Retrodictive Modeling of Modern Rugby Union
An extension of Bradley-Terry to multiple outcomes

Ranking vs Rugby
How do we produce a coherent ranking, respecting the scoring rules of a sport, when teams play different numbers of matches, against different strengths of opposition, with different proportions at home or away?

In modern rugby union a team earns:
- 4 for a win
- 2 for a draw
- 0 for a loss
- 1 bonus for losing by less than seven
- 1 bonus for scoring four or more
- A try bonus point.

This gives a system of five result outcomes: wide win, narrow win, draw, narrow loss, wide loss, plus a try bonus point.

The probability that \( i \) beats \( j \) is defined as
\[
P(i ≻ j) = \frac{\pi_i}{\pi_i + \pi_j + \nu \sqrt{\pi_i \pi_j}}
\]
\[
P(i ≼ j) = \frac{\pi_j}{\pi_i + \pi_j + \nu \sqrt{\pi_i \pi_j}}
\]
\[
P(i ≈ j) = \frac{\nu \sqrt{\pi_i \pi_j}}{\pi_i + \pi_j + \nu \sqrt{\pi_i \pi_j}}
\]
... extended by Davidson (1970) to allow for ties.

The Model
Given a set of rugby match results with teams \( i \) the home team and team \( j \) the away team, the model estimates a vector of ratings \( \pi \) such that
\[
P(i \text{ beats } j \text{ by wide}) \propto \pi_i^3
\]
\[
P(i \text{ beats } j \text{ by narrow}) \propto \pi_i^2 \pi_j
\]
\[
P(i \text{ draws with } j) \propto \pi_i \pi_j
\]
\[
P(j \text{ beats } i \text{ by narrow}) \propto \pi_i \pi_j
\]
\[
P(j \text{ beats } i \text{ by wide}) \propto \pi_i \pi_j
\]
and:
\[
P(i \text{ and } j \text{ both try bonus}) \propto \pi_i \pi_j
\]
\[
P(\text{only } i \text{ try bonus}) \propto \tau \pi_i
\]
\[
P(\text{only } j \text{ try bonus}) \propto \pi_j
\]
\[
P(\text{neither team try bonus}) \propto \phi
\]
with \( \nu, \kappa, \tau, \phi, \) and \( \phi \) structural parameters.

Calibrating Fairness
Which of these schools in the 2016/17 Daily Mail Trophy deserved to be ranked higher?

<table>
<thead>
<tr>
<th>School</th>
<th>P</th>
<th>W</th>
<th>D</th>
<th>L</th>
<th>Bonus</th>
<th>Points per Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manchester Grammar</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>4.75</td>
</tr>
<tr>
<td>Clifton College</td>
<td>10</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td></td>
<td>4.5</td>
</tr>
</tbody>
</table>

This is really a question for the tournament stakeholders rather than the statistician, but we can provide means by which a preference may be calibrated.

Ranking in practice
The Daily Mail Trophy is an annual competition between around a hundred of the best school teams in England. Schools play anywhere between five and thirteen matches, with the strength of opposition varying.

The tournament organisers have expressed a desire for a methodology that is more transparent to stakeholders than they believe this model to be. However the model provides insights on the design and calibration of an alternative.

This led us to propose the ‘Dapper’ (Damped and Adjusted Points Per match) family of models, which improves their current methodology by incorporating a per-match adjustment and a parameter that allows for record preference to be explicitly controlled (see Calibrating Fairness).

The Evolution of a Sports Ranking Model

1928
The probability that \( i \) beats \( j \) is defined as
\[
P(i > j) = \frac{\pi_i}{\pi_i + \pi_j + \nu \sqrt{\pi_i \pi_j}}
\]
where \( \pi \) is a positive-valued rating of \( i \). Originally proposed by Zermelo (1929) in the context of chess tournaments, before being rediscovered by Bradley and Terry (1952).

1970
... further extended by Davidson and Beaver (1977) to, in the context of sports ranking, account for home advantage.

1977

2017
... and updated by Firth (2017) to account for modern soccer, where teams are awarded three for a win and one for a draw.

Hint: some questions you might want to ask me....

- Hi, how are you?
- Is this your first poster?
- The \( \pi \)s are not a very intuitive measure - is there something more interpretable we can convert them to?
- Don’t you need a constraint on the \( \pi \)s?
- Why don’t you use a predictive machine learning based algorithm (similar to how they use NET in NCAA basketball)?
- Rugby is kind of like football, right?
- Would I ever want to calculate the parameters without using a prior?
- Isn’t there some sort of big rugby tournament on at the moment?
- How did you actually calculate the parameter estimates?
- I don’t suppose you have any business cards?

For more details please visit: warwick.ac.uk/IanHamilton