Estimating Binomial Index $N$ with Application to Bird and Bat Mortality at Wind and Solar Power Facilities

Lisa Madsen $^1$  Dan Dalthorp $^2$  Manuela Huso $^{2,1}$

$^1$ Oregon State University

$^2$ United States Geological Survey

TIES 2018
Bats and Wind Farms
Bats and Wind Farms
Bats and Wind Farms
Bats and Wind Farms
Bats and Wind Farms
Bats and Wind Farms
Bats and Wind Farms
Bats and Wind Farms
Bats and Wind Farms
Outline

Models and Estimators
   Known detection probability
   Unknown detection probability

Simulation Study

Wind Farm Example

R Package

Summary
Count Model

Notation and assumptions:

- $N$ carcasses total (unknown)
Count Model

Notation and assumptions:

- $N$ carcasses total (unknown)
- $x$ carcasses found
Count Model

Notation and assumptions:

- $N$ carcasses total (unknown)
- $x$ carcasses found
- Each carcass found independently with detection probability $p$
Count Model

Notation and assumptions:

- \( N \) carcasses total (unknown)
- \( x \) carcasses found
- Each carcass found independently with detection probability \( p \)

Then \( x \) is a realization of \( X \sim \text{Binomial}(N, p) \).
Count Model

Notation and assumptions:

- $N$ carcasses total (unknown)
- $x$ carcasses found
- Each carcass found independently with detection probability $p$

Then $x$ is a realization of $X \sim \text{Binomial}(N, p)$. 
Outline

Models and Estimators

Known detection probability

Unknown detection probability

Simulation Study

Wind Farm Example

R Package

Summary
Unbiased Estimator

If $p$ is known, then $E(X/p) = N$. 
Unbiased Estimator

If $p$ is known, then $E(X/p) = N$.

If $X$ is large, $X/p \overset{\text{approx}}{\sim} \text{Normal}(N, N(1 - p)/p)$. 
Unbiased Estimator

If $p$ is known, then $E(X/p) = N$.

If $X$ is large, $X/p \overset{\text{approx}}{\sim} \text{Normal}(N, N(1 - p)/p)$.

Rather than relying on asymptotic normality, we employ a **parametric bootstrap** to simulate the sampling distribution of the estimator.
Unbiased Estimator

If \( p \) is known, then \( E(X/p) = N \).

If \( X \) is large, \( X/p \overset{\text{approx}}{\sim} \text{Normal}(N, N(1 - p)/p) \).

Rather than relying on asymptotic normality, we employ a **parametric bootstrap** to simulate the sampling distribution of the estimator.

- Flexibility in inference
Known detection probability

Unbiased Estimator

If $p$ is known, then $E(X/p) = N$.

If $X$ is large, $X/p \approx \text{Normal}(N, N(1-p)/p)$.

Rather than relying on asymptotic normality, we employ a **parametric bootstrap** to simulate the sampling distribution of the estimator.

- Flexibility in inference
- Provides framework to augment model
Parametric bootstrap vs. non-parametric bootstrap

Non-parametric bootstrap draws samples from observed data (with replacement).
Parametric bootstrap vs. non-parametric bootstrap

Non-parametric bootstrap draws samples from observed data (with replacement).

Parametric bootstrap

- Assumes data drawn from a given distribution, e.g. binomial
Parametric bootstrap vs. non-parametric bootstrap

Non-parametric bootstrap draws samples from observed data (with replacement).

Parametric bootstrap

- Assumes data drawn from a given distribution, e.g. binomial
- Uses data to estimate parameters
Parametric bootstrap vs. non-parametric bootstrap

Non-parametric bootstrap draws samples from observed data (with replacement).

Parametric bootstrap

- Assumes data drawn from a given distribution, e.g. binomial
- Uses data to estimate parameters
- Draws samples from assumed distribution with parameters equal to estimates
Parametric Bootstrap of $X$

Given data $X = x$, simulate sampling distribution of $X$:
Parametric Bootstrap of $X$

Given data $X = x$, simulate sampling distribution of $X$:

1. Point estimate of $N$ is $x/p$. 
Parametric Bootstrap of $X$

Given data $X = x$, simulate sampling distribution of $X$:

1. Point estimate of $N$ is $x/p$.
2. Simulate $\tilde{X} \sim \text{Binomial}(x/p, p)$. 
Parametric Bootstrap of $X$

Given data $X = x$, simulate sampling distribution of $X$:

1. Point estimate of $N$ is $x/p$.
2. Simulate $\tilde{X} \sim \text{Binomial}(x/p, p)$.

Since $x/p$ is not necessarily integer-valued, simulate from the continuous binomial (CB) distribution [Ilienko, 2013, Dalthorp, 2018].
Known detection probability

Binomial Distribution

\[ X \sim B(N=5, p=0.4) \]
Continuous Binomial Distribution

$X \sim B(N=5, p=0.4)$

$X \sim CB(N=5, p=0.4)$
Continuous Binomial Distribution

\[ X \sim B(N=5, p=0.4) \]

\[ X \sim CB(N=5, p=0.4) \]

\[ X \sim CB(N=6.8, p=0.4) \]
Continuous Binomial Distribution

If \( X \sim CB(N, p) \), then
- Parameter space: \( N \geq 0 \) and \( 0 \leq p \leq 1 \)
Continuous Binomial Distribution

If \( X \sim CB(N, p) \), then

- Parameter space: \( N \geq 0 \) and \( 0 \leq p \leq 1 \)
- Support: \( 0 \leq X \leq N + 1 \)
Continuous Binomial Distribution

If $X \sim CB(N, p)$, then

- Parameter space: $N \geq 0$ and $0 \leq p \leq 1$
- Support: $0 \leq X \leq N + 1$
- $E(X) \equiv \mu_{CB}(N, p) \approx Np + 1/2$
Parametric Bootstrap of $X/p$

1. Estimate $N$ as $x/p$. 
Parametric Bootstrap of $X/p$

1. Estimate $N$ as $x/p$.
2. Simulate $\tilde{X} \sim CB(x/p, p)$
Parametric Bootstrap of $X/p$

1. Estimate $N$ as $x/p$.
2. Simulate $\tilde{X} \sim \text{CB}(x/p, p) - \mu_{CB}(x/p, p) + x$. 

Known detection probability
Parametric Bootstrap of $X/p$

1. Estimate $N$ as $x/p$.
2. Simulate $\tilde{X} \sim \text{CB}(x/p, p) - \mu_{\text{CB}}(x/p, p) + x$.
3. Iterate 2 to obtain
   $$\tilde{X}^{(1)}, \ldots, \tilde{X}^{(S)},$$
   for, say, $S = 1000$. 
Known detection probability

## Parametric Bootstrap of $X/p$

1. Estimate $N$ as $x/p$.
2. Simulate $\tilde{X} \sim \text{CB}(x/p, p) - \mu_{\text{CB}}(x/p, p) + x$.
3. Iterate 2 to obtain $\tilde{X}^{(1)}, \ldots, \tilde{X}^{(S)}$, for, say, $S = 1000$.
4. Form $\tilde{N}^{(s)} \equiv \tilde{X}^{(s)}/p$ to obtain $\tilde{N}^{(1)}, \ldots, \tilde{N}^{(S)}$. 


Parametric Bootstrap

Histogram of $X/p$

Histogram of $\tilde{N}$
Outline

Models and Estimators

Known detection probability

Unknown detection probability

Simulation Study

Wind Farm Example

R Package

Summary
\[ \hat{N} = \frac{X}{\hat{p}} \] must reflect uncertainty about \( \hat{p} \).
Assume we have an independent study to estimate detection probability:

- $M$ carcasses (known)
Modeling Detection Probability

Assume we have an independent study to estimate detection probability:

- $M$ carcasses (known)
- $y$ of these found
Modeling Detection Probability

Assume we have an independent study to estimate detection probability:

- $M$ carcasses (known)
- $y$ of these found
- Model $\logit(p_i) = \beta_0 + \beta_1 W_{i1} + \ldots + \beta_q W_{iq}$,
Modeling Detection Probability

Assume we have an independent study to estimate detection probability:

- $M$ carcasses (known)
- $y$ of these found

Model $\logit(p_i) = \beta_0 + \beta_1 W_{i1} + \ldots + \beta_q W_{iq}$,

- $p_i$ is the probability of detecting the $i$th carcass
Modeling Detection Probability

Assume we have an independent study to estimate detection probability:

- $M$ carcasses (known)
- $y$ of these found
- Model $\logit(p_i) = \beta_0 + \beta_1 W_{i1} + \ldots + \beta_q W_{iq}$,
  - $p_i$ is the probability of detecting the $i$th carcass
  - $W_i = (W_{i1}, \ldots, W_{iq})$ is a vector of covariates associated with the $i$th carcass (e.g. species, ground cover)
Unknown detection probability

## Estimating $p_i$

If $\beta = (\beta_0, \ldots, \beta_q)'$,

- Obtain $\hat{\beta}$ by logistic regression
Estimating $p_i$

If $\beta = (\beta_0, \ldots, \beta_q)'$,

- Obtain $\hat{\beta}$ by logistic regression
- $\hat{\beta}$ is asymptotically Normal($\beta, I^{-1}(\beta)$)
Unknown detection probability

Estimating $p_i$

If $\beta = (\beta_0, \ldots, \beta_q)'$, then:

- Obtain $\hat{\beta}$ by logistic regression
- $\hat{\beta}$ is asymptotically $\text{Normal}(\beta, I^{-1}(\beta))$
- $\hat{p}_i = \text{antilogit}(W_i\hat{\beta}) = 1/(1 + e^{-W_i\hat{\beta}})$
Estimating $p_i$

If $\bm{\beta} = (\beta_0, \ldots, \beta_q)'$,

- Obtain $\hat{\bm{\beta}}$ by logistic regression
- $\hat{\bm{\beta}}$ is asymptotically Normal$(\bm{\beta}, I^{-1}(\bm{\beta}))$
- $\hat{p}_i = \text{antilogit}(\bm{W}_i\hat{\bm{\beta}}) = 1/(1 + e^{-\bm{W}_i\hat{\bm{\beta}}})$

In the following, we suppress the subscript $i$ and suppose an intercept-only model.
Parametric Bootstrap of $\hat{p}$

For $s = 1, \ldots, S$,

1. Simulate $\hat{\beta}(s) \sim \text{Normal}(\hat{\beta}, I^{-1}(\hat{\beta}))$. 
Parametric Bootstrap of $\hat{p}$

For $s = 1, \ldots, S$,

1. Simulate $\hat{\beta}^{(s)} \sim \text{Normal}(\hat{\beta}, I^{-1}(\hat{\beta}))$.
2. Calculate

$$\hat{p}^{(s)} = \frac{1}{1 + e^{-\hat{\beta}^{(s)}}}.$$
Parametric Bootstrap of $\hat{p}$

For $s = 1, \ldots, S$,

1. Simulate $\hat{\beta}^{(s)} \sim \text{Normal}(\hat{\beta}, I^{-1}(\hat{\beta}))$.
2. Calculate

$$\hat{p}^{(s)} = \frac{1}{1 + e^{-\hat{\beta}^{(s)}}}.$$ 

Simulated sampling distribution: $\hat{p}^{(1)}, \ldots, \hat{p}^{(S)}$
Parametric Bootstrap of $\hat{N} = X/\hat{p}$

For $s = 1, \ldots, S$,

1. Simulate $\hat{\beta}^{(s)} \sim \text{Normal}(\hat{\beta}, I^{-1}(\hat{\beta}))$.
2. Calculate

$$\hat{p}^{(s)} = \frac{1}{1 + e^{-\hat{\beta}^{(s)}}}.$$
Parametric Bootstrap of $\hat{N} = X/\hat{p}$

For $s = 1, \ldots, S$,

1. Simulate $\hat{\beta}^{(s)} \sim \text{Normal}(\hat{\beta}, I^{-1}(\hat{\beta}))$.
2. Calculate $\hat{p}^{(s)} = \frac{1}{1 + e^{-\hat{\beta}^{(s)}}}$.
3. Simulate $\tilde{X}^{(s)} \sim \text{CB}(x/\hat{p}^{(s)}, \hat{p}^{(s)}) - \mu_{\text{CB}}(x/\hat{p}^{(s)}, \hat{p}^{(s)}) + x$. 
Unknown detection probability

**Parametric Bootstrap of \( \hat{N} = X/\hat{p} \)**

For \( s = 1, \ldots, S \),

1. Simulate \( \hat{\beta}^{(s)} \sim \text{Normal}(\hat{\beta}, I^{-1}(\hat{\beta})) \).
2. Calculate

\[
\hat{p}^{(s)} = \frac{1}{1 + e^{-\hat{\beta}^{(s)}}}.
\]

3. Simulate

\[
\tilde{X}^{(s)} \sim \text{CB}(x/\hat{p}^{(s)}, \hat{p}^{(s)}) - \mu_{CB}(x/\hat{p}^{(s)}, \hat{p}^{(s)}) + x.
\]
Parametric Bootstrap of $\hat{N} = X/\hat{p}$

For $s = 1, \ldots, S$,

1. Simulate $\hat{\beta}^{(s)} \sim \text{Normal}(\hat{\beta}, I^{-1}(\hat{\beta}))$.

2. Calculate

   $$\hat{p}^{(s)} = \frac{1}{1 + e^{-\hat{\beta}^{(s)}}}.$$ 

3. Simulate

   $$\tilde{X}^{(s)} \sim \text{CB}(x/\hat{p}^{(s)}, \hat{p}^{(s)}) - \mu_{CB}(x/\hat{p}^{(s)}, \hat{p}^{(s)}) + x.$$ 

4. Calculate $\tilde{N}^{(s)} \equiv \tilde{X}^{(s)}/\hat{p}^{(s)}$. 

Bootstrapped sampling distribution of $\hat{N}$ : $\tilde{N}^{(1)}, \ldots, \tilde{N}^{(S)}$. 

Unknown detection probability
Unknown detection probability

Parametric Bootstrap of $\hat{N} = X/\hat{p}$

For $s = 1, \ldots, S$,

1. Simulate $\hat{\beta}^{(s)} \sim \text{Normal}(\hat{\beta}, I^{-1}(\hat{\beta}))$.
2. Calculate

   $$\hat{p}^{(s)} = \frac{1}{1 + e^{-\hat{\beta}^{(s)}}}.$$ 

3. Simulate

   $$\tilde{X}^{(s)} \sim \text{CB}(x/\hat{p}^{(s)}, \hat{p}^{(s)}) - \mu_{\text{CB}}(x/\hat{p}^{(s)}, \hat{p}^{(s)}) + x.$$ 

4. Calculate $\tilde{N}^{(s)} \equiv \tilde{X}^{(s)}/\hat{p}^{(s)}$.

Bootstrapped sampling distribution of $\hat{N}$: $\tilde{N}^{(1)}, \ldots, \tilde{N}^{(S)}$.
Confidence Intervals

- \( y = 49 \) of \( M = 100 \) test carcasses found
- \( x = 56 \) search carcasses found
- \( S = 1000 \) realizations of \( \tilde{N} \)
Unknown detection probability

Confidence Intervals

- $y = 49$ of $M = 100$ test carcasses found
- $x = 56$ search carcasses found
- $S = 1000$ realizations of $\tilde{N}$
Confidence Intervals

- \( y = 49 \) of \( M = 100 \) test carcasses found
- \( x = 56 \) search carcasses found
- \( S = 1000 \) realizations of \( \tilde{N} \)

90% confidence interval for \( N \):

\( (90.4, 142.7) \)
## Simulation Scenarios

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>True binomial index</td>
<td>$N = 10, 100, 1000$</td>
</tr>
<tr>
<td>Detection probability</td>
<td>$p = 0.2, 0.5, 0.85$</td>
</tr>
<tr>
<td>Field trial index</td>
<td>$M = 30, 100, 1000$</td>
</tr>
</tbody>
</table>
## Simulation Scenarios

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>True binomial index</td>
<td>$N = 10, 100, 1000$</td>
</tr>
<tr>
<td>Detection probability</td>
<td>$p = 0.2, 0.5, 0.85$</td>
</tr>
<tr>
<td>Field trial index</td>
<td>$M = 30, 100, 1000$</td>
</tr>
<tr>
<td>Confidence levels</td>
<td>0.99, 0.95, 0.90, 0.80, 0.50</td>
</tr>
</tbody>
</table>
Boxplots of Medians

$p = 0.2$

$p = 0.5$

$p = 0.85$
Confidence Coverage

N = 10  N = 100  N = 1000

α = 0.01  α = 0.05  α = 0.1  α = 0.2  α = 0.5

CI coverage  Target
Confidence Coverage

![Confidence Coverage graph](image)

Problems when $P(X = 0)$ is large
Data

- Western EcoSystems Technology, Inc. study [Good et al., 2012]
**Data**

- Western EcoSystems Technology, Inc. study [Good et al., 2012]

- $x = 321$ carcasses found on roads and pads of 168 wind turbines at Fowler Ridge Wind Farm, Indiana, USA, fall, 2011
Data

- Western EcoSystems Technology, Inc. study [Good et al., 2012]
- \( x = 321 \) carcasses found on roads and pads of 168 wind turbines at Fowler Ridge Wind Farm, Indiana, USA, fall, 2011
Data

- Western EcoSystems Technology, Inc. study [Good et al., 2012]
- \( x = 321 \) carcasses found on roads and pads of 168 wind turbines at Fowler Ridge Wind Farm, Indiana, USA, fall, 2011

- We limit inferential scope to the 168 observed turbines (out of 355).
Field Trials

- $y = 84$ of $M = 104$ carcasses found.
Field Trials

- \( y = 84 \) of \( M = 104 \) carcasses found.
- Intercept-only logistic model: \( \hat{\beta} = 1.44 \) and \( \text{var}(\hat{\beta}) = 0.249 \).
Field Trials

- $y = 84$ of $M = 104$ carcasses found.
- Intercept-only logistic model: $\hat{\beta} = 1.44$ and $\text{var}(\hat{\beta}) = 0.249$.
- Simulate $\hat{\beta}(s) \sim \text{Normal}(1.44, 0.249), s = 1, \ldots, 50,000.$
Field Trials

- \( y = 84 \) of \( M = 104 \) carcasses found.
- Intercept-only logistic model: \( \hat{\beta} = 1.44 \) and \( \text{var}(\hat{\beta}) = 0.249 \).
- Simulate \( \hat{\beta}(s) \sim \text{Normal}(1.44, 0.249) \), \( s = 1, \ldots, 50,000 \).
- Calculate

\[
\hat{p}(s) = \frac{1}{1 + e^{-\hat{\beta}(s)}}.
\]
Field Trials

- $y = 84$ of $M = 104$ carcasses found.
- Intercept-only logistic model: $\hat{\beta} = 1.44$ and $\text{var}(\hat{\beta}) = 0.249$.
- Simulate $\hat{\beta}(s) \sim \text{Normal}(1.44, 0.249)$, $s = 1, \ldots, 50,000$.
- Calculate

$$\hat{p}(s) = \frac{1}{1 + e^{-\hat{\beta}(s)}}.$$ 

Simulated sampling distribution: $\hat{p}^{(1)}, \ldots, \hat{p}^{(50,000)}$
Varying Density and Incomplete Sampling
Varying Density and Incomplete Sampling
# Varying Density and Incomplete Sampling

Distance from turbine (m) | Fatalities (%) | Area (%)
---|---|---
$d_k$ | $f_k$ | $a_k$
---|---|---
0 - 10 | 0.0473 | 0.9999
10 - 20 | 0.225 | 0.3963
20 - 30 | 0.105 | 0.1405
30 - 40 | 0.129 | 0.0965
40 - 50 | 0.033 | 0.0484
50 - 60 | 0.03 | 0.0205
60 - 70 | 0 | 0.0171
70 - 80 | 0.006 | 0.0147

Proportion of carcasses on roads and pads is approximately $\frac{1}{\sum_{k=1}^{8} f_k/a_k} = 0.1904$. Therefore, $\approx 0.1904 \times p$ carcasses found.
Varying Density and Incomplete Sampling

<table>
<thead>
<tr>
<th>Distance from turbine (m)</th>
<th>Fatalities (%)</th>
<th>Area (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 10</td>
<td>0.0473</td>
<td>0.9999</td>
</tr>
<tr>
<td>10 - 20</td>
<td>0.225</td>
<td>0.3963</td>
</tr>
<tr>
<td>20 - 30</td>
<td>0.105</td>
<td>0.1405</td>
</tr>
<tr>
<td>30 - 40</td>
<td>0.129</td>
<td>0.0965</td>
</tr>
<tr>
<td>40 - 50</td>
<td>0.033</td>
<td>0.0484</td>
</tr>
<tr>
<td>50 - 60</td>
<td>0.03</td>
<td>0.0205</td>
</tr>
<tr>
<td>60 - 70</td>
<td>0</td>
<td>0.0171</td>
</tr>
<tr>
<td>70 - 80</td>
<td>0.006</td>
<td>0.0147</td>
</tr>
</tbody>
</table>

Proportion of carcasses on roads and pads is approximately

\[
1 / \sum_{k=1}^{8} \left( f_k / a_k \right) = 0.1904
\]
## Varying Density and Incomplete Sampling

<table>
<thead>
<tr>
<th>Distance from turbine (m)</th>
<th>Fatalities (%)</th>
<th>Area (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_k$</td>
<td>$f_k$</td>
<td>$a_k$</td>
</tr>
<tr>
<td>0 - 10</td>
<td>0.0473</td>
<td>0.9999</td>
</tr>
<tr>
<td>10 - 20</td>
<td>0.225</td>
<td>0.3963</td>
</tr>
<tr>
<td>20 - 30</td>
<td>0.105</td>
<td>0.1405</td>
</tr>
<tr>
<td>30 - 40</td>
<td>0.129</td>
<td>0.0965</td>
</tr>
<tr>
<td>40 - 50</td>
<td>0.033</td>
<td>0.0484</td>
</tr>
<tr>
<td>50 - 60</td>
<td>0.03</td>
<td>0.0205</td>
</tr>
<tr>
<td>60 - 70</td>
<td>0</td>
<td>0.0171</td>
</tr>
<tr>
<td>70 - 80</td>
<td>0.006</td>
<td>0.0147</td>
</tr>
</tbody>
</table>

Proportion of carcasses on roads and pads is approximately

\[ 1/ \sum_{k=1}^{8} (f_k/a_k) = 0.1904 \quad \therefore \approx 0.1904 \times p \text{ carcasses found.} \]
Parametric Bootstrap of $\hat{N}$

For $x = 321$ and $s = 1, \ldots, 50,000$,

1. Simulate

$$\tilde{X}^{(s)} \sim \text{CB}(x/(0.1904 \cdot \hat{p}^{(s)}), 0.1904 \cdot \hat{p}^{(s)})$$

$$- \mu_{CB}(x/(0.1904 \cdot \hat{p}^{(s)}), 0.1904 \cdot \hat{p}^{(s)}) + x.$$
Parametric Bootstrap of $\hat{N}$

For $x = 321$ and $s = 1, \ldots, 50,000,$

1. Simulate

\[
\tilde{X}^{(s)} \sim \text{CB}(x/(0.1904 \cdot \hat{p}^{(s)}), 0.1904 \cdot \hat{p}^{(s)}) - \mu_{\text{CB}}(x/(0.1904 \cdot \hat{p}^{(s)}), 0.1904 \cdot \hat{p}^{(s)}) + x.
\]

2. Calculate $\tilde{N}^{(s)} \equiv \tilde{X}^{(s)}/(0.1904 \cdot \hat{p}^{(s)}).$
Parametric Bootstrap of $\hat{N}$

For $x = 321$ and $s = 1, \ldots, 50,000$,

1. Simulate

\[
\tilde{X}^{(s)} \sim \text{CB}(x/(0.1904 \cdot \hat{p}^{(s)}), 0.1904 \cdot \hat{p}^{(s)}) - \mu_{\text{CB}}(x/(0.1904 \cdot \hat{p}^{(s)}), 0.1904 \cdot \hat{p}^{(s)}) + x.
\]

2. Calculate $\tilde{N}^{(s)} \equiv \tilde{X}^{(s)}/(0.1904 \cdot \hat{p}^{(s)})$.

This analysis supposes the 0.1904 is without error.
Confidence Interval

Histogram

Frequency

N
Confidence Interval

Histogram

95% confidence interval for $N$: $(1831.7, 2423.6)$
Confidence Interval

Histogram

95% confidence interval for $N$:

$(1831.7, 2423.6)$
Limitations

- Method collapses if $x = 0$
Limitations

- Method collapses if $x = 0$
  Use Evidence of Absence [Dalthorp et al., 2017]
Limitations

- Method collapses if $x = 0$
  - Use Evidence of Absence [Dalthorp et al., 2017]
- Confidence coverage is off when $P(X = 0)$ is large.
Limitations

- Method collapses if $x = 0$
  - Use Evidence of Absence [Dalthorp et al., 2017]
- Confidence coverage is off when $P(X = 0)$ is large.
- Field trial index $M$ must be adequate so that $y \neq 0$. 
R Package for Generalized Estimator

R package GenEst [Dalthorp et al., 2018] currently in beta review.
R Package for Generalized Estimator

R package GenEst [Dalthorp et al., 2018] currently in beta review.

- Implements sophisticated model to estimate total mortality at wind or solar facilities
R Package for Generalized Estimator

R package \texttt{GenEst} [Dalthorp et al., 2018] currently in beta review.

- Implements sophisticated model to estimate total mortality at wind or solar facilities
- Applicable when goal is to estimate population total when detection is imperfect (and modeled separately)
R Package for Generalized Estimator

R package GenEst [Dalthorp et al., 2018] currently in beta review.

- Implements sophisticated model to estimate total mortality at wind or solar facilities
- Applicable when goal is to estimate population total when detection is imperfect (and modeled separately)
- Inference by parametric bootstrap
R Package for Generalized Estimator

R package \texttt{GenEst} [Dalthorp et al., 2018] currently in beta review.

- Implements sophisticated model to estimate total mortality at wind or solar facilities
- Applicable when goal is to estimate population total when detection is imperfect (and modeled separately)
- Inference by parametric bootstrap
- Optional \texttt{shiny} [Chang et al., 2018] Graphical User Interface
Modeling detection probability

- Estimate **searcher efficiency**, $P(\text{carcass found}|\text{carcass is available for discovery})$
Modeling detection probability

- Estimate **searcher efficiency**, \( P(\text{carcass found}|\text{carcass is available for discovery}) \)
- Estimate **carcass persistence**, \( P(\text{carcass persists until a given time}) \)
Modeling detection probability

- Estimate **searcher efficiency**, $P(\text{carcass found|carcass is available for discovery})$
- Estimate **carcass persistence**, $P(\text{carcass persists until a given time})$
Modeling detection probability

- Estimate **searcher efficiency**, 
  \[ P(\text{carcass found}|\text{carcass is available for discovery}) \]

- Estimate **carcass persistence**, 
  \[ P(\text{carcass persists until a given time}) \]

- Estimate **arrival probabilities**, 
  \[ P(\text{carcass arrives in interval } (t_{j-1}, t_j]) \]
Modeling detection probability

- Estimate **searcher efficiency**, 
  \[ P(\text{carcass found}|\text{carcass is available for discovery}) \]

- Estimate **carcass persistence**, 
  \[ P(\text{carcass persists until a given time}) \]

- Estimate **arrival probabilities**, 
  \[ P(\text{carcass arrives in interval } (t_{j-1}, t_j]) \]

Bootstrap each of these quantities and combine to estimate sampling distribution of \( P(\text{carcass observed}) \).
Model selection

- Choose covariates for searcher efficiency model.
Model selection

- Choose covariates for searcher efficiency model.
- Choose covariates and survival model for carcass persistence model.
Splits

Total carcasses found: $x$. 
Splits

Total carcasses found: $x$.

Carcass $i$ has a potentially unique set of attributes.
Splits

Total carcasses found: $x$.

Carcass $i$ has a potentially unique set of attributes

- Covariates used in searcher efficiency and carcass persistence models
Splits

Total carcasses found: $x$.

Carcass $i$ has a potentially unique set of attributes

- Covariates used in searcher efficiency and carcass persistence models
- Other information *not* used in these models (e.g. arrival time)
Splits

Total carcasses found: $x$.

Carcass $i$ has a potentially unique set of attributes

- Covariates used in searcher efficiency and carcass persistence models
- Other information *not* used in these models (e.g. arrival time)

Would like to categorize mortality according to an arbitrary combination of attributes.
Splits

Earlier we simulated

$$\tilde{X}^{(s)} \sim \text{CB}(x/\hat{p}^{(s)}, \hat{p}^{(s)}) - \mu_{\text{CB}}(x/\hat{p}^{(s)}, \hat{p}^{(s)}) + x.$$ 

and calculated $$\tilde{N}^{(s)} \equiv \tilde{X}^{(s)}/\hat{p}^{(s)}.$$
Splits

Earlier we simulated

\[ \tilde{X}^{(s)} \sim CB\left(x/\hat{p}^{(s)}, \hat{p}^{(s)}\right) - \mu_{CB}\left(x/\hat{p}^{(s)}, \hat{p}^{(s)}\right) + x. \]

and calculated \( \tilde{N}^{(s)} \equiv \tilde{X}^{(s)}/\hat{p}^{(s)}. \)

Instead, take \( i \)th carcass individually to simulate

\[ \tilde{X}_i^{(s)} \sim CB\left(1/\hat{p}_i^{(s)}, \hat{p}_i^{(s)}\right) - \mu_{CB}\left(1/p_i^{(s)}, p_i^{(s)}\right) + 1, \]

and calculate \( \tilde{N}_i^{(s)} \equiv \tilde{X}_i^{(s)}/\hat{p}_i^{(s)}. \)
Splits

Earlier we simulated

$$\tilde{X}^{(s)} \sim CB(x/\hat{p}^{(s)}, \hat{p}^{(s)}) - \mu_{CB}(x/\hat{p}^{(s)}, \hat{p}^{(s)}) + x.$$ 

and calculated $$\tilde{N}^{(s)} \equiv \tilde{X}^{(s)}/\hat{p}^{(s)}.$$ 

Instead, take $i$th carcass individually to simulate

$$\tilde{X}_{i}^{(s)} \sim CB(1/\hat{p}_{i}^{(s)}, \hat{p}_{i}^{(s)}) - \mu_{CB}(1/p_{i}^{(s)}, p_{i}^{(s)}) + 1,$$

and calculate $$\tilde{N}_{i}^{(s)} \equiv \tilde{X}_{i}^{(s)}/\hat{p}_{i}^{(s)}.$$ 

$$\tilde{N}_{i}^{(1)}, \ldots, \tilde{N}_{i}^{(S)}$$ simulates the sampling distribution of the estimated total number of carcasses represented by $i$th carcass.
Splits

Suppose carcasses 1, \ldots, k represent category C for which we want to summarize estimated total mortality.

\[
\begin{bmatrix}
\tilde{N}_1^{(1)} & \cdots & \tilde{N}_1^{(S)} \\
\vdots & \ddots & \vdots \\
\tilde{N}_k^{(1)} & \cdots & \tilde{N}_k^{(S)}
\end{bmatrix}.
\]

Summing down the columns yields a parametric bootstrapped sampling distribution for category C,

\[
\left( \sum_{i=1}^{k} \tilde{N}_i^{(1)}, \ldots, \sum_{i=1}^{k} \tilde{N}_i^{(S)} \right).
\]
Summary

GenEst models important processes contributing to detection probability:

- Searcher efficiency
- Carcass persistence
- Arrival process

Models combined using parametric bootstrap to account for uncertainty from each model.
Summary

Parametric bootstrap for $\hat{p}$ combined with parametric bootstrap of $X$ to account for uncertainty in $\hat{p}$ and binomial sampling variability.

Result is simulated sampling distribution of $\hat{N}$, not just point estimate and standard error.
Thanks to Dan Dalthorp and Manuela Huso.


