Simulation and optimization of Penalty kicks and Shooting probability based on Monte Carlo method

Tiankuo Jiao

Department of Electrical & Electronic Engineering, University of Liverpool, 765 Brownlow Hill, Liverpool, Merseyside, L69 7ZX

E-mail: sgtjiao@liverpool.ac.uk

Abstract. Since there are few models that specialize Penalty kicks probability or shooting accuracy, this research innovatively uses the Monte Carlo method to simulate the situation of real football penalty kicks and shooting. Unifrom random generator and normal random generator are both used to better simulate the real situation, and finally the most suitable distribution model is chosen from these two distributions. Since the existence of goalkeepers is considered in a real football competition, this simulation experiment also divides the goalkeeper's actions into 5 types, and the corresponding probability of each action is also considered. Finally, specific shooting probability is collected from the Premier League in 2021 to compare with our simulated situation to obtain the most accuracy model. By comparing with real probability, I finally confirm the appropriate model and the goalkeeper's action probability. Similarly, the Monte Carlo method can also simulate some other real situations to determine the most suitable model.

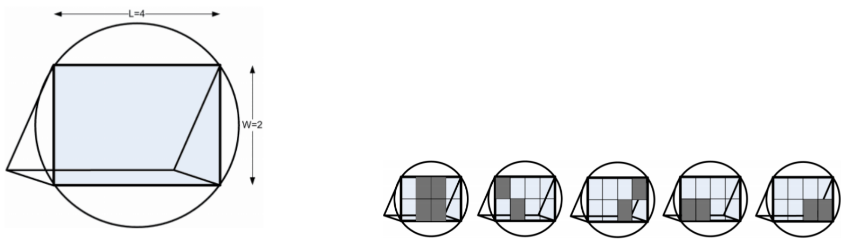
1. Introduction

Monte Carlo method, also known as statistical simulation method, is a kind of numerical calculation method. The Monte Carlo method was first proposed in the mid-1940s under the guidance of probability and statistics theory [1, 2, 3, 4]. Statistical estimates of the parameter can be simulated by setting a parameter of a hypothetical population and using random sequence of numbers to construct a sample of the population [3]. From a probabilistic point of view, Monte Carlo method can effectively reduce redundant operations in actual experiments and obtain results that are as close to real data as possible [5]. For problems of probability in real life, the occurrence of the event can be predicted to the greatest extent if choosing suitable the Monte Carlo method to simulate [4, 5, 6, 7]. In addition, when the sample size is large enough, the statistics obtained by the Monte Carlo method is relatively accurate and objective [1, 8]. Monte Carlo method can use different generators to generate different types of random numbers including uniform random generator [9, 10] and normal random generator [11, 12]. Each type of random generator will have a different impact on the simulation. Although the Monte Carlo method can solve the problems of image processing [13] and telecommunications [14], this article will mainly discuss Monte Carlo to optimize and improve the Penalty kicks and Shooting model.

In order to reduce the error of the simulation as much as possible, different types of random generators will be considered for simulation and the probability will be compared with the true probability of penalty kicks. In this simulation, the Monte Carlo method will be used to simulate the player's probability of scoring a penalty kick using uniform and normal random generators. The goalkeeper's actions and corresponding probability are also taken into account to simulate the true probability of a goal as much as possible. Finally, the most suitable model will be found and optimized through comparison to the situation of Penalty kicks and Shooting.

2. Simulated Experiment

The goal frame is assumed to be a rectangle with a length of 4m and a width of 2m. The player's shooting range is a circle which presents the total area of shoots. And the area of rectangular represents the range of successful shoots. The specific goal arrangement can be seen in Figure *1a*. In addition, the actions of goalkeeper and corresponding probability to jump in different position will be also considered. These actions include that the goalkeeper is defending in the middle of the goal frame, pounces to the upper right corner, pounces to the upper left corner, pounces to the lower right corner, pounces to the lower left corner. Therefore, the goalkeeper has the following 5 possible actions which can be also seen in Figure *1b*. In the process of simulation, the **plus** will be used to mark shooting successfully, the **blue circle** will be used to mark shooting unsuccessfully, and the **red dot** means that the point (football) appears outside the circular area.



(a) (b)

Figure 1. *(a) The goal arrangement, (b) Five possible actions*

I divided the rectangular area into eight grids. Because when the goalkeeper jumps up, his ability to defend the football will decrease, therefore, when the goalkeeper jumps up, he can only protect two grids. However, the goalkeeper will have a larger protection range when he stands in the center, which means that he can protect four grids. In the case of uniform distribution of random samples, the corresponding statistical principles will be based on **Geometric calculation method**. Therefore, in the process of shooting, random numbers will be uniformly distributed, so the specific calculation method will be:

$$P = \frac{S_{Number\ of\ shooting\ successfully}}{S_{Number\ of\ shooting}} \tag{1}$$

In this case of normal distribution, if it is assumed that the circle's radius will follow the normal distribution while the distribution to be centred at the centre of the shooting circle. The probability in this case will be:

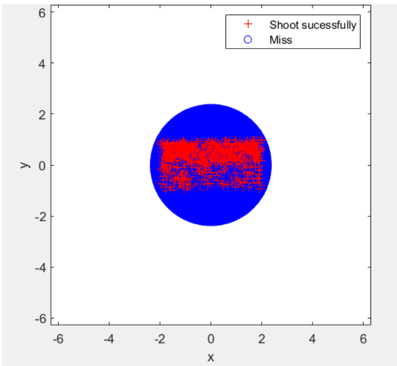
$$P_x = \frac{1}{\sigma * \sqrt{2\pi}} * e^{\frac{-(x-\mu)}{2\sigma^2}} \tag{2}$$

Because in the actual football competition, the goalkeeper has a small probability of staying in the middle (The first case). Generally, the probability of a goalkeeper standing in the middle will not exceed 20%. This is because the penalty player will choose the corner instead of the middle area to shoot, and the goalkeeper will also anticipate this situation in advance. Although the goalkeeper knows that he can protect a larger area of the grid by staying in the middle, the probability of a shot in the middle of the goal frame is too small. Therefore, in the simulation, the probability that the goalkeeper chooses to stay in the middle area will be set to 5%, 10%, 15% and 20%. After obtaining four sets of results corresponding to probability 5%, 10%, 15% and 20%, I will compare with the real probability to determine a model that is most suitable for simulating the goalkeeper's actions and the corresponding probability.

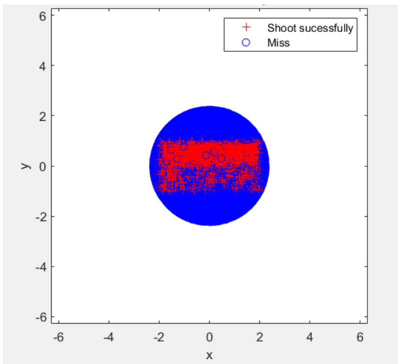
3. Results and discussion

Since the simulation experiment uses two different random generators, the experiment and discussion parts will also be divided into two different parts.

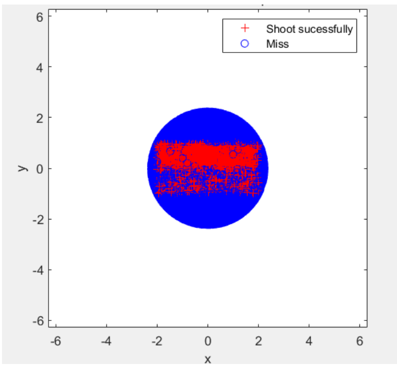
Each simulation result corresponds to the probability of the goalkeeper's different actions. Because the Monte Carlo method can only get a relatively accurate result when a large number of random numbers are used. Therefore, five different experiments are performed, and each experiment will simulate 10,000 shots. Finally, the average of the five experimental probabilities will be obtained to get the probability of the smallest error. To facilitate the simulation, I set the probability of the goalkeeper standing in the middle area of the goal frame as a variable. The following experimental statistics is the result of uniform random generator.



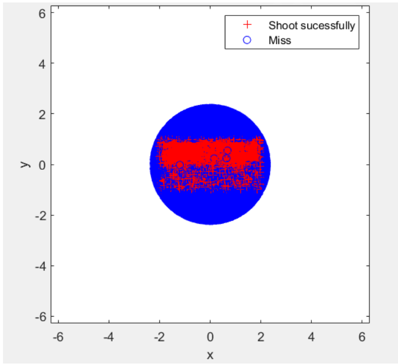
(a)



(b)



(c)



(d)

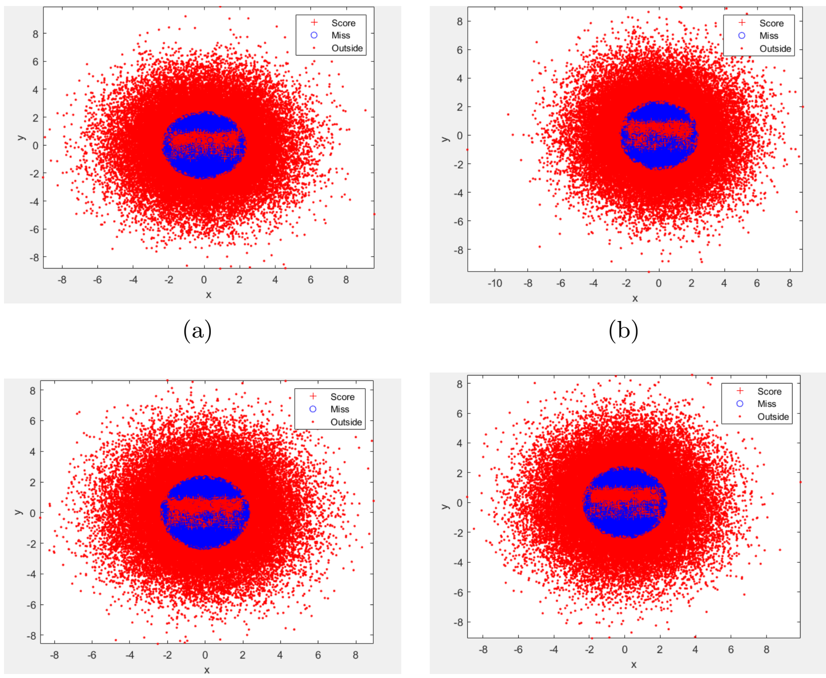
Figure 2. In the case of using a uniform generator, the probability that the goalkeeper chooses to stay in the middle area will be set to *(a) 20%, (b) 15%, (c) 10%, (d) 5%*

As the probability of the simulated goalkeeper standing in the center of the goal gradually decreases, it can be observed from the figures that the number of blue circles (Miss) in the lower half of the red (rectangular) area is gradually increasing. Due to the large number of samples, the differences between each result figure are not very obvious. Therefore, in order to better compare the differences, I also record the probability of each situation.

Table 1. Based on the situation under different probabilities, the simulation results of the Uniform random generator

Probability of staying in the middle	Probability of shoot successfully (Get a score)
`5% probability of standing in the middle`	37.58% probability of shooting successfully
`10% probability of standing in the middle`	37.76% probability of shooting successfully
`15% probability of standing in the middle`	37.35% probability of shooting successfully
`20% probability of standing in the middle`	37.22% probability of shooting successfully

It can also be seen from the above probability distribution that as the probability of the goalkeeper standing in the middle gradually decreases, the probability of getting a score (shoot successfully) gradually increases. This is to be expected, because when the goalkeeper is standing in the middle of the goal frame, he obviously has a larger sight range and better physical defense preparation. Therefore, I assume that the goalkeeper can prevent four grids in the center area, which is larger than the goalkeeper jumping to other sides. As the sample continues to increase, the goalkeeper can obviously protect more goals that shoot successfully, thereby reducing the accuracy of shooting. However, the simulation takes into account the uniform distribution, but in a real game, the penalty player does not randomly shoot in the direction of the goal, but selects the goalkeeper's physical and blind spots to shoot instead. The following experimental results are the results of using a normal random generator.



(a) (b) (c) (d)

Figure 3. In the case of using a normal random generator, the probability that the goalkeeper chooses to stay in the middle area will be set to *(a) 20%, (b) 15%, (c) 10%, (d) 5%*

For the convenience of simulation, I assume that the distribution to be centred at the centre of the shooting circle, and the standard deviation will be equal to the circle radius ($\sqrt{5}$). Due to the normal random generator used this time, there will be many points appearing outside the circular area. In the simulation, I mark the points outside the circle as **red dot**, mark the successful point as a **plus**, and mark the failed shoot as a **blue circle**. Compared with the case of a uniform random generator, since many points appear outside the circular area this time, the overall probability of shooting success will be lower than that in uniform generator. This can also be found in the probability statistics table below.

From the above results, I can find that the probability of the normal random generator is generally lower than that of the uniform random generator. And as the probability of the goalkeeper standing in the center decreases, the success rate of the player's shot will gradually increase. This is in line with expectations, and the specific reason is the same as that of a uniform random generator. After searching for information, I select the shoot accuracy of some

Table 2. Based on the situation under different probabilities, the simulation results of the Normal random generator

Probability of staying in the middle	Probability of shoot successfully (Get a score)
5% probability of standing in the middle	16.13% probability of shooting successfully
10% probability of standing in the middle	16.03% probability of shooting successfully
15% probability of standing in the middle	15.78% probability of shooting successfully
20% probability of standing in the middle	15.75% probability of shooting successfully

famous clubs as our reference data for selecting the most suitable model. All Shooting accuracy (%) comes from the 2021 Premier League [15].

Table 3. Based on the situation under different probabilities, the simulation results of the Normal random generator

Name of club	Shooting accuracy	Reference
Liverpool FC	34%	[15]
Manchester City	35%	[15]
Arsenal	36%	[15]
Chelsea	34%	[15]
Leicester City	35%	[15]
Manchester United	36%	[15]

It can be seen that the shooting accuracy of all clubs are in the range of 34%-36%, and the average accuracy is 35%. Since the probability generated by the normal random generator is much lower than 35%, it will be excluded when considering simulating the probability model. Instead, several probabilities generated by uniform random generator are more in line with the true probability of statistics (shooting accuracy). Additionally, when the probability of the goalkeeper standing in the middle is 20%, the error between the simulated situation and the real situation is the smallest, and the absolute error is 2.22%, the relative error is 6.3%, which is within an acceptable range. Therefore, the best model is when the probability of the goalkeeper standing in the middle is 20%, and the shooting distribution is the uniform distribution. Considering that the goalkeeper is impossible for a high probability of standing in the middle, therefore, 20% will be a relatively reasonable probability.

4. Conclusion

In the process of simulating the Penalty kicks and Shooting, Monte Carlo method is used to simulate the real situation in detail and verify which distribution is more suitable for this situation from uniform and normal distribution models. At the same time, the goalkeeper's pounce to prevent a successful shoot and corresponding probability are also taken into account. The area of goal frame is divided into eight grids. Considering that the goalkeeper has five actions, I assume that the goalkeeper can only protect four grids when standing in the middle, and the others four actions can only protect two grids from being shot successfully. After a series of simulation, the shooting accuracy of each distribution is obtained with corresponding simulated figures. At the same time, in order to find a more suitable model, the 2021 Premier League Statistics (shooting accuracy) has also been used as a reference. I find that the shooting accuracy of any club is around 35%, and the probability range of the normal random generator is much lower than this probability. Therefore, the uniform distribution is used to better simulate the shooting situation. By adjusting the probability of the goalkeeper standing in the center, I finally choose the probability model of 20% as the final suitable probability. This is because the error of the 20% model is the smallest relative to the actual shooting statistics, and the probability of 20% is also the maximum acceptable probability of the goalkeeper standing in the middle. In the future, we can also use the Monte Carlo method to simulate other realistic examples. Monte Carlo has a wide range of applicability when dealing with such problems and the simulation process is relatively simple. When the random number sample size is large enough, the data can be obtained accurately.

References

[1] Landau D P and Binder K 2005 *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge: Cambridge University Press)

[2] Liu S J 2008 *Monte Carlo Strategies in Scientific Computing* (Berlin: Springer)

[3] James F 1980 Monte Carlo theory and practice *Rep. Prog. Phys.* **43** 1145

[4] Malvin H K and Paula A W 2009 Monte Carlo methods *WILEY VCH*

[5] Alexander S 2005 Monte Carlo Sampling Methods *Handbooks in Operations Research and Management Science* **10** 353-425

[6] Robert L H 2010 Monte Carlo statistical methods *AIP conference proceedings* **1204** 17-21

[7] Ferrenberg A M and Swendsen R H 1989 Optimized monte carlo data analysis *Computers in Physics* **3** 101-104

[8] Kroese D P and Brereton T and Taimre T and Botev Z I 2014 Why the Monte Carlo method is so important today *Wiley Interdisciplinary Reviews: Computational Statistics* **6** 386-392

[9] Tezuka S 2012 Uniform random numbers: Theory and practice *Springer Science Business Media*

[10] MacLaren M D and Marsaglia G 1965 Uniform random number generators *Journal of the ACM (JACM)***12** 83-89

[11] Leva J L 1992 A fast normal random number generator *ACM Transactions on Mathematical Software (TOMS)* **18** 449-453

[12] Kinderman A J and Ramage J G 1976 Computer generation of normal random variables *Journal of the American Statistical Association* **71** 893-896

[13] Hiroshi W and Satoshi S and Yutaka O and Hwee K L 2011 Monte Carlo methods for optimizing the piecewise constant Mumford–Shah segmentation model *New Journal of Physics* **130** 023004

[14] Fishman G 2013 Monte Carlo: concepts, algorithms, and applications *Springer Science Business Media.*

[15] Premier league 2021 Club statistics *premierleague*