The Market for EPL Odds

Guanhao Feng

Booth School of Business, University of Chicago

R/Finance 2017

(Joint work with Nicholas Polson and Jianeng Xu)
Motivation Model Application

Soccermatics from David Sumpter

HOW I USED MATHS TO BEAT THE BOOKIES

A mathematician on how he made a 200% return from betting on football
Underdog Leicester defeated odds of 5000-1 to win the Premier League
The Soccer Betting Market

- In 2013, gambling on soccer is a global industry worth anywhere between $700 billion and $1 trillion a year.

- Now, odds (fixed-odd betting) are set via online and updated in real time during the game (Betfair, Bet365, Ladbrokes).

- Bets can be placed on any outcome from the game.

- Some investment firms have developed their sports trading team.
**Everton vs West Ham (March 5th, 2016)**

Original odds data from Ladbrokes before the game:

<table>
<thead>
<tr>
<th>Home \ Away</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>11/1</td>
<td>12/1</td>
<td>28/1</td>
<td>66/1</td>
<td>200/1</td>
<td>450/1</td>
</tr>
<tr>
<td>1</td>
<td>13/2</td>
<td>6/1</td>
<td>14/1</td>
<td>40/1</td>
<td>100/1</td>
<td>350/1</td>
</tr>
<tr>
<td>2</td>
<td>7/1</td>
<td>7/1</td>
<td>14/1</td>
<td>40/1</td>
<td>125/1</td>
<td>225/1</td>
</tr>
<tr>
<td>3</td>
<td>11/1</td>
<td>11/1</td>
<td>20/1</td>
<td>50/1</td>
<td>125/1</td>
<td>275/1</td>
</tr>
<tr>
<td>4</td>
<td>22/1</td>
<td>22/1</td>
<td>40/1</td>
<td>100/1</td>
<td>250/1</td>
<td>500/1</td>
</tr>
<tr>
<td>5</td>
<td>50/0</td>
<td>50/1</td>
<td>90/1</td>
<td>150/1</td>
<td>400/1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>100/1</td>
<td>100/1</td>
<td>200/1</td>
<td>250/1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>250/1</td>
<td>275/1</td>
<td>375/1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>325/1</td>
<td>475/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


What we can learn from the bookies?

▶ We provide a method for calibrating real-time market odds for the evolution of score difference for a soccer game.

▶ We rely on the odds market efficiency to calibrate a probability model and provide the market forecast of the final result.

▶ We provide an interpretation of the betting market and how it reveals the market expectation changes during the game.

▶ For future research, can we really beat the market (bookies)?
What we do?

- The Black-Scholes formula is a continuous-time model that describes the price of the option over time.

- We use a discrete-time probabilistic model to describe the evolution of score differences: Skellam process.

- Ex-ante, we show how to calibrate expected goal scoring rates using market-based odds information during the game.

- As the game evolves, we use the updated market odds to re-estimate the model.
We decompose the scores of each team as

\[
\begin{align*}
N_A(t) &= W_A(t) + W(t) \\
N_B(t) &= W_B(t) + W(t)
\end{align*}
\]

where \(W_A(t), \ W_B(t)\) and \(W(t)\) are independent processes with \(W_A(t) \sim \text{Poisson}(\lambda_A t)\), \(W_B(t) \sim \text{Poisson}(\lambda_B t)\).

Here \(W(t)\) is a non-negative integer-valued process to induce a correlation between the numbers of goals scored.

By modeling the score difference, we avoid having to specify the distribution of \(W(t)\) as the score difference is independent of \(W(t)\).
We define a Skellam process

\[ N(t) = N_A(t) - N_B(t) = W_A(t) - W_B(t) \sim Skellam(\lambda^A t, \lambda^B t), \]

where \( \lambda^A t \) is the cumulative expected scoring rate on the interval \([0, t]\).

At time \( t \), the conditional distributions for scores are

\[
\begin{align*}
W_A(1) - W_A(t) &\sim Poisson(\lambda^A (1 - t)) \\
W_B(1) - W_B(t) &\sim Poisson(\lambda^B (1 - t))
\end{align*}
\]

The conditional distribution for score difference is

\[ N(1) - N(t) \sim Skellam(\lambda^A (1 - t), \lambda^B (1 - t)). \]
Conditional Probability Calculation

- Specifically, conditioning on $N(t) = \ell$, we have the identity

$$N(1) = \ell + Skellam(\lambda_t^A, \lambda_t^B)$$

- The probability of home team A winning at time $t$ is

$$\mathbb{P}(N(1) > 0|\lambda_t^A, \lambda_t^B, N(t) = \ell) = \mathbb{P}(Skellam(\lambda_t^A, \lambda_t^B) > -\ell|\lambda_t^A, \lambda_t^B)$$

- The probability of a draw at time $t$ is

$$\mathbb{P}(N(1) = 0|\lambda_t^A, \lambda_t^B, N(t) = \ell) = \mathbb{P}(Skellam(\lambda_t^A, \lambda_t^B) = -\ell|\lambda_t^A, \lambda_t^B)$$
Dynamic Model Calibration

Use real-time market odds to calibrate parameters $\lambda^A_t$ and $\lambda^B_t$.

▶ Convert odds ratios to the implied probabilities of final scores

$$P(N_A(1) = i, N_B(1) = j) = \frac{1}{1 + \text{odds}(i, j)}.$$

▶ The unconditional moments are given by

$$E[N(1)] = E[W_A(1)] - E[W_B(1)] = \lambda^A - \lambda^B,$$

$$V[N(1)] = V[W_A(1)] + V[W_B(1)] = \lambda^A + \lambda^B.$$

▶ The conditional moments are given by

$$\begin{cases} 
E[N(1)|N(t) = \ell] = \ell + (\lambda^A_t - \lambda^B_t), \\
V[N(1)|N(t) = \ell] = \lambda^A_t + \lambda^B_t.
\end{cases}$$
Time-Varying Discussion

▶ Our approach: re-estimate \( \{ \lambda^A_t, \lambda^B_t \} \) dynamically through the real-time updated market odds.

▶ An alternative approach to time-varying \( \{ \lambda^A_t, \lambda^B_t \} \) is to use a Skellam regression with conditioning information such as possession percentages, shots (on goal), corner kicks, yellow cards, red cards, etc.

▶ We would expect jumps in the \( \{ \lambda^A_t, \lambda^B_t \} \) during the game when some important events happen. A typical structure takes the form

\[
\begin{align*}
\log(\lambda^A_t) &= \alpha_A + \beta_A X_{A,t-1} \\
\log(\lambda^B_t) &= \alpha_B + \beta_B X_{B,t-1},
\end{align*}
\]

(1)

estimated using standard log-linear regression.
We use the market information for score difference.

<table>
<thead>
<tr>
<th>Home \ Away</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>11/1</td>
<td>12/1</td>
<td>28/1</td>
<td>66/1</td>
<td>200/1</td>
<td>450/1</td>
</tr>
<tr>
<td>1</td>
<td>13/2</td>
<td>6/1</td>
<td>14/1</td>
<td>40/1</td>
<td>100/1</td>
<td>350/1</td>
</tr>
<tr>
<td>2</td>
<td>7/1</td>
<td>7/1</td>
<td>14/1</td>
<td>40/1</td>
<td>125/1</td>
<td>225/1</td>
</tr>
<tr>
<td>3</td>
<td>11/1</td>
<td>11/1</td>
<td>20/1</td>
<td>50/1</td>
<td>125/1</td>
<td>275/1</td>
</tr>
<tr>
<td>4</td>
<td>22/1</td>
<td>22/1</td>
<td>40/1</td>
<td>100/1</td>
<td>250/1</td>
<td>500/1</td>
</tr>
<tr>
<td>5</td>
<td>50/0</td>
<td>50/1</td>
<td>90/1</td>
<td>150/1</td>
<td>400/1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>100/1</td>
<td>100/1</td>
<td>200/1</td>
<td>250/1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>250/1</td>
<td>275/1</td>
<td>375/1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>325/1</td>
<td>475/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Market Implied Outcomes

**Probability of Score Difference – Before 1st Half**

- Draw = 19.50%
- West Ham Wins = 23.03%
- Everton Wins = 57.47%

**Probability of Score Difference – Before 2nd Half**

- Draw = 23.18%
- West Ham Wins = 15.74%
- Everton Wins = 61.08%

**Game Simulations – Before 1st Half**

**Game Simulations – Before 2nd Half**
Skellam Process Approximation I

Market implied and Skellam implied probabilities for score differences before the game: \((\lambda^A_0 = 2.33, \lambda^B_0 = 1.44)\)

<table>
<thead>
<tr>
<th>Score difference</th>
<th>-4</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Prob. (%)</td>
<td>1.70</td>
<td>2.03</td>
<td>4.88</td>
<td>12.33</td>
<td>21.93</td>
</tr>
<tr>
<td>Skellam Prob.(%)</td>
<td>0.78</td>
<td>2.50</td>
<td>6.47</td>
<td>13.02</td>
<td>19.50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Score difference</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Prob. (%)</td>
<td>22.06</td>
<td>16.58</td>
<td>9.82</td>
<td>4.72</td>
<td>2.23</td>
</tr>
<tr>
<td>Skellam Prob.(%)</td>
<td>21.08</td>
<td>16.96</td>
<td>10.61</td>
<td>5.37</td>
<td>2.27</td>
</tr>
</tbody>
</table>
Skellam Process Approximation II

Market Implied Probability vs Skellam Implied Probability

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>t = 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t = 0.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t = 0.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t = 0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t = 0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t = 0.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t = 0.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t = 0.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t = 0.94</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Score Difference

Probability (%)
Win/Draw/Lose Probability Evolution

PROBABILITY OF EITHER TEAM WINNING

- Everton goal
- Everton red card
- Everton goal
- West Ham goals

Everton wins
At the start: 57.5%

Draw
19.5%

West Ham wins
23%

Half time
We define a discrete version of the implied volatility of the games outcome as

$$\sigma_{IV,t} = \sqrt{\lambda_A^t + \lambda_B^t}.$$

Red line: the path of implied volatility: $\sigma_{red}^t = \sqrt{\hat{\lambda}_A(1 - t) + \hat{\lambda}_B(1 - t)}$.

Blue reference lines: constant volatility $\lambda_A + \lambda_B$: $\sigma_{blue}^t = \sqrt{(\lambda_A^t + \lambda_B^t) \times (1 - t)}$. 