# Assignment 2

Márton Kardos

October 2022

## Simulation

I simulated 100 studies according to the simulation process described in the task.

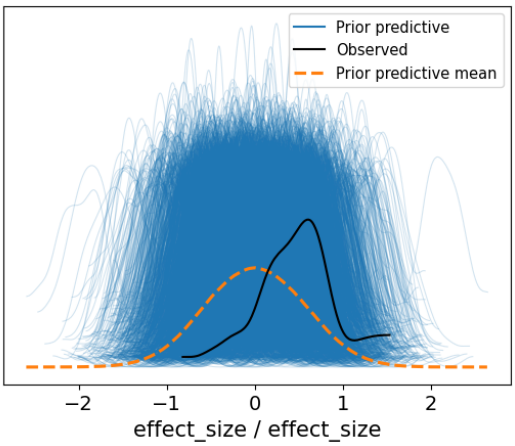
### Model

For analysing the data I used a probabilistic bayesian model, that assumes the effect size generation process to be the following:

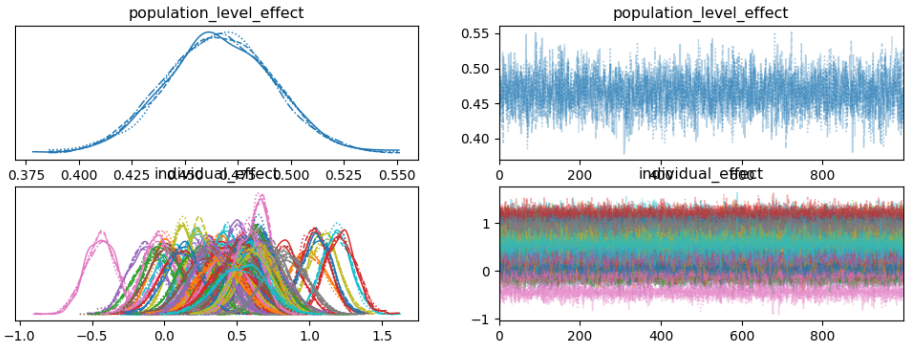
1. Draw a true underlying mean effect with prior: $\mu_{\beta_0} \sim \mathcal{N}(0, 0.5)$
2. For each study $i$ draw an individual effect size with prior: $\beta_{0_i} \sim \mathcal{N}(\mu_{\beta_0}, 0.2)$
3. Draw an observation of effect size: $y \sim \mathcal{N}(\beta_{0_i}, \sigma_i)$,
   where $\sigma_i$ is the standard error of the effect in the specific study

### Model fit

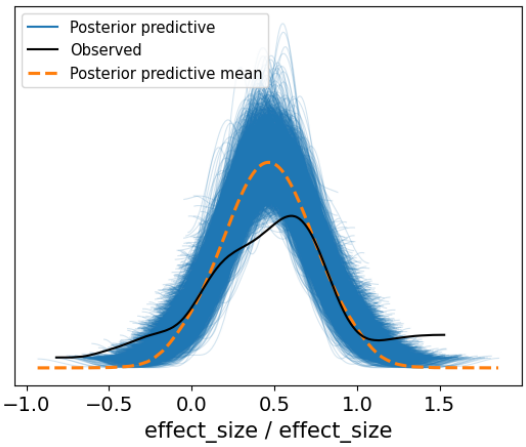
In order to assess whether the priors I set were realistic, I ran a prior predictive check:



The prior predictive definitely has a broader range than the actual values, but no unrealistically large or small values can be observed.
After sampling the posterior I inspected the trace plot to see whether it was hard for the sampler to explore the posterior.



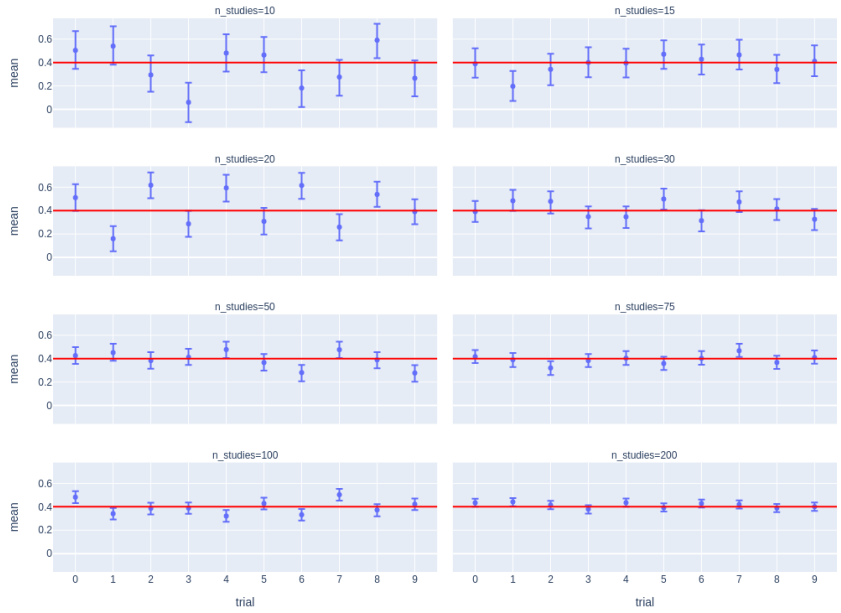

No divergences were encountered during sampling. The effective sample sizes were all above a thousand. To check whether the model captured the effects well, I had a look at the posterior predictive.




To me it seems that model captured the effects reasonably well, and the noise in the data is not as prominently observable, meaning that the model is relatively robust.

## Power analysis

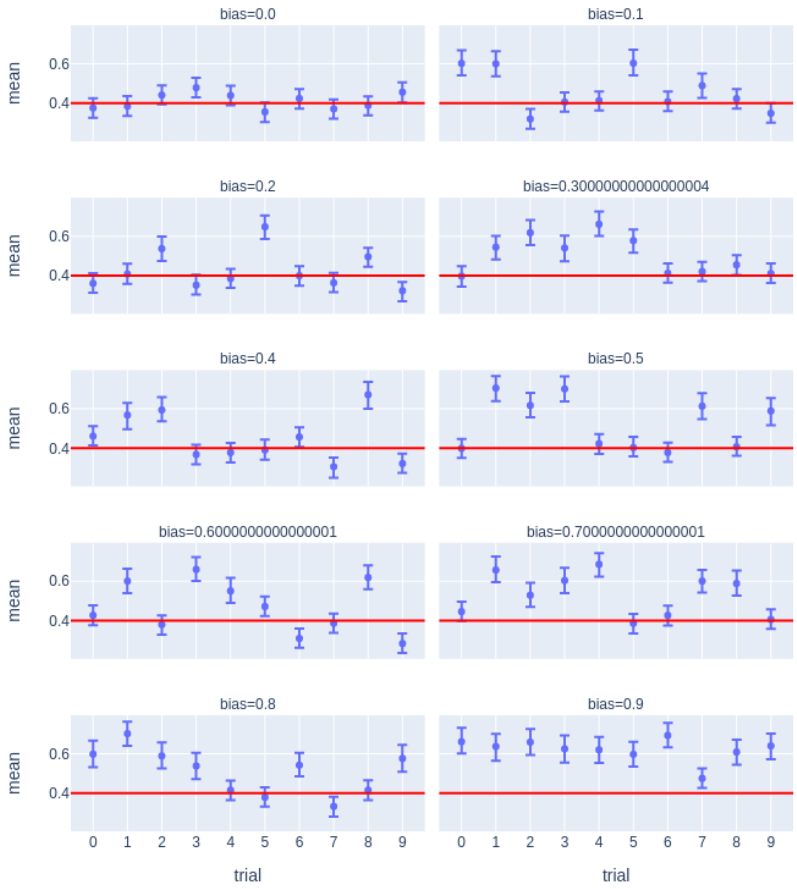
In order to assess how big of a sample I would need in order to capture the true underlying effect size, I generated simulated samples of different sizes. For each sample size I did 10 trials of sampling the model posterior on a freshly generated sample. I plotted the means of the posterior sample of the population level effect with error bars representing the 95% HDI estimates. I marked the true effect size with a horizontal red line.




Samples larger than or equal to 15 already yield significant results in the right direction, though the estimates only give a certain and stable approximation of the underlying effect size with a sample size of 30 or larger.

## Publication bias

In order to assess the influence of possible publication bias on my estimates I looped over different values of the publication bias $\pi$, where $1 - \pi$ represents the probability of a non-significant study being published from the simulated set of studies. All significant studies get published. I did 10 trials for each bias value, where each trial had a different simulated sample of sample size 200. I plotted the estimates in a similar fashion as with the power analysis. (excuse the floating point shenanigans)



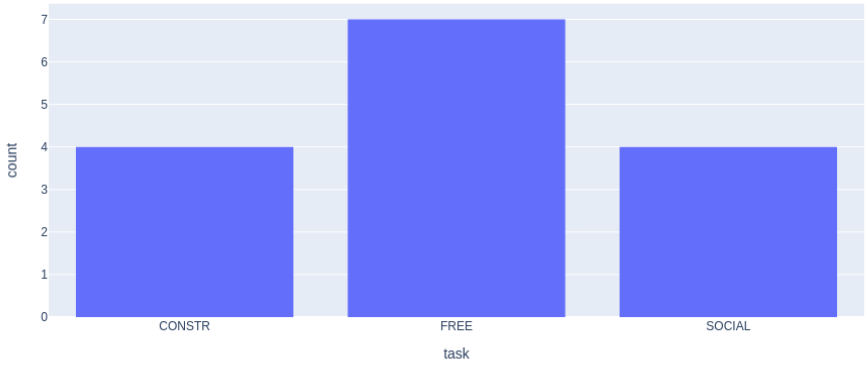

When $\pi <= 0.2$ the majority of posterior credible intervals still contain the true effect, thereby the estimates can be trusted. With bias values 0.4 and 0.5 around half of the estimates are accurate, and for all values of $\pi >= 0.6$ the effect size is systematically overestimated.

The empirical data should be examined for bias, as it can significantly influence the estimates of the underlying effect, the model produces.

# Analysis

## Empirical data

For the meta analysis a set of 57 studies were considered. Since we only examine the effect of schizophrenia in pitch variability, only those studies were kept, that had data on this, leaving us with a set of 15 studies. The remaining studies had a mean sample size of 77.4 ($SD$=38.74), which is way larger than expected, our simulation procedure had studies of smaller sample sizes.
The studies included different types of tasks, these being constrained, free conversation and conversation in a social context.

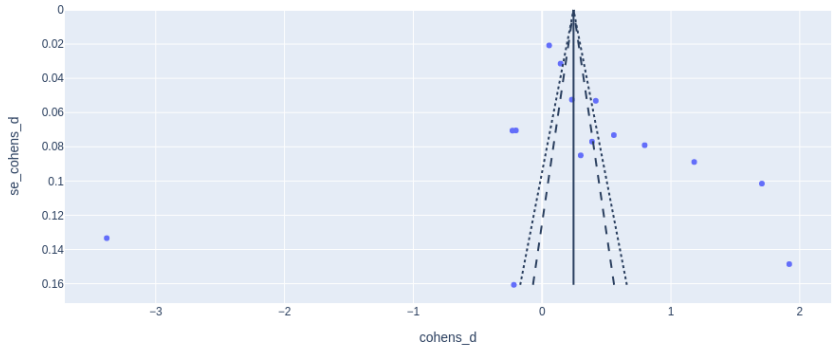



## Preprocessing

In order to standardize the effect sizes and errors of the different studies I calculated Cohen's D values for each study along with standard errors.

## Publication Bias

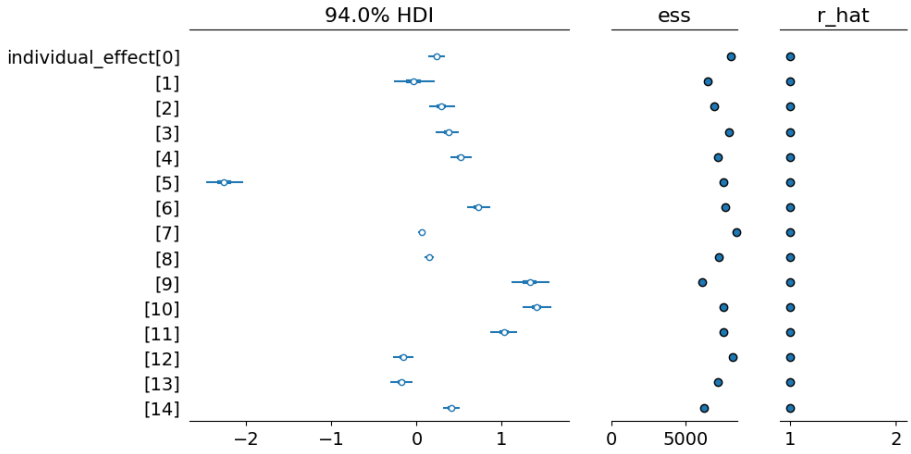
To assess whether the investigated studies suffer from publication bias I plotted the effect sizes and standard errors on a funnel plot.




It seems that the results of the studies are very heterogeneous, and heavily biased. It is also worth to note that due to the small sample size, there is a lot of data missing. This indicates, that we should be sceptical when interpreting the results of the meta-model.

## Results

Estimates of the model indicate that there is a significant positive effect in pitch variability, with a mean population level effect size of 0.257 ($SD$=0.056, $HDI\ 3\%$=0.154, $HDI\ 97\%$=0.363).
Estimates of study-level effects reflect, however, that the results are very heterogeneous.



One study in particular shows a very large effect in exactly the opposite direction.
Since schizophrenic individuals have reduced ability of emotional expression, it is quite perculiar that the meta-estimate suggests, that schizophrenic individuals would have higher variability in pitch of speech.

## Discussion

Due to the issues previously noted, I have to conclude, that the results of the meta-analysis cannot be trusted, and collection of further evidence would be needed to establish a firm connection between schizophrenia and pitch variability (if there is any at all).
Setting more restrictive priors on the meta-model could also have helped avoid such estimates, the model could only be conveyed by an overwhelming volume of evidence.