Sampling Design

Practical Applications

Sampling Design - Objectives

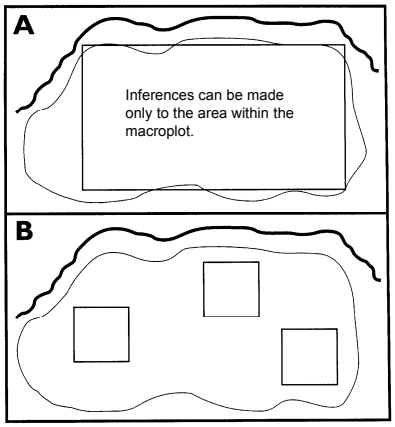
- List the six primary decisions common to all sampling designs
- List at least three issues affecting sampling unit size and shape
- Randomly position sample units
- Solve 4 sampling design problems
- Discuss advantages and disadvantages of permanent and temporary sampling units
- Use sample size equations for:
  - Obtaining an estimate of a population mean
  - Detecting a change over two time periods

6 primary sampling design decisions

- What is the population of interest?
- What is an appropriate sampling unit (S.U.)?
- What is an appropriate S.U. size and shape?
- How should S.U.s be positioned in the population?
- Should S.U. positions be temporary or permanent?
- How many S.U.s should be sampled?

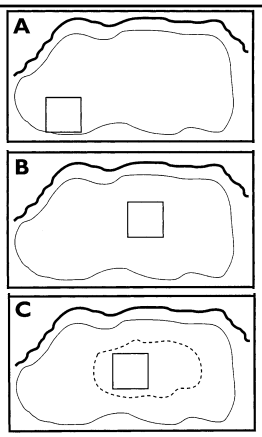
What is the population of interest?

- **Target population:** the area or whole population that you are interested in monitoring
- **Sampled population:** the statistical population that is sampled



Single macroplot covering most of the target population

3 Randomly positioned macroplots randomly positioned in the target population



random placement.

subjective placement.

random placement within a “representative” area, illustrated by the dotted line.

6 primary sampling design decisions

- What is the population of interest?
- What is an appropriate sampling unit (S.U.)?
- What is an appropriate S.U. size and shape?
- How should S.U.s be positioned in the population?
- Should S.U. positions be temporary or permanent?
- How many S.U.s should be sampled?

Common types of Sampling Units

- Quadrats

- Points

- Lines (transects)

- Individual Plants

What is an appropriate sampling unit?

Depends on:

- The attribute of abundance, condition, or population structure sampled as specified in the management objective
- Distribution of the entity sampled
- Size of the entity sampled
- Environmental / physical conditions at site

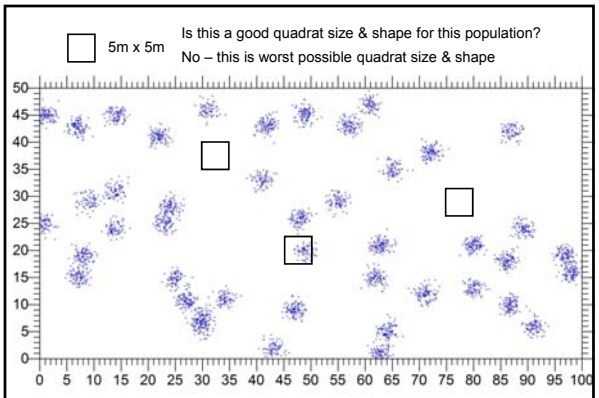
Covered under sampling methods (quadrat, point, line, individual)

6 primary sampling design decisions

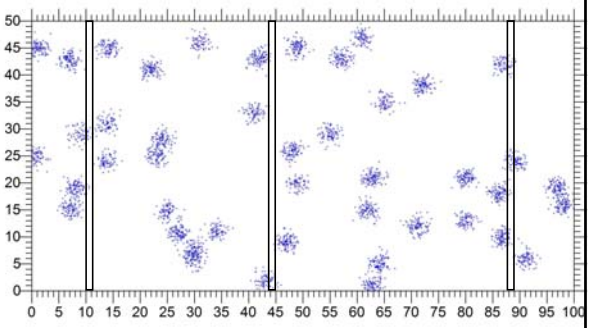
- What is the population of interest?
- What is an appropriate sampling unit (S.U.)?
- What is an appropriate S.U. size and shape?
- How should S.U.s be positioned in the population?
- Should S.U. positions be temporary or permanent?
- How many S.U.s should be sampled?

What is an appropriate S.U. size and shape?

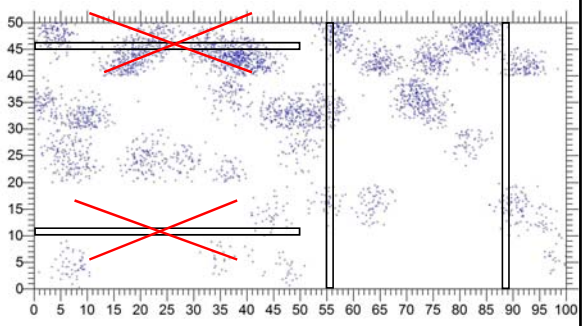
- Goal: greatest statistical precision for a given amount of resources
- Issues associated with quadrats:
  - Monitoring objective: parameter or pattern?
  - Logistics: minimize sample size or area?
  - Spatial distribution of individuals: minimize variation



A long narrow quadrat size performs much better



Quadrat orientation important when dealing with gradients



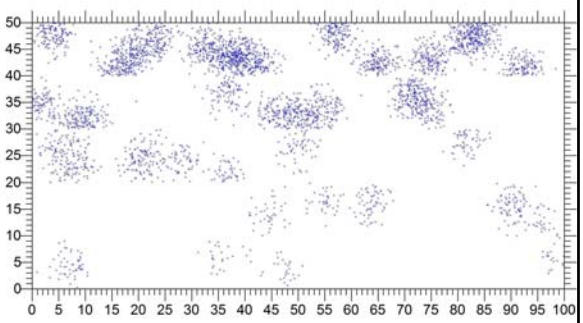
What is an appropriate S.U. size and shape?

- Goal: greatest statistical precision for a given amount of resources
- Issues associated with quadrats:
  - Monitoring objective: parameter or pattern?
  - Logistics: minimize sample size or area?
  - Spatial distribution of individuals: minimize variation
  - Edge effects
  - Density of the target population
  - Ease of sampling
  - Potential disturbance effects

Quadrat size & shape comparisons

- Long and skinny is better … but – how much better?
- Tested performance on actual & artificial populations
- Clumped-gradient population of 4000 plants

Clumped-gradient population of 4000 plants



Quadrat size & shape comparisons

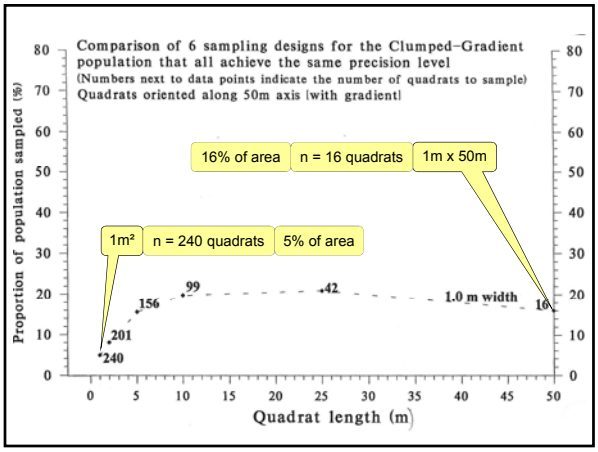
- Sampled population with 30 different quadrat length & width combinations

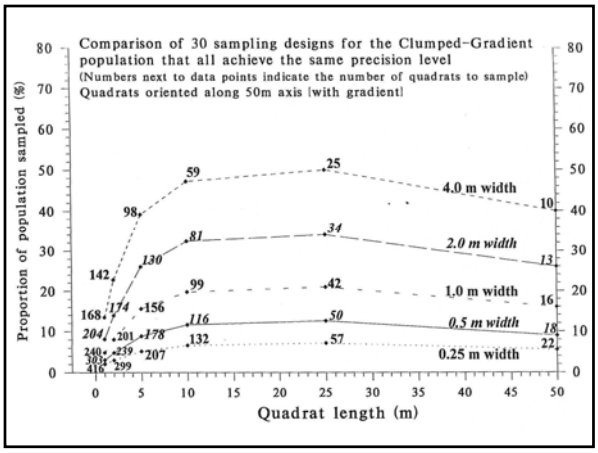
Widths	Lengths	Examples
0.25m	1m	0.25m x 1m
.5m	2m	0.25m x 2m
1m	5m	….
2m	10m	4m x 25m
4m	25m	4m x 50m
	50m	

Quadrat size & shape comparisions

- Calculated necessary sample size (and proportion of the population sampled) to achieve a targeted level of precision (95% CI ± 30%)?

Quadrat width	Quadrat length	Number of Quadrats	Proportion of pop. sampled
0.25m	1m	416	2.1%
1m	10m	99	19.8%
4m	50m	10	40%

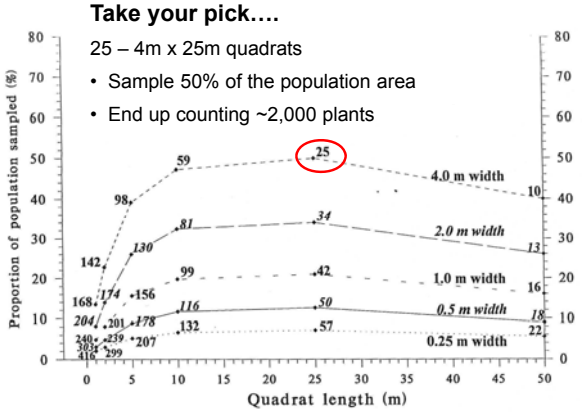




Take your pick….

25 – 4m x 25m quadrats

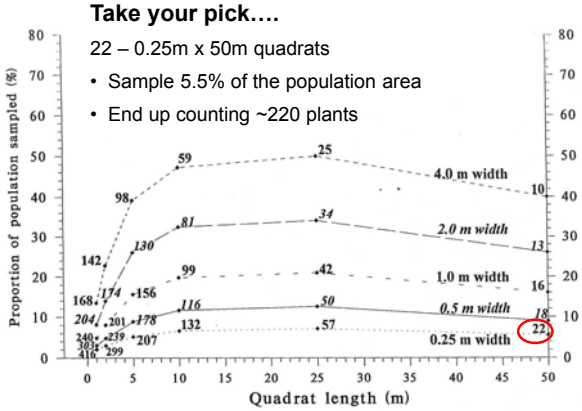
- Sample 50% of the population area
- End up counting ~2,000 plants



Take your pick….

22 – 0.25m x 50m quadrats

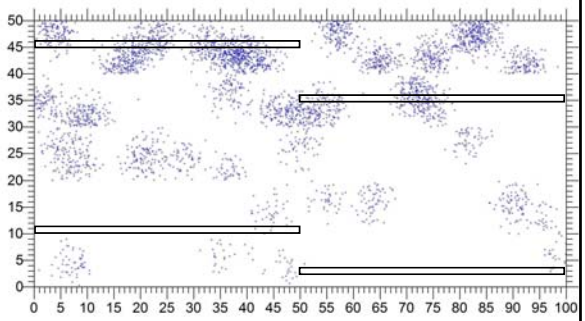
- Sample 5.5% of the population area
- End up counting ~220 plants

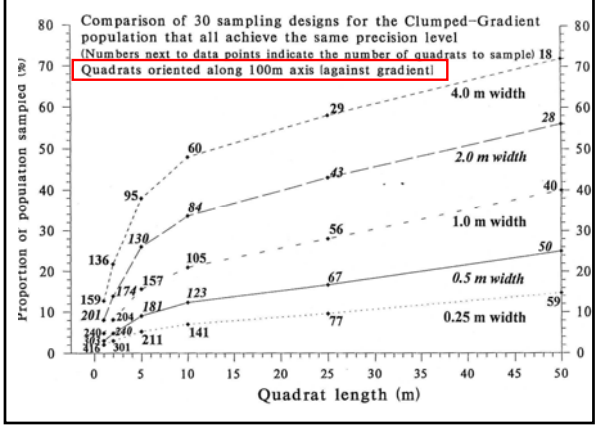


Webster's New World Dictionary

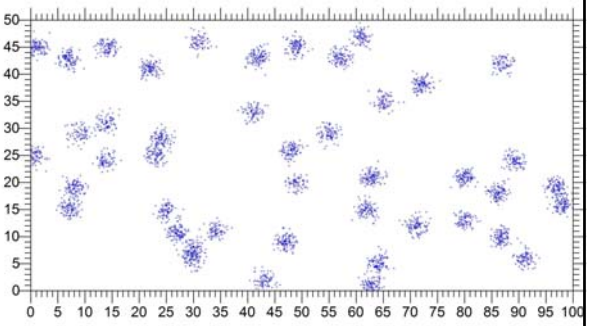
Quadrat – *Ecol.* a sampling plot, usually one square meter, used to study and analyze plant or animal life.

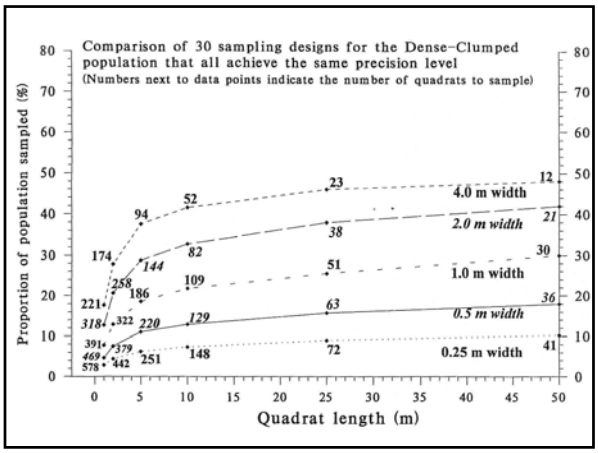
Consequence of quadrat orientation in relation to gradients

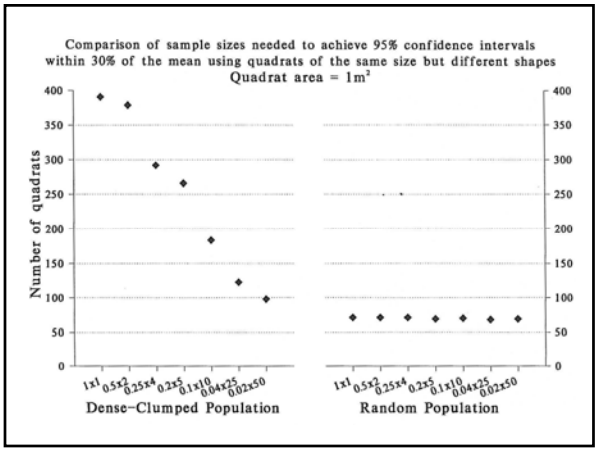




Dense-clumped population







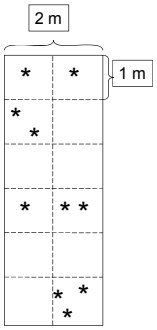
Site: _____ Date: _____ Page ____ of ____

Observer: _____ Total quadrat length: _____ Segment length: _____

Total quadrat width: _____ Width of left side: _____ Width of right side: _____

Plot #	Seg-ment #	Left	Right	Total
1	1	1	1	2
1	2	2	0	2
1	3	0	0	0
1	4	1	2	3
1	5	0	0	0
1	6	0	3	3

A sample field data sheet to use to collect pilot study data



Compare:

- 1m x 1m; 1m x 2m; 1m x 3m; 1m x 6m
- 2m x 1m; 2m x 2m; 2m x 3m; 2m x 6m

Use Coefficient of Variation (CV) to compare performance of different quadrat sizes and shapes

$$CV = \frac{\text{standard deviation}}{\text{mean}}$$

Ivesia aperta – pilot sampling

- Rare plant growing in Eastern Sierra Nevada floodplains
- Highly clumped distribution

- Sampling design
  - 0.25m x 100m quadrats

Sample statistics

- Mean = 65.67
- SD = 14.01
- CV = 0.21

STANDARDIZE THIS! THE FUTILITY OF ATTEMPTING TO APPLY A STANDARD QUADRAT SIZE AND SHAPE TO RARE PLANT MONITORING

Daniel W. Salzer

The Nature Conservancy of Oregon
821 Southeast 14th Avenue, Portland, Oregon 97214 USA
dsalzer@tnc.org (503) 230-0707

John W. Willoughby[1]

USDI Bureau of Land Management
2800 Cottage Way, Sacramento, California 95825 USA
John_Willoughby@ca.blm.gov (916) 978-4638

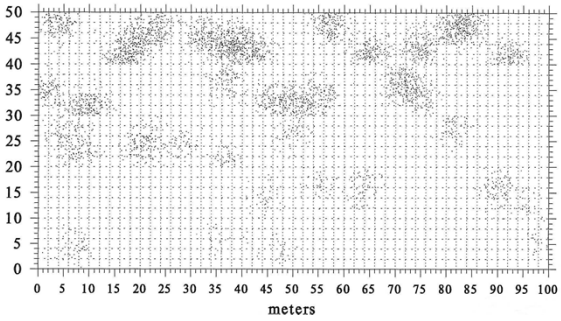
Abstract: Because of recommendations in some standard ecology texts, many botanists have attempted to employ 1 x 1 m quadrats to monitor changes in density or population size of herbaceous rare plant species. We will show that for most plant species, including common ones, the 1 x 1 m quadrat is one of the worst possible choices. For these plant species, long, narrow quadrats are much more efficient than square or circular quadrats for estimating density or population size, but the most efficient quadrat size and shape is species-specific. Because of this, agencies and organizations should resist the tendency to prescribe a standard quadrat size and shape to be applied to the monitoring of all rare plants. The quadrat size and shape employed in a particular monitoring study should be based on pilot sampling of the plant species that is the target of the study.

Sampling unit size & shape for other types of sampling

- Frequency – square quadrats work well

Frequency sampling – square quadrats are good!

There are 1250 grid cells in this population and 540 have ≥ 1 plant yielding a true percent frequency of 43.2%

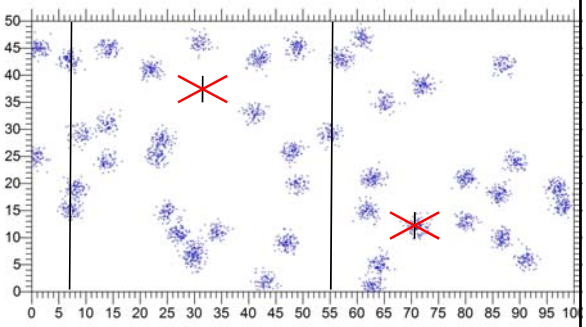


Sampling unit size & shape for other types of sampling

- Frequency – square quadrats work well
- % cover in quadrats – long and skinny is desirable but not practical
- Biomass in quadrats – long and skinny desirable
- Lines for line-intercept or point-intercept cover – make lines long enough to capture the spatial variability

Line-intercept or point-intercept cover along lines

Long lines will perform better than short lines



6 primary sampling design decisions

- What is the population of interest?
- What is an appropriate sampling unit (S.U.)?
- What is an appropriate S.U. size and shape?
- How should S.U.s be positioned in the population?
- Should S.U. positions be temporary or permanent?
- How many S.U.s should be sampled?

Take-home message…

- Use grid-cell method & systematic placement with random start

Positioning sampling units

- **2 main aims:**
  - Use random sampling
  - Achieve good interspersion of sampling units

Positioning sampling units

- Simple random sampling (with grid-cell)
- Stratified random sampling
- Systematic sampling
- Restricted random sampling
- Advanced designs
  - 2-stage sampling
  - Cluster sampling
  - Double sampling
  - Cluster sampling
- A simple random sample of individuals

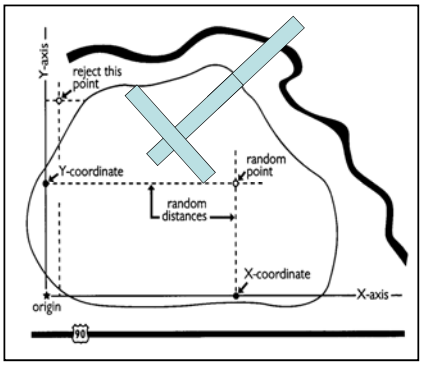
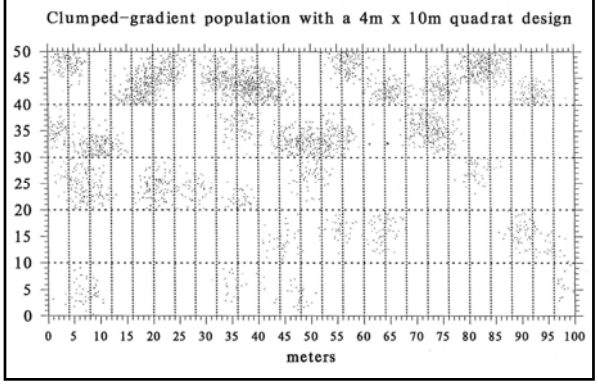
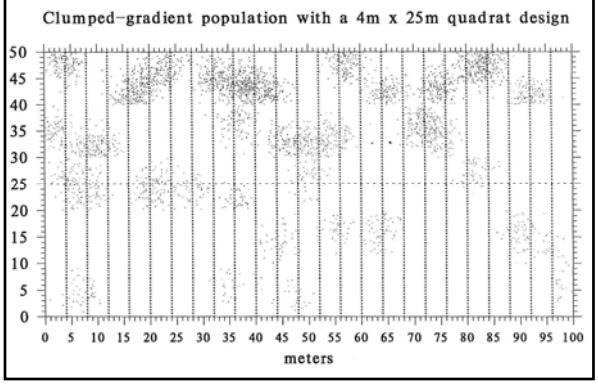


Figure 11. The simple random coordinate method. From Elzinga, et al. (1998).

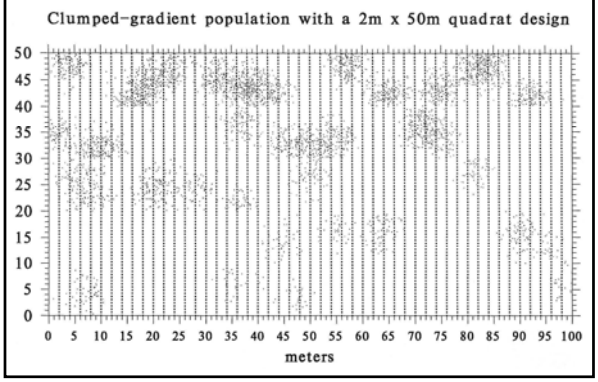
Grid-cell simple random sampling

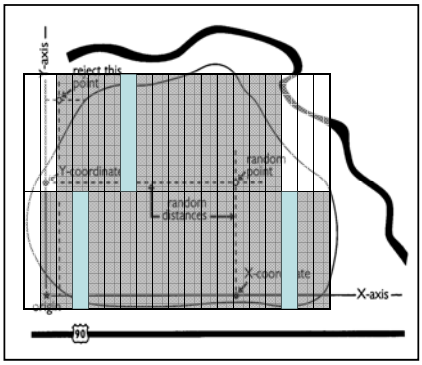


Grid-cell simple random sampling



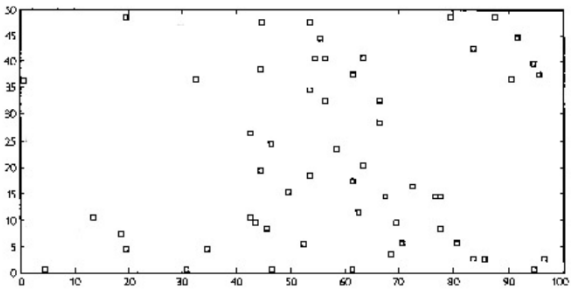
Grid-cell simple random sampling





Grid-cell approach applied to original example

Problems with good interspersion of sampling units with simple random sampling – 100 1m x 1m quadrats



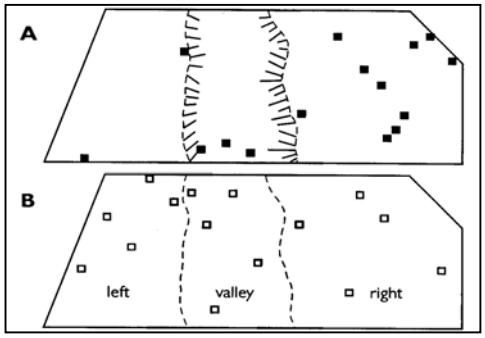
Simple Random Sampling

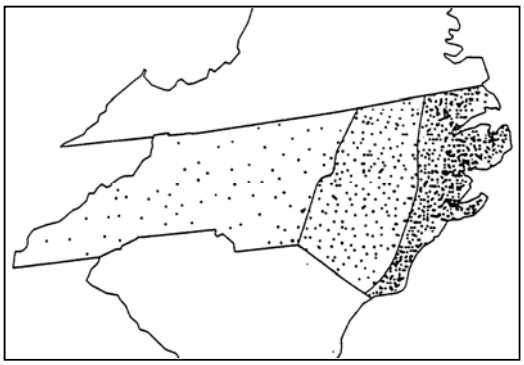
	Strengths		Weaknesses
+	Is conceptually the simplest type of random sampling	–	Use is limited to small geographic areas where a degree of homogeneity exists
+	Uses easier formulas to calculate means and standard errors than other types	–	When sample size is large, as with frequency or point intercept sampling, locating and traveling between quadrats can be time consuming
		–	By chance, some areas may not be sampled

Positioning sampling units

- Simple random sampling (with grid-cell)
- Stratified random sampling
- Systematic sampling
- Restricted random sampling
- Advanced designs
  - 2-stage sampling
  - Cluster sampling
  - Double sampling
  - Cluster sampling
- A simple random sample of individuals

Stratified Random Sampling (B)





A stratified random sampling scheme. From the National Wetlands Inventory, 1991.

Stratified Random Sampling

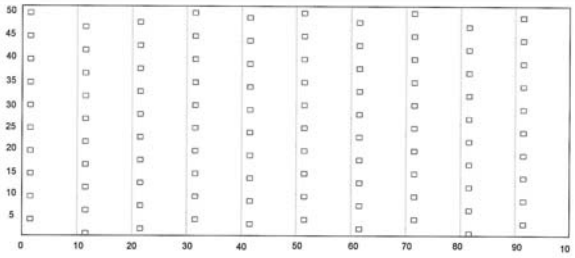
Strengths		Weaknesses	
+	Provides increased efficiency of population estimation when the attribute of interest responds differently to defined habitat features (strata)	–	Requires more complex formulas to calculate means and standard errors than simple random sampling
		–	Uses more complicated formulas to determine sample size allocation per stratum and estimate means and standard errors

Positioning sampling units

- Simple random sampling (with grid-cell)
- Stratified random sampling
- Systematic sampling
- Restricted random sampling
- Advanced designs
  - 2-stage sampling
  - Cluster sampling
  - Double sampling
  - Cluster sampling
- A simple random sample of individuals

Systematic sampling with a random start. Ten $1m^2$ quadrants are placed at 5m intervals along a 50m transect.

Systematic placement of quadrats/points along systematically positioned transects (good for frequency or point-intercept sampling)



What is the sampling unit? Each quadrat (n=100) or each transect (n=10)?

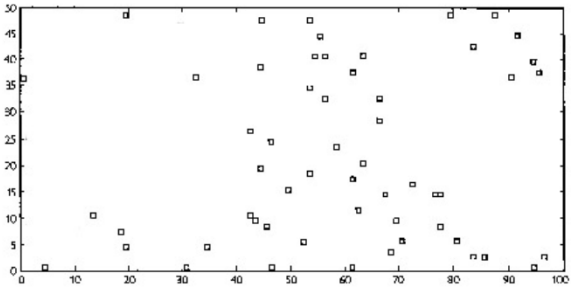
Could summarize either way

Usually better off treating each quadrat (or point) as the sampling unit

Rule of thumb to insure independence of quadrats/points:

Spacing of transects no more than 2X spacing of quadrats/points along transect

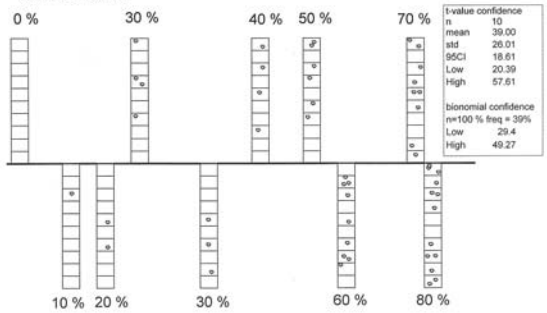
Simple random sampling – 100 1m x 1m quadrats



Frequency or Point-intercept Sampling: Binomial confidence intervals based on 95% confidence level and a sample size of 200 points

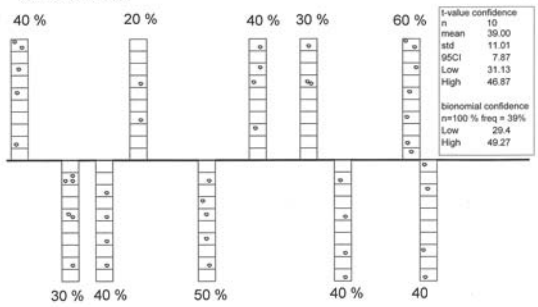
#	Percent Cover	lower (conf limit)	upper (conf limit)	length of lower C.I.	length of upper C.I.
1	0	0.00	1.83	0.00	1.83
2	5	2.42	9.00	2.58	4.00
3	10	6.22	15.02	3.78	5.02
4	15	10.35	20.72	4.65	5.72
5	20	14.69	26.22	5.31	6.22
6	25	19.16	31.60	5.84	6.60
7	30	23.74	36.87	6.26	6.87
8	35	28.41	42.05	6.59	7.05
9	40	33.15	47.15	6.85	7.15
10	45	37.98	52.18	7.02	7.18
11	50	42.87	57.13	7.13	7.13
12	55	47.82	62.02	7.18	7.02
13	60	52.85	66.85	7.15	6.85
14	65	57.95	71.59	7.05	6.59
15	70	63.13	76.26	6.87	6.26
16	75	68.40	80.84	6.60	5.84
17	80	73.78	85.31	6.22	5.31
18	85	79.28	89.65	5.72	4.65
19	90	84.98	93.78	5.02	3.78

Situation where variability among frequency values calculated for each transect and summarized with t-value confidence intervals yields a precision estimate **worse** than bionomial confidence intervals



What is the sampling unit in this case? 1m x 1m quadrat (n=100) or transect (n=10)

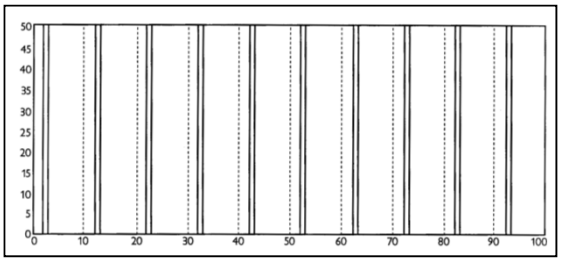
Situation where variability among frequency values calculated for each transect and summarized with t-value confidence intervals yields a precision estimate **better** than bionomial confidence intervals



Frequency or Point-intercept Sampling: Binomial confidence intervals based on 95% confidence level and a sample size of 200 points

#	Percent Cover	lower (conf limit)	upper (conf limit)	length of lower C.I.	length of upper C.I.
1	0	0.00	1.83	0.00	1.83
2	5	2.42	9.00	2.58	4.00
3	10	6.22	15.02	3.78	5.02
4	15	10.35	20.72	4.65	5.72
5	20	14.69	26.22	5.31	6.22
6	25	19.16	31.60	5.84	6.60
7	30	23.74	36.87	6.26	6.87
8	35	28.41	42.05	6.59	7.05
9	40	33.15	47.15	6.85	7.15
10	45	37.98	52.18	7.02	7.18
11	50	42.87	57.13	7.13	7.13
12	55	47.82	62.02	7.18	7.02
13	60	52.85	66.85	7.15	6.85
14	65	57.95	71.59	7.05	6.59
15	70	63.13	76.26	6.87	6.26
16	75	68.40	80.84	6.60	5.84
17	80	73.78	85.31	6.22	5.31
18	85	79.28	89.65	5.72	4.65
19	90	84.98	93.78	5.02	3.78

Systematic sampling that overly constrains possible samples



Ten 1m x 50m quadrats systematically positioned with a random starting point at the 3m position and quadrats positioned every 10m after that.

Since the position of all quadrats is fixed once the first quadrat is positioned, there are only 10 possible samples to choose from.

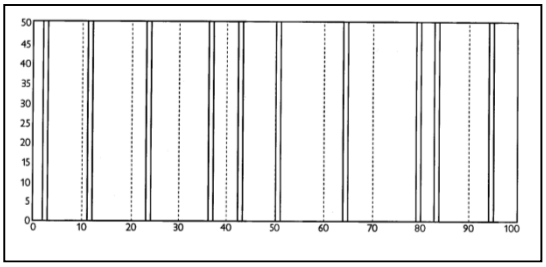
Systematic Sampling with a Random Start

Strengths		Weaknesses	
+	Typically faster to sample (reduced travel time between sampling units, fewer random coordinates)	–	Undesirable if sampling unit placement intersects some environmental pattern.
+	Results in more precise estimates than simple random sampling when sampling clumped distributions.	–	Sampling designs to estimate density can lead to questionable results.
+	Allows for even sampling across an area.		

Positioning sampling units

- Simple random sampling (with grid-cell)
- Stratified random sampling
- Systematic sampling
- Restricted random sampling
- Advanced designs
  - 2-stage sampling
  - Cluster sampling
  - Double sampling
- A simple random sample of individuals

Restricted random sampling



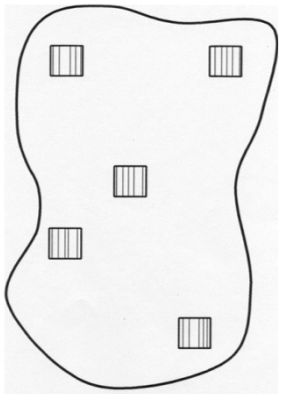
Restricted random sampling technique. Macroplot divided into 10 segments, a single quadrat randomly selected within each segment.

Restricted Random Sampling

	Strengths		Weaknesses
+	Results in good interspersion of sampling units throughout the target population.	–	Adds a bit more complexity and time to positioning sampling units (requires more random coordinates than systematic sampling)
+	Is more robust than systematic sampling because sampling designs don't severely constrain the number of possible samples		
+	Minimizes problem of intersecting repeating patterns in the environment		

Positioning sampling units

- Simple random sampling (with grid-cell)
- Stratified random sampling
- Systematic sampling
- Restricted random sampling
- Advanced designs
  - 2-stage sampling
  - Cluster sampling
  - Double sampling
- A simple random sample of individuals



2-Stage Design

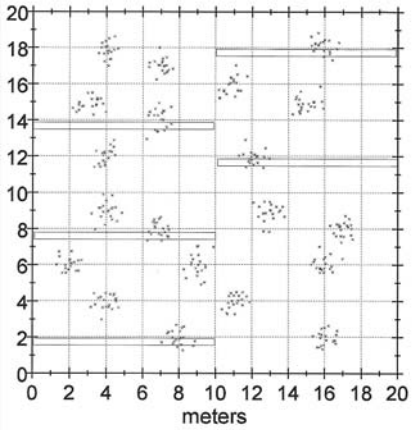
An example of two stage sampling for the line intercept method. Five macroplots are randomly positioned in the population. Five lines are then positioned within each macroplot.

Two-stage Sampling

Strengths		Weaknesses	
+	Reduced travel costs. It can be easier and less expensive to monitoring secondary units compared to the same number of sampling units placed using simple random sampling.	–	Requires more complex formulas to calculate means and standard errors than simple random sampling.
		–	May result in lower statistical precision.
		–	Uses more complicated formulas to determine sample size allocation for primary and secondary sampling units.

Positioning sampling units

- Simple random sampling (with grid-cell)
- Stratified random sampling
- Systematic sampling
- Restricted random sampling
- Advanced designs
  - 2-stage sampling
  - Cluster sampling
  - Double sampling
- A simple random sample of individuals



Cluster Sampling

An example of cluster sampling to estimate the mean height of plants in a population. Five clusters are randomly selected in the population and the heights of all plants within these quadrats are measured.

Cluster Sampling

	Strengths		Weaknesses
+	Costs less to sample elements in clusters than an equal number of elements selected at random.	–	Requires complex calculations to summarize and analyze data. Statistical software applications do not include these calculations.
+	Is most efficient when clusters are similar and together include much of the variability of the population.	–	Is less efficient than two stage sampling when each cluster contains large numbers of the target element.

Double-Sampling

- Rapid (less precise) measurement for large number of sampling units
- Detailed (more precise) measurement of a subset of all sampling units
- Examples:
  - Biomass measurements (clipped plots)
  - Hankin & Reeves riparian surveys

Double Sampling

	Strengths		Weaknesses
+	More efficient in estimating difficult to measure variables when the auxiliary variable is quick to measure and is highly correlated to the variable of interest.	–	Requires more complex formulas for data analysis and sample size determination than with simple random sampling.

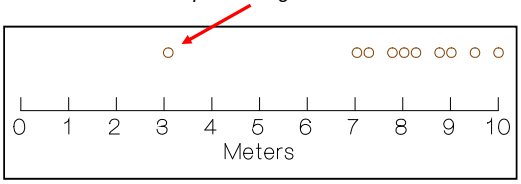
Positioning sampling units

- Simple random sampling (with grid-cell)
- Stratified random sampling
- Systematic sampling
- Restricted random sampling
- Advanced designs
  - 2-stage sampling
  - Cluster sampling
  - Double sampling
- A simple random sample of individuals

Common Approach - select random points then measure the nearest plant to that point

Introduces a bias towards isolated individuals

What is the probability of this plant being selected?



Randomly select a coordinate between 0-10 and then measure the nearest plant to that coordinate.

6 primary sampling design decisions

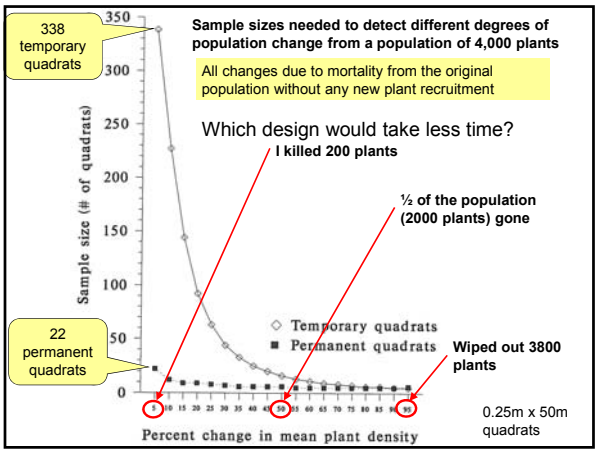
- What is the population of interest?
- What is an appropriate sampling unit (S.U.)?
- What is an appropriate S.U. size and shape?
- How should S.U.s be positioned in the population?
- Should S.U. positions be temporary or permanent?
- How many S.U.s should be sampled?

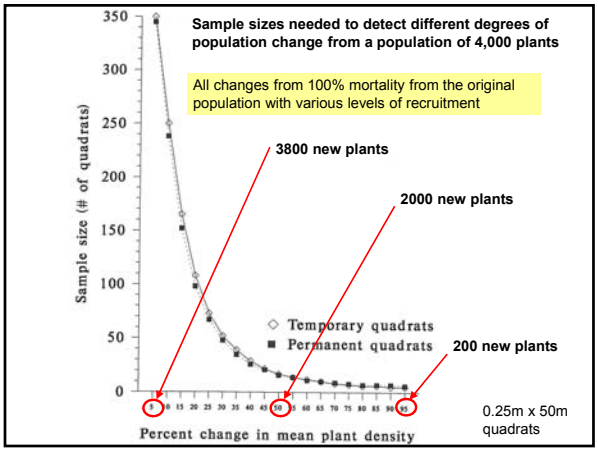
6 primary sampling design decisions

- What is the population of interest?
- What is an appropriate sampling unit (S.U.)?
- What is an appropriate S.U. size and shape?
- How should S.U.s be positioned in the population?
- Should S.U. positions be temporary or permanent?
- How many S.U.s should be sampled?

Permanent sampling units Advantages:

- Requires fewer sampling units for certain types of plants (e.g., long-lived perennials)
  - Sample size differences can be huge





Permanent sampling units Disadvantages:

- Cost (materials + installation)
- Susceptible to loss or damage
- Relocation difficulty
- Investigator impact
- Wildlife impacts

Tips for using permanent sampling units

- Use adequate marking system
  - Intermediate markers along long transects
  - If vandalism is risk – backup markers
  - GPS location
  - Photo documentation

Permanent sampling units – different sample size calculations

Quadrat number	# of plants/quadrat		Difference between 2009 & 2010
	2009	2010	
1	2	4	2
2	5	8	3
3	4	8	4
4	6	11	5
5	1	7	6
Mean of difference between 2009 and 2010			4
Standard deviation of difference between 2009 & 2010			1.58

Estimating the standard deviation of between year differences

$$s_{diff} = (s_1)\ (\sqrt{(2\ (1 - corr_{diff}))}\ )$$

Where:

s_{diff} = Estimated standard deviation of the differences between paired samples.

s_1 = Sample standard deviation among sampling units at the first time period.

$corr_{diff}$ = Correlation coefficient between sampling unit values in the first time period and sampling unit values in the second time period.

6 primary sampling design decisions

- What is the population of interest?
- What is an appropriate sampling unit (S.U.)?
- What is an appropriate S.U. size and shape?
- How should S.U.s be positioned in the population?
- Should S.U. positions be temporary or permanent?
- How many S.U.s should be sampled?

Determining how many sampling units to sample:

1. Should be driven by specific objectives:
   - How precise to you want your estimates to be?
   - What magnitude of change do you want to be able to detect?
   - What is the acceptable false-change error rate and desired Power?
2. Should be based on the amount of variability in the actual measurements (assess variability during pilot sampling)

Influence of sample size on level of precision

Standard formula for a confidence interval

$$\text{C.I.}_{\text{half-width}} = \text{SE}\left(\frac{s}{\sqrt{n}}\right) \times t_{\text{value}}$$

Influence of sample size on level of precision

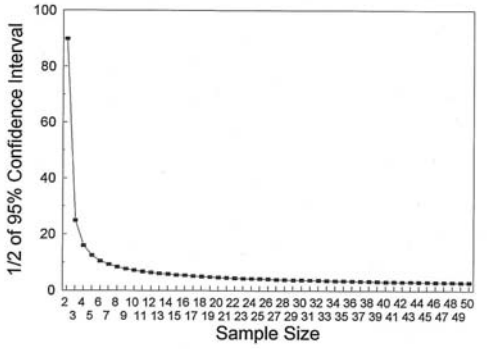
t-table

For confidence intervals, α = 1 - [confidence level] (e.g., use α=0.05 for 95% confidence intervals).
v = degrees of freedom (for confidence intervals, v = n-1)

v	α=0.20	α=0.10	α=0.05	α=0.01
1	3.078	6.314	12.706	63.657
2	1.886	2.920	4.303	9.925
3	1.638	2.353	3.182	5.841
4	1.533	2.132	2.776	4.604
5	1.476	2.015	2.571	4.032
6	1.440	1.943	2.447	3.707
7	1.415	1.895	2.365	3.499
8	1.397	1.860	2.306	3.355
9	1.383	1.833	2.262	3.250
10	1.372	1.812	2.228	3.169
11	1.363	1.796	2.201	3.106
12	1.356	1.782	2.179	3.055
13	1.350	1.771	2.160	3.012
14	1.345	1.761	2.145	2.977
15	1.341	1.753	2.131	2.947

Influence of sample size on level of precision

Using a t-distribution with a constant standard deviation (= 10)



Sample Size Equations

Equations for working with mean values (e.g., density, measurement data)

Target/Threshold — Equation #1: Determining the necessary sample size for estimating a **single population mean or a single population total** with a **confidence interval around the mean or total**.

Change/Trend — Equation #2: Determining the necessary sample size for detecting a difference between **two means** with **temporary sampling** units.

Change/Trend — Equation #3: Determining the necessary sample size for detecting a difference between **two means** with **permanent sampling** units.

Equations for working with proportion values (e.g., frequency, point-intercept data)

Target/Threshold — Equation #4: Determining the necessary sample size for estimating a **single population proportion** with a confidence interval around the proportion.

Change/Trend — Equation #5: Determining the necessary sample size for detecting a difference between **two proportions** with **temporary sampling** units.

Equation #1: Determining the necessary sample size for estimating a single population mean or a single population total with a confidence interval around the mean or total

This equation requires a 2 or 3 step process

STEP ONE: Calculate an initial sample size using the following equation:

$$n = \frac{(Z_\alpha)^2 (s)^2}{(B)^2}$$

n = The uncorrected sample size estimate.

Z_α = The standard normal coefficient from the table.

s = The standard deviation.

B = The desired precision level expressed as half of the maximum acceptable confidence interval width. This needs to be specified in absolute terms rather than as a percentage. For example, if you wanted your confidence interval half-width to be within 30% of your sample mean and your sample mean = 10 plants/quadrat then B = (0.30 x 10) = ± 3.0 (i.e., from 7-13).

EXAMPLE:
Management Objective: Restore the population of species Y in population Z to a density of at least 30 plants/quadrat by the year 2010.

Sampling Objective: Obtain estimates of the mean density and population size with **95% confidence intervals** that are **within 20%** of the estimated true value.

Results of pilot sampling:
Mean = 25 plants/quadrat.
Standard deviation (s) = 7 plants.

$$n = \frac{(Z_\alpha)^2 (s)^2}{(B)^2} \qquad n = \frac{(1.96)^2 (7)^2}{(B)^2}$$

Table of standard normal deviates (Z_α) for various confidence levels		
Confidence level	Alpha (α) level	(Z_α)
80%	0.20	1.28
90%	0.10	1.64
95%	**0.05**	**1.96**
99%	0.01	2.58

CALCULATE B:
B = The desired **confidence interval half-width** is 20% (0.20) of the estimated true value. Since the estimated true value is 25 plants/quadrat, the desired confidence interval half-width = (**B**) = 25 X 0.20 = **5** plants/quadrat.

$$n = \frac{(Z_\alpha)^2 (s)^2}{(B)^2} \rightarrow n = \frac{(1.96)^2 (7)^2}{(5)^2} = 7.5$$

Round 7.5 plots up to **8 plots** for the unadjusted sample size.

STEP 2. Adjust initial sample size with K-H table.

KUPPER & HAFNER CORRECTION TABLE:

Go to the sample size correction table on page 350 of the BLM TR and find n = 8 and the corresponding n* value in the 95% confidence level portion of the table. For n=8, the corresponding n* value = 15.

The corrected estimated sample size needed to be 95% confident that the estimate of the population mean is within 20% (± 5 plants) of the true mean = **15 quadrats.**

STEP 2 continued:

KUPPER & HAFNER CORRECTION TABLE

Table 1. Sample size correction

95% confidence level					
n	n*	n	n*	n	n*
1	5	51	66	101	121
2	7	52	67	102	122
3	8	53	68	103	123
4	10	54	69	104	124
5	11	55	70	105	125
6	12	56	71	106	126
7	14	57	72	107	128
8	**15**	58	74	108	129
9	16	59	75	109	130
10	18	60	76	110	131
11	19	61	77	111	132
12	20	62	78	112	133
13	21	63	79	113	134
14	23	64	80	114	135
15	24	65	81	115	136
16	25	66	83	116	137
17	26	67	84	117	138
18	28	68	85	118	139
19	29	69	86	119	141
20	30	70	87	120	142
21	31	71	88	121	143
22	32	72	89	122	144
23	34	73	90	123	145
24	35	74	91	124	146
25	36	75	92	125	147

STEP 3. Additional correction for sampling finite populations.

The above formula assumes that the population is very large compared to the proportion of the population that is sampled. **If you are sampling more than 5% of the whole population area then you should apply a correction to the sample size estimate that incorporates the finite population correction factor (FPC).**

$$n' = \frac{n^*}{(1 + (n^* / N))}$$

n' = The new sample size based upon inclusion of the finite population correction factor.
n* = The corrected sample size from the sample size correction table.
N = The total number of possible quadrat locations in the population. To calculate N, determine the total area of the population and divide by the size of the sampling unit.

If the pilot data described above was gathered using a 1m x 10m (10 m²) quadrat and the total population being sampled was located within a 20m x 50m macroplot (1000 m²)

N = 1000m²/10m² = 100
n* = 15

$$n' = \frac{15}{(1 + (15/100))} = 13.0$$

Equation #2: Determining the necessary sample size for detecting a difference between two means with temporary sampling units

This equation requires a 1 or 2 step process

STEP ONE: Calculate an initial sample size using the following equation:

$$n = \frac{2(s)^2 (Z_\alpha + Z_\beta)^2}{(MDC)^2}$$

s = Sample standard deviation.

$\mathbf{Z_\alpha}$ = Z-coefficient for the false-change (Type I) error rate from the table below.

$\mathbf{Z_\beta}$ = Z-coefficient for the missed-change (Type II) error rate from the table below.

MDC = Minimum detectable change size. This needs to be specified in absolute terms rather than as a percentage. For example, if you wanted to detect a 20% change in the sample mean from one year to the next and your first year sample mean = 10 plants/quadrat then:

MDC = (0.20 X 10) = 2 plants/quadrat.

Table of standard normal deviates for Z_α		Table of standard normal deviates for Z_β		
False-change (Type I) error rate (α)	Z_α	Missed-change (Type II) error rate (ß)	Power	Z_β
0.4	0.84	0.4	0.6	0.25
0.2	1.28	0.2	0.8	0.84
0.1	1.64	0.1	0.9	1.28
0.05	1.96	0.05	0.95	1.64
0.01	2.58	0.01	0.99	2.33

EXAMPLE: Management Objective: Increase the density of species F at Site Y by 20% between 2007 and 2012.

Sampling Objective: I want to be 90% certain of detecting a 20% change in mean plant density and I am willing to accept a 10% chance that I will make a false-change error (conclude that a change took place when it really did not).

Results from pilot sampling:

Mean = 25 plants/quadrat

Standard deviation = 7 plants.

$$n = \frac{2(s)^2(Z_\alpha + Z_\beta)^2}{(MDC)^2}$$

$$n = \frac{2(7)^2(1.64+1.28)^2}{(5)^2} = 33.4$$

Given:

The acceptable **False-change error rate (α) = 0.10** so the appropriate **Zα** from the table = **1.64**.

The desired **Power** is 90% (0.90) so the **Missed-change error rate (ß) = 0.10** and the appropriate **Zβ** coefficient from the table = **1.28**.

The **Minimum Detectable Change (MDC)** is 20% of the 2007 value or (0.20)x(25)= **5** plants/quadrat.

Round up 33.4 to 34 plots. The sample size correction table (Kupper & Hafner correction) **is not** needed for estimating sample sizes for detecting differences between two population means.

Correction for sampling finite populations

If you are sampling more than 5% of the whole population area then you should apply a correction to the sample size estimate that incorporates the finite population correction factor (FPC).

$$n' = \frac{n}{(1+(n/N))}$$

EXAMPLE:

If the pilot data described above was gathered using a 1m x 10m (10 m²) quadrat and the total population being sampled was located within a 20m x 50m macroplot (1000 m²) then **N = 1000m²/10m² = 100.**

$$n' = \frac{34}{(1+(34/100))} = 25.3$$

Round up 25.3 to 26.

The new, FPC-corrected estimated sample size needed to be 90% certain of detecting a change of 5 plants between 2007 and 2012 with a false-change error rate of 0.10 **= 26 quadrats**.

Sample size equation #4: Determining the necessary sample size for estimating a single population proportion with a specified level of precision.

The equation for determining the sample size for estimating a single proportion is:

$$n = \frac{(Z_\alpha)^2(p)(q)}{d^2}$$

Where:

n = Estimated necessary sample size.

Z_α = The coefficient from the table of standard normal deviates below.

p = The value of the proportion as a decimal percent (e.g., 0.45). If you don't have an estimate of the current proportion, use 0.50 as a conservative estimate.

q = 1 - p.

d = The desired precision level expressed as half of the maximum acceptable confidence interval width. This is also expressed as a decimal percent (e.g., 0.15) and this represents an *absolute* rather than a *relative* value. For example, if your proportion value is 30% and you want a precision level of ±10% this means you are targeting an interval width from 20% to 40%. Use 0.10 for the d-value and *not* 0.30 x 0.10 = 0.03.

Table of standard normal deviates (Z_α) for various confidence levels

Confidence level	Alpha (α) level	(Z_α)
80%	0.20	1.28
90%	0.10	1.64
95%	0.05	1.96
99%	0.01	2.58

Example:
Management objective:
Maintain at least a 40% frequency (in $1m^2$ quadrats) of species Y in population Z over the next 5 years.

Sampling objective:
Estimate percent frequency with 95% confidence intervals no wider than ± 10% of the estimated true value.

Results of pilot sampling:
The proportion of quadrats with species Z is estimated to be p = 65% (0.65).
Because q = (1-p), q = (1-0.65) = 0.35.

Given:
The desired confidence level is 95% so the appropriate Z_α from the table above = 1.96.
The desired confidence interval width (d) is specified as 10% (0.10).

Using the equation provided above:

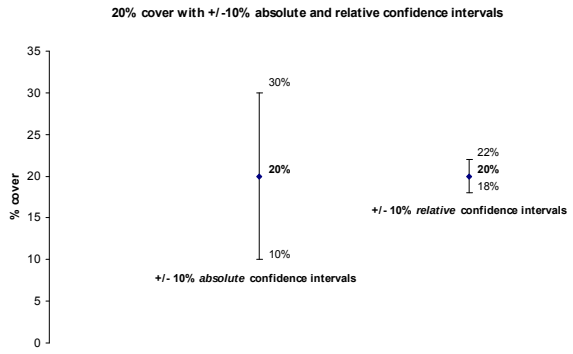
$$n = \frac{(Z_\alpha)^2(p)(q)}{d^2} \qquad n = \frac{(1.96)^2(0.65)(0.35)}{0.10^2} = 87.4$$

Round up 87.4 to 88.

The estimated sample size needed to be 95% confident that the estimate of the population percent frequency is within 10% (+/- 0.10) of the true percent frequency = 88 quadrats.

For proportion data (frequency, point-intercept cover)

- Be explicit about MDC and/or confidence interval width
- State desired confidence interval width & minimum detectable changes as **absolute** rather than **relative values**
- For example
  - Starting percent frequency = 20%
  - MDC = 10% (want to detect change from 20% to 10% or 20% to 30%) – NOT from 20% to 18%



Point Intercept Example

20% relative change

1st year	Abs MDC	2nd Year	sample size
100%	20%	80%	47
90%	18%	72%	91
80%	16%	64%	146
70%	14%	56%	216
60%	12%	48%	311
50%	***10%***	***40%***	***442***
40%	8%	32%	640
30%	6%	24%	970
20%	4%	16%	1629
10%	2%	8%	3601

10% absolute change

1st year	Abs MDC	2nd Year	sample size
100%	10%	90%	99
90%	10%	80%	237
80%	10%	70%	340
70%	10%	60%	408
60%	10%	50%	442
50%	***10%***	***40%***	***442***
40%	10%	30%	408
30%	10%	20%	340
20%	10%	10%	237
10%	10%	0%	99

Sample Size Equations

Equations for working with mean values (e.g., density, measurement data)

Target/Threshold

Equation #1: Determining the necessary sample size for estimating a **single population mean or a single population total** with a **confidence interval around the mean or total**.

Change/Trend

Equation #2: Determining the necessary sample size for detecting a difference between **two means** with **temporary sampling** units.

Equation #3: Determining the necessary sample size for detecting a difference between **two means** with **permanent sampling** units.

Equations for working with proportion values (e.g., frequency, point-intercept data)

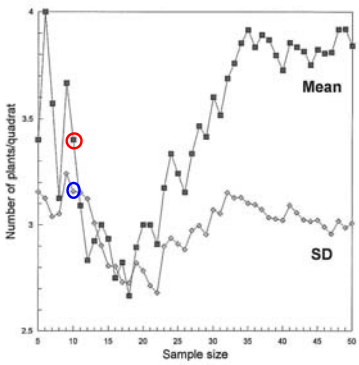
Target/Threshold

Equation #4: Determining the necessary sample size for estimating a **single population proportion** with a confidence interval around the proportion.

Change/Trend

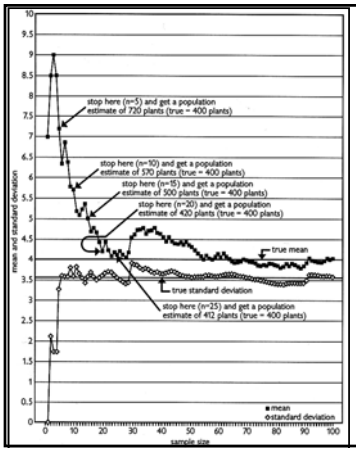
Equation #5: Determining the necessary sample size for detecting a difference between **two proportions** with **temporary sampling** units.

Sequential sampling graphs

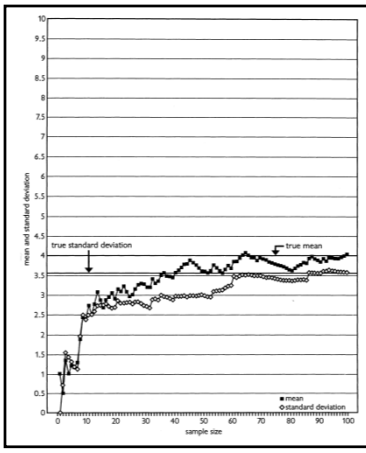


Running means and standard deviations

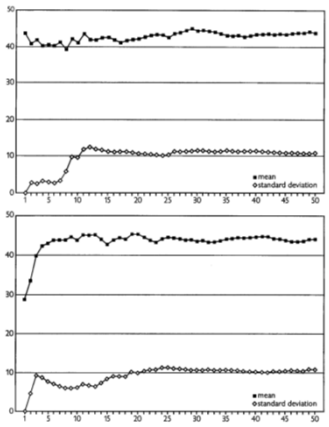
n	Plants	Mean	SD	n	Plants	Mean	SD
1	1	1.00	0.00	26	1	3.15	2.88
2	2	1.50	0.50	27	8	3.33	2.97
3	0	1.00	0.79	28	7	3.46	3.00
4	9	3.00	3.52	29	2	3.41	2.95
5	5	3.40	3.15	30	9	3.60	3.07
6	7	4.00	3.12	31	1	3.52	3.05
7	1	3.57	3.04	32	9	3.69	3.15
8	0	3.13	3.05	33	6	3.76	3.13
9	8	3.67	3.24	34	7	3.85	3.13
10	1	3.40	3.16	35	6	3.91	3.10
11	0	3.09	3.15	36	1	3.83	3.09
12	0	2.83	3.12	37	6	3.89	3.07
13	4	2.92	3.01	38	3	3.87	3.03
14	4	3.00	2.90	39	1	3.79	3.03
15	2	2.93	2.81	40	1	3.73	3.02
16	0	2.75	2.80	41	9	3.85	3.09
17	4	2.82	2.73	42	3	3.83	3.06
18	0	2.67	2.73	43	3	3.81	3.02
19	7	2.89	2.82	44	1	3.75	3.02
20	5	3.00	2.78	45	7	3.82	3.02
21	3	3.00	2.71	46	3	3.80	2.99
22	1	2.91	2.68	47	4	3.81	2.96
23	9	3.17	2.90	48	9	3.92	3.02
24	7	3.33	2.94	49	4	3.92	2.99
25	1	3.24	2.91	50	0	3.84	3.00



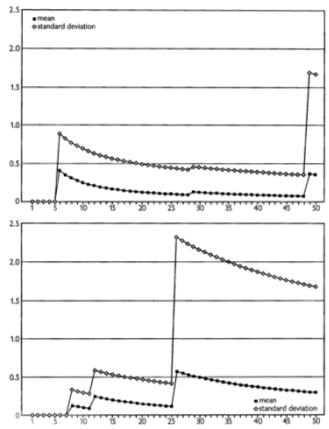
Sequential sampling graph for the 400 plant population described in the Introduction to Sampling session. Quadrat size = 0.4m x 10m = 100 (the entire population).



Sequential sampling graph for the 400 plant population described in the Introduction to Sampling session. Quadrat size = 0.4m x 10m, n + 100 (the entire population).



Sequential sampling graph for vegetation height measurements at Mt. Hebo.



Sequential sampling graph for bracken fern stem density at Mt. Hebo.

Sampling Design - Objectives

- List the six primary decisions common to all sampling designs
- List at least three issues affecting sampling unit size and shape
- Randomly position sample units
- Solve 4 sampling design problems
- Discuss advantages and disadvantages of permanent and temporary sample units
- Use sample size equations for:
  - Obtaining an estimate of a population mean
  - Detecting a change over two time periods

6 primary sampling design decisions

- What is the population of interest?
- What is an appropriate sampling unit (S.U.)?
- What is an appropriate S.U. size and shape?
- How should S.U.s be positioned in the population?
- Should S.U. positions be temporary or permanent?
- How many S.U.s should be sampled?