

Analiza gier Magnusa Carlsena na platformie lichess

405826, Maksymilian Sulima, środa 12:50

AGH, Wydział Informatyki Elektroniki i Telekomunikacji

Rachunek prawdopodobieństwa i statystyka 2021/2022

Kraków, 28 stycznia 2021

Ja, niżej podpisany własnoręcznym podpisem deklaruję, że przygotowałem przedstawiony do oceny projekt samodzielnie i żadna jego część nie jest kopią pracy innej osoby.

Maksymilian Sulima

1. Streszczenie raportu

Raport powstał w oparciu o analizę zapisów partii mistrza świata Magnusa Carlsena.

2. Opis danych

Dane do projektu pochodzą ze strony

<https://www.kaggle.com/zq1200/magnus-carlsen-lichess-games-dataset>.

Baza zawiera 2 foldery: csv i pgn. W folderze csv znajdują się 4 pliki .csv (dokładny opis poszczególnych kolumn jest niemożliwy ze względu na mnogość danych) :

- carlsen_games, zawiera on podstawowe informacje o partiach min. id_gry, format czasowy, nicki graczy, ELO graczy, rezultat partii, datę rozegrania partii
- carlsen_games_moves, zawiera spis wszystkich ruchów we wszystkich partiach
- eco_codes, zawiera informacje o użytych debituach
- user_names, zawiera 2 kolumny: user_name i real_name, które dopasowują nick na lichess do rzeczywistych imion graczy.

W folderze pgn znajdują się zapisy partii mistrza w formacie PGN dla jego poszczególnych nicków (miał ich co najmniej 10 na przestrzeni pomiaru danych). Podane dane nie posiadają nieścisłości i nie wymagają preprocessingu.

3. Opis danych

3.1 Wydobywanie podstawowych informacji z danych

Ze względu na strukturę danych wyliczenie tych parametrów wymaga sporo linii kodu dlatego po dokładny obraz procesu wyciągania danych zapraszam do lektury skryptu w R.

Liczba zarejestrowanych partii: 13 442

Liczba wygranych: 9 129

Procent wygranych: 67%

Liczba remisów: 1 068

Liczba przegranych: 3 245

Średnie ELO Carlsena: 3 058.7

Przedział ufności ELO Carlsena dla $\alpha = 0.1\%$: $< 3052.383, 3065.016 >$

Wariancja i odchylenie standardowe ELO Carlsena: 49 527.65, 222.5481

Średnie elo przeciwnika Carlsena: 2867.554

Wariancja ELO przeciwnika Carlsena: 80 117.04, 283.0495

Średnie elo przeciwnika w wygranych partiach: 2 840.885

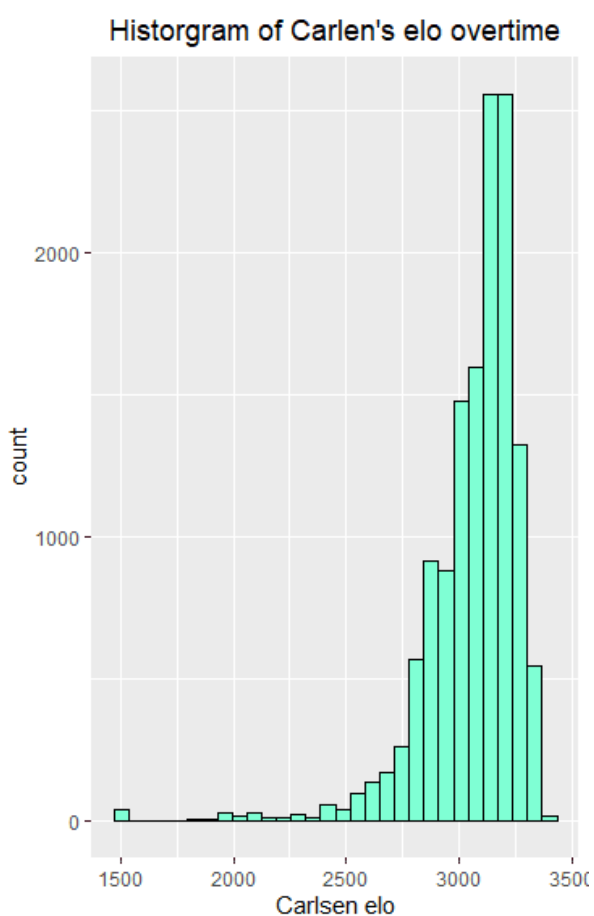
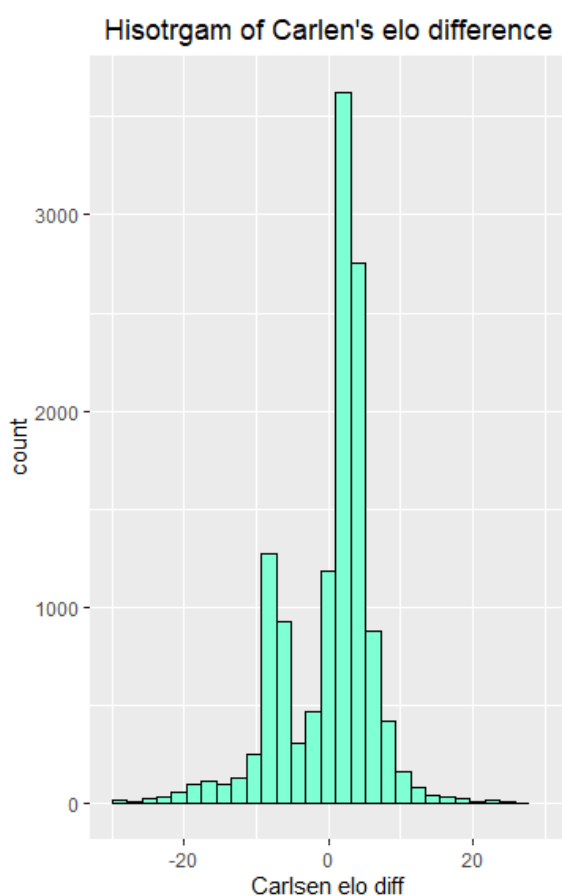
Średnie elo przeciwnika w partiach zakończonych remisem: 2 920.167

Średnie elo przeciwnika w partiach przegranych: 2 925.265

Korelacja między ELO przeciwnika, a rezultatem partii: -0.1326042

Średnia różnica ELO między Carlsenem a przeciwnikiem: 191.1457

3.2 Carlsen na wykresach



Współczynnik symetrii wartości ELO: -2.664667

Kurtoza wartości ELO: 15.17165

Analiza hipotezy: rozkłady zmiennych zmiany ELO oraz wartości ELO są rozkładami normalnymi. W celu zbadania hipotezy stworzono wykresy Q-Q ww. zmiennych oraz wykonano test normalności Shapiro-Wilka, ponieważ test na raz przyjmuje maksymalnie 5000 wartości wykonano 10 testów z których wyciągnięto średnie.

Wyniki testów dla rozkładu wartości ELO:

$$W = \{0.80161, 0.78163, 0.79725, 0.77146, 0.78759\}$$

$$p = \{2.2 \cdot 10^{-16}, 2.2 \cdot 10^{-16}, 2.2 \cdot 10^{-16}, 2.2 \cdot 10^{-16}, 2.2 \cdot 10^{-16}\}$$

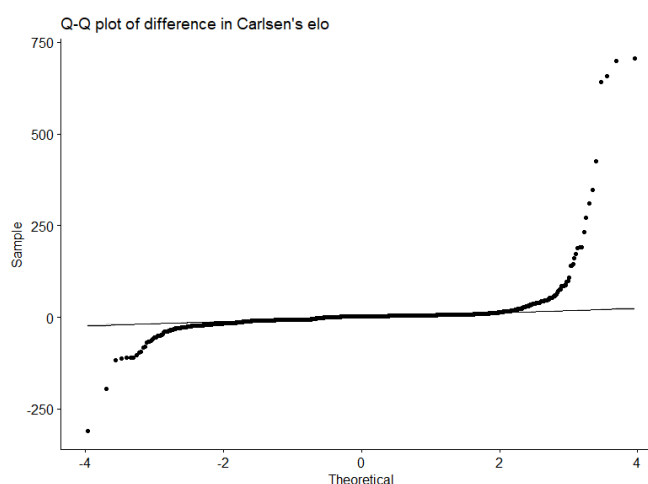
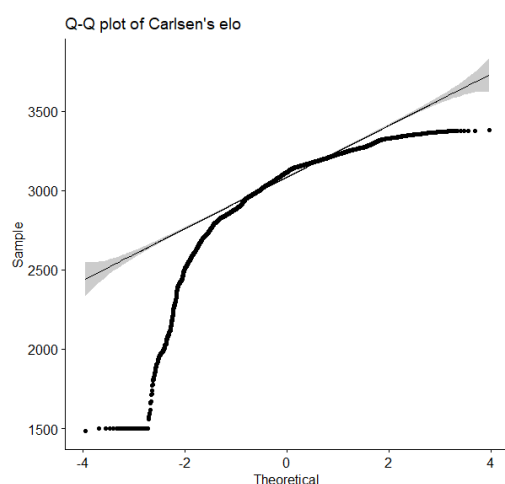
Wynik testu nie daje podstaw do odrzucenia hipotezy jeżeli wartość $p > 5\%$. Natomiast otrzymano wartości, które są nieporównanie mniejsze. Zatem można odrzucić hipotezę o rozkładzie normalnym zmiennej wartości ELO.

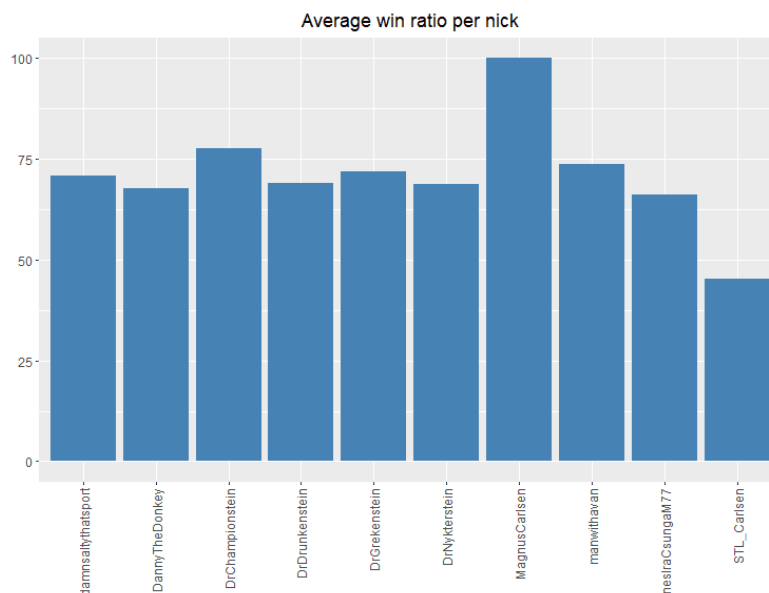
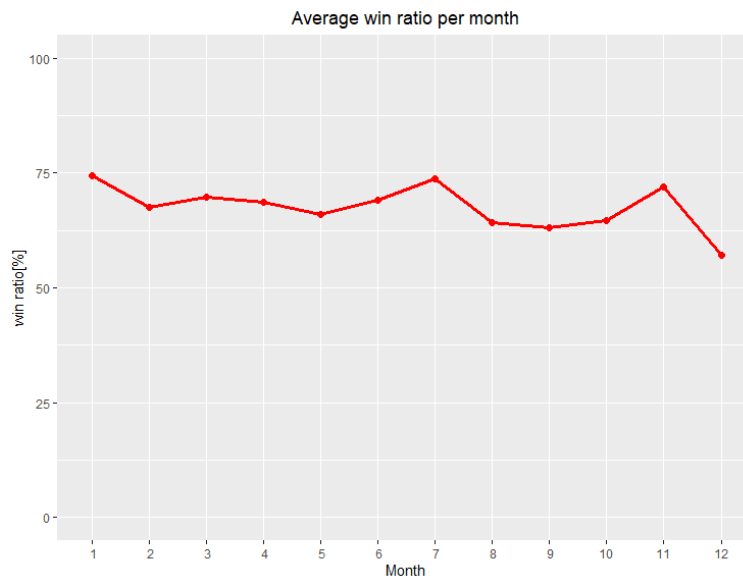
Wyniki testów dla rozkładu zmian ELO:

$$W = \{0.27506, 0.24222, 0.25567, 0.28048, 0.53423\}$$

$$p = \{2.2 \cdot 10^{-16}, 2.2 \cdot 10^{-16}, 2.2 \cdot 10^{-16}, 2.2 \cdot 10^{-16}, 2.2 \cdot 10^{-16}\}$$

Ponownie otrzymano wartości p znacząco mniejsze od 5% zatem odrzucono hipotezę o rozkładzie normalnym zmiennej zmiany ELO.





4. Wnioski

- Magnus Carlsen jest jednym z największych jeżeli nie największym szachistą współcześnie i być może nawet w historii
- Rozkład wartości ELO Magnusa na portalu lichess.org nie jest rozkładem normalnym
- Rozkład zmian ELO magnusa na ww. portalu również nie jest rozkładem normalnym
- Rozkład wartości ELO jest rozkładem lewostronnie skośnym i wysmukłym
- Magnus jest najlepszym graczem na platformie lichess
- ELO przeciwnika jest skorelowane z szansą na wygraną partii (im mniejsze ELO tym większa szansa na wygraną)
- Magnus powinien wrócić do nicu DrChampionstein ponieważ mają ten nick wygrywał najwięcej partii (poza Magnus Carlsen ale pod tym alisem rozegrał tylko 1 partię)
- Carlsen gra gorzej w grudniu