

AP Stats 1-6 Review**Multiple Choice**

Identify the choice that best completes the statement or answers the question.

- You measure the age, marital status and earned income of an SRS of 1463 women. The number and type of variables you have measured is
 - 14563.
 - four; two categorical and two quantitative.
 - four; one categorical and three quantitative.
 - three; two categorical and one quantitative.
 - three; one categorical and two quantitative.
- You open a package of plain M & M candies and count how many there are of each color. The distribution of the variable "candy color" is:
 - The colors: Red, Orange, Green, Yellow, Brown, and Blue.
 - The total number of candies in the package.
 - Six—the number of different colors the are in the package.
 - The six different colors and how many there are of each.
 - Since "color" is a categorical variable, it doesn't have a distribution.
- In a study of the link between high blood pressure and cardiovascular disease, a group of white males aged 35 to 64 was followed for 5 years. At the beginning of the study, each man had his blood pressure measured and it was classified as either "low" systolic blood pressure (less than 140 mm Hg) or "high" blood pressure (140 mm Hg or higher). The following table gives the number of men in each blood pressure category and the number of deaths from cardiovascular disease during the 5-year period.

Blood pressure	Deaths	Total
Low	10	2000
High	5	3500

Based on these data, which of the following statements is correct?

- These data are consistent with the idea that there is a link between high blood pressure and death from cardiovascular disease.
- The mortality rate (proportion of deaths) for men with high blood pressure is 5 times that of men with low blood pressure.
- These data probably understate the link between high blood pressure and death from cardiovascular disease, because men will tend to understate their true blood pressure.
- Although there were more deaths in the high blood pressure group, this is expected, because there were 1500 more men in that group.
- All of the above.

Scenario 1-1

A review of voter registration records in a small town yielded the following table of the number of males and females registered as Democrat, Republican, or some other affiliation.

	Male	Female
Democrat	300	600
Republican	500	300
Other	200	100

4. Use Scenario 1-1. The proportion of males that are registered as Democrats is
 - A. 300
 - B. 30
 - C. 0.33
 - D. 0.30
 - E. 0.15

5. Use Scenario 1-1 Your percentage from question number 12 is part of
 - A. The marginal distribution of political party registration.
 - B. The marginal distribution of gender.
 - C. The conditional distribution of gender among Democrats.
 - D. The conditional distribution of political party registration among males.
 - E. The conditional distribution of males within gender.

6. Use Scenario 1-1. The proportion of registered Democrats that are male is
 - A. 300
 - B. 33
 - C. 0.33
 - D. 0.30
 - E. 0.15

7. Use Scenario 1-1. Your percentage from question number 14 is part of
 - A. The marginal distribution of political party registration.
 - B. The marginal distribution of gender.
 - C. The conditional distribution of gender among Democrats.
 - D. The conditional distribution of political party registration among males.
 - E. The conditional distribution of males within gender.

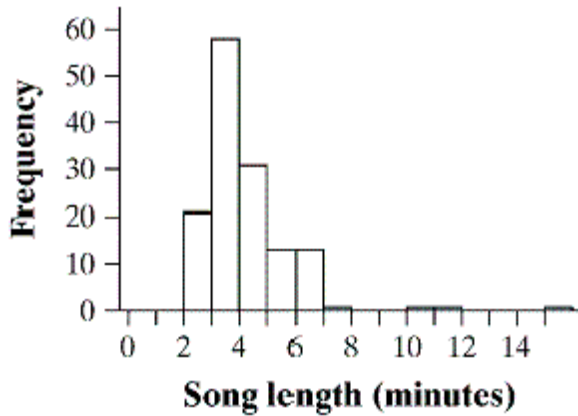
Scenario 1-2

Below is a two-way table summarizing the number of cylinders in selected car models manufactured in six different countries in the 1990's.

	Number of cylinders				Total
	4	5	6	8	
France	0	0	1	0	1
Germany	4	1	0	0	5
Italy	1	0	0	0	1
Japan	6	0	1	0	7
Sweden	1	0	1	0	2
U.S.A.	7	0	7	8	22
Total	19	1	10	8	38

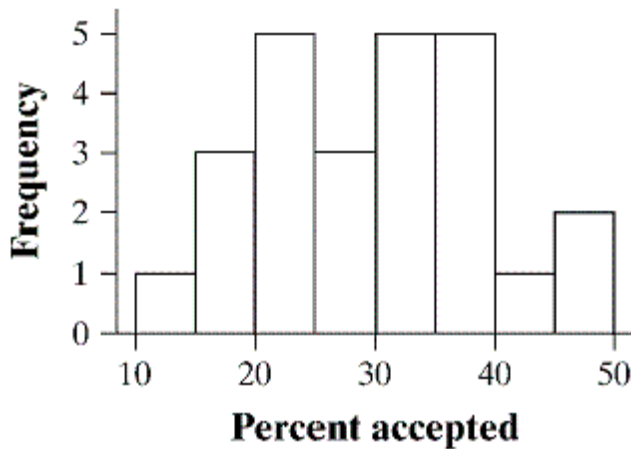
8. Use Scenario 1-2. The percentage of all cars listed in the table with 4-cylinder engines is
 - A. 19%.
 - B. 21%.
 - C. 50%.
 - D. 80%.
 - E. 91%.

9. The histogram below show the length (in minutes) of 140 songs recorded by the band Wilco.



Which of the following descriptions best fits this distribution?

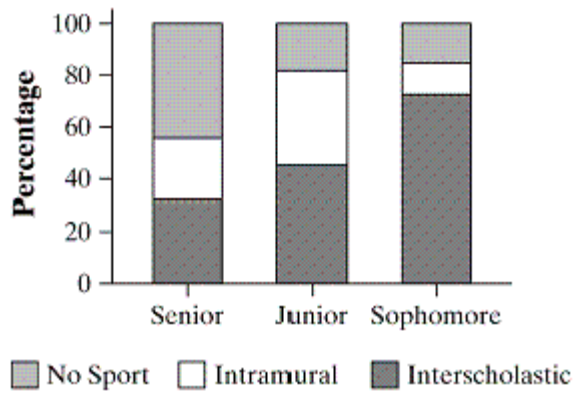
- A. Skewed right, centered at about 8, with several high outliers.
 - B. Skewed left, centered at about 8, with several high outliers.
 - C. Skewed right, centered at about 4.5, with several high outliers.
 - D. Skewed left, centered at about 4.5, with several high outliers.
 - E. Skewed left, centered at about 3.5, with several high outliers.
10. The following histogram represents the distribution of acceptance rates (percent accepted) among 25 business schools in 1997.



What percent of the schools have an acceptance rate of under 20%?

- A. 3%
- B. 4%
- C. 12%
- D. 16%
- E. 24%

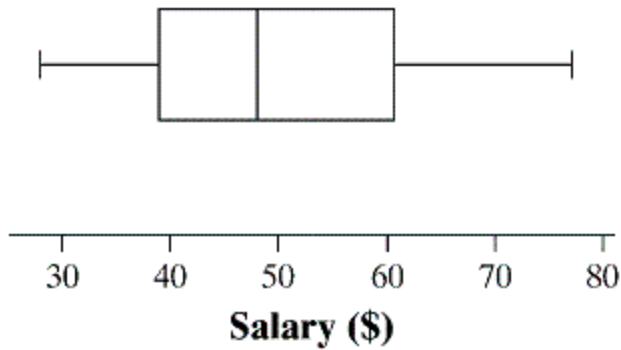
11. A random sample of 100 students in grades 10 through 12 were sampled and asked their year in school and whether they were involved in interscholastic sports, intramural sports, or no sports. The results are summarized in the segmented bar graph below.



- Based on this graph, which of the following statements is true?
- A. More seniors are involved in interscholastic sports than sophomores.
 - B. There is no association between year in school and whether students are involved in sports.
 - C. There were more seniors in the sample than juniors.
 - D. Juniors have the highest percentage participation in intramurals.
 - E. Less than half the seniors are involved in either interscholastic or intramural sports.

Scenario 1-5

A sample was taken of the salaries of 20 employees of a large company. The following boxplot shows the salaries (in thousands of dollars) for this year.



- 12. Use Scenario 1-5. Based on the 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 range of the middle half of the salaries is about \$20,000.
 - D. The median salary is about \$40,000.
 - E. 25% of the employees make more than \$70,000.
- 13. Use Scenario 1-5. Based on the boxplot, the five-number summary is
 - A. 28, 39, 48, 60.5, 77.
 - B. 28, 41, 48, 58, 77.
 - C. 28, 39, 51, 58, 77.
 - D. 28, 41, 51, 60.5, 77.
 - E. 26, 39, 48, 60.5, 81.

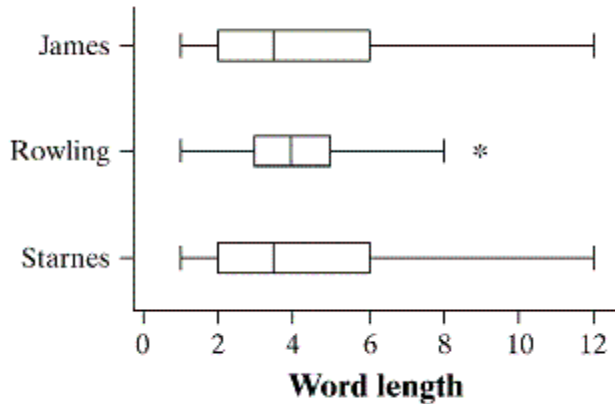
- 14. A lobster fisherman is keeping track of the productivity of a set of traps he has placed in a favorite location. Below are the numbers of lobsters in these traps over the course of 12 different hauls.

0 3 3 3 4 5 5 6 7 7 12 14

According to the 1.5 x IQR rule, which values in the above distribution are outliers?

- A. 0 only
- B. 14 only
- C. 12 and 14
- D. 0 and 14
- E. 0, 12, and 14

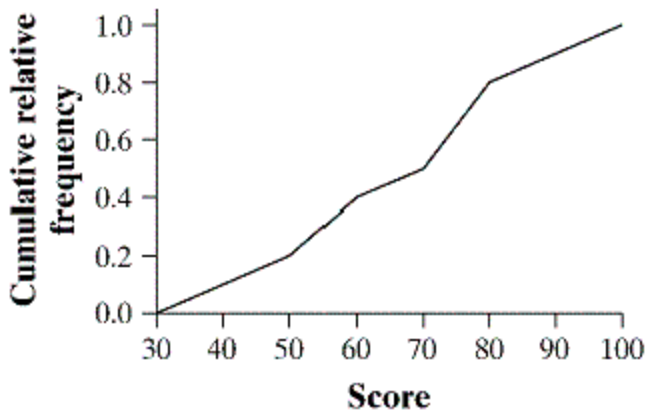
15. Different writers have different styles. One way to quantify this difference is to compare the distribution of word lengths in their work. Below are parallel boxplots describing the distributions of word lengths for the first 60 words in Henry James's *The Turn of the Screw*, J.K. Rowling's *Harry Potter and the Chamber of Secrets*, and Chapter 1 of your statistics textbook (labeled "Starnes" below).



Based on the graphs, which one of the following statements must be true?

- A. Dot plots of the distributions of James's word lengths and Starnes's word lengths are identical.
 - B. The longest word in the distribution of Rowling's word lengths is short than 25% of the word in the "James" distribution.
 - C. The range of Rowling's word lengths is smaller than the interquartile range of Starnes's word lengths.
 - D. The median word length for Rowling is longer than for either Starnes or James.
 - E. 75% of the words in Rowling's distribution are longer than the median word length in Starnes's distribution.
16. For the density curve below, which of the following is true?
-
- A. The median is 0.5.
 - B. The median is larger than 0.5.
 - C. The density curve is skewed right.
 - D. The density curve is Normal.
 - E. The density curve is symmetric.
17. You can roughly locate the mean of a density curve by eye because it is
- A. the point at which the curve would balance if made of solid material.
 - B. the point that divides the area under the curve into two equal parts.
 - C. the point at which the curve reaches its peak.
 - D. the point where the curvature changes direction.
 - E. the point at which the height of the graph is equal to 1.

18. You can roughly locate the median of a density curve by eye because it is
- the point at which the curve would balance if made of solid material.
 - the point that divides the area under the curve into two equal parts.
 - the point at which the curve reaches its peak.
 - the point where the curvature changes direction.
 - the point at which the height of the graph is equal to 1.
19. You are told that your score on an exam is at the 85 percentile of the distribution of scores. This means that
- Your score was lower than approximately 85% of the people who took this exam.
 - Your score was higher than approximately 85% of the people who took this exam.
 - You answered 85% of the questions correctly.
 - If you took this test (or one like it) again, you would score as well as you did this time 85% of the time.
 - 85% of the people who took this test earned the same score you did.
20. Here is a list of exam scores for Mr. Williams’s calculus class:
60 61 61 65 72 75 75 78 81 81 85 89 91 98
What is the percentile of the person whose score was 85?
- 15%
 - 21%
 - 29%
 - 71%
 - 85%
21. An ecologist studying starfish populations collected starfish of the species *Pisaster* was interested in the distribution of sizes of starfish on a certain shoreline. One measure of size is “arm length.” Below is a cumulative frequency distribution for the arm length of 102 *Pisaster* individuals.



The median and interquartile range of this distribution are approximately:

- Median is 15.2; Intequartile range is 12.5 to 16.8
- Median is 13; Interquartile range is 13 to 16.1
- Median is 13; Interquartile range is 3.1
- Median is 13; Intequartile range is 4.3
- Median is 15.2; Intequartile range is 4.3

22. The mean number of days that the midge *Chaoborus* spends in its larval stage is 14.1 days, with a standard deviation of 2.2 days. This distribution is skewed toward higher values. What is the z -score for an individual midge that spends 12.7 days in its larval stage?
- 1.11
 - 0.64
 - 0.64
 - 0.94
 - None of these, because z -score cannot be used unless the distribution is Normal.

Scenario 2-1

A sample was taken of the salaries of 20 employees of a large company. The following are the salaries (in thousands of dollars) for this year. For convenience, the data are ordered.

28	31	34	35	37	41	42	42	42	47
49	51	52	52	60	61	67	72	75	77

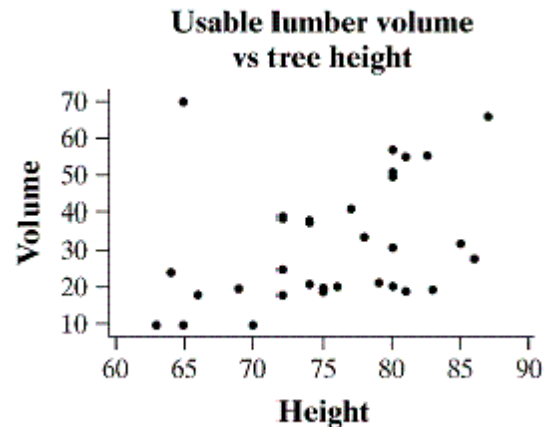
Suppose each employee in the company receives a \$3,000 raise for next year (each employee's salary is increased by \$3,000).

- Use Scenario 2-1. The mean salary for the employees will
 - be unchanged.
 - increase by \$3,000.
 - be multiplied by \$3,000.
 - increase by $\sqrt{\$3,000}$
 - increase by \$150.
- Use Scenario 2-1. The median salary for the employees working for the company will
 - be unchanged.
 - increase by \$3,000.
 - be multiplied by \$3,000.
 - increase by $\sqrt{\$3,000}$
 - increase by \$150.
- Use Scenario 2-1. The standard deviation of the salaries for the employees will
 - be unchanged.
 - increase by \$3,000.
 - be multiplied by \$3,000.
 - increase by $\sqrt{\$3,000}$
 - decrease by \$3,000.
- Use Scenario 2-1. The interquartile range of the salaries for the employees will
 - be unchanged.
 - increase by \$3,000.
 - be multiplied by \$3,000.
 - increase by $\sqrt{\$3,000}$
 - decrease by \$3,000.
- Use Scenario 2-1. The z -scores of the salaries for the employees will
 - be unchanged.
 - increase by \$3,000.
 - be multiplied by \$3,000.
 - increase by $\sqrt{\$3,000}$
 - decrease by \$3,000.
- IQs among undergraduates at Mountain Tech are approximately Normally distributed. The mean undergraduate IQ is 110. About 95% of undergraduates have IQs between 100 and 120. The standard deviation of these IQs is about
 - 5.
 - 10.
 - 15.
 - 20.
 - 25.

29. Let X denote the time taken for a computer link to be made between the terminal in an executive's office and the computer at a remote factory site. It is known that X has a Normal distribution with a mean of 15 seconds and a standard deviation of 3 seconds. On 90% of the occasions the computer link is made in less than
- 11.16 seconds.
 - 15.95 seconds.
 - 18.11 seconds.
 - 18.84 seconds.
 - 19.39 seconds.
30. Entomologist Heinz Kaefer has a colony of bongo spiders in his lab. There are 1000 adult spiders in the colony, and their weights are Normally distributed with mean 11 grams and standard deviation 2 grams. About how many spiders are there in the colony which weigh more than 12 grams?
- 117
 - 160
 - 310
 - 690
 - 840

Scenario 3-1

The height (in feet) and volume (in cubic feet) of usable lumber of 32 cherry trees are measured by a researcher. The goal is to determine if volume of usable lumber can be estimated from the height of a tree.



31. Use Scenario 3-1. In this study, the response variable is
- height of researcher.
 - volume of lumber.
 - height of tree.
 - the measuring instrument used to measure volume.
 - impossible to determine.
32. Use Scenario 3-1. Which of the following statements are supported by the scatterplot?
- There is a positive association between height and volume.
 - There is an outlier in the plot.
 - As the height of a cherry tree increases, the volume of useable lumber it yields increases.
- I only
 - II only
 - III only
 - I and II
 - I, II, and III

33. Use Scenario 3-1. If the data point (65,70) were removed from this study, how would the value of the correlation r change?
- r would be smaller, since there are fewer data points.
 - r would be smaller, because this point falls in the pattern of the rest of the data.
 - r would be larger, since the x and y coordinates are larger than the mean x and mean y , respectively.
 - r would be larger, since this point does not fall in the pattern of the rest of the data.
 - r would not change, since it's value does not depend which variable is used for x and which is used for y .
34. There is a positive correlation between the size of a hospital (measured by number of beds) and the median number of days that patients remain in the hospital. Does this mean that you can shorten a hospital stay by choosing to go to a small hospital?
- No – a negative correlation would allow that conclusion, but this correlation is positive.
 - Yes – the data show that stays are shorter in smaller hospitals.
 - No – the positive correlation is probably explained by the fact that seriously ill people go to large hospitals
 - Yes – the correlation can't just be an accident.
 - Yes – but only if r is very close to 1.
35. Which of the following best describes the correlation r ?
- The average of the products of each of the X and Y values for each point
 - The average of the products of the standardized scores of X and Y for each point.
 - The average of the squared products of the standardized scores of X and Y for each point.
 - The average of the differences between each X value and each Y value.
 - The average perpendicular distance between each data point and the least-squares regression line.
36. Which of the following statements are true about the least-squares regression line?
- The slope is the predicted change in the response variable associated with a unit increase in the explanatory variable.
 - The line always passes through the point (J, M) , the means of the explanatory and response variables, respectively.
 - It is the line that minimizes the sum of the squared residuals.
- I only.
 - II only.
 - III only.
 - I and III only.
 - I, II, and III are all true.
37. Suppose a straight line is fit to data having response variable y and explanatory variable x . Predicting values of y for values of x outside the range of the observed data is called
- contingency.
 - extrapolation.
 - causation.
 - correlation.
 - interpolation.
38. Which of the following statements describes what the standard deviation of residuals for a regression equation can be used for?
- It describes the typical vertical distance between an observed data point and the regression line.
 - It evaluate whether a linear model is appropriate for a set of data.
 - It measures the overall precision of predictions made using the regression equation.
- I only
 - II only
 - III only
 - Both I and II
 - Both I and III

Scenario 4-2

You want to know the opinions of American school teachers about establishing a national test for high school graduation. You obtain a list of the members of the National Education Association (the largest teachers' union) and mail a questionnaire to 2500 teachers chosen at random from this list. In all 1347 teachers return the questionnaire.

39. Use Scenario 4-2. The *sampling frame* is
- A. the 1347 teachers who mail back the questionnaire.
 - B. the 2500 teachers to whom you mailed the questionnaire.
 - C. all members of the National Education Association.
 - D. all American school teachers.
 - E. all American school students.
40. Use Scenario 4-2. The *sample* is
- A. the 1347 teachers who mail back the questionnaire.
 - B. the 2500 teachers to whom you mailed the questionnaire.
 - C. all members of the National Education Association.
 - D. all American school teachers.
 - E. all American school students.
41. A television station is interested in predicting whether voters in its viewing area are in favor of offshore drilling. It asks its viewers to phone in and indicate whether they support/are in favor of or are opposed to this practice. Of the 2241 viewers who phoned in, 1574 (70%) were opposed to offshore drilling. 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.
42. A simple random sample of size n is defined to be
- A. a sample of size n chosen in such a way that every unit in the population has the same chance of being selected.
 - B. a sample of size n chosen in such a way that every unit in the population has a known nonzero chance of being selected.
 - C. a sample of size n chosen in such a way that every set of n units in the population has an equal chance to be the sample actually selected.
 - D. a sample of size n chosen in such a way that each selection is made independent of every other selection.
 - E. all of the above. They are essentially identical definitions.
43. Simple random sampling
- A. reduces bias resulting from poorly worded questions.
 - B. offsets bias resulting from undercoverage and nonresponse.
 - C. reduces bias resulting from the behavior of the interviewer.
 - D. reduces variability.
 - E. None of the above.

Scenario 4-3

We wish to choose a simple random sample of size three from the following employees of a small company. To do this, we will use the numerical labels attached to the names below.

- | | | |
|--------------|------------|-----------|
| 1. Bechhofer | 4. Kesten | 7. Taylor |
| 2. Brown | 5. Kiefer | 8. Wald |
| 3. Ito | 6. Spitzer | 9. Weiss |

We will also use the following list of random digits, reading the list from left to right, starting at the beginning of the list.

11793 20495 05907 11384 44982 20751 27498 12009 45287 71753 98236 66419 84533

44. Use Scenario 4-3. The simple random sample is
- A. 117.
 - B. Bechhofer, Bechhofer again, and Taylor.
 - C. Bechhofer, Taylor, Weiss.
 - D. Kesten, Kiefer, Taylor.
 - E. Taylor, Weiss, Ito.
45. Use Scenario 4-3. Which of the following statements is true?
- A. If we use another list of random digits to select the sample, we would get the same result as that obtained with the list actually used.
 - B. If we use another list of random digits to select the sample, we would get a completely different sample than that obtained with the list actually used.
 - C. If we use another list of random digits to select the sample, we would get, at most, one name in common with that obtained with the list actually used.
 - D. If we use another list of random digits to select the sample, the result obtained with the list actually used would be just as likely to be selected as any other set of three names.
 - E. If we use another list of random digits to select the sample, the result obtained with the list actually used would be far less likely to be selected than any other set of three names.
46. Use Scenario 4-3. Which of these statements about the table of random digits is true?
- A. Every row must have exactly the same number of 0's and 1's.
 - B. In the entire table, there are exactly the same number of 0's and 1's.
 - C. If you look at 100 consecutive pairs of digits anywhere in the table, exactly 1 pair is 00.
 - D. All of these are true.
 - E. None of these is true.
47. A public opinion poll in Ohio wants to determine whether or not registered voters in the state approve of a measure to ban smoking in all public areas. They select a simple random sample of fifty registered voters from each county in the state and ask whether they approve or disapprove of the measure. This is an example of a
- A. systematic random sample.
 - B. stratified random sample.
 - C. multistage sample.
 - D. simple random sample.
 - E. cluster sample.
48. A marine biologist wants to estimate the mean size of the barnacle *Semibalanus balanoides* on a stretch of rocky shoreline. To do so, he randomly selected twenty 10-cm. square plots and measured the size of every barnacle in each plot. This is an example of
- A. convenience sampling.
 - B. cluster sampling.
 - C. stratified random sampling.
 - D. simple random sampling.
 - E. multistage sampling.

49. A local tax reform group polls the residents of the school district and asks the question, “Do you think the school board should stop spending taxpayers’ money on non-essential arts programs in elementary schools?” The results of this poll are likely to
- Underestimate support for arts programs because of undercoverage.
 - Underestimate support for arts programs because of nonsampling error.
 - Overestimate support for arts programs because of undercoverage.
 - Overestimate support for arts programs because of nonsampling error.
 - Accurately estimate support for arts programs.
50. Just before the presidential election of 1936, the magazine *Literary Digest* predicted—incorrectly, as it turned out—that Alf Landon would defeat Franklin Delano Roosevelt. Landon lost in a landslide. It turned out that the magazine had only polled its own subscribers, plus others from a list of automobile owners and a list of people who had telephone service. All three groups had higher than typical incomes during the Great Depression. This is an example of
- voluntary response bias.
 - bias resulting from question wording.
 - undercoverage.
 - nonresponse.
 - response bias.

Scenario 4-5

In order to assess the effects of exercise on reducing cholesterol, a researcher took a random sample of fifty people from a local gym who exercised regularly and another random sample of fifty people from the surrounding community who did not exercise regularly. They all reported to a clinic to have their cholesterol measured. The subjects were unaware of the purpose of the study, and the technician measuring the cholesterol was not aware of whether or not subjects exercised regularly.

51. Use Scenario 4-5. This is a(n)
- observational study.
 - experiment, but not a double blind experiment.
 - double blind experiment.
 - matched pairs experiment.
 - block design.
52. Use Scenario 4-5. Which of the following best describes the inferences the researcher can make based in his results?
- He can make inferences about cause and effect, but not about the populations from which the samples were taken.
 - He can make inferences about the populations from which the samples were taken, but not about cause and effect.
 - He can make inferences about both cause and effect and the populations from which the samples were taken.
 - He cannot make inferences about either cause and effect or the populations from which the samples were taken.
 - There is not enough information to make judgments about the scope of inference.

Scenario 4-6

Does caffeine improve exam performance? Suppose all students in the 8:30 section of a course are given a "treatment" (two cups of coffee) and all students in the 9:30 section are not permitted to have any caffeine before a mid-term exam.

53. Use Scenario 4-6. Unfortunately, any systematic difference between the two sections on the exam might be due to the fact that the 8:30 and 9:30 classes have different instructors. This is an example of
- A. placebo effect.
 - B. bias.
 - C. confounding.
 - D. observational study.
 - E. stratification.

Scenario 4-7

A farmer wishes to determine which of two brands of baby pig pellets, Kent or Moormans, produces better weight gains. Two of his sows each give birth to litters of 10 pigs on the same day, so he decides to give the baby pigs in litter A only Kent pellets, while the pigs in litter B will get only Moormans pellets. After four weeks, the average weight gain for pigs in litter A is greater than the average weight gain for pigs in litter B.

54. Use Scenario 4-7. If the farmer had fed Kent pellets to an SRS of 5 pigs from litter A and an SRS of 5 pigs from litter B, with the remaining 10 pigs getting Moormans pellets, then he would have been using
- A. a systematic random sample.
 - B. a convenience sample.
 - C. a matched-pairs design.
 - D. a block design.
 - E. a completely randomized design.

Scenario 4-8

Researchers wish to determine if a new experimental medication will reduce the symptoms of allergy sufferers without the side effect of drowsiness. To investigate this question, the researchers randomly assigned 100 adult volunteers who suffer from allergies to two groups. They gave the new medication to the subjects in one group and an existing medication to the subjects in the other group. Forty-four percent of those in the treatment group and 28% of those in the control group reported a significant reduction in their allergy symptoms without any drowsiness.

55. Use Scenario 4-8. The experimental units are the
- A. researchers.
 - B. 100 adult volunteers.
 - C. all the volunteers who reported a significant reduction in their allergy symptoms without any drowsiness.
 - D. all the volunteers who did not report a significant reduction in their allergy symptoms without any drowsiness.
 - E. pills containing the new experimental medication.
56. Use Scenario 4-8. Which of the following best describes the inferences the researchers can make based in his results?
- A. They can make inferences about cause and effect, but not about the populations from which the samples were taken.
 - B. They can make inferences about the populations from which the samples were taken, but not about cause and effect.
 - C. They can make inferences about both cause and effect and the populations from which the samples were taken.
 - D. They cannot make inferences about either cause and effect or the populations from which the samples were taken.
 - E. There is not enough information to make judgments about the scope of inference.

57. A study of elementary school children, ages 6 to 11, finds a high positive correlation between shoe size x and score y on a test of reading comprehension. The observed correlation is most likely due to
- the effect of a lurking variable, such as age.
 - a mistake, since the correlation must be negative.
 - cause and effect (larger shoe size causes higher reading comprehension).
 - "reverse" cause and effect (higher reading comprehension causes larger shoe size).
 - several outliers in the data set.
58. The principle reason for the use of *random assignment* in designing experiments is that it
- distinguishes a treatment effect from the effects of confounding variables.
 - allows double-blinding.
 - reduces sampling variability.
 - creates approximately equal groups for comparison.
 - eliminates the placebo effect.
59. One hundred volunteers who suffer from severe depression are available for a study. Fifty are selected at random and are given a new drug that is thought to be particularly effective in treating severe depression. The other fifty are given an existing drug for treating severe depression. A psychiatrist evaluates the symptoms of all volunteers after four weeks in order to determine if there has been substantial improvement in the severity of the depression. The study would be double blind if
- neither drug had any identifying marks on it.
 - all volunteers were not allowed to see the psychiatrist nor the psychiatrist allowed to see the volunteers during the session in which the psychiatrist evaluated the severity of the depression.
 - neither the volunteers nor the psychiatrist knew which treatment any person had received.
 - the patients were given a placebo.
 - all of the above.
60. An experiment compares the taste of a new spaghetti sauce with the taste of a commercially successful sauce readily available in grocery stores. Each of a number of tasters tastes both sauces (in random order) and says which tastes better. This is called a
- simple random sample.
 - stratified random sample.
 - completely randomized design.
 - matched pairs design.
 - double-blind design.
61. I toss a penny and observe whether it lands heads up or tails up. Suppose the penny is fair, i.e., the probability of heads is $1/2$ and the probability of tails is $1/2$. This means that
- every occurrence of a head must be balanced by a tail in one of the next two or three tosses.
 - if I flip the coin 10 times, it would be almost impossible to obtain 7 heads and 3 tails.
 - if I flip the coin many, many times the proportion of heads will be approximately $1/2$, and this proportion will tend to get closer and closer to $1/2$ as the number of tosses increases.
 - regardless of the number of flips, half will be heads and half tails.
 - all of the above.
62. When two coins are tossed, the probability of getting two heads is 0.25. This means that
- of every 100 tosses, exactly 25 will have two heads.
 - the odds against two heads are 4 to 1.
 - in the long run, the average number of heads is 0.25.
 - in the long run two heads will occur on 25% of all tosses.
 - if you get two heads on each of the first five tosses of the coins, you are unlikely to get heads the fourth time.

63. You read in a book on poker that the probability of being dealt three of a kind in a five-card poker hand is $1/50$. What does this mean?
- If you deal thousands of poker hands, the fraction of them that contain three of a kind will be very close to $1/50$.
 - If you deal 50 poker hands, then one of them will contain three of a kind.
 - If you deal 10,000 poker hands, then 200 of them will contain three of a kind.
 - A probability of 0.02 is somebody's best guess for a probability of being dealt three of a kind.
 - It doesn't mean anything, because $1/50$ is just a number.
64. You want to use simulation to estimate the probability of getting exactly one head and one tail in two tosses of a fair coin. You assign the digits 0, 1, 2, 3, 4 to heads and 5, 6, 7, 8, 9 to tails. Using the following random digits to execute as many simulations as possible, what is your estimate of the probability?
19226 95034 05756 07118
- $1/20$
 - $1/10$
 - $5/10$
 - $6/10$
 - $2/3$
65. The collection of all possible outcomes of a random phenomenon is called
- a census.
 - the probability.
 - a chance experiment
 - the sample space.
 - the distribution.
66. A basketball player shoots 8 free throws during a game. The sample space for counting the number she makes is
- $S =$ any number between 0 and 1.
 - $S =$ whole numbers 0 to 8.
 - $S =$ whole numbers 1 to 8.
 - $S =$ all sequences of 8 hits or misses, like HMMHHHMH.
 - $S = \{HMMMMMMM, MHMMMMMM, MMHMMMMM, MMMHMMMM, MMMMHMMM, MMMMMHMM, MMMMMMHM, MMMMMMMH\}$
67. An assignment of probabilities must obey which of the following?
- The probability of any event must be a number between 0 and 1, inclusive.
 - The sum of all the probabilities of all outcomes in the sample space must be exactly 1.
 - The probability of an event is the sum of the probabilities of outcomes in the sample space in which the event occurs.
 - All three of the above.
 - A and B only.
68. Event A has probability 0.4. Event B has probability 0.5. If A and B are disjoint, then the probability that both events occur is
- 0.0.
 - 0.1.
 - 0.2.
 - 0.7.
 - 0.9.
69. Event A has probability 0.4. Event B has probability 0.5. If A and B are independent, then the probability that both events occur is
- 0.0.
 - 0.1.
 - 0.2.
 - 0.7.
 - 0.9.

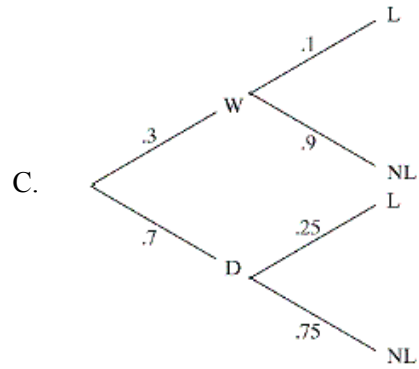
Scenario 5-2

If you draw an M&M candy at random from a bag of the candies, the candy you draw will have one of six colors. The probability of drawing each color depends on the proportion of each color among all candies made. The table below gives the probability that a randomly chosen M&M had each color before blue M & M's replaced tan in 1995.

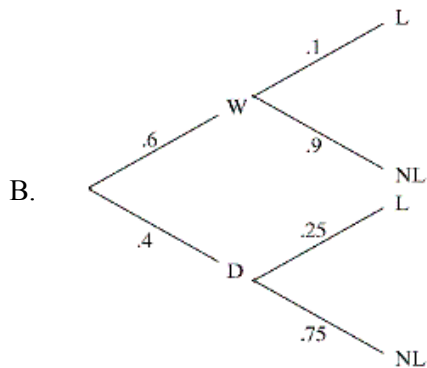
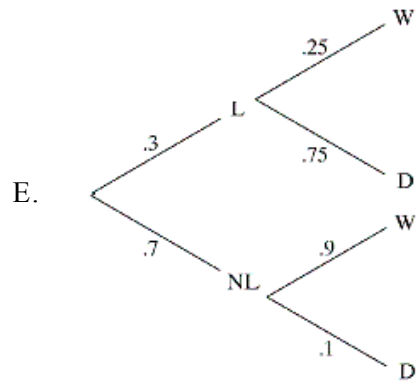
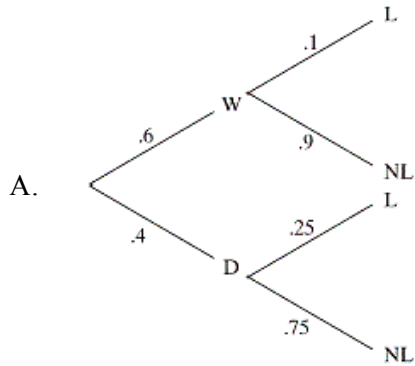
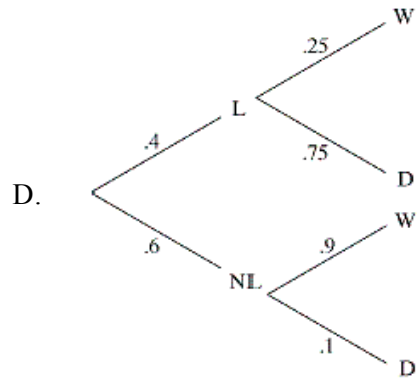
Color	Brown	Red	Yellow	Green	Orange	Tan
Probability	0.3	0.2	?	0.1	0.1	0.1

70. Use Scenario 5-2. The probability of drawing a yellow candy is
- 0.
 - .1.
 - .2.
 - .3.
 - impossible to determine from the information given.
71. Use Scenario 5-2. The probability that you do not draw a red candy is
- .2.
 - .3.
 - .7.
 - .8.
 - impossible to determine from the information given.
72. Use Scenario 5-2. The probability that you draw either a brown or a green candy is
- .1.
 - .3.
 - .4.
 - .6.
 - .7.
73. Event A occurs with probability 0.3. If event A and B are disjoint, then
- $P(B) \leq 0.3$.
 - $P(B) \geq 0.3$.
 - $P(B) \leq 0.7$.
 - $P(B) \geq 0.7$.
 - $P(B) = 0.21$.
74. In a certain town, 60% of the households have broadband internet access, 30% have at least one high-definition television, and 20% have both. The proportion of households that have neither broadband internet nor high-definition television is:
- 0%.
 - 10%.
 - 30%.
 - 80%.
 - 90%.
- Scenario 5-5**
- Suppose we roll two six-sided dice--one red and one green. Let A be the event that the number of spots showing on the red die is three or less and B be the event that the number of spots showing on the green die is three or more.
75. Use Scenario 5-5. The events A and B are
- disjoint.
 - conditional.
 - independent.
 - reciprocals.
 - complementary.
76. Use Scenario 5-5. $P(A \cap B) =$
- 1/6.
 - 1/4.
 - 1/3.
 - 5/6.
 - none of these.
77. Use Scenario 5-5. $P(A \cup B) =$
- 1/6.
 - 1/4.
 - 2/3.
 - 5/6.
 - 1.

78. The card game Euchre uses a deck with 32 cards: Ace, King, Queen, Jack, 10, 9, 8, 7 of each suit. Suppose you choose one card at random from a well-shuffled Euchre deck. What is the probability that the card is a Jack, given that you know it's a face card?
- A. $1/3$
 B. $1/4$
 C. $1/8$
 D. $1/9$
 E. $1/12$



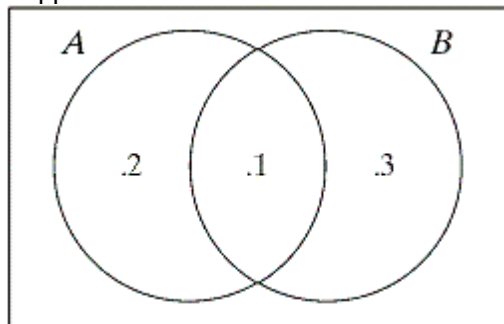
79. Each day, Mr. Bayona chooses a one-digit number from a random number table to decide if he will walk to work or drive that day. The numbers 0 through 3 indicate he will drive, 4 through 9 mean he will walk. If he drives, he has a probability of 0.1 of being late. If he walks, his probability of being late rises to 0.25. Let W = Walk, D = Drive, L = Late, and NL = Not Late. Which of the following tree diagrams summarizes these probabilities?



Scenario 5-10

The Venn diagram below describes the proportion of students who take chemistry and Spanish at Jefferson High School, Where A = Student takes chemistry and B = Students takes Spanish.

Suppose one student is chosen at random.



80. Use Scenario 5-10. Find the value of $P(A \cup B)$ and describe it in words.
- A. 0.1; The probability that the student takes both chemistry and Spanish.
 - B. 0.1; The probability that the student takes either chemistry or Spanish, but not both.
 - C. 0.5; The probability that the student takes either chemistry or Spanish, but not both.
 - D. 0.6; The probability that the student takes either chemistry or Spanish, or both.
 - E. 0.6; The probability that the student takes both chemistry and Spanish.

Scenario 5-13

One hundred high school students were asked if they had a dog, a cat, or both at home. Here are the results.

		Dog?		Total
		No	Yes	
Cat?	No	74	4	78
	Yes	10	12	22
Total		84	16	100

82. Use Scenario 5-13. If a single student is selected at random and you know she has a dog, what is the probability she also has a cat?
- A. 0.04
 - B. 0.12
 - C. 0.22
 - D. 0.25
 - E. 0.75
83. Use Scenario 5-13. If a single student is selected at random, what is the probability associated with the union of the events “has a dog” and “does not have a cat?”
- A. 0.04
 - B. 0.16
 - C. 0.78
 - D. 0.9
 - E. 0.94

84. Use Scenario 5-13. If two students are selected at random, what is the probability that neither of them has a dog or a cat?
- A. 0.37
 B. 0.540
 C. 0.548
 D. 0.655
 E. 0.74
85. Which of the following random variables should be considered continuous?
- A. The time it takes for a randomly chosen woman to run 100 meters
 B. The number of brothers a randomly chosen person has
 C. The number of cars owned by a randomly chosen adult male
 D. The number of orders received by a mail-order company in a randomly chosen week
 E. None of the above
86. Which of the following is *not* a random variable?
- A. The number of heads in ten tosses of a fair coin.
 B. The number of passengers in cars passing through a toll booth.
 C. The age of the driver in cars passing through a toll booth.
 D. The response of randomly-selected people to the question, "Did you eat breakfast this morning?"
 E. The response of randomly-selected people to the question, "How many hours of sleep did you get last night?"
87. A random variable is
- A. a hypothetical list of the possible outcomes of a random phenomenon.
 B. any phenomenon in which outcomes are equally likely.
 C. any number that changes in a predictable way in the long run.
 D. a variable used to represent the outcome of a random phenomenon.
 E. a variable whose value is a numerical outcome associated with a random phenomenon.
88. 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 two balls selected. Which of the following is the correct probability distribution for X ?

(A)

#	Prob
1	1/3
2	1/3
3	1/3

(B)

#	Prob
3	1/3
4	1/3
5	1/3

(C)

#	Prob
1	1/6
2	2/6
3	3/6

(D)

#	Prob
3	1/6
4	2/6
5	3/6

(E)

#	Prob
1	1/4
2	1/4
3	1/4

- A. A
 B. B
 C. C
 D. D
 E. E

Scenario 6-1

Flip a coin four times. If Z = the number of heads in four flips, then the probability distribution of Z is given in the table below.

Z	0	1	2	3	4
$P(Z)$	0.0625	0.2500	0.3750	0.2500	0.0625

89. Use Scenario 6-1. An expression that represents the probability of at least one tail is
- $P(Z \geq 3)$.
 - $P(Z \leq 3)$.
 - $P(Z < 3)$.
 - $P(Z > 3)$.
 - $P(Z \geq 1)$.
90. Use Scenario 6-1. The probability of at least one tail is
- 0.2500.
 - 0.3125.
 - 0.6875.
 - 0.9375.
 - none of these.

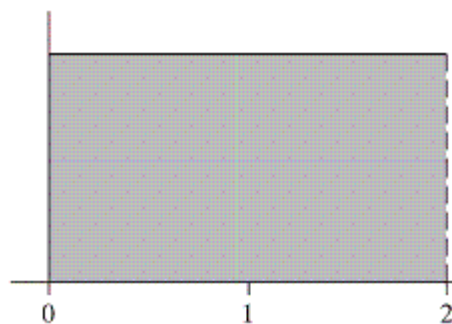
Scenario 6-2

In a particular game, a fair die is tossed. If the number of spots showing is either 4 or 5 you win \$1, if the number of spots showing is 6 you win \$4, and if the number of spots showing is 1, 2, or 3 you win nothing. Let X be the amount that you win.

91. Use Scenario 6-2. Which of the following is the expected value of X ?
- \$0.00
 - \$1.00
 - \$2.50
 - \$4.00
 - \$6.00
92. Use Scenario 6-2. Which of the following is the standard deviation of X ?
- \$1.00
 - \$1.35
 - \$1.41
 - \$1.78
 - \$2.00

Scenario 6-6

The probability distribution of a continuous random variable X is given by the density curve below.



93. Use Scenario 6-6. The probability that X is between 0.5 and 1.5 is
- $1/4$.
 - $1/3$.
 - $1/2$.
 - $3/4$.
 - 1.
94. Use Scenario 6-6. The probability that X is at least 1.5 is
- 0.
 - $1/4$.
 - $1/3$.
 - $1/2$.
 - $3/4$.
95. Use Scenario 6-6. The probability that $X = 1.5$ is
- 0.
 - very small; slightly larger than 0.
 - $1/4$.
 - $1/3$.
 - $1/2$.

96. The weight of written reports produced in a certain department has a Normal distribution with mean 60 g and standard deviation 12 g. The probability that the next report will weigh less than 45 g is
- .1056.
 - .3944.
 - .1042.
 - .0418.
 - .8944.

Scenario 6-8

Let the random variable X represent the profit made on a randomly selected day by a certain store. Assume X is Normal with a mean of \$360 and standard deviation \$50.

97. Use Scenario 6-8. The value of $P(X > \$400)$ is
- 0.2881.
 - 0.8450.
 - 0.7881.
 - 0.2119.
 - 0.1600.
98. Use Scenario 6-8. The probability is approximately 0.6 that on a randomly selected day the store will make less than which of the following amounts?
- \$330.00
 - \$347.40
 - \$361.30
 - \$372.60
 - \$390.00

Scenario 6-10

Your friend Albert has invented a game involving two ten-sided dice. One of the dice has threes, fours, and fives on its faces, the other has sixes, eights, and tens. He won't tell you how many of each number there are on the faces, but he does tell you that if X = rolls of the first die and Y = rolls of the second die, then $\mu_X = 3.6$, $\sigma_X = 0.8$, $\mu_Y = 8.0$, and $\sigma_Y = 0.9$. Let Z = the sum of the two dice when each is rolled once.

99. Use Scenario 6-10. What is the expected value of Z ?
- 1.7
 - 4.4
 - 8.8
 - 8.9
 - 11.6
100. Use Scenario 6-10. What is the standard deviation of Z ?
- 1.20
 - 1.30
 - 1.45
 - 1.70
 - 2.89
101. Use Scenario 6-10. Here's Albert's game: You give him \$10 each time you roll, and he pays you (in dollars) the amount that comes up on the dice. If P = the amount of money you gain each time you roll, the mean and standard deviation of P are:
- $\mu_P = -1.6$; $\sigma_P = 1.45$
 - $\mu_P = 1.6$; $\sigma_P = 1.45$
 - $\mu_P = 1.6$; $\sigma_P = 1.2$
 - $\mu_P = -1.6$; $\sigma_P = -13.8$
 - $\mu_P = 1.6$; $\sigma_P = 13.8$

102. Sulé's job is just a few bus stops away from his house. While it can be faster to take the bus to work, it's more variable, because of variations in traffic. He estimates that the commute time to work by bus is approximately Normally distributed with a mean of 12 minutes and a standard deviation of 4 minutes. The commute time if he walks to work is also approximately Normally distributed with a mean of 16 minutes with a standard deviation of 1 minute. What is the probability that the bus will be faster than walking?
- A. 0.8340
 B. 0.8485
 C. 0.8980
 D. 0.9756
 E. 0.9896
103. To pass the time, a toll booth collector counts the number of cars that pass through his booth until he encounters a driver with red hair. Suppose we define the random variable Y = the number of cars the collector counts until he gets a red-headed driver for the first time. Is Y a geometric random variable?
- A. Yes – all conditions for the geometric setting are met.
 B. No – “red-headed driver” and “non-red-headed driver” are not the same as “success” and “failure”.
 C. No – we can't assume that each “trial” (that is, each car) is independent of previous trials.
 D. No – the number of trials is not fixed.
 E. No – the probability of a driver being red-headed is not the same for each trial.

Scenario 6-12

There are twenty multiple-choice questions on an exam, each having responses a, b, c, or d. Each question is worth five points and only one option per question is correct. Suppose the student *guesses* the answer to each question, and the guesses from question to question are independent.

104. Use Scenario 6-12. The distribution of X = the number of questions the student will get correct, is
- A. binomial with parameters $n = 5$ and $p = 0.2$.
 B. binomial with parameters $n = 20$ and $p = 0.25$.
 C. binomial with parameters $n = 5$ and $p = 0.25$.
 D. binomial with parameters $n = 4$ and $p = 0.25$.
 E. none of these.
105. Use Scenario 6-12. Which of the following expresses the probability that the student gets no questions correct?
- A. $(0.25)^{20}$
 B. $(0.75)^{20}$
 C. $\binom{20}{1}(0.25)(0.75)^{19}$
 D. $\binom{5}{1}(0.25)(0.75)^4$
 E. $\binom{5}{1}(0.25)^4(0.75)$

Scenario 6-14

A worn out bottling machine does not properly apply caps to 5% of the bottles it fills.

106. Use Scenario 6-14. If you randomly select 20 bottles from those produced by this machine, what is the approximate probability that exactly 2 caps have been improperly applied?
- A. 0.0002
 B. 0.19
 C. 0.74
 D. 0.81
 E. 0.92

107. Use Scenario 6-14. If you randomly select 20 bottles from those produced by this machine, what is the approximate probability that between 2 and 6 (inclusive) caps have been improperly applied?
- A. 0.19
 - B. 0.26
 - C. 0.38
 - D. 0.74
 - E. 0.92
108. Use Scenario 6-14. In a production run of 800 bottles, what is the expected value for the number of bottles with improperly applied caps?
- A. 4
 - B. 8
 - C. 40
 - D. 50
 - E. 80
109. Use Scenario 6-14. In a production run of 800 bottles, what is the standard deviation for the number of bottles with improperly applied caps?
- A. 1.38
 - B. 6.16
 - C. 6.32
 - D. 6.89
 - E. 8.72
110. A college basketball player makes 80% of her free throws. Suppose this probability is the same for each free throw she attempts, and free throw attempts are independent. The probability that she makes all of her first four free throws and then misses her fifth attempt this season is
- A. 0.32768.
 - B. 0.08192.
 - C. 0.06554.
 - D. 0.00128.
 - E. 0.00032.

AP Stats 1-6 Review Answer Section

MULTIPLE CHOICE

- | | |
|------------|--|
| 1. ANS: E | TOP: Individuals and variables |
| 2. ANS: C | TOP: Definition of distribution |
| 3. ANS: A | TOP: Compare two categorical variables (not in two-way table). |
| 4. ANS: D | TOP: Conditional distribution--calculation |
| 5. ANS: D | TOP: Conditional distribution--identification |
| 6. ANS: C | TOP: Conditional distribution--calculation |
| 7. ANS: C | TOP: Conditional distribution--identification |
| 8. ANS: C | TOP: Marginal distribution-calculation |
| 9. ANS: C | TOP: Describing distribution; interpreting histogram |
| 10. ANS: D | TOP: Interpret histogram |
| 11. ANS: D | TOP: Interpret segmented bar graph |
| 12. ANS: C | TOP: Interpret boxplot |
| 13. ANS: A | TOP: Interpret boxplot |
| 14. ANS: B | TOP: 1.5 x IQR rule (from data) |
| 15. ANS: D | TOP: Compare parallel box plots |
| 16. ANS: B | TOP: Density curve features |
| 17. ANS: A | TOP: Mean of density curve |
| 18. ANS: B | TOP: Median of density curve |
| 19. ANS: B | TOP: Percentiles |
| 20. ANS: D | TOP: Percentiles (from data) |
| 21. ANS: E | TOP: Cumulative freq. graph |
| 22. ANS: B | TOP: Z-score calculation |
| 23. ANS: B | TOP: Impact of transformation on numerical summaries |
| 24. ANS: B | TOP: Impact of transformation on numerical summaries |
| 25. ANS: A | TOP: Impact of transformation on numerical summaries |
| 26. ANS: A | TOP: Impact of transformation on numerical summaries |
| 27. ANS: A | TOP: Impact of transformation on z scores |
| 28. ANS: A | TOP: 68-95-99.7 rule |
| 29. ANS: D | TOP: Inverse Normal Calculations |
| 30. ANS: C | TOP: Normal Calculations |
| 31. ANS: B | TOP: Explanatory/response |
| 32. ANS: E | TOP: Interpreting Scatterplot |
| 33. ANS: D | TOP: Impact of Outlier on r |
| 34. ANS: C | TOP: Causation |
| 35. ANS: B | TOP: How r is calculated |
| 36. ANS: E | TOP: Characteristics of LSRL |
| 37. ANS: B | TOP: Extrapolation |
| 38. ANS: E | TOP: Interpret s |
| 39. ANS: C | TOP: Identify sampling frame |
| 40. ANS: A | TOP: Identify sample |

41. ANS: A	TOP: Voluntary response
42. ANS: C	TOP: SRS definition
43. ANS: E	TOP: What a SRS doesn't do
44. ANS: C	TOP: Using a random digits table
45. ANS: D	TOP: Idea of random digits table
46. ANS: E	TOP: Idea of random digits table
47. ANS: B	TOP: Stratified random sample
48. ANS: B	TOP: What kind of sampling? (Cluster)
49. ANS: B	TOP: Question wording
50. ANS: C	TOP: What kind of bias? (Undercoverage)
51. ANS: A	TOP: Experiment vs. Observational study
52. ANS: B	TOP: Scope of inference
53. ANS: C	TOP: Confounding
54. ANS: D	TOP: Block design
55. ANS: B	TOP: Identify experimental units
56. ANS: A	TOP: Scope of inference
57. ANS: A	TOP: Confounding
58. ANS: D	TOP: Purpose of randomization
59. ANS: C	TOP: Double blind
60. ANS: D	TOP: Matched Pairs
61. ANS: C	TOP: Idea of probability
62. ANS: D	TOP: Idea of probability/Myths
63. ANS: A	TOP: Idea of probability/Myths
64. ANS: D	TOP: Simulation to estimate probability
65. ANS: D	TOP: Sample space
66. ANS: B	TOP: Sample space
67. ANS: D	TOP: Basic Probability Rules
68. ANS: A	TOP: Addition of disjoint events
69. ANS: C	TOP: Multiplication Rule, Independent events
70. ANS: C	TOP: Basic Probability Rules
71. ANS: D	TOP: Complement rule
72. ANS: C	TOP: Addition of disjoint events
73. ANS: C	TOP: Mutually exclusive events
74. ANS: C	TOP: General addition rule
75. ANS: C	TOP: Independent and mutually exclusive events
76. ANS: C	TOP: Multiplication Rule, Independent events
77. ANS: D	TOP: General addition rule (and multiplication of indep. events)
78. ANS: A	TOP: Conditional probability formula
79. ANS: A	TOP: Tree diagram from probabilities
80. ANS: D	TOP: Venn diagrams
81. ANS: D	TOP: Venn diagrams
82. ANS: E	TOP: Conditional probability from 2-way table
83. ANS: D	TOP: Conditional probability from 2-way table
84. ANS: C	TOP: Conditional probability from 2-way table
85. ANS: A	TOP: Continuous vs. Discrete random variables
86. ANS: D	TOP: Identifying random variables

87. ANS: E TOP: Idea of random variable
88. ANS: B TOP: Discrete random variables: probabilities from tables
89. ANS: B TOP: Discrete random variables: probabilities from tables
90. ANS: D TOP: Discrete random variables: probabilities from tables
91. ANS: B TOP: Mean of Discrete Random Variable
92. ANS: C TOP: Standard deviation of Discrete R.V.
93. ANS: C TOP: Continuous rand. vars.: probabilities from density curves
94. ANS: B TOP: Continuous rand. vars.: probabilities from density curves
95. ANS: A TOP: Continuous rand. vars.: probabilities from density curves
96. ANS: A TOP: Normal random variable probability
97. ANS: D TOP: Normal random variable probability
98. ANS: D TOP: Normal random variable probability
99. ANS: E TOP: Mean of sum of random variables
100. ANS: A TOP: Std. dev. of sum of random variables
101. ANS: C TOP: Linear transformation of random variable
102. ANS: A TOP: Combining normal random variables
103. ANS: A TOP: Geometric setting
104. ANS: B TOP: Binomial setting
105. ANS: B TOP: Binomial probability
106. ANS: B TOP: Binomial probability
107. ANS: B TOP: Binomial probability
108. ANS: C TOP: Binomial mean
109. ANS: B TOP: Binomial standard deviation
110. ANS: B TOP: Geometric probability