Computing the Invariant Measure for a Two-Player Elo Rating System

David Arturo Man Castillo

M1 - Master in Mathematical Engineering

University Côte d'Azur

Supervisor: Prof. Stéphane Junca

July 2, 2024

Acknowledgements

I would like to express my gratitude to my supervisor, Dr. Stéphane Junca, for his invaluable guidance, insightful feedback, and support throughout this research. His expertise and dedication were instrumental in shaping this thesis and driving it to completion.

I am profoundly thankful to my family for their continuous encouragement, love, and patience. Their belief in me has been a constant source of motivation and strength.

To my partner, María Camila, thank you for your unwavering support, understanding, and encouragement. Your love and patience have been a source of strength for me.

I would also like to extend my sincere thanks to my colleagues and friends who provided moral support and companionship during this journey. Your encouragement and assistance have been greatly appreciated.

A special thanks to the Université Côte D'Azur, for providing an excellent academic environment and the resources necessary for conducting this research.

Lastly, I am grateful to all those who, directly or indirectly, have contributed to the successful completion of this thesis. Your support and contributions have been deeply appreciated.

David Man

Abstract

The Elo rating system, developed by Arpad Elo, is a prominent method for evaluating player skill levels in two-player games. This master thesis investigates the stationary distribution of player ratings in the Elo model, explores the "Elo conjecture," and examines the distribution's characteristics, including symmetry and expected value - variance dependence or independence on K.

We developed a numerical algorithm to approximate the cumulative distribution function (CDF) of the Elo distribution. Through extensive simulations, we validated the algorithm's precision and applicability. Our study included a grid convergence analysis, symmetry tests,"Elo conjecture" analysis and comparisons with the normal CDF. The results confirm the algorithm's accuracy and suggest that the Elo's assumption about the expected match outcome and the rating difference may be accurate. Symmetry tests revealed that the Elo distribution might be symmetric around its mean for most configurations, aligning with theoretical expectations.

Investigating the parameter K, which influences rating adjustment speed, revealed that while the expected rating difference appears independent of K for certain probabilities, the variance shows a linear dependence on K. These insights enhance the understanding of the Elo rating system's dynamics and applicability.

This thesis highlights areas for further research, including more precise numerical techniques, stochastic algorithm comparisons, and extending the model to include ties. Our findings contribute to a deeper theoretical and numerical understanding of the Elo rating system and its robustness across various competitive settings.

Keywords: Elo rating system, stationary distribution, cumulative distribution function, symmetry, normal distribution, numerical algorithm, rating dynamics, grid convergence, Markov chains.

Contents

Chapter 1

Introduction

1.1 Background and Problem Statement

The Elo rating system, developed by Arpad Elo, is a widely used method for calculating the relative skill levels of players in two-player games such as chess. Elo's system has been adopted by the World Chess Federation (FIDE) and has seen applications in various other sports and competitive activities due to its simplicity and reliability [1]. The system updates players' ratings based on match outcomes, reflecting their current performance levels.

Despite its widespread use, the Elo rating system has certain limitations and unresolved problems. This study aims to explore the stationary distribution of player ratings in the Elo system and investigate Elo's assumption that the expected score depends on the difference of the rating [1]. The primary objective is to develop an algorithm that approximates the cumulative distribution function (CDF) of the Elo distribution starting from the Elo dynamic equations, and to evaluate the precision and applicability of this scheme.

1.2 Objectives

The primary objectives of this study are:

- To analyze the mathematical constructs of the Elo rating system and its update mechanism.
- To investigate the stationary distribution of player ratings within the Elo model.
- To numerically approximate the invariant distribution of the Elo rating system dynamics.

- To numerically demonstrate and verify Elo's assumption about the expected score and the difference of the ratings.
- To explore characteristics of the Elo distribution such as its symmetry, similarity to the normal distribution, and the dependence/independence of $E[\Delta_{ij}]$ and $V[\Delta_{ij}]$ with respect to K.

1.3 Structure of the Report

This report is structured as follows:

- **Literature Review** - Summarizes existing research and theories related to the Elo rating system and its applications, including works by Avdeev on the stationary distribution [2], and Krifa et al. on the convergence properties of the Elo rating system [3].
- **Chapter 2: Theoretical Framework** - Explains the Elo rating system dynamics, the definition and significance of the random variable Δ_{ij}^n, and how the dynamics can be viewed as a Markov chain process. It also details the original Elo conjecture and how it is framed in this study.
- **Chapter 3: Methodology** - Describes the research design and analytical techniques used in the study. This includes a method for numerically approximating the cumulative distribution function (CDF) of the Elo rating system starting from its dynamic equation, testing the convergence of the scheme, and analyzing properties of the numerical approximation.
- **Chapter 4: Results** - Presents the findings from the numerical approximation of the Elo rating system's stationary distribution and the numerical demonstration of the Elo conjecture. This includes graphical representations of the numerical approximations.
- **Chapter 5: Discussion and Conclusion** - Interprets the results, compares them with existing literature, discusses limitations, suggests areas for future research, and summarizes the study's key findings and their implications.

1.4 Literature Review

Foundational Work on the Elo Rating System Arpad Elo's seminal work, "The Rating of Chess Players, Past and Present," laid the groundwork for the rating system that bears his name [1]. Elo's system was designed to provide a more accurate and objective measure of a player's skill by updating ratings based on match outcomes. The key innovation of Elo's system was the dynamic adjustment of ratings, which allowed for continuous tracking of a player's performance over time.

Theoretical Developments Few theoretical studies have explored the mathematical properties of the Elo rating system. One key area of interest is the system's ability to achieve a stationary distribution, where player ratings stabilize over time. This stationary distribution is crucial for understanding the long-term behavior of the rating system and its ability to rank players accurately [2].

Krifa et al. investigated the convergence properties of the Elo rating system and demonstrated that the system correctly sorts players' strengths in round-robin tournaments [3]. Junca's work on contractions in the Elo rating update process further explored the conditions under which the ratings converge [6]. Aldous also highlighted the need for foundational research on the accuracy of Elo-type rating algorithms and questioned the common claim that 30 matches suffice to rank players accurately [4].

Applications and Extensions The Elo rating system has been adapted and extended for use in various other sports and competitive activities beyond chess. These adaptations often involve modifications to the basic rating update equations to account for specific characteristics of different games [4].

The Elo rating system have been applied to sports such as soccer, tennis, and online gaming. Each adaptation typically involves tweaking the rating update formulas to better reflect the scoring rules and match formats of the respective sports. Jabin and Junca developed a continuous model for ratings, providing a new perspective on the evolution of player ratings in large populations [7].

Challenges and Unresolved Questions Despite its widespread use, the Elo rating system faces several challenges and unresolved questions. One of the primary challenges is the lack of a comprehensive computation of the Elo distribution. While there have been efforts to approximate the stationary distribution of player ratings, a complete understanding of its form remains elusive [1]. Additionally, the symmetry properties of the Elo distribution are not fully understood, except for specific cases where the winning probability p is 0.5, as shown by Avdeev [2].

Another significant question pertains to Elo's assumption regarding the expected outcome of a match and its dependence on the rating difference [1]. This study aims to contribute by exploring the stationary distribution of ratings and examining whether the expected score can be accurately approximated by the difference of the rating.

Furthermore, the Elo rating system's dependence on parameters such as K, which controls the rate of rating adjustment, needs further exploration. Understanding the influence of these parameters on the distribution of player ratings is crucial for improving the accuracy and reliability of the rating system. Future research should also investigate the potential for using stochastic algorithms to provide a comparison with the deterministic approach developed in this work.

Chapter 2

Theoretical Framework

This chapter discusses the theoretical basis of the Elo rating system, its dynamics, and the mathematical constructs used in this thesis. We will explore the system's behavior, the concept of the rating difference as a random variable, and our reformulation of Arpad Elo's assumptions into what we term the Elo conjecture. The discussion is grounded in the works of Krifa, Spinelli, and Junca (2021) [3], Junca (2021) [6], and Jabin and Junca (2015) [7].

2.1 Elo Rating System Dynamics for Two Players

The Elo rating system for two players is designed to calculate the relative skill levels of players in competitor-versus-competitor games such as chess. The key equations that describe the update mechanism for players' ratings after a match are as follows:

2.1.1 Rating Update Equations

For two players, i and j, the ratings after a match are updated using the equations:

$$R_i^{n+1} = R_i^n + K\left(S_{ij}^n - b(R_i^n - R_j^n)\right) \tag{2.1}$$

$$R_j^{n+1} = R_j^n + K\left(S_{ji}^n - b(R_j^n - R_i^n)\right) \tag{2.2}$$

where $S_{ij}^n + S_{ji}^n = 1$ represents the outcome of the game, and b is the bonus-malus function, which is bounded by $b \in [0, 1]$ [3].

The modern bonus-malus function, primarily the logistic Elo function, used in practice is defined as:

$$b(x) = \frac{1}{1 + 10^{-x/400}} \tag{2.3}$$

2.1.2 Rating Difference

The difference in ratings and its update are defined as:

$$\Delta_{ij}^{n} = R_i^n - R_j^n \tag{2.4}$$

$$\Delta_{ij}^{n+1} = \Delta_{ij}^{n} + 2K\left(S_{ij}^n - b(\Delta_{ij}^n)\right) \tag{2.5}$$

The sum of Elo ratings is conserved:

$$2R_i^n = R_i^0 + R_j^0 + \Delta_{ij}^n \tag{2.6}$$

$$2R_j^n = R_i^0 + R_j^0 - \Delta_{ij}^n \tag{2.7}$$

The update process for the Elo ratings can be modeled as a Markov chain, where the state of the system at any given time is represented by the rating difference Δ_{ij}^n. The transition probabilities depend on the outcome of each game and the current rating difference [7].

Define the increasing function:

$$g(x) = x - 2Kb(x) \tag{2.8}$$

where $K > 0$ is a constant and $b(x)$ is the bonus-malus function. This function is key in updating the CDF for the next game. The inverse function $h(x) = g^{-1}(x)$ exists if $g(x)$ is strictly increasing, requiring $g'(x) > 0$ for all x. The derivative is:

$$g'(x) = 1 - 2Kb'(x) \tag{2.9}$$

Given $b'(x)$ from (2.3), at $x = 0$,

$$b'(0) = \frac{\ln(10)}{1600} \approx 0.00144 \tag{2.10}$$

Thus, for $g'(x) > 0$,

$$1 - 2K \cdot \frac{\ln(10)}{1600} > 0 \quad \Rightarrow \quad K < \frac{800}{\ln(10)} \approx 347.4 \tag{2.11}$$

This ensures $g(x)$ is strictly increasing and invertible [6].

2.2 Cumulative Distribution Function (CDF)

Let p and q be the probabilities of winning and losing respectively, with $p + q = 1$. The CDF of the Elo distribution after n games is:

$$F_n(x) = P(\Delta^n \leq x) \tag{2.12}$$

To update the CDF for the next game, we consider the law of total probabilities. By applying this law, we develop the following recurrence relation:

$$\begin{aligned} F_{n+1}(x) &= p \cdot P(\Delta_{ij}^{n+1} \leq x \mid S_{ij}^n = 1) + q \cdot P(\Delta_{ij}^{n+1} \leq x \mid S_{ij}^n = 0) \\ &= p \cdot P(\Delta_{ij}^n + 2K(1 - b(\Delta_{ij}^n)) \leq x) + q \cdot P(\Delta_{ij}^n + 2K(0 - b(\Delta_{ij}^n)) \leq x) \\ &= p \cdot P(g(\Delta_{ij}^n) \leq x - 2K) + q \cdot P(g(\Delta_{ij}^n) \leq x) \\ &= p \cdot F_n(h(x - 2K)) + q \cdot F_n(h(x)) \end{aligned} \tag{2.13}$$

$$\boldsymbol{F_{n+1}(x) = p \cdot F_n(h(x - 2K)) + q \cdot F_n(h(x))} \tag{2.14}$$

This equation is crucial for understanding the dynamics of the Elo rating system, as it describes the evolution of the CDF over successive games. The operator $\mathcal{L}$ is defined as:

$$\mathcal{L}(F_n) = p \cdot (F_n \circ h)(x - 2K) + q \cdot (F_n \circ h)(x) \tag{2.15}$$

Thus, the CDF at step $n + 1$ is:

$$F_{n+1}(x) = \mathcal{L}(F_n(x)) \tag{2.16}$$

The operator $\mathcal{L}$ leads to a recursive update process for the CDF. The fixed point of $\mathcal{L}$, a distribution F^* such that

$$\boldsymbol{\mathcal{L}(F^*) = F^*}, \tag{2.17}$$

represents the invariant distribution of the Elo ratings. While $\mathcal{L}$ is not generally contractible (as discussed in [6]), it has been suggested that $\mathcal{L}$ may be contractible on average for two players. The existence and uniqueness of the invariant distribution F^* have been established by Avdeev [2]. This implies that iterative application of $\mathcal{L}$ to any initial CDF F_0 will converge to the invariant distribution F^*.

In this Markovian process, the state of the system at any given time is represented by the rating difference Δ_{ij}^n. This invariant law for two players, resulting from the infinite recurrence, provides the steady-state behavior of the Elo rating system.

2.3 Symmetry of the Invariant Law

In this section, we aim to prove the symmetry of the cumulative distribution function $F^*(x)$ for the case where the probability $p = \frac{1}{2}$ and the mean difference $\mu = 0$. While this symmetry has already been established by Avdeev [2] using a Gaussian bonus-malus function, we will demonstrate it here using the logistic bonus-malus function symmetry. Our approach leverages the symmetry properties of the logistic function to show that the invariant distribution $F^*(x)$ satisfies $F^*(x) + F^*(-x) = 1$.

Proposition 1. *If $p = \frac{1}{2}$ and $F(x)$ is symmetric such that $F(x) + F(-x) = 1$, then the operator $\mathcal{L}(F)$ preserves this symmetry, i.e., $\mathcal{L}(F)(x) + \mathcal{L}(F)(-x) = 1$.*

Proof. First, we show the symmetry of $g(x)$:

$$g(x) = x - 2Kb(x)$$

$$g(-x) = -x - 2Kb(-x)$$

Adding these, we get:

$$g(x) + g(-x) = (x - 2Kb(x)) + (-x - 2Kb(-x)) = -2K(b(x) + b(-x))$$

Since $b(x) + b(-x) = 1$:

$$g(x) + g(-x) = -2K$$

Given $g(h(x)) = x$ and $g(h(-x)) = -x$. Next, we show the symmetry of $h(x)$:

$$g(h(x)) + g(-h(x)) = -2K$$

Since $g(h(x)) = x$, we have:

$$x + g(-h(x)) = -2K$$

Thus:

$$g(-h(x)) = -x - 2K$$

Since h is the inverse of g, we can write:

$$-h(x) = h(-x - 2K)$$

Similarly, evaluating $g(-h(-x))$:

$$g(h(-x)) = -x$$

$$g(-h(-x)) = x - 2K$$

Thus:

$$-h(-x) = h(x - 2K)$$

Finally, we prove the symmetry of the invariant law: Assume $F(x)$ is symmetric such that $F(x) + F(-x) = 1$ for $p = 0.5$ and $\mu = 0$:

$$\mathcal{L}(F)(x) = \frac{1}{2}F(h(x - 2K)) + \frac{1}{2}F(h(x))$$

Similarly:

$$\mathcal{L}(F)(-x) = \frac{1}{2}F(h(-x - 2K)) + \frac{1}{2}F(h(-x))$$

Adding these equations:

$$\mathcal{L}(F)(x) + \mathcal{L}(F)(-x) = \frac{1}{2}F(h(x - 2K)) + \frac{1}{2}F(h(x)) + \frac{1}{2}F(h(-x - 2K)) + \frac{1}{2}F(h(-x))$$

Using the symmetry properties of h:

$$h(x - 2K) = -h(-x) \quad \text{and} \quad h(-x - 2K) = -h(x)$$

Grouping the terms we get:

$$\frac{1}{2}F(-h(-x)) + \frac{1}{2}F(h(-x)) + \frac{1}{2}F(h(x)) + \frac{1}{2}F(-h(x))$$

Thus:

$$\frac{1}{2}[F(-h(-x)) + F(h(-x))] + \frac{1}{2}[F(h(x)) + F(-h(x))]$$

Since F is symmetric:

$$F(-h(-x)) + F(h(-x)) = 1 \quad \text{and} \quad F(h(x)) + F(-h(x)) = 1$$

Simplifying:

$$\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 1 = 1$$

Hence, we have shown that:

$$\mathcal{L}(F)(x) + \mathcal{L}(F)(-x) = 1$$

$\square$

As a consequence of the proposition, we can deduce that the invariant Elo distribution, for $p = \frac{1}{2}$, is also symmetric. This is formally stated in the following corollary:

Corollary 1.1. *If the hypotheses of the proposition are satisfied, then the unique fixed point F^* of $\mathcal{L}$, representing the invariant Elo distribution, is also symmetric. That is, $F^*(x) + F^*(-x) = 1$.*

Proof. Since F^* is a fixed point of the operator $\mathcal{L}$, we have:

$$F^* = \mathcal{L}(F^*)$$

Now, let F_0 be an initial symmetric distribution such that:

$$F_0(x) + F_0(-x) = 1$$

We apply the recursion operator $\mathcal{L}$ iteratively to F_0. Define the sequence $\{F_n\}$ by:

$$F_{n+1} = \mathcal{L}(F_n) \quad \text{for} \quad n \geq 0$$

By the hypothesis of the proposition, if F_n is symmetric, then $\mathcal{L}(F_n)$ is also symmetric. Hence, for all $n \geq 0$:

$$F_n(x) + F_n(-x) = 1$$

As n tends to infinity, the sequence $\{F_n\}$ converges to the unique fixed point F^* of the operator $\mathcal{L}$, as established by Avdeev [2].

Taking the limit as $n \to \infty$, we get:

$$\lim_{n\to\infty} F_n(x) = F^*(x) \quad \text{and} \quad \lim_{n\to\infty} F_n(-x) = F^*(-x)$$

Since each F_n is symmetric, we have:

$$\mathcal{L}(F_n)(x) + \mathcal{L}(F_n)(-x) = 1$$

Passing to the limit, we obtain:

$$F^*(x) + F^*(-x) = 1$$

Thus, the unique fixed point F^*, representing the invariant Elo distribution, is symmetric.

□

2.4 Elo Conjecture

Arpad Elo did not directly propose what we now refer to as the "Elo Conjecture." Instead, in his book [1], Elo suggested that the rating difference between two players is sufficient to estimate the probability of winning in an encounter between them.

In our model, the real strength of player i is represented by $E[R_i^\infty]$. This allows us to reformulate Elo's idea into a conjecture about the limit distribution of the random variable $\Delta_{ij}^\infty = R_i^\infty - R_j^\infty$, which represents the steady-state rating difference between two players i and j after an infinite number of games. For notation simplification, we will refer to Δ_{ij}^∞ simply as Δ_{ij} from now on.

We aim to verify numerically if:

$$E[b(\Delta_{ij})] = b(E[\Delta_{ij}])$$

Where b is the bonus-malus function (2.3).

This conjecture, inspired by Elo's original ideas, suggests that the long-term rating difference between two players can be used to accurately assess their relative strengths. By investigating this, we seek to provide a deeper understanding of the theoretical underpinnings of the Elo rating system and its applicability to estimating long-term player strengths.

From the proposition regarding symmetry, we can also deduce that the Elo conjecture holds true for $p = \frac{1}{2}$, as stated in the following corollary:

Corollary 1.2. *If $p = \frac{1}{2}$ and the hypotheses of the proposition are satisfied, then the Elo conjecture holds true. That is, $E[b(\Delta_{ij})] = b(E[\Delta_{ij}])$.*

Proof. First, recall that F^* is the fixed point of the operator $\mathcal{L}$, representing the invariant Elo distribution. For $p = \frac{1}{2}$, we have shown that F^* is symmetric, i.e., $F^*(x) + F^*(-x) = 1$.

We begin by noting that the bonus-malus function $b(x)$ respects the symmetry $b(x) + b(-x) = 1$ and thus, $\left(b(x) - \frac{1}{2}\right)$ is an odd function. Additionally, since F^* is symmetric, $dF^*(x)$ is an even function.

We then compute the expected value $E[b(\Delta_{ij})]$ as follows:

$$E[b(\Delta_{ij})] = \int_{-\infty}^{\infty} b(x) dF^*(x)$$

We can decompose the integral:

$$E[b(\Delta_{ij})] = \int_{-\infty}^{\infty} \left(b(x) - \frac{1}{2}\right) dF^*(x) + \int_{-\infty}^{\infty} \frac{1}{2} dF^*(x)$$

Since $\left(b(x) - \frac{1}{2}\right)$ is odd and $dF^*(x)$ is even, the integral of their product over the symmetric limits is zero:

$$\int_{-\infty}^{\infty} \left(b(x) - \frac{1}{2}\right) dF^*(x) = 0$$

The second term simplifies to:

$$\int_{-\infty}^{\infty} \frac{1}{2} dF^*(x) = \frac{1}{2} \int_{-\infty}^{\infty} dF^*(x) = \frac{1}{2} \cdot 1 = \frac{1}{2}$$

Thus:

$$E[b(\Delta_{ij})] = 0 + \frac{1}{2} = \frac{1}{2}$$

On the other hand, for $p = \frac{1}{2}$, the expected value of the rating difference is zero due to symmetry:

$$E[\Delta_{ij}] = 0$$

Hence:

$$b(E[\Delta_{ij}]) = b(0) = \frac{1}{2}$$

Therefore, we have shown that:

$$E[b(\Delta_{ij})] = b(E[\Delta_{ij}]) = \frac{1}{2}$$

This completes the proof. $\square$

Chapter 3

Methodology

3.1 Invariant Law Computational Algorithm

The goal of this section is to compute the invariant cumulative distribution function (CDF) of the Elo rating system's rating differences. This involves discretizing the space of rating differences and linearized update operator $\mathcal{L}(\cdot)$. The following steps outline the methodology employed:

Discretization and Mesh Generation To begin, we discretize the domain of possible rating differences, x, into a finite set of points:

$$\{x_0, x_1, \ldots, x_{N+1}\} \tag{3.1}$$

where N is the number of discretization points chosen based on the desired resolution and computational constraints. The grid spacing δx is set to $\delta x = \frac{2K}{m}$, ensuring that the point $x - 2K$ falls into the grid. This precise spacing facilitates accurate computations and aligns with the dynamic transformations of the system.

Using the Newton-Raphson method, we directly compute the inverse transformations dictated by the system's dynamics for $h(x)$ and $h(x - 2K)$:

$$h(x_i) \quad \text{and} \quad h(x_i - 2K) \quad \text{for each} \quad i = 0, 1, \ldots, N+1. \tag{3.2}$$

This approach allows us to obtain the necessary points for our computations without generating much complexity.

Interpolation and Matrix Construction Given the monotonic nature of the cumulative distribution function F^*, and the transformations $h(x)$ derived from $g(x) =$

$x - 2Kb(x)$ with $b(x) = \frac{1}{1+10^{-x/400}}$, we can estimate F^* at any non-mesh point through linear interpolation. The bonus-malus function $b(x)$, specific to the Elo context, ensures that $g(x)$ and hence $h(x)$ are well-behaved and bounded within predictable limits. We choose a K value that makes $g'(0) > 0$ and maintains its increasing properties. Specifically, we know:

$$x - 2K < g(x) < x \quad \text{and} \quad x < g(x + 2K) < x + 2K, \tag{3.3}$$

implying that:

$$x < h(x) < x + 2K \quad \text{and} \quad x - 2K < h(x - 2K) < x \tag{3.4}$$

This property allows us to interpolate values of F at $h(x)$ and $h(x - 2K)$ using neighbouring mesh points. For each x_i in our discrete mesh, we identify points x_a and x_b such that $x_a \leq h(x_i) \leq x_b$. The weights for interpolation are then calculated as:

$$w_i = \frac{h(x_i) - x_a}{x_b - x_a}, \tag{3.5}$$

Similarly, we determine $x_{a'}$ and $x_{b'}$ for $h(x_i - 2K)$, and compute corresponding weights.

The matrix $\mathbf{A}$ for the linear approximation of $F(x)$ is constructed as follows, taking into account the correct indices for each case based on the interpolation:

$$\mathbf{A}_{ij} = \begin{cases} p(1 - w_i') & \text{if } j = a', \\ pw_i' & \text{if } j = b', \\ q(1 - w_i) & \text{if } j = a, \\ qw_i & \text{if } j = b, \\ 0 & \text{otherwise.} \end{cases} \tag{3.6}$$

This effectively linearizes the operator $\mathcal{L}$ such that:

$$\begin{aligned} F_{n+1}(x) &= pF_n(h(x - 2K)) + qF_n(h(x)) \\ &\approx p(1 - w_i')F_n(x_{a'}) + pw_i'F_n(x_{b'}) + q(1 - w_i)F_n(x_a) + qw_iF_n(x_b) \\ &\approx \mathbf{A}F_n \end{aligned} \tag{3.7}$$

Here are schematic representations of how A looks, where the rows correspond to discrete points x_i and columns correspond to indices j which may be a, a', b, b' based on the interpolation:

$$
\mathbf{A} = \begin{pmatrix}
\ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
\cdots & p(1-w'_{i-1}) & pw'_{i-1} & 0 & \cdots & 0 & q(1-w_{i-1}) & qw_{i-1} & \cdots \\
\cdots & 0 & p(1-w'_i) & pw'_i & 0 & \cdots & 0 & q(1-w_i) & \cdots \\
\cdots & 0 & 0 & p(1-w'_{i+1}) & pw'_{i+1} & 0 & \cdots & 0 & \cdots \\
\ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
$$

This first matrix shows the center of A, highlighting the structure and interpolation weights at a given point. It focuses on a central portion of the matrix, illustrating how entries are populated based on the interpolation weights $p(1-w'_i)$, pw'_i, $q(1-w_i)$, and qw_i.

$$
\mathbf{A} = \begin{pmatrix}
q(1-w_0) & qw_0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
\ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
p(1-w'_m) & pw'_m & 0 & \cdots & 0 & q(1-w_m) & qw_m & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
\ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\ddots & 0 & p(1-w'_{i-1}) & pw'_{i-1} & 0 & \cdots & 0 & q(1-w_{i-1}) & qw_{i-1} & 0 & \cdots & \cdots & \ddots \\
\ddots & 0 & 0 & p(1-w'_i) & pw'_i & 0 & \cdots & 0 & q(1-w_i) & qw_i & 0 & \cdots & \ddots \\
\ddots & 0 & 0 & 0 & p(1-w'_{i+1}) & pw'_{i+1} & 0 & \cdots & 0 & q(1-w_{i+1}) & \ddots & \cdots & \ddots \\
\ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & \cdots & \cdots & \cdots & 0 & p(1-w'_{N-m}) & pw'_{N-m} & 0 & \cdots & 0 & q(1-w_{N-m}) & qw_{N-m} \\
\ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & p(1-w'_0) & pw'_0
\end{pmatrix}
$$

The second, larger matrix representation includes the entire structure of $\mathbf{A}$, capturing the beginning and the end of the matrix. It demonstrates how the interpolation weights apply not only at the center but across all discrete points x_i from the start to the end of the matrix. This comprehensive view provides a full schematic, detailing how each entry in $\mathbf{A}$ is populated with non-zero entries at indices related to $x_{a'}, x_{b'}, x_a, x_b$ for each x_i.

3.1.1 Finding the Invariant Law

Once convergence is reached, F_∞ approximates the fixed point F^*. To directly find F^*, we solve:

$$F^* = \mathbf{A}F^*, \tag{3.8}$$

$$(\mathbf{I} - \mathbf{A})F^* = 0, \tag{3.9}$$

which involves finding the null space of the matrix $\mathbf{I} - \mathbf{A}$, where $\mathbf{I}$ is the identity matrix. This solution represents the stationary distribution F^* of the system.

Theoretically, if the matrix $\mathbf{A}$ could be as large as needed, we could compute the null space directly to find F^*. However, due to the linearisation process, we must truncate the computational space. This truncation results in incomplete information at the boundaries, as the infinite nature of the problem is approximated by a finite grid.

To address the missing information caused by this truncation, we introduce a source term $\mathbf{b}$. The system can be written as:

$$F^* = \mathbf{A}F^* + \mathbf{b}, \tag{3.10}$$

Solving this system gives us the invariant cumulative distribution for the linearized model:

$$(\mathbf{I} - \mathbf{A})F^* = \mathbf{b} \tag{3.11}$$

The source term $\mathbf{b}$ is constructed to approximate the boundary values at the mesh limits, ensuring that the solution adheres to the theoretical behavior of the cumulative distribution function (CDF) at infinity. Specifically, $\mathbf{b}$ is filled corresponding to terms that are outside the matrix $\mathbf{A}$, which happens when $h(x)$ is out of the mesh (either at the beginning or the end).

These points typically correspond to m points within the first and last $2K$ values, aligning with the bounds of $h(x)$ and $h(x-2K)$. At the beginning of the grid, where x_0 represents $-\infty$, we have $h(x_0)$ and $h(x_0-2K)$. Given x < h(x) < x + 2K and $h(x_0-2K) \approx x_0-2K$, the first $2K$ values fall outside the grid, representing m points. This is because $\Delta x = 2K/m$, where m is the number of points dividing $2K$.

Similarly, at the other extreme, where x_{N+1} represents $+\infty$. Given $x < h(x) < x + 2K$ and $h(x_{N+1}) \approx x_{N+1} + 2K$, the values for $2K$ again fall outside the grid, representing m points.

Specifically, terms $F^* \circ h(x)$ corresponding to x values that effectively represent ∞ (corresponding to values close to 1) in our grid are set to 1, ensuring the appropriate boundary conditions at the upper bound. Similarly, terms $F^* \circ h(x-2K)$ corresponding to x values that effectively represent $-\infty$ (corresponding to values close to 0) in our grid are set to 0, ensuring the appropriate boundary conditions at the lower bound.

A reference of the $\mathbf{b}$ source term is:

$$\mathbf{b} = \begin{pmatrix} p \cdot F^* \circ h(x_0 - 2K) \\ \vdots \\ p \cdot F^* \circ h(x_{m-1} - 2K) \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ q \cdot F^* \circ h(x_{N-m+1}) \\ \vdots \\ q \cdot F^* \circ h(x_{N+1}) \end{pmatrix} \approx \begin{pmatrix} p \cdot 0 \\ \vdots \\ p \cdot 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ q \cdot 1 \\ \vdots \\ q \cdot 1 \end{pmatrix}$$

This ensures that the invariant cumulative distribution F^* correctly approximates the theoretical CDF within the chosen grid limits. By carefully selecting the grid boundaries, we ensure that the extremities of the grid reflect the behavior of the CDF at $-\infty$ and ∞, allowing for an accurate and stable solution.

Box Size Discussion To balance computational capacity and precision, we choose a box where $b(x)$ is close to 0 and 1 with three figures of precision. This box corresponds to the range $[-1200, 1200]$, ensuring that the values outside this range are appropriately approximated by the boundaries set in the source term $\mathbf{b}$.

3.2 Grid Convergence Study

A grid convergence study is crucial in numerical analysis to verify the accuracy of a numerical scheme by evaluating its behavior as the grid resolution is refined. Given the absence of an analytical solution, this study assesses the reliability and order of convergence of our numerical model, as described in [8]. The principles underlying this method are extensively discussed in numerical analysis literature, including works such as Schatzman's book [9].

The primary objective is to determine the order of convergence of our numerical scheme. We solve the system for different grid resolutions, denoting the solutions with grid spacings δx and $\delta x/2$ as $F_{\delta x}$ and $F_{\delta x/2}$, respectively. The maximum absolute error between these solutions at corresponding points is measured using the L_∞ norm:

$$\|F_{\delta x} - F_{\delta x/2}\|_\infty.$$

Assuming the numerical scheme has an error of the form $C(\delta x)^n$, we perform a logarithmic regression to determine the order of convergence n. By testing various values of δx, we plot the error on a logarithmic scale against the grid spacing and fit a linear regression model:

$$\log(\|F_{\delta x} - F_{\delta x/2}\|_\infty) = n\log(\delta x) + \log(C).$$

The slope of the log-log plot provides the order of convergence n of the numerical scheme. This rigorous method allows us to evaluate the convergence properties and accuracy of our model, ensuring its reliability for practical applications.

Exact order of convergence computation The idea of grid convergence is simple yet powerful. It helps in verifying that as the grid is refined, the numerical solution converges to a consistent result.

To illustrate how grid convergence works, consider a solution F that we approximate using numerical methods. Let $F_{\delta x}$ and $F_{\delta x/2}$ be the numerical solutions obtained with grid spacings δx and $\delta x/2$, respectively. Assuming our numerical method is of order n, the errors in these approximations can be expressed as:

$$F_{\delta x} = F + C(\delta x)^n + \mathcal{O}((\delta x)^{n+1}),$$
$$F_{\delta x/2} = F + C\left(\frac{\delta x}{2}\right)^n + \mathcal{O}((\delta x)^{n+1}),$$

where C is a constant and $\mathcal{O}((\delta x)^{n+1})$ represents higher-order terms.

The difference between these solutions can be used to estimate the error:

$$F_{\delta x} - F_{\delta x/2} = (F + C(\delta x)^n) - \left(F + C\left(\frac{\delta x}{2}\right)^n\right) + \mathcal{O}((\delta x)^{n+1}).$$

Simplifying, we get:

$$F_{\delta x} - F_{\delta x/2} = C(\delta x)^n - C\left(\frac{\delta x}{2}\right)^n.$$

Factoring out $C(\delta x)^n$:

$$F_{\delta x} - F_{\delta x/2} = C(\delta x)^n \left(1 - \frac{1}{2^n}\right).$$

This expression shows that the difference between the solutions at two different grid resolutions is proportional to $(\delta x)^n$, and by analyzing this difference, we can estimate the order of convergence n.

To determine the order of convergence n, we can use the ratio of errors at different grid spacings. Suppose we have solutions at three grid spacings: δx, $\delta x/2$, and $\delta x/4$. Denote the corresponding solutions as $F_{\delta x}$, $F_{\delta x/2}$, and $F_{\delta x/4}$.

The errors between these solutions can be written as:

$$E_1 = \|F_{\delta x} - F_{\delta x/2}\|_\infty,$$

$$E_2 = \|F_{\delta x/2} - F_{\delta x/4}\|_\infty.$$

Given the proportional relationship between the errors and the grid spacings, we have:

$$E_1 \approx C(\delta x)^n \left(1 - \frac{1}{2^n}\right),$$

$$E_2 \approx C\left(\frac{\delta x}{2}\right)^n \left(1 - \frac{1}{2^n}\right).$$

Taking the ratio of these errors:

$$\frac{E_1}{E_2} \approx \frac{C(\delta x)^n \left(1 - \frac{1}{2^n}\right)}{C\left(\frac{\delta x}{2}\right)^n \left(1 - \frac{1}{2^n}\right)} = 2^n.$$

Solving for n:

$$n \approx \log_2\left(\frac{E_1}{E_2}\right).$$

Therefore, by calculating the errors E_1 and E_2 at different grid resolutions and taking their ratio, we can estimate the order of convergence n.

3.3 Symmetry Study

To examine the symmetry properties of the Elo distribution, we perform a symmetry test. For a symmetric distribution, the CDF should satisfy the condition:

$$F_X(\mu - y) + F_X(\mu + y) = 1$$

where μ is the mean of the distribution, and y is any value on the x-axis. For the Elo distribution, if p and q are not equal, the mean μ will not be zero. Thus, we calculate the mean of F_{Elo} and use it in a symmetry test.

Symmetry test To examine the symmetry properties of the Elo distribution, we perform a symmetry test where we first calculate the mean μ of F_{Elo}. With this mean value, we then verify the symmetry around μ by checking if the condition $F_X(\mu-y)+F_X(\mu+y) = 1$ holds for various values of y within our selected box. The second part of this symmetry test involves taking the norm infinity ($\|\cdot\|_\infty$) of the difference for different grid spacings. This allows us to observe how the symmetry behavior evolves as the grid becomes finer. By analyzing the symmetry property of F_{Elo}, we can further understand the characteristics of the Elo invariant law.

3.4 Expected Value Computation Using the CDF

To compute the expected value $E[\phi(X)]$ using the cumulative distribution function (CDF) $F_X(x)$, we generalize our approach to cover various functions $\phi(x)$ because this allows us to derive important statistical measures like the expected value, the second moment, and the expected value of the bonus-malus function from a single unified framework. By doing so, we can apply the same methodology to different functions of X without repeating the entire derivation process each time. This not only simplifies our computations but also ensures consistency in the approach.

We leverage the fact that we can split the integral into two parts and apply integration by parts to simplify the computation. Since we do not have the density function, we use the CDF directly.

The expected value $E[\phi(X)]$ [5] is defined as:

$$E[\phi(X)] = \int_{-\infty}^{\infty} \phi(x)\, dF_X(x)$$

We split the integral into two parts:

$$E[\phi(X)] = \int_{-\infty}^{0} \phi(x)\, dF_X(x) + \int_{0}^{\infty} \phi(x)\, dF_X(x)$$

Using the complementary CDF $G_X(x) = 1 - F_X(x)$, and noting that $dF_X(x) = -dG_X(x)$, we proceed with integration by parts.

For the integral over $(-\infty, 0)$:

$$\int_{-\infty}^{0} \phi(x)\, dF_X(x) = [\phi(x)F_X(x)]_{-\infty}^{0} - \int_{-\infty}^{0} F_X(x)\phi'(x)\, dx$$

Evaluating at the boundaries, we get:

$$[\phi(x)F_X(x)]_{-\infty}^{0} = \phi(0)F_X(0) - \left(\lim_{x\to-\infty} \phi(x)F_X(x)\right)$$

Assuming $\phi(x)F_X(x) \to 0$ as $x \to -\infty$, this simplifies to:

$$\int_{-\infty}^{0} \phi(x)\, dF_X(x) = \phi(0)F_X(0) - \int_{-\infty}^{0} F_X(x)\phi'(x)\, dx$$

For the integral over $(0, \infty)$:

$$\int_{0}^{\infty} \phi(x)\, dF_X(x) = -\int_{0}^{\infty} \phi(x)\, dG_X(x)$$

Applying integration by parts:

$$-\int_{0}^{\infty} \phi(x)\, dG_X(x) = -\left[\phi(x)G_X(x)\right]_{0}^{\infty} + \int_{0}^{\infty} G_X(x)\phi'(x)\, dx$$

Evaluating at the boundaries, we get:

$$-\left[\phi(x)G_X(x)\right]_{0}^{\infty} = -\left(0 - \phi(0)G_X(0)\right) = \phi(0)G_X(0)$$

Thus:

$$-\int_{0}^{\infty} \phi(x)\, dG_X(x) = \phi(0)G_X(0) + \int_{0}^{\infty} G_X(x)\phi'(x)\, dx$$

Combining both integrals, we get:

$$E[\phi(X)] = \phi(0)F_X(0) - \int_{-\infty}^{0} F_X(x)\phi'(x)\, dx + \phi(0)G_X(0) + \int_{0}^{\infty} G_X(x)\phi'(x)\, dx$$

Since $G_X(0) = 1 - F_X(0)$, we can further simplify:

$$E[\phi(X)] = \phi(0)\left(F_X(0) + 1 - F_X(0)\right) - \int_{-\infty}^{0} F_X(x)\phi'(x)\, dx + \int_{0}^{\infty} G_X(x)\phi'(x)\, dx$$

Thus:

$$E[\phi(X)] = \phi(0) - \int_{-\infty}^{0} F_X(x)\phi'(x)\, dx + \int_{0}^{\infty} (1 - F_X(x))\phi'(x)\, dx$$

Expected Value Computation For the expected value [5] $E[X]$, we set $\phi(x) = x$:

$$E[X] = -\int_{-\infty}^{0} F_X(x)\,dx + \int_{0}^{\infty} (1 - F_X(x))\,dx \tag{3.12}$$

Variance Computation Using the Moments Using the first moment $E[X]$ and the second moment $E[X^2]$, the variance [5] is computed as:

$$V[X] = E[X^2] - (E[X])^2 \tag{3.13}$$

For the second moment $E[X^2]$, we set $\phi(x) = x^2$:

$$E[X^2] = -\int_{-\infty}^{0} 2xF_X(x)\,dx + \int_{0}^{\infty} 2x(1 - F_X(x))\,dx \tag{3.14}$$

Expected Value of the Bonus-Malus Function For the expected value $E[b(X)]$, we set $\phi(x) = b(x)$ and with $b(0) = \frac{1}{2}$:

$$E[b(X)] = \frac{1}{2} - \int_{-\infty}^{0} F_X(x)b'(x)\,dx + \int_{0}^{\infty} (1 - F_X(x))b'(x)\,dx \tag{3.15}$$

3.5 Elo Conjecture Study

We test the Elo conjecture by analyzing the differences between $E_{p,q}[b(\Delta_{ij})]$ and $b(E_{p,q}[\Delta_{ij}])$ using the equations we developed for computing the expected value (3.12) and (3.15). Specifically, we test the condition $E_{p,q}[b(\Delta_{ij})] = b(E_{p,q}[\Delta_{ij}])$:

$$\text{difference} = E_{p,q}[b(\Delta_{ij})] - b(E_{p,q}[\Delta_{ij}]) \tag{3.16}$$

We compute the L_1 norm of the differences obtained in Equation (3.16) across different grid spacings and values of p and q, examining how these differences behave as the grid becomes finer.

This approach allows us to thoroughly evaluate the Elo conjecture, offering a detailed understanding of the relationship between $E_{p,q}[b(\Delta_{ij})]$ and $b(E_{p,q}[\Delta_{ij}])$.

3.6 Comparing the Elo CDF to the Normal CDF

In addition to testing the Elo conjecture, we compare the computed Elo's cumulative distribution function (CDF) $F_{Elo}(x)$ to the CDF of a normal distribution to assess the distributional properties of X.

We compute the mean μ and variance σ^2 of Δ_{ij} using the CDF $F_{Elo}(x)$ and equations (3.12), (3.13):

$$E[\Delta_{ij}] = \mu = -\int_{-\infty}^{0} F_{Elo}(x)\,dx + \int_{0}^{\infty} (1 - F_{Elo}(x))\,dx$$

$$V[\Delta_{ij}] = \sigma^2 = E[\Delta_{ij}^2] - (E[\Delta_{ij}])^2$$

Using these parameters, we generate the CDF of a normal distribution $\mathcal{N}(\mu, \sigma^2)$ and compare it to $F_{Elo}(x)$.

We assess the differences between $F_{Elo}(x)$ and the normal CDF computing the infinity norm ($\|\cdot\|_\infty$) for different grid spacings and various values of p and q.

Normal Residual Error Analysis We analyze the residual error by inserting the normal distribution into the Elo dynamic equation:

$$\text{residual}_x = p \cdot F_{\text{Normal}}(h(x - 2K)) + q \cdot F_{\text{Normal}}(h(x)) - F_{\text{Normal}}(x)$$

Then we compute the infinity norm ($\|\cdot\|_\infty$) of these residuals to quantify the discrepancies. This norm is determined for different grid spacings δx and values of p (and q). By performing this residual error analysis, we can evaluate the adherence of the normal distribution to the Elo dynamic equation.

3.7 $\mathbb{E}[\Delta_{ij}]$ and $\mathbb{V}[\Delta_{ij}]$ vs K

To investigate the dependence of the expected value $\mathbb{E}[\Delta_{ij}]$ and variance $\mathbb{V}[\Delta_{ij}]$ on the parameter K in the Elo rating system, we used numerical methods to compute these quantities for various configurations (values of p) and values of K.

We varied K from 10 to 40, and for each configuration and value of K, we computed the expected value $\mathbb{E}[\Delta_{ij}]$ and variance $\mathbb{V}[\Delta_{ij}]$.

To determine if $\mathbb{E}[\Delta_{ij}]$ is independent of K, we performed a linear regression on the plot of $\mathbb{E}[\Delta_{ij}]$ vs. K and calculated the slope and the coefficient of determination R^2 value to assess the fit. For the variance, we similarly performed linear regression on the plot of $\mathbb{V}[\Delta_{ij}]$ vs. K.

The slope and R^2 value were calculated to provide insights into the relationship between $\mathbb{E}[\Delta_{ij}]$ and K, and $\mathbb{V}[\Delta_{ij}]$ and K. By analyzing the slope and R^2 value, we gain insights into the dependence of $\mathbb{E}[\Delta_{ij}]$ and $\mathbb{V}[\Delta_{ij}]$ on the parameter K.

Chapter 4

Results

In this chapter, we present the results of our numerical simulations and analyses related to the Elo rating system. The key findings are organized as follows:

4.1 Invariant Law Computational Algorithm

Figure 4.1 shows the computed cumulative distribution function (CDF) for different combinations of probabilities p and q. These graphs demonstrate that the numerical scheme correctly approximates a CDF, confirming the validity of our approach.

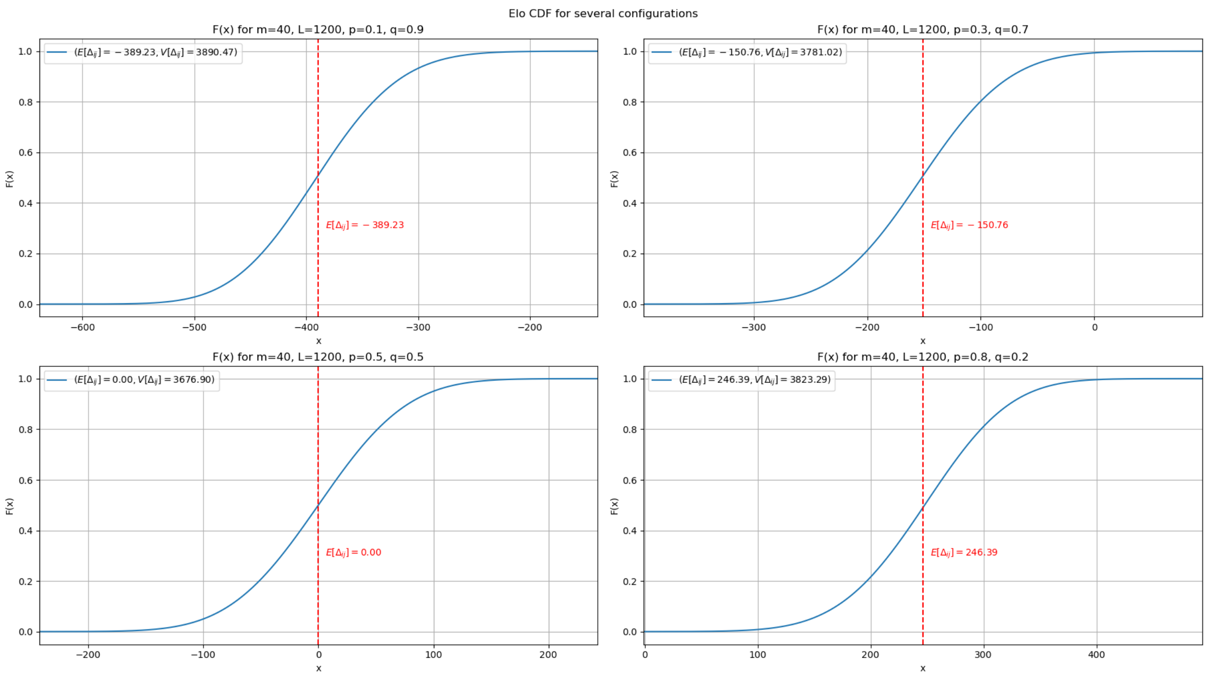

Figure 4.1: Computed CDF for different p and q values. Each plot represents the distribution function derived from the Elo rating system for different winning probabilities.

The figure also shows the expected value $E[\Delta_{ij}]$ and the variance $V[\Delta_{ij}]$ for the different configurations. For the case $p = 0.5$, our simulation shows that $E[\Delta_{ij}] = 0$, a result that aligns with the theoretical expectation that both opponents have the same strength. Additionally, the definition of Δ_{ij} indicates that when the probability of winning is greater than 0.5, the expected value is positive, and when it is less than 0.5, the value is negative. This consistency with real-world expectations validates the model's reliability.

4.2 Grid Convergence Study

The results of the grid convergence study are shown in Figure 4.2. The study, presented on a log-log scale, indicates an approximate second-order convergence with a regression slope of 1.98. This demonstrates the reliability and accuracy of the numerical scheme as the grid resolution is refined.

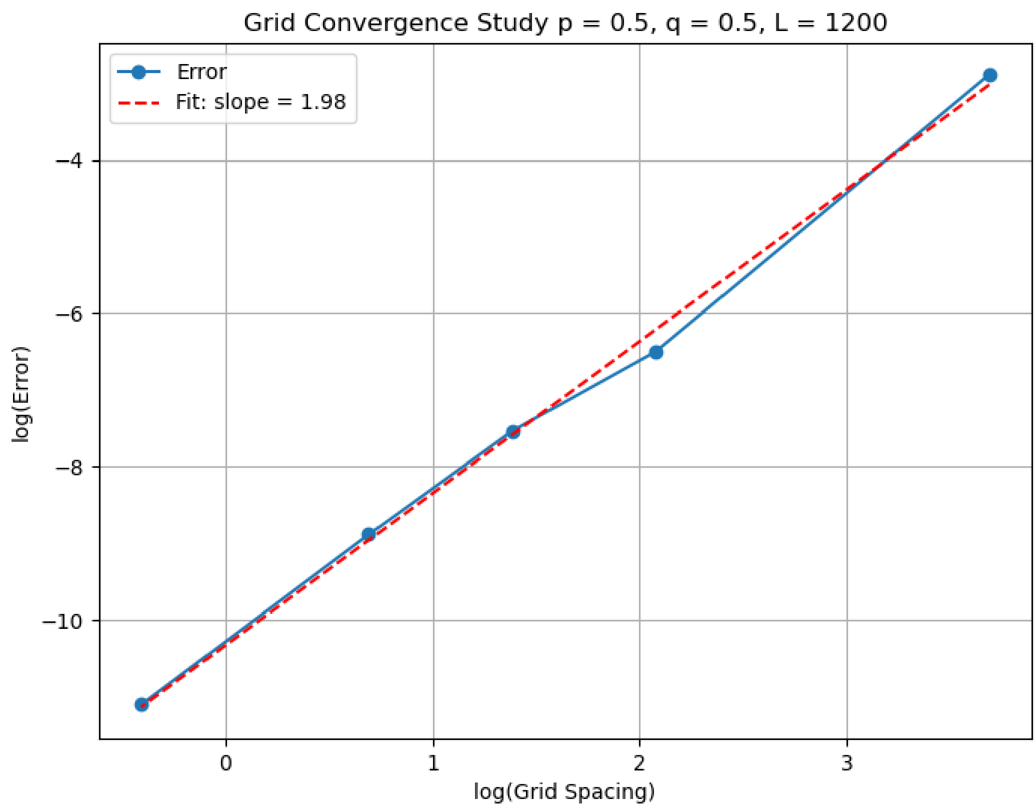

Figure 4.2: Grid convergence study showing second-order convergence with a regression slope of 1.98. The graph plots the log of the error against the log of the grid spacing.

To further validate the convergence behavior, an exact computation of the order of convergence was performed. The errors between solutions at different grid spacings were calculated, and the ratio of these errors was used to estimate the order of convergence, n. This rigorous approach involved solving the system for several grid spacings and taking the ratio of errors at each step. The average order of convergence was found to be 1.98, as shown in Figure 4.3.

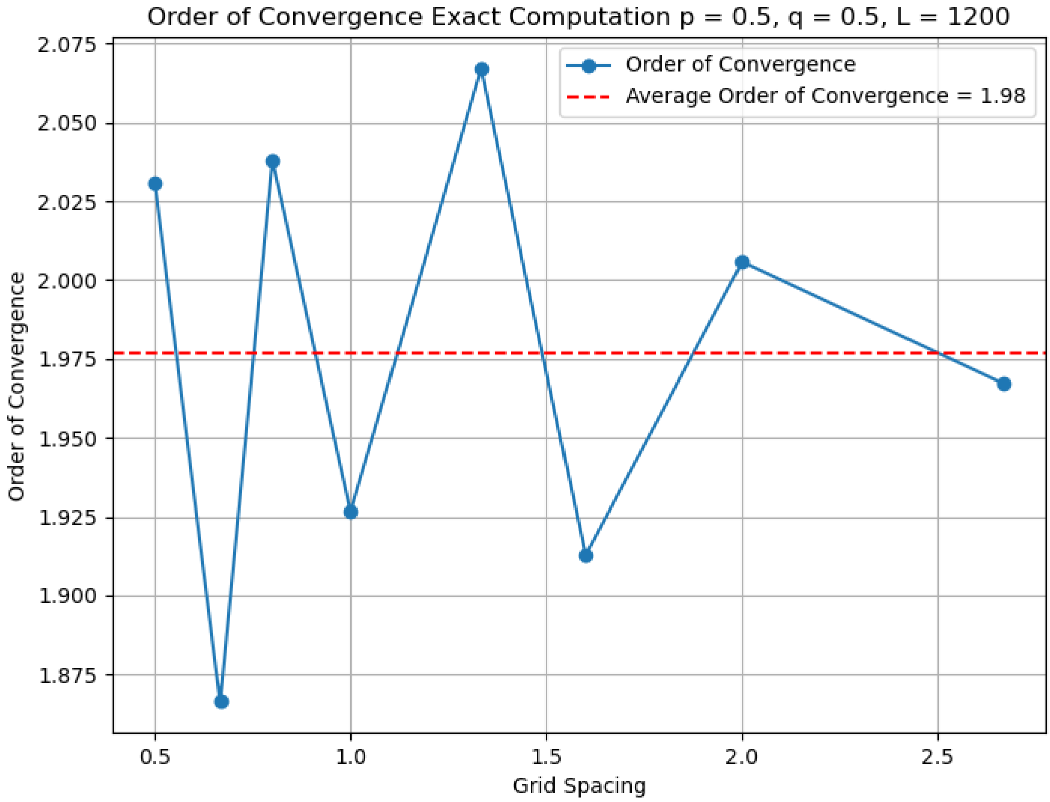

Figure 4.3: Exact computation of the order of convergence for different grid spacings. The average order of convergence is 1.98.

The results confirm the numerical scheme's robustness and highlight its accuracy as the grid spacing is refined. This order of convergence may be attributed to the sparsity of the matrix used in our numerical scheme and the careful selection of grid spacings.

4.3 Symmetry Study

To examine the symmetry properties of the Elo distribution, we performed a symmetry test. Figure 4.4 verifies the symmetry condition $F_X(\mu - y) + F_X(\mu + y) = 1$ for various y values. The results confirm that symmetry holds for most values, except for those close to the symmetric point of the CDF.

Although we observe some values that are not exactly 1, this can be attributed to numerical errors. These differences are relatively small compared to the expected error of the scheme. Specifically, for $p = 0.5$, the difference is minimal, confirming the theory that the Elo distribution is symmetric for this case [2] and as we saw in Corollary 1.1.

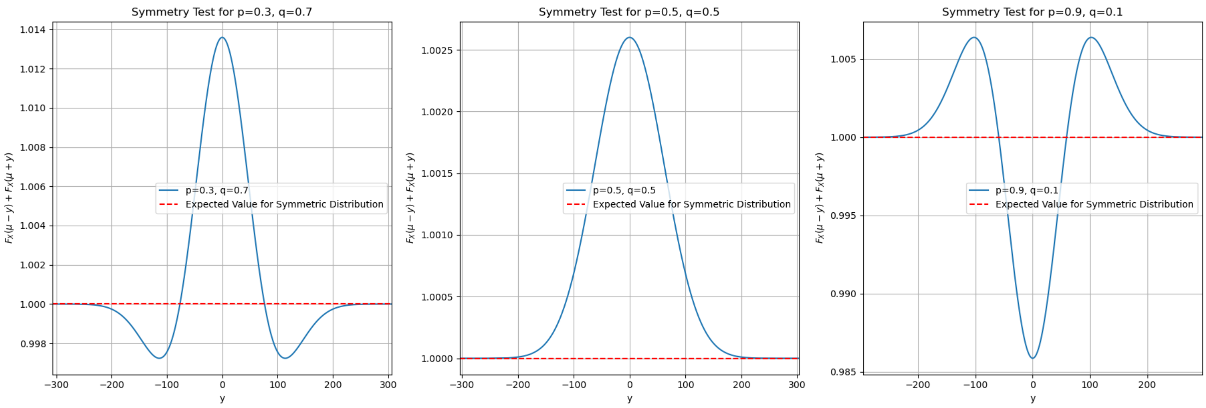

Figure 4.4: Symmetry test for different p and q values. The results confirm the expected symmetry of the distribution for most values.

L_∞ Symmetry Test Figure 4.5 presents the difference symmetry analysis. The L_∞ norm difference for different p and q combinations shows that the differences are small, indicating a good fit with the expected symmetry condition.

The differences decrease as the grid spacing becomes finer. For $p = 0.5$, the symmetry is clear, while for the other cases, the results suggest but do not confirm that the Elo distribution might be also symmetric for cases $p \neq 0.5$.

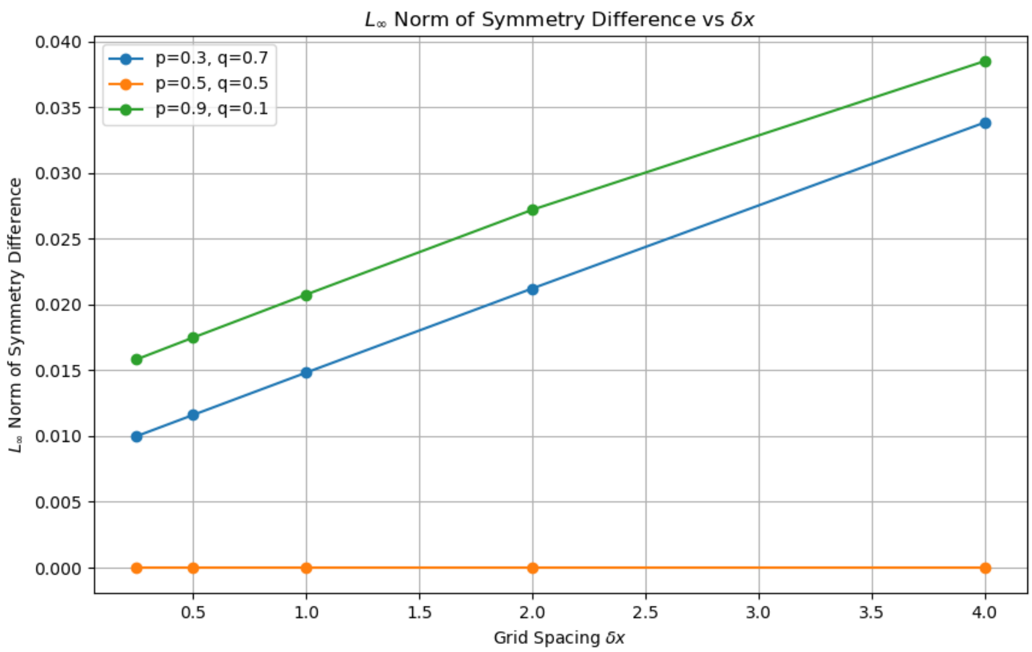

Figure 4.5: Difference symmetry analysis for different p and q values.

4.4 Elo Conjecture Study

We tested the Elo conjecture by analyzing the condition $E_{p,q}[b(\Delta_{ij})] = b(E_{p,q}[\Delta_{ij}])$. The analysis is presented in Figure 4.6, which shows the difference between $E[b(X)]$ and $b(E[X])$ as a function of the grid spacing. The results indicate that the difference decreases as the grid spacing becomes finer, suggesting that the Elo conjecture may hold with some level of imprecision.

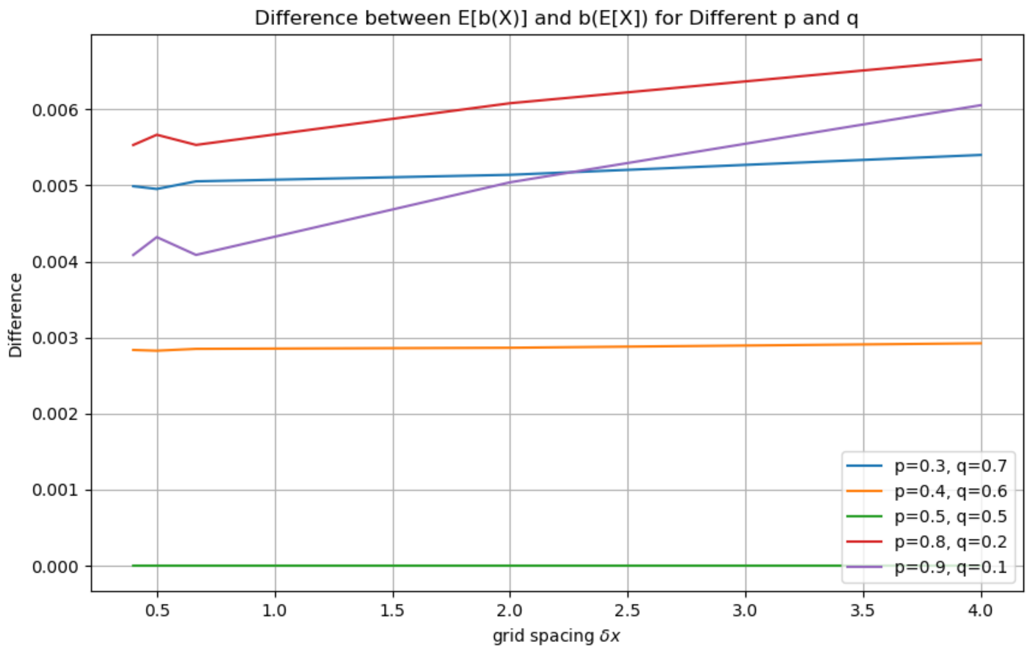

Figure 4.6: Error of the Elo conjecture vs. grid spacing. The decrease in error as the grid spacing becomes finer suggests that the conjecture may hold approximately.

Our numerical scheme confirms the Elo conjecture for $p = 0.5$ as the difference is very close to zero. This result is consistent with theory, as we proved the conjecture for this case in Corollary 1.2. For other cases, the differences are also small and less than the supposed numerical error, suggesting that the results indicate but do not confirm that the conjecture holds.

4.5 Comparing the Elo CDF to the Normal CDF

In Figure 4.7, we compare the computed Elo CDF with the CDF of a normal distribution, calculated using the expected value and variance. The comparison indicates that the Elo CDF closely approximates the normal CDF, suggesting that the distribution of player ratings in the Elo system may closely follow a normal distribution under certain conditions.

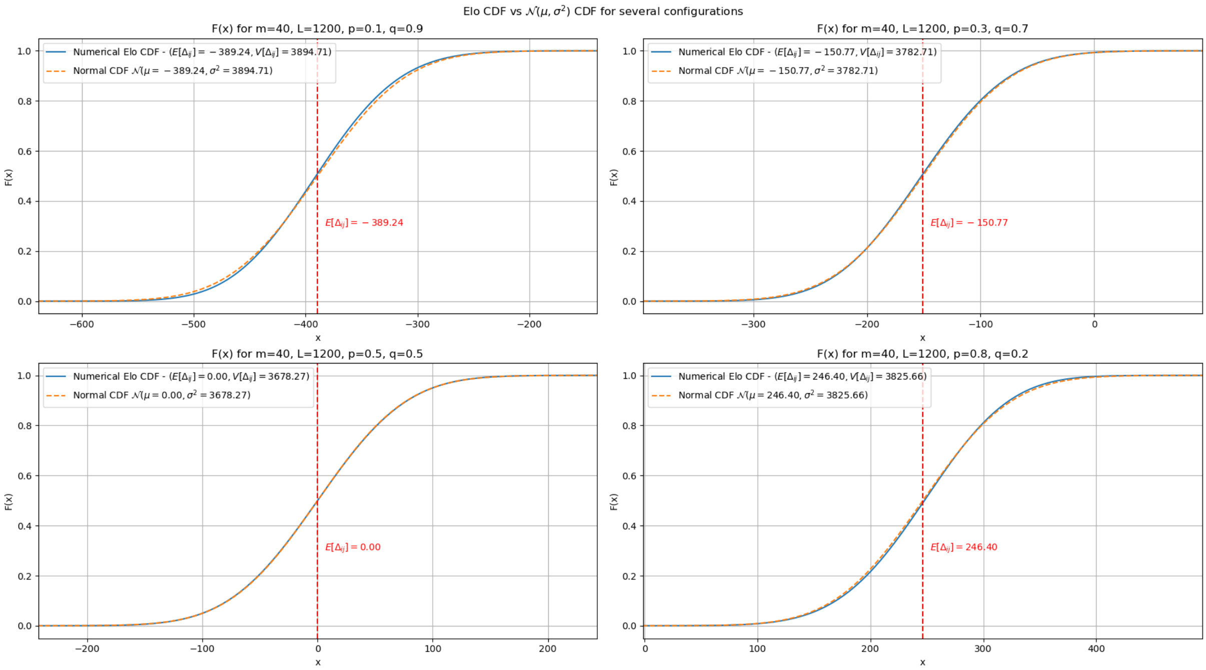

Figure 4.7: Comparison between the computed Elo CDF and the normal CDF for different p and q values. The close fit suggests that the Elo rating system approximates a normal distribution.

The figure shows the Elo CDF in blue and the normal CDF in orange dashed lines. For $p = 0.5$, the fit is almost perfect, while for more extreme cases like $p = 0.1$, the differences are larger. This suggests that the normal distribution closely models the Elo dynamics for $p = 0.5$ but not as well for more extreme probabilities.

Normal Residual Error Analysis The residual error analysis for different p and q combinations is shown in Figure 4.8. We analyze the residual error by inserting the normal distribution into the Elo dynamic equation:

$$\text{residual}_x = p \cdot F_{\text{Normal}}(h(x - 2K)) + q \cdot F_{\text{Normal}}(h(x)) - F_{\text{Normal}}(x)$$

Figure 4.8 shows the residual error analysis for this scenario. The results indicate that the error decreases as the grid spacing becomes smaller, and for $p = 0.5$, the residual error is almost zero.

The residual error decreases as the grid spacing becomes finer. However, further investigation into this type of analysis indicates that consistency does not imply convergence. Since $2K$ is approximately 40, which is small relative to the scheme, we can approximate $h(x) \approx x$. Using the identity developed earlier, $-h(-x) = h(x - 2K)$, we substitute in the Elo equation at the fixed point:

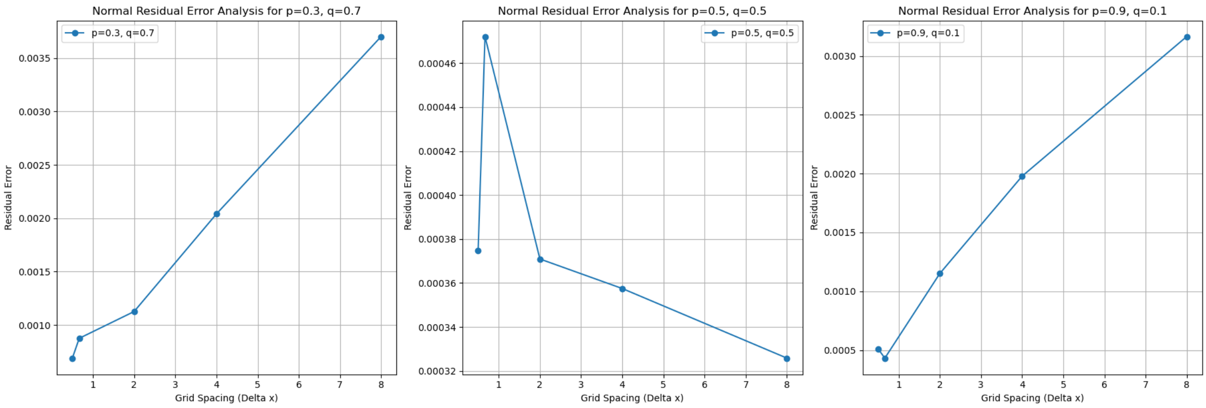

Figure 4.8: Residual error analysis for the normal distribution in the Elo dynamic equation.

$$F(x) = p \cdot F(h(x - 2K)) + q \cdot F(h(x)) = p \cdot F(-h(-x)) + q \cdot F(h(x))$$

Approximating $h(x) \approx x$, we get:

$$F(x) = p \cdot F(x) + q \cdot F(x) = (p + q)F(x) = (1)F(x)$$

This holds for all $F(x)$, indicating that this test is not useful for confirming or denying that the Elo distribution is a normal distribution.

4.6 $\mathbb{E}[\Delta_{ij}]$ and $\mathbb{V}[\Delta_{ij}]$ vs K

Figure 4.9 shows the results of the expected value $\mathbb{E}[\Delta_{ij}]$ for different values of K. The analysis of the slope and R^2 value provides insights into the relationship between $\mathbb{E}[\Delta_{ij}]$ and K.

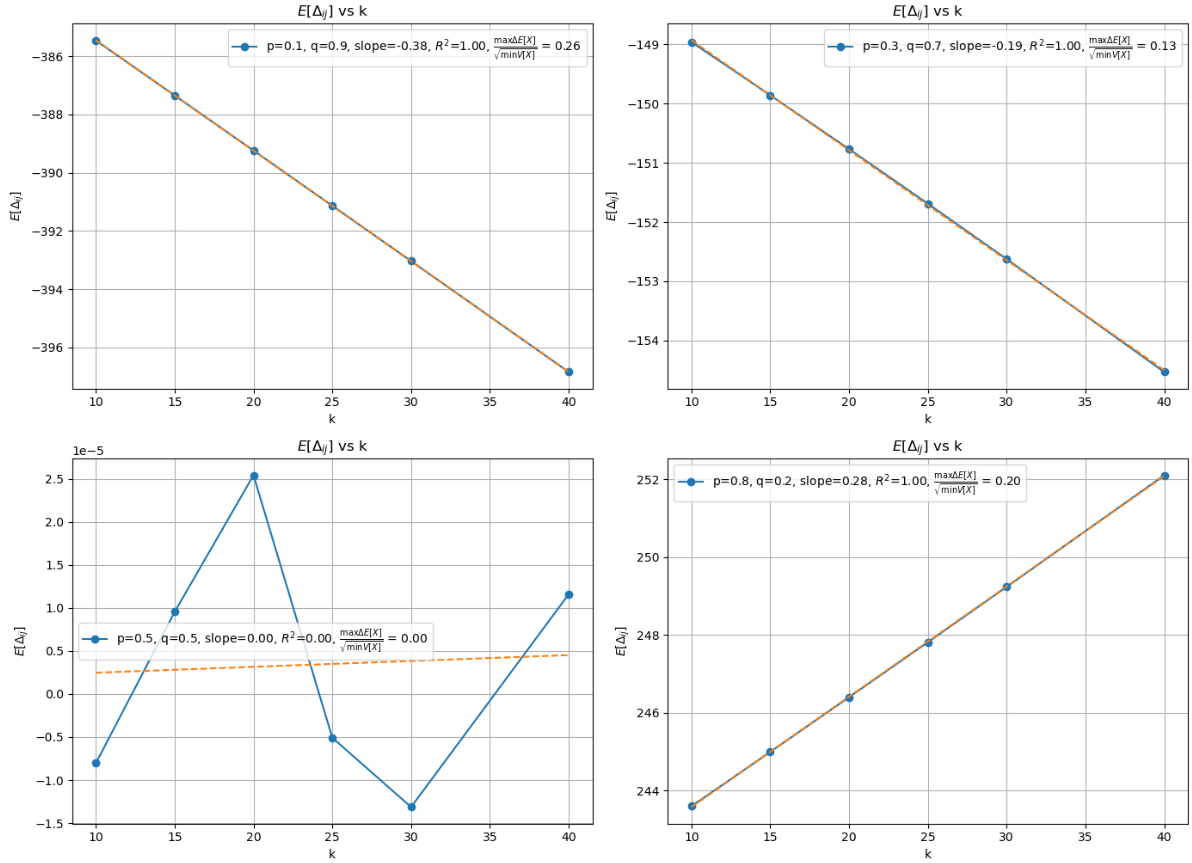

Figure 4.9: Expected value $\mathbb{E}[\Delta_{ij}]$ vs. parameter K. The analysis shows the independence of this statistical measure on the parameter K for different values of p.

For $p = 0.5$, it is clear that $\mathbb{E}[\Delta_{ij}]$ and K are independent. For other cases, there appears to be a good linear correlation. However, computing the difference between the maximum and minimum expected values and dividing by the standard deviation shows that the difference is minimal. Thus, we can say (without confirming) that the numerical scheme indicates that the expected value is independent of K.

Figure 4.10 shows the results of the variance $\mathbb{V}[\Delta_{ij}]$ for different values of K. The analysis of the slope and R^2 value provides insights into the relationship between $\mathbb{V}[\Delta_{ij}]$ and K.

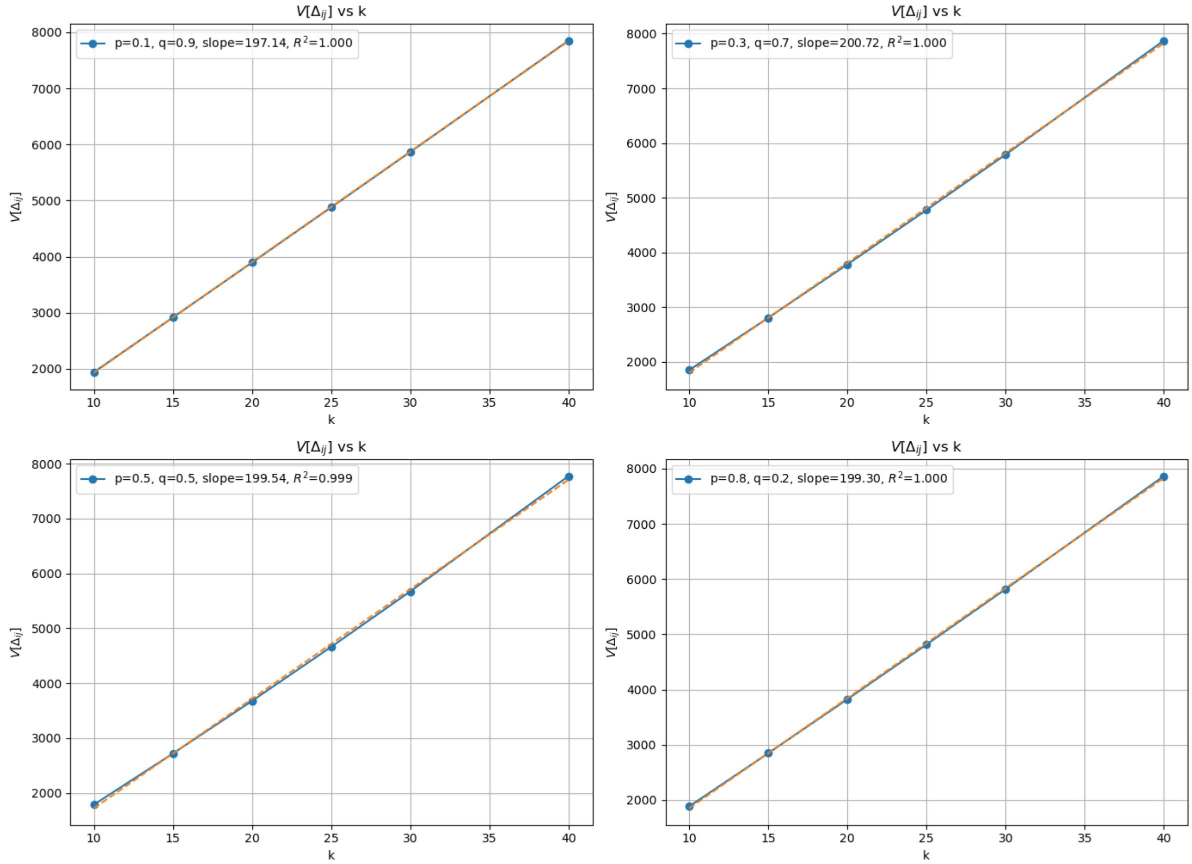

Figure 4.10: Variance $\mathbb{V}[\Delta_{ij}]$ vs. parameter K. The analysis shows the dependence of this statistical measure on the parameter K for different values of p.

In contrast, the variance shows a significant change and a high linear correlation (with R^2 near 1). The change is substantial, as the variance ranges from 2000 to nearly 8000. This suggests that the numerical scheme indicates a linear dependence of the variance on K with a slope of around 200.

Chapter 5

Discussion and Conclusion

In this chapter, we analyze and interpret the results presented in the previous section. Our discussion will cover the implications of the findings, compare them with existing literature, and identify potential limitations and areas for future research.

Invariant Law Computational Algorithm The computed cumulative distribution function (CDF) for different p and q values demonstrated that our numerical scheme accurately approximates the Elo rating system's invariant distribution. This finding aligns with theoretical expectations and confirms the validity of our approach. The expected values and variances observed for different configurations further validate the model's consistency with real-world scenarios. For instance, the result $E[\Delta_{ij}] = 0$ for $p = 0.5$ reflects the theoretical balance when both players have equal skill levels.

Grid Convergence Study The grid convergence study revealed an approximate second-order convergence with a regression slope of 1.98. Additionally, an exact computation of the order of convergence confirmed this result, with an average order of 1.98 across multiple grid spacings. This result suggests that our numerical scheme is both reliable and accurate as the grid resolution is refined. The observed order of convergence, attributed to the sparsity of the matrix and the careful selection of grid spacings, supports the robustness and precision of the numerical method.

Symmetry Study The symmetry study confirmed that the Elo rating system's distribution is symmetric for most values, particularly for $p = 0.5$, where the symmetry was almost perfect. This aligns with Avdeev's theoretical predictions and indicates that our numerical approach accurately captures the underlying symmetry of the Elo distribution. Although some minor deviations from symmetry were observed, these can be attributed to numerical errors and are relatively insignificant compared to the overall accuracy of the scheme.

Elo Conjecture Study The analysis of the Elo conjecture revealed that the difference between $E_{p,q}[b(\Delta_{ij})]$ and $b(E_{p,q}[\Delta_{ij}])$ decreases as the grid spacing becomes finer. For $p = 0.5$, the difference was very close to zero, suggesting that the conjecture holds for this case. For other probabilities, the differences were also small, indicating that the conjecture may hold approximately. However, the presence of numerical errors prevents a definitive confirmation, highlighting the need for further investigation with more refined numerical techniques.

Comparing the Elo CDF to the Normal CDF The comparison between the computed Elo CDF and the normal CDF showed a close fit, especially for $p = 0.5$. This suggests that the distribution of player ratings in the Elo system closely follows a normal distribution under certain conditions. For extreme probabilities like $p = 0.1$, the differences were larger, indicating that the normal distribution may not perfectly model the Elo dynamics in such cases. Nevertheless, the overall close fit supports the hypothesis that the Elo rating system's distribution approximates a normal distribution.

The residual error analysis further confirmed that the normal distribution fits well within the Elo dynamic equation, particularly for $p = 0.5$. However, it is important to note that consistency in residuals does not imply convergence, and this type of analysis may not be the best test for confirming the Elo distribution's normality.

$\mathbb{E}[\Delta_{ij}]$ and $\mathbb{V}[\Delta_{ij}]$ vs K Our investigation into the dependence of $\mathbb{E}[\Delta_{ij}]$ and $\mathbb{V}[\Delta_{ij}]$ on the parameter K revealed interesting insights. For $p = 0.5$, the expected value $\mathbb{E}[\Delta_{ij}]$ appeared to be independent of K, while for other probabilities, a linear correlation was observed. However, the differences in expected values were minimal, suggesting that K does not significantly affect the expected value of Δ_{ij}.

In contrast, the variance $\mathbb{V}[\Delta_{ij}]$ showed a clear linear dependence on K, with a high R^2 value indicating a strong correlation. This finding suggests that K significantly influences the spread of the Elo rating differences, with higher values of K leading to greater variance.

Limitations and Future Research While our numerical approach has provided valuable insights into the Elo rating system's behavior, several limitations must be acknowledged. The presence of numerical errors and the need for more refined grid resolutions indicate that further optimization of the numerical scheme is necessary. Additionally, our findings suggest that while the Elo distribution closely approximates a normal distribution, this may not hold under all conditions, particularly for extreme probabilities.

Future research should focus on developing more precise numerical techniques, such as Richardson extrapolation, to minimize errors and provide a more definitive confirmation of the Elo conjecture. Investigating the tail behavior of the distribution and comparing it to the normal distribution could yield additional insights. Additionally, stochastic algorithms should be explored to provide a comparison with the deterministic approach developed in this work, offering a broader validation of the results. Our research has been limited to two players and binary outcomes (win or lose). It would be interesting to extend the model to include ties and analyze its implications.

5.1 Conclusion

In conclusion, our study has validated the accuracy and reliability of the numerical scheme used to approximate the Elo rating system's invariant distribution. The results confirmed the system's consistency with theoretical expectations like symmetry and its close approximation to a normal distribution. While the Elo conjecture holds approximately, further refinement of numerical methods is needed for a more definitive confirmation. Our findings contribute to a deeper understanding of the Elo rating system and its potential applications across various competitive domains.

Bibliography

[1] Elo, A., *The Rating of Chess Players, Past and Present*, Arco Pub, 1978.

[2] Avdeev, M., "Stationary Distribution of the Elo Rating System," 2020.

[3] Adrien Krifa, Florian Spinelli, Stéphane Junca. "On the convergence of the Elo rating system for a Bernoulli model and round-robin tournaments." [Research Report] Université Côte D'Azur, 2021. ⟨hal-03286065⟩

[4] Aldous, David. "Elo ratings and the sports model: a neglected topic in applied probability?" Statist. Sci. 32 (2017), no. 4, 616–629.

[5] Barbe, P. and Ledoux, M., *Probabilité*, Collection Enseignement sup. Mathématiques, EDP Sciences, 2007. ISBN: 9782868839312.

[6] Stéphane Junca. "Contractions to update Elo ratings for round-robin tournaments." 2021. ⟨hal-03286591⟩

[7] Pierre-Emmanuel Jabin, Stéphane Junca. "A Continuous Model For Ratings." SIAM Journal on Applied Mathematics, 2015, 75 (2), pp.420-442. ⟨10.1137/140969324⟩. ⟨hal-01143609⟩

[8] S. S. Y. Wang, P. J. Roache, R. A. Schmalz, Y. Jia, P. E. Smith, *Verification and Validation of 3D Free-Surface Flow Models*, American Society of Civil Engineers, 2009.

[9] Schatzman, M., *Numerical Analysis: A Mathematical Introduction*, Clarendon Press, 2002. Translation of *Analyse Numérique* by Michelle Schatzman originally published in French by Inter Editions, Paris 1991. xix, 496 pages.

Appendix A

Code

`Code Available on GitHub:  https://github.com/K10K30/Elo-CDF-Computation.git`