Xi'an Jiaotong-Liverpool University

西交利物浦大學

EXISTENCY OF GAMBLER'S FALLACY AND HOT OUTCOME ON GOALKEEPER DECISION MAKING

赌徒谬误和热门结果于守门员决策

Name: **Nico Septianus**

Student ID: **1510228**

Supervised By: **Dr. Cihangir Kan**

17th April 2019

Financial Mathematics

Department of Mathematical Science

ABSTRACT

In the world full of uncertainty, many assumption and doubtfulness has been occurred on people's decision making. Eventually, this might lead to some undesirable results which shows people try to gamble with his life choice. However, to minimize the risk of getting the unwanted outcome, many strategies and procedure have been design by researcher and geniuses. On each penalty shootout, goalkeeper tries to find the best way to defend the goal from 5 independent kickers through random process. Some beliefs such as Gambler's fallacy and Hot outcome are expected pattern to benefit goalkeeper when facing penalty shootout. The objective of this study is to check whether those fallacies are exists for goalkeeper on the shootout following streaks of correct and incorrect prediction. Another objective is to observe if more correct prediction will be brought from either of fallacy in the real world or it is just as a mere fallacy. The penalties data were obtained from 3 different football tournament such as World Cup, AFC Asian Cup, and UEFA Europa League from 1998 to 2018. After being extracted, there were total of 600 penalties that need to be observed. To check the existence of either fallacy (first objective), binomial test was used and compared to the empirical evidence. As the increase length of correct prediction, the goalkeeper tends to dive to opposite direction (Gambler's fallacy) significantly. Whereas, there were no clue for either fallacy on continuing incorrect prediction, which shows goalkeepers did not dive with the pattern significantly for streaks of incorrect prediction. Furthermore, Fisher's exact test was tested to check which fallacy was better. Even though, it shows in favour more to Gambler's fallacy, yet the test did not reveal any significant dependency between Gambler's fallacy and Hot outcome. Hence, the fallacies cannot be accepted in the real world or in fact it was only a mere fallacy. Some psychological effect needs to be considered further on this case as well.

Keywords: Gambler's fallacy, Hot outcome, Decision Theory, Fallacy biases

摘要

在充满不确定性的世界中，人们的决策制定了许多假设和怀疑。最终，这可能会导致一些不良结果，这表明人们试图用他的生活选择进行赌博。然而，为了最大限度地降低获得不必要结果的风险，研究人员和天才设计了许多策略和程序。在每次点球大战中，守门员都试图从五个独立的踢球者随机过程中找到防守球门的最佳方法。一些信仰，如赌徒的谬误和热门结果，预计将在面对点球大战时受益于守门员。这项研究的目的是检查在正确和不正确的预测条件下，守门员在枪战中是否存在这些谬误。另一个目标是观察是否会从现实世界中的任何一个谬误中得出更正确的预测，或者它仅仅是一个谬误。罚球数据来自于 1998 年至 2018 年的三个不同的足球锦标赛，如世界杯，亚足联亚洲杯和欧足联欧洲联赛。在被提取后，总共需要观察六百个罚球。为了检查是否存在任何谬误（第一个目标），使用二项式检验并与经验证据进行比较。随着正确预测的增加长度，守门员倾向于显着地向相反方向（赌徒谬误）倾斜。然而，对于持续不正确预测的任何谬误都没有任何线索，这表明守门员没有因为错误预测的条纹而显着地潜入模式。此外，Fisher 的精确测试进行了测试，以检查哪个谬误更好。尽管如此，它更有利于赌徒谬误，但测试并没有揭示赌徒谬误和热门结果之间存在任何重大依赖关系。因此，这些谬论在现实世界中是不可接受的，或者实际上它只是一种谬论。在这种情况下，还需要进一步考虑一些心理影响。

TABLE OF CONTENT

1. Introduction 1
2. Literature Review 2
3. Methodology 4
   3.1 Data 4
   3.2 Method Analysis 4
4. Data Results and Analysis 6
   4.1 Descriptive Statistics 6
   4.2 Gambler Fallacy and Hot Outcome Testing 6
   4.3 Fallacy Biases Analysis 9
5. Discussion 12
6. Conclusion 13
7. Appendix 14
8. References 17

LIST OF TABLES AND FIGURES

TABLES

Table 1. Numbers and Weights of 3 different tournaments penalty shootouts dataset 4

Table 2. Example of 2 consecutive correct/incorrect prediction 4

Table 3. Frequency of Goalkeepers dive to opposite or same direction in the next kick after several sequence streak length of correct prediction (1,2 and streaks) 6

Table 4. Frequency of Goalkeepers dive to opposite or same direction in the next kick after several sequence streak length of incorrect prediction (1,2 and 3 streaks) 8

FIGURES

Figure 1. Ball positioning based on left/right 5

Figure 2. Percentage of Goalkeepers dive to opposite or same direction in the next kick after several sequence streak length of correct prediction (1,2 and streaks) 7

Figure 3. Percentage of Goalkeepers dive to opposite or same direction in the next kick after several sequence streak length of incorrect prediction (1,2 and streaks) 8

Figure 4. Proportion of Goalkeepers dive correctly predicted in the next kick after several sequence streak length of correct prediction (1,2 and 3 streaks) 9

Figure 5. Proportion of Goalkeepers dive correctly predicted in the next kick after several sequence streak length of incorrect prediction (1,2 and 3 streaks) 10

1. Introduction

Football or as some people know as soccer has been a massive truly enjoyable sports for all not only young and adult but a fair amount of elders as well. Since this profound sport from 1863s has reached more than 100 more years of game span (FIFA, 2018). Therefore, there is no doubt this popular sport is very familiar for every ethnic, gender, age and it is also being acknowledged by the world. Regardless that, the improvement of new technology and regulation is needed as well which might result in more interesting and fairer game for players and spectators. For instance, FIFA (2015) stated about 15 minutes of each period of extra time instead of golden goal after both teams tied the scores on 90 minutes gameplay. Furthermore, if the result still the same after both periods on extra time, the referee will conduct a match-winning procedure called penalty shootout. A penalty shootout is multiple succession of penalty kick where the winning team is with greater goal determine after five kicks for each team or 10 kicks in total (FIFA, 2015). Otherwise, the shootout will keep going until there is a goal deficit with one team that failed to score. Apart from the tense of the penalty shootout, many football observers and fans think whether penalty requires less skill and it is just a matter of luck. However, according to Arrondel, Duhautois, and Laslier (2018) research, psychology state and skills are also importantly needed for the kicker to execute penalty. High skilled cap player is often prioritized to take the penalty because they tend to be calmer, more confident and well experience. Nevertheless, psychology is still the utmost role to consider when taking penalty kick due to issues like pressure, bad health-condition (fatigue, injury) and the personal issue might influence failure on scoring (Arrondel, Duhautois, and Laslier, 2018). Additionally, Dalton, Guillon, and Naroo (2015) also concluded whether to optimize the chance of success on taking a penalty, coach requires to boost player technically and psychologically beforehand.

On the other hand, from the goalkeeper perspective to stop the kicker from scoring, factors that they require are absolutely more or less similar to the kicker. Nevertheless, goalkeepers are on disadvantage since they have less information about the ball positioning compare to the kicker, they might guess direction before the ball kicked (Misirlisoy & Haggard, 2014). Moreover, there is no correlation between each sequence of a penalty kick, they tend to be independent or i.i.d (independent identically distributed) random variable (Palacios-Huerta, 2002). Likewise, the coin toss where both sides have fifty-fifty percent chance with independent random sequence if being toss couple times (Stein and Wapner, 2017). Thus, there is randomness from the penalty kicker and goalkeeper only need to guess either right or left if the line is drawn in the middle of the goal and only set left and right direction to choose. In spite of independence, researchers and observers still seek patterns to deny this false independent state and gain benefits from it. For instance, Nesbitt and Barrass (2004) found out by identifying a new pattern in the stock market might gain an edge over other competitors, that means it can predict the future outcome through the pattern. Since this is what the researchers looking for. Hence, they invented two beliefs about penalty shootout case which are gambler fallacies and hot hand fallacies. Both fallacies are not simply contrasting to each other, but one can give the opposite view of each other and vice versa (Croson and Sundali, 2005).

The aim of this paper is to check the presence of either one of the fallacies in the penalty shootout data and determine if it can be helpful in the real world or just a theory of mere fallacy.

2. Literature Review

Gambler fallacy and hot hands fallacy can be interpreted as two behavioral biases, which is believed can mislead the perception of every person although he/she has already followed the pattern. Sundali and Croson (2005) define gambler fallacy as the belief of negative autocorrelation of an i.i.d random sequence. The believers of this fallacy admit whether a streak will be more highly possible to break as the sequence streak getting longer. For instance, on flipping a fair coin and having 3 streaks of the head in a row, the gambler fallacy person will definitely choose tail on the fourth flip. They tend to believe the probability of tail appears in the fourth round are higher than 50% although they are independent sequence between each flip. Hot hand fallacy is defined as the belief of positive autocorrelation of an i.i.d random sequence. Opposite to gambler fallacy, hot hand believers admit the outcome of the same streaks tends to be increasingly likely for longer streaks. For example, the hot hand person will still pursue playing after winning 3 streaks of flip coin game in a row. From the "hot" term in hot hand, they believe the capability of winning is over 50% although the sequence between each flip is independent (more detail can be searched at Sundali and Croson, 2005). Therefore, two biases can be adaptive or just a mere fallacy which leads to misconception.

Despite both fallacies, some psychology side needs to be pointed out as well. Since the occurrence of fallacy is dependent on the outcome from each shootout. Moreover, if recall the psychology part on introduction, each sequence outcome is determined from the player/goalkeeper behavior as well. Therefore, fallacy and player behavior are connected. It is proved from a book written by Welsh (1980) where on penalty shootout goalkeeping, goalkeeper decision making is crucial for winning the game. Hunter, Murphy, Angilletta Jr. and Wilson (2018) paper also added goalkeeper anticipation which also part of behavior is important in the penalty shootout. Another early study by Kahneman and Miller (1986) regarding psychology on a penalty shootout, stated a case about goalkeeper decision making, whether they should be more active (move left or right) or inactive (stay in middle). Furthermore, based on their research, after a goal occurs, fans and observers will psychologically perceive a worse habit of the goalkeeper when they just stay in the middle. This is because less action being put by goalkeeper. Consequently, due to this psychology effect, most of the goalkeeper tend to be dive right/left rather than stay in the middle or this psychology theory can be called as norm theory (Kahneman and Miller, 1986). This also supported by Zeelenberg et al. (2002) and claimed it is natural to choose action instead of inaction especially in losing condition. Likewise, from the research, after a football team losing a game, the coach will act more (changing strategy and looking for weakness) rather than doing nothing (keep the same strategy). Another research, Apesteguia, and Palacios-Huerta (2010) observed kicking order is another factor that can affect the player and team psychology. They said random psychological pressure may be generated between shoots. This means each shoot and save may give either morality boost or mental breakdown to the corresponding team.

The reason behind using some examples about coin toss on the previous part is due to the originality of gambler fallacy and hot hands. Laplace (1825), a French mathematician firstly discovered the concept of gambler fallacy through some experiment of a coin toss and later then Laplace did expand his study with some similar area like on gambling and lottery. Subsequently, some conclusion like independence and probability illusion are drawn. This is written on the book on "Illusions in the Estimation of Probabilities" part. Despite theory, some

early real-life application about gambler fallacy was finally declared by Hans Reichenbach, Reichenbach (1949) found out a certain pattern on human behavior through events such as roulette and coin toss, where bettors tend to shift their wagers after similar outcomes appeared frequently in a row. As a result, research on this behavior getting progressively studied and argued. Nowadays, this behavior is known as gambler fallacy. Similarly, Clotfelter and Cook (1993) researched on the behavior for all 52 consecutive winning sample data on pari-mutuel lottery. Pari-mutuel lottery works by gathering all bets in a pool which later deducted by tax and will be shared equally among the winning bet eventually. Later then the research showed a unique pattern behavior likewise the probability of the same number will be picked for following or next month lottery was reduced after that number win today. Three years later, Terrell and Farmer (1996) were inspired by Clotfelter and Cook paper. They found similar results on dog and hound racing. Bettors will less likely bet on the first-place dog on previous race. This implied the gambler fallacy affect betting patterns and odds. Other study related by Ayton and Fischer (2004), they examined whether gambler fallacy exists on a roulette wheel. Subsequently, they found there is negative recency (gambler fallacy) following streak outcome of 1 to 5. Croson and Sundali (2005) also supported them, they observed despite the statistical illusion in gambling, a fair number of gamblers still fell into the fallacy illusion after seeing the same outcome appears in 5 or more consecutively. In contrast, Xu and Harvey (2014) found no evidence of gambler fallacy streaks on data of online gambling. However, they learned if gamblers are more into safer bet after winning and be riskier when losing consecutive bets.

Opposite to Gambler fallacy, Hot hands fallacy is getting publicly known in the early 60s. Edwards (1961) observed there is positive recency during probability matching hypothesis experiments. It showed the longer the event occurred, the more likely contestant to predict the event. Tversky and Gilovich (1989) also learned the trend of hot outcomes in basketball. They believed each successful shoot might increase player momentum, confidence and morality boost. Therefore, players are more likely to points between the sequence of successes shoot or this player known as “on fire”. Keren and Lewis (1994) discovered the lab results on gambling experiment favoring in the hot outcome. Later after, Ayton and Fischer (2004) found a hot result on success and failure outcome of the roulette wheel. Since both results could give different psychological behavior to the gamblers (safer or riskier bet). Sundali and Croson (2005) studied similar evidence on hot outcome regarding gambling, they analyzed 18 hours of roulette wheel data. The data showed consistency on bettors’ performance where they tend to play more when winning the game since they assume it is their ‘lucky day’. However, when having a bad day and losing too much game, they will prefer to back off (leave the seat/change seat).

The objective is to check the existence of gamblers fallacy or hot outcome pattern on the real-world data and determine whether this pattern will be useful for the goalkeeper in the future. The data will consist of penalty shootouts from World Cup, AFC Asian Cup, and UEFA Europa League from the last 20 years (1998 – 2018). Moreover, keeping the same frequency of total kicks limit is needed here, therefore penalty shootouts that is more than 10 kicks in total will be eliminated. The reason is to avoid more error and some unwanted outcome.

3. Methodology

3.1 Data

The data will be gathered from some reliable sport resources likewise theadgalternative.com, YouTube and Goal.com. First, all penalty shootouts from 3 corresponding competitive tournaments over 20 years will be extracted from theadgalternative.com and Goal. From the result there are a total of 67 penalty shootouts. Those 67 shootouts video will be searched on YouTube, so it can be examined the goalkeeper dive direction. Lastly, conclusion need to be made to answer the objectives.

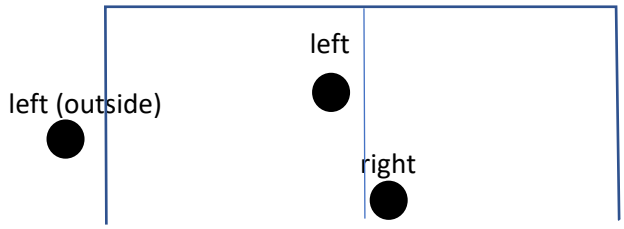
Table 1. *Numbers and Weights of 3 different tournaments penalty shootouts dataset*

Tournaments	Number of Penalty	Weights
UEFA Europa League	40	60%
World Cup	19	28%
AFC Asian Cup	8	12%
Total	**67**	**100%**

3.2 Method Analysis

As there are 3 possible shooting direction for penalty shootouts (left, right and centre). To ease the analysis on gambler fallacy and hot outcomes, a line will be drawn in the middle of the goal to split the directions. So, the direction will be either left or right only. Another reason is due to not much data of goalkeeper stays in the middle. Therefore, when goalkeeper freeze in the middle, the data will simply be omitted. This can be referred from the literature review whether goalkeeper is favour more on action than inaction.

Figure 1. *Ball positioning based on left/right*



The subject of this research is to analyse the goalkeeper adjustment. Moreover, the aim is to check whether both fallacies exists and efficient in real life. Therefore, goalkeeper guess (correct or incorrect) needs to be examined. Goalkeeper same and opposite direction of correct and incorrect decision will be checked. This to obtain the decision behaviour of real-world goalkeeper when facing numbers of different condition. For example, the decision where to dive right after correctly predicted 1, 2 and 3 shoots in a row are all different. Afterward, the behaviours data can determine which fallacy they used. For more detail, when goalkeeper choose opposite direction next (e.g. after right) after at least one or more same correct prediction (e.g. previous left-correct), this will be stated as gambler fallacy. Otherwise if the goalkeeper still maintains previous correct guess (e.g. previous left-correct) and choose same direction next (e.g. after left), this means hot outcome. However, for incorrect guess is reversed, it will count as gambler fallacy when goalkeeper choose same direction (e.g. after left) from the previous incorrect direction (e.g. previous left-incorrect), whereas hot outcome

is when goalkeeper change direction (e.g. after right) from previous incorrect direction (e.g. previous left-incorrect).

Table 2. *Example of 2 consecutive correct/incorrect prediction*

Streak length	Previous kick	Previous dive	Prediction Result	Next Kick Prediction	
				Gambler Fallacy	Hot Outcome
2	Left (2 in a row)	Left (2 in a row)	Correct	Right	Left
	Right (2 in a row)	Right (2 in a row)	Correct	Left	Right
	Left (2 in a row)	Right (2 in a row)	Incorrect	Right	Left
	Right (2 in a row)	Left (2 in a row)	Incorrect	Left	Right

In total there are 4 kinds of table following the 1,2,3,4 streaks from each 5 shoots from each team. As both left and right are binomially equal distributed with probability of 0.5 each, therefore the method that is going to use to process this data is Binomial distribution. This method might determine whether there is gambler fallacy or hot outcome exist on the data (Croson & Sundali, 2005). Another reason is there is independency for each subsequent kick, this might tell if binomial is the perfect test. One-tailed and Two-tailed test will be needed to test the goalkeeper dive behaviour (same or opposite way) after different streaks.

To check whether the fallacy is not a mere fallacy, it needs to show whether the number of correct predictions of gambler fallacy and hot outcome (follow the table 2) is significantly more than the non-gambler fallacy and non-hot outcome dive. In addition, length of streak also goes along with the correct prediction (positive correlation). This means, the higher the streaks length it tends to be better prediction by using the fallacy. Therefore, theoretically the longer the streaks, goalkeeper more likely to predict the outcome using the fallacy. On the other hand, a test needs to be run to prove whether the empirical data is right or wrong. Since, the sample sizes are small, fisher's exact test might work better instead of chi-square and F-distribution test (Sarkar, 2017). By checking at 5% significance level, whether the test is significantly difference or not.

$$H_0: x_1 \text{ is independent to } x_2$$

$$H_1: x_1 \text{ is dependent to } x_2$$

Where x_1 and x_2 is the probability of gambler fallacy and hot outcomes respectively. The test will be run on correct prediction of gambler fallacy and hot outcomes after each streaks of correct and incorrect prediction. So, the outcome will be separated into two different cases (correct and incorrect) which will be combined as one in the conclusion. If H_1 is true, or there is statistically significant difference between two fallacy (dependent means one increases and the other will go down), the empirical data will be supported. That means, when streak goes on there is fallacy that can be relied on.

4. Data Results and Analysis

4.1 Descriptive Statistics

From 67 data of penalty shootouts, 600 penalties have been obtained to be extracted and processed. Without considering the middle part (when goalkeeper freeze), 52% of kicks went to the left side of goalkeeper, 48% of kicks went to the right side whereas the percentage that goalkeeper dove to the left was 52.5% and 47.5% to the right. To be certain about this probability and check whether there is no biasness on this probability result, a test such as binomial distribution test will be used. Since, they are independent between each other kicks and only two categories tested (left and right) therefore binomial test is a perfect for this case. Two tailed tests will be brought, where L is the probability of choosing left and R is the probability of choosing right,

H_0: L is likely to occur as R (L = R)

H_1: L is not likely to occur as R (L ≠ R)

The binomial distribution shows there are equally likely for kickers to shoot left and right (p-value= 0.3478, for α=0.05, p-value>α), it is not significant so H_0 is not rejected. Similarly, the goalkeepers dive equally on either direction (p-value=0.2364, for α=0.05, p-value>α), since H_0 is not rejected and it is not significant. Both results correspondingly describe the data for left and right are not bias and it is safe to proceed using this data.

4.2 Gambler Fallacy and Hot Outcome Testing

Recall Table 2 from methodology part, by doing similar concept which is analysing goalkeeper gambler's fallacy and hot outcome for each penalty varying on a streak length. Some possible results can be obtained below.

Firstly, Table 3 shows the prediction of goalkeeper's dive for opposite or same direction in the next kick after certain correct prediction streak. For instance, the outcome will be placed on the table as 2 sequence length and same direction when goalkeeper choose left after successfully predict previous 2 kickers direction which is also left (first 2 kickers shoot to the left).

Table 3. *Frequency of Goalkeepers dive to opposite or same direction in the next kick after several sequence streak length of* ***correct prediction*** *(1,2 and 3 streaks)*

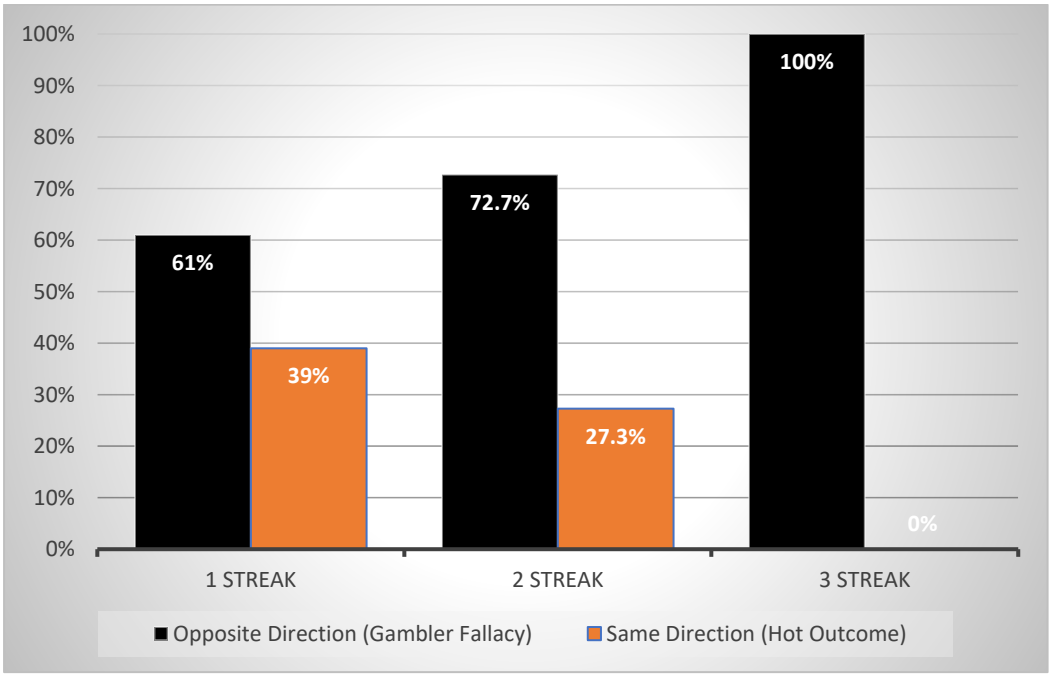
Sequence Length	Opposite Direction	Same Direction	Total
1	138	88	226
2	24	10	33
3	3	0	3

According to the chapter 3.1, goalkeeper dives on the opposite direction for previous correct prediction will be regarded as Gambler's fallacy, whereas same direction will be considered as Hot outcome. Moreover, some results from Table 3 can be extracted and interpreted such as, goalkeeper dove 138 (61%) times in the opposite direction and 88 (39%) times in the same direction after one streak of correct prediction in any direction (left or right). As can be seen majority goalkeeper go for opposite dive after a streak which conclude they

followed Gambler's fallacy. Following two streaks of predictions in any direction (left or right), goalkeeper dove to the opposite direction 24 (72.7%) times and same direction 10 (27.3%) times. At the most of 3 streaks, with the total only 3 dives, all 3 (100%) dives accounted to opposite direction, while no goalkeeper (0%) dove in the same direction after 3 consecutive streaks of correct prediction.

The empirical evidence can be viewed in the Figure 2, where the bar graph shows the pattern of Gambler's fallacy instead of Hot outcome, since the numbers of goalkeeper dove to opposite direction is monotonically increasing as streak length increases.

Figure 2. *Percentage of Goalkeepers dive to opposite or same direction in the next kick after several sequence streak length of* ***correct prediction*** *(1,2 and 3 streaks)*



To check whether the empirical result is eligible, one-tailed and two-tailed binomial test for each and every length streak from the data need to be performed. Using 0.5 hypothesized probability of success and 5% of significance level (α), hypothesis testing can be constructed as below,

$$\text{One-tailed: } H_0: p <= 0.5 \quad \& \quad H_1: p > 0.5$$

$$\text{Two-tailed: } H_0: p = 0.5 \quad \& \quad H_1: p \neq 0.5$$

Where p is the probability of goalkeeper dive to opposite direction. Using the R-Gui program to calculate the p-value, for streak length of one (two tailed p-value= 0.001071; one tailed p-value = 0.0005355), it is significant so the null hypothesis can be rejected at 5% significance level (α). Similarly, the null hypothesis was also rejected at 5% significance level for streaks length of two (two tailed p-value= 0.01353; one tailed p-value = 0.006765). Whereas, for streaks length of three, the p-value is not significant (two tailed p-value= 0.25; one tailed p-value = 0.125) so the null hypothesis cannot be rejected at 5% significance level. The reason why the length of three streaks was differ from other two is due to lack of sampling data which result sampling bias. Thus, it is not included to further analysis and can be omitted as well. On

the other hand, there are strong significant evidence for length of one and two streaks which show the presence of Gambler's fallacy through rejecting null hypothesis of one-tailed and two-tailed test (so probability of dive to opposite direction is > 0.5 and ≠ 0.5).

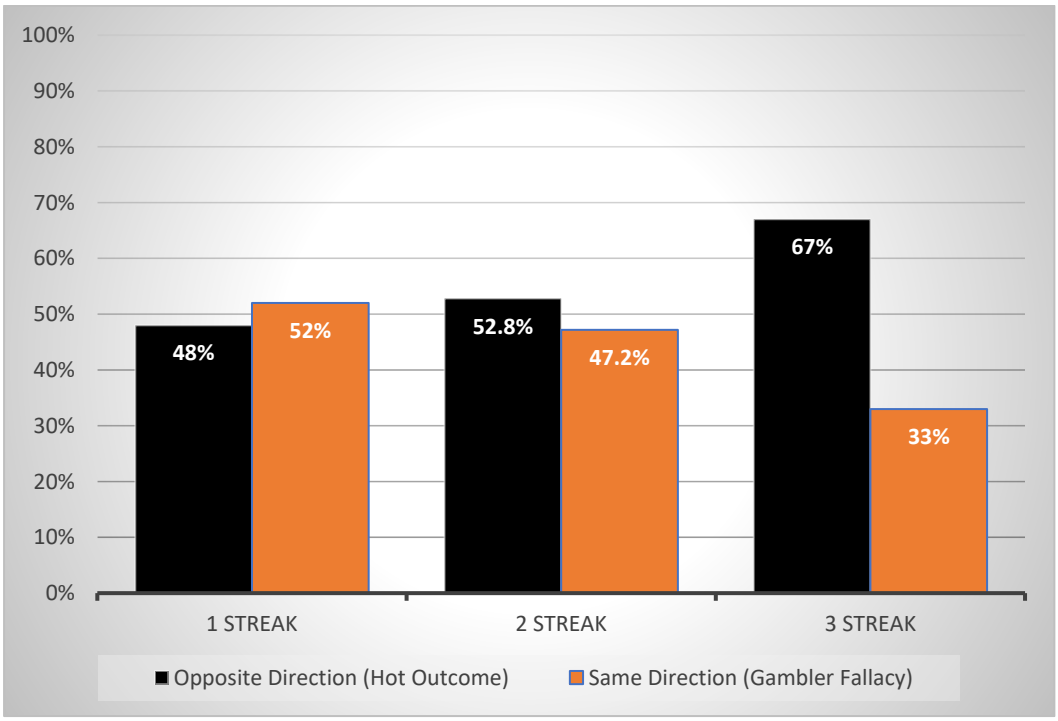
Secondly, Table 4 describes the frequency of goalkeeper's dive for opposite or same direction in the following kicks after incorrect streaks of prediction. For instance, the outcome will be placed on the table as 2 sequence length and same direction when goalkeeper choose right after wrongly predict previous 2 kickers direction which are left (first 2 kickers shoot to the left).

Table 4. *Frequency of Goalkeepers dive to opposite or same direction in the next kick after several sequence streak length of* ***incorrect prediction*** *(1,2 and 3 streaks)*

Sequence Length	Opposite Direction	Same Direction	Total
1	80	87	167
2	19	17	36
3	2	1	3

Unlike the correct prediction, the incorrect prediction will regard Gambler's fallacy when goalkeeper dives in the same direction as the previous dives, otherwise it will be a Hot outcome (Goalkeeper dives to the opposite direction). As shown in the Table 4, from the total of 167 dives after a streak length, 80 (48%) dove to the opposite direction, while 87 (52%) dove in the same direction. Following prediction on length of two streaks in any direction (left or right), goalkeeper dove to the opposite direction 19 (52.8%) times and same direction 17 (47.2%) times. At last, after the length of incorrect three streaks, only 2 (67%) were hot outcome dives (Opposite) and 1 (33%) were Gambler's fallacy dive (Same direction. Figure 3 illustrates the proportion of the data from Table 4 in term of bar graph.

Figure 3. *Percentage of Goalkeepers dive to opposite or same direction in the next kick after several sequence streak length of* ***incorrect prediction*** *(1,2 and 3 streaks)*



The empirical evidence can be viewed in the Figure 3, where the bar graph is favour more in the pattern of Hot outcome instead of Gambler's Fallacy, since the numbers of goalkeeper dove to opposite direction is monotonically increasing as streak length increases. To check whether the empirical result is eligible, one-tailed and two-tailed binomial test for each and every length streak from the data need to be performed. However, the one-tailed test for 1 streak is differ from the 2, 3 streaks due to different sampling ratio which lead to distinct null hypothesis. Moreover, using 0.5 hypothesized probability of success and 5% of significance level (α), hypothesis testing can be constructed as below,

One-tailed: H_0: $p >= 0.5$ & H_1: $p < 0.5$ (for 1 streak)

One-tailed: H_0: $p <= 0.5$ & H_1: $p > 0.5$ (for 2,3 streak)

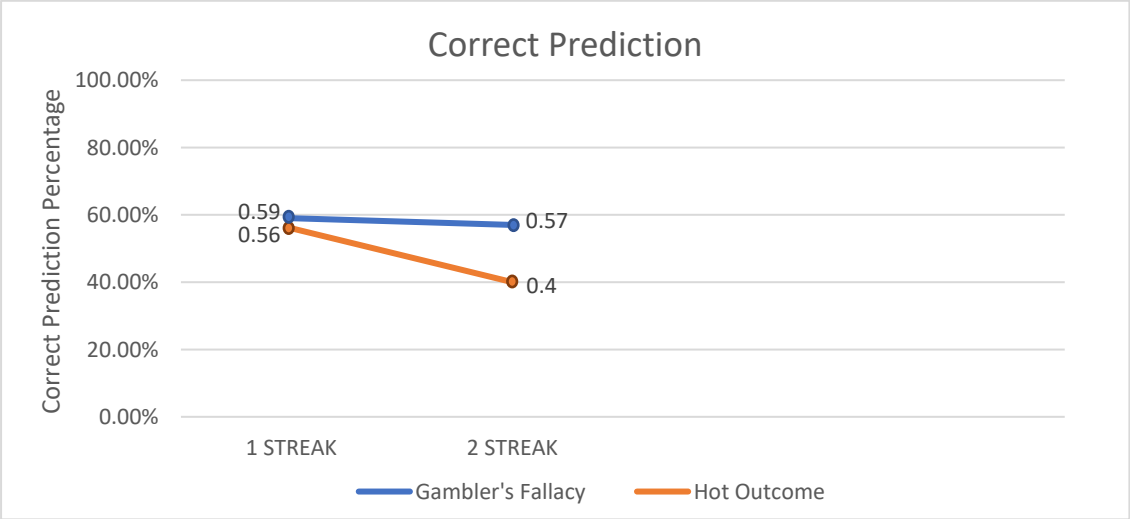
Two-tailed: H_0: $p = 0.5$ & H_1: $p \neq 0.5$

Where p is the probability of goalkeeper dive to opposite direction. Using the R-Gui program to calculate the p-value, the null hypothesis is not rejected at 5% significance level for the length of one streak (two tailed p-value= 0.6338; one-tailed p-value= 0.3169). Similarly, it is also insignificant for length of two streaks (two-tailed p-value= 0.8679; one-tailed p-value= 0.434) and three streaks (two-tailed p-value= 1; one-tailed p-value= 0.5), so the null hypothesis is not rejected at 5% significance level as well. The data for three streaks can be omitted and not considered for further analysis due to small sampling bias. Although, the empirical data show positively hot outcome result, but the hypothesis test did not show any significant pattern towards Gambler's fallacy and Hot outcome for goalkeeper data.

4.3 Fallacy Biases Analysis

This analysis is made to check whether the pattern for both fallacies can be useful in the real world. By using the previous data, it can be improved by adding some condition such as the next dive after certain length streak should be correct. Therefore, both the correct and incorrect streaks should be checked and tested. First, Figure 4 shows the percentage of correct prediction dive after certain correct streak length (left or right). The data for 3 streaks is excluded since not adequate for the test.

Figure 4. *Proportion of Goalkeepers dive correctly predicted in the next kick after several sequence streak length of* ***correct prediction*** *(1,2 and 3 streaks)*



Of the total 138 goalkeepers who dove in regards of Gambler's fallacy (Opposite direction) for one length of correct streak prediction, 82 (59%) of them predict correctly. On the other hand, 49 (56%) out of 88 goalkeepers correctly predict the subsequent kick of one streak for the Hot outcome (Same direction). For goalkeeper who correctly guessed two consecutive kicks in the same direction, the number of Gambler's fallacy and Hot outcome are 24 and 10, respectively. Correspondingly, 13(57%) and 4(40%) of goalkeepers managed to predict it right.

To examine whether Gambler's fallacy is independent to Hot outcome or vice versa. Fisher's exact test need to be brought using the correct probabilities that have been obtained. Using 5% of significance level (α), hypothesis testing can be constructed as below,

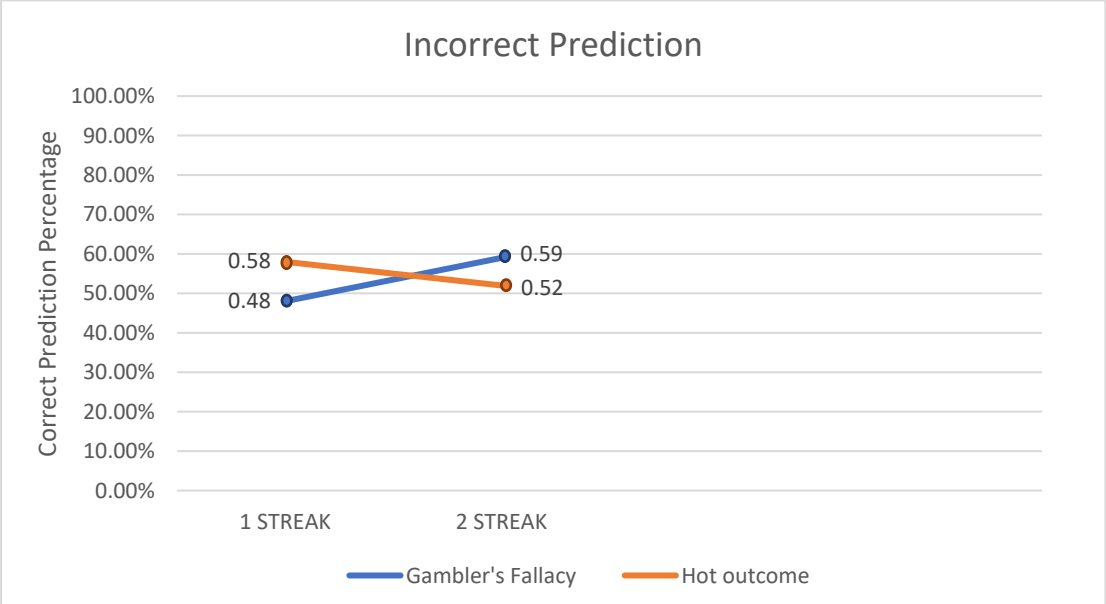
$$H_0: x_1 \text{ and } x_2 \text{ are independent}$$

$$H_1: x_1 \text{ and } x_2 \text{ are dependent}$$

Where x_1 is the probability of Gambler's fallacy and x_2 is the probability of Hot outcome. Using the R-Gui program to calculate the p-value, both the null hypothesis for subsequent kick following one (p-value= 0.5838) and two correct (p-value= 0.4569) prediction are not rejected, since p-value> α. Hence, the test indicates whether both Gambler's fallacy and Hot outcome are independent to each other. This also can be seen from the line graph where both fallacies are reduced as the length of streaks increases.

Same method is being applied to dives made by goalkeeper following subsequent incorrect length of streaks. Figure 5 illustrates the percentage of correct prediction dive after certain incorrect streak length (left or right). Similar to previous data, the streaks of 3 is not tested due to lack of sample size.

Figure 5. *Proportion of Goalkeepers dive correctly predicted in the next kick after several sequence streak length of* ***incorrect prediction*** *(1,2 and 3 streaks)*



Following one incorrect prediction in any direction (left or right), 80 goalkeepers dove based on the Hot outcome (Opposite direction), and 46 (58%) of them predicted it right. Meanwhile, goalkeepers dove according to Gambler's fallacy (Same direction) has accounted for 87, and the number correctly predict reach 42 (48%). For goalkeeper who incorrectly guessed two kicks in a row, the number of Hot outcome (Opposite) and Gambler's fallacy (Same) are 19 and 17, respectively. Correspondingly, 10(52%) and 10(59%) of goalkeepers managed to predict it right.

Same approach to examine whether Gambler's fallacy is independent to Hot outcome or vice versa. Fisher's exact test need to be performed using the correct probabilities that have been obtained from consecutive incorrect prediction. Using 5% of significance level (α), hypothesis testing can be constructed as below,

H_0: x_1 and x_2 are independent

H_1: x_1 and x_2 are dependent

Where x_1 is the probability of Gambler's fallacy and x_2 is the probability of Hot outcome. Using the R-Gui program to calculate the p-value, both the null hypothesis for length of streak following one (p-value =0.8751) and two incorrect (p-value= 0.7486) prediction are not rejected, since p-value> α. Thus, it shows independency between Gambler's fallacy and Hot outcome. However, the line graph shows if it is more favour in Gambler's fallacy than Hot outcome as the streaks increases.

5. Discussion

The objective of this study can be found through the analysis, empirical data and hypothesis test that have been obtained. Recall, the aims are to check the existence of either fallacy and effectiveness in the real world. For a goalkeeper who dove in correct prediction for any length of streaks (1,2 and 3), the empirical evidence shows in favour of Gambler's fallacy as it tends to increase for subsequent streaks compare to Hot outcome. The goalkeepers tend to dive to opposite direction although the kicks are i.i.d (Independent and Identical Distributed) random process between kickers. The binomial test also supported the results, where it shows significantly strong evidence towards Gambler's fallacy. These results were also similar to Reichenbach (1949) research where bettors more likely to shifts their wagers after the same results that appeared frequently in a row on roulette. This is true since the goalkeepers also have a tendency to change direction after suffering from the same outcome in a row. Another previous study regarded to penalty shootout also stated whether the outcome (by considering previous successive kicks) is positively into Gambler's fallacy as well (Misirlisoy & Haggard, 2014). According to Ayton and Fischer (2004), they believed goalkeepers decision on choosing for Gambler's fallacy are naturally based on behaviour and feeling. Therefore, by considering psychology and another factor, some bias might have occurred. For instance, Palacios-Huerta (2002) stated minimax such as prisoner's dilemma from game theory might dwell with the outcome. On the other hand, although the empirical evidence for subsequent incorrect goalkeepers' dives shows Hot outcome (tends to dive to the opposite direction as the streaks increase), however the test did not show any significant towards the Hot outcome. It did not work in the way of Gambler's fallacy as well, which can be expressed more goalkeeper behaviour and psychology affects the outcome. This further analysis should be brought by other professional since this study only aims the objectives. In general, the goalkeeper prediction did not exhibit Hot outcome but Gambler's fallacy does exist for streaks of correct guesses.

Another investigation of whether adopting both fallacies might give a higher chance for the goalkeeper to dive in the right direction. Some outcomes have been attained using Fisher's exact test. For correct prediction from Figure 4, although both Gambler's fallacy and Hot outcome are falling off corresponding to the higher length of streaks, goalkeeper still in dominant favour of Gambler's fallacy since it has higher frequency compare to Hot outcome. Furthermore, after checking the p-value, the results show statistically significant independent between both fallacies, as this might explains how the test results are identical to the empirical evidence. In similar, the incorrect prediction in Figure 5, even though the fallacies show contradiction between each other (In theory when Gambler's fallacy dives increase, Hot outcome dives will decrease since both probabilities will be 1 in total), goalkeepers tend to have more correct prediction in Gambler's fallacy when streak length increases. However, the test also describes there is no dependency between fallacies. Therefore, both outcomes explain following increasing streak length of correct and incorrect guesses in a certain direction (left or right), it does not guarantee whether Gambler's fallacy dives will generate more correct prediction than Hot outcome dives and vice versa. Thus, these fallacies are just a mere illusion which might lead to mistakenly results. Moreover, professional goalkeepers cannot rely too much on these fallacies, since it is not 100% statistically proven better than without following the patterns.

6. Conclusion

In conclusion, for a subsequent correct prediction, goalkeepers tend to dive based on Gambler's fallacy. Whereas, no preference pattern on goalkeeper dives for streak of an incorrect prediction. Hence, the study shows only Gambler's fallacy influence the real-world goalkeeper's decision. Moreover, the are no statistically significant evidence for one pattern is better than another for both subsequent correct and incorrect prediction which means Gambler's fallacy and Hot outcome are just mere fallacies (not always working). Therefore, professional goalkeepers need to adopt certain strategies to predict kick direction such as, combining two fallacies or using one of them only. Nevertheless, there are risks likewise predictable pattern by kickers set to happen as well. Hence, the psychology part also needs to be considerate into the decision-making process. Furthermore, limitation due to small data which results in some error on some outcomes such as, length of 3 streaks is being omitted eventually. Therefore, this study can only determine the pattern up to a maximum length of two streaks and cannot be determined beyond the 2 streaks.

7. Appendix

```
#Left and Right bias between goalkeeper and kicker test
# L is the probability of choosing left
# R is the probability of choosing right

# two sided Ho: L likely to occur as R
# two sided H1: L not likely to occur as R

#Kicker
binom.test (312, 600, p=0.5, alternative=c("two.sided"))
#p-value = 0.3478, not significant so we don't reject Ho

#Goalkeeper
binom.test (315, 600, p=0.5, alternative=c("two.sided"))
#p-value = 0.2364, not significant so we don't reject Ho

### CORRECT PREDICTION
# p is the probability of choosing opposite direction

# one sided   Ho: p<=0.5
# one sided   H1: p>0.5

# two sided   Ho: p equals 0.5
# two sided   H1: p not equals 0.5

binom.test (138, 226, p=0.5, alternative=c("two.sided"))
# significant p-value 0.001071 < 5%(alpha) so reject Ho

binom.test (138, 226, p=0.5, alternative=c("greater”))
# significant p-value 0.0005355 < 5%(alpha) so reject Ho
```

```
binom.test (24, 33, p=0.5, alternative=c("two.sided"))
# not significant p-value 0.01353 < 5%(alpha) so reject Ho

binom.test (24, 33, p=0.5, alternative=c("greater"))
# significant p-value 0.006765 < 5%(alpha) so reject Ho

binom.test (3, 3, p=0.5, alternative=c("two.sided"))
# p-value 0.25 > 5%(alpha), we not reject Ho

binom.test (3, 3, p=0.5, alternative=c("greater"))
# p-value 0.125 > 5%(alpha), we not reject Ho

### INCORRECT PREDICTION
# p is the probability of choosing opposite direction

# one sided   Ho: p>=0.5 for 1 streak
# one sided   H1: p<0.5 for 1 streak

# one sided    Ho: p<=0.5 for 2,3 streaks
# one sided    H1: p>0.5 for 2,3 streaks

# two sided    Ho: p equals 0.5
# two sided    H1: p not equals 0.5

binom.test (52, 110, p=0.5, alternative=c("two.sided"))
# p-value 0.6338> 5%(alpha), not significant so we do not reject Ho
binom.test (52, 110, p=0.5, alternative=c("less"))
# p-value 0.3169> 5%(alpha), not significant we do not reject Ho

binom.test (19, 36, p=0.5, alternative=c("two.sided"))
# p-value 0.8679> 5%(alpha), not significant we do not reject Ho
```

```
binom.test (19, 36, p=0.5, alternative=c("greater"))
# p-value 0.434> 5%(alpha), not significant we do not reject Ho

binom.test (2, 3, p=0.5, alternative=c("two.sided"))
# p-value 1 > 5%(alpha), not significant we do not reject Ho
binom.test (2, 3, p=0.5, alternative=c("greater"))
# p-value 0.5 > 5%(alpha), not significant we do not reject Ho

### Fisher test for CORRECT prediction
# x1 is gambler fallacy, x2 is hot outcome

# Ho: x1 and x2 are independent
# H1: x1 and x2 are dependent

one_streak= matrix (c(82 ,56, 49, 39),nrow=2)
fisher.test(one_streak) # p-value 0.5838>5%(alpha) so we dont reject Ho

two_streak= matrix (c(14 ,10, 4, 6),nrow=2)
fisher.test(two_streak) # p-value 0.4569>5%(alpha) so we dont reject Ho

### Fisher test for INCORRECT prediction
# x1 is gambler fallacy, x2 is hot outcome

# Ho: x1 and x2 are independent
# H1: x1 and x2 are dependent

IC_one_streak= matrix (c(46 ,34, 52, 35), nrow=2)
fisher.test (IC_one_streak) # pvalue 0.8751>5%(alpha) so we do not reject Ho

IC_two_streak= matrix (c(10 ,9, 10, 7), nrow=2)
fisher.test (IC_two_streak) # pvalue 0.7486>5%(alpha) so we do not reject Ho
```

8. References

Apesteguia, J. and Palacios-Huerta, I. (2010) 'Psychological Pressure in Competitive Environments: Evidence from a Randomized Natural Experiment', American Economic Review, 100(5), pp. 2548-2564. (Accessed: 23 November 2018)

Arrondel, L., Duhautois, R. & Laslier, J.F. (2018) 'Decision Under Psychological Pressure: The Shooter's Anxiety at the Penalty Kick', l'Homme et de la Société [Online]. Available at: https://halshs.archives-ouvertes.fr/halshs-01785306/document (Accessed: 2 November 2018)

Ayton, P. & Fischer, I. (2004) 'The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness?', Memory & Cognition, 32(8), pp. 1369-1378. (Accessed: 25 November 2018)

Clotfelter, C.T. & Cook, P.J. (1993) 'The Gambler's Fallacy in Lottery Play', National Bureau of Economic Research [Online]. Available at: https://www.researchgate.net/publication/5191744_The_Gambler's_Fallacy_in_Lottery_Play (Accessed: 25 November 2018)

Croson, R. & Sundali, J. (2005) 'The Gambler's Fallacy and the Hot Hand: Empirical Data from Casinos', The Journal of Risk and Uncertainty, 30(3), pp. 195-209. (Accessed: 9 November 2018)

Dalton, K., Guillon, M. & Naroo, S.A. (2015) 'Importance of the Penalty Kick in Football', Aston University [Online]. Available at: https://research.aston.ac.uk/portal/files/20041282/Penalty_kicks_in_elite_football_post_1997.pdf (Accessed: 2 November 2018)

Edwards, W. (1961) 'Probability Learning in 1000 Trials', Journal of Experimental Psychology, 62(4), pp. 385-394. (Accessed: 26 November 2018)

History of Football - The Origins (2018) [Online] FIFA. Available at: https://www.fifa.com/about-fifa/who-we-are/the-game/index.html (Accessed: 2 November 2018)

Hunter, A.H., Murphy, S.C., Angilletta Jr, M.J. & Wilson, R.S. (2018) 'Anticipating the Direction of Soccer Penalty Shots Depends on the Speed and Technique of the Kick', MDPI [Online]. Available at: https://www.mdpi.com/2075-4663/6/3/73/pdf (Accessed: 22 November 2018)

Kahneman and Miller (1986) 'Norm Theory: Comparing Reality to Its Alternatives', Psychological Review, 93(2), pp. 136-153. (Accessed: 22 November 2017)

Keren, G & Lewis, C (1994) 'The Two Fallacies of Gamblers: Type I and Type II', Organisational Behaviour and Human Decision Processes, 60(1), pp.75-89. (Accessed: 26 November 2018)

Laplace, P.S. (1825) *Philosophical Essay on Probabilities*, New York, Springer-Verlag, Available from:https://books.google.co.kr/books?id=vDZzuGcM4DUC&pg=PA91&dq=Illusions+in+the+Estimation+of+Probabilities,%E2%80%9D&hl=en&sa=X&ved=0ahUKEwjhrvvoiO_eAhUKXLwKHUdFDkMQ6AEIKDAA#v=onepage&q=independent&f=false (Accessed: 25 November 2018)

Laws of the Game 2015/2016 (2016) [Online] FIFA. Available at: https://img.fifa.com/image/upload/datdz0pms85gbnqy4j3k.pdf (Accessed: 2 November 2018)

Misirlisoy, E. & Haggard, P. (2014) 'Asymmetric Predictability and Cognitive Competition in Football Penalty Shootouts', Current Biology, 24(16), pp. 1918-1922. (Accessed: 4 November 2018)

Nesbitt, K. & Barrass, S. (2004) 'Finding Trading Patterns in Stock Market Data', IEEE Computer Graphics and Application, 24(5), pp. 45-55. (Accessed: 4 November 2018)

Palacios-Huerta, I. (2002) 'Professionals Play Minimax: Appendix', Brown University [Online]. Available at: http://www.palacios-huerta.com/docs/PeleAppendix.pdf (Accessed: 4 november 2018)

Reichenbach, H. (1949) *The Theory of Probability*, Berkeley & Los Angeles, University of California Press, Available from: https://archive.org/details/in.ernet.dli.2015.234003/page/n21 (Accessed: 25 November 2018)

Sarkar, A. (2017) 'The Gamblers Fallacy and Hot Outcome: Cognitive Biases or Adaptive Thinking for Goalkeepers Decisions on Dive Direction During Penalty Shootouts', Bowling Green State University [Online]. Available at: https://etd.ohiolink.edu/!etd.send_file?accession=bgsu1483529030818181&disposition=inline (Accessed: 26 November 2018)

Stein, J. & Wapner, L.M. (2017) 'A Coin-Tossing Conundrum', arXiv.org [Online]. Available at: https://arxiv.org/ftp/arxiv/papers/1710/1710.01298.pdf (Accessed: 4 November 2018)

Terrell, D. & Farmer, A. (1996) 'Optimal Betting and Efficiency in Parimutuel Betting Markets with Information Costs', The Economic Journal, 106(437), pp.846-868. (Accessed: 25 November 2018)

Tversky, A. and Gilovich, T. (1989) 'The Cold Facts About the Hot Hand in Basketball', Chance, 2(1), pp. 16-21. (Accessed: 26 November 2018)

Welsh, A. (1990) *The Soccer Goalkeeping Handbook*, London, Bloomsbury Publishing Plc, Available from: https://books.google.co.kr/books?hl=en&lr=&id=ukOkAgAAQBAJ&oi=fnd&pg=PA6&dq=welsh+goalkeeping&ots=1YBa6QA4bi&sig=qVfgfeXISIykw82ztOU33BXFLUo#v=onepage&q=welsh%20goalkeeping&f=false (Accessed: 22 November 2018)

Xu, J. & Harvey, N. (2014) 'Carry on Winning: The Gamblers' Fallacy Creates Hot Hand Effects in Online Gambling', Cognition, 131(2), pp. 173-180. (Accessed: 25 November 2018)

Zeelenberg, M. et al. (2002) 'The Inaction Effect in the Psychology of Regret', Journal of Personality and Social Psychology, 82(3), pp. 314-327. (Accessed: 23 November 2018)