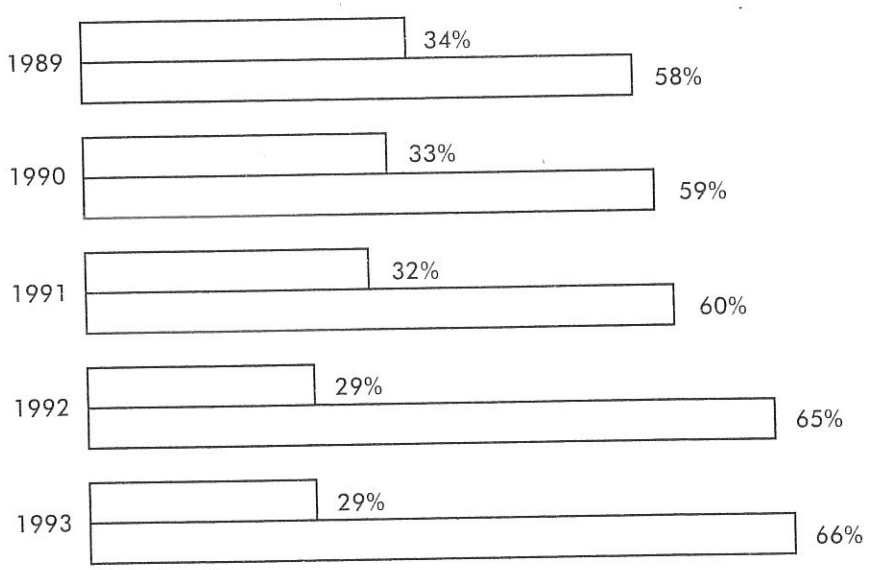
Questions on Topic Three: Comparing Distributions

Multiple-Choice Questions

Directions: The questions or incomplete statements that follow are each followed by five suggested answers or completions. Choose the response that best answers the question or completes the statement.

1. The following double bar graph shows the choice of heating fuel for new homes completed during the indicated years. In each pair the top bar indicates electricity while the lower bar indicates gas.

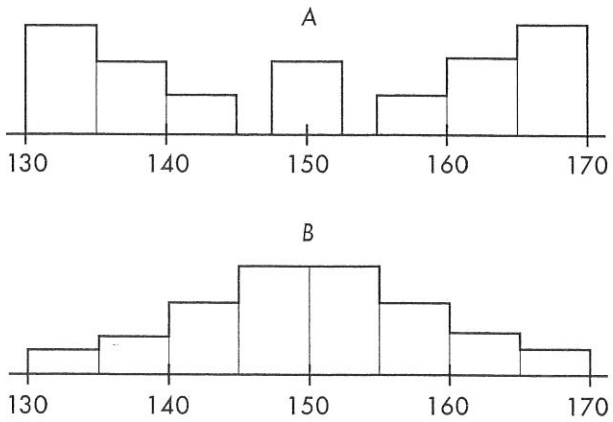


Which of the following are true statements?

I. The percentage of new homes using electricity has never risen from one year to another.
II. The percentage of new homes using gas has increased every year.
III. The percentage of new homes using fuels other than electricity and gas has never risen from one year to another.

(A) I and II
(B) II and III
(C) I and III
(D) I, II, and III
(E) None of the above gives the complete set of true responses.

For Questions 2 and 3 consider the following two histograms:



2. Which of the following statements are true?

   I. Both sets have the same mean.
   II. Both sets have the same range.
   III. Both sets have the same variance.

   (A) I only
   (B) I and II
   (C) I and III
   (D) I, II, and III
   (E) None of the above gives the complete set of true responses.

3. Which of the following statements are true?

   I. The empirical rule applies only to set *A*.
   II. You can be sure that the standard deviation of set *A* is greater than 5.
   III. You can be sure that the standard deviation of set *B* is greater than 5.

   (A) I only
   (B) II only
   (C) I and II
   (D) I and III
   (E) II and III

4. Consider the following back-to-back stemplots comparing car battery lives (in months) of samples of two popular brands.

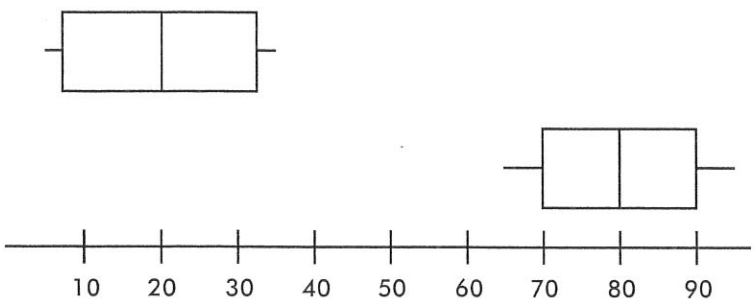
Brand A		Brand B
	3	7
7	4	2 3 4 8 8
3 2	5	1 4 5 6 7 8 9 9
8 7 5 4	6	3 4 6 6 8
9 6 5 3 3 0	7	6
6 5 4 3 3 3 1	8	

Which of the following are true statements?

I. The sample sizes are the same.
II. The ranges are the same.
III. The variances are the same.
IV. The means are the same.
V. The medians are the same.

(A) I and II
(B) I and IV
(C) II and V
(D) III and V
(E) I, II, and III

5. Consider the following parallel boxplots illustrating the daily temperatures (in degrees Fahrenheit) of an upstate New York city during January and July.

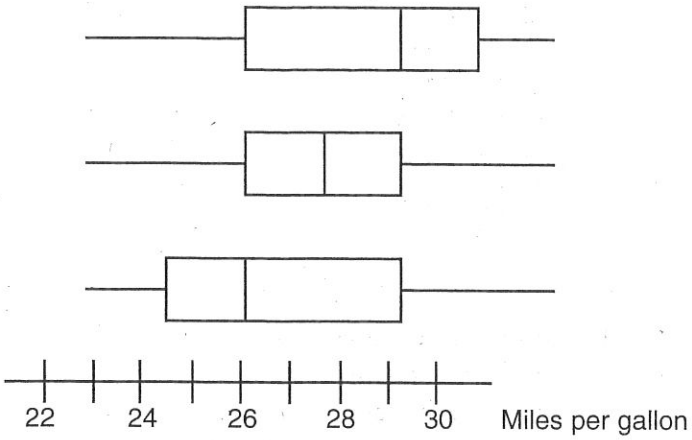


Which of the following are true statements?

I. The ranges are the same.
II. The interquartile ranges are the same.
III. Because of symmetry, the medians are the same.

(A) I only
(B) II only
(C) I and II
(D) I and III
(E) II and III

For Questions 6 and 7 consider the following parallel boxplots of gasoline mileage for three car makes:



6. Which of the following are true statements?

I. All three have the same range.
II. All three have the same interquartile range.
III. The difference in the medians between the first and third distributions is equal to the interquartile range of the second distribution.

(A) I and II
(B) I and III
(C) II and III
(D) I, II, and III
(E) None of the above gives the complete set of true responses.

7. Which of the following are true statements?

I. All three are symmetric.
II. The first is skewed to the left while the third is skewed to the right.
III. The second is skewed on both sides.

(A) I only
(B) II only
(C) III only
(D) II and III
(E) None of the above gives the complete set of true responses.

8. Consider the following back-to-back stemplot:

73	2	
642	3	37
7	4	246
9300	5	7
9920	6	0039
943	7	0299
8	8	349
	9	8

Which of the following are true statements?

I. The distributions have the same mean.
II. The distributions have the same range.
III. The distributions have the same variance.

(A) II only
(B) I and II
(C) I and III
(D) II and III
(E) I, II, and III

Answer Key

1. **D**	3. **E**	5. **A**	7. **B**
2. **B**	4. **A**	6. **B**	8. **D**

Answers Explained

1. **(D)** The percentage of homeowners choosing electricity has gone down from 34% to 33% to 32% to 29%, where it has settled for another year. The percentage of homeowners using gas has risen steadily from 58% to 66%. Together, electricity and gas have captured from 92% to 95% of the market, leaving 8%, 8%, 8%, 6%, and 5% of the market for other heating fuels such as oil.

2. **(B)** Both sets are symmetric about 150 and so have the same mean. Both sets have the range 170–130 = 40. Set *A* is much more spread out than set *B*, and so set *A* has the greater variance.

3. **(E)** The empirical rule applies to bell-shaped data like those found in set *B*, not in set *A*. For bell-shaped data, 95% of the values fall within two standard deviations of the mean, and 99.7% within three. However, in the histogram for set *B* one sees that 95% of the data are not between 140 and 160 and 99.7% are not between 135 and 165. Thus the standard deviation for set *B* must be greater than 5. The standard deviation for set *A* is even larger, and so it too must be greater than 5.

4. **(A)** Both sets have 20 elements. The ranges, 76–37 = 39 and 86–47 = 39, are equal. Brand A clearly has the larger mean and median, and with its skewness it also has the larger variance.

5. **(A)** Both ranges are 30. January has an interquartile range of 25, while July has an interquartile range of 20 (but to compare the interquartile ranges, you need only visually compare the lengths of the boxes). The median temperatures for January and July are 20 and 80, respectively.

6. **(B)** The range is the distance between the tips of the two whiskers, and so all three ranges are equal. The interquartile range is the length of the box, and so these three values are not all equal. The median of the first distribution is equal to Q_3 of the second, while the median of the third distribution is equal to Q_1 of the second, and so the difference of the medians is $Q_3–Q_1$, which is the interquartile range of the second.

7. **(B)** In the first distribution the median is far to the right, and so the scores are concentrated there, while both the lower whisker and between the median and first quartile are spread out; thus it is skewed to the left. Similarly the third distribution is skewed to the right. Only the middle distribution looks symmetric around its median. There is no such thing as being skewed to both sides.

8. **(D)** The two sets of data have different means, but they have identical shapes and thus the same variability, including both range and variance (and standard deviation). Note that adding 10 to each score in the left set results in the right set. Thus the means differ by 10, but the measures of variability remain the same.