Inefficiencies in the Soccer Betting Market

Owen Pruniskis Akhil Ketkar

April 9, 2015

1 Introduction

Soccer is the most popular sport in the world, particularly in Europe. The popularity of the sport combined with the legal status of gambling has led to the formation of a very deep and active betting market for soccer. A number of different bookmakers allow people to bet large sums of money on various aspects of the game. One can obviously bet on the outcomes of games but also other scenarios like the score at half-time, number of penalties, etc. The objective our project is to find actionable inefficiencies in the soccer betting market.

2 Previous work

A fair amount of work has been done in this area. But it pales in comparison to the amount of research done on American sports, especially baseball. The most important work for the purposes of our project is the Dixon-Coles 1997 paper [1].

3 Methodology

Dixon and Coles propose a bayesian model that uses a bi-variate Poisson distribution for number of goals scored by each team. We will use this model as the basis of our inquiry and improve upon it as needed. In particular we might look at differences in scoring in the first half of a game compared to the second half, the effect of yellow and red cards, the effect of multiple games within the same week, of injuries to certain players, of coaches etc.

We will train the model on available data and then use MCMC methods to sample from the distributions of the model parameters. Finally we will test the model against bookmakers' odds to see if they can be systematically beaten.

4 Data

The data we will use for this project will primary come from the various divisions of English Football focusing on the Premier League which is the top tier of competition. Although one might suspect the betting market in smaller teams to be less efficient. The data comes from the website Football Data [2] which has data for all divisional games for multiple years along with the bookmakers odds for each game.

References

[1] Mark J. Dixon, Stuart G. Coles *Modeling Association Football Scores and Inefficiencies in the Football Betting Market* Applied Statistics, Volume 46, Issue 2 (1997) 265-280

[2] http://www.football-data.co.uk/englandm.php