VOCABULARY

Statistics Numerical values used to summarize and compare sets of data

Measure of central tendency A number used to represent the center or middle of a set of data values. This is represented by the mean, median, and mode.

Measure of dispersion A statistic that tells you how dispersed, or spread out, data values are

Standard deviation A measure that describes the typical difference (or deviation) between a data value and the mean

Outlier A value that is much greater than or much less than most of the other values in a data set

MEASURES OF CENTRAL TENDENCY

- The mean, or average, of n numbers is the sum of the numbers divided by n. The mean is denoted by $\bar{x}$, which is read as "x-bar." For the data set $x_1, x_2, \cdots, x_n$, the mean is
$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$
- The median of n numbers is the middle number when the numbers are written in order. (If n is even, the median is the mean of the two middle numbers.)
- The mode of n numbers is the number or numbers that occur most frequently. There may be one mode, no mode, or more than one mode.

STANDARD DEVIATION OF A DATA SET

The standard deviation σ (read as "sigma") of $x_1, x_2, \cdots, x_n$ is:

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n}}$$

ADDING A CONSTANT TO DATA VALUES

When a constant is added to every value in a data set, the following are true:

- The mean, median, and mode of the new data set can be obtained by adding the same constant to the mean, median, and mode of the original data set.
- The range and standard deviation are unchanged.

MULTIPLYING DATA VALUES BY A CONSTANT

When each value of a data set is multiplied by a constant, the new mean, median, mode, range, and standard deviation can be found by multiplying each original statistic by the same constant.

VOCABULARY

Normal distribution A probability distribution modeled by a bell shaped curve that is symmetric about the mean

Normal curve The bell shaped curve that models a normal distribution

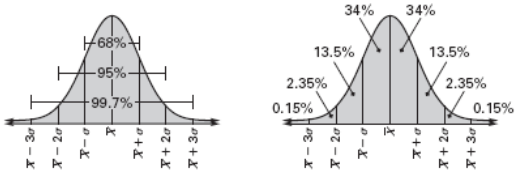
Standard normal distribution The normal distribution with mean 0 and standard deviation 1

z-score The z-value for a particular x-value and is the number of standard deviations the x-value lies above or below the mean $\bar{x}$

AREAS UNDER A NORMAL CURVE

A normal distribution with mean $\bar{x}$ and standard deviation σ has these properties:

- The total area under the related normal curve is 1.
- About 68 % of the area lies within 1 standard deviation of the mean.
- About 95 % of the area lies within 2 standard deviations of the mean.
- About 99.7 % of the area lies within 3 standard deviations of the mean.



VOCABULARY

Population A group of people or objects that you want information about

Sample A subset of the population being studied

Unbiased sample A sample that is representative of the population you want information about

Biased sample A sample that overrepresents or underrepresents part of the population

Margin of error The number that gives a limit on how much the responses of the sample would differ from the responses of the population

MARGIN OF ERROR FORMULA

When a random sample of size n is taken from a large population, the margin of error is approximated by:

$$\text{Margin of error} = \pm \frac{1}{\sqrt{n}}$$

This means that if the percent of the sample responding a certain way is p (expressed as a decimal), then the percent of the population that would respond the same way is likely to be between $p - \frac{1}{\sqrt{n}}$ and $p + \frac{1}{\sqrt{n}}$.