Are ELO Rating tables false ?

Jean-Marc Alliot

Abstract

According to an article in the French magazine "Europe Echecs", tables used for calculating ELO ratings are false. I myself made the same constatations a few years ago. But is there really an error ?

1 Introduction

A few years ago, I had to write a program for my chess club for calculating ELO ratings. I had only a few notions about the way ELO ratings were calculated. Hopefully some persons were able to help me on the net. Mr Blair sent me Dr Arpad Elo's book reference, and professor Hans Van Staveren sent me its program which did exactly what I wanted to do. I use this opportunity to thank them for their help. But working on the problem, I noticed some strange things in the values used in ELO tables. An article in the last issue of "Europe Echecs" shows exactly the same result.

2 How to calculate ELO ratings

I'm first going to explain how ratings are calculated. Chess ratings are calculated according to a standard Gauss distribution, i.e if the difference (in ELO points) of two players is D, then the probability of winning a game is:

$$P(D) = \frac{1}{\sqrt{2}\,\pi} \int_{-\infty}^{\frac{D}{\sqrt{2}\,\sigma}} e^{-t^2/2}\,dt$$

σ is 400 (by convention) in the ELO system.

It is interesting here to put: $u^2 = t^2/2$. The expression is then:

$$P(D) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\frac{D}{\sigma}} e^{-u^2}\,du$$

It is impossible to calculate $I(x) = \int_0^x e^{-u^2}\,du$, but we can calculate $I = \int_0^{+\infty} e^{-u^2}\,du$ quite easily. We just use polar coordinates and express that (x, y) with $0 < x < +\infty$ and $0 < y < +\infty$ is the same as ρ, θ with $0 < \rho < +\infty$ and $0 < \theta < \pi/2$ which both represent the upper left quarter of the (x, y) space:

$$I \times I = \int_0^{+\infty} e^{-x^2}\,dx \times \int_0^{+\infty} e^{-y^2}\,dy = \int_{x=0}^{+\infty}\int_{y=0}^{+\infty} e^{-(x^2+y^2)}\,dx\,dy$$

If we wanted a real demonstration, we would have to be a little longer on this. And now, we just pose $x = \rho\cos\theta, y = \rho\sin\theta$ and we have:

$$\int_{\rho=0}^{+\infty}\int_{\theta=0}^{\pi/2} e^{-\rho^2}\,\rho\,d\rho\,d\theta = [\theta]_0^{\pi/2}[-e^{-\rho^2}/2]_0^{+\infty} = \pi/4$$

So: $I \times I = \pi/4$ and $I = \sqrt{\pi}/2$.

$$\int_{-\infty}^{0} e^{-u^2}\,du = \int_{0}^{+\infty} e^{-u^2}\,du = \frac{\sqrt{\pi}}{2}$$

You can easily now check that when $D \to \infty$, then $P(D) \to 1$: P is a valid distribution.

We can rewrite the expression:

$$P(D) = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\frac{D}{\sigma}} e^{-u^2}\,du = 1/2 + \frac{1}{\sqrt{\pi}}\int_{0}^{\frac{D}{\sigma}} e^{-u^2}\,du$$

Now, to calculate $P(D)$, we can develop e^{-u^2} (we all remember that $e^x = \sum_{n=0}^{+\infty}\frac{x^n}{n!}$):

$$P(D) = 1/2 + \frac{1}{\sqrt{\pi}}\int_0^{\frac{D}{\sigma}}\sum_{n=0}^{+\infty}(-1)^n\frac{u^{2n}}{n!} = 1/2 + \frac{1}{\sqrt{\pi}}\sum_{n=0}^{+\infty}(-1)^n\frac{(D/\sigma)^{2n+1}}{(2n+1)n!}$$

Using modern computers, it is easy to calculate very accuratly $P(D)$ when D is given. But a few years ago things were not that easy. So Dr Arpad Elo gave in his book a table which made the association between D and $P(D)$:

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.5	3	10	17	25	32	39	46	53	61	68
0.6	76	83	91	98	106	113	121	129	137	145
0.7	153	162	170	179	188	197	206	215	225	235
0.8	245	256	267	278	290	302	315	328	344	357
0.9	374	391	411	432	456	484	517	559	619	735

Using this table is very easy. If the difference between two players is, for example, 240 ELO points, you look in the table for the number n such as n is the *smallest number* verifying $n > 240$. Here you find $n = 245$, so the expectancy in this game is $0.8 + 0.00 = 0.80$. **This table is the one used in all Chess federations all over the World**.

But problems arise when you want to calculate the table yourself using the formula given above.

I wrote a program which does the calculation. It is in appendix. This is the output of this program:

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.5	3	10	17	24	31	39	46	53	60	68
0.6	75	82	90	97	105	112	120	128	136	144
0.7	152	160	169	177	186	195	204	213	223	233
0.8	243	253	264	275	287	299	311	325	339	354
0.9	370	388	407	428	452	479	512	554	613	728

As you can see, the two tables are different. The first time I saw the result, I could not believe it, and I thought that I had made a mistake somewhere and gave up. But, this month appeared an article in “Europe Echecs”. The author (Jean-François Hunon, 6 avenue Franklin Roosevelt, 94300 Vincennes, France) publishes an ELO table that is, according to him, corrected, and this table is exactly the table I calculate.

The question is: **Are standard tables really false?**

3 Consequences

I'm going to show how this error (if any) can modify seriously the ratings of all players.

When you participate in a tournament, the modification of you ELO ranking is given by the formula:

$$\text{NEWELO} = \text{OLDELO} + K(P_s - N * P(D))$$

K is a constant which depends on the tournament, among other things, P_s is the number of points you scored, N is the number of games you played and $P(D)$ is your expectancy in this tournament according to the average of the ELO ranking of all your opponents. $P(D)$ is calculated using Dr Elo's table.

Let's have a look at an example: suppose you participated in a 10 games tournaments. You scored 6 points, your ELO was 1900 and the average ELO of your opponents was 1600 (bad week-end, isn't it?). Let's suppose $K = 25$. According to the old table, you should lose:

$$25 \times (6 - 10 \times 0.85) = 25 \times -2.5 = -62.5$$

According to the new table:

$$25 \times (6 - 10 \times 0.86) = 25 \times -2.6 = -65$$

So high rated players lose less when they do bad performances.

Now, suppose you participated in a 10 games tournaments. You scored 9.5 points, your ELO was 1900 and the average ELO of your opponents was 1600 (Better than Ivantchouk against Youdassine!). Let's suppose $K = 25$. With the old table:

$$25 \times (9.5 - 10 \times 0.85) = 25 \times 1 = 25$$

With the new table:

$$25 \times (9.5 - 10 \times 0.86) = 24 \times 1 = 24$$

And high rated players win more when they do a good performance. Of course, this error is cumulative for all tournaments and all players.

Consequences are easy to understand. ELO ranking of top players are getting higher and higher each year, and the real values of ELO ratings are perverted.

I would be very interested in other people's opinions on this subject. (BTW: my interest is purely theorytical, I gave up tournament chess years ago!)

A A program to compute a new table

```
#define PI 3.141592654
#include <stdio.h>
#include<math.h>

  double proba(d)
double d;
{
  double s,t,t1;
  long int i;
  i=1;t1=d/400.0;s=t1;
  do
    {
      t1= t1*d/400.0*d/400.0/i; t = t1/(2*i+1);
      if (i%2==0) s+=t; else s-=t;
      i=i+1;
    }
  while(t>0.0001);
  s=s/sqrt(PI)+0.5;return s;
}

main()
{
  double s=0.505,res,delta;
  long int d;
  for (d=0;d<740;d+=1)
    {
      res=proba((double)d);delta=s-res;
      if (delta < 0.0)
        {
          printf("%ld;",d-1);
          s=s+0.01; if (s>=1.0) break;
        }
    }
}
```