ELO numbers for random teams

Maximilian Schmidt

March 17, 2015

Contents

1 Calculation of the ELO number

Say we have 4 players $p_i, i \in \{a, b, c, d\}$ with individual ELO numbers x_i which form two teams $t_0 = [p_a, p_b], t_1 = [p_c, p_d]$ with team strenghts X_1, X_0. Then the expectation valye of the game outcome reads

$$E = \frac{1}{1 + 10^{\frac{(X_1 - X_0)}{s}}} , \tag{1}$$

where $s = 400$ in chess (In table tennis, 200 was chosen as a value.). E represents the average outcome of the game after a sufficient amount of samples, where "1" correponds to a victory of team 0 and "0" a victory of team 1. Then, the team strengths are updated as follows:

$$X'_0 = X_0 + k \cdot (S - E) \tag{2}$$

$$X'_1 = X_1 + k \cdot (E - S) \tag{3}$$

where k is a constant which has to be chosen, in chess, $k = 20$. The game thus leads to changes of team strenghts:

$$\delta_i = X'_i - X_i \tag{4}$$

Overall, after the teams have played, the strenghts of the 4 players are updated as follows:

$$
\begin{aligned}
x'_a &= x_a + \delta_0^{(a)} && (5)\\
x'_b &= x_b + \delta_0^{(b)} && (6)\\
x'_c &= x_c + \delta_1^{(c)} && (7)\\
x'_d &= x_d + \delta_1^{(d)} && (8)
\end{aligned}
$$

2 Geometric vs. arithmetic mean

Say we have 4 players $p_i, i \in \{a, b, c, d\}$ with individual ELO numbers x_i which form two teams $t_0 = [p_a, p_b], t_1 = [p_c, p_d]$. The first question is: How to derive team strengths X_0, X_1 for the two teams.

1. Arithmetic mean: The simplest way would be to compute the team strengths as the arithmetic means of the indivual player strengths:

$$
\begin{aligned}
X_0 &= \tfrac{1}{2}\left(x_a + x_b\right) && (9)\\
X_1 &= \tfrac{1}{2}\left(x_c + x_d\right) && (10)
\end{aligned}
$$

2. Geometric mean: Alternatively, one could use the geometric mean.

$$
\begin{aligned}
X_0 &= \sqrt{x_a\, x_b} && (11)\\
X_1 &= \sqrt{x_c\, x_d} && (12)
\end{aligned}
$$

What is the difference between the two methods. 1 shows that for teams with very different individual strengths, the geometric mean is lower than the arithmetic mean while they are exactly equal for equal player strengths. One could argue for two different hypotheses:

- In a kicker team, equal players are more stronger than a team of one strong and one weak player $\Rightarrow$ Geometric mean
- It is more benefitial to have one strong player in the team $\Rightarrow$ Arithmetic mean.

3 How to distribute the change of team strength among the players

After calculating team strengths, we can perform the update rule presented in (17) and get a certain δ_i for each team. The question is: How do we distribute this δ_i among the players? Two different options are possible:

- Distribute δ_i equally among the players:

$$\delta_i^{(j)} = \frac{1}{2}\delta_i \qquad (13)$$

  This ensures that the difference between the players' strenghts remains constant, which seems sensible because in the calculation of expectation

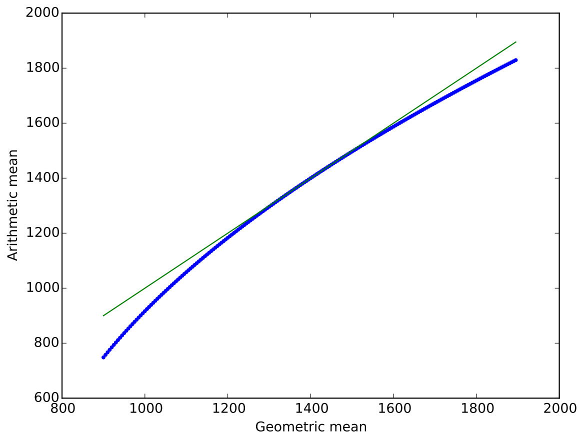

Figure 1: Comparison of geometric vs. arithmetic mean for a team of $x_a = 1400$ and x_b varying from 400 to 2400.

values, it is the difference of ELO numbers which matters and if 2 players play in the same team, the result of the team game does not gives us any information on their relative strengths.

- On the other hand, we can expect that the players' responsabilities for the outcome of the game are not equal. Typically, you would expect that the stronger players is more influential on the outcome, i.e., if the team wins, he should gain more credit for the victory. We can provide for this by conserving not the difference of ELO strengths but their proportion, i.e. by distributing the team increase/decrease δ_i according to the individual strenghts that they brought into the team strength:

$$\delta_i^{(j)} = \frac{x_j}{\sum_j x_j} \delta_i \tag{14}$$

Fig. 2 shows the difference between the two methods. For the first option, both player always get the same change in strengths, no matter whether they are equally strong. For the 2nd option, the stronger player gets a higher increase of strength if the team wins but also a higher decrease if the team loses.

4 Setting of parameters k, s

- s is a simple scale parameter, which was chosen to be 400 for chess for historical reason. Thus, we can simply set it to 1. We adjust the starting value of the ELO numbers to $3.5 = 1400/400$.
- k controls the size of fluctuations from one to another game, where $k = 20/400 = 0.05$ seems to be the most reasonable value, cf. figure 3 on the following page. Note that k is not simply a 'kernel width' because the whole ELO process is a 2nd-order process.

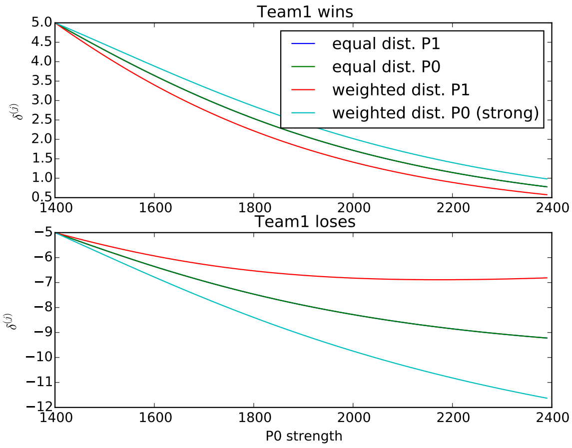

Figure 2: Change of ind. player strengths after their team wins (top) of loses (bottom). The strenght of P1 is always 1400., while the strength of P0 varies from 1400. to 2400., i.e., he is stronger than P1.

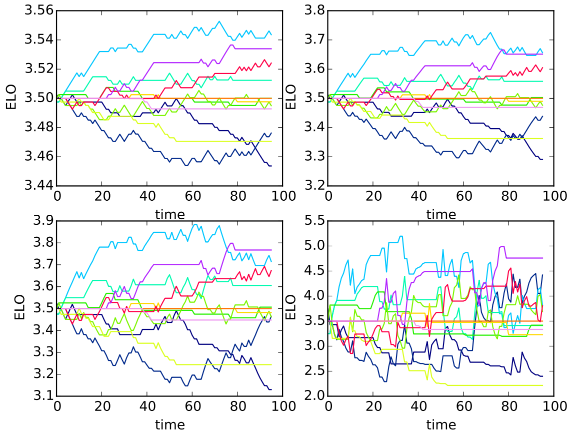

Figure 3: Time development of the ELO numbers for different valyes of k: 0.01 (top left), 0.05 (top right), 0.1 (bottom left) and 1.0 (bottom right)

5 Extension of the ELO update rule

A kicker game does not only yield binary information about winning or losing, but it ends with a goal score. If we don't want to lose this additional information, we can extend the ELO update rule introduced in (equation (17)). If we set the game result to 1 for a win and to 0 for a loss, we basically map the number of scored goals with a heaviside function to 1 or 0. We can extend the information, by mapping the scored goals g as follows:

$$S(g0, g1) = c \cdot \Theta(g0 - g1) + (1 - c) \begin{cases} \frac{1}{6}\,(g0 - g1) & g0 > g1 \\ (1 - \frac{1}{6})\,(g0 - g1) & g0 < g1 \end{cases} \tag{15}$$

$$X_0' = X_0 + k \cdot (S - E) \tag{16}$$

$$X_1' = X_1 + k \cdot (E - S) \tag{17}$$