Bayesian Statistics
Estimation of a Single Mean and Variance
MCMC Diagnostics and Missing Data

Michael Anderson, PhD
Hélène Carabin, DVM, PhD

Department of Biostatistics and Epidemiology
The University of Oklahoma Health Sciences Center

May 19, 2016

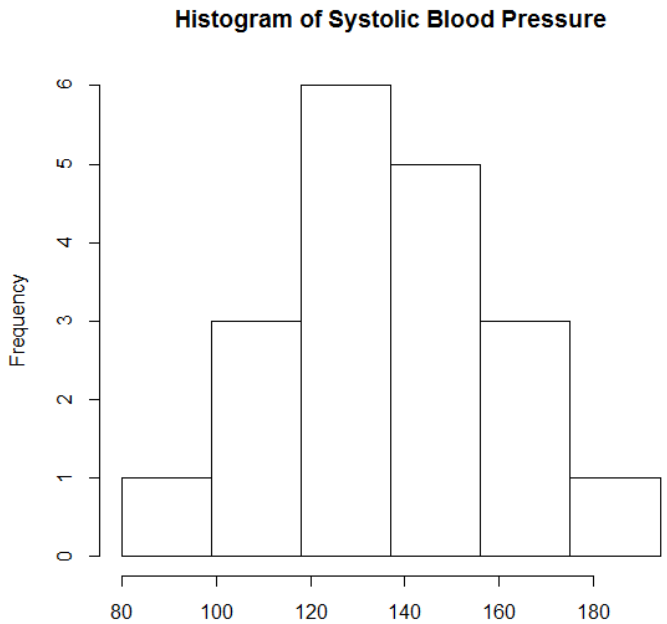
Outline

1. Motivating Example
2. Likelihood and Prior
3. MCMC Diagnostics
4. MCMC Diagnostics
5. Missing Data

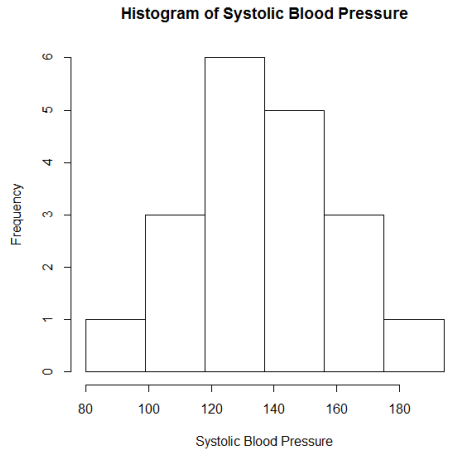
High Blood Pressure Treatment

Suppose a study examines the systolic blood pressure (SBP) of hypertensive subjects (SBP $>$ 140) after 3 months of using blood pressure medication. The SBP for 19 subjects using this medication for 3 months is given below.

```
list(N=19,sbp=c(121,94,119,122,142,168,116,172,155,
     107,180,119,157,101,145,148,120,147,125))
```



High Blood Pressure Treatment



- Q: What seems like a reasonable distribution of the **data**?
  - SBP is a continuous measure.
  - Histogram above shows rough symmetric bell-shaped form.
- A: Normal distribution seems to be a reasonable fit.
  - Shape of Normal is determined by two parameters: μ and σ^2.

We might reasonably conclude $p(x_1, \ldots, x_{19}|\mu, \sigma^2)$ is Normal(μ,σ^2).
We will seek to obtain the posterior $p(\mu, \sigma^2|x_1, \ldots, x_{19})$.
This requires specification of a joint prior, $p(\mu, \sigma^2)$.

Specification of the Likelihood and Prior

$$p(\mu, \sigma^2 | x_1, \ldots, x_n) = \frac{p(x_1, \ldots, x_n | \mu, \sigma^2) p(\mu, \sigma^2)}{\int \int p(x_1, \ldots, x_n | \mu, \sigma^2) p(\mu, \sigma^2) d\mu d\sigma^2}$$

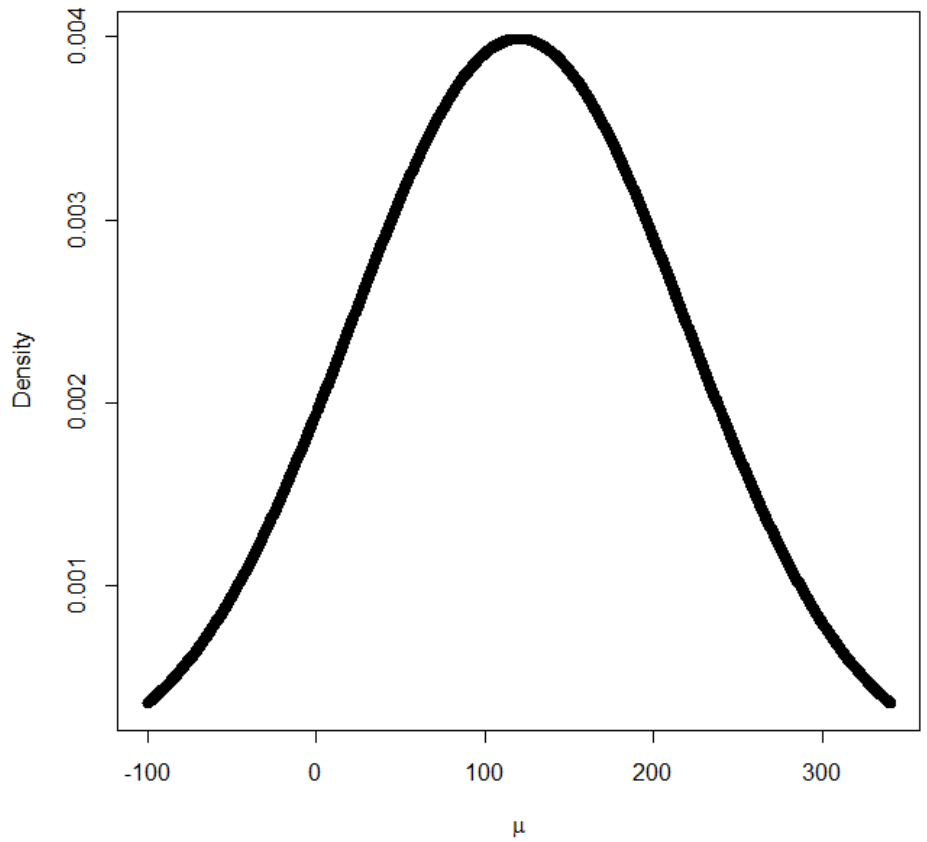
- Note if μ and σ^2 are independent then
  - $p(\mu, \sigma^2) = p(\mu)p(\sigma^2)$.
  - This means we specify a prior for μ and a separate prior for σ^2

$$p(\mu, \sigma^2 | x_1, \ldots, x_n) = \frac{p(x_1, \ldots, x_n | \mu, \sigma^2) p(\mu) p(\sigma^2)}{\int \int p(x_1, \ldots, x_n | \mu, \sigma^2) p(\mu) p(\sigma^2) d\mu d\sigma^2}$$

Specification of the Likelihood and Prior

Q: What makes a reasonable prior for μ?

A: Diffuse prior on μ could be $Normal(120, 100^2)$

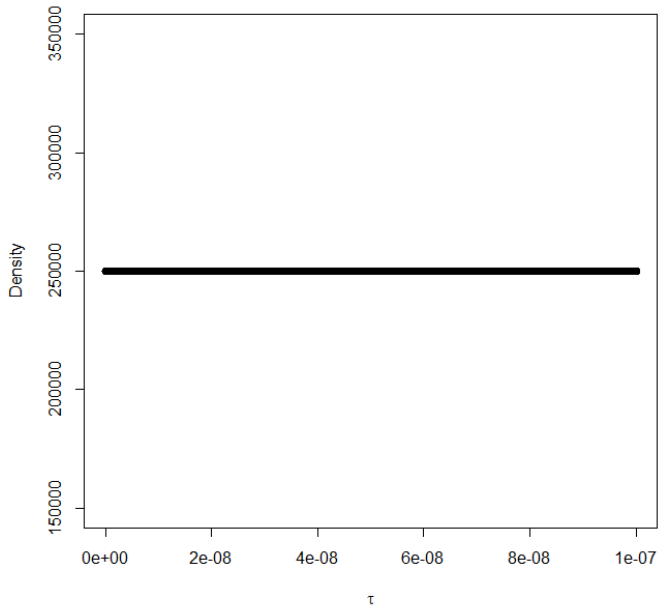


Specification of the Likelihood and Prior

Q: What makes a reasonable prior for σ^2?

A: Diffuse prior on σ could be $Unif(0, 500)$ but another popular option is a Gamma with wide variance.

- Mean of this uniform is 250.
- Variance of this uniform is $500^2/12 = 20833$.
- NOTE: WinBUGS requires precision, τ where $\tau = 1/\sigma^2$.



Specification of the Likelihood and Prior

Putting this altogether we have:

- $p(x_1, \ldots, x_n|\mu, \sigma^2) \sim dnorm(\mu, \tau)$
- $p(\mu) \sim dnorm(120, 0.0001)$.
- $p(\sigma) \sim unif(0, 500)$ and $\tau = 1/\sigma^2$

Let's do this for the Systolic Blood Pressure Example

one mean and variance sbp example

```
model
{

  #likelihood
  for(i in 1:N)
  {
      sbp[i]~dnorm(mu,tau)  #note this uses precision not variance
  }

  #priors

  mu~dnorm(120,0.0001)      #diffuse prior on mean
  sd~dunif(0,500)           #diffuse prior on sd
  tau<-1/(sd*sd)            #compute precision from sd
  var<-sd*sd
}

#Data
list(N=19,sbp=c(121,94,119,122,142,168,116,172,155,107,180,119,
157,101,145,148,120,147,125))

#Initial Values
list(mu=50,sd=20)
```

In Class Practice Problems

See file named "Systolic Blood Pressure Example.odc"

MCMC Diagnostics

There are a few diagnostic tools in WinBUGS to assess posterior samples that have been drawn.

- History (sequential posterior samples).
- Trace (similar to history but can drill down to fine samples).
- bgr diag (Also known as Brooks-Gelman-Rubin diagnostic).
- auto corr (Checks the correlation among posterior samples).

There are additional diagnostic tools but they require the use of other software (R) to implement.
We will focus on those "built in" to WinBUGS.
The trick here will be to run 2 or more chains and see if they get to the same place.

MCMC Diagnostics

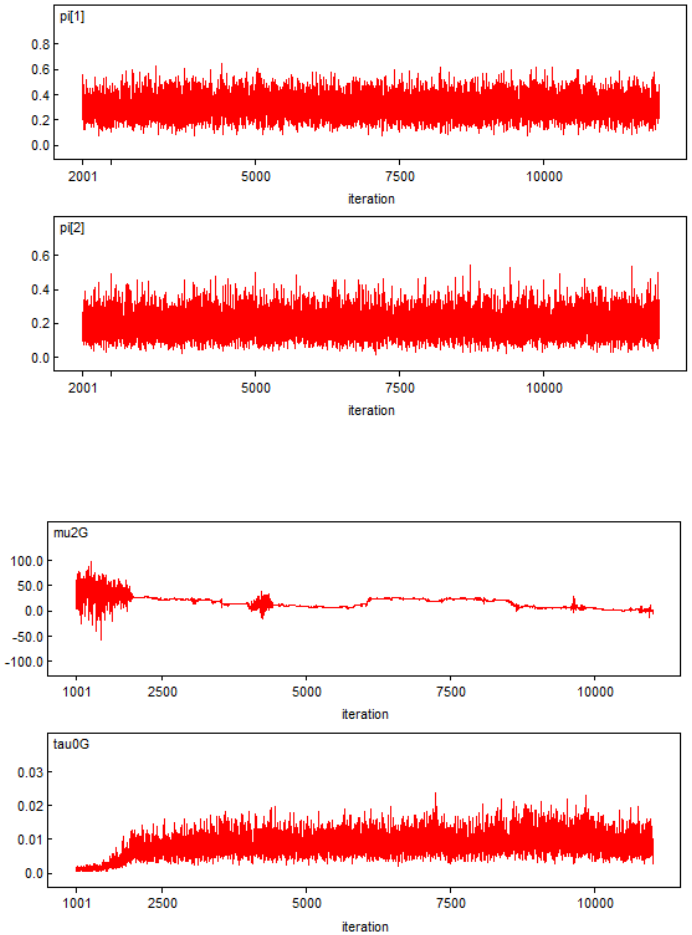
Here are a few terms that will be helpful when discussing MCMC diagnostics

- Thinning-utilizing fewer of posterior samples for analysis in a systematic way.
- Chain length-the number of posterior samples requested for MCMC.
- Burn-in-The walk the MCMC chain takes prior to arriving at the true posterior.

Occasionally, we will use these, separately or in combination, to "fix" an markov chain obtained through Gibbs sampling.

MCMC Diagnostics

Trace Plots: `History` in WinBugs
Patterns are bad.

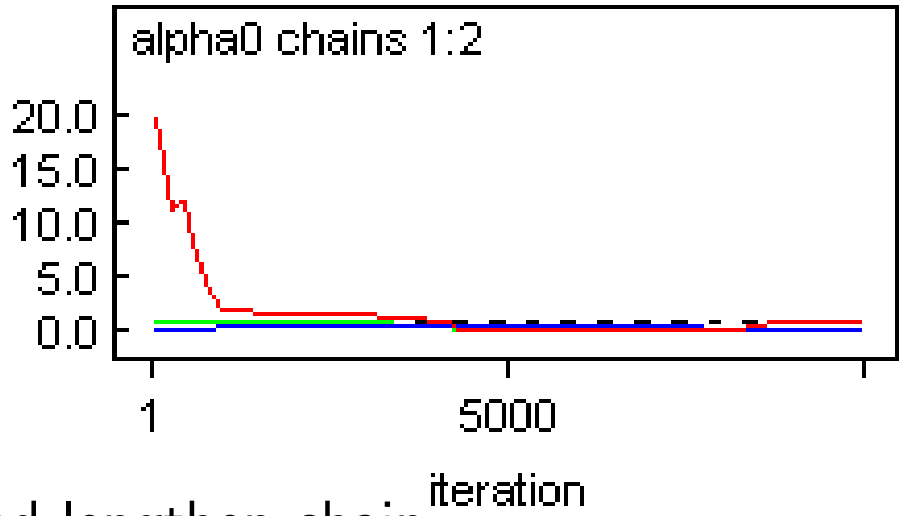


- Increase burn-in period and lengthen chain.

MCMC Diagnostics

Brooks-Gelman-Rubin: `bgr plot in WinBugs`

- Each chain is subset into overlapping sets.
- For each set, an average width, W, of $100(1-\alpha)$ intervals is computed.
- Between chain interval widths, B, are computed.
- The ratio $\hat{R}$=B/W is computed.
- See when this ratio converges to 1 (should start larger than 1).



- Increase burn-in and lengthen chain.

MCMC Diagnostics

auto corr:

- Consecutive Gibbs samples will be correlated.
- Too much auto correlation is bad.
- MC standard error reflects accuracy of Monte Carlo process to estimate true posterior *mean* with dependent samples.
- Increase thinning and lengthen chain.

MCMC Diagnostics for SBP Example

one mean and variance sbp example

```
model
{

  #likelihood
  for(i in 1:N)
  {
      sbp[i]~dnorm(mu,tau)  #note this uses precision not variance
  }

  #priors

  mu~dnorm(120,0.0001)      #diffuse prior on mean
  sd~dunif(0,500)           #diffuse prior on sd
  tau<-1/(sd*sd)            #compute precision from sd
  var<-sd*sd
}

#Data
list(N=19,sbp=c(121,94,119,122,142,168,116,172,155,107,180,119,
157,101,145,148,120,147,125))

#Initial Values
#chain 1
list(mu=0,sd=5)

#chain 2
list(mu=200,sd=400)
```

MCMC Diagnostics for SBP Example

See file named "Systolic Blood Pressure Diagnostics Example.odc"

Dealing with Missing Data the Bayesian Way

Missing data are common in practice and there are many alternatives for handling it.
A Bayesian perspective would view missing data in the same way it views unknown parameters.

- Just need to specify the joint model for the missing and observed data and model parameters.
- MCMC can be used to generate a predicted value for the missing data in the usual way.
- The reason for the missingness (mechanism) will dictate the appropriateness of the joint model.

Dealing with Missing Data the Bayesian Way

Three missing data mechanisms and how to handle them in WinBUGS are outlined below

- Missing Completely At Random (MCAR)-Probability of missingness does not depend on the observed or unobserved quantities.
  - Do nothing, just be sure the data value is NA.
  - WinBUGS will generate a predicted value from the posterior.
  - Missing data mechanism is assumed to be *ignorable*.
- Missing At Random (MAR)-Probability for the missingness depends only on the observed data.
  - Do nothing, just be sure the data value is NA.
  - WinBUGS will generate a predicted value from the posterior.
  - Missing data mechanism is assumed to be *ignorable*.
- Missing Not At Random (MNAR)-Neither MCAR or MAR hold.
  - Model the missing data from the observed and prior knowledge.
  - Need to specify additional likelihood and prior terms for missing data.
  - Missing data mechanism is assumed to be *informative*.

Missing Data for SBP Example

See file named "Systolic Blood Pressure Missing Data Example.odc"