Computational Modeling – Assignment 4

Applying Meta-Analytic Priors

Caroline Casey

3/16/2018

Link to Github: https://github.com/Carolinecasey17/ComputationalModels.Portfolio4.git

In this assignment we reproduced the meta-analysis of pitch SD from last semester in a Bayesian framework using both a conservative and a meta-analytic prior. We assessed the difference in model quality and estimates using two priors.

We started reproducing the meta-analysis of pitch SD from previous studies of voice in schizophrenia and calculated a meta-analytic effect size using the package bromance, and the function brms, which automatically picks the best prior.

The ouput gave an Estimate of -0.54, with an Estimated error of 0.24, which we used later as our meta-analytic prior.

Then, we prepared the pitch SD data from last semester by simplifying the dataset by gropuing the participant ID across trials and taking the mean. Furthermore, we standardized the dataset to make it compatible with our meta-analytic prior.

Next, we made a regression model predicting pitch SD from Diagnosis. We can see from our density plot (*Figure 2*), illustrating the likelihood distribution, that both distributions have long tails. However, the distribution for the schizophrenia condition (displayed as the red line in the density plot), has a taller peak that occurs at slightly lower values than the non-schizophrenic condition. Our alpha and beta are normally distributed, we conditioned these on Diagnosis only.

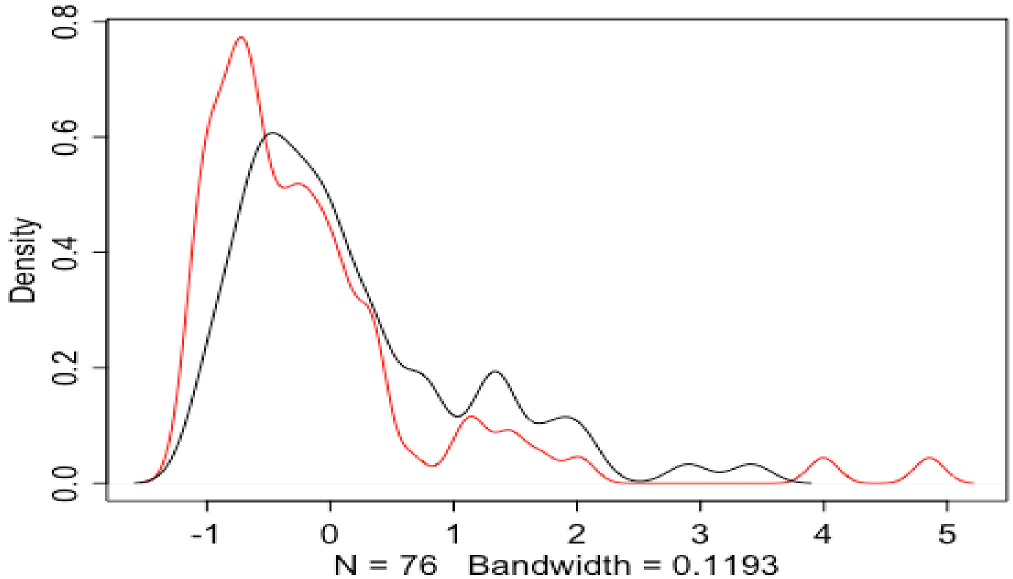


Figure 2 - plot of the likelihood distribution

We decided to use a skeptical prior in our model to avoid overfitting, as a false positive in the diagnosis of schizophrenia has serious consequences for the participant and clinical practice. To see the effect of a conservative prior, we made 4 models to test the skeptical prior: One with sd = 1 (loose prior), one with sd = 0.5 (moderate prior), one with sd = 0.25 (moderately conservative prior), and one with sd = 1 (very conservative prior) with a mean = 0 for all. This was to compare the differences in the posterior distribution.

Model	**Mean**	**Standard Deviation**	**Confidence Interval** **5.5% - 94.5%**
Model 1 – prior = 1	a – 0.15 b - -0.30 sigma – 0.98	a – 0.11 b – 0.16 sigma – 0.06	a – (-0.03 ; 0.33) b – (-0.55 ; -0.04) sigma – (0.89 ; 1.08)
Model 2 – prior = 0.5	a – 0.14 b - -0.28	a – 0.11 b – 0.15	a – (-0.04 ; 0.31) b – (-0.52 ; -0.03)

	sigma – 0.99	sigma – 0.06	sigma – (0.89 ; 1.08)
Model 3 – prior = 0.25	a – 0.11 b - -0.22 sigma – 0.99	a – 0.10 b – 0.13 sigma – 0.06	a – (-0.06 ; 0.28) b – (-0.43 ; 0.00) sigma – (0.90; 1.08)
Model 4 – prior = 0.10	a – 0.04 b - -0.08 sigma – 0.99	a – 0.09 b – 0.09 sigma – 0.06	a – (-0.10 ; 0.19) b – (-0.22 ; 0.05) sigma – (0.90 ; 1.08)

To evaluate model quality, we made posterior predictive plots for the different priors. From the plot presented below, we can see that accurate prediction is not very good, as the two conditions overlap too much. We plotted the posteriors from sampled and actual data, and they don't look too good. The distributions of the actual data's posteriors are much longer tailed and lower, and they almost overlap entirely.

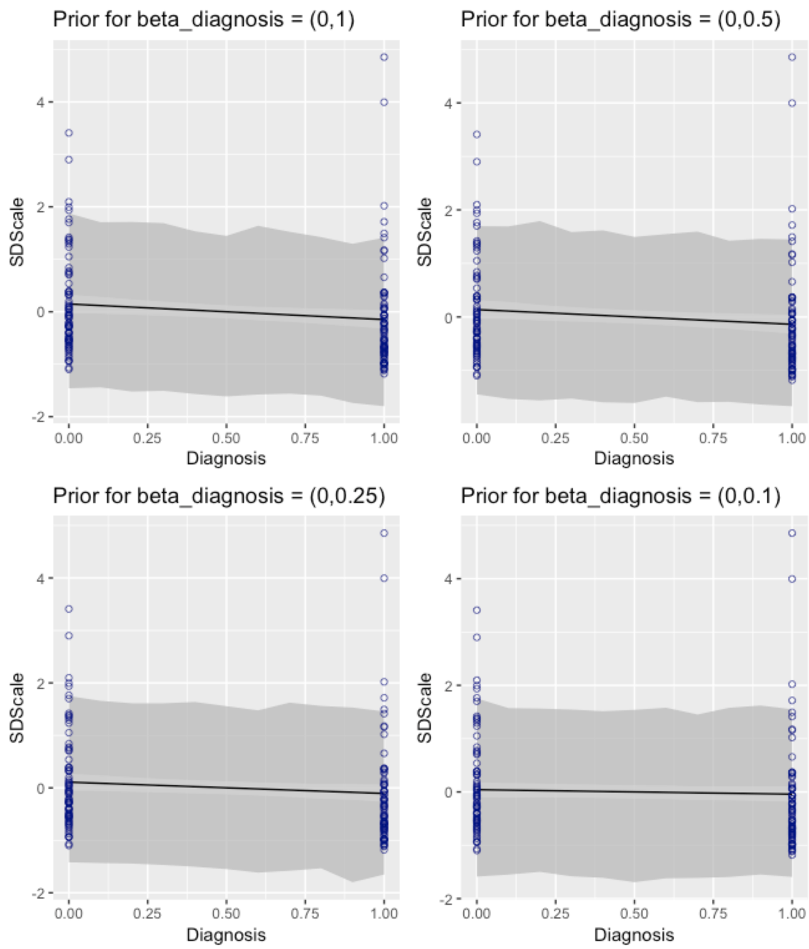


Figure 3 – Posterior predictive plots with uncertainty for 4 models with different priors

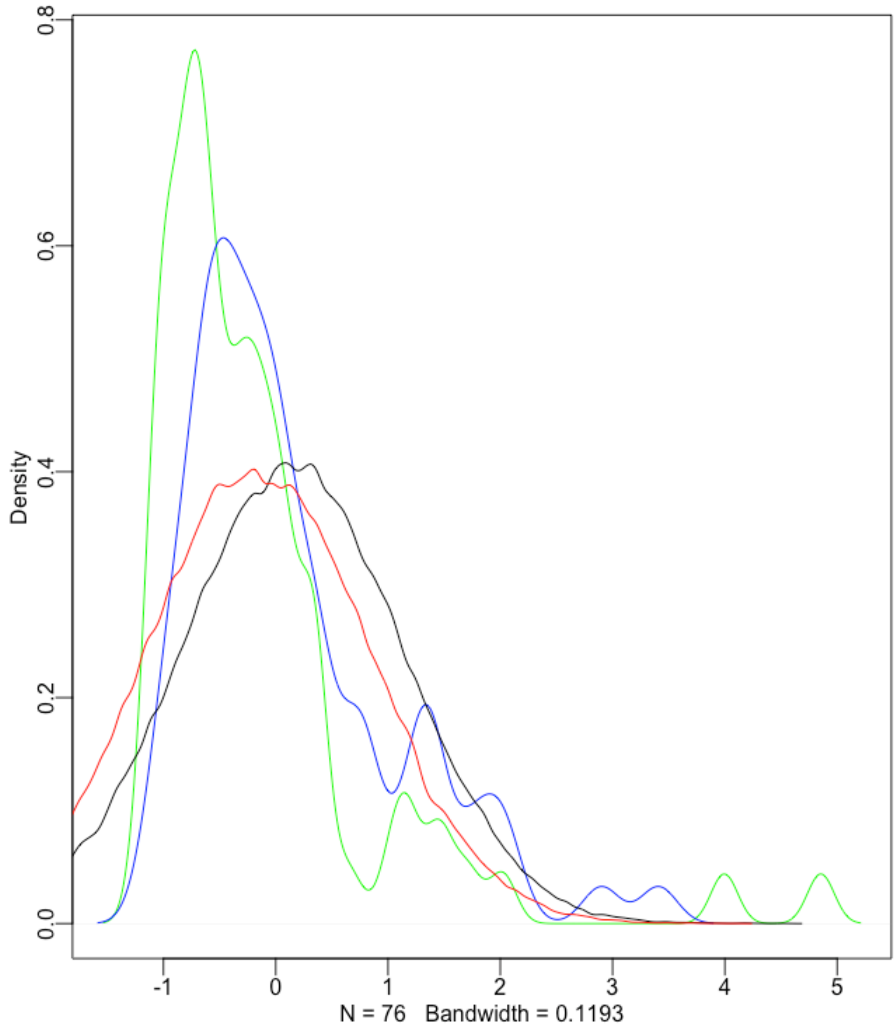


Figure 4 - Posterior predictive plots

The green and blue lines are the distributions of the posterior of the actual data, and the red and black lines are the predicted posterior distributions. We can see they don't match too well.

For the model we ran with a meta-analytic prior, we used a mean effect size of Diagnosis of -0.54, with a standard error of 0.24. We chose to use this, as the standard error is the area of uncertainty of the intercept of diagnosis, and we are interested in the true effect of this variable.

Model	**Mean**	**Standard Deviation**	**Confidence Interval** **5.5% - 94.5%**
Meta-analytic model-prior = (-0.54, 0.24)	a – 0.19 b - -0.38 sigma – 0.99	a – 0.10 b – 0.13 sigma – 0.06	a – (0.02 ; 0.35) b – (-0.59 ; -0.16) sigma – (0.89 ; 1.08)

When plotting the estimates of the models, we see that they are all very similar despite changing the priors. We can se that the meta-analytic prior (the plot lowest to the right) is a little more sloped than the others, and most similar to the model with a prior of 1.

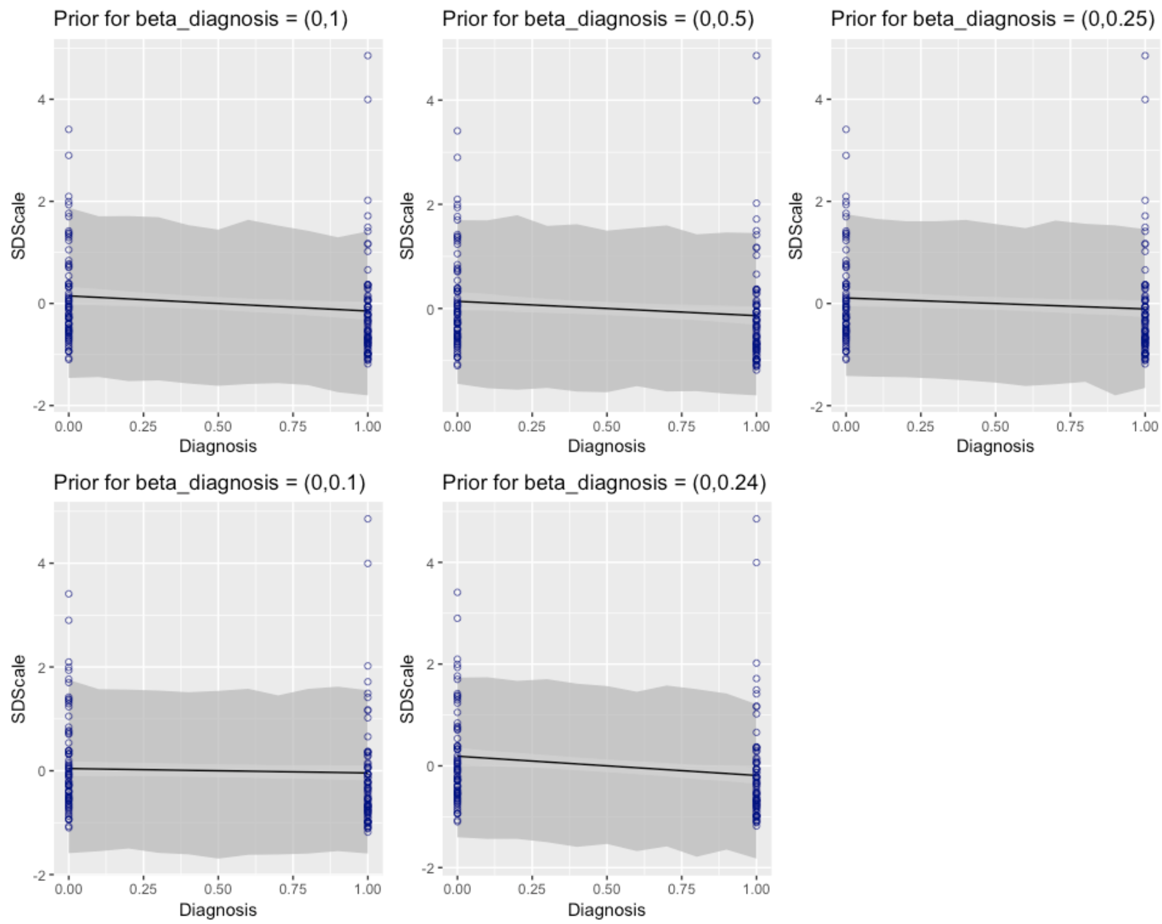


Figure 5 - Plot estimates for the model with the meta-analytic prior compared to the other models

Step 5: Compare the models - Plot priors and posteriors of the diagnosis effect in both models - Compare posteriors between the two models - Compare their relative distance from truth (WAIC) - Discuss how they compare and whether any of them is best.

When we compare the WAIC, we see that all the models have almost the same WAIC, and are close together in weight. Whenever the compare is re-run the order of models changes, as they are so similar, so it is not useful to use WAIC to compare the models.

```
       WAIC pWAIC dWAIC weight    SE  dSE
m1    433.5   5.9   0.0   0.24 37.41   NA
m4    433.6   5.0   0.1   0.23 35.67 3.26
mMeta 433.9   6.2   0.4   0.19 38.04 1.17
m2    434.0   6.3   0.5   0.18 37.87 0.50
m3    434.2   6.1   0.7   0.16 37.35 0.95
```

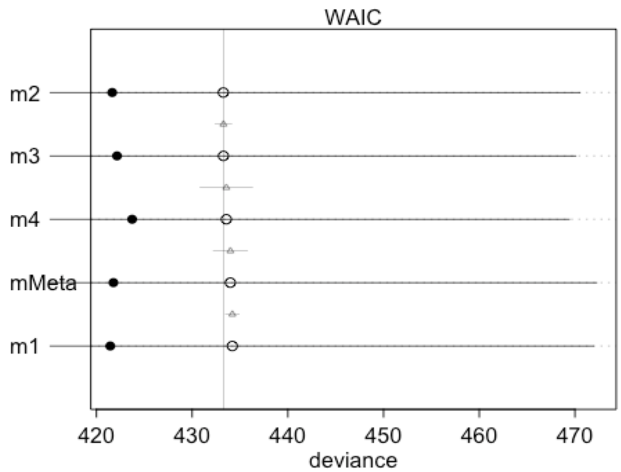


Figure 6 – WAIC comparisons of models

When we compare the posteriors, we see that model 4 has a much smaller beta (B = -0.08) than the other models, due to having a conservative prior (0.1). We know the priors affect the beta values of the models, which explains why the betas get systematically smaller. We can see the meta-analytic model has an even lower beta than model 1, and a prior that is similarly distributed to the loose prior of model 1. So the meta-analytic models prior pulls the mean of the distribution to -0.54.

	m1	m2	m3	m4	mMeta
a	0.15	0.14	0.11	0.04	0.19
b	-0.30	-0.28	-0.22	-0.08	-0.38
sigma	0.98	0.99	0.99	0.99	0.99
nobs	151	151	151	151	151

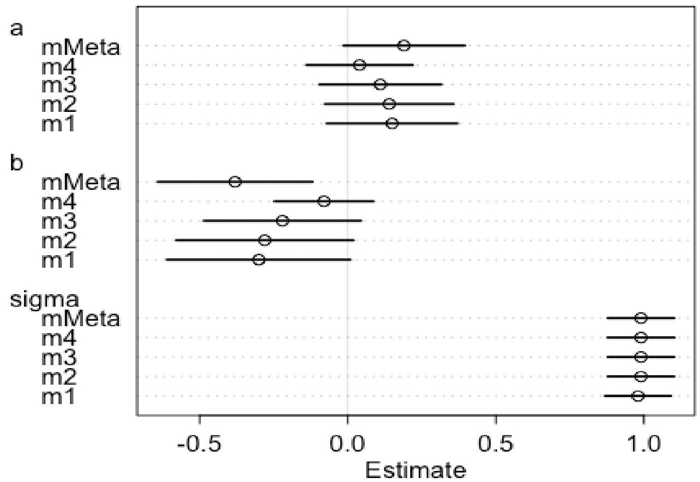


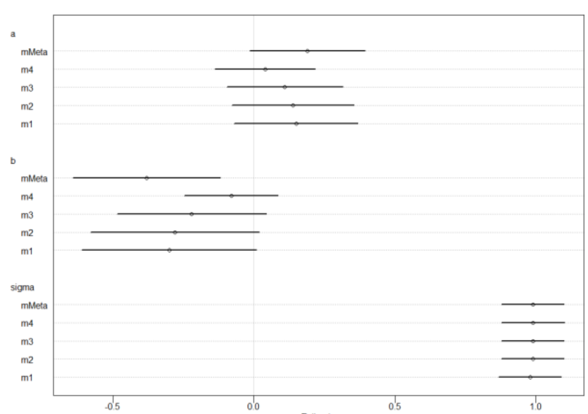
Figure 7 – Comparison of estimates using WAIC

Step 6: Prepare a nice write up of the analysis and answer the questions at the top.

The questions you need to answer are:

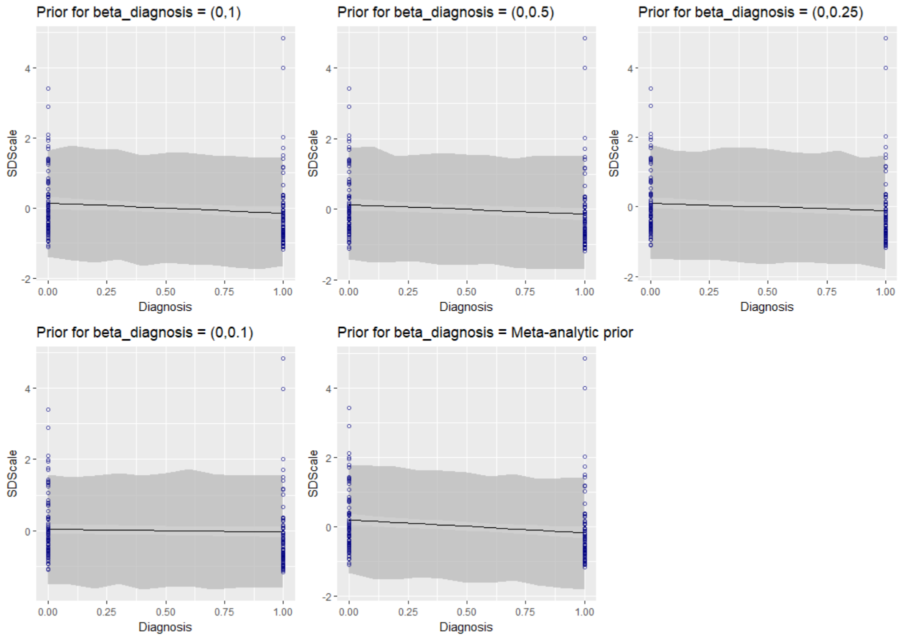
1) **What are the consequences of using a meta-analytic prior?**
   The consequence of the meta-analytic prior is distribution of posteriors that is shifted more towards -0.54, which is the weighted mean of the effect size of the previous studies combined. This gives a larger effect size than if we were to use a normal regularising prior.

2) **Evaluate the models with conservative and meta-analytic priors.**



When we compare the posteriors, we see that model 4 has a much smaller beta (B = -0.08) than the other models, due to having a conservative prior (0.1). We know the priors affect the beta values of the models (insert plot of pretty priors), which explains why the betas get systematically smaller. We can see the meta-analytic model has an even lower beta than model 1, and a prior that is similarly distributed to the loose prior of model 1. So the meta-analytic models prior pulls the mean of the distribution to -0.54.

3) Discuss the effects on estimates.



When plotting the estimates of the models we see that they are all very similar despite changing the priors. We can see the meta-analytic prior is a little more sloped than the others, and most similar to the model with a prior of 1.

The meta-analytic estimates are less conservative than the conservative model 4, and shows more of an effect.

4) Discuss the effects on model quality. Discuss the role that meta-analytic priors should have in scientific practice.

Effects on model quality – We can see in model comparisons that the meta-analytic prior doesn't have much positive effect on model quality.

Role of meta-analytic – Meta-analytic priors are useful in science because they allow for previously accumulated knowledge to be fused together and integrated into current analysis and studies. This gives the current analysis a guide for the estimates of the beta on the back of previous studies.

5) Should we systematically use them?

Yes, we should. Science is an accumulative practice, and it would be redundant to have to start from the bottom each time. We should integrate as much data as possible (this is the best doctrine in statistical practice), and if we have previous knowledge to build upon and integrate we have a higher chance of accurate results.

6) **Do they have drawbacks?**
Yes. We must be wary of publication bias. As we tend to only see studies with significant effect sizes in the literature, and we can see the meta-analytic priors bias larger effect sizes. This bias can accumulate if we always use meta-analytic priors, and cause false-positive errors in statistical analysis.

7) **Should we use them to complement more conservative approaches?**
Yes, we can always use a conservative prior and compare to the meta-analytic prior. It can regularise for both false-positives and false-negatives, given that the previous studies the meta-analysis is built on are high in validity and reliability.

8) **How does the use of meta-analytic priors you suggest reflect the skeptical and cumulative nature of science?**
As previously mentioned, science is an accumulative practice, and it would be redundant to have to start from the bottom each time. We should integrate as much data as possible (this is the best doctrine in statistical practice), and if we have previous knowledge to build upon and integrate we have a higher chance of accurate results.