Detectability, not detection

Bill Link

USGS Patuxent Wildlife Research Center

Laurel, Maryland

“Abundance”

- Count surveys: $C = N p$

“Abundance”

- Count surveys: $C = N p$
  - non-constant p implies $\left(C_1/C_2\right) \neq \left(N_1/N_2\right)$

“Abundance”

- Count surveys: $C = N\,p$
  - non-constant p implies $\left(C_1/C_2\right) \neq \left(N_1/N_2\right)$
- Solution! $\hat{N} = C\,/\,\hat{p}$

"Abundance"

- Count surveys: $C = N\,p$
  - non-constant p implies $\left(C_1/C_2\right) \neq \left(N_1/N_2\right)$
- Solution! $\hat{N} = C\,/\,\hat{p}$
- Questions
  - is there really an N ?
  - couldn't we simply analyze C ?

Analysis of *C*

$$\log\left(E(C)\right) = \log\left(N_1 p_1\right) + \log\left(N/N_1\right) + \log\left(p/p_1\right) + Noise$$

Baseline

Relative Abundance

Relative Detectability

Analysis of *C*

$$\log\left(E(C)\right) = \log\left(N_1 p_1\right) + \log\left(N/N_1\right) + \log\left(p/p_1\right) + Noise$$

Baseline

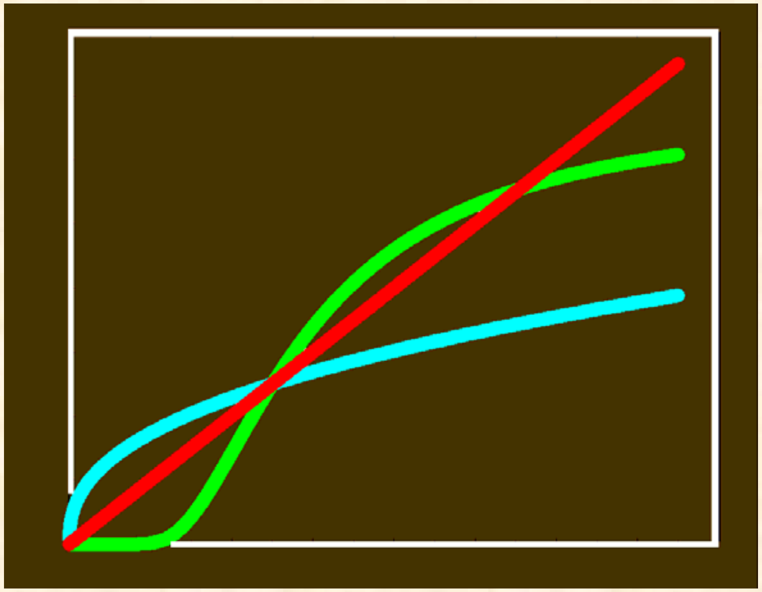
Relative Abundance

Relative Detectability

$$\log\left(E(C_t)\right) = \beta_0 + f\left(t;\theta\right) + q\left(X_t;\psi\right) + Noise$$

Analysis of *C*

$$\log\left(E(C_t)\right) = \beta_0 + f\left(t;\theta\right) + q\left(X_t;\psi\right) + Noise$$



CBC effort

Analysis of *C*

$$\log\left(E(C)\right) = \log\left(N_1 p_1\right) + \log\left(N/N_1\right) + \log\left(p/p_1\right) + Noise$$

Baseline

Relative Abundance

Relative Detectability

Unpalatable model assumptions: we want *N* !

Can we have *N* ?

Can we have *N* ?

- Sure, provided we have a good model for *p*

Can we have *N* ?

- Sure, provided we have a good model for *p*
  - $\boldsymbol{f}(p) = X\beta$, homogeneous covariate model

Can we have *N* ?

- Sure, provided we have a good model for *p*
  - $f(p) = X\beta$, homogeneous covariate model
  - $p \sim g(p; X\beta, \theta)$, heterogeneous covariate model

Can we have *N* ?

- Sure, provided we have a good model for *p*
  - $f(p) = X\beta$, homogeneous covariate model
  - $p \sim g(p; X\beta, \theta)$, heterogeneous covariate model

- The simplest heterogeneity model: M_h
  - *T* sampling occasions
  - individual specific, time invariant $p \sim g(p;\theta)$
  - sufficient statistics: f_j = # individuals sighted *j* times,

$$n = \sum_{j=1}^{T} f_j \;, \quad N = n + f_0$$

A simple example

- $T = 4$; $n = 195$
  - $f_0 = ??$, $f_1 = 84$, $f_2 = 54$, $f_3 = 36$, $f_4 = 21$

A simple example

- $T = 4$; $n = 195$
  - $f_0 = ??$, $f_1 = 84$, $f_2 = 54$, $f_3 = 36$, $f_4 = 21$
  - $e_0 = 20$, $e_1 = 66$, $e_2 = 79$, $e_3 = 42$, $e_4 = 8$
  - huge evidence of heterogeneity

A simple example

- $T = 4$; $n = 195$
  - $f_1 = 84$, $f_2 = 54$, $f_3 = 36$, $f_4 = 21$
  - huge evidence of heterogeneity

- Perfect fit using 2-point mixture distribution
  - 75% mass on ¼, 25% mass on ¾
  - indicated $f_0 = 61$

A simple example

- $T = 4$; $n = 195$
  - $f_1 = 84$, $f_2 = 54$, $f_3 = 36$, $f_4 = 21$
  - huge evidence of heterogeneity

- Perfect fit using 2-point mixture distribution
  - 75% mass on ¼, 25% mass on ¾
  - indicated $f_0 = 61$

- Perfect fit using $\beta(1/2, 3/2)$ distribution

A simple example

- $T = 4$; $n = 195$
  - $f_1 = 84$, $f_2 = 54$, $f_3 = 36$, $f_4 = 21$
  - huge evidence of heterogeneity

- Perfect fit using 2-point mixture distribution
  - 75% mass on ¼, 25% mass on ¾
  - indicated $f_0 = 61$

- Perfect fit using $\beta(1/2, 3/2)$ distribution
  - indicated $f_0 = 189$

A simple example

- $T = 4$; $n = 195$
  - $f_1 = 84$, $f_2 = 54$, $f_3 = 36$, $f_4 = 21$
  - huge evidence of heterogeneity

- Perfect fit using 2-point mixture distribution !
  - 75% mass on ¼, 25% mass on ¾
  - indicated $f_0 = 61$

- Perfect fit using $\beta(1/2, 3/2)$ distribution !
  - indicated $f_0 = 189$

- Ouch

N not identifiable

- Not a quirk of small *T*
- Model averaging?
- Arbitrary choice of heterogeneity distribution unacceptable!
- Solutions:
  - having a good reason to choose a particular family for $g(p)$
  - identification of covariates sufficient to account for heterogeneity

Perspectives

- Given heterogeneity in detectability (beyond covariates) detection estimation can have as serious problems as count surveys.
- The real issue is identifying and controlling for sources of variation in detectability
- Detection methods generally superior, simply because better data are collected
- Still … ought not to disparage count surveys
  - attempts to standardize
  - identification of sources of variation in detectability
  - worst case: detectability associated with abundance

Seasonal components of population trend

- Carolina wren
  - Nonmigratory
  - susceptible to severe winters
- BBS data (June)
  - observer effects (among, within)
- CBC data (December)
  - effort effects
- Joint hierarchical analysis
  - 70 strata
  - Proportion of variation in population size associated with winters ~ $N(0.82, 0.05^2)$