

AP Statistics – Problem Set #1

There are 10 multiple choice and 1 “FRAPPY” on this problem set. Do all of your work within this packet. Circle the letter corresponding to the best answer for the multiple choice questions (worth 1 point each). Answer the AP Free Response question completely, but concisely (worth 4 points).

1. A particularly common question in the study of wildlife behavior involves observing contests between “residents” of a particular area and “intruders.” In each contest, the “residents” either win or lose the encounter (assuming there are no ties). Observers might record several variables, listed below. Which of these variables is categorical?
- (A) The duration of the contest (in seconds)
 - (B) The number of animals involved in the contest
 - (C) Whether the “residents” win or lose
 - (D) The total number of contests won by the “residents”
 - (E) None of these

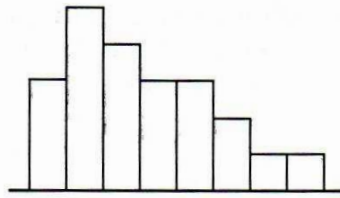
Use the following to answer questions 2 and 3:

During the early part of the 1994 baseball season, many sports fans and baseball players noticed that the number of home runs being hit seemed to be unusually large. Below are the team-by-team statistics on home runs hit through Friday, June 3, 1994 (from the *Columbus Dispatch* Sports Section, Sunday, June 5, 1994), in the form of separate stemplots for the number of home runs by American and National League teams.

American League	National League
2	2 9
3 5	3 1
4 0 3 9	4 2 6 7 8 8
5 1 4 7 8 8	5 3 5 5 5
6 4 8 8	6 3 3 7
7 5 7	7

2. The median for the number of home runs for the American League is:
- (A) lower than that for the National League teams.
 - (B) 45.
 - (C) 50.
 - (D) 57.
 - (E) 57.5.
3. Which of the following is a correct statement?
- (A) The American League plot is reasonably symmetric.
 - (B) The National League plot is slightly skewed to the left.
 - (C) The median number of home runs hit by American League teams was higher than that of National League teams.
 - (D) The 29 home runs in the National League is not considered an outlier.
 - (E) All of the above.

4. Which of the following is more likely to be true of this distribution?



- (A) Mean = 3 Median = 3 Mode = 3
- (B) Mean = 3.5 Median = 4 Mode = 3
- (C) Mean = 4 Median = 3.5 Mode = 3
- (D) Mean = 3.5 Median = 3.5 Mode = 5
- (E) Mean = 3 Median = 2 Mode = 5

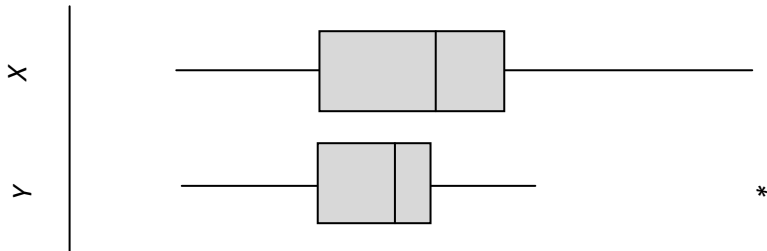
5. A random sample of 25 birth weights (in ounces) is taken yielding the following summary statistics:

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Birthwt	25	129.40	129.00	128.35	17.41	3.48
Variable	Minimum	Maximum	Q1	Q3		
Birthwt	96.00	187.00	119.50	135.50		

What can be said about the number of outliers for this data set?

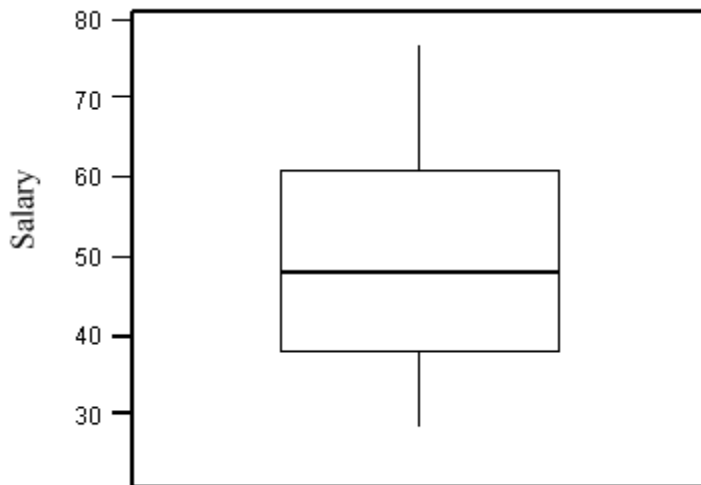
- (A) 0
- (B) At least 1
- (C) No more than 1
- (D) At least 2
- (E) No more than 2

6. The boxplots below summarize two data sets, X and Y . Which of the following MUST be true?



- (A) Set X and set Y have the same number of data points.
- (B) The box of set X contains more data points than the box of set Y .
- (C) The data in set X have a larger range than the data in set Y .
- (D) About 50% of the values in set X are greater than about 75% of the values in set Y .
- (E) The median of set X is less than the median of set Y .

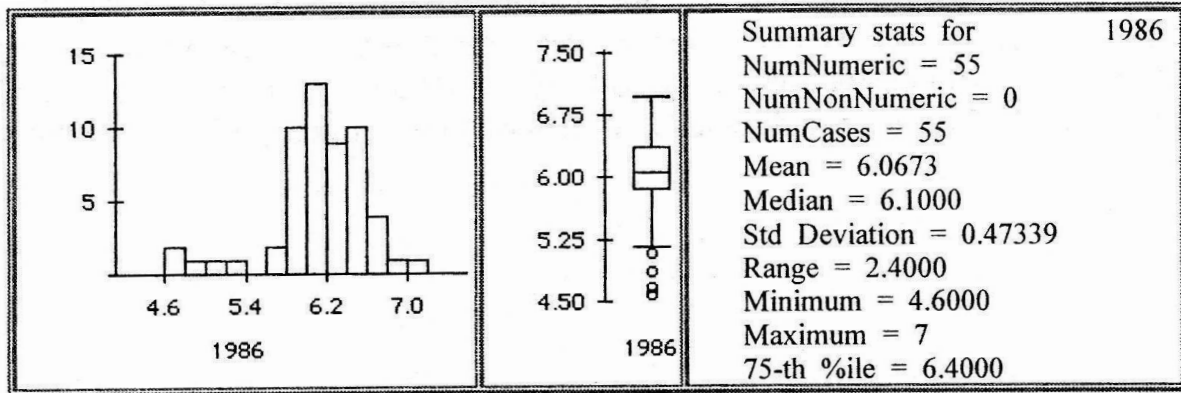
7. Which of the following is likely to have a mean that is smaller than the median?
- (A) the number of candy bars sold in a movie theater on a randomly selected day
 - (B) the scores of students (out of 100 points) on a very easy exam in which most score perfectly, but a few do very poorly
 - (C) the prices of homes in a large city
 - (D) the scores of students (out of 100 points) on a very difficult exam on which most score poorly, but a few do very well
 - (E) the salaries of all National Football League players
8. A sample of 99 distances has a mean of 24 feet and a median of 24.5 feet. Unfortunately, it has just been discovered that an observation which was erroneously recorded as “30” actually had a value of “35”. If we make this correction to the data, then:
- (A) The mean remains the same, but the median is increased.
 - (B) The mean and median remain the same.
 - (C) The median remains the same, but the mean is increased.
 - (D) The mean and median are both increased.
 - (E) We do not know how the mean and median are affected without further calculations, but the variance is increased.
9. A sample was taken of the salaries of 20 employees of a large company. The following is a boxplot of the salaries (in thousands of dollars) for this year.



Based on this boxplot, which of the following statements is true?

- (A) The maximum salary is between \$60,000 and \$70,000.
- (B) The minimum salary is \$20,000.
- (C) The interquartile range is about \$20,000.
- (D) The data is skewed to the lower end.
- (E) The median salary is about \$40,000.

10. Consider the following output analyzing pH values of some 1986 data on precipitation events.



Which of the following is NOT correct?

- (A) The 25th percentile is about 5.9.
- (B) Some outliers appear to be present below a pH of 5.4.
- (C) About 95% of the observations have pH values in the approximate range 6 ± 1 .
- (D) About 10% of the values are in the range 5.8 to 6.0.
- (E) About 75% of the values are less than 6.4.