Settling Scores and Gambling on Goals*

Matthew Cotton AMIMA and **Brady Metherall** AMIMA, University of Oxford

Matthew and Brady are struggling to settle a dire dispute. Having exhausted all other means of resolution, they are forced to rely on chance and agree to let Lady Luck lay the hatchet to rest. Yet in today's cashless society, neither of them has a coin to toss and so they decide to bet on the result of a football match.[1] They cannot simply bet on the score, since both of them may guess incorrectly, and they cannot bet on a winner, since the result may be a draw or one team may be a clear favourite. Instead, they decide to wager on whether there is an even or odd number of goals scored in the match – the 'evenness' of the end result. If Matthew bets on even and Brady bets on odd, who is more likely to win?

Winners evaluate infinite sums

As with all good mathematical models, a useful first step in describing our system is to abstract away the details. We assume that goals are scored at a constant rate and that each goal occurs independently, so that we model goal scoring as a Poisson process. The rate at which each team scores is a reflection of their ability and will likely differ for each team. Since in our model goals are independent, the overall scoring rate – the rate we care about – will be the sum of the two teams' rates and will remain Poisson.

Therefore, the evenness of the final score is simply the probability of there being an even or odd number of events in a Poisson process. If we expect λ goals in a match, then the probability of scoring n goals is $p(n) = \mathrm{e}^{-\lambda}\lambda^n/n!$, where λ is also the variance of the distribution.

The probability of the game ending with an even number of goals, p_{even}, is then the sum of the probabilities of all the even results, $p_{\text{even}} = p(0) + p(2) + p(4) + \dots$, which we can write as the infinite sum

$$p_{\text{even}} = \sum_{n \text{ even}} \frac{\mathrm{e}^{-\lambda}\lambda^n}{n!} = \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{(2k)!},$$

and for the odd case

$$p_{\text{odd}} = \sum_{n \text{ odd}} \frac{\mathrm{e}^{-\lambda}\lambda^n}{n!} = \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{2k+1}}{(2k+1)!}.$$

We can compare the expressions to standard series expansions and spot that the sums are simply the hyperbolic cosine and sine functions! That is, the probability of an even or odd number of goals in a game is

$$p_{\text{even}} = \mathrm{e}^{-\lambda}\cosh(\lambda) = \frac{1+\mathrm{e}^{-2\lambda}}{2}, \tag{1}$$

$$p_{\text{odd}} = \mathrm{e}^{-\lambda}\sinh(\lambda) = \frac{1-\mathrm{e}^{-2\lambda}}{2}. \tag{2}$$

If the teams are so bad that they never score, $\lambda = 0$, the end result will always be 0–0, an even number of goals. This is predicted by (1), since $p_{\text{even}}(\lambda = 0) = 1$. Conversely, when goals are scored at a high rate, the distribution of $p(n)$ broadens and the difference $p(n) - p(n+1)$ tends to 0. As more and more goals are scored, the advantage for the even player vanishes, and in the limit $\lambda \to \infty$, $p_{\text{even}} = p_{\text{odd}} = 1/2$. However, since $\sinh(x) < \cosh(x)$ for all x, the smart money always bets on an even number of goals. *It is even better to be an even better.*

Just as Matthew and Brady are placing their bets, their officemate wants to join the action. Our duelling duo becomes a trio in trouble, so now we need to add a third player to our betting game! This is easy to do if we change the win condition from depending on evenness to depending on the remainder when dividing the total number of goals scored in the match by three.

It is easiest to label ourselves players 0, 1 and 2, where player 0 wins if the number of goals scored is congruent to 0 modulo 3 (that is, if 0, 3, 6, etc. goals are scored), player 1 wins if the number of goals scored is congruent to 1 modulo 3 (1, 4, 7, etc.) and player 2 wins if the number of goals scored is congruent to 2 modulo 3 (2, 5, 8, etc.). As we saw in the two-player case, we can sum over all the states in which each player wins to calculate their overall probability of winning, and so

$$p_0 = \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{3k}}{(3k)!},$$

$$p_1 = \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{3k+1}}{(3k+1)!},$$

$$p_2 = \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{3k+2}}{(3k+2)!}.$$

More generally, in a game of N players labelled 0 to $N-1$, the winner is determined by the remainder after dividing the total number of goals scored by N. Then

$$p_j = \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{Nk+j}}{(Nk+j)!} \tag{3}$$

is the probability that player j wins. Unlike in the even vs odd case, the sums given by (3) are non-standard and cannot be found in tables. They are less recognisable and require more work than simply reading off the expression. So our simple strategy to settle scores for the two-player case is not so easy to find with additional arguing agents. However, by modelling the time-dependent dynamics of the system, we can find p_j via a different method.

* Graham Hoare Prize 2023 winning article

Time evolution of the tally

For illustration, let us return to the two-player, even–odd case. Assuming that goal scoring is a Poisson process, then at every moment throughout the game, goals are scored at a constant rate $\tilde{\lambda} = \lambda / t_{\text{match}}$, where t_{match} is the duration of the match. Each time a goal is scored, the evenness of the score changes, and so we also switch the current bet winner with rate $\tilde{\lambda}$. If at some time t, there is a $p_{\text{even}}(t)$ chance of there being an even number of goals and a $p_{\text{odd}}(t)$ chance of there being an odd number of goals, then we can write the differential equations describing the evolution of these probabilities, namely

$$\dot{p}_{\text{even}}(t) = -\tilde{\lambda} p_{\text{even}}(t) + \tilde{\lambda} p_{\text{odd}}(t), \tag{4}$$
$$\dot{p}_{\text{odd}}(t) = \tilde{\lambda} p_{\text{even}}(t) - \tilde{\lambda} p_{\text{odd}}(t). \tag{5}$$

Both (4) and (5) have a term describing the probability flux out of the current state and into the opposite state and a term describing the probability flux out of the opposite state and into the current state.

We can write this coupled system as the matrix equation

$$\underbrace{\begin{pmatrix} \dot{p}_{\text{even}} \\ \dot{p}_{\text{odd}} \end{pmatrix}}_{\dot{\mathbf{p}}(t)} = \underbrace{\begin{pmatrix} -\tilde{\lambda} & \tilde{\lambda} \\ \tilde{\lambda} & -\tilde{\lambda} \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} p_{\text{even}} \\ p_{\text{odd}} \end{pmatrix}}_{\mathbf{p}(t)}, \tag{6}$$

where $\mathbf{p}$ is the vector of probabilities and M is the transition matrix. We can find the solution via the matrix exponential to give $\mathbf{p}(t) = \exp(Mt)\mathbf{p}(0)$, where $\mathbf{p}(0)$ is the probability distribution at the start of the match. Each match starts with an even number of goals, and so $\mathbf{p}(0) = (1, 0)^{\mathrm{T}}$. After evaluating the matrix exponential at $t = t_{\text{match}}$, we find that

$$\begin{aligned} \mathbf{p}(t_{\text{match}}) &= \frac{1}{2} \begin{pmatrix} 1 + \mathrm{e}^{-2\lambda} & 1 - \mathrm{e}^{-2\lambda} \\ 1 - \mathrm{e}^{-2\lambda} & 1 + \mathrm{e}^{-2\lambda} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 + \mathrm{e}^{-2\lambda} \\ 1 - \mathrm{e}^{-2\lambda} \end{pmatrix}, \end{aligned}$$

which gives the same values as in (1) and (2).

With more than two players, each time a goal is scored, the current winner iteratively cycles through all the players. Updating the coupled system for the three-player game, (6) becomes

$$\begin{pmatrix} \dot{p}_0 \\ \dot{p}_1 \\ \dot{p}_2 \end{pmatrix} = \begin{pmatrix} -\tilde{\lambda} & 0 & \tilde{\lambda} \\ \tilde{\lambda} & -\tilde{\lambda} & 0 \\ 0 & \tilde{\lambda} & -\tilde{\lambda} \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix}.$$

The diagonal elements describe the rate out of each probability state, and the off-diagonal M_{ij} terms describe the rate of transitioning from winner i to winner j, which is non-zero only for the $j-1 \to j$ transition.

Notice that each row of M is the same, but shifted one entry to the right (with wrapping), as illustrated by Figure 1. This special kind of matrix is called a *circulant* matrix and has the convenient property that the eigenvectors are the same *for any circulant matrix* [1]! Moreover, the eigenvectors are given by the columns of the Fourier matrix

$$\mathcal{F} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \cdots & \omega^{(N-1)^2} \end{pmatrix}, \tag{7}$$

where

$$\omega = \exp\left(-\frac{2\pi \mathrm{i}}{N}\right)$$

is a primitive Nth root of unity. Due to the cyclic nature of our problem, it is unsurprising that the eigenvectors are related to the Fourier transform.

Our matrix can be diagonalised as [1]

$$M = \mathcal{F}^* D \mathcal{F},$$

where

$$D = \operatorname{diag}(\boldsymbol{\mu}),$$

and * denotes the conjugate transpose. Here, the entries of $\boldsymbol{\mu}$ are the eigenvalues of M, given by [1]

$$\mu_j = \tilde{\lambda}\left(\omega^{(N-1)j} - 1\right),$$

for $j = 0, 1, \ldots, N-1$.

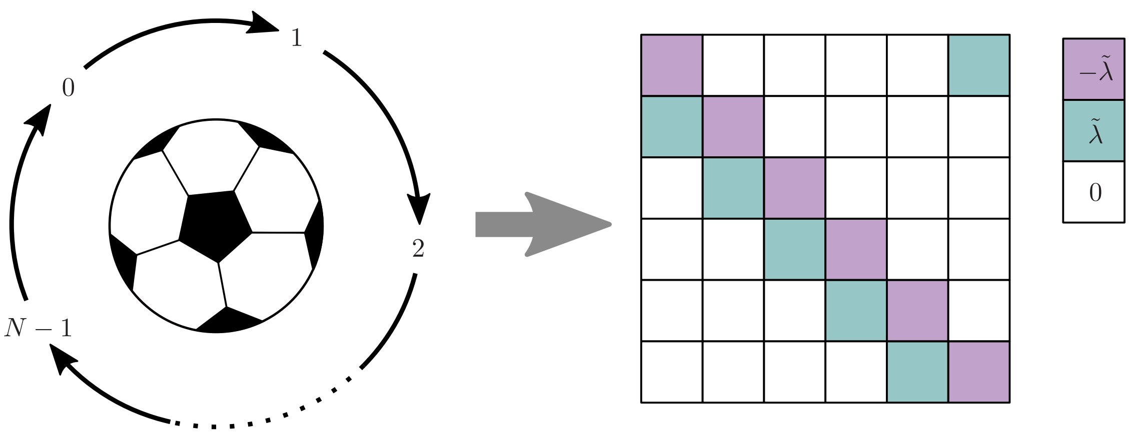
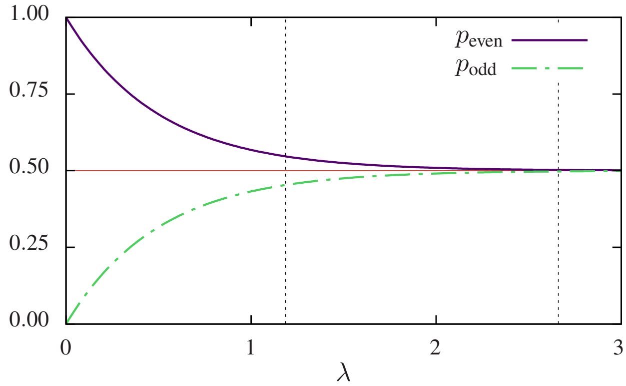
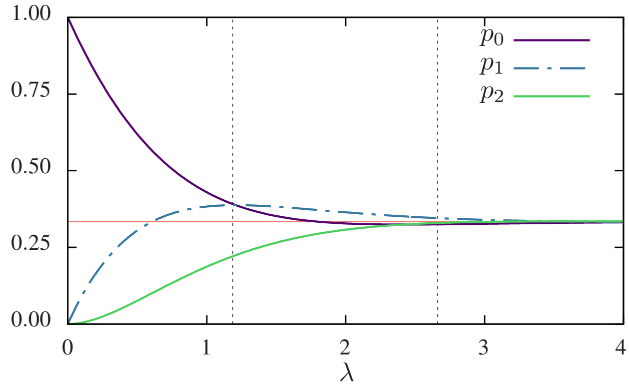


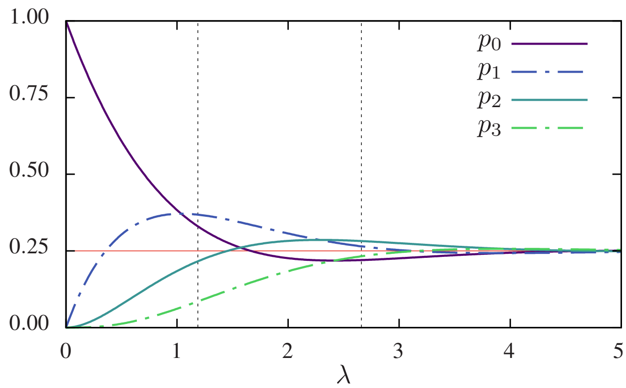
Figure 1: Each goal scored changes the current winner to cycle through all players. This gives the graph on the left, which is described by the transition matrix on the right.



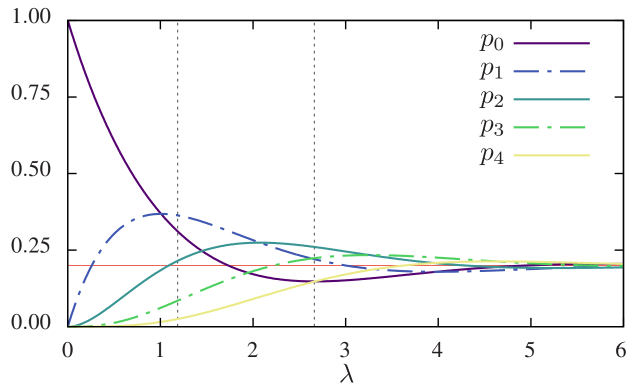
(a) Two players



(b) Three players



(c) Four players



(d) Five players

Figure 2: Probability of each player winning as a function of average number of goals. In the Premier League, an average of 1.19 goals are scored in the first half and 2.66 overall, as shown by the vertical dashed lines.

We may now evaluate the matrix exponential and solve our system, so that

$$\begin{aligned} \mathbf{p}(t) &= \exp(Mt)\mathbf{p}(0) \\ &= \mathcal{F}^* \exp(Dt)\mathcal{F}\mathbf{p}(0). \end{aligned}$$

A match always starts at score 0–0, with player 0 winning, so $\mathbf{p}(0) = (1, 0, \cdots, 0)^{\mathrm{T}}$. Notice from (7) that $\mathcal{F}\mathbf{p}(0)$, is just a column vector of 1's normalised by $\sqrt{N}$, and so

$$\mathbf{p}(t) = \frac{1}{\sqrt{N}} \mathcal{F}^* \exp(\boldsymbol{\mu} t).$$

Therefore, we can find the probabilities, and also determine closed-form expressions for the sums in (3), with a single matrix–vector multiplication! In the three-player case, at the end of the match, the probabilities are

$$\mathbf{p}(t_{\text{match}}) = \frac{1}{3} \begin{pmatrix} 1 + 2\,\mathrm{e}^{-\frac{3\lambda}{2}} \cos\left(\frac{\sqrt{3}\lambda}{2}\right) \\ 1 + \mathrm{e}^{-\frac{3\lambda}{2}} \left(\sqrt{3}\sin\left(\frac{\sqrt{3}\lambda}{2}\right) - \cos\left(\frac{\sqrt{3}\lambda}{2}\right)\right) \\ 1 - \mathrm{e}^{-\frac{3\lambda}{2}} \left(\sqrt{3}\sin\left(\frac{\sqrt{3}\lambda}{2}\right) + \cos\left(\frac{\sqrt{3}\lambda}{2}\right)\right) \end{pmatrix}.$$

We show how the likelihood of each player winning depends on λ in Figure 2 for games with 2–5 players. In the two-player case, we find that the even player is more likely to win, regardless of λ. However, the situation is not as clear cut when we add additional players. In the three-player case, $p_0 < p_1$ for $\lambda > 1.2$. However, all players are essentially equally likely to win for $\lambda > 3$. As we continue to add additional players, the behaviour becomes more complex when the goal scoring rate is low. When the goal scoring rate is high, each player is equally likely to win. However, the goal scoring rate required for a fair game increases with the number of players.

Gambling on goals

Having thought about the theory and made our model, it is wise to check that our strategy provides an advantage when applied to real-world football results. We found Premier League results [2] of $10\,189$ football matches across 26 seasons (1995/96 to 2021/22), with an average of 2.66 goals per match.

We show a histogram of the goals scored in the matches in Figure 3, and this appears to be approximated well by our Poisson prediction with $\lambda = 2.66$ (the mean goals scored). However, the advantage for the even player is more pronounced in the data than in the predictions of our model. In 51.19% of real matches, an even number of goals were scored, leaving 48.81% with an odd number of total goals, compared to our model predictions of 50.25% and 49.75%. Possible reasons for this difference include that the constant rate assumption made in the model may break down near the end of a match, for example, due to substitutions, fatigue or tactical changes made by the teams, such as trying to secure a draw or push for a win.

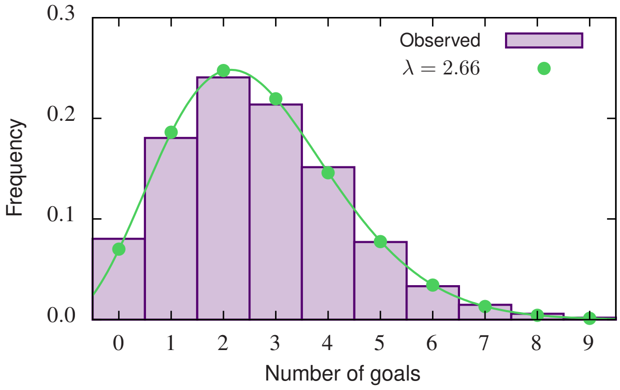


Figure 3: Full-time distribution comparing observed and predicted number of goals.

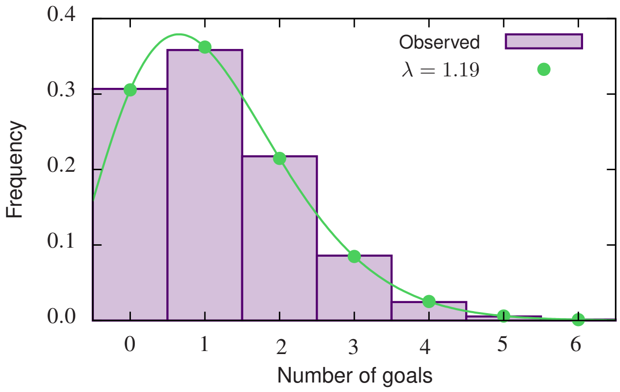


Figure 4: Half-time distribution comparing observed and predicted number of goals.

As a further investigation, we study the distribution of goals at half-time. The assumptions made are less likely to be affected during the first half of a match as fewer substitutions are typically made, players will not be as tired and the pressure of the full-time whistle is still a distant thought. The dataset [2] also includes the scores at half-time, with an average of 1.19 goals in the first half. Comparing the first-half data to the Poisson model, shown in Figure 4, the frequencies matched up almost exactly. The model predicts that 54.66% of matches have an even number of goals in the first half, compared to the observed frequency of 55.01% of an even number of goals.

We can continue our analysis by considering the full-time three-player game. From the data, player 0 wins 32.92%, player 1 wins 34.71% and player 2 wins 32.37% of the time. Our model predicts $\mathbf{p} = (0.3251, 0.3454, 0.3295)^{\mathrm{T}}$. We find a slight advantage for player 1 in both the data and the model. However, at full time, the probabilities are fairly uniform. On the other hand, at the end of the first half, player 0 and player 1 win 39.39% and 38.32% of the time, with player 2 winning only 22.27% of the time. Our model closely predicts these probabilities with $\mathbf{p} = (0.3915, 0.3876, 0.2208)^{\mathrm{T}}$, and we can clearly see the disadvantage of being player 2 when gambling on goals in the first half.

In conclusion, evaluating infinite sums can identify the best strategy when betting with friends on football or other systems that are governed by Poisson processes. For the two-player game, Matthew might have made the right move betting on an even number of goals. However, Brady still has a good chance to win since the advantage is small. The clever play is to bet on an even number of goals in the first half. With more than two players, the strategy is more nuanced, as the player most likely to win depends on λ and the number of players. So, the next time you want to settle a dispute with your friends, consider using the goal parity of a football match.

Notes

1. This article is intended to illustrate a mathematical analysis. It is not intended, in any way, to recommend any form of betting. If you are worried about how gambling makes you or someone else feel, see `www.begambleaware.org`.

References

1 Davis, P.J. (1979) *Circulant Matrices*, Wiley, New York.

2 Kaggle (2023) English Premier League (EPL) results, `www.kaggle.com/datasets/irkaal/english-premier-league-results`.