Choosing a Detection function
Overview

Formal definition
Criteria for a good detection function model
Key functions and adjustment terms
Fitting models in Distance
Choosing the number of parameters
Introduction to truncation
Formal definition

The **detection function** describes the relationship between distance and the probability of detection.

Formally denoted by $g(x)$ (usually referred to as ‘g of x’)

$$g(x) = \text{the probability of detecting an animal, given that it is at distance x from the line}$$

Key to the concept of distance sampling
The detection function, $g(x)$

We assume $g(0) = 1$.

Histogram bars are scaled.

\[
\hat{P}_a = \frac{\text{area under curve}}{\text{area under rectangle}} = \frac{\int_0^w \hat{g}(x)dx}{w}
\]
Modelling $g(x)$

$g(x)$ represents the **underlying** relationship between detection probability and distance

However, the true form of $g(x)$ is unknown to us

We need to **estimate** $g(x)$ by fitting a **model** to our data

i.e., we need to find a curve that will approximate the underlying relationship
Criteria for robust estimation

Four main criteria for a good model:

1. **Model robustness** – use a model that will fit a wide variety of plausible shapes for $g(x)$
2. **Shape criterion** – use a model with a ‘shoulder’ – i.e. $g'(0)=0$
3. **Pooling robustness** – use a model for the average detection function, even when many factors affect detectability
4. **Estimator efficiency** – use a model that will lead to a precise estimator of density
Key functions

The first step in constructing a model for $g(x)$ is to choose a **key function**

This determines the basic model shape

Four key functions available in Distance:

1. Uniform
2. Half normal
3. Hazard rate
4. Negative exponential
Key functions (cont.)

- Model formula:
  \[ g(x) = 1, \quad x \leq w \]
- Parameters = 0
- Shape criterion?
  Yes
- Model robust?
  No
Key functions (cont.)

- Model formula:
  \[ g(x) = \exp\left(\frac{-x^2}{2\sigma^2}\right), \quad x \leq w \]
- Parameters = 1
- Shape criterion?
  Yes
- Model robust?
  No
Key functions (cont.)

- Model formula:
  \[ g(x) = 1 - \exp\left[-\left(\frac{x}{\sigma}\right)^\beta\right], \quad x \leq w \]
- Parameters = 2
- Shape criterion?
  Yes
- Model robust?
  Yes
Key functions (cont.)

- Model formula:
  \[ g(x) = \exp\left( -\frac{x}{\sigma}\right), \; x \leq w \]

- Parameters = 1

- Shape criterion?
  No

- Model robust?
  No
Key functions in Distance
Adjustment terms

Models can be made more robust by adding a series of adjustment terms (also called series expansion or series adjustment) to the key function.

Key function $\times (1 + \text{Series})$

Series $= \alpha_1 \times \text{term}_1 + \alpha_2 \times \text{term}_2 + \ldots \ldots$ etc.

The $\alpha_i$ parameters must be estimated.

Resulting curve model is scaled so that $g(0)=1$

The number of adjustment terms needs to be chosen.
Adjustment terms

Distance allows the selection of three types of series (one type per model)

<table>
<thead>
<tr>
<th>Key function</th>
<th>Series adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform*</td>
<td>Cosine*</td>
</tr>
<tr>
<td>Half normal†</td>
<td>Hermite polynomial†</td>
</tr>
<tr>
<td>Hazard rate</td>
<td>Simple polynomial</td>
</tr>
<tr>
<td>Negative exponential</td>
<td></td>
</tr>
</tbody>
</table>
How adjustment terms work

E.g. Cosine series (for different values of $\alpha$)

(1st order only used for uniform)
How adjustment terms work

E.g. Uniform + 1 Cosine adjustment term:

The effect of the adjustment terms depends on the value of their parameters.
How adjustment terms work

E.g. Half normal + 1 or 2 Cosine terms:
Adjustments in Distance
Adjustments in Distance
Adjustment terms – how many?

<table>
<thead>
<tr>
<th>Half normal</th>
<th>Half normal</th>
<th>Half normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 adjustment terms</td>
<td>1 adjustment term</td>
<td>5 adjustment terms</td>
</tr>
<tr>
<td>1 parameter</td>
<td>2 parameters</td>
<td>6 parameters</td>
</tr>
<tr>
<td>( \hat{P}_a = 0.65 )</td>
<td>( \hat{P}_a = 0.72 )</td>
<td>( \hat{P}_a = 0.63 )</td>
</tr>
<tr>
<td>( CV(\hat{P}_a) = 5.8% )</td>
<td>( CV(\hat{P}_a) = 11.6% )</td>
<td>( CV(\hat{P}_a) = 19.9% )</td>
</tr>
</tbody>
</table>

Note: There is a monotonicity constraint in Distance that is switched on by default to prevent detection functions from increasing. The constraint had to be turned off to produce the third plot. The third plot is for demonstration only – it would not be a good detection function to choose (unless there was a biological reason why detection probability would increase at those distances).
How many parameters?

Models with too few parameters will not be flexible enough to describe the underlying relationship.

Adding parameters will improve the fit.

But models with too many parameters will be too flexible and will also describe the random noise in the data.

We generally require models with an intermediate number of parameters.
How many parameters?

This problem can also be expressed as a trade-off between bias and variance.

Models with too few parameters tend to produce estimates with low variance and high bias.

Models with too many parameters tend to produce estimates with low bias and high variance (note the increasing CV for the estimate of $P_a$ on the previous slide).
How many parameters?

Need an objective way of choosing the ‘best’ model...
Truncation

$$\hat{N} = \frac{nA}{2wLP_a}$$

Need to choose the value of $w$ (right truncation)

Large distances contribute little to estimating the shape of $g(x)$ at small distances (i.e., the shoulder) and may lead to poor fit and high variance.

Typically we might truncate around 5% of observation for line transects (perhaps nearer 10% for point transects).

Can truncate in the field or at the analysis stage.