MAT-150 Statistics Final Exam Review

Name_____

A large state university conducted a survey among their students and received 300 responses. The survey asked the students to provide the following information:

* Age
* Year in School (Freshman, Sophomore, Junior, Senior)
* Gender
* GPA

      1) What type of graph would you use to describe the variables Gender and Year in School?
          A) A side-by-side histogram should be used since these are two numerical variables.
          B) A side-by-side bar chart should be used since these are two categorical variables.
          C) A side-by-side bar chart should be used since these are two numerical variables.
          D) A side-by-side histogram should be used since these are two categorical variables.

A large state university conducted a survey among their students and received 400 responses. The survey asked the students to provide the following information:

* Age
* Year in School (Freshman, Sophomore, Junior, Senior)
* Major

      2) What type of graph would you use to describe the variable Major?
          A) A bar chart because Major is a categorical variable.
          B) A histogram because Major is a numerical variable.
          C) A bar chart because Major is a numerical variable.
          D) A histogram because Major is a categorical variable.

Assume that the average Systolic blood pressure for adults age 50-54 is 125 mmHg with a standard deviation of 5 mmHg. It is known that Systolic blood pressure is not normally distributed. A sample of 25 adult Systolic blood pressure measurements are taken from the population.

      3) What is the approximate z-value with interpretation for the probability that the average Systolic blood pressure
          will be less than 122 mmHg? Round to the nearest hundredth.
          A) z = -3.00 which is less than or equal to 3 standard deviations which is not a significant result.
          B) This probability cannot be determined because we do not know the distribution of the population.
          C) z = 3.00 which is greater than or equal to 3 standard deviations which is a significant result.
          D) z = -3.00 which is less than or equal to 3 standard deviations which is a significant result.
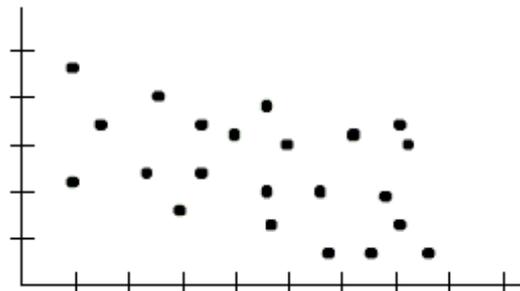
Choose the scatterplot that matches the given correlation coefficient.
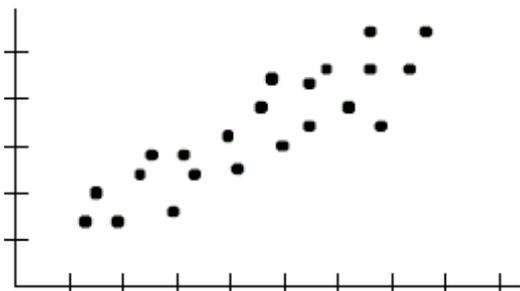
4) $r = 0.8787$

A)

B)

C)

Determine if the following scenario is an observational study or a controlled experiment.

5) A doctor is interested in determining whether a certain medication reduces migraines. She randomly selects 100 people for his study - 50 who will take the medication, and 50 who will take a placebo. The patients are examined once a week for six weeks.

A) Observational study   B) Neither   C) Controlled experiment

Feature movie lengths (in hours) were measured for all movies shown in the past year in the U.S. The mean length of all feature length movies shown was 1.80 hours with a standard deviation of 0.15 hours. Suppose the length of a random sample of 20 movies was recorded from all movies released this year. The mean length of the feature length movies was found to be 1.72 hours with a standard deviation of 0.18 hours.

6) If we create a sampling distribution of sample means, what would be the mean and standard deviation of that distribution given the sample size of 20?

A) The mean length would be 1.80 hours with a standard deviation of 0.15 hours.
B) The mean length would be 1.80 hours with a standard deviation of 0.18 hours.
C) The mean length would be 1.80 hours with a standard deviation of 0.034 hours.
D) The mean length would be 1.72 hours with a standard deviation of 0.18 hours.

7) In this example, the numerical values of 1.80 hours and 0.15 hours are _____ .
    A) statistics   B) unbiased estimators
    C) parameters   D) estimates

Fill in the blank to complete the statement.
    8) Researchers are interested in learning more about the age of men when they marry for the first time so they survey 500 married or divorced men and ask them how old they were when they first married. The mean of age of the 500 men when they married for the first time would be a _____ .
    A) Parameter   B) Population   C) Sample   D) Statistic

Match one of the histograms with its description.
    9) The distribution of male heights is displayed in histogram _____ .
    A)   B)   C)

    10) The distribution of scores on an easy test is displayed in histogram _____ .
    A)   B)   C)

Solve the problem.
    11) Which of the following would likely show a bimodal distribution in a histogram?
    A) The midterm exam scores for an introduction to Spanish course.
    B) The ages of students who attend a local high school.
    C) The price of college tuition, including both public and private schools.
    D) The number of hours a college student spends on homework per night.

12) According to the following data table, which variable(s) is(are) categorical?

Age	Gender	Weight	Ethnicity
23	1	180	1
18	0	126	0
20	0	139	2
19	1	154	1
20	1	202	3

    A) Gender
    B) Gender and ethnicity
    C) None are categorical because there are only numbers in the table
    D) Age, gender, and ethnicity

13) According to the following two-way table, why are percentages more useful than counts to compare the amount of males and females who take naps?

	Male	Female
Naps	25	30
Does not nap	35	10

    A) There are more people who take naps than people who do not in the sample.
    B) There are more males than females in the sample.
    C) You should only use counts in a two-way table.
    D) You should only use percentages in a two-way table.

14) According to the following two-way table, why are percentages more useful than counts to compare pet preferences between males and females?

	Male	Female
Dog	40	25
Cat	25	10

    A) There are more people who prefer dogs than cats in the sample.
    B) There are more males than females in the sample.
    C) You should only use counts in a two-way table.
    D) You should only use percentages in a two-way table.

15) What does it mean for an experiment to be double-blinded?
    A) The participants do not know who is in the treatment and control groups.
    B) Neither the researcher nor the participants know who is in the treatment and control groups.
    C) The researcher and the participants know which group they are in because it is unethical to keep this information from them.
    D) The researcher does not know which participants are in the treatment and control groups.

16) In Los Angeles, juice cleansing is very popular. Some people have claimed that the cleanses are beneficial for weight loss, body detoxification, and treatment and prevention of illnesses. Can we conclude that juice cleansing causes these health benefits?
    A) No, the claims are anecdotes and do not give us a true comparison group to find health differences.
    B) Yes, the claims are true stories, so we do have evidence of the health benefits.
    C) Yes, the claims are anecdotes and give us a good comparison group to find health differences.
    D) No, the claims are lies, so we do not have evidence of the health benefits.

17) Suppose temperatures and monthly ice cream sales were measured for 100 randomly selected cities. Which variable is the explanatory variable and which one is the response?

    A) Temperature is the explanatory variable and ice cream sales is the response because a city's temperature could explain the amount of ice cream sales.

    B) Ice cream sales is the explanatory variable and temperature is the response because the ice cream sales could explain the city's temperature.

    C) Ice cream sales is the explanatory variable and temperature is the response because a city's temperature could explain the amount of ice cream sales.

    D) Temperature is the explanatory variable and ice cream sales is the response because the ice cream sales could explain the city's temperature.

18) Complete the statement by filling in the blanks.

When constructing a confidence interval, if the level of confidence increases, the margin of error must _____ and the confidence interval will be _____.

    A) Increase; narrower.   B) Increase; wider.

    C) Decrease, narrower.   D) Decrease, wider.

19) A polling agency wants to estimate the proportion of U.S. citizens who support the president's domestic policies. They surveyed 2500 U.S. citizens and found a 95% confidence interval for the difference in proportions between men and women who support the president's domestic policies as (-0.025 to 0.050) where population 1 is men and population 2 is women. Select the correct interpretation of this result.

    A) The interval contains zero which shows that men are more likely than women to disagree with the president's foreign policies.

    B) The interval does not contain zero which shows that there is no significant difference in the proportions between men and women.

    C) The interval contains zero which shows that there is no significant difference in the proportions between men and women.

    D) The interval contains zero which shows that women are more likely than men to disagree with the president's foreign policies.

20) Complete the statement by filling in the blanks.

The null hypothesis H_0 is the statement of _____ and always has a _____ sign. The alternative hypothesis H_a is the _____ hypothesis. It is a statement about the value of a _____ that we intend to test.

    A) change; = ; no change; parameter   B) no change; = ; research; parameter

    C) no consequence; > ; research; sample   D) no change; = ; research; sample

21) Choose the statement that best describes what is meant when we say that the sample mean is *unbiased* when estimating the population mean.

    A) The sample mean will always equal the population mean.

    B) The standard deviation of the sampling distribution (also called the standard error) and the population standard deviation are equal.

    C) On average, the sample mean is the same as the population mean.

    D) None of these.

22) Which of the following statements is true about the t-distribution?

    A) Since population standard deviation is usually unknown, the standard error uses the sample standard deviation to estimate population standard deviation.

    B) Like the Normal distribution, the t-distribution is symmetric for small n.

    C) For small sample sizes, the t-distribution has the same properties as the normal curve.

    D) For large sample sizes, the t-distribution has the same properties as the normal curve.

23) Suppose a consumer product researcher wanted to find out whether a Sharpie lasted longer than the manufacturer's claim that their Sharpies could write continuously for a mean of 14 hours. The researcher tested 40 Sharpies and recorded the number of continuous hours each Sharpie wrote before drying up. Test the hypothesis that Sharpies can write for more than a mean of 14 continuous hours. Following are the summary statistics: $\bar{x} = 14.5$ hours, $s = 1.2$ hours At the 5% significance level, $t = 2.635$; $p = 0.006$. State your conclusion about the original claim.

   A) Reject the null hypothesis; there is strong evidence to suggest that the Sharpies last longer than a mean of 14 hours.

   B) There needs to be more data to determine if the Sharpies last longer than a mean of 14 hours.

   C) Do not reject the null hypothesis; there is not strong enough evidence to suggest that the Sharpies last longer than a mean of 14 hours.

   D) Reject the alternative hypothesis; there is strong evidence to suggest that the Sharpies last longer than a mean of 14 hours.

24) Suppose a consumer product researcher wanted to find out whether a highlighter lasted less than the manufacturer's claim that their highlighters could write continuously for 14 hours. The researcher tested 40 highlighters and recorded the number of continuous hours each highlighter wrote before drying up. Test the hypothesis that the highlighters wrote for less than 14 continuous hours. Following are the summary statistics:
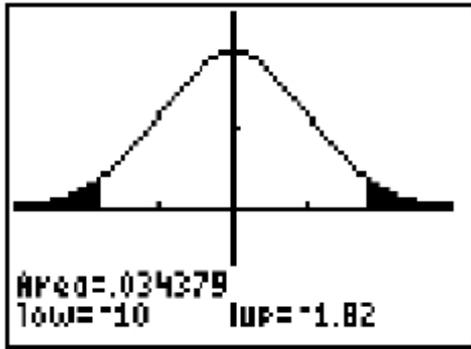
   $\bar{x} = 13.6$ hours, s = 1.3 hours

Report the test statistic, p-value, your decision regarding the null hypothesis, and your conclusion about the original claim. Round all values to the nearest thousandth.
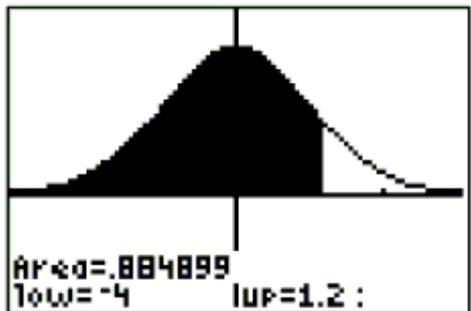
   A) $t = -1.946$; $p = 0.029$; Reject the null hypothesis; there is strong evidence to suggest that the highlighters last less than 14 hours.

   B) $t = -1.946$; $p = 0.029$; Fail to reject the null hypothesis; there is not strong evidence to suggest that the highlighters last less than 14 hours.

   C) $z = 1.946$; $p = 0.974$; Fail to reject the null hypothesis; there is not strong evidence to suggest that the highlighters last less than 14 hours.

   D) $z = 1.946$; $p = 0.029$; Reject the null hypothesis; there is strong evidence to suggest that the highlighters last less than 14 hours.

25) From the TI-84 graphing calculator screenshots below, choose the screenshot whose shaded area depicts a p-value for a two-tailed test.
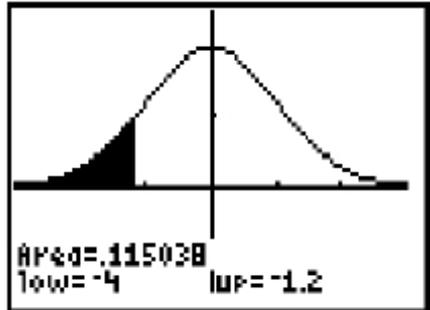
A)

```
Area=.034379
low=-10     lup=-1.82
```

B)

```
Area=.884899
low=-4      lup=1.2 :
```

C)

```
Area=.115038
low=-4      lup=-1.2
```

26) A quality control manager thinks that there is a higher defective rate on the production line than the advertised value of $p = 0.025$. She does a hypothesis test with a significance level of 0.05. Symbolically, the null and alternative hypothesis are as follows: H_0: $p = 0.025$ and H_a: $p > 0.025$. She calculates a p-value for the hypothesis test of defective light bulbs to be approximately 0.067. Choose the correct interpretation for the p-value.

 A) The p-value tells us that if the defect rate is 0.025, then the probability that she would observe the percentage she actually observed or higher is 0.067. At a significance level of 0.05, this would not be an unusual outcome.

 B) The p-value tells us that the result is significantly higher than the advertised value using a significance level of 0.05.

 C) The p-value tells us that the probability of concluding that the defect rate is equal to 0.025, when in fact it is greater than 0.025, is approximately 0.067.

 D) The p-value tells us that the true population rate of defective light bulbs is approximately 0.067.

27) Which of the following is not a condition that must be checked before proceeding with a two-sample test?
   A) The samples must be independent of each other.
   B) Both samples must be large enough so that the product of each sample size (n_1 and n_2) and the pooled

   estimate, $\hat{p}$, is greater than or equal to 10.
   C) Each sample must be a random sample.
   D) All of these are conditions that must be checked to proceed with a two-sample test.

28) A researcher believes that the proportion of women who exercise with a friend is greater than the proportion of men. He takes a random sample from each population and records the response to the question, "Have you exercised with a friend at least once in the last seven days?" The null hypothesis is $H_0: p_{women} = p_{men}$.

   Choose the correct alternative hypothesis.
   A) $H_a: p_{women} > p_{men}$
   B) $H_a: p_{women} \neq p_{men}$
   C) $H_a: p_{women} < p_{men}$
   D) $H_a: p = 0$

29) A researcher believes that children who attend elementary school in a rural setting are more physically active then children who attend elementary school in an urban setting. The researcher collects a random sample from each population and records the proportion of children in each sample who reported participating in at least one hour of rigorous activity a day. The data is summarized in the table below. Assume the all conditions for proceeding with a two-sample test have been met.

Rural	Urban
$n_1 = 90$	$n_2 = 78$
$x_1 = 74$	$x_2 = 55$

   Find the z-statistic (rounded to the nearest hundredth) and p-value (rounded to the nearest thousandth) for this hypothesis test. Using a 5% significance level, state the correct conclusion regarding the null hypothesis, H_0:

   $p_{rural} = p_{urban}$.
   A) $z = -1.79$, $p = 0.037$. There is insufficient evidence to reject the null hypothesis.
   B) $z = 0.71$, $p = 0.073$. There is sufficient evidence to reject the null hypothesis.
   C) $z = 0.82$, $p = 0.073$. There is sufficient evidence to accept the null hypothesis.
   D) $z = 1.79$, $p = 0.037$. There is sufficient evidence to reject the null hypothesis.

30) A medical study examined data on patients with cardiovascular disease who were currently non-smokers and those who were current smokers. Population 1 were smokers and population 2 were non-smokers. After data analysis, the 95% confidence interval for the difference in proportions is 0.015 +/- 0.011. The most accurate interpretation is...
   A) We are 95% confident that the difference in the proportion of smokers compared to nonsmokers is between 0.004 and 0.026. There is a significance difference indicating higher cardiovascular disease amongst smokers.
   B) We are 95% confident that the interval of the difference in the proportions contains zero. There is not a significance difference between smokers and non-smokers.
   C) We are 95% confident that the proportion of smokers compared to non-smokers is between 0.004 and 0.026.
   D) We are 95% confident that the difference in the proportion of smokers compared to nonsmokers is between -0.004 and 0.026. There is not a significance difference in the proportions.

31) Confidence intervals can be used to determine whether different sample proportions reflect a "real" difference in the population. The basic approach is to...
   A) find the margin of error for each proportion and see if the difference is less than zero.
   B) find a confidence interval at the significance level desired for the difference in proportions.
   C) find the difference in the proportions and see if the difference is less than zero.
   D) find the difference in the proportions and see if the difference is greater than zero.

32) Complete the statement by filling in the blank. When constructing a confidence interval, if the level of confidence increases the margin of error will _____ and the confidence interval will be _____ . A larger sample size will improve the accuracy of the confidence interval, therefore margin of error will _____ and the confidence interval will be _____ .
   A) Decrease, wider. Increase, narrower   B) Increase, narrower. Decrease, wider.
   C) Increase, wider. Decrease, narrower.   D) Decrease, narrower. Increase, wider.

33) Is it plausible that more than 10% of Americans believe in aliens? A random sample of 2000 adult Americans were surveyed and 15% of them said that they believed in aliens. Find the 95% confidence interval for the proportion of Americans who believe in aliens then choose the correct interpretation. (Round to the nearest tenth of a percent)
   A) The population proportion of Americans who believe in aliens is between 10% +/- 1.6% with a confidence level of 95%. The interval includes 10% and therefore, it is plausible that at least 10% of Americans believe in aliens.
   B) The population proportion of Americans who believe in aliens is between 15% +/- 1.6% with a confidence level of 95%. The interval includes 10% and therefore, it is plausible that at least 10% of Americans believe in aliens.
   C) The population proportion of Americans who believe in aliens is between 15% +/- 0.8% with a confidence level of 95%. The interval does not include 10% and therefore, it is not plausible that at least 10% of Americans believe in aliens.
   D) The population proportion of Americans who believe in aliens is between 15% +/- 1.6% with a confidence level of 95%. The interval is higher than 10% and therefore, it is plausible that more than 10% of Americans believe in aliens.

34) If 20 babies are born, how often are there 12 or more female babies? Assume that the gender of a baby is a random event. Which of the following experiments would not simulate this situation?
   A) Flip a coin twenty times. Designate a head to mean "female" and a tail to mean "male."
   B) Choose the first twenty digits from a row in the random number table. Designate odd numbers to mean "female" and even numbers to mean "male."
   C) Roll a die twenty times. Designate a 1, 2, or 3 to mean "female" and a 4, 5, or 6 to mean "male."
   D) All of these will simulate the gender of twenty babies.

35) Which of the following statements is true about the "law of large numbers" (LLN)?
   A) If an experiment with a random outcome is repeated a large number of times, the empirical probability that is observed is consistently different from the theoretical probability.
   B) If you repeat a random experiment many, many times, your outcomes should on average approach the theoretical average.
   C) If you simulate or conduct an experiment or simulation enough times the empirical probability observed will always match the theoretical probability that is expected.
   D) If you repeat a random experiment many, many times, your outcomes should be a unique value that is separate from the theoretical average.

36) Choose the best statement to summarize the association shown between hat size and IQ in the scatterplot below.
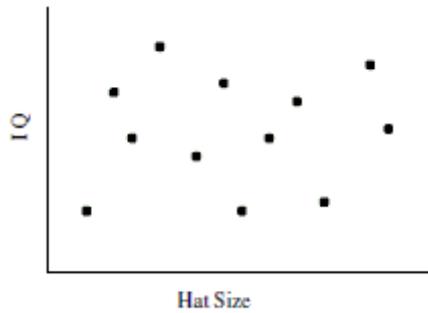
Hat Size

A) The scatterplot does not show a trend that would indicate an association between hat size and IQ scores.
B) Hat size causes IQ to increase.
C) As hat size increases, IQ scores tend to increase.
D) As hat size increases, IQ scores tend to decrease.

37) The following calculator screenshots show the scatterplot and the correlation coefficient between car weight and car length for a sample of 2009 model year cars.


```
LinRegTTest
y=a+bx
β≠0 and ρ≠0
↑b=.0099418867
s=4.308740467
r²=.4435295792
r=.6659801643
```
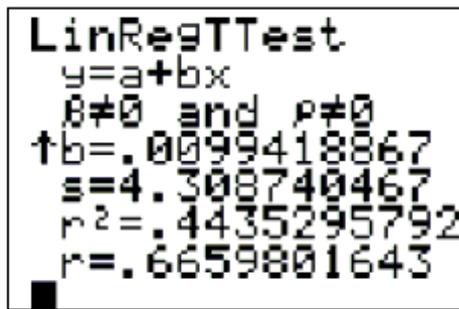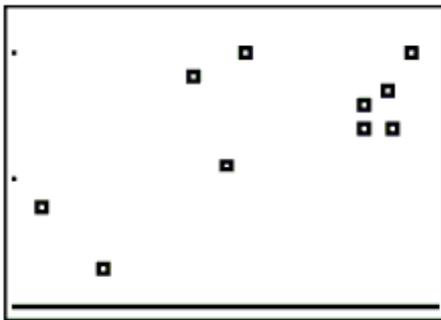
The relationship between "car length" and "car weight" can be described as
A) A strong negative linear relationship   B) A moderate positive linear relationship
C) A weak negative relationship   D) A strong positive linear relationship

38) The data in the table represent the amount of raw material (in tons) put into an injection molding machine each day (x), and the amount of scrap plastic (in tons) that is collected from the machine every four weeks (y). Also shown below are the outputs from two different statistical technologies (TI-83/84 Calculator and Excel). A scatterplot of the data confirms that there is a linear association. Report the equation for predicting scrap from raw material using words such as scrap, not x and y. State the slope and intercept of the prediction equation. Round all calculations to the nearest hundredth.

x	y
2.71	3.61
2.33	2.80
2.33	2.77
2.21	2.34
2.11	2.15
2.08	2.06
1.98	2.02
1.95	1.95
1.84	1.73
1.84	1.68

LinReg
$$y = a + bx$$
$$a = -2.376991175$$
$$b = 2.192699333$$
$$r^2 = .9806834986$$
$$r = .9902946524$$

	Coefficients
Intercept	-2.376991175
X Variable 1	2.192699333

A) scrap = 2.19 - 2.38(raw material); slope = 2.19 and the intercept is -2.38.
B) scrap = -2.38 + 2.19(raw material); slope = -2.38 and the intercept is 2.19.
C) scrap = -2.38 + 2.19(raw material); slope = 2.19 and the intercept is -2.38.
D) scrap = 2.19 - 2.38(raw material); slope = -2.38 and the intercept is 2.19.

39) The boxplots below represent movie runtimes (length of a movie in minutes) for movies that have been rated by the Motion Picture Association of America as R, PG-13, PG, and G. List ratings according to their median runtimes, from longest to shortest.
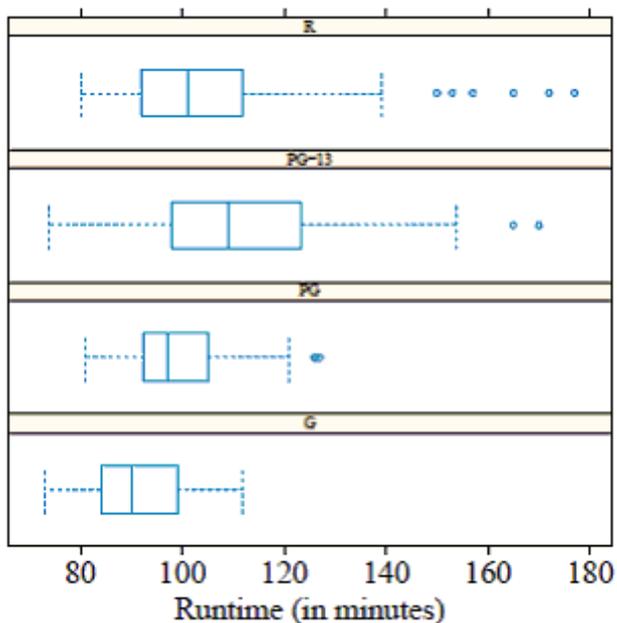
A) G, PG, PG-13, R   B) R, PG-13, PG, G   C) PG-13, R, PG, G   D) G, PG, R, PG-13

Suppose that a recent poll of American households about car ownership found that for households with a car, 39% owned a sedan, 33% owned a van, and 7% owned a sports car. Suppose that three households are selected randomly and with replacement.

40) What is the probability that at least one of the three randomly selected households own a sports car? (Round to the nearest thousandth)

    A) 0.627   B) 0. 200   C) 0.800   D) 0.003

Suppose that a recent poll of American households about pet ownership found that for households with pets, 45% owned a dog, 34% owned a cat, and 10% owned a bird. Suppose that three households are selected randomly and with replacement and the ownership is mutually exclusive.

41) What is the probability that at least one of the three randomly selected households own a bird? (Round to the nearest hundredth)

    A) 0.73   B) 0.08   C) 0.92   D) 0.27

The following side-by-side bar graph shows the level of post-secondary education achieved ten years after high school for graduates from the years 1999 and 2001. Use the bar graph to answer the question.
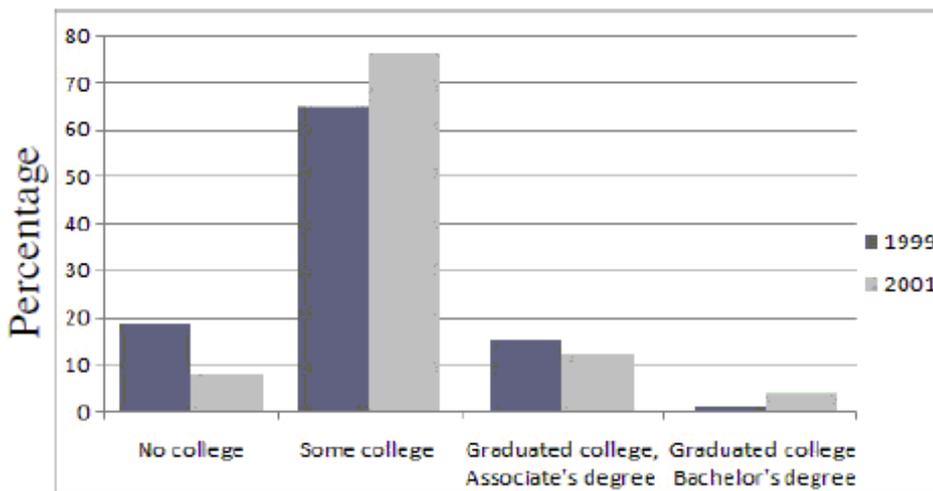
42) In which category was there more variability?
    A) Graduated College, Associate's Degree   B) No College
    C) Some College   D) Graduated College, Bachelor's Degree

The two-way table below shows the survey results when sixty adults were asked whether they had made a clothing purchase in the last thirty days.

	Male	Female
Purchased clothing in the last thirty days.	10	29
Had not purchased clothing in the last thirty days.	10	11

43) Of the adult males surveyed, what percentage had made a clothing purchase in the last thirty days?
    A) 65%   B) 35%   C) 50%   D) 33%

Use the following information for the question. The average travel time to work for a person living and working in Kokomo, Indiana is 17 minutes. Suppose the standard deviation of travel time to work is 4.5 minutes and the distribution of travel time is approximately normally distributed.

44) Approximately what percentage of people living and working in Kokomo have a travel time to work of at least 20 minutes? Round to the nearest whole percent.

    A) 25%   B) 15%   C) 75%   D) None of these.

Use the following information to answer the question. A janitor at a large office building believes that his supply of light bulbs has a defect rate that is different than the defect rate stated by the manufacturer. The janitor's null hypothesis is that the supply of light bulbs has a defect rate of $p = 0.09$ (the light bulb manufacturer's stated defect rate). Suppose we do a hypothesis test with a significance level of 0.01. Symbolically, the null and alternative hypothesis are as follows: H_0: $p = 0.09$ and H_a: $p > 0.09$.

45) Choose the statement that best describes the significance level in the context of the hypothesis test.
    A) The significance level of 0.01 is the probability of concluding that the defect rate is equal to 0.09 when in fact it is greater than 0.09.
    B) The significance level of 0.01 is the probability of concluding that the defect rate is different than 0.09 when in fact the defect rate is equal to 0.09.
    C) The significance level of 0.01 is the z-statistic that we will use to compare the observed outcome to the null hypothesis.
    D) The significance level of 0.01 is the defect rate we believe is the true defect rate.

Use the following information to answer the question. A pescatarian is a person who eats fish and seafood but no other animal. An event planner does some research and finds that approximately 2.75% of the people in the area where a large event is to be held are pescatarian. Treat the 250 guests expected at the event as a simple random sample from the local population of about 150,000.

46) Suppose the event planner assumes that only 1.6% of the guests will be pescatarian so he orders 4 pescatarian meals. What is the approximate probability that he will have too many pescatarian meals? Round to the nearest thousandth.

    A) 0.245   B) 0.113   C) 0.387   D) 0.613

Use the following information to answer the question. According to the website www.costofwedding.com, the average cost of flowers for a wedding is $698. Recently, in a random sample of 40 weddings in the U.S. it was found that the average cost of the flowers was $734, with a standard deviation of $102. On the basis of this, a 95% confidence interval for the mean cost of flowers for a wedding is $701 to $767.

47) Does the confidence interval provide evidence that the mean cost of flowers for a wedding has increased?
    A) Yes   B) No

Use the following information to answer the question. Here is a table recording the number of deaths for the top thirteen worst U.S. tornados since 1925. A histogram showing the distribution is also included.

689	268
454	150
102	181
208	114
142	115
271	110
315	
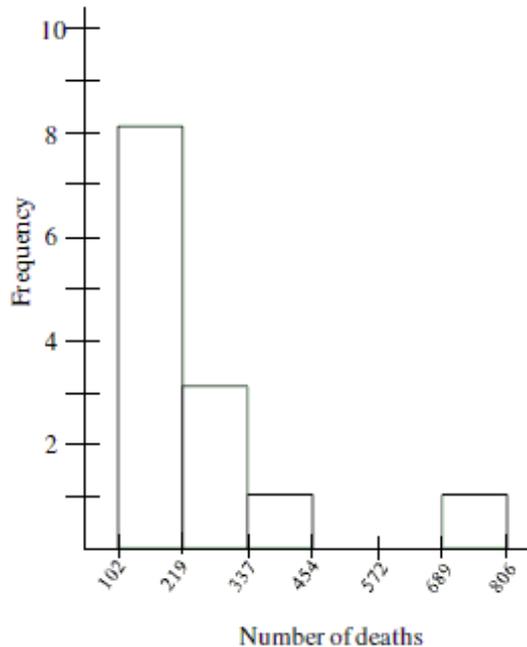
48) Estimate the most appropriate measure of variability.
   A) Standard Deviation; 169.4
   B) Standard Deviation; 178.5
   C) IQR; 156
   D) IQR; 574

Use the following information to answer the question. Many couples believe that it is getting too expensive to host an "average" wedding in the United States. According to the website www.costofwedding.com, the average cost of a wedding in the U.S. in 2009 was $24,066. Recently, in a random sample of 40 weddings in the U.S. it was found that the average cost of a wedding was $23,224, with a standard deviation of $2,903. On the basis of this, a 95% confidence interval for the mean cost of weddings in the U.S. is $22,296 to $24,152.
   49) Choose the statement that is the best interpretation of the confidence interval.
      A) We are extremely confident that the mean cost of a U.S. wedding is between $22,296 and $24,152.
      B) In about 95% of all samples of 40 U.S. weddings, the resulting confidence interval will contain the mean cost of all weddings in the U.S.
      C) That probability that a U.S. wedding will cost more than $24,152 is less than 3%.

Use the following information to answer the question. The distribution of the number of hours of sleep people get per night is unimodal and symmetric with a mean of 6 hours and a standard deviation of 1.5 hours.
   50) If James had a z-score of 1.2, explain what this value means in terms of the number of hours of sleep he gets.
      A) James sleeps 1.2 hours less than the average person.
      B) The number of hours James sleeps is 1.2 standard deviations below the mean.
      C) James sleeps 1.2 hours more than the average person.
      D) The number of hours James sleeps is 1.2 standard deviations above the mean.

14

Use the following information to answer the question. The following linear regression model can be used to predict ticket sales at a popular water park.

Ticket sales per hour = -631.25 + 11.25(current temperature in °F)

51) In this context, does the intercept have a reasonable interpretation?
A) Not enough information available
B) Yes, it is reasonable for people to go to a water park when it is 0°F, so park managers might want to know how many tickets they would sell on average on a 0°F day.
C) No, at a temperature of 0°F, ticket sales would be -631.25 and it is not reasonable (or possible) to have negative ticket sales.

Use the following information to answer the question. The mean age of lead actresses from the top ten grossing movies of 2010 was 29.6 years with a standard deviation of 6.35 years. Assume the distribution of the actresses' ages is approximately unimodal and symmetric.

52) In 2010, popular actress Jennifer Aniston was 41-years-old. What is Jennifer Aniston's age if it is standardized? Would it be unusual for a 41-year-old actress to be in a top-grossing film of 2010? Assume the Empirical Rule applies and round to the nearest hundredth.
A) $z = -1.80$; It would not be unusual.
B) $z = 1.80$; It would not be unusual.
C) $z = -1.80$; It would be unusual.
D) $z = 1.80$; It would be unusual.

Use the following information to answer the question. The mean age of lead actors from the top ten grossing movies of 2007 was 36.4 years with a standard deviation of 9.87 years. Assume the distribution of the actors ages is approximately unimodal and symmetric.

53) Between what two values would you expect to find about 95% of the lead actors ages?
A) 26.53 and 46.27 years
B) 6.87 and 66.01 years
C) 16.66 and 56.14 years
D) None of these

Use the following information to answer the question.
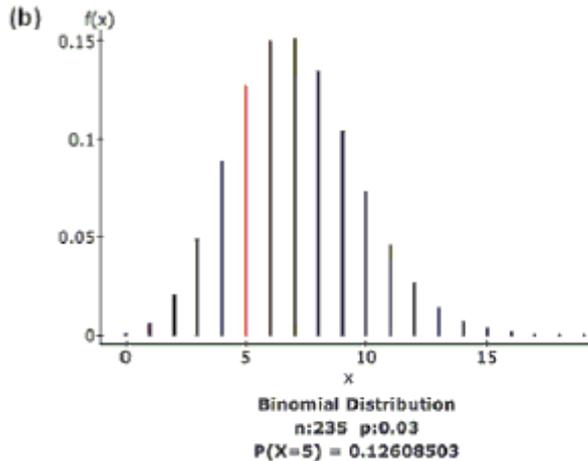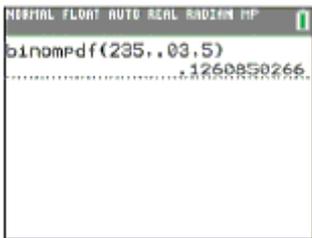
54) What is the approximate probability that x is equal to 5 or less? Round to nearest hundredth place.
A) 0.13
B) 0.17
C) 0.71
D) 0.29

Use the following regression equation regarding car mileage to answer the question.

$\widehat{\text{Highway}} = 0.892 + 1.337 \cdot (\text{City})$

Note that City is the estimated miles per gallon (mpg) a car gets while driving on city streets, and Highway is the estimated miles per gallon (mpg) a car gets while driving on highways.

55) Interpret the slope in the context of the data.
   A) The slope is 0.892. If a car gets 0 mpg in the city, it will get 0.892 mpg on the highway.
   B) The slope is 1.337. If a car gets 0 mpg in the city, it will get 1.337 mpg on the highway.
   C) The slope is 0.892. For every additional mpg a car gets in the city, its highway mpg is predicted to increase by 0.892.
   D) The slope is 1.337. For every additional mpg a car gets in the city, its highway mpg is predicted to increase by 1.337.

Use the following table to answer the question. A random sample of college students was asked to respond to a survey about how they spend their free time on weekends. One question, summarized in the table below, asked each respondent to choose the one activity that they are most likely to participate in on a Saturday morning. The activity choices were homework, housework, outside employment, recreation, or other.

	Homework	Housework	Outside Employment	Recreation	Other	Total
Male	29	15	20	23	9	96
Female	18	17	26	39	4	104
Total	47	32	46	62	13	200

56) If one student is randomly chosen from the group, what is the probability that the student is female or chose "homework" as their most likely activity on a Saturday morning?
   A) 0.755   B) 0.900   C) 0.665   D) None of these

Use the side-by-side boxplots below to answer the question. The boxplots summarize the number of sentenced prisoners by state in the Midwest and West.

Midwest

1,435   4,322   9,891   29,928   50,648

West

2,114 3887.5   6,887   15,706   22,662   34,581
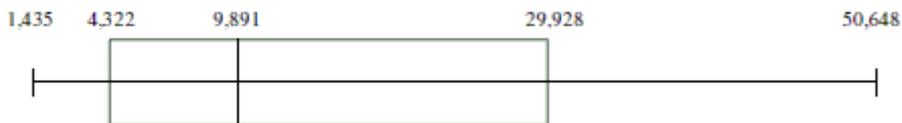
57) Pick the statement that best describes the shape of the distribution for the states in the West.
   A) The data appears to be left-skewed with large variability.
   B) The data appears to be roughly symmetrical with a possible outlier.
   C) The data appears to be right-skewed with a possible outlier.

A marble manufacturer advertises that its bags of marbles will contain 25% "milky-white" marbles. Suppose that a bag containing 80 marbles is inspected.

58) Use your answers to fill in the blanks: We expect _____% milky-white marbles, give or take _____%.

A random sample of car buyers was asked to respond to a survey about what was the most important quality of the car they purchased. This question is summarized in the table below. The important contributors were fuel efficiency, looks, manufacturer reputation, price or other.

	Fuel Efficiency	Looks	Manufacturer Reputation	Price	Other	Total
Male	29	15	20	23	9	96
Female	18	17	26	39	4	104
Total	47	32	46	62	13	200

59) If one car buyer is randomly chosen from the group, what is the probability that the buyer is male and chose "manufacturer reputation" as their most important factor for the purchase?

60) If one car buyer is randomly chosen from the group, what is the probability that the buyer is female and chose "fuel efficiency" or "other" as their most important factor for the purchase?

A random sample of college students was asked to respond to a survey about how they spend their free time on a week night. One question, summarized in the table below, asked each respondent to choose the one activity that they are most likely to participate in on a Wednesday afternoon/evening. The activity choices were homework, housework, outside employment, recreation, or other.

	Homework	Housework	Outside Employment	Recreation	Other
Male	40	2	40	11	19
Female	37	3	30	10	8

61) Using this example, state two events that are mutually exclusive.

An event planner does some research and finds that in the area where a large children's event is to be held, approximately 1.75% of the children are lactose intolerant. Treat the 250 children expected at the event as a simple random sample from the local population of about 100,000 children.

62) Suppose the event planner assumes that only 0.8% of the children attending the event will be lactose intolerant so he orders 2 lactose-free meals. What is the approximate probability that he will have too many lactose-free meals? Round to the nearest thousandth.

Answer the question about the Law of Large Numbers.

63) A fair coin is tossed 1000 times. What can you say about getting the outcome of exactly 500 tails?
a. Since the probability of a tail is 0.5 for each toss, you should expect exactly 500 tails in 1000 tosses.
b. You should not expect exactly 500 tails in 1000 tosses, but the proportion of tails should approach 0.5 as the number of tosses increases.
c. You should expect between 400 and 600 tails in 1000 tosses.
d. Getting 500 tails is no more likely than getting any other number of tails in 1000 tosses.

Answer the question.

64) A magazine publisher always mails out a questionnaire six months before a subscription ends. This questionnaire asks its subscribers if they are going to renew their subscriptions. On average, only 7% of the subscribers respond to the questionnaire. Of the 7% who do respond, an average of 47% say that they will renew their subscription. This 7% who respond to the questionnaire are known as what?

Determine the interquartile range.

65) Determine the interquartile range.

2, 3, 5, 8, 9, 12, 2, 3, 5, 8, 9, 12

66) The test scores of 19 students are listed below. Find the interquartile range.

91 45 86 71 58
64 97 55 90 79
82 83 52 88 75
43 92 94 67

Determine whether the following is a probability distribution. If not, identify the requirement that is not satisfied.
67) In a certain town, 40% of adults have a college degree. The accompanying table describes the probability distribution for the number of adults (among 4 randomly selected adults) who have a college degree.

x	P(x)
0	0.1296
1	0.3456
2	0.3456
3	0.1536
4	0.0256

Find the indicated probability.
68) In a homicide case 6 different witnesses picked the same man from a line up. The line up contained 5 men. If the identifications were made by random guesses, find the probability that all 6 witnesses would pick the same person.

Find the indicated probability. Round to three decimal places.
69) Assume that male and female births are equally likely and that the birth of any child does not affect the probability of the gender of any other children. Find the probability of exactly two girls in ten births.

Find the standard deviation for the given sample data. Round your answer to one more decimal place than is present in the original data.
70) Christine is currently taking college astronomy. The instructor often gives quizzes. On the past seven quizzes, Christine got the following scores:

44 20 37 28 19 52 55

For the given hypothesis test, explain the meaning of the kind of error, as requested.
71) A statistics student has heard that about 21% of the students on his campus attend sporting events weekly. He wants to know if statistics students attend events in the same proportions as the general student body. Explain what the second type of error would be in this case (where the student fails to reject a null hypothesis that is actually false).

From a random sample of workers at a large corporation you find that 58% of 200 went on a vacation last year away from home for at least a week.
72) An approximate 95% confidence interval is (0.50, 0.66). Which of the following statements is a correct interpretation?

Provide an appropriate response.
73) The city council wants to know what percentage of people in their town have health insurance. They take a poll, and of 88 adults selected randomly from the town, 63 have health insurance. Should the council use a hypothesis test or confidence interval to answer their question?

74) What generally happens to the sampling error as the sample size is decreased?

75) Decide whether the experiment is a binomial experiment. If it is not, explain why. Testing a cough suppressant using 120 people to determine if it is effective. The random variable represents the number of people who find the cough suppressant to be effective.

76) The data below are the ages and annual pharmacy bills (in dollars) of 9 randomly selected employees. Calculate the linear correlation coefficient.

Age, x	34	37	41	44	47	49	53	57	61
Pharmacy bill ($), y	113	117	120	128	139	142	145	147	149

Shade the approximate area that would represent the p-value for the alternative hypothesis and z-score, and then calculate the p-value. Round to the nearest thousandth.

77) The alternative hypothesis is a two-tailed with a z-score = −1.88
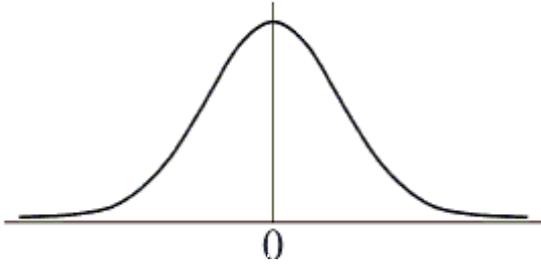
Solve the problem.
78) The weights at birth of five randomly chosen baby giraffes were 111, 115, 120, 103, and 106 pounds. Assume the distribution of weights is normally distributed. Find a 95% confidence interval for the mean weight of all baby giraffes. Use technology for your calculations. Give the confidence interval in the form "estimate ± margin of error." Round to the nearest tenth of a pound.

79) A bicycle manufacturer produces four different bicycle models. Information is summarized in the table below:

Model	Series Number	Weight	Style
Ascension	A120	32	Mountain
Road Runner	B640	20	Road
All Terrain	C300	28	Hybrid
Class Above	D90	14	Racing

Identify the variables and determine whether each variable is numerical or categorical.

80) The bar charts below depict the MPAA movie ratings of 489 movies, separated by high and low critic scores. Which bar chart shows more variability in MPAA movie ratings? Why?
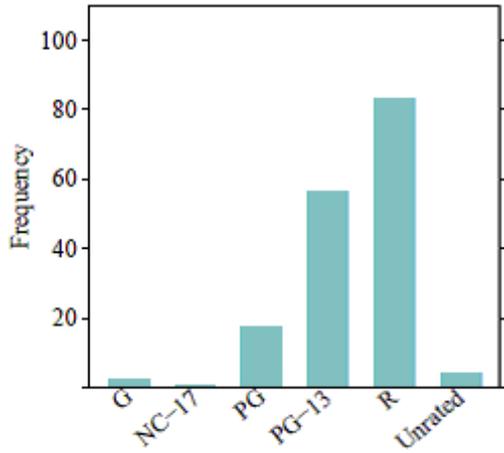
MPAA Movie Rating – High Critic Scores
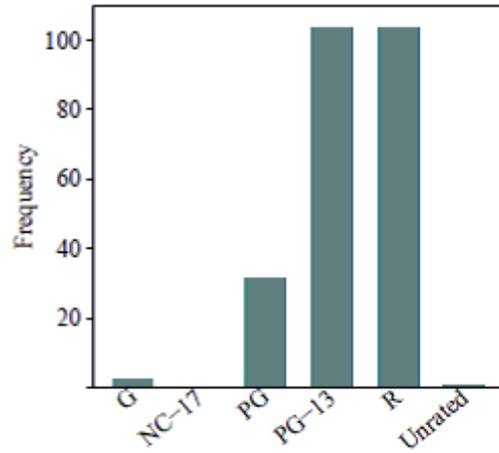
MPAA Movie Rating – Low Critic Scores

81) What is the difference between a bar chart and a histogram?

82) A statistics student collected data from other students in her class who ride a bike to school. The following table shows data about their bikes:

Color	Series Number	Weight (lbs)	Road Bike	Average Speed (mph)
Black	A120	32	0	16
Blue	B640	20	1	24
Green	C300	29	0	14
Black	D90	15	1	23

How many variables are there?

83) The quality engineer at a paint manufacturer conducted a hypothesis test to test the claim that the mean volume of paint cans had changed after an adjustment in the manufacturing process. Mean volume in paint cans before the adjustment was 1.02 gallons. Assume that all conditions for testing have been met. She used technology to complete the hypothesis test. Following is the null and alternative hypothesis and the output from her graphing calculator.

H_0: μ = 1.02 gallons

H_a: $\mu \neq$ 1.02 gallons

T-Test
$\mu \neq 1.02$
$t = -5.89255651$
$p = 3.4248373E - 7$
$\bar{x} = .97$
$Sx = .06$
$n = 50$
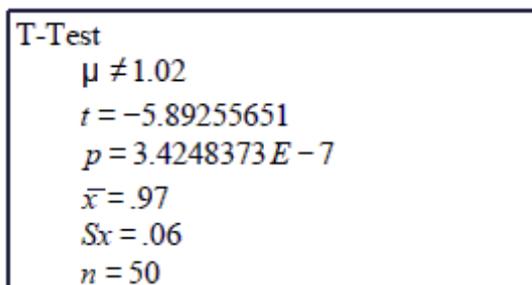
Write a statement explaining what her decision regarding the null hypothesis should be and a statement summarizing her conclusion regarding the claim that average volume of paint cans had changed. Has the adjustment in the manufacturing process changed the average volume of paint cans?

20

84) Swinging Sammy Skor's batting prowess was simulated to get an estimate of the probability that Sammy will get a hit. Let 1 = HIT and 0 = OUT. The output from a simulation was as follows.

1 0 0 0 1 0 0 1 0 0 1 1 1 0 0 0 0 1 1 1 1 0 0 0 0 1 1 0 0 1 1 1 1 0 0 0 0 0 1 1 1 1

Estimate the probability that he gets a hit.

85) Suppose data were collected on neighborhoods about the number of crimes committed and the number of police who patrol the area. Which variable is the explanatory variable and which one is the response? Explain your reasoning.

86) For the following scatterplot, what effect would the outlier have on the slope of the regression equation? Explain your reasoning.
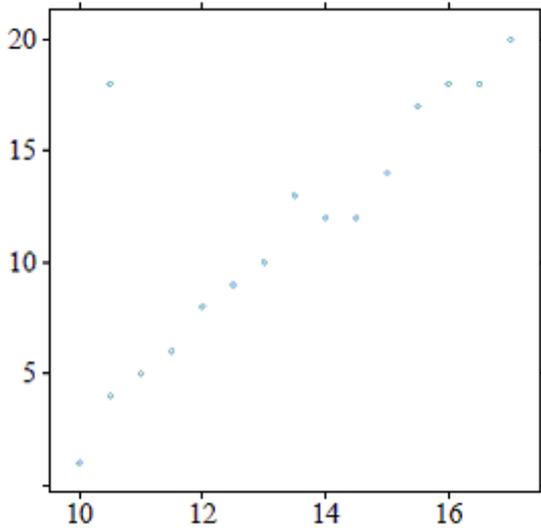
87) Two algebra classes at University High School took the same quiz. Mr. Athens had 25 students in his class with a mean score of 80. Mrs. Sutton's class of 30 students had a mean score of 75. Overall, what was the mean score for all students on the quiz?

88) Suppose we have a data set of the number of car accidents per day in Los Angeles during the year 2013. The data was input into a spreadsheet manually by an assistant at the Department of Transportation. For one day in July 2013, he input that there were 1500 car accidents; but there were actually only 150 that day. How will this error affect the measures of center for this data?

89) For a left-skewed distribution, how will the median value compare to the mean value?

90) Ten parents were asked the ages of their youngest child. The results are shown below.

   6   25   12   20   6   2   17   22   23   10

What is the IQR for this set of data?

91) Suppose that the following is to be tested: H_0: p = 0.72 and H_a: $p \neq 0.72$. Calculate the observed z-statistic for the following sample data: Sixty-eight out of ninety test subjects have the characteristic of interest. Round to the nearest thousandth.

Solve the problem. Round monetary amounts to the nearest dollar.

92) Packages of a certain candy vary slightly in weight. Here are the measured weights of nine packages, in ounces:
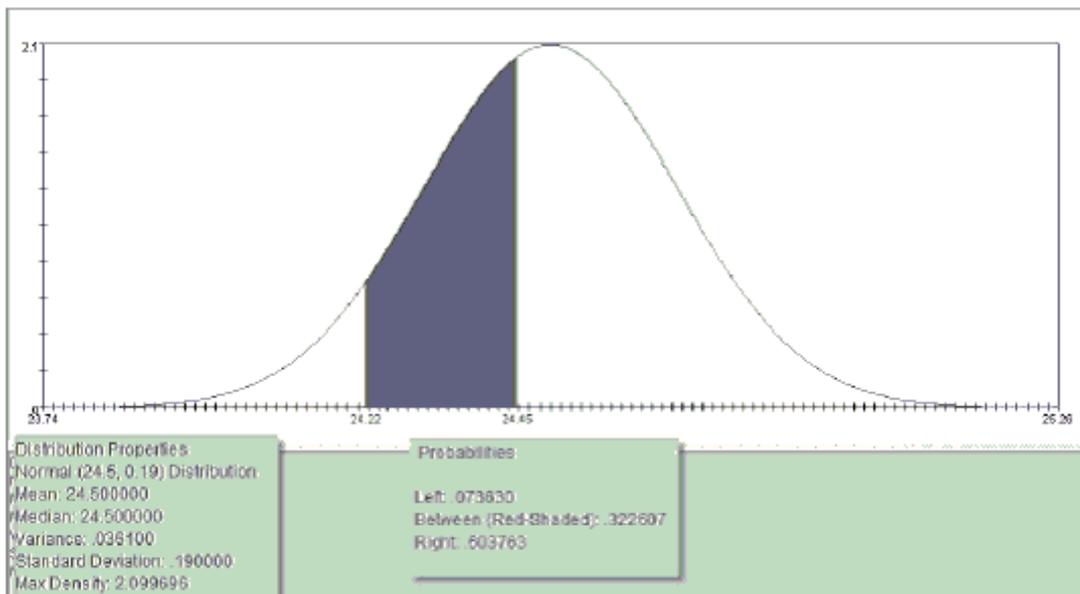
1.6823   1.6844   1.6851   1.6866   1.6562

1.6848   1.6829   1.6835   1.6858

a. Find the mean and the median of these weights.

b. Which, if any, of these weights would you consider to be an outlier?

c. What are the mean and median weights if the outlier is excluded?

Suppose that a recent poll of single people over the age of thirty-five were asked about their living arrangements. The poll found that 34% rented a house or apartment, 21% owned a house, and 17% owned a condominium. Suppose that four single people are selected randomly and with replacement.

93) What is the probability that none of the four randomly selected people rent a house or apartment? Show your work and round to the nearest thousandth.

Suppose that weights of jars of Puff brand marshmallow cream has a population mean of 24.5 ounces and a population standard deviation of 0.19 ounces and are approximately normally distributed. Use the figure below to find determine the specified probabilities.

Distribution Properties
Normal (24.5, 0.19) Distribution
Mean: 24.500000
Median: 24.500000
Variance: .036100
Standard Deviation: .190000
Max Density: 2.099696

Probabilities
Left: .073630
Between (Red-Shaded): .322607
Right: .603763

94) If a large random sample of Puff marshmallow cream jars were weighed, approximately what percentage of the jars would weigh between 24.22 and 24.45 ounces? Round to the nearest tenth of a percent.

The following scatterplot shows the relationship between heights (in cm) and weights (in kg) of 100 Americans. The coefficient of determination was found to be 37.9%.
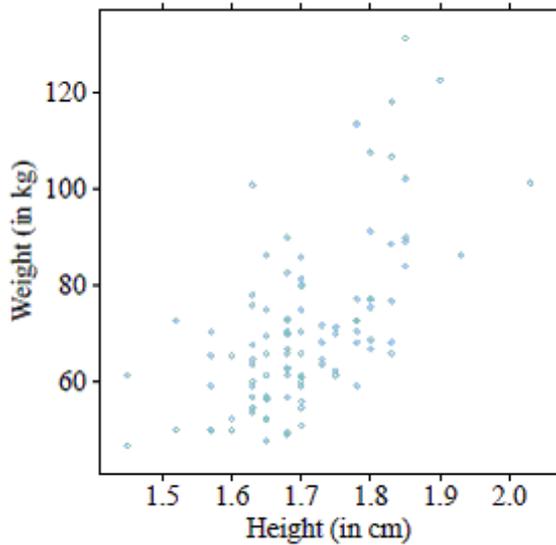
95) Calculate the value of the correlation coefficient between height and weight.

Use the empirical rule to solve the problem.

96) At one college, the GPA's are Normally distributed with a mean of 3 and a standard deviation of 0.5. What percentage of students at the college have a GPA between 2.5 and 3.5?

97) Solar energy is considered by many to be the energy of the future. A recent survey was taken to compare the cost of solar energy to the cost of gas or electric energy. Results of the survey revealed that the distribution of the amount of the monthly utility bill of a 3-bedroom house using gas or electric energy had a mean of $98 and a standard deviation of $14. If the distribution is normal, what percentage of homes will have a monthly utility bill of more than $84?

98) At one college, GPA's have a distribution that is unimodal and symmetric with a mean of 2.7 and a standard deviation of 0.5. Is a GPA of 3.8 more than one standard deviation above the mean?

Use the following regression equation regarding airline tickets to answer the question.

$\widehat{Price} = 49 + 0.22 \cdot (Distance)$

Note that Distance is the amount of miles between the departure and arrival cities, and Price is the cost of an airline ticket.

99) Interpret the slope of the regression equation in the context of the data.

Using the following probability density curve, answer the question.
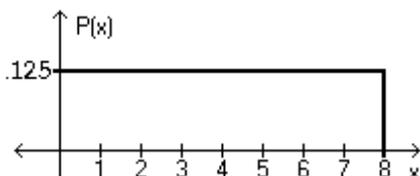
100) What is the probability that the random variable has a value greater than 2?

23

When exposed to heat, the reaction time of a certain chemical always occurs after thirteen minutes, but before 17 minutes. Reaction times for this chemical can be modeled by a uniform distribution, that is, the reaction time is just as likely to occur at thirteen minutes as it is to occur at seventeen minutes.

101) Find the probability that the reaction will happen after fifteen minutes. Shade the appropriate area and calculate the numerical value of the probability.
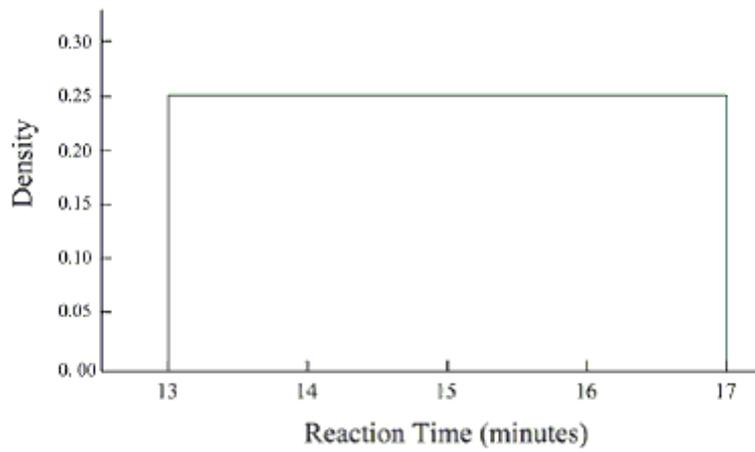
1) B
2) A
3) D
4) C
5) C
6) C
7) C
8) D
9) C
10) A
11) C
12) B
13) B
14) B
15) B
16) A
17) A
18) B
19) C
20) B
21) C
22) D
23) A
24) A
25) A
26) A
27) D
28) A
29) D
30) A
31) B
32) C
33) D
34) D
35) B
36) A
37) B
38) C
39) C
40) B
41) D
42) B
43) C
44) A
45) B
46) B
47) A
48) C
49) B
50) D

51) C
52) B
53) C
54) D
55) D
56) C
57) C
58) 25%; 4.8%
59) 0.100
60) 0.110
61) Various. Any combination of the events "student doing homework – type x" and "student engaged in outside employment – type y".
62) 0.126
63) b
64) The sample
65) 6
66) 28
67) Yes
68) 0.00032
69) 0.044
70) 14.7
71) The second kind of error would be saying that there is no difference in the attendance of statistics students and the student body as a whole at sporting events, even though there really is.
72) We are 95% confident that the proportion of coworkers who went on a vacation last year away from home for at least a week is between 50% and 66%.
73) Confidence interval.
74) It gets larger.
75) Binomial experiment
76) 0.960
77) Both tails of the curve should be shaded and should approximately represent the p-value of 0.060.
78) 111.0 ± 8.5 pounds
79) series number: categorical; weight: numerical; style: categorical
80) The "high critic scores" bar chart shows more variability because there are more observations in the different categories than there are for the "low critic scores."
81) A bar chart represents a categorical variable and a histogram represents a numerical variable.
82) 5
83) Reject the null hypothesis; there is strong evidence to suggest that average volume of paint cans is different than 1.02 gallons. The adjustment in the manufacturing process has affected volume of paint cans.
84) 0.476
85) There are two possible answers here: (1) The number of crimes committed is the explanatory variable and the number of police on patrol is the response because the amount of crime can explain a need for more or less police officers. (2) The number of police on patrol is the explanatory variable and the amount of crime is the response because the more police that are on patrol could explain a reduction in the amount of crime.
86) The outlier would decrease the slope because it would influence the lower left hand side of the regression line to move closer to it. This would make the line less steep, which results in a smaller slope value.
87) Mean Score $= \dfrac{(25 \cdot 80) + (30 \cdot 75)}{25 + 30} = \dfrac{4250}{55} = 77.27$.
88) The mean will be higher than it should be, but the median will not be affected. This is because the median is resistant to outliers, but the mean is not.
89) For a left-skewed distribution, the median value will be larger than the mean.

90) When the values are arranged in order (2, 6, 6, 10, 12, 17, 20, 22, 23, 25), we find that Q3 = 22 and Q1 = 6. Therefore,
IQR = Q3 - Q1 = 22 - 6 = 16.

91) $z = 0.751$

92) a. Mean: 1.6813; median: 1.6844
    b. 1.6562 is an outlier.
    c. Mean: 1.6844; median: 1.6846

93) 0.190

94) 32.6%

95) $r = \sqrt{r^2} = \sqrt{0.379} = 0.6156$. Since the plot shows a positive relationship between height and weight, r = 0.6156.

96) 68%

97) approximately 84%

98) Yes

99) The slope is 0.22. For every additional mile of flight travel, the price of the airline ticket is predicted to increase by
$0.22.

100) 0.750

101) 0.50, the right half of the distribution should be shaded.