Chapter 7 - Sampling and Sampling Distributions

Chapter 7 - Sampling

An element is the entity on which data are collected.

A population is the set of all the elements of interest.

A sample is a subset of the population.

A frame is a list of the elements that the sample will be selected from.

The reason we select a sample is to collect data to answer a research question about a population.

Slide 1

Introduction

The sample results provide only estimates of the values of the population characteristics.

The reason is simply that the sample contains only a portion of the population.

With proper sampling methods, the sample results can provide "good" estimates of the population characteristics.

This chapter provides the basis for determining how large the estimation error might be.

Slide 2

Sampling from a Finite Population

- Finite populations are often defined by lists such as:
  - Organization membership roster
  - Credit card account numbers
  - Inventory product numbers
- A simple random sample of size n from a finite population of size N is a sample selected such that each possible sample of size n has the same probability of being selected.
- Replacing each sampled element before selecting subsequent elements is called sampling with replacement.
- Sampling without replacement is the procedure used most often.
- In large sampling projects, computer-generated random numbers are often used to automate the sample selection process.

Slide 3

Sampling from a Finite Population Using Excel

- Example: St. Andrew's College

St. Andrew's College received 900 applications for admission in the upcoming year from prospective students. The applicants were numbered, from 1 to 900, as their applications arrived. The Director of Admissions would like to select a simple random sample of 30 applicants.

Step 1: Assign a random number to each of the 900 applicants.

The random numbers generated by Excel's *RAND* function follow a uniform probability distribution between 0 and 1.

Step 2: Select the 30 applicants corresponding to the 30 smallest random numbers.

Slide 4

Sampling from a Finite Population Using Excel

- Excel Value Worksheet
- Use Rand to assign random number to obervation

	A	B
1	Applicant Number	Random Number
2	1	=RAND()
3	2	=RAND()
4	3	=RAND()
5	4	=RAND()
6	5	=RAND()
7	6	=RAND()
8	7	=RAND()
9	8	=RAND()

	A	B
1	Applicant Number	Random Number
2	1	0.61021
3	2	0.83762
4	3	0.58935
5	4	0.19934
6	5	0.86658
7	6	0.60579
8	7	0.80960
9	8	0.33224

Note: Rows 10-901 are not shown.

Slide 5

Sampling from a Finite Population Using Excel

- Put Random Numbers in Ascending Order
- Set calculation option to manual

Step 1 Select any cell in the range B2:B901

Step 2 Click the **Home** tab on the Ribbon

Step 3 In the **Editing** group, click **Sort & Filter**

Step 4 Choose **Sort Smallest to Largest**

	A	B
1	Applicant Number	Random Number
2	12	0.00027
3	773	0.00192
4	408	0.00303
5	58	0.00481
6	116	0.00538
7	185	0.00583
8	510	0.00649
9	394	0.00667

Slide 6

Sampling from a Process

- Populations are often defined by an ongoing process whereby the elements of the population consist of items generated as though the process would operate indefinitely.
- Some examples of on-going processes, with infinite populations, are:
  - parts being manufactured on a production line
  - transactions occurring at a bank
  - telephone calls arriving at a technical help desk
  - customers entering a store

Slide 7

Sampling from a Process

- In the case of infinite populations, it is impossible to obtain a list of all elements in the population.
- The sampled population is such that a frame cannot be constructed.
- The random number selection procedure cannot be used for infinite populations.
- A random sample from an infinite population is a sample selected such that the following conditions are satisfied.
  - Each of the sampled elements is independent.
  - Each of the sampled elements follows the same probability distribution as the elements in the population.

Slide 8