Predicting the past (and the future) in Sport using the Bradley-Terry model

Ian Hamilton

19th January 2022

Bradley Terry

In the context of tournaments, the probability that team i beats team j is given by

$$P(i \succ j) = \frac{\pi_i}{\pi_i + \pi_j}$$

where π_i is positive-valued, and can be thought of as a parameter reflecting the strength of team i.

Zermelo (1929), Bradley & Terry (1952)

Why the Bradley-Terry model?

For example:

- Unique entropy maximiser subject to retrodictive criterion
- Unique likelihood maximiser subject to retrodictive criterion
- Wins as a sufficient statistic
- Simplicity maximiser
- Luce's Choice Axiom
- Transitivity of odds
- Game scenarios e.g. Poisson scoring, Sudden death, Accumulated win ratio, Continuous time state transition

Extension to include ties

$$P(i \succ j) = \frac{\pi_i}{\pi_i + \pi_j + \nu\sqrt{\pi_i \pi_j}}$$

$$P(i \approx j) = \frac{\nu\sqrt{\pi_i \pi_j}}{\pi_i + \pi_j + \nu\sqrt{\pi_i \pi_j}}$$

Davidson (1970)

Extension to account for home advantage (order effects)

$$P(i \succ j) = \frac{\pi_i}{\pi_i + \gamma\pi_j + \nu\sqrt{\pi_i\pi_j}}$$
$$P(i \prec j) = \frac{\gamma\pi_j}{\pi_i + \gamma\pi_j + \nu\sqrt{\pi_i\pi_j}}$$
$$P(i \approx j) = \frac{\nu\sqrt{\pi_i\pi_j}}{\pi_i + \gamma\pi_j + \nu\sqrt{\pi_i\pi_j}}$$

Davidson & Beaver (1977)

Applying to 3 for a win, 1 for a draw

$$P(i \succ j) = \frac{\pi_i}{\pi_i + \pi_j + \nu(\pi_i \pi_j)^{\frac{1}{3}}}$$

$$P(i \approx j) = \frac{\nu(\pi_i \pi_j)^{\frac{1}{3}}}{\pi_i + \pi_j + \nu(\pi_i \pi_j)^{\frac{1}{3}}}$$

See: alt-3.uk

Firth (2017)

Retrodictive modelling of modern rugby union

Joint work with Professor David Firth

19th January 2022

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated?

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated?

A: Rugby union!

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin?

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin?

A: Schools rugby!

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin, and there was an actual tournament based on the results of these matches?

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin, and there was an actual tournament based on the results of these matches?

A: Daily Mail Trophy!

Motivation

Q: Wouldn't it be nice (for me, at least) if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin, and there was an actual tournament based on the results of these matches, and the methodology they currently use could do with some serious improvement?

Motivation

Q: Wouldn't it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin, and there was an actual tournament based on the results of these matches, and the methodology they currently use could do with some serious improvement?

A: Full house!

Rugby union scoring rule

League Points:

4 points for a win

2 points for a draw

0 points for a loss

1 bonus point for losing by less than seven points

1 bonus point for scoring four or more tries

Summary

Model	B-T	Davidson	Firth	Rugby
Points - win	1	2	3	4
Points - draw	NA	1	1	2
Points - other	NA	NA	NA	1 (try,losing)
Model - i win	π_i	π_i	π_i	???
Model - draw	NA	$(\pi_i \pi_j)^{1/2}$	$(\pi_i \pi_j)^{1/3}$	???
Model - other	NA	NA	NA	???

RASR (pronounced 'razor') - Ranking Algorithm for Schools Rugby

Part one: result outcome

$$
\begin{aligned}
P(\text{team } i \text{ beats team } j \text{ by wide margin}) &\propto \tau^4 \pi_i^4 \\
P(\text{team } i \text{ beats team } j \text{ by narrow margin}) &\propto \kappa \tau^3 \pi_i^4 \pi_j \\
P(\text{team } i \text{ draws with team } j) &\propto \nu \pi_i^2 \pi_j^2 \\
P(\text{team } j \text{ beats team } i \text{ by narrow margin}) &\propto \frac{\kappa \pi_i \pi_j^4}{\tau^3} \\
P(\text{team } j \text{ beats team } i \text{ by wide margin}) &\propto \frac{\pi_j^4}{\tau^4}
\end{aligned}
$$

A principle-based approach

Maximise entropy

$$S(p) = -\sum_{i,j} \sum_{a,b} p^{ij}_{a,b} \log p^{ij}_{a,b} \quad ,$$

subject to conditions,

$$\sum_{a,b} p^{ij}_{a,b} = 1 \quad , \tag{1}$$

and

$$\sum_{j} \sum_{a,b} a p^{ij}_{a,b} = \sum_{j} \sum_{a,b} a m^{ij}_{a,b} \quad , \tag{2}$$

where $p^{ij}_{a,b}$ is the probability that i gains a points and j gains b points, and $m^{ij}_{a,b}$ is the number of matches that have resulted with i gaining a points and j gaining b points.

A principle-based approach

Taking the Lagrangian and differentiating wrt $p^{ij}_{a,b}$ we have

$$\log p^{ij}_{a,b} = -\lambda_{ij} - a\lambda_i - b\lambda_j - 1 \quad , \tag{3}$$

which gives us that

$$p^{ij}_{a,b} \propto \pi_i^a \pi_j^b \quad , \tag{4}$$

where the $\pi_i = \exp(-\lambda_i)$, may be used to rank the teams, and $\exp(-\lambda_{ij} - 1)$ is the constant of proportionality.

Potential models

Examples:

- Try bonus dependent on result outcome and opposition
- Try bonus independent of result outcome but dependent on opposition
- Try bonus independent of result outcome and opposition
- Offensive-defensive strengths
- Home-away strengths

RASR (pronounced 'razor') - Ranking Algorithm for Schools Rugby

Part two: try bonus outcome

$$
\begin{aligned}
P(\text{team } i \text{ and team } j \text{ both gain try bonus point}) &\propto \theta\pi_i\pi_j \\
P(\text{only team } i \text{ gains try bonus point}) &\propto \tau\pi_i \\
P(\text{only team } j \text{ gains try bonus point}) &\propto \frac{\pi_j}{\tau} \\
P(\text{neither team gains try bonus point}) &\propto \phi
\end{aligned}
$$

Intuitive Measure

Projected Points per Match

$$\text{PPPM}_i = \frac{1}{n-1}\sum_j \sum_{a,b} a p_{a,b}^{ij}$$

Intuitive measure that converges to the rating in round robin

To prior or not to prior?

Introduce a dummy $team_0$ against whom each other team wins one and loses one, then decide how much weight to give these matches.

Pros:

- Ensures connectedness therefore rating from start of season
- Explicitly controls fairness in situations of varying fixture numbers
- Allows for estimation of structural parameters even with existence of 100% record

Cons:

- Might not match intuition / round robin outcomes

The effect of a prior

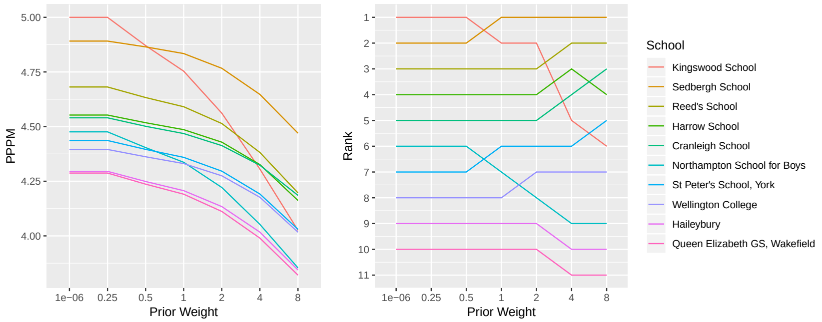


Figure: Top10 PPPM and Rank variation with prior weight for Daily Mail Trophy 2017/18

Daily Mail Trophy

League Points per Match + Additional Points

Additional Points in the Daily Mail Trophy are awarded based on the ranking of the current season's opponents in the previous season's tournament:

Rank 1 to 25:	0.3
Rank 26 to 50:	0.2
Rank 51 to 75:	0.1
Otherwise:	0

Results 2015/16

School	DMT Rank	DMT	PPPM Rank	PPPM
Wellington College	1	6.46	7	3.73
Kirkham	2	6.44	1	4.41
Bedford	3	6.35	2	4.37
Bromsgrove	4	6.21	4	4.15
Sedbergh	5	6.10	5	3.99
Woodhouse Grove	6	5.65	19	3.31
Millfield	7	5.21	13	3.64
Clifton College	8	5.11	8	3.73
Solihull	9	5.10	11	3.67
St Paul's	9	5.10	14	3.58

Results 2016/17

School	DMT Rank	DMT	PPPM Rank	PPPM
Wellington College	1	7.22	3	4.37
Sedbergh	2	6.50	2	4.43
Harrow	3	6.34	6	4.22
St Peter's, York	4	6.23	8	4.06
Kirkham	5	6.15	1	4.61
Canford	6	6.10	9	4.02
Clifton College	7	6.00	5	4.25
Rugby	8	5.96	7	4.06
Brighton College	9	5.90	4	4.29
Woodhouse Grove	10	5.81	12	3.93

Results 2017/18

School	DMT Rank	DMT	PPPM Rank	PPPM
Sedbergh	1	7.41	1	4.65
Wellington College	2	7.18	7	4.18
Cranleigh	3	6.33	4	4.32
Harrow	4	6.20	3	4.33
Cheltenham College	5	6.16	8	4.07
St Peter's, York	6	5.83	6	4.19
Brighton College	7	5.63	20	3.59
Reed's	8	5.50	2	4.38
Clifton College	8	5.50	16	3.72
Haileybury	10	5.49	10	4.02

Euro 2020 Prediction Competition

Joint work with David Selby and Stefan Stein

19th January 2022

Competition details

1. Predict the outcome of all group and knock-out matches in Euro 2020.
2. Outcomes for group matches win/draw/loss, for knock-out matches win/loss.
3. All predictions to be in before kick-off of first match of the tournament.
4. Winner determined by minimum negative log-loss

Further details: `https://github.com/mberk/rss-euro-2020-prediction-competition`

Motivating ideas

1. A significant proportion of the competition outcome would be luck.

Motivating ideas

1. A significant proportion of the competition outcome would be luck.
2. We don't have much time!

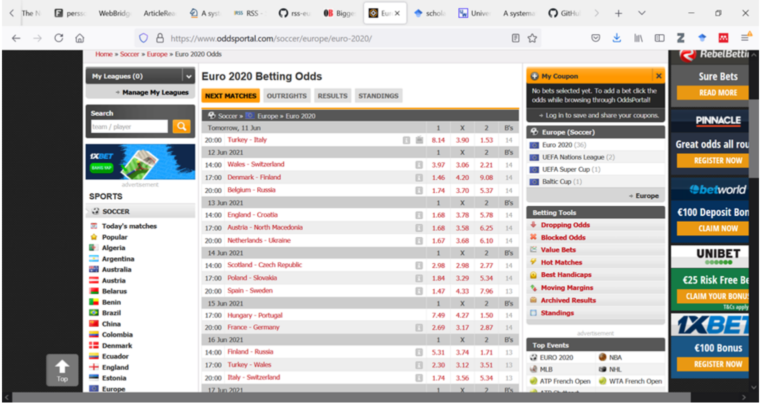
Motivating ideas

1. A significant proportion of the competition outcome would be luck.
2. We don't have much time!
3. Markets are good at evaluation.

Motivating ideas

1. A significant proportion of the competition outcome would be luck.
2. We don't have much time!
3. Markets are good at evaluation.
4. Markets get things wrong in (somewhat) predictable ways.

The data - match probabilities



The model - Bradley-Terry

Probability that i beats j

$$p_{ij} = \frac{\pi_i}{\pi_i + \pi_j},$$

where π_i is the 'strength' of i.

As a generalised linear model

$$\text{logit}(p_{ij}) = \lambda_i - \lambda_j,$$

where $\lambda_i = \log(\pi_i)$

Zermelo (1929); Bradley and Terry (1952)

Bradley-Terry - typical use

Bradley-Terry model applied to a set of results, for the purpose of prediction or ranking e.g. alt-3.uk

Parameters estimated by maximum likelihood estimation

$$L(\boldsymbol{\lambda}) = \prod_{i<j} \binom{m_{ij}}{c_{ij}} p_{ij}^{c_{ij}} (1 - p_{ij})^{m_{ij} - c_{ij}},$$

where c_{ij} is the number of time i beats j and $m_{ij} = c_{ij} + c_{ji}$ is the number of matches between i and j.

Bradley-Terry - issues

But:

1. not enough recent useful results to estimate strengths reliably
2. market prices are likely to be more informative
3. draws in the group stages

Bradley-Terry - dealing with draws

Extension to draws (alt-3.uk, Davidson (1970))

$$\mathbb{P}(i \text{ beats } j) = \frac{\pi_i}{\pi_i + \pi_j + \nu(\pi_i \pi_j)^{\frac{1}{3}}}$$

$$\mathbb{P}(i \text{ draws with } j) = \frac{\nu(\pi_i \pi_j)^{\frac{1}{3}}}{\pi_i + \pi_j + \nu(\pi_i \pi_j)^{\frac{1}{3}}}$$

Note even with draws:

$$\frac{p_{ij}}{p_{ji}} = \frac{\pi_i}{\pi_j} \quad \text{or} \quad \text{logit}(p_{ij}) = \lambda_i - \lambda_j$$

Intra-group strength estimation

Can estimate the intra-group log-strengths $r_i = \log s_i$ by linear regression:

$$\log\left(\frac{p_{ij}}{p_{ji}}\right) = r_i - r_j,$$

since p_{ij} are known from market odds.

But how do we compare strengths between groups?

Overall strength estimation

Assumptions:

1. Team i's overall strength π_i is a scaling of its intra-group strength s_i by a factor dependent on its group $\gamma_{G(i)}$

$$\pi_i = \gamma_{G(i)} s_i \qquad \text{or equivalently} \qquad \lambda_i = \log \gamma_{G(i)} + r_i$$

2. The strength of every team's unknown final opponent is the same

$$p_{io} = \mathbb{P}(i \text{ winning tournament} \mid i \text{ reaches final}) = \frac{\pi_i}{\pi_i + \pi_o},$$

where π_o is the strength of the unknown final opponent.

Overall strength estimation

We can calculate p_{io} from market odds since

$$p_{io} = \frac{\mathbb{P}(i \text{ winning tournament})}{\mathbb{P}(i \text{ reaches final})}$$

Then we have that

$$\log\left(\frac{p_{io}}{p_{oi}}\right) = \lambda_i - \lambda_o = \log\gamma_{G(i)} + r_i - \lambda_o,$$

and we can estimate $\log\gamma_{G(i)}$ and λ_o through linear regression.

Knock-out prediction

Now we can calculate the strengths of each team

$$\pi_i = \gamma_{G(i)} s_i,$$

and apply these through the Bradley-Terry model to predict the KO match results

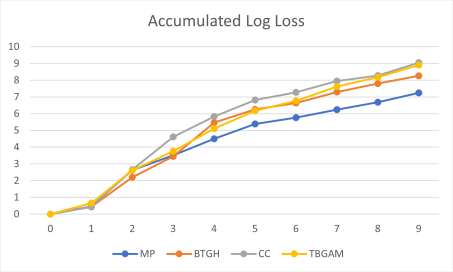
$$p_{ij} = \frac{\pi_i}{\pi_i + \pi_j}.$$

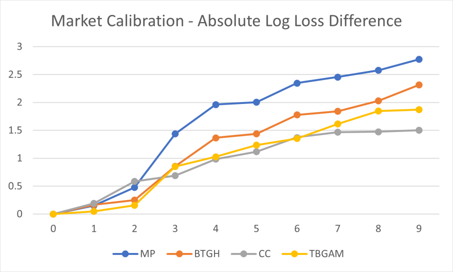
Miscellaneous notes

1. Parsimonious model - 120 data points (72 group stage match probabilities + 24 reach final + 24 tournament win); two linear regressions; two days
2. What happened to market being wrong in predictable ways?
3. Did we do well just because of taking market odds for the group stage?

How much was luck?

Performance graphs for KO stages alone based on the nine competition updates





Bibliography

Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, 39(3/4):324–345.

Davidson, R. R. (1970). On extending the bradley-terry model to accommodate ties in paired comparison experiments. *Journal of the American Statistical Association*, 65(329):317–328.

Zermelo, E. (1929). Die berechnung der turnier-ergebnisse als ein maximumproblem der wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 29(1):436–460.

Resources

Talks: RSS Merseyside Local Group: Statistics and Football https://www.youtube.com/channel/UChNo0mvmV9KzB8KCxP2n9_w

Books: Who's #1? by Langville & Meyer; Contest Theory (ch 9,10) by Vojnovic

Conferences: http://www.nessis.org/index.html

Competitions: https://rss.org.uk/news-publication/news-publications/2021/section-group-reports/sports-section-euro-2020-prediction-competition/

Football prediction: https://mathematicalfootballpredictions.com/dixon-coles/

Others: https://alt-3.uk/; www.warwick.ac.uk/IanHamilton