Chapter 11 Sampling and the Sampling Distribution of the Means

I. **Inferential statistics uses sample statistics to estimate population parameters. This chapter will explore how a sample mean ($\bar{x}$) is used to predict its population mean (μ).**

II. **Why use sample data**
- A. The cost of a census is prohibitive.
- B. The time required to take a census is not available.
- C. Measuring a parameter destroys the item being tested (measuring the mean lifetime in hours of light bulbs).
- D. A sample will yield adequate results.

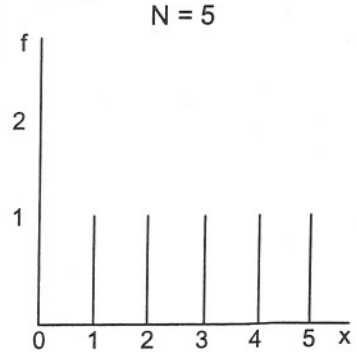
III. **Probability samples**
- A. A probability sample is one in which the likelihood of an item being chosen is known.
- B. Probability sampling methods
  1. **Simple random samples**
     - a. Each population member has an equal chance of being chosen.
     - b. Put an identification (name, serial number, etc.) into a hat, mix, and select.
     - c. A table of random digits or a computer program often replace the hat.
     - d. To sample 30 out of 485 students using their ID numbers from 1 to 485:
       1) Arbitrarily choose a starting point on a table of random digits.
       2) Working in some direction (horizontally, vertically, or diagonally), and using the first or last three digits, choose 30 student numbers ignoring those over 485.

Partial Table of Random Digits			
1318	7677	9619	2786
2122	8297	1190	1379
0037	6355	4717	5184
4788	9044	5583	0292

  2. **Systematic random samples**
     - a. Use every nth item beginning at some random point on a list of population members.
     - b. This method could be biased because population members at the beginning of a list (Mr. Abbot or employee 0001) and end of a list (Ms. Zona or employee 9999) might not have an equal chance of being chosen.
  3. **Stratified random samples**
     - a. Divide population into homogeneous subgroups and sample each subgroup.
     - b. This type of sample can be more representative than a simple random sample because a small diverse section of a population might not be chosen with a simple random sample.
- C. Sampling and nonsampling error
  1. **Sampling error** exists because a nonrepresentative sample was used in place of a census.
  2. **Nonsampling error**, which occurs with any survey, exists primarily because of poor collection techniques. High nonsampling error can make a census less accurate than a sample. Why? Limited funds and having to survey all population members cause poor collection techniques and high nonsampling error.

IV. **Sampling distribution of the means**
- A. The sampling distribution of the means consists of all the possible sample means of size n that may be drawn from a population size N. It is important. Taking one sample is really taking one out of many possible samples. The sampling distribution is the key to why accurate predictions can be made with inferential statistics.
- B. Population members A,B,C,D, and E have 1,2,3,4, and 5 children respectively. The sampling distribution of the means, its mean and standard deviation, for a sample of 3 out of 5 has been calculated and demonstrated below.

Population Distribution
N = 5



Number of Children

$$\mu = \frac{\sum x}{N} = \frac{1+2+3+4+5}{5} = 3$$

σ given at = 1.414

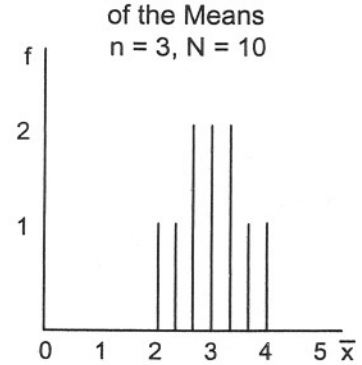
10 possible samples result from a sample of 3 out of 5

	x	$\bar{x}$
ABC	1,2,3	2.00
ABD	1,2,4	2.33
ABE	1,2,5	2.67
ACD	1,3,4	2.67
ACE	1,3,5	3.00
ADE	1,4,5	3.33
BCD	2,3,4	3.00
BCE	2,3,5	3.33
BDE	2,4,5	3.67
CDE	3,4,5	4.00
		$\Sigma \bar{x}$ = 30.00

Sampling Distribution of the Means

$\bar{x}$	f
2.00	1
2.33	1
2.67	2
3.00	2
3.33	2
3.67	1
4.00	1

Sampling Distribution of the Means
n = 3, N = 10



$$\mu_{\bar{x}} = \frac{\sum \bar{x}}{N} = \frac{30}{10} = 3$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.414}{\sqrt{3}} = .82$$