

A Statistical Analysis On Bowling Performance Of Player In Professional Cricket

HarshitaKhangarot¹, Alok Kumar²

Department of Computer Science and Engineering

Institute of Engineering and Technology

JK Lakshmipat University, Jaipur, India

{¹harshitakhangarot, ²alokkumar}@jklu.edu.in

Abstract

Measuring the performance of the bowler according to wickets taken is the natural way of quantifying their performance. Apart from it, the situation at which they scored is the match-winning event. The performance of a bowler could evaluate by determining the ability to take wickets against strong batting line up. In this paper, we present an analysis of the performance of a bowler. Many parameters play a key role in analyzing the bowling performance. These include venue conditions, extras runs, the number of dismissals, type of dismissal of a batsman, especially during power-play and death-over. The testing supports the hypothesis with 95% confidence interval for the first, death and power-play over and also not found any evidence in the data for finding the equal probability for fall of wickets in both the innings.

Keywords: Bowling-performance, death-over, wicket, dismissal, sports analytics.

1. Introduction

Batting as well as bowling performance is vital for the team. The winning probability of the match is dependent on both of them. But with the emergence of the new version of cricket i.e. T20 format, it has become a challenging task for bowlers. It has forced them to think more for the benefit of their team. Average speed, variation in ball speed, age, type of bowler, matches played, country, bowling rate are the major parameters considered for analyzing their performance [7]. The researchers in this field are using various statistical techniques for analyzing the bowler's performance in the game. Generative as well as discriminative methods are used in analyzing the rising stars of the bowling area [1]. Team selection is an upcoming application using a Self-Adapting Intelligent Optimized Analytical Model for evaluating the players [2]. The network analysis technique quantifies the quality of wickets taken by the bowler, using [3]. The physical property of the ball that is the weight of the ball also affects the performance of the fast bowler. Accuracy and bowling mechanics were not adversely affected when using underweight and overweight balls [4]. The factor analysis approach in performance analysis is used for systematic covariation among various dimensions of batting and bowling capabilities [5].

The research paper is concerned with identifying the effects of first over, power play, venue, extra runs, wickets fallen, dismissal on the bowlers' performance.

1.1. Background

'Cricket' also known as 'Gentlemen's game' is a bat and ball team sport played between two teams of eleven players each. The game is played on ground at the center of which is a 22-yard pitch with three wooden stumps at both ends known as wickets. The wickets have also had two bails balanced on them. The game has a fixed pitch size and variable ground size. One of the team act as a batting side in what is called an inning while the other team at the same time act as the bowline or fielding side. The batting side score runs by striking the ball bowled to them using a bat, and the fielding side tries to catch and stop the ball. Team scoring more runs in the specified balls wins the game. The possible scoring options in cricket include runs ranging from one to six and the possible dismissal options include clean bowled, run out, caught, leg before wickets and stumped. The fielding side tries to restrict the batting side to minimum

runs by taking wickets and exhausting the specified balls. A cluster of six such balls is known as an over. The batting side always plays in pair, in which one of the batsmen known as striker stands at the further side of the pitch to the bowler and other known as non-striker stand near the bowler. When ten players have been dismissed, the innings ends and the teams swap roles. The game has three umpires, two on the ground and one in a room with the ability to give decisions with the help of cameras. There is also a match referee in a cricket match.

The game has three popular formats the Twenty20, the One Day International and Test comprising of 20, 50 and unlimited over respectively. Test matches are played for five days, where each team gets batting twice.

2. The Cricket Dataset

The dataset has been considered for analysis is obtained from the cricsheet website. It has ball by ball data of nearly 4000 matches played 2006-2019. This data is processed to create a scorecard file and match the information file. The file comprises parameters such as match date, venue umpire, winner, winning runs/wickets, etc. along with scorecard of each match.

3. Evaluation Parameter

The evaluation of various parameters that will help in determining the probability of finding the evidence in support of the hypothesis. Chi-square testing is used for finding whether the distribution of categorical variables has a relationship between them. If it results in a null hypothesis, means the variables have no dependency. The initial step is to create a contingency table of the known variables and find the χ^2 (chi-square) statistics value. Observed values are the values that we get after the calculations done on the table, which are the experimental values and the other one is the expected values (Eq 1).

$$\chi^2 = \sum (\text{Observed} - \text{Expected})^2 / \text{Expected} \quad (1)$$

4. Results and Discussion

4.1. Effect of First over

The first over of every inning had a crucial rule in the whole match. It acts as motivation if it goes according to the pre-planning of the team. To measure the impact of wickets fallen first over of both the innings are equal in T20 and test cricket. The null hypothesis made is rejected with 0.034 and 0.003. The property of ball changes in the two different innings. In the first over, the maximum wickets fallen are caught as depicted in figure 1.

4.2. Death over and Power play

Death over, the final few overs left in a limited over game. Power-play is the duration of over's in which restriction is set to the fielding positions having only players outside the inner circle. These two sets of over's are the crucial time for both the teams in a match. However, bowlers have a good opportunity to create pressure on the opponent team. Total wickets (WO) per over during these two intervals of time can be calculated as given in eq. 2, where DO is the wickets fallen in o of death over and PO is the wickets fallen in o of power play.

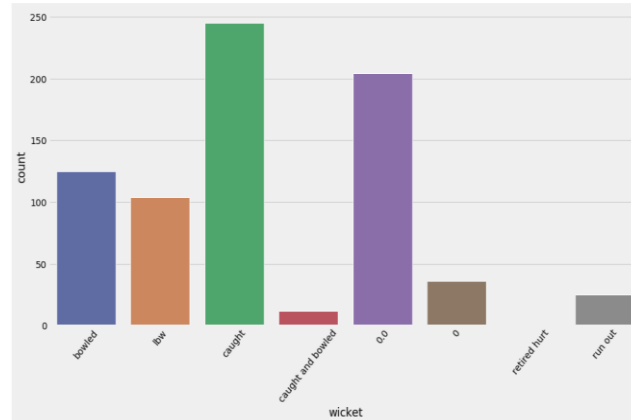


Figure 1. Wicket fell

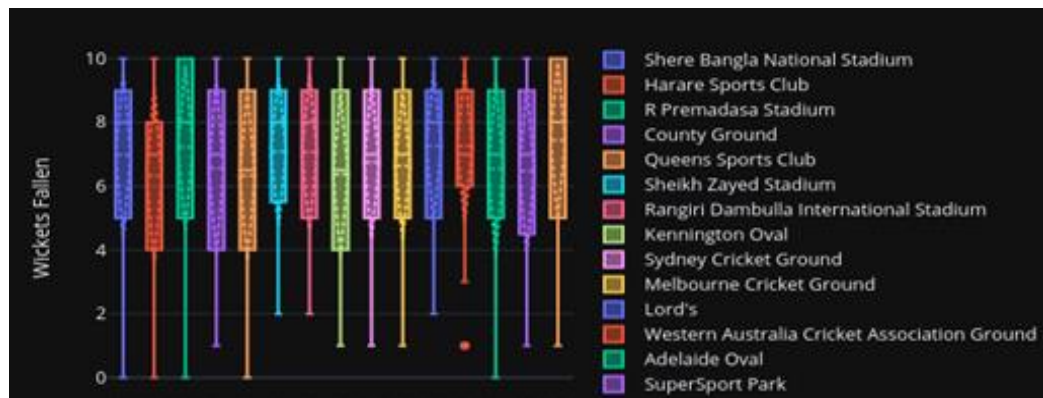


Figure 2. Venue wise wickets fallen

$$WO = \Sigma(DO + PO) \quad (2)$$

It has been assumed here that the death over and power-play has an equal number of wickets fall per over in T20 format. The testing rejects this hypothesis as these two do not have any dependency on equal wickets fallen with 0.0002 p-values.

4.3. Effect of venue

The ground also plays a crucial role in taking wickets by the bowler. If a bowler performs well at one ground, it is expected to perform better next time again [6]. Figure 2 shows that the grounds have nearly 7 wickets fallen in every match for an equal number of matches at that venue. The wickets taken by different bowlers (W_b) are summed up to concerning every venue (V_t).

$$V_t = \Sigma(W_b) \quad (3)$$

4.4. Effect extra runs given

The extra runs given by the bowler benefits the opponent team. So, there is a need to control those runs if a player wants to improve their teams' performance. West Indies gives most extras with the mean extras given equally to 8.18 among the prominent cricket playing nations (fig 3). The stress is one such factor due to which bowlers bowl across the line, as can be seen from figure 4. The more extra runs are gifted at

the end of the game. The probability of extra runs for a team $\rho_{e_r}(t)$ is evaluated using the eq(4) where $\Sigma E_R(t)$ is the total extra runs given by the team to the total runs $\Sigma T_R(t)$.

$$\rho_{e_r}(t) = \Sigma E_R(t) / \Sigma T_R(t) \quad (4)$$

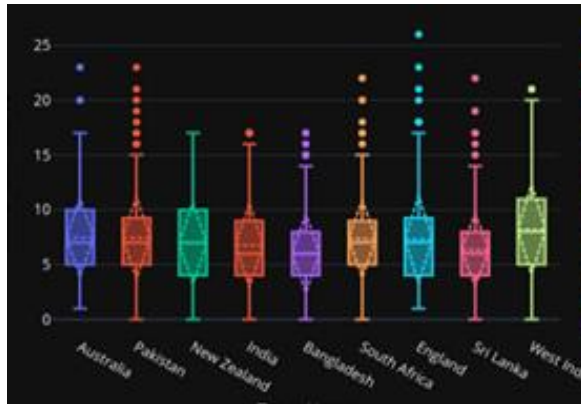


Figure 3. Team wise extra runs

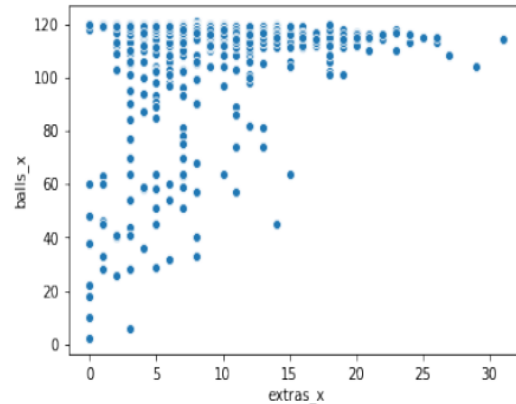


Figure 4. Maximum extras concerning balls

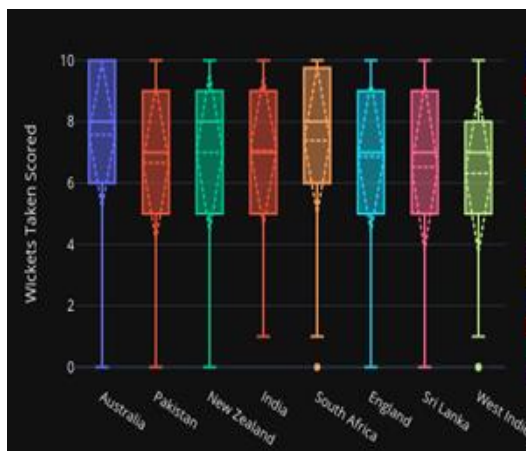


Figure 5. Team-wise wickets taken.

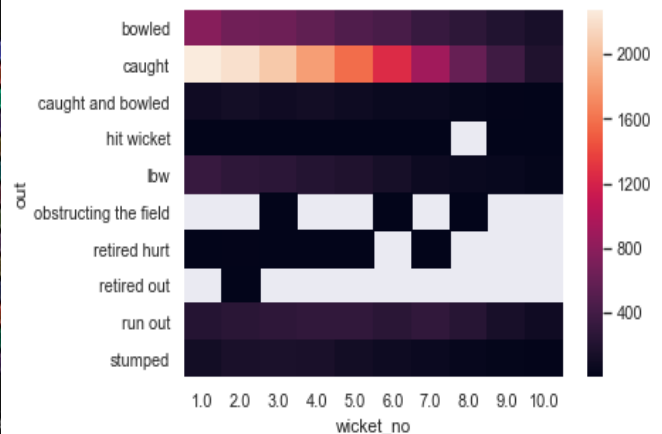


Figure 6. Dismissals across Wicket Number

4.5.Effect of wickets

Wickets fell in a match adversely effects the winning probability of the team. From the box plot of figure 5, it has depicted that Australia has taken the most number of average wickets per match. The different dismissal types majorly considered in cricket are run out, hit wicket, obstructing the field, retired out, stumped, caught, bowled, LBW, caught and bowled. Among these, Caught is the dismissal type which is majorly attempted by the bowler to make the batsmen move out of the ground (fig 6).

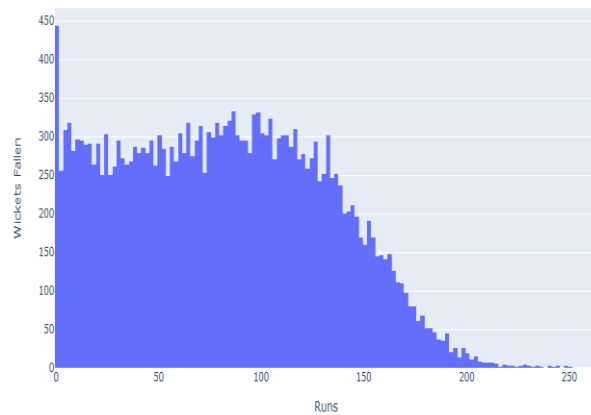


Figure 7. Fall of wicket by runs

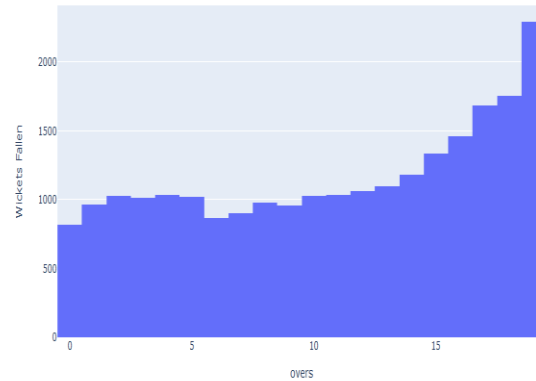


Figure 8. Fall of wickets by over.

The fall of wickets concerning runs shows a negative correlation (figure 7). As the player can score more runs, the possibility of losing the wicket decreases as they will set on the pitch. Whereas wickets fallen concerning the over shows positive correlation (figure 8). The increase in over impacts the performance of the player which creates pressure on the making more and more runs which give benefits to the bowler.

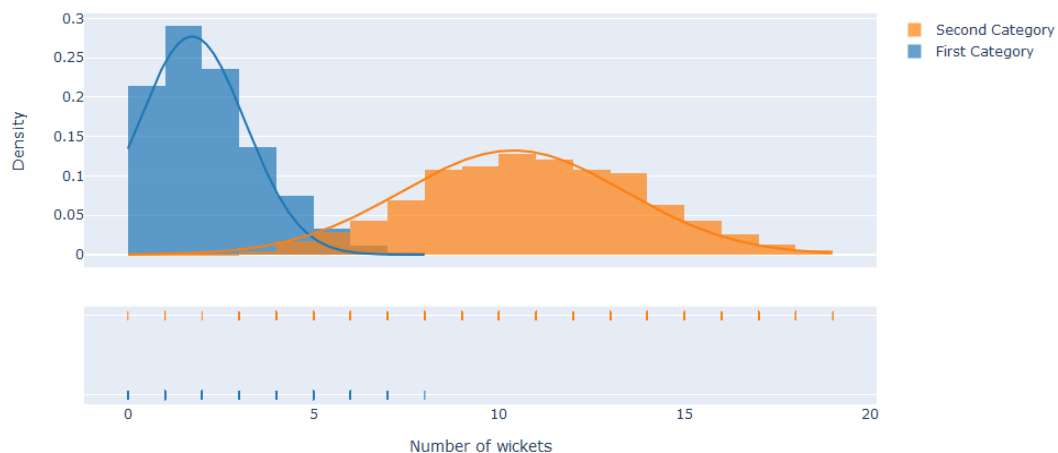


Figure 9: Density of wickets fallen by the first and second category of dismissal methods.(T20)

The dismissal types are categorized into two. The first category includes run out, hit wicket, obstructing the field, retired out, stumped and the second category includes caught, bowled, LBW, caught and bowled. Fig shows the density curve of both of these categories. The curve shows the distribution of density under its area (fig9). The horizontal axis shows the range of the curve, in which we can see the range of the second category is more than the first category. The curve of the second category has symmetrical and the first one is positively skewed behavior.

4.6. Effect of dismissal type

Dismissal of batsman occurs in cricket when its batting duration has come to an end by the bowler by using any of the dismissal methods. After that, the batsman doesn't get the opportunity to score runs in that inning. So, It is a method of controlling the count of runs by the opponent team. In this section, all the hypothesis is made on the T20 and ODI format of cricket. For determining the dependency of the batsman getting bowled and lbw, the mean value of both the methods is determined. The null hypothesis

says that both the mean are equal but is not so, it gets rejected with $1.075e-215$ (T20) and $8.85e-83$ (ODI) values.

H₀:Toss decision affectsthe number of dismissals.

H_A:Toss decision does not affectthe number of dismissals.

Other null hypothesis stats, the toss decisions affect the number of dismissals. The hypothesis gets rejected as it gives a p-value 0.0004. There is no association between the number of dismissals and the toss decision. In choosing either bat or field, Chances of getting caught are maximum for the team. There is a very minute difference seen in both chosen decision (fig 10).

The null hypothesis is assumed regarding wickets fallen in the first and second innings that are tested both for T20 and ODI. In these, different probability combinations at the starting and the end of the match are considered (Table 1). The hypothesis for both T20 and ODI gets rejected with 0.0001 and 0.020 for identifying the equal probability of the first 70per of the first innings to the wickets fallen in the last 30per of the first innings. The wickets fallen probability at different time intervals of a match is tested and the entire null hypothesis gets rejected, which indicates there is no association between them.

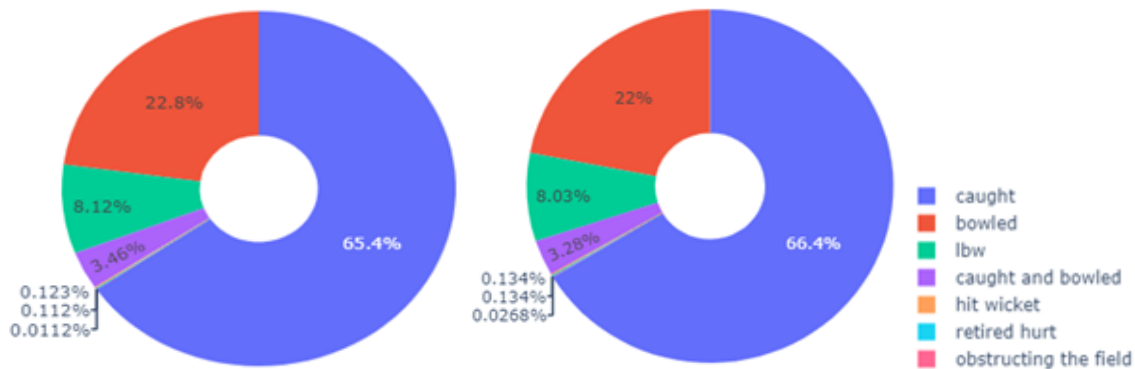


Figure 10. Dismissals by choosing Bat first and choosing field first

Table 1: Null hypothesis results of T20 and ODI format of matches

S.No	Null Hypothesis	T20	Result	ODI	Result
1	Wickets fallen in the first 70per of the first innings is equal to the wickets fallen in the last 30per of the first innings	0.0001	Reject H ₀	0.020	Reject H ₀
2	Wickets fallen in the first 71per of the first innings is equal to the wickets fallen in the last 29per of the first innings	6.923e-16	Reject H ₀	0.25	Accept H ₀
3	Wickets fallen in the first 50per of the first innings is equal to the wickets fallen in the last 50per of the first innings	3.374e-16	Reject H ₀	6.43e-225	Reject H ₀
4	Wickets fallen in the first 10 overs of the first innings is equal to the wickets fallen in the last 10 overs of the first innings	3.322e-188	Reject H ₀	8.92e-289	Reject H ₀
5	An equal probability of wicket by the first category of dismissal and second category of dismissal	3.718e-187	Reject H ₀	0.01	Reject H ₀

5. Conclusion

The dismissal related decisions made acts as a boost for those who are making money by predicting in the betting business. Among all the dismissal types, Caught is found to be the maximum for controlling the

runs scored by the opposition team and making the batsmen leave the ground. Batting team takes advantage of runs when bowlers miss their line of length and decrease the performance graph. Even the fall of wickets by runs and overacts differently, show different properties concerning the correlations. Dismissal of wickets does not show any association with the toss decision taken by the captain. Apart from it, there is a need to identify the dependency of the bowler's performance on the parameters that remain static during the match.

Reference

1. Haseeb Ahmad, Ali Daud, Licheng Wang, Haibo Hong, Hussain Dawood, And Yixian Yang, "Prediction Of Rising Stars In The Game Of Cricket", IEEE., Volume 5, (2017).
2. Sarbani Roy, ParamitaDey, And DebajyotiKundu, "Social Network Analysis Of Cricket Community Using A Composite Distributed Framework: From Implementation Viewpoint", Ieee Transactions On Computational Social Systems, Volume 5, No. 1, March (2018).
3. Satyam Mukherjee, "Quantifying individual performance in Cricket — A network analysis of batsmen and bowlers", Elsevier, Physica A, (2014).
4. K.L. Wickington, N.P. Linthorne, "Effect of Ball Weight on Speed, Accuracy, and Mechanics in Cricket Fast Bowling" Sports,(2017).
5. Bond, Nicholas, "A technique analysis of the fast bowler action in cricket", Cardiff Metropolitan University, MS thesis, (2014).
6. KalpdrumPassi and Niravkumar Pandey, "Increased Prediction Accuracy In The Game Of Cricket Using Machine Learning", International Journal of Data Mining & Knowledge Management Process (IJDMP) Volume 8, No.2, March (2018).
7. Dibyojyoti Bhattacharjee, Darshan G Pahinkar, "Analysis of Performance of Bowlers using Combined Bowling Rate", International Journal of Sports Science and Engineering, Volume 6, (2012).