

# Inference from MCMC

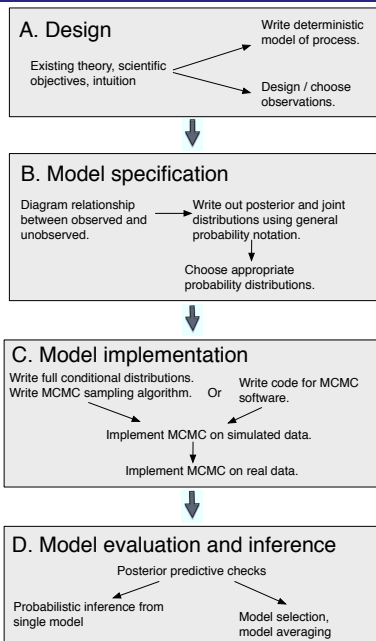
## ESS 575 Models for Ecological Data

N. Thompson Hobbs

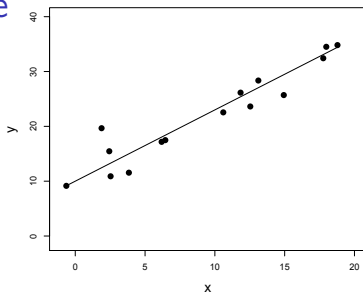
February 28, 2019



# The Bayesian method



## A simple example



$$g(\boldsymbol{\beta}, x_i) = \beta_0 + \beta_1 x_i$$

$$\begin{aligned} [\boldsymbol{\beta}, \sigma^2 | \mathbf{y}] &\propto \prod_{i=1}^{15} \text{normal}(y_i | g(\boldsymbol{\beta}, x_i), \sigma^2) \text{normal}(\beta_0 | 0, 1000) \times \\ &\quad \text{normal}(\beta_1 | 0, 1000) \times \\ &\quad \text{inverse gamma}(\sigma^2 | .001, .001) \end{aligned}$$

The same principles apply to models with any number of parameters and latent states.

# Predictions

We want to know the distribution of the mean of the response when the predictor variable equals  $x_4$

$$\mu_4 = \beta_0 + \beta_1 x_4.$$

We also want to know the distribution of a new observation at  $x_4$

$$y_4^{new} \sim \text{normal}(\mu_4, \sigma^2).$$

# Output from JAGS

```
> summary(jc)
```

```
Iterations = 35001:45000  
Thinning interval = 1  
Number of chains = 1  
Sample size per chain = 10000
```

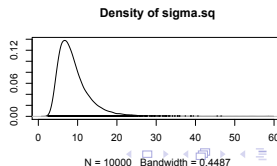
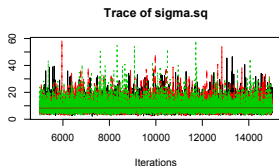
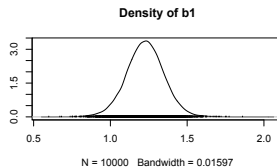
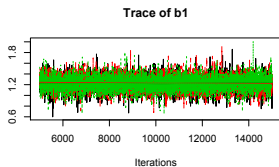
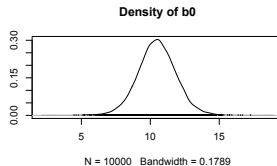
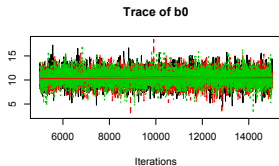
1. Empirical mean and standard deviation for each variable,  
plus standard error of the mean:

	Mean	SD	Naive SE	Time-series SE
b0	10.486	1.3992	0.013992	0.032676
b1	1.227	0.1239	0.001239	0.002812
sigma.sq	9.065	4.4158	0.044158	0.055806

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
b0	7.6559	9.607	10.489	11.371	13.264
b1	0.9832	1.148	1.226	1.305	1.476
sigma.sq	3.9819	6.206	8.054	10.666	20.128

# Output from JAGS



## Output from JAGS

```
> y.new[,4]
      2.5%      50%      97.5%
7.146763 13.618321 19.975031
> mu[,4]
      2.5%      50%      97.5%
11.30646 13.60504 15.92383
>
```

## Where do these come from?

The marginal distribution of  $\beta_0$  is

$$[\beta_0|\mathbf{y}] = \int_{\beta_1} \int_{\sigma^2} [\beta_0, \beta_1, \sigma^2|\mathbf{y}] d\beta_1 d\sigma^2. \quad (1)$$

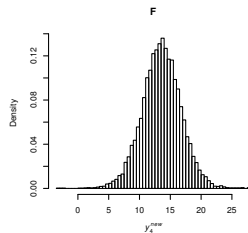
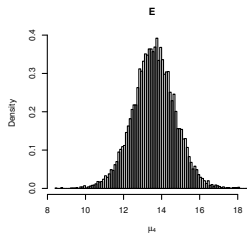
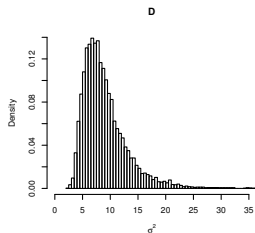
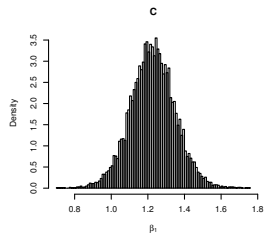
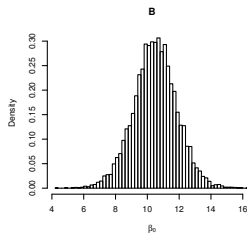
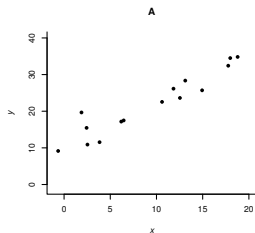
The marginalization property of MCMC allows us to estimate it as follows.



# MCMC output

$k$	5001	5002	5003	5004	5005	...	9996	9997	9998	9999	10000
$\beta_0$	9.84	10.9	10.6	10.7	11.9	...	12.5	9.84	11.1	10.9	11.7
$\beta_1$	1.38	1.22	1.25	1.12	1.14	...	1.04	1.23	1.27	1.06	1.14
$\sigma^2$	10.8	6.32	4.96	4.57	5.76	...	12.4	9.77	8.19	6.88	13.1
$\mu_4$	13.3	14	13.8	13.6	14.7	...	15.1	12.9	14.3	13.6	14.6
$y_4^{new}$	9.29	14.3	9.44	15.3	18.5	...	11.4	16.8	12.8	14.5	15.8

# Marginal distributions



## Monte Carlo Integration

We can also use the row for  $\beta_0$  to approximate moments of its marginal posterior distribution. For example, the mean is given analytically by the integral

$$E(\beta_o | \mathbf{y}) = \int_{\beta_o} \beta_0 [\beta_0 | \mathbf{y}] d\beta_0, \quad (2)$$

which is approximated, simply enough, using

$$E(\beta_0 | \mathbf{y}) \approx \frac{1}{K} \sum_{k=1}^K \beta_0^{(k)}. \quad (3)$$

Similarly, the variance is  $\text{var}(\beta_0 | \mathbf{y}) \approx \frac{\sum_{k=1}^K \left( \beta_0^{(k)} - \frac{1}{K} \sum_{k=1}^K \beta_0^{(k)} \right)^2}{K}$ , the MCMC sample variance of the elements of the  $\beta_0$  row. We can obtain other statistics of interest (e.g., medians, coefficients of variation, quantiles, highest posterior density intervals) by applying the appropriate function to the row.

## Bayesian credible intervals

- ▶ Equal tailed: The upper and lower quantiles, conventionally .025 and .975. Fine for symmetric marginal posteriors distributions.
- ▶ Highest posterior density: Best for asymmetric or multimodal distributions or when you want the credible interval to be able to contain zero. Probably best to use all the time.
- ▶ Illustrate on board.

Median height of willows in dammed and fenced plots was 184 cm (highest posterior density interval, HPDI = 173, 192)

Median height of willows in dammed and fenced plots was 184 cm (equal-tailed Bayesian Credible Interval, BCI = 174, 194)

## Predicting a new observation

We need to integrate

$$[y_4^{new}|\mathbf{y}] = \int_{\beta_0} \int_{\beta_1} \int_{\sigma^2} [y_4^{new}|\beta_0, \beta_1, \sigma^2] [\beta_0, \beta_1, \sigma^2|\mathbf{y}] d\beta_0 d\beta_1 d\sigma^2,$$

which is called the *posterior predictive distribution* of  $y_4$ . To approximate this formidable integral, we make a draw from

$$y_4^{new(k)} \sim \text{normal}(\beta_0^{(k)} + \beta_1^{(k)} x_4, \sigma^{2(k)})$$

at each MCMC iteration. The posterior predictive distribution of  $y^{new}$  is approximated by the elements in its row, a process called *composition sampling*. This is the basis for posterior predictive checks, a topic we will cover soon.

## Predicting the mean at a given $x$

The posterior predictive distribution of the mean:

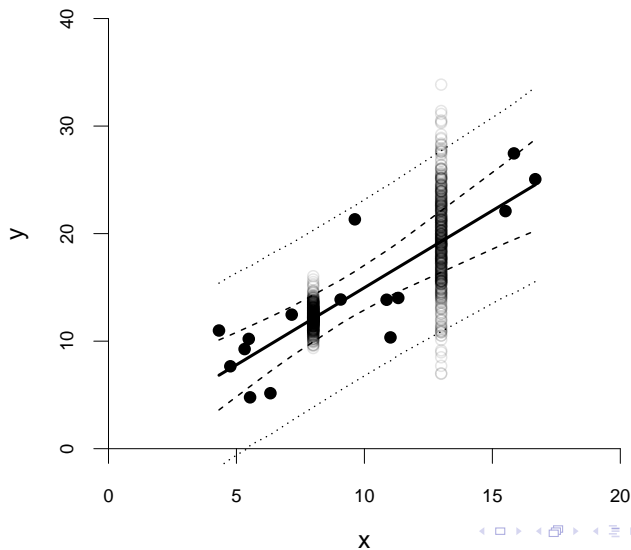
$$[\mu_4|\mathbf{y}] = \int_{\beta_0} \int_{\beta_1} \int_{\sigma^2} \left[ E(y_4^{new(k)}) | \beta_0, \beta_1, \sigma^2 \right] [\beta_0, \beta_1, \sigma^2 | \mathbf{y}] d\beta_0 d\beta_1 d\sigma^2,$$

which we approximate by calculating

$$\mu_4^{(k)} = \beta_0^{(k)} + \beta_1^{(k)} x_4 \quad (4)$$

at each iteration of the MCMC algorithm. Statistics can be calculated from the  $\mu_4$  row.

## Predictions of $y_i^{new}$ and $E(y_i^{new})$



## Derived quantities

The *equivariance* property of MCMC means that any quantity calculated from random variables becomes a random variable with its own posterior distribution. We simply calculate its value at each iteration in the chain based on the current values of other parameters at that iteration. This allows inference on any numeric function of parameters or latent states, for example,

1. Differences between means
2. Ratios of parameters or means
3. Forecasts in time series models
4. Eigen analysis
5. Indices (e.g., Shannon diversity index)

This can be done within JAGS code, as you will learn, or it can be done after the analysis by sampling from the CODA object.



# Main points

- ▶ The marginal distribution of a single unknown quantity is a multi-dimensional integral over the other unknowns.
- ▶ We can approximate these distributions and summarize them using the MCMC output in a process called Monte Carlo integration.