Topic 5.1

Introducing Statistics: Why Is My Sample Not Like Yours?

Why Is My Sample Not Like Yours?

Unit 5 applies the probabilistic reasoning introduced in Unit 4 to the sampling distributions of statistics that we will use to make statistical inferences in Units 6 and 7 (for categorical and quantitative data). A sample statistic (ex. sample mean) calculated from a random sample taken for a population can be used to estimate the population parameter (ex. population mean). The mean and standard deviation of the sampling distribution of a sample statistic (ex. sample mean) can be determined directly from the population parameters when certain sampling conditions are met. For sufficiently large samples taken from any population (normal or not normal), the sampling distribution of the sample statistic will have an approximately normal distribution. By simulating sampling distributions, we can understand how the values of sample statistics vary in repeated random sampling from populations with known parameters.

In this unit, we will revisit the normal distribution and extend its properties to calculate the probability of sample statistics (ex. sample mean) when certain sampling conditions are met. We will study the Central Limit Theorem and how to apply the theorem to justify the shape of sampling distributions when appropriate. We will learn the reasons why some estimators (ex. sample mean) are unbiased and how the size of the sample affects the variability of estimators. These concepts will be used to determine parameters, approximate shapes, and calculate probabilities for different sampling distributions. Specifically, we will study sampling distributions for sample proportions and differences in sample proportions, and for means and differences in sample means. The sampling distributions for proportions will become building blocks for statistical inference about population proportions in Unit 6, and the sampling distributions for means will become building blocks for statistical inference about population means in Unit 7.

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

- For the normal distribution, be able to:
  - Calculate the probability that a particular value lies within a given interval of a normal distribution.
  - Determine the interval associated with a given area in a normal distribution.
  - Determine the appropriateness of using the normal distribution to approximate probabilities for unknown distributions.
- For estimators and the sampling distributions for sample proportions, differences in sample proportions, sample means, and differences in sample means, be able to:
  - Understand the Central Limit Theorem and how to justify the shape of a sampling distribution.
  - Understand what makes an estimator unbiased.
  - Determine parameters of sampling distributions.
  - Determine whether sampling distributions can be described as approximately normal.
  - Interpret probabilities and parameters for sampling distributions.

Note: It is not required to memorize the formulas for the parameters of the different sampling distributions. The AP exam provides these formulas. However, specific conditions for the approximation to the normal distribution must be memorized.