

Larval h2 analysis: priors and posteriors

E Schlatter

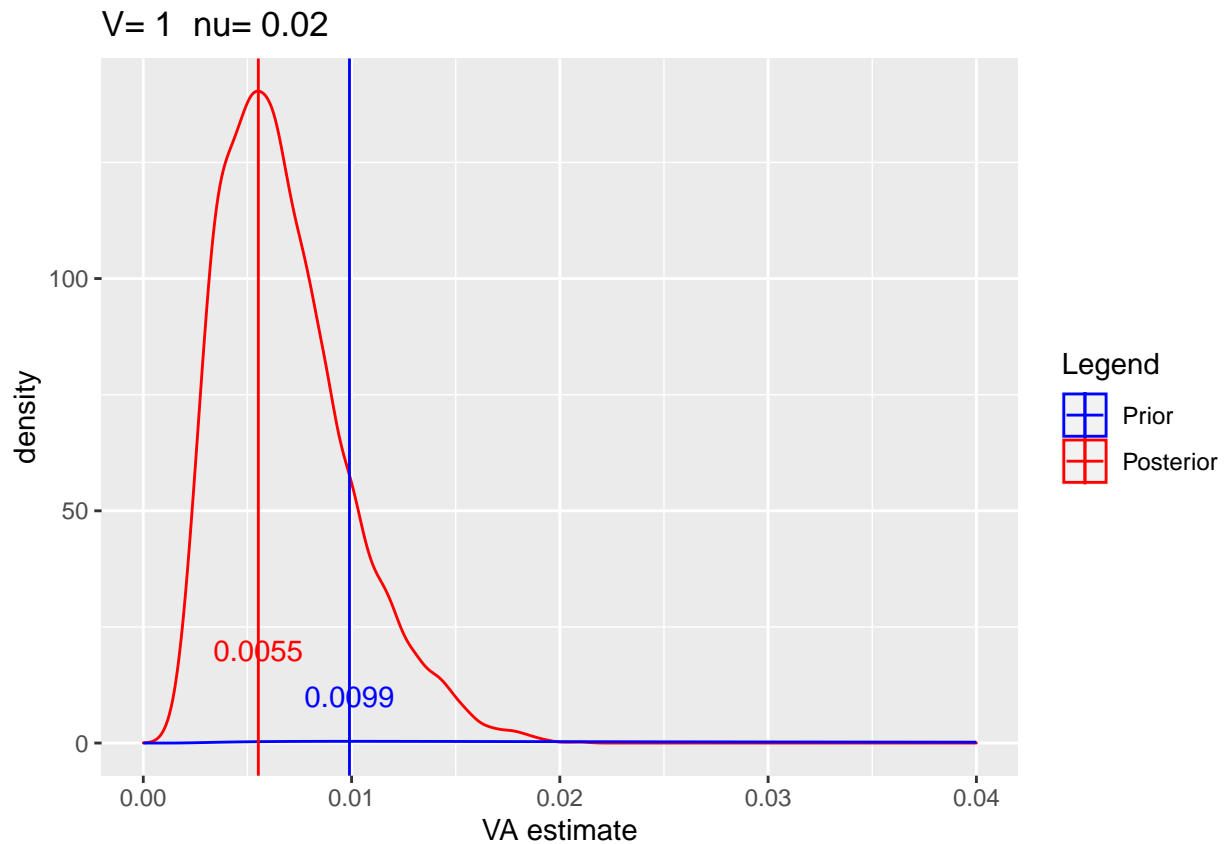
6/14/2022

Prior sensitivity

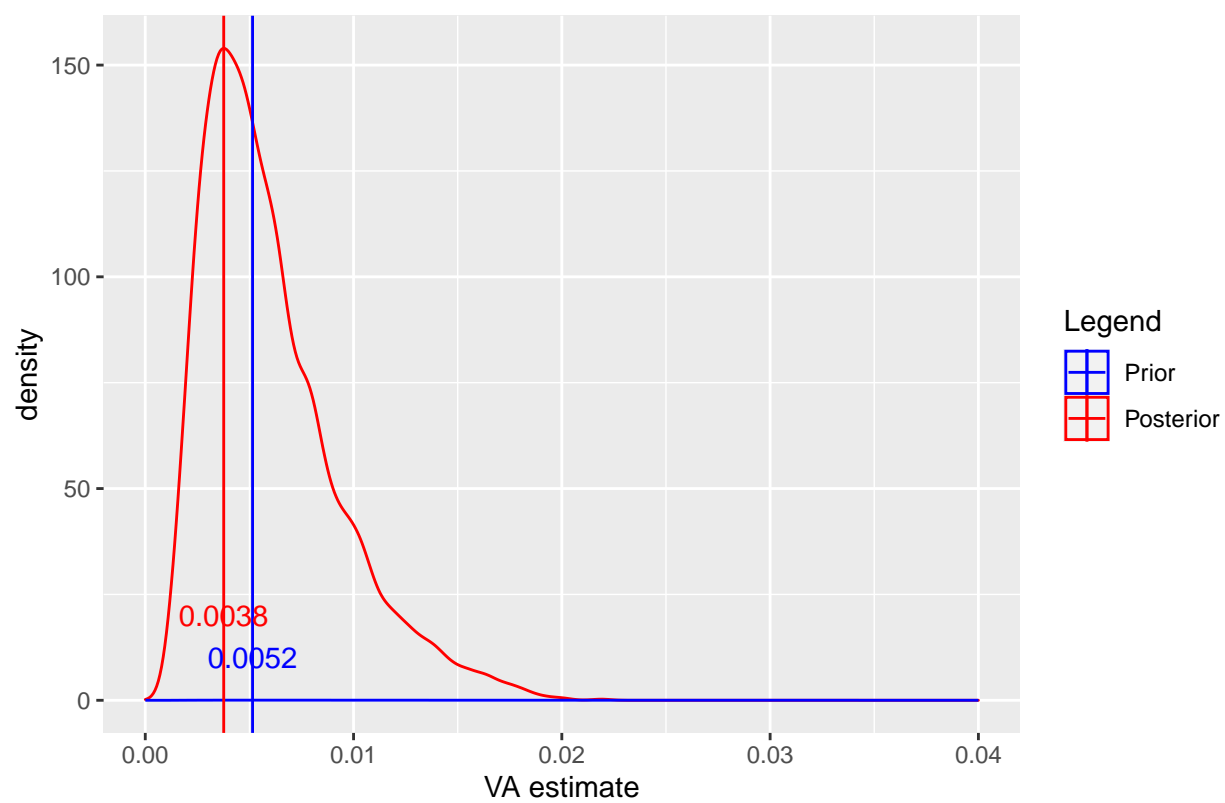
I've been looking at the sensitivity of estimates of variance components to what prior distribution is used. For example, let's look at estimating the additive genetic component of variance for standard length (in a model w/o parental effects). Here are the results for three different prior distributions:

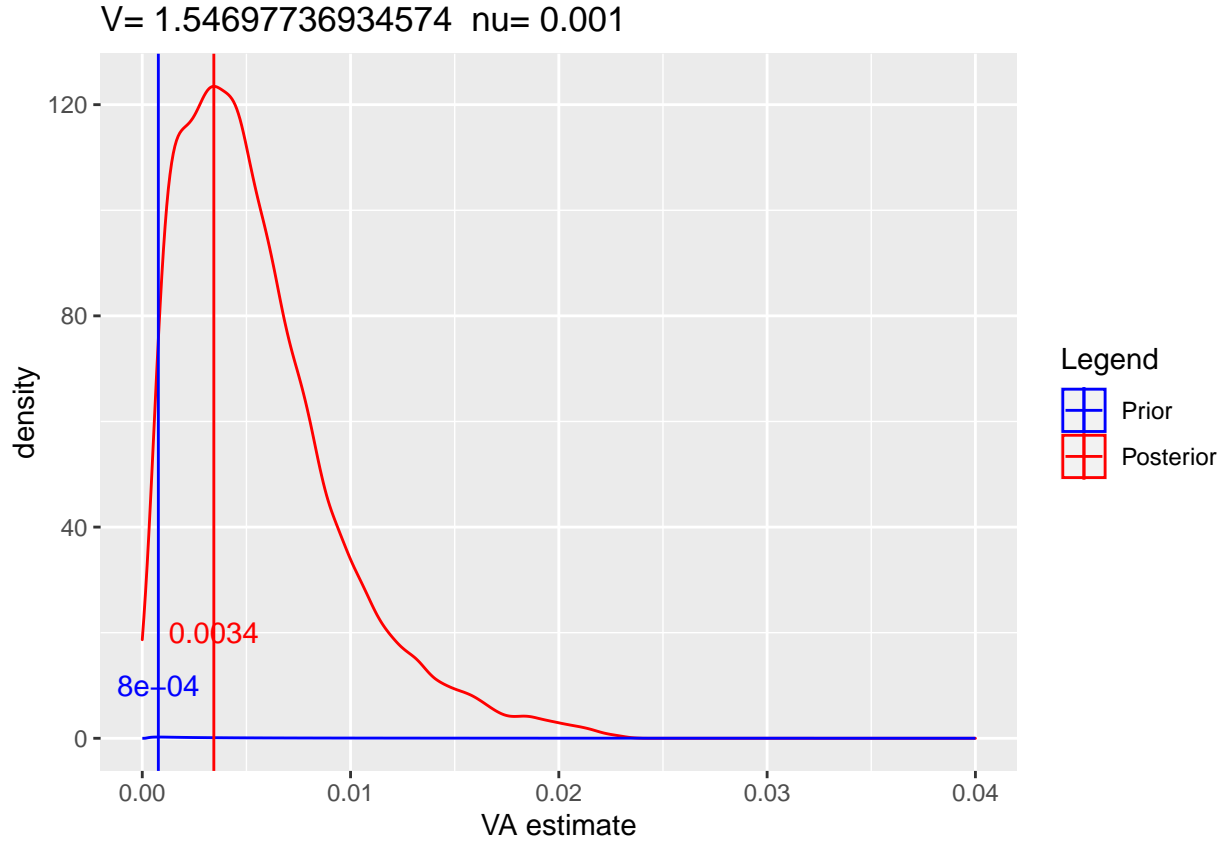
1. What I used in the manuscript ($V = 1$, $\nu = 0.02$; mode = 0.01)
2. Suggestion from reviewer ($V = 10.31$, $\nu = 0.001$; mode = 0.005)
3. One with a much smaller mode, chosen somewhat arbitrarily ($V = 1.55$, $\nu = 0.001$; mode = 0.0008)

Each plot shows the prior distribution in blue, and the posterior distribution in red. Vertical lines are the mode of each distribution. The priors look like a line at 0 because they're very flat, compared to the posteriors which have most of their probability concentrated in a very small range of values.



V= 10.3131824623049 nu= 0.001



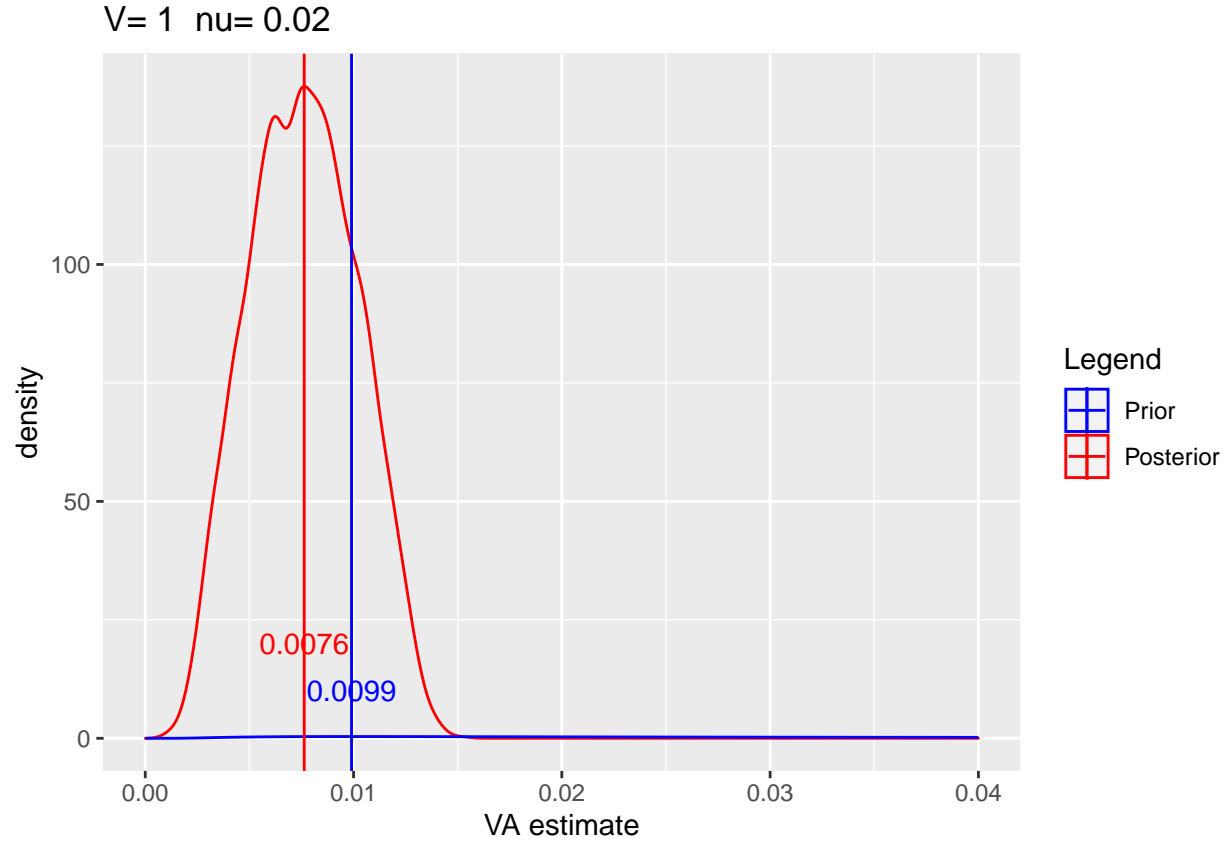


In cases 1 and 2, the posterior estimates of additive genetic variance are smaller than the priors. In case 3, when we start with a very small prior, the posterior estimate is larger than the prior. Taken together, these three examples seem to indicate that the posterior is moving away from the prior in an appropriate direction to converge around some intermediate value (~ 0.0033 - 0.0055), but that the prior does have some influence (the posterior estimate isn't exactly the same in all cases). I think this is basically what we want to see – although I'm not really sure what constitutes “too much” sensitivity to the prior.

I didn't include the plots for VE and VC here, but VE seems to vary with the priors to about the same extent as VA, while the estimate for VC is much lower but a little more consistent.

Null expectation for posterior

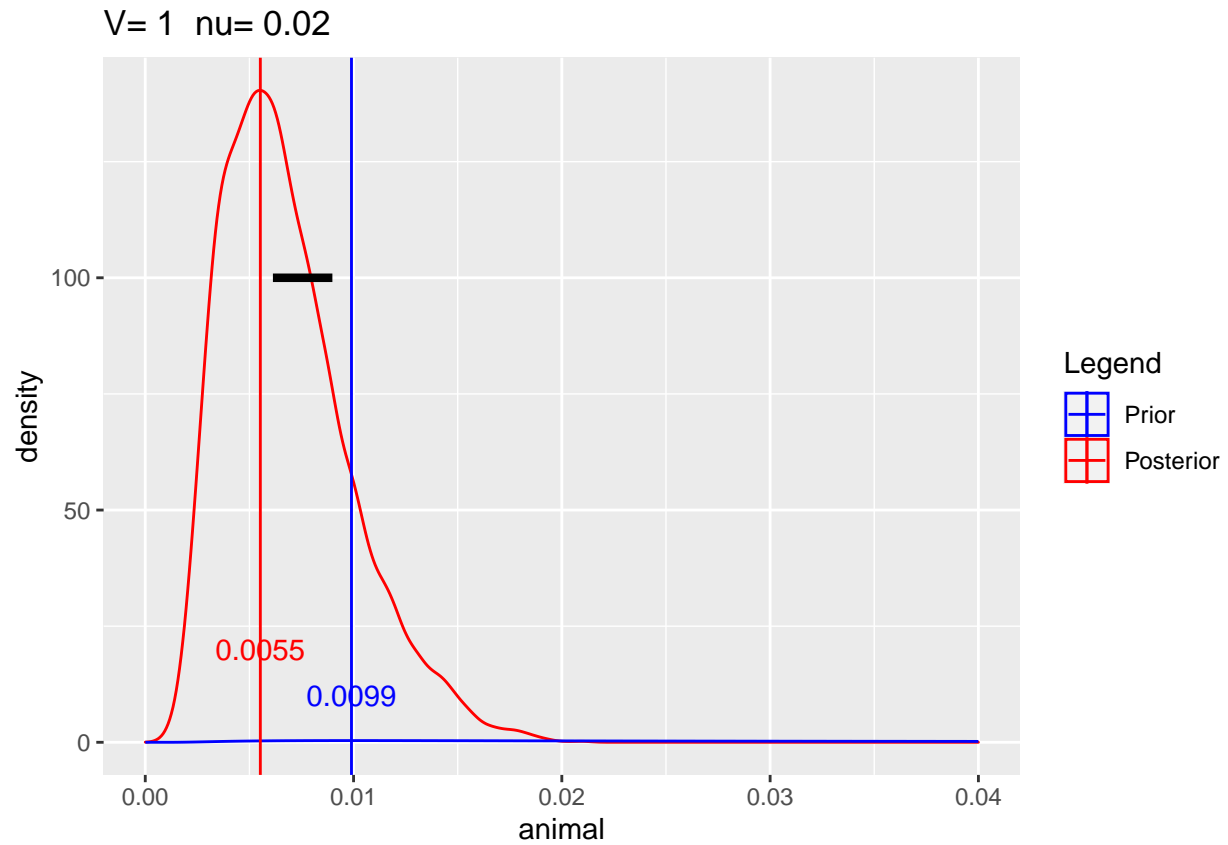
I also wanted to know how this estimate deviates from what we'd expect if there truly was no additive genetic variance. I broke up any causal relationship between standard length and relatedness by shuffling SL's randomly among larvae. Using prior 1 from above ($V=1$, $\nu=0.02$), the results look like this:



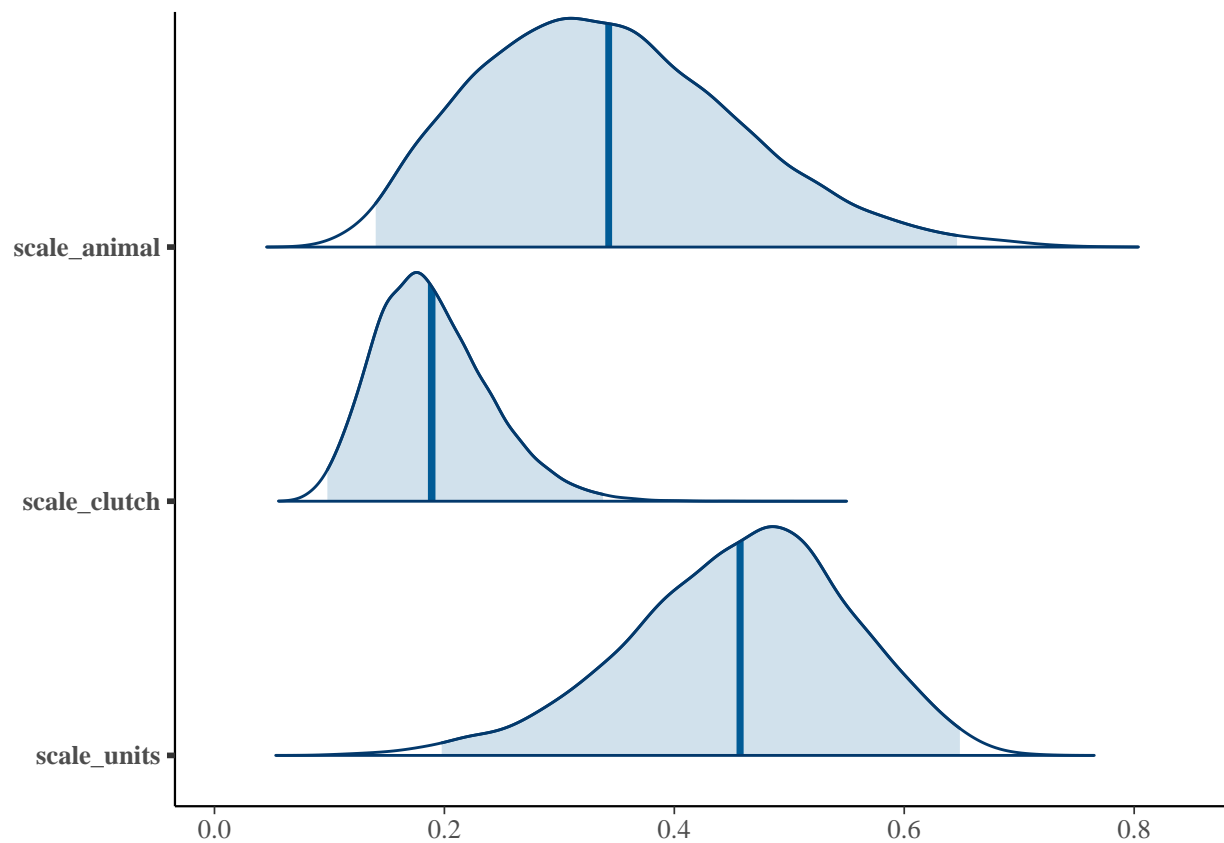
[Note that for this and most of the results presented here, the model convergence isn't terrible, but it also isn't ideal – that's why you see the bumpy peak above. If I ran the models for longer, I'd expect that peak to even out. But I think we can draw conclusions from the general patterns.]

The result shown in this plot seems like a concern – the estimate of additive genetic variance here, where there should be none, is larger than any of the estimates we got from using the actual data!

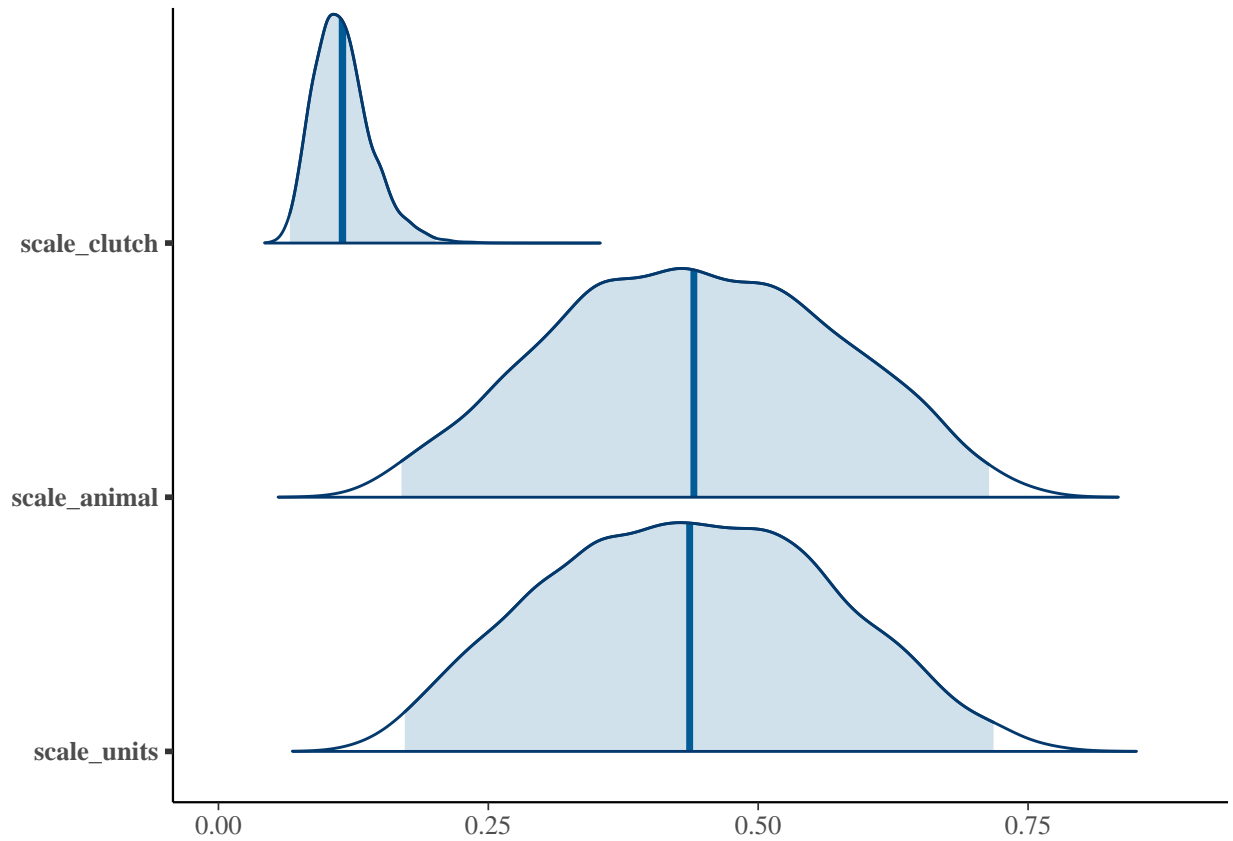
This isn't an artifact; when we reshuffle the SL's among larvae 100 times, we get the following: below is the same prior vs posterior plot from case 1 above (the prior used in the manuscript); the black bar represents the range of estimates of the mode of the posterior distribution of VA from the 100 random trials. None of them had a posterior mode as low as the one we get from the actual data.



So far we've been looking at estimates of V_A ; it's also worth looking at the estimates of heritability (i.e., V_A /total phenotypic variance). I don't have prior distributions for heritability (although I think I could theoretically derive them with some effort), but we can look at the posteriors. Here's the one using the prior from the manuscript:



And here are the corresponding posterior estimates for the shuffled data – that is, the null expectations for the variance components, if lengths were distributed randomly among larvae. (Note that clutch is the top row in this plot instead of the middle row)



This pattern is very consistent, no matter which prior is used, etc. Using the shuffled data, clutch is estimated to account for a small proportion of the variance, and the estimates for VA and VE are basically the same. That seems to me to say that, with the study design we have and the data we collected, if there is no underlying relationship between genetic relatedness and SL, the model apportions variance equally between VA and VE.

- This doesn't seem right. I'm not sure yet if it's a limitation of the study design, or if there's some error in the way I'm coding the model, or something else.
- If I haven't made some computational error, then what does it mean that our estimate of VA is actually lower than the null expectation?