University of Nevada, Reno

Lecture 4: Linking Models to Data

Perry Williams, PhD

NRES 779

Bayesian Hierarchical Modeling in Natural Resources

Today

- How we link deterministic models to data modeling

Linking Models to Data

Probability model

Observations (data)

Deterministic model

Idea!

What is the probability that I would observe the data if my model is a faithful representation of the processes that gave rise to the data?

Layout of next few lectures

- Today: Linking data to deterministic models
- Friday: Basic laws of probability
- Wednesday: Probability distributions
- Friday: Likelihood and Bayes' Theorem

Learning Objectives

- Introduce first ideas about *support*
- Distinguish between purely empirical models and models symbolizing processes.
- Introduce a set of functional forms useful for composing deterministic models.
- Cross cutting themes
  - A relatively small set of functions can be used to describe a broad array of (ecological) processes.
  - The same process can be represented by different functional forms.
  - The same functional form can be used to represent different processes.

$f(x_i, \boldsymbol{\theta})$

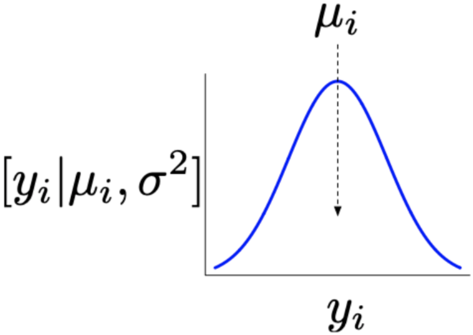
Any type of mathematical function

- linear models
- non-linear models
- systems of differential equations
- systems of difference equations
- integral-projection models
- state-transition models
- matrix models

Any equation or system of equations making a prediction that can be compared with an observation.

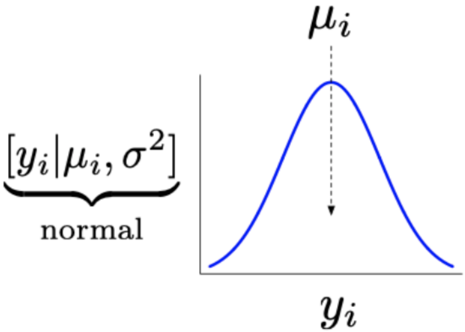
Linking Models to Data

$\boldsymbol{\mu}_i = f(x_i, \boldsymbol{\theta})$



Example

$\boldsymbol{\theta} = (\beta_0, \beta_1)'$
$\boldsymbol{\mu}_i = f(x_i, \boldsymbol{\theta}) = \beta_0 + \beta_1 x_1$



This model is often a poor choice in biology. Why?

Support

Support refers to the range of values that a variable can realize. A more formal definition will come soon. Describe the support for the following variables:

- Soil organic matter content (gm OM/gm dry matter)
- Observed survival of an individual
- Species richness
- Carbon flux
- Above ground biomass of grassland

Example functional forms for $f(x_i, \boldsymbol{\theta})$

- Additive effects
- Asymptotic processes
- Power functions

Additive, nonlinear models

- Additive models contain linear functions of coefficients and predictor variables, e.g., $\beta_0 + \beta_1 x_{1,i} + \ldots + \beta_d x_{d,i}$
- Often referred to as generalized linear models family because transforming the left hand side results in a linear model.
- Are usually *empirical* - very useful for modeling correlation between predictors and responses.

Generalized linear models

What if response variable is between 0 and 1

- Proportion of plots with invasive species
- Nitrogen content of soil (gN/gOM)
- Proportion of landscape burned
- Survival probability of juveniles
- Prevalence of a disease in a population

Inverse logit function

Let $\mu =$ variable that can take on values between 0 and 1.

$\text{logit}(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$ converts μ to values between -∞ and ∞.

$\text{logit}(\mu) = \log\left(\frac{\mu}{1-\mu}\right) = \beta_0 + \beta_1 x_1 + \ldots + \beta_d x_d$

$\mu = \text{logit}^{-1}\left(\beta_0 + \beta_1 x_1 + \ldots + \beta_d x_d\right)$

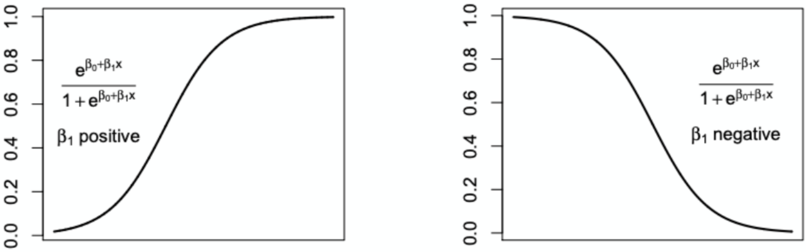
$\mu = \frac{e^{\beta_0 + \beta_1 x_1 + \ldots + \beta_d x_d}}{1 + e^{\beta_0 + \beta_1 x_1 + \ldots + \beta_d x_d}}$

You will also see $\mu = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \ldots + \beta_d x_d)}}$

But, be careful about the minus in the exponent!

You can include powers and products of the xs

Inverse logit function



What Not to Do

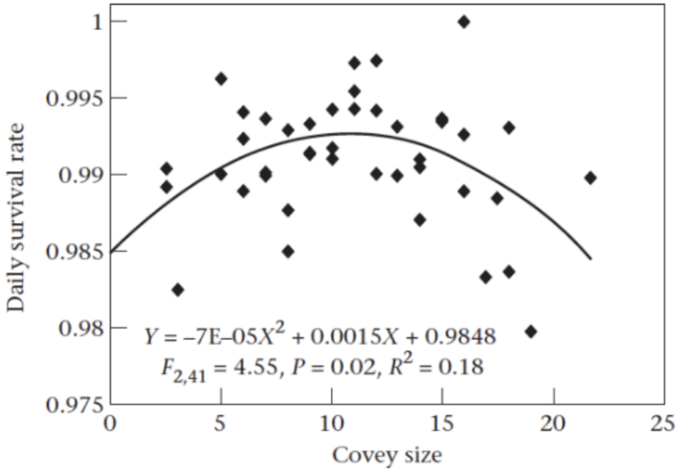


Figure 10. The influence of covey size on individual daily survival between 9 November and 31 January 1997–2000 in east-central Kansas.

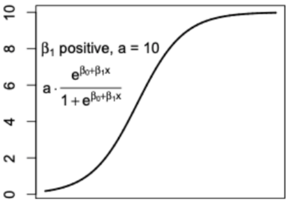
Generalize Linear Models

What if a response is between 0 and *a*?

Multiply by a:

$$\frac{a\,e^{(\beta_0+\beta_1 x_i....+\beta_n x_n)}}{1+e^{(\beta_0+\beta_1 x_i....+\beta_n x_n)}}$$

Always non-negative and does not reach excessively large values



Generalize Linear Models

What if a response must be $\geq$0?

For example, we want to model μ_t as an additive function of covariates:

$$\begin{aligned} N_{t+1} &= \mu_t N_t \\ \mu_t &= f(x_t, \boldsymbol{\beta}) \end{aligned}$$

Other example responses that must be non-negative:

- biomass
- energy expenditure
- nitrogen mineralization
- population density
- species richness
- ground water flow

Generalize Linear Models

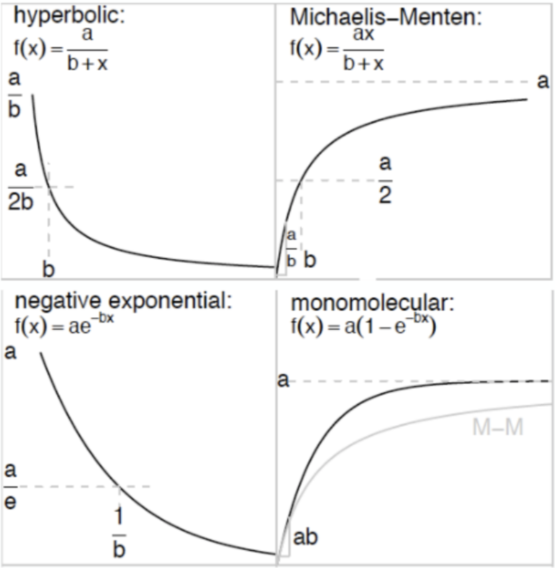
Exponential model

$$\mu_t = \exp\big(\beta_0 + \beta_1 x_{1,t} + \ldots + \beta_d x_{d,t}\big)$$

which is also written as

$$\log(\mu_t) = \beta_0 + \beta_1 x_{1,t} + \ldots + \beta_d x_{d,t}$$

Asymptotic Functions



Figures courtesy of
Bolker, B. 2008. Ecological Models and Data in R. Princeton University Press, Princeton, N. J USA.

Deterministic Functions From Past Classes

Meredith Brehob, Lambert-Beer Law (depth of light penetration)
Negative exponential function:

$$I_z = I_0 e^{-k_d z}$$
$$\mu_i = \beta_0 e^{-\beta_1 x},$$

Deterministic Functions From Past Classes

Mia Goldman, Discrete Logistic Growth

$$N_{t+1} = N_t + rN_t(1 - \frac{N_t}{K})$$

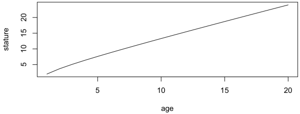
Deterministic Functions From Past Classes

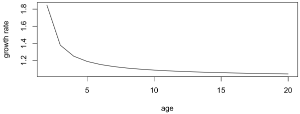
Chris Wolfe, human structural growth

$$\mu_i = \beta_0 + \beta_1 x_i + \beta_2 \log(x_i),$$

where

- μ_i is stature of individual i.
- x_i is the age of individual i





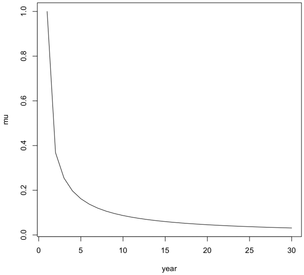
Deterministic Functions From Past Classes

Meghan Keating, Ricker model

$$\mu_t = \beta_0 \mu_{t-1} e^{-\beta_1 \mu_{t-1}},$$

where

- μ_t is population size at time t.



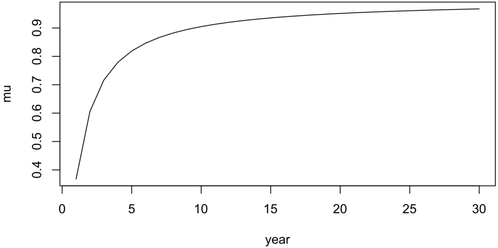
Deterministic Functions From Past Classes

Steve Hromada, Arrhenius equation

$$\mu_t = \beta_0 e^{-\frac{\beta_1}{\sigma^2 t}},$$

where

- μ_t thermal rate constant at temperature t.
- β_0 the "pre-exponential factor" (called a scale parameter in statistics)
- β_1 is the activation energy for the reaction
- σ^2 The universal gas constant.



Deterministic Functions From Past Classes

Elaine Chu, Logistic equation

$$\text{logit}(\mu_i) = \beta_0 + \beta_1 x_{i,1} + \ldots + \beta_d x_{i,d},$$

where

- μ_t is the "probability" of being female.
- βs are the "weights" associated with morphological features x_i
- x_i are morphological features typical of males and females from the pelvis and/or cranium.

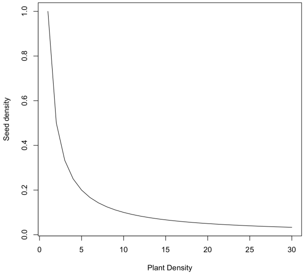
Deterministic Functions From Past Classes

Sage Ellis, Hyperbolic function

$$\mu_i = \frac{\beta_0}{x_{i,1}},$$

where

- μ_t is, for example, seed density.
- β_0 is some constant.
- x_i is plant density.



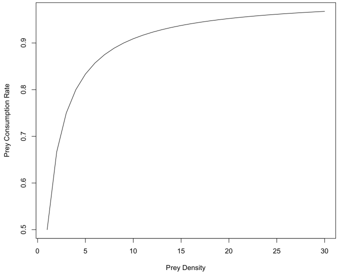
Deterministic Functions From Past Classes

Madeleine Lohman, Holling Type II Functional Response

$$\mu_t = \frac{\beta_0 x_t}{1 + \beta_0 \beta_1 x_t},$$

where

- μ_t is the prey consumption rate.
- β_0 is the attack rate.
- x_i is prey density.
- β_1 is the handling time.



Deterministic Functions From Past Classes

Jason Gundlach, Holling Type III Functional Response

$$\mu_t = \frac{\beta_0 \beta_1 x_t}{1 + \beta_0 \beta_2 x_t},$$

where

- μ_t is the prey consumption rate.
- β_0 is the attack rate.
- β_1 total time spent.
- β_2 is the handling time.
- x_t is prey density.