Advances and Applications in Mathematical Sciences
Volume 21, Issue 8, June 2022, Pages 4605-4613

APPLICATIONS OF MARKOV CHAIN IN THE GAMES

D. SARGUNASUNDARI

Assistant Professor
Department of Mathematics
Annai Vailankanni Arts and Science College
Thanjavur, India
Email: sargunamaths@gmail.com

Abstract

As a result of random events, many things have happened in our lives. A Markov Chain is just a "time-varying sequence of random variables". The Markov chain, which is a sort of 'stochastic process', is the topic of this study. Markov chains are also a set of fields that have created is applied to govern random processes. Markov Chains are used to create sequences of random variables connected by a simple dependency relationship. Transitions between states take place in the Markov Chain's underlying system. It's frequently used to make a stochastic process's future state more understandable. Markov Chains are utilized in "physics, biology, statistics, and the social sciences", among other subjects. The goal of this paper is to introduce you to "Absorbing and Ergodic Markov Chains", are the applications of Markov chains in games.

1. Introduction

The Markov Chain is a system that performs state transitions and is thought to be independent of the past depending on the current [5]. The Markov Chain is based on the current state events rather than the order in which they occur. The Markov property [6] is a form of "memoryless" feature of the past. Transitions are defined as changes that occur within the states of the system, as well as transition probabilities connected with those changes.

A state space, entries of a transition matrix expressing 'transition probabilities between states', and a starting distribution throughout the state space make up a Markov chain. In a random process, a 'state space' is a collection of all the values that occur. Furthermore, states are the basics

2020 Mathematics Subject Classification: 60J27.

Keywords: Markov chains, Transition probabilities, Absorbing and Ergodic Markov chains.

Received March 24, 2022; Accepted April 11, 2022

elements in a state space, and they are a key component in the creation of Markov Chain models. These three parts, as well as the Markov property, can be used to create a Markov chain to depict the evolution of a random process over time.

Baseball will be the application on which we will concentrate in this paper, as well as the various game aspects that can be evaluated using Markov chains. We'll look at two forms of Markov chains utilized in 'Markov chain models' in this section: "absorbing Markov chains and ergodic Markov chains". We'll look at how a Markov chain can be used to describe baseball, as well as some of the game's techniques. We want to figure out how many runs there are using Markov chains.

2. Methodology

Definition 2.1. If $\{X_1, X_2, \ldots,\}$ is a Markov process with state space S, then

$$P(X_{n+1} = s_{n+1} \mid X_n = s_n, \ldots, X_2 = s_2, X_1 = s_1)$$
$$= P(X_{n+1} = s_{n+1} \mid X_n = s_n) \tag{1.1}$$

holds for any $n = 1, 2, \ldots$ and any $s_1, s_2, \ldots, s_{n+1}$ with $s_k \in S$ for $1 \leq k \leq n+1$.

Definition 2.2. "A Markov chain is a Markov process that has a finite number of states with the Markov property" (1.1). The main components of stochastic processes are $S = \{s_1, s_2, \ldots, s_r\}$ [3], [1]. The term "state" refers to all of the possible values for each X_n.

Definition 2.3. If A and B represent the two events of state space S, then $P(B)$ is not equal to 0, the Conditional Chance is if B is known to happen or has already happened, a given B is the chance that A will occur.

The formula is $P(A/B) = \dfrac{P(A \cap B)}{P(B)}$ [8].

If 'A and B are independent events', then $P(A/B) = P(A)$.

Definition 2.4. Let $S = \{1, 2, \ldots, r\}$ be a Markov chain with a 'finite state space' in discrete time, $\{X_k\}$. Consider the P_{jk} transition probabilities, where $j = 1, 2, \ldots, r$ and $k = 1, 2, \ldots, r$. As a result, r^2 transition probabilities exist [9].

These values are entered into a transition matrix P, to make grouping the transition probabilities easier. The probability of travelling from the j^{th} state to the k^{th} state of S is represented by the matrix P, where the j^{th} row and k^{th} column of P reflect the probabilities of travelling from the j^{th} state to the k^{th} state of S".

$$P = \begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,r} \\ P_{2,1} & P_{2,2} & \cdots & P_{2,r} \\ P_{3,1} & P_{3,2} & \cdots & P_{3,r} \\ \vdots & \vdots & \vdots & \vdots \\ P_{r,1} & P_{r,2} & \cdots & P_{r,r} \end{bmatrix}$$

Definition 2.5. Absorbing Markov Chain. If it is not possible to leave a state S_j of a Markov chain (i.e., $P = jj = 1$), it is called Absorbing. If a Markov chain has at least one absorbing state and it is possible to travel from any state to an absorbing state, although not necessarily in one step, it is said to be absorbing [7].

Definition 2.6. "A state that is not absorbing in an absorbing Markov chain" is known as Transient [2].

Assume that the state space of an 'absorbing Markov chain' is $S = \{s_1, s_2, \ldots, s_r\}$. If there are m transient states, then the 'set of transient states' is $\{t_1, t_2, \ldots, t_m\}$, with each transient state labelled as $t_j \in S$. The set of absorbing states is therefore $\{a_1, a_2, \ldots, a_{r-m}\}$. by naming each absorbing state as $a_j \in S$. As a result, $\{t_1, t_2, \ldots, a_m\} \cup \{a_1, a_2, \ldots, a_{r-m}\} = S$.

Definition 2.7. Let P be an "absorbing Markov chain, then the Fundamental matrix for P is the matrix $N = (I - Q)^{-1}$ where Q consists only of transition probabilities between transient states of P".

Definitions 2.8. An Ergodic Markov chain is one in which it is 'possible to reach every state from any other state,' but not necessarily in a single step.

If all of the elements in a Markov chain's transition matrix are bigger than 0, the chain is said to be regular. In other words, for some natural number n, a Markov chain is regular if it can "start at any state and reach any state in exactly n steps". A typical Markov chain is clearly ergodic.

Canonical Form

Assuming we have an arbitrary absorbing Markov chain, here we have to "renumber the states so that the transient states are listed first and the absorbing states are listed last".

Let $\{X_n\}$ n be a finite-state Markov chain with states $0, 1, 2, 3, \ldots, N$. Assume that states $0, 1, 2, 3, \ldots, r-1$ are transitory in the sense that $P_{jk}^{(n)} \to 0$ as $n \to 0$ for $0 \leq j, k < r$, whereas states $r, \ldots, N$ are absorbing $(P_{ij} = 1)$ for $r \leq j \leq N$.

The transition matrix has the form

$$P = \begin{Vmatrix} Q & R \\ 0 & I \end{Vmatrix} \tag{1.2}$$

where 0 is an $(N - r + 1) \times r$ matrix all of whose entries are zero, I is an $(N - r + 1) \times (N - r + 1)$ identity matrix and $Q_{jk} = P_{jk}$ for $0 \leq j, k < r$. We previously demonstrated that the entry $P_{jk}^{(n)}$ of the transition matrix P^n gives the probability of reaching state s_k after n steps, given that the chain begins in state s_j.

From matrix algebra, we find that P^n is of the form $P^n = \begin{Vmatrix} Q^n & * \\ 0 & I \end{Vmatrix}$ where the symbol * stands for a $r \times (N - r + 1)$ Components from both Q and R are represented in this matrix. The "sub-matrix Q^n gives the odds of transitioning from one transitory state to another after n steps, which we may learn from this matrix". The theory provided in this work will benefit from this matrix.

3. Theoretical Framework

Setting Up the Model

A baseball team on offence can be in one of 20 states at any time during an inning [4]. There are eight distinct ways to distribute runners to the bases in a baseball game. They're all empty bases with "one man on first, second, or third base, men on first and second, men on first and third, men on second and third, men on second and third, and men on first, second, and third". At any moment throughout an inning, there could be zero, one, or two outs. We get a total of 20 states when the number of outs and the number of occupied bases is combined together. We also have to factor in the half-inning concluding with three outs. For the purposes of this analysis, we'll refer to the 21st state as two outs and no runners on base at the end of the inning. The 22nd, 23rd, and 24th states, respectively, will signal inning endings with one, two, or three runners on base [11].

The process must satisfy the Markov property (see equation 1.1), "the number of states must be finite, and the transition matrix must contain positive entries and row sums of at least one to form a Markov chain. For one inning of baseball, we can simply create a Markov chain that fits these criteria. The transitory states with zero, one, or two outs, as well as the absorbing states with three outs, are the states of a Markov chain's Absorbing Chain. A list of every temporary condition with its consistent number of bases occupied and outs within an inning. As before, we'll put the transition matrix into canonical form by listing the transitory states first, followed by the absorption states [10]. As a result, the first 20 states have fewer than three outs, while the final four are absorbing states, each with a different number of runners on base when the third out is made".

"Q is a 20×20 matrix, R is a 20×4 nonzero matrix, 0 is a 4×20 zero matrix, and I_4 is a 4×4 identity matrices in canonical form. We can start entering the transition probabilities for the baseball Markov chain into the matrix once we have the transition matrix in canonical form".

The initial purpose of this Markov chain study is to" determine the number of runs the baseball team should have scored that year. We'll achieve this with the use of statistics, and we'll need to collect data on baseball events that occurred over the season in order to calculate the chances of shifting from one state to another utilizing these occurrences".

The first assumption we make when developing this model is that after each transition, the current batter will be replaced by the next batter in the lineup. That is, "either the current batter advances to second base safely or is struck out. As a result, plays like stealing a base or advancing on a passed ball aren't counted".

Furthermore, triple plays, sacrifice flies, errors, plays in which a runner advances from first to third base on a single, and other rare and infrequent baseball game plays are not included. We make these assumptions to simplify the model, but we don't have access to the majority of the probability of these events occurring. We can assume that these events do not occur during the model's development because they are uncommon during a baseball season, and we will still have a very accurate run output model. The following list describes additional assumptions made for each baseball event used in the model. [11]:

"**Single.** A first-base runner moves to second base, other base runners score, and the batter moves to first base".

"**Walk or hit.** Runners advance one base only if forced to do so (for example, a runner on first base will advance to second base, whereas a runner on second base will not advance), and the batter advances to first base".

"A first-base runner advances to third base, other runners score, and the batter moves to second base".

"All runners score in a triple, and the hitter advances to third base".

"All runners and batters score on a home run".

Out. The number of outs increases by one since the runners do not advance.

"**Double Play.** This is only possible if there is a force out at second base. If there are more than one base runner and a force at second base, only the batter and the runner on first base are out, and the other runners advance one base or score from third if no three outs are recorded".

4. Results

Calculating Expected Number of Runs

At the end of any inning of a baseball game, "every batter who comes to the plate either makes an out, scores a run, or is left on base" [12].

Let C denote the number of batters faced in an inning, S the number of runs scored, and E the number of runners left on base at the end of the inning.

Then, $C = 3 + S + E$.

This can be equivalently written as $S = C - E - 3$.

When we take 'the expected value of this equation', we get $\text{EXP(S)} = \text{EXP(C)} - \text{EXP(E)} - 3$.

As a result, "the expected number of runs for an inning can be calculated" by multiplying the expected number of batters in the inning by the expected number of runners left on base. There aren't any three outs.

An absorbing Markov chain's fundamental matrix N is a 20×20 matrix with entries n_{jk} denoting that "the number of times the chain has been in state's k since it began in state s_j." We also know that the sum of the j^{th} row's elements of the matrix N equals the expected number of steps from a transient state s_j. Until absorption. In baseball, this signifies that the sum of the j^{th} row of N is the projected number of batters who will come to bat for the rest of the inning, starting from state s_j. We need the sum of the first-row value in the fundamental matrix N since we're seeking for the predicted number of batters to come to the plate starting at the start of an inning. As a result, EXP(C) equals the total of N's first row.

"The fundamental matrix N of an absorbing Markov chain is a 20×20 matrix with entries n_{jk} denoting the number of times the chain has been in state s_k since it began in state s_j. We also know that the sum of the elements of the matrix N in the j^{th} row equals the expected number of steps from a transient state s_j. Until absorption occurs. In baseball, this means that the sum of the j^{th} row of N is the projected number of batters who will come to bat for the remainder of the inning, beginning with state s_j. The probabilities of the chain being absorbed with zero, one, two, or three runners left on base are provided by s_j the j^{th} row of the matrix B. As a result, "beginning at the top of the inning, the first row of B will be used to calculate the estimated number of runners still on base".

The equation is

$$EXP(E) = 0 * P(0\,left) + 1 * P(1\,left) + 2 * P(2\,left) + 3 * P(3\,left)$$

Where $P(0\,left)$ is the chance of concluding that "the inning with no runners on base, $P(1\,left)$ is the likelihood of ending the inning with one runner on base", and so on. As a result, before we can determine $EXP(S)$ we must first find $EXP(C)$ and $EXP(E)$.

5. Conclusion

We studied key aspects of Markov chain theory throughout this paper and discussed both absorption and ergodic Markov chains, both of which have numerous real-world applications. "Absorbing Markov chains" were found to have a significant impact on baseball development. Markov chain theory, "the fundamental matrix" was used to "calculate the expected number of batters in a half-inning of play".

References

[1] Baseball Statistics and Analysis, Fan Graphs Baseball. http://www.fangraphs.com

[2] Jyoti prasad Medhi, Stochastic processes, third edition, 2009.

[3] Frank E. Beichelt and L. Paul Fatti. Stochastic processes and their applications, Taylor and Francis Ltd., London, 2002, Translated and revised from the 1997 German original.

[4] Bruce Bukiet, Elliotte Rusty Harold and Jose Luis Palacios, A Markov Chain Approach to Baseball, Operations Research 45(1) (1997), 14-23.

[5] Richard M. Felman and Ciriaco Valdez-Flores, Applied Probability and Stochastic Processes, chapter Markov Chains, page 141, Springer Berlin Heidelberg, 2010.

[6] Charles M. Grinstead and J. Laurie Snell, Introduction to probability, American Mathematical Society, Providence, RI, 2nd rev. edition, 1997.

[7] Dean L. Isaacson and Richard W. Madsen, Markov chains. John Wiley and Sons, New York-London-Sydney, Theory and Applications, Wiley Series in Probability and Mathematical Statistics, 1976.

[8] Lawrence M. Leemis, Probability, s. n., S.L., 2011.

[9] Andrei Andreyevich Markov, Markov Biography.http://www-history.mcs.st-and.ac.uk/Biographies/Markov.html.

[10] Howard M. Taylor and Samuel Karlin, An introduction to Stochastic Modeling Third Edition.

[11] Mark D. Pankin, Markov chain models, Theoretical background, 2007.

[12] Thomas Polaski, A markov chain model for run production in baseball, Winthrop University, February 2013.

[13] Tom M. Tango, Mitchel G. Lichtman and Andrew E. Dolphin, The Book: Playing the Percentages in Baseball, Potomac Books, Inc., 2007.