

### Week 3: Determining Normal Distribution

**When trying to determine probabilities, one must first assess whether the variable would have a normal distribution. Using the tools from this course, what are some methods that could be used to determine whether a variable has a normal distribution?**

**Answer:**

The distribution function of a random variable is a function used in statistics to show the probability that an observed value will be found in a particular region of possible values. The normal distribution is commonly used. A normal distribution is a probability that is symmetric about the mean. This means that data around the mean are more frequent in occurrence than those far from the mean. Box plot, histogram, and A P-P plot are some of the methods used to find out whether a variable has a normal distribution.

The box is a common method used to assess the normality of the data. In this method, the median is a flat line within the box while the interquartile range is the stretch of the box. A line running from the bottom and top of the box is used to represent the maximum and minimum

values within  $1\frac{1}{2}$  times the interquartile range from any end of the box. Values high than  $1\frac{1}{2}$  times and 3 times the interquartile range are placed outside the box and said to be outliers. A variable is said to have normal distribution if a box plot is symmetric with the median line at the middle of the box.

A histogram is used to assess the probability distribution of a continuous variable. A variable is said to have normal distribution if its histogram graph is symmetric about the mean and is bell-shaped. A P-P is another method of assessing the distribution of the data. Percent-Percent Plot or Probability-Probability Plot (A P-P) is a graphical method for examining how closely to sets of data (expected and observed) agree. If the data has normal distribution the sets will form a straight line.

Departures from a straight line show that the variable does not have a normal distribution.

References

Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of cardiac anaesthesia*, 22(1), 67.