An Introduction to Bayesian Analysis Using Latent Variables

Sarah Hunter

University of Colorado, Boulder

sarah.hunter@colorado.edu

October 19, 2021

Overview

1. Introduction
2. Latent Concepts to Latent Variables
3. Probability and Hypothesis Testing
4. Bayes' Rule
5. Applied Bayesian Analysis
6. Latent Variable Estimation

Learning Objectives

By the end of this workshop, you will learn:

- The purpose of latent variables
- The basic logic of Bayesian Analysis
- How to estimate and interpret Bayesian linear models
- How to estimate and interpret Bayesian Latent Variable models

Latent Concepts

How do we measure complex concepts such as:

- Development
- Democracy
- Ideology
- Gender Equality

Latent Variables

Defined

Latent variables are those concepts that cannot be directly observed, but can be inferred through observable indicators.

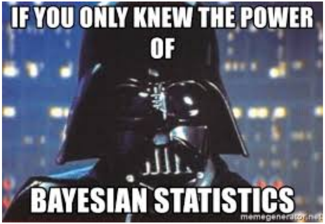
Ways to Estimate Latent Concepts

- Classical Factor Analysis
- Item Response Theory
- Bayesian Latent Variable Models

Ways to Estimate Latent Concepts

- Classical Factor Analysis
- Item Response Theory
- **Bayesian Latent Variable Models**

Why use Bayes?

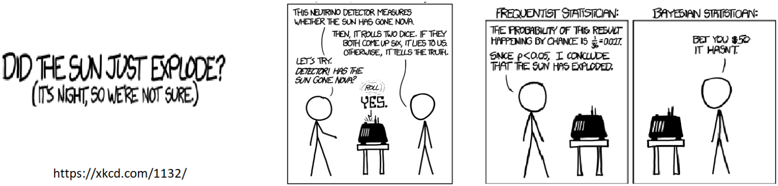


- Ability to include prior findings
- Built-in uncertainty (No Bootstrap required)
- Seamlessly incorporate latent variable into other model, while including uncertainty

A Bayes Crash Course

- Defining Probability
- Hypothesis Testing with Bayes
- Priors and Posteriors
- MCMC

The Difference Between Bayesians and Frequentists



https://xkcd.com/1132/

Defining Probability

Frequentist Definition

Long run frequency of events.

Bayesian Definition

Expectations of events based on subjective beliefs.

Objective Probability and Hypothesis Testing

- Asymptotic Properties and Sampling
- Confidence Intervals and P-values
- Null Hypothesis Significance Testing

Limitations of the Frequentist Approach

- Sample Size
- Populations v. Samples
- Relies on the Central Limit Theorem for inference

Switching the Conditionality

Frequentist Statistics Tell us the Probability of the data, given the parameters

$$L(\theta|Y) = \prod_{i=1}^{n} p(Y_i|\theta)$$

Bayes' Law Tells us the Probability of the Parameters, given the data

$$\pi(\theta|X) = \frac{L(\theta|X) * p(\theta)}{\int p(\theta)L(\theta|X)d\theta}$$

Hypothesis Testing with Bayesian Analysis

The Bayes Mantra

$$\pi(\theta|X) \propto L(\theta|X) * p(\theta)$$

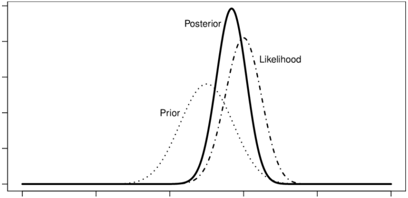


Figure: Source: Etz 2017

Priors

$$\underbrace{\pi(\theta|X)}_{\text{Posterior}} \propto \underbrace{L(\theta|X)}_{\text{Likelihood}} * \underbrace{p(\theta)}_{\text{Prior}}$$

Defined Probability statements about unknown quantities of interest.

- Otherwise known as the prior belief about the probability of the unknown quantity (i.e. parameter)
- A way to include prior research and expectations directly into a model.
- Also necessary for Bayes' Law.

What are Priors in a Model?

The Frequentist

A way for the research to inject subjective beliefs into a regression

The Bayes-or-Bust Bayesian

A way to include prior information AND uncertainty into the model; a way to not pretend ignorance

The Practical Bayesian

A mathematical necessity to do inference without NHST

Types of Priors

Uninformative Priors

- Intentionally add little new information about the unknown parameter.
- Useful when little is known, or to satisfy frequentist criticism of subjectivity

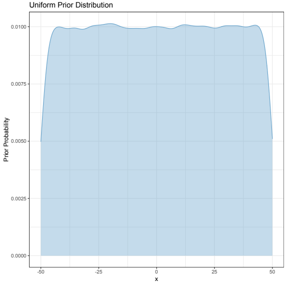
Informative/ Elicited Priors

- Adds information based on previous research directly into the model.
- These priors give more weight to certain values than others.

Uninformative Priors

Any prior distribution can be uninformative, but the usual choice is the **Uniform Prior**

- A prior based on a uniform distribution
- Every value within the range of the uniform distribution has an equal probability.
- Two types:
    - Proper: Integrates to 1
    - Improper: Does not integrate to 1



Informative Priors

Informative Priors (AKA Elicited Priors) fall into several categories:

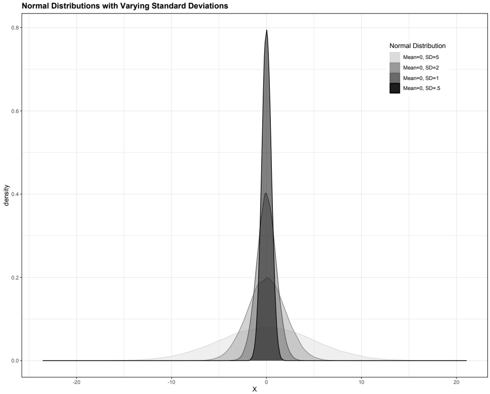
- **Clinical Priors**: use information form experts working on the project
- **Skeptical Priors**: Assumes the hypothesized effect does not exist
- **Enthusiastic Priors**: Opposite of the skeptical prior

Choosing an Appropriate Prior

Keep in mind:

- The function of the prior
- Strong prior knowledge
- Know your audience
- Some priors give weird results
- Some priors make estimation easier
- Some priors actually make estimation impossible
- Publication

Factors that Control the Impact of the Prior



- Sample size
- Strength of relationship being tested
- **Standard Deviation** of the prior
- Prior distribution

The Posterior Distribution

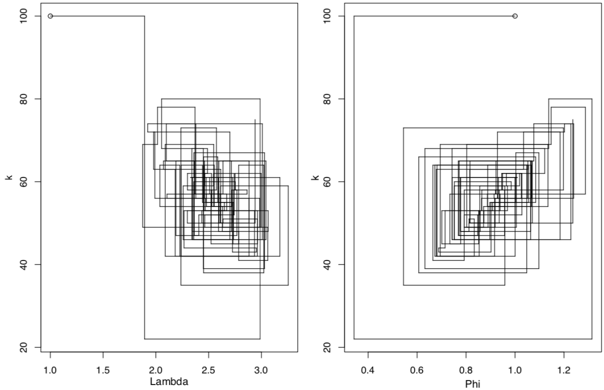
Instead of point estimates of parameters (e.g. β), Bayesian models estimate a **posterior distribution**.

Advantages of Posteriors:

- No Central Limit Theorem
- No assumptions of normality
- Inference
- Inference in small samples

MCMC

Simple Gibbs Sampler Example



Software

Software that does Bayesian Analysis:

- JAGS
- OpenBUGS
- STAN

Software

Software that does Bayesian Analysis:

- **JAGS**
- BUGS
- STAN

An Example: Bayesian Linear Regression

$$y_i = \beta_0 + \beta_1 * x1_i + \beta_2 * x2_i$$

where:

$$y \sim \mathcal{N}(\mu, \tau)$$
$$\beta_0 \sim \mathcal{N}(1, 10)$$
$$\beta_1 \sim \mathcal{N}(1, 10)$$
$$\beta_2 \sim \mathcal{N}(1, 10)$$
$$tau \sim Gamma(.1, .1)$$

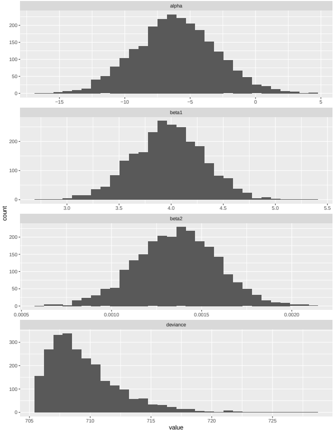
Example

```
jags.model<-function(){
for(i in 1:N){
prestige[i]~dnorm(mu[i], tau)
mu[i]<-alpha + beta1*education[i]
+ beta2*income[i]
}
alpha~dnorm(0, 1)
beta1~dnorm(0, 1)
beta2~dnorm(0, 1)
tau~dgamma(.1,.1)
}
```

Implementing Bayesian Models in JAGS

1. Write your model function.
2. Choose your priors.
3. Convert your data to list format.
4. Choose the parameters you want to monitor.
5. Estimate the model.
6. Conduct convergence diagnostics.

An Example: Bayesian Linear Regression



Example

```
fit<-jags(data=jagsdata,
    inits=NULL,
    params, n.chains=2,
    n.iter=4000, n.burnin=400,
    model.file=
    prestige.model.jags)

print(fit)
```

Interpreting Bayesian Models

	mean	sd	2.5%	97.5%
Constant	-5.84	3.14	-11.93	0.45
Education	4.03	0.34	3.36	4.71
Income	0.001	0.001	0.001	0.002
deviance	709.81	2.82	706.17	716.71

Credible Intervals

95% of the posterior distribution for β_1 falls between 1.136 and 14.043.

You know the shape of the posterior, just plot it.

The Bayesian Latent Variable Model

$$X^*_{ij} = \Lambda_j \phi_i + \epsilon_{ij}$$

where: X^*_{ij} = the observed indicator j for observation i

ϕ_i = the estimate of the latent variable for observation i
Λ_j = the factor loading for observed indicator j
ϵ_{ij} = the errors

Bayesian Latent Variables: An Example

Research Question

How does the mobility of a population influence the allocation of UN Emergency Funds?

Measuring Mobility

Mobility is not observed directly, but the combination of legal restrictions and restrictions caused by infrastructure are our observed indicators.

Bayesian Latent Variables: An Example

Research Question

How does the mobility of a population influence the allocation of UN Emergency Funds?

Measuring Mobility

```
lat.mob<-function(){
for(i in 1:N){
mobility[i]~dnorm(0, 1)

outb.tour[i]~dnorm(mu2[i], tau[1])
mu2[i]<- b[1]*mobility[i]

inb.tour[i]~dnorm(mu3[i], tau[2])
mu3[i]<- b[2]*mobility[i]

ports[i]~dnorm(mu4[i], tau[3])
mu4[i]<- b[3]*mobility[i]

air.pass[i]~dnorm(mu5[i], tau[4])
mu5[i]<- b[4]*mobility[i]

ffm[i]~dnorm(mu6[i], tau[5])
mu6[i]<- b[5]*mobility[i]
}
for (j in 1:3){
b[j]~dnorm(0, .1)}
b[4]~dnorm(0, .1);T(0,)
for(j in 1:5){
tau[j]~dgamma(1, .1)
}
b[5]~dnorm(0, .1)
}
```

Some Warnings

- Identification Restrictions
- Convergence
- Time
- Different tools for different jobs

Further Reading

Armstrong II, David A., Ryan Bakker, Royce Carroll, Christopher Hare, Keith T. Poole, and Howard Rosenthal. 2014 *Analyzing spatial models of choice and judgment with R.* Chapman and Hall/CRC.

Gill, Jeff. 2015. *Bayesian Methods: A Social and Behavioral Science Approach, 3rd ed.*. CRC Press.

The End

Special thanks to Ryan Bakker and Johannes Karreth