**Section 1.3 Simple Random Sampling**

**Objective**

* 1. Obtain a Simple Random Sample

Introduction, Page 1

Observational studies can be conducted by administering a survey. When administering a survey, the researcher must first identify the population that is to be targeted.

1. Define: Random sampling **Random sampling** is the process of using chance to select individuals from a population to be included in the sample.

For the results of a survey to be reliable, the characteristics of the individuals in the sample must be representative of the characteristics of the individuals in the population. The key to obtaining a sample representative of a population is to let *chance* or *randomness,* rather than convenience, play a role in dictating which individuals are in the sample. **If convenience is used to obtain a sample, the results of the survey are meaningless.**

For the results of a survey to be reliable, the characteristics of the individuals in the sample must be representative of the characteristics of the individuals in the population.

The key to obtaining a sample representative of a population is to let chance or randomness play a role in dictating which individuals are in the sample, rather than convenience.

If convenience is used to obtain a sample, the results of the survey are meaningless.

Introduction, Page 2

2.Why are the survey results from the sample taken outside Fenway Park not likely to be reliable?  Clearly, the individuals in the sample do not accurately reflect the makeup of the entire population, CONVENIENCE.

3.Why are the results of a survey of students in your statistics class likely to be misleading when trying to determine what proportion of students on your campus work? The convenient sample is not representative of the population, which means that any results reported from your survey are misleading. Does your class mirror the gender, grade, or day/ evening of the entire campus, probably not.

Introduction, Page 3

List the four basic sampling techniques.

1. Simple random sampling
2. Stratified sampling
3. Systematic sampling
4. Cluster sampling

***Objective 1: Obtain a Simple Random Sample***

Objective 1, Page 1

1. What is a simple random sample? A sample of size n from a population of size N is obtained through **simple random sampling** if every possible sample of size n has an equal chance of occurring. The sample is then called a **simple random sample.**

The number of individuals in the sample is always less than the number of individuals in the population.

Objective 1, Page 2

**Example 1** ***Illustrating Simple Random Sampling***

Sophie has four tickets to a concert. Six of her friends, Yolanda, Michael, Kevin, Marissa, Annie, and Katie, have all expressed an interest in going to the concert. Sophie decides to randomly select three of her six friends to attend the concert.

1. List all possible samples of size *n* = 3 from the population of size *N* = 6. Once an individual is chosen, he/she cannot be chosen again. 20 total
2. Comment on the likelihood of the sample containing Michael, Kevin, and Marissa. 1 in 20

Objective 1, Page 5

How do we select the individuals in a simple random sample?

Typically, each individual in the population is assigned a unique number between 1 and *N*, where *N* is the size of the population. Then *n* distinct random numbers are selected, where *n* is the size of the sample. To number the individuals in the population, we need a frame**–** a list of all the individuals within the population. ***How do we select the individuals in a simple random sample?***

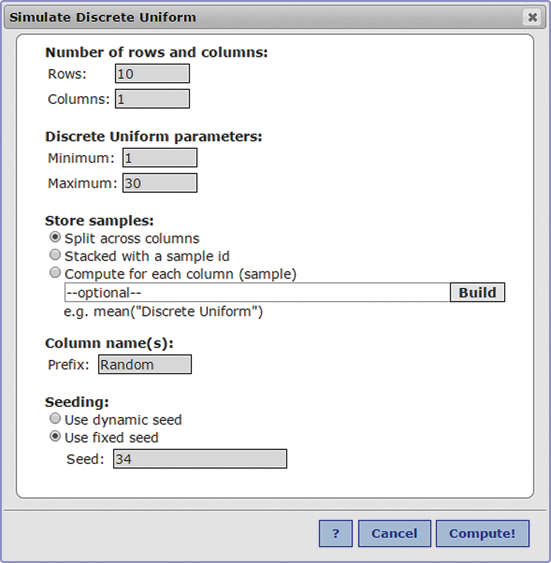
We could write the names of the individuals in the population on different pieces of paper and then select names from a hat. Often, however, the size of the population is so large that performing simple random sampling in this fashion is not practical.

Typically, each individual in the population is assigned a unique number between 1 and N, where N is the size of the population. Then n distinct random numbers are selected, where n is the size of the sample.

**To number the individuals in the population, we need a frame—a list of all the individuals within the population.**

#### **Obtaining a Simple Random Sample**

1. Select **Data**, highlight **Simulate**, and then highlight **Discrete Uniform**.
2. Fill in the window with the appropriate values. To obtain a simple random sample of size n=5 from a population of size N=30, enter the values shown in the figure. The reason we generate 10 rows of data (instead of 5) is in case any of the random numbers repeat. Click Compute! and the random numbers will appear in the spreadsheet. *Note*: You could also select the single dynamic seed radio button to set the seed.



Objective 1, Page 6

*Answer the following after watching the animation.*

1. What is the frame in this animation? Students in class

Explain why a second sample of 5 students will most likely be different than the first sample of 5 students? **The individuals differ from sample to sample**

**because chance is used to select the individuals.**

Explain why inferences based on samples vary. **Inferences based on samples will vary**

**because the individuals in the samples vary.**

Objective 1, Page 8

**Example 2 *Obtaining a Simple Random Sample***

The accounting firm of Senese and Associates has grown. To make sure their clients are still satisfied with the services they are receiving, the company decides to send a survey out to a simple random sample of 5 of its 30 clients.

**TABLE 3**

* + 1. ABC Electric
    2. Brassil Construction
    3. Bridal Zone
    4. Casey's Glass House
    5. Chicago Locksmith
    6. DeSoto Painting
    7. Dino Jump
    8. Euro Car Care
    9. Farrell's Antiques
    10. First Fifth Bank
    11. Fox Studios
    12. Haynes Hauling
    13. House of Hair
    14. John's Bakery
    15. Logistics Management, Inc.
    16. Lucky Larry's Bistro
    17. Moe's Exterminating
    18. Nick's Tavern
    19. Orion Bowling
    20. Precise Plumbing
    21. R&Q Realty
    22. Ritter Engineering
    23. Simplex Forms
    24. Spruce Landscaping
    25. Thors, Robert DPS
    26. Travel Zone
    27. Ultimate Electric
    28. Venetian Gardens Restaurant
    29. Walker Insurance

Worldwide Wire

Suppose Professor Cummings wishes

to estimate the average travel time to school for his class.

Rather than surveying each of the 33 students enrolled,

he decides to obtain a simple random sample of 5 students.

Professor Cummings already has a frame.

Remember, a "frame" is a list of all the individuals

in the population of interest.

In this case, the frame is the list

of students enrolled in the class.

To obtain a simple random sample,

Professor Cummings first assigns a unique number

to each student in the class.

So the first student in the class

is assigned number 1, the second student is assigned the number

2, and so on until we reach the last student who

is assigned the number 33.

5 different numbers will be randomly selected.

The students corresponding to these numbers

are the individuals in the sample.

This is sampling without replacement,

which means that an individual who is selected to be

in the sample from the population cannot be selected

again.

On the screen, we have our 33 students--

numbered 1 through 33.

To get the random numbers used to generate

our simple random sample, we could draw the numbers out

of an urn.

We select "Sample Students," and we draw our 5 numbers.

So the students that are in our sample

are student 8, student 24, student 14, student 15,

and student 32.

Student 8, student 24, student 14, student 15, and student 32.

We find that the average time of their commute is 25 minutes.

Let's obtain a second random sample of size 5

from our student population.

The procedure is the same.

We're going to randomly select 5 unique numbers from an urn.

The individuals corresponding to the numbers are in the sample.

Student 15, student 29, student 1, student 17, and student 6.

We then ask these students to report their travel time

to school.

19 minutes, 18 minutes, 4 minutes, 3 minutes,

and 10 minutes.

We get an average travel time to school of 10.8 minutes.

Notice that the individuals in our first sample

are different from the individuals

in the second sample.

Our first sample had Megan, Uri, Adam, Suman, and Keith.

Our second sample had Alizandro, Matt, Adam, Crystal, and Megan.

For this reason, each sample results

in different descriptive statistics.

The first sample had an average commute time of 25 minutes.

The second sample had an average commute time of 10.8 minutes.

Therefore, any inference based on each sample

may result in different conclusions regarding

the population.

This is the very nature of statistics.

Inferences based on samples will vary

because the individuals in the samples vary.

The individuals differ from sample to sample

because chance is used to select the individuals.

Of course, for large populations--

such as all the students enrolled

in a particular college or university--

the approach just taken in using an urn

would be difficult to use.

Imagine a school with 15,000 students.

It would be quite a large urn that

would require 15,000 balls.

Instead, in practice, random number generators--

such as those found on TI graphing calculators

or StatCrunch-- are used to obtain the individuals

in the simple random sample.