



- [Enter to Win the Bookdown Contest](#) - Deadline Oct 1 2018 - Yihui Xie
- [Announcing RStudio Package Manager Beta](#)



A brand new product from RStudio to organize and centralize R packages across your team, department, or entire organization.

Bookdown contest submission: Odds Modelling and Testing Inefficiency of Sports Bookmakers



englianhu

6 3h

Title : Odds Modelling and Testing Inefficiency of Sports Bookmakers

by [@y0r, Lian Hu](#)



Abstract

I learn R language since decade and finish my first research which was a backtest on the soccer betting. Today I tried to summarise it. Below two links are my research and submitted for the contest.

- [Odds Modelling and Testing Inefficiency of Sports Bookmakers](#) 4
- [Application of Kelly Criterion model in Sportsbook Investment](#) 1

1. Data

1.1 Odds Modelling and Testing Inefficiency of Sports Bookmakers

I collected the pre-match odds price (open and close price) of 1x2 of 40 bookmakers, 29 bookmakers among them have Asian-Handicap and Over-Under via soccer betting information website.

1.2 Application of Kelly Criterion model in Sportsbook Investment

I collected the Asian-Handicap and Over-Under odds price (whole price movement from open price until close price) of 13 bookmakers.

2. Odds Modelling

I used the *Dimitris Karlis and Ioannis Ntzoufras (2005)* and then fit the weighted function from *Mark Dixon and Stuart Coles (1996)* into it. All research are using Markov-Chain method to test every single match in 3 years soccer seasons. However, my model using 1st season to model for 2nd season, and then based on the Markov basic model in 2nd season to measure the weighted function, only the 3rd season started a weighted model, therefore 2 soccer season dataset required in order to predict coming soccer match.

There are 4 soccer seasons dataset and 1st and 2nd seasons is the datasets for modelling. The staking model started from 3rd season to 4th season.

2.1 Weighted Function

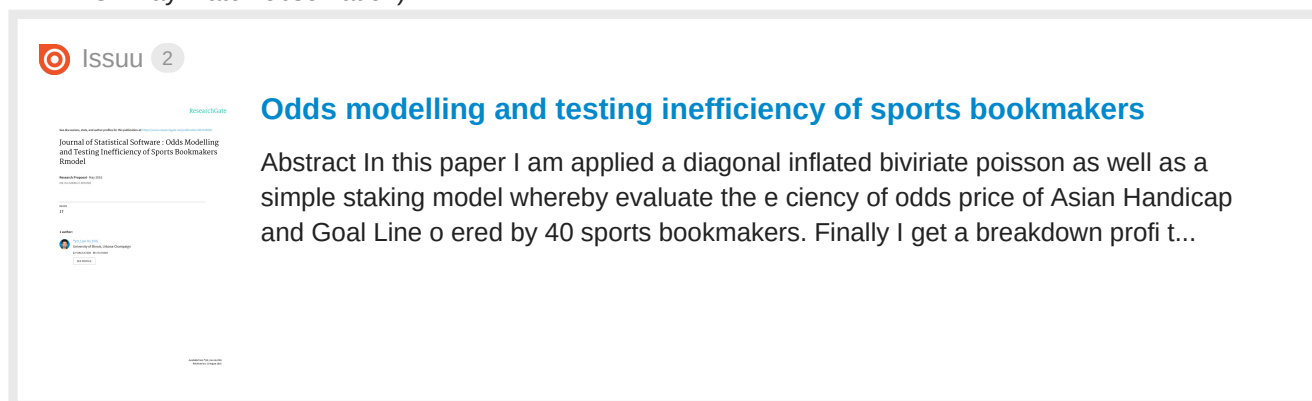
There has few adjustment onto the model.

- A constant weighted function over over whole moving windows for a soccer season based on last soccer season
- A weekly constant Markov-Chain weighted function
- A daily constant Markov-Chain weighted function
- An observation dynamic weighed function

2.2 Data Constrant

Secondly, due to constrant restricted on to the model, I tried to compare (Might similar with hidden factor for the Markov Chain).

- All observation weighted function (for example, a soccer team played with another team more than once within a soccer season, all soccer matches will be counted)
- Unique observation weighed function (for example, a soccer team played with another team more than once within a soccer season, only latest soccer matches will be counted. There will be no any duplicated Home vs Away match observation)



3. Model Comparison

Finally the model with below adjustment came with highest ROI.

- An observation dynamic weighed function
- Unique observation weighed function (for example, a soccer team played with another team more than once within a soccer season, only latest soccer matches will be counted. There will be no any duplicated Home vs Away match observation)

4. Betting Strategy

4.1 Odds Price

- probability with vigorish (directly convert from the odds price to probability)
- probability without vigorish 1 (the vigorish part divided to 2 to be 50% for both home and team)
- probability without vigorish 2 (the vigorish part divided based on the portion of odds price)

4.2 Betting Strategy

I took below probability as the baseline of bookmakers. After that applied a full Kelly Criterion betting model.

- probability with vigorish (directly convert from the odds price to probability)
- probability without vigorish 2 (the vigorish part divided based on the portion of odds price)

5. Profit and Loss

Due to the paper above using the normal betting strategy by refer to *Mark Dixon and Stuart Coles (1996)* is not profitable, here I try to use Kelly Criterion (staking on 13 bookmakers) and the ROI was outperformed.

No	Category	With Spreads (2011/12)	Ratio (%)	Without Spreads (2011/12)	Ratio (%)	With Spreads (2012/13)	Ratio (%)	Without Spreads (2012/13)	Ratio (%)
1	No of Matches	4,896	1	4,896	1	5,514	1	5,514	1
2	Total PL	\$353.96966	55.57%	\$381.06299	53.98%	\$448.8993	59.35%	\$488.91841	58.60%
3	No of Bets	2,268	46.32%	2,404	49.10%	2,570	46.61%	2,697	48.91%
4	No of Won Bets	1,531	31.27%	1,584	32.35%	1,765	32.01%	1,824	33.08%
5	No of Voided Bets	128	2.61%	143	2.92%	192	3.48%	196	3.55%
6	No of Lose Bets	609	12.44%	677	13.83%	613	11.12%	677	12.28%
7	Staking	\$636.98372	1	\$705.89203	1	\$756.2979	1	\$834.32032	1
8	Won Bets Stakes	\$453.43724	71.19%	\$496.09555	70.28%	\$563.6685	74.53%	\$614.24795	73.62%
9	Voided Stakes	\$19.13296	3.00%	\$22.67241	3.21%	\$27.1151	3.59%	\$32.12999	3.85%
10	Lose Bets Stakes	-\$99.46758	-15.62%	-\$115.03256	-16.30%	-\$114.7691	-15.18%	-\$125.32954	-15.02%

table 4.1 : Staking breakdown and result of the bets settlement.

Company	PL (2011/12)	Ratio (%)	PL2 (2011/12)	Ratio (%)	PL (2012/13)	Ratio (%)	PL2 (2012/13)	Ratio (%)
Ladbrokes	\$33.772411	9.54%	\$38.25184	10.04%	\$44.53507895	9.92%	\$46.7763362	9.57%
Bet365	\$34.120624	9.64%	\$37.19263	9.76%	\$33.53743752	7.47%	\$40.6766948	8.32%
Macauslot	\$35.740062	10.10%	\$40.23454	10.56%	\$1.76408658	0.39%	\$1.9329171	0.40%
X10Bet	\$37.538487	10.61%	\$41.64034	10.93%	\$33.62892077	7.49%	\$40.6921976	8.32%
X188Bet	\$36.579289	10.33%	\$38.25589	10.04%	\$41.05668234	9.15%	\$46.1156378	9.43%
SBOBET	\$40.392461	11.41%	\$41.90898	11.00%	\$43.50915478	9.69%	\$47.8308212	9.78%
Mansion88	\$31.219547	8.82%	\$32.38999	8.50%	\$42.44403404	9.46%	\$43.9521665	8.99%
YSB88	\$13.167746	3.72%	\$14.34128	3.76%	\$45.92688667	10.23%	\$46.8393963	9.58%
X12BET	\$36.802466	10.40%	\$38.19015	10.02%	\$36.03065656	8.03%	\$36.2449956	7.41%

VictorChandler	\$24.391917	6.89%	\$25.95763	6.81%	\$45.13420638	10.05%	\$46.2154620	9.45%
Canbet	\$10.347393	2.92%	\$10.97516	2.88%	\$41.04019224	9.14%	\$46.9617659	9.61%
Betinternet	\$10.286812	2.91%	\$11.18731	2.94%	\$40.20594752	8.96%	\$44.5435274	9.11%
Titanbet	\$9.610441	2.72%	\$10.53726	2.77%	\$0.08604952	0.02%	\$0.1364944	0.03%

table 4.2 : Breakdown of Operators - Profit & Loss on the Odds Price with/without Overrounds.

6. Conclusion

Finally, the model is proof that profitable. However the odds price was collected from the information website but the from bookmakers website.

Niko Marttinen (2006) introduced a multinomial model which is more accurate than the bivariate poisson model. The bivariate poisson model by *Niko Marttinen (2006)* added the bookmakers odds price as the hidden effects where I need to upgraded beyond the future.

7. Reference

1. [Modelling association football scores 1982 by M.J Maher](#)
2. [Modelling Association Football Scores and Inefficiencies in the Football Betting Market. 1996 by Mark Dixon and Stuart Coles](#)
3. [A Birth Process Model for Association Football Matches. 1997 by Mark Dixon and Michael Robinson](#)
4. [Dynamic Modelling and Prediction of English Football League Matches for Betting. 2002 by Martin Crowder, Mark Dixon, Anthony Ledford and Mike Robinson](#)
5. [The value of statistical forecasts in the UK association football betting market. 2004 by Mark Dixon and Peter Pope](#)
6. [Statistical Modelling for Soccer Games: The Greek League. 1998 by Dimitris Karlis and Ioannis Ntzoufras](#)
7. [Bayesian modelling of football outcomes \(using Skellam's Distribution\). 2007 by Dimitris Karlis and Ioannis Ntzoufras](#)
8. [Bivariate Poisson and Diagonal Inflated Bivariate Poisson Regression Models in R. 2005 by Dimitris Karlis and Ioannis Ntzoufras](#)
9. [John Goddard and Ioannis Asimakopoulos 2004 by John Goddard and Ioannis Asimakopoulos](#)
10. [Statistical Methodology for Profitable Sports Gambling 2012 by Fabián Enrique Moya](#)
11. [Creating a Profitable Betting Strategy for Football by Using Statistical Modelling by Niko Marttinen \(2006\)](#)

Powered by - Copyright® Intellectual Property Rights of  Scibrokes® 個人の経営企業



[Reply](#)

Suggested Topics

Topic	Category	Replies	Activity
-------	----------	---------	----------

Different output for Notebook and knitted html (flexdashboard, tabset) rmarkdown	R Markdown	3	Apr 10
Is there a place we can put non-blog files (pdf files) in blogdown? blogdown	R Markdown	4	Jun 27

Want to read more? Browse other topics in [R Markdown](#) or [view latest topics](#).