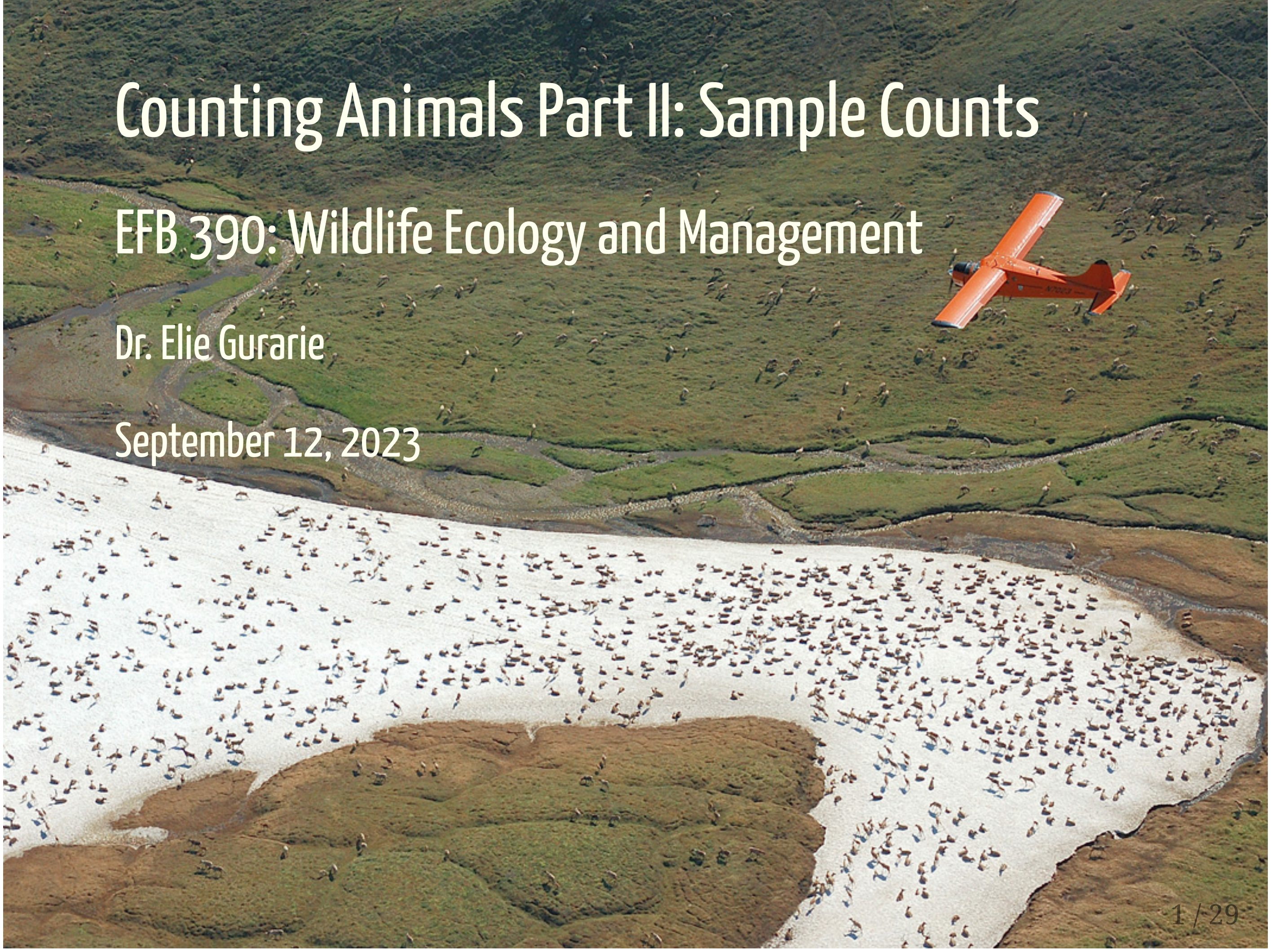
Counting Animals Part II: Sample Counts

EFB 390: Wildlife Ecology and Management

Dr. Elie Gurarie

September 12, 2023



Drawbacks of total counts / censusing

Expensive & labor-time intensive

Impractical for MOST species / systems

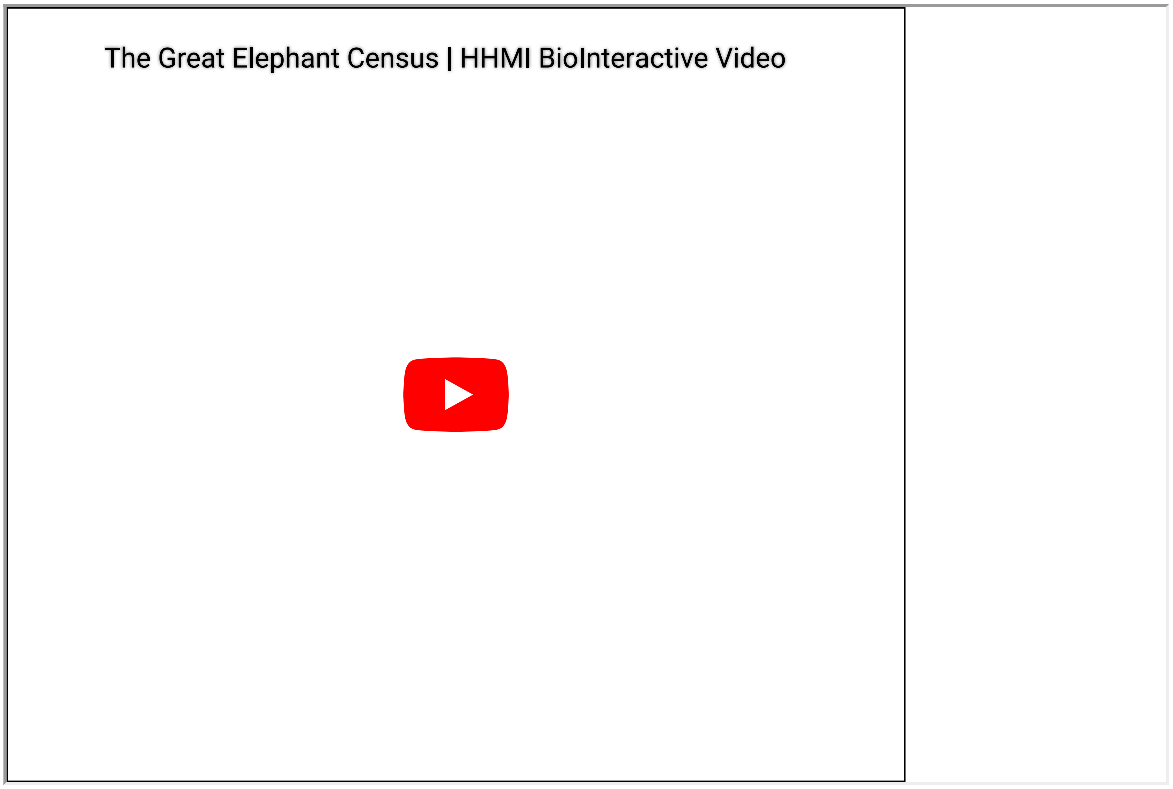
- need to ALL be **visible**
- the **ENTIRE** study area needs to be survey-able

Hard to assess precision

Hippos

(Marc Mol/Mercury Press/Caters)

Is the great Elephant Census a Census?



Sample counts

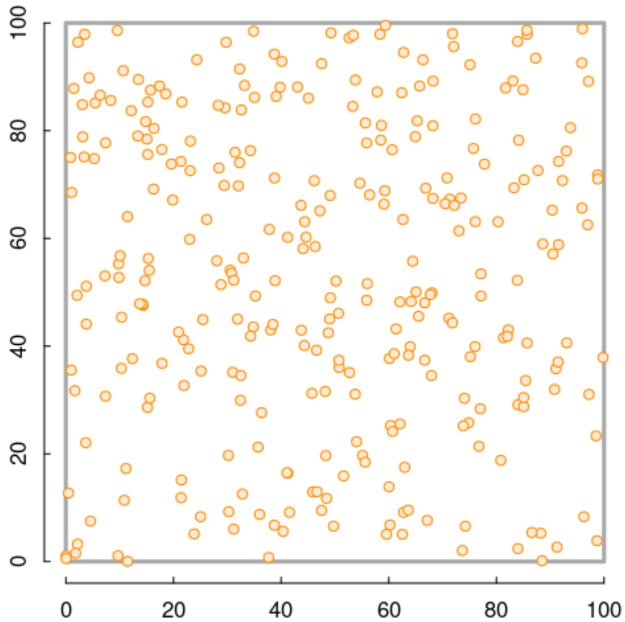
Simple idea:

- count *some* of the individuals
- extrapolate!

In practice:

- Involves some tricky statistics and modeling!
- Necessarily - less *precise* due to *sampling error*.
- BUT … if properly done … more *accurate* and **much less effort**.

A random population

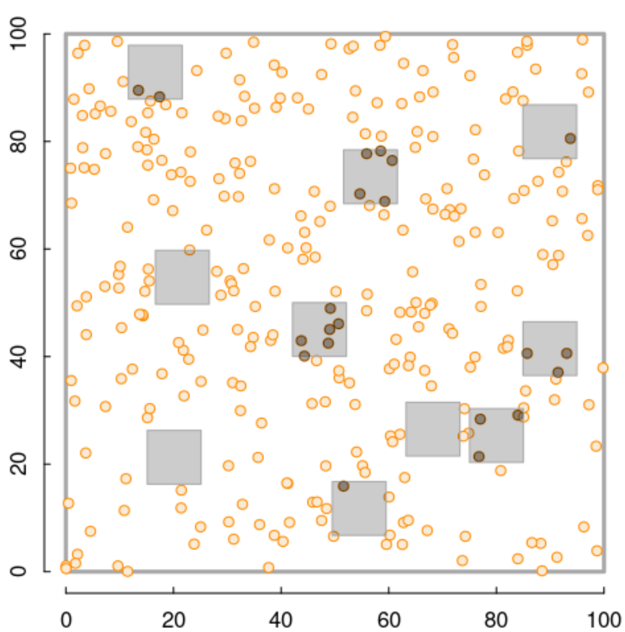


Population density

$$N = A \times D$$

- N - total count
- A - total area
- D - overall density

Sampling from the population



Sample density:

$$n_{sample} = \sum_{i=1}^{k} n_i$$

$$a_{sample} = \sum_{i=1} a_i$$

$$d_{sample} = \frac{n_{sample}}{a_{sample}}$$

Squares, aka, **quadrats**

Sample vs. Population

	Population	Sample
size	N	n_s
area	A	a_s
density	D	d_s

Note: sample density is an *estimate* of total density. So $\widehat{D} = d_s$.

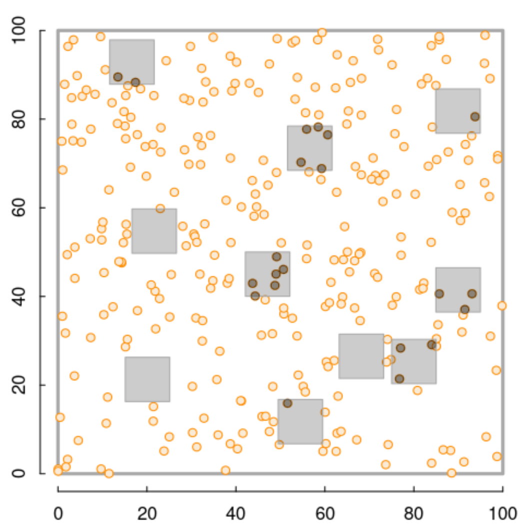
True population:

$$N = A \times D$$

Population **estimate** (best guess for N): just replace true (unknown) density D with *sampling estimate* of density d_s:

$$\widehat{N} = A \times \widehat{D} = A \times \frac{n_s}{a_s}$$

Example



Data

10 quadrats; 10x10 km each

`n = {0,0,5,0,3,1,2,3,6,1}`

note: *variability / randomness!*

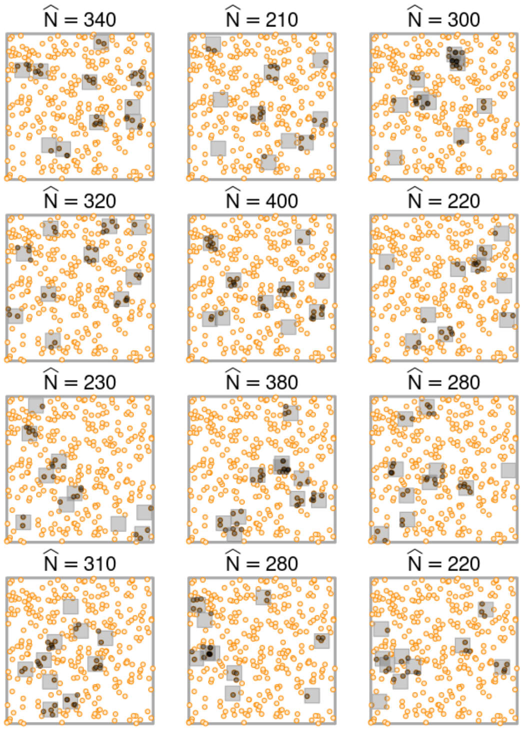
Analysis

$$n_s = \sum n_i = 21$$
$$d_s = \widehat{D} = \frac{35}{10 \times 10 \times 10} = 0.021$$
$$A = 100 \times 100$$

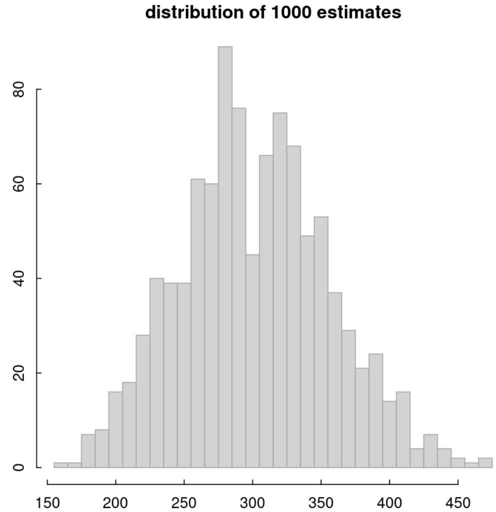
final estimate:

$$\widehat{N} = \widehat{D} \times A = 100 \times 100 \times 0.021 = 210$$

What happens when we do this many times?



Every time you do this, you get a different value for $\widehat{N}$.



Statistics

Mean of estimates:

$$\widehat{N} = 301.5$$

S.D. of estimate:

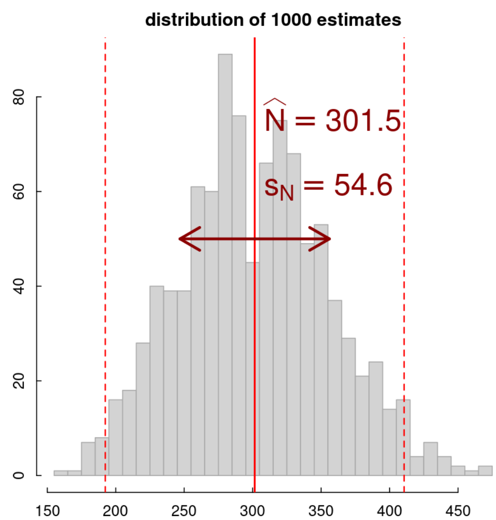
$$s_{\widehat{N}} = 54.6$$

important: the *standard deviation* of an *estimate* = **standard error**, SE

95% Confidence Interval:

$$\widehat{N} \pm 1.96 \times SE = \{195 - 408\}$$

note: the 1.96 is the number of standard deviatinos that captures 95% of a Normal distribution.



General principle: The bigger the sample, the smaller the error.

1. If $a_s \ll A$ (i.e. low sampling intensity)

$$SE(\widehat{N}) = \frac{A}{a}\sqrt{\sum n_i}$$

remember: $n_s = \sum n_i$ is the total sample count

in our example: $SE = 100^2/(10 \times 10^2)\sqrt{30} = 54.8$

2. If you are NOT resampling previously sampled locations:

$$SE(\widehat{N}) = \frac{A}{a}\sqrt{\sum n_i(1 - a_s/A)}$$

This is the Finite Area Correction. If $a = A$ - you sampled everything - SE goes to 0 as expected.

in our example: $SE = 54.5$ ... Almost no difference (because $a \ll A$).

Some more complex formulae

from Fryxell book Chapter 12:

Table 13.3 Estimates and their standard errors for animals counted on transects, quadrats, or sections. The models are described in the text.

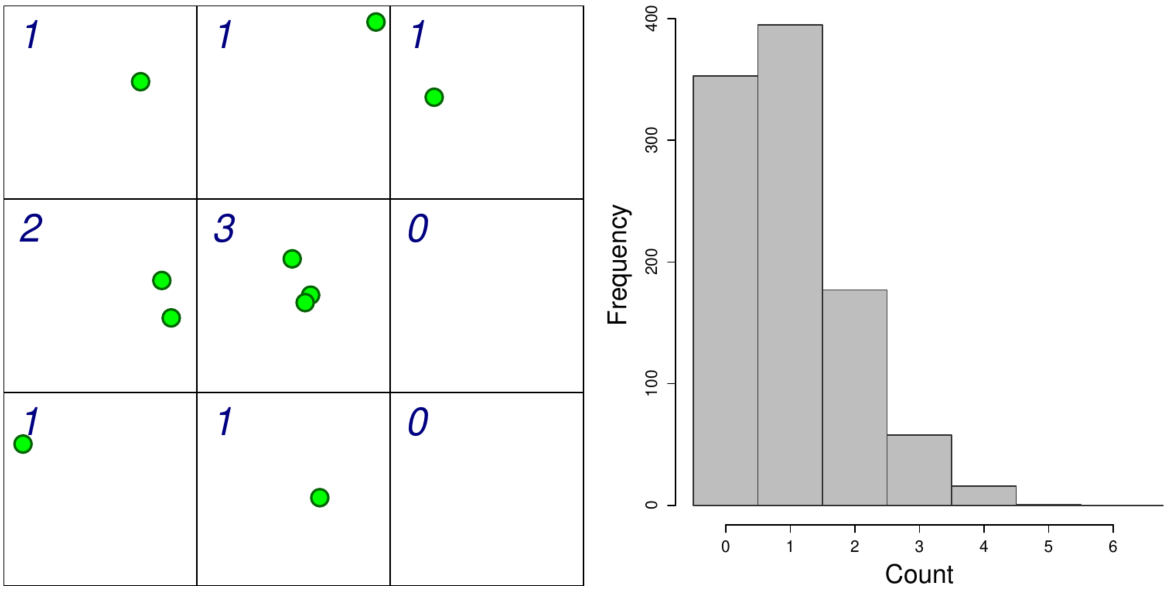
Model	Density	Numbers
Simple		
Estimate	$D = \Sigma y/\Sigma a$	$Y = A \times D$
Standard error of estimate (SWR)	$SE(D)_1 = 1/a \times \sqrt{[(\Sigma y^2 - (\Sigma y)^2/n)/(n(n-1))]}$	$SE(Y) = A \times SE(D)_1$
Standard error of estimate (SWOR)	$SE(D)_2 = SE(D)_1 \times \sqrt{[1 - (\Sigma a)/A]}$	$SE(Y) = A \times SE(D)_2$
Ratio		
Estimate	$D = \Sigma y/\Sigma a$	$Y = A \times D$
Standard error of estimate (SWR)	$SE(D)_3 = n/\Sigma a \times \sqrt{[(1/n(n-1))(\Sigma y^2 + D^2\Sigma a^2 - 2D\Sigma ay)]}$	$SE(Y) = A \times SE(D)_3$
Standard error of estimate (SWOR)	$SE(D)_4 = SE(D)_3 \times \sqrt{[1 - (\Sigma a)/A]}$	$SE(Y) = A \times SE(D)_4$
PPS		
Estimate	$d = 1/n \times \Sigma(y/a)$	$Y = A \times d$
Standard error of estimate (SWR)	$SE(D) = \sqrt{[(\Sigma(y/a)^2 - (\Sigma(y/a))^2/n)/(n(n-1))]}$	$SE(Y) = A \times SE(d)$

SWR, sampling with replacement; SWOR, sampling without replacement. Notation is given in Section 13.5.1.

These are used when **sampling areas** are unequal, and account for differences when sampling **with replacement** or **without replacement**.

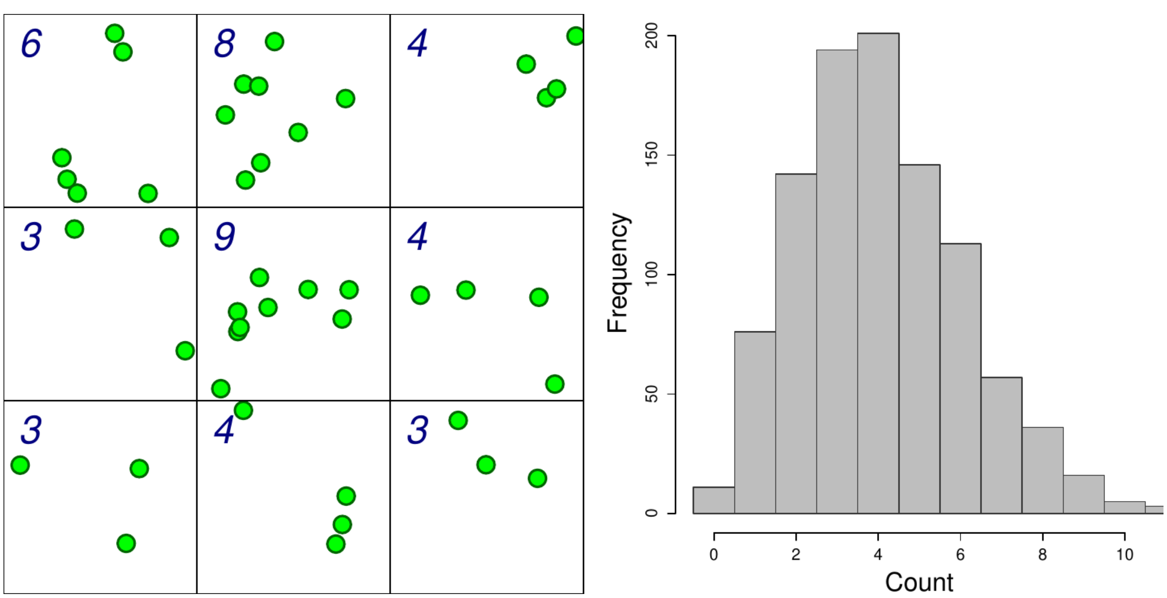
Poisson process

Models *counts*. If you have a perfectly random process with mean *density* (aka *intensity*) 1, you might have some 0 counts, you might have some higher counts. The *average* will be 1:



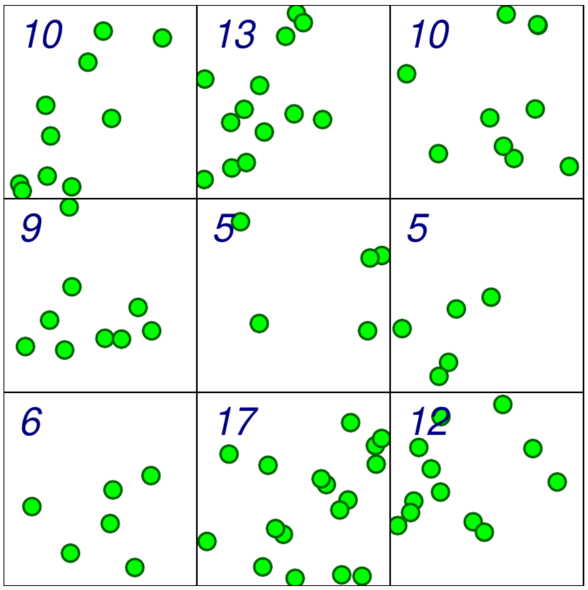
Poisson process

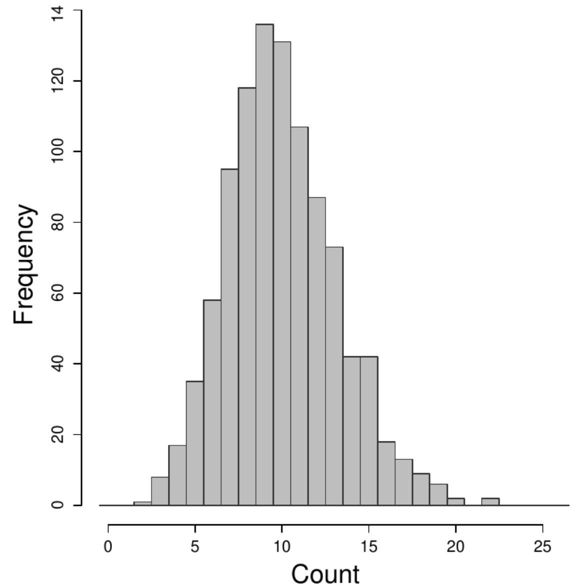
Here, the intensity is 4 ...



Poisson process

... and 10. Note, the bigger the intensity, the more "bell-shaped" the curve.

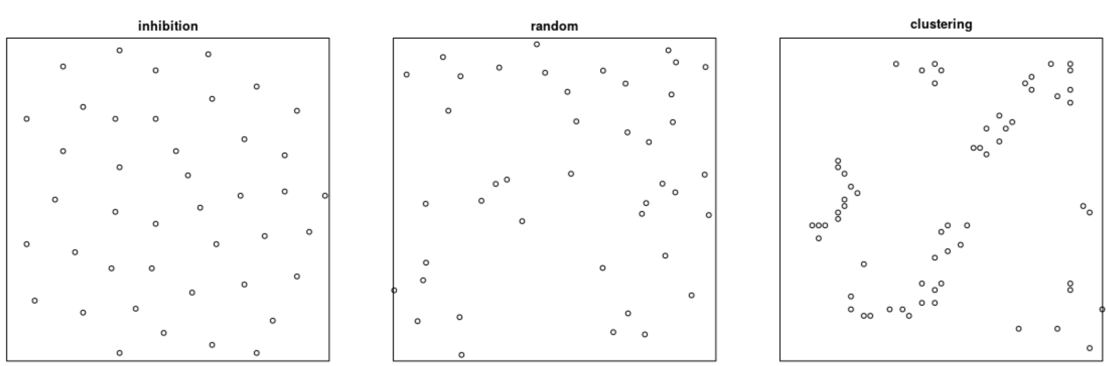




Here's the formula of the Poisson Distribution: $f(k; \lambda) = \Pr(X{=}k) = \frac{\lambda^k e^{-\lambda}}{k!}$

Poisson distribution holds if process is truly random

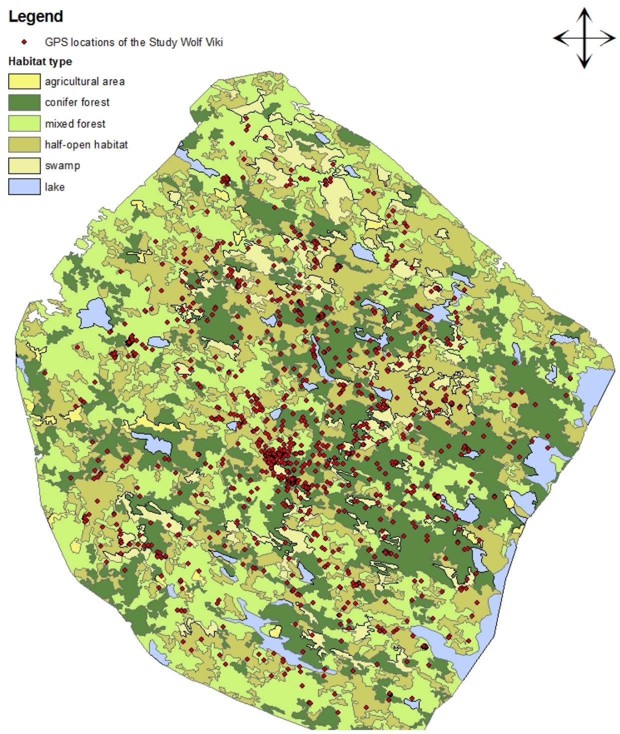
... not **clustered** or **inhibited**



If you **sample** from these kinds of spatial distributions, your standard error might be smaller (*inhibited*) or larger (*clustering*). This is called *dispersion*.

Also … densities of animals can depend on habitat

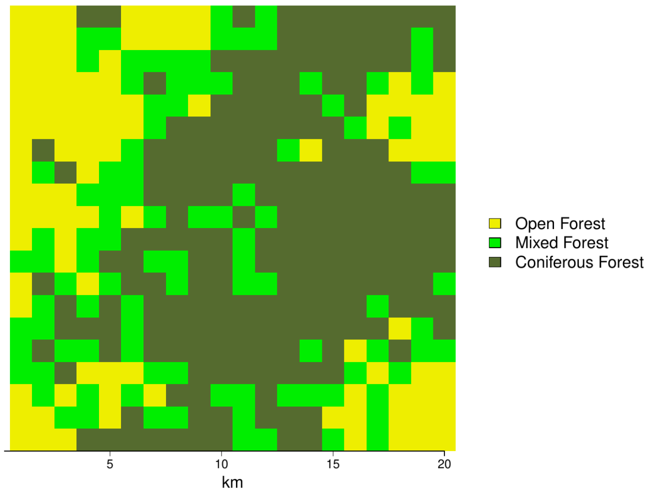
Wolf habitat use



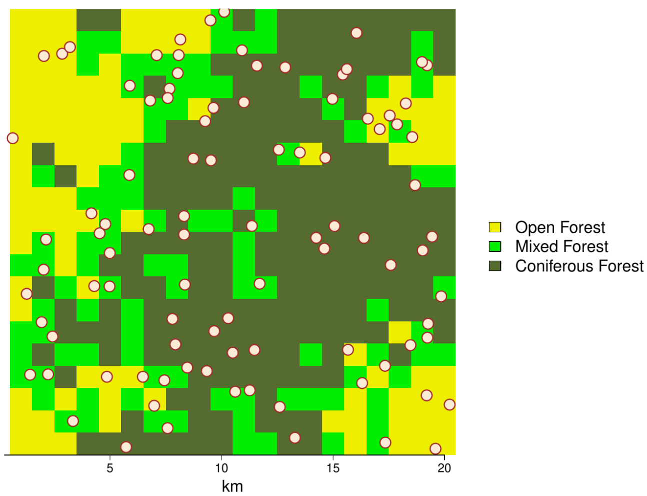
If you look closely:

- No locations in lakes
- Relatively few in bogs / cultivated areas.
- Quite a few in mixed and coniferous forest

Imagine a section of forest …



... with observations of moose

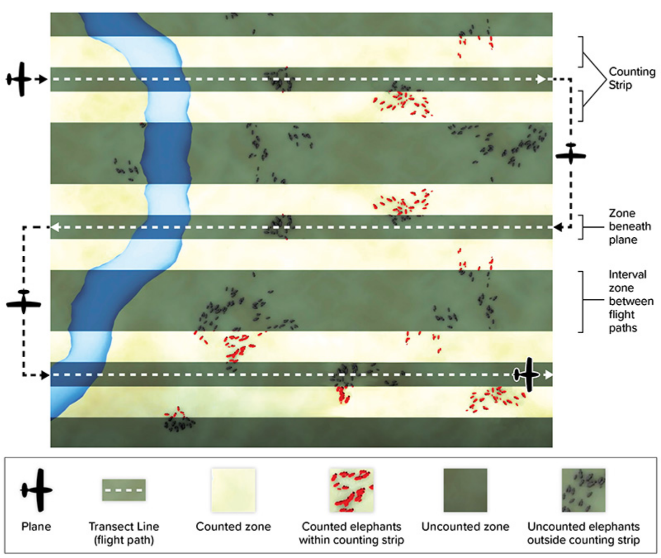


How can we tell what the moose prefers?

Habitat	Area	n	Density
open	100	21	0.21
mixed	100	43	0.43
dense	200	31	0.17
total	400	95	0.24

Knowing how densities differ as a function of **covariates** can be very important for generating estimates of abundances, increasing both **accuracy** and **precision**, and informing **survey design**.

Sample frames need not be **squares**



Transects

Linear strip, usually from an aerial survey.

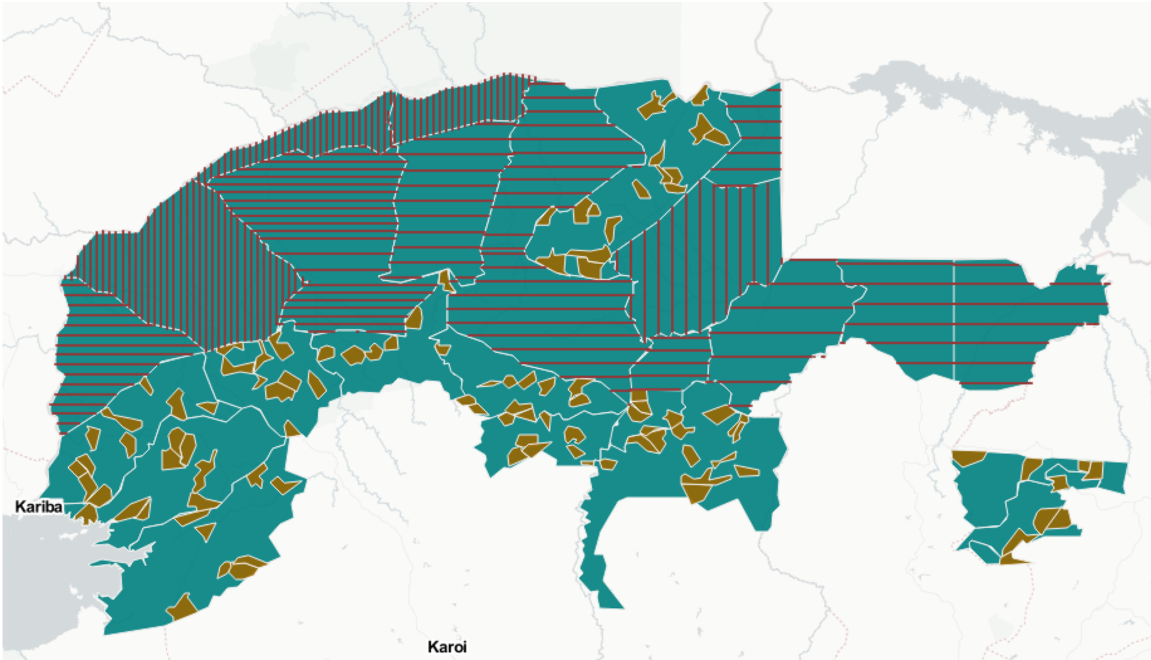
Efficient way to sample a lot of territory.

If "perfect detection", referred to as a **strip transect**.

Statistics - essentially - identical to quadrat sampling.

https://media.hhmi.org/biointeractive/click/elephants/survey/survey-aerial-surveys-methods.html

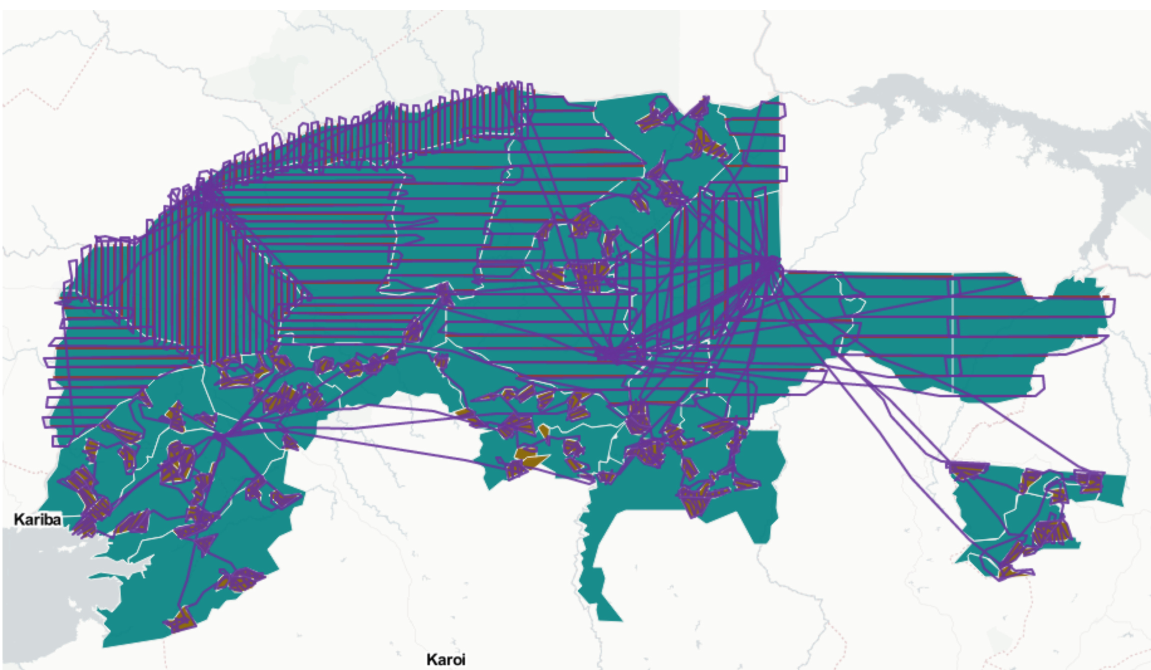
Stratified sampling for more efficient estimation



Sample more intensely in those habitats where animals are more likely to be found. Intensely survey **blocks** where detection is more difficult.

https://media.hhmi.org/biointeractive/click/elephants/survey/survey-aerial-surveys-methods.html

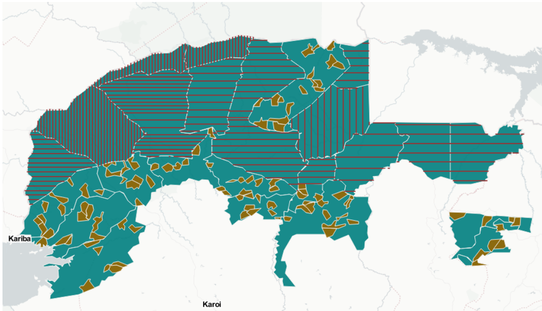
Stratified sampling for more efficient estimation

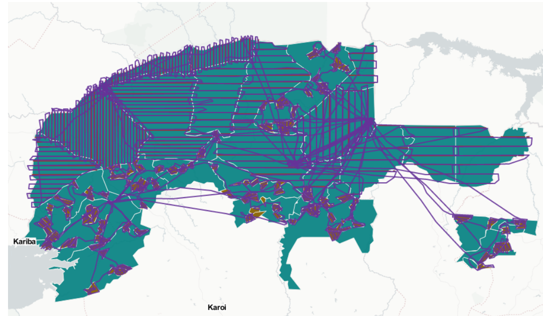


Actual elephant flight paths,

https://media.hhmi.org/biointeractive/click/elephants/survey/survey-aerial-surveys-methods.html

Stratified sampling

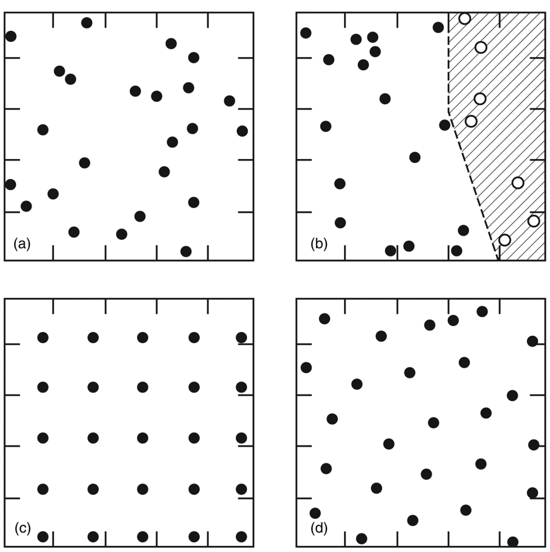




Stratification is used to optimize **effort** and **precision**. Aircraft cost thousands of dollars per hour!

(In all of these comprehensize surveys - *design* takes care of **accuracy**).

Sampling strategies



(a) simple random,

(b) stratified random,

(c) systematic,

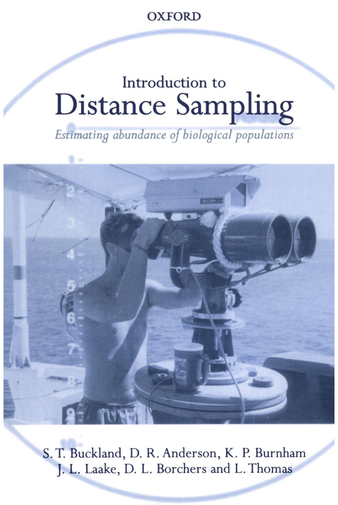
(d) pseudo-random (systematic unaligned).

Each has advantages and disadvantages.

See also: *Adaptive Sampling*

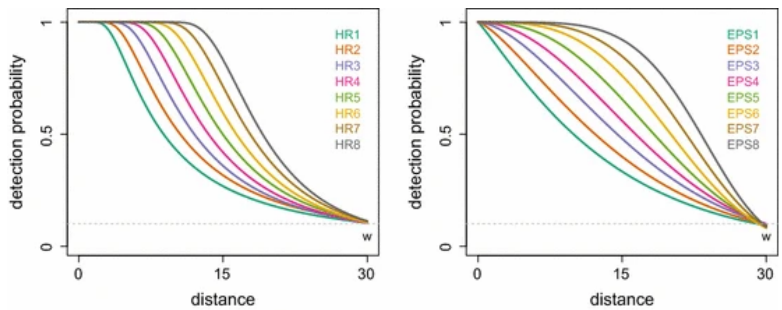
Detections usually get *worse* with distance!

Detection function	Form
Uniform	$1/w$
Half-normal	$\exp(-y^2/2\sigma^2)$
Hazard-rate	$1-\exp(-(y/\sigma)^{-b})$
Negative exponential	$\exp(-ay)$

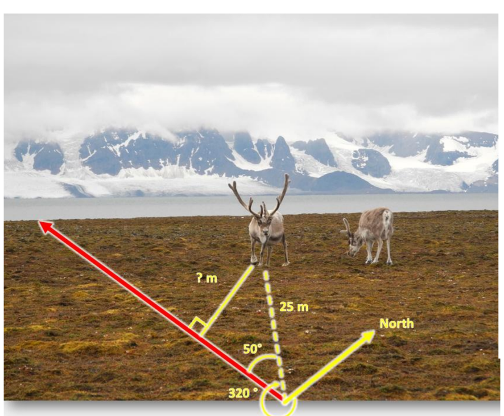


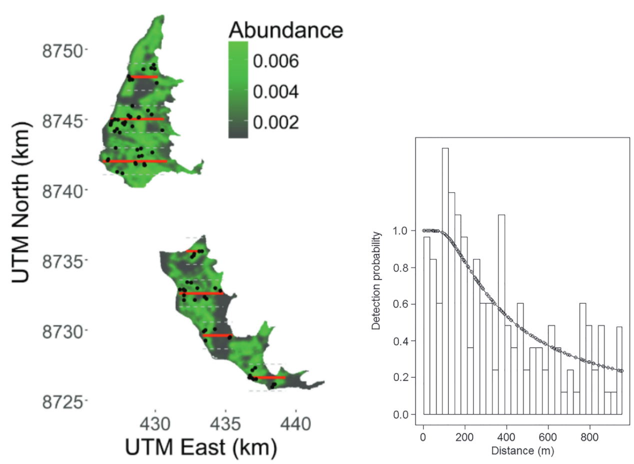
Distance Sampling

The statistics of accounting for visibility decreasing with distance



Example reindeer in Svalbard





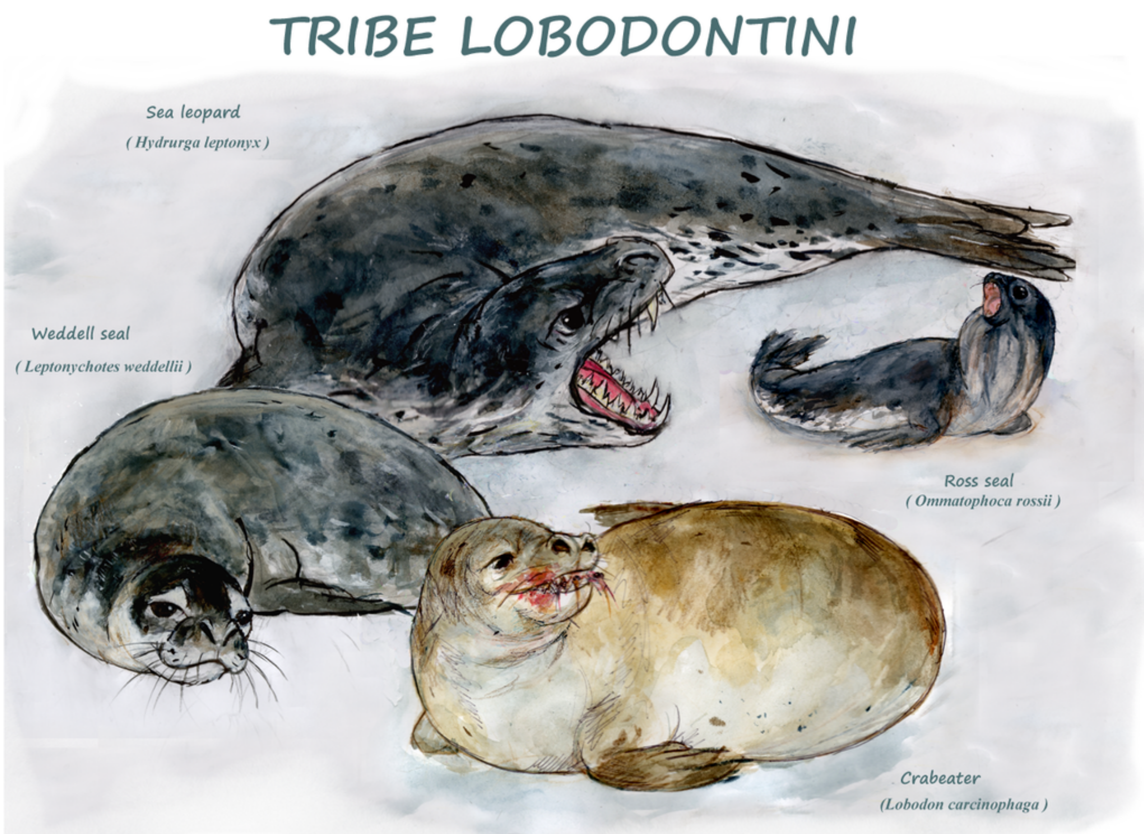
Ungulate population monitoring in an open tundra landscape: distance sampling versus total counts

Wildlife Biology 2017: wlb.00299
doi: 10.2981/wlb.00299

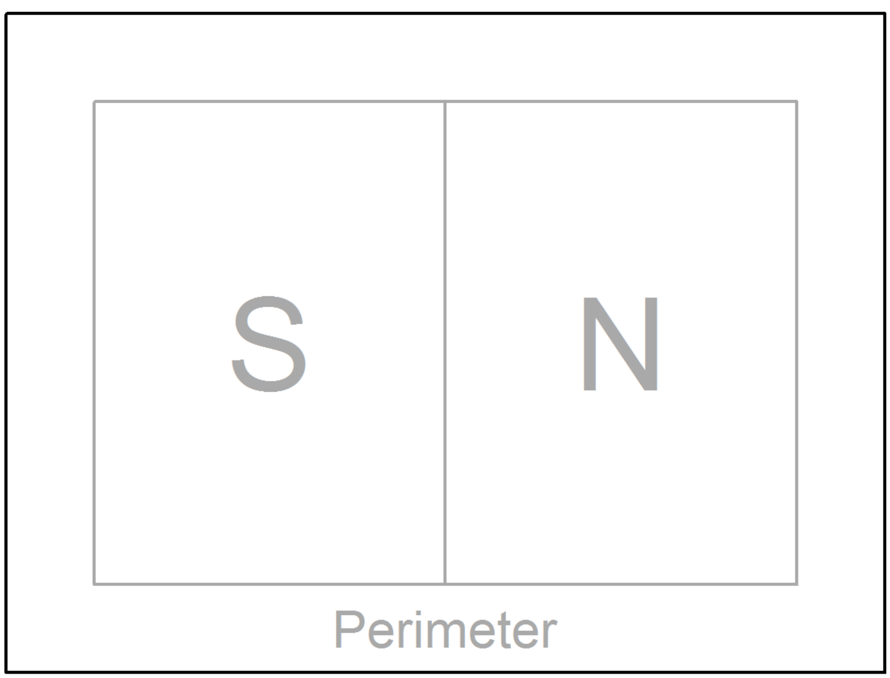
Mathilde Le Moullec, Åshild Ønvik Pedersen, Nigel Gilles Yoccoz, Ronny Aanes, Jarle Tufto and Brage Bremset Hansen

Estimated detection distance, compared to **total count**, incorporated **vegetation modeling**, computed **standard errors**, concluded that you can get a 15% C.V. for 1/2 the cost.

Example Ice-Seals



Example: Flag Counting at Baker



Nice video on counting caribou

https://vimeo.com/471257951