Introduction to Some Experimental Design Terminology

Experiment An investigation that establishes a particular set of circumstances under a specified protocol to observe and evaluate the implications of the resulting observations (*This is Kuehl's definition. It is hard to understand unless you already know what an experiment is.*)

Treatments The set of circumstances created for an experiment

Experimental Unit A physical entity or subject exposed to the treatment independently of the other units

Response Variable A characteristic of an experimental unit measured after treatment and analyzed to address the objectives of the experiment

Describe the treatments, experimental units, and response variables for each of the experiments described below.

1. A field experiment is conducted to compare the yield of three varieties of corn. A field containing 30 plots of land is used for the experiments. Each variety is planted on 10 randomly selected plots. The yield for each plot is measured at the time of harvest.

2. An experiment is conducted to compare the effect of three drugs on the lean percentage of hogs. A total of 30 hogs are assigned to the three drugs in a completely randomized fashion so that 10 hogs will receive weekly injections with each drug. The lean percentage of each hog is recorded at the time of slaughter.

Experimental Error Accounts for the fact that experimental units treated independently and identically will not have identical response variable measurements

Experiment (*definition 2*) An investigation in which the investigator applies some treatments to experimental units and then proceeds to observe the effect of the treatments on the experimental units by measuring one or more response variables

Observational Study An investigation in which the investigator observes units and measures one or more response variables without imposing treatments on the individuals

A researcher studying hospital records finds that the average recovery time for heart attack survivors who own dogs is significantly shorter than average the recovery time for heart attack survivors who do not own a pet of any kind. The researcher concludes that owning a dog reduces recovery time for heart attack survivors.

1. Is this an experiment or an observational study? Explain.

2. Is the researchers conclusion justified? Explain.

3. Describe an experiment that could be used to determine if owning a dog reduces recovery time for heart attack survivors.

4. Explain why it might be inappropriate to conduct an experiment in this case.

Factor An explanatory variable studied in an investigation that can be set at any one of two or more values

Levels The different values of a factor

Full Factorial Treatment Design The treatments consist of all possible combinations of the levels of the factors of interest

Suppose we are conducting an experiment to determine the optimal time and temperature combination for baking a cake. The response variable of interest is taste. We are willing to consider three baking temperatures ($275°F$, $300°F$, and $325°F$) and two baking times (25 and 30 minutes). Four batches of cake will be baked separately at each combination of baking time and temperature.

1. What are the experimental units?

2. What are the factors in this experiment?

3. State the levels of each factor?

4. List the treatments for this full factorial treatment design.

We would like to make treatment comparisons among similar experimental units. Often there will be great variation among the experimental units available for an experiment. The effect of this variation can be controlled by grouping similar experimental units together in groups called **blocks**. Treatments are evaluated by comparing the response variable measurements for experimental units within the same block. This general design strategy is known as **blocking**.

A **completely randomized design** makes no use of blocking. For a given number of experimental units per treatment, all possible assignments of treatments to experimental units are equally likely.

A **randomized complete block design** is the simplest design that uses blocking. Treatments are randomly assigned to experimental units separately within each block in such a way that all treatments occur an equal number of times in each block.

Suppose a researcher would like to compare the effects of five diets (A, B, C, D, and E) on the weight gain of rats. A total of 150 rats are available for an experiment. The rats are housed in 15 cages with 10 rats per cage. For logistical reasons, all rats within any given cage must be fed the same diet. The cages are stored on 3 rectangular tables (with 5 cages per table) in a single rectangular room. The room has three windows, and these windows are equally spaced along the room's south wall. The room's one door is located in the center of the room's north wall. The rectangular tables run parallel to the north and south walls of the room.

1. Why might a researcher consider using a randomized complete block design in this case?

2. What are the experimental units in this case?

3. What is the natural choice for blocks in this case?

4. Sketch the layout of the room and provide one example of an assignment of treatments to experimental units that could arise from a randomized complete block design.

An experiment contains **replication** if at least one treatment is applied independently to two or more experimental units.

An **observational unit** is a unit on which the response variable is measured. Observational units are often the same as experimental units, but that is not always true.

Suppose an experiment is conducted to gauge the effect of soil temperature on the number of seeds produced by soybean plants. Six plants are grown in each of three baths. Three temperatures (low, medium, and high) are assigned at random to the three baths. Each bath is kept at its assigned temperature for the duration of the experiment. At the end of the growing season the number of seeds produced by each of the 18 plants is recorded.

1. Name the treatments in this experiment.

2. How many factors are used in this experiment?

3. Name the levels of each factor.

4. What is the response variable?

5. What are the experimental units in this example?

6. What are the observational units in this example?

7. Does this experiment have replication?

8. Answer the questions 1 through 7 above for the experiment to study the effect of five diets on the weight gain of rats. Assume that the amount of weight gained by each of the 150 rats will be measured at the end of the experiment. How would your answers change if the weight gains were measured on a per cage rather than per rat basis?