



Politechnika Wrocławska

Wydział Matematyki

Kierunek studiów: Matematyka Stosowana

Specjalność: –

Praca dyplomowa – inżynierska

# ANALIZA STATYSTYCZNA CZASÓW NA WYKONYWANIE RUCHÓW W SZACHACH

Piotr Rogula

słowa kluczowe:  
analiza statystyczna, szachy, korelacja, sta-  
tystyka, probabilistyka, błędy szachowe

krótkie streszczenie:

W pracy przedstawiona zostanie analiza zależności między czasami wykonania poszczególnych posunięć w szachach, ich numerem oraz jakością, biorąc jako miarę zgodność z silnikiem szachowym. Analiza statystyczna zostanie wykonana ze szczególnym uwzględnieniem różnic wynikających z poziomu rankingowego graczy oraz formatu czasowego poszczególnych partii.

Opiekun pracy dyplomowej	Prof. dr hab. inż. Marcin Magdziarz	.....	.....
	Tytuł/stopień naukowy/imię i nazwisko	ocena	podpis

*Do celów archiwalnych pracę dyplomową zakwalifikowano do:\**

*a) kategorii A (akta wieczyste)*

*b) kategorii BE 50 (po 50 latach podlegające ekspertyzie)*

*\* niepotrzebne skreślić*

pieczęćka wydziałowa

Wrocław, rok 2022





Wrocław University  
of Science and Technology

Faculty of Pure and Applied Mathematics

Field of study: Applied Mathematics

Specialty: –

Engineering Thesis

## STATISTICAL ANALYSIS OF TIMES BETWEEN CHESS MOVES

Piotr Rogula

keywords:

statistical analysis, chess, correlation, statistics, probabilistics, chess blunder

short summary:

The paper presents an analysis of the relationship between the times of taking individual chess moves, their number and quality, taking as a measure compliance with the chess engine. Statistical analysis will be performed with particular emphasis on the differences resulting from the ranking level of players and the time format of individual games.

Supervisor	Prof. dr hab. inż. Marcin Magdziarz	.....	.....
	Title/degree/name and surname	grade	signature

*For the purposes of archival thesis qualified to:\**

*a) category A (perpetual files)*

*b) category BE 50 (subject to expertise after 50 years)*

*\* delete as appropriate*

stamp of the faculty

Wrocław, 2022



# Spis treści

<b>Wstęp</b>	<b>3</b>
<b>1 Zagadnienia teoretyczne dotyczące gry w szachy</b>	<b>5</b>
1.1 Zasady gry w szachy . . . . .	5
1.2 System rankingowy Glicko-2 . . . . .	6
1.3 Funkcja oceny . . . . .	6
1.3.1 Stockfish . . . . .	6
1.3.2 Ewaluacja . . . . .	7
1.3.3 Rodzaje błędów szachowych . . . . .	7
<b>2 Zagadnienia teoretyczne dotyczące analizy statystycznej</b>	<b>9</b>
2.1 Metoda najmniejszych kwadratów . . . . .	9
2.2 Regresja liniowa . . . . .	9
2.2.1 Estymacja parametrów prostej regresji metodą najmniejszych kwadratów . . . . .	9
2.3 Metoda momentów . . . . .	10
2.3.1 Estymacja parametrów rozkładu log-normalnego metodą momentów . . . . .	10
<b>3 Sformułowanie i rozwiązanie problemów analitycznych</b>	<b>11</b>
3.1 Sformułowanie problemów . . . . .	11
3.2 Dane . . . . .	11
3.2.1 Odfiltrowanie danych . . . . .	12
3.3 Analiza problemu . . . . .	13
3.3.1 Wstępny przegląd danych . . . . .	13
3.3.2 Zależność między indeksem ruchu, a poświęconym czasem . . . . .	14
3.3.3 Zależność między czasem poświęconym na ruch, a jego oceną . . . . .	21
<b>4 Podsumowanie</b>	<b>25</b>



# Wstęp

Wraz z rozwojem technologii komputerowej, rozpoczęła się nowa era szachów. Technologia, korzystając z dużej mocy obliczeniowej, bezpowrotnie wyprzedziła człowieka w różnorodnych grach deterministycznych, do których zaliczają się szachy. Profesjonalni szachiści zaczęli wykorzystywać nowe strategie korzystając z coraz nowszych pomocy technologicznych, w związku z czym rozgrywki na najwyższym poziomie są zdecydowanie dokładniejsze, niż te przed komputerową rewolucją. Jeden z najsilniejszych obecnie silników, *Stockfish* [1] posiada zdolność oceniania wprowadzonej pozycji pod kątem przewagi jednej ze stron, co pozwala na oznaczenie błędnych ruchów w ciągu kilku sekund i bez potrzeby angażowania ekspertów.

W dobie internetu gra w szachy stała się dużo wygodniejsza i znacznie przystępniejsza niż przed laty. Ludzie grają w różnych miejscach i praktycznie o każdej porze. W związku z tym, coraz większą popularnością zaczęły cieszyć się szachy szybkie, czyli takie, w których każdy z zawodników ma relatywnie mało czasu na wykonanie wszystkich ruchów. Wiąże się to ze znacznie większym znaczeniem zarządzania czasem w trakcie gry. Za każdym razem zawodnik musi ustalić równowagę pomiędzy dokładnością posunięcia, a czasem, który jest w stanie na nie poświęcić.

Przedmiotem badań tej pracy jest analiza zależności między dokładnością posunięcia, czasem na niego poświęconym oraz jego numerem. Całość zostanie oparta na ogólnodostępnej bazie danych wszystkich gier rozgrywanych na platformie internetowej *Lichess.com* [2]. Rozważania zostaną podzielone ze względu na format czasowy danej partii oraz na poziom umiejętności zawodników. Dodatkowo przedstawiona będzie konfrontacja wyników statystycznych poszczególnych grup rankingowych, by móc ustalić wiodące różnice wynikające z poziomu gry zawodnika oraz wydłużania, bądź skracania formatu czasowego. Powstało już kilka badań, które opierając się na tych samych danych sprawdzało pokrewne zależności. Jedną z ciekawszych, w mojej opinii, prac [3] przedstawia analizę wpływu pozostałego czasu na prawdopodobieństwo popełnienia błędu, pokazując widoczną anomalię będącą skutkiem wpływu dźwięku alarmującego zawodnika o kończącym się czasie.

W pierwszej części pracy przedstawione i dość obszernie wyjaśnione zostaną podstawowe zagadnienia teoretyczne związane z grą w szachy, systemami rankingowymi i silnikami szachowymi. Przedstawione zostaną zasady gry w szachy ze szczególnym uwzględnieniem konkretnych elementów, których niewiedza nie pozwoliłaby na pełne przyswojenie przedstawionej później analizy.

W następnym kroku opisany zostanie system *Elo* [4], a także wzorujący się na nim system *Glicko-2* [5], którego pierwsza wersja została opisana w 1999 roku przez Marka Glickmana. Używany jest on do przypisania rankingu każdemu z graczy znajdujących się w bazie danych używanej do analiz. Ostatnim elementem będzie przedstawienie działania silnika szachowego *Stockfish* i jego wkładu w – kluczową dla analizy statystycznej – funkcję oceny posunięcia.

Kolejna część pracy opowiada o zagadnieniach teoretycznych z dziedziny statystyki,

zastosowanych w analizie przedstawionych problemów oraz odtwórcze wyprowadzenie wzorów używanych później.

Główna część pracy będzie się składać z trzech elementów. W pierwszym z nich odnaleźć można dokładne sformułowanie problemów, które są przedmiotem badań niniejszej pracy. Następnie opisane zostaną dane, na których bazuje późniejsza analiza, uwidoczniony sposób ich pozyskania oraz stworzenia na ich podstawie bazy ruchów szachowych. Kluczowym elementem tego etapu będzie analiza problemu. Po wstępnym przeglądzie danych i wyznaczeniu podstawowych statystyk w celu nabrania intuicji wobec kolejnych kroków badań, sprawdzona zostanie różnica pomiędzy średnim czasem posunięcia w poszczególnych ruchach i dla różnych przedziałów rankingowych. Pokazana będzie również różnica w zarządzaniu pozostałym czasem w zależności od formatu, a także prawdopodobieństwo popełnienia błędu w konkretnym ruchu. Dodatkowo, przy pewnych założeniach, określona zostanie szansa na popełnienie błędu w  $n$  pierwszych ruchach dla różnych formatów i przedziałów rankingowych. Ostatnim elementem tej części będzie sprawdzenie zależności pomiędzy czasem na wykonaniu ruchu, a jego oceną wystawioną przez silnik *Stockfish*. Podjęta zostanie również próba dopasowania rozkładu teoretycznego dla zbioru czasów na wykonanie posunięcia dla poszczególnych ocen silnika.

W końcowej części pracy zawarte zostaną najważniejsze wnioski wynikające z dokonanej analizy, a także rozważania na temat dalszej pracy w temacie analizy czasów wykonania posunięć w szachach.



# Rozdział 1

## Zagadnienia teoretyczne dotyczące gry w szachy

Niniejszy rozdział poświęcony zostanie zagadnieniom teoretycznym dotyczącym szachów, używanych systemów rankingowych oraz działaniu silników szachowych.

### 1.1 Zasady gry w szachy

Początki szachów nie są znane, jednak ich historia trwa już ok. 1500 lat i zaczyna się w Indiach. Na przestrzeni wieków zasady gry w szachy były wielokrotnie zmieniane. Powszechnie stosowane przepisy pochodzą z roku 1851<sup>1</sup>.

Gra odbywa się na kwadratowej planszy o wymiarach 8 na 8 pól. Każdy gracz posiada 16 figur, ustawionych w pozycji startowej i stojących po przeciwnych stronach szachownicy. Zawodnicy wykonują na przemian ruchy dowolną ze swoich figur, zgodnie z jej zasadami poruszania się. W trakcie tury zawodnikowi upływa czas ustalony przed grą. W niektórych wariantach gracz otrzymuje też niewielką ilość czasu za wykonanie każdego ruchu.

Wygrana następuje, gdy król jednego z graczy jest atakowany i nie można w legalny sposób nim ruszyć, ani zasłonić jedną ze swoich figur przed atakiem. Pozycję taką nazywa się „matem”. Drugim sposobem na wygraną jest skończenie się czasu jednego z graczy, niezależnie od sytuacji na szachownicy. Rozrywka może także zakończyć się remisem. Następuje on, gdy żaden z graczy nie ma na planszy figur, które mogą pozwolić na „zamatowanie” przeciwnika. Inną możliwością jest trzykrotne powtórzenie się na planszy tej samej pozycji. Do remisu doprowadza też sytuacja, w której jednemu z graczy zakończył się czas, a jego przeciwnik nie ma figur pozwalających na wygraną lub gdy jeden z graczy nie ma możliwości wykonania żadnego legalnego ruchu, a jego król nie jest atakowany.

W związku z możliwością przegranej poprzez upłynięcie czasu na zegarze, zawodnicy muszą indywidualnie określić podczas gry, ile czasu są w stanie poświęcić danemu ruchowi, tak by nie stracić na niego zbyt wiele czasu, ale też, żeby ruch był jak najlepszy, co jest przedmiotem badań niniejszej pracy.

---

<sup>1</sup>Pełne aktualne przepisy można znaleźć w książce: FIDE. *The Official Laws of Chess and Other Fide Regulations* (Macmillan Chess Library). Collier Book, 1990

## 1.2 System rankingowy Glicko-2

System rankingowy ELO - pierwowzór systemu Glicko-2, został zaprezentowany w latach pięćdziesiątych dwudziestego wieku przez Węgierskiego fizyka i szachistę Arpada Elo (1903-1992) (patrz Rozdział 1 w [4]). Początkowo był używany jedynie w szachach, jednak wraz ze wzrostem jego popularności zaczął być stosowany również w innych rozgrywkach z możliwością pomiaru poziomu zawodników. System ten jest pierwszym systemem mającym podłoże probabilistyczne i jest oparty na rozkładzie normalnym z ustaloną średnią. Przyznaje odpowiednią liczbę punktów zwycięzcy rozgrywki i odbiera przegranemu bazując na różnicy między ich aktualnym rankingiem.

System Glicko-2 używany przez stronę *Lichess.com*, na danych której oparta jest niniejsza praca, opracowany został przez Marka Glickmana jako ulepszenie systemu ELO. Podstawową zmianą jest uwzględnienie historycznych wyników każdego z zawodników w celu ustalenia wariancji aktualnego rankingu. Glickman w swojej pracy z roku 1999 [5] przedstawia problem dwóch graczy o takim samym rankingu, z których jeden gra regularnie, a drugi wrócił do gry po długiej przerwie. System Glicko-2 przyznając punkt za grę bierze pod uwagę wiarygodność każdego z rankingów. Zawodnikowi grającemu regularnie zostanie przyznane bądź odebrane mniej punktów ze względu na duże potencjalne odchylenie rankingu przeciwnika od zadeklarowanej wartości. Innymi słowy, w miarę zwiększania się liczby partii gracza, przedział ufności dla jego realnego rankingu zawęży się i przypisany mu ranking zbiega do realnego poziomu i wiarygodność przypisanego rankingu jest uwzględniana w przyznawaniu i odbieraniu punktów zawodnikom po zakończeniu partii.

## 1.3 Funkcja oceny

Przed przystąpieniem do opisania funkcji oceny, należy przybliżyć działanie silnika szachowego, który wykorzystywany jest do określenia aktualnej pozycji.

### 1.3.1 Stockfish

Silnik szachowy Stockfish [1] jest jednym z najlepszych i najpopularniejszym obecnie używanym silnikiem szachowym, którego zaprojektowali Joost VandeVondele, Joon Kiiski, Marco Costalb, Tord Romstad, Gary Linscott, Stefan Geschwentner i Stéphane Nicolet. Jest stale ulepszany jako oprogramowanie typu open-source. Strona *Lichess.com*, tak jak większość stron internetowych poświęconych szachom on-line, wykorzystuje go do analizy i oceny aktualnej pozycji [6].

Stockfish poprzez przeszukiwanie według strategii mini-max<sup>2</sup> z odcięciem, za pomocą algorytmu alfa-beta, analizuje legalne (czyli następujące po ruchu zgodnym z zasadami gry) pozycje, które mogą wynikać z aktualnej sytuacji na szachownicy. Dobierają na podstawie najlepszego możliwego zestawu ruchów (zakłada się, że każdy z graczy wykona najlepszy w ocenie silnika ruch) pozycje, które wystąpią dla określonej głębokości (głębokość 18 oznacza 18 ruchów białych i 18 czarnych wykonanych zaczynając z analizowanej pozycji) i na ich podstawie ocenia aktualną pozycję, określając przewagę jednego z graczy.

---

<sup>2</sup>Dokładniej opisana w artykule: E. Vilkas. *Axiomatic definition of the value of a matrix game* Theory Probabl. Appl. , 8 (1963) pp. 304–307

### 1.3.2 Ewaluacja

Wspomniana wcześniej ewaluacja, wyliczana przez silnik szachowy jest wynikiem liniowej funkcji ważonej sumy cech, na którą składają się między innymi:

- $f_b, f_c$  oznaczających wartość figur odpowiednio białych i czarnych.
- $k_b, k_c$  oznaczających bezpieczeństwo króla odpowiednio białych i czarnych.
- $m_b, m_c$  oznaczających mobilność figur odpowiednio białych i czarnych.
- $z_b, z_c$  oznaczających potencjalne zagrożenia wykonane odpowiednio białych i czarnych.

Funkcję można dla zapewnienia intuicji zapisać w uproszeniu:

$$f(f_b, f_c, k_b, k_c, m_b, m_c, \dots) = c_1(f_b - f_c) + c_2(k_b - k_c) + c_3(m_b - m_c) + \dots \quad (1.1)$$

gdzie:  $c_i$  są stałymi określającymi wagę danej pary zmiennych.

Wraz ze wzrostem wartości funkcji zwiększa się przewaga białych, natomiast wraz z jej spadkiem, przewaga czarnych. Wartość wynosząca 0 oznacza stan równowagi, czyli sytuacji, w której żadna ze stron nie ma przewagi lub gracz z gorszą sytuacją na szachownicy jest w stanie przy odpowiedniej kombinacji ruchów doprowadzić do remisu. Dodatkowo, w przypadku nieuniknionego zwycięstwa jednej ze stron w  $n$  ruchach, wynikiem funkcji zamiast odpowiedniej wartości liczbowej jest tekst  $\#-n$  w przypadku wygranej czarnych lub  $\#n$  w przypadku wygranej białych, oznaczający nieuchronną wygraną jednego z graczy po wykonaniu  $n$  ruchów, przy najlepszej obronie przeciwnika.

### 1.3.3 Rodzaje błędów szachowych

W notacji szachowej obok zapisanego ruchu mogą pojawić się symbole określające jakość danego ruchu. Dla analizowanych danych, ruch oceniany jest przez silnik szachowy za pomocą skomplikowanych algorytmów, których głównym składnikiem jest zmiana oceny pozycji na szachownicy przez silnik. Przykładowo, gdy przed ruchem czarnych ocena pozycji wynosiła -2, a po ruchu 3, silnik oceni ruch jako błąd, jednakże jeśli wcześniejsza ocena wynosiła -18, a po ruchu -13, silnik nie oznaczy ruchu jako błędny, ponieważ przewaga czarnych i tak jest bardzo znacząca.

Wśród licznych oznaczeń silnika, w pracy analizowane będą następujące oznaczenia [7]:

- ?? – błąd (ang. *blunder*)
- ? – pomyłka (ang. *mistake*), posunięcie błędne w mniejszym stopniu niż „błąd”
- ?! – niedokładność (ang. *innacuracy*), posunięcie, które można zastąpić zdecydowanie lepszym.

W części pracy na wykresach pojawiać się będzie oznaczenie „duży błąd”. Ma ono służyć lepszemu rozróżnieniu ruchów oznaczonych jako „błąd” i „pomyłka”



# Rozdział 2

## Zagadnienia teoretyczne dotyczące analizy statystycznej

Niniejszy rozdział poświęcony zostanie zagadnieniom teoretycznym z dziedziny statystyki, zastosowanych w analizie przedstawionych problemów. Pokróćce umówione zostaną metody wykorzystywane przy rozważaniach.

### 2.1 Metoda najmniejszych kwadratów

Weźmy zbiór punktów  $(x_i, y_i)$  dla  $0 \leq i \leq n$ .

Metoda najmniejszych kwadratów [8] polega na znalezieniu takich parametrów  $\theta_1, \theta_2, \dots, \theta_k$  pewnej określonej funkcji  $\hat{y}_i = f(x_i, \theta_1, \theta_2, \dots, \theta_k)$  estymującej wartość  $y_i$  w zadanym punkcie  $x_i$ , dla których suma kwadratów odległości  $\hat{y}_i$  od rzeczywistej wartości  $y_i$  dla każdego punktu ze zbioru jest najmniejsza. Innymi słowy, metoda polega na minimalizacji funkcji

$$L = \sum_{i=0}^n (y_i - f(x_i, \theta_1, \theta_2, \dots, \theta_k))^2$$

biorąc jako zmienne parametry  $\theta_1, \theta_2, \dots, \theta_k \in \mathbb{R}$

### 2.2 Regresja liniowa

Model regresji liniowej [9] (strony 265 – 267) polega na dopasowaniu parametrów  $\beta_0$  i  $\beta_1$  prostej  $y = \beta_0 + \beta_1 x$  do zestawu danych  $(x_i, y_i)$  dla  $0 \leq i \leq n$  za pomocą opisanej wyżej metody najmniejszych kwadratów. Prosta regresji pomaga ustalić trend liniowy w danych.

#### 2.2.1 Estymacja parametrów prostej regresji metodą najmniejszych kwadratów

Problem polega na minimalizacji funkcji  $L(\beta_0, \beta_1) = \sum_{i=0}^n (y_i - (\beta_0 + \beta_1 x_i))^2$  ze względu na  $\beta_0, \beta_1$ .

$$0 = \frac{\partial L}{\partial \beta_0} = -2 \sum_{i=0}^n (y_i - (\beta_0 + \beta_1 x_i)) \quad (2.1)$$

$$0 = \frac{\partial L}{\partial \beta_1} = -2 \sum_{i=0}^n x_i (y_i - (\beta_0 + \beta_1 x_i)) \quad (2.2)$$

Stąd

$$\sum_{i=0}^n y_i - n\beta_0 - \beta_1 \sum_{i=0}^n x_i = 0 \quad (2.3)$$

$$\sum_{i=0}^n x_i y_i - \beta_0 \sum_{i=0}^n x_i - \beta_1 \sum_{i=0}^n x_i^2 = 0 \quad (2.4)$$

Z równania 2.3 otrzymujemy

$$\beta_0 = \frac{1}{n} \left( \sum_{i=0}^n y_i - \beta_1 \sum_{i=0}^n x_i \right) = \bar{y} - b_1 \bar{x}, \text{ gdzie } \bar{x} = \frac{1}{n} \sum_{i=0}^n x_i, \bar{y} = \frac{1}{n} \sum_{i=0}^n y_i \quad (2.5)$$

Po wstawieniu powyższego wyrażenia do równania 2.4 oraz po kilku prostych przekształceniach dostajemy

$$\sum_{i=0}^n x_i (y_i - \bar{y}) = \beta_1 \left( \sum_{i=0}^n x_i^2 - n\bar{x}^2 \right) \quad (2.6)$$

Korzystając z faktu, że  $\sum_{i=0}^n x_i^2 - n\bar{x}^2 = \sum_{i=0}^n (x_i - \bar{x})^2$  otrzymujemy ostateczny wynik estymacji:

$$\beta_1 = \frac{\sum_{i=0}^n x_i (y_i - \bar{y})}{\sum_{i=0}^n (x_i - \bar{x})^2} \quad (2.7)$$

$$\beta_0 = \bar{y} - b_1 \bar{x} \quad (2.8)$$

## 2.3 Metoda momentów

Celem metody momentów [10] jest estymacja parametrów rozkładu prawdopodobieństwa  $X(\theta_1, \theta_2, \dots, \theta_k)$ . Metoda polega na stworzeniu, a następnie rozwiązaniu układu równań, w którym przyrównywane są empiryczne wartości pierwszych  $k$  momentów z ich wartościami teoretycznymi  $M(\theta_1, \theta_2, \dots, \theta_k)$ . Metody można użyć jedynie, gdy istnieje pierwsze  $k$  momentów danego rozkładu.

### 2.3.1 Estymacja parametrów rozkładu log-normalnego metodą momentów

Dla rozkładu log-normalnego  $Lognormal(\mu, \sigma^2)$  momenty zadane są równaniem [11]

$$E[X^n] = \exp\left(n\mu + \frac{(n\sigma)^2}{2}\right).$$

Po rozwiązaniu układu równań zawierających 2 pierwsze momenty, wyznaczone zostają parametry  $\mu$  i  $\sigma^2$ .

$$\mu = \log \left( \frac{E[X]^2}{\sqrt{D^2[X] + E[X]^2}} \right)$$

$$\sigma = \log \left( \frac{D^2[X]}{E[X]^2} + 1 \right)$$

gdzie  $E[X]$  jest średnią empiryczną rozkładu, a  $D^2[X] = E[X^2] - E[X]^2$  jego empiryczną wariancją.

## Rozdział 3

# Sformułowanie i rozwiązanie problemów analitycznych

### 3.1 Sformułowanie problemów

Podstawowym problemem podjętym w pracy jest analiza zależności pomiędzy czasem poświęconym na wykonanie ruchu, a jego dokładnością. Główny nacisk kładziony zostanie na ruchy oznaczone przez silnik jako błąd, co można uargumentować tym, że właśnie te ruchy mają największy wpływ na grę, szczególnie biorąc pod uwagę szachy szybkie, gdzie czasu na ruch jest zdecydowanie mniej, czego wynikiem jest to, że błędne ruchy pojawiają się znacznie częściej. Dodatkową osią ujętą w rozważaniach będzie ranking zawodników i jego wpływ na rozkłady danych. Sprawdzona i wzięta pod uwagę zostanie również zależność między numerem ruchu, a czasem na wykonanie oraz jego dokładnością.

### 3.2 Dane

Dane potrzebne do analizy zostały pobrane z platformy Lichess [2], która zapisuje i udostępnia wszystkie rozgrywane gry. Brane pod uwagę będą wszystkie gry (ponad 35 milionów) rozegrane na platformie w maju 2019 roku.

W pracy analizowane będą dwa najpopularniejsze formaty czasowe grane na platformie Lichess. Są one przedstawione w tabeli 3.1. Pierwszy z nich - „600+0” oznacza grę, w której każdy z graczy posiada 600 sekund na wykonanie wszystkich swoich ruchów, bez inkrementu (dodawanego czasu po zagranii posunięcia). Analogicznie wygląda format 300+0, z tym, że czas zmniejszony jest o połowę - do pięciu minut.

Tabela 3.1: Baza gier na portalu Lichess z maja 2019, 6 najpopularniejszych formatów

Format czasowy sekundy+sekundy przyznawane po wykonaniu ruchu	procent gier w bazie
600+0	18,63%
300+0	14,48%
60+0	12,85%
900+15	11,34%
180+0	10,86%
300+3	9,67%

Po zakończeniu gry na platformie, każdy gracz ma możliwość zbadać swoją grę przy użyciu silnika szachowego Stockfish. Około 7% gier z bazy zawiera taką analizę. Posiadają one dane punktowe o nazwie *Eval*, określające unormowaną przewagę jednego z graczy w danym momencie gry. Dodatkowo silnik określa jakość wykonanego ruchu według własnych kryteriów (opisane w sekcji 1.3.3 na stronie 7). Przykładowy zapis jednej takiej gry z pliku został zaprezentowany na Rysunku 3.1. Informacje potrzebne do analizy to:

- WhiteElo – ranking białych
- BlackElo – ranking czarnych
- TimeControl – czas na wykonanie ruchów każdego z graczy w formacie „sekundy + sekundy dodane za wykonanie ruchu”
- % eval – aktualna przewaga jednej ze stron
- zapis partii ze wskazaniem błędnych ruchów w ocenie silnika
- % clk – pozostały czas w formacie „godziny : minuty : sekundy”
- result – wynik partii

```
[event "Rated Bullet game"]
[site "https://lichess.org/oxdyeh38"]
[date "2019.04.30"]
[round "-"]
[white "daengtiro59"]
[black "7..."]
[result "0-1"]
[urcdate "2019.04.30"]
[urctime "22:06:31"]
[whiteelo "1879"]
[blackelo "1871"]
[whitekatingdiff "-30"]
[blackkatingdiff "+30"]
[eco "B00"]
[opening "Queen Defense"]
[timecontrol "0+0"]
[termination "time forfeit"]

1. e4 { [Eval 0.12] [Clk 0:01:00] } 1... b6? { [Eval 0.62] [Clk 0:01:00] } 2. d4 { [Eval 0.64] [Clk 0:00:57] } 2... Bb7 { [Eval 0.62] [Clk 0:01:00] } 3. Bd3 { [Eval 0.47] [Clk 0:00:57] } 3... g6 { [Eval 0.54] [Clk 0:01:00] }
4. c3 { [Eval 0.27] [Clk 0:00:55] } 4... Bg7 { [Eval 0.68] [Clk 0:00:57] } 5. h3 { [Eval 0.62] [Clk 0:00:51] } 5... e6 { [Eval 0.8] [Clk 0:00:55] } 6. f4? { [Eval -0.15] [Clk 0:00:48] } 6... Ne7 { [Eval 0.27] [Clk 0:00:54] }
7. Nf3 { [Eval 0.46] [Clk 0:00:40] } 7... d5 { [Eval 0.58] [Clk 0:00:53] } 8. Qc2? { [Eval -5.18] [Clk 0:00:45] } 8... dxe4 { [Eval -5.27] [Clk 0:00:52] } 9. Bxe4 { [Eval -5.76] [Clk 0:00:41] } 9... Bxe4 { [Eval -5.76] [Clk 0:00:51] }
10. Ng5? { [Eval -6.57] [Clk 0:00:48] } 10... Bb7 { [Eval -6.42] [Clk 0:00:47] } 11. Ne7? { [Eval -6.64] [Clk 0:00:38] } 11... e4? { [Eval -8.5] [Clk 0:00:45] } 12. Qe2? { [Eval -9.21] [Clk 0:00:33] } 12... Qd6 { [Eval -8.83] [Clk 0:00:42] }
13. Nc2 { [Eval -8.84] [Clk 0:00:29] } 13... Qc5 { [Eval -8.39] [Clk 0:00:29] } 14. Rf2 { [Eval -8.62] [Clk 0:00:24] } 14... R5 { [Eval -8.45] [Clk 0:00:37] } 15. Qd3 { [Eval -8.77] [Clk 0:00:23] } 15... R5 { [Eval -8.75] [Clk 0:00:35] }
16. Qd3 { [Eval -8.8] [Clk 0:00:18] } 16... Nc7 { [Eval -8.27] [Clk 0:00:27] } 17. Nc4? { [Eval -16.19] [Clk 0:00:14] } 17... Nf6? { [Eval -8.67] [Clk 0:00:24] } 18... Nf5? { [Eval -8.84] [Clk 0:00:12] }
18... Ke7? { [Eval -8.12] [Clk 0:00:20] } 19. Bc2? { [Eval -8.99] [Clk 0:00:08] } 19... Rxd8? { [Eval -8.3] [Clk 0:00:15] } 20. Re1 { [Eval -8.23] [Clk 0:00:06] } 20... Kf8? { [Eval -7.32] [Clk 0:00:12] }
21. Nxe4 { [Eval -7.3] [Clk 0:00:04] } 21... Kf7 { [Eval -6.56] [Clk 0:00:10] } 22. Rxd8+? { [Eval -7.87] [Clk 0:00:02] } 22... Rxd8 { [Eval -7.81] [Clk 0:00:08] } 0-1
```

Rysunek 3.1: Przykładowy zapis jednej partii zawierającej ocenę silnika Stockfish

Istotnym problemem związanym z danymi jest dokładność pozostałego czasu jedynie do sekund, co zaburza lekko ich istotę, szczególnie podczas analizy gier w krótszym formacie czasowym (jak np. „60+0”), gdzie ułamki sekundy są znaczące. Dodatkowo ruch, który oznaczony został jako  $n$  sekund w rzeczywistości jest ruchem z przedziału obustronnie otwartego  $(n-1, n+1)$  sekund. Przy dużej liczbie danych nie ma to jednak większego znaczenia, ponieważ zgodnie z Prawem Wielkich Liczb można takie dane uśrednić do  $n$  *cite*???. Jedynym problemem, który może się pojawić jest późniejsze sprawdzenie zgodności danych z rozkładem.

### 3.2.1 Odfiltrowanie danych

Pierwszym krokiem potrzebnym do wykonania analizy jest odfiltrowanie danych. Po rozdzieleniu pliku tekstowego na kolejne gry, pozostawione zostały jedynie te, które zostały wcześniej ocenione przez silnik oraz grane były w formacie „300+0” lub „600+0”. Następnie przygotowany wcześniej program zbadał każdy ruch każdej gry i przypisał do niego następujące atrybuty:

- game\_ID – numer gry



- score – ocena ruchu według silnika Stockfish
- delta\_time – czas na wykonanie danego ruchu
- WhiteElo – ranking białych
- BlackElo – ranking czarnych
- TimeControl – czas na wykonanie ruchów każdego z graczy w formacie „sekundy + sekundy dodane za wykonanie ruchu”
- color – gracz, wykonujący dany ruch
- move – numer ruchu w danej partii
- result - wynik partii

Stworzona baza zawiera 17 524 012 posunięć ze 275 994 gier, co daje średnią 31,75 ruchu na grę (jeden ruch oznacza posunięcie białych i czarnych). Fragment bazy przedstawiony został na Rysunku 3.2.

	game_ID	score	delta_time	WhiteElo	BlackElo	TimeControl	color	move	Result
<b>510</b>	9	0	0	1192	1204	300+0	w	23	0-1
<b>511</b>	9	0	9	1192	1204	300+0	b	23	0-1
<b>512</b>	9	0	2	1192	1204	300+0	w	24	0-1
<b>513</b>	9	0	18	1192	1204	300+0	b	24	0-1
<b>514</b>	9	0	7	1192	1204	300+0	w	25	0-1
<b>515</b>	9	0	10	1192	1204	300+0	b	25	0-1
<b>516</b>	9	blunder	8	1192	1204	300+0	w	26	0-1
<b>517</b>	9	blunder	8	1192	1204	300+0	b	26	0-1
<b>518</b>	9	mistake	4	1192	1204	300+0	w	27	0-1
<b>519</b>	9	0	2	1192	1204	300+0	b	27	0-1
<b>520</b>	9	0	29	1192	1204	300+0	w	28	0-1

Rysunek 3.2: Fragment bazy zawierającej ruchy z gier o formatach czasowych „300+0” i „600+0”.

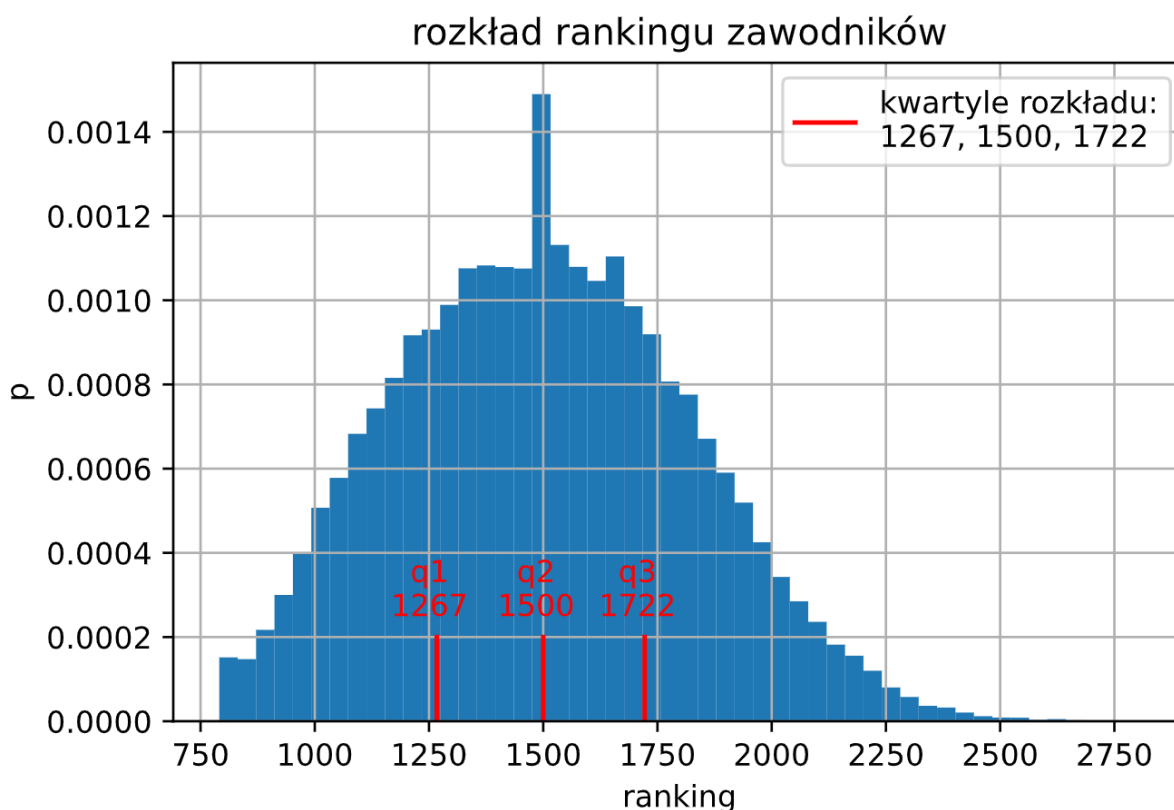
## 3.3 Analiza problemu

### 3.3.1 Wstępny przegląd danych

Pierwszym problemem, który zostanie poruszony jest zbadanie statystycznej zależności jakości wykonanego ruchu według oceny silnika Stockfish od czasu potrzebnego na jego wykonanie. W celu uzyskania bardziej precyzyjnych wyników pierwsze 4 ruchy (po 4 posunięcia białych i czarnych) nie będą brane pod uwagę. Są one elementem teorii otwarc

szachowych <sup>1</sup>, w związku z czym wykonywane są zazwyczaj bardzo szybko i ze znikomą szansą popełnienia błędu, co może powodować zaburzenie danych. Dodatkowo na stronie *Lichess.com* pierwszy ruch obydwu zawodników zawsze zabiera 0 sekund.

Przy analizie z uwzględnieniem rankingu graczy, zastosowane będą podziały kwartylowe. Rozkład rankingu wraz z zaznaczonymi kwartylami pokazany jest na Rysunku 3.3. Jak widać zbiega on do rozkładu normalnego. Anomalia dla rankingu 1500 związana jest z tym, że jest to ranking przyznawany każdemu graczowi przy jego pierwszej rozgrywce na platformie.



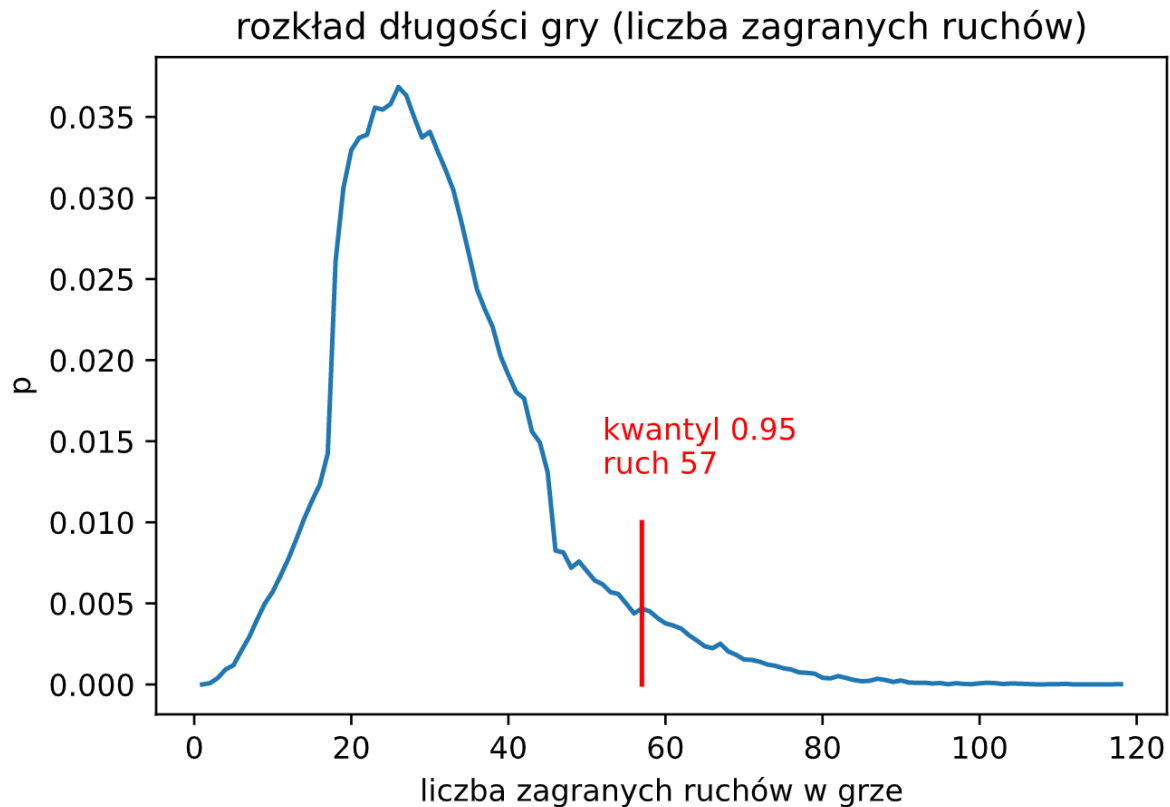
Rysunek 3.3: Rozkład rankingu zawodników wraz z zaznaczonymi kwartylami

W późniejszej części pracy pojawią się wykresy i analizy, w których dziedziną będzie numer wykonywanego ruchu. W związku z bardzo małą liczbą danych dla późniejszych ruchów, ich analiza jest zaburzona i niewiarygodna. Dlatego analizowane będą jedynie ruchy z indeksem poniżej 59, jako, że wynik 95% gier pojawia się maksymalnie w 58 ruchach. Rozkład długości gier z bazy danych, który obrazuje przedstawioną sytuację zamieszczony jest na Rysunku 3.4.

### 3.3.2 Zależność między indeksem ruchu, a poświęconym czasem

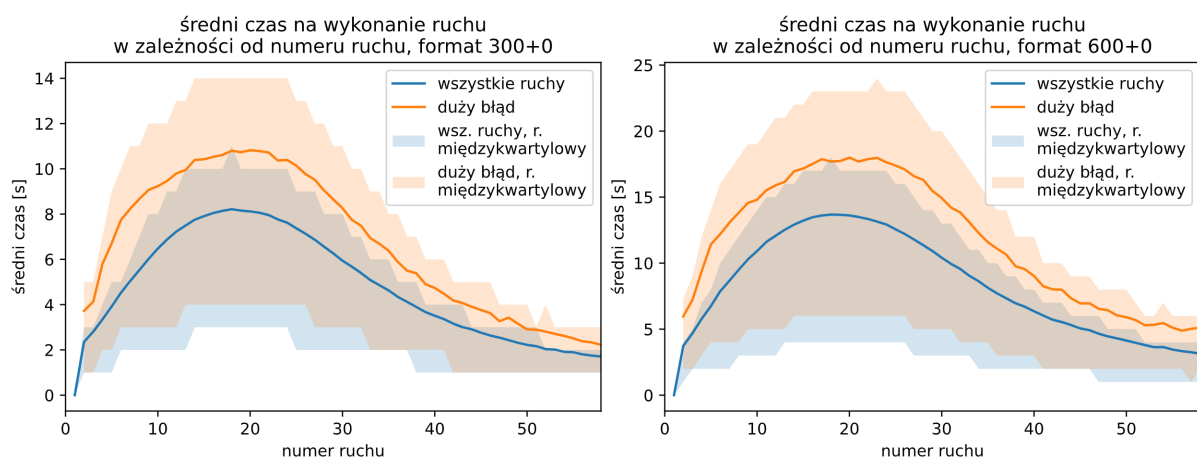
Na początku tej części trzeba ustalić pewne założenia dotyczące zestawu danych. W pojedynczej grze szachowej często występuje sytuacja, w której gracz długo myśli nad konkretnym posunięciem, rozpatrując możliwe następujące po nim sekwencje. W określonych okolicznościach, takich sekwencji jest na tyle mało, że każdy z graczy ma przez kilka

<sup>1</sup>Jedynie otwarcia, które składają się z mniej niż czterech ruchów są wariacjami otwarcia *Fool's Mate*, które prowadzi do zakończenia gry w dwóch lub trzech ruchach, bądź sekwencjami posunięć, które nie mają logicznych podstaw.



Rysunek 3.4: Rozkład długości gier wraz z zaznaczonym kwantylem rzędu 0.95, dla gier z formatu „300+0” oraz „600+0”.

kolejnych ruchów tylko jedno nieprzegrywające, łatwe do znalezienia posunięcie. Czas na wykonanie takich ruchów jest więc od siebie w pewien sposób zależny, jednakże sytuacje takie pojawiają się na bardzo nieoczekiwanie i przez mnogość zmiennych są trudne do znalezienia i uwzględnienia. Dlatego w celu ułatwienia analizy, w dalszych rozważaniach następujące po sobie ruchy będą uznawane za niezależne względem siebie.



Rysunek 3.5: Średni czas na wykonanie ruchu dla analizowanych formatów czasowych wraz z zaznaczonym rozstępem międzykwartylowym. Osobno wszystkie ruchy i ruchy oznaczone przez silnik jako błąd.

Pierwszym elementem analizy, który pozwoli na określenie wstępnej różnicy między

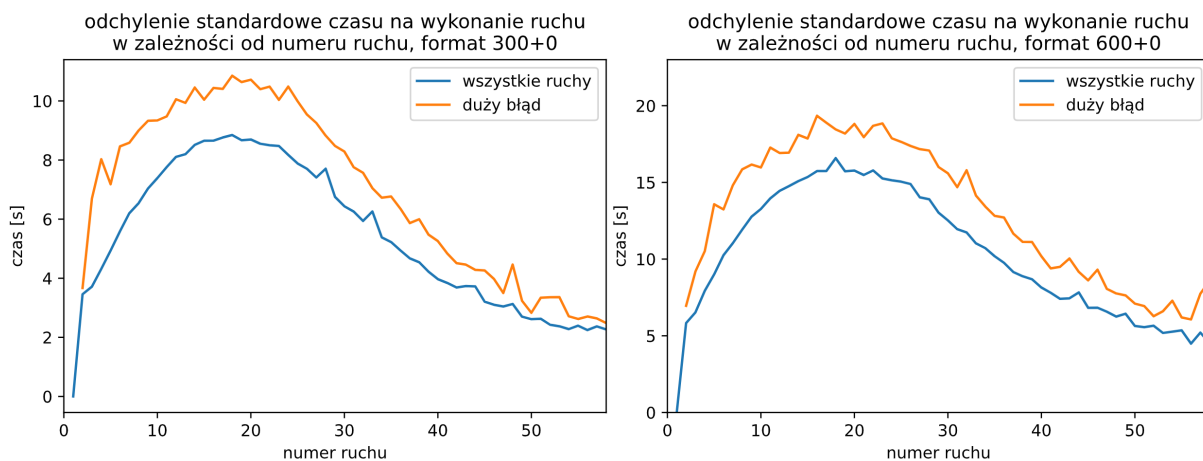
ruchami oznaczonymi jako błąd, a zbiorem wszystkich ruchów jest sprawdzenie zależności między indeksem ruchu, a średnim czasem na jego wykonanie dla obydwu zbiorów. Intuicja wskazuje, że błąd popełniamy poświęcając mniej czasu na ruch. Jak widać na Rysunku 3.5, dla każdego analizowanego ruchu, ruch oznaczony jako błędny zajął statystycznie więcej czasu - zarówno dla formatu 300+0 jak i 600+0, co przeczy intuicji. Przyczyną tego może być zaliczanie do zbioru wszystkich ruchów posunięć oczywistych, zajmujących niewielką ilość czasu, następujących dla pozycji, w której gracz na każdym poziomie zaawansowania rzadko kiedy popełnia błędy. Kolejną przyczyną takiego stanu rzeczy może być oczywista zależność skomplikowania pozycji od szansy na popełnienie błędu. W trudnych sytuacjach na szachownicy błędów pojawia się po prostu więcej, a trudnych sytuacji najwięcej jest w środkowej części partii, przez mnogość figur i możliwych legalnych ruchów. Trudno też znaleźć takie pozycje, ponieważ zależą one od przyjętej miary, o czym pisze D. H. Holding w swojej pracy zatytułowanej „*Evaluation of Chess Positions*” [12]. Dlatego też, co dobrze widoczne jest na wykresie, w środkowej części gry posunięcia zajmują najwięcej czasu.

Oznaczając przez  $D$  zmienną określającą skomplikowanie pozycji,  $T$  - zmienną określającą czas na wykonanie ruchu w danej pozycji,  $B$  - zdarzenie polegające na tym, że posunięcie jest błędne i przy założeniu, że  $T$  jest silnie dodatnio skorelowane z  $D$ , według prawdopodobieństwa otrzymujemy

$$P(B|D > d_0) > P(B|D \leq d_0) \rightarrow P(B|T > t_0) > P(B|T \leq t_0), \quad (3.1)$$

gdzie  $d_0$  i  $t_0$  określają punkty, od których według wybranej miary można określić, że pozycja jest skomplikowana ( $D > d_0$ ) i czas na wykonanie posunięcia jest długi ( $T > t_0$ ).

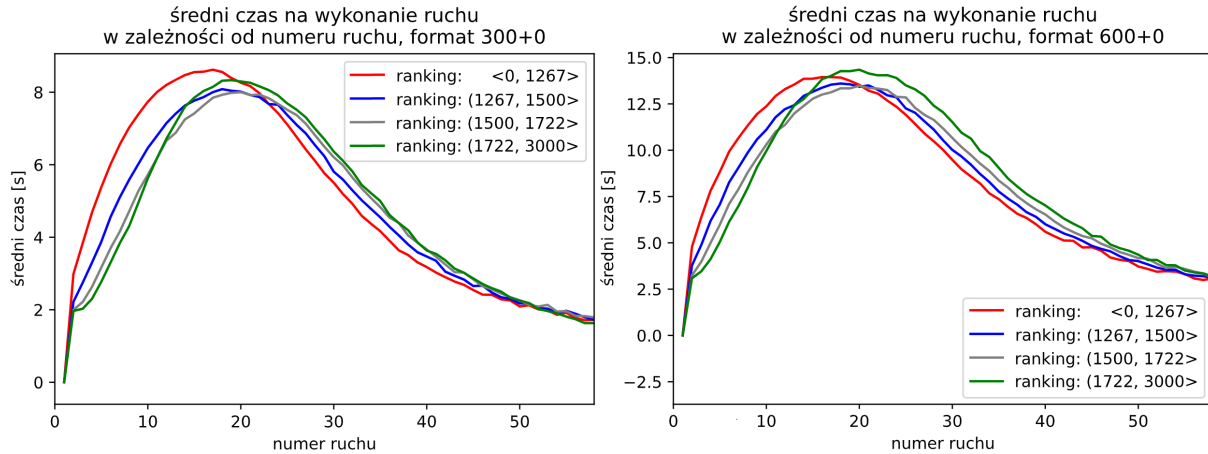
Może to tłumaczyć wyższą średnią czasu na wykonanie błędnego posunięcia w porównaniu do zbioru wszystkich posunięć.



Rysunek 3.6: Odchylenie standardowe czasu na wykonanie ruchu dla analizowanych formatów czasowych. Osobno wszystkie ruchy i ruchy oznaczone przez silnik jako błąd.

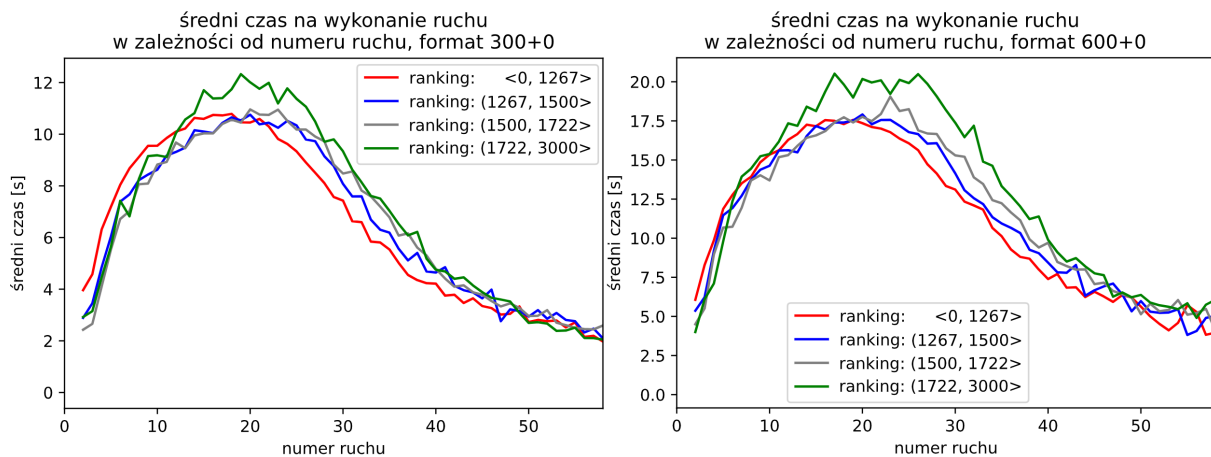
Badając w dalszym ciągu wykresy na Rysunku 3.5 można zauważyć, że dla obydwu zbiorów posunięć początkowo niska średnia osiąga swój szczyt w okolicach dwudziestego ruchu, by później stopniowo się zmniejszać. Ma to związek z tym, że początkowe posunięcia są elementem teorii szachowej. Spora część graczy nauczyła się kilku, kilkunastu, bądź w przypadku gry na najwyższym poziomie - nawet ponad dwudziestu pierwszych ruchów niektórych otwarć. W związku z tym początkowe ruchy grane są relatywnie szybko. Końcowe posunięcia natomiast nie mogą zająć wiele czasu, ponieważ jest go zazwyczaj zbyt

mało na zegarze, by gracz mógł sobie na to pozwolić. Dlatego też od pewnego momentu, wraz ze wzrostem indeksu średni czas poświęcony na ruch zbiega do zera. Zaznaczone na wykresie rozstępy międzykwartylowe dla każdego ruchu wskazują na znacznie większą wariancję czasu na wykonanie błędnego ruchu porównując do zbioru wszystkich ruchów. Potwierdza to Rysunek 3.6 pokazujący odchylenie standardowe dla kolejnych ruchów.



Rysunek 3.7: Średni czas na wykonanie ruchu dla graczy z różnych przedziałów rankingowych. Wszystkie posunięcia.

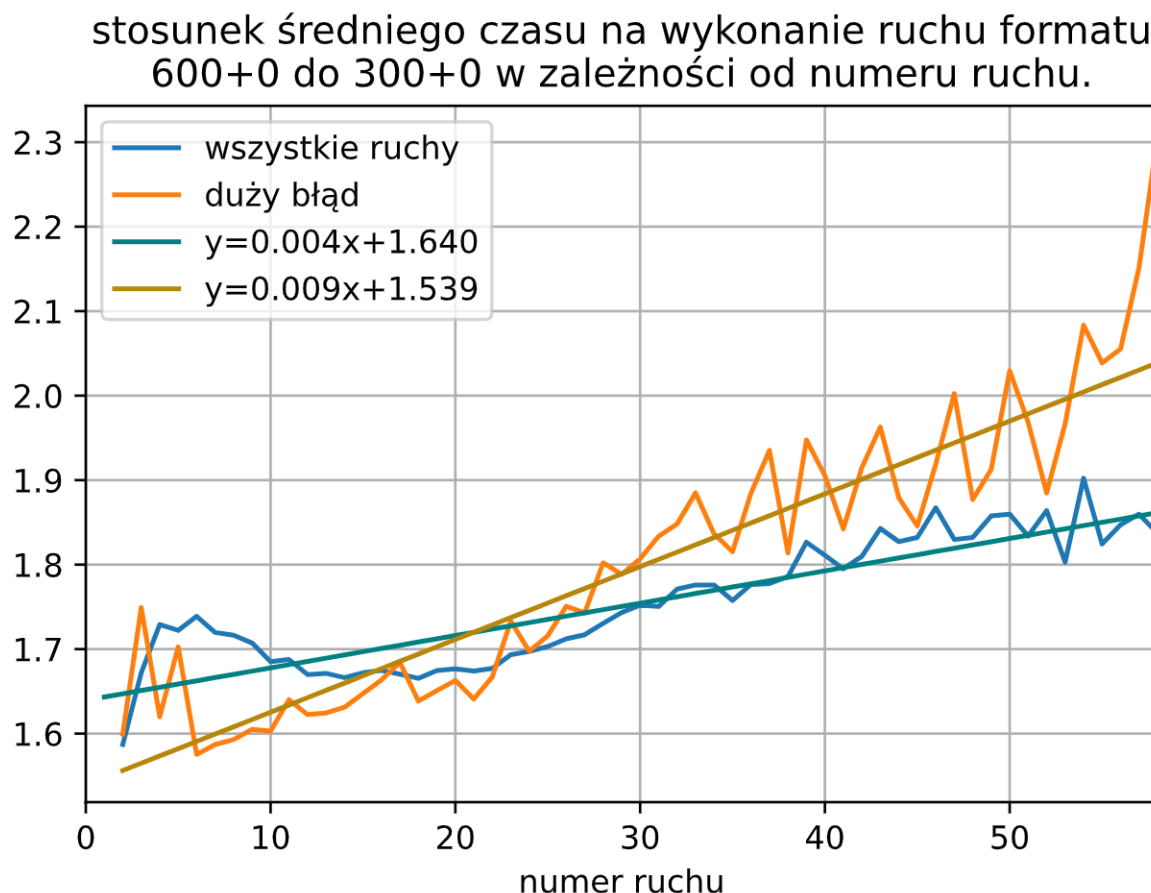
Rysunki 3.7 i 3.8 wprowadzają do rozważań różnicę między średnim czasem na wykonanie ruchu dla graczy z poszczególnych przedziałów rankingowych. Można zauważyć, że zarówno w przypadku zbioru wszystkich ruchów (Rysunek 3.7) jak i pozdbioru ruchów błędnych (Rysunek 3.8), słabsi gracze mają tendencję do dłuższych rozważań nad początkowymi ruchami. Wraz ze wzrostem rankingu wydłuża się średni czas na wykonanie ruchu dla ruchów z indeksem od 20 do 30 - czyli tych, gdzie średnia jest najwyższa. Przy końcowych posunięciach, posunięcia zajmują podobną ilość czasu dla wszystkich rankingów. Warto zwrócić uwagę, że odsetek 25% najniżej notowanych graczy przez zbyt długie początkowe ruchy zmuszona jest ruszać się szybciej w późniejszych posunięciach. Szczególnie widoczne jest to na pierwszym wykresie Rysunku 3.8.



Rysunek 3.8: Średni czas na wykonanie ruchu dla dla graczy z różnych przedziałów rankingowych. Posunięcia oznaczone przez silnik jako błąd.

Niezależnie od rankingu, średnia poświęconego czasu dla posunięć oznaczonych przez

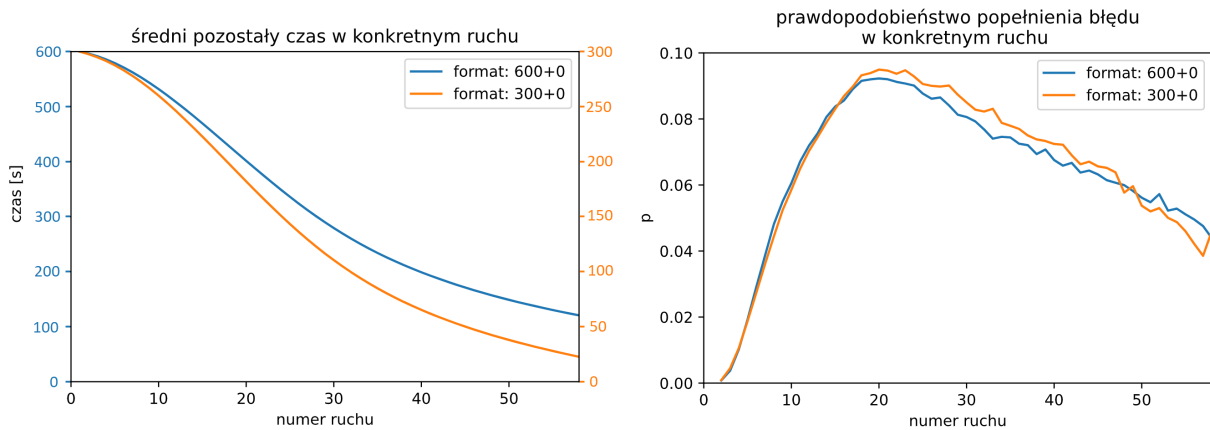
silnik jako błąd jest w każdym momencie wyższa niż dla zbioru wszystkich ruchów. Warto również zauważyć, że wykresy dla obydwu formatów czasowych wyglądają podobnie. By



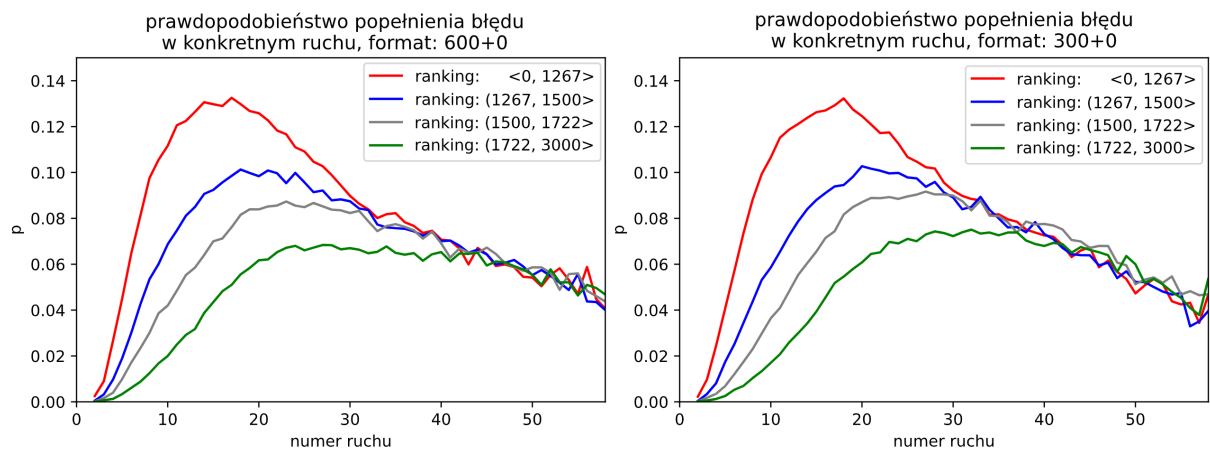
Rysunek 3.9: Stosunek średniego czasu na wykonanie ruchu w formacie 600+0 do czasu w formacie 300+0. Zaznaczone proste regresji obrazujące trend.

móc przeanalizować zależność między rozgrywanym formatem, a wzrostem średniego czasu poświęconego na ruch, powstał wykres umieszczony na Rysunku 3.9. Dopasowane metodą najmniejszych kwadratów opisanej w rozdziale 2 proste regresji obrazują różnice w zachowanych trendach. Czas w formacie 600+0 jest dokładnie 2 razy dłuższy niż w formacie 300+0, jednak średni czas na wykonanie ruchu jest w pierwszym przypadku około 1,7 raza wyższy. Wartość stosunku jednak nie jest stała i wzrasta wraz z indeksem ruchu. Dla zbioru wszystkich posunięć dopasowana prosta regresji wskazuje na stosunek około 1,65 dla tych początkowych, do 1,85 przy końcowych. Znacznie większa rozpiętość w stosunku widoczna jest dla posunięć oznaczonych jako błąd. Stosunek wynosi około 1,55 dla początkowych wartości i powyżej 2 dla ruchów z indeksem powyżej 55. Warto zauważyć, że dla obu zbiorów w formacie 600+0 gracze nie są zmuszeni do podejmowania decyzji pod dużą presją czasu. Szczególnie różnicę tę da się zauważyć dla ruchów błędnych, nad którymi gracz musi się statystycznie dłużej zastanowić (Rysunek 3.5 na stronie 15). Gracz posiadając 2 razy więcej czasu w tym formacie i myśląc nad każdym ruchem średnio od 1.55 do 1.95 raza dłużej, ma w każdym momencie gry do dyspozycji statystycznie większy ułamek początkowego czasu niż w przypadku gry w formacie 300+0. Zestawienie pozostałego czasu pokazane jest na pierwszym wykresie na Rysunku 3.10. Widać dość sporą różnicę w średnim pozostałym czasie narastającą podczas przebiegu gry. Drugi wykres ukazuje różnicę w prawdopodobieństwie popełnienia błędu w konkretnym ruchu

dla obu formatów. Warto zauważyć, że mimo znacznie większego czasu do dyspozycji w przypadku formatu „600+0”, do 20 ruchu szansa na błąd jest niemalże identyczna jak w przypadku formatu „300+0” – Krzywe odbiegają od siebie nieznacznie. W okolicach od 20 do 40 ruchu, szansa na popełnienie błędu w formacie „300+0” jest około 5-10% większa. Dobrze jest w tym miejscu dodać, że jest to przedział, w którym błędów występuje najwięcej.



Rysunek 3.10: Pierwszy wykres – zestawienie średniego pozostałego czasu w formatach 600+0 oraz 300+0. W celu lepszego porównania zastosowane osobne osie dla każdego formatu. Drugi wykres – empiryczne prawdopodobieństwo popełnienia błędu w konkretnym ruchu.



Rysunek 3.11: Prawdopodobieństwo popełnienia błędu w konkretnym ruchu dla różnych przedziałów rankingowych.

Różnica w prawdopodobieństwie popełnienia błędu pomiędzy poszczególnymi formatami nie jest tak duża jak ta pomiędzy rankingiem graczy. Na Rysunku 3.11 można zauważyć, że dla zarówno dla formatu „600+0” jak i „300+0” słabsi gracze mają tendencję do popełniania większej liczby błędów w początkowej fazie gry, niż ich wyżej notowani koledzy. Statystycznie, będąc w grupie 25% najniżej notowanych graczy ma się około 6 razy większe prawdopodobieństwo na popełnienie błędu w 10 posunięciu, niż osoba z grupy 25% najwyżej notowanych graczy. Format gry ma na to niewielki wpływ. Można wysunąć wniosek, że w przypadku gry pomiędzy dwoma takimi zawodnikami, dwukrotnie zwiększenie czasu niżej notowanego gracza z 5 do 10 minut będzie miało minimalny wpływ



na finalny wynik gry (przy założeniu, że liczba błędów jest mocno skorelowana z wynikiem gry). Krzywe zaczynają do siebie zbiegać od okolic 20 ruchu, by stać się praktycznie nierozróżnialne w okolicach 40 ruchu.

Wykres zależności prawdopodobieństwa od ruchu jest tym bardziej „spłaszczony” im wyższa jest notacja elo zawodników. Brak szybkiego wzrostu dla początkowych posunięć sprawia, że statystyczne prawdopodobieństwo popełnienia błędu w pierwszych 10 ruchach niesamowicie różni się dla poszczególnych grup rankingowych. Po oznaczeniu przez  $X_i$  zmiennej losowej określającej popełnienie błędu w  $i$ -tym posunięciu:

$$X_i = \begin{cases} 1, & \text{gdy w } i\text{-tym ruchu został popełniony błąd} \\ 0, & \text{gdy w } i\text{-tym ruchu nie został popełniony błąd} \end{cases}$$

oraz założeniu, że  $X_i$  oraz  $X_j$  są niezależne dla  $i \neq j$  (drugie przejście w równaniu poniżej [13, str. 43]) możemy zbadać empiryczne prawdopodobieństwo popełnienia przynajmniej jednego błędu w  $n$  pierwszych ruchach.

$$\begin{aligned} P(X_1 = 1 \vee X_2 = 1 \vee \dots \vee X_n = 1) &= 1 - P(X_1 = 0, X_2 = 0, \dots, X_n = 0) = \\ &= 1 - P(X_1 = 0)P(X_2 = 0) \dots P(X_n = 0) = 1 - \prod_{i=1}^n P(X_i = 0) \end{aligned} \quad (3.2)$$

Przy czym  $P(X_i = 0)$  można wyznaczyć empirycznie na podstawie danych.

Tabela 3.2 przedstawia różnicę w szansie na popełnienie przynajmniej jednego błędu w pierwszych 10, 20 i 30 ruchach dla graczy z określonych przedziałów rankingowych i gier w określonym formacie. Przykładowo, dla osoby z grupy 25% najniżej notowanych graczy grającej w formacie „600+0”, szansa na popełnienie błędu w 10 pierwszych ruchach wynosi aż 43,46%, natomiast dla grupy 25% najlepszych graczy jest to jedynie 6,76%. Różnica ta zmniejsza się znacząco, gdy weźmie się pod uwagę większą liczbę pierwszych ruchów. Dla 20, jest to kolejno 85,50% i 40,78%, a dla 30 – 95,32% i 70,27%. Wartości te wyglądają podobnie dla formatu 300+0.

Tabela 3.2: Szansa na popełnienie błędu w pierwszych  $n$  ruchach dla gracza z określonego przedziału rankingowego z rozdzieleniem na formaty „600+0” i „300+0”

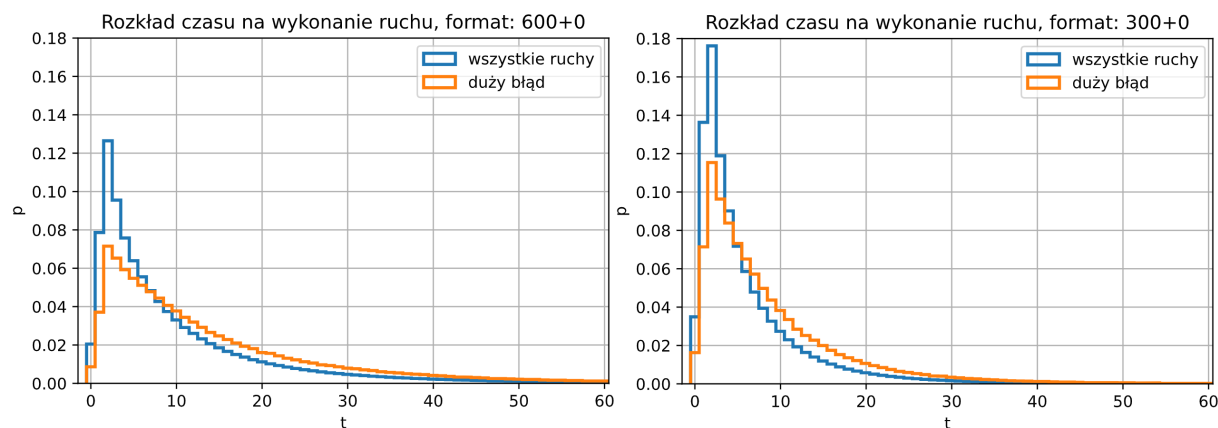
$n$ pierwszych ruchów	$n = 10$		$n = 20$		$n = 30$	
centyl \ format	600+0	300+0	600+0	300+0	600+0	300+0
0-25%	43,36%	40,72%	85,50%	84,33%	95,32%	94,96%
25%-50%	25,52%	22,09%	71,56%	68,63%	89,38%	88,66%
50%-75%	15,58%	12,42%	58,86%	56,27%	83,08%	82,93%
75%-100%	6,76%	5,63%	40,78%	38,66%	70,27%	70,40%
0%-100%	23,79%	22,81%	67,95%	67,58%	87,11%	87,50%

Co ciekawe, w pierwszej części tabeli, dla 10 pierwszych ruchów prawdopodobieństwo popełnienia błędu w dłuższym formacie czasowym jest wyższe niż w krótszym, niezależnie od rankingu zawodnika. Dla większej liczby pierwszych ruchów ta różnica się zmniejsza, jednak cały czas dłuższy format oznacza statystycznie większą szansę na choć 1 popełniony błąd, co wydaje się być sprzeczne z intuicją, ponieważ gracze mają 2 razy więcej czasu na wszystkie posunięcia.



### 3.3.3 Zależność między czasem poświęconym na ruch, a jego oceną

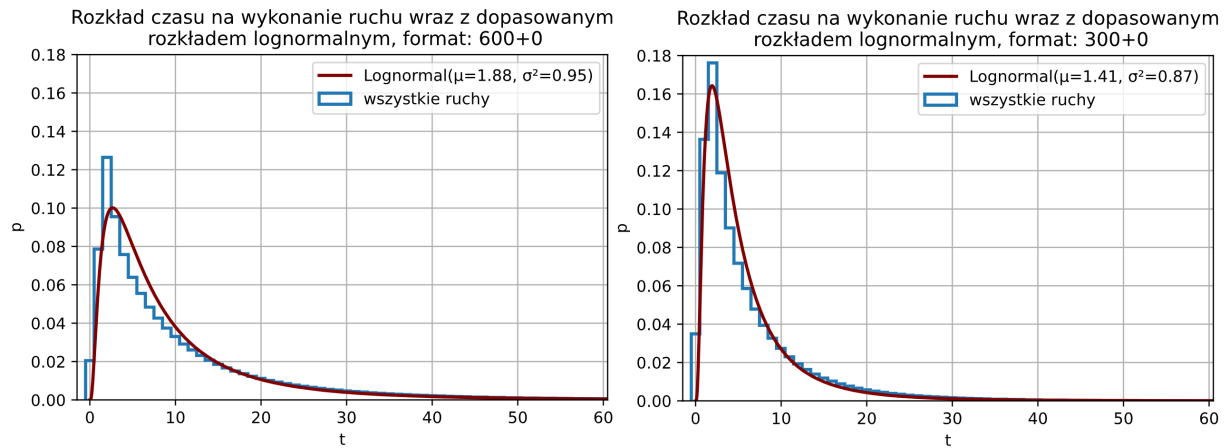
Kolejnym problemem, który zostanie zbadany jest zależność pomiędzy czasem poświęconym na ruch, a jego oceną. Tak jak i we wcześniejszych częściach wartość oceny silnika rozróżniana będzie jedynie na ruchy oznaczone jako błąd i te bez takiego oznaczenia.



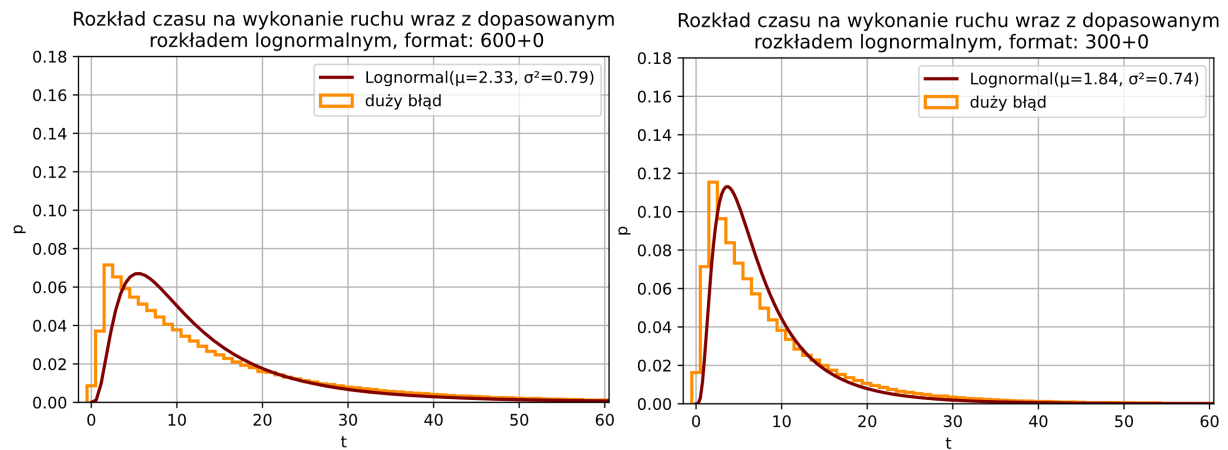
Rysunek 3.12: Rozkład czasu poświęconego na wykonanie posunięcia dla formatów „600+0” i „300+0”. Osobno wszystkie posunięcia i posunięcia błędne

Pierwszym spojrzeniem na wymienioną wyżej zależność będzie przeanalizowanie rozkładu czasu na wykonanie ruchu osobno dla zbioru wszystkich ruchów i dla tych błędnych. Jak można zauważyć na Rysunku 3.12, w każdym rozważanym formacie czasowym różnica w rozkładzie dla zbioru wszystkich ruchów i zbioru ruchów błędnych jest zauważalna. Oba zbiory mają zachowane podobne tendencje – najwięcej posunięć zagrywana jest w ciągu od jednej do pięciu sekund. W każdym ze zbiorów wraz ze wzrostem czasu na wykonanie posunięcia, empiryczna gęstość rozkładu się zmniejsza. Ruchów błędnych jest widocznie mniej w porównaniu do zbioru wszystkich ruchów w przedziale od 0 do ok. 8 sekund. Natomiast w przedziale od ok. 8 do 30 sekund, posunięcia błędne są spotykane częściej. Po przyjrzeniu się danym da się zauważyć, że mogą one pochodzić z rozkładu logarytmicznie-normalnego, co można uwarunkować tym, że w zdecydowanej większości pozycji zawodnik potrzebuje krótkiej chwili na odpowiedź na posunięcie przeciwnika, ponieważ samo zarejestrowanie zmian na planszy i fizyczne wykonanie posunięcia (przesunięcie figury na konkretne pole za pomocą myszki dla komputerów lub przeciągnięcie palcem dla smartfonów) zajmuje pewien czas, wykluczając sytuacje, w których zawodnik dokonuje tzw. „premove”, czyli podejmuje decyzję o wykonaniu konkretnego posunięcia zajmującego 0 sekund jeszcze przed zagranieniem przeciwnika. Dla formatów „300+0” i „600+0” jest to jednak rzadkością – szczególnie podczas środkowej części gry.

Za pomocą metody momentów opisaną w rozdziale na stronie 2 do każdego zbioru danych został dopasowany rozkład log-normalny. Na Rysunku 3.13 przedstawiającym znormalizowany histogram dla zestawu wszystkich ruchów dla formatów „300+0” i „600+0” rozkład teoretyczny jest dobrze dopasowany i zgodny z rozkładem empirycznym. Dla zestawu ruchów oznaczonych jako błędne, dopasowanie (Rysunek 3.14) w dużym stopniu odbiega od rozkładu empirycznego. Ruchy błędne zajmują statystycznie mniej czasu, niż wskazuje rozkład. Ze względu na wielkość prób wynoszących od ponad 500 tysięcy obserwacji dla najmniejszego z analizowanych zbiorów (format „300+0”, ruchy błędne) do aż 8,6 miliona (format „600+0”, wszystkie ruchy), najpopularniejsze testy statystyczne,



Rysunek 3.13: Rozkład czasu poświęconego na wykonanie posunięcia dla formatów „600+0” i „300+0” z dopasowanym rozkładem log-normalnym.



Rysunek 3.14: Rozkład czasu poświęconego na wykonanie posunięcia dla formatów „600+0” i „300+0” z dopasowanym rozkładem log-normalnym. Zestaw ruchów błędnych.

wrażliwe na liczbę danych odrzuciły hipotezę ich pochodzenia z rozkładu log-normalnego. Problematiczną kwestią jest tutaj również digitalizacja danych, które są zaokrąglane do pełnych sekund. W związku z powyższym zastosowane zostało kilka poprawek, by móc lepiej ustalić tożsamość danych. Pierwszą z nich jest zmiana rozkładu log-normalnego na dyskretny rozkład log-normalny [14]. Jest to drugi z badanych rozkładów w tegorocznej pracy „Discrete lognormal distribution with application to insurance data”, którą napisali Jiahang Lyu oraz Saraless Nadarajah. Funkcja prawdopodobieństwa przedstawionego rozkładu opisana jest następującym wzorem

$$p_X(i) = \Phi\left(\frac{\log i - \mu}{\sigma}\right) - \Phi\left(\frac{\log(i-1) - \mu}{\sigma}\right) \quad (3.3)$$

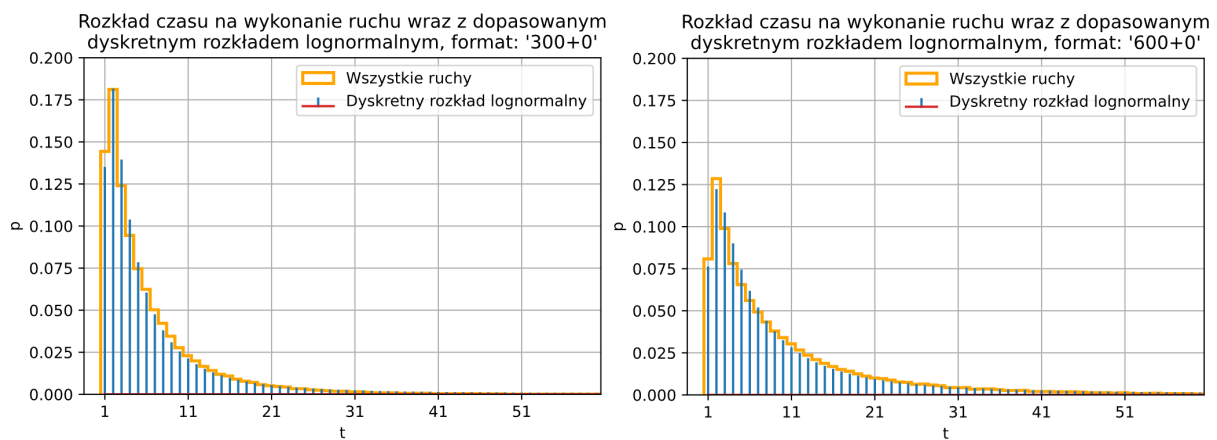
dla  $i \in \mathbb{N}_+$ , gdzie  $\Phi(x)$  jest wartością dystrybucyjną standardowego rozkładu normalnego w punkcie  $x$ . W celu estymacji parametrów  $\mu$  i  $\sigma$ , z każdego ze wzorów wzięta została losowa próbka  $(x_1, \dots, x_n)$  100 tysięcy obserwacji, a następnie numerycznie znaleziona optymalna wartość funkcji największej wiarygodności opisanej równaniem

$$L(\mu, \sigma) = \sum_{i=1}^n \log \left[ \Phi\left(\frac{\log x_i - \mu}{\sigma}\right) - \Phi\left(\frac{\log(x_i - 1) - \mu}{\sigma}\right) \right] \quad (3.4)$$

W równaniu 3.3 można zauważyć, że najniższą obserwacją należącą do dziedziny jest  $i = 1$  sekunda. W związku z tym, przed estymacją parametrów rozkładu, z danych usunięte zostały ruchy, których zaokrąglony czas wykonania wynosi 0 sekund. Dodatkowo wartość  $\log(0)$  traktowana jest jako  $\lim_{x \rightarrow 0^+} \log(x) = -\infty$ . W związku z tym

$$p_X(1) = \Phi\left(\frac{\log 1 - \mu}{\sigma}\right) - \Phi\left(\frac{\log(0) - \mu}{\sigma}\right) = \Phi\left(\frac{-\mu}{\sigma}\right).$$

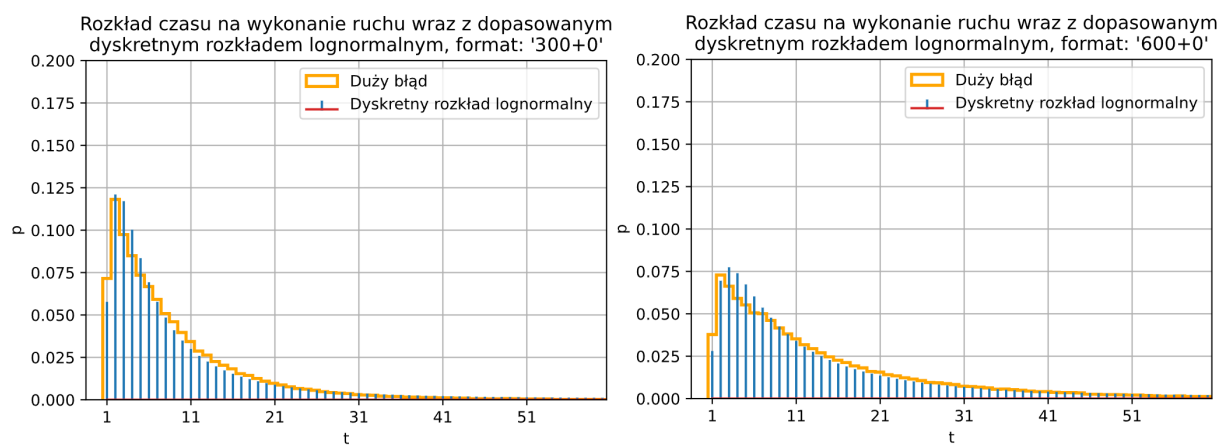
Wyniki obserwacji przedstawione są na Rysunkach 3.15 oraz 3.16. Na pierwszym z nich, zawierającym zestaw wszystkich ruchów, dyskretny rozkład normalny jest dobrze dopasowany w obydwu formatach czasowych. Także w przypadku zestawu ruchów błędnych (rys. 3.16) dopasowanie poprawiło się w znacznym stopniu względem rozkładu ciągłego. Wykonany został test Chi-kwadrat badający zgodność częstości występujących w danych z tymi zadanymi przez rozkład. Dla losowych zestawów danych liczących 1000 próbek i po wykonaniu 100 prób Monte-Carlo, średnia p-wartość dla każdego z badanych zbiorów na poziomie istotności  $\alpha = 0,05$  wskazuje na przyjęcie hipotezy zerowej mówiącej o pochodzeniu danych z dyskretnego rozkładu log-normalnego. Szczegółowe wyniki, wraz z wyznaczonymi wcześniej za pomocą metody największej wiarygodności parametrami rozkładu umieszczone są w tabeli 3.3. Na podstawie średniej p-wartości da się zauważyć, że pomimo zgodności dopasowania do każdego ze zbiorów, zestaw wszystkich ruchów jest lepiej opisany przez zadany rozkład teoretyczny, niż zestaw ruchów błędnych. Potwierdzają to odchylenia w okolicy ruchów z indeksem od 3 do 8 na Rysunku 3.16, które są zauważalnie większe, niż odchylenia w zbiorze wszystkich ruchów na Rysunku 3.15.



Rysunek 3.15: Rozkład czasu poświęconego na wykonanie posunięcia dla formatów „600+0” i „300+0” z dopasowanym dyskretnym rozkładem log-normalnym.

Tabela 3.3: Średnie wyniki testu zgodności Chi-kwadrat, średnia ze 100 prób Monte-Carlo dla losowych próbek o długości 1000. Poziom istotności  $\alpha = 0,05$ .

dane	parametry rozkładu	p-wartość	wynik
„300+0”, wszystkie ruchy	$\mu = 1,220, \sigma = 1,091$	22,28%	przyjęcie hipotezy
„600+0”, wszystkie ruchy	$\mu = 1,705, \sigma = 1,188$	20,40%	przyjęcie hipotezy
„300+0”, ruchy błędne	$\mu = 1,673, \sigma = 1,059$	7,14%	przyjęcie hipotezy
„600+0”, ruchy błędne	$\mu = 2,158, \sigma = 1,123$	6,49%	przyjęcie hipotezy



Rysunek 3.16: Rozkład czasu poświęconego na wykonanie posunięcia dla formatów „600+0” i „300+0” z dopasowanym dyskretnym rozkładem log-normalnym. Zestaw ruchów błędnych.

# Rozdział 4

## Podsumowanie

Na podstawie bazy danych ze strony *Lichess.com* dotyczących 17 524 012 posunięć ze 275 994 gier rozgrywanych na każdym poziomie rankingowym w formatach „300+0” oraz „600+0” zbadane zostało kilka statystyk i zależności związanych z szachami szybkimi. Pokróćce opisane zostaną teraz najważniejsze analizy, a także wnioski z nich płynące.

Po wstępnym przygotowaniu danych i wyznaczeniu pierwszych statystyk, w pracy został określony empiryczny rozkład długości gry. Niecałe 5% gier z rozpatrywanych formatów czasowych zawiera więcej niż 57 ruchów.

Przy badaniu zależności pomiędzy indeksem ruchu, a czasem poświęconym na posunięcie określony został średni czas na wykonanie posunięcia w zależności od jego indeksu. W obydwu formatach dla każdego numeru ruchu błędne posunięcia zajmowały średnio więcej czasu porównując z zestawem wszystkich ruchów, a także miały większą wariancję. Zaproponowana została przyczyna takiego stanu, mianowicie zarówno szansa na popełnienie błędu jak i czas na wykonanie posunięcia mogą być silnie zależne od stopnia skomplikowania pozycji, a wśród zestawu wszystkich ruchów średnia może być zaniżana przez ruchy oczywiste i forsowne (tylko jeden możliwy legalny ruch), które zajmują statystycznie mniej czasu. Analiza zależności pomiędzy numerem ruchu, a średnim poświęconym czasem została również przeprowadzona z podziałem na poszczególne grupy rankingowe. Dla każdego analizowanego formatu czasowego gracze z wyższym rankingiem mają tendencję do szybszego wykonywania początkowych posunięć, będących zazwyczaj elementem teorii otwarć szachowych, dla których szansa na popełnienie błędu jest najmniejsza. Zauważalnie dłużej natomiast wykonują posunięcia z indeksem od 20 do 30 - czyli te, gdzie prawdopodobieństwo popełnienia błędu jest największe, a nieznacznie dłużej ruchy z indeksem od 30. Dla zestawu ruchów błędnych, wyżej notowani gracze poświęcają na posunięcie zdecydowanie więcej czasu. Czas wydzielony na ruchy z największym prawdopodobieństwem na popełnienie błędu sprawia, że różnica w częstotliwości popełniania błędów pomiędzy lepszymi, a gorszymi graczami dla gry środkowej jest bardzo duża. Przy posunięciach z indeksem powyżej 40, szansa na pomyłkę jest podobna dla każdego rankingu. 25% Najslabszych graczy ma tendencję do popełniania nawet kilkukrotnie większej liczby błędów w ruchach z indeksem poniżej 10, niż ich wyżej notowani rywale. Nasuwa to wniosek, że nauka otwarć szachowych (wiążąca się z mniejszą liczbą błędów w początkowej fazie gry) może pozwolić początkującym zawodnikom na skuteczniejsze zdobycie punktów rankingowych.

Badanie zależności pomiędzy formatem „300+0” i „600+0” pokazało kilka ciekawych cech. W formacie „600+0” gracze mają do dyspozycji dokładnie dwa razy więcej czasu, niż w dla „300+0”, jednakże czas poświęcony na konkretny ruch nie jest prawie nigdy

dwukrotnie większy. Dla zbioru wszystkich ruchów w początkowej fazie gry posunięcia w dłuższym formacie trwają średnio około 1,65 raza dłużej z tendencją rosnącą liniowo wraz ze wzrostem indeksu, aż do okolic 1,8 raza dłużej przy ruchach z indeksem powyżej 50. Natomiast stosunek czasu dla ruchów oznaczonych przez silnik jako błędne zachowuje podobne wartości do zestawu wszystkich ruchów w początkowej fazie gry, w przypadku gry końcowej jest jednak znacznie większy i dochodzi 2. Wiąże się to z większą średnią pozostałą procentową ilością czasu w formacie „600+0”, dzięki której zawodnik może pozwolić sobie na dłuższe zastanowienie się w późniejszej fazie meczu.

Prawdopodobieństwo popełnienia błędu w konkretnym ruchu jest bardzo zbliżone w obydwu analizowanych formatach w początkowej fazie gry. Od ruchu 20, szansa na błąd jest nieznacznie większa w krótszym formacie. Różnica prawdopodobieństwa na wykonanie choć jednego błędu w pierwszych 10, 20 lub 30 ruchach, przy założeniu niezależności ruchów jest mała. Mimo większej ilości czasu w formacie „600+0”, błędów w początkowej fazie gry pojawia się statystycznie nieznacznie więcej niż w krótszym formacie. Znacząca różnica jest natomiast widoczna pomiędzy poszczególnymi przedziałami rankingowymi. 25% najgorzej radzących sobie graczy ma ponad 40% szansy na błąd w pierwszych 10 ruchach, natomiast w przypadku graczy z rankingiem plasującym się pomiędzy 25, a 50 centylem jest to o prawie połowę mniej.

Ostatnią analizowaną zależnością jest rozkład czasu poświęconego na wykonanie posunięcia, osobno dla zbioru wszystkich ruchów i dla oznaczonych przez silnik jako błędne. Dla obu formatów najwięcej posunięć zajmuje od jednej do pięciu sekund, by stopniowo maleć dla większej ilości czasu. Dominanta dla wszystkich zestawów wynosi 2, jednakże dla ruchów błędnych, histogram częstotliwości wystąpienia jest bardziej spłaszczony. Rozkład danych dla każdego zestawu jest wizualnie dobrze opisany przez rozkład log-normalny, a jeszcze lepiej - w związku z tym, że dane są dyskretne, przez dyskretny rozkład log-normalny z parametrami estymowanymi metodą największej wiarygodności. Potwierdza to test zgodności Chi-kwadrat.

Powyższe analizy i rozważania mogą zostać uzupełnione w dalszej pracy o inne formaty czasowe, ze szczególnym uwzględnieniem tych z czasem dodawanym po posunięciu. Dodatkowo, można wprowadzić miarę skomplikowania pozycji i uwzględnienia jej wpływu na wyniki analiz, wymaga to jednak dużej mocy obliczeniowej i dokładnego zbadania poszczególnych posunięć w każdej z analizowanych partii.

# Bibliografia

- [1] “Repozytorium zawierające kod źródłowy silnika Stockfish.” <https://github.com/official-stockfish/Stockfishr>. Dostęp: 11/10/2021.
- [2] “Lichess database.” <https://database.lichess.org/>. Accessed: 30/05/2021.
- [3] “Does the low time alarm make people play worse?.” <https://github.com/Antiochian/chess-blunders/blob/main/README.md>. Dostęp: 22/12/2021.
- [4] A. E. Elo, *The rating of chessplayers, Past & Present (second edition)*. New York, United States: FIDE, 1986.
- [5] M. Glickman, “The glicko system,” *Boston University*, 1999.
- [6] “Lichess computer engine used from 2014.” <https://lichess.org/blog/U4mtoEQAAEEAgZRL/strongest-chess-player-ever>. Dostęp: 30/05/2021.
- [7] *chesscenter.net*. C&O Family Chess Center, “Chess analysis symbols.” <https://web.archive.org/web/20180116135828/http://www.shatranj.us/files/AnalysisSymbols.pdf>, 2014. Dostęp: 17/10/2021.
- [8] F. Dekking, C. Kraaikamp, H. Lopuhaä, and L. Meester, *A Modern Introduction to Probability and Statistics: Understanding Why and How*. Springer Texts in Statistics, Springer, 2005.
- [9] J. Koronacki, J. Mielniczuk, *Statystyka dla studentów kierunków technicznych i przyrodniczych*. Warszawa: Wydawnictwa Naukowo-Techniczne, 2001.
- [10] K. O. Bowman and L. R. Shenton, “Estimator: Method of moments,” *Encyclopedia of statistical sciences*, 1998.
- [11] *mathworld.wolfram.com*. Weisstein, Eric W., “Log normal distribution.” <https://mathworld.wolfram.com/LogNormalDistribution.html>, 2020. Dostęp: 17/10/2021.
- [12] D. H. Holding, “The evaluation of chess positions,” *Simulation & Gaming*, vol. 10, pp. 207 – 221, 1979.
- [13] J. Jakubowski, R. Sztencel, *Wstęp do teorii prawdopodobieństwa*. Warszawa: Script, 2001.
- [14] J. Lyu and S. Nadarajah, “Discrete lognormal distributions with application to insurance data,” *Springer*, 7 2021.