

Chapter 13: Data Collection

Simple Observation

Surveys

Designed Experiments

Simple Observation (page 348)

Observing outcomes as they become available

No intervention

No selecting of elements by randomization

Examples:

Photocopy data in textbook

Ages and Heights of Young Girls

Natural gas consumption, Temperature, and Wind speed

Gasoline consumption and Weight of car

Chap. 13 - page 1

Chap. 13 - page 2

Surveys (page 350)

Surveys are “frozen in time.”

- A survey is done during a brief *observation period*
- We usually observe a large number of elements
- High-quality surveys are difficult to carry out

Probability surveys have the following additional characteristics:

- The elements are selected according to a probability sampling design
- Probability sampling permits rigorous, objective inferences about the characteristics of the universe (population) sampled

Surveys (and polls) are cross-sectional studies.

Chapters 14, 15, and 16 deal with the practical and technical aspects of surveys

Chap. 13 - page 3

Chap. 13 - page 4

Experiments (page 353)

In an experiment we attempt to learn what would be true of a process or population **if it were changed**.

terminology:

factors—variables that we control or manipulate to see their effect, if any, on the response.

Sometimes factors are continuous and sometimes categorical.

Frequently factors that are continuous are collapsed into categories.

levels—the values of the factors to be used in the experiment.

treatment—one combination of the levels of the factors to be considered.

Chap. 13 - page 5

A **replication** is a single run (repetition) of the experiment in which all the planned treatments are run at all the planned levels. (page 353)

When working with human subjects, blinding and double blinding are important techniques of experimental design.

blinding—make sure the subject does not know which treatment they are receiving.

double blinding—make sure neither the subject nor the evaluator knows which treatment the subject is receiving

Chap. 13 - page 7

Example from textbook (page 353):

response is quality of baked cake

One factor is the time taken to bake the cake.

This factor was run at three levels, 55, 60, and 65 minutes.

Another factor is the temperature at which the cake is baked.

This factor was run at three levels also—300 degrees, 325 degrees, and 350 degrees Fahrenheit.

A treatment is a combination of one level of the first factor with one level of the second factor.

One treatment is a 65 minute baking time at 300 degrees temperature.

Chap. 13 - page 6

factorial experiment

A **factorial experiment** is an experiment in which all possible combinations of treatment levels are run in each replication.

Examples:

The age guessing experiment is the simplest factorial experiment with one factor at two levels.

The cake baking experiment has two factors each with two levels. There are $3 \times 3 = 9$ different treatments. One replication involves baking 9 cakes.

To eliminate bias, these 9 cakes should be baked in random order.

Chap. 13 - page 8

Comparisons (page 356)

stratification—separating observations into groups or categories in order to make comparisons

control group—receives no treatment or a standard treatment.

experimental group—the group that receives the “real” treatment.

Example:

With human subjects, the control group may receive a placebo.

placebo—a nontreatment disguised as a “real” treatment (a sugar pill)

Chap. 13 - page 9

Three basic principles of designed experiments: (page 359)

- **control**
- **randomization**
- **replication**

Chap. 13 - page 10

Steps in Good Data Collection (page 361)

Data Design Quality—Are the data relevant to the problem we wish to solve?

operational definitions

measurement methods

Data Production Quality—Were the data collected with sufficient skill and care?

Primary Data—data collected by the person or organization that also analyzes the data

Secondary Data—data collected by one person or organization but analyzed by another person or organization

Chap. 13 - page 11

Step 1: State the Purpose for Collecting the Data

Step 2: Determine Sources

Step 3: Determine Data Capture and Presentation Methods

Step 4: Train Personnel

Step 5: Collect the Data Accurately

Step 6: Document the Work

Chap. 13 - page 12