Classroom Voting Questions: Elementary Statistics

Describing Distributions with Numbers

1. In a certain university there are three types of professors. Their salaries are approximately normally distributed within each of the following types:

   - Assistant Professors make a median salary of \$50K, with a minimum of \$40K and a maximum of \$60K.
   - Associate Professors make a median salary of \$65K per year, with minimum of \$57K and a maximum of \$80K.
   - Full Professors make a median salary of \$90K per year, with a minimum of \$70K and a maximum of \$110K.

   There are 1600 total Professors at this University, with the following distribution: 50% of all Professors are Assistants, 30% are Associates, and 20% are Fulls.

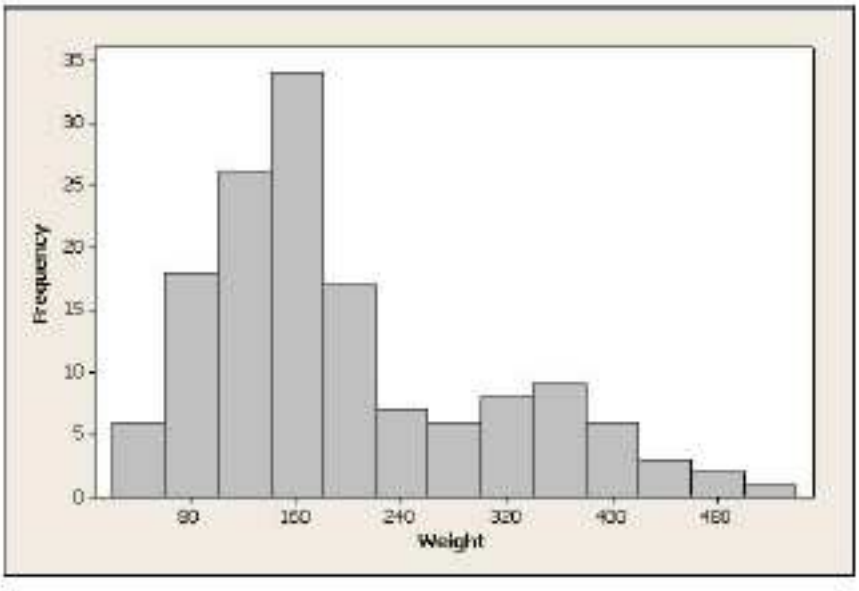
   What can we say about the average salary at this university?

   (a) mean $<$ median
   (b) mean $=$ median
   (c) mean $>$ median
   (d) insufficient information

2. Many individuals, after the loss of a job, receive temporary pay unemployment compensation until they are re-employed. Consider the distribution of time to re-employment as obtained in an employment survey. One broadcast reporting on the survey said that the average time until re-employment was 4.5 weeks. A second broadcast reported that the average was 9.9 weeks. One of your colleagues wanted a better understanding of the situation and learned (through a Google search) that one report was referring to the mean and the other to the median and also that the standard deviation was about 14 weeks. Knowing that you are a statistically-savvy person, your colleague asked you which is most likely the mean and which is the median?

   (a) 4.5 is the mean and 9.9 is the median.
   (b) 4.5 is the median and 9.9 is the mean.
   (c) Neither (A) nor (B) is possible given the SD of the data.
   (d) I am not a statistically-savvy person, so how should I know?

3. For the data set displayed in the following histogram, which would be larger?



(a) mean

(b) median

(c) Can't tell from the given histogram.

4. Why is the term $(n-1)$ used in the denominator of the formula for sample variance?

(a) There are $(n-1)$ observations.

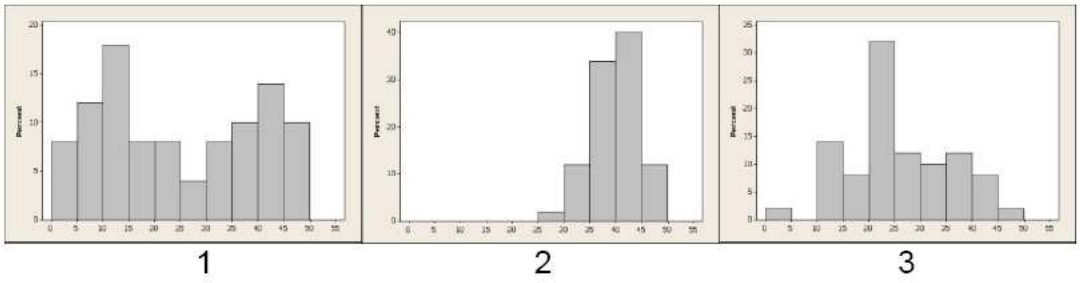
(b) There are $(n-1)$ uncorrelated pieces of information.

(c) The $(n-1)$ term gives the correct answer.

(d) There are $(n-1)$ samples from the population.

(e) There are $(n-1)$ degrees of freedom.

5. Which of the three histograms shown summarizes the data set with the smallest standard deviation?

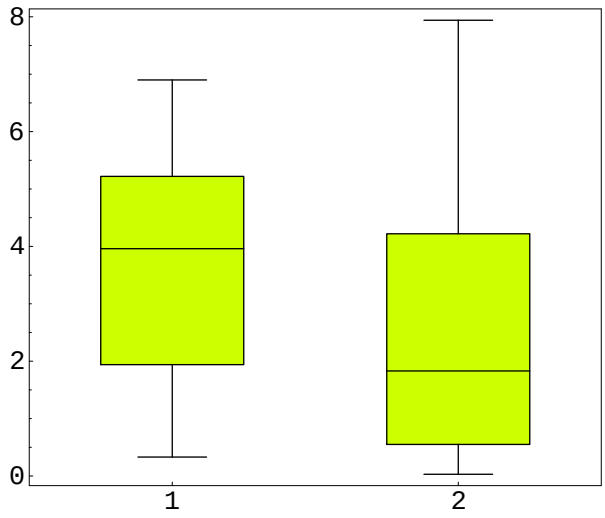


6. Suppose your statistics instructor tells you that you scored 70 on an exam and that the class mean was 74. You should hope that the standard deviation of exam scores was _____.

   (a) Small

   (b) Large

7. Below are boxplots for two data sets.



   TRUE or FALSE: There is a greater proportion of values outside the box for the set on the right than for the set on the left.

   (a) True, and I am very confident.

   (b) True, and I am not very confident.

   (c) False, and I am not very confident.

   (d) False, and I am very confident.

8. The five-number summary for all student scores on an exam is 29, 42, 70, 75, 79. Suppose 200 students took the test. How many students had scores between 42 and 70?

   (a) 25

   (b) 28

   (c) 50

   (d) 100

9. The five-number summary for all student scores on an exam is 40, 60, 70, 75, 79. Suppose 500 students took the test. What should you conclude about the distribution of scores?

   (a) Skewed to the left.

   (b) Skewed to the right.

   (c) Not skewed.

   (d) Not enough information given to determine skew.

10. Jack uses a calculator to find the sample standard deviation of a data set and ends up getting a negative number as the result. This implies that

    (a) on average, the deviations from the mean are negative.

    (b) the mean of the deviations is negative.

    (c) the mean is negative.

    (d) Jack made a mistake.

11. Which set of two observations would you expect to have the smaller standard deviation, the weights of two randomly-selcted professional ballerinas, or the weights of two randomly-selected sumo wrestlers?

    (a) The weights of the ballerinas, because ballerinas are lighter.

    (b) The weights of the sumo wrestlers, because sumo wrestlers are heavier.

    (c) The weights of the ballerinas; their weights are more likely to be closer together.

    (d) The weights of the ballerinas; their weights are more likely to be farther apart.

    (e) The weights of the sumo wrestlers; their weights are more likely to be closer together.

    (f) The weights of the sumo wrestlers; their weights are more likely to be farther apart.

12. When a professional sumo wrestler joined a group of people, the standard deviation of the weights of the new group members was substantially less than the standard deviation of the weights of the original group members. Which of the following is most likely?

    (a) The original group consisted of 3 professional sumo wrestlers.

    (b) The original group consisted of 100 professional sumo wrestlers.

    (c) The original group consisted of 3 professional ballerinas.

    (d) The original group consisted of 100 professional ballerinas.

13. A multi-billionaire decides to retire back in the small town in which she grew up. All of the houses in this town are modest and inexpensive. On the outskirts of town, she builds a huge, luxurious mansion. Consider house prices in the town before and after she builds her mansion. Which of the following measures of central tendency changes the most?

    (a) mean
    (b) median
    (c) mode

14. Which of the following measures is resistant to the influence of outliers?

    (a) mean
    (b) median
    (c) standard deviation
    (d) Q_3
    (e) interquartile range
    (f) two out of (a) through (e)
    (g) three out of (a) through (e)
    (h) four out of (a) through (e)

15. In a history class with over 500 students, a professor gave a very easy test, so that the distribution of scores was highly left-skewed. Which measure of central tendency and measure of variation should be used to summarize the scores?

    (a) median and standard deviation
    (b) median and interquartile range
    (c) mean and standard deviation
    (d) mean and interquartile range

16. Consider a data set that consists of the following four numbers: 2, 5, 6, and a certain number that is less than negative one million. For this data set, rank the following from least to greatest: mean, median, standard deviation.

    (a) mean, median, standard deviation
    (b) mean, standard deviation, median
    (c) median, mean, standard deviation
    (d) median, standard deviation, mean
    (e) standard deviation, mean, median
    (f) standard deviation, median, mean