, then dividing by $\pi(1)\pi(2)\pi(3)$ gives (2.1.1). To prove the reverse direction, let $\pi(3) = c$ for some c to be determined later, $\pi(2) = \pi(3)\frac{p_{3,2}}{p_{2,3}}$ and $\pi(1) = \pi(2)\frac{p_{2,1}}{p_{1,2}}$. Note that the definitions of $\pi(3), \pi(2)$ satisfy the second equality, and those of $\pi(2)$ and $\pi(1)$ satisfy the first equation. Finally,

$$\pi(1) = \underbrace{\pi(3) \times \frac{p_{3,2}}{p_{2,3}}}_{=\pi(2)} \times \frac{p_{2,1}}{p_{1,2}} = \pi(3) \times \underbrace{\frac{p_{3,2}p_{2,1}p_{1,3}}{p_{2,3}p_{1,2}p_{3,1}}}_{=1}\frac{p_{3,1}}{p_{1,3}},$$

satisfying the third equation. All that remains is to normalize, i.e. divide π by $\sum_{i=1}^{n} \pi(i)$ □

Corollary 2.1.5. *The stationary distribution* $\pi = \begin{bmatrix} \pi(1) & \pi(2) & \pi(3) \end{bmatrix}$ *of a 3-state discrete* *time reversible Markov chain is given by*

$$\pi(1) = \frac{1}{\frac{p_{1,2}}{p_{2,1}} + \frac{p_{1,3}}{p_{3,1}} + 1},$$
$$\pi(2) = \frac{p_{1,2}}{p_{2,1}}\pi(1),$$
$$\pi(3) = \frac{p_{3,1}}{p_{1,3}}\pi(1).$$

This follows directly from Definition 2.1.1

2.2 Symmetric and Reversible

Readers who are interested in the connection between symmetric matrices and reversible transition functions may find the following section noteworthy. Our main work will resume in Section 2.3.

We've stated that Markov chain with transition function p is called reversible if there exists a positive function π such that

$$\pi(i)p_{i,j} = \pi(j)p_{j,i} \text{ for all } i, j. \tag{2.2.1}$$

Condition (2.2.1) is known as the detailed balance condition. Every irreducible Markov chain on two states is reversible, but this is not the case for three-state Markov chains as demonstrated in Section

Clearly every symmetric matrix satisfies the ance condition with constant π. As for a partial converse, suppose that p satisfies the ance condition with π, and let $D = diag(\pi)$

$$\begin{aligned}(D^{1/2}pD^{-1/2})(i,j) &= \pi^{1/2}(i)p_{i,j}\pi^{-1/2}(j)\\ &= \pi^{-1/2}(i)p_{j,i}\pi^{1/2}(j)\\ &= (D^{-1/2}p^tD^{1/2})(i,j)\\ &= (D^{1/2}pD^{-1/2})^t(i,j).\end{aligned}$$

In other words, $D^{1/2}pD^{-1/2}$ is a nonnegative symmetric matrix, but not necessarily a transition function. In addition, $\pi^{1/2}$ is a positive eigenvector (both left and right due to symmetry) with eigenvalue 1. Reversing the process, starting with a positive symmetric matrix q with a strictly positive eigenvector ρ corresponding to eigenvalue 1, one obtains a transition function p satisfying the ance condition by letting $E = diag(\rho)$

$$p = E^{-1}qE.$$

Indeed, $p(i,j) = \rho^{-1}(i)q(i,j)\rho(j)$, so $\sum_j p(i,j) = 1$, and

$$\rho^2(i)p(i,j) = \rho(i)q(j,i)\rho(j) = \rho^2(j)p(j,i),$$

therefore p satisfies the ance condition with $\pi = \rho^2$.

Summarizing, we have shown the following correspondence between symmetric matrices with nonnegative entries and reversible transition functions. This is a theoretical tool that can be used for generating reversible transition functions from the larger class of symmetric matrices. Due to the specificity of our main problem we did not use it, and rather obtained a concrete construction described in Section 2.4.

Proposition 2.2.1. *1. Suppose p and π satisfy the detailed balance condition* (2.2.1). *Let $D = diag(\pi)$.*

2. *Suppose that q is a symmetric matrix with nonnegative entries and a strictly positive eigenvector ρ corresponding to the eignevalue 1. Let $E = diag(\rho)$. Then $p = E^{-1}qE$ is a reversible transition function and satisfies the ance condition with $\pi = \rho^2$.*

2.3 Our Model

Our goal is to sample a reversible transition function uniformly. This is obtained as follows. Let $\mathbb{P}$ be the uniform measure on transition functions. If we condition $\mathbb{P}$ on the event $R = \{\text{reversible}\}$, then the conditional measure $\mathbb{Q}$, given by $\mathbb{P}(\cdot|R)$ is the uniform measure on the set of reversible matrices.

We first explain how we construct the measure $\mathbb{P}$, the uniform measure on transition matrices. Next we restrict $\mathbb{P}$ to R, and then normalize.

The transition matrix for a 3-state reversible Markov chain has 5 degrees of freedom. The transition matrix for any 3-state discrete time Markov chain has 9 entries. Because each row must sum to 1, this fixes the value of exactly one entry in each row if the other two entries per row are chosen randomly. We can assume without loss of generality that these determined entries are the diagonal. This gives six total degrees of freedom.

Of the six remaining entries, they must be chosen to satisfy the detailed balance condition (2.1.1). Once five of those entries are chosen randomly, the sixth entry is determined. Without loss of generality, we can assume the determined entry to be $p_{3,2}$.

Thus a transition matrix for a 3-state discrete time Markov chain has five degrees of freedom. And we must assign five random variables to these degrees of freedom to successfully sample a random reversible transition matrix. Our algorithm for uniformly sampling a reversible transition matrix is described in Section 2.4

2.4 Algorithm for Sampling a Reversible Transition Matrix Uniformly

Here is the algorithm we constructed. A code for MATLAB^{TM} implementing the algorithm is given in Chapter 3.

1. Randomly chooses $p_{1,2}$ and, $p_{2,1}$ from a uniform distribution on (0,1).

2. Randomly chooses $p_{1,3}$ and $p_{2,3}$ from a uniform distribution on (0, $p_{1,2}$) and (0, $p_{2,1}$) respectively. This is such that $p_{1,2}$ and $p_{1,3}$ are uniformly distributed over all the possible values in a valid transition matrix, i.e. $\{(x, y) \in (0, 1)^2 : x + y < 1\}$.

3. Sets $p_{1,1}$ to 1 - $p_{1,2}$ - $p_{1,3}$ and $p_{2,2}$ to 1 - $p_{2,1}$ - $p_{2,3}$ so that the row sums in row 1 and 2 are equal to 1.

4. We recall that this transition matrix has five total degrees of freedom. We have already set four (i.e. $p_{1,2}, p_{1,3}, p_{2,1}, p_{2,3}$)

   It is important to note that $p_{3,1}$ cannot be chosen uniformly distributed on (0,1), as was $p_{1,2}$ and $p_{2,1}$

   We need that $0 < p_{3,1} + p_{3,2} < 1$.

   But by (2.1.1), Let $\alpha := \frac{p_{1,2}p_{2,3}}{p_{2,1}p_{1,3}} = \frac{p_{3,2}}{p_{3,1}}$ Thus $p_{3,1}\alpha = p_{3,2}$. Substituting this back in to the row sum restriction, we have

   $0 < p_{3,1} + \alpha p_{3,1} < 1$

   or $0 < p_{3,1} < \frac{1}{1+\alpha}$

   And we see that the necessary distribution of $p_{3,1}$ is uniform on (0, $\frac{p_{2,3}p_{1,2}}{p_{2,3}p_{1,2}+p_{1,3}p_{2,1}}$) and we set $p_{3,2}$, which is determined by the previous entries to be $\frac{p_{1,2}p_{2,3}p_{3,1}}{p_{2,1}p_{1,3}}$

5. Sets $p_{3,3}$ to 1 - $p_{3,1} - p_{3,2}$

6. Calculates the stationary distribution π according to Corollary 2.1.5.

7. The entries in the third row are distributed more strictly than the entries in the second and first row. To overcome this, we generate a random permutation π' of the entries in π

8. Repeats steps 1-8 for a given number of times. In our simulation, we generated 10^5 such matrices.

9. Plots various charts with the frequencies of $\pi(r)', \pi(y)'$, and $\pi(g)'$

We also did a similar simulation with unrestricted, randomly generated transition functions chosen uniformly from the set of 3×3 transition functions as a control. The results below are computed with 10^{15} such transition functions.

2.5 Data

FIGURE 2.5.1

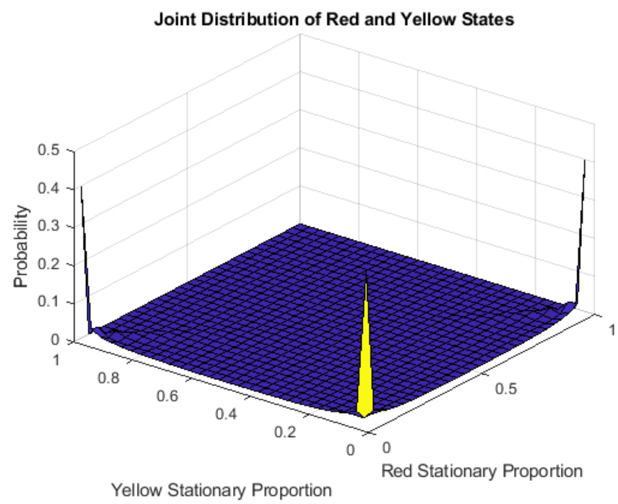
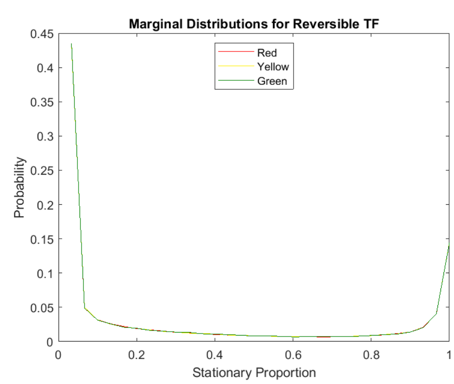


FIGURE 2.5.2



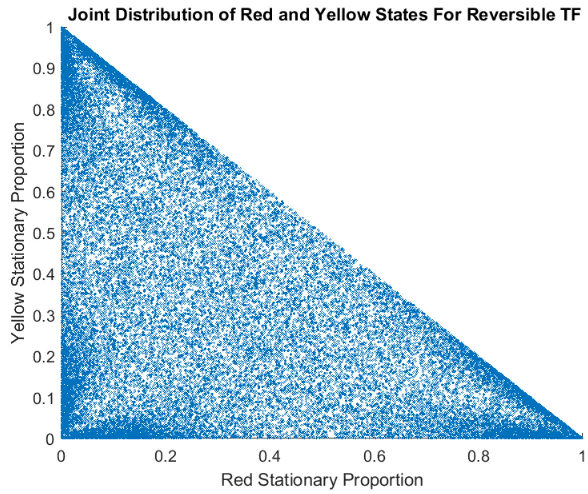


FIGURE 2.5.3

FIGURE 2.5.4

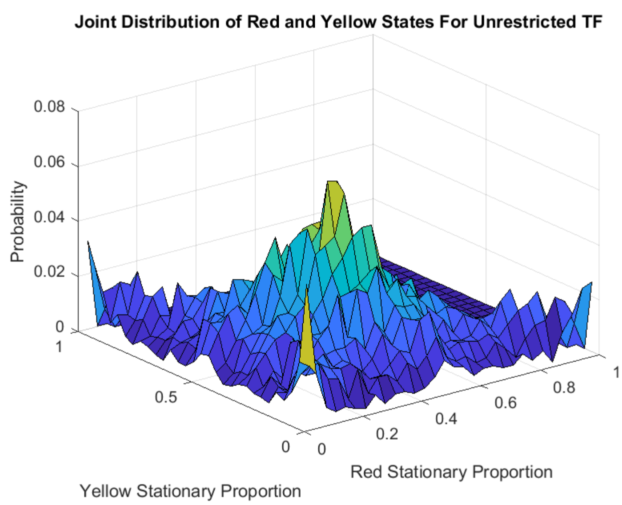
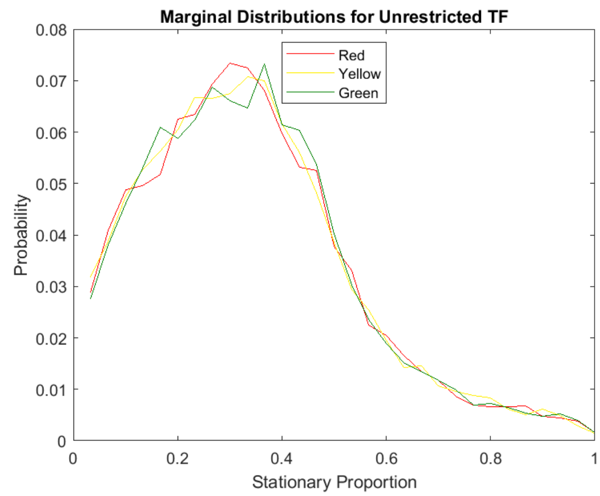


FIGURE 2.5.5



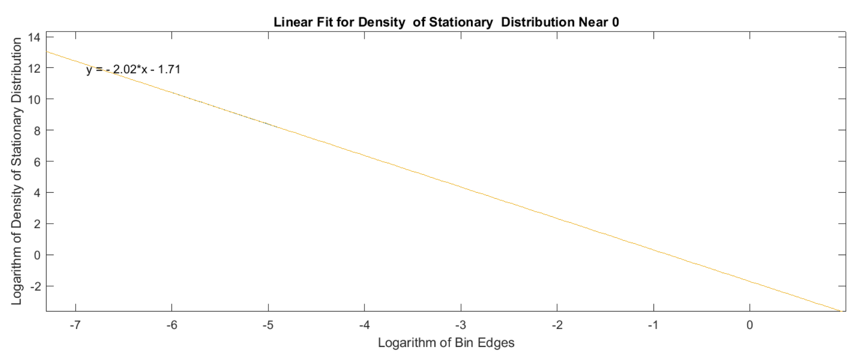


FIGURE 2.5.6

2.6 Observations

Figure 2.5.1 is a plot of the empirical joint distribution of $\pi(2)$ and $\pi(3)$. The distribution is not uniform, and more dense around near the corners (0,0), (0,1) and (1,0). Figure 2.5.2 is the empirical density for $\pi(1), \pi(2), \pi(3)$ all of which are identical. The distribution is not uniform and is most dense near the extreme values 0 and 1, with highest density near zero. Figure 2.5.3 is a scatter plot with the same information as Figure 2.5.2.

Figures 2.5.4 and 2.5.5 show the same data for the unrestricted transition functions. The convergence to the true PDF of each $\pi(i)$ was slower, but appears to be more concentrated near 0.3.

In the context of the leveling policy, the empirical probability that in the long term at least 10% of the students are in the red state is around 49%, much lower than previous results, due high density near zero. The probability of the same event for the unrestricted transition function was found to be approximately 89%.

2.7 PDF of Marginal Distribution of $\pi(i)$

As demonstrated by Figure 2.5.2, the PDF of each $\pi(i)$ appears to tend to infinity near 0. To verify this, we assume that between 10^{-6} and 10^{-2} the PDF has the form $f(x) = kx^{-\beta}$ for some $k, \beta \in (0, \infty)$. The CDF is therefore $F(x) = kx^{1-\beta}/(1-\beta)$, where x represents a certain value for a stationary distribution, and $F(x)$ represents the proportion of stationary distributions less than that value (for a certain color category i). Taking the logarithm of both sides yields $\ln(F(x)) = (k + \beta - 1) + (1 - \beta)\ln(x)$ which is a linear relationship. A least-squares regression line for $\ln(x)$ vs. $\ln(F(x))$ obtained from the algorithm in Section 2.4 is shown in Figure 2.5.6. By calculating the least-squares linear model, we see that the value of β is 3.02, which explains that the PDF of the marginal distribution of each $\pi(i)$ tends to ∞ as predicted.

2.8 Statistics

We shall assume that the true long term proportion of students in the red category is some value $\pi(r_0)$. Because of the large sample size (10^5), we are justified in assuming that the value $\pi(r)$ realized in the algorithm (in this case 0.49) above is sampled from a normal distribution with mean $\pi(r_0)$ and standard deviation $\sqrt{\pi(r_0)(1-\pi(r_0))10^{-5}}$. To estimate the true value $\pi(r_0)$, we construct a 95% confidence intervals centered around $\pi(r_0)$ with 0.49 as a left endpoint and solve for the appropriate value of $\pi(r_0)$ (which is calculated as solving $\pi(r_0) + \Phi^{-1}(0.025)\sqrt{\pi(r_0)(1-\pi(r_0))10^{-5}} = 0.49$ for $\pi(r_0)$. Here $\Phi(t)$ represents the cumulative distribution function of a standard normal random variable. We do the same with 0.49 as a right endpoint (which is calculated as solving $\pi(r_0) + \Phi^{-1}(0.975)\sqrt{\pi(r_0)(1-\pi(r_0))10^{-5}} = 0.49$ for $\pi(r_0)$. Doing this we observed that with 95% confidence, $0.4868 \leq \pi(r_0) \leq 0.4932$. This process was also done for the value 0.89 realized under the unrestricted transition function assumption. The true value of $\pi(r_0)$ in this case was found within machine precision to be exactly 0.89, due to the high sample size of 10^{15}.

2.9 Discussion

This model is again, one of many possible models to use to describe the movement of Pittsburgh Arlington PreK-8 students. The results obtained from this algorithm do demonstrate that there exists a non-negligible probability that a non-neglibible proportion of students in the long run will be placed in the red category, in both the reversible and unrestricted models.

Another model to investigate is a Polya's Urn Model, which is often described as a self-reinforcing stochastic process, in that it has the property that if a certain event is observed, the conditional probability that the same event is observed again is higher than if the event had not been observed. This model is useful if we would like to make the assumption that the longer an individual spends in a given colored category, the higher the probability that that individual stays in that category. It can be shown that the relative proportion of each color in the urn approaches a limiting distribution, with this distribution being randomly distributed. Thus an individual seeking to model the Pittsburgh Arlington leveling policy may wish to find the expectation of a certain distribution, and draw conclusions about the efficacy of the policy from that.

Polya's Urn can be rigorously defined by the following:

Let

- $N, k \in \mathbb{N}$
- $B_N = \{1, 2, ...N\}, C_k = \{1, 2, ...k\}$
- $f : B_N \to C_k$ be a surjection
- $c_i := \#\{x \in B_N : f(x) = i, i \in C_k\}$, i.e. the cardinality of the pre-image of i.

Polya's Urn scheme on N balls and k colors is a Markov chain $\{\mathbf{X}_n\}_{n\in\mathbb{N}} \in \mathbb{Z}_+^k - \{(0, 0, \ldots, 0)\}$, with the following transition probabilities:

$$P(X_{n+1}(i) = m|X_n) = \frac{X_n(i)}{N + n}\delta_{m,X_n(i)+1}.$$

Where $\delta_{m,X_n(i)+1} = 1$ if $X_n(i) + 1 = m$, or 0 if $X_n(i) + 1 \neq m$

Chapter 3

Code

This is the MATLAB^{TM} code implementing the algorithm in Section 2.4. THe code contains three files listed in the following sections. It can be downloaded from the link below.

`https://www.dropbox.com/sh/5jys7cpru0us396/AAA-pu0S-sT40oVR_tkvosTza?dl=0`

Samples reversible transition functions

File: `revmatrix.m`

```
function [percentage, pir] = revmatrix(maxit, nbins, takelog)

%initial counters
it=1;

%initial graph slots
plots=zeros(3,maxit);

while it<maxit

%build transition matrix T using random variables x1-x9
```

```
x2=rand(1,1);
x4=rand(1,1);
x3=rand(1,1)*(1-x2);
x6=rand(1,1)*(1-x4);
x7=rand(1,1)*x6*x2/(x6*x2+x3*x4);

x8=x3*x4/(x2*x6);

x9=1-x7-x8;
x5=1-x4-x6;
x1=1-x2-x3;


%build stationary distribution
p2= x3*x8/(x6*x7);
p3 =x6*x2/(x4*x8);

p1 = 1+ p2 + p3;
p2 = p2 / p1;
p3 = p3/p1;
p1 = 1/p1;

pi=[p1 p2 p3];
pi=pi(randperm(length(pi)));

%assembles pi's into a matrix

plots(:,it)=pi';




it=it+1;
end

%plots
```

```
%finds percent of students in the red category

counter = 0;
for i = 1:length(plots(1,:))
    if plots(1,i)>0.10
        counter = counter + 1;
    end
end
percentage = counter/length(plots(1,:));

pir = plots(1,:)



%separates the total stationary distributions into red, green, and yellow

xaxis = 1/nbins*(1:nbins);

datared = histogram(plots(1,:),nbins);
reds = datared.Values;

datayellow = histogram(plots(2,:),nbins);
yellows = datayellow.Values;

datagreen = histogram(plots(3,:),nbins);
greens = datagreen.Values;

redyellow = hist3(plots(1:2,:)',[nbins nbins]);

%You can take the logarithm of the data if it seems to spread out

if takelog == 1
    reds = log(reds+ones(nbins,1)');
    yellows = log(yellows+ones(nbins,1)');
    greens = log(greens+ones(nbins,1)');
    redyellow = log(redyellow+ones(nbins,nbins));
```

```
end

reds = reds/sum(reds);
yellows = yellows/sum(yellows);
greens = greens/sum(greens);
redyellow = redyellow/sum(max(redyellow));

%marginal distritbutions

figure;

plot(xaxis,reds,'color','red')
hold on
plot(xaxis,yellows, 'color',[1 0.9375 0])
hold on
plot(xaxis,greens, 'color', [0 0.5 0])

xlabel('Stationary Proportion')
ylabel('Probability')
legend({'Red','Yellow','Green'},'Location','North')
title('Marginal Distributions for Reversible TF')
hold off


%surface plot

figure;

surf(xaxis,xaxis,redyellow)

title('Logarithmic Joint Distribution of Red and Yellow States For Reversible TF')
xlabel('Red Stationary Proportion')
ylabel('Yellow Stationary Proportion')
zlabel('Probability')
hold off

%scatter plot
```

```
figure;
scatter(plots(1,:), plots(2,:),1)
title(' Joint Distribution of Red and Yellow States For Reversible TF')
xlabel('Red Stationary Proportion')
ylabel('Yellow Stationary Proportion')
hold off
end
```

Samples unrestricted transition functions

File: `unres.m`

```
function percentage=unres(MAXIT, nbins)

MAXIT=10000;
bank=zeros(3,MAXIT);

for t=1:MAXIT

r = rand(3,2);

x11 = r(1,1);
x12 = (1- x11)*r(1,2);
x13 = (1-x11-x12);

x21 = r(2,1);
x22 = (1- x21)*r(2,2);
x23 = (1-x21-x22);

x31 = r(3,1);
x32 = (1- x31)*r(3,2);
x33 = (1-x31-x32);

r1 = [x11 x12 x13];
r2 = [x21 x22 x23];
```

```
r3=  [x31  x32  x33];

r1=r1(randperm(numel(r1)));
r2=r2(randperm(numel(r2)));
r3=r3(randperm(numel(r3)));

p = transpose([r1;r2;r3]);

[V,D]=eig(p);
d=diag(D);
[~,idx]=max(d);



sdist1=real(V(:,idx));
sdist1=sdist1/sum(sdist1);

bank(:,t)=sdist1;

end



counter = 0;
for i = 1:length(bank(1,:))
    if bank(1,i)>0.10
        counter = counter + 1;
    end
end
percentage = counter/length(bank(1,:));

%separates the total stationary distributions into red, green, and yellow

xaxis = 1/nbins*(1:nbins);

datared = histogram(bank(1,:),nbins);
reds = datared.Values;

datayellow = histogram(bank(2,:),nbins);
```

```
yellows = datayellow.Values;

datagreen = histogram(bank(3,:),nbins);
greens = datagreen.Values;

redyellow = hist3(bank(1:2,:)',[nbins nbins]);

reds = reds/sum(reds);
yellows = yellows/sum(yellows);
greens = greens/sum(greens);
redyellow = redyellow/sum(max(redyellow));

%marginal distritbutions

figure;

plot(xaxis,reds,'color','red')
hold on
plot(xaxis,yellows, 'color',[1 0.9375 0])
hold on
plot(xaxis,greens, 'color', [0 0.5 0])

xlabel('Stationary Proportion')
ylabel('Probability')
legend({'Red','Yellow','Green'},'Location','North')
title('Marginal Distributions for Unrestricted TF')
hold off


%surface plot

figure;

surf(xaxis,xaxis,redyellow)

title('Joint Distribution of Red and Yellow States For Unrestricted TF')
xlabel('Red Stationary Proportion')
```

```
ylabel('Yellow Stationary Proportion')
zlabel('Probability')
hold off

end
```

Analyzes the marginal pdf of the reversible stationary distributions near zero

File: `cumulative.m`

```
function [x,y]=cumulative(pir)
dd=histogram(pir,200);
x=dd.BinEdges(1:200)+dd.BinWidth/2;
y=dd.Values;

m=find(x>10^(-5) & x<10^(-2));
A=log(x(m));
S=log(y(m));

%[a,s]=edf(pir);
%A=log(a);
%S=log(s);
%m=find(A>-6);

axis([-10 0 -60,0]);
plot(A,S);
% plot(x,y);

end
```

Bibliography

[Dur11] R. Durrett, *Essentials of stochastic processes*, New York City, 2011.

[Mur18] T Murphy, *School policy evaluated with markov chain.*