Bayesian Regression
Models for Socio-Environmental Data

Chris Che-Castaldo, Mary B. Collins, N. Thompson Hobbs

December 17, 2020

Outline

- Be able to write models for different types of data.
- Appreciate one-to-one relationship between math and JAGS code.
- Be able to interpret coefficients of general linear models.

The general Bayesian set-up

Recall that the posterior distribution is proportional to their joint distribution:
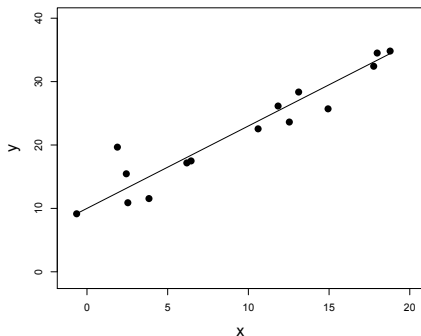
$$[\theta|y] \propto [\theta, y].$$

The joint distribution can be factored into a likelihood and priors:

$$\left[\theta, \sigma^2\right] = \left[y \mid \theta, \sigma^2\right] \left[\theta\right] \left[\sigma^2\right]$$

We often wish to model the mean response to y. To do this we embed some ecological or socio-environmental deterministic function $g(\beta, x)$ into the likelihood like this. . . .

$$\left[\boldsymbol{\theta}, \sigma^2\right] \propto \left[y \mid g(\beta, x), \sigma^2\right] \left[\beta\right] \left[\sigma^2\right]$$

A simple example with an old friend

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

$$\left[\boldsymbol{\beta}, \sigma^2 | \boldsymbol{y}\right] \propto \prod_{i=1}^{15} \text{normal}\left(y_i | g(\boldsymbol{\beta}, x_i), \sigma^2\right) \text{normal}\left(\beta_0 | 0, 1000\right) \times$$
$$\text{normal}\left(\beta_1 | 0, 1000\right) \times$$
$$\text{inverse gamma}\left(\sigma^2 | 0.001, 0.001\right)$$
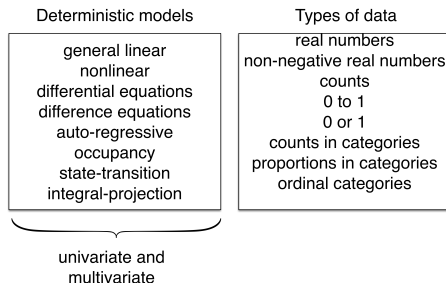
Lots of flexibility as a modeler

Bayesian regression models are building blocks for more complex models. For example, the continent-wide Adélie penguin population dynamics:

$$z_{s,y} \sim \text{lognormal}\big(z_{s,y} \mid g(\beta_1, \beta_2, z_{s,y-1}), \sigma^2\big)$$
$$g(\theta) = \log\big(z_{s,y-1} e^{\beta_1 + \beta_2 \text{wsic}_{s,y} + \beta_3 \text{ssic}_{s,y}}\big)$$
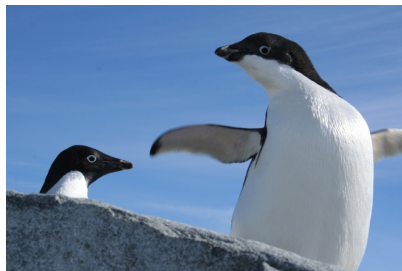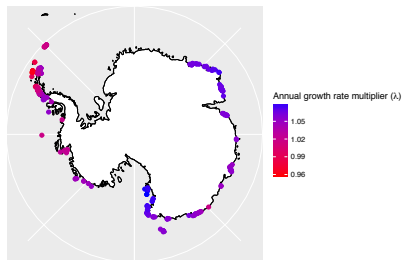
{Photo c/o Heather J. Lynch}

What about other kinds of data?

Deterministic models	Types of data
general linear	real numbers
nonlinear	non-negative real numbers
differential equations	counts
difference equations	0 to 1
auto-regressive	0 or 1
occupancy	counts in categories
state-transition	proportions in categories
integral-projection	ordinal categories

univariate and
multivariate

Two considerations:

- Moment matching (which you are familiar with)
- Mapping our deterministic model $g(\beta, x) = \beta_0 + \ldots + \beta_n x_n$ from $(-\infty, \infty)$ to a range appropriate for the likelihood in our model

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	$\mu = \frac{\exp(\beta_0 + \beta_1 x)}{\exp(\beta_0 + \beta_1 x) + 1}$	$\mathrm{logit}(\mu) = \log\left(\frac{\mu}{1-\mu}\right) = \beta_0 + \beta_1 x$
$[0,1]$	beta	$\mu = \frac{\exp(\beta_0 + \beta_1 x)}{\exp(\beta_0 + \beta_1 x) + 1}$	$\mathrm{logit}(\mu) = \log\left(\frac{\mu}{1-\mu}\right) = \beta_0 + \beta_1 x$
continuous, strictly positive, variance $\uparrow$ as a f(mean)	lognormal	$\mu = e^{\beta_0 + \beta_1 x}$	$\log(\mu) = \beta_0 + \beta_1 x$
continuous, strictly positive, constant variance	gamma	$\mu = e^{\beta_0 + \beta_1 x}$	$\log(\mu) = \beta_0 + \beta_1 x$

Normal data, continuous and real valued

$$
\begin{aligned}
[\beta_0, \beta_1, \sigma \mid \mathbf{y}] \;\propto\; & \prod_{i=1}^{N} \text{normal}\big(y_i \mid g(\beta_0, \beta_1, x_i), \sigma^2\big) \\
& \times \text{normal}\big(\beta_0 \mid 0, 1000\big) \, \text{normal}\big(\beta_1 \mid 0, 1000\big) \\
& \times \text{uniform}\big(\sigma \mid 0, 100\big) \\
g(\beta_0, \beta_1, x_i) \;=\; & \beta_0 + \beta_1 x_i
\end{aligned}
$$

```
b0 ~ dnorm(0, .001)
b1 ~ dnorm(0, .001)
sigma ~ dunif(0, 100)
tau <- 1/sigma^2
for (i in 1:length(y)){
  mu[i] <- b0 + b1 * x[i]
  y[i] ~ dnorm(mu[i], tau)
}
```

Counts, discrete and non-negative

You have collected some count data ($y = 12, 17, 1, 0, 31, \ldots, 25$) at n sites, along with a covariate (x) which you believe is likely to affect these counts. Write a model regressing y on x.

- Choose a specific deterministic and stochastic model.
- Write out the DAG, posterior distribution, and joint distribution for your model.
- Interpret the coefficients of your model.
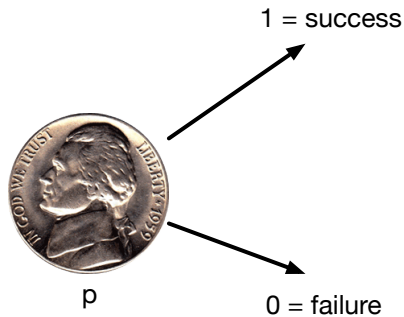
Poisson, discrete and non-negative

$$
\begin{aligned}
[\beta_0, \beta_1 \mid \mathbf{y}] &\propto \prod_{i=1}^{N} \text{Poisson}\big(y_i \mid g(\beta_0, \beta_1, x_i)\big) \\
&\quad \times \text{normal}\big(\beta_0 \mid 0, 1000\big)\,\text{normal}\big(\beta_1 \mid 0, 1000\big) \\
g(\beta_0, \beta_1, x_i) &= e^{\beta_0 + \beta_1 x_i}
\end{aligned}
$$

```
b0 ~ dnorm(0, .001)
b1 ~ dnorm(0, .001)
for(i in 1:length(y)){
  log(mu[i]) <- b0 + b1 * x[i]
  y[i] ~ dpois(mu[i])
}
```

or

```
mu[i] <- exp(b0 + b1 * x[i])
y[i] ~ dpois(mu[i])
```

Poisson, discrete and non-negative

$$
\begin{aligned}
[\beta_0, \beta_1 \mid \mathbf{y}] &\propto \prod_{i=1}^{N} \text{Poisson}(y_i \mid g(\beta_0, \beta_1, x_i)) \\
&\times \text{normal}(\beta_0 \mid 0, 1000)\, \text{normal}(\beta_1 \mid 0, 1000) \\
g(\beta_0, \beta_1, x_i) &= e^{\beta_0 + \beta_1 x_i}
\end{aligned}
$$

Exponentiate β_0 and β_1 and report counts or multiplicative changes in counts.
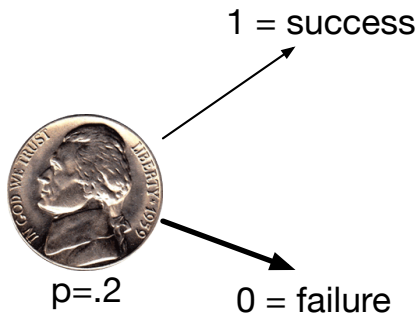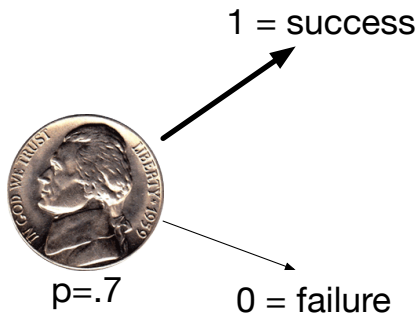
$e^{\beta_1} = \frac{e^{\beta_0 + \beta_1 x}}{e^{\beta_0}}$ which is the multiplicative change in the mean count per unit change in x.

For example: "Mean western toad juvenile abundance is reduced by a factor of 5.1 (95% CI: 3.4, 10.8) per unit change in UV-B radiation."

Binary data, 0 or 1 (aka logistic)

1 = success

0 = failure

p

Binary data, 0 or 1 (aka logistic)

$1 = success$

$1 = success$

p=.7

$0 = failure$

p=.2

$0 = failure$

Binary data, 0 or 1 (aka logistic)

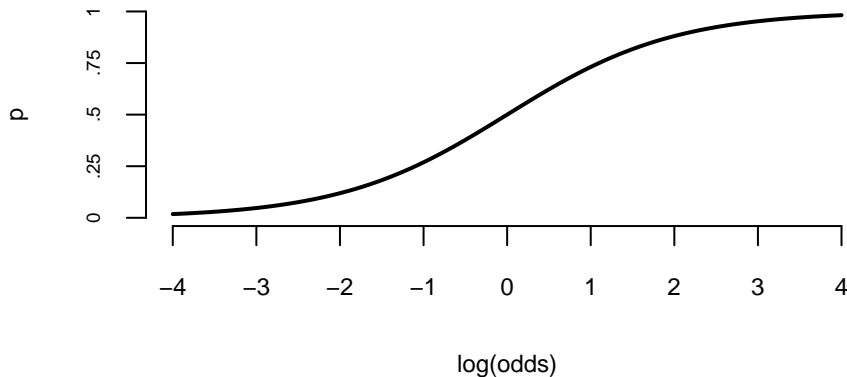What happens when we want to make p a function of a continuous predictor?

- probability: p $[0, 1]$
- odds: $\frac{p}{1-p}$ $[0, \infty)$
- log odds: $\log(\text{odds})$ $(-\infty, \infty)$

Moving between probabilty and log odds

- $\text{logit}() = \log\left(\frac{p}{1-p}\right)$: input is probability, output is log odds
- inverse $\text{logit}() = \frac{e^{\log(\text{odds})}}{e^{\log(\text{odds})}+1}$: input is log(odds), output is probability

Binary data, 0 or 1 (aka logistic)

inverse logit mapping: input is log(odds), output is probability

Binary data, 0 or 1 (aka logistic)

You have collected some binary data ($y = 1, 0, 0, 1, 1, 0, 1, \ldots, 1$) at n sites, along with a covariate (x) which you believe is likely to affect these counts. Write a model regressing y on x.

- Choose a specific deterministic and stochastic model.
- Write out the DAG, posterior distribution, and joint distribution for your model.
- Interpret the coefficients of your model.

Bernoulli, 0 or 1 (aka logistic)

$$[\beta_0, \beta_1 \mid \mathbf{y}] \propto \prod_{i=1}^{N} \text{Bernoulli}(y_i \mid g(\beta_0, \beta_1, x_i))$$
$$\times \text{normal}(\beta_0 \mid 0, 2.7)\,\text{normal}(\beta_1 \mid 0, 2.7)$$
$$g(\beta_0, \beta_1, x_i) = \frac{e^{\beta_0 + \beta_1 x_i}}{e^{\beta_0 + \beta_1 x_i} + 1}$$

```
b0 ~ dnorm(0, 0.37)
b1 ~ dnorm(0, 0.37)
for(i in 1:length(y)){
  logit(p[i]) <- b0 + b1 * x[i]
    y[i] ~ dbern(p[i])
 }
```

or

```
p[i] <- inv.logit(b0 + b1 * x[i])
y[i] ~ dbern(p[i])
```
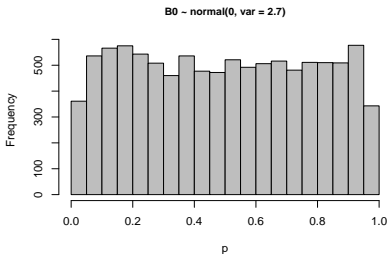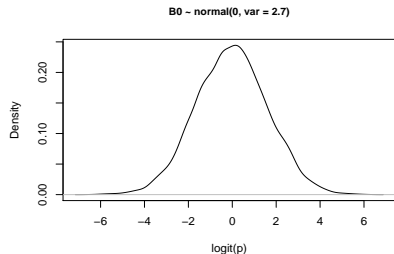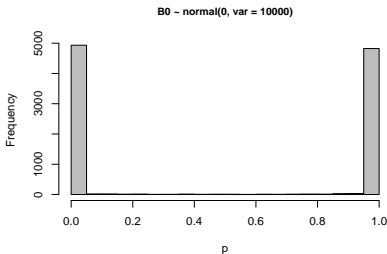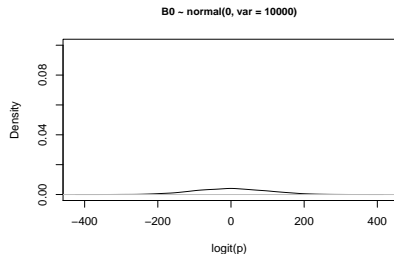
Interpreting logit coefficients

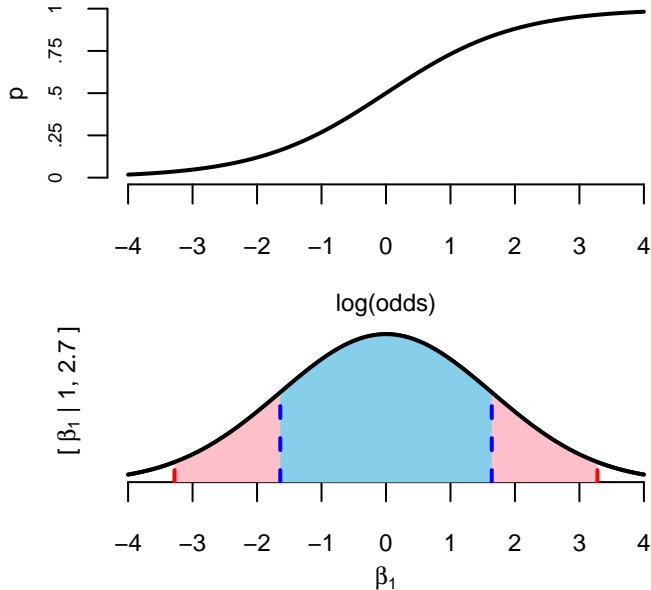Exponentiate β_0 and β_1 and report odds and odds ratios.

$e^{\beta_1} = \frac{e^{\beta_0 + \beta_1 x}}{e^{\beta_0}}$ which is an odds ratio, i.e. the change in the odds.

For example: "The odds of detecting weevils in upland willow stems were 3.2 (95% CI: 2.3, 4.8) times greater than detecting them in riparian willow stems."

Choosing reasonable flat priors on logit intercept

Choosing reasonable flat priors on logit effects

lognormal, data continuous and > 0 (log link)

$$
\begin{aligned}
[\beta_0, \beta_1, \sigma \mid \mathbf{y}] &\propto \prod_{i=1}^{N} \text{lognormal}\big(y_i \mid \log\big(g(\beta_0, \beta_1, x_i)\big), \sigma^2\big) \\
&\quad \times \text{normal}\big(\beta_0 \mid 0, 1000\big) \, \text{normal}\big(\beta_1 \mid 0, 1000\big) \\
&\quad \times \text{uniform}\big(\sigma \mid 0, 5\big) \\
g(\beta_0, \beta_1, x_i) &= e^{\beta_0 + \beta_1 x_i}
\end{aligned}
$$

Talk about the interpretation of σ.

```
b0 ~ dnorm(0, .001)
b1 ~ dnorm(0, .001)
sigma ~ dunif(0, 5)
tau <- 1/sigma^2
for(i in 1:length(y)){
  mu[i] <- exp(b0 + b1 * x[i])
  y[i] ~ dlnorm(log(mu[i]), tau)
}
```

lognormal, data continuous and > 0 (not log link)

$$
\begin{aligned}
[\beta_0, \beta_1, \sigma \mid \mathbf{y}] \quad &\propto \quad \prod_{i=2}^{N} \text{lognormal}\big(y_i \mid \log\big(g(\beta_0, \beta_1, y_{i-1})\big), \sigma^2\big) \\
&\quad \times \text{normal}\big(\beta_0 \mid 0, 1000\big) \, \text{normal}\big(\beta_1 \mid 0, 1000\big) \\
&\quad \times \text{uniform}\big(\sigma \mid 0, 5\big) \, \text{uniform}\big(y_1 \mid 1, 1E6\big) \\
g(\beta_0, \beta_1, y_{i-1}) \quad &= \quad y_{i-1} e^{\beta_0 + \beta_1 y_{i-1}}
\end{aligned}
$$

```
b0 ~ dnorm(0, .001)
b1 ~ dnorm(0, .001)
sigma ~ dunif(0, 5); tau <- 1/sigma^2
y[1] ~ dunif(1, 1E6)
for(i in 2:length(y)){
  mu[i] <- y[i-1] * exp(b0 + b1 * y[i-1])
  y[i] ~ dlnorm(log(mu[i]), tau)
}
```