## Module 4, Unit 9 Lesson 23 – Experiments and Random Assignment

### *Sampling Review – Bad Sampling*

a. If people are not randomly selected, misleading conclusions from the data may be drawn. Perhaps the most famous case was in 1936 when *The Literary Digest* magazine predicted that Alf Landon would beat incumbent President Franklin Delano Roosevelt by *370* electoral votes to *161*. Roosevelt won *523* to *8*. Ten million questionnaires were sent to prospective voters (selected from the magazine's subscription list, automobile registration lists, phone lists, and club membership lists), and over two million questionnaires were returned. Surely such a large sample should represent the whole population. How could *The Literary Digest* prediction be so far off the mark?

b. To determine if the potato chips made in a factory contain the desired amount of salt, a sample of chips would be selected randomly so that the sample can be considered to be representative of the population of chips. Discuss two different ways a random sample of 100 chips might be selected from a conveyor belt of chips.

**Goals of a Study**

- The objective of an **observational study** and a **survey** is to learn about characteristics of some population, so the data should be collected in a way that would result in a representative sample. This speaks to the importance of **random *selection*** of subjects for the study.

- The objective of an **experiment** is to answer such questions as "What is the effect of treatments on a response variable?" Data in an experiment need to be collected in a way that does not favor one treatment over another. This demonstrates the importance of **random *assignment*** of subjects in the study to the treatments. **Random *selection*** of subjects, discussed last class, is also important here.

### Random Assignment

In an experiment we wish to randomly assign subjects to treatment groups. The goal of random assignment is to create two (or more) groups that are as similar to each other as possible in all ways except for the treatment received. This allows us to conclude that any observed differences between the groups are likely due to the difference in treatments.

### *Example 1 – Weed Killers*

A University biologist wants to compare the effects of two weed killers on pansies. She chooses 24 plants. She is considering applying weed killer A to the 12 healthiest plants and B to the remaining 12 plants. What is wrong with this approach?

Randomly selecting 12 plants to receive weed killer A and then assigning the rest to B would help ensure that the plants in each group are fairly similar. Could she be sure to get exactly 12 plants assigned to weed killer A and 12 plants to weed killer B by tossing a fair coin for each plant and assigning "heads up" plants to weed killer A and "tails up" to weed killer B?

How might the biologist go about randomly assigning 12 plants from the 24 candidates to receive weed killer A? (The ways we randomly assign treatments are similar to methods we use for gathering a simple random sample.

**Hat Method** **Random Number Generator Method**

***Dogs and Texts Part A***

Two studies are described below. One is an observational study, while the other is an experiment.

*Study A:*
A new dog food, specially designed for older dogs, has been developed. A veterinarian wants to test this new food against another dog food currently on the market to see if it improves dogs' health. Thirty older dogs were randomly assigned to either the "new" food group or the "current" food group. After they were fed either the "new" or "current" food for six months, their improvement in health was rated.

*Study B:*
The administration at a large school wanted to determine if there was a difference in the mean number of text messages sent by ninth-grade students and by eleventh-grade students during a day. Students in a random sample of 30 ninth-grade students were asked how many text messages they sent per day. Students in another random sample of 30 eleventh-grade students were asked how many text messages they sent per day. The difference in the mean number of texts per day was determined.

1. Which study is the experiment? Explain.

2. Describe what a subject is in an experiment. Who/what are the subjects in this experiment?

3. What is a treatment in an experiment? What is the treatment in this experiment?

4. Describe what a response variable is in an experiment. What is it in this experiment?

**Random Selection versus Random Assignment**

People often confuse these two terms. The term *random sample* implies that a sample was randomly selected from a population. The terms *random selection* and *random assignment* have very different meanings.

***Random selection*** refers to randomly *selecting a sample* from a population. Random selection allows generalization to a population and is used in well-designed studies.

***Random assignment*** refers to randomly *assigning the subjects* in an experiment to treatments. Random assignment allows for cause-and-effect conclusions and is used in well-designed experiments.

***Dogs and Texts Part B***

In study B, the data were collected from two random samples of students.

5. Can the results of the survey be generalized to all ninth-grade and all eleventh-grade students at the school? Why or why not?

6. Suppose there really is a difference in the mean number of texts sent by ninth-grade students and by eleventh-grade students. Can we say that the grade level of the students is the cause of the difference in the mean number of texts sent? Why or why not?

In study A, the dogs were randomly assigned to one of the two types of food.

7. Suppose the dogs that were fed the new food showed improved health. Can we say that the new food is the cause of the improvement in the dogs' health? Why or why not?

8. Can the results of the dog food study be generalized to all dogs? To all older dogs? Why or why not?

The table below summarizes the differences between the terms *random selection* and *random assignment.*

9. For each statement, put a check mark in the appropriate column(s), and explain your choices.

| | **Random Selection** | **Random Assignment** |
|---|---|---|
| **Used in Experiments** | | |
| **Used in Observational Studies** | | |
| **Allows Generalization to the Population** | | |
| **Allows a Cause-and-Effect Conclusion** | | |

What is the purpose of random assignment in experiments? To answer this, consider the following investigation:

A researcher wants to determine if the yield of corn is different when the soil is treated with one of two different types of fertilizers, fertilizer A and fertilizer B. The researcher has 16 acres of land located beside a river that has several trees along its bank. There are also a few trees to the north of the 16 acres. The land has been divided into 16 one-acre plots. These 16 plots are to be planted with the same type of corn but can be fertilized differently. At the end of the growing season, the corn yield will be measured for each plot, and the mean yields for the plots assigned to each fertilizer will be compared.

10. For the experiment, identify the following, and explain each answer:
    a. Subjects

    b. Treatments

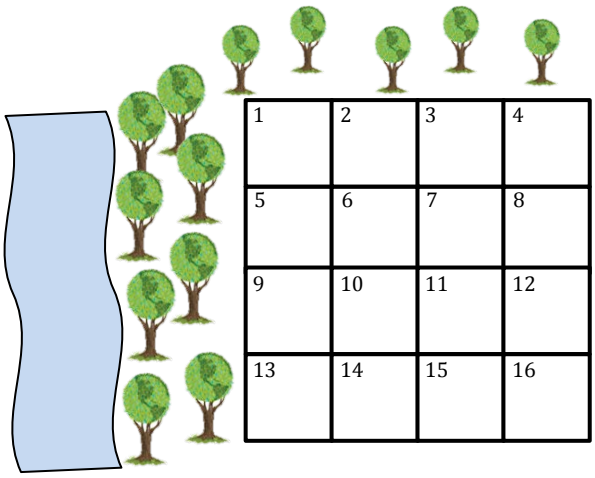
    c. Response variable

***Dogs and Texts Part C***

Next, you need to assign the plots to one of the two treatments. We are going to use two methods of random assignment: Slip Method and Random Number Generator.

11. Write the 8 numbers chosen from the slips of paper. Those plots will get fertilizer A. Write A (for fertilizer A) or B (for fertilizer B) in each of the 16 squares in the diagram so that it corresponds to your random assignment of fertilizer to plots.

    a. On the diagram to the right, draw a vertical line down the center of the 16 plots of land. Count the number of plots on the left side of the vertical line that will receive fertilizer A. Count the number of plots on the right side of the vertical line that will receive fertilizer A.

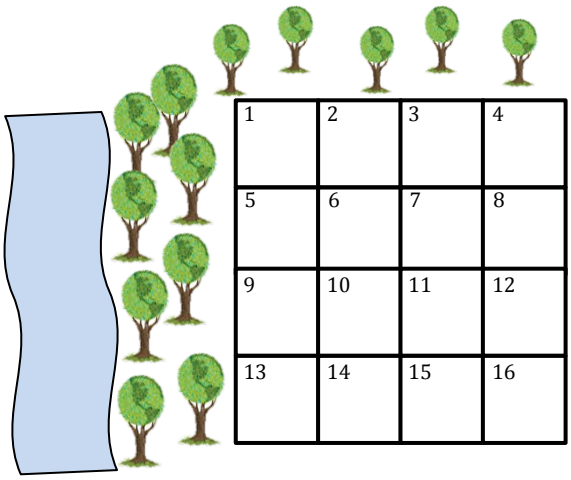



Left _____ Right _____

    b. On the diagram above, draw a horizontal line using highlighter through the center of the 16 plots of land. Count the number of plots above the horizontal line that will receive fertilizer A. Count the number of plots below the horizontal line that will receive fertilizer A.

Above _____ Below _____

12. Write the 8 numbers chosen using the random digit generator. Those plots will get fertilizer A. Write A (for fertilizer A) or B (for fertilizer B) in each of the 16 squares in the diagram so that it corresponds to your random assignment of fertilizer to plots.




    a. On the diagram to the right, draw a vertical line down the center of the 16 plots of land. Count the number of plots on the left side of the vertical line that will receive fertilizer A. Count the number of plots on the right side of the vertical line that will receive fertilizer A.

    Left _____ Right _____

    b. On the diagram above, draw a horizontal line using highlighter through the center of the 16 plots of land. Count the number of plots above the horizontal line that will receive fertilizer A. Count the number of plots below the horizontal line that will receive fertilizer A.

    Above _____ Below _____

In experiments, random assignment is used as a way of ensuring that the groups that receive each treatment are as much alike as possible with respect to other factors that might affect the response.

13. Explain what this means in the context of this experiment, and why the methods we used are better than just choosing 8 plots for each on your own.

14. Suppose that, at the end of the experiment, the mean yield for one of the fertilizers is quite a bit higher than the mean yield for the other fertilizer. Explain why it would be reasonable to say that the type of fertilizer is the cause of the difference in yield and not the proximity to the river or to the northern trees.

## Lesson Summary

- An *experiment* is an investigation designed to compare the effect of two or more treatments on a response variable.
- A *subject* is a participant in the experiment.
    - The *response variable* is a variable that is not controlled by the experimenter and that is
- measured as part of the experiment.
- The *treatments* are the conditions to which subjects are randomly assigned by the experimenter.
- *Random selection* refers to randomly selecting a sample from a population.
    - Random selection allows for generalization to a population.
- *Random assignment* refers to randomly assigning subjects to treatment groups.
    - Random assignment allows for cause-and-effect conclusions.
    - The purpose of random assignment in an experiment is to create similar groups of subjects for each of the treatments in the experiment.