Journal of Tourism, Hospitality and Sports www.iiste.org
ISSN (Paper) 2312-5187 ISSN (Online) 2312-5179 An International Peer-reviewed Journal
Vol.5, 2015

Predicting ODI Cricket Result

Parag Shah[1*] Mitesh Shah[2]
1.H L College of Commerce, Ahmedabad 380009, Gujarat, India
2.S V Commerce College, Ahmedabad, Gujarat, India
* E-mail of the corresponding author: pbs1374@yahoo.co.in

Abstract
Result of one Day International (ODI) cricket match depends on various factors related to scoring as well as the strengths of the two teams. While some of these factors have been well analyzed in the literature, others have yet to be investigated. In this paper, statistical significance for a range of variables that could explain the outcome of an ODI cricket match is explored. In particular, home field advantage, game plan (batting first or fielding first), match type (day or day & night), past performance of team will be key interests in our investigation. For purposes of model-building, logistic regression is applied retrospectively to data already obtained from previously played matches.
Keywords: logistic regression, log likelihood, winning margin.

1. Introduction

Cricket began to emerge as a very popular worldwide game around the 13th century. Moreover, cricket was one of the first sports to use statistics as a tool for illustration and comparison. Although dating back to 13th century, compared to sports like baseball, there has not been much statistical modeling work done for cricket. 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. Chedzoy (1997) addressed the issue of the effect of umpiring errors in cricket. 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 (Bailey and Clarke, 2006).

Clarke (1988) used a dynamic programming model to calculate the expected score for games with rain interruptions, so that both teams have the same chance of winning the game. Popular article, Duckworth and Lewis (1998), introduced a technique for revising the target for games that are shortened due to weather interruptions. Factors such as winning the toss and the home team advantage affecting the results of ODI games have been studied in the literature. In this paper, statistical significance for a range of variables that could explain the outcome of an ODI cricket match is explored. In particular, home field advantage, winning the toss, game plan (batting first or fielding first), match type (day or day & night) are key interests in our investigation. Logistic regression is applied to historical data for purposes of model-fitting. Our data consists of a set of the ODIs played between nations for the time period starting from October 2002 to July 2011. Some of the matches were deleted from the analysis due to certain reasons such as abundance of bad weather or when the one team was much superior to the other (ranked teams playing non-ranked teams). Tied games were also deleted from the analysis. Due to the continuous update of cricket rules, we chose, in particular, to use this most recent data and to collect a significant amount of information for 9 years.

2. Data

Variables like Pitch, Avg. First innings Score at this venue, Avg. second innings Score at this venue, ICC points, number of experienced players, previous match results, % of winning in last 5 matches, team batting first and more were considered for 1169 ODI matches from Oct 2002 to July 2011 from www.crickinfo.com.

3. Modeling

We model match outcome at the start of a match. We considered all possible factors affecting the outcome at the start of a match. Team strength, ground effect and home field advantage were found to be important (based on AIC and Nagelkerke R2). For team strength, we have used win percentage differences and the ICC rating differences and find that the rating differences have better explanatory power. This may be because the ICC rating takes account of result (win, draw, loss), along with the win margin, wickets and opponent rating. Winning the toss was also considered in the model fitting but was found to be unimportant. The playing conditions vary from ground to ground and country to country. For example, playing conditions in Wankheda at Mumbai are quite different than in Leeds at Headingley.

Results of fitting the logistic regression model to 1169 match outcomes for various sets of predictors: log-likelihood, AIC and Nagelkerke R2. Covariates here are win percentage difference, the ICC rating difference, home factor, ground effect. We can calculate the probability of win and loss given the position at the start of a

match. This will help team captains and management to consider their batting and bowling strategy for the match. Matches with no results or tie results were deleted. Logistic regression was used to model the outcome of the cricket match.

4. Results

Out of the total 20 variables used for modeling, variables : teams playing, ICC points of each team , previous match result of each team, Day or Night match, Home team or not, team batting first were significant variables for the model.

Prob. =Exp(Y)/ (1+Exp(Y)) = 0.264 - 0603 *Team 1 is not a Home team
-0.8 *Day/ Night
+ 0.851* Team 1 losing previous match
+ 0.978 *Team 1 winning previous match
-0.54 *Team 2 losing previous match
-0.61* Team 2 winning previous match
- 0.764*Team 1 batting first
+0.029* ICC points difference+ 0.101* No. of Exp. Player difference
+0.684*(Day/Night*Team 1 batting first)
+0.646* (Day/Night* Team 1 not a Home team * Team 1 batting first)

Team 1 (+) 1.063 *Aus + 0.931 *Bangladesh - 0.303*Canada + 0.966*Eng + 0.904*Ind + 0.834*Ireland - 0.534*Kenya + 0.457*Netherlands + 1.287*NZ + 0.754*Pak + 0.917* Scotland+1.001*SA+0.879*SL+ 0.219*WI

Team2 (-) 1.015 *Aus-0.253 *Bangladesh+0.531*Canada-0.695*Eng-0.56*Ind - 0.446*Ireland + 0.329*Kenya - 0.19*Netherlands - 0.235*NZ -0.543*Pak + 0.206* Scotland-0.341*SA - 0.413*SL-0.594*WI

The classification of the model is given below:

MODEL ACCURACY

Classification Table

	Actual Result		Predicted		
			Winning team 1		Correct %
			0	1	
	Winning team 1	0	**360**	173	67.5
		1	121	**515**	81
	Overall Percentage				**74.9**

Note:

Model Accurately predicts the results in 74.9% of matches.
Winning Team is predicted winning in 81% of matches

References

Chedzoy O.B., (1997) The effect of umpiring errors in cricket, The Statistician, 46, 4, 529-540.
Clarke, S. R. (1998) Dynamic programming in one-day cricket-optimal scoring rates, Journal the Operational Research Society, 39, 331-337.
De Silva, B. M., and Swartz, T.B. (2001) Estimation of the magnitude of the victory in one-day cricket. Austr. New Zeal. J. Statist., 43, 1369-1373.
De Silva, B. M., and Swartz, T.B. (1997) Winning the coin toss and the home team advantage in one-day international cricket. Austr. New Zeal. J.Statist., 32, 16-22.
Duckworth, F. and Lewis, T. (1998) A fair method for resetting the target in interrupted one-day cricket matches. J. Oper. Res. Soc. , 49, 22-28.
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.
Wood, G.H. (1945), Cricket scores and geometrical progression, Journal of the Royal Statistical Society Series A, 108, 12-22. www.cricket.org.

The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: http://www.iiste.org/journals/ All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

Academic conference: http://www.iiste.org/conference/upcoming-conferences-call-for-paper/

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar