BBS Trend Summaries for Grassland Birds

BBS Case Study from Link and Barker's book

- Data from the North American Breeding Bird Survey (BBS)
- Goal of the yearly survey is to monitor bird population at large geographic scales
- Data collected for over 400 species (focus on 28 grassland species)
- Interested in measuring trend from 1966 to 1999
- Focus on summaries of temporal trend parameters β_i for species $1, ..., K$.
- For each of 28 species, have trend estimate $\hat{\beta}_i$, the corresponding standard error SE_i, and the number of sites n_i.

Read in the data

```
library(tidyverse)
(d <- read_csv("BBS_survey.csv"))
```

```
## # A tibble: 28 x 4
##    Species_Name               Trend    SE N_Site
##    <chr>                      <dbl> <dbl>  <int>
##  1 Upland Sandpiper            0.76  0.39    582
##  2 Long-billed Curlew         -0.77  1.01    222
##  3 Mountain Plover            -1.05  2.24     37
##  4 Greater Prairie-Chicken -2.54  2.33     33
##  5 Sharp-tailed Grouse        -0.92  1.43    128
##  6 Ring-necked Pheasant       -1.06  0.32   1239
##  7 Northern Harrier           -0.8   4       935
##  8 Ferruginous Hawk            3.52  1.31    200
##  9 Common Barn Owl            -2     2.14     32
## 10 Short-eared Owl            -6.23  4.55    140
## # ... with 18 more rows
```

Derived Parameters of Interest

- How many of these species have negative trends?
- That is, we are interested in the derived parameter

$$D(\beta_1, ..., \beta_{28}) = \sum_{i=1}^{28} I(\beta_i < 0)$$

Derived Parameters of Interest

- What is the actual rank of the 27th species?

$$R_{27} = \sum_{i=1}^{28} I(\beta_{27} \geq \beta_i)$$

The Model

Sampling Model:

1. Assume the estimates $\hat{\beta}_i$ are independent
2. $\hat{\beta}_i \sim N(\beta_i, 1/\tau(\hat{\beta}_i))$
3. $SE_i \sim Gamma(n/2, (n/2)\tau(\hat{\beta}_i))$

The Prior (Multilevel)

1. Assume $\beta_1, ..., \beta_{28}$ are iid $N(\mu, 1/\tau)$
2. Assign (μ, τ) a vague prior

Fitting the Model using JAGS

Write a script defining the model.

```
model_string <- "model{
  for (s in 1:28) {
      varhat[s] ~ dgamma(p[s], lam[s])
      p[s] <- n[s] / 2
      lam[s] <- p[s] * tau.betahat[s]
      tau.betahat[s] ~ dgamma(0.001, 0.001)
      sd.betahat[s] <- 1 / sqrt(tau.betahat[s])
      betahat[s] ~ dnorm(beta[s], tau.betahat[s])
      beta[s] ~ dnorm(mu, tau.beta)
  }
  mu ~ dnorm(0.0, 1.0E-6)
  tau.beta ~ dgamma(0.001, 0.001)
  sd.beta <- 1 / sqrt(tau.beta)
  another.beta ~ dnorm(mu, tau.beta)
}"
```

Define the data.

```r
d <- read_csv("BBS_survey.csv")
betahat <- d$Trend
varhat <- d$SE ^ 2
n <- d$N_Site
```

Compile the model.

```
library(rjags)
model <- jags.model(textConnection(model_string),
                data = list(varhat = varhat,
                            betahat = betahat,
                            n = n))
```

```
## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
##    Observed stochastic nodes: 56
##    Unobserved stochastic nodes: 59
##    Total graph size: 262
##
## Initializing model
```

Running JAGS

Burn-in for 10,000 iterations and run for additional 10,000 iterations, saving the output.

```
update(model, 10000, progress.bar="none")
samp <- coda.samples(model,
        variable.names=c("beta","mu", "sd.beta"),
          n.iter=10000, progress.bar="none")
```

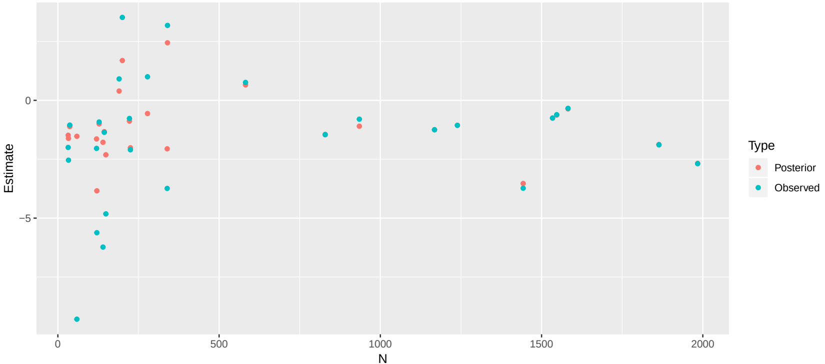
Running JAGS

Create a matrix of simulated draws.

```
S <- data.frame(as.matrix(samp))
```

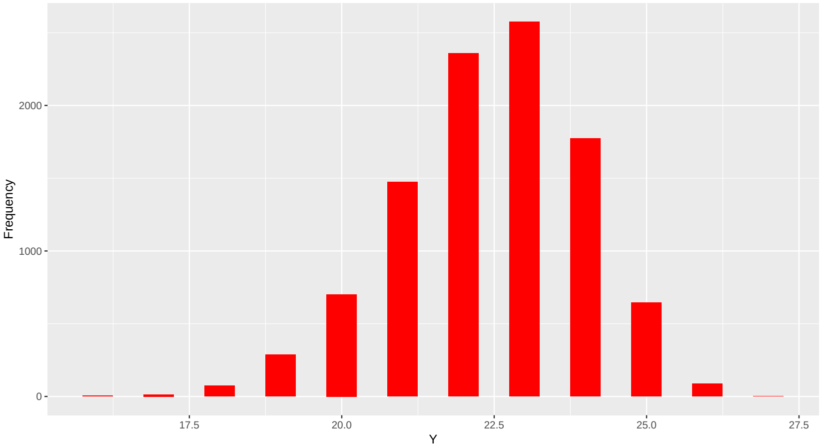
Compare the observed and multilevel estimates of trend

```
Beta <- S[, 1:28]
d1 <- data.frame(Type = "Posterior", N = n,
                 Estimate = apply(Beta, 2, mean))
d2 <- data.frame(Type = "Observed", N = n,
                 Estimate = betahat)
ggplot(rbind(d1, d2), aes(N, Estimate, color=Type)) +
  geom_point()
```



Posterior of D

```
sum.negative <- function(y) {sum(y < 0)}
D <- apply(Beta, 1, sum.negative)
bar_plot(D)
```



Posterior of rank of R_27 ?

```
Ranks <- t(apply(Beta, 1, rank))
bar_plot(Ranks[, 27]) + xlim(0, 13)
```