Bayesian methods for ecological and environmental modelling

Trainers:

Lindsay Banin, David Cameron, Pete Henrys & Peter Levy

UK Centre for Ecology & Hydrology

Hierarchical modelling

Part 1: Introducing the ‘whys’ and ‘hows’

Lindsay Banin

UK Centre for Ecology & Hydrology

What we will cover for the remainder of today: Hierarchical Modelling

Session 5a

What are hierarchical models and why do we use them?

Extending our learning from simple linear model to hierarchical model (rstanarm)

Practical part 1

Session 5b

Considering different data types and their distributions

Practical part 2

Tomorrow - Session 6a

Visualising and specifying our models

Implementing our model using MCMC (JAGS and R – rjags, coda)

Practical part 3

What we will cover in Session 5a

- Introduction to Hierarchical Modelling - What are hierarchical models and why do we use them?
- Framework for conducting Bayesian analyses
- Extending our learning from simple linear model to hierarchical model (rstanarm)
- Practical part 1

What is a hierarchical model?

Motivation for using hierarchical Bayes…

Why Bayes?

- All scientific models are abstractions

- Because models are abstractions and reduce the dimensions of a problem, we must deal with the elements we choose to leave out

- so assessing uncertainty is fundamental to science:
    - “process uncertainty”
    - “observation uncertainty”

Why hierarchical models?

- Allow us to decompose complex, high dimensional problems into parts that can be thought about and analysed individually

- Broad and flexible approach, allowing us to tackle virtually any ecological problem

Motivation for using hierarchical Bayes…

"Hierarchical Bayes transformed computational statistics in the 1990s providing **a framework that can accommodate nearly all high-dimensional problems** (Gelfand & Smith 1990; Carlin & Louis 2000). The structure is **so flexible** that it not only opens doors to **complex problems informed by messy data.** The intuitive approach facilitates a deeper understanding of the processes and the **importance of treating unknown elements** in appropriate ways." Clark (2005)

Ecological research commonly has to deal with issues such as:

- <u>Variation among individuals</u> (e.g., location or genotype; repeated measures and longitudinal studies)

- Ecological processes operating at <u>more than one spatial scale</u> (plot → habitat type) or <u>level of ecological organisation</u> (individuals → populations → communities)

- The need to accommodate <u>uncertainty arising from *modelling a process* as well as *uncertainty derived from imperfect observations*</u>

- Dealing with <u>changes in state that cannot be observed directly</u> (age transitions of individuals that are hard to observe)

- Compiling <u>datasets from different sources</u> (e.g. different research teams, different methods)

Hierarchical models provide a natural framework for addressing all of these issues. It provides a very flexible approach

Ecological variation – levels of biological organisation

Individuals

Populations

Communities and Ecosystems

Spatial and temporal heterogeneity

Scale

Ecological variation – spatial scales

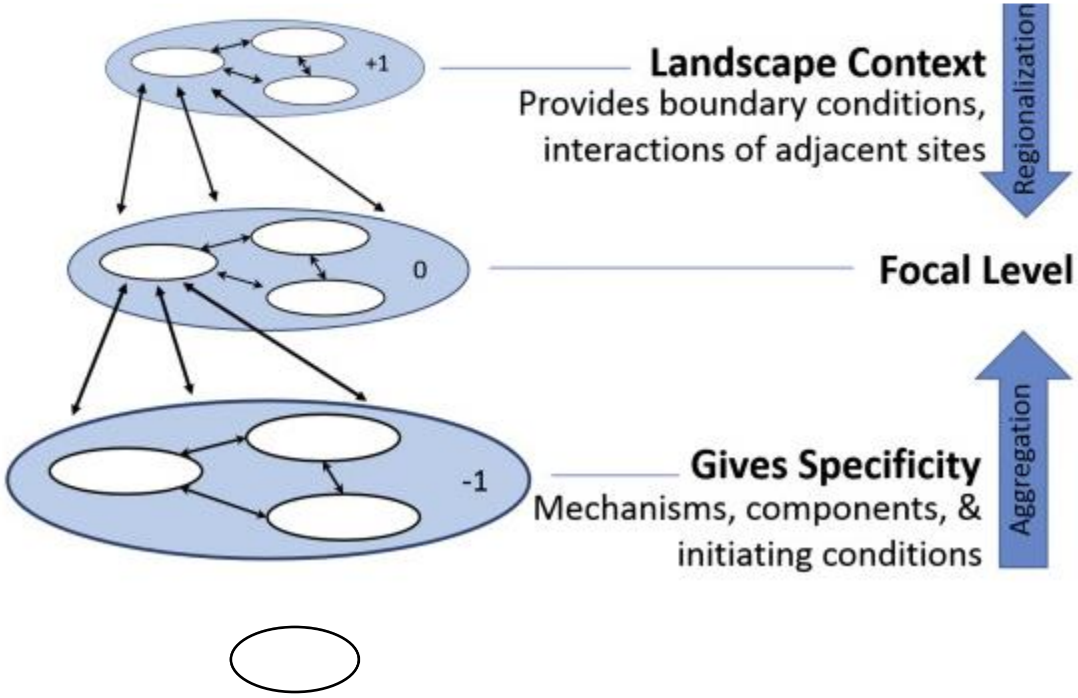


Figure 1. Scales of organization showing feedback and feedforwards from the immediate higher and lower scale in a hierarchical system. Adapted from Urban et al.[8] and O'Neill et al.[10]

Landscape

Habitat type

Plots

Individual plants

Adapted from Salley et al. 2016. Rangelands 38(6):313—317 doi: 10.1016/j.rala.2016.10.003

Motivation for using hierarchical Bayes…

Clark (2005) Ecol. Letts.

Structure in space, time, among individuals/groups

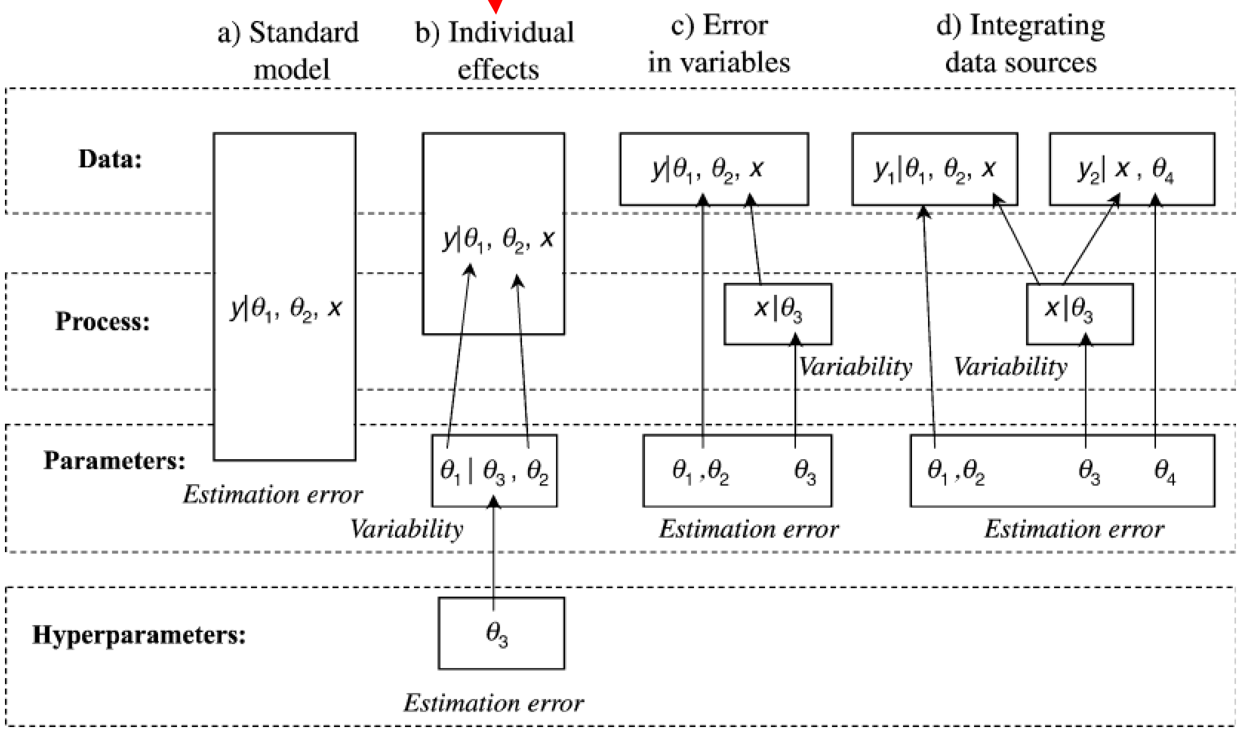


Figure 1 Four examples of how the Bayesian framework admits complexity (see text). Models can be viewed as networks of components, some of which are known and many unknown. The stages shown here, Data, Process, Parameters, and Hyperparameters, represent an overarching structure that admits complex networks. A model might include structure in space, time, or among individuals or groups (b), hidden processes (c,d), and multiple sources of information that bear on the same process (d). Acknowledging variability in a 'parameter' θ_1 (b) is accomplished by conditioning on additional parameters (θ_3). Now θ_1 occupies a middle stage and is truly variable, not just uncertain; θ_3 is asymptotic. Acknowledging variability in a predictor variable x (c) is accomplished in the same fashion.

A hierarchical model with grouping structure

"An analysis that **disregards between-group heterogeneity can yield parameter estimates that are wrong** if there is between-group heterogeneity but would be relatively precise if there actually were no between-group heterogeneity.

Group-by-group analyses, on the other hand, are valid but produce estimates that are relatively imprecise. While complete pooling or no pooling of data across groups is sometimes called for, **models that ignore the grouping structures in the data tend to underfit or overfit** (Gelman et al., 2013).

Hierarchical modeling provides a compromise by **allowing parameters to vary by group at lower levels of the hierarchy while estimating common parameters at higher levels.** Inference for each group-level parameter is informed not only by the group-specific information contained in the data but also by the data for other groups as well. This is commonly referred to as *borrowing strength* or *shrinkage*."

Gabry & Goodrich (2021) Vignette: Estimating Generalized (Non-)Linear Models with Group-Specific Terms with rstanarm; https://mc-stan.org/rstanarm/articles/glmer.html

A hierarchical model with grouping structure

Steve Midway online book: https://bookdown.org/steve_midway/BHME/Ch1.html#hierarchical-models

Hierarchical model:
"partial pooling"

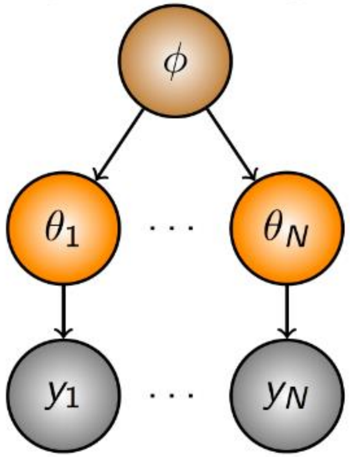


Figure 2.5: Diagrammatic representation partial pooling, a model structure in which different observations inform different latent parameters, which are then governed by additional parameters.

Identical parameters:
"complete pooling"

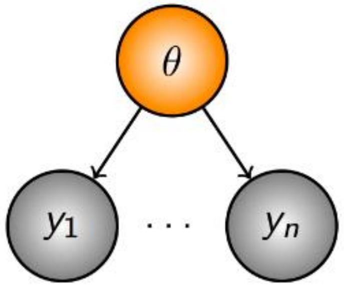


Figure 2.3: Diagrammatic representation complete pooling, a model structure in which all observations inform one parameter.

Fixed effect:
"no pooling"

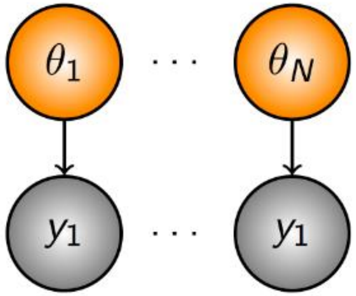


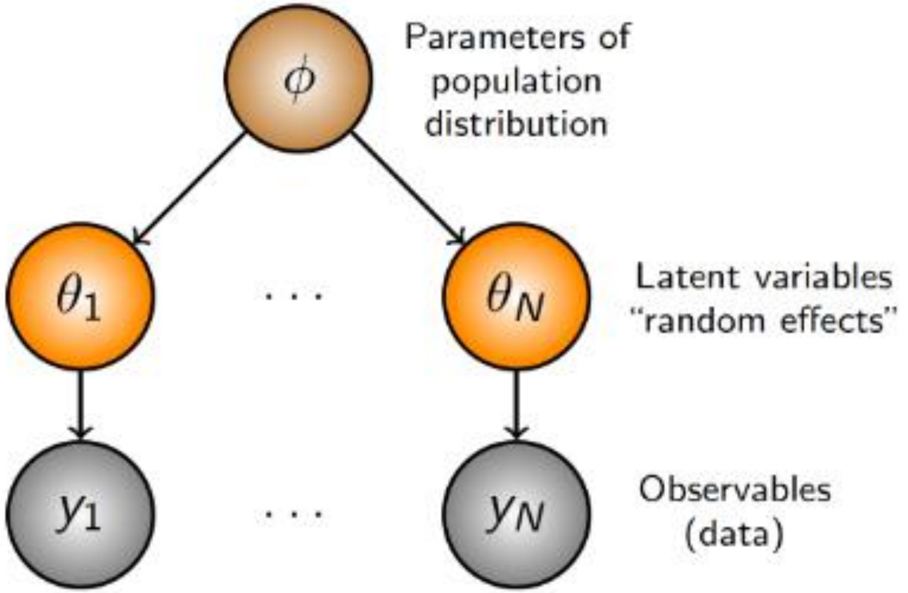
Figure 2.4: Diagrammatic representation no pooling, a model structure in which separate observations inform separate parameters.

A hierarchical model with grouping structure

$$\theta_i \sim N(\mu, \sigma_2)$$

$$\mu \sim N(.,.); \sigma \sim unif(.,.)$$

Hierarchical model



Each group has its own parameter(s) value, drawn from a distribution – a random variable. This distribution has *hyperparameters*

We learn about θ_i not only directly through y_i, but also indirectly through information from the other y_j via the population distribution parameterized by ϕ

Steve Midway online book: https://bookdown.org/steve_midway/BHME/Ch6.html

A hierarchical model with grouping structure

Steve Midway online book: https://bookdown.org/steve_midway/BHME/Ch1.html#hierarchical-models

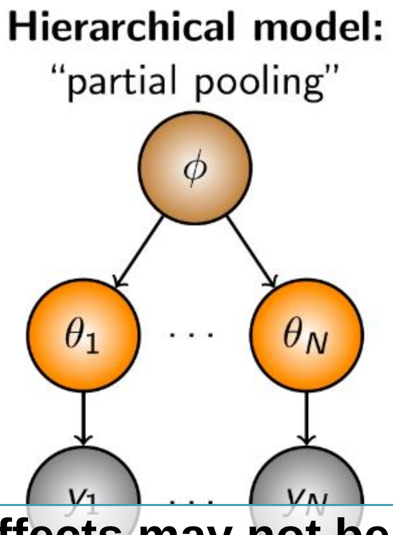


Figure ... representation partial pooling, a model structure in which different observations inform different latent parameters, which are then governed by additional parameters.

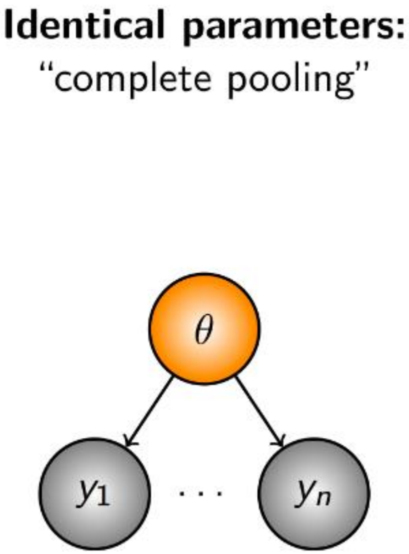


Figure 2.3: Diagrammatic representation complete pooling, a model structure in wh... inform one parameter.

Random effects may not be appropriate when…

Factor levels are very low (e.g., 2); it's hard to estimate distributional parameters with so little information (but still little risk).
Or when you don't want factor levels to inform each other.
e.g., male and female growth (combined estimate could be meaningless and misleading)

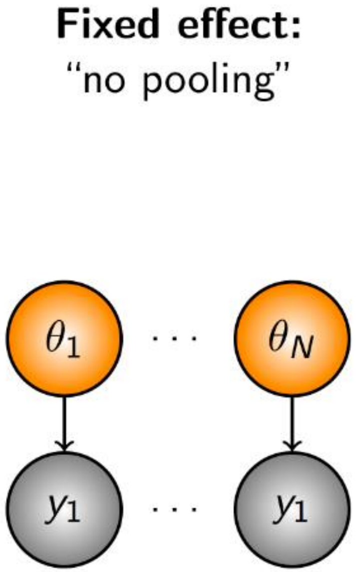


Figure 2.4: Diagrammatic representation no pooling, a model structure in which separate observations inform separate parameters.

A hierarchical model with grouping structure

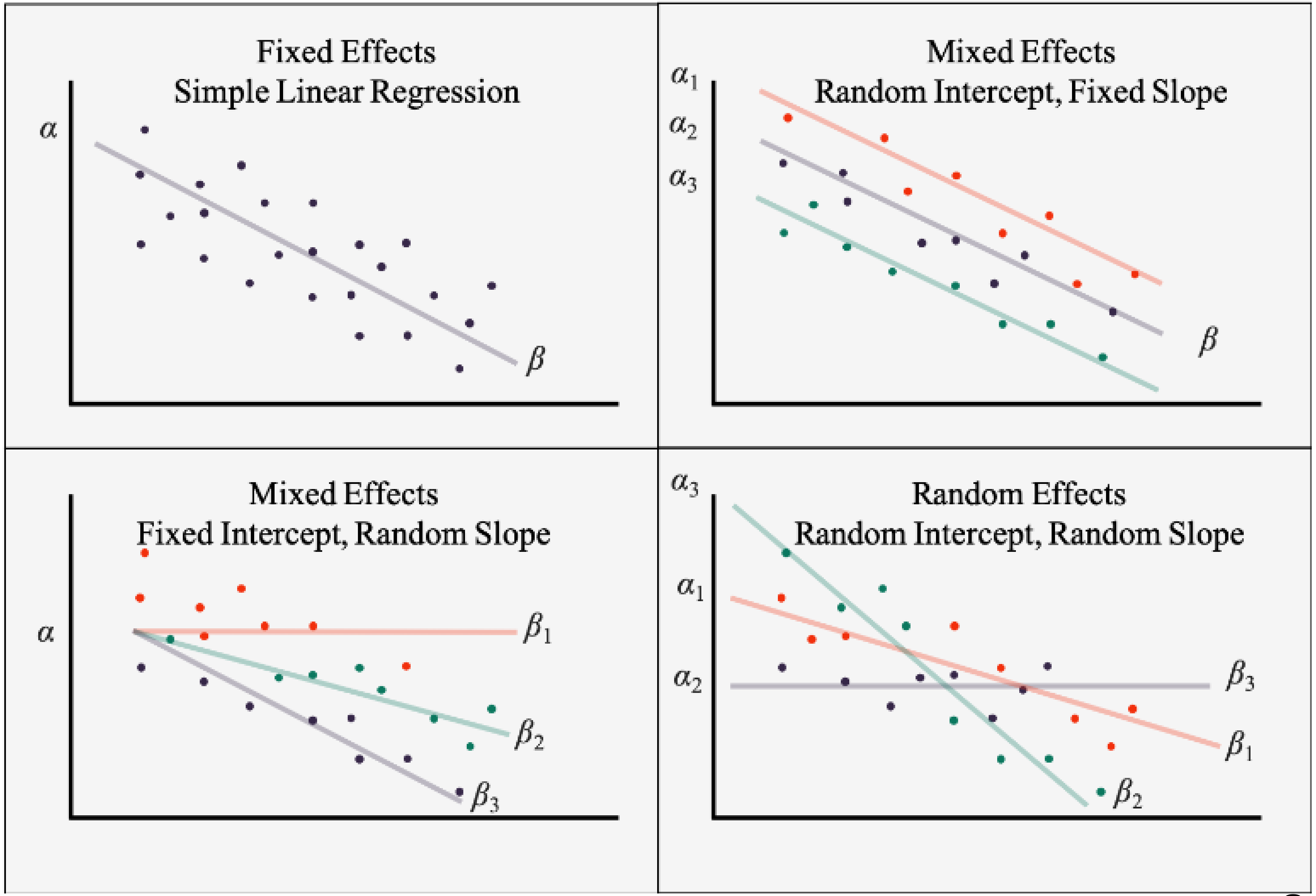


Figure 6.4: Examples of different varying coefficient models

Steve Midway online book: https://bookdown.org/steve_midway/BHME/Ch6.html

A hierarchical model with grouping structure

- SLR, no RE:

$$y_i = \alpha + \beta \times x_i + \epsilon_i$$

- SLR, random intercept, fixed slope:

$$y_i = \alpha_j + \beta \times x_i + \epsilon_i$$

- SLR, fixed intercept, random slope:

$$y_i = \alpha + \beta_j \times x_i + \epsilon_i$$

- SLR, random coefficients:

$$y_i = \alpha_j + \beta_j \times x_i + \epsilon_i$$

- MLR, random slopes:

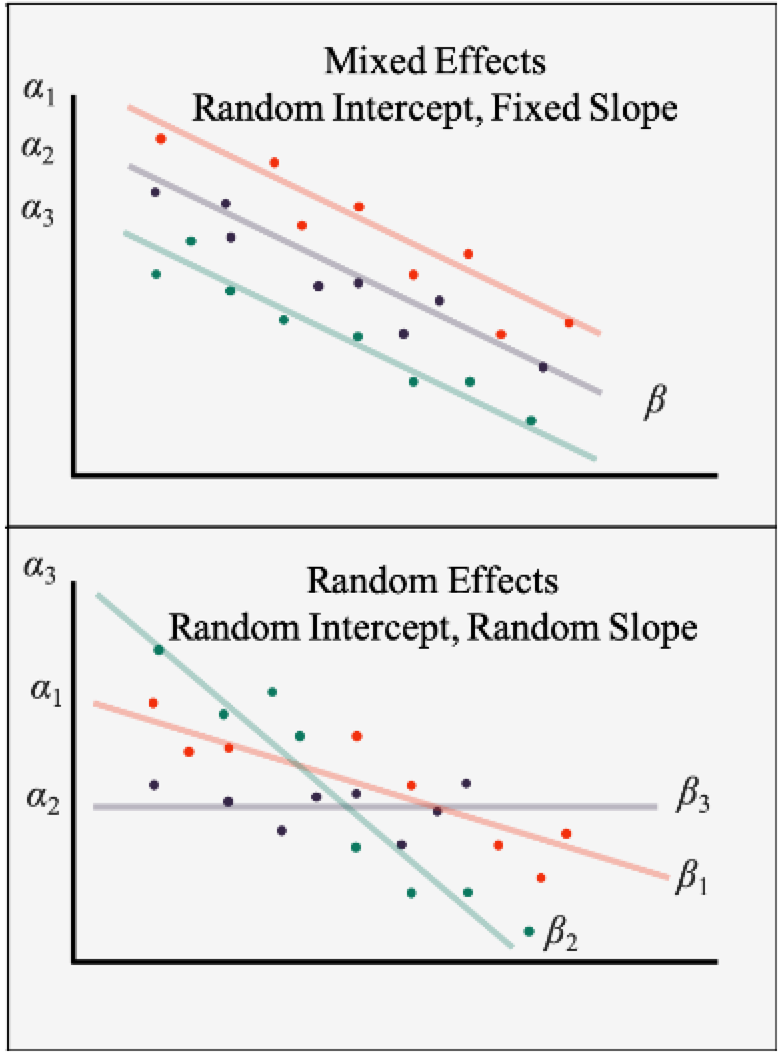
$$y_i = \alpha + \beta_j^1 \times x_i^1 + \beta_j^2 \times x_i^2 + \epsilon_i$$

For the most part, we will use subscript i to index the observation-level (data) and subscript j to index groups (and indicate a random effect).

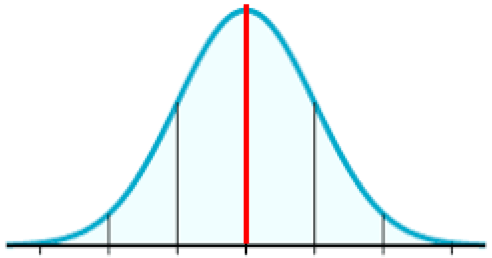
Steve Midway online book:
https://bookdown.org/steve_midway/BHME/Ch6.html

The parameters of our groups also form random variables

A hierarchical model with grouping structure

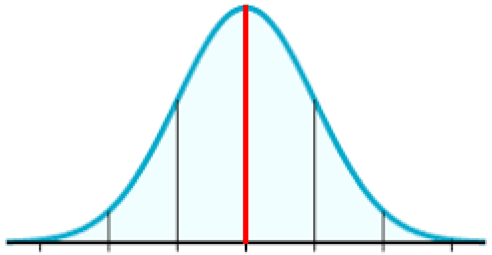


α

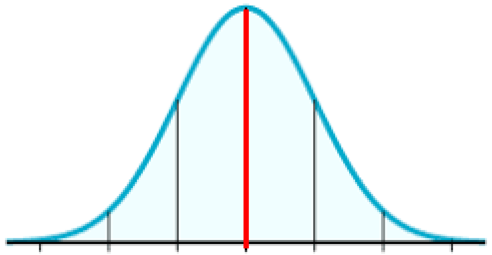


Values of α are drawn from a random variable. This distribution is described by it's *hyperparameters*

α



β



Steve Midway online book:
https://bookdown.org/steve_midway/BHME/Ch6.html

A hierarchical model with grouping structure

Simple linear model

The model from a Bayesian point of view

$$y_i \sim N(\alpha + \beta x_i, \sigma^2)$$

for $i, \dots n$

Priors:

$$\alpha \sim N(0, 0.001)$$

$$\beta \sim N(0, 0.001)$$

$$sigma \sim U(0, 10)$$

Steve Midway online book:
https://bookdown.org/steve_midway/BHME/Ch6.html

Varying intercepts, fixed slope model

This model might look like:

$$y_i \sim N(\alpha_{j(i)} + \beta x_i, \sigma^2)$$

for $i, \dots n$

$$\alpha_j \sim N(\mu_\alpha, \sigma_\alpha^2)$$

Overall mean, between-group variance

for $j, \dots J$

Group term

$$\mu_\alpha \sim N(0, 0.001)$$

$$\beta \sim N(0, 0.0001)$$

$$\sigma_\alpha^2 \sim U(0, 10)$$

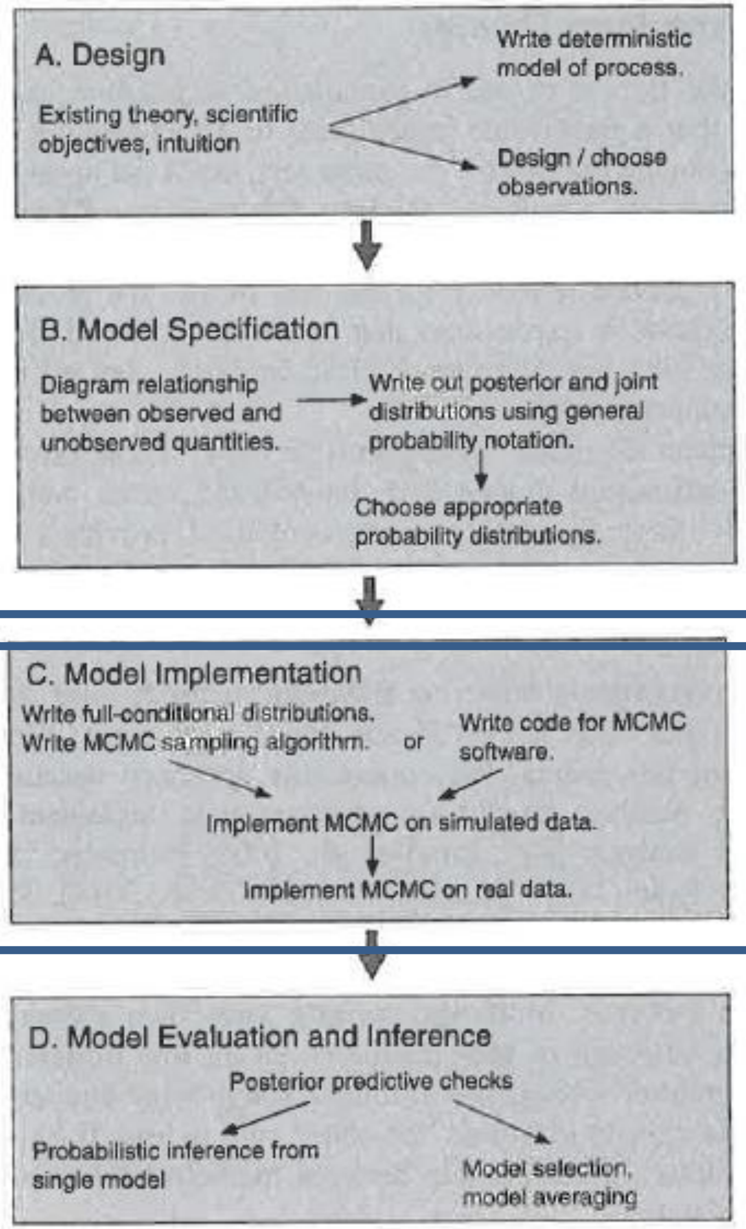
$$\sigma \sim U(0, 10)$$


Are you working with hierarchical data?
© Lindsay Banin

A framework for thinking about and fitting Bayesian hierarchical models in ecology

Four steps of a Bayesian analysis

1. Specify a joint distribution for the outcome(s) and all the unknowns, which typically takes the form of a marginal prior distribution for the unknowns multiplied by a likelihood for the outcome(s) conditional on the unknowns. This joint distribution is proportional to a posterior distribution of the unknowns conditional on the observed data
2. Draw from posterior distribution using Markov Chain Monte Carlo (MCMC).
3. Evaluate how well the model fits the data and possibly revise the model.
4. Draw from the posterior predictive distribution of the outcome(s) given interesting values of the predictors in order to visualize how a manipulation of a predictor affects the outcome(s).



Hobbs & Hooten (2015)

Define question, think about
ecological process and observations

Define the model to represent the ecological process and observations

- Diagram it
- Write out posterior and joint distributions
- Choose probability distributions

Implement the model using your expression for the posterior and joint distributions

Perform model checking

Practical

- We'll shortly fit a hierarchical model in 'rstanarm', working through it together
- Formulae based on 'lme4' (mixed-effects models; GLMMs) package
- Sticking with the tree allometry theme!

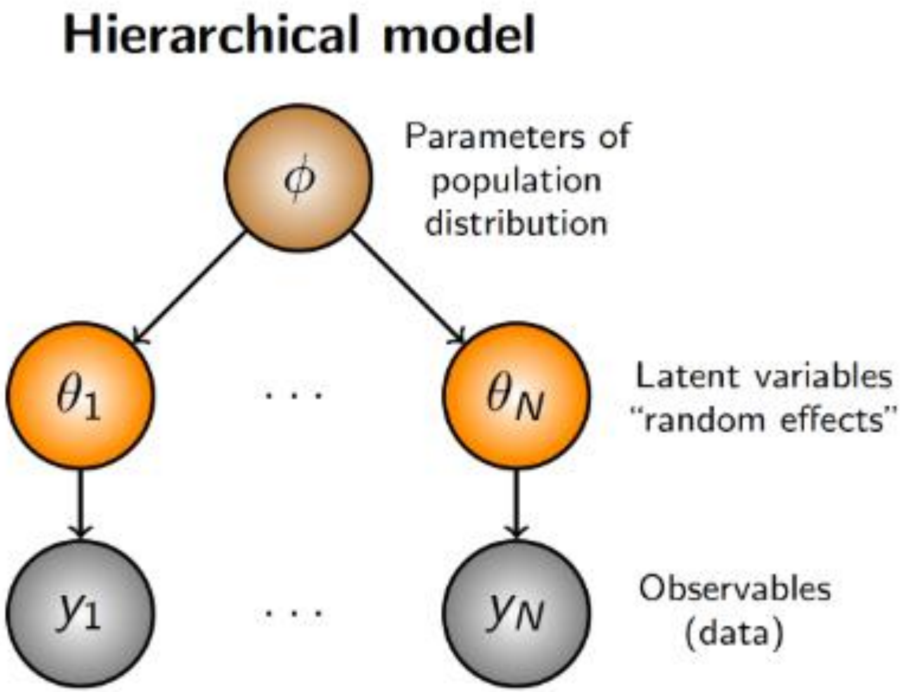
Practical factors to remember

- Choices as to how we handle our grouping terms
- Priors
- Number of chains
- Number of iterations
- Chain convergence (Rhat)
- Evaluation – does our model seem reasonable?

A hierarchical model with grouping structure

$$\theta_i \sim N(\mu, \sigma_2)$$

$$\mu \sim N(.,.); \sigma \sim unif(.,.)$$



Each group has its own parameter(s) (e.g. mean) drawn from a distribution – a random variable. This distribution has *hyperparameters*

We learn about θ_i not only directly through y_i, but also indirectly through information from the other y_j via the population distribution parameterized by ϕ

Steve Midway online book: https://bookdown.org/steve_midway/BHME/Ch6.html

What we covered in this session:

- **Hierarchical models** provide a **flexible framework** for addressing common features of our data: **structure** at different organisational levels and spatial scales (next we will generalise to unknowns such as latent variables and processes)
- We can follow **key steps in our Bayesian analyses**: define our model and break it down into components, draw from posterior distribution by applying MCMC algorithm, interrogate, visualise and evaluate the outputs
- We can fairly simply follow familiar (to some!) approaches to mixed-effects models using **rstanarm**…but sometimes we might want more flexibility or control.

Any Questions?

This concludes

Hierarchical modelling Part 1

Please be back & ready to go at 4 pm!

UK Centre for Ecology & Hydrology

Image © Pixaby CC01

What is a hierarchical model?

Steve Midway online book:
https://bookdown.org/steve_midway/BHME/Ch1.html#hierarchical-models

"…hierarchical structure, in which response variables are measured at the lowest level of the hierarchy and modelled as a function of predictor variables measured at that level and higher levels of the hierarchy." Wagner, Hayes, and Bremigan (2006)

"Multilevel (hierarchical) modeling is a generalization of linear and generalized linear modelling in which regression coeffcients are themselves given a model, whose parameters are also estimated from data." Gelman and Hill (2006)

"The existence of the process model is central to the hierarchical modelling view. We recognize two basic types of hierarchical models. First is the hierarchical model that contains an explicit process model, which describes realizations of an actual ecological process (e.g., abundance or occurrence) in terms of explicit biological quantities. The second type is a hierarchical model containing an implicit process model. The implicit process model is commonly represented by a collection of random effects that are often spatially and or temporally indexed. Usually the implicit process model serves as a surrogate for a real ecological process, but one that is diffcult to characterize or poorly informed by the data (or not at all)." Royle and Dorazio (2008)