Profiting from the English Premier League:

Objective Predictive Elicitation, the Kelly Criterion and Black Swans

Andrés Ramírez Hassan*

Mateo Graciano Londoño†

June 28, 2016

Abstract

We present a novel *Objective* Predictive Elicitation approach for a Dirichlet–Categorical model that provides the inputs to define the Kelly criterion for betting in football matches. In particular, we use betting odds to elicit the hyperparameters of our prior distribution, define the optimal bet size based on the Kelly criterion, and establish a stopping rule for betting, based on atypical positive returns (black swans). We apply our methodology to the English Premier League, obtaining profits of 14.14% and 67% in a period of 10 and 2 weeks betting on *Bet365* for the 2013/2014 and 2014/2015 seasons, respectively.

Keywords: Betting Odds, Black Swans, English Premier League, Predictive Elicitation, The Kelly Criterion.

JEL: C11, C53.

*Department of Economics, School of Economics and Finance, Universidad EAFIT, Medellín, Colombia. email: aramir21@eafit.edu.co

†Mathematical Engineering Department, School of Sciences, Universidad EAFIT, Medellín, Colombia. email: mgracian@eafit.edu.co

1 Introduction

Football has gained increasing popularity since its contemporary introduction in 1863 in England. Nowadays, it is the most popular sport around the world [Dunning et al., 1993]. As a consequence of this fact, betting on football through the online bookmarkers is by far the biggest sport in terms of turnover [Finnigan and Nordsted, 2010; Constantinou and Fenton, 2013b]. Therefore, the main objective in this paper is to propose a football betting strategy based on a simple *Objective* Predictive Elicitation approach, the Kelly criterion and atypical positive returns (black swans).

Although most of the literature on football forecasting has focused on various scoring rules to determine the performance of different methodologies [Spann and Skiera, 2009; Constantinou and Fenton, 2012, 2013a], it is natural to determine their forecast accuracy based on the ability to generate profits against market odds [Rue and Salvesen, 2000; Cain et al., 2000; Kuypers, 2000b; Crowder et al., 2002; Goddard and Asimakopoulos, 2004; Dixon and Pope, 2004; Forrest et al., 2005; Graham and Stott, 2008; Hvattum and Arntzen, 2010; Constantinou et al., 2012, 2013]. However, the latter approach depends on the model's forecasting ability relative to the market odds and a betting strategy. We propose to forecast football match outcomes using a simple Predictive Elicitation approach [Kadane, 1980; Garthwaite et al., 2005] based on a novel objective strategy rather than subjective expert knowledge, where the hyperparameters of a Dirichlet–Categorical model are elicited using betting odds from different bookmakers. Regarding the betting strategy, we use the Kelly criterion, which defines the optimal size of a series of bets that maximizes the wealth growth rate in the long run [Kelly, 1956], and a stopping rule based on atypical positive returns (black swans), which are defined based on historical information.

We tested our procedure in the English Premier League for the 2013/2014 and 2014/2015 seasons betting on *Bet365*, one of the most major online bookmaker Štrumbelj [2014b]. Our methodology obtained profitable outcomes: 14.14% and 67% in a period of 10 and 2 weeks, respectively.

The English Premier League is one of the most important football leagues in the world. It was founded in 1992 after the Football League First Division members decided to break away from the Football League, which was originally founded in 1888. The EPL is composed of 20 clubs where each one plays 38 matches in a regular season that runs from August to May, totalling 380 matches. Due to being the most-watched football league in the world, and its excellent available historical records, it is very attractive for testing the ability of betting strategies to generate profits against market odds [Rue and Salvesen, 2000; Goddard and Asimakopoulos, 2004; Constantinou et al., 2012, 2013]. In general, the English Football League System has been considered by many researchers [Cain et al., 2000; Kuypers, 2000b; Crowder et al., 2002; Goddard and Asimakopoulos, 2004; Dixon and Pope, 2004; Forrest et al., 2005]. Some methodologies have been shown to generate returns [Rue and Salvesen, 2000; Cain et al., 2000; Goddard and Asimakopoulos, 2004; Vlastakis et al., 2009; Constantinou et al., 2012, 2013], which sometimes are associated with black swans [Rue and Salvesen, 2000]. However, others have not found that fact [Forrest et al., 2005; Graham and Stott, 2008; Spann and Skiera, 2009; Hvattum and Arntzen, 2010]. Many of these methodologies have a Bayesian flavor [Rue and Salvesen, 2000; Constantinou et al., 2012, 2013] that permits easily introducing prior information from expert knowledge as well as sample information such as the teams' skills [Rue and Salvesen, 2000; Crowder et al., 2002; Goddard and Asimakopoulos, 2004; Forrest et al., 2005]. On the other hand, there are some authors who have preferred to use frequentist approaches, such as Poisson or Negative Binomial models [Cain et al., 2000; Crowder et al., 2002], and Categorical Ordered models [Kuypers, 2000a,b; Goddard and Asimakopoulos, 2004; Graham and Stott, 2008; Hvattum and Arntzen, 2010]. Research proposing betting strategies tries to exploit market inefficiencies, such as favorite longshot or home–away team biases [Cain et al., 2000], where some authors have focused on particular bookmakers such as Intertops [Rue and Salvesen, 2000] and William Hill [Cain et al., 2000; Graham and Stott, 2008].

After this Introduction, we develop in the following section our methodological framework. Section (3) exhibits the results of our application. Lastly, we will make some concluding remarks.

2 Methodology

Our Bayesian approach to predicting the possible outcomes of football matches is the Categorical-Dirichlet model [Constantinou et al., 2013]. In particular, we assume that the likelihood is given by a categorical distribution with three possible outcomes: $\{Win, Draw, Loss\}$. That is,

$$p(x = i) = p_i, \sum_i p_i = 1,$$

where $i = \{Win, Draw, Loss\}$.

On the other hand, prior information is summarized in a Dirichlet distribution such that $\pi(\boldsymbol{p}) \sim \mathcal{D}(\boldsymbol{\alpha})$, that is,

$$\pi(\boldsymbol{p}) = \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1} p_3^{\alpha_3 - 1}. \tag{1}$$

Following Bayes's theorem, the posterior distribution is $\pi(\boldsymbol{p}|Data) \sim \mathcal{D}(\boldsymbol{\alpha} + \boldsymbol{c})$ where $\boldsymbol{c} = (c_1, c_2, c_3)$ is the vector with the number of occurrences of category i, $c_i = \sum_{j=1}^{N} [x_j = i]$.

$$\pi(\boldsymbol{p}|Data) = \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3 + N)}{\Gamma(\alpha_1 + c_1)\Gamma(\alpha_2 + c_2)\Gamma(\alpha_3 + c_3)} p_1^{\alpha_1 + c_1 - 1} p_2^{\alpha_2 + c_2 - 1} p_3^{\alpha_3 + c_3 - 1}. \tag{2}$$

One question that always arises in application of Bayesian analysis is how to obtain the hyperparameters of the prior distributions. In particular, because the concept of probability from a Bayesian point of view is associated with degrees of belief, under this scenario, the experts' knowledge about an event can be tackled from either a subjective or objective perspective. The construction of prior distributions based on the subjective approach should be adopted in scenarios where it is tenable [Berger, 2006]. However, this methodology is strongly influenced by the experts' perception of reality [Garthwaite et al., 2005]; unfortunately, experimental exercises have shown that human beings use heuristic strategies to make statistical statements which lead to biased affirmations [Kahneman, 2011]. Therefore, we use betting odds to obtain the hyperparameters of our prior distributions due to their predictive power [Forrest et al., 2005; Spann and Skiera, 2009; Štrumbelj and Sikonja, 2010] and because bookmakers have financial incentives to correctly assess betting odds [Štrumbelj and Robnik, 2010; Štrumbelj, 2014a]. So, we use a novel *Objective* Predictive Elicitation procedure where we use betting odds to obtain the hyperparameters of the prior Dirichlet distribution. In particular, we take the betting odds associated with the matches from different online bookmakers, and transform them into outcome probabilities using the methodology developed by Shin [1991, 1993] but using the procedure of Jullien and Salanié [1994]. Therefore, we try to mitigate the implicit biases that are present in betting odds using a theoretical framework. The reason for this decision is the well known argument that betting odds are inherently biased due to bookmakers' profits from their service, thus they offer unfair odds. More specifically, the booksum (the sum of the inverse odds) is greater than 100%, and as a consequence, the quoted odds overestimate the probability of every possible outcome [Štrumbelj and Robnik, 2010; Štrumbelj, 2014a]. A common practice is to use a procedure called "basic normalization"; however, Štrumbelj and Robnik [2010] shows that the probabilities deduced from betting odds using the Shin [1993] procedure are more accurate than the probabilities produced by basic normalization. Then, we apply Maximum Likelihood based on Dirichlet models [Thomas and Jacob, 2006; Hizaji and Jernigan, 2009] using the outcome probabilities to estimate the hyperparameters.

Specifically, Shin [1993] develops a model where bookmakers quote odds such that they maximize their expected profit. Therefore, following the methodology proposed by Shin [1993], the probability estimates from the betting odds for a specific match $\theta = \{\theta_{Win}, \theta_{Draw}, \theta_{Loss}\}$ are given by

$$p(z)_i^{betting\ odds} = \frac{\sqrt{z^2 + 4(1-z)\frac{(1/\theta_i)^2}{\sum_l (1/\theta_i)}} - z}{2(1-z)},\ i = \{Win, Draw, Loss\}, \tag{3}$$

so that $\sum_{i=1}^{3} p_i^{betting\ odds} = 1$.

Following Jullien and Salanié [1994], we find the proportion of insider traders z,

$$Arg \min_z \left\{ \sum_{i=1}^{3} p(z)_i^{betting\ odds} - 1 \right\}^2. \tag{4}$$

In addition, we use historical information about every football match to build the categorical likelihood. Thus, we count, for each pair of contenders, their numbers of wins, draws and losses when they met in previous matches.[1] Finally, we obtain the posterior means from the Dirichlet distributions. These are the probabilities associated with the outcomes (*Win*, *Draw* , and *Loss*) for each match.

Once we have designed a forecast methodology, we propose a betting strategy based on the Kelly criterion and black swans. Regarding the former, we follow Hvattum and Arntzen [2010], who use the Kelly criterion to define the optimal size of a series of bets that maximizes the wealth growth rate in the long run based on the following equation [Kelly, 1956]:

$$f^* = \frac{\theta p - 1}{\theta - 1}, \tag{5}$$

where θ is the betting odds from a particular bookmaker, p is the probability of the outcome, which we obtain using our Bayesian procedure for each match outcome, and f^* is the fraction of our budget that we should bet according to this criterion.

Note that there is no restriction on f^*, so when $p < \frac{1}{\theta}$, f^* becomes a negative fraction, but there is no possibility of betting a negative amount of money, so we only bet when $p > \frac{1}{\theta}$.

Regarding a stopping rule for our betting strategy, we follow Rue and Salvesen [2000] who identify that some specific outcomes drive remarkable returns in the football betting market. Therefore, we determine to stop betting when there is a positive return that is higher than four standard deviations of the historical returns associated with betting in the market. We found after some preliminary analysis that betting returns in our betting league (EPL) follow a Gaussian distribution, therefore a positive return higher than four standard deviations has a very small probability, and we refer to them as "black swans."

3 Results

We tested our procedure in the English Premier League for the 2013/2014 and 2014/2015 seasons. First, we gather betting odds associated with the outcome (Win, Draw and Loss) of each

[1] This information is available at http://www.soccerbase.com/.

the 760 matches in these seasons from different online bookmakers, namely: bet365, BetWin, Interwetten, Ladbrokes, Pinnacle Sports, William Hill and Bet home. This information is available at http://www.football-data.co.uk/englandm.php. Second, we implement Shin's procedure to obtain the experts' implicit probabilities associated with each match for each bookmaker. Third, we elicit the hyperparameters of the prior Dirichlet distributions of each football match using Maximum Likelihood with the bookmakers' probabilities obtained in the previous stage. Finally, we calculate the posterior parameters using the hyperparameters, and all the historical records available since 1893 for each match. So, the posterior means are the probabilities for each outcome (p in Equation 5).

We can see the Mean Squared Error and Mean Absolute Error for the 2013/2014 and 2014/2015 seasons in Table 1. In particular, we compare the predictive power of our approach with the probabilities obtained from *Bet365* using Shin's procedure and basic normalization, and a naive forecast that assigns $p = 0.33$ to each possible outcome. As we show in this table, there are no significant differences in the average predictive performance associated with the posterior probabilities, Shin's procedure, and basic normalization. However, these three approaches surpass a naive forecast.

Table 1: Comparing different methodologies.

		Model			
Season	Measure	Posterior	Shin	Normalization	Naive
	MSE	0.63	0.63	0.63	0.81
	MAE	1.02	1.01	1.02	1.33
	MSE	0.71	0.72	0.70	0.81
	MAE	1.15	1.15	1.12	1.33

We implemented our betting strategy on *Bet365*, one of the most major online bookmaker Štrumbelj [2014b]. In particular, we assume an initial capital of US$100, and define the optimal bet size using the Kelly criterion (Equation 5). Remember that each match has three possible choices for betting (Win, Draw and Loss), and we always bet in each of these football matches because we found that $p > \frac{1}{\theta}$ for some specific outcome in each of the 760 matches.

We obtained on average US$124.9, US$98.2 and US$100.7 in the 2013/2014 season using the posterior Dirichlet parameters, and the average probabilities from Shin's procedure, and basic normalization, respectively. Regarding the 2014/2015 season, we obtained US$103.2, US$92.7 and US$62.8, respectively. So, the average outcome of our procedure is better than using Shin's procedure and basic normalization. However, we found a higher risk level, measured through the standard deviation, using our approach. These are US$17.4, US$3.9 and US$10.9 using the posteriors, Shin's procedure, and basic normalization in the 2013/2014 season, and US$17.5, US$2.3 and US$22.9 in the 2014/2015 season, respectively.

Betting over the entire length of the season is not a good strategy because empirical evidence suggests that bookmakers gain predictive power as the season progresses [Peter F. Pope, 1989; Forrest et al., 2005]. Therefore, it is difficult to find market inefficiencies, and as a consequence to overcome the betting odds. So, we define a stopping rule base on the appearance of atypical positive returns (black swans), which have been shown to guide remarkable returns [Rue and Salvesen, 2000]. In particular, we found that there is no statistical evidence to reject the null hypothesis that the percentage returns follow a Gaussian distribution. Therefore, we define a black swan as a positive return higher than four standard deviations, because its probability is approximately zero.

We calculate the standard deviation of the percentage returns of betting in a moving window that is fixed at the opening match, and updated each match day. We stop betting as soon as a black swan occurs, taking into consideration that some matches take place simultaneously. The black swan criterion suggests to stop betting at matches 103 and 26 for the 2013/2014 and 2014/2015 seasons, respectively.

The first black swan happened on the match between Newcastle and Chelsea (02/11/13, 77 days after the opening day), when Newcastle obtained a 2–0 home win, and the second when Stoke City beat Manchester City at the Etihad stadium (30/08/14, 14 days since the opening day). The associated probabilities with the events which actually happened can be seen in Table

2.

Table 2: Black swan probabilities.

Match	Posterior	Shin	Normalization
Newcastle vs. Chelsea	0.1916	0.1642	0.1697
Manchester City vs. Stoke City	0.0902	0.0465	0.0594

The odds for these events in *Bet365* were 6 and 18 respectively, so using the Kelly criterion and the proposed probability we should bet on both events. The Kelly fractions are 2.9% and 3.4% which lead to percentage returns equal to 14.96% ($2.9\% \times (6-1)$) and 57.8% ($3.4\% \times (18-1)$) in these matches, respectively. We should take into consideration that there is a negative fraction using the Shin's probabilities, and very small fractions using the basic normalization procedure (0.3% and 0.4% for each match respectively). This creates remarkable differences in the returns despite the fact that the average predictive performance is similar using these three methodologies. In addition, we did not find any presence of black swans using the Shin and basic normalization probabilities.

After using the betting outcomes since the opening match in our stopping rule, taking into consideration simultaneous matches, our percentage returns are 14.14% and 67% for the 2013/2014 and 2014/2015 seasons respectively.

The differences in the probabilities associated with these matches between the three methodologies is a matter of history. Our approach takes into consideration two sources of information to forecast the probabilities: *Objective* elicitation from experts (bookmakers) and historical records, whereas Shin's procedure and the basic normalization just take into consideration the bookmakers. Seeing the history of these matches in Table 3, the expected result according to history was different from the odds given by *Bet365*. In particular, historical records indicate a Newcastle home win probability equal to 50/153=0.32, whereas Shin's procedure and basic normalization indicate approximately 0.17. The posterior mean (0.19) is higher than the latter due to the influence of historical records. The same pattern is present in the match Manchester City versus Stoke City, where the historical records indicate a probability 36/107=0.33 for the event away win, whereas bookmakers information estimates a probability as high as 0.06 for this event. On the other hand, our Bayesian methodology estimates a probability equal to 0.09 due to sample information. However, we can see from these outcomes that the influence of prior information (bookmakers) exceeds sample information (historical records). This is a desirable characteristic in our approach because betting odds better indicate the present performance of a football team.

Table 3: History and black swans.

Match	Home wins	Draws	Away wins
Newcastle vs. Chelsea	50	39	64
Manchester City vs. Stoke City	47	24	36

4 Concluding Remarks

There is no academic consensus regarding the exploitation of betting market inefficiencies to obtain returns. Some authors have found profitable betting strategies whereas others have not. So, we proposed a novel betting strategy based on *Objective* Predictive Elicitation, the Kelly criterion, and "black swans." The first stage is a forecast methodology based on a very simple natural conjugate family, the Dirichlet–Categorical model, where the hyperparameters are elicited using the betting odds from different bookmakers. The second stage defines an optimal bet size, and the third stage is a stopping rule.

Despite the fact that our methodology does not have a better forecast performance than that of Shin's procedure or basic normalization, it obtains profitable outcomes due to its use of expert knowledge and sample information. This strategy allows identifying differences between these sources of information that can be exploited to obtain good betting returns in specific matches. In particular, with this simple framework we obtained profitable outcomes, 14.14% and 67% in a period of 10 and 2 weeks, betting on *Bet365* in the English Premier League for the 2013/2014 and 2014/2015 seasons. However, a future research agenda should consider other betting markets.

References

Berger, J. (2006). The case for objective bayesian analysis. *Bayesian Analysis*, 1(3):385–402.

Cain, M., Law, D., and Peel, D. (2000). The favourite–longshot bias and market efficiency in UK football betting. *Scottish Journal of Political Economy*, 47:25–36.

Constantinou, A. and Fenton, N. (2012). Solving the problem of inadequate scoring rules for assessing probabilistic football forecast models. *Journal of Quantitative Analysis in Sports*, 8(1):Article 1.

Constantinou, A., Fenton, N., and Neil, M. (2012). pi-football: a Bayesian network model for forecasting Association Football match outcomes. *Knowledge-Based Systems*, 36:322–339.

Constantinou, A. C. and Fenton, N. E. (2013a). Determining the level of ability of football teams by dynamic ratings based on the relative discrepancies in scores between adversaries. *Journal of Quantitative Analysis in Sports*, 9(1):37–50.

Constantinou, A. C. and Fenton, N. E. (2013b). Profiting from arbitrage and odds biases of the european football gambling market. *The Journal of Gambling Business and Economics*, 7:41–70.

Constantinou, A. C., Fenton, N. E., and Martin, N. (2013). Profiting from an inefficient association football gambling market: Prediction, risk and uncertainty using bayesian networks. *Knowledge-Based Systems*, 50:60–86.

Crowder, M., Dixon, M., Ledford, A., and Robinson, M. (2002). Dynamic modelling and prediction of english football league matches for betting. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 51(2):157–168.

Dixon, M. and Pope, P. (2004). The value of statistical forecasts in the UK association football betting market. *International Journal of Forecasting*, 20:697–711.

Dunning, E., Joseph, R., and Maguire, A. (1993). *The Sports Process: A Comparative and Developmental Approach*. Human Kinetics Publishers.

Finnigan, M. and Nordsted, P. (2010). *The Premier Football Betting Handbook 2010/11*. Harriman House, Hampshire, Great Britain.

Forrest, D., Goddard, J., and Simmons, R. (2005). Odds-setters as forecasters: the case of english football. *International Journal of Forecasting*, 21:551–564.

Garthwaite, P., Kadane, J., and O'Hagan, A. (2005). Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, 100(470):680–701.

Goddard, J. and Asimakopoulos, I. (2004). Forecasting football results and the efficiency of fixedâodds betting. *Journal of Forecasting*, 23(1):51–66.

Graham, I. and Stott, H. (2008). Predicting bookmaker odds and efficiency for uk football. *Applied Economics*, 40:99–109.

Hizaji, R. and Jernigan, R. (2009). Modelling compositional data using Dirichlet regression models. *Journal of Applied Statistics & Probability*, 4:77–91.

Hvattum, L. and Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. *International Journal of Forecasting*, 26:460–470.

Jullien, B. and Salanié, B. (1994). Measuring the incidence of insiders trading: A comment on Shin. *The Economic Journal*, 104(427):1418–1419.

Kadane, J. B. (1980). Predictive and structural methods for eliciting prior distributions. In Zellner, A., editor, *Bayesian Analysis in Econometrics and Statistics: Essays in honor of Harold Jeffreys,*, pages 89–93. North–Holland Publishing Company,, Amsterdam.

Kahneman, D. (2011). *Thinking Fast and Slow*. Farrar Straus Giroux.

Kelly, J. L. (1956). A new interpretation of information rate. *Bell System Technical Journal*, 35(4):917–926.

Kuypers, T. (2000a). Balance in competition in dutch soccer. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 49(3):419–431.

Kuypers, T. (2000b). Information and efficiency: an empirical study of a fixed odds betting market. *Applied Economics*, 32(11):1353–1363.

Peter F. Pope, D. A. P. (1989). Information, prices and efficiency in a fixed-odds betting market. *Economica*, 56(223):323–341.

Rue, H. and Salvesen, O. (2000). Prediction and retrospective analysis of soccer matches in a league. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 49(3):399–418.

Shin, H. (1991). Optimal betting odds against insider trading. *The Economic Journal*, 101:1179–1185.

Shin, H. (1993). Measuring the incidence of insiders trading in a market for state-contingent claims. *The Economic Journal*, 103(420):1141–1153.

Spann, M. and Skiera, B. (2009). Sports forecasting: a comparison of the forecast accuracy of prediction markets, betting odds and tipsters. *Journal of Forecasting*, 28(1):55–72.

Štrumbelj, E. (2014a). A comment on the bias of probabilities derived from bookmaker odds and their use in measuring outcome uncertainty. *Journal of Sport Economics*, pages 1–15.

Štrumbelj, E. (2014b). On determining probability forecasts from betting odds. *International Journal of Forecasting*, 30(4):934–943.

Štrumbelj, E. and Robnik, M. (2010). Online bookmakers' odds as forecasts: The case of European football league. *International Journal of Forecasting*, 26:482–488.

Štrumbelj, E. and Sikonja, M. R. (2010). Online bookmakers' odds as forecasts: The case of European soccer leagues. *International Journal of Forecasting*, 26(3):482–488.

Thomas, S. and Jacob, J. (2006). A generalized Dirichlet model. *Statistics and Probability Letters*, 76:1761–1767.

Vlastakis, N., Dotsis, G., and Markellos, R. N. (2009). How efficient is the European football betting market? Evidence from arbitrage and trading strategies. *Journal of Forecasting*, 28(5):426–444.