Predicting Outcome of Live Cricket Match Using Duckworth- Lewis Par Score

Parag Shah

Department of Statistics, H L College of Commerce, Ahmedabad, India

Abstract: **Cricket is the second most watched sport in the world after soccer, and enjoys a multi-million dollar industry. There is remarkable interest in simulating cricket and more importantly in predicting the outcome of cricket match which is played in three formats namely test match, one day international and T20 match. The complex rules prevailing in the game, along with the various natural parameters affecting the outcome of a cricket match present significant challenges for accurate prediction. Several diverse parameters, including but not limited to cricketing skills and performances, match venues and even weather conditions can significantly affect the outcome of a game. There are number of research paper on pre-match prediction of cricket match. Many papers on building a prediction model that takes in historical match data as well as the instantaneous state of a match, and predict match results. We know in the cricket match with shorter version match result keep on changing every ball. So, it is important to predict the outcome of the match on every ball. In this paper, I have developed a model that predicts match result on every ball played. Using Duckworth- Lewis formula match outcome will be predicted for live match. For every ball bowled a probability is calculated and probability figure is plotted. For betting industry this model and the probability figure will be very useful for bettor in deciding which team to on and how much to bet.**

Keywords: **Simulating, Duckworth-Lewis, Prediction Model, Probability Figure, Betting**

I. INTRODUCTION

Cricket was one of the first sports to use statistics as a tool for illustration and comparison. Although compared to other sports, there has not been much statistical modeling work done for cricket.

For baseball, Ganeshapillai and Guttag (2013) developed a prediction model that decides when to change the starting pitcher as the game progresses. It is very much similar to our work-flow, where they used the combination of previous data and in game data to predict a pitchers performance. Tulabandhula and Rudin (2014) were designed a real time prediction and decision system for professional car racing. Model makes the decision of when is the best time for tire change and how many of them.

Wood (1945) used the geometric distribution to model the total score, while Kimber and Hansford (1993) proposed a nonparametric approach based on runs scored for assessing batting performance. The most common and popular form of cricket is the One Day International (ODI), where over 50-overs per side are played. As is typical in games of sport, winning is the ultimate goal. Some studies, (De Silva, 2001), analyze the magnitude of the victory, but most consider the factors affecting winning. There are cases where the magnitude of the victory is important; and, in fact, large sums of money are routinely wagered when it comes to betting on the outcomes of ODI games. Duckworth and Lewis (1998) introduced a technique for revising the target for games that are shortened due to weather interruptions. This method was well received by the cricket-playing community, and it has been using for more than 10 years.

Kaluarachchi and et al (2010) takes into account various factors affecting the game including home team advantage, day/night effect and toss, etc., and uses the Bayesian classifier to predict the outcome of the match. Sankaranarayanan et al (2014) used machine learning approach to predict the result of a one day match depending on the previous data and in game data. Sohail Akhtar and Philip Scarf (2012) have forecasts match outcomes in test cricket in play, session by session. Match outcome probabilities at the start of each session are forecast using a sequence of multinomial logistic regression models. These probabilities can facilitate a team captain or management to consider an aggressive or defensive batting strategy for the coming session. M. J. Bailey & S. R. Clarke (2006) has predicted the match outcome in one day international cricket matches, while the match is in progress. Madan Gopal Jhawar and Vikram Pudi(2016) has predicting the outcome of ODI cricket matches with a team composition based approach.

II. METHODOLOGY

We need to study how the outcome probabilities (of a win, draw, and a loss) vary ball by ball and how the covariate effects vary ball by ball. The covariates fall into two categories, pre-match effects (strengths of teams, a ground effect, home field advantage, outcome of the toss) and in-play effects (score or lead, overs used, overs-remaining, run-rate, and wicket resources). Results indicate that lead has a small effect on the match outcome early on but it dominates later; pre-match team strengths, ground effect and home field advantage are important predictors of a win early on; wicket resources remaining are important throughout a match.

The Duckworth - Lewis method works using the notion that teams have two resources with which to make as many runs as they can - these are the number of overs they have still to receive and the number of wickets they have in hand. From any stage in their innings, their further run-scoring capability depends on both these two resources in combination. The single table gives the percentage of these combined resources that remain for any number of overs left and wickets lost. A ball-by-ball version of the table has also been given by them to enable scorers to deal with instances when play is interrupted mid-over.

In this paper we have defined a ball by ball probability of winning or losing using ball by ball version of Duckworth-Lewis resource table such that if the match was evenly placed, the probability will be 0.5. The ball-by-ball probability could also be plotted graphically to produce a 'probability figure'. This probability figure provides a brilliant instantaneous assessment of how the ODI match went.

From ball by ball data of a particular match we calculate Duckworth-Lewis resources left form its table and from it Par score is calculated. Using this probability is calculated ball by ball.

The formula for calculating probability is given below:

If Runs scored > Par Score, Team Batting wins with probability $= \frac{Runs\ scored - Par\ Score}{Par\ Score}$

If Runs scored < Par Score, Team Bowling wins with probability $= \frac{Par\ Score - Runs\ scored}{Runs\ scored}$

If Runs scored = Par Score, Any team wins with probability = 0.5

III. RESULTS

To explain how probability and its figure is calculated ball by ball, we have considered an Carlton Mid One-Day International Tri-Series, 2nd Match: Australia v India at Melbourne, Jan 18, 2015. India batted first and scored 267 in their innings. So the target for the Australia was 268.

Table1 shows ball by ball probability calculation for 60 balls and similar for rest of the balls can be calculated. We reproduce below the ball by ball probability figure of Australia in this match in Figure 1.

Table 1. Ball by Ball probability of winning.

Balls	Runs	Wicket	Runs	Par Score	Winning Team	Probability
0	0		0			
1	0		0	0	Both	0.5
2	0		0	0.534	Bowling	1
3	0		0	0.801	Bowling	1
4	0		1	1.335	Bowling	0.25
4	1		2	1.602	Batting	0.199
5	0		2	2.136	Bowling	0.06
6	3		5	2.403	Batting	0.5194
7	2		7	2.937	Batting	0.58042857
8	0		7	3.471	Batting	0.50414286
9	1		8	3.738	Batting	0.53275
10	4		12	4.272	Batting	0.644
11	0		13	4.806	Batting	0.63030769
11	4		17	5.073	Batting	0.70158824
12	0		18	5.607	Batting	0.6885
13	0		18	5.874	Batting	0.67366667
14	0		18	6.408	Batting	0.644
15	2		20	6.942	Batting	0.6529
16	0		20	7.209	Batting	0.63955
17	0		20	7.743	Batting	0.61285
18	0		20	8.277	Batting	0.58615
19	4		24	8.811	Batting	0.632875
20	0		25	9.078	Batting	0.63688
20	0		25	9.612	Batting	0.61552
21	4		29	10.146	Batting	0.65013793
22	0		29	10.413	Batting	0.64093103
23	0		29	10.947	Batting	0.62251724
24	0		29	11.481	Batting	0.60410345
25	2		31	12.015	Batting	0.61241935
26	0		31	12.282	Batting	0.60380645
27	0		31	12.816	Batting	0.58658065
28	0		31	13.35	Batting	0.56935484
29	0		31	13.884	Batting	0.55212903

| Balls	Runs	Wicket	Runs	Par Score	Winning Team	Probability
30	1		32	14.418	Batting	0.5494375
31	0		32	14.685	Batting	0.54109375
32	0		32	15.219	Batting	0.52440625
33	1		33	15.753	Batting	0.52263636
34	1		34	16.287	Batting	0.52097059
35	0		34	16.821	Batting	0.50526471
36	0		34	17.355	Batting	0.48955882
37	0		34	17.622	Batting	0.48170588
38	1		35	18.156	Batting	0.48125714
39	1		36	18.69	Batting	0.48083333
40	0		36	19.224	Batting	0.466
41	1		37	19.758	Batting	0.466
42	1		38	20.292	Batting	0.466
43	4		42	20.826	Batting	0.50414286
44	0		42	21.36	Batting	0.49142857
45	1		43	21.894	Batting	0.49083721
46	4		47	22.161	Batting	0.52848936
47	0		47	22.695	Batting	0.51712766
48	2		49	23.229	Batting	0.52593878
49	0		49	23.763	Batting	0.51504082
50	0		49	24.297	Batting	0.50414286
51	0		49	24.831	Batting	0.4932449
52	1		50	25.365	Batting	0.4927
53	0		50	25.899	Batting	0.48202
54	1		51	26.433	Batting	0.48170588
55	0	1	51	26.967	Batting	0.47123529
56	1		52	41.118	Batting	0.20926923
57	0		52	41.652	Batting	0.199
58	0		52	42.186	Batting	0.18873077
59	4		56	42.72	Batting	0.23714286
60	0		56	42.72	Batting	0.23714286
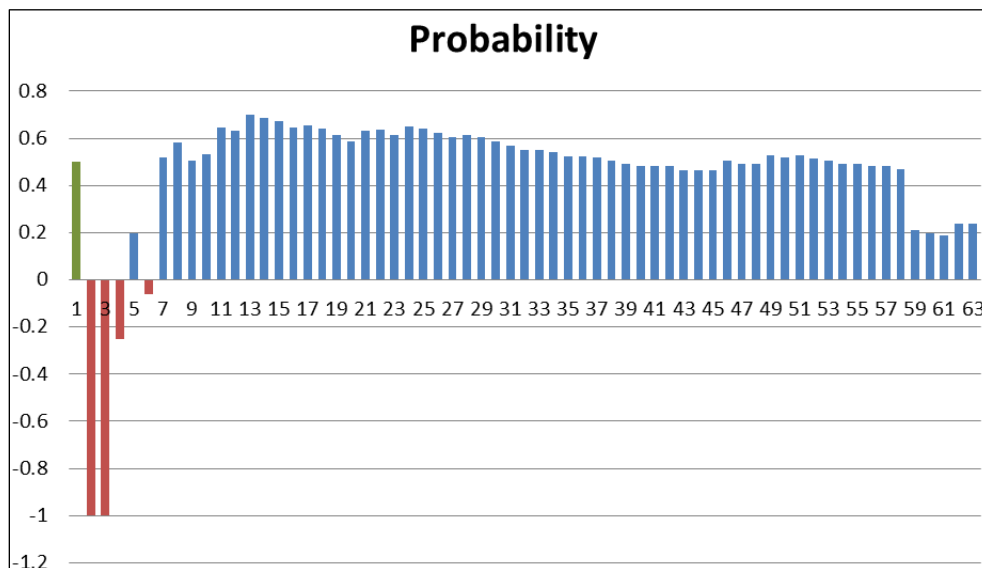
Figure 1. Ball by ball Probability.

For each ball bowled, par score is calculated using Duckworth- Lewis resources used table. Runs scored is compared with the par score and based on that probability is calculated. We can see that for ball number 2, 3, 4 and 6, the runs scored is less than par score. For these balls the team bowling has chance of winning and corresponding probability of bowling team winning is calculated. In figure it is shown in negative side with red colour. For remaining balls team batting has chance of winning and corresponding probability of batting team winning is calculated. In graph it is shown in

positive side with blue colour. If the par score and runs scored are equal we shall consider that both teams have equal chance of winning and probability for each team will be 0.5. In the figure for the first ball it can be seen that probability is 0.5 and shown in green.

IV. CONCLUSION

Using par score concept given by Duckworth & Lewis, probability we calculated provides a clear situation of which team has chance of winning. This probability considers balls faced, balls left, runs scored, runs left, wicket, wickets left. The probability figure perfectly depicts the team having upper hand at every ball bowled in the innings. This will help the team coaches, management to analyze at which situation team needs to improve and can also decide the batting order and bowling changes. If a team bowling had an upper hand in most of the stages of the match but lost at the end, team may need to work on the death overs of the bowling. If the team batting were on losing side in the first stage of the match and then could pull on the victory, team has players who bats well under pressure and also they need to work on the upper order batsmen. It will also be of great help in the betting industry as at every ball, we get a probability of team winning. Amount of bet and team can be decided based on this probability and lot of money can be made.

REFERENCES

[1] Akhtar S and Scarf PA (2012): "Forecasting test cricket match outcomes in play". International Journal of Forecasting, 28(3), 632–643.

[2] Bailey & Clarke (2006): "Predicting the match outcome in one day international cricket matches, while the match is in progress". Journal of Science and Sports Medicine, 5, 480–487.

[3] De Silva, B. M., and Swartz, T. B. (2001): "Estimation of the magnitude of the victory in one-day cricket". Australian & New Zealand Journal of Statistics, 43, 1369-1373.

[4] Duckworth, F. and Lewis, T. (1998): "A fair method for resetting the target in interrupted one-day cricket matches". Journal of Operation Research Society, 49, 22-28.

[5] Ganeshapillai G, Guttag J (2013): "A data-driven method for in game decision making in MLB: When to pull a starting pitcher". In: Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 973–979.

[6] Kaluarachchi, Amal, and S. Varde Aparna(2010): "CricAI: A classification based tool to predict the outcome in ODI cricket". 2010 Fifth International Conference on Information and Automation for Sustainability. IEEE.

[7] Kimber, A. C. and Hansford A. R. (1993): "A statistics analysis of batting in cricket". Journal of the Royal Statistical Society Series A, 156, 443-455.

[8] Madan Gopal Jhawar, Vikram Pudi(2016): "Predicting the Outcome of ODI Cricket Matches: A Team Composition Based Approach". European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD 2016).

[9] Sankaranarayanan VV, Sattar J. and Lakshmanan LVS (2014): "Autoplay - A data mining approach to ODI cricket simulation and prediction". In: Proceedings of the 2014 SIAM International Conference on Data Mining, 1064– 1072.

[10] Theja Tulabandhula and Cynthia Rudin (2014): "Tire Changes, Fresh Air, And Yellow Flags: Challenges in Predictive Analytics For Professional Racing". Massachusetts Institute of Technology, Cambridge, Massachusetts.

[11] Wood, G. H. (1945): "Cricket scores and geometrical progression". Journal of the Royal Statistical Society Series A, 108, 12-22.