ESTIMATION OF ANIMAL ABUNDANCE

ST 506 2006 Lecture 2

Kenneth H. Pollock

Professor of Zoology, Biomathematics, and Statistics

231 David Clark Labs

919-515-3514

pollock@unity.ncsu.edu

Tentative Office Hours

MW 1.30-2.30

Tu 2.00-3.00

231 DCL

Alternately by Appointment

pollock@unity.ncsu.edu

515-3514

An Cute Example of Biased Sampling (☺)

The story goes that in the Second World War a group was studying planes returning from bombing Germany. They drew a rough diagram showing where the bullets holes were and recommending those areas be reinforced.

However a statistician pointed out that there was "missingness" in the sample they were studying. That is, what about the planes that didn't return from Germany because they were shot down?The two areas that had almost no bullet holes were where the wings and where the tail joined the fuselage. These turn out to be crucial and planes with bullet holes there likely don’t make it back.

Moral of the Story: The sample you have may not tell you much about the remainder of the population unless you use probability sampling (simple random sampling, stratified random sampling). In ecology we are in a very tricky situation like this because we cannot always get a good sample that we know is free of bias.

Mangel &Samaniego (1984). Abraham Wald's work

on aircraft survivability. J A Statistical Assoc, 79, 259-267.

OVERVIEW OF METHODS OF ASSESSING POPULATIONS

Direct methods of monitoring

Census Method

Count all animals in the population
(in usually unrealistic practice)

Sampling Methods

Count animals in sampling units (or areas)

Absolute Abundance

Estimate popn size by adjusting for ‘unseen’ animals and only part of area being sampled

Relative Abundance

Use incomplete count as an ‘index’ of popn size

ABSOLUTE ABUNDANCE ESTIMATION

Not all area sampled and not all animals ‘seen’

$$\hat{N} = C/\alpha\hat{\beta}$$

C = count of animals seen

α = fraction of area sampled

$\hat{\beta}$ = estimate of the fraction of animals seen or caught

Note: there are many ways to estimate β . For example, capture-recapture, line transects etc.

REMOVAL MODELS (TWO SAMPLE)

Fish in a stream

n_1 animals captured and removed in sample 1
n_2 animals captured and removed in sample 2

$$\hat{N} = n_1^2 / (n_1 - n_2)$$

$$\hat{p} = \hat{\beta} = (n_1 - n_2) / n_1$$ estimate of detectability

$$\hat{N} = n_1 / \hat{p}$$

Note: Generalizes to other more general removal and catch effort models.

REMOVAL MODELS (TWO SAMPLE)

Example of Model Based Sampling

Model Assumptions (Validity of Estimator, Bias)

1. Closed Population
2. Equal capture probs for all animals
3. Equal capture probs at time 1 and time 2.

Sample Sizes (Precision of Estimator, Standard error)

Larger sample sizes will give better precision
(i.e. a smaller Standard Error)

Note: Bias and Precision discussed next

Simulation Removal Estimator

Computer Simulation can be used to study properties of estimators like bias and precision.

First I used N=100, p=0.3 and 2 removals

Did a very small simulation and generated 10 values.

56, 73, 82, 89, 96, 146, 169, 392, Fails, Fails

One cant formally calculate the bias and variance of this estimator because it sometimes fails

Repeated with N=100 but now p=0.5.

Simulation Removal Estimator

Removal sample of 25 values with N=100, p=0.5

104	
91	
96	
106	
92	
126	
97	
205	
95	
84	
121	
85	
113	
129	
81	
101	
101	
142	
161	
96	
82	
89	
97	
93	
78	
106.6	**Mean**
+6.6	**Bias**
28.61	**St Error**

SAMPLING ANIMAL POPULATIONS

Statistical Concepts (Chapter 4)

Methods of Estimation

Method of Moments

Method of Least Squares

Method of Maximum Likelihood

SAMPLING ANIMAL POPULATIONS

Statistical Concepts (Chapter 4)

Methods of Estimation

Method of Moments

- Simplest method and quite intuitive often.
- It involves equating sample moments to population moments.
- For biology students this will become clearer when I illustrate for the removal model.

SAMPLING ANIMAL POPULATIONS

Statistical Concepts (Chapter 4)

Methods of Estimation

Method of Least Squares

Used in regression problems in a course like 511,512.Will not discuss further here.

SAMPLING ANIMAL POPULATIONS

Statistical Concepts (Chapter 4)

Methods of Estimation

Method of Maximum Likelihood (See Also Text)

- Derive the probability distribution of the observed data as a function of the parameters $p(x;\theta)$

- View this as a function of the parameters, this is the likelihood function

$$L(\theta)$$

- Find the values of the parameters which maximise this function.These are the maximum likelihood estimators (MLEs) $\hat{\theta}$

Method of Maximum Likelihood

Tag Retention Example

R Retained tag L Lost tag

p - prob tag retained, (1 - p) - prob tag lost

$$P(RRLRLRRRRL) = p^7 (1 - p)^3$$

$$L(p) = p^7 (1 - p)^3$$

The likelihood is maximised at

$$\hat{p} = x/n = 7/10 = 0.7$$

See plot on whiteboard.

Method of Maximum Likelihood

Methods of Maximising a function

•Set partial derivatives equal to 0. Solve the resulting equations. This gets very complicated as we often have a lot of parameters to estimate.

•Use a computer package which may use a variety of algorithms to find the values which maximise the function. These are the MLES. The programs also give the variances and standard errors of the estimators.

Method of Maximum Likelihood

Removal Model

Can find MLEs mathematically and I will talk about this in class. The Likelihood is

$$L(N, p; n_1, n_2)$$

$$= C_{n_1}^{N} p^{n_1} (1-p)^{(N-n_1)}.$$

$$C_{n_2}^{(N-n_1)} p^{n_2} (1-p)^{(N-n_1-n_2)}$$

$$Then\,Use$$

$$L(N, p) / L(N-1, p) = 1$$

$$and$$

$$dLogL(N, p) / dp = 0$$

SAMPLING ANIMAL POPULATIONS

Method of Maximum Likelihood

Removal Model

Note that the method of moment estimators are also MLEs although I don't show this.

$$\hat{N} = n_1^2/(n_1 - n_2)$$

$$\hat{p} = (n_1 - n_2)/n_1$$

SAMPLING ANIMAL POPULATIONS

Maximum Likelihood Estimators Properties

For “large samples” (asymptotics) approximately:

• Unbiased

•Normally distributed

•No other estimator has smaller variance. So “best” in

terms of precision (efficiency).

• Possible to derive expressions for the large sample variances and SEs of the MLEs

SAMPLING ANIMAL POPULATIONS

Maximum Likelihood Estimators Properties

Possible to derive expressions for the large sample variances and SEs of the MLEs

Removal Model

$$\hat{S}E(\hat{N}) = n_1 n_2 (n_1 + n_2)^{1/2} / (n_1 - n_2)^2$$

SAMPLING ANIMAL POPULATIONS

Confidence Intervals(4.2.3)

Using Asymptotic Normality-for example 95% CI is

$$\hat{N} \pm 1.96\, SE(\hat{N})$$

Using Profile Likelihood Approach

Will not discuss but will be computed in some packages

REMOVAL MODELS (TWO SAMPLE)

Example-Seber (1982) P324.

n_1=49
n_2=26

$\hat{N}$ **= n_1^2/ (n_1 - n_2) = 49^2 / 23 = 104 (SE=20)**

$\hat{p} = \hat{\beta}$ **= (n_1 - n_2) / n_1= 23/49 =0.47**

Approx 95%CI is (104-1.96x20, 104+1.96x20)
or (popn size between 65 and 143 animals)

REMOVAL MODELS
(TWO SAMPLE)

Example-Seber (1982) P324.

Wide Confidence interval even though p large here.With small populations you need a very large removal probability for this method to work well.

Note if p is small the estimate may "fail"! Why? Discuss by referring back to the structure of the equation.

SAMPLING ANIMAL POPULATIONS

Goodness of Fit Tests (4.3.3)

Chi-Square Goodness of Fit Tests are the simplest approach is to use

$$\chi^2 = \sum (O_i - E_i)^2 / E_i$$

O_i – observed value for cell i

E_i – expected value for cell i

We will illustrate when we fit more complex models later on.

SAMPLING ANIMAL POPULATIONS

WE WILL POSTPONE THESE SECTIONS (4.3.4 and 4.4) UNTIL WE HAVE FITTED SOME MORE COMPLEX MODELS WITH MORE PARAMETERS

Choosing Between Models

Likelihood Ratio Tests

Formal Model Selection Methods

Akaike Information Criteria

Text Reading:

1. Chapter 4

   4.1-4.3.3 Important now and 4.3.4-4.4 Important later

   4.5 Skip

2. Chapter 5

   5.1-5.3, 5.5-5.6 Very important

   5.4.2 Very important

   5.4.1 & 5.4.3 Less important

   5.4.4 Not covered at all

3. Chapter 12

   12.1-12.4 Very important

   12.5- skip totally

   12.6- May be difficult to read- I shall give examples to make it simpler.

   12.7-12.8 Very important

SAMPLING ANIMAL POPULATIONS

Overview of Material

1. Counting Methods of Estimating Animal Abundance (Ch 12)

- not all the area can be covered (i.e., sample, not a census)
- usually not all animals will be seen, but the book starts with the assumption of all detected and then extend the results to when not all detected

2. Finite Population Sampling Overview (Ch 5 will not cover in class)

- Simple Random Sampling
- Stratified Random Sampling
- Systematic Random Sampling
- Other Sampling Designs

3. Uses of Finite Population Sampling in Ecology

- To estimate abundance (with extensions-Ch 12 and beyond)
- In Angler and Hunter Surveys (will not be covered in this class)

Some Other References:

1. “Sampling” Steven K. Thompson, Wiley 1992.
2. “Angler Survey Methods” Pollock et al. 1994. AFS Special Publication 25
3. Text book, Chapters 5 and 12

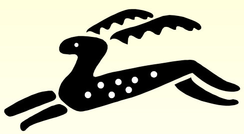
NOTE: I spend a lot of time on this material in my design of ecological studies class and have given you some of that lecture material on the website.

Formulation

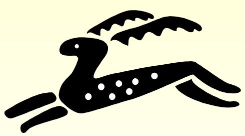
C – count of animal in a samples of area fraction α.

$$\hat{N} = \frac{C}{\alpha}$$

Example: Count 161 gazelles in an aerial survey in open habitat- Area is 15% of total area chosen at random.



$$\hat{N} = \frac{161}{0.15} = 1074$$ animals estimated



Suppose that there is a visibility bias so that only some ($\beta < 1$) detected.

$$\hat{N} = \frac{C}{\alpha\hat{\beta}}$$

<u>Example:</u> Count 161 gazelles in aerial survey of 15% of area and we estimate somehow that 0.81 is detection probability

$$\hat{N} = \frac{161}{0.15x0.81} = \frac{1074}{0.81} = 1326$$ animals estimated

Note: May convert to density

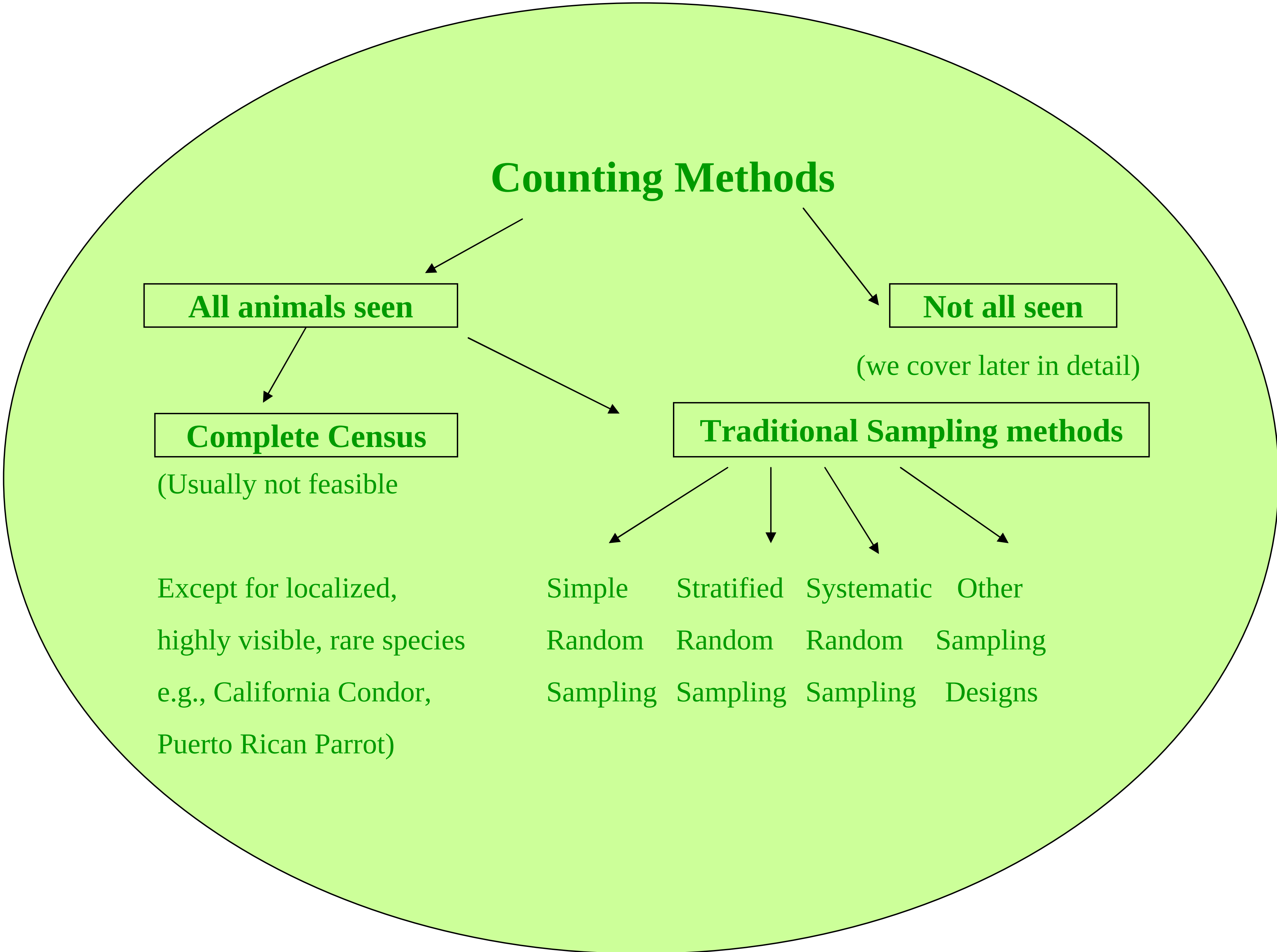
$$\hat{D} = \frac{\hat{N}}{A} = \frac{1326}{1218} = 10.9$$ gazelles/square mile

Questions

1. How to design the sampling study to pick the areas to be sampled.

   That is do we use simple random sampling, systematic random sampling or stratified random sampling.

2. How to estimate the fraction of animals detected ($\hat{\beta}$). What method is the best one for our situation. We shall discuss many!


Counting Methods
All animals seen
Not all seen
(we cover later in detail)
Complete Census
(Usually not feasible
Except for localized,
highly visible, rare species
e.g., California Condor,
Puerto Rican Parrot)
Traditional Sampling methods
Simple Random Sampling
Stratified Random Sampling
Systematic Random Sampling
Other Sampling Designs

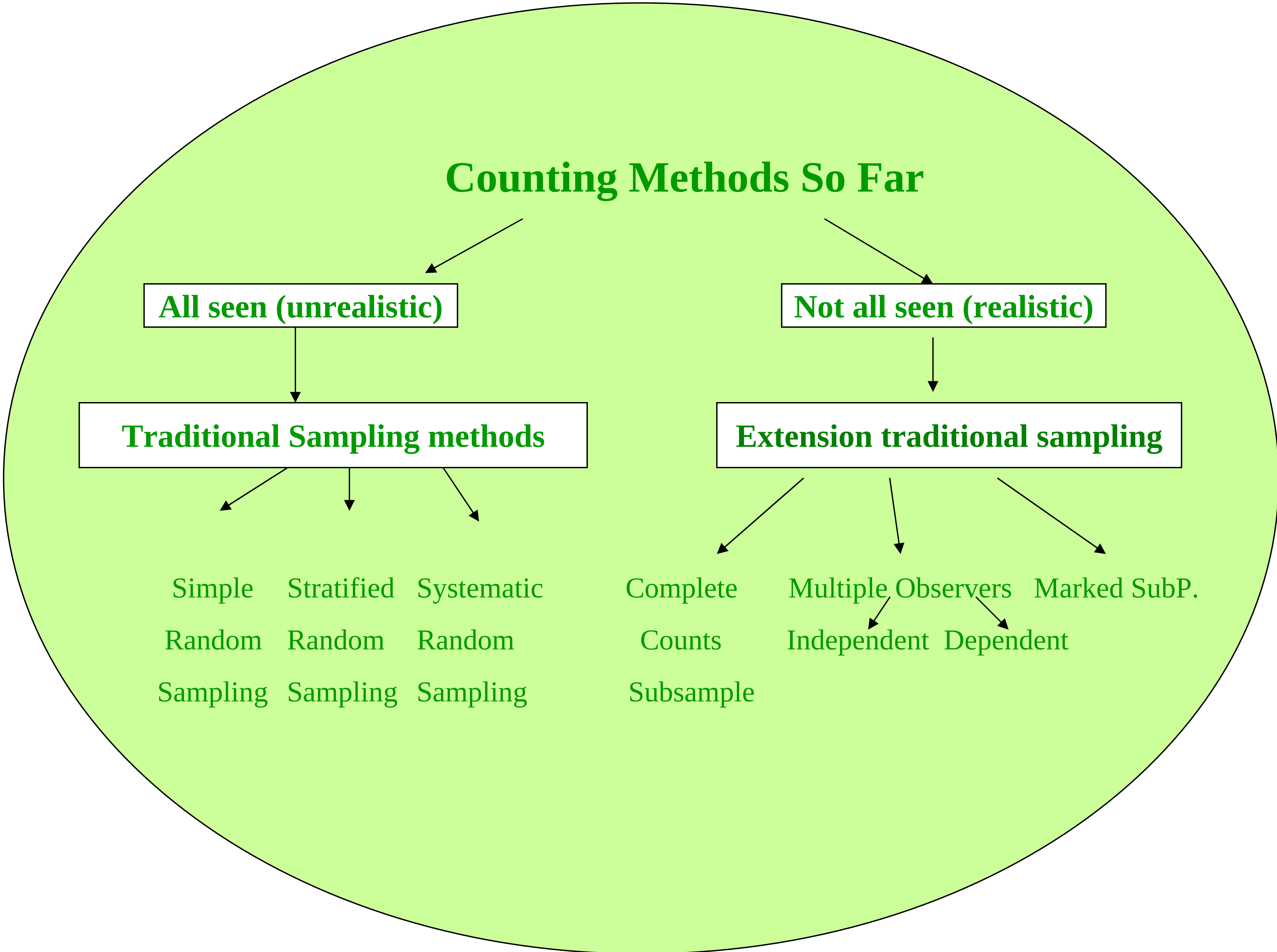
ABUNDANCE ESTIMATION (Ch 12)

• Complete Detectability of Individuals on Sample Plots of Equal Area (12.4)

(we will not cover this section but the supplemental material is relevant to it)

• Complete Detectabilty of Individuals on Sample Plots of Unequal Area (12.5)

(we shall not cover this, but results are generalizations of results in 12.4)


Counting Methods So Far
All seen (unrealistic)
Not all seen (realistic)
Traditional Sampling methods
Extension traditional sampling
Simple Random Sampling
Stratified Random Sampling
Systematic Random Sampling
Complete Counts Subsample
Multiple Observers
Independent
Dependent
Marked SubP.

ABSOLUTE ABUNDANCE ESTIMATION

Not all area sampled and not all animals 'seen'

$$\hat{N} = C/\alpha\hat{\beta}$$

C = count of animals seen
α = fraction of area sampled
$\hat{\beta}$ = estimate of the fraction of animals seen or caught

This would need extension to stratified sampling if that was used to pick the sample units

ABSOLUTE ABUNDANCE FORMULA EXTENSION TO STRATIFIED SAMPLING

$$\hat{N} = \sum_{1}^{L} C_k / \alpha_k \hat{\beta}_k$$

There are L strata and in each one there would be an area sampled and a detection probability estimator.

In the rest of the lecture, we do not consider this aspect of the problem although I emphasize that what we do could be extended to this more complex case.

Partial Detectability on Sample Units (12.6 and 13)

Methods for Count Data

- Complete Counts on a Subsample of Plots
- Multiple Observers
  - Independent
  - Dependent
- Marked Subpopulation of Animals
- Distance Sampling (Ch 13) Later Lectures
  - Line Transects
  - Point Counts

Note: In all methods the objective is to obtain an estimate of detectability. All methods have strong assumptions

$\hat{\beta}$

Complete Counts on a Subsample of Plots

Example 1: Aerial Survey with Complete Ground Counts on a Subsample of Plots

Description:

Key Assumption:

The "complete" count is a truly complete, accurate count on the same plot, at the same time as the aerial count.

Key Questions:

1. Is it possible to do an accurate, complete count? In aerial survey applications it is often impossible to do a complete ground count even in a few plots.

2. If it is possible, is it possible at reasonable expense for a reasonable size sample? Otherwise $\hat{\beta}$ based on a very small subsample, will be very imprecise.

Complete Counts on a Subsample of Plots

Example 1: Aerial Survey with Complete Ground Counts on a Subsample of Plots

Numerical Example : Thompson (1992) and p. 251-252 text

Aerial Survey alone

$M = 100$ plots

$m' = 20$ plots 240 moose detected

$$\hat{X} = \frac{M}{m'} \sum_{i=1}^{m'} x_i = \frac{100}{20} x240 = 1200 \text{ moose estimated}$$

$$\alpha = \frac{20}{100} = 0.2$$

Ground Survey on Subsample of 5 plots

$$\hat{\beta} = \frac{56}{70} = \frac{\text{air count}}{\text{ground count}} = 0.80$$

Estimation of Population Size

$$\hat{N} = \frac{1200}{\hat{\beta}} = \frac{1200}{0.80} = 1500$$

or

$$\hat{N} = \frac{C}{\alpha\hat{\beta}} = \frac{240}{0.2x0.8} = 1500$$

Complete Counts on a Subsample of Plots

Example 2: Arctic Breeding Bird Survey with some plots searched twice.(Bart Paper).

All plots – standard search

Some plots – more thorough search, problem is if they are really complete counts

DETOUR!!! NEEDED FOR MULTIPLE OBSERVERS WORK

CAPTURE-RECAPTURE MODELS

LINCOLN-PETERSEN MODEL

N - Population size

n_1 - No. of marked animals in the population

n_2 - Sample size

m_2 - No. of marked animals in the sample

Sample Population

$$(m_2/n_2) \approx (n_1/N)$$

$$\hat{N} = n_1 n_2 / m_2$$

CHAPMAN'S MODIFICATION

TO REDUCE BIAS

-

$$\frac{m_2 + 1}{n_2 + 1} = \frac{n_1 + 1}{N + 1}$$

$$\hat{N}_c = \frac{(n_1 + 1)(n_2 + 1)}{(m_2 + 1)} - 1$$

This estimator is approximately unbiased.

Unbiased: If you repeated the study many times and took the average N, it would be equal to N.

PETERSEN MODEL ASSUMPTIONS

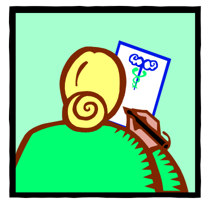
1. Closure

2. Equal Catchability

3. Zero Mark Loss (Marking is definitive)

Multiple Observers

Two independent observers



Apply the Lincoln-Petersen model!!

n_1 - seen by first observer

n_2 - seen by second observer

m_2 - seen by both observers

$$\hat{N} = \frac{n_1 n_2}{m_2}$$

$$\hat{N} = \frac{n_1}{\hat{p}_1}$$

$$\hat{p}_1 = \frac{m_2}{n_2} = \hat{\beta}$$

Multiple Observers

Two independent observers

Key Assumptions

1. No matching errors that is n_1,n_2 and m_2 are accurate.
2. Closed population
3. The observers really are independent

Multiple Observers

Two independent observers

Examples: aerial surveys

-Two observers in same plane

-One observer in plane, one on ground

Examples: Bird point counts

-Two observers at the same points

Multiple Observers

Two independent observers

Example: Henny and Burnham (1971) JWM.

Osprey nests- $\alpha = 1$ i.e., total area is sampled , here there is *one air observer* and one *ground observer*.

Data

$n_1 = 51$ seen by air observer

$n_2 = 63$ seen by ground observer

$m_2 = 41$ seen by both observers

Multiple Observers

Two independent observers

Estimation

$$\hat{N} = \frac{(n_1 + 1)(n_2 + 1)}{(m_2 + 1)} - 1$$

$$= \frac{52x64}{42} - 1 = 78.24$$

$$\text{SE}(\hat{N}) = 3.11$$ therefore a precise estimate

$$\hat{\beta}_{\text{air}} = \frac{n_1}{\hat{N}} = \frac{51}{78.24} = 0.65$$

Note: this is the estimated probability of detection of an osprey nest from the air

(35% missed)

Note: Ground observer here also misses nests

$$\hat{\beta}_{ground} = \frac{n_2}{\hat{N}} = \frac{63}{78.24} = 0.81$$

Multiple Observers

Two independent observers

Examples: aerial surveys

-Two observers in same plane unlike Osprey example where one on the ground.

Caughley, G. and Grice, D. (1982). A correction factor for counting emus from the air, and its application to counts in Western Australia. Australian Wildlife Research 9, 253-259.

Marsh, H. and Sinclair, D. F. (1989a). Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. J. Wildlife Manage. 53, 1017-24.

Pollock,K. H., Marsh, H. D., Lawler, I. R., and Alldredge, M. W. (2006). Estimating animal abundance in heterogeneous environments: an application to aerial surveys for dugongs. Journal of Wildlife Management 70, 255-262.

Multiple Observers

- Primary and Secondary Observers - then part way through the survey they switch roles.

- Cook and Jacobson (1979) applied to aerial surveys

- Nichols et al. (2000) Auk paper -applied to fixed radius point counts.

-We will not discuss.

Marked Subpopulation (Usually Telemetry Tags)

P 255-256

- Rice and Harder(1977) JWM used this approach with deer and tags visible from the air.

- Many later authors used radio tags.

- Only simplest case presented p255-256.

Marked Subpopulation (Usually Telemetry Tags)

P 255-256

M - radio-tagged animals present in the population.

m - radio-tagged animals counted from the air.

$$\hat{\beta} = \frac{m}{M}$$

n - **animals counted from the air.**

$$\hat{N} = \frac{n}{\hat{\beta}}$$

NOTE- This is basically another version of the Lincoln-Petersen method. Assumes all the area sampled or α would also have to be included.

Marked Subpopulation (Usually Telemetry Tags)

Numerical Example

$\mathrm{M} = 50$ - radio-tagged animals.

$m = 45$ - radio-tagged animals counted from the air.

$$\hat{\beta} = \frac{m}{M} = \frac{45}{50} = 0.9$$

$n = 465$ **- all animals counted from the air.**

$$\hat{N} = \frac{n}{\hat{\beta}} = \frac{465}{0.90} = 516.67$$

Extensions to More Complex Sampling Designs

- Note in our examples we basically assumed that our plots were chosen by simple random sampling.
- We can extend them to cases where we have stratified random sampling. In that case we would have to decide whether we wanted to estimate the detection probability in each strata separately or just use one overall detection probability.

Next Lecture

- The use and misuse of population indicies and comparison with the use of absolute abundance estimates.
- This material is presented in Section 12.7 of the text