1 Introduction

We consider the game of *foosball.* A foosball match is between two players, or two teams of two players each, and customarily continues until one of the sides has reached k points (typically $k = 7$ or $k = 10$, depending on the table in use), at which point the game ends.

We are interested in developing a Bayesian rating system, similar to Glicko, but handling the peculiar distribution of the end results of a foosball game. First, we model foosball as a Poisson process, where each player has an unknown *scoring rate* λ, and goals for that player are treated as independent, memoryless events with that rate. (Although modeling the scoring rate only seems to rate the *offensive* skill of a player, it will later be seen that the system is scale-invariant, and as such defensive skill is also naturally incorporated in this rate.) The Poisson process has earlier been shown to be a good model for similar games, such as "real" (non-table) soccer or basketball.

The aim of the rating system will be to estimate the true skill λ for each player. As it is only possible to measure relative skill, however, a different output representation is defined, similar to the Elo rating system in chess.

Assume that a player's true skill is estimated by a Gaussian distribution:

$$\theta \sim \mathcal{N}(\mu, \sigma)$$

Arbitrarily, we define an unknown player's μ to be 1500 and σ to be 350, in line with the Glicko rating system. (Actually measuring the most accurate initial σ is a problem that will not be discussed here.) These ratings are to be understood differentially, in that a player that is about 455 rating points over another player is expected to have twice the scoring rate of the weaker player. This might seem like an arbitrary calibration; however, it shall be seen that this calibration results in a player rated 400 points higher than his/her opponent will win 10/11 of all games on average, which is consistent with the calibration and Bradley-Terry assumption used in Glicko.

Assume that a player with known strength parameters μ_1, σ_1 plays one game (FoosRank does not yet model multiple games in one rating period) against another player with known strength parameters μ_2, σ_2, and the end score is s. (s is here a result such as 10-5 or 6-10 if $k = 10$; two integers, of which one must be k and the other one must be in the range $[0, k-1]$.) Before the game, player 1's probability density function is simply

$$f(\theta_1|\mu_1, \sigma_1) = \Phi(\theta_1|\mu_1, \sigma_1)$$

where Φ is the usual Gaussian pdf with parameters μ_1 and σ_1, that is (unnormalized, which will be the convention from here on as the constant factor does not matter), $e^{-\frac{(x-\mu_1)^2}{2\sigma_1{}^2}}$.

After the game, the marginal posterior probability density function for player 1 (the other player's pdf can be worked out in exactly the same way) is, generally,

$$f(\theta_1|r,\mu_1,\sigma_1,\mu_2,\sigma_2) = \Phi(\theta_1|\mu_1,\sigma_1)\int\limits_{-\infty}^{+\infty} L(s|\theta_1,\theta_2)\Phi(\theta_2|\mu_2,\sigma_2)d\theta_2$$

where $L(s|\theta_1,\theta_2)$ is the likelihood of the result s given a performance of θ_1 and θ_2 from the two players, and $\Phi(\theta_2|\mu_2,\sigma_2)$ is the probability of the performance θ_2 from the other player given the prior skill distribution of $N(\mu_2,\sigma_2)$.

Estimating the likelihood $L(s|\theta_1,\theta_2)$ will be the concern of the following section.

2 Likelihood estimation

As noted before, we choose to model the foosball game as a Poisson process, with the two players having scoring rate λ_1 and λ_2, respectively. (The scoring rate is assumed to be constant throughout the game.) Assume without loss of generality that the first player wins, ie. the score is k to a, where $a < k$. The game length t (the time of the tenth goal from player 1) is governed by the Erlang-k distribution with rate λ_1:

$$P(t = x) = \frac{{\lambda_1}^k x^{k-1} e^{-\lambda_1 x}}{(k-1)!}$$

In t seconds, how many goals (a) does the other player score? The answer is given by the Poisson distribution with rate $\lambda_2 t$, with the probability mass function (pmf)

$$P(a = k) = \frac{e^{-\lambda_2 t}(\lambda_2 t)^k}{k!}$$

Integrating over all possible game lengths t yields the likelihood of the result $k - a$ given the two rates:

$$\begin{aligned} L(a|\lambda_1,\lambda_2) &= \int\limits_0^{\infty} \frac{{\lambda_1}^k t^{k-1} e^{-\lambda_1 t}}{(k-1)!} \frac{e^{-\lambda_2 t}(\lambda_2 t)^a}{a!} dt \\ &= \binom{k+a-1}{k-1} \frac{\left(\frac{\lambda_1+\lambda_2}{\lambda_1}\right)^{-a} {\lambda_1}^{k-a} {\lambda_2}^a}{(\lambda_1+\lambda_2)^k} \end{aligned}$$

As can be readily seen, this likelihood is scale-invariant, in that multiplying the scoring rates λ_1 and λ_2 by a fixed constant C will not affect the expression. Thus, we can arbitrarily set $\lambda_1 = 1$, yielding:

$$L(a|\lambda_1) = \binom{k+a-1}{k-1} \frac{{\lambda_2}^a}{(\lambda_2+1)^{k+a}}$$

which is a *Pascal distribution* with $p = \frac{\lambda_2}{\lambda_2+1}$, recognizable as the classical Bradley-Terry assumption. (The Pascal distribution is also known as the discrete form of the *negative binomial* or *Gamma-Poisson mixture* distribution; the Gamma distribution is the continuous form of the Erlang distribution we started with, so it is easy to see why the latter name was chosen.)

We now apply our initial assumption that $\lambda_2 = \lambda_1 2^{d/455} = 2^{d/455}$, where d is the difference in (performance) rating points between player 1 and player 2 (ie. $d = \theta_2 - \theta_1$):

$$L(a|d) = \binom{k+a-1}{k-1} \frac{(2^{d/455})^a}{(2^{d/455}+1)^{k+a}}$$

However, noting that L only depends on $d = \theta_2 - \theta_1$ allows us to rewrite the original formulation of the posterior probability density functions, now:

$$f(\theta_1|r,\mu_1,\sigma_1,\mu_2,\sigma_2) = \Phi(\theta_1|\mu_1,\sigma_1) \int_{-\infty}^{+\infty} L(s|\theta_2-\theta_1)\Phi(\theta_2|\mu_2,\sigma_2)d\theta_2$$

as a convolution, since $(f \star g)(t) = \int_{-\infty}^{+\infty} f(\tau)g(t-\tau)d\tau$, giving:

$$f(\theta_1|r,\mu_1,\sigma_1,\mu_2,\sigma_2) = \Phi(\theta_1|\mu_1,\sigma_1)(\Phi \star L')(\theta_1)$$

where $L'(d) = L(-d)$. Such a convolution can be computed efficiently by means of the *Fourier transform* of the two involved functions:

$$f(\theta_1|r,\mu_1,\sigma_1,\mu_2,\sigma_2) = \Phi(\theta_1|\mu_1,\sigma_1)(\mathcal{F}^{-1}(\mathcal{F}(\Phi)\mathcal{F}(L')))(\theta_1)$$

Unfortunately, the Fourier transform for $L'(d)$ is not expressible with primitive functions (and likewise, the integral itself). However, the relationship also exists for discrete functions, where the *Fast Fourier Transform* (FFT) can be used to calculate *circular convolutions.* With appropriate zero padding, the two parts of the convolution can be computed numerically, FFTed, multiplied together, and finally IFFTed to give the entire right-hand factor of the expression as a function of θ_1. For calculating the entire interesting segment of $f(\theta_1)$ to the resolution of n points, this technique requires $O(n \log n)$ operations, compared to $O(n^2)$ for naive evaluation of the integral (assuming n points were used to approximate the integral). This saves considerable processing time in the actual estimation of the integral, especially when it is used as part of a larger integral (see below).

3 Team ratings

For teams, we assume $\mu_T = \mu_A + \mu_B$ (the team's total rating is the sum of its two members), which implies $\lambda_T = \lambda_A\lambda_B$. Note that in this process, we double the rating constant to 910, or equally, we introduce a factor 0.5 into μ_T (so the team's total rating is the average instead of the sum).

For players 1 and 2 playing against 3 and 4 (whose such combined skill is termed μ_T), the posterior probability density function for player 1 given the score s becomes

$$f(\theta_1|s, \mu_{1\ldots4}, \sigma_{1\ldots4}) = \Phi(\theta_1|\mu_1, \sigma_1) \int\limits_{-\infty}^{+\infty} \Phi(\theta_2|\mu_2, \sigma_2) \int\limits_{-\infty}^{+\infty} L(s|\theta_1, \theta_2, \theta_T)\Phi(\theta_T|\mu_T, \sigma_T) d\theta_T d\theta_2$$