**Statistics Assignment 2**

1. How can we figure out what the interquartile range is?

The interquartile range is a measure of where the “middle fifty” is in a data set. Where a range is a measure of where the beginning and end are in a set, an interquartile range is a measure of where the bulk of the values lie.

The interquartile range formula is the first quartile subtracted from the third quartile:

IQR = Q3 – Q1.

Step 1: Put the numbers in order.

1, 2, 5, 6, 7, 9, 12, 15, 18, 19, 27.

Step 2: Find the median.

1, 2, 5, 6, 7, 9, 12, 15, 18, 19, 27.

Step 3: Place parentheses around the numbers above and below the median.

Not necessary statistically, but it makes Q1 and Q3 easier to spot.

(1, 2, 5, 6, 7), 9, (12, 15, 18, 19, 27).

Step 4: Find Q1 and Q3

Think of Q1 as a median in the lower half of the data and think of Q3 as a median for the upper half of data.

(1, 2, 5, 6, 7), 9, ( 12, 15, 18, 19, 27). Q1 = 5 and Q3 = 18.

Step 5: Subtract Q1 from Q3 to find the interquartile range.

18 – 5 = 13.

1. What exactly is the value of the 5-number theory?

There are a variety of descriptive statistics. Numbers such as the mean, median, mode, skewness, kurtosis, standard deviation, first quartile and third quartile, to name a few, each tell us something about our data. Rather than looking at these descriptive statistics individually, sometimes combining them helps to give us a complete picture. With this end in mind, the five-number summary is a convenient way to combine five descriptive statistics.

It is clear that there are to be five numbers in our summary, but which five? The numbers chosen are to help us know the center of our data, as well as how spread out the data points are. With this in mind, the five-number summary consists of the following:

* The minimum – this is the smallest value in our data set.
* The first quartile – this number is denoted Q1 and 25% of our data falls below the first quartile.
* The median – this is the midway point of the data. 50% of all data falls below the median.
* The third quartile – this number is denoted Q3 and 75% of our data falls below the third quartile.
* The maximum – this is the largest value in our data set.

The mean and standard deviation can also be used together to convey the center and the spread of a set of data. However, both of these statistics are susceptible to outliers. The median, first quartile, and third quartile are not as heavily influenced by outliers.

1. What is the relationship between standard deviation and variance?

Standard deviation is the square root of the arithmetic mean of the squares of the deviations measured from the arithmetic mean of the data.

Variance is the mean of the squares of the deviations from the mean.

Standard deviation is the square root of variance or variance is the square of standard deviation.

S.D = {Var(X)}1/2

or

Var(X) = (S.D)2

4. What does the difference between variance and standard deviation mean?

Variance is referred to as the variability measure to represent the members of the group. It is the distance of observations corresponding to a higher variance value. It is a squared measure.

Variance is either zero or positive but never negative because of the squared value.

Variance Formula

σ2 = ∑ (x – M)2/n

M=mean

X= value of data set

N= number of observations

Standard Deviation A straightforward dispersion measure is the standard deviation. Also, it defines the method of data value spread around the mean in a data set. We can refer to it as the closeness between the data set values and mean.

The standard deviation always measures according to the original data, and it is still positive.

Standard Deviation

σ = √∑ (x – M)2/n

M=mean

X= value of data set

N= number of observations in the given data set

Differences Between Variance And Standard Deviation

It is essential to know the definition and understand the basic concepts whenever you need to differentiate two terms in any subject. Here we have differentiated both the measures. Difference Between Variance and Standard Deviation In Statistics Variance Variance is a numeric value, and it is a squared value. It defines the variability of observations. The average value of squared deviation is referred to as variance. Its value is always squared. Denominate as (σ2) in mathematics It is the individual data spread out indicator in a group Standard Deviation Measured observations that are the result of measuring by dispersion in the given data set are standard deviation. It is merely the root square value of the mean square deviation The unit is derived from the group of original data only. Denominates as (σ) in mathematics It is referred to as observations' perfect indicator in the available data set.

1. When is it appropriate to refer to a skewed data distribution?

While normal distribution is one of the most common forms of distribution, not all data sets follow this basic bell shaped curve. In other words, some histograms are skewed to the right or left.

If one tail is longer than another, the distribution is skewed. These distributions are sometimes called asymmetric or asymmetrical distributions as they don’t show any kind of symmetry. Symmetry means that one half of the distribution is a mirror image of the other half.

A left-skewed distribution has a long left tail. Left-skewed distributions are also called negatively-skewed distributions. That’s because there is a long tail in the negative direction on the number line. The mean is also to the left of the peak.

A right-skewed distribution has a long right tail. Right-skewed distributions are also called positive-skew distributions. That’s because there is a long tail in the positive direction on the number line. The mean is also to the right of the peak.

On a right-skewed histogram, the mean, median, and mode are all different. In this case, the mode is the highest point of the histogram, whereas the median and mean fall to the right of it (or, visually, the right of the peak).