Accelerated Pre-Calculus
Unit 8: Inferences & Conclusions from Data

References

Helpful Links:

Measures of Central Tendency - video

Box and Whisker Plots - video

Variance and Standard Deviation - video

Normal Distribution - video

Types of Statistical Studies - video

Comparing and Sampling Populations- video

Sampling Distribution - video

Confidence Intervals - video

Margin of Error - video

Hypothesis Testing - video

Dear Parents,

In this unit, students see how the visual displays and summary statistics they learned in earlier grades relate to different types of data and to probability distributions. They identify different ways of collecting data— including sample surveys, experiments, and simulations—and the role that randomness and careful design play in the conclusions that can be drawn.

Concepts Students will Use & Understand

- Construct appropriate graphical displays (dot plots, histogram, and box plot) to represent sets of data values.
- Describe a distribution using shape, center and spread and use the correct measure appropriate to the distribution.
- Compare two or more different data sets using center and spread.
- Recognize data that is described well by a normal distribution.
- Estimate probabilities for normal distributions using area under the normal curve using calculators, spreadsheets and tables.
- Design a method to select a sample that represents a variable of interest from a population.
- Design simulations of random sampling and explain the outcomes in context of population and known proportions or means.
- Use sample means and proportions to estimate population values and calculate margins of error.
- Read and explain in context data from real-world reports.

Vocabulary

- **Center:** measures of center such as median & mean.
- **Central Limit Theorem:** the mean of a sample of data will be closer to the mean of the overall population as the sample size increases.
- **Confidence Interval**: the probability that the interval reveals the true parameter value.
- **Empirical Rule**: in a normal distribution, the data will fall into 3 standards deviations of the mean.
- **Margin of Error:** the value in the confidence interval that shares how accurate the parameter may be.
- **Parameters:** numerical values that describe a population.
- **Sample Mean:** a statistic measuring the average of the observation in the sample.
- **Sample Proportion:** a statistic indicating the proportion of successes in a particular sample.
- **Sampling Distribution:** the distribution of values taken by a statistic in all possible same size and population samples.
- **Sampling Variability:** the fact that the value of a statistic varies in repeated random sampling.
- **Shape:** distribution described by symmetry, number of peaks, direction of skew, or uniformity.
- **Spread:** variability of the data.
- **Standard Deviation:** the square root of the variance.
- **Variance:** the average of the squares of the deviations of the observations from their mean.

Sample Practice Problems

Data Collection and Design

1. Adam rolled a six sided die 4 times and obtained the following results: 5, 5, 3, and 4. He computed the mean of the 4 rolls and used the result to estimate the mean of the population. Identify the parameter, sample, and statistic of interest. Calculate the identified statistic.

2. The Bennett family believes that they have a special genetic makeup because there are 5 children in the family and all of them are girls. Create a simulation of 100 families with 5 children using coins. Determine the percent of families in which all 5 children are girls and interpret your results.

3. Your classmate Jimmy presents a project to your class about carcinogens, substances that can cause cancer in living cells. When Jimmy said during his presentation that some soda ingredients may be carcinogens, you nearly spit out your root beer. Now you can't rest until you know whether soda consumption is linked to developing cancer. How would you go about investigating whether soda and cancer are linked? What are the benefits and drawbacks of using surveys, experiments, or observational studies?

Summarizing and Interpreting Data

1. Two rival basketball teams each have ten players on a team. The total points scored by each player in the first five games of the season are shown below.
Cougars: 21, 30, 8, 41, 11, 21, 26, 28, 32, 30   Knights: 27, 15, 22, 31, 26, 22, 93, 29, 5, 20
The coaches want to compare the points scored by a typical player on each team. What statistics should the coaches use? Create a box plot and compare those statistics. Then compare any other statistics that are appropriate so that center and spread are compared for both data sets. Identify any outliers and explain their effects.

2. Two small start-up companies are hiring. Josefina, who is interviewing for jobs at both companies, is comparing the salaries of the companies' current employees. The representative for Company A says her company's typical salary is $42,000 per year. The company B representative says his company's typical salary is $63,000 per year. The actual salaries, in thousands of dollars are shown below.
Company A: 21, 33, 35, 40, 42, 45, 45, 49, 160   Company B: 31, 31, 33, 38, 41, 44, 48, 238
Do the figures given by the company representatives really represent the typical salaries for each company? Based on the current employee's salaries, at which company is Josefina likely to earn more money. Explain.

Normal Distribution Curve and its Applications

1. If a population of human body temperatures is normally distributed with a mean of $98.2 \degree$ F and a standard deviation of $0.7 \degree$ F, estimate the percent of temperatures between $98.0 \degree$ F and $99.0 \degree$ F.

Conclusions from Data

1. In 2011, the average salary for a sample of NCAA Division 1 head football coaches was $1.5 million per year, with a standard deviation of $1.07 million. If there are 100 coaches in this sample, what is the standard error of the mean? What can your predict about the population mean based on the sample mean and its standard error?

2. A group of marine biologists placed tracking tags on 100 fish in Lake Erie one summer. The weight of each fish was recorded at the beginning and end of the summer. The average weight gain for all of the tagged fish was 1.2 pounds, with a standard deviation of 0.4 pound. What is the margin of error with 90% confidence for this study?