Use the following to answer question 1.

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.

<table>
<thead>
<tr>
<th>American League</th>
<th>National League</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3 5</td>
<td>3 1</td>
</tr>
<tr>
<td>4 0 3 9</td>
<td>4 2 6 7 8 8</td>
</tr>
<tr>
<td>5 1 4 7 8 8</td>
<td>5 3 5 5 5</td>
</tr>
<tr>
<td>6 4 8 8</td>
<td>6 3 3 7</td>
</tr>
<tr>
<td>7 5 7</td>
<td>7</td>
</tr>
</tbody>
</table>

1. The median for the number of home runs for the American League teams is
   A) lower than that for the National League teams.
   B) 45.
   C) 50.
   D) 57.
   E) 57.5.
Use the following to answer question 2.

The timeplot below gives the number of burglaries committed each month for a city in Ohio. The plot is for the three-year period January 1987–December 1989.

2. Which of the following is a true statement?
   A) The number of burglaries in each month of 1988 was lower than the number of burglaries in each month of 1989.
   B) The median number of burglaries for a month in 1988 was a little over 25.
   C) The total number of burglaries in 1989 was higher than in 1988.
   D) The maximum number of burglaries is bi-modal.
   E) None of the above.

3. A set of data has a median that is much larger than the mean. Which of the following statements is most consistent with this information?
   A) A stemplot of the data is assymetrical.
   B) A stemplot of the data is skewed left.
   C) A stemplot of the data is skewed right.
   D) The data set must be so large that it would be better to draw a histogram than a stemplot.
   E) A stemplot of the data is symmetric.
4. This is a standard deviation contest. Which of the following sets of four numbers has the largest possible standard deviation?
   A) 7, 8, 9, 10.
   B) 1, 2, 9, 10.
   C) 0, 0, 10, 10.
   D) 0, 1, 2, 3.
   E) 5, 5, 5, 5

5. Items produced by a manufacturing process are supposed to weigh 90 grams. However, the manufacturing process is such that there is variability in the items produced and they do not all weigh exactly 90 grams. The distribution of weights can be approximated by a normal distribution with a mean of 90 grams and a standard deviation of 1 gram. Using the 68–95–99.7 rule, what percentage of the items will either weigh less than 87 grams or more than 93 grams?
   A) 0.3%.
   B) 3%.
   C) 6%.
   D) 94%.
   E) 99.7%.

6. Increasing the frequency of observations in the tails of a distribution will
   A) not affect the standard deviation as long as the increases are balanced on either side of the mean.
   B) not affect the standard deviation under any circumstances.
   C) increase the standard deviation.
   D) decrease the standard deviation.
   E) skew the standard deviation.

7. The time it takes for students to complete a standardized exam is approximately normal with a mean of 70 minutes and a standard deviation of 10 minutes. Using the 68–95–99.7 rule, what percentage of students will complete the exam in under an hour?
   A) 68%.
   B) 47.5%.
   C) 32%.
   D) 16%.
   E) 5%.
Use the following to answer question 8.

The distribution of actual weights of 8.0-ounce chocolate bars produced by a certain machine is normal with a mean of 8.1 ounces and a standard deviation of 0.1 ounces.

8. The proportion of chocolate bars weighing between 8.2 and 8.3 ounces is
   A) 0.819.
   B) 0.636.
   C) 0.477.
   D) 0.136.
   E) 0.022.

Use the following to answer question 9.

A researcher measures the height (in feet) and volume of usable lumber (in cubic feet) of 32 cherry trees. The goal is to determine if the volume of a tree's usable lumber can be estimated from the height of the tree. The results are plotted below.

9. In the study above, the response variable is
   A) number of trees.
   B) volume.
   C) height or volume; it doesn't matter which is considered the response variable.
   D) neither height nor volume; the measuring instrument used to measure height is the response variable.
   E) height.
Use the following to answer question 10.

I wish to determine the correlation between the height (in inches) and weight (in pounds) of 21-year-old males. To do this, I measure the height and weight of two 21-year-old men. The measured values are

<table>
<thead>
<tr>
<th></th>
<th>Male #1</th>
<th>Male #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td>Weight</td>
<td>160</td>
<td>200</td>
</tr>
</tbody>
</table>

10. Referring to the information above, which of the following units would the correlation coefficient $r$ have?
   A) Inches.
   B) Pounds.
   C) Pounds per inch.
   D) None, because $r$ has no units.
   E) Inches-pounds.
11. The profits (in multiples of $100,000) versus the sales (in multiples of $100,000) for a number of companies are plotted below. The correlation between profits and sales is 0.814. Suppose we removed the point that is circled from the data represented in the plot. The correlation between profits and sales would then be

A) 0.814.  
B) significantly larger than 0.814.  
C) significantly smaller than 0.814.  
D) slightly larger than 0.814.  
E) slightly smaller than 0.814.

12. A researcher wishes to study how the average weight $Y$ (in kilograms) of children changes during the first year of life. He plots these averages versus the children's age $X$ (in months) and decides to fit a least-squares regression line to the data with $X$ as the explanatory variable and $Y$ as the response variable. He computes the following quantities.

$r = \text{correlation between } X \text{ and } Y = 0.9$

$\bar{X} = \text{mean of the values of } X = 6.5$

$\bar{Y} = \text{mean of the values of } Y = 6.6$

$s = \text{standard deviation of the values of } X = 3.6$

$s_m = \text{standard deviation of the values of } Y = 1.2$

The slope of the least-squares line is

A) 0.30.  
B) 0.88.  
C) 1.01.  
D) 2.7.  
E) 3.0.
Use the following to answer question 13.

A television station is interested in predicting whether voters in its viewing area are in favor of federal funding for abortions. It asks its viewers to phone in and indicate whether they support/are in favor of or are opposed to this policy. Of the 2241 viewers who phoned in, 1574 (70.24%) were opposed to federal funding for abortions.

13. Referring to the information above, the viewers who phoned in are
A) a voluntary response sample.
B) a convenience sample.
C) a probability sample.
D) a population.
E) a simple random sample.

14. A simple random sample of 1200 adult Americans is selected, and each person is asked the following question:
In light of the huge national deficit, should the government at this time spend additional money to establish a national system of health insurance?
Only 39% of those responding answered yes. This survey
A) is reasonably accurate since it used a large, simple random sample.
B) needs to be larger since only about 24 people were drawn from each state.
C) probably understates the percentage of people that favor a system of national health insurance.
D) is very inaccurate, but neither understates nor overstates the percentage of people that favor a system of national health insurance. Since simple random sampling was used, it is unbiased.
E) probably overstates the percentage of people that favor a system of national health insurance.

Use the following to answer question 15.

A study of human development showed two types of movies to groups of children. Crackers were available in a bowl, and the investigators compared the number of crackers eaten by children watching the different kinds of movies. One kind of movie was shown at 8 a.m. (right after the children had breakfast) and another at 11 a.m. (right before the children had lunch). It was found that during the movie shown at 11 a.m., more crackers were eaten than during the movie shown at 8 a.m.. The investigators concluded that the different types of movies had an effect on appetite.
15. The results cannot be trusted because
   A) the study was not double-blind. Neither the investigators nor the children should have
      been aware of which movie was being shown.
   B) children are usually too sleepy early in the morning to watch movies.
   C) the investigators should have used several bowls, with crackers randomly placed in
      each.
   D) the time the movie was shown is a confounding variable.
   E) the investigators were biased. They knew beforehand what they hoped the study would
      show.

Use the following to answer question 16.

To simulate a toss of a coin we let the digits 0, 1, 2, 3, and 4 correspond to a head and the
digits 5, 6, 7, 8, and 9 correspond to a tail. Consider the following game: We are going to
 toss the coin until we either get a head or we get two tails in a row, whichever comes first. If
 it takes us one toss to get the head we win $2, if it takes us two tosses we win $1, and if we
 get two tails in a row we win nothing. Use the following sequence of random digits:
12975 13258 45144

16. The estimated number of tosses in a single trial of the game is
   A) 2.0.
   B) 15/9.
   C) 15/11.
   D) 11/7.
   E) 7/11.

17. A game consists of drawing three cards at random from a deck of playing cards. You win $3
    for each red card that is drawn. It costs $2 to play. For one play of this game, the sample
    space $S$ for the net amount you win (after deducting the cost of play) is
    A) $S = \{0, 1, 2, 3\}$
    B) $S = \{-6, -3, 0, 6\}$
    C) $S = \{-2, 1, 4, 7\}$
    D) $S = \{-2, 3, 6, 9\}$
    E) $S = \{0, 3, 6, 9\}$
Use the following to answer questions 18-19.

Ignoring twins and other multiple births, assume that babies' births at a hospital represent independent events, with the probability that a baby is a boy and the probability that a baby is a girl both being equal to 0.5.

18. Referring to the information above, the probability that the next three babies are all of the same sex is
   A) 1.
   B) 0.125.
   C) 0.250.
   D) 0.333.
   E) 0.500.

19. Referring to the information above, the events
   \( A = \) the next two babies are boys
   \( B = \) at least one of the next two babies is a boy
   are
   A) disjoint.
   B) complements.
   C) independent.
   D) dependent.
   E) none of the above.

Use the following to answer question 20.

An event \( A \) will occur with probability 0.5. An event \( B \) will occur with probability 0.6. The probability that both \( A \) and \( B \) will occur is 0.1.

20. Referring to the information above, the conditional probability of \( A \) given \( B \)
   A) is 0.3.
   B) is 0.2.
   C) is 1/6.
   D) is 0.1.
   E) cannot be determined from the information given.
Use the following to answer question 21.

The probability density curve of a random variable $X$ is given in the figure below.

21. Referring to the information above, the probability that $X = 1.5$ is
   A) 0.
   B) 1/4.
   C) 1/3.
   D) 1/2.
   E) 3/4.
22. Consider the following probability histogram for a discrete random variable $X$.

![Probability Histogram]

This probability histogram corresponds to which of the following probability distributions for $X$?

A)  

<table>
<thead>
<tr>
<th>$X$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X)$</td>
<td>0.06</td>
<td>0.25</td>
<td>0.38</td>
<td>0.25</td>
<td>0.06</td>
</tr>
</tbody>
</table>

B)  

<table>
<thead>
<tr>
<th>$X$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X)$</td>
<td>0.10</td>
<td>0.25</td>
<td>0.30</td>
<td>0.20</td>
<td>0.15</td>
</tr>
</tbody>
</table>

C)  

<table>
<thead>
<tr>
<th>$X$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X)$</td>
<td>0.10</td>
<td>0.25</td>
<td>0.30</td>
<td>0.25</td>
<td>0.10</td>
</tr>
</tbody>
</table>

D)  

<table>
<thead>
<tr>
<th>$X$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X)$</td>
<td>0.10</td>
<td>0.35</td>
<td>0.65</td>
<td>0.85</td>
<td>1.0</td>
</tr>
</tbody>
</table>

E) None of the above.

Use the following to answer question 23.

Suppose there are three balls in a box. On one of the balls is the number 1, on another is the number 2, and on the third is the number 3. You select two balls at random and without replacement from the box and note the two numbers observed. The sample space $S$ consists of the three equally likely outcomes $\{(1, 2), (1, 3), (2, 3)\}$. Let $X$ be the sum of the numbers on the two balls selected.
23. Referring to the information above, the variance of $X$ is
   A) $1/3$.
   B) $2/3$.
   C) $0.816$.
   D) $1$.
   E) $4$.

Use the following to answer question 24.

A psychologist studied the number of puzzles subjects were able to solve in a five-minute period while listening to soothing music. Let $X$ be the number of puzzles completed successfully by a subject. The psychologist found that $X$ had the following probability distribution.

<table>
<thead>
<tr>
<th>Value of $X$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.2</td>
<td>0.4</td>
<td>0.3</td>
<td>0.1</td>
</tr>
</tbody>
</table>

24. Referring to the information above, if three subjects solve puzzles for five minutes each and the numbers of puzzles solved by the subjects are independent of each other, then the mean of the total number of puzzles solved by the three subjects is
   A) $1.8$.
   B) $2.3$.
   C) $2.5$.
   D) $6.9$.
   E) $7.5$.

25. I roll a fair die and count the number of spots on the upward face. A fair die is one for which each of the outcomes 1, 2, 3, 4, 5, and 6 are equally likely. According to the law of large numbers
   A) seeing several (four or five) consecutive rolls for which the outcome 1 is observed is impossible in the long run. If such an event did occur, it would mean the die is no longer fair.
   B) after rolling a 1, you will usually roll nearly all the numbers at least once before rolling a 1 again.
   C) in the long run, a 1 will be observed about every sixth roll and certainly at least once in every 8 or 9 rolls.
   D) a histogram of the results of a large number of rolls will show 6 bars of equal height.
   E) none of the above are true.
Answer Key - 1st Semester Final Exam

1. E
2. C
3. B
4. C
5. A
6. C
7. D
8. D
9. B
10. D
11. C
12. A
13. A
14. C
15. D
16. C
17. C
18. C
19. D
20. C
21. A
22. B
23. B
24. D
25. E