Markov's Monopoly

R. Dougherty-Bliss

January 17, 2017

Abstract

We give historical background on Markov and the creation of his chains as they relate to the weak law of large numbers. We then provide a brief introduction to Markov chain analysis in the finite case. Chain theory is used to create a model of the boardgame Monopoly and a limiting distribution is given, with code available using the library Pykov.

1 Introduction

Games of chance have long fascinated humans, often dictating large portions of our lives. From letting a coin toss decide where to eat, to sitting at a blackjack table, praying that the dealer busts, everyone has had experience with these moments of chance. For Americans, there may be no game of chance that has taken up more of our lives than the boardgame Monopoly.

Monopoly is a simple game conceived during the Great Depression. Players move around a square board by rolling dice, purchasing iconic American property and developing hotels on them as they go. The objective is ultimately bankrupt other players; that is, to be the last player with any money. The main way to affect the money of other players is to have them land on a property that you own. If this happens, they must pay rent. The more hotels are built on a property, the more expensive rent is. This creates a balancing act where players must spend money purchasing and developing property to force other players out, but must also maintain enough cash to ensure their own safety. What would be beneficial is if there was someway to know a prior what the "best" spaces to develop were.

As it turns out, there is indeed a way to determine what the "best" spaces are—if we quantify what we mean by "best." The "best" spaces in Monopoly are those that players are most likely to land on. Finding these spaces is a problem that can be solved mathematically. As we will see in Section 5, Monopoly can be modeled as a mathematical object known as a Markov chain. Markov chains are a very well-understood tool from probability developed by the Russian mathematician Markov in the early 20th century. As we will discuss in Section 4, Markov chains model a specific type of random behavior, and are applicable wherever this type of randomness occurs. They have been used to model systems in physics, psychology, economics, and various other applied fields [5]. They are interesting in their own right, but we will explore the power they possess as modeling tools, seeing first-hand why they have become mainstays of the modeling world for a hundred years.

First we will discuss the origins of Markov chains in more detail.

2 Markov in Context

Though Markov chains are some of the most well-known and understood types of stochastic processes, the background of their creation is often not discussed. In Section 3 we will discuss the technical goals of Markov chains, but here we will briefly describe Markov himself—the mathematician behind the discovery.

Markov's life in Russia straddled the nineteenth and twentieth centuries. The political climate of early twentieth-century Russia was tense, to say the least. There was a revolution brewing that would change the world's political landscape for decades to come. Despite this, there is little mention of politics in the works of Markov. Theorems are not directly concerned with who controls the government, so this is not surprising. However, the lack of political commentary in Markov's work should not be taken as disinterest. By the early twentieth-century Markov was an older, established academic. His opinions were freely given—usually unsolicited. Markov wrote so many letters of protest to newspapers that he was dubbed "Andrew the Furious" and "the militant academician" [4, p. 5]. In 1913 the Romanov dynasty celebrated 300 years of rule over Russia. Being displeased with them, Markov held a counter-celebration for 200 years of the Law of Large Numbers [3, chap. 10]. In 1902 the "socialist realist" Russian writer Maxim Gorky was elected to the Russian Academy of Sciences. His writings displayed a socialist viewpoint not in line with government position, so his election was met with significant resistance. The Minister of Education wrote that Gorky's election had a "most distressing effect" upon all "right-thinking Russians," and ordered that it be withdrawn [13, p. 37]. Markov protested this loudly and refused honors from the tzar the following year[1]. In 1912 the Russian Orthodox Church excommunicated Leo Tolstoy. In response, Markov submitted a request to the Church that he himself be excommunicated. The Church happily fulfilled this request: "[Markov] has seceded from God's Church and [we] expunged him from the lists of Orthodox believers" [4, p. 5].

In his time as a student, graduate student, and eventually professor at the St. Petersburg school of mathematics, Markov studied under or collaborated with many prominent Russian mathematicians. Markov's graduate advisor Chebyshev made contributions to probability and number theory, proving the eponymous Chebyshev's Inequality, which bounds the probability that a random variable can fall certain distances away from its mean, and the Bertrand-Chebyshev Theorem, which states that there exists a prime between n and $2n$ for all integers $n > 1$. Chebyshev's Inequality leads to a proof of the Weak Law of Large Numbers (stated in Section 3), and the Bertrand-Chebyshev Theorem is an important step in discovering the distribution of the primes, which will eventually lead mathematicians to the Riemann-Zeta function, and thus to the famed Riemann Hypothesis. Lyapunov, slightly younger than Markov, shared Chebyshev as an advisor. Lyapunov's thesis *The general problem of the stability of motion* introduced stability theory, opening the doors to a rigorous study of dynamical systems. Among many other things, Lyapunov applied his stability theory to show that movement of particles in "pear-shaped" bodies of liquid was unstable, refuting an early theory that the formation of orbital satellites, e.g. the Moon, was due to rotating bodies of liquid [8, pg. 276]. In short, Markov was surrounded by groundbreaking mathematicians. Interactions with his contemporaries will prove crucial to Markov's own groundbreaking work.

True to his character, the interactions that spurred Markov's work were not the pleasant collaborations—they were the drawn out, public arguments with mathematicians outside of St. Petersburg. A notable case of this is the tension between the schools of mathematics of St. Petersburg and Moscow. Many mathematicians in the Moscow school were members of the Orthodox Church, and used their research to argue for the validity of the Judeo-Christian concept of "free will" [10, p. 255]. Chuprov wrote that "[t]he 'Moscow School' decidedly insists that free will is the *conditio sine qua non* of statistical laws governing everyday life" [10, p. 257]. The St. Petersburg school, and Markov in particular, did not agree.

Markov and the Moscovian mathematician Nekrasov were the most prominent figures in the bickering between the St. Petersburg and Moscow schools. In many ways, there were similar. Both outspoken, both established mathematicians, and they both hated each other. Chuprov published a book that mentioned Nekrasov

[1] Gorky was eventually reinstated to the Academy due to the efforts of Markov and other prominent Russians.

in a positive light. Markov told Chuprov that Nekrasov's work was an "abuse of mathematics" [10, p. 257]. Nekrasov published a paper on probability that contained no proofs, and dedicated it to Chebyshev. Markov took great offense to his advisor even being mentioned in a work without proof, and one of many lengthy disputes between Nekrasov and Markov began. Out of the Moscow school, Nekrasov was the most adamant in applying probability and statistics to his Orthodox faith. Markov often worked to specifically refute these points. In fact, as we shall soon see, Markov's study of his chains was initiated almost entirely to provide a counterexample to a claim by Nekrasov about the "logical underpinnings" of the Weak Law of Large Numbers.

3 Origins of Markov Chains

Despite their wide use in modeling today, Markov chains were not invented as modeling techniques. They were invented primarily as a counterexample to claims Nekrasov made about independence in the Weak Law of Large Numbers. We will take the liberty to describe these claims in detail.

The Weak Law of Large Numbers (WLLN) is the formalization of a simple intuition: if we draw a large enough sample from a distribution, then the average value of the sample will approach the mean of the distribution. More accurately, the probability that the average value will be far away from the mean approaches zero as the sample size grows.

Theorem 1 (Weak Law of Large Numbers). *Let $X_1, X_2, \ldots$ be an infinite sequence of independent and identically distributed (iid) random variables with finite expectation $E[X]$. Define the random variable*

$$\bar{X}_n := \frac{1}{n}\sum_{k=1}^{n} X_n$$

to be the "mean" of the first n variables in the sequence. Then, $\bar{X}_n$ converges in probability to $E[X]$. That is,

$$\lim_{n\to\infty} P(|\bar{X}_n - E[X]| < \epsilon) = 1$$

for all $\epsilon > 0$.

As stated, this theorem applies to sequences of *independent* variables. Nekrasov noticed that the WLLN held if this condition was relaxed so that the variables were only *pairwise independent*. This was an interesting result, but Nekrasov had larger designs. He believed that this pairwise independence was implied by the philosophical concept of free will. He concluded that the fact that average measurements in the "real world" seemed to follow the WLLN implied that pairwise independence was a *necessary* condition for the WLLN to hold. That is, if $\bar{X}_n$ converges in probability to $E[X]$ for some sequence $X_1, X_2, \dots$, then the sequence must be pairwise independent. It is this claim that Markov set out to disprove with his chains [10].

As a counterexample to Nekrasov's claims, Markov invented a special class of discrete stochastic processes that satisfies the consequence of the WLLN with dependent random variables. Today, this construction would be called a finite Markov chain with positive transition probabilities. In addition to creating a tidy counterexample, in a single paper Markov creates a new subfield of probability.

4 Introduction to Markov Chains

Informally, the Markov chains that we will consider are collections of states equipped with probabilities of transitioning from one state to the other. To describe them formally, we must make use of stochastic processes.

A discrete time stochastic process is a collection of random variables $\{X_n\}$, $n = 0, 1, 2, \dots$, where each random variable X_n is discrete and takes values in a common, countable set, called the state space. We will generally assume that the state space is $\{1, 2, 3, \dots\}$, but the particular set is not important. We can think of the indexing variable n as "time" and the value X_n as the current state of the process at time n. The distribution of the random variable X_n completely describes the process at time n from a stochastic perspective. Thus, a discrete stochastic process is a collection of states and random variables that describe the distribution of the states of the process at any given time.

A simple example of a discrete stochastic process is the (infinite) random walk. Suppose that we begin at a state labeled 0 and flip a coin. If the coin is heads, then we move to state 1. If it is tails, then we move to state -1. We continue this in an indefinitely, where our position at time n is given by the random variable X_n.

In the random walk example, the state of the process at time n only depends on the state of the process at time $n-1$. To determine the possible states at time n, we only need to know what state we were in at time $n-1$, so we can flip our coin. This property is called the Markov property, and will characterize Markov chains.

Definition 1. A discrete time stochastic process $\{X_n\}$ has the *Markov property* iff

$$P(X_n = i_n \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) = P(X_n = i_n \mid X_{n-1} = i_{n-1})$$

for all sequences of states $\{i_k\}_{k=1}^{n}$. Any discrete time stochastic process that satisfies the Markov property is called a *Markov chain*.

Every Markov chain possesses a set of probabilities that describe the probability of transitioning to state i from state j at time n. These are called *transition probabilities* and are denoted

$$p_{ij}(n) = P(X_{n+1} = i \mid X_n = j).$$

We will consider the case where the transition probabilities are *stationary*, or do not depend on n.

Definition 2. If the transition probabilities are independent of n, that is, if $p_{ij}(n) = p_{ij}(n')$ for all times n and n', then they are called *stationary*, *time homogeneous*, or *homogeneous*. In this case, the chain is called time homogeneous as well. The stationary transition probabilities are denoted p_{ij}. That is, p_{ij} is the probability of moving to state i from state j at any time.

From here, by "Markov chain" we mean "time homogenous Markov chain," unless noted otherwise.

Given a Markov chain, we can form a matrix of its transition probabilities. This matrix is the defining aspect of a Markov chain.

Definition 3. The matrix

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots \\ p_{21} & p_{12} & p_{23} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

is the *transition matrix* of a Markov chain. This matrix is *stochastic* in that its columns sum to one. (It is certain that we will always transition to a state.) Conversely, any stochastic square matrix defines a Markov chain.

Every transition matrix can be thought of as an adjacency matrix for a directed graph. Thus, Markov chains can be represented as directed graphs with weighted edges. An example for the matrix

$$P = \begin{bmatrix} 0 & 0.1 & 0 & 0.4 \\ 0.6 & 0.9 & 0.3 & 0 \\ 0.4 & 0 & 0 & 0 \\ 0 & 0 & 0.7 & 0.6 \end{bmatrix}$$

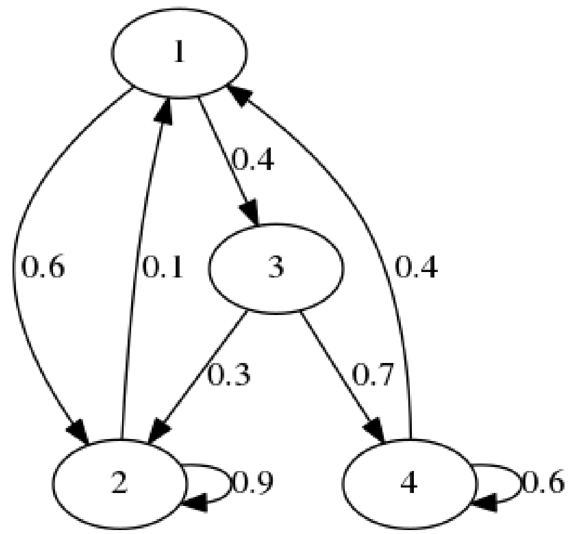

Figure 1: Graphic representation of a Markov chain. Each directed edge $j \to i$ is labeled with the transition probability p_{ij}.

is shown in Figure 1.

We have discussed the one-step transition probabilities p_{ij}. The next definition considers multiple steps. That is, we consider the probability that we will transition to state i from state j in n steps as opposed to one step.

Definition 4. The n-step transition probabilities are defined as

$$p_{ij}^{(n)} = P(X_n = i \mid X_0 = j).$$

Thus, $p_{ij}^{(n)}$ is the probability of moving to state i from state j in n steps. The matrix of n-step transition probabilities is denoted $P^{(n)}$.

There is a connection between the transition matrix and the n-step transition probabilities, namely that the matrix of n-step transition probabilities is the transition matrix P raised to the nth power.

Theorem 2. *The n-step transition matrix $P^{(n)}$ is the transition matrix raised to the nth power. That is,*

$$P^{(n)} = P^n.$$

Given an initial distribution of states $\mathbf{p}(0)$, the probability distribution at time n is simply

$$\mathbf{p}(n) = P^n \mathbf{p}(0).$$

By the long-term behavior of a Markov chain, we generally mean the behavior of $\mathbf{p}(n)$ as n tends to infinity. This behavior is governed by limiting distributions, which are the stochastic analog to stable equilibrium points in differential equations.

Definition 5. The probability distribution $\boldsymbol{\pi}$ is a *limiting distribution* of a Markov chain iff $P\boldsymbol{\pi} = \boldsymbol{\pi}$ and

$$\lim_{n\to\infty} p(n) = \lim_{n\to\infty} P^n\mathbf{p}(0) = \boldsymbol{\pi}$$

for all initial distributions $\mathbf{p}(0)$.

The existence of limiting distributions for arbitrary Markov chains is a nuanced discussion. A relatively simple theorem will suffice for our purposes.

Theorem 3. *Let P be a finite stochastic matrix such that $P^n > 0$ for some positive integer n (P is said to be regular). Then, the Markov chain defined by P has a unique limiting distribution. Further, P has a dominant eigenvalue of 1 associated with a positive eigenvector. This limiting distribution is the eigenvector normalized so that its components sum to one.*

For a more thorough discussion on Markov chains, see Allen [2] or Brémaud [5].

5 Monopoly as a Markov Chain

In the past few decades Markov chains have been used to model board games, with the general goal being to analyze long term behavior of the game through limiting distributions. In this context the limiting distribution describes how likely a game is to be in a certain state as the game goes on. Players armed with this knowledge have the advantage of knowing what states are more or less likely, and can act on this to increase their reward.

Osborne [7] models combat in the board game RISK as a Markov chain. In RISK, two players engage in combat by assigning a certain number of armies to battle over a territory. In each round, the probabilities of moving from one state to another (the number of armies on each side) depends only on how many armies each side has at one moment, i.e., the current state. Thus, battles can be modeled as Markov chains, and [7] presents the expected losses a player will suffer to their army given initial army counts.

Relevant to our current goal, Abbot and Richey [1] model Monopoly as a Markov chain. The goal of Monopoly is to be the only player to not go bankrupt, i.e., to lose all of their money. The main way to do this is to purchase and develop property, since players must pay rent for landing on property owned by another player. Knowing what spaces were most frequently visited would provide an advantage in purchasing decisions, thus a limiting distribution of Monopoly movement is of interest.

If we ignore the more complicated movement rules for a moment, Monopoly is a very simple Markov chain. There are forty spaces on a Monopoly board. At each space on the board, two dice are rolled, and a player advances forward by the sum of the roll. The probability of each sum is well-known, and any spaces not covered by the sum trivially have a transition probability of zero. For example, the (transpose of the) column corresponding to the first space would look like

$$\begin{bmatrix} 0 & 0 & \frac{1}{36} & \frac{2}{36} & \cdots & \frac{2}{36} & \frac{1}{36} & 0 & 0 & \cdots \end{bmatrix}.$$

It turns out that even after introducing the more complicated movement rules, the transition matrix of Monopoly is regular, so Monopoly has a unique limiting distribution. We will walk through the development of the chain, then present the limiting distribution.

5.1 Constructing the Monopoly Chain

For a full discussion of the rules of Monopoly, see [12].

We will model the following movement rules of Monopoly:

- Monopoly has 40 spaces, numbered 0–39. Space 0 is GO, and 39 is the space just before GO.
- Jail is at space 10. The rules of Jail are as follows:
  - When sent to Jail, a player may escape by rolling doubles (probability $6/36 = 1/6$). They attempt this for three turns, after which they are released no matter what. That is, they attempt to roll doubles twice, then roll normally from jail on their third turn. After a successful escape, a player is placed on Jail at space 10.
  - The three turns will be modeled as spaces 40, 41, and 42. The probability of escape for space 40 and 41 is $1/6$, and the probability of advancing to the next jail state is $5/6$. From space 42, the transition probabilities are as if the player was on Jail, but not *in* Jail.

  - Space 30 is 'Goto Jail,' and ending a turn there immediately moves a player to Jail at space 10. That is, when there is a probability of landing on space 30, this will be treated as the probability of landing on space 40.

- In a game of Monopoly, there are chance and community chest cards which have various effects on the game. These cards are drawn when landing on a Chance space, located at spaces 7, 22, and 36. The edition of Monopoly consulted by the author contained 32 of these cards, twelve of which had special movement rules. To simplify matters, we will assume landing on a Chance space will, with probability 20/32, do nothing, and with probability 12/32, move a player to a random space.

  This random space includes Goto Jail at 30 and excludes Jail first turn at 40. This decision is made for technical reasons. If there was no probability of reaching Goto Jail from anywhere, then the transition matrix would not be regular, making our analysis less concrete. The effects of this slight change are minimal from a probabilistic perspective.

5.2 Implementation

Fortunately, we do not have to compute the eigensystem of a 43×43 matrix by hand. Using numeric libraries, we *could* compute the eigensystem of our transition matrix and use this to find the limiting distribution, but we do not even need to do this. There exist efficient algorithms and implementations of them that exploit graph theory to compute limiting distributions of finite Markov chains. For our project, we will make use of the Python library `Pykov` [9], which can compute limiting distributions of regular Markov chains. For a discussion of the methods used by `Pykov`, see Stewart [11].

The code used to find our results can be found in [6]. At the expense of not following a strictly correct Monopoly board, our modeling decisions allow for more general boards. For instance, we can model arbitrary sized boards, number of dice, Jail and Goto Jail locations, and chance spaces. These features are not strictly useful when modeling standard Monopoly, but they are interesting generalizations.

5.3 Results

Our top ten most probable spaces are given in Table 1. (For a full list of results, refer to the code in [6].) Our results closely mirror those in [1], with a few differences due

Table 1: Ten most likely spaces in the Monopoly limiting distribution.

Space	Limiting Probability
Jail (10)	0.098797483065
Community Chest (17)	0.0275048069999
Tennessee Avenue (18)	0.0271924035748
New York Avenue (19)	0.0270245750067
Free Parking (20)	0.0269659953248
Kentucky Avenue (21)	0.0267537693537
B. & O. Railroad (25)	0.0264649765049
Atlantic Avenue (26)	0.0263975595999
Illinois Avenue (24)	0.0263950140309
St. James Place (16)	0.0263043157387

to technical decisions. In particular, reporting the four different jail spaces as one space, we see that most likely place to be is Jail. On the surface this is not useful, but knowing this makes properties near, but after Jail more valuble since players will have to move there after being in Jail. Sure enough, we see that the next most likely spaces after Jail are within twelve spaces of Jail itself.

5.4 Future Work

We have essentially repeated the analysis given by Abbot and Richey [1] on Monopoly, adding an implementation in Python. The most interesting way that this analysis could be extended is to introduce money into the model. That is, have states be tuples of (s, m), where s is the current space and m is the current amount of money a player has. Care would have to be taken to ensure that such a state space is actually finite. Even ignoring payments to other players, this becomes a very interesting model.

Perhaps the most practical future work would be to put this analysis to the test. We argue that the most likely spaces are the most valuable. By using a Monopoly simulation, we could write clients that prioritize investing in the most likely spaces and see how they fare against “naïve” clients.

Our Monopoly model breaks slightly from traditional rules to allow for more flexibility in constructing a board. Future work may attempt to bring our model more in line with standard Monopoly without sacrificing this flexibility.

References

[1] Stephen D. Abbot and Matt Richey. Take a Walk on the Boardwalk. *The College Mathematics Journal*, 28(3):162–171, 1997. ISSN 07468342, 19311346. URL http://www.jstor.org/stable/2687519.

[2] Linda JS Allen. *An introduction to stochastic processes with applications to biology*. CRC Press, 2010.

[3] K. Baclawski. *Introduction to Probability with R*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press, 2008. ISBN 9781420065220. URL https://books.google.com/books?id=Kglc9g5IPf4C.

[4] Gely P. Basharin, Amy N. Langville, and Valeriy A. Naumov. The life and work of AA Markov. *Linear Algebra and its Applications*, 386:3–26, 2004.

[5] Pierre Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer, 1998.

[6] Robert Dougherty-Bliss. boardwalk. https://github.com/rwbogl/boardwalk, 2016.

[7] Jason A. Osborne. Markov Chains for the RISK Board Game Revisited. *Mathematics Magazine*, 76(2):129–135, 2003. ISSN 0025570X, 19300980. URL http://www.jstor.org/stable/3219306.

[8] Patrick C Parks. AM Lyapunov's stability theory—100 years on. *IMA journal of Mathematical Control and Information*, 9(4):275–303, 1992.

[9] Riccardo Scalco. Pykov. https://github.com/riccardoscalco/Pykov, 2016.

[10] Eugene Seneta. Markov and the birth of chain dependence theory. *International Statistical Review/Revue Internationale de Statistique*, pages 255–263, 1996.

[11] Williams J Stewart. *Introduction to the numerical solutions of Markov chains*. Princeton Univ. Press, 1994.

[12] Richard Wilding. Official Monopoly® Game Rules. http://richard_wilding.tripod.com/monorules.htm, 2006.

[13] T. Yedlin. *Maxim Gorky: A Political Biography*. Praeger, 1999. ISBN 9780275966058. URL https://books.google.com/books?id=Zxt5_vI_ozcC.