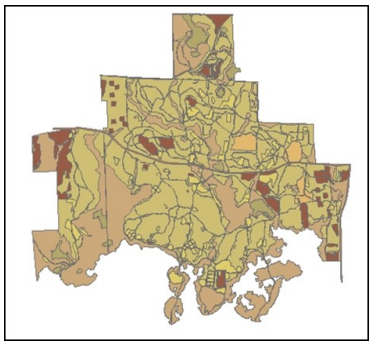
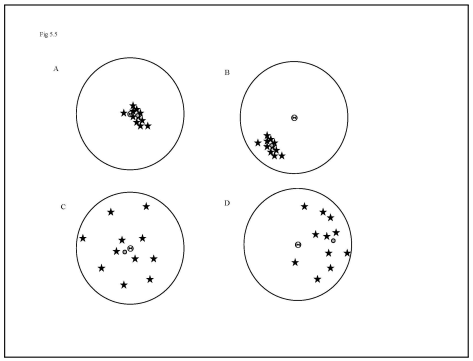
Sampling and Estimation





Model parameters

Key concepts

- All models include parameters (e.g., N, r, K, etc. . .)
- Parameters are almost never known. Why?
  - We usually have to sample
  - Animals are hard to detect
- Good sampling designs yield accurate estimates of unknown parameters

Parameters, estimates, and statistics

Parameter
A characteristic of a population

Statistic
A characteristic of a dataset (often used to estimate a parameter)

	Population parameter	Parameter estimate
Population size	N_t	$\hat{N}_t$
Growth rate	r	$\hat{r}$
Occurrence probability	ψ	$\hat{\psi}$

How do you get accurate estimates?

Properties of a good design:

(1) Clearly defined objective, in terms of:
- Parameter(s) that will be estimated
- Population of interest
- Criteria for reliability (e.g., precision)
- Practical constraints (e.g., costs)

(2) Replication

(3) Randomization

(4) Controls (when conducting an experiment)

Defining the population

Target population
The population of interest. Should be defined in terms of time and space.

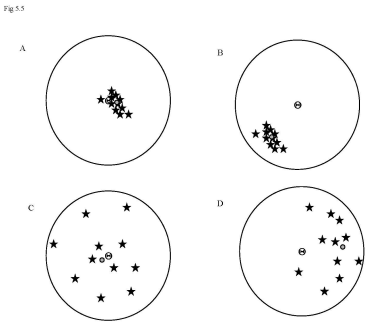
Sampled population
The sampled portion of the population of interest, usually defined in terms of the sample units (such as plots, quadrats, etc.).

Accuracy

Accuracy has two components:

(1) **Bias**: The difference between the average estimate and the true parameter

(2) **Variance**: The variability of the estimates.

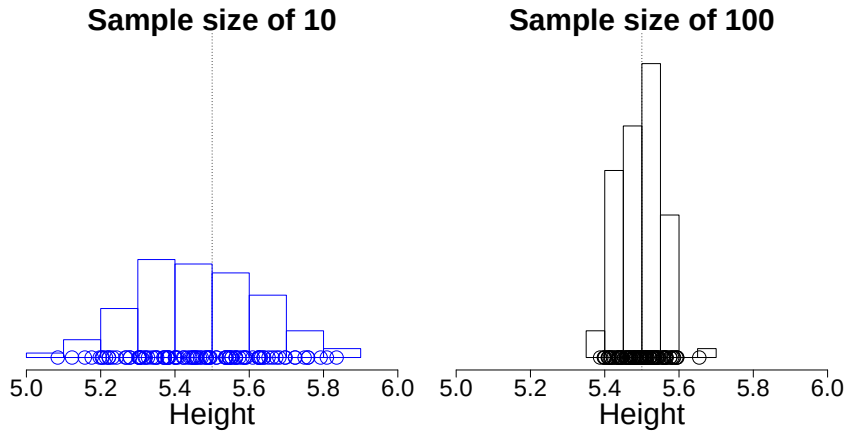


How do we reduce variance?

Huge sample size

Effect of sample size

Suppose we want to estimate the height of students on campus, and we have enough resources to repeat a survey many times. Each point below is an estimate.



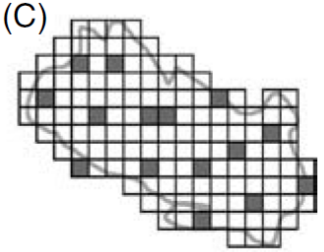
The standard deviation of the sampling distribution is called the standard error (SE)

Randomization

Simple random sampling

- All sample units have the same inclusion probability
- Easiest and most reliable method
- But not always cost effective

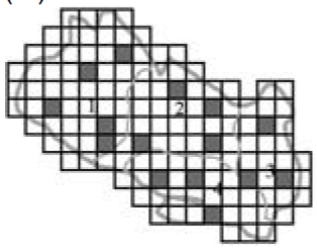
(C)



Stratified random sampling

- Useful when study area is characterized by several homogeneous regions
- Regions with higher variability should be sampled more intensively than regions with low variability
- Often more cost effective than simple random sampling

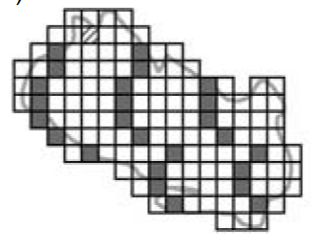
(D)



Systematic sampling

- Sample units are selected according to a regular, ordered scheme with the first unit being sampled randomly.
- Easy to implement in the field
- Potentially dangerous because sample unit spacing could coincide with natural spacing of environmental features

(E)



Main points

- We have to estimate model parameters
- Reliable estimates require good sampling design
- Replication reduces variance
- Randomization reduces bias