www.simiode.org

SIMIODE Systemic Initiative for Modeling Investigations and Opportunities with Differential Equations

STUDENT VERSION

Rating Chess Players with Elo's Method

Hope McIlwain
Department of Mathematics
Mercer University
Macon GA USA

STATEMENT

Activity Instructions

This activity provides an example of writing a difference equation as a mathematical model, using qualitative techniques to interpret the model, and solving the model with simulated data. At the end of the activity, each group should submit one paper with answers to the questions. Answers to questions should be labeled clearly and should be in complete sentences with proper grammar. All mathematical calculations should be explained.

Part 1: Developing a Model for Elo's Method

The process of rating and then using those ratings to rank objects is becoming increasingly important and many mathematical methods have been developed to rate and then rank objects. For example, search engines such as Google use rating and ranking methods to determine the type and ordering of web pages appear in search results [2].

One such method is Elo's Method. Developed by physics professor and chess player Arpad Elo, the method was originally devised to rate chess players. His ratings are called **Elo ratings**.

Chess Tournaments and Elo Ratings

The goal of Elo ratings is to provide a measure for a player's chess-playing ability. In a typical chess tournament, each player enters the tournament with an Elo rating based on previous game performance. As we will see later, the ratings are based on game scores. For each game, the winning player scores 1 and the losing player scores 0. If a game ends in a tie, then each player scores $\frac{1}{2}$. A player's performances (or score) can be measured with be a single game score or by amalgamating the score acquired during tournament play. Elo ratings range from 0 to 2800 or above [1].

Tournament play can be set up several different ways. A round-robin tournament involves each player playing each other player a game of chess with each other player. For example, if 8 players enter a round robin tournament, each player plays 7 games. If there are an odd number of players, one player gets a bye in each round. In these tournaments, each set of opponents plays exactly one game. A second type of tournament is a Swiss (or Swiss-pairing) system. Although specific rules vary from tournament to tournament, this format is generally played as a non-elimination tournament but in fewer rounds than a round-robin tournament. The players are initially sorted from highest rating to lowest rating where unrated players are listed at the bottom in alphabetical order. In the first round the players are divided into two groups by ranking. Each player in the top half of the rankings plays against a player in the bottom half of the rankings in the same relative position. For example, if there are 8 players in a tournament the players are placed into a ranked list and divided into two groups of 4. Player # 1 then plays Player #5, Player #2 then plays Player #6, etc. For subsequent rounds, players are then re-ranked and sorted into pairs for the next rounds. [4].

Modeling Elo's Method

For purposes of our model, we first list our players in alphabetical order.

- Let t be a unit of time. Let k be used to index the games played. Then t_k is the time at the start of game k and t_{k+1} is the time at the completion of game k (and the start of game $k+1$).
- Let $r_i = r_i(t)$ be the ranking of a player at time t.
- Let $S_{ij} = S_{ij}(t_{k+1})$ to be the score of player i against player j in game k. More details about S_{ij} are given after Question 2.
- Let $\mu_{ij} = \mu_{ij}(t_k)$ be the expected performance of player i against player j at the start of game k (where game k is a game between player i and player j). After Question 2 we will give a specific expression for $\mu_{ij}(t_k)$.

When players i and j play a game against each other, their ratings at the beginning of a game at an initial time t_k can be given by $r_i(old) = r_i(t_k)$ and $r_j(old) = r_j(t_k)$. After the game is played an the results are known, the ratings at the end of the game at time t_{k+1} are updated to become $r_i(new) = r_i(t_{k+1})$ and $r_j(new) = r_j(t_{k+1})$. Elo thought that the difference in ratings $r_i(new) - r_i(old) = r_i(t_{k+1}) - r_i(t_k)$ should be proportional to the difference between the actual performance $S_{ij}(t_{k+1})$ of player i against player j in game k and the expected performance μ_{ij} of player i against player j.

Question 1:

1. Using Elo's assumption that the difference in ratings is proportional to how a player's performance in game k varies from his or her expected performance, write a difference equation to

express the relationship $r_i(\text{new}) - r_i(\text{old})$ when player i plays a game against player j. Use K as your (positive)constant of proportionality.

2. Given the initial ratings $r_i(t_k) = r_{ik}$ and $r_j(t_k) = r_{jk}$ for players i and j at the beginning of game k respectively, write initial value problems that models the change in player i's rating and player j's ratings.

Qualitative Analysis of the Model

Question 2: For this question, use the difference equation from Question 1.

1. Suppose that at the conclusion of game k, $r_i(\text{new}) - r_i(\text{old}) = 0$. What can you say about the player i's performance against player j in game k?
2. Suppose that at the conclusion of game k, $r_i(\text{new}) - r_i(\text{old}) > 0$. What can you say about the player i's performance against player j in game k?
3. Suppose that at the conclusion of game k, $r_i(\text{new}) - r_i(\text{old}) < 0$. What can you say about the player i's performance against player j in game k?

Part 2: A System of Difference Equations

The elegance of Elo's Method is that it rewards a weak player for beating a strong player more than it rewards a strong player for defeating a weak player [3]. We will explore this result in Question 3.

In a game between player i and player j, the change in ratings of each player i and j is given by separate difference equations, $r_i(\text{new}) - r_i(\text{old}) = f(S_{ij}, \mu_{ij})$ and $r_j(\text{new}) - r_j(\text{old}) = f(S_{ji}, \mu_{ji})$ where $f(S_{ij}, \mu_{ij})$ is the function from the first part of Question 1.

In chess ratings, the performance function $S_{ij}(t)$ is defined by

$$S_{ij}(t) = \begin{cases} 1 & \text{if } i \text{ beats } j \\ 0 & \text{if } i \text{ loses to } j \\ \frac{1}{2} & \text{if } i \text{ and } j \text{ tie} \end{cases} \tag{1}$$

An analysis of chess scores determined that the values of $\mu_{ij}(t)$ are a logistic function of the difference between the previous rating of player i and the previous rating of his or her opponent. In fact, for a game between players i and j at time $t = t_k$, $\mu_{ij}(t_{k+1})$ is often given by

$$\mu_{ij}(t_{k+1}) = \frac{1}{1 + 10^{-\frac{(r_i(t_k) - r_j(t_k))}{400}}}. \tag{2}$$

where $r_i(t_k)$and $r_j(t_k)$ represent the ratings of player i and player j at $t = t_k$ [3].

An Example of Player Interaction

Question 3: Suppose Player A is a strong player with a rating $r_A(t_k)$ at time $t = t_k$ of 1900. Suppose Player B is a weak player with a rating $r_B(t_0)$ at time $t = t_k$ of 1500. The goal of this question to demonstrate how Elo's Method rewards a stronger player less for defeating a weaker player than Elo's Method rewards a weaker player for defeating a stronger player.

1. (a) What is $\mu_{AB}(t_{k+1})$? Round your answer to two decimal places.
   (b) If player A defeats player B at time $t = t_k$, what is $S_{AB}(t_{k+1})$?
   (c) If player A defeats player B at time $t = t_k$, what is $r_A(\text{new}) - r_A(old)$ at $t = t_{k+1}$? Your answer should be in terms of K.
2. (a) What is $\mu_{BA}(t_{k+1})$? Round your answer to two decimal places.
   (b) If player B defeats player A at time $t = t_{k+1}$, what is $S_{BA}(t_1)$?
   (c) If player B defeats player A at time $t = t_{k+1}$, what is $r_B(\text{new}) - r_B(\text{old})$ at $t = t_{k+1}$? Your answer should be in terms of K.
3. Explain how your answers in the previous parts demonstrate how Elo's Method rewards a stronger player less for defeating a weaker player than Elo's Method rewards a weaker player for defeating a stronger player.

Tournament Simulation

Question 4: Starting with eight players, do a tournament simulation using one of the available pieces of software described below.

- **Excel**: The accompanying Excel worksheet, 1-95-T-Excel-RatingChessPlayers-TeacherVersion, which is in the Supplemental Docs for this modeling scenario models a Round Robin tournament with eight players. On this question, your instructor will give a list of 8 players with an initial set of Elo ratings.In this simulation, the K value is 15. The results of each game are generated randomly. Enter the initial set of ratings onto the front tab of the file. The ratings will be automatically generated.

  Generate a plot of the ratings for each player with the round played on the horizontal axis and the ratings on the vertical axis. In a brief paragraph, describe the outcome of the simulation. Pay particular attention to when the ordering of the ratings (and thus rankings) change.
- **Mathematica**: The accompanying Mathematica worksheet gives a more thorough simulation. There are two examples in this simulation. In the first part, 8 players are randomly assigned initial ratings from 10 to 2000. They then play 50 games in which the player with the stronger rating always wins. In this simulation, the K value is 30. At the end of the first part, the progression of the player's ratings through 50 games is plotted. The simulation is then repeated with a narrower band of initial ratings from 1500 to 2000.

In a brief paragraph, describe the outcome of the simulation. Pay particular attention to when the ordering of the ratings (and thus rankings) change.

REFERENCES

[1] Daniel Ross. 2007. Arpad Elo and the Elo Rating System. *Chess News*. ChessBase GmbH. 16 December 2007. `http://en.chessbase.com/post/arpad-elo-and-the-elo-rating-system` Accessed 16 August 2016.

[2] Amy K. Langville and Carl D. Meyer. 2006. *Google's PageRank and Beyond: the Science of Search Engines and Beyond*. Princeton NJ: Princeton University Press.

[3] Amy K. Langville and Carl D. Meyer. 2012. *Who's #1? The Science of Rating and Ranking*. Princeton NJ: Princeton University Press.

[4] World Chess Federation. 2016. `www.fide.com`.Accessed 16 August 2016.