AP Statistics | Univariate Data Analysis Review

1. Which of the following are NOT true?
(a)   In a symmetric distribution, the mean and the median are equal
(b)   About fifty percent of the scores in a distribution are between the first and third quartiles
(c)   In a symmetric distribution, the median is halfway between the first and third quartiles
(d)   The median is always greater than the mean
(e)   The range is the difference between the largest and the smallest observation in the data set

2. A set of data has a mean that is much larger than the median. Which of the following
    statements is **most** consistent with this information?
(a)   the distribution is symmetric
(b)   the distribution is skewed left
(c)   the distribution is skewed right
(d)   the distribution is bimodal
(e)   the data set probably has a few low outliers

3.
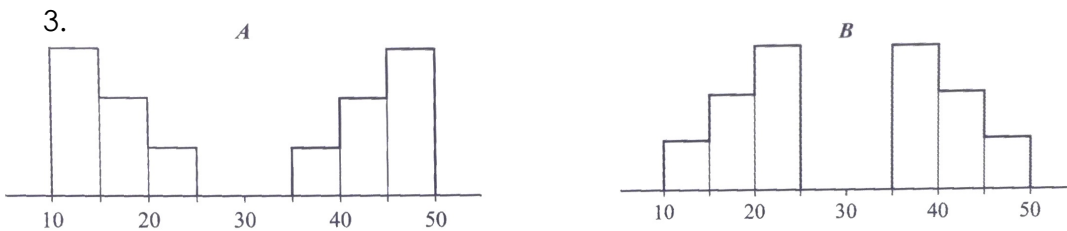
Given the above two histograms, which of the following statements is *incorrect?*
(a) Both sets have approximately the same mean
(b) Both sets have approximately the same median
(c) Both sets have approximately the same range
(d) Set A has a greater variance than does set B
(e) Each set has approximately 12 elements (6 under and 6 over the median)

4. The sixteenth percentile of a Normally distributed variable has a value of 25 and the 97.5
    percentile has a value of 40. Which of the following is the best estimate of the mean and
    standard deviation of the variable?
(a) $\mu \approx 32.5$; $\sigma \approx 2.5$
(b) $\mu \approx 32.5$; $\sigma \approx 5$
(c) $\mu \approx 32.5$; $\sigma \approx 10$
(d) $\mu \approx 30$; $\sigma \approx 2.5$
(e) $\mu \approx 30$; $\sigma \approx 5$

5. The mean age (at inauguration) of all U.S. Presidents is approximately Normally distributed
    with a mean of 54.6 years. Barack Obama was 47 when he was inaugurated, which is the 11th
    percentile of the distribution. George Washington was 57. What percentile was he in?
(a) 6.17
(b) 65.07
(c) 62.92
(d) 34.83
(e) 38.9

6. High school textbooks don't last forever. The lifespan of all high school statistics textbook is approximately Normally distributed with a mean of 9 years and a standard deviation of 2.5 years. What percentage of the books last more than 10 years?
(a) 11.5%
(b) 34.5%
(c) 65.5%
(d) 69.0%
(e) 84.5%

7. Suppose the average score on a national exam is 500 with a standard deviation of 100. If each score is increased by 20 and the result is increased by 10 percent, what are the new mean and standard deviation?
(a) $\mu = 570, \sigma = 100$
(b) $\mu = 570, \sigma = 110$
(c) $\mu = 572, \sigma = 100$
(d) $\mu = 572, \sigma = 110$
(e) $\mu = 572, \sigma = 132$

8. Given that the median is 270 and the interquartile range is 20, which of the following statements is true?
   I. Fifty percent of the data are greater than or equal to 270
   II. Fifty percent of the data are between 260 and 280
   III. Seventy-five percent of the data are less than or equal to 280
(a)   I only
(b)   II only
(c)   III only
(d)   I and II
(e)   I, II, and III

9. A data set includes two outliers, one at each end. If both these outliers are removed, which of the following is a possible result?
(a) Both the mean and standard deviation remain unchanged
(b) Both the median and standard deviation remain unchanged
(c) Both the standard deviation and variance remain unchanged
(d) Both the mean and median remain unchanged
(e) Both the mean and standard deviation increase

10. When a set of data has suspect outliers, which of the following are preferred measures of central tendency and of variability?
(a) mean and standard deviation
(b) mean and variance
(c) mean and range
(d) median and range
(e) median and interquartile range

1. The Graduate Record Examinations are widely used to help predict the performance of applicants to graduate schools. The range of possible scores on a GRE is 200-900. They psychology department at a university finds that the scores of its applicants on the quantitative GRE are approximately Normal with a mean of 544 and a standard deviation of 103.

(a) What percent of applicants scored between 500 and 700?

(b) What minimum score would a student need in order to score better than 77% of those taking the test?

2. A company is considering implementing one of two quality control plans for monitoring the weights of automobile batteries that it manufactures. If the manufacturing process is working properly, the battery weights are approximately Normally distributed with a specified mean and standard deviation.

Quality Control Plan A calls for rejecting a battery as defective if its weight falls more than 2 standard deviations below a specified mean.

Quality Control Plan B calls for rejecting a battery as defective if its weight falls more than 1.5 interquartile ranges below the lower quartile of the specified population.

(a) What proportion of batteries will be rejected by Plan A assuming the manufacturing process is under control?

(b) What proportion of batteries will be rejected by Plan B assuming the manufacturing process is under control?

3. A consumer advocate conducted a test of two popular gasoline additives, A and B. There are claims that the use of either of these additives will increase gasoline mileage in cars. A random sample of 30 cars was selected. Each car was filled with gasoline and the cars were run under the same driving conditions until the gas tanks were empty. The distance traveled was recorded for each car.

Additive A was randomly assigned to 15 of the cars and additive B was randomly assigned to the other 15 cars. The gas tank of each car was filled with gasoline and the assigned additive. The cars were again run under the same driving conditions until the tanks were empty. The distance traveled was recorded and the difference in the distance with the additive minus the distance without the additive for each car was calculated.

The following table summarizes the calculated differences. Note that negative values indicate less distance was traveled (in miles) with the additive than without the additive.

Additive	Values below Q1	Q1	Median	Q3	Values above Q3
A	-10, -8, -2	1	3	4	5, 7, 9
B	-5, -3, -3	-2	1	25	35, 37, 40

(a) On the grid below, display parallel boxplots (showing outliers, if any) of the differences of the two additives.
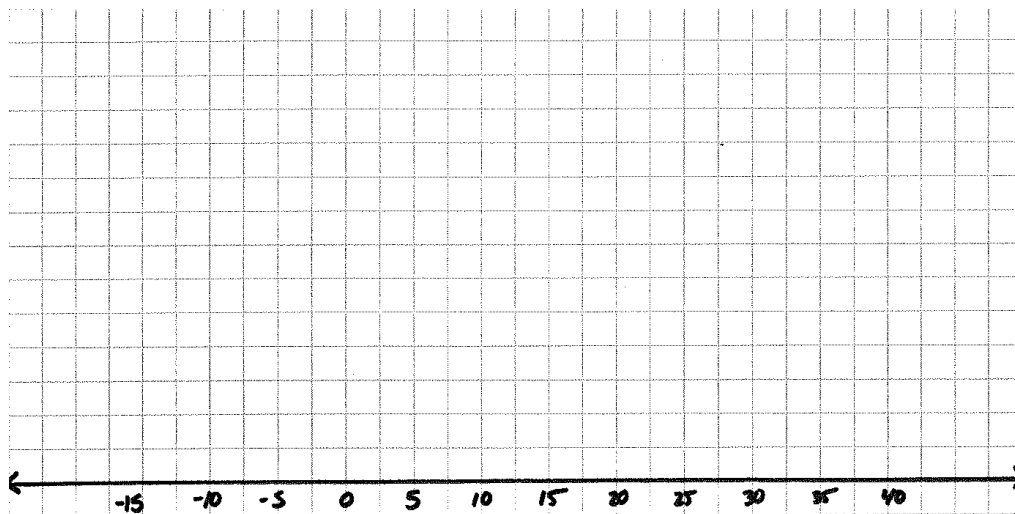
(b) Two ways that the effectiveness of a gasoline additive can be evaluated are by looking at either
- The proportion of cars that have increased gas mileage when the additive is used in those cars **OR**
- The mean increase in gas mileage when the additive is used in those cars

(i) Which additive, A or B, would you recommend if the goal is to increase gas mileage in the highest proportion of cars? Explain your choice.

(ii) Which additive, A or B, would you recommend if the goal is to have the highest mean increase in gas mileage? Explain your choice.