# Writing Bayesian Hierarchical Models

ESS 575 Models for Ecological Data

N. Thompson Hobbs

March 7, 2019

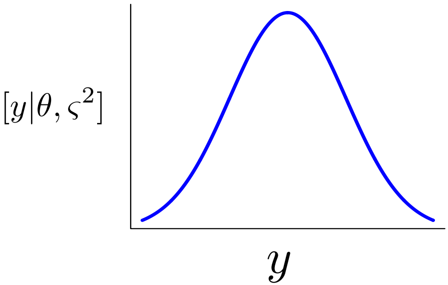
Graduate Degree Program in Ecology
Department of Ecosystem Science and Sustainability
Colorado State University®

# Housekeeping

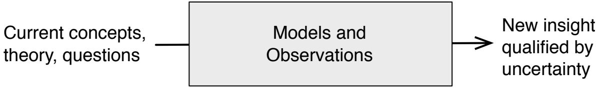
- Debrief MCMC lab
- Model Building exercise and answers is in `ESS_575_2019/Labs/Lab7ModelBuilding/`.Due on March 22

# Data simulation

$$\widetilde{\mathbf{y}} \longrightarrow \text{MCMC} \longrightarrow [\theta|\widetilde{\mathbf{y}}],\ [\varsigma^2|\widetilde{\mathbf{y}}]$$



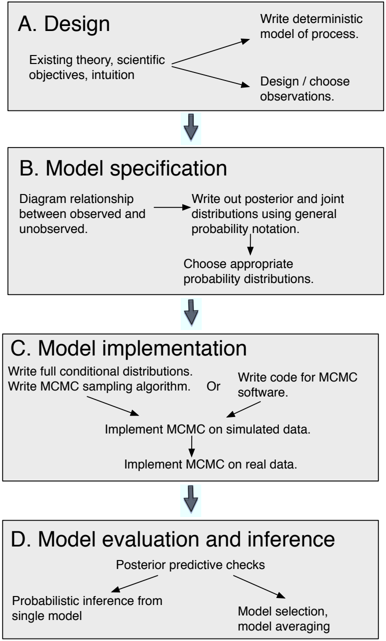
# What is this course about?



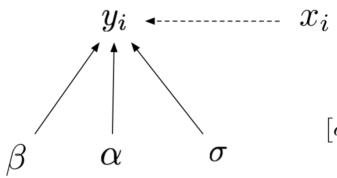
# All modeling problems have idiosyncrasies

- Different types of data
- Different deterministic models
- Sampling error in the predictors or responses
- Calibration error for predictors or responses
- Prior knowledge of parameters
- Missing data
- Multiple scales of data (group level effects)
- Prediction and forecasting
- Spatial or temporal dependence
- Derived quantities

# The Bayesian method



# Cross cutting theme




$$\mu_i = \frac{m x_i^a}{h^a + x_i^a}$$

$$[a, h, m, \sigma^2 \mid y] \propto \prod_{i=1}^{n} [y_i|\mu_i, \sigma^2][a][h][m][\sigma^2]$$

```
model{

     for(i in 1:length(y)){

          mu[i] <- (m*x[i]^a)/(h^a+x[i]^a)
          y[i] ~ dgamma(mu[i]^2/sigma^2,mu[i]/sigma^2)

     }

a ~ dnorm(0,.0001)
m ~ dgamma(.01,.01)
h ~ dgamma(.01,.01)
sigma ~ dunif(0,5)
}
```

# Things to watch for today

- Multi-level models (aka random effects)
- Sampling error in x's and y's
- Calibration error in y's
- Derived quantities (in JAGS lab)

All of these will appear in the exercises.

# Things to watch for today

Partitioning uncertainty

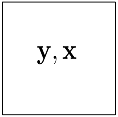
- Process variance
- Sampling variance
- Calibration variance (aka observation variance)
- Group level variance

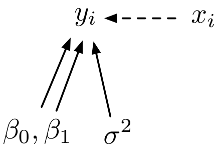
# Steps in writing Bayesian models

1. Write your deterministic model. Be careful about support.
2. Draw Bayesian network (DAG) describing relationships between observed and unobserved quantities.
3. Use the Bayesian network to write proportionality between posterior and joint distributions using bracket notation $[\,|\,]$.
   3.1 Posterior distribution: $[\text{unobserved quantities} \,|\, \text{data}]$
   3.2 Joint distribution
      3.2.1 All nodes in Bayesian network at the heads of arrows (children) must be on the left hand side of a conditioning symbol.
      3.2.2 All nodes in Bayesian network at the tails of arrows (parents) must be on the right hand side of a conditioning symbol $|$.
      3.2.3 All nodes at the end of an arrow with no arrow coming into them must be expressed unconditionally, i.e., they must have numeric arguments.
4. Assign specific PDF or PMF to each of the brackets.
5. Choose numeric values for parameters of prior distributions. Do this sensibly! Do not default to vague priors. (Do as I say, not as I do.)

# The problem we treated with simple Bayesian regression

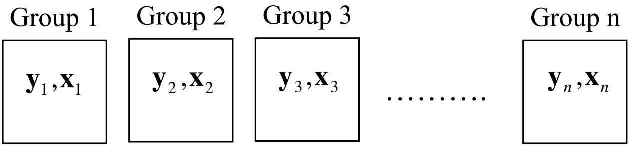
A single "group"



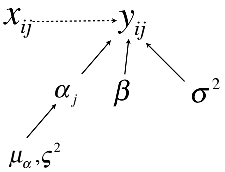




$$
\begin{aligned}
g(\beta_0, \beta_1, x_i) &= \beta_0 + \beta_1 x_i \\
[\beta_0, \beta_1, \sigma^2 \mid y_i] &\propto [\beta_0, \beta_1, \sigma^2, y_i] \\
\text{factoring rhs using DAG:} & \\
[\beta_0, \beta_1, \sigma^2 \mid y_i] &\propto [y_i \mid g(\beta_0, \beta_1, x_i), \sigma^2][\beta_0], [\beta_1][\sigma^2] \\
\text{joint for all data :} & \\
[\beta_0, \beta_1, \sigma^2 \mid \boldsymbol{y}] &\propto \prod_{i=1}^{n} [y_i \mid g(\beta_0, \beta_1, x_i), \sigma^2][\beta_0][\beta_1][\sigma^2] \\
\text{choose specific distributions:} & \\
[\beta_0, \beta_1, \sigma^2 \mid \boldsymbol{y}] &\propto \prod_{i=1}^{n} \text{normal}(y_i \mid g(\beta_0, \beta_1, x_i), \sigma^2) \\
&\times \text{normal}(\beta_0 \mid 0, 10000)\text{normal}(\beta_1 \mid 0, 10000) \\
&\times \text{uniform}(\sigma^2 \mid 0, 500)
\end{aligned}
$$

# The problem

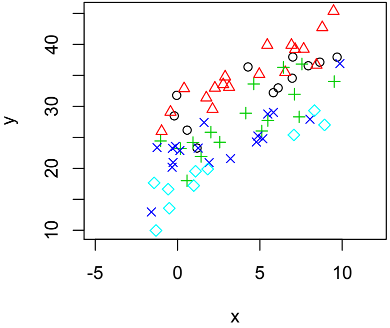

Group 1
Group 2
Group 3
Group n
$\mathbf{y}_1, \mathbf{x}_1$
$\mathbf{y}_2, \mathbf{x}_2$
$\mathbf{y}_3, \mathbf{x}_3$
..........
$\mathbf{y}_n, \mathbf{x}_n$

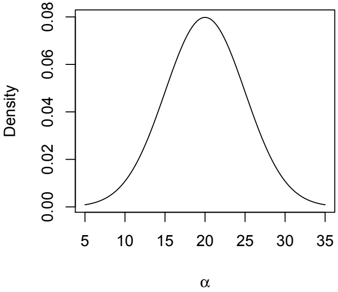
# We can model the intercept (or slope):




$$
\begin{aligned}
&[\beta, \boldsymbol{\alpha}, \sigma^2, \mu_\alpha, \varsigma^2, |\mathbf{y}] \propto \prod_{i=1}^{n_j} \prod_{j=1}^{J} \text{normal}\left(y_{ij} | \alpha_j + \beta x_{ij}, \sigma^2\right) \\
&\times \text{normal}\left(\alpha_j | \mu_\alpha, \varsigma^2\right) \\
&\times \text{normal}\left(\beta | 0, 10000\right) \text{normal}\left(\mu_\alpha | 0, 1000\right) \\
&\times \text{inverse gamma}\left(\sigma^2 | .001, .001\right) \text{inverse gamma}\left(\varsigma^2 | .001, .001\right)
\end{aligned}
$$

# We seek to understand the distribution of intercepts.



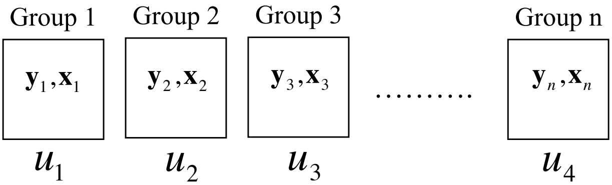
# Some notation

$$\mu_{ij} = \beta_0 + \beta_1 x_{ij} + \alpha_j$$
$$y_{ij} \sim \text{normal}(\mu_{ij}, \sigma^2)$$
$$\alpha_j \sim \text{normal}(0, \varsigma^2)$$

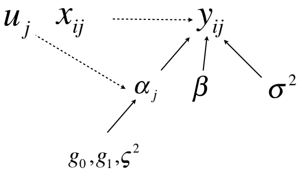
is identical to:

$$\mu_{ij} = \alpha_j + \beta_1 x_{ij}$$
$$y_{ij} \sim \text{normal}(\mu_{ij}, \sigma^2)$$
$$\alpha_j \sim (\mu_\alpha, \varsigma^2)$$

# Include data on groups.



# We can model the intercept (or slope) as a function of group level data:




$$
\begin{aligned}
\left[\boldsymbol{\alpha}, \beta, \boldsymbol{\sigma}^2, \mathbf{g}, \varsigma^2, |\mathbf{y}\right] \propto & \prod_{i=1}^{n_j} \prod_{j=1}^{J} \text{normal}\left(y_{ij} | \boldsymbol{\alpha}_j + \boldsymbol{\beta} x_{ij}, \boldsymbol{\sigma}^2\right) \\
& \times \text{normal}\left(\boldsymbol{\alpha}_j | g_0 + g_1 u_j, \boldsymbol{\varsigma}^2\right) \\
& \times \text{normal}\left(\boldsymbol{\beta} | 0, .001\right) \text{normal}\left(g_0 | 0, 1000\right) \text{normal}\left(g_1 | 0, 1000\right) \\
& \times \text{inverse gamma}\left(\boldsymbol{\sigma}^2 | .001, .001\right) \text{inverse gamma}\left(\boldsymbol{\varsigma}^2 | .001, .001\right)
\end{aligned}
$$

Board work on light limitation of trees for errors in x's and y's