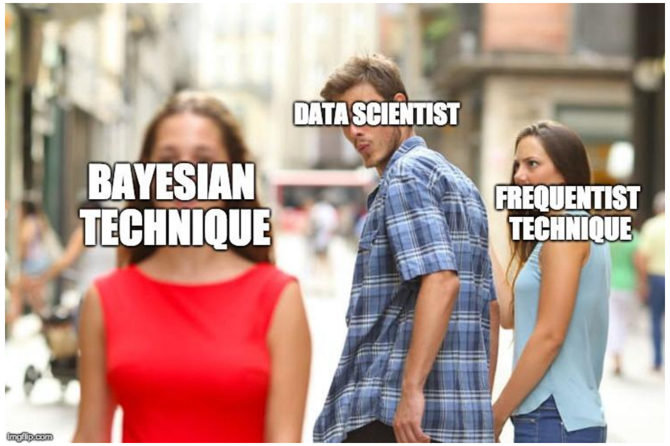
Intro to Bayesian Inference

Nick Luymes – Learning Lab Meetings

April 17, 2024

Frequentist vs. Bayesian inference



Fixed Parameters

Parameters with a Distribution

Frequentist vs. Bayesian inference

Frequentist: Probability of event based on its **frequency** in random trials

Bayesian: Probability of event based on **beliefs** that are **updated** based on new data

Frequentist vs. Bayesian inference

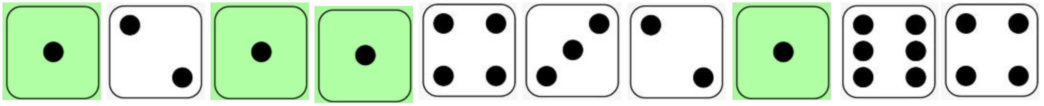
Example: probability of rolling a **1** using a weighted dice

Frequentist: frequency of roll outcomes

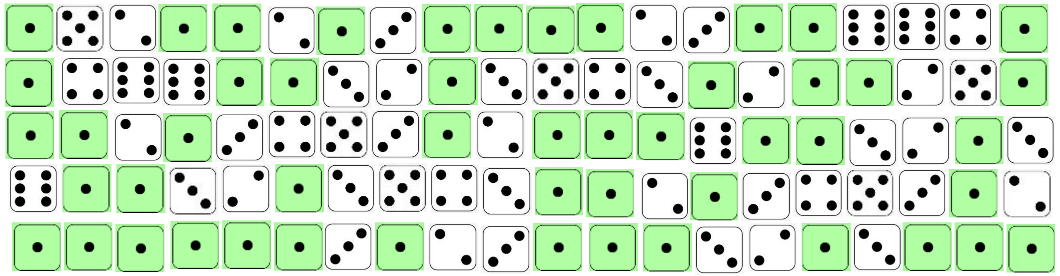
Bayesian: prior belief of 1/6 ≈ 17%

Frequentist vs. Bayesian inference

10 rolls: 4 **1s**



100 rolls: 50 **1s**



Frequentist vs. Bayesian inference

10 rolls: 4 **1s**

Frequentist: 40%

100 rolls: 50 **1s**

Frequentist: 50%

Frequentist vs. Bayesian inference

10 rolls: 4 **1s**

Frequentist: 40%

Bayesian: ~17%

100 rolls: 50 **1s**

Frequentist: 50%

Bayesian: ~50%

Bayes Theorem

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

Bayes Theorem

- Example: probability of bird in some area given the presence of preferred host tree in that area

$$P(bird|tree) = \frac{P(tree|bird) * P(bird)}{P(tree)}$$

Bayes Theorem

- Example: probability of bird in some area given the presence of preferred host tree in that area

$$P(bird|tree) * P(tree)$$
$$= P(tree|bird) * P(bird)$$

Bayes Theorem

$$\overbrace{P(\theta|model)}^{\text{posterior}} = \frac{\overbrace{P(model|\theta)}^{\text{likelihood}} * \overbrace{P(\theta)}^{\text{prior}}}{\underbrace{P(model)}_{\text{marginal}}}$$

Frequentist vs. Bayesian inference

Prior: 17%

Likelihood: Binomial

10 rolls: 4 **1s**

Posterior: ~17%

100 rolls: 50 **1s**

Posterior: ~50%

$$\overbrace{P(\theta|model)}^{\text{posterior}} = \frac{\overbrace{P(model|\theta)}^{\text{likelihood}} * \overbrace{P(\theta)}^{\text{prior}}}{\underbrace{P(model)}_{\text{marginal}}}$$

Bayesian Modelling

- Problem: Posterior distribution is often difficult or impossible to derive
- Solution:
  1. Specify prior distributions
  2. Define model
  3. Generate samples of parameters from the posterior distribution using Markov Chain Monte Carlo algorithms

$$\overbrace{P(\theta|model)}^{\text{posterior}} = \frac{\overbrace{P(model|\theta)}^{\text{likelihood}} * \overbrace{P(\theta)}^{\text{prior}}}{\underbrace{P(model)}_{\text{marginal}}}$$

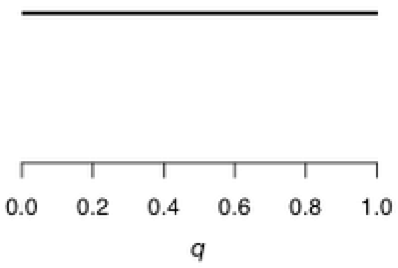
Example: Flipping a coin

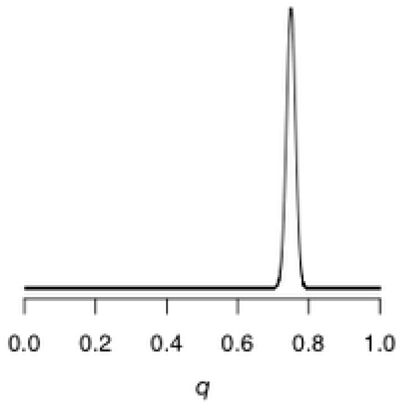
The brms Package

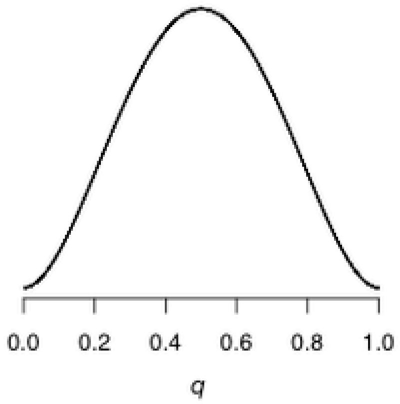
- Uses STAN
- Similar syntax to popular frequentist functions/packages (e.g. lm(), glm(), lme4)
- Allows for many different linear and non-linear model structures

Types of priors

- **Flat prior**: all potential values are equally likely – *data determines posterior*
- **Strong prior:** fairly certain about the parameter value – *prior determines posterior*
- **Weak prior:** partial information about the parameter value – *data and prior determine posterior*







Example: Salamander larvae abundance

Pros and Cons

Pros:

- Incorporate prior knowledge
- Flexible model specification
- Robust estimates of uncertainty

Cons:

- Computationally intensive
- Subjectivity
- Easy for models to be misspecifed

Further learning

- Deeper intro to Bayesian inference: https://statswithr.github.io/book/
- Brms vignettes: https://paul-buerkner.github.io/brms/
- Choosing priors: https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations
- Avoiding bad habitats: https://www.nature.com/articles/s43586-020-00001-2