SOCIETY OF ACTUARIES

Article from

Forecasting and Futurism

Month Year July 2015
Issue Number 11

Simple Rating Systems: Entry-level sports forecasting

By Doug Norris

As actuaries, we typically focus our predictive efforts in a relatively small niche area—for instance, I primarily focus on commercial health care pricing, reserving, and strategy. However, I would speculate that most of us learned our love for mathematics and forecasting long before we were formally trained in actuarial techniques.

Growing up in the suburbs of Seattle, I was a sports fan. In particular, I was fascinated by sports statistics. I would invent baseball games using my card collection and a set of oddball dice (for anyone looking to follow in my footsteps, 10-sided dice are incredibly handy). I pored over Bill James's annual *Baseball Abstract* editions. I tracked the statistics for my Little League team. I played APBA[1], Strat-O-Matic, and SherCo simulation games (and still play in a Strat-O-Matic hockey league to this day[2]). In 1994, I started one of the first sports websites, The Goaltender Home Page,[3] dedicated to preserving the history and numbers of hockey's unsung heroes.

A seminal moment in my actuarial career came with Bill James's 1985 *Baseball Abstract*, where James develops a rudimentary predictive model called "Brock-2." Given a baseball player's statistics to date, this model attempted to "complete" the player's remaining career. I dutifully reproduced the formulas in my parents' Apple II+ (fortunately, we had the model with 64k of memory, which was almost enough to reproduce the model), and thus began my first foray into predictive modeling. As an 11-year-old, I had a very powerful thought—how cool would it be if we could predict everything in sports? (I now realize that not only are sports inherently not perfectly predictable, but that those unpredictable aspects are the things that make sports the most fun.)

One of the most basic elements of sports forecasting involves predicting the winner of an upcoming game. Many are interested in being able to do this well, with billions of dollars bet on just the most recent Super Bowl alone.[4] This article describes a simple method for forecasting sports outcomes; in fact, the name itself has an appealing simplicity to it. However, the method is flexible enough to incorporate personal touches and improvements based on your own experience and judgment.

WHAT IS A SIMPLE RATING SYSTEM?

Nearly everyone who attempts to predict the outcome of sporting events realizes that past performance is a key indicator of future success. For instance, when an undefeated team plays a winless team, the undefeated team usually wins.

Predictions incorporating each team's point totals involve a trade-off—instead of focusing on what we are truly interested in (wins), we emphasize a proximate measure (points are not identical to wins, but points represent a "currency" that is used to purchase wins). Therefore, although a team's overall success is intimately intertwined with how well they produce points and prevent their opponents from producing points, counting points (instead of counting wins) results in a loss of specificity. However, this loss is offset by a gain in data—although most sporting events produce only one winner and one loser, each event produces many more points (or goals, or runs, or whatever translation your sport of choice uses). The increase in data helps to offset small sample size variation to some degree, and the trade-off typically results in increased predictive ability.[5] The Simple Rating System (SRS) method incorporates point totals, but takes things one step further.

Consider a six-team hockey league, with franchises named the Alligators, the Badgers, the Conquistadors, the Dragons, the Eagles, and the Falcons. So far in the season, each team has played three games, as shown in Figure 1.

Figure 1: League Outcomes to Date

Road Team	Goals	Home Team	Goals
Alligators	2	Badgers	5
Conquistadors	1	Dragons	0
Eagles	2	Falcons	2
Badgers	3	Alligators	0
Dragons	1	Falcons	4
Eagles	1	Conquistadors	3
Badgers	5	Dragons	1
Eagles	0	Alligators	3
Conquistadors	1	Falcons	1

The Falcons will next visit the Alligators, the Dragons will visit the Eagles, and the Badgers will visit the Conquistadors. Our goal is to provide our best estimate of who will win each game (and by how much). A good first step would be based on how each team has performed so far, so let's look at that to guide us, in Figure 2.

Figure 2: League Performance to Date

	Games	Wins	Losses	Ties	Win %	Goals Scored	Opponent Goals	Avg. Margin Victory
Badgers	3	3	0	0	100%	13	3	+3.33
Conquistadors	3	2	0	1	83%	5	2	+1.00
Falcons	3	1	0	2	67%	7	4	+1.00
Alligators	3	1	2	0	33%	5	8	-1.00
Eagles	3	0	2	1	17%	3	8	-1.67
Dragons	3	0	3	0	0%	2	10	-2.67

As you can see, the Badgers are playing very well, and the Dragons are playing very poorly, with the other teams spread out in between. The Falcons are outscoring their opponents by one goal per game, while the Alligators are being outscored by one goal per game, so we might reasonably predict that the Falcons will beat the Alligators by two goals in their next contest. Similarly, we might predict the Eagles (-1.67 goals/game) to defeat the Dragons (-2.67 goals/game) by one goal, and the Badgers (+3.33 goals/game) to outscore the Conquistadors (+1.0 goal/game) by 2.33 goals.

But wait—the Badgers have played a pretty weak schedule thus far (facing the Alligators twice and the Dragons once). Could their observed dominance be merely a reflection of their strength of schedule, and not their true ability? The Badgers' typical opponent has lost games by an average of 1.56 goals, so if playing against a truly "average" opponent, we would expect the Badgers to win by (+3.33 goals) + (-1.56 goals) = +1.78 goals. Let's revisit all six teams, focusing on their average margins of victory along with their strength of schedule (opponents' average margins of victory), as shown in Figure 3.

Figure 3: Simple Rating System: First Iteration

	Avg. Margin	Schedule Strength	Adj. Avg. Margin
Badgers	3.33	-1.56	1.78
Conquistadors	1.00	-1.11	-0.11
Falcons	1.00	-1.11	-0.11
Alligators	-1.00	1.67	0.67
Eagles	-1.67	0.33	-1.33
Dragons	-2.67	1.78	-0.89

Note: Adjusted average margin of victory = average margin of victory + schedule strength)

CONTINUED ON **PAGE 38**

We can see that the teams with worse records have generally played a stronger schedule, facing stronger opponents (and vice versa)—this makes sense intuitively for two reasons: first, we are measuring average margins of victory, and teams with losing records necessarily have given their opponents more wins than losses. Second, teams with losing records don't get to play themselves. Typically, these disparities are stronger when teams have only played a few games (and have a disproportionate share of their games against one team).

Revisiting our upcoming games, and noting that the adjusted average margin of victory (AAMV) represents how a team might fare against an "average" opponent, we would adjust our predictions such that the Alligators (+0.67 AAMV) would be favored by 0.78 goals over the Falcons (-0.11 AAMV), the Dragons (-0.89 AAMV) would be favored by 0.44 goals over the Eagles (-1.33 AAMV), and the Badgers (+1.78 AAMV) would be favored by 1.89 goals over the Conquistadors (-0.11 AAMV). Note that the predicted outcomes of our three games have changed considerably (with the overall winner changing in two of the three predictions).

At this point, you may be wondering—if we believe that the AAMV values represent a more accurate "team strength" metric, why aren't we using them to determine each team's schedule strength? Yes, we should be using the AAMV to develop an updated strength of schedule (SOS) estimate for each team, which in turn produces an improved estimate of AAMV (and so forth). In the end, we're looking for AAMV estimates that, when used to compute schedule strength estimates, produce the same AAMV estimates in return. In linear algebra parlance:

$AAMV_O$ = initial average margin of victory for each team
SOS_n = average $AAMV_n$ of each team's opponents (weighted by times played)
$AAMV_1 = AAMV_O + SOS_O$
$AAMV_{n+1} = AAMV_O + SOS_n$

We would like to find values for $AAMV_n$ such that $AAMV_n$ equals $AAMV_{n+1}$. If S (short for "schedule") represents the matrix where $S_{x,y}$ counts the proportion of times that team *x* has played team *y*, we know that

$$AAMV_{n+1} = AAMV_O + S * AAMV_n$$

For $AAMV_n$ to equal $AAMV_{n+1}$, we must satisfy:

Solving for $AAMV_n$:

$$AAMV_n - S * AAMV_n = AAMV_O$$
$$(I - S) * AAMV_n = AAMV_O$$
$$AAMV_n = (I - S)^{-1} * AAMV_O$$

Where *I* is the $n_x n$ identity matrix. Those of us who have taken linear algebra are happy to see the end point; however, in this case, the (I – S) matrix proves to be singular (and therefore non-invertible).[6] However, we can solve the problem numerically, and compare the differences of successive iterations; our hope is that the sum of the absolute value of these differences becomes sufficiently small after a large number of iterations, in which case we have found a convergent solution.[7] For our mythical hockey league, Figure 4 shows the unique convergent solution (and final Simple Rating System margins of victory for each team).

Ultimately, our SRS algorithm predicts the Alligators (+0.55 SRS) to be favored by 0.72 goals over the Falcons (-0.17

Figure 4: Simple Rating System: First Iteration

	Avg. Margin Victory	Avg. Opp. SRS Margin	SRS Margin
Badgers	3.33	-0.21	3.12
Conquistadors	1.00	-1.17	-0.17
Falcons	1.00	-1.17	-0.17
Alligators	-1.00	1.55	0.55
Eagles	-1.67	0.07	-1.60
Dragons	-2.67	0.93	-1.74

SRS), the Eagles (-1.60 SRS) to be favored by 0.14 goals over the Dragons (-1.74 SRS), and the Badgers (+3.12 SRS) to be favored by 3.29 goals over the Conquistadors (-0.17 SRS).

HOW CAN WE IMPROVE UPON THE SIMPLE RATING SYSTEM?

First and foremost, the SRS algorithm is not guaranteed to converge, particularly when the network of games played is sparse. For instance, when each team has only played one game, then an infinite number of convergent solutions exist. Related to this, until the SRS algorithm has enough data to work with, the credibility of the predictions suffers. Similar to pricing an insurance product, sports forecasters will typically blend experience data with a "manual rate" until the experience data can stand on its own legs. This manual rate could be based upon prior years' data, or built using other information (there are some interesting agent-based approaches to this), and then massaged by knowledge of the participants.

Speaking of credibility, one flaw of the SRS algorithm (as presented here) is that it considers all data fed into it to be of equal credibility. In reality, a team with sufficient sample size is more likely to perform at the level of its recent performance than at the level of earlier events. It is a simple matter to tweak the SRS algorithm to allow for different outcome weights (as for what those weights should be, that's where art meets science).

Similarly, there are many things that are "known" about sports. First, teams typically perform better in their home environment (this probably makes sense intuitively, even if you aren't a sports fan). Second, outlier performances, where one team dominates an opponent to an excessive degree—such as a football game with a score of 55-0, or a baseball game with a score of 14-1—can have a disproportionate effect on SRS algorithms, because in games where the outcome is decided early, teams do not necessarily finish the game at their "true" ability level. Third, team composition can change throughout the course of a season, which is due to trades, promotions and demotions, coaching changes, injuries, and other factors (which can also affect individual, i.e. non-team, athletes, such as tennis players and golfers). All of these can be accounted for, using judgment and experience, in SRS algorithms. One additional modification (that shows some predictive "lift") considers offensive and defensive contributions separately, as (for instance) a team that scores proficiently against a good defense might deserve more credit than would be expected by comparing against the opponent's overall AAMV.

Sports fans reading this article are probably already thinking of additional improvements that could be made to the SRS algorithm, including sport-specific nuances to improve the predictive nature of the methodology. Of course, this is the fun of predictive models, and in sports forecasting, the truly brilliant modifications are proprietary and confidential. With that said, if you come up with anything compelling, I'd love to hear more about your efforts. Remember that illegal gambling is illegal (hence the term "illegal gambling"), and that this article is for entertainment purposes only. ▼

ENDNOTES

1 A game company once named American Professional Baseball Association.

2 See the National Strat-O-Matic Hockey League at http://nshl.org.

3 See http://hockeygoalies.org. Clearly, I have invested more in data improvements than in aesthetics.

4 American Gaming Association (January 22, 2015). Illegal Super Bowl bets to total $3.8 billion this year. Retrieved March 27, 2015, from http://www.americangaming.org/newsroom/press-releases/illegal-super-bowl-bets-to-total-38-billion-this-year.

5 Baseball-Reference.com (February 20, 2015). Pythagorean Theorem of Baseball. Retrieved March 27, 2015, from http://www.baseball-reference.com/bullpen/Pythagorean_Theorem_of_Baseball.

6 This is a rather fun proof left for the reader. First, prove that each row of (I – S) sums to zero. What does this imply about the triangularized matrix?

7 In this case, all of the linear algebra holds up to (but not including) the matrix inversion step, meaning that the solution (if it exists) is not necessarily unique.

Doug Norris

Doug Norris, FSA, MAAA, Ph.D., is a principal and consulting actuary at Milliman in Denver, Colo. He can be reached at doug.norris@milliman.com.