Homework 2

I downloaded a dataset containing chess games with columns:

- Game ID;
- Rated (T/F);
- Start Time;
- End Time;
- Number of Turns;
- Game Status;
- Winner;
- Time Increment;
- White Player ID;
- White Player Rating;
- Black Player ID;
- Black Player Rating;
- All Moves in Standard Chess Notation;
- Opening Eco
- Opening Name;
- Opening Ply (Number of moves in the opening phase)

I dropped the unwanted (most of them are NaN values) columns and manipulated column "winner" with the rule, if white won the value is 0, if black won the value is 1 and 2 for draw. I created two new columns using existing data, which are:

- Point difference (White Points – Black Points)
- Estimated White Points regarding performance (for more details and mathematical background of how Elo rating system works, check https://en.wikipedia.org/wiki/Elo_rating_system )

  If Player A has a rating of $R_A$ and Player B a rating of $R_B$, the exact formula (using the logistic curve)[11] for the expected score of Player A is

  $$E_A = \frac{1}{1 + 10^{(R_B - R_A)/400}}.$$

  Similarly the expected score for Player B is

  $$E_B = \frac{1}{1 + 10^{(R_A - R_B)/400}}.$$

Now, remaining columns are:

1. Turns (Number of turns played)
2. Point difference (White Points – Black Points)
3. Estimated White Points (in terms of Elo and performance)

```
(base) csoylu@Cagatay-MacBook-Pro qmbu-hw2 % /Users/csoylu/opt/anaconda3/bin/python /Users/csoylu/Workspace/qmbu-hw2/main.py
Result table:
   lower_bound_for_estimates  estimates  upper_bound_for_etimates  standard_errors
0                   0.002630   0.002866                  0.003103         0.000118
1                  -0.001383  -0.001351                 -0.001319         0.000016
2                   0.697773   0.731146                  0.764519         0.016687
```

By looking at the table above, we can see that:

1. Number of turns played is positively correlated with winner value. Beta value is 0.002866
2. Point difference is negatively correlated with winner value. Beta value is -0.001351

3. Estimated White Points is positively correlated with winner value. Beta value is 0.731146

---