

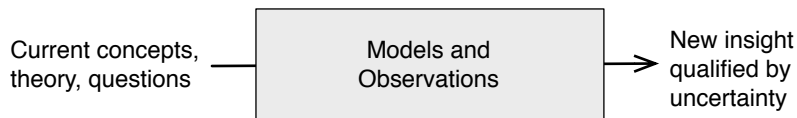
Bayesian Regression

ESS 575 Models for Ecological Data

N. Thompson Hobbs

March 3 2019

Where are we?



Where are we?

Deterministic models

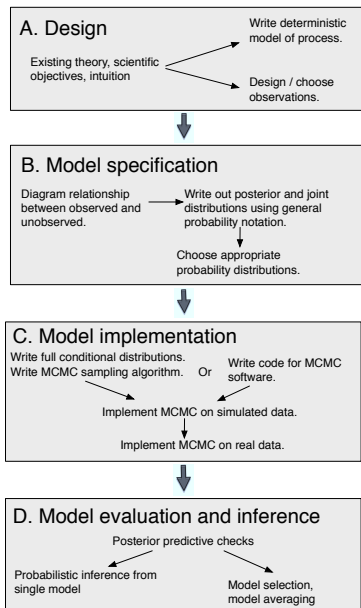
general linear
nonlinear
differential equations
difference equations
auto-regressive
occupancy
state-transition
integral-projection

Types of data

real numbers
non-negative real numbers
counts
0 to 1
0 or 1
counts in categories
proportions in categories
ordinal categories

univariate and
multivariate

Where are we?



Roadmap

- Today
 - ▶ Overview of simple Bayesian models
 - ▶ Exercise in writing regression models
 - ▶ Multi-level models
- Thursday: Introduction to hierarchical models
- Next two weeks: Practice writing models, learning JAGS
- Tuesday after spring break: debrief model writing problems
- Thursday after spring break: Nuts and bolts of regression
 - ▶ Standardizing
 - ▶ Interpreting coefficients
 - ▶ Matrix notation
 - ▶ Missing data
 - ▶ Some coding tricks

Learning outcomes for regression models

- Be able to write proper Bayesian, regression models for all types of data.
- Appreciate one-to-one relationship between math and JAGS code.
- Be able to interpret coefficients of general linear models.
- Know how and why to center or standardize data.
- Be able to translate scalar linear equations into matrix equations.
- Understand options for dealing with missing data
- Understand how to formulate multi-level models for slopes and intercepts (aka “random effects”)

A great follow-up

This book should be in your library:



Bayesian regression models are building blocks for more complex models

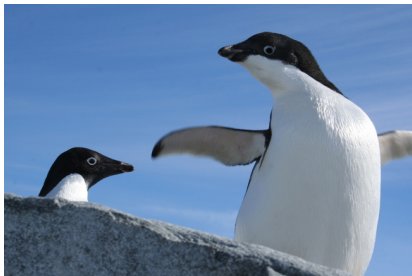
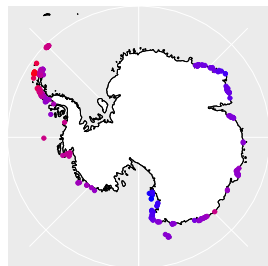
Continent-wide Adélie penguin population dynamics

$$g(\beta, z_{i,t}) = \log(z_{i,t-1} e^{(\beta_{0,i} + \beta_1 \text{wsic}_{i,t} + \beta_2 \text{ssic}_{i,t} + \beta_3 \text{krill}_{i,t}) \Delta t})$$

$$y_{i,t} \sim \text{Poisson}(z_{i,t})$$

$$z_{i,t} \sim \text{lognormal}(z_{i,t} \mid g(\beta_{0,i}, \beta_1, \beta_2, \beta_3, z_{i,t-1}), \sigma_{\text{process}}^2)$$

$$\beta_{0,i} \sim \text{normal}(\mu_{\text{site}}, \varsigma_{\text{site}}^2)$$



The general Bayesian set-up

Recall that the posterior distribution of the unobserved quantities conditional on the observed ones is proportional to their joint distribution:

$$[\theta|y] \propto [\theta, y].$$

The joint distribution can be factored into a likelihood and priors for simple Bayesian models:

$$[\theta, \sigma^2] = [y | \theta, \sigma^2] [\theta] [\sigma^2]$$

A deterministic model of an ecological process is embedded in the likelihood like this...

$$[\theta, \sigma^2] \propto [y | g(\theta, x), \sigma^2] [\theta] [\sigma^2]$$

Simple Bayesian regression models

As always, we start with a deterministic model,

$$\mu_i = \underbrace{g(\beta, x_i)}_{\text{deterministic model}}$$

where β is a vector of regression coefficients and \mathbf{x}_i is a vector of predictor variables corresponding to observation y_i . We use likelihood to connect the predictions of our model to data:

$$\underbrace{[y_i \mid \mu_i, \sigma^2]}_{\text{stochastic model}}$$

$$[\beta, \sigma^2 \mid \mathbf{y}] \propto \prod_{i=1}^n [y_i \mid g(\beta, x_i), \sigma^2] [\boldsymbol{\theta}] [\sigma^2]$$

We choose appropriate deterministic functions (linear or non-linear) and appropriate probability distributions to compose a specific model. Simple and flexible.

Identical notation

$$y_i = g(\beta, x_i) + \epsilon_i$$

$$\epsilon_i \sim \text{normal}(0, \sigma^2)$$

is the same as

$$y_i \sim \text{normal}(g(\beta, x_i), \sigma^2),$$

but the second notation is much more flexible because it doesn't require additive errors.

You don't have to be normal!

Data (y-values)	Distribution	Mean function	Link
continuous, real valued	normal	$\mu = \beta_0 + \beta_1 x$	NA
discrete, strictly positive	Poisson	$\mu = e^{\beta_0 + \beta_1 x}$	$\log(\mu) = \beta_0 + \beta_1 x$
0 or 1	Bernoulli	$\mu = \frac{\exp(\beta_0 + \beta_1 x)}{\exp(\beta_0 + \beta_1 x) + 1}$	$\text{logit}(\mu) = \log\left(\frac{\mu}{1-\mu}\right) = \beta_0 + \beta_1 x$
0 – 1	beta	$\mu = \frac{\exp(\beta_0 + \beta_1 x)}{\exp(\beta_0 + \beta_1 x) + 1}$	$\text{logit}(\mu) = \log\left(\frac{\mu}{1-\mu}\right) = \beta_0 + \beta_1 x$
continuous, strictly positive	lognormal or gamma	$\mu = e^{\beta_0 + \beta_1 x}$	$\log(\mu) = \beta_0 + \beta_1 x$