

HW 2

Shravan Vasishth

Exercise 1: Check for parameter recovery in a linear model using simulated data.

Generate some simulated independent and identically distributed data with $n = 100$ data points as follows:

```
y <- rnorm(100, mean = 500, sd = 50)
```

Next, fit a simple linear model with a normal likelihood:

$$y_n \sim \text{Normal}(\mu, \sigma) \quad (1)$$

Specify the following priors:

$$\begin{aligned} \mu &\sim \text{Uniform}(0, 60000) \\ \sigma &\sim \text{Uniform}(0, 2000) \end{aligned} \quad (2)$$

Generate posterior predictive distributions of the parameters and check that the true values of the parameters $\mu = 500, \sigma = 50$ are recovered by the model. What this means is that you should check whether these true values lie within the range of the posterior distributions of the two parameters. This is a good sanity check for finding out whether a model can in principle recover the true parameter values correctly.

Exercise 2: A simple linear model.

- Fit the model `fit_press` with just a few iterations, say 50 iterations (set warmup to the default of 25, and use four chains). Does the model converge?
- Using normal distributions, choose priors that better represent **your** assumptions/beliefs about finger tapping times. To think about a reasonable set of priors for μ and σ , you should come up with your own subjective assessment about what you think a reasonable range of values can be for μ and how much variability might happen. There is no correct answer here, we'll discuss priors in depth in a later chapter. Fit this model to the data. Do the posterior distributions change compared to the model fit in the chapter/lecture notes?

Exercise 3: Revisiting the button-pressing example with different priors.

- Can you come up with very informative priors that influence the posterior in a noticeable way (use normal distributions for priors, not uniform priors)? Again, there are no correct answers here; you may have to try several different priors before you can noticeably influence the posterior.
- Generate and plot prior predictive distributions based on this prior and plot them.
- Generate posterior predictive distributions based on this prior and plot them.

Exercise 4: Posterior predictive checks with a log-normal model.

- a. For the log-normal model `fit_press_ln`, change the prior of σ so that it is a log-normal distribution with location (μ) of -2 and scale (σ) of 0.5 . What does such a prior imply about your belief regarding median button-pressing times in milliseconds? Is it a good prior? Generate and plot prior predictive distributions. Do the new estimates change compared to earlier models that were fit in the chapter/lecture notes?
- b. For the log-normal model, what is the mean (rather than median) time that takes to press the space bar, what is the standard deviation of the finger tapping times in milliseconds? You have to read the chapter to find out the answer.

OPTIONAL Exercise 5: A skew normal distribution.

Would it make sense to use a “skew normal distribution” instead of the log-normal? The skew normal distribution has three parameters: location ξ (this is the lower-case version of the Greek letter Ξ , pronounced “chi”, with the “ch” pronounced like the “ch” in “Bach”), scale ω (omega), and shape α . The distribution is right skewed if $\alpha > 0$, is left skewed if $\alpha < 0$, and is identical to the regular normal distribution if $\alpha = 0$. For fitting this in `brms`, one needs to change `family` and set it to `skew_normal()`, and add a prior of `class = alpha` (location remains `class = Intercept` and scale, `class = sigma`).

- a. Fit this model with a prior that assigns approximately 95% of the prior probability of `alpha` to be between 0 and 10.
- b. Generate posterior predictive distributions and compare the posterior distribution of summary statistics of the skew normal with the normal and log-normal.