# A Note on Estimating Elo

Isak Ellmer

December 2023

## 1 Introduction

Suppose that two players are playing a game of chess. We want to make some inference on the difference in strength between the two players. Let $l$, $d$ and $w$ be the probabilities that player 1 loses, draws or wins respectively. It is not hard to see that player 1 is stronger than player 2 if and only if $w > l$. In fact, $d$ does not matter at all, if we only want to know which of the two players is stronger.

Consider however the two following cases. In the first case, $l = d = 0$ and $w = 1$. And in the second case, $l = 0$, $d = 0.99$ and $w = 0.01$. We see that player 1 is stronger than player 2 in both of these cases, but much more so in the first.

To quantify this strength difference we naturally look at the expected score of a game. If a loss is worth 0 points, a draw $\frac{1}{2}$ and a win 1, the expected score is given by $s = \frac{1}{2}d + w$. The closer $s$ is to $\frac{1}{2}$ the more even the games are. And if $s$ is close to 1 then player 1 wins always every game. We see that in the first case $s = \frac{1}{2} \cdot 0 + 1 \cdot 1 = 1$, and in the second case $s = \frac{1}{2} \cdot 0.99 + 1 \cdot 0.01 = 0.505$.

We see that player 1 is stronger than player 2 if and only if $s > \frac{1}{2}$. This is because $l + d + w = 1$ and we then have

$$s = \frac{1}{2}d + w = \frac{1}{2}(d + w) + \frac{1}{2}w = \frac{1}{2} + \frac{1}{2}(w - l).$$

The relationship between expected score and Elo difference [2] is given by

$$\Delta E = S^{-1}(s)$$

where

$$S(x) = \frac{1}{1 + 10^{-x/400}}.$$

The inverse of $S$ is given by

$$S^{-1}(s) = -\frac{400}{\log 10} \log\left(\frac{1}{s} - 1\right).$$

## 2 Estimating Elo

We make the language a little more rigorous. Let $X$ be a random variable taking the values 0, $\frac{1}{2}$ and 1 with probabilities $l$, $d$ and $w$. Suppose that $x_1, ..., x_N$ are independent observations of $X$. The maximum likelihood estimator $\hat{s}$ of $s = E(X) = \frac{1}{2}d + w$ is given by

$$\hat{s} = \frac{1}{N}\sum_{j=1}^{N} x_j.$$

If $\sigma^2$ is the variance of $X$, then by the Central Limit Theorem [1], we have

$$\frac{\hat{s} - s}{\sigma/\sqrt{N}} \to N(0, 1).$$

We want to construct a confidence interval for $s$ of confidence grade $1 - \alpha$. Let $\lambda_{\alpha/2}$ be the upper $\alpha/2$ quantile of $N(0, 1)$. Then

$$\frac{\hat{s} - s}{\sigma/\sqrt{N}} \in \left(-\lambda_{\alpha/2}, \lambda_{\alpha/2}\right)$$

with probability $1 - \alpha$. Equivalently

$$s \in \left(\hat{s} - \frac{\lambda_{\alpha/2}\sigma}{\sqrt{N}}, \hat{s} + \frac{\lambda_{\alpha/2}\sigma}{\sqrt{N}}\right)$$

with the same probability. Since $S$ is an increasing function we easily derive the confidence interval

$$\Delta E \in \left(S^{-1}\left(\hat{s} - \frac{\lambda_{\alpha/2}\sigma}{\sqrt{N}}\right), S^{-1}\left(\hat{s} + \frac{\lambda_{\alpha/2}\sigma}{\sqrt{N}}\right)\right)$$

for $\Delta E = S^{-1}(s)$. Note however that unless $\hat{s} = 0$, $\Delta\hat{E}$ will not lie in the middle of this interval.

### 2.1 Centered Approximate Confidence Intervals

We will take two approaches. The first approach assumes that the Elo difference is small and that $N$ is large. The second approach only assumes that $N$ is large.

We begin with the first approach. We have $s = S(\Delta E)$. We expand $S$ in its first order Taylor polynomial around 0. We get

$$s = S(\Delta E) = S(0) + \Delta E S'(0) = \frac{1}{2} + \Delta E \frac{\log 10}{1600}.$$

Solving for $\Delta E$ gives

$$\Delta E = \frac{1600}{\log 10}\left(s - \frac{1}{2}\right).$$

Starting with the confidence interval

$$s \in \left( \hat{s} - \frac{\lambda_{\alpha/2}\sigma}{\sqrt{N}}, \hat{s} + \frac{\lambda_{\alpha/2}\sigma}{\sqrt{N}} \right)$$

from Section 2, we subtract $\frac{1}{2}$ and multiply by the constant $\beta = \frac{1600}{\log 10}$. This gives

$$\beta \left( s - \frac{1}{2} \right) \in \left( \beta \left( \hat{s} - \frac{1}{2} \right) - \frac{\beta\lambda_{\alpha/2}\sigma}{\sqrt{N}}, \beta \left( \hat{s} - \frac{1}{2} \right) + \frac{\beta\lambda_{\alpha/2}\sigma}{\sqrt{N}} \right).$$

The assumption that $\Delta E$ is small now gives

$$\Delta E \in \left( \Delta\hat{E} - \frac{\beta\lambda_{\alpha/2}\sigma}{\sqrt{N}}, \Delta\hat{E} + \frac{\beta\lambda_{\alpha/2}\sigma}{\sqrt{N}} \right).$$

For the second approach we start from

$$\Delta E \in \left( S^{-1} \left( \hat{s} - \frac{\lambda_{\alpha/2}\sigma}{\sqrt{N}} \right), S^{-1} \left( \hat{s} + \frac{\lambda_{\alpha/2}\sigma}{\sqrt{N}} \right) \right).$$

We now instead expand $S^{-1}$ in its first order Taylor polynomial around $s = \hat{s}$. We simply get

$$S^{-1} \left( \hat{s} \pm \frac{\lambda_{\alpha/2}\sigma}{\sqrt{N}} \right) = S^{-1} \left( \hat{s} \right) + \left( \hat{s} \pm \frac{\lambda_{\alpha/2}\sigma}{\sqrt{N}} - \hat{s} \right) S^{-1\prime} \left( \hat{s} \right) =$$

$$\Delta\hat{E} \pm \frac{\lambda_{\alpha/2}\sigma}{\sqrt{N}} \frac{1}{S' \left( \Delta\hat{E} \right)}.$$

This implies that

$$\Delta E \in \left( \Delta\hat{E} - \frac{\lambda_{\alpha/2}\sigma}{\sqrt{N}} \frac{1}{S' \left( \Delta\hat{E} \right)}, \Delta\hat{E} + \frac{\lambda_{\alpha/2}\sigma}{\sqrt{N}} \frac{1}{S' \left( \Delta\hat{E} \right)} \right).$$

# References


[1] Wikipedia contributors. Central limit theorem — Wikipedia, the free encyclopedia, 2023. `https://en.wikipedia.org/w/index.php?title=Central_limit_theorem&oldid=1188038646`.

[2] Wikipedia contributors. Elo rating system — Wikipedia, the free encyclopedia, 2023. `https://en.wikipedia.org/w/index.php?title=Elo_rating_system&oldid=1185663969`.