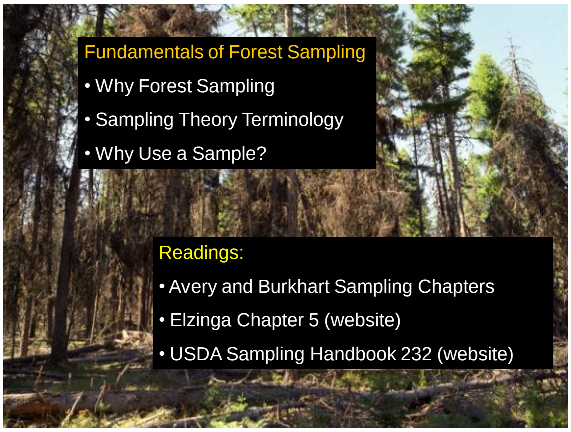
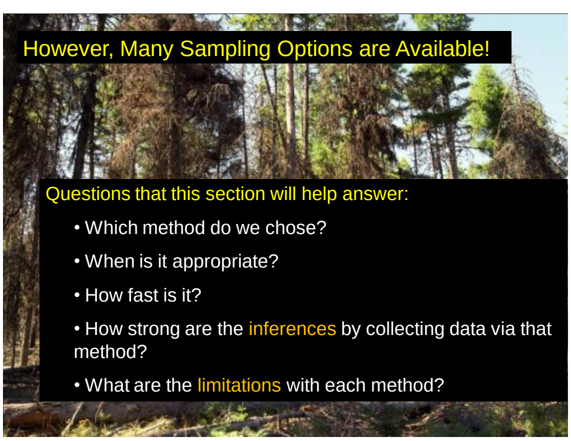

Fundamentals of Forest Sampling
• Why Forest Sampling
• Sampling Theory Terminology
• Why Use a Sample?
Readings:
• Avery and Burkhart Sampling Chapters
• Elzinga Chapter 5 (website)
• USDA Sampling Handbook 232 (website)


Why do we Care About Forest Sampling?
Forest managers are required to make important decisions relating to numerous (sometimes competing) resources.
Ideally such decisions would be based off an inventory of all the organisms, but cost, feasibility, and timing make this rare
As such, we must make decision based off only a small portion, or sample, of the data.


However, Many Sampling Options are Available!
Questions that this section will help answer:
• Which method do we chose?
• When is it appropriate?
• How fast is it?
• How strong are the inferences by collecting data via that method?
• What are the limitations with each method?

Sampling Theory: Terminology

Common Terms Include:

- Statistical Study
- Sampling
- Data
- Sampling Unit
- Frame
- Population
- Parameters
- Statistics
- Accuracy and Bias
- Distributions
- Confidence

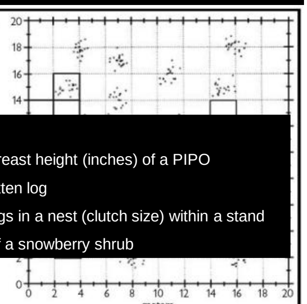
A "statistical study": A process that always involves measurements to produce numbers

Sample information

Coordinates X	Coordinates Y	# of plants
2	2	4
6	4	0
16	4	3
12	6	2
14	6	5
6	8	10
0	12	0
2	12	6
14	12	0
2	14	20

Population parameters
Total population size:
400 plants
Mean # plants/quadrat:
$\mu = 4$
Standard deviation:
$\sigma = 5.005$

interval = ± 361 plants



Examples:

- The diameter at breast height (inches) of a PIPO
- The length of a rotten log
- The number of eggs in a nest (clutch size) within a stand
- The height (cm) of a snowberry shrub

Elzinga et al Chapter 5

FIGURE 5.1. Population of 400 plants distributed in 20 clumps of 20 plants. This figure shows a simple random sample of ten 2m x 2m quadrats, along with sample statistics and true population parameters.

Sampling: process of inferring properties of a population from only a sample of that population

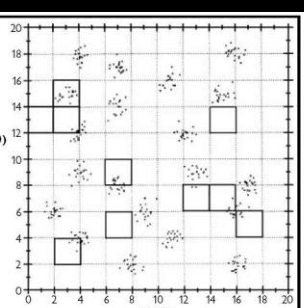
Sample information

Coordinates X	Coordinates Y	# of plants
2	2	4
6	4	0
16	4	3
12	6	2
14	6	5
6	8	10
0	12	0
2	12	6
14	12	0
2	14	20

Population parameters
Total population size:
400 plants
Mean # plants/quadrat:
$\mu = 4$
Standard deviation:
$\sigma = 5.005$

Sample statistics (n = 10)
Mean # plants/quadrat:
$\bar{x} = 5.0$
Standard deviation:
$s = 6.146$

Population estimate
Estimated population size = 500 plants
95% confidence interval = ± 361 plants



Elzinga et al Chapter 5

FIGURE 5.1. Population of 400 plants distributed in 20 clumps of 20 plants. This figure shows a simple random sample of ten 2m x 2m quadrats, along with sample statistics and true population parameters.

Data: Collected from measuring a sample, which is made up of sampling units.

Sample information

Coordinates X	Coordinates Y	# of plants
2	2	4
6	4	0
16	4	3
12	6	2
14	6	5
6	8	10
0	12	0
2	12	6
14	12	0
2	14	20

Population parameters
Total population size: 400 plants
Mean # plants/quadrat: $\mu = 4$
Standard deviation: $\sigma = 5.005$

Sample statistics (n = 10)
Mean # plants/quadrat: $\bar{x} = 5.0$
Standard deviation: $s = 6.146$

Population estimate
Estimated population size = 500 plants
95% confidence interval = ± 361 plants

Elzinga et al Chapter 5

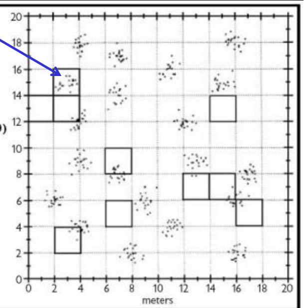


FIGURE 5.1. Population of 400 plants distributed in 20 clumps of 20 plants. This figure shows a simple random sample of ten 2m x 2m quadrats, along with sample statistics and true population parameters.

Sampling Unit: The part that the population of interest is divided into. This is what we infer from

Sample information

Coordinates X	Coordinates Y	# of plants
2	2	4
6	4	0
16	4	3
12	6	2
14	6	5
6	8	10
0	12	0
2	12	6
14	12	0
2	14	20

Population parameters
Total population size: 400 plants
Mean # plants/quadrat: $\mu = 4$
Standard deviation: $\sigma = 5.005$

Sample statistics (n = 10)
Mean # plants/quadrat: $\bar{x} = 5.0$
Standard deviation: $s = 6.146$

Population estimate
Estimated population size = 500 plants
95% confidence interval = ± 361 plants

Elzinga et al Chapter 5

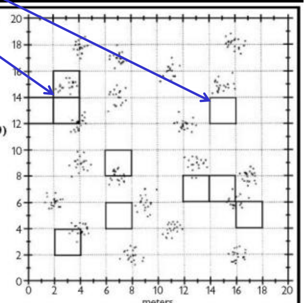


FIGURE 5.1. Population of 400 plants distributed in 20 clumps of 20 plants. This figure shows a simple random sample of ten 2m x 2m quadrats, along with sample statistics and true population parameters.

Sampling Unit: This is often called the experimental unit and is the physical unit from which the measurement is obtained

Sample information

Coordinates X	Coordinates Y	# of plants
2	2	4

Population parameters
Total population size: 400 plants
Mean # plants/quadrat:

Examples:

- An individual PIPO
- A fuels inventory plot
- A woodpeckers nest
- A shrub monitoring plot

Estimated population size = 500 plants
95% confidence interval = ± 361 plants

Elzinga et al Chapter 5

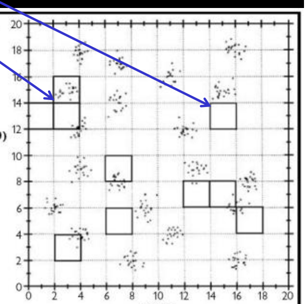
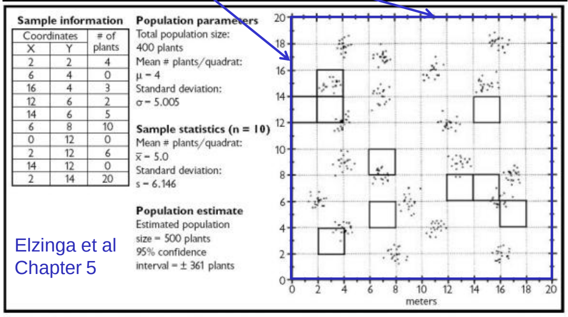


FIGURE 5.1. Population of 400 plants distributed in 20 clumps of 20 plants. This figure shows a simple random sample of ten 2m x 2m quadrats, along with sample statistics and true population parameters.

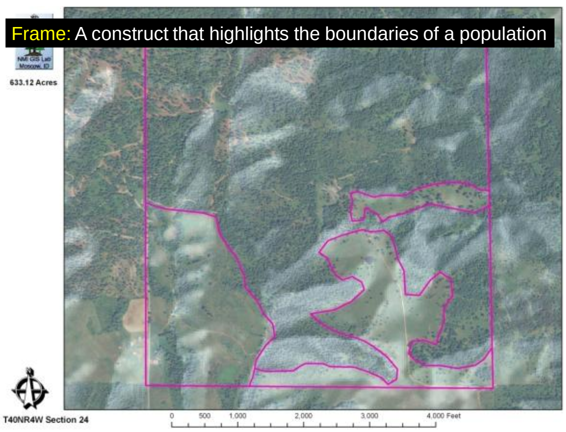
Frame: A construct that highlights the boundaries of a population



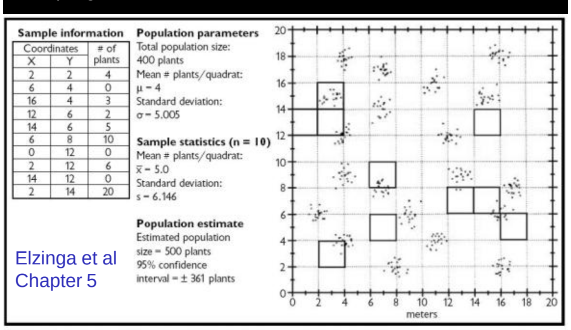
Elzinga et al Chapter 5

FIGURE 5.1. Population of 400 plants distributed in 20 clumps of 20 plants. This figure shows a simple random sample of ten 2m x 2m quadrats, along with sample statistics and true population parameters.

Frame: A construct that highlights the boundaries of a population



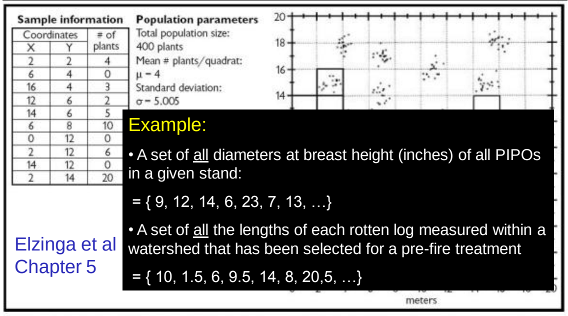
Population (or Universe/Universal Set): The collection of all the sampling units = what we want to learn about



Elzinga et al Chapter 5

FIGURE 5.1. Population of 400 plants distributed in 20 clumps of 20 plants. This figure shows a simple random sample of ten 2m x 2m quadrats, along with sample statistics and true population parameters.

Population (or Universe/Universal Set): The collection of all the sampling units = what we want to learn about



Sample information

Coordinates X	Coordinates Y	# of plants
2	2	4
6	4	0
16	4	3
12	6	2
14	6	5
6	8	10
0	12	0
2	12	6
14	12	0
2	14	20

Example:

- A set of all diameters at breast height (inches) of all PIPOs in a given stand:

= { 9, 12, 14, 6, 23, 7, 13, …}

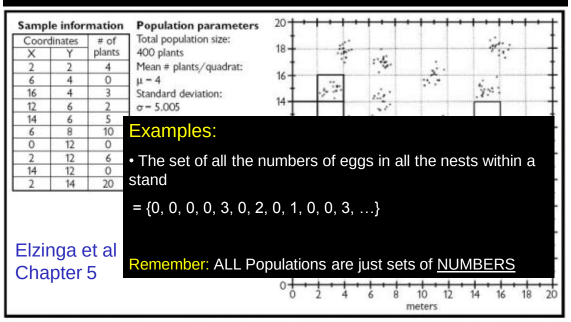
- A set of all the lengths of each rotten log measured within a watershed that has been selected for a pre-fire treatment

= { 10, 1.5, 6, 9.5, 14, 8, 20,5, …}

Elzinga et al Chapter 5

FIGURE 5.1. Population of 400 plants distributed in 20 clumps of 20 plants. This figure shows a simple random sample of ten 2m x 2m quadrats, along with sample statistics and true population parameters.

Population (or Universe/Universal Set): The collection of all the sampling units = what we want to learn about



Sample information

Coordinates X	Coordinates Y	# of plants
2	2	4
6	4	0
16	4	3
12	6	2
14	6	5
6	8	10
0	12	0
2	12	6
14	12	0
2	14	20

Examples:

- The set of all the numbers of eggs in all the nests within a stand

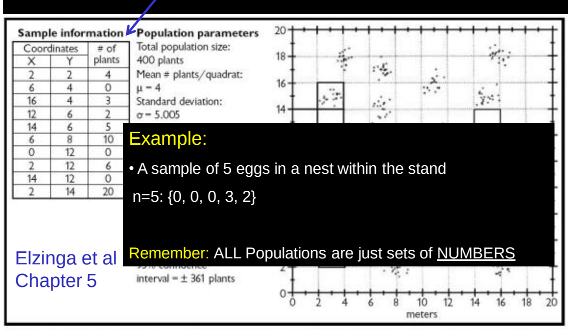
= {0, 0, 0, 0, 3, 0, 2, 0, 1, 0, 0, 3, …}

Remember: ALL Populations are just sets of NUMBERS

Elzinga et al Chapter 5

FIGURE 5.1. Population of 400 plants distributed in 20 clumps of 20 plants. This figure shows a simple random sample of ten 2m x 2m quadrats, along with sample statistics and true population parameters.

A Sample: is any subset of the population



Sample information

Coordinates X	Coordinates Y	# of plants
2	2	4
6	4	0
16	4	3
12	6	2
14	6	5
6	8	10
0	12	0
2	12	6
14	12	0
2	14	20

Example:

- A sample of 5 eggs in a nest within the stand

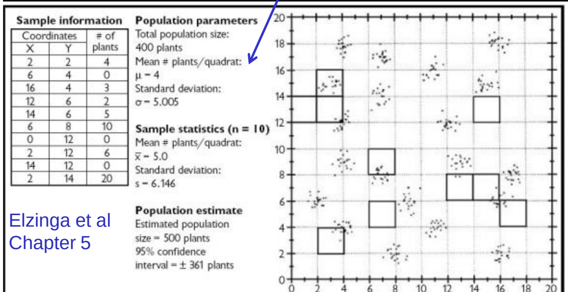
n=5: {0, 0, 0, 3, 2}

Remember: ALL Populations are just sets of NUMBERS

Elzinga et al Chapter 5

FIGURE 5.1. Population of 400 plants distributed in 20 clumps of 20 plants. This figure shows a simple random sample of ten 2m x 2m quadrats, along with sample statistics and true population parameters.

Parameters: Characteristics of a population (using all the data), such as total number of measures per sampling unit

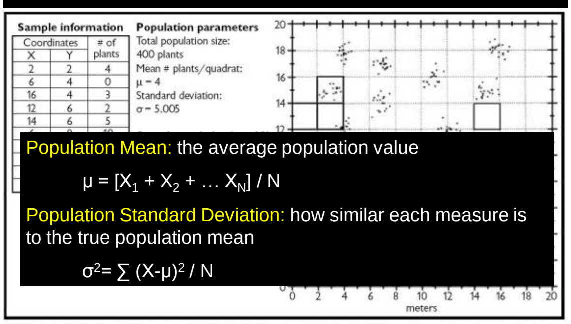


Coordinates X	Coordinates Y	# of plants
2	2	4
6	4	0
16	4	3
12	6	2
14	6	5
6	8	10
0	12	0
2	12	6
14	12	0
2	14	20

Elzinga et al Chapter 5

The aim of sampling is to measure a representative sample of the population in to obtain a "reasonable estimate of your parameter"

Population Parameters: Fixed, unknown, and only change if the population changes



Population Mean: the average population value

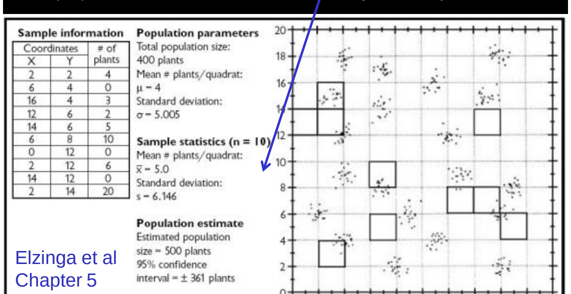
$$\mu = [X_1 + X_2 + \ldots X_N] / N$$

Population Standard Deviation: how similar each measure is to the true population mean

$$\sigma^2 = \sum (X-\mu)^2 / N$$

FIGURE 5.1. Population of 400 plants distributed in 20 clumps of 20 plants. This figure shows a simple random sample of ten 2m x 2m quadrats, along with sample statistics and true population parameters.

Statistics: Estimates of your parameters population (using a sample), and their errors calculated from your samples



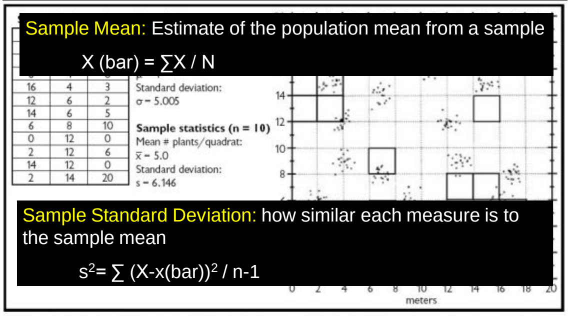
Elzinga et al Chapter 5

Note: Greek letters (μ, σ) are used for parameters and Roman (x bar, s) for Statistics

Sample Statistics: Derived descriptors derived from a sample that provide estimates of the population parameter

Sample Mean: Estimate of the population mean from a sample

$\bar{X}$ (bar) = $\sum X / N$



Sample Standard Deviation: how similar each measure is to the sample mean

$s^2 = \sum (X - \bar{x})^2 / n-1$

FIGURE 5.1. Population of 400 plants distributed in 20 clumps of 20 plants. This figure shows a simple random sample of ten 2m x 2m quadrats, along with sample statistics and true population parameters.

Distributions: The Law of Large Numbers states that as the sample size increases the sample statistics tend toward the population parameters, regardless of the population distribution

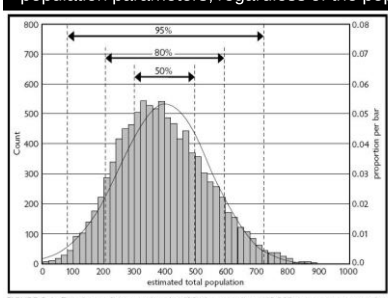
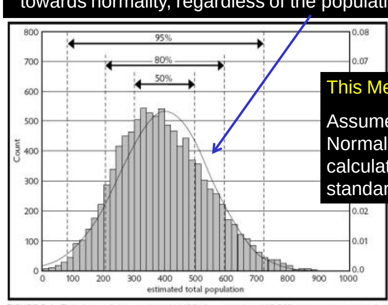


FIGURE 5.4. Distribution from sampling the 400-plant population 10,000 times using ten samples of 2m x 2m quadrats. The 95%, 80%, and 50% confidence intervals around the true population of 400 plants are shown. The smooth line shows a normal, bell-shape curve fit to the data.

Elzinga et al Chapter 5

Distributions: The Central Limit Theorem states that as the sample size increases the sample mean distribution tends towards normality, regardless of the population distribution



This Means that:

Assume a distribution is Normal and then can easily calculate its mean and standard deviation

FIGURE 5.4. Distribution from sampling the 400-plant population 10,000 times using ten samples of 2m x 2m quadrats. The 95%, 80%, and 50% confidence intervals around the true population of 400 plants are shown. The smooth line shows a normal, bell-shape curve fit to the data.

Elzinga et al Chapter 5

A Note of Confidence: This is the measure of precision about the mean, such that Y% of the time our sample value will lie within a given range on the distribution

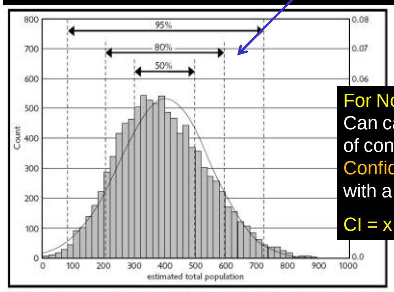


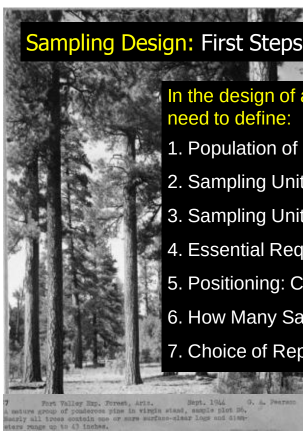
FIGURE 5.4. Distribution from sampling the 400-plant population 10,000 times using ten samples of 2m x 2m quadrats. The 95%, 80%, and 50% confidence intervals around the true population of 400 plants are shown. The smooth line shows a normal, bell-shape curve fit to the data.

For Normal Distributions: Can calculate this range of confidence or Confidence Interval (CI) with a students t-table

CI = x ± t*SE

Elzinga et al Chapter 5

Sampling Design: First Steps

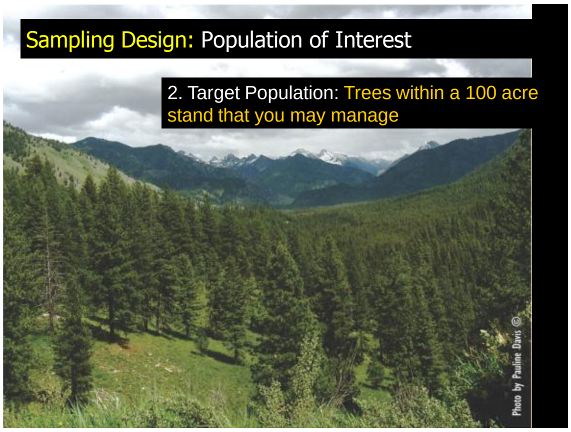


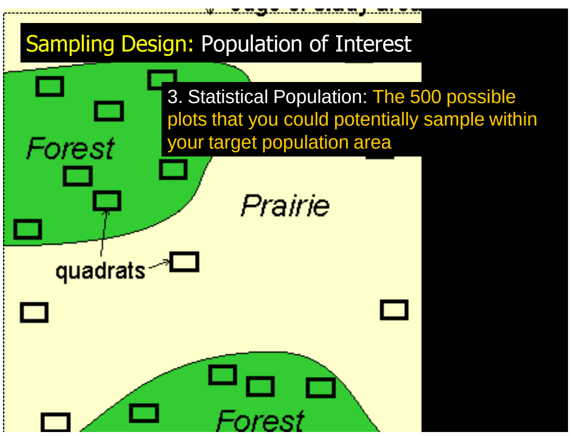
In the design of a sampling process we need to define:

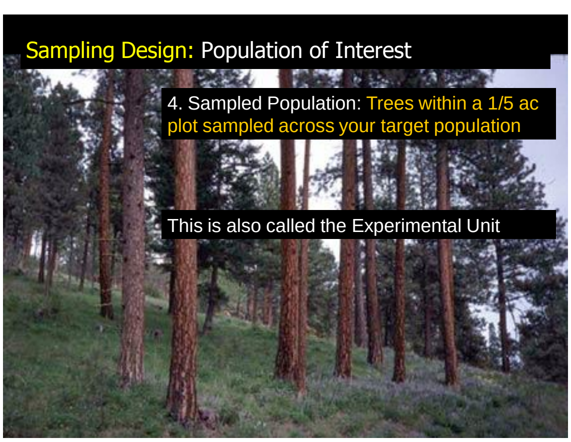
1. Population of Interest
2. Sampling Unit
3. Sampling Unit Size and Shape
4. Essential Requirements
5. Positioning: Choice of Sampling Design
6. How Many Sampling Units (next lecture)
7. Choice of Representative Measures

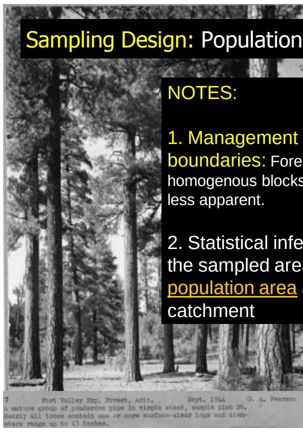
Sampling Design: Population of Interest

1. Biological Population


Sampling Design: Population of Interest
2. Target Population: Trees within a 100 acre stand that you may manage
Photo by Pauline Davis ©


Sampling Design: Population of Interest
3. Statistical Population: The 500 possible plots that you could potentially sample within your target population area
Forest
Prairie
quadrats
Forest


Sampling Design: Population of Interest
4. Sampled Population: Trees within a 1/5 ac plot sampled across your target population
This is also called the Experimental Unit

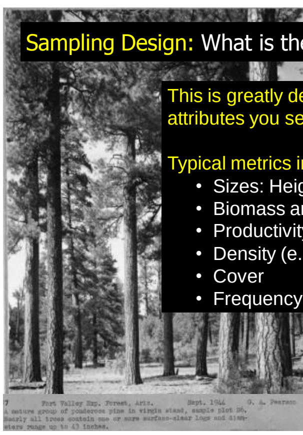


Sampling Design: Population of Interest

NOTES:

1. Management usually ignores biological boundaries: Forest stands are often defined via homogenous blocks but over time this may become less apparent.

2. Statistical inferences call only be made to the sampled areas within your targeted population area and not to the wider catchment

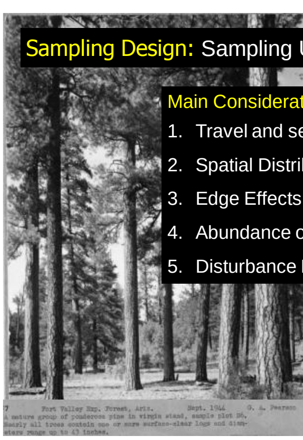


Sampling Design: What is the Sampling Unit

This is greatly dependant on the type of attributes you seek to measure.

Typical metrics in forestry include:

- Sizes: Heights and Diameters
- Biomass and Volume (e.g., Bd ft)
- Productivity (e.g., NPP)
- Density (e.g., SDI)
- Cover
- Frequency

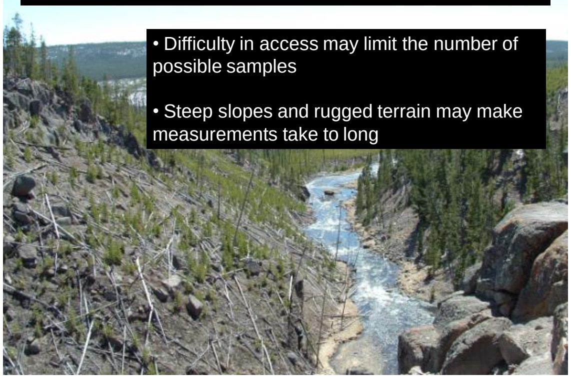


Sampling Design: Sampling Unit Size and Shape

Main Considerations:

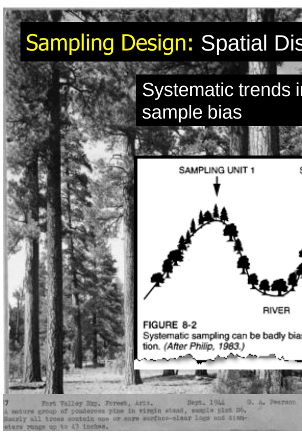
1. Travel and set-up time
2. Spatial Distribution
3. Edge Effects
4. Abundance of Trees of Interest
5. Disturbance Impacts

Sampling Design: Travel and set-up time



- Difficulty in access may limit the number of possible samples
- Steep slopes and rugged terrain may make measurements take to long

Sampling Design: Spatial Distribution



Systematic trends in topography may lead to sample bias

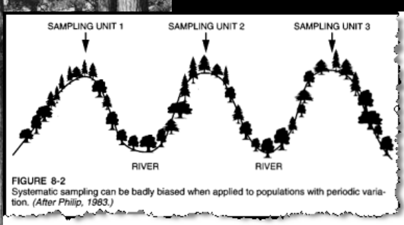
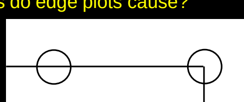


FIGURE 8-2
Systematic sampling can be badly biased when applied to populations with periodic variation. *(After Philip, 1983.)*

Fixed Area Plots: Stand Boundaries

What issues do edge plots cause?



- Plots for which the plot center is outside the boundary will not be measured. So, trees close to the boundary are less likely to be sampled and are under-represented.
- Portions of our plots may land outside the population. If we count such plots as being full-sized, we bias our statistics.

Why might the edge trees differ from the central trees?

Fixed Area Plots: Stand Boundaries

Edge trees can exhibit:
- Less competition
- More wind impacts

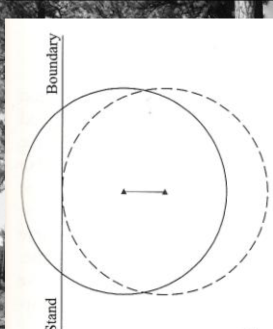
Fixed Area Plots: Stand Boundaries

Solution 1: Ignore it.

When large stands are cruised with small circular plots – the bias can be considered negligible

But when cruised tracts are narrow and long – i.e. more likely to have edge plots there are several methods that can be used

Fixed Area Plots: Stand Boundaries



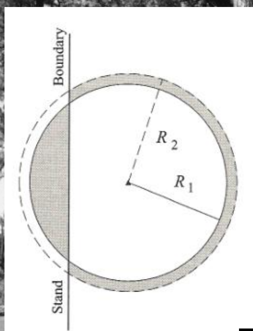
Solution 2: Move plot so it falls within boundary

Worst Method!

- Edge trees will be under sampled
- Can lead to significant bias if stand has lots of edges!

Source: Husch Beers and Kershaw

Fixed Area Plots: Stand Boundaries



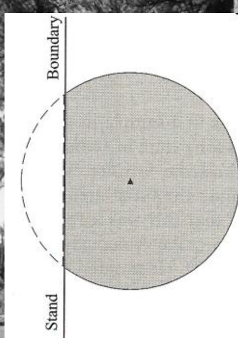
Solution 3: Add additional radius to account for lost area

Intermediate Method

• Edge trees will be under sampled

Source: Husch Beers and Kershaw

Fixed Area Plots: Stand Boundaries



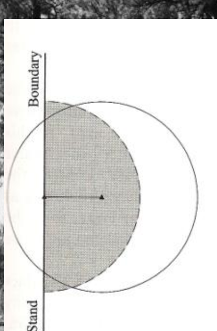
Solution 4: Re-calculate area and only measure within stand

Intermediate Method

• Very time consuming as need to infer samples under correct areas

Source: Husch Beers and Kershaw

Fixed Area Plots: Stand Boundaries



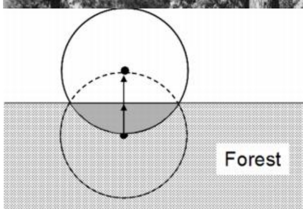
Solution 5: Establish exactly half a plot at the stand edge → Then double count

Intermediate Method

• Edge trees will be over sampled leading to bias

Source: Husch Beers and Kershaw

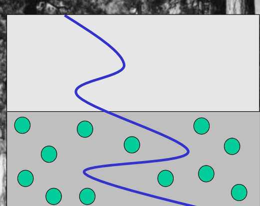
Fixed Area Plots: Stand Boundaries



Solution 6: Mirage Plots
Intermediate method

- Lay out your plot and measure all the “in” trees
- Imagining the stand edge a mirror, lay out the mirror image of your plot with the plot centre outside the stand and measure the “in” trees
- Edge trees will be over sampled leading to bias

Fixed Area Plots: Stand Boundaries



Solution 7: Don’t Place a Plot at the edge in the first place!

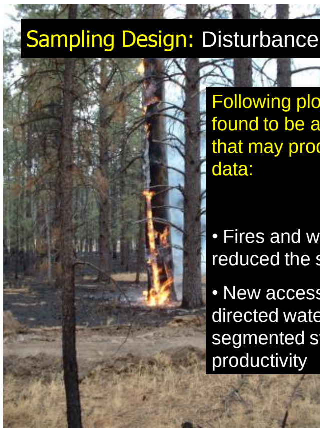
Use Buffers around edges, roads, and rivers

- Biases against edge trees

Sampling Design: Abundance of Trees of Interest

We may wish to increase our number of samples if we are interested in evaluating rare species.

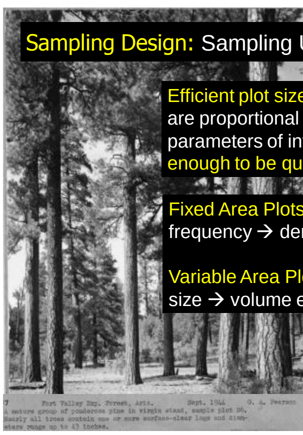
We may also stratify our samples if we have a priori knowledge of what environmental conditions might affect the spatial distribution of those species



Sampling Design: Disturbance Impacts

Following plot selection the area may be found to be affected by a disturbance that may produce unrepresentative data:

- Fires and wind events may have reduced the stocking
- New access roads, power lines, re-directed water channels etc may have segmented stands or adversly affected productivity

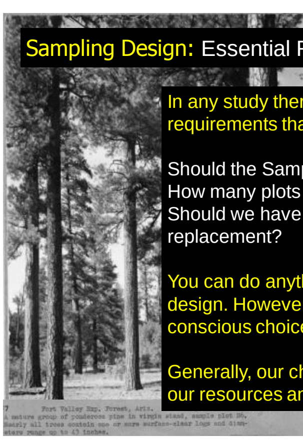


Sampling Design: Sampling Unit Size and Shape

Efficient plot sizes have sample sizes that are proportional to the variance of the parameters of interest, but that are small enough to be quickly collected

Fixed Area Plots: samples proportional to frequency → density estimates

Variable Area Plots: samples proportional to size → volume estimates



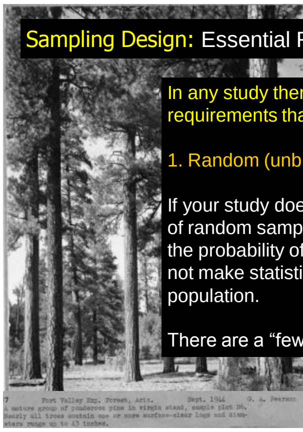
Sampling Design: Essential Requirements

In any study there are three general requirements that should be considered

Should the Sampling Design be Random?
How many plots (samples) are needed?
Should we have sampling with or without replacement?

You can do anything in your sampling design. However, you need to make a conscious choice for everything you do

Generally, our choices should be guided on our resources and objectives



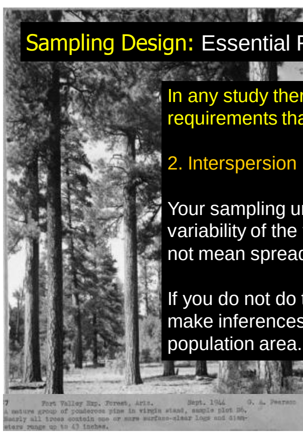
Sampling Design: Essential Requirements

In any study there are three general requirements that should be considered:

1. Random (unbiased) Sampling

If your study does not contain some manner of random sampling you can not determine the probability of selection and therefore can not make statistical inferences about your population.

There are a "few" exceptions …



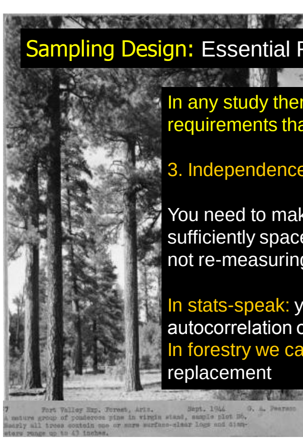
Sampling Design: Essential Requirements

In any study there are three general requirements that should be considered:

2. Interspersion

Your sampling units should capture the variability of the target population (this does not mean spread out over the entire area).

If you do not do this it is very difficult to make inferences of the entire target population area.



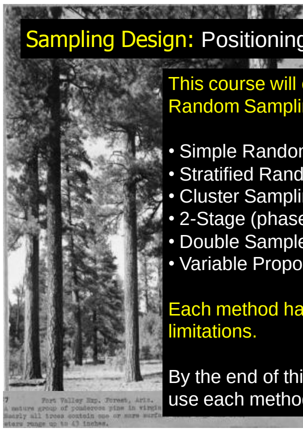
Sampling Design: Essential Requirements

In any study there are three general requirements that should be considered:

3. Independence

You need to make sure that your plots are sufficiently spaced apart such that you are not re-measuring the same information.

In stats-speak: you need to avoid spatial autocorrelation or "over sampling"
In forestry we call this: sampling without replacement



Sampling Design: Positioning

This course will cover several types of Random Sampling:

- Simple Random Sampling
- Stratified Random Sampling
- Cluster Sampling
- 2-Stage (phase) Sampling
- Double Sample
- Variable Proportion Sampling

Each method has its advantages and limitations.

By the end of this course you will be able to use each method appropriately