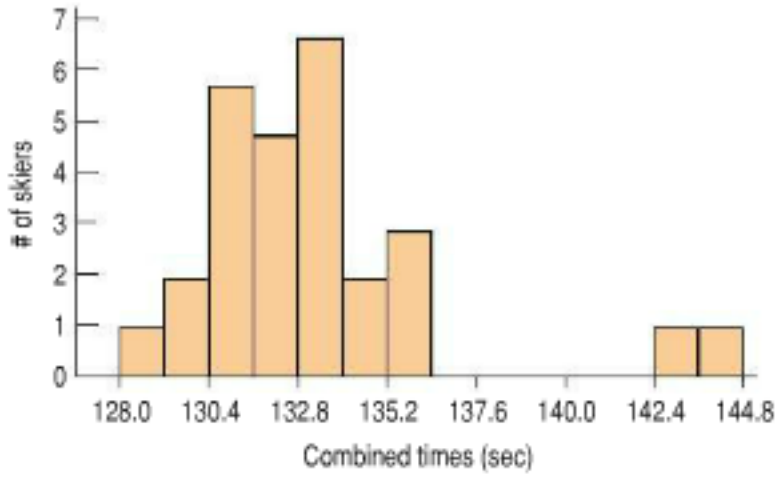
201-H02-HR **Fake Final Exam** **Fall 2016**

1. From the following histogram, discuss the modality, skew, and outliers of the data distribution. Would the mean and standard deviation be appropriate measures of centre and spread, respectively, for this distribution? Explain.

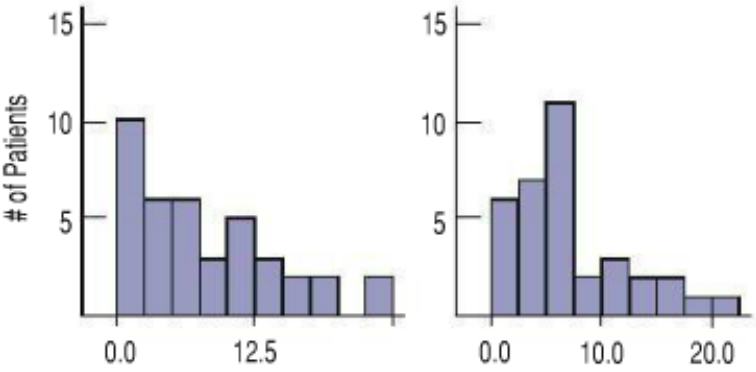


2. The magnitudes of the last 7 Tsunami causing earthquakes are given below (rounded to the nearest integer).

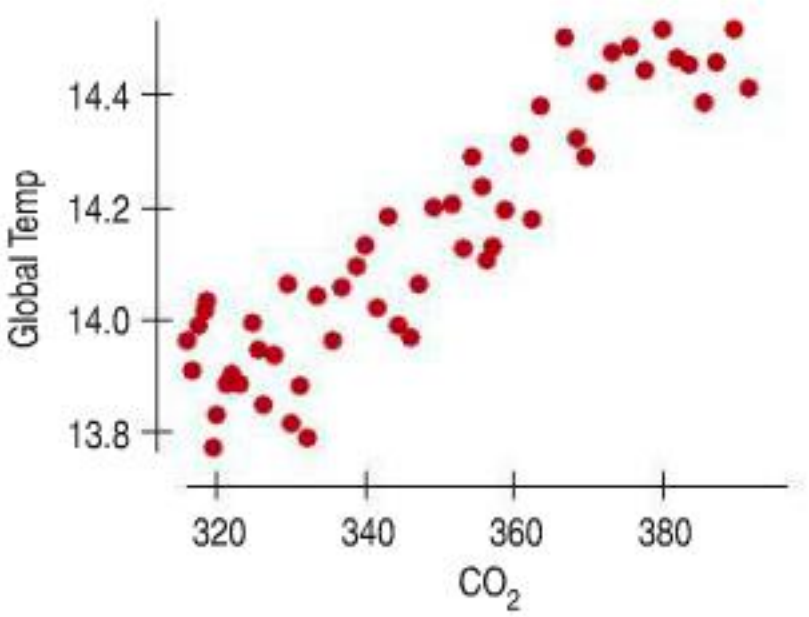
4, 3, 10, 7, 8, 9, 3

Compute the mean, median, first quartile, third quartile, range, and standard deviation of this data set.

3. Describe the shapes of the distributions depicted by each of the following histograms separately. Then compare the distributions portrayed by these histograms.



4. Consider the following scatterplot of atmospheric CO_2 levels vs. Global Temperature from 1959 to 2011.



   a. Describe the direction, form, and strength of this scatterplot.

   b. Comment on each of the Correlation Conditions for this scatterplot.

5. Consider a scatterplot which meets the Correlation Conditions and which has the following attributes:

$$\bar{x} = 15$$
$$\bar{y} = 35$$
$$\sum (x - \bar{x})^2 = 75$$
$$\sum (y - \bar{y})^2 = 130$$
$$\sum (x - \bar{x})(y - \bar{y}) = -90$$

   a. Calculate and interpret the Correlation Coefficient.

   b. Construct the Linear Model and use it to approximate the value of the response variable when *x* = 25.

   c. Calculate and interpret the Coefficient of Determination.

6. Calculate each of the following probabilities.

    a. Calculate the probability that a randomly-generated 6-digit number will have no 2s or 3s. For this question, a 6-digit number cannot start with a 0.

    b. Six men and five women are living at an alien outpost as what they believe is the first wave of colonizers to a new world. In fact, prior to leaving Earth, four of them were randomly and secretly infected with Bovine spongiform encephalopathy so the government could research the implications of this disease in small isolated social groups. If only four of them are randomly infected, what is the probability that two men and two women will have the disease?

7. In a survey of 216 men and 167 women, 112 women reported that they believe Cylons look like us now and live amongst humans in secret waiting for the day to overthrow their human masters, while 97 men reported that they do not believe Cylons look like us now, and are in fact just fancy toasters.

    a. Construct and complete a contingency table that displays the numbers of men and women that do and do not believe Cylons look like us now. Be sure to have a row and a column for the totals.

    b. Compute the probability that a person surveyed is male given that the person believes Cylons look like us now.

    c. Are gender and believing Cylons looks like us now dependent? Justify your answer by comparing two relevant probabilities.

8. Consider the random variable *X,* whose possible values, and their corresponding probabilities, are listed below.

X	0	1	2	3
P(X=x)	0.1	0.3	0.4	0.2

    Compute the expected value of *X* and compute the standard deviation of *X*.

9. We know that there are 12 human Cylon models but we do not know what they look like. There can be multiple copies of each model and thus we can assume that 12% of "humans" are actually Cylons. You can randomly select 200 individuals to test if they are human. What is the probability that you find:
    a. At least 20 Cylons?
    b. 30 to 40 Cylons?
    c. No more than 40 Cylons?

10.
    a. List the three properties that define every set of Bernoulli trials.

    b. If the population modeled by a set of trials is finite, then the trials cannot, technically, be Bernoulli trials if we draw without replacement. We still draw from finite populations for Bernoulli trials provided that they meet the *10% Condition*. Explain what this condition is and which of the 3 properties in part a) is violated when using a finite population.

    c. Why do we sometimes approximate Binom *(n,p)* with a Normal Model? Explain what extra condition must we check in order to make this approximation.

    d. When would we need to use a Poisson Model to approximate a Normal Model? What conditions must be met?

11. An airline, believing that 5% of passengers fail to show up for flights, overbooks (sells more tickets than there are seats). Suppose a plane will hold 265 passengers, and the airline sells 275 tickets. What is the probability that the airline will not have enough seats, so someone gets bumped? Be sure to show the check of the appropriate condition you will need to verify in order to proceed with your approximation.

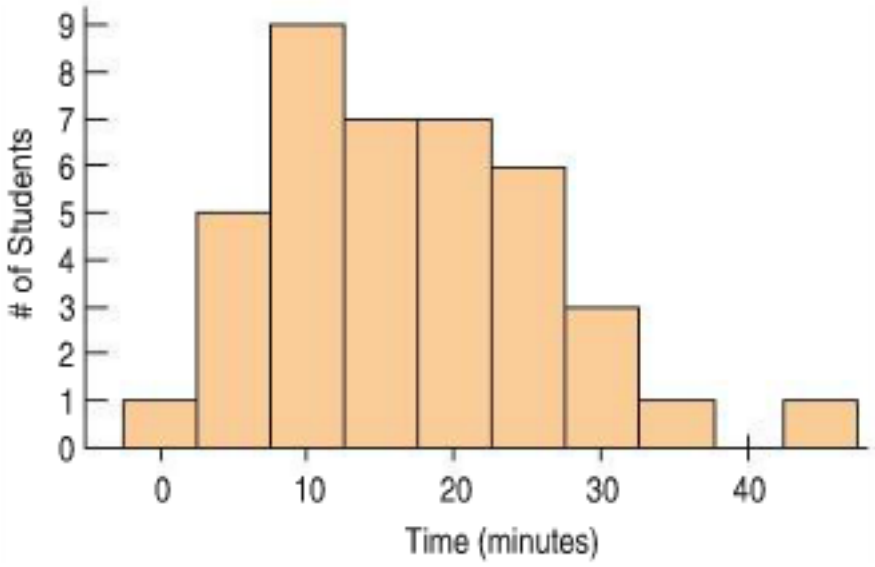
12. Calculate the confidence you would have in a confidence interval based on a sample proportion of $\hat{p} = 0.75$ with a margin of error of 3 percentage points and a sample size of 500.

13. A study of 118 men and 123 women found that 78 of the men and 89 of the women responded that they believe the moon landing was faked.

    a. Assuming that the survey of the men and the survey of the women each meet all the criteria for a set of Bernoulli trials separately, what extra condition must be met in order for us to compare the proportions of men and women who responded that they believe the moon landing was faked?

    b. Assuming that the condition from part a) is met, perform a hypothesis test to determine whether there is a difference between the proportions of men and women who responded that they believe the moon landing was faked. For an alternative hypothesis, claim that the two proportions are different. Be sure to properly state the null hypothesis, use a 5% significance level, and draw some conclusions. Do not forget to pool the data when it is called for in your solution.

14. The distribution of travel times, in minutes, for 40 Ontario secondary school students is shown in the histogram below.



    a. Does this distribution satisfy the Nearly Normal Condition? Explain.

    b. Assuming that the distribution does satisfy all the required conditions, construct a 95% confidence interval for the mean given that the sample mean is 17.00 minutes and the sample standard deviation is 9.66 minutes. Write a concluding sentence interpreting this confidence interval.