Journal homepage http://revistas.unitru.edu.pe/index.php/SSMM

SELECCIONES MATEMÁTICAS
Universidad Nacional de Trujillo
ISSN: 2411-1783 (Online)
2022; Vol. 9(1): 53-78.

Markov chains and applications

Cadenas de Markov y aplicaciones

Mississippi Valenzuela

Received, Feb. 05, 2022 *Accepted, Jun. 29, 2022*

How to cite this article:
Valenzuela M. *Markov chains and applications*. Selecciones Matemáticas. 2022;9(1):53–78. http://dx.doi.org/10.17268/sel.mat.2022.01.05

Abstract

This work has three important purposes: first it is the study of Markov Chains, the second is to show that Markov chains have different applications and finally it is to model a process of this behaves. Throughout this work we will describe what a Markov chain is, what these processes are for and how these chains are classified. We will describe a Markov Chain, that is, analyze what are the primary elements that make up a Markov chain, among others.

Keywords . Markov chains, Poisson process, continuous-time Markov chain.

Resumen

Este trabajo tiene tres propósitos importantes: primero es el estudio de las Cadenas de Markov, el segundo es mostrar que las cadenas de Markov tienen diferentes aplicaciones y por último es modelar como se comporta un proceso de este tipo. A lo o largo de este trabajo describiremos que es una cadena de Markov, para que sirven estos procesos y como se clasifican dichas cadenas. Veremos además como se conforman estos procesos, es decir, analizar cuales son los elementos primordiales que conforman una cadena de Markov, entre otras.

Palabras clave. Cadena de Markov, proceso de Poisson, cadenas de Markov de tiempo continuo.

1. Introduction. The basic concepts of Markov chains were introduced by Andrew A. Markov [15], from Markov's work is when the development of stochastic processes formally begins. N. Wiener [22] was the first to rigorously treat the continuous case of the Markov chain and it was A. N. Kolmogorov during the 1930's who developed the general theory of stochastic processes [13]. From this moment on, a large number of mathematicians got involved giving it a great boom. The importance of studying chains as a study of random variables is that a large number of applications have the Markov property, this led to a large amount of research in the theory of stochastic processes [13, 6]. Markov chains are useful in certain branches of Physics such as Thermodynamics, Quantum Mechanics, in Meteorology it helps to have more accurate predictions in the change of time from one day to another, in Biological Sciences Epidemiological models are explained, in Game theory, Finance, or Social Sciences, Statistics and Mathematics [6][9]. The concept of the Markov Chain was undoubtedly one of Andrew A. Markov's greatest contributions, and it has been recognized over time.

This work will be structured as follows:

In Section 2 we give a brief review of Probability theory, going through conditional probability and using Bayes' formula.

Throughout section 3 we will define what a Markov Chain is, the states, their classification and we will see what happens with these processes in the long term. For this, an important result known as the Convergence Theorem will be demonstrated and we will apply this result. It is worth mentioning that we

*Universidad Nacional Autónoma de México, Instituto de Física, Departamento de Física Teórica, Apartado Postal 20-364, Cd. México 01000, México. (mvalenzuelalumat@uaz.edu.mx).

will analyze a simple way to see a string by using matrices (Transition Matrix). As Markov chains are sequences of random variables with a certain structure, they can be continuous or discrete, for our purposes we will only analyze discrete random variables.

In section 4, we will study Markovian processes. To do this, we analyze a population as a set whose elements have a common characteristic. The number of elements in a given population at time t is denoted by $N(t)$. The states of a growth process are the different ones that $N(t)$ can take; these are generally non-negative integers. is denoted by $p_n(t)$ to the probability that $N(t)$ takes a non-negative integer value n. In addition, a solution will be given to the Markovian process of birth-death with the help of the Kolmogorov equations.

In section 5 we will analyze some interesting applications; first the well-known board game Monopoly. This can be modeled using a Markov chain and we will study the long-term behavior with the Transition Matrix. We will analytically investigate whether the stationary distribution exists and simulate the chain to estimate its limit distribution using Python. In the second application we use Excel to solve the problem of how the weather changes from one day to the next. In particular we will use logical conditions and generate random numbers that will be the distribution of the random variable (time) and we will compare these values with the transition probabilities of the different states.

Finally, the conclusions of the work are presented.

2. Probability Review. The degree of randomness can be quantified with the concept of probability. The mathematical theory of probability has a history that goes back to the 17th century, and different definitions of probability have been developed. The definition of probability will be used in terms of set theory as formulated in 1933 by Kolmogorov [13]. When considering a set S called the sample space consisting of a certain number of elements, each subset A of S is assigned a real number $P(A)$ defined by the following three axioms:

1. For every subset $A \subset S$, $P(A) \geq 0$.
2. The probability assigned to the sample space is 1, $P(S) = 1$.
3. If two subsets A and B they are disjoint (i.e., mutually exclusive, $A \cap B = \emptyset$), then $P(A \cup B) = P(A) + P(B)$.

Suppose we have a space S of the sample that contains the subsets A y B. Provided $P(B) \neq 0$, one defines the conditional probability $P(A|B)$ (P of A given B) as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{2.1}$$

Two subsets A y B are said to be independent if

$$P(A \cap B) = P(A)\, P(B). \tag{2.2}$$

For A y B independient, it follows from the definition of conditional probability that $P(A|B) = P(A)$ and $P(B|A) = P(B)$. From equation (2.1) we have the probability of B given A (assuming $P(A) \neq 0$)

$$P(B|A) = \frac{P(B \cap A)}{P(A)}. \tag{2.3}$$

Since $A \cap B$ is the same as $B \cap A$, by combining equations (2.1) and (2.2) one has

$$P(B \cap A) = P(B|A)\, P(B) = P(A|B)\, P(A) \tag{2.4}$$

or

$$P(A|B) = \frac{P(B|A)\, P(A)}{P(B)}. \tag{2.5}$$

Equation (2.5), which relates the conditional probabilities $P(A|B)$ and $P(B|A)$, is called Bayes' theorem [18].

Suppose the sample space S can be broken into disjoint subsets A_i; i. e., $S = \cup_i A_i$ with $A_i \cap A_j = \emptyset$ for $i \neq j$. . Assume further that $P(A_i) \neq 0$ for all i. Then, an arbitrary subset B can be expressed as $B = B \cap S = B \cap (\cup_i A_i) = \cup_i (B \cap A_i)$. Since the subsets $B \cap A_i$ are disjoint, their probabilities add, giving

$$P(B) = P(\cup_i (B \cap A_i)) = \sum_i P(B \cap A_i) = \sum_i P(B|A_i)\, P(A_i). \tag{2.6}$$

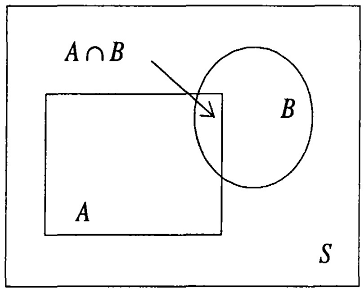

Figure 2.1: Relationship between the sets A, B and S in the definition of conditional probability.

The last line comes from the equation (2.3) for the case $A = A_i$. Equation (2.5) is called the law of total probability. it is useful, for example, if one can break the sample space into subsets A_i for which the probabilities are easy to calculate. It is often combined with Bayes' theorem (2.5) to give

$$P(A|B) = \frac{P(B|A)\,P(A)}{\sum_i P(B|A_i)\,P(A_i)}. \tag{2.7}$$

Of the possible ways to interpret a probability, the one most commonly found in the physical sciences is as a limiting frequency. That is, we interpret the elements of the sample space as possible outcomes of a measurement, and we take a $P(A)$ to mean the fraction of times that the outcome is in the subset A in the limit where we repeat the measurement an infinite number of times under "identical" conditions:

$$P(A) = \lim_{n\to\infty} \frac{\text{times outcomes is in } A}{n}. \tag{2.8}$$

Use of probability in this way leads to what is called the frequentist approach to statistics [**?**]. To define what is called subjective probability one interprets the elements of the set S as hypotheses, i.e., statements that are either true or false, and one defines

$$P(A) = \text{degree of believe that } A \text{ is true}. \tag{2.9}$$

Use of subjective probability leads to what is called Bayesian statistics, owing to its important use of Bayes' theorem described below.

In Bayesian (as opposed to frequentist) statistics, one uses subjective probability to describe one's degree of belief in a given theory or hypothesis. The denominator in Eq. (2.7) can be regarded as a constant of proportionality and therefore Bayes' theorem can be written as

$$P(\text{theory}|\text{data}) \propto P(\text{data}|\text{theory})\,P(\text{theory}), \tag{2.10}$$

where "theory" represents some hypothesis and "data" is the outcome of the experiment. Here $P(\text{teoría})$ is the prior probability for the theory, which reflects the experimenter's degree of belief before carrying out the measurement, and $P(\text{datos}|\text{teoría})$ is the probability to have gotten the data actually obtained, given the theory, which is also called the likelihood.

Bayesian statistics [7][10][14][11] provides no fundamental rule for obtaining the prior probability; in general this is subjective and may depend on previous measurements, theoretical prejudices, etc. Once this has been specified, however, Eq. (2.10) tells how the probability for the theory must be modified in the light of the new data to give the posterior probability, $P(\text{theory}|\text{data})$. As Eq. (2.10) is stated as a proportionality, the probability must be normalized by summing (or integrating) over all possible hypotheses.

3. Markov chains . Markov chains form a very important part of stochastic processes and probability theory, since they have a wide theory and a number of applications. There are different types of Markov chains and we will focus on the study of homogeneous chains, where they do not depend on time, having said this, we will begin by analyzing the formal definition of Markov chain as well as its different components. Taking into account that for the study of this theory a probability space $(\Omega, \mathfrak{I}, \mathbb{P})$ is necessary, where Ω is the sample space, $\mathfrak{I}$ is the family of all subsets of Ω and $\mathbb{P}$ is a probability function.

3.1. Definition and Markov property. There are different processes that can be modeled using a Markov chain and thanks to these we can see how our model changes as time progresses. Some classic examples of this type of model are: the ruin of the player [4], the Wright-Fisher model [5], the chain of

Ehrenfest [8], the time of a certain city, although it is not modeled in a very exact way we can approximate its behavior through this concept [6], there are a number of applications of this type of model, first we will define what a state is, as well as a space of states and later we will give the formal definition of a Markov chain [16][3].

Definition 3.1.

Let be X a random variable and $\mathcal{S}$ a set of numbers, let $i \in \mathcal{S}$ say that i is a state, if the random variable X takes the value of i, that is, $\mathbb{P}(X = i) > 0$, and the set $\mathcal{S}$ is known as the state space.

From definition 3.1 we can deduce that a state is the possible value that a random variable takes. A state space is the set of all possible states that a random variable can pass through. For practical purposes we will consider our state space to be finite or countable, furthermore we will take sets of integers, so that our random variables are discrete. In this way we will analyze the fundamental idea of this work, this is the formal definition of a Markov chain.

Definition 3.2.

Let $\{X_n, n \geq 0\}$ be a sequence of random variables, $\mathcal{S}$ is a states space, $i_0, i_1, ..., i_{n-1}, i, j \in \mathcal{S}$, if

$$\mathbb{P}(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, ..., X_0 = i_0) = \mathbb{P}(X_n = j | X_{n-1} = i). \tag{3.1}$$

We say that the sequence $\{X_n, n \geq 0\}$, is a Markov's chain.

Equation (3.1) is known as the Markov property, it tells us that the probability of a future event only depends on the immediate previous event and not on the evolution of the system, which implies that Markov chains are processes without memory [8]. Now that we understand what a Markov chain is, we can focus on certain chains that will be the object of our study, which leads us to the following definition:

Definition 3.3. *Let $\{X_n, n \geq 0\}$ a Markov chain with state space $\mathcal{S}$, we say that $\{X_n, n \geq 0\}$, is a homogeneous Markov chain, if*

$$\mathbb{P}(X_n = j | X_{n-1} = i) = \mathbb{P}(X_1 = j | X_0 = i), \ \forall i, j \in \mathcal{S}. \tag{3.2}$$

In other words, definition 3.3 refers to the fact that if a Markov chain does not depend on time, that is, it does not depend on n, it is considered a homogeneous chain [9], once we have already defined our object of study, we will use a function to see how a string evolves or moves between different states. This is the transition function, and it will allow us to see it more easily.

Definition 3.4.

Let $\{X_n, n \geq 0\}$ be a Markov chain with state space $\mathcal{S}$. We define the function $p(i,j)$ by

$$p(i,j) = \mathbb{P}(X_1 = j | X_0 = i). \tag{3.3}$$

Where $p(i,j)$ is the transition function from state i to state j, and has the following properties:

1. *$p(i,j) \geq 0 \, \forall i, j \in \mathcal{S}$.*
2. *$\sum_j p(i,j) = 1$, since when $X_n = i$, X_{n+1} goes somewhere j.*

It is necessary to mention that $p(i,j)$ is known as the one-step transition from state i to state j. In other words, if we are at time i and we want to get to j in a single step, we denote it by $p(i,j)$, from definition 3.4 it is clear to see that 1 is true since $p(i,j)$ is a probability, while property 2 refers to the fact that the sum of the probabilities of going from i to j is one.

3.2. Transition Matrix. Now that we know more about the transition function, which we will define in a simple way, that is, see how we can go from one state to another, better said, how the states of the chain interact with each other. That is why we will rely on some concepts of linear algebra to make it possible, which leads us to the following definition.

Definition 3.5.

Since $p(i,j)$ is the transition function of a Markov chain, we say that $p(i,j)$ is the $ij-th$ entry of the matrix P, and we will say that this is the transition matrix of the chain. Since P is formed by each of the transitions of the chain, we can affirm that P is a non-negative matrix, which implies that each of its entries is positive, this is true for 1 of definition 3.4.

From the above it is easy to deduce that throughout this work we will use non-negative matrices. In addition to being non-negative, each of the rows of these matrices adds up to 1 and we know that random variables go through different states as they move, but what will happen if the chain reaches a specific state and does not leave it [8], the state becomes an absorbing state, resulting in the following definition.

Definition 3.6.

Let $k \in \mathcal{S}$ be a state of the Markov chain, we say that k is an absorbing state if $p(k,k) = 1$.

3.3. Transition of m steps. Because we have seen this part of the theory, we wonder, what other use these Markov chains have? that is, now that we know a little about them, what more can we know about them apart from the information that each one gives us? model, we will dedicate ourselves to answering a simple question: What happens in the long term with these chains? [8], it is easy to deduce this since the model is given in the form of a matrix, we will relate this to what happens in more than one step. Basically analyze the power of the transition matrix [18]. Which brings us to the next definition.

Definition 3.7.

Let $\{X_n, n \geq 0\}$, a Markov chain, $p(i,j) = \mathbb{P}(X_{n+1} = i | X_n = i)$ the probability of going from i to j in one step, we define the probability of going from i to j in m steps, with $m > 1$, by

$$p^m(i,j) = \mathbb{P}(X_{n+m} = j | X_n = i). \tag{3.4}$$

It is necessary to see that this property is valid, for this let us suppose that we want to calculate

$$p^2(i,j) = \mathbb{P}(X_{n+2} = j | X_n = i).$$

We see that to arrive at $X_{n+2} = j$ but X_{n+1} must go through some state k, that is,

$$p^2(i,j) = \mathbb{P}(X_{n+2} = j | X_n = i),$$

$$= \sum_k \mathbb{P}(X_{n+2} = j, X_{n+1} = k | X_n = i),$$

$$= \sum_k \frac{\mathbb{P}(X_{n+2} = j, X_{n+1} = k, X_n = i)}{\mathbb{P}(X_n = i)},$$

$$= \sum_k \frac{\mathbb{P}(X_{n+2} = j, X_{n+1} = k, X_n = i)}{\mathbb{P}(X_n = i)} \cdot \frac{\mathbb{P}(X_{n+1} = k, X_n = i)}{\mathbb{P}(X_{n+1} = k, X_n = i)},$$

$$= \sum_k \frac{\mathbb{P}(X_{n+2} = j, X_{n+1} = k, X_n = i)}{\mathbb{P}(X_{n+1} = k, X_n = i)} \cdot \frac{\mathbb{P}(X_{n+1} = k, X_n = i)}{\mathbb{P}(X_n = i)},$$

$$= \sum_k \mathbb{P}(X_{n+2} = j | X_{n+1} = k, X_n = i) \cdot \mathbb{P}(X_{n+1} = k, X_n = i),$$

$$p^2(i,j) = \sum_k p(i,k)\, p(k,j). \tag{3.5}$$

We can notice that the last row is the $(i,j) - th$ entry of the matrix P^2. If followed by mathematical induction we can conclude this in the following.

Theorem 3.1.

The transition probability of m steps

$$p^m(i,j) = \mathbb{P}(X_{n+m} = j | X_n = i), \tag{3.6}$$

is the m-th power of the transition matrix P, that is, $P^m = P \cdots P$.

Proof: This proof follows by induction, once established that to go from i to j in m steps it is necessary to calculate the $m - th$ power of the transition matrix, the importance of this is to be able to prove the Chapman-Kolmogorov equation [8], the which is given in the following proposition. □

Lemma 3.1.

Let $n, m \in \mathbb{Z}^+$, then the probability of going from i to j in $m+n$ steps we have to go from i to j in $m+n$ steps is

$$p^{m+n}(i,j) = \sum_k p^m(i,k)\, p^n(k,j). \tag{3.7}$$

Proof:

Next we will dedicate ourselves to proving that equation (3.7) is correct, that is, we will show its truth taking into account that going from i to j in $m+n$ steps we have to go from i to k in m steps and from k to j in n steps, that is [8],

$$\mathbb{P}\left(X_{n+m}=j|X_0=i\right)=\sum_k \mathbb{P}\left(X_{m+n}=j,X_m=k|X_0=i\right),$$

$$=\sum_k \mathbb{P}\left(X_{m+n}=j,X_m=k|X_0=i\right),$$

$$=\sum_k \frac{\mathbb{P}\left(X_{m+n}=j,X_m=k,X_0=i\right)}{\mathbb{P}\left(X_0=i\right)},$$

$$=\sum_k \frac{\mathbb{P}\left(X_{m+n}=j,X_m=k,X_0=i\right)}{\mathbb{P}\left(X_0=i\right)}\cdot\frac{\mathbb{P}\left(X_m=k,X_0=i\right)}{\mathbb{P}\left(X_m=k,X_0=i\right)},$$

$$=\sum_k \frac{\mathbb{P}\left(X_{m+n}=j,X_m=k,X_0=i\right)}{\mathbb{P}\left(X_{m+n}=k,X_0=i\right)}\cdot\frac{\mathbb{P}\left(X_m=k,X_0=i\right)}{\mathbb{P}\left(X_0=i\right)},$$

$$=\sum_k \mathbb{P}\left(X_{m+n}=j|X_m=k,X_0=i\right)\cdot\mathbb{P}\left(X_m=k,X_0=i\right),$$

$$=\sum_k p^m\left(i,k\right)p^n\left(k,j\right).$$

□

3.4. Initial distribution. Once we have already analyzed some concepts, we can ask ourselves what happens if our first state is random, that is, consider the possibility that the first state of our Markov chain is a randomly generated state [8], if this were, we will take into account the next

$$\mathbb{P}\left(X_n=j\right)=\sum_k \mathbb{P}\left(X_0=i,X_n=j\right),$$

or

$$\mathbb{P}\left(X_n=j\right)=\sum_i \mathbb{P}\left(X_0=i\right)\mathbb{P}\left(X_n=j|X_0=i\right). \tag{3.8}$$

From equation (3.8) we can see that $\mathbb{P}\left(X_n=j|X_0=i\right)=p^n\left(i,j\right)$, which assuming that $\mathbb{P}\left(X_0=i\right)=q\left(i\right)$ it transforms in

$$\mathbb{P}\left(X_n=j\right)=\sum_i q\left(i\right)p^n\left(i,j\right). \tag{3.9}$$

In other words, equation (3.9) tells us to multiply the transition matrix by the initial probabilities from the left, in other words, for this operation to be well defined, since the transition matrix is of size $k\times k$ we need to multiply by a matrix of size $1\times k$ (row matrix), which leads us to the following definition.

Definition 3.8.

Let $\{X_n,n\geq 0\}$ be a Markov chain, we will call the initial distribution the row vector whose entries are the probability that the random variable starts in a state and will be denoted by π_0, that is,

$$\pi_0=\left(q\left(0\right),q\left(1\right),...,q\left(k\right)\right), \tag{3.10}$$

where $0,1,2,...,k\in\mathcal{S}$, $q\left(i\right)=\mathbb{P}\left(X_0=i\right)$ y $\sum_i q\left(i\right)=1$.

3.5. Stationary distribution. It would be interesting to know what happens when the initial distribution is equal to the distribution of the random variable X_1 [8], which leads us to the definition.

Definition 3.9.

If the distribution at time 0 is the same as at time 1, the Markov property assures us that it would be the distribution at all times n, and it will be called the stationary distribution, that is,

$$\pi\cdot P=\pi. \tag{3.11}$$

3.6. Limit distribution. As we have already seen throughout the previous sections, the long-term behavior of a Markov chain is important, because we see how the chain will behave. Like the initial distribution, we can define a row vector for each instant of it. The vector has as components the probability of starting in one of the states at time n [18], that is,

$$\pi_n = (\pi_n(1), ..., \pi_n(k)), \tag{3.12}$$

with

$$\pi_n(j) \geq 0, \quad \textstyle\sum_{j=0}^{k} \pi_n(j) = 1.$$

From where we get the following relationship

$$\pi_n(j) = \mathbb{P}(X_n = j) = \sum_{i=1}^{k} \mathbb{P}(X_0 = i)\,\mathbb{P}(X_n = j|X_0 = i)$$

or

$$\pi_n(j) = \sum_{i=1}^{k} \pi_0(i)\,\mathbb{P}(X_n = j|X_0 = i). \tag{3.13}$$

If we use matrix theory for distribution calculations, we can view equation (3.13) above as follows

$$\pi_n = \pi_0 P^n.$$

As already mentioned, every transition matrix P determines a succession of distributions $\pi_0, \pi_1, ...$ over the space of states $\mathcal{S}$ [18], and this is given by expression

$$\pi_n = \pi_{n-1}P = ... = \pi_0 P^n, \text{ with } n \geq 1. \tag{3.14}$$

Under certain conditions the previous sequence is convergent to a probability distribution π, suppose then that:

$$\pi = \lim_{n\to\infty} \pi_0. \tag{3.15}$$

Having said this, we will analyze the properties of the π distribution. Taking the limit as $n \to \infty$ in equality (3.14) we have

$$\pi = \pi P$$

and

$$\pi = \pi_0 \left(\lim_{n\to\infty} P^n\right). \tag{3.16}$$

This leads us to the analysis of several intuitive results:

1. Equation (3.11) tells us that the limit distribution is a stationary distribution.
2. Equation (3.11) indicates that the limit distribution does not depend on the initial distribution.
3. Equation (3.16) implies that the limit distribution is given by the n-th power of the matrix P.
4. From equation (3.16) the limit of the powers P is a matrix with all its rows equal and the entries of said matrix will be the elements of the limit distribution.

Now, we will analyze the formal definition of the limit distribution.

Definition 3.10.

We consider a Markov chain with transition matrix P and initial distribution π_0. In the Markov chain we will call the limit distribution of this chain the row matrix

$$\pi = \lim_{n\to\infty} \pi_0 P^n = \lim_{n\to\infty} \pi_0 p^n(i,j). \tag{3.17}$$

Next we will consider a couple of definitions about periodicity.

Definition 3.11.

Let $\{X_n, n \geq 0\}$ be a Markov chain with transition matrix P, we say that P is irreducible, if for each i and j we can get from i to j , that is, $p^m(i,j) > 0$ for some $m \geq 1$.

Definition 3.12.

Let i be a state of a Markov chain, we will say that i is an aperiodic state, if the greatest common divisor of $J_i = \{n \geq 1 | p^n(i,i) > 0\}$ is 1, that is:

$$\gcd(J_i) = 1.$$

Suppose the greatest common divisor of J_i is k, in other words, $k = \gcd(J_i)$ we will say that k is the period of the state i.

3.7. Convergence theorem.

Theorem 3.2 (Convergence theorem). *If P is irreducible and has an aperiodic state, then there is a unique stationary distribution π for any i and j, that is:*

$$\lim_{n\to\infty} p^n(i,j) = \pi(j). \tag{3.18}$$

Proof:
Let $\{Y_n, n \geq 0\}$, a independent Markov chain of the chain $\{X_n, n \geq 0\}$, but with the same transition matrix. In this way we define $\{Z_n, n \geq 0\}$. Since $Z_n = (X_n, Y_n)$ is a Markov chain with transition probability, then

$$\mathbb{P}(Z_n = (x_{n+1}, y_{n+1}) | Z_n = (x_n y_n)) = p(x_n, x_{n+1}) p(y_n, y_{n+1}). \tag{3.19}$$

We can easily verify that Z_n has a stationary distribution, that is,

$$\pi_Z = \pi_X \pi_Y. \tag{3.20}$$

To see that this is indeed true, let us take $(x_0, y_0) \in \mathcal{S}'$, where $\mathcal{S}'$ is the state space of $\{Z_n, n \geq 0\}$, thus

$$\sum_{(x_0,y_0)} \pi_Z((x_0,y_0)) p((x_0,y_0),(x,y)) = \sum_{x_0}\sum_{y_0} \pi_X(x_0)\pi_Y(y_0) p(x_0,x) p(y_0,y),$$

rearranging terms

$$\sum_{(x_0,y_0)} \pi_Z((x_0,y_0)) p((x_0,y_0),(x,y)) = \sum_{x_0} \pi_X(x_0) p(x_0,x) \sum_{y_0} \pi_Y(y_0) p(y_0,y)$$

or

$$\sum_{(x_0,y_0)} \pi_Z((x_0,y_0)) p((x_0,y_0),(x,y)) = \pi_X(x)\pi_Y(y) = \pi_Z((x,y)).$$

We see that, as P is irreducible then there exists a natural n_0, such that

$$p^n(x_0,x) > 0 \quad y \quad p^n(y_0,y) > 0 \text{ para toda } n \geq n_0.$$

So that

$$p^n((x_0,y_0),(x,y)) = p^n(x_0,x) p^n(y_0,y) \quad \forall n \geq n_0 \ .$$

The above is true because X_n and Y_n are aperiodic, so Z_n is recurrent positive and in particular is recurrent.
Let j be a state of the original chain (X_n), we define the first moment in which the chain $Z_n, n \geq 0$, first time visit to the state (j,j) like $\tau_j = \min\{n \geq 1 : Z_n = (j,j)\}$, let $\tau = \min\{n \geq 1 : X_n = Y_n\}$, and τ will be the first time the two chains match, like $Z_n, n \geq 0$, is recurring then $\mathbb{P}(\tau < \infty) = 1$, further $\tau \leq \tau_j$, by the Markov property

$$\mathbb{P}(X_n = x, \tau \leq n) = \sum_j \sum_{r=1}^{n} \mathbb{P}(X_n = x, X_r = j, \tau = r).$$

Through a transformation of the second member of the preceding equation, we have

$$\sum_j \sum_{r=1}^{n} \mathbb{P}(X_n = x | X_r = j, \tau = r) \mathbb{P}(X_r = j, \tau = r)$$

in equivalent form

$$\sum_j \sum_{r=1}^{n} \mathbb{P}\left(Y_n = x | Y_r = j, \tau = r\right) \mathbb{P}\left(Y_r = j, \tau = r\right),$$

or

$$\sum_j \sum_{r=1}^{n} \mathbb{P}\left(Y_n = x | Y_r = j\right) \mathbb{P}\left(Y_r = j, \tau = r\right).$$

Finally, we concluded that

$$\mathbb{P}\left(X_n = x, \tau \leq n\right) = \mathbb{P}\left(Y_n = x, \tau \leq n\right). \tag{3.21}$$

That is, about the event $(\tau \leq n)$, random variables X_n and Y_n have the same probability distribution, on the other hand

$$\mathbb{P}\left(X_n = j\right) = \mathbb{P}\left(X_n = j, \tau \leq n\right) + \mathbb{P}\left(X_n = j, \tau > n\right),$$

and by using equation (3.21) we can write

$$\mathbb{P}\left(X_n = j\right) = \mathbb{P}\left(Y_n = j, \tau \leq n\right) + \mathbb{P}\left(X_n = j, \tau > n\right),$$

or

$$\mathbb{P}\left(X_n = j\right) \leq \mathbb{P}\left(Y_n = j\right) + \mathbb{P}\left(\tau > n\right). \tag{3.22}$$

While

$$\mathbb{P}\left(Y_n = j\right) = \mathbb{P}\left(Y_n = j, \tau \leq n\right) + \mathbb{P}\left(Y_n = j, \tau > n\right),$$

and by virtue of equation (3.21) we can write

$$\mathbb{P}\left(Y_n = j\right) = \mathbb{P}\left(X_n = j, \tau \leq n\right) + \mathbb{P}\left(Y_n = j, \tau > n\right),$$

or

$$\mathbb{P}\left(Y_n = j\right) \leq \mathbb{P}\left(X_n = j\right) + \mathbb{P}\left(\tau > n\right). \tag{3.23}$$

From equations (3.22) and (3.23) we conclude that

$$\left|\mathbb{P}\left(X_n = j\right) - \mathbb{P}\left(Y_n = j\right)\right| \leq \mathbb{P}\left(\tau > n\right) \to 0, \tag{3.24}$$

when $n \to \infty$. Taking $X_0 = i$ with probability one, we can write

$$\mathbb{P}\left(X_n = j\right) = \mathbb{P}\left(X_n = j | X_0 = i\right) \mathbb{P}\left(X_0 = i\right) = p^n\left(i, j\right) \mathbb{P}\left(X_0 = i\right),$$

or

$$\mathbb{P}\left(X_n = j\right) = p^n\left(i, j\right). \tag{3.25}$$

If we take Y_0 with the stationary distribution π, then:

$$\mathbb{P}\left(Y_n = j\right) = \sum_i \mathbb{P}\left(Y_n = j | Y_0 = i\right) \pi\left(i\right) = \sum_i \pi\left(i\right) p^n\left(i, j\right)$$

too

$$\mathbb{P}\left(Y_n = j\right) = \pi\left(j\right). \tag{3.26}$$

Substituting in (3.24) we can conclude that

$$\left|p^n\left(i, j\right) - \pi\left(j\right)\right| \to 0. \tag{3.27}$$

This theorem is known as the convergence theorem [18]. □

Corollary 3.1.
If for n, $p^n\left(i, j\right) > 0$ for all i and j then there is a unique stationary distribution π and

$$\lim_{n \to \infty} p^n\left(i, j\right) = \pi\left(j\right). \tag{3.28}$$

Proof:
Like P is irreducible, then we can get to any state in n steps, that is, $p^n\left(i, j\right) > 0$, like i and j are arbitrary, then all states are aperiodic, thus $p^{n+1}\left(i, j\right) > 0$ so that $n, n+1 \in J$ and $\gcd J_i = 1$. □

3.8. Doubly Stochastic Chains. Since we have analyzed the fundamental concepts of Markov chains, now we will introduce a new idea, although related to the previous ones and based on it we will see if the different properties that we already have are still fulfilled.

Definition 3.13.

Let $\{X_n, n \geq 0\}$ be a Markov chain, where $\sum_j p(i,j) = 1$, further assuming that the chain satisfies the condition that $\sum_i p(i,j) = 1$, then $X_n, n \geq 0$, is a doubly stochastic Markov chain.

Proposition 3.1.

Let $\{X_n, n \geq 0\}$ be a doubly stochastic chain and if the chain has N states, then the stationary distribution is $\pi(i) = 1/N$, as

$$\sum_i \pi(i)\, p(i,j) = \frac{1}{N}. \tag{3.29}$$

Proof:

To see that (3.29) is valid, $\pi(i) = 1/N$ we have to $\sum_i \frac{1}{N} p(i,j) = \frac{1}{N} \sum_i p(i,j)$ and by definition 3.13 we know that $\sum_i p(i,j) = 1$ which leads us to affirm that (3.29) is true.

3.9. Continuous Time Chains . Throughout this work we have discussed the Markov chains in which the changes between the different states occur in a discrete manner, that is, in a fixed period of time. These chains are known as discrete time chains, however, in general the time periods are not necessarily fixed, that is, there is the possibility that the changes occur continuously over time, in which case we will call it a process of Markov. There is also the possibility that the periods are continuous random variables. We will call this type of process continuous-time Markov Chains. These chains are quite useful for solving models of queuing systems, manufacturing systems and re-manufacturing systems [6].

To get a clearer idea of this, we will consider the study of a process that is a clear example of a Markov chain of continuous time, this process is known as a Poisson process.

3.10. Poisson process. Next we will study one of the most important processes within the theory of continuous-time Markov chains. First we will analyze the definition of a Poisson process, having said this we will give the main characteristics of the Poisson process [6].

A process is called a Poisson process if:

1. The probability of occurrence of an event in the time interval $(t + \delta t)$ is $\lambda \delta t + o(\delta t)$. Where λ is a positive constant and $o(\delta t)$ is such that
$$\lim_{\delta t \to 0} \frac{o(\delta t)}{\delta t} = 0.$$
2. The probability of occurrence of none event in time interval $(t, t + \delta t)$ is $1 - \lambda t + o(\delta t)$.
3. The probability of occurrence of more than one event is $o(\delta t)$.

In this way, an event of this process can describe the arrival of a bus or a change of client [6]. From 1, 2 and 3, we can observe the Poisson distribution.

Proposition 3.2.

The following statements are equivalent to each other: (1) The arrival process of a Poisson process with rate λ. (2) Let $N(t)$ be the number of arrivals in the time interval $[0, t]$, then

$$P_n(N(t) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t},\ n = 0, 1, 2, ...$$

(3) The arrival time follows the exponential distribution with mean $-\lambda$.

Proof:

Let $P_n(t)$ be the probability that event n occurs in the interval $[0, t]$, we suppose that $P_n(t)$ is differentiable, then we can get a relation between $P_n(t)$ and $P_{n-1}(t)$ by

$$P_n(t + \delta t) = P_n(t)(1 - \lambda \delta t - o(\delta t)) + P_{n-1}(t)(\lambda \delta t + o(\delta t)) + o(\delta t).$$

By taking the limit when $\delta t \to 0$, we have

$$\lim_{\delta t \to 0} \frac{P_n(t + \delta t) - P_n(t)}{\delta t} = -\lambda P_n(t) + \lambda P_{n-1}(t) + \lim_{\delta t \to 0} \left[P_{n-1}(t) + P_n(t)\right] \frac{o(\delta t)}{\delta t},$$

or

$$\lim_{\delta t \to 0} \frac{P_n(t + \delta t) - P_n(t)}{\delta t} = -\lambda P_n(t) + \lambda P_{n-1}(t) + 0.$$

So we get a differential equation

$$\frac{dP_n(t)}{dt} = -\lambda P_n(t) + \lambda P_{n-1}(t), \; n = 0, 1, 2, ... \tag{3.30}$$

If we take $n = 0$ in (3.30) as $P_{-1}(t) = 0$ we get the following differential equation for $P_0(t)$

$$\begin{cases} \frac{dP_n(t)}{dt} = -\lambda P_0(t), \\ P_0(t) = 1, \end{cases}$$

where $P_0(0)$ is the probability that no event occurred in the interval $[0, 1]$ that's why it must be 1. Solving the equation for $P_0(t)$ we get

$$P_0(t) = e^{-\lambda t}. \tag{3.31}$$

Expression (3.31) is the probability that none event occurs in the interval $[0, t]$. Thus

$$1 - P_0(t) = 1 - e^{-\lambda t} \tag{3.32}$$

is the probability that at least one event occurred in the time interval $[0, t]$, so the probability density distribution $f(t)$ (for the waiting time and the first event occurs) is given by the exponential distribution, well known as

$$f(t) = \frac{d\left(1 - e^{-\lambda t}\right)}{dt} = \lambda e^{-\lambda t}, \quad t \geq 0.$$

Let's mention that

$$\begin{cases} \frac{dP_n(t)}{dt} = -\lambda P_n(t) + \lambda P_{n-1}(t), \; n = 0, 1, 2, ... \\ P_0(t) = e^{-\lambda t} \\ P_n(0) = 0, \; n = 1, 2, ... \end{cases}$$

We will start by solving this differential equation for $n = 1$. In this case

$$\frac{dP_1(t)}{dt} + \lambda P_1(t) = \lambda P_0(t),$$

or

$$\frac{dP_1(t)}{dt} + \lambda P_1(t) = \lambda e^{-\lambda t}.$$

Multiplying both sides by $e^{\lambda t}$ and integrating we get

$$\int \frac{d}{dt}\left[e^{\lambda t} P_1(t)\right] dt = \lambda \int t dt,$$

thus

$$P_1(t) = \lambda t e^{-\lambda t}.$$

For $n = 2$, we have

$$\frac{dP_2(t)}{dt} + \lambda P_2(t) = \lambda P_1(t),$$

or

$$\frac{dP_2(t)}{dt} + \lambda P_2(t) = \lambda^2 t e^{-\lambda t}.$$

Again if we multiply both sides of the above equation, then we can write

$$e^{-\lambda t} \frac{dP_2(t)}{dt} + \lambda e^{\lambda t} P_2(t) = \lambda^2 t.$$

In this way it is found

$$P_2(t) = \frac{(\lambda t)^2}{2!} e^{-\lambda t}.$$

For $n = 3$, we consider the differential equation

$$\frac{dP_3(t)}{dt} + \lambda P_3(t) = \lambda P_2(t)$$

or

$$\frac{dP_3(t)}{dt} + \lambda P_3(t) = \frac{\lambda^3 t^2}{2} e^{-\lambda t}.$$

If we multiply this equation by $e^{\lambda t}$

$$e^{\lambda t} \frac{dP_3(t)}{dt} + \lambda e^{\lambda t} P_3(t) = \frac{\lambda^3 t^2}{2},$$

whose solution is

$$P_3(t) = \frac{(\lambda t)^3}{3!} e^{-\lambda t}.$$

In general

$$P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

With this we can say that the Poisson process, the Poisson distribution and the exponential distribution are related to each other [6]. □

Now, we can conclude that the Poisson process is a clear example of a Continuous-time Markov chain.

3.11. A Two-State Continuous Markov Chain. We consider a server queue system that has two possible states: 0 (idle) and 1 (busy). Assume that the client arrival process is a Poisson process with mean rate λ and the server time follows the exponential distribution with mean rate μ. Let $P_0(t)$ be the probability that the server is down at time t, and $P_1(t)$ be the probability that the server is busy at time t. If we use the same argument as in the Poisson process [6], then we have

$$P_0(t + \delta t) = P_0(t)(1 - \lambda \delta t - o(\delta t)) + P_1(t)(\mu \delta t + o(\delta t)) + o(\delta t).$$

$$P_1(t + \delta t) = P_1(t)(1 - \mu \delta t - o(\delta t)) + P_0(t)(\lambda \delta t + o(\delta t)) + o(\delta t).$$

Rearranging the terms of the above equations, we find

$$\frac{P_0(t + \delta t) - P_0(t)}{\delta t} = -\lambda P_0(t) + \mu P_1(t) + [P_1(t) - P_0(t)] \frac{o(\delta t)}{\delta t},$$

$$\frac{P_1(t + \delta t) - P_1(t)}{\delta t} = \lambda P_0(t) - \mu P_1(t) + [P_0(t) - P_1(t)] \frac{o(\delta t)}{\delta t}.$$

If we take the limit when $\delta t \to 0$, then

$$\lim_{\delta t \to 0} \frac{P_0(t + \delta t) - P_0(t)}{\delta t} = -\lambda P_0(t) + \mu P_1(t) + \lim_{\delta t \to 0} [P_1(t) - P_0(t)] \frac{o(\delta t)}{\delta t}$$

$$\lim_{\delta t \to 0} \frac{P_1(t + \delta t) - P_1(t)}{\delta t} = \lambda P_0(t) - \mu P_1(t) + \lim_{\delta t \to 0} [P_0(t) - P_1(t)] \frac{o(\delta t)}{\delta t},$$

which leads to

$$\lim_{\delta t \to 0} \frac{P_0(t + \delta t) - P_0(t)}{\delta t} = -\lambda P_0(t) + \mu P_1(t)$$

$$\lim_{\delta t\to 0}\frac{P_1\left(t+\delta t\right)-P_1\left(t\right)}{\delta t}=\lambda P_0\left(t\right)-\mu P_1\left(t\right),$$

or

$$\frac{dP_0\left(t\right)}{dt}=-\lambda P_0\left(t\right)+\mu P_1\left(t\right),$$

$$\frac{dP_1\left(t\right)}{dt}=\lambda P_0\left(t\right)-\mu P_1\left(t\right).$$

Consequently, we obtain a system of differential equations, with $P_1\left(0\right)=1$, of the form

$$\begin{pmatrix}\frac{dP_0(t)}{dt}\\ \frac{dP_1(t)}{dt}\end{pmatrix}=\begin{pmatrix}-\lambda & \mu\\ \lambda & -\mu\end{pmatrix}\begin{pmatrix}P_0\left(t\right)\\ P_1\left(t\right)\end{pmatrix},$$

where

$$A=\begin{pmatrix}-\lambda & \mu\\ \lambda & -\mu\end{pmatrix}.$$

By calculating the eigenvalues for A, so $\det\left(A-\pi\mathbf{I}\right)=0$, that is,

$$\begin{vmatrix}-\lambda-\pi & \mu\\ \lambda & -\mu-\pi\end{vmatrix}=\pi^2+\left(\lambda+\mu\right)\pi=0.$$

When solving $\pi^2+\left(\lambda+\mu\right)\pi=0$, we see that the eigenvalues are $\pi_1=0$ y $\pi_2=-\left(\lambda+\mu\right)$, now we will calculate the respective eigenvectors to each of the eigenvalues, in this way for $\pi_1=0$, we have $\left(A-0\mathbf{I}\right)X=0$, therefore

$$\begin{pmatrix}-\lambda & \mu\\ \lambda & -\mu\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}.$$

So we get

$$\begin{pmatrix}-\lambda & \mu\\ \lambda & -\mu\end{pmatrix}\to\begin{pmatrix}1 & -\mu/\lambda\\ \lambda & -\mu\end{pmatrix}\to\begin{pmatrix}1 & -\mu/\lambda\\ 0 & 0\end{pmatrix}.$$

Whereby $x_1=\frac{\mu}{\lambda}x_2$, if we do $x_2=\lambda$then we see that $v_1=\begin{pmatrix}\mu\\ \lambda\end{pmatrix}$. Now we compute the eigenvector for $\pi_2=-\left(\lambda+\mu\right)$, that is, $\left(A+\left(\lambda+\mu\right)I\right)X=0$, therefore

$$\begin{pmatrix}\lambda & \mu\\ \lambda & \mu\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}.$$

So, we find

$$\begin{pmatrix}\lambda & \mu\\ \lambda & \mu\end{pmatrix}\to\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}.$$

Like $x_1=-x_2$, if $x_2=-1$, then for $\pi_2=-\left(\lambda+\mu\right)$, is obtained $v_2=\begin{pmatrix}1\\ -1\end{pmatrix}$. In this way we can say that the solution of the system is

$$\begin{pmatrix}P_0\left(t\right)\\ P_1\left(t\right)\end{pmatrix}=C_1\begin{pmatrix}\mu\\ \lambda\end{pmatrix}+C_2\begin{pmatrix}1\\ -1\end{pmatrix}e^{-(\lambda+\mu)t}.$$

In virtue of $P_1(0) = 1$, we can write

$$C_1 \begin{pmatrix} \mu \\ \lambda \end{pmatrix} + C_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

where $C_1 = \frac{1}{\lambda+\mu}$ and $C_2 = -\frac{\mu}{\lambda+\mu}$, whereby

$$P_0(t) = \frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t},$$

$$P_1(t) = \frac{\mu}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t}.$$

Since the probabilities of the stable states are given by

$$\lim_{t\to\infty} P_0(t) = \frac{\mu}{\lambda+\mu},$$

$$\lim_{t\to\infty} P_1(t) = \frac{\mu}{\lambda+\mu}.$$

With this we can say that it is not necessary to solve the system of differential equations to find the probability distribution of the stable state, that is, it can be seen that both $P_0(t)$ and $P_1(t)$ when $t \to \infty$ are constant, that is, they do not depend on t, if we take $P_0(t) = p_0$ and $P_1(t) = p_1$ then

$$\frac{P_0(t)}{dt} = \frac{dp_0}{dt} = 0,$$

$$\frac{P_1(t)}{dt} = \frac{dp_1}{dt} = 0.$$

By reducing the problem to calculate the following system of linear equations, the probability of the stable state is calculated, that is,

$$\begin{pmatrix} -\lambda & \mu \\ \lambda & -\mu \end{pmatrix} \begin{pmatrix} P_0 \\ P_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Finally, taking into account that $p_0 + p_1 = 1$, we have

$$\left(\begin{array}{cc|c} \lambda & \mu & 0 \\ 1 & 1 & 1 \end{array}\right) \to \ldots \to \left(\begin{array}{cc|c} 1 & 0 & \mu/(\lambda+\mu) \\ 0 & 1 & \lambda/(\lambda+\mu) \end{array}\right).$$

Whereby $p_0 = \mu/(\lambda+\mu)$ and $p_1 = \lambda/(\lambda+\mu)$. With this it can be affirmed that the distribution of the stable state of the Markov chain is due to the fact that the indicators of the system; such as the expected number of customers, and the mean waiting time, can be written in terms of the steady-state probability distribution [6].

4. Markovian birth-death processes. A population is a set whose elements have a common characteristic. The number of elements in a given population at time t is denoted by $N(t)$. The states of a growth process are the different ones that $N(t)$ can take; these are generally non-negative integers. $p_n(t)$ denotes the probability that $N(t)$ takes a non-negative integer value n.

A birth occurs each time a new element joins the population; a death occurs each time a member leaves the population. A pure process of birth is one in which there are only births and no deaths; a pure process of death is one in which there are only deaths and no births.

4.1. Markov birth-death processes, generalized. A population growth process is a Markovian process if the transition probabilities to change from one state to another depend only on the current state and not on the past history of the process reaching the current state. More formally, a generalized Markovian birth-death process [12][1][23] satisfies the following criteria:

1. Probabilistic distributions that control for the number of births and deaths during a time interval depend on the length of the interval, but not on its initial point.

2. The probability of exactly one birth in the interval of duration δt, given a population of n elements at the beginning of the interval, is $\lambda_n \Delta t + o(\Delta t)$, where λ_n is a possibly different constant for different values of n.
3. The probability of exactly one death in an interval of duration Δt, given a population of n elements at the start of the interval, is $\mu_n \Delta t + o(\Delta t)$, where μ_n is a possibly different constant for different values of n.
4. The probability of more than one birth and the probability of more than one death in an interval of duration Δt are both $o(\Delta t)$.

These criteria imply, in the limit when $\Delta t \to 0$, the Kolmogorov equations [13] for the state probabilities:

$$\frac{dP_n(t)}{dt} = -(\lambda_n + \mu_n) p_n(t) + \mu_{n+1} P_{n+1}(t) + \lambda_{n-1} P_{n-1}(t), \tag{4.1}$$

$$\frac{dP_0(t)}{dt} = -\lambda_0 p_0(t) + \mu_1 P_1(t), \tag{4.2}$$

where $n = 1, 2, ...$

4.2. Markovian processes of birth, linear. A linear Markovian birth process is a pure Markovian birth process [20], in which the probability of a birth in a small interval is proportional to both the actual number of elements in the population and the length of the interval. That is, for all n, $\mu_n = 0$ and $\lambda_n = n\lambda$. The proportionality constant λ is the birth rate or arrival rate. The solution to equations (4.1) and (4.2), for a population of one element, is:

$$p_n(t) = \begin{cases} [1 - \exp(-\lambda t)]^{n-1} \exp(-\lambda t) & (n = 1, 2, ...) \\ 0 & (n = 0) \end{cases} \tag{4.3}$$

The expected size of the population at the time t is $E[N(t)] = \exp(\lambda t)$. If the initial population has $N(0)$ elements, then its expected size at time t is:

$$E[N(t)] = N(0) \exp(\lambda t). \tag{4.4}$$

4.3. Markovian processes of death, linear. A linear Markovian death process is a pure Markovian death process [20], in which the probability of a death in a small interval is proportional to both the actual number of elements in the population and the length of the interval. That is, for all n, $\mu_n = n\mu$ y $\lambda_n = 0$. The proportionality constant μ is the birth rate or death rate. The solution to equations (4.1) and (4.2), for an initial population $N(0)$, is:

$$p_n(t) = \begin{cases} \begin{pmatrix} N(0) \\ n \end{pmatrix} [1 - \exp(-n\mu t)]^{N(0)-n} \exp(-n\mu t) & [n \le N(0)] \\ 0 & [n > N(0)] \end{cases} \tag{4.5}$$

The expected size of the population at time t is $E[N(t)] = \exp(\lambda t)$. If the initial population has $N(0)$ elements, then its expected size at time t is:

$$E[N(t)] = N(0) \exp(-\mu t). \tag{4.6}$$

4.4. Markovian birth-death processes, linear. A linear Markovian birth-death process [20] is a Markovian birth-death process in which, for all n, $\lambda_n = n\lambda$ and $\mu_n = n\mu$. The solution to equations (4.1) and (4.2), for an initial population of one element, is:

$$p_n(t) = \begin{cases} [1 - r(t)][1 - s(t)][s(t)]^{n-1} & (n = 1, 2, ...) \\ r(t) & (n = 0) \end{cases} \tag{4.7}$$

where

$$r(t) = \frac{\mu [\exp(\lambda - \mu) t - 1]}{\lambda \exp[(\lambda - \mu) t] - \mu}$$

and

$$s(t) = \frac{\lambda\left[\exp(\lambda-\mu)t-1\right]}{\lambda\exp\left[(\lambda-\mu)t\right]-\mu}.$$

The expected size of the population at the time t is $E[N(t)] = \exp[(\lambda-\mu)t]$. If the initial population has $N(0)$ elements, then its expected size at time t is:

$$E[N(t)] = N(0)\exp[(\lambda-\mu)t]. \tag{4.8}$$

It is clear that the linear birth-death process includes the linear birth process and the linear death process as the special cases $\mu = 0$ and $\lambda = 0$, respectively.

4.5. Poissonian processes of birth. A Poisson birth process [19] is a pure Markovian birth process, in which the probability of a birth in a small interval is proportional to both the actual number of elements in the population and the length of the interval. That is, for all n, $\mu_n = 0$ and $\lambda_n = n\lambda$. In this type of process, new arrivals in the population are not a result of the current number of elements. New elements can arrive in the population even when the current state is 0, which represents a marked difference in relation to the linear Markov processes of birth.

The solution of the equations (4.1) and (4.2) for an initial population of 0 is:

$$p_n(t) = \frac{(\lambda t)^n}{n!}\exp(-\lambda t) \quad (n = 0, 1, 2, 3, ...) \tag{4.9}$$

If the population starts with $N(0)$ elements, the solution to equations (4.1) and (4.2) is:

$$p_n(t) = \begin{cases} \frac{(\lambda t)^{n-N(0)}\exp(-\lambda t)}{[n-N(0)]!} & [n \geq N(0)] \\ 0 & [n < N(0)] \end{cases} \tag{4.10}$$

The expected size of the population at time t is:

$$E[N(t)] = N(0) + \lambda t. \tag{4.11}$$

We can condense (4.9) and (4.10), in a Poissonian birth process with birth rate λ, $N(t) - N(0)$ has a Poisson distribution, with parameter λt. Furthermore, in such a case, the interarrival time, which is the time between successive births, has an exponential distribution with expected value $1/\lambda$.

4.6. Poissonian processes of death. A Poissonian death process [19] is a pure Markovian death process, in which the probability of a death in a small interval is proportional to both the actual number of elements in the population and the length of the interval. That is, for all n, $\mu_n = \mu$ and $\lambda_n = 0$. The solution to equations (4.1) and (4.2), for an initial population, is:

$$p_n(t) = \begin{cases} 0 & [n > N(0)] \\ \frac{(\mu t)^{N(t)-n}\exp(-\mu t)}{[N(0)-n]!} & [1 \leq n \leq N(0)] \\ 1 - \sum_{n=1}^{N(0)} p_n(t) & (n = 0) \end{cases} \tag{4.12}$$

4.7. Poissonian birth-death processes, linear. A Poissonian birth-death process is a Markovian birth-death process in which both the probability of a birth and a death in any small interval is independent of population size. That is, for all n, $\lambda_n = \lambda$ and $\mu_n = \mu$. These are the basis of queuing theory [19] [2] [17].

4.8. Solution to the Markovian processes of birth-death. With $\lambda_n = n\lambda$ and $\mu_n = n\mu$, Kolmogorov equations (4.1) and (4.2), become

$$\frac{dP_n(t)}{dt} = -n(\lambda+\mu)p_n(t) + (n+1)\mu P_{n+1}(t) + (n-1)\lambda P_{n-1}(t), \tag{4.13}$$

for $n = 1, 2, ...$ and

$$\frac{dP_0(t)}{dt} = \mu P_1(t). \tag{4.14}$$

One way to solve these equations is by substituting them as a single partial differential equation for the probabilistic generating function

$$F(z,t)=\sum_{n=0}^{\infty}p_n(t)\,z^n. \tag{4.15}$$

The procedure is the next. Multiply (4.13) by z^n, add for all n, and add the result to (4.14), leaving after sorting,

$$\sum_{n=0}^{\infty}\frac{dP_n(t)}{dt}z^n=-(\lambda+\mu)\sum_{n=1}^{\infty}np_n(t)\,z^n+\mu\sum_{n=0}^{\infty}(n+1)\,P_{n+1}(t)\,z^n+\lambda\sum_{n=0}^{\infty}(n-1)\,P_{n-1}(t)\,z^n. \tag{4.16}$$

But differentiating (4.15),

$$\sum_{n=0}^{\infty}\frac{dp_n(t)}{dt}z^n=\frac{\partial F(z,t)}{\partial t},$$

$$\sum_{n=0}^{\infty}np_n(t)\,z^n=z\frac{\partial F(z,t)}{\partial z},$$

$$\sum_{n=0}^{\infty}(n+1)\,p_{n+1}(t)\,z^n=\frac{\partial F(z,t)}{\partial z},$$

$$\sum_{n=0}^{\infty}(n-1)\,p_{n-1}(t)\,z^n=z^2\frac{\partial F(z,t)}{\partial z}.$$

Then, (4.16) becomes,

$$\frac{\partial F(z,t)}{\partial t}=-(\lambda+\mu)\,z\frac{\partial F(z,t)}{\partial z}+\mu\frac{\partial F(z,t)}{\partial z}+\lambda z^2\frac{\partial F(z,t)}{\partial z}. \tag{4.17}$$

Solving this partial differential equation by separation of variables, one solution is found to be:

$$e^t\left(\frac{z-1}{z-\mu/\lambda}\right)^{1/(\mu-\lambda)}.$$

The general solution to equation (4.17) is:

$$F(z,t)=g\left[e^t\left(\frac{z-1}{z-\mu/\lambda}\right)^{1/(\lambda-\mu)}\right], \tag{4.18}$$

where g is an arbitrary function of one variable. To determine g, note that for an initial population of one element: $p_1(0)=1$ and $p_n(0)=0$ $(n\neq 0)$; then,

$$F(z,0)=\sum_{n=0}^{\infty}p_n(0)\,z^n=z. \tag{4.19}$$

Applying this initial condition to (4.18), we obtain:

$$z=g\left[\left(\frac{z-1}{z-\mu/\lambda}\right)^{1/(\lambda-\mu)}\right]. \tag{4.20}$$

Making

$$y=\left(\frac{z-1}{z-\mu/\lambda}\right)^{1/(\lambda-\mu)}$$

we have inversely

$$z=\frac{(\mu/\lambda)\,y^{\lambda-\mu}-1}{y^{\lambda-\mu}-1}$$

and (4.20) can be written as

$$g(y) = \frac{(\mu/\lambda)\, y^{\lambda-\mu} - 1}{y^{\lambda-\mu} - 1}. \tag{4.21}$$

So (4.18) will be expressed by

$$F(z,t) = \frac{z\left[e^{t}\left(\frac{z-1}{z-(\mu/\lambda)}\right)^{1/(\lambda-\mu)}\right]^{\lambda-\mu} - 1}{\left[e^{t}\left(\frac{z-1}{z-(\mu/\lambda)}\right)^{1/(\lambda-\mu)}\right]^{\lambda-\mu} - 1}, \tag{4.22}$$

which simplifies to

$$F(z,t) = \frac{\mu\left[e^{t(\lambda-\mu)} - 1\right] + z\left[-\mu e^{t(\lambda-\mu)} + \lambda\right]}{\left[\lambda e^{t(\lambda-\mu)} - \mu\right] - z\lambda\left[e^{t(\lambda-\mu)} - 1\right]}. \tag{4.23}$$

Finally, we need to expand $F(z,t)$ into powers of z, thus obtaining $p_n(t)$ as a coefficient of z^n. Be done

$$r(t) = \frac{\mu\left[\exp(\lambda-\mu)\,t - 1\right]}{\lambda\exp\left[(\lambda-\mu)\,t\right] - \mu}$$

$$s(t) = \frac{\lambda\left[\exp(\lambda-\mu)\,t - 1\right]}{\lambda\exp\left[(\lambda-\mu)\,t\right] - \mu}$$

and

$$m(t) = \frac{\lambda - \mu e^{t(\lambda-\mu)}}{\lambda e^{t(\lambda-\mu)} - \mu}.$$

Then

$$F(z,t) = \frac{r(t) + z m(t)}{1 - z s(t)}. \tag{4.24}$$

For the geometric series, $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$ ($|x| < 1$), equation (4.24) gives:

$$F(z,t) = r(t) + \sum_{n=1}^{\infty}\left[r(t)\,s(t) + m(t)\right]\left[s(t)\right]^{n-1} z^n.$$

It can be verified algebraically that

$$r(t)\,s(t) + m(t) = \left[1 - r(t)\right]\left[1 - s(t)\right].$$

Then

$$F(z,t) = r(t) + \sum_{n=1}^{\infty}\left\{\left[1 - r(t)\right]\left[1 - s(t)\right]\left[s(t)\right]^{n-1}\right\} z^n. \tag{4.25}$$

The coefficients in (4.25) give equation (4.7).

It can be verified that any power of a solution to (4.16) is itself a solution. In particular,

$$\Phi(z,t) = \left[F(z,t)\right]^{N(0)},$$

where $F(z,t)$ is given by (4.23) or (4.25), it is a solution; and this solution satisfies the initial condition

$$\Phi(z,0) = \left[F(z,0)\right]^{N(0)} = z^{N(0)}.$$

Then, $\Phi(z,t)$ is the generating function of the state probabilities for a population that starts with $N(0)$ elements. The fact that Φ is equal to $F^{N(0)}$ implies that the random variable corresponding to Φ can be expressed as the sum of $N(0)$ independent random variables, each one corresponding to F.

5. Applications of Markov chains .

5.1. Monopoly Chain. This section is dedicated to the board game known as Monopoly. This game is played on a board consisting of 40 boxes (see Figure 5.1), each box has its name. For practical purposes we will consider labeling them from 0 to 39, the game is very simple as it is played with two dice and we move around the board adding the number on the face of the dice after rolling them. For this work we will omit some small details to facilitate the construction of the Markov chain for this game: First we will omit the fact that if a person falls in jail they stay in it until they get pairs, or three turns pass. Second, we will consider the Communal Ark and Fortuna slots as common slots, just like the rest of the slots.

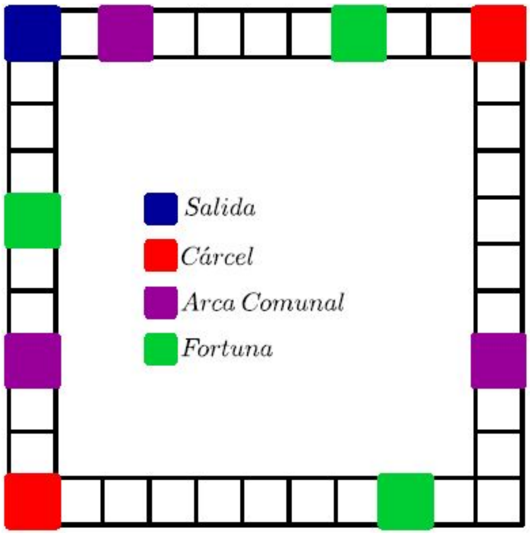

Figure 5.1: Monopoly board.

The first thing we have to notice is that the chain that we will model has as state space $\mathcal{S} = \{0, 1, 2, 3, ..., 38, 39\}$, let's take into account that, as we play with two dice, the minimum sum that can be obtained by throwing a pair of these is 2, while the maximum number of squares that we can advance is 12, so we will call r_k the probability that the sum of the dice is k, in other words, $r_2 = 1/36$, $r_3 = 2/36$,...,$r_{11} = 2/36$, $r_3 = 1/36$. It is easy to see that $\sum_{k=2}^{12} r_k = 1$, in this way we consider that $X_n, n \geq 0$, is the possibility that a player is in square n in the n-th turn $X_n, n \geq 0$ is a Markov chain. Now we will define the transition probability, that is, the probability of going from being in square i to square j in one turn, with $i, j \in \mathcal{S}$, and we will define it as follows:

$$p\left(i,j\right) = r_k \text{ si } j + k \bmod (40) . \tag{5.1}$$

5.1.1. Transition Matrix. To get an idea of this distribution function, suppose we are in box number 38, and when we roll a die we get an 8, so $38 + 8 = 46$ and $46\text{mod}\left(40\right) = 6$, so $p\left(38, 6\right) = r_8 = 5/36$. So the transition matrix is given by

$$P = \begin{pmatrix} 0 & 0 & 0.028 & . & . & . & 0 & 0 & 0 \\ 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ . & . & . & . & & & . & . & . \\ . & . & . & & . & & . & . & . \\ . & . & . & & & . & . & . & . \\ 0.056 & 0.083 & 0.111 & . & . & . & 0 & 0 & 0.028 \\ 0.028 & 0.056 & 0.083 & . & . & . & 0 & 0 & 0 \\ 0 & 0.028 & 0.056 & . & . & . & 0 & 0 & 0 \end{pmatrix} .$$

To see what a Markov chain is, let's take into account that $\sum_i p\left(i,j\right) = 1$, for each i, as well as $p\left(i,j\right) \geq 0$ for all $i, j \in \mathcal{S}$, the interesting thing about a Markov chain is to see how it behaves in the long run. With the help of Python we will do the necessary calculations to see how this chain behaves, using the necessary commands to avoid doing the calculations by hand, since the transition matrix has size 40×40, thus calculating P^2 can be found

$$P^2 = \begin{pmatrix} 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ . & . & . & . & & & . & . & . \\ . & . & . & & . & & . & . & . \\ . & . & . & & & . & . & . & . \\ 0 & 0.00784 & 0.003136 & . & . & . & 0 & 0 & 0 \\ 0 & 0 & 0.000784 & . & . & . & 0 & 0 & 0 \\ 0 & 0 & 0 & . & . & . & 0 & 0 & 0 \end{pmatrix}.$$

Again with the help of the software, we calculate P^4

$$P^4 = \begin{pmatrix} 3.62\times10^{-3} & 1.96\times10^{-3} & 1.00\times10^{-3} & . & . & . & 1.50\times10^{-2} & 9.89\times10^{-3} & 6.15\times10^{-3} \\ 6.18\times10^{-3} & 3.59\times10^{-3} & 1.98\times10^{-3} & . & . & . & 2.16\times10^{-2} & 1.50\times10^{-2} & 9.89\times10^{-3} \\ 9.94\times10^{-3} & 6.15\times10^{-3} & 3.62\times10^{-3} & . & . & . & 2.97\times10^{-2} & 2.16\times10^{-2} & 1.50\times10^{-2} \\ . & . & . & . & & & . & . & . \\ . & . & . & & . & & . & . & . \\ . & . & . & & & . & . & . & . \\ 4.63\times10^{-4} & 1.89\times10^{-4} & 6.87\times10^{-5} & . & . & . & 3.59\times10^{-3} & 1.96\times10^{-3} & 9.92\times10^{-4} \\ 1.00\times10^{-3} & 4.57\times10^{-4} & 1.90\times10^{-4} & . & . & . & 6.15\times10^{-3} & 3.59\times10^{-3} & 1.96\times10^{-3} \\ 1.98\times10^{-3} & 9.92\times10^{-4} & 4.61\times10^{-4} & . & . & . & 9.89\times10^{-3} & 6.15\times10^{-3} & 3.59\times10^{-3} \end{pmatrix}.$$

From P^4 we can see that all the entries in the transition matrix are positive. With this we can conclude that P is irreducible and aperiodic, this is even easier to see since the maximum sum of the faces of the dice is 12, if we throw the dice 4 times, the minimum we can advance is 8 squares, while the maximum squares that can be advanced is 48, and basically with 4 launches we can go around the board completely. In the case that in the four releases we obtained quantities very close to 12. Continuing with the calculations to see the long-term behavior, we have

$$P^8 = \begin{pmatrix} 0.0038 & 0.0053 & 0.0073 & . & . & . & 0.0016 & 0.0020 & 0.0027 \\ 0.0027 & 0.0038 & 0.0054 & . & . & . & 0.0014 & 0.0016 & 0.0029 \\ 0.0020 & 0.0027 & 0.0038 & . & . & . & 0.0015 & 0.0014 & 0.0016 \\ . & . & . & . & & & . & . & . \\ . & . & . & & . & & . & . & . \\ . & . & . & & & . & . & . & . \\ 0.0098 & 0.0128 & 0.0164 & . & . & . & 0.0038 & 0.0053 & 0.0073 \\ 0.0073 & 0.0098 & 0.0098 & . & . & . & 0.0027 & 0.0038 & 0.0053 \\ 0.0054 & 0.0073 & 0.0098 & . & . & . & 0.0020 & 0.0027 & 0.0038 \end{pmatrix}.$$

Squaring the above matrix, we find

$$P^{16} = \begin{pmatrix} 0.0210 & 0.0206 & 0.0204 & . & . & . & 0.0359 & 0.0339 & 0.0317 \\ 0.0317 & 0.0293 & 0.0206 & . & . & . & 0.0377 & 0.0359 & 0.0339 \\ 0.0340 & 0.0317 & 0.0210 & . & . & . & 0.0391 & 0.0377 & 0.0359 \\ . & . & . & . & & & . & . & . \\ . & . & . & & . & & . & . & . \\ . & . & . & & & . & . & . & . \\ 0.0220 & 0.0196 & 0.0175 & . & . & . & 0.0293 & 0.0268 & 0.0244 \\ 0.0244 & 0.0219 & 0.0197 & . & . & . & 0.0317 & 0.0293 & 0.0268 \\ 0.0269 & 0.0244 & 0.0220 & . & . & . & 0.0339 & 0.0317 & 0.0293 \end{pmatrix}.$$

After 16 stages, or turns, we will calculate for 32, and all of the above is possible thanks to the Chapman-Kolmogorov equation

$$P^{32} = \begin{pmatrix} 0.0210 & 0.0206 & 0.0204 & . & . & . & 0.0227 & 0.0220 & 0.02149 \\ 0.0317 & 0.0210 & 0.0206 & . & . & . & 0.0234 & 0.0227 & 0.02207 \\ 0.0221 & 0.0214 & 0.0210 & . & . & . & 0.0241 & 0.0234 & 0.02272 \\ . & . & . & . & & & . & . & . \\ . & . & . & & . & & . & . & . \\ . & . & . & & & . & . & . & . \\ 0.0202 & 0.0201 & 0.0202 & . & . & . & 0.0210 & 0.0206 & 0.02034 \\ 0.0203 & 0.0201 & 0.0201 & . & . & . & 0.0214 & 0.0210 & 0.02062 \\ 0.0206 & 0.0203 & 0.0202 & . & . & . & 0.0220 & 0.0214 & 0.02101 \end{pmatrix}.$$

If we continue squaring, we can see that

$$P^{64} = \begin{pmatrix} 0.0253 & 0.0253 & 0.0255 & . & . & . & 0.0250 & 0.0251 & 0.0252 \\ 0.0252 & 0.0253 & 0.0256 & . & . & . & 0.0250 & 0.0250 & 0.0251 \\ 0.0251 & 0.0252 & 0.0256 & . & . & . & 0.0249 & 0.0250 & 0.0250 \\ . & . & . & . & & & . & . & . \\ . & . & . & & . & & . & . & . \\ . & . & . & & & . & . & . & . \\ 0.0255 & 0.0255 & 0.0256 & . & . & . & 0.0253 & 0.0253 & 0.0254 \\ 0.0254 & 0.0255 & 0.0256 & . & . & . & 0.0252 & 0.0253 & 0.0253 \\ 0.0254 & 0.0254 & 0.0255 & . & . & . & 0.0251 & 0.0253 & 0.0253 \end{pmatrix}.$$

5.1.2. Stationary Distribution. Each entry of our matrix tends to 0.025, now we will calculate the stationary distribution. As we note that the matrix is irreducible and aperiodic, it can be affirmed by the Convergence Theorem that (for the Monopoly chain) the stationary distribution exists. We will do this calculation by taking the first 39 columns of matrix P and we will subtract 1 from the diagonal and we will replace the last column with a column that consists only of ones, in this way we have

$$A = \begin{pmatrix} -1 & 0 & 0 & . & . & . & 0 & 0 & 1 \\ 0 & -1 & 0 & . & . & . & 0 & 0 & 1 \\ 0 & 0 & -1 & . & . & . & 0 & 0 & 1 \\ . & . & . & . & & & . & . & . \\ . & . & . & & . & & . & . & . \\ . & . & . & & & . & . & . & . \\ 0.056 & 0.083 & 0.111 & . & . & . & -1 & 0 & 1 \\ 0.028 & 0.056 & 0.083 & . & . & . & 0 & -1 & 1 \\ 0 & 0.028 & 0.056 & . & . & . & 0 & 0 & 1 \end{pmatrix}.$$

By calculating the inverse matrix with the help of Python, we find

$$A^{-1} = \begin{pmatrix} -1.0035 & 0.0244 & 0.0244 & . & . & . & -0.0036 & -0.0035 & 0.9964 \\ -0.0070 & -1 & 0.0488 & . & . & . & -0.0072 & -0.0071 & 0.9929 \\ -0.0106 & 0.0173 & -0.9547 & . & . & . & -0.0108 & -0.0108 & 0.9893 \\ . & . & . & . & & & . & . & . \\ . & . & . & & . & & . & . & . \\ . & . & . & & & . & . & . & . \\ -0.0488 & -0.0486 & -0.0509 & . & . & . & -0.9928 & 0.0070 & 0.9790 \\ -0.0244 & -0.0244 & -0.0241 & . & . & . & 0.0035 & -0.9964 & 1.0035 \\ 0.0250 & 0.0249 & 0.0250 & . & . & . & 0.0250 & 0.0250 & 0.0249 \end{pmatrix}.$$

If we look at the last row of matrix A^{-1}, then

$$\pi = \begin{pmatrix} 0.0250 & 0.0249 & 0.0250 & . & . & . & 0.0250 & 0.0250 & 0.0249 \end{pmatrix}$$

which is very close to

$$\pi = \begin{pmatrix} \frac{1}{40} & \frac{1}{40} & \frac{1}{40} & . & . & . & \frac{1}{40} & \frac{1}{40} & \frac{1}{40} \end{pmatrix}.$$

5.1.3. Limit distribution. Let us now consider the row vector of size 1×40 where the first entry is a 1, and the rest of the entries are 0, and we will say that it will be the initial distribution, since in this game all the players start in the starting square, then:

$$\pi_0 = \begin{pmatrix} 1 & 0 & 0 & . & . & . & 0 & 0 & 0 \end{pmatrix}.$$

It is interesting to see what happens with the distribution of this Markov chain, if we start with the previous initial distribution, having said that we will calculate the limit distribution for this chain, notice that

$$\pi_1 = \pi_0 P = \begin{pmatrix} 0 & 0 & 0.028 & 0.056 & 0.083 & . & . & . & 0 & 0 \end{pmatrix}.$$

Calculated $\pi_2 = \pi_1 P = \pi_0 P^2$

$$\pi_2 = \pi_0 P^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0.000784 & 0.003136 & 0.007784 & ... & 0 & 0 \end{pmatrix}.$$

Analogously for $n = 4$

$$\pi_4 = \pi_0 P^4 = \begin{pmatrix} 3.6235 \times 10^{-3} & 1.96513 \times 10^{-3} & 1.00042 \times 10^{-3} & . & . & . & 6.15312 \times 10^{-3} \end{pmatrix}.$$

For $n = 8$

$$\pi_8 = \pi_0 P^8 = \begin{pmatrix} 0.003891 & 0.00539 & 0.00737 & 0.00984 & 0.01286 & . & . & . & 0.002042 & 0.002789 \end{pmatrix}.$$

If we continue with the process we can see that

$$\pi_{16} = \pi_0 P^{16} \begin{pmatrix} 0.02939 & 0.02688 & 0.02450 & 0.02198 & 0.01286 & . & . & . & 0.03396 & 0.031732 \end{pmatrix}.$$

For $n = 32$ we can see where the limit distribution is headed

$$\pi_{32} = \pi_0 P^{32} = \begin{pmatrix} 0.02103486 & 0.02062202 & 0.02018032 & 0.02018032 & 0.02013923 & . & . & . & 0.02149927 \end{pmatrix}.$$

A clearer idea of the above is seen in the following calculation

$$\pi_{64} = \pi_0 P^{64} = \begin{pmatrix} 0.0253434 & 0.02538735 & 0.02552678 & 0.0255072 & 0.02555667 & . & . & . & 0.02524536 \end{pmatrix}.$$

With the calculations made above, we can say that

$$\lim_{n\to\infty} \pi_n = \begin{pmatrix} \frac{1}{40} & \frac{1}{40} & \frac{1}{40} & . & . & . & \frac{1}{40} & \frac{1}{40} & \frac{1}{40} \end{pmatrix}.$$

In fact, it is the stationary distribution. It is clear that we would arrive at the stationary distribution, we can affirm this by the convergence theorem.

5.2. Weather Chain (Climate). In this section we will study a Markov chain using EXCEL, this to see how the chain behaves and how it evolves, that is, to see what states it visits. Consider the weather chain studied in the previous section, remember that its state space is $\mathcal{S} = \{1, 2, 3\}$ where, 1 is rainy, 2 is cloudy without rain, and 3 is sunny, we know that this chain has transition matrix P, given by

$$\begin{pmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.6 \\ 0.7 & 0.2 & 0.1 \end{pmatrix}.$$

5.2.1. Transition function for X_n. Suppose that $X_0 = 1$, the objective of this study is to construct the sequence $\{X_n : n \geq 1\}$, or to generate said sequence we have three possibilities:

1. If $X_n = 1$, then

$$\begin{aligned} P\left(X_{n+1} = 1\right) &= 0.2 \\ P\left(X_{n+1} = 2\right) &= 0.5 \\ P(X_{n+1} = 3) &= 0.3 \end{aligned}.$$

2. If $X_n = 2$, then

$$\begin{aligned} P\left(X_{n+1} = 1\right) &= 0.1 \\ P\left(X_{n+1} = 2\right) &= 0.3 \\ P(X_{n+1} = 3) &= 0.6 \end{aligned}.$$

3. If $X_n = 3$, then

$$\begin{aligned} P\left(X_{n+1} = 1\right) &= 0.7 \\ P\left(X_{n+1} = 2\right) &= 0.2 \\ P(X_{n+1} = 3) &= 0.6 \end{aligned}.$$

5.2.2. Distribución de X_n. In EXCEL we can generate random numbers in $[0, 1]$ with the command $U = ALEATORIO()$, in this way we will generate the distribution of the random variable for case 1, as follows [6]

$$X_{n+1} = \begin{cases} 1 & if & U \in [0, 0.2) \\ 2 & if & U \in [0.2, 0.7) \\ 3 & if & U \in [0.7, 1] \end{cases}$$

Similarly the distribution for case 2

$$X_{n+1} = \begin{cases} 1 & if & U \in [0, 0.1) \\ 2 & if & U \in [0.1, 0.4) \\ 3 & if & U \in [0.4, 1] \end{cases} .$$

Analogously, we do it for case 3

$$X_{n+1} = \begin{cases} 1 & if & U \in [0, 0.7) \\ 2 & if & U \in [0.7, 0.9) \\ 3 & if & U \in [0.9, 1] \end{cases} .$$

5.2.3. Study in EXCEL. In the following table we explain the function that each cell will perform in EXCEL to simulate our Markov chain, to be more specific, our model will tell us how our chain evolves over 40 stages, that is, X_n, $n = 0, 1, 2, 3, ..., 40$. That said, let's see how this program works in EXCEL [6].

Q2	1,2,3
B3	=ALEATORIO()
C3	=SI(B3¡0,2,1,-1)
D3	=SI(Y(B3¿0.2,B3¡0.5),2,-1)
E3	=SI(B3¿0.5,3,-1)
F3	=MAX(C3,D3,E3)
G3	=ALEATORIO()
H3	=SI(G3¡0.2,1,-1)
I3	=SI(Y(G3¿0.2,G3¡0.5),2,-1)
J3	=SI(G3¿0.5,3,-1)
K3	=MAX(H3,I3,J3)
L3	=ALEATORIO()
M3	=SI(L3¡0.2,1,-1)
N3	=SI(Y(L3¿0.2,L3¡0.5),2,-1)
O3	=SI(L3¿0.5,3,-1)
P3	=MAX(M3,N3,O3)
Q3	=MAX(SI(Q2=1,F3,-1),SI(Q2=2,K3,-1),SI(Q2=3,P3,-1))

In the table we can see that Q2 is X_0, this implies that the chain can start at 1,2 or 3, from B3 – Q3 it will be the simulation for X_1, we will have to do each of the steps of the table up to $B_i - Q_i$, $i = 3, 4, 5, ..., 42$ which will be the value 1,2 or 3 that the chain takes in the 40 steps. This is how we show how our studio looks. For the case in which $X_0 = 1$ and figure 5.2 shows us a graph of said study, while figure 5.3 shows us the case in which $X_0 = 2$ in the same way that figure 5.4 shows us show what happens when $X_0 = 3$.

Figure 5.2: Study of a Markov chain in EXCEL when $X_0 = 1$. We can see that we will have exactly 18 days with rain, 9 cloudy days without rain, and 14 sunny days.

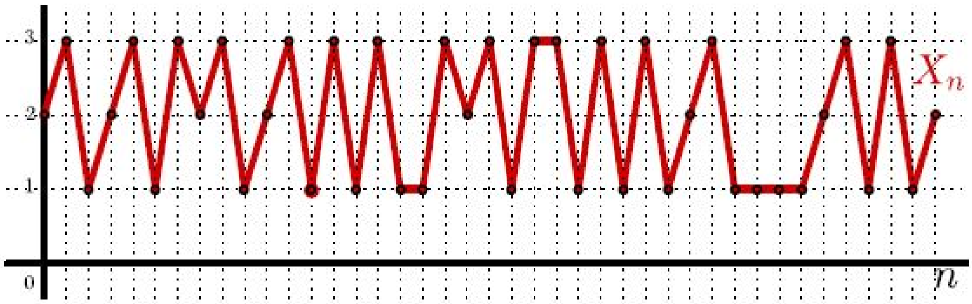

Figure 5.3: Study of a Markov chain in EXCEL when $X_0 = 2$. We can see that we will have exactly 17 days with rain, 8 cloudy days without rain, and 16 sunny days.

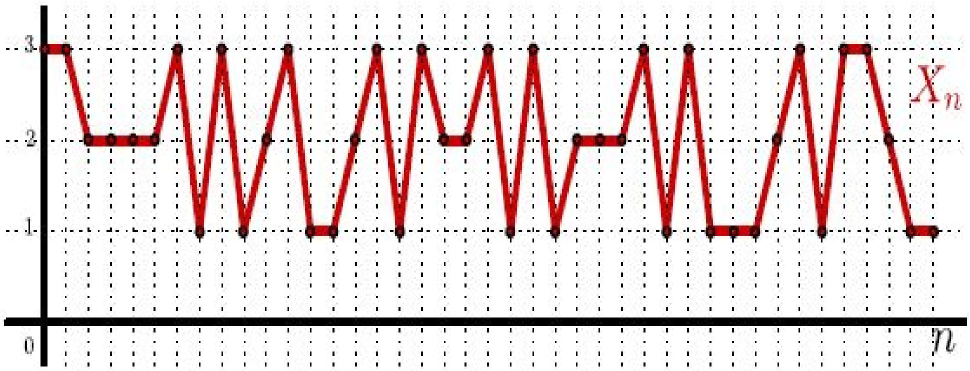

Figure 5.4: Study of a Markov chain in EXCEL when $X_0 = 3$. We can see that we will have exactly 14 days with rain, 13 cloudy days without rain, and 14 sunny days.

6. Conclusions. Throughout this work, basic concepts of probability theory were analyzed, such as conditional probability, the property of total probability and the Bayes Formula. We study homogeneous Markov chains using some concepts and reaffirming, the most important part of this work consisted in analyzing discrete-time Markov chains and continuous-time chains related to Markov processes. Then we prove the Convergence Theorem and for this we study important concepts such as: the transition function, the limit distribution, the reducible and irreducible chains, finally we saw the stationary distribution and the periodicity. We also study chains in continuous time such as the Poisson process. With the help of Python we calculate the limit distribution and the stationary distribution of the Monopoly game chain, since the chain has a limit distribution and a stationary distribution, our chain is irreducible. In addition, with the help of Excel we studied a chain of time (climate), this allowed us to see the behavior of the chain throughout 40 stages. This we think about the endless applications that Markov chains have, as well as the importance that they give us by making a more detailed analysis.

Later we can continue studying the following lines of research related to Markov chains, particularly in queuing theory, Markov chains in continuous time a little more complex. While in the part of discrete time chains, processes such as martingales, birth and death chains, absorbing Markov chains could be studied. Furthermore, we can go into the study of Markov decision processes, or a bit of stochastic calculus, manufacturing and re-manufacturing systems (which are widely related to queuing theory and continuous-time processes), Markov chains of Monte Carlo, in short a number of theories in which this work could be continued.

Markov chains are useful in certain branches of Physics such as Thermodynamics, Quantum Mechanics and gauge field theories [21], in Meteorology it helps to have more accurate predictions in the change of time from one day to another, in Biological Sciences Epidemiological models are explained, in Game theory, Finance, or Social Sciences, Statistics and Mathematics.

Acknowledgements. The autor would like the IF-UNAM for being an associate doctoral student, while this work has been prepared.

ORCID and License

Mississippi Valenzuela https://orcid.org/0000-0002-8532-229X

References

[1] Artalejo JR, Gómez-Corral A. Generalized birth and death processes with applications to queues with repeated attempts and negative arrivals. OR Spektrum. 1998; 20:5-14.
[2] Artalejo JR, Phung-Duc T. Markovian retrial queues with two way communication. Journal of Industrial and Management Optimization. 2012; 8:781-806.
[3] Artalejo JR, Gómez-Corral A, He QM. Markovian arrivals in stochastic modeling: A survey and some new results. Statistics and Operations Research Transactions. 2010; 34:101-144.
[4] Caballero ME, Rivero VM. Cadenas de Markov, Un enfoque elemental. Sociedad Matemática Mexicana. 2004.
[5] Castellanos-Rodríguez V. Forma de la Distribución cuasiestacionaria de una cadena de Markov bivariada: Aplicación a un modelo genético. M. Sc. thesis. Centro de Investigación en Matemáticas. September 2006.
[6] Ching W, Huang X. Markov Chains: Models, Algorithms and Applications. Springer. 2013.
[7] Devore JL. Probabilidad y estadística para ingenieria y ciencias. Seventh edition. Cengage Learning. 2008.
[8] Durrett R. Elementary Probability for Applications. Cambrige. 2009.
[9] Häggströng O. Finite Markov Chains and Algorithmic Applications. Cambrige University Press, 2003.
[10] Isaac R, The pleasures of Probability. Undergraduate Texts in Mathematics. Springer. USA. 1995.
[11] James F. Statistical Methods in Experimental Physics. Second edition. World Scientific. Singapur. 2006.
[12] Karlin S, Mcgregor J. The classification of birth and death processes. Transactions of the American Mathematical Society. 1957; 86(2):366-400.
[13] Kolmogorov AN, Grundbegriffe der Wahrscheinlichkeitsrechnung. Springer. 1933; Fundations of the Theory of Probability. second edition. Chelsea. 1956.
[14] Maibaum G. Teoría de probabilidades y estadística matemática. Pueblo y educación. 1976.
[15] Márkov AA. Rasprostranenie zakona bol'shih chiselna velichiny, zavisyaschie drug ot druga. Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete. 1906; 2(15):135-156.
[16] Markov AA. Extension of the limit theorems of probability theory to a sum of variables connected in a chain. Reprinted in Appendix B of: R. Howard. Dynamic Probabilistic Systems, volume 1: Markov Chains. John Wiley and Sons, 1971.
[17] Rabiner LR. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Procedings of the IEEE. 1989; 77(2):257-286.
[18] Rincon L. Introducción a los Procesos Estocásticos. Departamento de Matemáticas. UNAM. 2021.
[19] Salinas Santamaría L. Procesos de nacimiento y muerte y sus aplicaciones a colas[B. Sc. thesis]. Escuela Superior de Física y Matemáticas. Instituto Politécnico Nacional. June 2009.
[20] Si Tung Ho L, Xu J, Crawford FW, Minin VN, Suchard MA. Birth/birth-death processes and their computable transition probabilities with biological applications. August 2017. ArXiv:1603.03819 [stat.CO].
[21] Valenzuela M, Pruebas Estadísticas en la Física del bosón de Higgs[Tesis de Maestría], Matemáticas Aplicadas, Universidad Autónoma de Zacatecas, Mexico, february 2018. DOI: 10.13140/RG.2.2.23007.00167
[22] Wiener N. Differential-Space. Studies in applied mathematics, (1923); 2(4):131-174.
[23] Wrenn-Crawford F. General birth-death processes:probabilities, inference, and applications[Ph. D. thesis]. University of California; Los Angeles. 2012.