Stat 534: formulae referenced in lecture, week 8b: Occupancy, N-mixture and Spatial models

Occupancy modeling

- Notation:
  - X_i: characteristics of site i, may be a vector
  - Y_i: species is seen at site i,
    0 = not seen, 1 = seen
  - Z_i: species occurs at site i,
    0 = doesn't occur there, 1 = does
- The model:

$$\begin{aligned} Z_i &\sim \text{Bern}(\pi_i) \\ \pi_i &= f(\beta, X_i) \\ Y_i \mid Z_i &\sim \text{Bern}(Z_i \times g(\alpha, X_i)) \end{aligned}$$

N-mixture models:

- Estimate N and p without marking animals
- Notation:
  - N_i: # individuals at site i
  - S sites, each sampled T times
  - n_{ij}: # individuals in j sample at site i
  - p: constant, or function of known covariates
- The model (Royle 2004; Kéry, Royle and Schmid, 2005)

$$\begin{aligned} n_{ij} &\sim \text{Bin}(N_i, p) \\ N_i &\sim f(\theta) \end{aligned}$$

- log-likelihood:

$$\text{lnL}(\theta, p \mid \{n_{ij}\}) = \prod_{i=1}^{S} \left\{ \sum_{N_i=k_i}^{\infty} \left(\prod_{i=1}^{T} \text{Bin}(n_{ij} \mid N_i, p) \right) f(N_i \mid \theta) \right\}$$

- Why multiple $T > 1$ times are needed:
  - $T = 1$: $N_i \sim \text{Poiss}(\lambda_i)$, $n_i \mid N_i \sim \text{Bin}(N_i, p) \Rightarrow n_i \sim \text{Poiss}(\lambda_i p)$
  - Can't estimate λ_i or N_i, only relative abundance: λ_i/λ_j
  - $T = 2$: $n_{ij} \sim \text{Poiss}(\lambda_i p)$.
  - n_{i1} and n_{i2} are correlated
  - That correlation $\Rightarrow \hat{\lambda}_i$
  - But inference from correlations is fragile

Distance sampling: just pictures in hand-written notes

Spatial capture-recapture:

- Notation:
  - K: # capture occasions, J: # traps, N: # individuals
  - Y_{ij}: # times animal i caught in trap j $\sim \text{Bin}(K, P_{ij})$
  - p_{ij}: P[animal i caught in trap j]
  - S_i: activity center for animal i
  - X_j: location of trap j
  - α: parameters of home range / activity distributions
- The model, omitting some details
  - p_{ij} depends on home-range "shape" and distance between S_i and X_j
  - Gaussian home ranges:
    $p_{ij} = \alpha_0 \exp(-\alpha_1 \|S_i - X_j\|)$
    * α_0: capture probability when trap at activity center
    * α_1: describes "spread" of home range
  - Big assumption: home range is isotropic
    only distance matters, not direction
  - For animal i:

$$f(Y_{ij} \mid S_i) = \prod_{traps} \text{Bin}(k, p_{ij})$$

- Want $f(Y_{ij})$:

$$f(Y_{ij}) = \int f(Y_{ij} \mid S_i) f(S_i) dS_i$$

- requires a distribution for activity centers, $f(S_i)$
  * Usual assumption: activity centers equally likely throughout area A
  * $f(S_i) = 1/A$
  * integral usually unwieldy
  * approximated by numerical approximation using a grid of possible activity centers

- Really nice thing: get estimate of density: $\hat{N}/A$