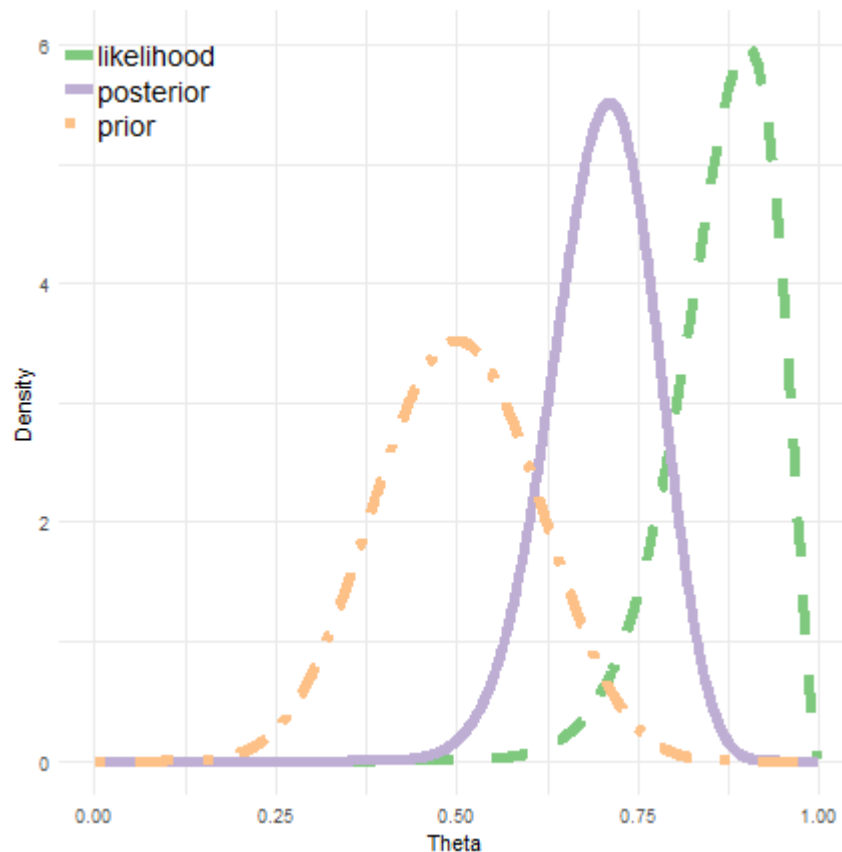# Taster on Bayesian Inference

Glenn Williams

2019-02-01

Bayes Rule (Does it?)

General approach is to use **prior information** with your observed data to determine the most plausibile parameter estimates conditional on your model. Probability is interpreted as a **degree of belief**.

Frequentist vs. Bayesian Statistics

Some differences between Frequentist (left) and **Bayesian** statistics (right):

1. Finds the best model (amongst many others) to fit your data.

2. 95%[1] **confidence interval**: Using this method to calculate CI, 95% of CIs will contain the true parameter value in the long run.

3. Naive models: Assumes flat priors where all values are equally likely. (Including crazy-large effects; i.e. not those in most psychological research.)

1. Gives a distribution of models that are (more or less) likely given the data.

2. 95% **credible interval**: 95% probability the CI contains the true parameter value[2].

3. (Can be) less naive models: Allows flat priors on parameters, as well as **weakly informative, regularising priors**, or even **strongly informative priors**.

[1] Other intervals are available in all major supermarkets.
[2] Conditional on the data and model.

Mother-flippin' Priors

How does a prior work? Define your expectations:

$IQ \sim Normal(\mu_i, \sigma)$ [likelihood]

$\mu_i = \alpha + \beta x_i$ [linear model]

$\alpha \sim Normal(100, 15)$ [alpha prior]

$\beta \sim Normal(0, 10)$ [beta prior]

$\sigma \sim Uniform(0, 20)$ [SD prior]

- A data generation process for IQs
- Linear model to predict outcomes
- Intercept (control condition here)
- Slope (difference between control & treatment here)
- Variation (SD) from the mean

Remember, 95% probability of IQs between $\pm SD \times$ 1.96.

This model expects a wide range of differences between the control and treatment conditions, but puts **most probability around 0 (small effects are more likely)**.

Effects of Priors on Results

Priors can be split into a few broad categories:

- **Flat Priors**: So-called "uninformative priors"[1]. All values receive equal probability prior to fitting. Generally, a bad idea and doesn't reflect your beliefs.

- **Informative Priors**: Specific predictions based on prior experience (e.g. knowing effect sizes from the literature).

- **Weakly Informative Priors**: Vague predictions that act to "regularise" the data. Often, extreme effect sizes are downweighted, and smaller effect sizes receive more weighting. Many "default" priors in Bayesian analysis packages use these types of priors.

Both types of informative priors work well with small data sets. Priors can skew predictions (more so the stronger they are), but given enough data all priors are dominated by the likelihood.

[1] They are informative, though, as they still make a statement of belief.
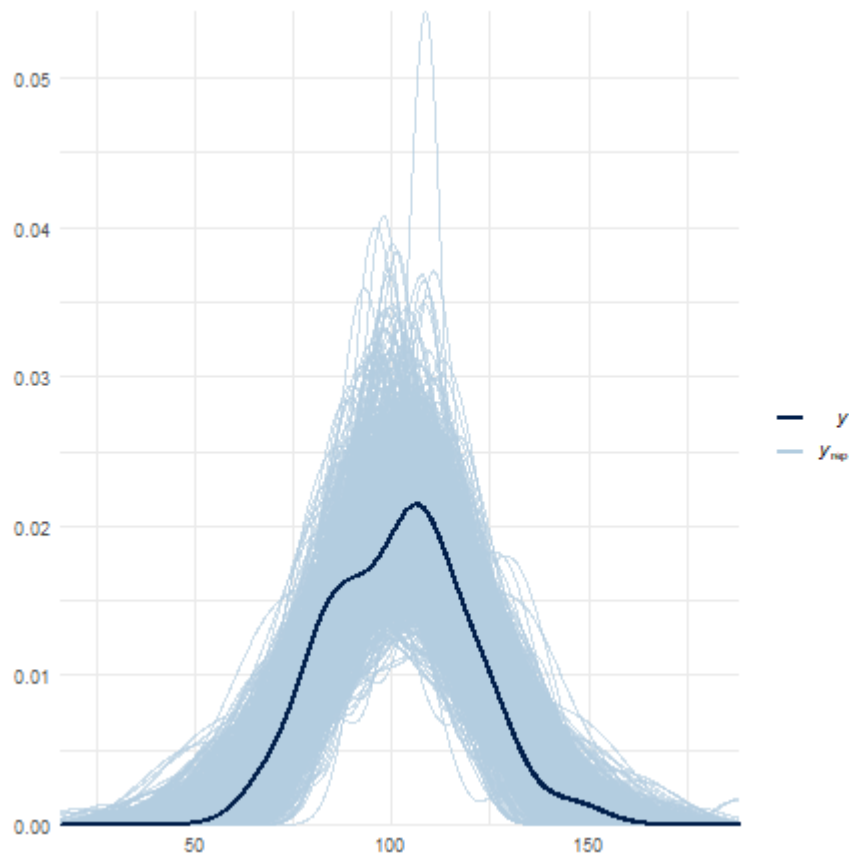
Effects of Priors on Results: Examples
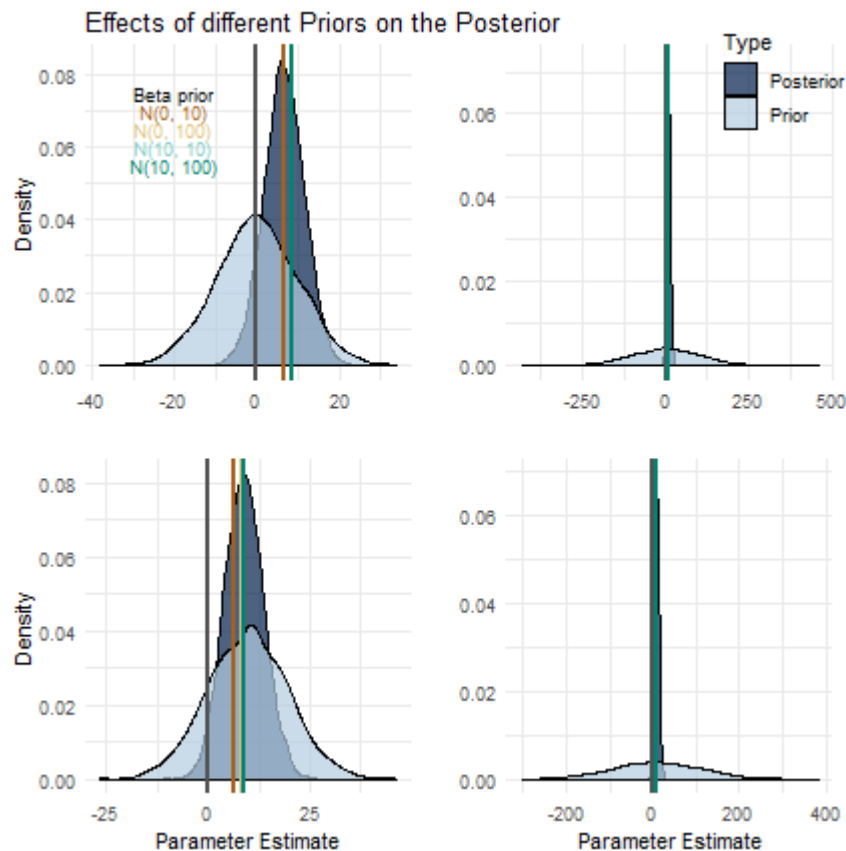
The wider the prior, the less impact it has on the posterior regardless of its position. More specific priors have a larger impact on the posterior.

Bayesian Inference

Many options, but often involves sampling from the posterior to determine plausible parameter values and intervals of values.

Bayesian Inference (Continued)

Compare the Bayesian estimate with the wide prior (*SD* = 100) centred on 0 for the effect of intervention to that of the frequentist estimate.

Bayesian

term	estimate	error	l-95%	u-95%
Intercept	97.9	3.9	90.2	106
Intervention	8.5	5.4	-2.1	19

Frequentist

term	estimate	error	l-95%	u-95%	t-value	p-value
Intercept	97.9	3.8	90.2	106	25.6	0.00
Intervention	8.6	5.4	-2.4	19	1.6	0.12

Bayes Factors

The Bayes factor provides evidence for one hypothesis[1] against another (e.g. null over alternative).

But, use this with caution. Just because one model is "better" than another, it still might be terrible.

hypothesis	estimate	error	l-95%	u-95%	evid.ratio	post.prob
Intervention = 0	8.5	5.4	-2.1	19	5.3	0.84

We calcaulted the Bayes factor using the **Savage-Dickey density ratio** in the `brms` package. This calculates the ratio of evidence for one model over another at some parameter value of interest (e.g. effect size = 0).

Thus, BF_{01} = 5.27. The data are approximately **5 times more likely under the null than the alternative hypothesis**.

[1] Conditional on the model and the data.

Ways to Perform Bayesian Analyses (Programmy)

Stan

Statistical computing languge. Very involved, but very efficient and flexible.

R (packages)

1. `brms` package covers multilevel modelling with arbitrarily complex models. It is an `lme4`-like interface to Stan (a language for Bayesian analysis). Most flexible, but most involved. Has sensible default priors, but requires you to specify priors for Bayes factors.

2. `BayesFactor` package covers calculating Bayes factors. Again has sensible default priors. All effects/priors are on a standardised scale. More limited use, but covers most design cases (except, e.g. logistic regression).

Ways to Peform Bayesian Analyses (Clicky)

JASP

Free, open source software that provides a nice, intuitive GUI for doing frequentist and Bayesian statistics. Relies on the BayesFactor R package as its Bayesian engine.

SPSS (boooooooo!)

I'm not going to pretend I know anything about this. But, they now have a Bayesian Statistics submenu in Analyze.