

Generalized Linear Mixed Models: Binomial

Quantitative Analysis of Vertebrate Populations

Generalized Linear Model - Binomial

=Logistic Regression

- Remember the binomial distribution
- Series of “coin-flip” trials
- Count the total number of an outcome (“heads”) in a series of N trials (“coin flips”)
- Estimate the $\text{Pr}(\text{heads})$ and factors that influence that
- Presence-Absence or detection-nondetection data

Logistic Regression

1. Distribution: $C_i \sim \text{Binomial}(N, p_i)$
2. Link Function: $\text{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \text{linearpredictor}$
3. Linear Predictor: $\text{logit}(p_i) = \alpha + \beta X_i$

Let's try it

Data Generation

Two species (Cross-leaved gentian & Chiltern gentian) with P/A data from 50 sites

```
N <- 50
p.cr <- 13/50 # avg prob of finding
p.ch <- 29/50 # avg prob of finding

C.cr <- rbinom(1, 50, prob = p.cr) # add binomial noise
C.ch <- rbinom(1, 50, prob = p.ch) # add binomial noise
C <- c(C.cr, C.ch)
species <- factor(c(0,1), labels = c("Cross-leaved", "Chiltern"))
```

Analysis

Check if a significant difference in the distribution (freq of occurrence)

```
summary(glm(cbind(C, N-C) ~ species, family = "binomial"))

##
## Call:
## glm(formula = cbind(C, N - C) ~ species, family = "binomial")
##
## Deviance Residuals:
## [1] 0 0
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.2657     0.3414  -3.707 0.000209 ***
## speciesChiltern  1.5068     0.4447   3.389 0.000702 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1.2466e+01  on 1  degrees of freedom
## Residual deviance: 2.4425e-15  on 0  degrees of freedom
## AIC: 12.363
##
## Number of Fisher Scoring iterations: 3
predict(glm(cbind(C, N-C) ~ species, family = "binomial"), type = "response")

##      1      2
## 0.22 0.56
```

Binomial GLMM

Random intercept and slope

1. Distribution: $C_i \sim \text{Binomial}(N_i, p_i)$
2. Link Function: $\text{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \text{linearpredictor}$
3. Linear Predictor: $\text{logit}(p_i) = \alpha_{j(i)} + \beta_{j(i)} X_i$

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

$$\beta_j \sim \mathcal{N}(\mu_\beta, \sigma_\beta^2)$$

Example: Generate Data

Red-backed Shrike: nest success (or failure) rather than counts

```
n.groups <- 16
n.years <- 30
n <- n.groups * n.years
pop <- gl(n = n.groups, k = n.years)
```

Uniform covariate as an index of spring precipitation (0 = little rain, 1 = lots)

```
precip <- runif(n, 0, 1)
```

Binomial total = number of nest attempts in survey year i

```
N <- round(runif(n, 10, 50))
```

Build the design matrix

```
Xmat <- model.matrix(~pop*precip-1-precip)
```

Choose parameter values and hyperparameters

```
intercept.mean <- 1          # Select hyperparams
intercept.sd <- 1
slope.mean <- -2
slope.sd <- 1
```

```

intercept.effects<-rnorm(n = n.groups, mean = intercept.mean, sd = intercept.sd)
slope.effects <- rnorm(n = n.groups, mean = slope.mean, sd = slope.sd)
all.effects <- c(intercept.effects, slope.effects) # Put them all together

```

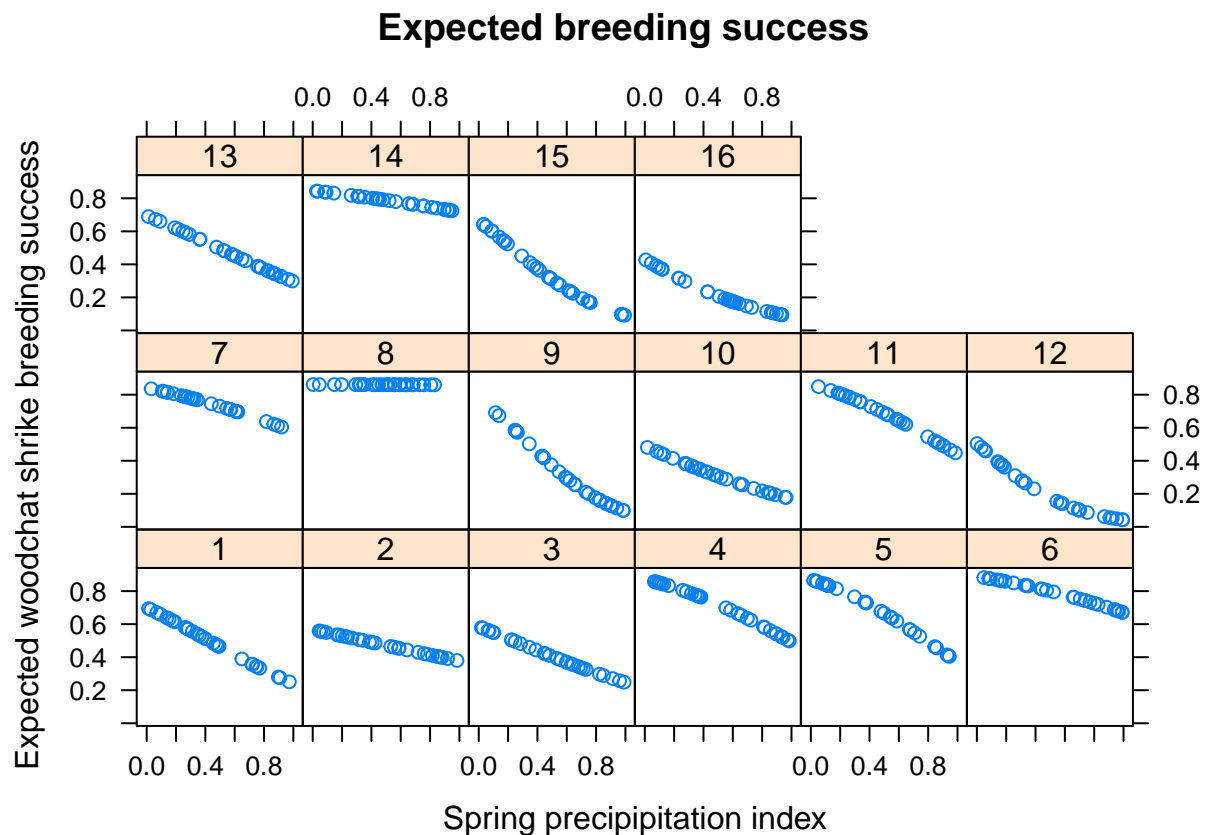
Assemble to make the linear predictor and visualize

```

lin.pred <- Xmat %*% all.effects # Value of lin.predictor
exp.p <- exp(lin.pred) / (1 + exp(lin.pred)) # Expected proportion

library("lattice")
xyplot(exp.p ~ precip | pop, ylab = "Expected woodchat shrike breeding success ",
xlab = "Spring precipitation index", main = "Expected breeding success")

```

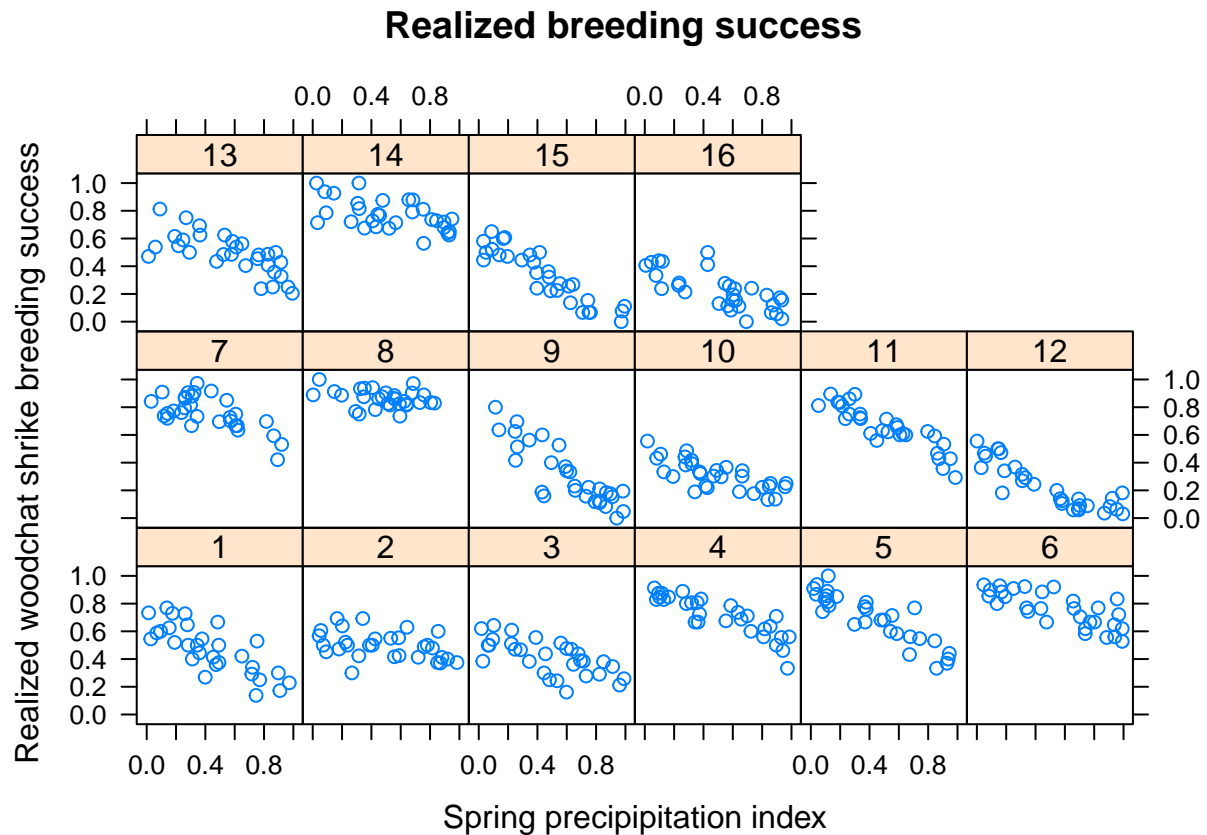


Add random binomial variation

```

C <- rbinom(n = n, size = N, prob = exp.p) # Add binomial variation
xyplot(C/N ~ precip | pop, ylab = "Realized woodchat shrike breeding success ",
xlab = "Spring precipitation index", main = "Realized breeding success")

```



Analysis under a random-coefficients model

Assume no correlation between intercept and slope

```
library('lme4')
```

```
## Loading required package: Matrix
```

```
glmm.fit <- glmer(cbind(C, N-C) ~ precip + (1 | pop) + (0 + precip | pop), family = binomial)
summary(glmm.fit)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: cbind(C, N - C) ~ precip + (1 | pop) + (0 + precip | pop)
##
```

```
##      AIC      BIC    logLik deviance df.resid
##  2266.4   2283.1  -1129.2   2258.4     476
##
```

```
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.48045 -0.69571  0.01177  0.72728  2.72686
##
```

```
## Random effects:
##  Groups Name      Variance Std.Dev.
##  pop    (Intercept) 0.7786   0.8824
##  pop.1  precip      0.6069   0.7790
```

```

## Number of obs: 480, groups:  pop, 16
##
## Fixed effects:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.0624      0.2244   4.735 2.19e-06 ***
## precip      -1.8768      0.2077  -9.036 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr)
## precip -0.055

```

Activity

Use the `DeerEcervi.txt` data and run a binomial GLMM including random intercepts by farm and fixed (non-random) independent predictor variables for Length, Sex, and Length x Sex.