Modelling the Dunning Kruger effect

Introduction

It is observed that poor performers tend to be less calibrated in their ability to judge their performance after the task in comparison to the high performers. Commonly known as Dunning-Kruger effect, a cognitive bias in which people with low capability overestimate their abilities and similarly, high performers tend to underestimate the skills.
For example, if in a panel discussion of cognitive science, a new student, a phd student, and a professor is sitting, the new student will be more probable to be more confident than the phd student in front of the professor because he the new student thinks whatever he knows is sufficient.

Model the Dunning Kruger effect

Generally, the self assessment model involve confidence and error detection to measure self assessment. This model's parameters are 'sensitivity' and 'bias', where sensitivity is their ability to discriminate between correct and incorrect performance, and bias is how confident the participant is.

The self assessment is mainly dependent on 3 factors;

1. Perceived Ability (related to belief on own skills)
2. Discrimination Ability (related to technical knowledge/ subject knowledge)
3. Difficulty of the task

In our model, we introduced 3 parameters, the perceived ability of person p (θp) difficulty of an item i (βi) and error of participant ϵ.
Since researchers generally use, the one-parameter item response theory (IRT) model, known as Rasch model, to judge the capability of the student.
The IRT function, without including the error parameter is given by:

$$P(X_{p,i} = 1 | \theta_p, \beta_i) = \frac{1}{1 + e^{-(\theta_p - \beta_i)}} \qquad \text{...............1}$$

In the model , we include 2 factors: one is making error, and other is guessing factor, since the participants can also guess (task is QnA, otherwise no need)

$$P(X_{p,i} = 1|\theta_p, \beta_i; \epsilon) = (1-\epsilon)\cdot\left(g + \frac{1-g}{1+e^{-(\theta_p-\beta_i)}}\right) + \epsilon\cdot\left(1 - \left(g + \frac{1-g}{1+e^{-(\theta_p-\beta_i)}}\right)\right) \quad \text{…………2}$$

The probability of the person responding correctly is the probability of answering correctly and recognizing that one is correct plus the probability of answering incorrectly but erroneously believing one is correct. Also when one guesses and one not guesses, and the answer is correct. Usually g is given by 1/N (no. of options).

The model mostly uses the bayesian approach, i.e we assume priors (gaussian distribution, here) on difficulty on item βi and the perceived ability θp of the person.

The posterior distribution of θp and βi is given by:

$$P(\theta_p, \beta_i | X_{p,i} = 1) \propto P(X_{p,i} = 1 | \theta_p, \beta_i; \epsilon) \cdot P(\theta_p) \cdot P(\beta_i)$$

Then we marginalize over βi, and then transform the distribution to person's performance expectation and then calculated the probability being correct (put $\epsilon = 0$) using equation (2) on average question ($\beta i = 0$). This probability is multiplied by maximum score and the expectation is computed.

Now, its obvious that error won't be same for every participant, so we also used another approach, performance dependent estimation, where everything is same, but the error is measured as given below, where xi is correct response and n is maximum score, slope is $-\alpha$, intercept $\epsilon 0$ are parameters

$$\epsilon_p = \epsilon_0 - \alpha \cdot \frac{\sum_i x_i}{n},$$

Model methods

In the model, a study done on around 4000 participants giving MCQ questions on grammar with 5 questions. The maximum score they can attain is 20 and minimum 0. Then the participants were told to mark their own work with actual checking. The mean and standard deviation for actual score was 10.17 and 3.4 respectively, and for the estimated score was 12.49 and 3.91 respectively.

MCMC method were used to calculate the posterior distribution over beliefs about the ability. 3000 iterations were done but and 10% of iterations is burnout.

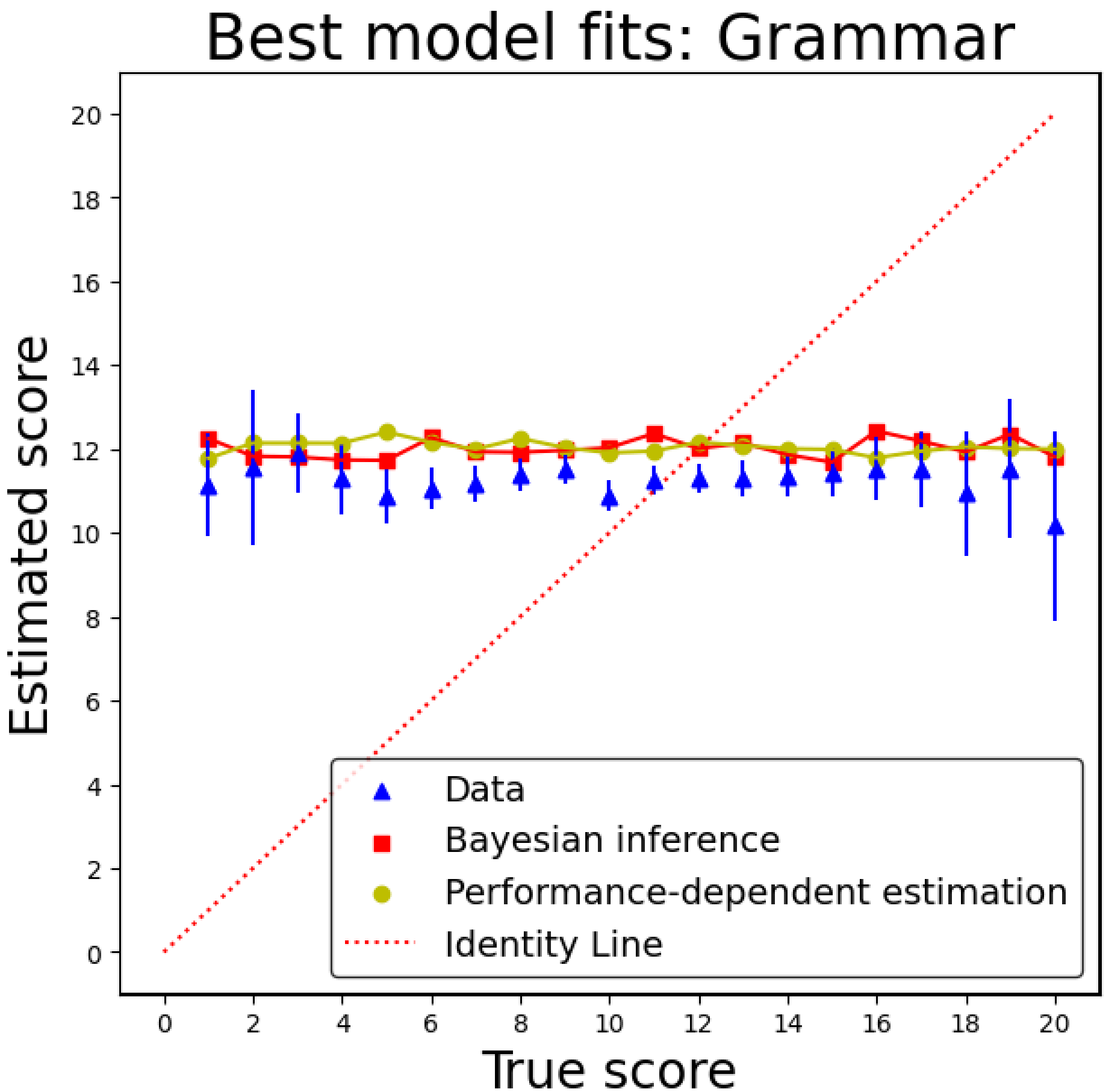
For both the models, we choose many different values of parameters and for each set of parameters, the predicted score was calculated and squared error was calculated with the actual guess score and set with minimum error was taken.

Link of the Google Drive:
https://drive.google.com/drive/folders/15pQ-f2DQ8Q91jcWYZxSdBb06geyTSGUM?usp=drive_link

The code was run on google colab, so please upload all the files on the google colab to run the output.

Conclusions and Discussions

The results are not appropriate because the no. of MCMC iterations were kept less due to limited computational powers and, similarly for the permutations of parameters. The BIC score was calculated and it showed the performance dependent model gives a better approach. The quadratic model can also be used to generate the better results for the error.



Best model fits: Grammar
Estimated score
True score
0
2
4
6
8
10
12
14
16
18
20
Data
Bayesian inference
Performance-dependent estimation
Identity Line