

1 2 points

Suppose dataset A has a sample size of 1500, while dataset B has a sample size of 150. In all other ways, the datasets are very similar. All other things being equal, for which dataset will the posterior be more affected by the choice of the prior

☐ Dataset A

☒ Dataset B

The smaller the sample size, the more the prior plays a role

2 1 point

Suppose you are fitting a model with `stan_glmr()` and model formula $y \sim x + (1 | \text{group})$. Suppose you would like to change the prior distribution for the regression coefficient of x .

The name of the argument you would use is

prior
type your answer...

3 2 points

Suppose you are fitting a hierarchical model and get the warning "There were 7 divergent transitions." Please indicate all of the possible arguments that you could add that would decrease the chance of getting such a warning.

☐ `adapt_delta = .01`
☒ `adapt_delta = .99`
☐ `adapt_delta = 5`
☒ `prior_covariance = decov(5)`

adapt_delta refers to the jump proposal acceptance rate. It has to be less than 1.0, but the larger the value, the smaller the step size, the lower the chance of a big ~~step~~ surprise (a divergent transition)

- This pulls the posterior for ^{the} group-level effects away from zero and often is very helpful.

4

3 points

Suppose you are going to fit a Bayesian regression model $y \sim x$ to the dataset `df1`. You plan to use the default priors.

Suppose that as part of your model you would like the observations y_i to be distributed around their means according to Student's t distribution.

Please write an appropriate line of code to fit such a model. (That is, assume that the dataset `df1` exists and that all necessary packages have been loaded. Write the function, including all necessary arguments, to fit the described model.)

type your answer...

`brm(y ~ x, data = df1, family = student)`

this uses Student's t distribution instead of a normal
for $y_i \sim t(\mu_i, \sigma)$

you have to
use `brm()` and not `stan-glm()`
because `stan-glm()` does not
allow for student in the family
argument.

5

2 points

In a posterior predictive check, please describe what characteristic of the plot would make you concerned about the model fit. One or two sentences will be sufficient.

← → B I U A ▾ A ▾ Ix ≡ ≡ ≡ x^2 x_2 ≡ ≡ 12pt ▾ Paragraph ▾ ▢ ▣ ▤ 🔗 fr

I would be concerned if the black curve, the empirical probability density of the actual outcomes, looks quite different from the blue curves, " " " densities of simulated data. If they are quite different, it suggests that the model is inadequate in describing the actual data.