Reviewers: Rafał Guzek, Jaromir Węgrzyn

1. Problem formulation [5 / 5 pts]:

- **Is the problem clearly stated [1 / 1 pt]**
  The problem has been clearly defined, and some of the specialized terms related to the issue under study have also been explained.
- **What is the point of creating model, are potential use cases defined [1 / 1 pt]**
  The purpose of this project is to analyze the impact of different chess openings on the number of errors made during the game. Many use cases are included in the project, for example: chess training and coaching, opponent analysis etc.
- **Where do data comes from, what does it containt [1 / 1 pt]**
  Dataset comes from the Kaggle website and it's called „Lichess Chess Game Dataset". The dataset contains diverse information about played games, such as moves made by players, the time it takes to make a move, the ranking of players, etc.
- **DAG has been drawn [1 / 1 pt]**
  Yes.
- **Confoundings (pipe, fork, collider) were described [1 / 1 pt]**
  All possible confoundings are listed and described.

2. Data preprocessing [2 / 2 pts]:

- **Is preprocessing step clearly described [1 / 1 pt]**
- **Reasoning and types of actions taken on the dataset have been described [1 / 1 pt]**
  The entire pre-processing operation has been described in detail step by step, and all actions regarding data selection have also been justified.

3. Model [4 / 4 pts]

- **Are two different models specified [1 / 1 pt]**
- **Are difference between two models explained [1 / 1 pt]**
- **Is the difference in the models justified (e.g. does adding aditional parameter makes sense?) [1 / 1 pt]**
- **Are models sufficiently described (what are formulas, what are parameters, what data are required) [1 / 1 pt]**
  The project presents two models that use the same Poisson distribution, although they use different numbers of predictors. The selection of an additional predictor for the second model makes sense and is intended to improve the precision of the model. The formulas used to create the model are provided and described.

4. Priors [4 / 4 pts]

- **Is it explained why particular priors for parameters were selected [1 / 1 pt]**
  Yes. The explanation of the chosen priors parameters was precise.
- **Have prior predictive checks been done for parameters (are parameters simulated from priors make sense) [1 / 1 pt]**
  Prior predictive checks have been created. Histograms of simulated parameters from the prior confirm obtaining the expected values.
- **Have prior predictive checks been done for measurements (are measurements simulated from priors make sense) [1 / 1 pt]**

Yes, the simulated measurements make sense and have been visualized in the form of a histograms.

- **How prior parameters were selected [1 / 1 pt]**
  As prior parameters, those values were chosen for which the model results were most similar to the real data.

5. Posterior analysis (model 1) [3 / 4 pts]

- **Were there any issues with the sampling? if there were what kind of ideas for mitigation were used [1 pt]**
  No problems occurred during sampling.
- **Are the samples from posterior predictive distribution analyzed [0.5 / 1 pt]**
  The results of posterior predictive distribution were generally considered.
- **Are the data consistent with posterior predictive samples and is it sufficiently commented (if they are not then is the justification provided) [1 / 1 pt]**
  Yes, the data is consistent with the posterior predictive samples. The data do not overlap completely, but they are quite similar to each other. The reason why the charts may differ has also been considered.
- **Have parameter marginal disrtibutions been analyzed (histograms of individual parametes plus summaries, are they diffuse or concentrated, what can we say about values). [0.5 / 1 pt]**
  Histograms of marginal distributions were analyzed in general. More detailed analysis was lacking. Parameter values were concentrated and quite similar to those in the earlier analysis.

6. Posterior analysis (model 2) [3 / 4 pts]

- **Were there any issues with the sampling? if there were what kind of ideas for mitigation were used [1 / 1 pt]**
  No problems occurred during sampling.
- **Are the samples from posterior predictive distribution analyzed [0.5 / 1 pt]**
  The results of posterior predictive distribution were generally considered.
- **Are the data consistent with posterior predictive samples and is it sufficiently commented (if they are not then is the justification provided)**
  Yes, the data is consistent with the posterior predictive samples. Moreover, the fact of the high similarity of the data was also pointed out.
- **Have parameter marginal disrtibutions been analyzed (histograms of individual parametes plus summaries, are they diffuse or concentrated, what can we say about values) [0.5 / 1 pt]**
  Histograms of marginal distributions were briefly analyzed. More detailed analysis was missing. Parameter values were concentrated and quite similar to those in the earlier analysis.

7. Model comaprison [4 / 4 pts]

- **Have models been compared using information criteria [1 / 1 pt]**
  Yes.
- **Have result for WAIC been discussed (is there a clear winner, or is there an overlap, were there any warnings) [1 / 1 pt]**
  The results obtained for the WAIC comparison have been clearly considered. A description of all variables returned by the az.compare function has been added. Based on these, it has been clearly considered that the second model is superior. No warnings appeared for the WAIC comparison.

- **Have result for PSIS-LOO been discussed (is there a clear winner, or is there an overlap, were there any warnings) [1 / 1 pt]**
  The results obtained for the LOO comparison have been clearly considered. A description of all variables returned by the az.compare function has been added. Based on these, it has been clearly considered that the second model is superior. For the LOO comparison, a warning appeared for the first model.
- **Whas the model comparison discussed? Do authors agree with information criteria? Why in your opinion one model better than another [1 / 1 pt]**
  Yes, the model comparison has been discussed. The authors agree with the results obtained by using presented information criteria. According to the assumptions, due to the fact that the second model used more predictors it was characterized by a better fit and higher accuracy.

Total grade [25/27 pts] – Percentage: 92,6%