

Copyright © 2006 Donald Forbes, Stephen Clarke and Denny Meyer.

This is the author's version of the work. It is posted here with the permission of the publisher for your personal use. No further distribution is permitted. If your Library has a subscription to these conference proceedings, you may also be able to access the published version via the library catalogue.
AFL FOOTBALL – HOW MUCH IS SKILL AND HOW MUCH IS CHANCE?

Forbes, D, Clarke, S. and Meyer, D.
School of Life Sciences, Swinburne University

ABSTRACT

It is widely accepted that luck has some influence on all sporting events, whether it be an injury at a crucial time, an unfortunate bounce of the ball or a controversial decision by an official. But just how much does the result of a match come down to bad or good luck i.e. the random events within a game? A Markov process model with 18 states has been developed to approximate AFL football using match statistics. The models have been applied to matches from the 2003 and 2004 AFL seasons and displayed very encouraging results with minimal errors. Using transition probabilities derived from these matches, simulated goal distributions and probabilities of victory have been calculated. Using these distributions and the standard errors for each team’s match scores, the randomness that occurs in AFL matches is identified. Some of these matches have been investigated with the results presented in this paper.

KEY WORDS
Markov process, Poisson distribution, computer simulation, random events

INTRODUCTION

In almost all sports played worldwide, there is an element of luck associated with performance. The game of Australian Rules football is no different. The fact that the ball used is oval in shape and often does not bounce as expected probably exacerbates the importance of being on the right side of lady luck. Historically, the issue of chance in sport and the role it plays has not been a focal point of academic research. Investigation into chance playing a part in sporting outcomes has mostly been concentrated on soccer.

Early work on modelling sporting events using statistical distributions found that the best descriptor of goals scored in soccer was the negative binomial distribution (Moroney, 1956). The author expressed surprise that weather conditions and the strength of the competing teams did not exert as great an effect as is often supposed. Later work also found the negative binomial distribution to give the best fit due to it being generated by random or chance mechanisms that underlie the conclusion that soccer is a game dominated by chance. It was suggested that, due to this notion, a team who recognised this random element would be able to develop a successful style of play that harnessed the importance of chance on the game (Reep and Benjamin, 1968). A direct style of play was deemed to be the best method by which to harness the chance occurrences that appear in soccer. By putting the ball into goal scoring situations as often as possible, a team was more likely to take advantage of these opportunities when they arose. Recent work has traced this direct and successful style of play back to the Arsenal side of the 1930s, but does acknowledge the part Reep played in bringing it to the successful Norwegian side of the 1990s (Larson, 2001).
The negative binomial distribution was applied to a number of facets of different sports (Pollard, Benjamin and Reep, 1977). Events looked at included passing chains in soccer (and goal scoring), points scored in gridiron, runs in a baseball half-inning, goals scored in ice-hockey, strokes per rally in tennis and runs scored per partnership in cricket. It was found that the negative binomial produced a good fit where there was an occurrence of infrequent events in a team environment e.g. soccer goals. However, when individual performances were looked at, such as in the tennis or cricket examples, there was not a close fit, indicating that individual skill was more significant than chance in these instances.

These discoveries lend weight to chance playing a significant part in matches played in the Australian Football League (AFL), where the negative binomial has been found to be the best approximation for goal scoring in the competition, and the Poisson is the best approximation for individual team returns (Forbes, 2006). This paper uses a Markov process model to investigate matches from the 2005 AFL season, highlighting the importance of chance and making the most of it to win AFL games. Furthermore, simulated matches give a better understanding of the amount of variation and randomness that can be expected in a team’s score for a match.

THE MODEL AND SIMULATION PROCESS

An 18-state Markov process model has been developed to approximate Australian Rules football (Forbes, 2006). This model builds on an earlier, cruder model that was made up of only seven states (Forbes and Clarke, 2004). The accuracy of this model in reflecting the events of a match is impressive. It gives rise to a number of applications that can be used by AFL clubs, as well as media outlets and sports bookmakers. Some of the applications of the model will be presented in the case studies that follow, investigating the role of chance and random events in AFL football.

A simulation program has been developed to utilise the model in a post match environment. The underlying idea of the program is to generate random numbers that determine which state the model will move into next. The process has been broken down into four parts to reflect more accurately an AFL match, which is divided into four quarters. The starting state for each quarter is state seven (centre bounce) and the simulation is run according to the number of transitions contained in the match divided by four. The simulation is run 10,000 times in general, but there is no restriction, save CPU processing time, to the number of times it can be run. To give an example of speed, on a notebook with 256MB of RAM, a season comprising 185 matches can have each match simulated 10,000 times in roughly 20 minutes.

Upon running a simulation of a match, or set of matches, there is a variety of information that can be gleaned from the process. Perhaps the most important and useful data are the projected score of each team for each simulated match. Goals are calculated for each team by counting the number of times either Team A or B had possession directly before a centre bounce, excluding the centre bounce at the start of each quarter. These occurrences are then multiplied by six and added to the number of
times the model entered the behind state for each team to come up with a match score, which is then used to ascertain the match outcome. The probability of victory for each team and the likelihood of a draw is then a simple calculation according to the expected outcome. Other information that can be extracted from a simulation is the proportion of time the ‘match’ is likely to spend in any one state and the number of occurrences of each state within a simulated match.

The simulation process was validated by the results obtained under different scenarios. Home advantage in the AFL competition between 1998 and 2003 amounted to 12.3 points on average. When the 2004 season was simulated using a transition matrix for each match, the average margin in favour of the home side amounted to a very similar 11.8 points. A matrix was derived for the 2005 season based on the 185 matches and used to simulate an ‘average’ match. The score line that resulted from the simulation was 99-90 in Team A’s favour, whereas for the actual data from the 2005 season, the average score was 102-87, again suggesting a favourable comparison between the simulated and actual results. Further results from the simulation program will be presented in the following case studies.

**CASE STUDY 1: SYDNEY V WEST COAST**

The 2005 AFL Grand Final was a classic, not so much for the quality of football played, but more so due to the commitment of both teams and the knife edge atmosphere that was created by the closeness of the score line. In the end Sydney won the match by just four points, 58-54, and no doubt West Coast spent the summer licking their wounds and pondering upon just how close they came. Some analysis has been done on the match to investigate how much of a part randomness played in the final result and to determine whether this was a game that could have been won by the bounce of the ball.

The transition matrix for the match was derived using the 18-state zone model. The match contained 885 transitions, which amounts to 221 a quarter. Using the observed transition matrix and the number of transitions for a quarter, the match was simulated 10,000 times. From the simulation, expected scores for each team were obtained, along with probabilities of victory and these are contained in Table 1.

<table>
<thead>
<tr>
<th>Team</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Probability of Victory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sydney</td>
<td>58.4</td>
<td>16.9</td>
<td>0.53</td>
</tr>
<tr>
<td>West Coast</td>
<td>56.0</td>
<td>16.0</td>
<td>0.45</td>
</tr>
</tbody>
</table>

What is evident from Table 1 is that, with the transition probabilities as they were in the match, the West Coast were expected to score, on average, two more points than they did. They have a mean score of 56.0 points from the simulation results, yet in the Grand Final they scored 54 points. Do we then consider West Coast unlucky not to get the return on the scoreboard that was expected from the way the match was played? They had a 45% chance of winning the match and a 2% chance of the draw.

To continue on from this match, the distribution of goals for each team has been calculated from the simulated matches. It has been discovered in other research that the
distribution of goals for teams in the AFL is best approximated by the Poisson distribution (Forbes, 2006). The assumptions that must be met for the Poisson distribution to apply are that the probability of an event within a certain interval does not change over different intervals and that the probability of an event in one interval is independent of the probability of an event in any other non-overlapping interval. The first assumption indicates that club scoring rates in the AFL do not vary from week to week and therefore, are not affected by factors such as weather conditions, venue size, strength of the opposition or influential players missing. Although this is somewhat of a surprising result, the distributions presented below for both Sydney and the West Coast, highlight the variability associated with how many goals a team can be expected to score in an AFL match.

---

Figure 1: Sydney goal distribution from simulated Grand Final

Figure 2: West Coast goal distribution from simulated Grand Final

The final analysis of the Grand Final involves what turned out to be the ultimate play of the match. West Coast surged into attack and West Coast players were queuing up to mark the ball when Sydney’s Leo Barry came from the side for a spectacular contested...
mark. Barry’s mark was the kind of mark that a player might pull off a couple of times during a career and further indicates the importance of taking your chances when they arise. The ball could quite easily have been marked by a West Coast player, who would have had a relatively straightforward shot on goal to win a premiership.

To investigate the importance of Barry’s mark, the last three plays of the game have been simulated starting from West Coast possession in the midfield, with the next transition resulting in an Inside 50 play for the West Coast. Using the matches from the simulation where these conditions were satisfied, with Sydney leading the match by four points at the time, the West Coast were found to have an 11% chance of winning the match. Even though this is still a relatively low chance of winning, compared to Sydney’s chances, it indicates the importance of the play by Barry. Paul Roos, the coach of Sydney, would not like to endure those final few seconds again, knowing that he had an 11% chance of having the Premiership snatched away from him if Barry spills the mark.

**CASE STUDY 2: KANGAROOS V RICHMOND**

The 2005, round 12 match between the Kangaroos and Richmond, saw the Roos walk away with a comfortable 29 point victory, 109 - 80. On paper, the issue of chance does not seem to be a concern. A five goal victory is comprehensive in anyone’s language, but upon drilling deeper into the match statistics, chance played a big part in such an easy win. The transition matrix for the match, when used in the simulation program produced an expected score line of 111 – 78, which is actually four points more than the observed margin, indicating that perhaps Richmond were ‘lucky’ to get as close as they did. Not surprisingly, the Kangaroos were an 85% chance of winning the match.

That aside, the goal kicking of both teams was the crucial difference, with the Kangaroos kicking 17 goals, 7 behinds and Richmond 10 goals, 20 behinds. Richmond had six more scoring shots than their opponents, yet they were soundly beaten by nearly five goals. The possession profiles of the teams were very similar and this is highlighted by how they entered their attacking zones. Table 2 contains the profile for each team when they put the ball into their attacking zone.

*Table 2: Distribution of ball into attacking zones, Kangaroos v Richmond, round 12 2005*

<table>
<thead>
<tr>
<th>Team</th>
<th>BUBO</th>
<th>THIN</th>
<th>DISP</th>
<th>APOS</th>
<th>BPOS</th>
<th>BEHI</th>
<th>CBO</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kangaroos</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>17</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>46</td>
</tr>
<tr>
<td>Kangaroos</td>
<td>%</td>
<td>0.0</td>
<td>0.0</td>
<td>45.7</td>
<td>37.0</td>
<td>10.9</td>
<td>0.0</td>
<td>6.5</td>
</tr>
<tr>
<td>Richmond</td>
<td>0</td>
<td>0</td>
<td>22</td>
<td>10</td>
<td>20</td>
<td>3</td>
<td>0</td>
<td>55</td>
</tr>
<tr>
<td>Richmond</td>
<td>%</td>
<td>0.0</td>
<td>0.0</td>
<td>40.0</td>
<td>18.2</td>
<td>36.4</td>
<td>5.5</td>
<td>0.0</td>
</tr>
</tbody>
</table>

There is no clear indication from Table 2 as to why the Kangaroos were such easy victors. In their favour was the ability to kick long goals from outside 50m. Aside from this, the profiles are very similar. Deeper investigation has to be carried out on what happened when the ball was in the attacking zones to see why there was such a discrepancy on the scoreboard. Richmond did better than the Kangaroos at getting the ball from dispute in the Kangaroos attacking zone. In Richmond’s attacking zone the teams broke even in this area. The lopsided scoreboard can only be put down to Richmond’s inaccuracy both from inside and outside 50. Table 2 shows that the
Kangaroos kicked three goals from beyond 50m whilst Richmond could manage only three behinds. When the ball got into the attacking zones, the Kangaroos kicked a goal 40.0% of the time and a behind only 20.0% of the time, whereas, Richmond could manage a goal only 16.8% of the time and a behind 30.4% of the time. This is a clear case where the relationships in the game are relatively equal; however, poor conversion by one side has led to a comprehensive defeat.

To investigate the random effect of goal kicking accuracy, adjustments have been made to the transition matrix so that each side has a similar conversion rate of 50%. Using this adjusted matrix in the simulation program, the expected score line is now 111 – 104 in the Kangaroos favour, but they are now only a 59% chance of winning. This is a powerful example of the importance of accurate kicking in the AFL. Richmond, cruelled their chances of victory with poor conversion and this is a purely random occurrence as the week before they kicked 15 goals, 13 behinds and the week after 9 goals, 8 behinds.

CONCLUSION

Clearly, randomness is a big part of success in the AFL competition. With the aid of an 18 state Markov process model and a simulation program, two case studies have been presented that identify the importance of chance in AFL matches. The first game looked at was the 2005 grand final and the simulated goal distributions gave a clear indication of the variability one can expect from game to game in goals kicked. Furthermore, the expected score line was closer than the actual score, highlighting that perhaps the luck did not run with the West Coast. Finally, a once in a lifetime mark at the end of the game saved the premiership for Sydney, when a different outcome produced an 11% chance of victory for the opposition with only three plays left in the match.

The second case study highlighted the random effect of goal kicking in an AFL match; Richmond had more shots than their opposition, but got beaten by five goals due to their poor conversion. With an adjusted conversion rate of 50%, surpassed by Richmond in the week before and the week after the match, their chances of victory improved by over 25%. This is another example of the effect of chance in the competition and, for Richmond, the match against the Kangaroos was obviously one of those days.

REFERENCES


