Topic 3.1

Introducing Statistics: Do the Data We Collected Tell the Truth?

Do the Data We Collected Tell the Truth?

Unit 3 introduces the concepts of generalization and causality. Generalization refers to the extrapolation of the results of a study to a larger population, whereas causality refers to the establishment of causal relationships from the results of a study. We may or may not be able to generalize results to the larger population or establish evidence of causal relationships in a study. It all depends on how the data are collected. For example, the results of a study generalize to the larger population only when the sample is randomly selected from the population. Similarly, the results of a study may establish that changes in the response are caused by the effects of the treatments (a cause-effect relationship) only in well-designed experiments where there is random assignment of treatments.

In this unit, we will learn important principles of sampling and study designs. Specifically, we will study random and nonrandom sampling methods, as well as different types of bias that may result when nonrandom sampling methods are used to select a sample from a population. We will describe different types of observational studies and experimental designs, and based on the type of study, determine if it is appropriate to generalize the results to the larger population or to establish causal relationships. For experimental designs, we will focus on the elements of a well-designed experiment, and on explaining why a particular experimental design is more appropriate than others for a given situation.

To prepare for the AP exam, we suggest focusing on developing the following skills:

- Differentiate between random and nonrandom sampling methods, and specifically between stratified random sampling and cluster sampling.
- Differentiate between observational studies and experiments.
- Identify the different sources of bias in sampling methods.
- Identify the experimental units, response variable, explanatory variable, and potential confounding variable in a given experiment.
- Describe the elements of a well-designed experiment.
- Compare the different experimental designs.
- Determine whether a randomized complete block design is more appropriate than a completely randomized design in a given situation.
- Determine whether a matched pairs design is more appropriate than other experimental designs in a given situation.
- Interpret the results of a well-designed experiment in terms of generalizability, statistically significant results, and causal relationships.