Module 5: Relating Data and Decisions

TOPIC 1: INTERPRETING DATA IN NORMAL DISTRIBUTIONS

This topic begins by using student knowledge of relative frequency histograms to introduce normal distributions. Students build their knowledge of normal distributions using the Empirical Rule for Normal Distributions. They are introduced to z-scores and use a z-score table and technology to determine the percent of data in given intervals that are bounded by non-integer multiples of the standard deviation from the mean. Finally, students integrate their knowledge of probability and normal distributions to analyze scenarios and make decisions.

Where have we been?

Since middle school, students have created and analyzed data in a variety of distributions, and they have compared different displays. They know that mean and standard deviation are two ways to quantify center and spread, and they know how to calculate each. Students may be familiar with percentiles from growth charts and standardized tests, and this topic makes sense of these values and provides a process to calculate them.

Where are we going?

Students are inundated with data in school, on the news, and through social media. They need to recognize the structure of data sets and which measures of center and spread can be used to interpret them. Developing the ability to analyze data helps students to interpret test scores, make good consumer decisions, and draw their own sound conclusions using data.

The Empirical Rule for Normal Distributions

The Empirical Rule for Normal Distributions states:
- Approximately 68% of the data in a normal distribution for a population is within 1 standard deviation of the mean.
- Approximately 95% of the data in a normal distribution for a population is within 2 standard deviations of the mean.
- Approximately 99.7% of the data in a normal distribution for a population is within 3 standard deviations of the mean.
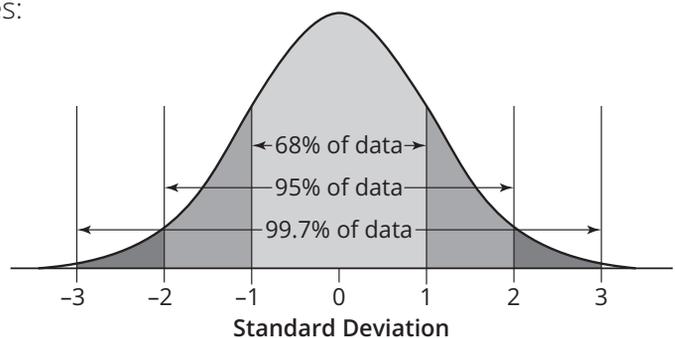
The Empirical Rule applies most accurately to population data rather than sample data. However, the Empirical Rule is often applied to data in large samples.

Black Monday

On October 19, 1987, stock markets around the world fell into sharp decline. In the United States, the Dow Jones Industrial Average dropped 508 points—a 22% loss in value. Black Monday, as the day came to be called, represented at the time the largest one-day decline in the stock market ever.

According to some economic models, the crash that occurred on Black Monday represented an event that was 20 standard deviations away from the normal behavior of the market. Mathematically, the odds of a Black Monday-type event occurring were 1 in 10^{50}.

Talking Points

Normal distributions can be an important topic to know about for college admissions tests.

Here is an example of a sample question:

Suppose SAT scores are approximately normally distributed. In recent years, the SAT mathematics scores have averaged around 480 with a standard deviation of 100. What percent of students will have a math score less than 550?

Calculate the z-score for a math score of 550: $z = \frac{550 - 480}{100} = 0.7$. Then consult a z-score table for 0.7.

Approximately 75.8% of students will score less than 550.

Key Terms

normal distribution
A normal curve is a bell-shaped curve that is symmetric about the mean of the data. A normal curve models a theoretical data set that is said to have a normal distribution.

standard deviation
The standard deviation of data is a measure of how spread out the data are from the mean.

z-scores
A z-score is the number of standard deviations that a data point is from the mean of a normal distribution.

percentile
A percentile is a data value for which a certain percentage of the data is below the data value in a normal distribution.