Bayesian Estimation

PSYC 575

September 8, 2020 (updated: 27 September 2020)

Week Learning Objectives

- Describe the benefits of Bayesian estimation methods for MLM
- Use relevant statistical vocabulary to describe Bayesian MLM
  - Prior & posterior distributions
  - Credible intervals
- Assess model convergence from MCMC sampling

Bayesian/MCMC

- A different approach to statistics by treating the parameters as random with a distribution
- Treat probability in a <u>subjective</u> sense
  - As opposed to “frequentist” interpretation of probability

Flipping a Coin

- P(head | fair coin)
- Frequentist
  - Flip it many (e.g., R = 10,000) times, and record the proportion of heads
  - The limit (R → Infinity) is probability
- Bayesian
  - A subjective, usually rational belief of getting a head when the coin is flipped

Probability Distribution

- Quantify uncertainty
- Frequentist
  - Uncertainty only in data and estimates, not parameter
  - Distribution of data, sampling distribution of estimates
- Bayesian
  - Uncertainty in the subjective belief of everything, including the parameter
  - Distribution of data, posterior distribution of parameters

Why Bayesian?

Benefits of Bayesian

1. Not relying on asymptotic results
   - Better small sample properties
2. You can say "95% chance that γ is inside an interval"
3. Easy to get estimates and CIs for derived quantities
   - E.g., R^2, Cohen's *d*, predicted probability (converted from log odds), ICC, etc
4. Handle complex models
   - E.g., with 2+ random slopes
   - Getting more popular as more complex models are proposed

Downsides

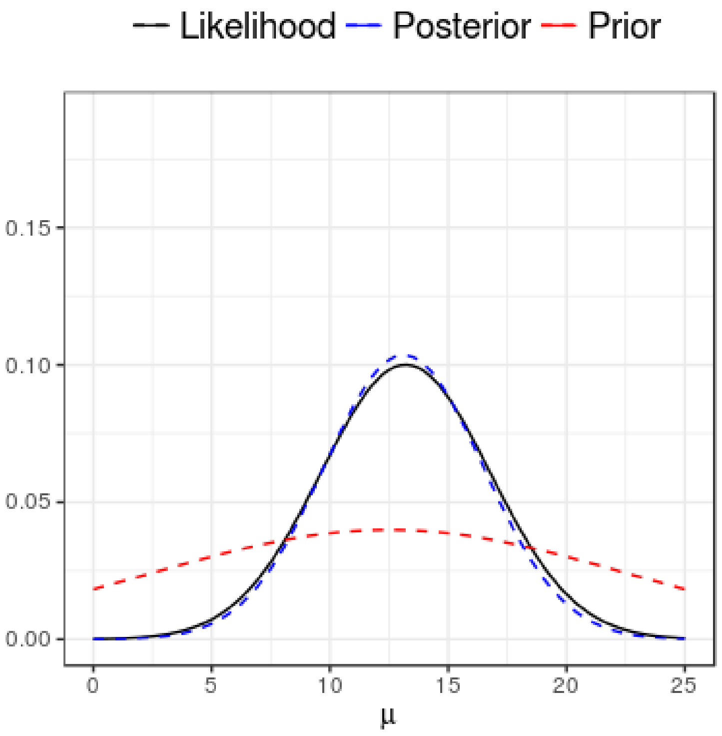
- Computationally intensive
  - However, fitting multiple frequentist models and figuring out how to get things work may end up taking more time
- Need specification of priors
  - For this class we'll mostly use the default in brms

Bayesian/MCMC (Markov Chain Monte Carlo)

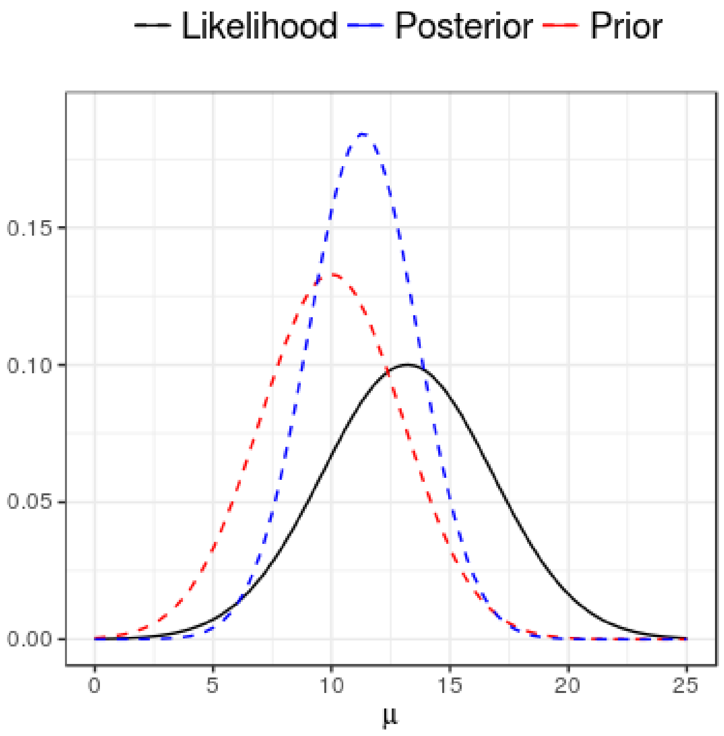
- *Prior distribution*
  - Belief on what values of the parameter is plausible, before looking at the data
- Noninformative prior
- Weakly informative prior
  - Does it make sense that 1 unit of SES is associated with a 20 point difference in MATHACH?
- Informative prior
  - What does previous study say about SES → MATHACH

Bayes's Theorem: $P(\theta|y) \propto P(y \mid \theta)\ P(\theta)$

Weak Prior

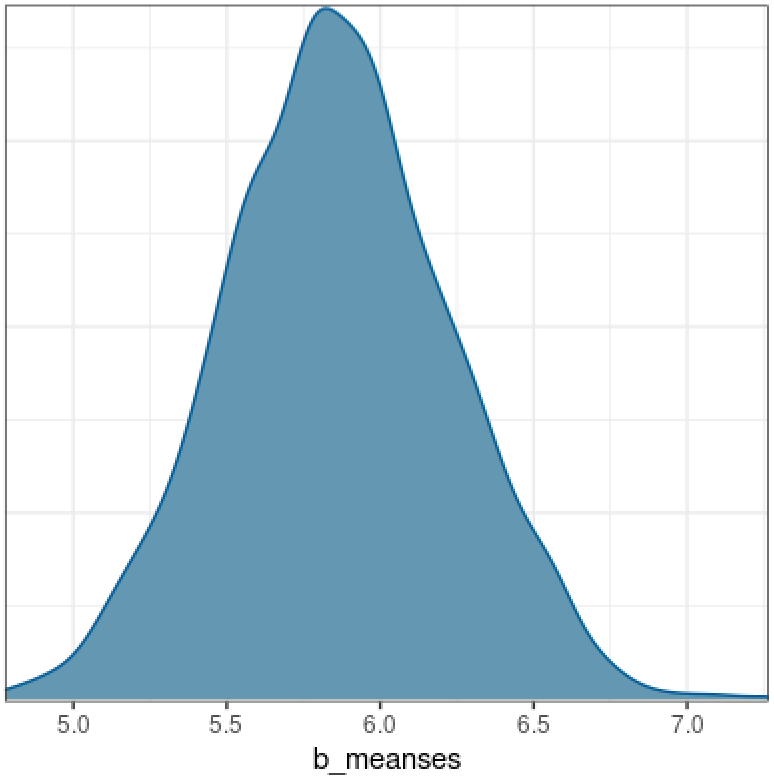


Strong Prior



Bayesian/MCMC

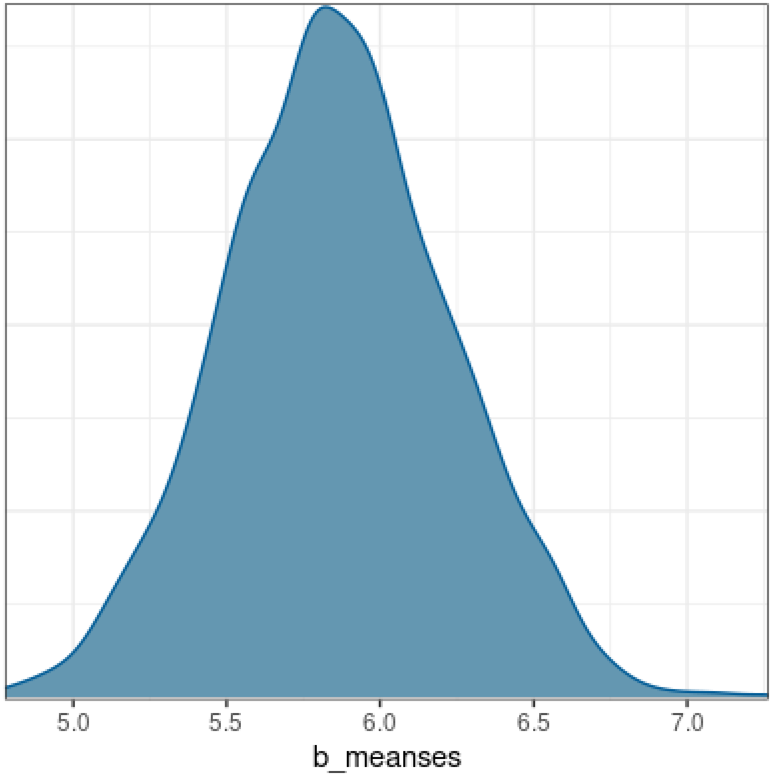
- *Posterior distribution*
  - Updated belief after considering the data
- Bayes's theorem:
  - Posterior ∝ Prior × Likelihood (Data)



Bayesian/MCMC

- Bayes estimate: posterior mean
- Uncertainty estimate:
  - Posterior *SD* (analogous to *SE*)
  - Credible interval (analogous to CI)

There is a 95% chance that the slope of `meanses` is in the interval [5.15, 6.58]



Bayesian/MCMC

- Similar results to ML/REML with simple models with non-informative/vague or weakly informative priors
- More stable estimates and less convergence issues with complex models
- Flexible to fit more complex models
  - Unequal variances
  - Outlier-robust models (Student-*t* model)