Introduction to Bayesian Statistics

Bayesian Modeling in brms

Julia Haaf

September, 2022

Disclaimer

- Bayesian statistics is complicated.

Disclaimer

- Bayesian statistics is complicated.
- Hierarchical modeling is complicated.

Disclaimer

- Bayesian statistics is complicated.
- Hierarchical modeling is complicated.
- Cognitive modeling is complicated.

Disclaimer

- Bayesian statistics is complicated.
- Hierarchical modeling is complicated.
- Cognitive modeling is complicated.
- This is an extremely brief primer.

Disclaimer

- Bayesian statistics is complicated.
- Hierarchical modeling is complicated.
- Cognitive modeling is complicated.
- This is an extremely brief primer.
- Check out resources here: https://github.com/jstbcs/ESCOP2022-WS.

An Example

- This is Frank.

An Example

- This is Frank.
- Frank likes to eat but he might be a tad picky.

An Example

- This is Frank.
- Frank likes to eat but he might be a tad picky.
- We want to model how often Frank eats his food in a month.

An Example

Is Frank a picky eater?

- $Y \sim \text{Binomial}(\theta, 30)$, modeling how often out of 30 Frank eats his food.

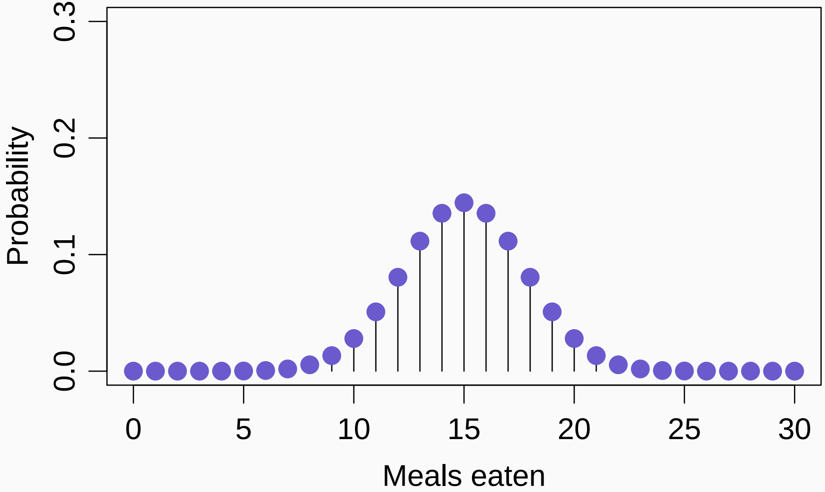
An Example

Is Frank a picky eater?

- $Y \sim \text{Binomial}(\theta, 30)$, modeling how often out of 30 Frank eats his food.
- $\theta = .5$, assuming the probability of eating is 50/50.

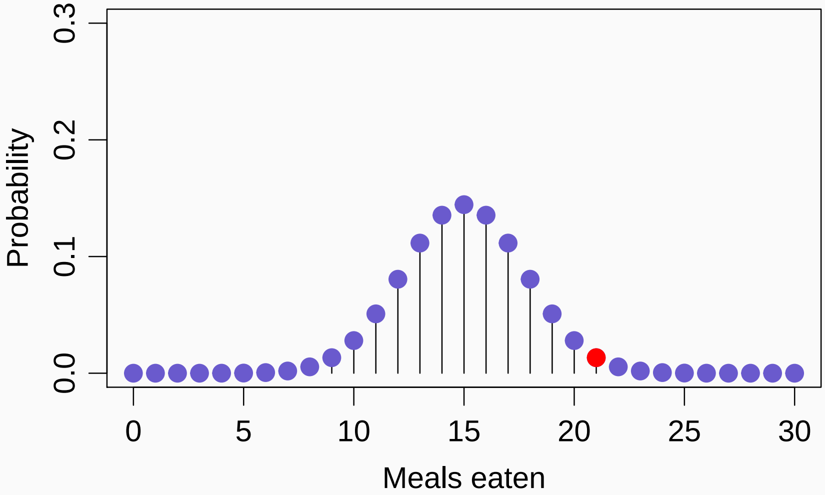
Models, an Example

Predictions on data, based on the model $Y \sim \text{Binomial}(0.5, 30)$.



Data

Let's say he ate 21 out of 30 meals. $Y = 21$.



Priors

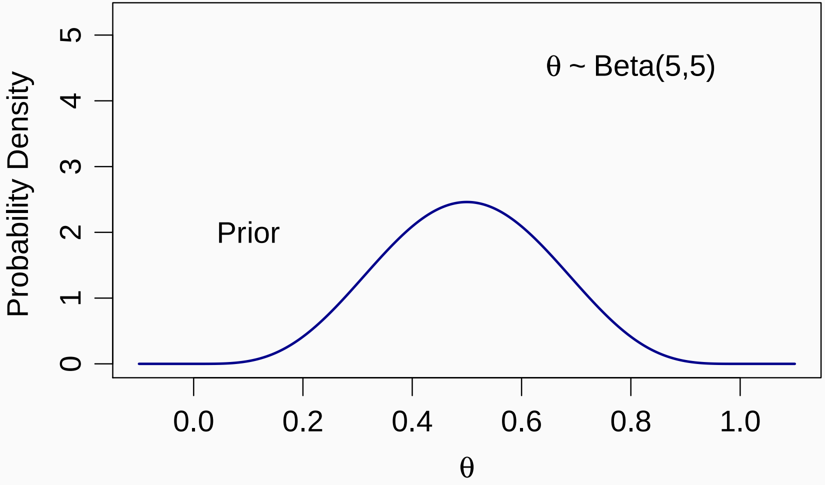
In Bayesian statistical analysis we typically would use a prior *distribution* for parameters.

$$
\begin{aligned}
Y &\sim \text{Binomial}(\theta, N),\\
\theta &\sim \text{Beta}(a, b).
\end{aligned}
$$

If we assume Frank will most likely eat 5 out of 10 meals we may use $a = 5$ and $b = 5$.

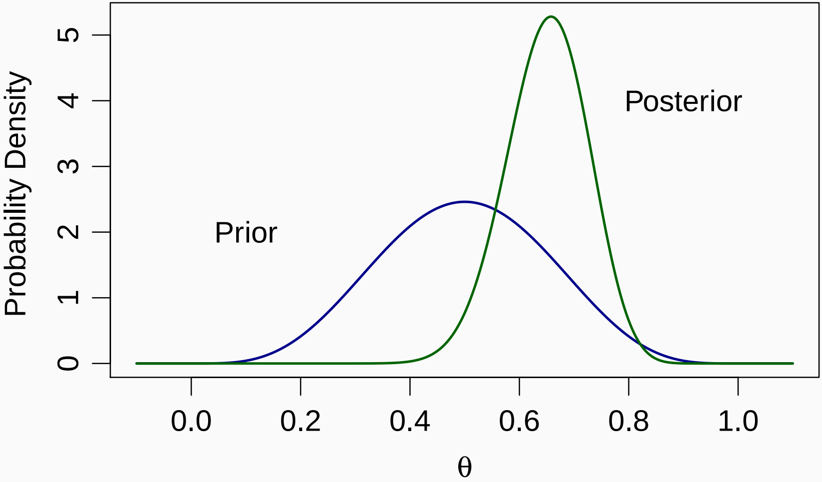
Posterior Updating

from $Pr(\theta)\ldots$



Posterior Updating

from $Pr(\theta)$... to $Pr(\theta|Y)$.



Bayes' Rule:

$$Pr(\theta|Y) = Pr(\theta)\frac{Pr(Y|\theta)}{Pr(Y)},$$

- $Pr(\theta|Y)$ is the *posterior distribution* of θ.

Bayes' Rule:

$$Pr(\theta|Y) = Pr(\theta)\frac{Pr(Y|\theta)}{Pr(Y)},$$

- $Pr(\theta|Y)$ is the *posterior distribution* of θ.
- $Pr(\theta)$ is the *prior distribution* of θ.

Bayes' Rule:

$$Pr(\theta|Y) = Pr(\theta)\frac{Pr(Y|\theta)}{Pr(Y)},$$

- $Pr(\theta|Y)$ is the *posterior distribution* of θ.
- $Pr(\theta)$ is the *prior distribution* of θ.
- $Pr(Y|\theta)$ is the probability distribution of the data.

Bayes' Rule:

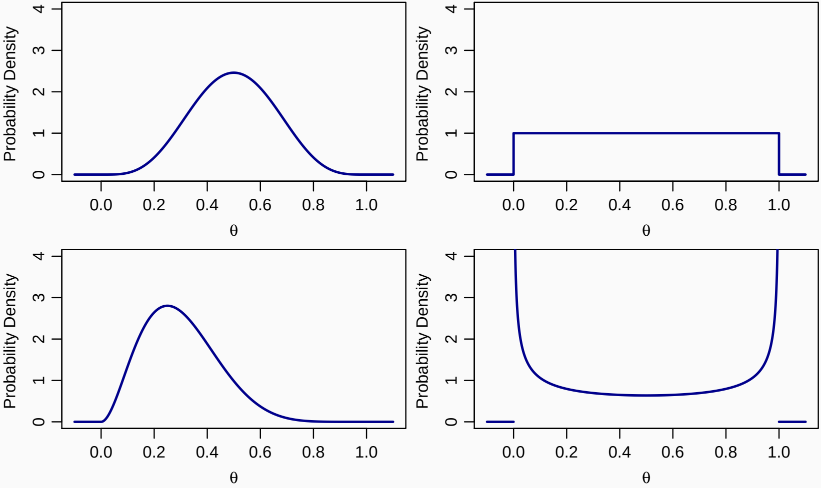
$$Pr(\theta|Y) = Pr(\theta)\frac{Pr(Y|\theta)}{Pr(Y)},$$

- $Pr(\theta|Y)$ is the *posterior distribution* of θ.
- $Pr(\theta)$ is the *prior distribution* of θ.
- $Pr(Y|\theta)$ is the probability distribution of the data.
- $Pr(Y)$ is the prediction for the data.

Priors

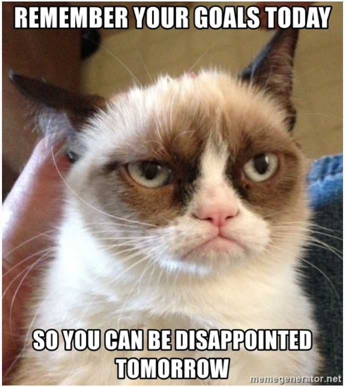
Did we choose a good prior?

What should the prior on Frank's eating habit look like?



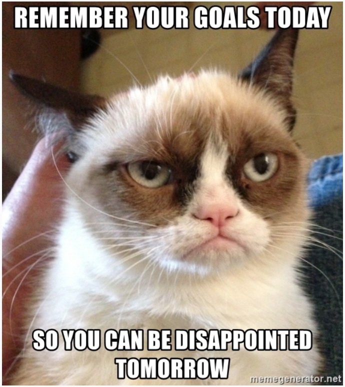
Matching priors to goals of analysis

- There are priors that are most suitable for estimation.



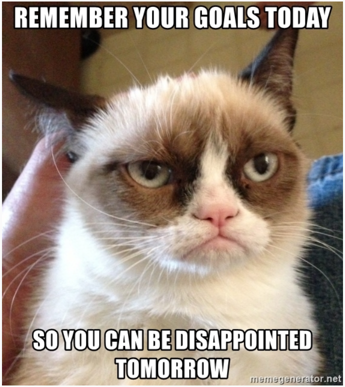
Matching priors to goals of analysis

- There are priors that are most suitable for estimation.
- And there are priors most suitable for model comparison.



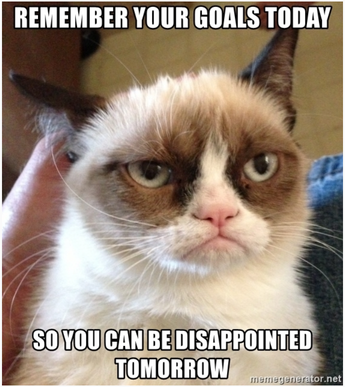
Matching priors to goals of analysis

- There are priors that are most suitable for estimation.
- And there are priors most suitable for model comparison.
- And there are priors that are pretty good for both.



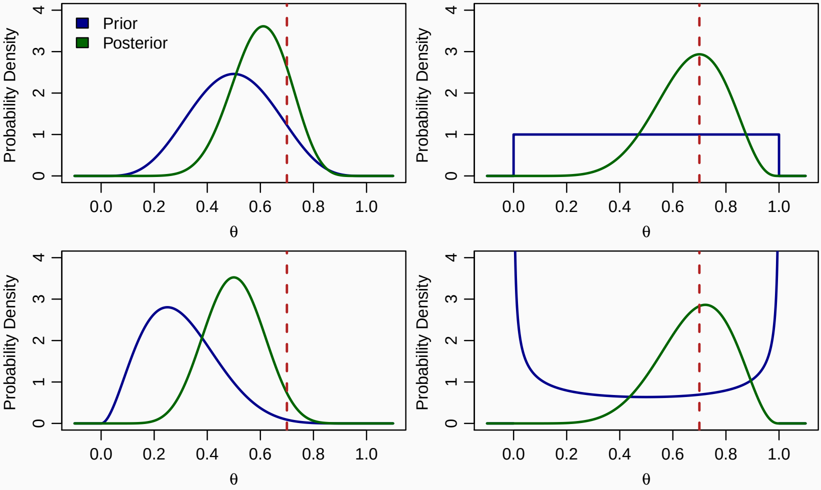
Matching priors to goals of analysis

- There are priors that are most suitable for estimation.
- And there are priors most suitable for model comparison.
- And there are priors that are pretty good for both.
- Oh, and not everyone agrees on this classifications (or what "good means").



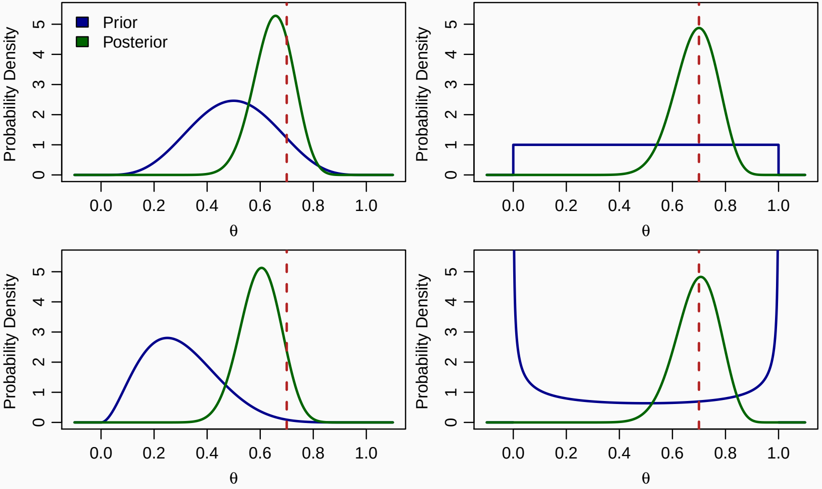
Why Priors for Estimation Don't Matter That Much

Frank eats his food 7 out of 10 times.



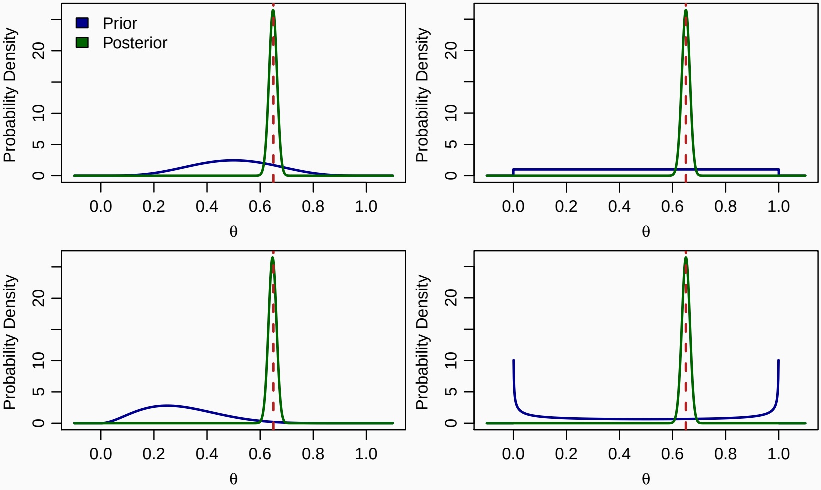
Why Priors for Estimation Don't Matter That Much

Frank eats his food 21 out of 30 times.



Why Priors for Estimation Don't Matter That Much

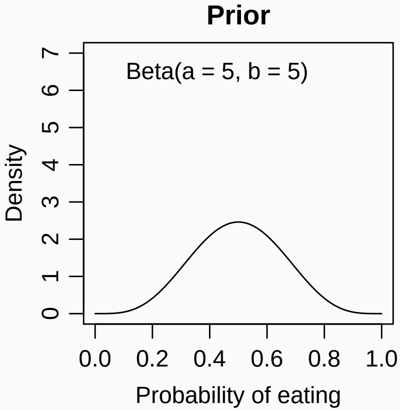
Frank eats his food 650 out of 1000 times.

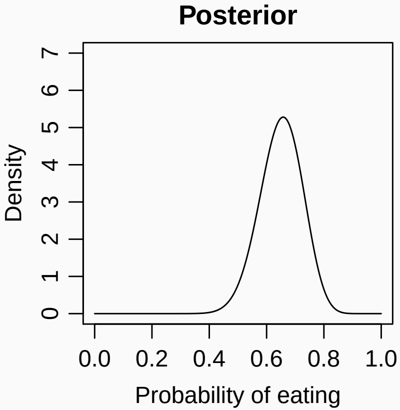


Posterior-based summaries of results

Posterior-based summaries of results

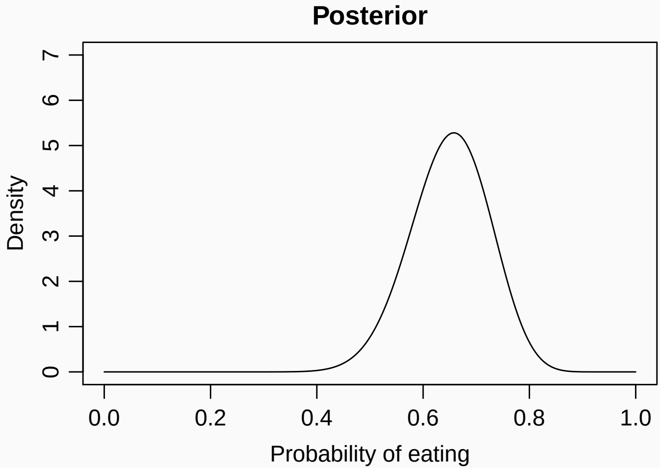
Once we have obtained a posterior distribution, how can we summarize the results?





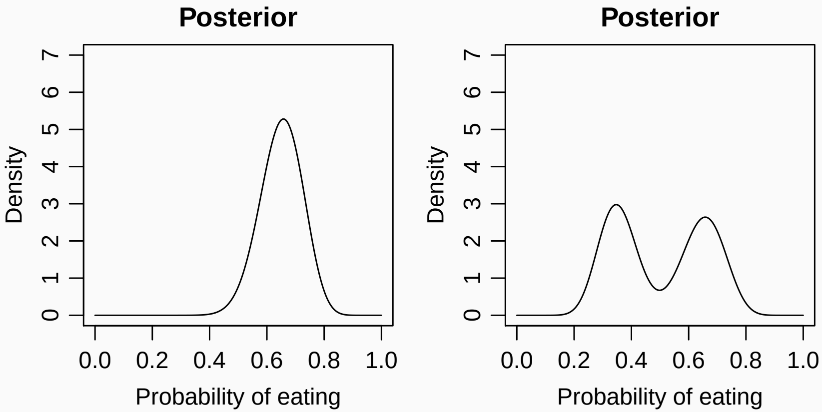
Posterior-based summaries of results

Mean or Median?



Posterior-based summaries of results

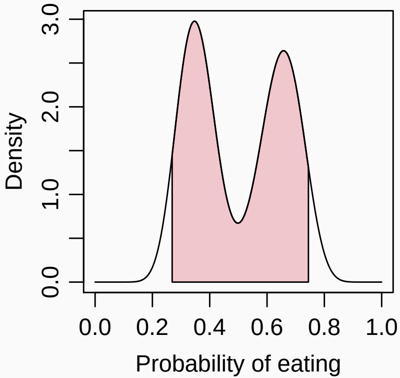
Reporting uncertainty.

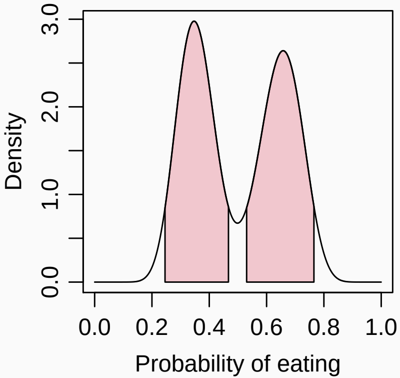


Posterior-based summaries of results

Estimation intervals

- Credible interval.
- Highest density interval.





Questions?