7.0 Sampling Designs in Stand Inventory

Most of the time it's impractical to conduct a 100% inventory; sampling can provide all the necessary information in less time and for lower cost

7.1 Simple Random Sampling

Simple Random Sampling (SRS) is *the* fundamental sample selection method.

All other sample selection methods are some modification of SRS designed to be more efficient or less costly, or both.

In SRS, each combination of n sample units (tree, plot, transect, etc.) has an equal chance of being selected from the population.

Selection of each sample unit must be free from deliberate or subjective choice and may not influence the selection of any of the remaining sample units

The definition of the sample unit defines the size, N, of the population

Primary control – based on section corners; some permanent fixed object

Secondary control – usually established by hand compass and pacing

SRS has nice properties

- The sample mean $\bar{x}$ and sample variance s^2 are *unbiased* estimates of the population mean μ_x and variance σ^2, respectively
- Central Limit Theorem indicates that when $n \geq 30$, sample mean is normally distributed

Analysis of resulting data is summarized as follows:

Given:
- n = number of sampling units measured
- N = total number of sampling units in the population
- x_i = quantity of X measured on i^{th} sampling unit
- $\bar{x}$ = sample mean of X per sampling unit; estimates pop'n mean
- s = standard deviation of sample
- $s_{\bar{x}}$ = standard error of mean
- $\hat{X}$ = estimated total of X for the population
- CV = Coefficient of Variation
- $E\%$ = allowable error as a percent of the mean
- t = Student's t for desired level of confidence

Then,

Sample mean, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Sample variance, $s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$, or, $s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}{n-1}$

Variance of the mean, $s_{\bar{x}}^2 = \frac{s^2}{n}\left(\frac{N-n}{N}\right)$

Estimate of the total, $\hat{X} = N\bar{x}$

Variance of the total, $s_{\hat{X}}^2 = N^2 s_{\bar{x}}^2$

Example. We are interested in knowing the number of cords per acre on our 100-ac mixed-wood forest. We've got a budget for 20, 1-acre fixed area plots. We inventory the forest with the following results:

Plot	cords	Plot	cords	Plot	cords	Plot	cords
1	10	6	41	11	29	16	57
2	17	7	39	12	33	17	54
3	19	8	37	13	34	18	47
4	15	9	41	14	40	19	39
5	9	10	27	15	22	20	36

Data Analysis

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{20}(10+17+19+\ldots+47+39+36) = 32.3 \text{ cds / acre}$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}{n-1} = \frac{24{,}398 - \frac{1}{20}(646)^2}{19} = 185.90 \text{ (cds / acre)}^2$$

$$s_{\bar{x}}^2 = \frac{s^2}{n}\left(\frac{N-n}{N}\right) = \frac{185.90}{20}\left(\frac{100-20}{100}\right) = 7.436 \text{ (cds / acre)}^2$$

$$\hat{X} = N\bar{x} = 100 \bullet 32.30 = 3230 \text{ cds}$$

$$s_{\hat{X}}^2 = N^2 s_{\bar{x}}^2 = 100^2 \bullet 7.436 = 74{,}360 \text{ cds}^2$$

Confidence Interval

On the mean …

$$\bar{x} \pm t_{\alpha,n-1} \bullet s_{\bar{x}}, \qquad \text{where } s_{\bar{x}} = \sqrt{s_{\bar{x}}^2}$$

So for 95% CI, in this case n - 1 = 20 – 1 = 19, and $t_{0.05,19}$ = 2.093

$$32.30 \pm 2.093 \bullet \sqrt{7.436} \quad \Rightarrow \quad (26.59, 38.01) \text{ cds/acre}$$

On the total …

$$\hat{X} \pm t_{\alpha,n-1} \bullet s_{\hat{X}}, \qquad \text{where } s_{\hat{X}} = \sqrt{s_{\hat{X}}^2}$$

So for 95% CI, same t-value as before, i.e., $t_{0.05,19}$ = 2.093

$$3230 \pm 2.093 \bullet \sqrt{74360} \quad \Rightarrow \quad (2659, 3801) \text{ cds/acre}$$

Sample Size Calculation

Infinite population: $$n = \frac{t^2 (CV)^2}{E\%^2}$$

Finite population: $$n = \frac{Nt^2 (CV)^2}{NE\%^2 + t^2 (CV)^2},$$

all variables as defined before.

NOTE: this is an *iterative* formula, because you need to know degrees of freedom, $df = n - 1$, before you can choose t. To make it work, take a guess at n (guess high, say 31 or more) then get a t-value. Plug in N, t, and CV to compute the next sample size, n. If it is higher or lower than your guess, use that new n to obtain new df to choose a new t. Do the calculation again. When the sample size you compute agrees with the sample size you put in, you're done – it shouldn't take more than 2 or 3 iterations.

Example. How many samples are needed to estimate the mean with 10% Error at 95% confidence using fixed area 1/10-acre plots in a 40-acre forest with a CV = 33.2%?

Guess n_0 = 31, giving df = 30, and $t_{0.05,30}$ = 2.042

$$n_1 = \frac{400(2.042)^2(33.2)^2}{400(10)^2 + (2.042)^2(33.2)^2} = 41.2 \sim 42$$

Next iteration with n_1 = 42, giving df = 41, and $t_{0.05,41}$ = 2.020

$$n_2 = \frac{400(2.020)^2(33.2)^2}{400(10)^2 + (2.020)^2(33.2)^2} = 40.4 \sim 41$$

Next iteration with n_2 = 41, giving df = 40, and $t_{0.05,40}$ = 2.021

$$n_3 = \frac{400(2.021)^2(33.2)^2}{400(10)^2 + (2.021)^2(33.2)^2} = 40.5 \sim 41$$ STOP.

We need 41 sample plots for the desired inventory parameters. Note that we always increase n to the next largest integer – whole plots only.

Sample Size Shortcut Formulas

- For SRS infinite populations (or sampling with replacement)

$$n = k + \frac{z^2(CV)^2}{E\%^2},$$ where

k = correction term to avoid iterating between *t*-values

Confidence level	z-value	k
80%	1.282	1.31
90%	1.645	1.87
95%	1.960	2.44
99%	2.576	3.79

z = standard normal deviate

all other symbols as before

- For SRS finite populations (or sampling without replacement)

$$n = k + \frac{Nz^2(CV)^2}{NE\%^2 + z^2(CV)^2}$$

– Rules of thumb:

For ~ 1/10 acre plots in highly variable populations:

Area (in acres)	number of samples
Up to 10	10
11 – 40	1 per acre
41 – 80	20 + 0.5 (area in acres)
81 – 200	40 + 0.25(area in acres)
200 +	Use sample size formulas

Disadvantages of SRS

Difficult to map the stand concurrent with cruise (plots are located randomly)

Spend a disproportionate amount of time locating plots vs. measuring them

May end up with an unrepresentative sample

Systematic Sampling

Sampling units are spaced at fixed intervals throughout the population

Since the units occur at regular intervals, the number of outcomes, is fixed, i.e., if the sampling interval is every k units, there are only k possible outcomes

For the mean to be unbiased, one of the k outcomes has to be randomly chosen

A sample unit can be chosen at random from the entire population; usually one of the first k is chosen – systematic sampling w/ random start

If the population itself is randomly distributed, a systematic sample is equivalent to a random sample; statistical formulas for random sampling can be used

If some pattern in variation is present in the population, then there is no way to determine sampling error,

Real problems with bias can occur if the sampling interval, k, happens to coincide with the pattern of variation in the population

The larger the sampled area, the more systematic sampling results agree with simple random sampling results

Over the long run, treating a systematic sample as if it were random, provides a larger standard error compared to using SRS, leading to wider confidence intervals compared to SRS

Hypothetical 40-ac. forest divided into 0.1-acre, square, fixed-area sample plots

22	26	26	19	34	18	17	25	20	28	0	0	2	0	6	0	3	0	0	4
21	18	23	22	28	24	33	36	23	15	17	0	2	11	15	0	17	5	2	8
30	28	23	21	29	18	14	30	25	28	20	14	8	1	15	2	5	0	0	4
28	19	21	20	26	20	38	23	20	27	24	11	6	4	0	5	5	9	2	5
17	14	20	26	25	22	22	19	15	20	25	26	15	9	12	0	0	16	8	5
38	42	37	39	22	44	47	17	25	29	34	39	20	24	14	10	12	1	0	0
43	34	23	46	47	46	39	35	31	30	24	35	23	26	18	25	21	12	5	1
36	45	47	36	35	29	49	44	31	42	33	47	31	28	15	18	20	23	9	12
38	48	42	51	17	54	47	52	30	34	30	46	24	12	21	12	32	29	27	16
17	24	45	47	52	28	43	45	46	27	40	32	51	28	25	41	20	27	14	21
47	56	43	37	30	60	56	43	29	33	41	30	30	20	15	18	20	23	13	28
38	36	27	38	24	31	48	32	25	31	35	31	22	15	10	24	22	19	18	16
46	54	34	46	37	43	44	34	25	44	43	40	28	14	26	33	18	34	17	31
47	47	35	40	39	39	50	28	50	36	44	27	16	21	36	17	27	21	33	31
22	37	11	29	28	33	29	35	53	18	26	20	29	12	23	25	15	17	26	21
37	22	20	22	14	29	25	21	30	31	27	16	21	36	17	27	21	17	14	12
18	25	26	20	27	26	26	22	19	30	15	2	7	3	4	15	9	7	0	3
22	34	37	17	24	28	26	25	30	33	18	6	0	3	0	0	6	5	7	6
13	28	27	27	31	20	21	21	16	27	12	0	17	3	1	1	0	0	0	5
32	31	30	22	16	26	8	32	19	22	3	3	0	0	0	2	0	5	10	2

Advantages to Sys Sampling (over SRS)

- Stand can be mapped easily concurrently with sampling
- More efficient trade-off between locating plots & measuring them
- Representative sample is guaranteed (provided there are no identifiable cyclical patterns in population, i.e., its spatial distribution is random)

EXHIBIT "A"

REPRODUCTION SURVEY SPECIFICATIONS

1. Work shall start in the Grays Harbor Tree Farm on the 1st day of October and will be completed no later than in the Grays Harbor Tree Farm.

2. A 10% performance bond will be withheld from each payment.

3. Payment is made only after RTOC's representative has determined units have been done properly.

4. Failure to do survey plots properly will result in no payment for the entire unit until the problem is corrected.

5. Each plot center will be marked with a flag or ribbon provided by RTOC.

FIELD PROCEDURES

The goal of the regeneration survey is to establish one observation point (plot) on each clearcut acre to determine if adequate reproduction of desirable trees is being attained.

1.) Plot Location:

To achieve this sampling density, lines will be run through the clearcut units in the cardinal directions at five chain (330') intervals. Plots will then be placed along the lines at 2 chain (132') intervals. Direction will be determined by hand compass and distances measured by pacing. The direction of the lines, either East-West or North-south, and location of the first plot are left to the discretion of RTOC's representative.

East-west lines will be numbered starting with the most northerly line as number one. The next line to the south will be line 2 and so on consecutively numbering the lines to the southern most line in the unit. The most westerly plot in the unit is numbered one with consecutive numbering proceeding to the East. North-South lines will be consecutively numbered from West to East. Plots will be consecutively numbered from North to South.

If a plot falls along a type line, a stream, or a road make the appropriate notation on the regeneration survey form and go forward or backward one half chain along the line and take the plot.

Areas 10 acres or smaller may be run with a random line. Its direction and length will be pre-determined before visiting the area and will be mapped on the 1" = 400' maps showing starting point, ending point, and bearings run. All plots will be considered on line one, and will be numbered consecutively from the starting to ending point.

Figure 1

East-West Line Configuration

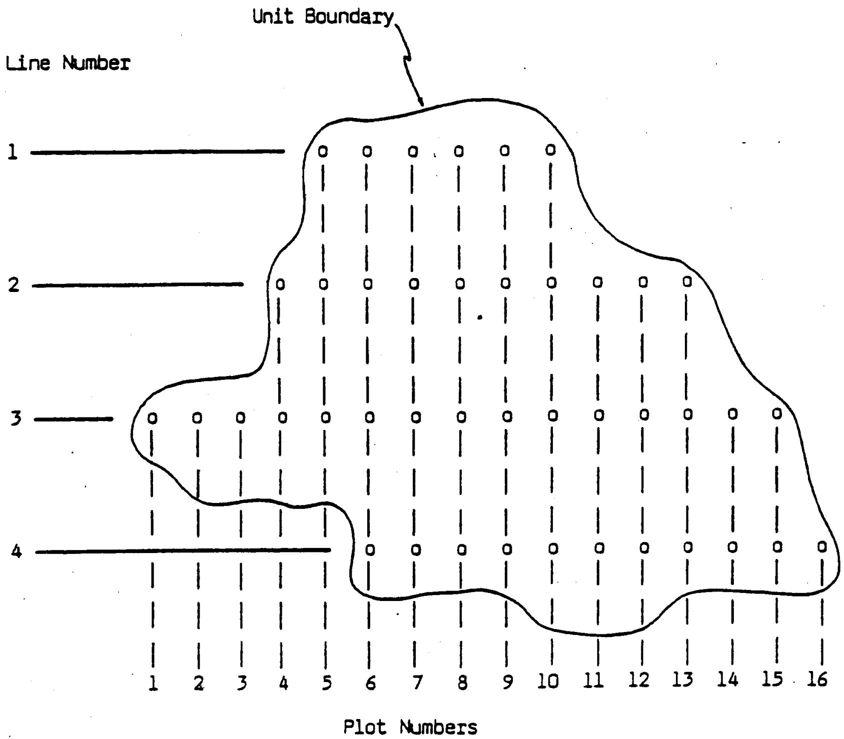

Figure 2

North-South Line Configuration