FUNDAMENTALS OF ENGINEERING STATISTICAL ANALYSIS

ISE/DSA 5013

Assignment 1

Show your work for calculation problems. You will receive no credit if you only provide the answer. If the problem asks you to construct a graphical depiction, you should use software as much as possible. Make sure you label your axes. *As with all homework this semester, spend time to be neat and organized. Any disorganized submissions are subject to a zero grade.*

Problem 1

Imagine a database from the Gallogly College of Engineering that contained the starting salaries for a sample of Spring 2018 graduating engineers. Recreate the following table and fill out at the intersections what happens to each of the sample statistics when every salary is altered by the specified amount.

	Due to an increase in every salary by	
Change in	$1000	5%
Mean		
Median		
Standard deviation		

Problem 2

In Florida, civil engineers are designing roads with the latest safety-oriented construction methods in response to the fact that an increasing people in Florida were killed by bad roads. One year, a total of 135 traffic accidents that occurred was attributed to poorly constructed roads. A breakdown of the poor road conditions that caused the accidents is found in the *Assignment 1* workbook. I recommend checking out the COUNTIF function in Excel to assist.

a. Produce a bar graph of the data.
b. Produce a pie chat of the data.
c. What is your opinion on how effective one graphical method is relative to the other?

Problem 3

Classify each of the following as *nominal*, *ordinal*, *interval*, or *ratio* data (adapted from [Black 2012]).

a. The time required to produce each tire on an assembly line
b. The number of quarts of milk a family drinks in a month
c. The ranking of four machines in your plant that have been designated as excellent, good, satisfactory, and poor
d. The telephone area code of clients in the US
e. The age of each of your employees
f. The dollar sales at the local pizza shop each month
g. An employee's identification number
h. The response time of an emergency unit

Problem 4

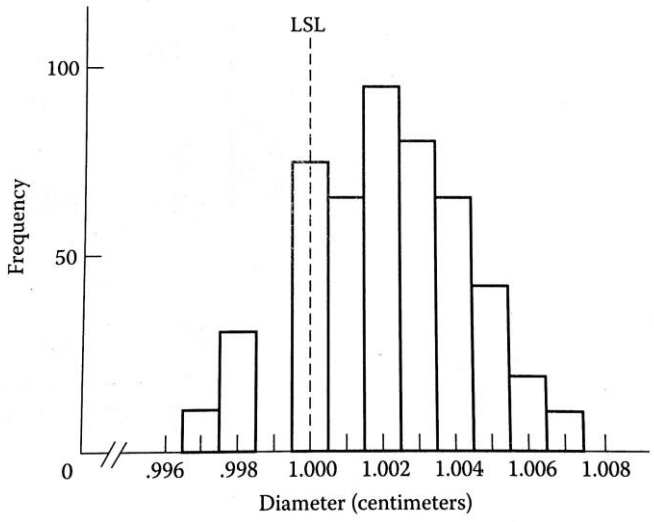
Industrial engineers periodically conduct "work measurement" analyses to determine the time required to produce a single unit of output. At a large processing plant, the number of total worker-hours required per day to perform a certain task was recorded for 50 days. The data shown in the *Assignment 1* workbook.

a. Compute the mean, median, and mode of the data set. What do the differences in these measures, if any, suggest about the data set?
b. Find the range, variance, and standard deviation of the data set.
c. Construct a box plot for the data. Do you detect any outliers?
d. Find the 70^{th} percentile for the data on total daily worker-hours and interpret its value.

Problem 5

In his essay "Making Things Right," W. Edwards Deming considered the role of statistics in the quality control process for a manufacturer of steel rods. Rods produced with diameters smaller than 1 centimeter fit too loosely in their bearings and ultimately must be rejected (thrown out). To determine whether the diameter setting of the machine that produces the rods is correct, 500 rods are selected from the day's production and their diameters are recorded. The distribution of the 500 diameters for one day's production is shown in the figure below. Note that the symbol LSL in the figure represents the 1-centimeter lower specification limit of the steel rod diameters.

a. What type of data, quantitative or qualitative, does the figure portray?
b. What type of graphical method is being used to describe the data?
c. Use the figure to estimate the proportion of rods with diameters between 1.0025 and 1.0045 centimeters.
d. There has been speculation that some of the inspectors are unaware of the trouble that tan undersized rod dimeter would cause later in the manufacturing process. Consequently, these inspectors may be passing rods with diameters that were barely below the lower specification limit and recording them in the interval centered at 1.000 centimeter. According to the figure, is there any evidence to support this claim? Explain.



Problem 6

Why is the formula $\sum_{i=1}^{n}(x - \bar{x})$ *not* used as a measure of dispersion? Show mathematically (i.e., prove) for a set of observations $x_1, x_2, \ldots, x_n$ why this formula fails to give useful information about the dispersion of the data. (Hint: You may want to calculate the formula for a few data points, but you still must show it mathematically for a general set of n observations.)