Bayes Factors

Bayesian Data Analysis

Steve Buyske

Bayes Factors

- You may well come across the use of what are known as Bayes factors.
- I don't recommend their use in general, but you should know about them.
- The idea is to compare the probability of a *model*, conditional on the data, to the probability of another *model*, conditional on the data.
- In short hand, compare $\text{Prob}(M_1 \mid \text{data})$ to $\text{Prob}(M_2 \mid \text{data})$.

Bayes' rule for models

- Imagine that instead of using Bayes' rule for parameters, we used it for an entire model (including parameters):

$$\text{Prob}(M_1 \mid \text{data}) = \frac{\text{Prob}(\text{data} \mid M_1)\text{Prob}(M_1)}{\text{Prob}(\text{data})}.$$

- What if we wanted to look at the ratio of $\text{Prob}(M_2 \mid \text{data})$ to $\text{Prob}(M_1 \mid \text{data})$?

$$\frac{\text{Prob}(M_2 \mid \text{data})}{\text{Prob}(M_1 \mid \text{data})} = \frac{\frac{\text{Prob}(\text{data}|M_2)\text{Prob}(M_2)}{\text{Prob}(\text{data})}}{\frac{\text{Prob}(\text{data}|M_1)\text{Prob}(M_1)}{\text{Prob}(\text{data})}}$$

$$= \frac{\text{Prob}(\text{data} \mid M_2)}{\text{Prob}(\text{data} \mid M_1)} \times \frac{\text{Prob}(M_2)}{\text{Prob}(M_1)}.$$

Bayes factor definition

- The first expression is the *Bayes factor* (BF) for comparing model 2 to model 1:

$$\text{BF}(M_1, M_2) = \frac{\text{Prob}(\text{data} \mid M_2)}{\text{Prob}(\text{data} \mid M_1)}.$$

- You can think of the Bayes factor as multiplying the ratio of prior probabilities of the models based on new data.
- What is a large Bayes factor? Two of its proponents, Kass and Raferty, provided the following table

$\log_{10}(BF)$	BF of M_2 to M_1	Evidence for M_2 compared to M_1
0 to 0.5	1 to 3.2	barely worth a mention
0.5 to 1.0	3.2 to 10	substantial evidence
1 to 2	10 to 100	strong evidence
greater than 2	greater than 100	decisive evidence

Shortcomings

- The mathematics of Bayes factors is indisputable and lovely. However, …
- Practical problem: calculating $\text{Prob}(\text{data} \mid M)$ can be difficult.
  - $\int \text{Prob}(\text{data} \mid M, \theta)\text{Prob}(\theta \mid M)\, d\theta$
    - can be high dimensional but
    - is particularly difficult to evaluate because of the way the two terms are related to each other.

- Theoretical problem: The Bayes factor is very sensitive to the priors (on the parameters, not the models).
- Consider a toy example: let's say that you have some data, a normal model, and a uniform prior on the mean μ.
  - If the prior has width a, then the density will be $1/a$, while if the prior has width $2a$, then the density will be $1/(2a)$.
  - That means the height of the likelihood $\times$ prior will be only half as high.
    - This doesn't affect the posterior, since we divide through by $\text{Prob}(\text{data})$.
- If you use Bayes factors you should be using informative priors based on previous experience.

Implementation

- To calculate Bayes factors starting with `stan_glm()` or `stan_glmer()`, you have to add an argument for a temporary file (notice each fit has a file with a different name).
- With `brm()`, you just need to add the argument `save_pars = save_pars(all = TRUE)`.
- In either case, you should have much longer chains, such as 40,000 in total.

```
citibike3_fit1c_bf <- stan_glmer(rides ~ high_temp + lrainfall + covid_cases + (1 | day_of_the_week),
                            data = citibike3 %>% filter(year == "2020"), adapt_delta = 0.999,
                            chains = 8, prior_covariance = decov(2),
                            diagnostic_file = file.path(tempdir(), "df1c.csv"))


citibike3_fit2_bf <- stan_glmer(rides ~ high_temp + lrainfall + (1 | day_of_the_week),
                            data = citibike3 %>% filter(year == "2020"), adapt_delta = 0.999,
                            chains = 8, prior_covariance = decov(2),
                            diagnostic_file = file.path(tempdir(), "df2.csv"))
```

Implementation cont.

- To get the Bayes factor, use the `bayesfactor_models()` function.
- The denominator goes first, the numerator second.

```
bayesfactor_models(citibike3_fit2_bf, citibike3_fit1c_bf)
```

```
## Warning in .bayesfactor_models_stan(..., denominator = denominator): Bayes factors might not be precise.
## For precise Bayes factors, it is recommended sampling at least 40,000 posterior samples.


## Computation of Bayes factors: estimating marginal likelihood, please wait...


## # Bayes Factors for Model Comparison
##
##   Model                                                                BF
##   [2] high_temp + lrainfall + covid_cases + (1 | day_of_the_week) 28041.621
##
## * Against Denominator: [1] high_temp + lrainfall + (1 | day_of_the_week)
## *   Bayes Factor Type: marginal likelihoods (bridgesampling)
```

- Here we see that the evidence is decisive in favor of the model with `high_temp + lrainfall + covid_cases + (1 | day_of_the_week)`.
- It is so strong, the choice of priors wouldn't have had much effect.