Cihan University-Erbil Scientific Journal (CUESJ)

RESEARCH ARTICLE

Analyzing the Football Sports Outcomes by the Two Types Branching Markov Process: Real Madrid Football Club as a Case Study

Bestoon M. Abdullkarim[1], **Azhin M. Khudhur**[1], **Dler H. Kadir** [1'2]

[1]*Department of Statistics and Informatics, College of Administration and Economics, Salahaddin University-Erbil, Iraq*

[2]*Department of Business Administration, Cihan University-Erbil, Kurdistan Region, Iraq*

ABSTRACT

In this research the football sport outcomes are analyzed by the two-type branching Markov process, assuming that the outcomes occur as a Poisson process and are described by the two-state Markov process. For this study, the number of losses, wins, and trophy results of the Real Madrid football club for 24 years (2000–2024) are obtained from the KOOORA Website. This study aimed to determine the probabilities of two state outcomes, the time the process stays in each state and the expected number of the two types for successive periods. In conclusion, the expected number of Real Madrid football sports outcomes for successive game times are continually decreasing until the number of games.

Keywords: Two-state Markov process, two-type branching process, probability generating function, trophy, win and lose

INTRODUCTION

The stochastic branching process played a great role in various scientific applications and sports activities. In this work, the two types of branching Markov process have been applied to football sport outcomes. This study is in three sections, the first one includes the concept of the Markov process, the transition probabilities of two states, the expected time interval the process stays in each state, and the two types of branching process with their probability generating function. The second section specified for application, involves describing and analyzing the data about the game outcomes of Real Madrid football club during (2000–2024). The final section dedicated to conclusions, indicates the deduces which are found from the results of the application.

METHODOLOGY

In this study, the Markov process, the transition probabilities of two states, two types of branching process, and the probability generating function for two types.

Markov Processes

A stochastic process with continuous time parameter $\{X(t); t > 0\}$ is called the Markov process if for any set of periods $t_1 < t_2, <\ldots, <t_n$ if the conditional probability distribution of $X(t_n)$ given the values of $\{X(t_0), X(t_1), \ldots, X(t_{n-1})\}$ depends on $X(t_{n-1})$ only, as defined.[1,2]

$$P_r[X(t_n) \le X_n | X(t_0) = x_0, X(t_1) = x_1, \ldots \;, X(t_{n-1}) = x_{n-1}]$$

$$= P_r[X(t_n) \le X_n | X(t_{n-1}) = X_{n-1}]$$

Hence, the two-state Markov process {0, 1}, in which the process changes its state from 0 to 1 with the rate (λ) and from 1 to 0 with the rate (μ). The transition probabilities for the state Markov process $P_{ij}(t)(i, j = 0,1)$ are defined as[3,4]

Both rates (μ, λ) represent the waiting time between transitions in exponential distribution. The waiting time in state 0 is referred to as μ while the λ rate follows an exponential distribution with state 1. These rates are expressed as memoryless properties.

$$P_{ij}(t) = P_r[X(t) = j | X(0) = i \;] \tag{1}$$

The probabilities are given by

$$P_{00}(t) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t}$$

Corresponding Author:
Azhin M. Khudhur, Department of Statistics and Informatics, College of Administration and Economics, Salahaddin University-Erbil, Iraq Region, Iraq. E-mail: azhin.khudhur@su.edu.krd

Received: November 20, 2024
Accepted: December 26, 2024
Published: January 20, 2025

DOI: 10.24086/cuesj.v9n1y2025.pp8-13

$$P_{01}\left(t\right)=\frac{\mu}{\lambda+\mu}-\frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t}$$

$$P_{10}\left(t\right)=\frac{\mu}{\lambda+\mu}-\frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t}$$

$$P_{11}\left(t\right)=\frac{\mu}{\lambda+\mu}+\frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t}$$

Farther, as t → ∞

$$\pi_0=\lim_{t\to\infty}P_{00}\left(t\right)=\lim_{t\to\infty}P_{10}\left(t\right)=\frac{\mu}{\lambda+\mu} \quad (2)$$

$$\pi_1=\lim_{t\to\infty}P_{01}\left(t\right)=\lim_{t\to\infty}P_{11}\left(t\right)=\frac{\mu}{\lambda+\mu}$$

The Expected Length of Time in Two-State Markov Processes

The expected length of time μ_{ij} (t) in (0, t) that the process spends in state (j), having initially started from the state (i = 0,1), is given by[1,5]

$$\mu_{00}\left(t\right)=\frac{\mu t}{\lambda+\mu}+\frac{\lambda}{(\lambda+\mu)^2}[1-e^{-(\lambda+\mu)t}] \quad (3)$$

$$\mu_{01}\left(t\right)=\frac{\lambda t}{\lambda+\mu}-\frac{\lambda}{(\lambda+\mu)^2}[1-e^{-(\lambda+\mu)t}]$$

$$\mu_{10}\left(t\right)=\frac{\mu t}{\lambda+\mu}-\frac{\mu}{(\lambda+\mu)^2}[1-e^{-(\lambda+\mu)t}]$$

$$\mu_{11}\left(t\right)=\frac{\lambda t}{\lambda+\mu}+\frac{\mu}{(\lambda+\mu)^2}[1-e^{-(\lambda+\mu)t}]$$

Furthermore, limiting the fraction of time (t) the process spent in any one of the states (0 or 1) during each transition is given by

$$\pi_0=\lim_{t\to\infty}\frac{\mu_{00}(t)}{t}=\lim_{t\to\infty}\frac{\mu_{10}(t)}{t}=\frac{\mu}{\lambda+\mu} \quad (4)$$

$$\pi_1=\lim_{t\to\infty}\frac{\mu_{01}(t)}{t}=\lim_{t\to\infty}\frac{\mu_{11}(t)}{t}=\frac{\lambda}{\lambda+\mu}$$

Thus, the time spent by the process in each state during every visit is an exponential distribution.

$f_0(t) = \lambda e{-}\lambda t \qquad t > 0$

$f_1(t) = \mu e{-}\mu t \qquad t > 0$

Where the win and loss (λ and μ) rates are estimated by the maximum likelihood estimate (MLE), and they equal $1/\bar{t}$. In addition, the average time ($\bar{t}$) represents the average of spending in a specific state during each transition to that state in a two-state Markov process such as:

$$\bar{t}_0=E\left(T_0\right)=\frac{1}{\lambda},\quad \bar{t}_1=E\left(T_1\right)=\frac{1}{\mu}$$

They related the process parameters λ and μ to the directly observable characteristics of the process and the expected lengths of visits to each state {0,1}.

$$E\left(0\right)=\frac{1}{\lambda}\text{ and }E\left(1\right)=\frac{1}{\mu}$$

Thus, crude estimates of the process parameters are obtained.

$$\lambda=\frac{1}{E(0)},\quad \mu=\frac{1}{E(1)}$$

Two Type Branching Processes

Considering a branching Galton–Watson process where two different types may be distinguished, either type will produce or possibly both types independently.

Let U_n and V_n be a random variable representing type (1) and type (2), respectively, in the nth period,[6,7] which can be written as

$$U_{n+1}=\sum_{j=1}^{U_n}X_j^1+\sum_{j=1}^{V_n}X_j^2$$

$$V_{n+1}=\sum_{j=1}^{U_n}Y_j^1+\sum_{j=1}^{V_n}Y_j^2$$

When the branching process is one of the applications of Markov sequent then the transition probability law of the Markov process is:

$P\ (X_j^i = K,\ Y_j^i = L) = P_i\ (k, l)$

k, l = 0,1,2, 3… for j = 1,2,3,… and i = 1,2

Where X_j^i, Y_j^i are independent identically distributed (iid) random vectors with probability mass functions P_i (k, l) as above.

The simplest situation assumes the process begins with a single type then we define for initial conditions.

$U_0 = 1 \quad$ *and* $\quad V_0 = 0$

Or

$U_0 = 0 \quad$ *and* $\quad V_0 = 1$

Moreover, the probability extinction of the two types (Π^1, Π^2) is determined as follows:

For type (1)

$$\neq^1=P[U_n=0,V_n=0|U_n=1,V_n=0]$$

And for type (2)

$$\neq^2=P\ \ [U_n=0,V_n=0\,|\,U_0=0,V_n=1]$$

Probability Generating Function Relations for Two Types of Branching Process

In this section, we show some relation for the two-dimensional probability generating function of (U_n) and (V_n) of the number of types of n^{th} periods,[8,9] as

$$\varphi_n^{(i)}\left(q_1,q_2\right)=E_n^i\left[q_1^k,q_1^l\right]=\sum_{k,l=0}^{\infty}P_i(k,l)q_1^k q_2^l \tag{5}$$

$0 \le q_1,\ q_2 \le 1$ *and* $i = 1,2$

Where:

1. q_1 is the auxiliary variable representing the weight assigned to the number of type (1) individuals (U_n)
2. Moreover, q_2 is the auxiliary variable representing the weight assigned to the number of type (2) individuals (V_n).

Moreover, their multiple-step generalizations

$$\varphi_n^{(1)}\left(q_1,q_2\right)=\sum_{k,l=0}^{\infty}p(U_n=k,V_n=l\,|\,U_0=1,V_0=0)q_1^k q_2^l \tag{6}$$

$$\varphi_n^{(2)}\left(q_1,q_2\right)=\sum_{k,l=0}^{\infty}p(U_n=k,V_n=l\,|\,U_0=0,V_0=1)q_1^k q_2^l$$

With $n \ge 0$, it is clear that the generation function for type (1) is:

$$\varphi_0^{(1)}\left(q_1,q_2\right)=q_1$$

$$\varphi_1^{(1)}\left(q_1,q_2\right)=\varphi^{1}(q_1,(q_2)$$

Moreover, for type (2) is

$$\varphi_0^{(2)}\left(q_1,q_2\right)=q_2$$

$$\varphi_1^{(2)}\left(q_1,q_2\right)=\varphi^{2}(q_1,q_2)$$

In general, the method can be used as

$$\varphi_{n+m}^{(i)}\left(q_1,q_2\right)=\varphi_m^i(\varphi_n^1\left(q_1,q_2\right),\varphi_n^2\left(q_1,q_2\right))$$

The Expected Numbers of the Two Types of Branching Process

The method of generating function is the most important tool in the study of branching processes, it is useful to determine the mean matrix of the n^{th} period.

When (U_n) and (V_n) the number of type 1 and 2 of the n^{th} period,[9,10] then the expected numbers are given by

$$\frac{\partial\varphi_n^{(1)}\left(q_1,q_2\right)}{\partial q_1}\Big|_{q_1=q_2=1}=E[U_n|U_0=1,V_0=0]=m_{11}^{(n)}$$

$$\frac{\partial\ \ ^{(2)}\left(q_1\ \ q_2\right)}{\partial}\Big|_{q_1=q_2=}=E[U\ \ |U_0=1,V_0=0]=m_{12}^{(\ )} \tag{7}$$

$$\frac{\partial\varphi_n^{(1)}\left(q_1,q_2\right)}{\partial q_1}\Big|_{q_1=q_2=1}=E[V_n|U_0=0,V_0=1]=m_{21}^{(n)}$$

$$\frac{\partial\varphi_n^{(2)}\left(q_1,q_2\right)}{\partial q_2}\Big|_{q_1=q_2=1}=E[V_n|U_0=0,V_0=1]=m_{22}^{(n)}sss$$

Then the expected matrix for the n^{th} period for two types of branching process is defined as,

$$M=\begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \tag{8}$$

So

$$M^{(n)}=\begin{bmatrix} m_{11}^{(n)} & m_{12}^{(n)} \\ m_{21}^{(n)} & m_{22}^{(n)} \end{bmatrix}$$

$$m_{ij}=\frac{\varphi_{(1,1)}^{1}}{\partial q_j} \quad j=1,2$$

$$=\lambda_{1j}$$

And

$$m_{ij}=\frac{\varphi_{(1,1)}^{2}}{\partial q_j} \qquad j=1,2$$

$$=\propto_{2j}$$

Here

$$M=\begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \mu_{21} & \mu_{22} \end{bmatrix}$$

According to the distribution of the two types (1) and (2) the probability generation function is given as

$$\varphi\left(1\right)\left(q_1,q_2\right)=\varphi\left(q_{11}\right).\varphi\left(q_{12}\right)$$

$$=e^{\lambda_{11}(q_1-1)+\lambda_{12}(q_2-1)}$$

$$\varphi\left(2\right)\left(q_1,q_2\right)=\varphi\left(q_{21}\right).\varphi(q_{22})$$

$$=e^{\mu_{21}(q_1-1)+\mu_{22}(q_2-1)}$$

In addition, the expected number of the two types (1) and (2) for successive periods is

$$G_n=\left(U_n\,V_n\right) \tag{9}$$

Then

$$E(G_{n+s}|G_n)=G_n M^s$$

Application

This section is specified for analyzing the football sports outcomes of Real Madrid football club by the two-type branching Markov process, where the football outcomes (loss and win) are denoted by the two states (0 and 1), respectively. Thus, the transition probabilities of the two-state Markov process and the expected length of time that the process spends in each state are defined, also the matrix of the expected numbers to four possible transitions of the two states (0,1) is prepared.

Data Description

For this study, the data about the football outcomes of the Real Madrid team, which has played annually with 19 other Spanish teams in periodic competitions for 24 years have been taken from the KOOORA website.

Table 1 shows the total outcome of 748 games played between 2000 and 2024, which included 142 losses with a probability of 0.2, and 606 wins with a probability of 0.8. Among these, 14 losses with a probability of 0.58 and 10 wins with a probability of 0.42 were associated with trophy outcomes.

Estimation of the Loss and Win Rates

The loss and win rates of the football outcomes are estimated concerning the trophy results, hence the loss rate ($\hat{\lambda}$)is estimated from the total number of losses, which is 101 from 448 games of 14 losses trophies in the average of 32 games annually

$$\hat{\lambda} = \frac{101}{448} = 0.225 \quad / game\,time$$

$$0.225 \times 32 = 7.2 \quad games \quad / annum$$

Furthermore, the win rate ($\hat{\mu}$) is estimated from the total number of wins, which is 259 from 300 games of 10 wins trophies in an average of 30 games annually

$$\hat{\mu} = \frac{259}{300} = 0.863 \quad / game\,time$$

$$0.863 \times 30 = 25.9\,games \quad / annum$$

The Transition Probabilities of the Two-state Football Sport Outcomes

The transition probabilities of the two-state football sport $P_{ij}(i, j = 0,1)$ for successive game times (t_0) are referred to in (1).

Table 2 shows the transition probabilities between the two state football sports outcomes (loss and win) for successive game times (t = 1,2,…,10).

Hence, the stationary distribution of the loose (π_0) and win (π_1) is defined in (2).

$$\pi_0 = \frac{0.863}{0.225 + 0.863} = 0.79$$

$$\pi_1 = \frac{0.225}{0.225 + 0.863} = 0.21$$

The Expected Time the Football Sport (Process) Stays in Each Outcome (State)

The expected length of time $\mu_{ij}(t) \;\; (i, j = 0,1)$ the football sport spent in any one of the states (0,1) for game time (t) is defined in (3).

Table 3 shows the expected length of time the process spends in any football state during each transition for given game times (t = 5,10,15,20,25,30,35).

The Expected Numbers of the Two-state Football Sport Outcomes

The four expected numbers of the two-state (0,1) football sports m_{00}, m_{01}, m_{10}, m_{11} are computed from the transition counts of the number of loss and win games according to the loss (state 0) and win (state 1) trophy of Real Madrid which has got during (2000–2024), as defined:

$$Trophy : 101000110001000010011101$$

Table 1: The number of loss and win outcomes with the trophy results

Years	Number of losses	Number of wins	Total	Trophy
2000–2001	6	24	30	Win
2001–2002	10	19	29	Loss
2002–2003	4	22	26	Win
2003–2004	10	21	31	Loss
2004–2005	9	24	33	Loss
2005–2006	8	20	28	Loss
2006–2007	8	23	31	Win
2007–2008	7	27	34	Win
2008–2009	10	25	35	Loss
2009–2010	4	31	35	Loss
2010–2011	4	29	33	Loss
2011–2012	2	32	34	Win
2012–2013	5	26	31	Loss
2013–2014	5	27	32	Loss
2014–2015	6	30	36	Loss
2015–2016	4	28	32	Loss
2016–2017	3	29	32	Win
2017–2018	6	22	28	Loss
2018–2019	12	21	33	Loss
2019–2020	3	26	29	Win
2020–2021	3	21	24	Win
2021–2022	4	26	30	Win
2022–2023	8	24	32	Loss
2023–2024	1	29	30	Win

The transitions are as noted in Table 4.

Table 4 shows the transition counts of the losses and wins of the Real Madrid football club.

Then the four expected numbers m_{ij} (i,j = 0,1) of the football sports outcomes are represented in the following matrix, defined in (8).

$$M_{ij} = \begin{pmatrix} m_{00} & m_{01} \\ m_{10} & m_{11} \end{pmatrix}$$

$$\propto_{ij} = \frac{m_{ij}}{\sum_i m_{ij}} \qquad i, j = 0,1$$

$$M_{ij} = \begin{pmatrix} 0.58 & 0.69 \\ 0.42 & 0.31 \end{pmatrix}$$

The Expected Number of Football Sport Outcomes for Successive Games

The expected number of two state football outcomes (loss and win) with their probabilities for successive game times (t) is defined as in (9).

Table 2: The transition probabilities of football sports outcomes with game times

Time	$P_{00}(t)$	$P_{01}(t)$	$P_{10}(t)$	$P_{11}(t)$
1	0.860746816	0.139253184	0.523857216	0.476142784
2	0.813833866	0.186166134	0.700339264	0.299660736
3	0.798029382	0.201970618	0.759794231	0.240205769
4	0.792705015	0.207294985	0.77982399	0.22017601
5	0.790911291	0.209088709	0.786571808	0.213428192
6	0.790307005	0.209692995	0.788845078	0.211154922
7	0.790103427	0.209896573	0.789610919	0.210389081
8	0.790034843	0.209965157	0.789868923	0.210131077
9	0.790011738	0.209988262	0.789955841	0.210044159
10	0.790003955	0.209996045	0.789985123	0.210014877

Table 3: The expected time interval of the process spent in each state with game time

Time	$\mu_{00}(t)$	$\mu_{01}(t)$	$\mu_{10}(t)$	$\mu_{11}(t)$
5	4.139175498	0.860824502	3.223167823	1.776832177
10	8.089996422	1.910003578	7.170013747	2.829986253
15	12.03999998	2.960000016	11.12000006	3.87999994
20	15.99	4.01	15.07	4.93
25	19.94	5.06	19.02	5.98
30	23.89	6.11	22.97	7.03
35	27.84	7.16	26.92	8.08

Table 4: The transition counts of the football sport outcomes

State	0	1	Total
0	261	187	448
1	186	84	270
Total	447	271	718

Table 5: The expected number of football sports outcomes for successive games

Time (t)	$E(G_{n+t} \mid G_n) = G_n M^{(t)}$		Pr. of loss game	Pr. of win game
	Loss	Win		
1	337	286	0.54	0.46
2	154	126	0.55	0.45
3	73	65	0.53	0.47
4	35	37	0.49	0.51
5	17	24	0.41	0.59
6	8	16	0.33	0.67
7	4	11	0.27	0.73
8	2	7	0.22	0.78
9	1	5	0.2	0.8

$$E(G_{n+t} \mid G_n) = G_n \; M^{(t)}$$

$$= \begin{pmatrix} 142 & 606 \end{pmatrix} \begin{pmatrix} 0.58^{(t)} & 0.69^{(t)} \\ 0.42^{(t)} & 0.31^{(t)} \end{pmatrix}$$

Table 5 shows the expected number of losses and wins with their probabilities of Real Madrid football outcomes for successive game times ($t = 1, 2, \ldots, 9$).

CONCLUSIONS

From the results of the application, the following conclusions are found:

1. For given some successive game times, the transition probabilities $P_{00}\ (t)$ and $P_{11}\ (t)$ of the football sports outcomes are decreasing, but $P_{01}\ (t)$ and $P_{10}\ (t)$ are increasing
2. The expected time interval the process (football sport) spends in any one of the outcomes during each transition is continually increasing
3. The stationary distribution of the football sports outcomes shows that the Real Madrid team will win a game with the lowest probability and lose a game with the highest probability in a long time
4. The expected number of Real Madrid football sports outcomes for successive game times are continually decreasing till the number of games becomes five, then the mentioned team has won four games against one lost game.

REFERENCES

1. J. Medhi. *Stochastic Processes*. 2nd ed. New Age International, New Delhi, 2004.
2. B. Sericola. *Markov Chains: Theory and Applications*. John Wiley & Sons, United States, 2013.

3. U. N. Bhat and G. K. Miller. *Elements of Applied Stochastic Processes*. *Wiley Series in Probability and Statistics*. 3rd ed. Wiley-Interscience, Hoboken, NJ, p. 488, 2002.
4. R. Douc, E. Moulines, P. Priouret and P. Soulier. *Markov Chains: Basic Definitions.* (*Springer Series in Operations Research and Financial Engineering*). Springer International Publishing, Cham, 2018.
5. S. M. Ross. *Introduction to Probability Models*. 11th ed. Elsevier Inc, Amsterdam, Netherlands, 2014.
6. M. González, I. M. Puerto, R. Martínez, M. Molina, M. Mota and A. Ramos. *Workshop on Branching Processes and Their Applications*. Springer Science & Business Media, Berlin, Germany, 2010.
7. D. Axelrod and M. Kimmel. *Branching Processes in Biology*. Springer-Verlag, Berlin, Germany, 2015.
8. J. R. Kirkwood. *Markov Processes*. 1st ed. CRC Press, United States, 2015.
9. M. Lefebvre. *Applied Stochastic Processes*. Springer Science & Business Media, Berlin, Germany, 2007.
10. P. Haccou, P. Jagers and V. A. Vatutin. *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge University Press, United Kingdom, 2005.