DAN SCHUMACHER, HDD249

DA 6823

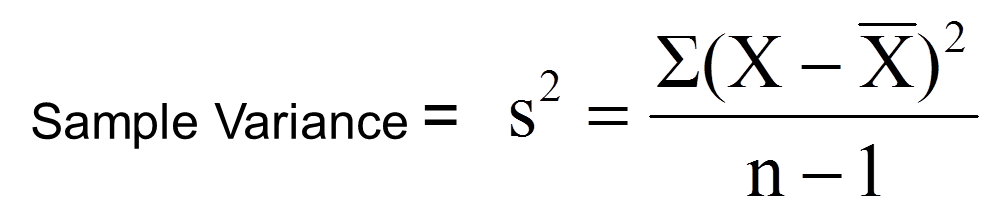
Kilger

Module 1: Part #2 (35 points)

**The Power of Statistics + the Levels of Measurement + the Different Classes of Variables and Determining Appropriate Statistical Technique + Basic Descriptive Measures**

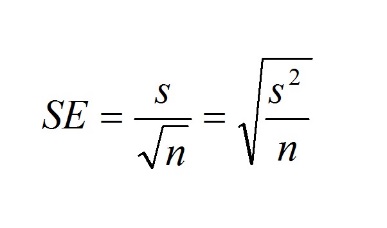
**General Instructions:** In your own words, answer each of the following questions - don’t copy (e.g. cut and paste) some definition out of a book word for word. This is not a group project – you are expected to complete this module on your own. You may refer to text books, online or other sources but not your fellow classmates. If you don’t understand the question, feel free to ask the instructor in class, in office hours or in an email.

1. The first couple of questions deal with the concepts of population and sample.
   1. What is a population? (3 points)
      1. The whole group
   2. What is a sample? (3 points)
      1. A selection of the population
   3. What is the objective of inferential statistics in terms of sample and population? (4 points)
      1. To make *inferences* about the population using the sample.
      2. To make *inferences* about the future using past data.
2. Name two common measures of central tendency (4 points)
   1. Mean
   2. median
3. Describe one situation where one measure of central tendency might be better than another measure (2 points)
   1. If the data is heavily skewed (has at least one large outlier), the median will do a better job of measuring central tendency than the mean.
4. Variance and standard deviation are two of the most used measures of variability. Take a look at the formula for variance below:



* 1. Looking at the variance formula, if the data points (X) are closely packed around the sample mean, what happens to the sample variance? (3 points)
     1. If the data points are close to the sample mean, there will be a small sample variance.
  2. The sample standard deviation is closely related to the sample variance. How? (2 points)
     1. The sample standard deviation is the square root of the sample variance.

1. Many sample statistics you encounter have standard errors associated with them. Imagine that you are studying the heights of the undergraduate student body of UTSA. The total number of students is 30,000 (e.g. the population) and you randomly pull 10 samples of 100 students each from that population.
   1. Will the sample means from each of the 10 samples be the same? ( 2 points)
      1. No, each sample will have unique individuals whose heights vary from one another, therefore there will be *variation* in the sample means.
   2. The standard error of the mean represents the variation in sample means that you find in different samples. The formula below is the standard error of the mean.



1. What happens to the standard error of the mean if there is a lot of variation in the data? (3 points)
   1. The standard error will be larger.

ii. What happens to the standard error of the mean as sample size increases? (3 points)

1. The standard error of the mean gets smaller as the sample size increases (if the variation stays the same).
2. The sum of squares is a statistical concept that measures variation in data that you will find in many different statistical techniques. Here is the formula for sum of squares below:



* 1. This formula should look familiar. What other statistical measures of variation are related to this? (3 points)
     1. Sample variance
  2. As the data points (X) get spread out farther and farther from the sample mean, what happens to the sum of squares? (3 points)
     1. As X gets farther from the sample mean, the sum of squares becomes larger.