

Advanced Data Analysis in R

Bayesian Modeling in R

Michael DeWitt

2018-03-17 (Updated 2019-04-14)

Bayesian Modeling in R

A Thought Exercise

You are already Bayesian!

You just didn't know it!

What is the probability a given coin is fair?

If you didn't answer 100% or 0% you're Bayesian!

What is Bayes?

- Named after Rev. Thomas Bayes

```
knitr::include_graphics(here::here("bayesian_modeling", "figs",
```

Just derived from identities of probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Prior, Likelihood, Posterior Distribution

Prior is *subjective* and represents a range of potential values

Specified by a distribution

Informative

Restricts range of likely values to small range

More informative adds more “weight” to the prior vs the data

Noninformative/ Uninformative

Wider range of possibilities

More closely approximates the Maximum likelihood estimates

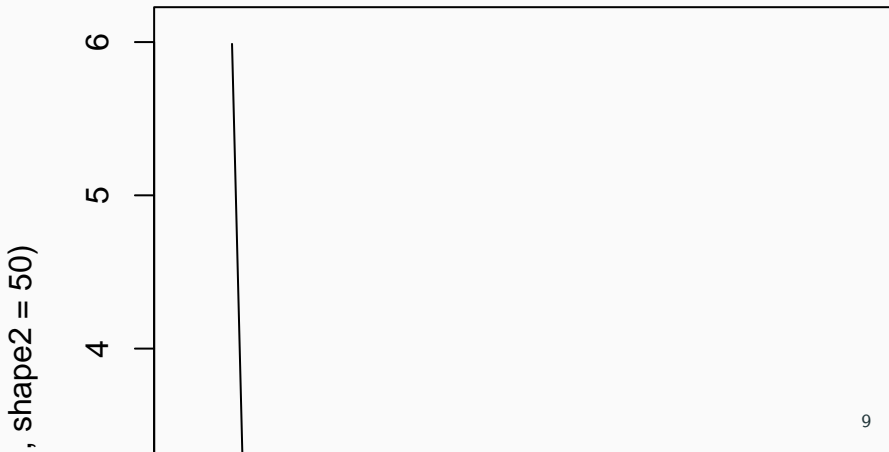
Likelihood is the distribution for the data generating process

Examples

- Poisson process -> poisson likelihood function
- Binomial process -> binomial likelihood function
- Normal distribution -> normal likelihood function
- Ordered categorical -> ordered categorical likelihood function

$$P(A|B) \sim \text{Prior} * \text{Likelihood}$$

```
curve(expr = dbeta(x, shape1 = .1, shape2 = 50), from = 0, to =
```



Frequentist vs Bayesian Inferences

Bayesian Workflow

Create A Data Generating Process

Define your data generating process

Generate some fake data from it

Write Your Model

Test it! Write a model, run it and test it.

Domain Specific Languages

BUGS

JAGS

Stan

Hand coded samplers

brms makes Bayesian Modeling Easy

Utilises Hamiltonian Monte Carlo with a No U-Turn Sampler

Specifying a Model in brms

Run the Model

- Convergence
 - Trace Plots `tracplot`
 - Rhat metrics
 - Effective Sample Size
- Posterior Predictive Intervals
 - Was there a good fit between the model and the data

Advantages of Bayesian Analysis

- Takes advantage of expert opinion
 - Especially helpful with small samples size studies
 - Reduces possibility of wildly odd results
- Easier communications (more intuitive to discuss probabilities)
- Studies can build on one another
 - Results from one study can be supplied directly as a prior into a replication or another study

Drawbacks of Bayesian Inference

- Not as widely utilised in major publications
- Computationally intensive
- Picking a prior

