# Chapter 3: Conceptual Issues in Bayesian Inference

*Bayesian Psychometric Modeling, Lecture 6*

## Highlights of Chapter

- Concepts that reappear in Bayesian inference
- Features where Bayesian methods differ from maximum likelihood methods

## Issue Number 1: Influence of Prior vs. Data

- Covered in each of the past few lectures
- To see an example, see Lecture 05: JAGS introduction
  - See how inferences change when prior distribution changes
- Key to understanding: As the likelihood function is at the core of Bayesian theorem, properties of the likelihood from ML apply in Bayesian
  - For instance, limiting distributions
    - For continuous parameters, the posterior distribution becomes normally distributed as N approaches infinity
- Some exceptions do apply:
  - Models with priors outside the sample space of the parameter
  - When estimates are in a suboptimal area of the likelihood (for models with multiple modes)

## Issue Number 2: Motivating the Selection of Prior Distributions

- The easy parts to understand
  - Priors should quantify uncertainty about parameters
  - Choosing flexible priors seems better than inflexible ones (e.g., Beta vs. Uniform)
- Harder:
  - Picking priors from previous research
  - Interpretability: How do you choose a good prior for parameters that are difficult overall?
  - Having input from "experts"
    - "Experts" is often an easy way to say "uncertainty is not easily quantifiable"
    - "Experts" are often wrong, especially with respect to determining numeric quantities (see Judgment and Decision Making literature about eliciting human judgement about stochastic processes)
- In between:
  - Choosing priors that lead to easily obtainable posteriors
    - Used to mean picking conjugate priors (Normal-Normal or Beta-Binomial)
    - Now, with software like JAGS or STAN, it is a matter of waiting time

## Issue Number 3: Bayesian vs. ML

- Pro ML: Priors are somewhat (possibly entirely) subjective
  - Counter: Data can overwhelm prior
  - Counter, part deux: Choice of a model is a type of prior present in ML
- Pro Bayes: ML is "hard" – difficult to calculate derivatives for using numerical optimization of likelihood function
  - Counter: Many new Bayesian techniques need same derivatives (or other quantities)
- Remember: Sample size is key to determining many differences

- ML beliefs: Estimators are random, so statements aren't about probabilities of parameters
- Bayesian beliefs: Parameters are random, so statements are more direct
  - Bayes theorem makes inductive reasoning transparently about the parameters
- Pro Bayes: Models are explicit – so more transparent
  - Counter: Not unique to Bayes, probably more about didactic traditions in teaching statistics

## Issue Number 4: Exchangability and Conditional Independence

- Exchangablility: belief that quantities (data or parameters) can have labels erased and be treated the same by the model
  - Shows up in ML and in Bayes
  - de Finetti's theorems (which show conditional independence)
- Conditional independence:

$$p(x_1, x_2) = p(x_1)p(x_2)$$

- Conditional independence is important for building efficient (fast) algorithms
  - Things that are exchangeable are conditionally independent
  - Independent processes can be split computationally (more processors; Hello GPUs)
- Text sidebar (p. 66): The use of a uniform prior $U(-\infty, infty)$ is indeed improper, however, when using computers to derive results, $\infty$ isn't a number.
  - All integer and floating point numbers have a maximum and minimum value – so such a prior would be proper
  - Bigger issue: Belief that uniform priors may be "bad"
    - Yes, some results say they can be bad – often these are extreme cases
- Managing and propagating error: important in both Bayes and ML…can be done easier in Bayes
- Updating results based on new information: Bayes is built for it
  - Today's posterior is tomorrow's prior
  - Probably one of the easier ways to justify the use of a prior

## Conceptulizations of Bayesian Modeling

- Several conceptualizations used in book
  1. Bayesian inference is a belief updating process
     - Prior beliefs are updated when they meet data and become a posterior distribution
  2. Bayesian methods augment information in the data
  3. Joint distribution of data is primary conceptual basis for analysis (consistent with ML)
  4. Model building/expansion